DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE
NATIONAL INSTITUTES OF HEALTH

NATIONAL CENTER FOR RESEARCH RESOURCES
BIOMEDICAL TECHNOLOGY AREA

ANNUAL PROGRESS REPORT (8/1/06 – 7/31/07)

1. PHS GRANT NUMBER: P41RR05969
2. NAME OF RECIPIENT INSTITUTION: Macromolecular Modeling and Bioinformatics
3. HEALTH PROFESSIONAL SCHOOL (If applicable):
4. REPORTING PERIOD:
   A. FROM (Month, Day, Year): August 1, 2006
   B. TO (Month, Day, Year): July 31, 2007
5. PRINCIPAL INVESTIGATOR:
   A. NAME: Klaus J. Schulten
   B. TITLE: Professor
   C. SIGNATURE:
6. DATE SIGNED (Month, Day, Year):
7. TELEPHONE NUMBER (Include Area Code): (217) 244-2212
8. FACSIMILE NUMBER (Include Area Code): (217) 244-6078
9. ELECTRONIC MAIL ADDRESS: `kschulte@ks.uiuc.edu`
10. URL: `http://www.ks.uiuc.edu`
11. Was Patent or Copyright awarded this grant year: N
12. Total % effort related to AIDS research: 0%

Contents

General Description of Resource Operation:

The NCRR Resource for Macromolecular Modeling and Bioinformatics develops new methodological solutions for NIH researchers and others in the field of computational biomedicine. The researchers are offered atomic–level microscopic views of cellular processes that guide clinical research, as well as pharmacological and biotechnological development. The views provided, of static structures and dynamic processes, are obtained computationally by combining different experimental modalities, mainly from crystallography, NMR structure analysis, electron microscopy, atomic force microscopy, and single–molecule fluorescence, along with knowledge from physics and chemistry. The "computational microscope" is developed with very advanced concepts from computer science and introduces, through cooperation with various manufacturers and national facilities, the most recent computer technology to biomedicine as soon as it becomes available.

The computational technologies developed and provided combine structural and sequence data with mathematical and computational modeling. They are made available in the form of computer software that runs on a wide range of popular commodity computers as well as on the most advanced computers at leading National Science Foundation centers or in development at manufacturers. Examples range from laptop computers running Windows or Mac OS X, to commodity clusters running Linux, to the IBM Blue Gene machine with 40,000 processors. The three Resource software programs, with over 100,000 registered users combined, are the molecular graphics and sequence analysis program VMD (Visual Molecular Dynamics), the molecular dynamics program NAMD (Nanoscale Molecular Dynamics), and the grid computing and workgroup program BioCoRE (Biological Collaborative Environment). Users of the software receive extensive training in hands-on workshops and responsive service through email. Clinicians, bench scientists, and advanced modelers are served. Software and training material of the Resource are distributed free of charge through a much–visited web site.

The Resource has a long tradition in working closely with biomedical scientists at clinical and biomedical institutions. Without exception, all of the scientific projects conducted by the Resource are collaborations with experimental groups, most of them at medical institutions. Currently active collaborations of the Resource include: investigations into the mechanism of ion channels in collaboration with researchers at the Institute of Molecular Pediatric Sciences, University of Chicago; investigations of virus infection, of proteins involved in hearing, and of inner cellular membrane channels involved in trafficking of proteins all with researchers at Harvard Medical School; investigations of blood clotting factors with researchers at Mayo Clinic, Rochester, MN. Several of the Resource collaborators are Howard Hughes Medical Institute investigators.

During the past funding period, including the last year, core activities (1-3) focused on

the technological development of the three main software programs of the Resource:

(1.) The program VMD for displaying static and dynamic structures, for sequence information, for structure generation and dynamic analysis is continuously enhanced and adapted to the needs of NIH researchers. Improvements to VMD over the past year include new publication rendering features such as export capability for Adobe Acrobat3D and NVIDIA Gelato, reduced file sizes and support for ambient occlusion lighting in the Tachyon ray tracer, and improved scene rendering quality with PIXAR RenderMan. Internal data structure improvements have cut memory used for large data structures in half, which with other improvements now allows VMD to load single structures of up to 72 million atoms, and collections of up to 30,000 molecules at a time for batch mode analysis of molecular structures. Researchers can now visualize electrostatic field lines and animate biomolecular simulations involving time-varying topology using new VMD graphical representations. Plans for VMD development include supporting interactive visualization, analysis, and modeling of large size and long timescale biomolecular simulations, integration of evolutionary, sequence, and genetic information with structural and dynamical information, integration of more tools and databases into the graphical interface, and improved graphics quality.

(2.) The program NAMD is a parallel molecular dynamics code designed for high–performance simulation of large biomolecular systems, and is used on massively parallel computers and other computer clusters by experimentalists and advanced modelers for both large- and small-scale modeling purposes. NAMD provides an "imaging tool" for biomolecular systems that is accessible to novice users such as experimentalists, triggers insights through hands-on interaction with the simulation, provides fast results by scaling to the fastest supercomputers available, and is readily adaptable to the unique requirements of novel simulations such as those employing a mix of all-atom and coarse-graining techniques. The latest release added binaries for Mac OS X on Intel processors, and source code to build on the Cray XT3 and IBM Blue Gene/L supercomputers. Support for reading new CHARMM 31 stream files (combining topology and parameters) and the new CMAP crossterm (dihedral-dihedral) potential function were also added over the last year. Plans for NAMD development include extending simulation size and time scale via petascale computing, incorporation of new simulation methods, simplifying the process through which scientists can develop novel methods, and accelerating simulations with emerging performance technologies.

(3.) The program BioCoRE is a web-based collaborative environment designed to enhance biomedical research and training by facilitating grid computing and research management, and is integrated with both VMD and NAMD. Over the past year functions have been added which allow VMD users to engage in BioCoRE-mediated collaborations (data sharing, chatting, sharing graphics sessions) directly from the graphical interface. BioCoRE

technology now offers VMD users access to software packages running on remote servers, helping researchers analyze more complex problems than possible with common desktop computers. Furthermore, running analysis on a remote server ensures that researchers always has access to the most up-to-date version of the analysis software.

During the past year, the Resource continued to place uniquely strong emphasis on Collaborations, Service, Training and Dissemination.

Collaborations applied the Resource's most advanced modeling capabilities to medically relevant cellular systems investigated by leading intramural and extramural experimentalists making its computational infrastructure and expertise in the field of molecular simulations, structure and sequence analysis available. The Resource has completed 17 joint publications through these collaborations with experimentalists last year (in addition to 10 joint publications with extramural computational biologists). Currently, the Resource is engaged in 17 different collaborations with experimental groups. The Resource adds on average one collaboration each month, completing collaborations also in a timely fashion.

Service is provided for the Resource software VMD, NAMD, BioCoRE through responses to user inquiries, support of user groups, maintenance of program libraries, and provision of a visitor and training center as well as an advanced computer laboratory. Registrations over the last year for VMD increased by 19,000, for NAMD by 4,700, and for BioCoRE by 605. User support continued, with for example, 7,800 exchanges sent to the VMD support email address. Over the last year there were 121 users of the Resource computational facilities, 17 visitors received training at the Resources visitor center, and 16 seminars were organized by the Resource. The Resource continues to offer technical advice, e.g., on building computer clusters and visualization facilities, to both external users and users of our major software packages, and will maintain an excellent seminar series. The Resource will also overhaul its computational infrastructure, in order to offer significantly more disk space to local and remote users. Further, the web and mail servers will be replaced with faster, clustered servers, offering improved reliability and performance.

Training has been and will continue to be available through on-site and online hands-on workshops, tutorials, and case studies. Over the past year, three workshops on computational biology, exploring physical models and computational approaches used for the simulation of biological systems and the investigation of their function at an atomic level were provided, two on-site in Pittsburgh, PA and Talca, Chile, and one online from the Resource web site. Two cluster-building workshops, teaching the design, construction, and deployment of computational clusters, were also provided, each hosted at the Beckman Institute (where the Resource is housed). Five new tutorials were developed on the following topics: bionanotechnology; ion conduction, permeation, and electrostatic maps; VMD images and movies; scripting forces in NAMD; and, membrane proteins. A new

case study on ion channels was also developed, and a case study on water and ice was updated to make use of the new Adobe Acrobat3D technology (readers can manipulate loaded images in 3D space). Future plans for training include continuing the face-to-face computational biophysics workshops, researching new technologies for enhancing online workshops, writing new tutorials and case studies, and compiling training information into a series of printed an online books.

Dissemination is achieved primarily through the Resources highly visited web site, where the biomedical community can download software, access a variety of training materials, get electronic copies of the majority of Resource publications, and view research summaries and exemplary modeling projects. Over the past year, the Resource had 808,000 unique visitors to its web site, resulting in 1.7 terabytes of information transfer; these visitors downloaded over 23,600 publications from the Resources online publications database. Dissemination also prospered in more traditional academic activities, including 41 articles in refereed journals or other publications, 56 talks by Resource PIs and 32 presentations by others; and, 34 stories about the Resource posted in various media outlets. The challenge for the Resource over the next funding period will be to maintain its already high level of dissemination, without devoting more than the present (already extensive) level of resources to these pursuits.

Changes in Resource Direction:

The Resource submitted in October 2006 a proposal for its five–year renewal and had its site visit in April 2007, receiving a very outstanding priority score. In its proposal the Resource has set itself several broad goals: (i) Extend simulation size and timescale through new concepts and technologies; (ii) provide expert and novice modelers with a comprehensive suite of largely automated tools for model building, simulation, visualization, and analysis, integrated with an intuitive graphical interface; (iii) integrate evolutionary, genetic, structural, and dynamics information; and, (iv) develop modeling for bioengineering of medical devices and sensors. While the Resource will continue Collaboration, Service, Training, and Dissemination activities as well as the development of its VMD and NAMD software along the prior successful directions, having received indeed exemplary scores from the site visit panels for its record and plans, the Resource integrated the BioCoRE technology development into VMD and two new Core activities. The two new core activities focus on “structural systems biology” and on “biotechnology”. The VMD and NAMD software are both receiving a huge speed advance through adoption of a new acceleration technology based on a new generation of graphics processing units.

In the past seventeen years, the Resource has developed parallel computing to boost the speeds of its software, in particular, of the program NAMD. The Resource will continue this development, but is also working with great intensity on acceleration of common computational biology tasks by means of graphics processing units. This involves a brand-new

technology and the Resource is working closely with a leading manufacturer (NVIDIA) and electrical and intramural electrical and computer engineers to harness this technology for biomedical research. The solutions being developed will enhance the speed of software by a factor 10–100 at no cost since graphics processors available in workstations and laptops can be used, saving researchers investments into expensive computer clusters. The large speed increase offers also entirely new opportunities in computational biomedicine.

In the new "structural systems biology" core development activity the Resource acknowledges that the very nature of living systems lies in the harmonious hierarchical assembly, regulation, and function of their biomolecular building blocks. In the past, the Resource focused its research and modeling tool development mainly on the building blocks, but is now shifting its research to the assembly level, developing modeling tools for structural systems biology. This shift requires existing tools to become more efficient in order to handle much larger structures, more automated in order to assist the modeler in building and analyzing models, and more comprehensive in order to address the modelers' wider range of tasks. New approaches and tools will also be needed to deal with the larger size and longer time scales required in investigations. In response, the Resource will build a comprehensive "Biomolecular Modeling Suite" to manage and automate large-scale modeling and to extend the use of modeling to a wider community; for very large-scale and long-time simulations of solvent, lipids, proteins, nucleic acids, and their assemblies, it will develop strategies for coarse-grained and multiscale modeling with associated tools and force fields; and, to establish a close link between supramolecular structure simulation and observation, the Resource will develop also tools for combining multimodal data into models and for predicting such data from models.

In the new "Biotechnology" core development activity the Resource will assist in particular NIH researchers in bionanotechnology. This new technology exploits biological processes to create novel technological solutions, either through manufacturing synthetic devices that can directly interact with cellular machinery or by altering the design of biological molecules, such as proteins. Biotechnology has the potential to revolutionize medicine, with computer modeling greatly accelerating the process of designing synthetic biodevices or engineering biological ones. However, modeling tools and methods are significantly less developed than those available in the mainstream life sciences. The Resource aims to dramatically improve the instrumentation of computer modeling in several areas of biotechnology by providing a comprehensive set of modeling tools and methods. Goals include the development of (i) methods and software for modeling silicon bionanodevices; (ii) methods for modeling carbon nanotube-biomolecular systems; (iii) of coarse-grained descriptions of synthetic materials; (iv) methods to accelerate the design of artificial proteins, focusing on altering gas migration rates; and(v) tools for high-throughput management of simulations.

Impact of Resource on Biomedical Research

The impact of the Resource on biomedical research is evident in the number of registered users of its software, the number of citations of its publications, its publication record, and the number of accesses of its web site.

The Resource is known for producing high-quality computational technologies that combine structural and sequence data with mathematical and computational modeling, specifically its software applications VMD, NAMD, and BioCoRE. Acknowledgment from the biomedical community of the utility and quality of Resource software is provided by the number of registered downloads over the lifetime of the software and over the last year (figures are rounded): 19,000 new registered VMD users (90,000 total); 4,700 new registered NAMD users (20,000 total); 600 new registered BioCoRE users (2,000 total). Repeat users, or those users who have downloaded more than one version of a software application, run at around 25% per application. Data across recent user surveys (2005 - 2006) indicate that the majority of users are affiliated with academic institutions (88%), use Resource software for research (86%), and just over one-fourth (26%) are NIH-funded researchers. Survey results further indicate a high degree of satisfaction, with majorities of VMD (94%), NAMD (77%), and BioCoRE (73%) users agreeing or strongly agreeing with the statement "I am satisfied" with the indicated software.

Citations by others of Resource publications are another indication of impact on the biomedical community. A search of the Thompson ISI citations database in March 2007 returned 7,000 citations of all papers published by the Resource since 1989 with 1,330 citations in 2006. The VMD and NAMD source papers received 1,800 and 540 citations, respectively, with 504 and 193 citations, respectively, in 2006. Journals with citations of Resource publications include in 2006 alone Cell, Structure, Nature, EMBO Journal, PNAS, and others. The Resource has published 41 articles over the last year.

That the biomedical community turns to the Resource for software, research, and information is evident in the number of visitors to its web site, which received over 800,000 unique visitors (i.e., non-redundant IP addresses) over the last year. The Resource develops and maintains a broad and active training program, consisting of workshops, tutorials, and case studies, in addition to a visitor program. Over the last year, the Resource conducted five workshops serving 128 members of the biomedical community, via one online and two on-site workshops (in Pittsburgh, PA, and Talca, Chile) on computational biophysics, and two workshops on cluster-building (both held at the Beckman Institute, UIUC). Certainly, the impact of the Resource is also due to its highly developed training through hands-on workshops and through web-based training materials.

Highlights

Structural transitions of HDL and Nanodiscs

Cholesterol is a lipid compound used by animals in cellular membranes and in the synthesis of several endocrine signaling molecules. If not properly managed by the body, however, cholesterol can build up in blood vessels, contributing to heart disease. High density lipoprotein (HDL)* is one of several classes of particles used to absorb and transport cholesterol. HDL, sometimes known as "good cholesterol", lowers the risk of heart disease because of its role in removing excess cholesterol from the body [1–3]. HDL particles are initially created from two copies of a protein, apolipoprotein A-I (apo A-I), which wrap equatorially around a small patch of lipid bilayer, forming a disc-like assembly [4–6]. As these particles absorb cholesterol, they swell and form spherical particles with esterified cholesterol at the core [7]. Treatment using agents which raise HDL levels, or supplementation with artificial HDL-like particles, have shown significant potential for heart disease management [8] and are currently undergoing clinical trials.

In addition to their medical relevance, HDL particles have inspired the creation of artificial constructs known as nanodiscs for the study of membrane proteins [9]. Nanodiscs are composed of a truncated form of apo A-I assembled around a membrane. They have the advantage that unlike HDL, the sizes of nanodiscs may be precisely controlled by the conditions under which they are synthesized [10], and they can also be assembled in the presence of membrane proteins to incorporate those proteins into the nanodisc's bilayer [11–13]. Nanodiscs also provide a more physiological environment than micelles to embed and solubilize membrane proteins. Thus, nanodiscs are useful both as a laboratory model for HDL, and as scaffolds for studying membrane proteins in a realistic environment.

Understanding the assembly of HDL, and its structural transitions during absorption would both aid in the treatment of heart disease and allow the usage of nanodiscs to be optimized. All-atom molecular dynamics (MD) simulations provide the needed information, but the relevant structural transitions occur over timescales on the order of several microseconds that are much longer than can be simulated successfully. To meet this challenge, the Resource developed a new methodology, coarse-grained (CG) molecular dynamics.

Indeed, recent simulations of randomized mixtures of the lipid and protein components of nanodiscs using the Resource's CG simulation techniques revealed the complete assembly mechanism of these particles [14, 15]. These results now guide efforts to tune experimental protocols used in assembling nanodiscs for specific applications, such as incorporating membrane proteins. In addition, they are used to investigate the absorption of cholesterol by HDL. The Resource's results on nanodisc/HDL assembly mechanisms were confirmed

*http://www.ks.uiuc.edu/Research/Lipoproteins

Grant Number: P41RR05969 (8/1/06 – 7/31/07)

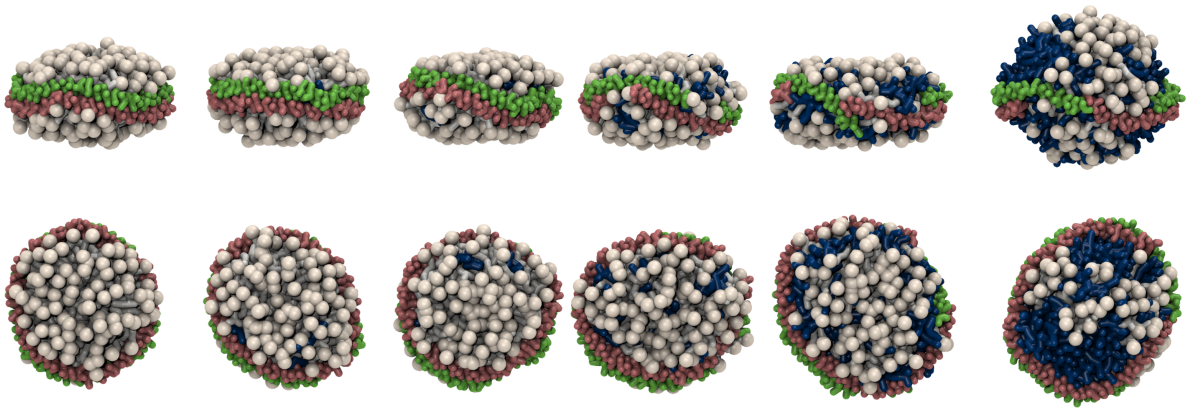

Figure 1: Detergent-induced structural transition of a nanodisc (progressing from left to right), shown in both side (above) and top views. Lipids are shown in white, scaffold proteins in pink and green, and detergent in blue. A period of 2 μs of simulation is shown.

through small-angle x-ray scattering experiments [16, 17]. As an illustration, the expansion of the nanodisc into a spherical particle upon the addition of detergent is shown in Fig. 1. The Resource's current findings on HDL structural transitions, as well as ongoing work involving the interaction of nanodiscs with cholesterol, will guide therapeutic strategies targeting HDL.

The Mechanical Strength of a Blood Clot

The ability of tissues to heal and regenerate is essential for multicellular living organisms. In the vertebrate cardiovascular system, healing of wounded blood vessels proceeds in several steps and relies on complex mechanisms preventing bleeding and pathogen invasion. Blood clots (solid yet elastic clumps of blood cells mixed with fibrin proteins) form part of the emergency response to an injured blood vessel. Blood clots surround the damaged tissue to prevent bleeding; even a minor cut in the skin would result in death if blood clots were not formed to seal the wound. On the other hand, blood clots can restrict essential and normal blood flow if they form at the wrong place, or break free from a larger vessel only to later block a smaller one. The formation of a clot which obstructs the normal blood flow is called thrombosis, and usually results in a severe, if not lethal, cardiovascular disease [18].

The viscoelastic properties of blood clots must be finely tuned: clots must be stiff enough to prevent bleeding, yet flexible to prevent breakage and subsequent translocation to a smaller vessel which can be completely blocked [19]. The origin of blood clot elasticity will naturally depend on the elasticity of its cellular and molecular components. Blood clots are built from red blood cells and a protein called fibrinogen, which, in its monomeric form, consists of three pairs of polypeptide chains assembled as sets of coiled-coils. Fibrinogen, when converted by thrombin into its active form called fibrin, polymerizes into a branched network to form a hemostatic plug in combination with platelets and blood clotting factors [20, 21]. The mechanical properties of fibrin networks are highly dependent on both the network architecture and the mechanical properties of the individual fibers [21].

Interactions between paired chains of fibrin and fibrinogen have been described by several recent studies in which these molecules were stretched using both optical tweezers [23–25] and atomic force microscopy (AFM) [26]. However, the elastic properties of single fibrinogen molecules and their coiled-coil helices, which are the predominant structures along the length of the molecule, remain unclear.

A first step in the characterization of fibrinogen's elasticity at the molecular level has been recently carried out by the Resource. Dr. Bernard Lim (Cardiology, Mayo Clinic) approached the Resource requesting assistance in the interpretation of force–distance profiles obtained using AFM on single fibrinogen molecules. The data required an interpretation in terms of the protein's molecular architecture which could only be obtained through simulation.

Using steered molecular dynamics simulations, a technique pioneered by the Resource [27], the mechanical strength and unfolding pathway of fibrinogen was probed. The simulations involved an elongated, hexameric fibrinogen molecule solvated in a water box large enough to accommodate fibrinogen's stretched conformations. The simulated system comprised

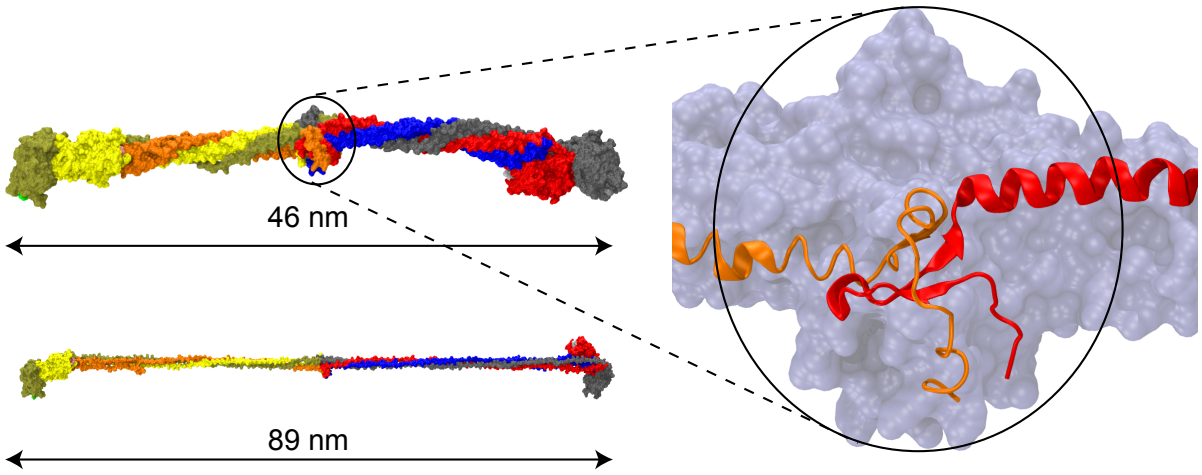

Figure 2: Stretching fibrinogen, a blood clot forming protein. The left panel shows the relaxed fibrinogen conformation (as depicted by crystallography [22]) and the stretched fibrinogen conformation obtained using steered molecular dynamics. The right panel shows a detail view of fibrinogen's central region, where chain entanglement favors a "lock-and-key" mechanism that gives rise to a characteristic elastic response in fibrinogen molecules.

over one million atoms; simulations could only be carried out using the NAMD molecular dynamics package developed by the Resource [28] and especially designed to harness the power of parallel computers. NAMD allowed Resource scientists to identify, for the first time, the specific interactions along fibrinogen's coiled-coil helix region that confer elasticity to this principle blood clot component. The simulation results permitted definitive interpretation of the atomic force microscopy data obtained by the Resource's collaborator, offering a fascinating and unprecedented view of how fibrinogen molecules derive their viscoelastic properties through both coiled-coil interactions and a novel lock-and-key mechanism at their central domain. A manuscript describing this work is currently in preparation.

Grant Number: P41RR05969 (8/1/06 – 7/31/07)

Artificial Design of an O_2-Tolerant Hydrogenase

With the world's oil reserves dwindling, the development of an alternative energy fuel has become an urgent priority. Hydrogen gas (H_2), a renewable resource which boasts zero pollution, is a promising alternative to gasoline. One method of producing H_2 is by means of the unicellular green algae *Chlamydomonas reinhardtii*. Because *Chlamydomonas* has the natural ability to couple photosynthetic water oxidation to the generation of H_2 by means of its hydrogenase enzyme* (see Fig. 3), it can be used to produce H_2 [29–31]. Such a means of H_2 production could be made affordable and efficient, requiring only water and sunlight, with up to 10 % of incident sunlight energy being converted.

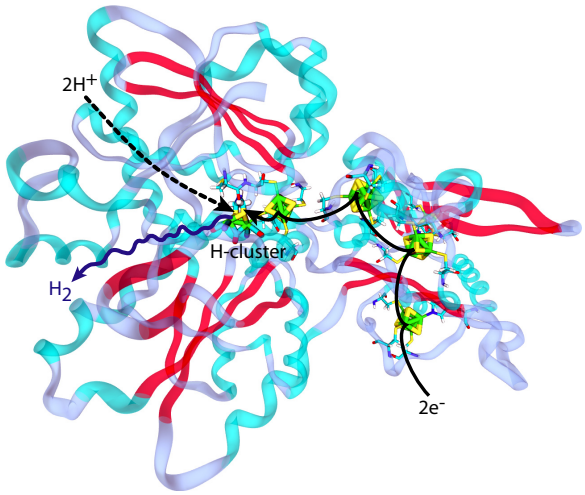

Figure 3: The H_2 production reaction. The hydrogenase enzyme (shown as ribbons) generates H_2 from hydrogen ions (H^+) and electrons (e^-) coming from photosynthetic water oxidation. O_2 inactivates this reaction by binding to the H-cluster's active site irreversibly.

While hydrogenase has much potential as a source of H_2, it unfortunately does not operate under high concentration of oxygen gas (O_2), such as exists in ambient air. In fact, a single O_2 molecule can irreversibly bind to the hydrogenase's buried active site, inhibiting its activity, and leading to enzyme degradation. For this reason, the Resource is collaborating with the National Renewable Energy Lab, towards conferring O_2-tolerance to hydrogenase, by engineering the protein in such a way as to prevent O_2 from being able to access hydrogenase's active site.

Molecular dynamics (MD) simulation is the ideal complement to experiment for making hydrogenase O_2-resistant. MD has been instrumental in identifying the pathways taken by O_2 and H_2 inside hydrogenase. Previous work done by the Resource described the

*URL: `http://www.ks.uiuc.edu/Research/hydrogenase/`

mechanisms of gas permeation in hydrogenase and found the two main pathways by which O_2 reaches the active site of hydrogenase [32–35]. Recently, a method called implicit ligand sampling [36] has been developed, which dramatically increases the usefulness of computational analysis in assessing O_2-tolerance, by providing comprehensive maps of *all* the pathways and energy barriers for gas migration inside any protein. Already, the method has been applied to hydrogenase [37], aquaporin [38], myoglobin [36], copper amine oxidases [39], and monomeric globins [40].

In order to create an O_2-resistant hydrogenase mutant, an effective mutation strategy needs to be developed. Work has been done that investigated how O_2 is transported inside naturally-occurring proteins, specifically globins [40]. Since all globins share a common protein fold, by looking at differences in sequence between various globins we can infer the influence of specific residues on O_2 migration. It was found from this study that flexible and bulky hydrophobic residues promote the formation of pathways, as opposed to blocking them, as was previously assumed. Future mutation studies of hydrogenase will therefore focus on replacing such residues with residues that have been found to statistically block O_2 pathways.

Using this mutation strategy, hydrogenase mutants have been created *in silico*, and tested *in vitro* on recombinant hydrogenase purified from an *E. coli* expression system. One of the mutations, L283W, located in one of the newly discovered pathways, was measured to partially decrease hydrogenase's sensitivity to O_2, both *in silico* and *in vitro* [35, 37]. Given that O_2 has been found by simulation to reach the active site through two independent pathways, work is ongoing to characterize double-mutants, experimentally and computationally, using scenarios in which site-directed mutations target both pathways simultaneously, yet without blocking the H_2 product's outward diffusion. The end product of this work will result in a hydrogenase mutant that can be used for commercial H_2 production using sunlight in the presence of ambient air.

Accelerating Molecular Modeling with Graphics Processors

Biological simulations, such as those performed and supported by the Resource, provide high-resolution information on the cellular functions of biomolecules and are thus useful in designing treatments for diseases. Recent trends have pushed such simulations toward larger and more health relevant biomolecular assemblies such as viruses. However, such large simulations require ever-increasing amounts of computational effort. Currently the most useful and health-relevant simulations can only be performed on extremely costly supercomputers. In this respect, it is of interest to note that there exists an inexpensive source of tremendous compute power in the form of graphics processing units (GPUs).

Previously usable only for graphics processing, GPUs have now evolved to the point that they can be used for general purpose scientific computations. For some tasks, a single state-of-the-art graphics processor can achieve levels of performance hundreds of times greater than a typical computer's central processing unit (CPU). At the same time, GPUs have become increasingly common even in standard consumer PCs, and thus are relatively inexpensive, particularly considering their potential computing power. Thus, if the capabilities of GPUs could be successfully harnessed for scientific applications, some tasks currently requiring supercomputers might be transferred to commonly available desktop computers. To this end, the Resource has transferred key molecular modeling tasks to GPUs and, indeed, achieved performance levels of ten to one hundred times that of traditional CPU implementations. This development, realized in close collaboration with a leading GPU manufacturer (NVIDIA), should save biomedical researchers millions of dollars.

Realistic simulations of biological systems require not only the structures of the biomolecules, but also the placement of a solution of water and ions around them to mimic physiological conditions. At present, most methods for ion placement are too computationally expensive for practical use with large systems, such as viral particles, the ribosome, and other macromolecular assemblies. A new GPU-accelerated ion placement tool has been developed, achieving performance levels up to 100 times faster than CPU-based versions. This tool allows precise ion placement that would have required a day of simulation on a supercomputer to now be performed in a few hours on a desktop computer. Figure 4 shows the ribosome with 1,308 ions placed using GPUs.

Seeing that simulation setup tasks can be aided by GPU acceleration, it is of interest to also apply this technique to the simulations themselves, as these consume a far greater amount of computer time. A GPU-accelerated algorithm for the evaluation of non-bonded interatomic forces has been developed and integrated into NAMD,* the Resource's parallel molecular dynamics program, which is used by thousands of biomedical researchers. This

*`http://www.ks.uiuc.edu/Research/namd/`

Grant Number: P41RR05969 (8/1/06 – 7/31/07)

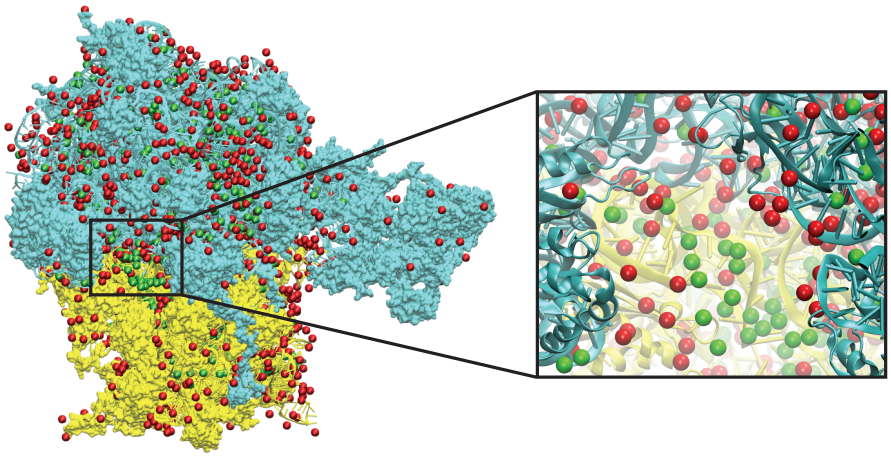

Figure 4: Ions placed on the ribosome structure using GPUs. The inset shows a close-up view. Placed ions are shown in red and ions present in the initial structure are shown in green.

initial implementation reduces the time needed for non-bonded force calculation by 90% compared to a CPU, tripling overall performance. Porting the remaining CPU-based algorithms to the GPU should allow simulations to be performed on small clusters of GPU-containing machines at speeds currently only achievable on supercomputers. Much work lies ahead, but GPU acceleration holds the promise of moving routine molecular dynamics simulations from dedicated compute clusters in the hands of a lucky few, to the desktop workstations of all biomedical researchers.

Towards Petascale Simulations with NAMD

NAMD (Nanoscale Molecular Dynamics),* is a parallel molecular dynamics code designed for high performance simulation of large biomolecular systems [28, 41]. NAMD employs the prioritized message-driven execution capabilities of the Charm++ parallel runtime system,† allowing excellent parallel scaling on both massively parallel supercomputers and commodity workstation clusters. NAMD is distributed free of charge to over 20,000 registered users as both source code and convenient precompiled binaries.

DOE, NSF, and DARPA have all launched aggressive programs to create first-generation petascale supercomputers, achieving petaflop/s theoretical performance, and second-generation machines that will sustain petaflop/s performance on a variety of applications. For example, the NSF solicitation *Leadership-Class System Acquisition—Creating a Petascale Computing Environment for Science and Engineering*‡ states that the computer must demonstrate sustained petaflop/s performance on a suite of applications rather than just a peak-performance requirement. NAMD, running a 100-million-atom simulation of a BAR domain vesicle, is specified as one of the three benchmarks that must achieve sustained petaflop/s performance on the NSF machine. Proteins containing BAR domains play an important role in essential cellular processes (such as vesicle endocytosis at synaptic nerve terminals) by inducing or sensing membrane curvature.

The reasons that NAMD was singled out as a target application for the NSF petascale supercomputer are clear. NAMD is generally acknowledged as the leading cross-platform parallel program for molecular dynamics simulations of biomolecules, and has been since winning a Gordon Bell award in 2002. For example, NAMD was used as early as 2003 for a multi-million-atom, thousand-processor simulation of the ribosome at Los Alamos National Laboratory. More importantly, NAMD is a program that is usable by average biomedical researchers without a background in high-performance computing. The user experience is identical on a laptop or supercomputer, allowing scientists with important applications to move rapidly from learning molecular modeling at one of the many workshops taught by the Resource to running on the most powerful supercomputers available.

NAMD has already shown excellent scaling to thousands of processors on large parallel supercomputers. Figure 5 shows NAMD scaling to 32,768 processors of a Blue Gene/L machine and to 4,000 processors of a Cray XT3. The Blue Gene/L machine at IBM T. J. Watson has 20,480 nodes, each with two 700 MHz PowerPC 440 cores and 512 MB of memory, connected in a 3D torus. The "Big Ben" Cray XT3 at the Pittsburgh

*URL: `http://www.ks.uiuc.edu/Research/namd/`

†URL: `http://charm.cs.uiuc.edu/`

‡`http://www.nsf.gov/pubs/2006/nsf06573/nsf06573.html`

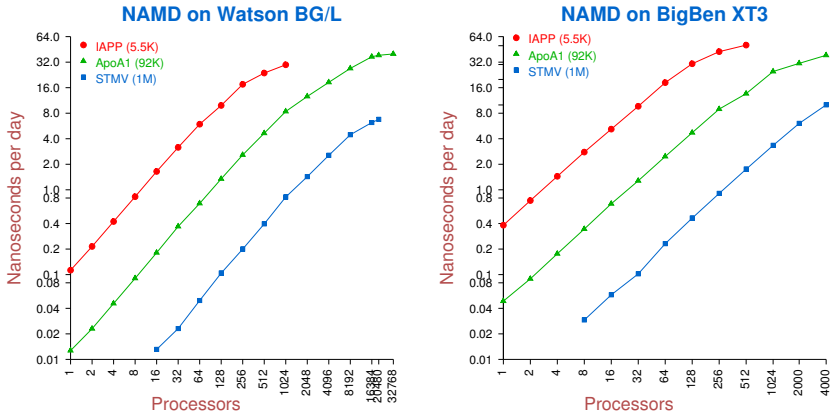

Figure 5: NAMD performance on Blue Gene/L and Cray XT3. The simulations shown represent the full range of simulation sizes currently found in biomedical applications.

Supercomputing Center has 2068 nodes, each with a dual-core 2.6 GHz AMD Opteron processor and 2 GB of memory, again connected by a 3D torus network.

The NAMD developers eagerly accept the challenge of petascale simulations. The NSF BAR domain benchmark represents the upper range of simulation sizes. NAMD has recently been adapted for such large simulations through the compression of the molecular data structure and a two-dimensional decomposition of the particle-mesh Ewald method for long-range electrostatics. Future work will include the parallelization of input, output, and load-balancing to remove the remaining memory and performance bottlenecks.

For more common simulations of 100,000 to 1,000,000 atoms, the goal is to achieve simulation rates of 100 ns per day and beyond. One target application, potassium channels, requires simulations of only 350,000 atoms, but the voltage-sensitive gating events of interest occur on the millisecond timescale. This goal is more difficult than enabling larger simulations, and will require the systematic minimization of the serial calculations and parallel overhead in the program, as well as optimization of parallel communication and load balance. Through this process, NAMD will become an even more powerful tool for biomedical researchers to probe the mechanisms of molecular machines on the world's fastest computers.

Subprojects

BTA UNIT: T

TITLE: VMD

KEYWORDS:

INVEST1: John Stone
DEGREE1: M.S.
DEPT1: Beckman Institute
NONHOST1:

INVEST2: Michael Bach
DEGREE2: B.S.
DEPT2: Beckman Institute
NONHOST2:

INVEST3: Peter Freddolino
DEGREE3: B.S.
DEPT3: Center for Biophysics and Computational Biology
NONHOST3:

INVEST4: Zaida Luthey-Schulten
DEGREE4: Ph.D.
DEPT4: School of Chemical Sciences
NONHOST4:

INVEST5: Elijah Roberts
DEGREE5: B.S.
DEPT5: Center for Biophysics and Computational Biology
NONHOST5:

INVEST6: John Eargle
DEGREE6: B.A.
DEPT6: Center for Biophysics and Computational Biology
NONHOST6:

INVEST7: Dan Wright
DEGREE7: B.S.
DEPT7: School of Library and Information Sciences
NONHOST7:

% BTA $: 10 %

ABSTRACT: VMD [42] is a molecular visualization program that provides interactive biomolecular display and analysis capabilities. VMD incorporates scripting and plugin interfaces for user-extensibility and automation of complex tasks.
(`http://www.ks.uiuc.edu/Research/vmd/`)

VMD runs on all major operating systems and supports computers ranging from laptops to supercomputers. VMD utilizes advanced hardware technologies including stereoscopic displays, six-degree-of-freedom input devices with haptic feedback, multiprocessor and clustered rendering systems, OpenGL, programmable shading, and 64-bit processors.

In the past year, VMD has been improved with many new publication rendering features. These include newly added molecular scene export capabilities for Adobe Acrobat3D and NVIDIA Gelato, greatly reduced file sizes and support for ambient occlusion lighting in the Tachyon ray tracer, and improved scene rendering quality with PIXAR RenderMan. Improvements to VMD internal data structures have decreased the amount of memory used for large structures down to roughly half of that used by prior versions. These and other efficiency improvements enable VMD to load single structures of up to 72 million atoms and collections of over 30,000 molecules at a time for batch mode analysis of structural databases. New graphical representations and visualization tools allow researchers to visualize electrostatic potential field lines and animate biomolecular simulations involving time-varying topology. New and updated plugins provide additional molecular and volumetric file format support.

More than 10,400 new users registered and downloaded VMD 1.8.4 since the previous progress report. Over 17,500 users have registered for VMD 1.8.5 since it was released on August 25, 2006. The latest version, VMD 1.8.6, was released on April 7, 2007.

Ongoing VMD developments include additional tools for structure building, multiple sequence alignment, simulation analysis, new and improved graphical representations, and multiprocessor performance improvements. The next release of VMD is planned for the summer of 2007.

BTA UNIT: T

TITLE: Acceleration of Molecular Modeling Applications with Graphics Processors

KEYWORDS:

INVEST1: Peter Freddolino
DEGREE1: B.S.
DEPT1: Center for Biophysics and Computational Biology
NONHOST1:

INVEST2: David Hardy
DEGREE2: Ph.D.
DEPT2: Beckman Institute
NONHOST2:

INVEST3: Jim Phillips
DEGREE3: Ph.D.
DEPT3: Beckman Institute
NONHOST3:

INVEST4: John Stone
DEGREE4: M.S.
DEPT4: Beckman Institute
NONHOST4:

INVEST5: Leonardo Trabuco
DEGREE5: B.S.
DEPT5: Biophysics
NONHOST5:

% BTA $: BTA %

ABSTRACT: Over the past several years, the hardware and software architecture of graphics processing units (GPUs) has evolved to the point that they can now be used for general purpose scientific computations. State-of-the-art graphics processors include hundreds of individual arithmetic units and can perform up to 500 billion floating point operations per second, a level of performance far above that available with current generation CPU cores. The Resource has implemented several GPU-accelerated computational kernels for key molecular modeling tasks which achieve performance levels of ten to one hundred times that of traditional CPU implementations.

Current methods for ion placement are computationally too expensive for practical use with large systems. Since MD simulations of very large systems are becoming increasingly common, the development of efficient ion-placement methods is highly desirable. A new GPU-accelerated ion placement tool (*cionize*) has been developed, achieving performance levels up to 100 times faster than CPU-based versions using a direct summation Coulomb potential calculation, and an even higher performance method using an approximate multilevel summation of Coulomb potentials on the GPU.

Ensemble-averaged electrostatic potentials of biomolecules provide a more useful description than potentials calculated from static structures. For example, a time-averaged potential is much more realistic for regions where ions bind transiently, *i.e.*, where a static picture would either overestimate or underestimate the ionic contribution when compared to the potential averaged over relevant timescales. Even in the absence of large center of mass motion, one may want to calculate the electrostatic potential experienced by a subsystem in the simulation, which requires that every frame is fitted to a reference structure with respect to the subsystem. However, such fitting is not possible if one is to employ periodic boundary conditions. The use of direct Coulomb summation eliminates this shortcoming, but is prohibitively expensive for the calculation of time-averaged electrostatic potentials using CPUs. However, performing the calculations on GPUs makes the use of direct Coulomb summation feasible. A GPU-accelerated direct Coulomb summation has been implemented in the volumetric data processing framework of VMD.
(`http://www.ks.uiuc.edu/Research/vmd/`)

An initial GPU-accelerated algorithm for evaluation of non-bonded interatomic forces has been developed and integrated into NAMD.
(`http://www.ks.uiuc.edu/Research/namd/`) This initial implementation achieves performance levels up to 10 times faster than a corresponding CPU implementation, but since non-bonded force computations only account for some of the overall work, the overall net performance increases only by a factor of 3.5.

BTA UNIT: T

TITLE: NAMD: Scalable Molecular Dynamics Software

KEYWORDS: molecular dynamics simulation, high-performance computing, parallel programming

INVEST1: James Phillips
DEGREE1: Ph.D.
DEPT1: Beckman Institute
NONHOST1:

INVEST2: Laxmikant Kale
DEGREE2: Ph.D.
DEPT2: Beckman Institute
NONHOST2:

INVEST3: Abhinav Bhatele
DEGREE3: B.S.
DEPT3: Computer Science
NONHOST3:

INVEST4: David Kunzman
DEGREE4: M.S.
DEPT4: Computer Science
NONHOST4:

INVEST5: Chee Wai Lee
DEGREE5: M.S.
DEPT5: Computer Science
NONHOST5:

INVEST6: Chao Mei
DEGREE6: B.S.
DEPT6: Computer Science
NONHOST6:

% BTA $: BTA %

ABSTRACT: NAMD (Nanoscale Molecular Dynamics, http://www.ks.uiuc.edu/Research/namd/) is a parallel molecular dynamics code designed for high performance simulation of large biomolecular systems [28, 41]. NAMD employs the prioritized message-driven execution capabilities of the Charm++/Converse parallel runtime system (http://charm.cs.uiuc.edu/), allowing excellent parallel scaling on both massively parallel supercomputers and commodity workstation clusters. NAMD is distributed free of charge to over 20,000 registered users as both source code and convenient precompiled binaries.

NAMD 2.6 was released in August 2006 and has been downloaded by over 4400 users, 800 of whom are NIH-funded. This release added binaries for Mac OS X on Intel processors and source code to build on the Cray XT3 and IBM Blue Gene/L (neither platform supports dynamic linking and experience shows that binaries would be quickly out of date, so only source code is released). The replica exchange method is implemented as a set of Tcl scripts that use socket connections to drive a set of NAMD jobs, exchanging temperatures (or any other scriptable parameter) based on energy. Both NAMD and the psfgen structure-building tool read and interpret the new CHARMM 31 stream files (combining topology and parameters) and the new CMAP crossterm (dihedral-dihedral) potential function.

The new 5000-core Dell Infiniband cluster at the Texas Advanced Computing Center (TACC) became available for NAMD porting in November, 2006. This was the first major Infiniband-based cluster available to the Resource. NAMD was rapidly ported to the machine and serial performance was excellent, but parallel scaling was poor until recently due to issues with the MPI library—NAMD now performs well on the machine. More aggressive parallel performance tuning has targeted the Cray XT3 and IBM Blue Gene/L platforms, including optimizations to map communication efficiently to the Blue Gene/L toroidal network. The performance of larger simulations has been improved by a two-dimensional decomposition of the particle-mesh Ewald method, and compression of the molecular structure to reduce memory usage allows multi-million-atom simultions to fit in the 256–512 MB of memory available per-core on the Blue Gene/L.

BTA UNIT: T

TITLE: NAMD-Lite and Methods Development

KEYWORDS: molecular dynamics simulation, methods development

INVEST1: David Hardy
DEGREE1: Ph.D.
DEPT1: Beckman Institute
NONHOST1:

INVEST2: Aleksei Aksimentiev
DEGREE2: Ph.D.
DEPT2: Beckman Institute
NONHOST2:

INVEST3: Zhongzhou Chen
DEGREE3: B.S.
DEPT3: Physics
NONHOST3:

INVEST4: Jeffrey Comer
DEGREE4: B.S.
DEPT4: Physics
NONHOST4:

INVEST5: Eduardo R. Cruz-Chu
DEGREE5: B.Sc.
DEPT5: Biophysics
NONHOST5:

% BTA $: BTA %

ABSTRACT: NAMD-Lite (http://www.ks.uiuc.edu/Development/MDTools/namdlite/) is a rapid prototyping framework for developing new simulation methods for biomolecules, consisting of sequential C language code with a modular design. The intention is to separate the development of methods from the additional complication due to parallel implementation, providing a simpler way to test and validate new methods before bringing them into NAMD (http://www.ks.uiuc.edu/Research/namd/). The source code is distributed under the University of Illinois/NCSA Open Source License

(http://www.opensource.org/licenses/UoI-NCSA.php) to allow scientists complete freedom over all of their code modifications.

New NAMD-Lite capabilities include the particle-mesh Ewald method for electrostatics, temperature control methods (Berendsen, Nosé-Hoover), rigid bond constraints for water (the SETTLE method), fixed atoms, and harmonic atomic restraints.

NAMD-Lite also offers several innovative methods. An energy-conserving correction to the force that conserves linear momentum when using grid-based methods for electrostatics [43] was initially validated in NAMD-Lite before being implemented in NAMD. The multilevel summation method [44,45] offers fast electrostatic evaluation for both periodic and nonperiodic boundary conditions. This method has been used to speed up ionization of a molecular system, recently implemented in the Cionize plugin to VMD (http://www.ks.uiuc.edu/Research/vmd/) and accelerated using graphics processors [46]. Methods for polarizable force fields improve accuracy by modeling electron density redistribution due to an electric field; new development includes fast methods for induced point dipoles solved to self-consistency [47] and the dual-thermostat integration of classical Drude oscillators [48]. Methods for modeling the annealing of amorphous silica enable simulations of biotechnology devices; new development includes multiple force field parameterizations of the Buckingham potential for silica.

Future work includes parallelizing multilevel summation for NAMD, extending the Drude implementation to include polarizable ions solvated using a five-point polarizable water model [49], implementing the fluctuating charge model for polarization, and implementing implicit solvent models.

BTA UNIT: T, D, S

TITLE: BioCoRE

KEYWORDS: web-based collaboratory, software engineering, internet, evaluation, collaborative research environment

INVEST1: Laxmikant V. Kalé
DEGREE1: Ph.D.
DEPT1: Computer Science
NONHOST1:

INVEST2: Robert Brunner
DEGREE2: B.S.
DEPT2: Beckman Institute
NONHOST2:

INVEST3: Kirby Vandivort
DEGREE3: M.S.
DEPT3: Beckman Institute
NONHOST3:

INVEST4: Michael Bach
DEGREE4: B.S.
DEPT4: Beckman Institute
NONHOST4:

% BTA $: BTA %

ABSTRACT: BioCoRE [50] (http://www.ks.uiuc.edu/Research/biocore/) is a web-based collaborative environment designed to enhance biomedical research and training. By using a standard web-browser (on a desktop or laptop computer or hand held PDA) scientists create projects in which all private data is secure and shared only within the specific project team. Researchers use BioCoRE to create input files for supercomputer runs, submit jobs to remote sites including supercomputers, and share the visualization of molecular systems across distances. BioCoRE features a synchronous and asynchronous chat, a project-wide "bookmarks" file for sharing web links, as well as a web-based filesystem. Summary pages within BioCoRE regularly inform the project team of the project status.

In the past year the interactions between BioCoRE and VMD have been greatly enhanced. Directly from the VMD graphical user interface, users can now create

BioCoRE accounts, join projects, and chat with collaborators using a VMD version of the BioCoRE Control Panel. Additionally, the NAMDEnergy and APBSRun Plugins within VMD have been extended to use BioCoRE's supercomputer job submission. As a direct result of the above efforts, the generic programming interfaces for BioCoRE have been enhanced, which benefits researchers wishing to write their own interfaces to the BioCoRE environment.

Future BioCoRE efforts will involve the development of key pieces of infrastracture needed for the modeling suite of biomolecular tools.

BTA UNIT: S

TITLE: Computational Facility

KEYWORDS: parallel computing, visualization, network

INVEST1: Tim Skirvin

DEGREE1: B.S.

DEPT1: Theoretical and Computational Biophysics

NONHOST1:

% BTA $: 10%

ABSTRACT: The last year has seen three major improvements to the Resource's computational facility (http://www.ks.uiuc.edu/Development/Computers/): the addition of several serial workstations (two of which take advantage of GPU technology), the addition of 25 TB of disk space, and the roll-out of a wiki-based internal web site for management.

In order to simulate and analyze ever-larger molecular systems, our users need access to fast, large-memory workstations. To provide this, the Resource has purchased five Sun Ultra 40 workstations, each with four compute cores and 16 GB of memory. These systems have proven to be extremely stable and resilient, and allow researchers to utilize long-running analysis scripts. Additionally, the Resource has deployed a pair of GPU-accelerated serial machines, offering as much as a 300x speed boost for tools such as cIonize.

The Resource has recently increased the amount of available data storage on the local network from 20 TB to 45 TB. This was implemented through the purchase of a SunFire X4500 server, containing 48 500 GB hard drives and capable of serving the data directly to the entire network. Additionally, through the use of Sun's ZFS file system, this space is significantly more flexible than what was previously available.

Finally, in order to better maintain its web-based materials, the Resource has recently replaced its internal Group Manual with a wiki-based system, based on MediaWiki. This system allows for significantly faster distribution of information throughout the Resource"s user community, and can be maintained by all Resource members and users. At the same time, formal management of other sections of the web site is handled by the newly-formed web management team.

A large part of our computational growth has again come from the national supercomputing centers. The total number of raw Service Units awarded to us by the Large Resource Allocations Committee

(http://www.ks.uiuc.edu/Development/Computers/nrac.html) has more than doubled compared to last year, increasing to a total of 8.4 million service units (963 CPU-Years). This time is supplemented by the 480 processors on our local compute clusters, which remain unchanged from last year.

BTA UNIT: C

TITLE: Gas and Ion Conduction in Aquaporins

KEYWORDS: aquaporins, water channels, membrane proteins, gas conduction, ion conduction

INVEST1: Emad Tajkhorshid
DEGREE1: Ph.D.
DEPT1: Beckman Institute
NONHOST1:

INVEST2: Jin Yu
DEGREE2: M.S.
DEPT2: Physics
NONHOST2:

INVEST3: Yi Wang
DEGREE3: B.S.
DEPT3: Biophysics
NONHOST3:

INVEST4: Andrea Yool
DEGREE4: Ph.D.
DEPT4: Physiology and Pharmacology
NONHOST4: University of Arizona

INVEST5: Walter F. Boron
DEGREE5: M.D., Ph.D.
DEPT5: Cellular and Molecular Physiology
NONHOST5: Yale University

% BTA $: BTA %

ABSTRACT: Aquaporins (AQPs) (URL: `http://www.ks.uiuc.edu/Research/aquaporins/`) are a family of membrane channels specializing in rapid water conduction across biological membranes [51, 52]. They are widely distributed in all forms of life. Impaired functions of AQPs have been directly linked to diseases such as cataracts and diabetes insipidus [53–56]. Aquaporins form tetramers in the membrane, with each of the four monomers functioning as an independent water pore. A fifth pore is formed in the middle of the tetramer, known as the central pore. MD simulations of membrane-embedded, fully hydrated models of various AQPs with NAMD [28]

have revealed the role of the central pore in gas and ion conduction [38, 57]. A proper description of these proteins in their natural environment of lipid and water requires system sizes of 100,000 atoms or more, and simulation times on the order of tens of nanoseconds.

In close collaboration with W. Boron, the Resource has investigated the permeation of O2 and CO2 through both pure lipid bilayers and membrane-embedded models of AQP1 using 140 ns of MD simulations [38]. The free energy profiles (PMFs) associated with gas permeation through AQP1 were calculated through both explicit gas diffusion simulations and the implicit ligand sampling, a method developed by the Resource to study gas migration pathways inside proteins [36]. The central pore of AQP1, being completely hydrophobic and empty, is found to be a gas reservoir. Furthermore, the PMFs associated with gas permeation revealed an energy well in the middle of the AQP1 central pore, and an energy barrier of 3.6 to 4.6 kcal/mol on the periplasmic side. The low energy barrier suggests that the central pore of AQP1 could indeed serve as a pathway for gas permeation across the membrane [38]. In contrast, the monomeric water pores of aquaporins were found to be much less gas-permeable, likely because of the strongly hydrogen-bonded water molecules in the water pores.

Using MD simulations of a membrane-embedded model of AQP1, the Resource has also investigated the ion conductivity of the AQP1 central pore. By constraining a sodium ion at various locations along the pore axis and simulating the conduction process, energy barriers against ion permeation have been identified [57]. An interesting and somewhat unexpected behavior of the protein is the conformational coupling of the pore-lining residues with a conserved cytoplasmic loop. This loop was found to bind to the gating signal molecule cGMP. The binding initiated conformational changes on the cytoplasmic half of the protein, which in turn resulted in the opening of the central pore. The involvement of this cytoplasmic loop in the central pore gating was further verified by our collaborator A. Yool through experiments, in which the mutation of two arginines from the loop resulted in an almost complete loss of cGMP-activated ion conduction as reported in a joint publication [57].

BTA UNIT: C

TITLE: Magnetic Field Effects in Cryptochrome

KEYWORDS: cryptochrome, magnetoreception, radical pair, arabidopsis thaliana

INVEST1: Danielle Chandler
DEGREE1: B.S.
DEPT1: Physics
NONHOST1: University of Illinois at Urbana-Champaign

INVEST2: Ilia Solov'yov
DEGREE2: M.S.
DEPT2: Physics
NONHOST2: Frankfurt Institute for Advanced Studies, Johann Wolfgang Goethe University

% BTA $: BTA %

ABSTRACT: It has been suggested that the "magnetic compass sense" of migratory birds and some other magnetoreceptive animals may be mediated by the blue-light-receptor protein cryptochrome. Recent experiments on plant seedlings have shown that the activity of cryptochrome in *Arabidopsis thaliana* is enhanced by the presence of a weak external magnetic field, confirming the ability of cryptochrome to harbor magnetic field responses. Additionally, cryptochrome has been found in retinal cells of birds known to be active during orientation behavior. Cryptochrome's signaling is regulated by the photoreduction of an internally bound chromophore, flavin adenine dinucleiotide (FAD). The spin chemistry of this photoreduction process, which involves electron transfer from a chain of three tryptophans, is modulated by the presence of a magnetic field via the radical pair mechanism. Computational studies of the magnetic field dependence of cryptochrome's signaling activity may give insights as to how or whether this radical pair mechanism could be involved in the avian magnetic compass (URL: `http://www.ks.uiuc.edu/Research/cryptochrome/`).

The Resource and its collaborator, I. Solov'yov, modeled the photoreduction pathway of *Arabidopsis thaliana* cryptochrome, using realistic hyperfine coupling constants and reaction rate constants. This model was used to investigate the activation behavior of cryptochrome in an external magnetic field by studying the effects of the magnetic field on the FAD reduction pathway. It was found that the activation yield of the protein varied by as much as 10% over the range 0-5 Gauss [58]. This result, combined with the consistency of a radical-pair-based magnetic compass with observed features of migratory bird behavior, strengthens the case for the feasibility of a radical pair effect in cryptochrome as a mechanism for the avian

magnetic compass. The Resource plans to refine its model of the photoreduction pathway in addition to exploring the possibility of radical pair effects in other parts of the cryptochrome protein.

BTA UNIT: C

TITLE: The Mechanical Strength of a Blood Clot

KEYWORDS: fibrinogen, protein mechanics, thrombosis, stroke, cardiovascular disease, cerebrovascular accident

INVEST1: Eric H. Lee
DEGREE1: B.S.
DEPT1: Medicine and Biophysics
NONHOST1:

INVEST2: Marcos Sotomayor
DEGREE2: M.S.
DEPT2: Physics
NONHOST2:

INVEST3: Bernard B. Lim
DEGREE3: M.B.B.Ch., Ph.D.
DEPT3: Cardiology
NONHOST3: Mayo Clinic

% BTA $: BTA %

ABSTRACT: Regeneration and protection of the cardiovascular system is essential for vertebrates. Blood clots (solid yet elastic clumps of blood cells mixed with fibrin proteins) form part of the emergency response to an injured blood vessel; they surround the damaged tissue stopping bleeding and blocking invasion by foreign pathogens. On the other hand, blood clots can restrict essential and normal blood flow if they form at the wrong place, or break free from a larger vessel only to later block a smaller one (thrombosis) [18]. Therefore, blood clots must be stiff enough to seal wounded vessels, yet flexible to prevent breakage and subsequent blockage of small vessels [19].

Blood clots are built from red blood cells and a protein called fibrinogen. In its active form, fibrinogen is converted by thrombin into fibrin, which polymerizes into a branched network to form a hemostatic plug in combination with platelets and blood clotting factors [20, 21]. The mechanical properties of blood clots are highly dependent on both the network architecture of fibrin and the mechanical properties of fibrin's individual components [21].

Interactions between paired chains of fibrin and fibrinogen have been described by several recent studies that stretched these molecules using both optical tweezers [23–

25] and atomic force microscopy (AFM) [26]. However, the elastic properties of single fibrinogen molecules and their coiled-coil helices, the predominant structures along the length of the molecule, remain unclear.

A first step in the characterization of fibrinogen's elasticity at the molecular level has been recently carried out by the Resource. Using steered molecular dynamics (SMD) simulations, a technique pioneered by the Resource [27], the mechanical strength and unfolding pathway of fibrinogen was probed. The simulations involved two independent systems containing trimeric and hexameric fibrinogen molecules, respectively. Each system was solvated in water boxes large enough to accommodate fibrinogen's stretched conformations. The resulting systems comprised 365,000 and 1,008,000 atoms and were simulated with NAMD for over 10 nanoseconds each. Long-range electrostatic interactions were computed using a novel parallel implementation of the particle mesh Ewald method (pencil decomposition), which permitted efficient use of 1024 processors in parallel for the largest system simulated. After equilibration, fibrinogen was stretched to over twice its resting length in each case, permitting the visualization of the structures within the molecule that elongate due to the stretching force. The simulation results permitted definitive interpretation of atomic force microscopy data obtained by the Resource collaborator (Dr. B. Lim, Mayo Clinic) and explained, at the molecular level, how fibrinogen is capable of buffering significant mechanical force. A manuscript describing this work is currently in preparation.

BTA UNIT: C

TITLE: Molecular Basis of Bacterial Motility

KEYWORDS: Bacterial motility, Flagellum, Polymorphism, Protein aggregates, Large simulations, Coarse-grained simulations

INVEST1: Anton Arkhipov
DEGREE1: M.S.
DEPT1: Physics
NONHOST1:

INVEST2: Peter Freddolino
DEGREE2: B.S.
DEPT2: Center for Biophysics and Computational Biology
NONHOST2:

INVEST3: Keiichi Namba
DEGREE3: Ph.D.
DEPT3: Graduate School of Frontier Biosciences
NONHOST3: Osaka University, Japan

% BTA $: BTA %

ABSTRACT: Many types of bacteria propel themselves through liquid media using whip-like structures known as flagella (`http://www.ks.uiuc.edu/Research/flagellum/`). The bacterial flagellum is a huge (several micrometers long, 20 nm wide), multiprotein assembly built of three domains: a basal body, acting as a motor; a hook, acting as a joint; and a filament, which makes up the bulk of the length of the flagellum and interacts with solvent to propel the bacterium. Depending on the direction of the torque applied by the basal body, the filament assumes different helical shapes. Under counter-clockwise rotation (as viewed from the exterior of the cell), several flagella form a single helical bundle which propels the cell along a straight line, referred to as running mode [59]. Under clockwise rotation, the individual flagella dissociate from the bundle and form separate right-handed helices, causing the cell to tumble. Varying the duration of running and tumbling, bacteria can move up or down a gradient of an attractant or repellent by a biased random walk.

One of the unresolved questions about the flagellum is how the reversal of torque applied by the motor results in a switching between the helical shapes of the filament. This switching is a result of polymorphic transitions in the filament, when individual protein units slide against each other [60], but its molecular mechanism

remains poorly understood despite extensive experimental work [60–62] and a recent computational study [63]. All-atom simulations presently cannot reach the time and length scales relevant to the structural transitions of the filament; therefore, Resource scientists applied the shape-based coarse-graining (CG) method [64, 65] to simulate the filament, using NAMD [28]. In this method, monomer proteins that the filament is built of were represented by 15 CG particles each (about 500 atoms per CG particle) mapped on the structure using a neural networking algorithm. The effective potentials for interactions between CG particles were derived from all-atom MD simulations. Resource scientists simulated a 0.5-micrometer-long flagellar filament for 100 microseconds; it was found that viscosity of the solvent plays a key role in inducing polymorphic transitions between filament coiling states [65].

Ongoing efforts in this project focus on studying the hook and the hook-filament connection. Resource collaborator K. Namba was able to obtain a low-resolution cryo-EM structure of the hook, and a high-resolution X-ray structure of a part of the hook unit protein. The missing domains of the unit protein have been built by Resource researchers using the ab initio structure prediction software Rosetta [66, 67], after which the complete units were fitted into the cryo-EM map and their coordinates refined using VMD [42]. The shape-based CG model of the hook has been built and connected with the CG model of the filament. The resulting complex is currently being simulated on time scales of hundreds of microseconds, to study torque propagation through the entire hook-filament assembly.

BTA UNIT: C

TITLE: Molecular Mechanism of PcrA Helicase

KEYWORDS: helicase, molecular motor, ATP hydrolysis, DNA, MD, QM/MM

INVEST1: Jin Yu
DEGREE1: M.S.
DEPT1: Physics
NONHOST1:

INVEST2: Markus Dittrich
DEGREE2: Ph.D.
DEPT2: Physics
NONHOST2:

INVEST3: Taekjip Ha
DEGREE3: Ph.D.
DEPT3: Physics
NONHOST3: HHMI

% BTA $: BTA %

ABSTRACT: DNA helicases (http://www.ks.uiuc.edu/Research/helicase/) are ATP-driven molecular motors involved in all aspects of DNA metabolism, such as replication, transcription, and repair. They catalyze the separation of double stranded DNA into its constituent single stranded DNA (ssDNA) components. PcrA is a monomeric 3' to 5' helicase for which several atomic resolution X-ray structures have been reported [68]. The DNA unwinding action of PcrA is caused by helicase translocation along ssDNA, which is driven by the hydrolysis of ATP at a single catalytic site. The translocation proceeds at a rate of approximately 50 nucleotides per second [69].

The Resource has used complementary approaches to investigate PcrA from the electronic to the functional level [70–73]. At the catalytic level, combined quantum mechanical/molecular mechanical (QM/MM) simulations were used to study the ATP hydrolysis reaction pathway and its coupling to protein conformational changes. The simulation system consisted of about 20,000 atoms, 77 of which were treated quantum mechanically at the B3LYP/6-31G level of theory. The simulation results [70, 72] reveal a close similarity in a structure and reaction pathway of PcrA to those of F1-ATPase which the Resource previously studied [74, 75]. These similarities include a proton relay mechanism important to efficient ATP hydrolysis and an "arginine finger" residue that is key to the coupling of the chemical reaction

to protein conformational changes. Additional *in silico* mutation studies also reveal that the residue Q254 is crucial for the coupling of ssDNA translocation to the actual catalytic event [70, 72].

The ATP hydrolysis powered ssDNA translocation of PcrA was first studied by molecular dynamics (MD) simulations [71,72]. Using the program NAMD [28], MD simulations were performed on a fully solvated PrcA-DNA complex with and without ATP bound to the catalytic site. The system contained approximately 110,000 atoms and the simulations lasted for about ten nanoseconds. On the basis of these simulations, effective potential energies governing individual domain movement of PcrA were derived and utilized in millisecond Langevin stochastic simulation [71,72]. The calculations support a domain stepping mechanism in which, during one ATP hydrolysis cycle, the pulling together and pushing apart of two translocation domains is synchronized with alternating domain mobilities such that PcrA moves unidirectionally along ssDNA [71, 72]. In order to substantiate this mechanism, further analyses based on MD simulations, elastic network theory, and multiple sequence alignment have been conducted [73]. These analyses provide further physical evidence for the alternating domain mobility mechanism and reveal essential residues that are coevolutionarily coupled [76]. Insight obtained from these studies is expected to guide *in vitro* mutational studies to achieve the reverse of PcrA translocation.

BTA UNIT: C

TITLE: Gas Migration Pathways in Proteins

KEYWORDS: gas migration, hydrogen production, implicit ligand sampling

INVEST1: Jordi Cohen
DEGREE1: Ph.D.
DEPT1: Beckman Institute
NONHOST1:

INVEST2: Maria Ghirardi
DEGREE2: Ph.D.
DEPT2:
NONHOST2: National Renewable Energy Laboratory, CO

INVEST3: Paul King
DEGREE3: Ph.D.
DEPT3:
NONHOST3: National Renewable Energy Laboratory, CO

INVEST4: Kwiseon Kim
DEGREE4: Ph.D.
DEPT4:
NONHOST4: National Renewable Energy Laboratory, CO

INVEST5: Richard J. Kufrin
DEGREE5: M.S.
DEPT5:
NONHOST5: National Center for Supercomputing Applications

INVEST6: Michelle Gower
DEGREE6: M.S.
DEPT6:
NONHOST6: National Center for Supercomputing Applications

% BTA $: BTA %

ABSTRACT: Many proteins perform their function by interacting with gas molecules such as oxygen and carbon monoxide. Unlike in the case of most other ligands, the protein reactive sites for these gases are often buried deep inside the proteins, with no obvious entry pathway. Describing the locations of the gas pathways and the mechanism of gas migration is an essential step towards understanding how proteins such as globins, oxygenases, oxidases work. The knowledge can then be directly applied to solve concrete and relevant problems. One example is the commercial production of hydrogen gas for use as an energy fuel, using hydrogenase: knowing the gas migration pathways inside hydrogenase provides crucial guidance to the re-design of the enzyme through targeted mutations, which can then be used to affordably produce H2.

To map oxygen pathways in proteins, the Resource has developed a novel algorithm – implicit ligand sampling – for accurately predicting gas migration pathways inside any protein [36]. This methodology was then applied to the case of myoglobin, a protein whose gas conduction properties have been well characterized both experimentally and theoretically over decades. The methodology was also used to explain oxygen migration inside a family of copper-amine oxidases [39] for which results have been corroborated with experimentally-determined xenon binding sites, and to determine the mechanism of gas conduction across an aquaporin water channel [77]. Furthermore, a set of closely related studies also investigated oxygen and hydrogen gas migration inside hydrogenase from Clostridium pasteurianum [32–35, 37]. These studies confirmed the function of a previously hypothesized pathway and also discovered the existence of a second major gas pathway inside hydrogenase, thus providing a target for mutations that would benefit the engineering of hydrogenase into a commercially-viable hydrogen-producing enzyme.

Finally, the Resource completed the development of the NAMD-G grid submission engine [78]. NAMD-G is a software packages that automaties the process of running jobs, generating and recovering files, on remote TeraGrid supercomputing resources. The Resource has used NAMD-G to farm out the computations of oxygen migration maps for 12 monomeric globins. This study revealed the surprising result that oxygen migration pathways are not conserved within proteins families [40]. Furthermore, by investigating the correlation between residues and pathways, certain specific residue types were found to directly influence the formation of oxygen pathways in proteins, and can be taken into consideration when designing mutations that affect oxygen migration in proteins [40].

BTA UNIT: C

TITLE: Genetic Regulation by Proteins

KEYWORDS: multiscale, coarse-grained, elastic rod, DNA, lac repressor, molecular dynamics, gene control, genetic switch, protein-DNA interaction

INVEST1: Alexander Balaeff
DEGREE1: Ph.D.
DEPT1: Chemistry
NONHOST1: Duke University

INVEST2: L. Mahadevan
DEGREE2: Ph.D., Professor
DEPT2: Applied Mathematics and Mechanics
NONHOST2: Harvard University

INVEST3: Rob Phillips
DEGREE3: Ph.D., Professor
DEPT3: Physics
NONHOST3: Caltech

INVEST4: Elizabeth Villa
DEGREE4: B.S.
DEPT4: Center for Biophysics and Computational Biology
NONHOST4:

% BTA $: BTA %

ABSTRACT: The mechanical manipulation of DNA is central to all aspects of genetic expression in cells [79]. In order to express or repress certain genes, proteins mechanically manipulate DNA away from its equilibrium configuration by introducing bends, loops, twists and supercoils [80] against mechanical strain arising from the DNA. In some cases, this is done to preclude other proteins from binding DNA, *e.g.*, to prevent gene expression. Little is known about how regulatory proteins are able to handle and withstand the forces stemming from DNA. Often, the changes in DNA structure are known to occur, but determination of the structure eludes experimental techniques when the changes are too large or when the DNA becomes disordered. This is the case when regulatory proteins force DNA into loops.

The *lac* repressor (LacI) protein, a paradigm of genetic regulation, controls the function of the *lac* operon in *E. coli*, a set of genes involved in lactose catabolism [81]

by bending the DNA into a loop. The all-atom structure of LacI bound to its DNA binding sites has been determined, but with the DNA loop missing [82, 83]. It is therefore unclear what are the mechanisms that the protein uses to resist the strain from the connecting DNA loop, which likely changes the structure of LacI.

The Resource developed a multiscale method for simulating protein-DNA complexes [84] that combines molecular dynamics simulation with an elastic rod model of DNA. The elastic rod model accounts for the physical properties of DNA to build the structure of the missing DNA loop and computes the forces that this loop exerts on the protein [85–88]. Molecular dynamics simulation of the protein incorporates these forces using the SMD method [89], revealing the structural dynamics of the protein-DNA interaction. The geometry of the DNA loop and the forces stemming from it are updated and exchanged every 10 ps of MD simulation. The multiscale method was applied to LacI in complex with a 76 base pair elastic loop [84]. The MD part of the simulation encompassed 320,000 atoms, simulated with the NAMD program on 256 processors.

The results of the simulations have revealed the preferred loop that LacI induces on DNA and the mechanical properties of the protein. Contrary to what was believed, the simulations suggest that the protein maintains its overall conformation by allowing its flexible DNA binding domains to control the ability of LacI to enforce a DNA loop by "wrestling" DNA. Furthermore, the multiscale method permitted the study of all the possible loops that the protein can induce in DNA, revealing a preference of LacI to form a so-called antiparallel loop. Such a loop has the same supercoiling as the cell's genome, facilitating its formation, and the forces stemming from this loop further stabilize the conformation of the protein. The computationally modeled behavior has provided interpretation for the experimental data from our collaborator [90] and is in good agreement with other existing data [91], complementing it with unprecedented atomistic detail of the process of gene repression.

BTA UNIT: C

TITLE: Transport Mechanism of Lactose Permease

KEYWORDS: co-transporter, membrane proten, conformational changes, salt bridge, proton gradient

INVEST1: Ying Yin
DEGREE1: B.S
DEPT1: Physics
NONHOST1:

INVEST2: Morten Jensen
DEGREE2: Ph.D
DEPT2: Biophysics
NONHOST2: Roskilde University, Denmark

INVEST3: Emad Tajkhorshid
DEGREE3: Ph.D
DEPT3: Biophysics
NONHOST3:

% BTA $: BTA %

ABSTRACT: Lactose permease (LacY) is an integral membrane protein (`http://www.ks.uiuc.edu/Research` that uses the cell's electrochemical proton gradient to actively transport substrates across the cell membrane [92–94]. These proteins play a critical role in transmembrane traffic, and, therefore, are critical for a healthy metabolism of a wide range of living organisms, including humans. Malfunction of these transporters is associated with various pathophysiological conditions, such as diabetes and depression [95,96]. After the first functional characterization of LacY in 1956 [97], biochemical, biophysical, and structural biological studies have addressed the transport mechanism of this enzyme [96,98,99]. These studies resulted in a transport model that involves two main conformational states of the transporter protein: an *outward open* state in which the substrate is accessible only from the periplasmic side, and an *inward open* state in which the periplasmic entrance is closed, but the cytoplasmic half channel is open, thus, allowing the substrate to diffuse into the cell. The crystal structure of substrate-bound LacY from *E. coli*, solved by our collaborator at 3.5 Å resolution, captured LacY in its *inward open* state [95]. In this structure, the 12 transmembrane helices of LacY form a hydrophilic cavity in the middle, where the substrate is bound to its binding pocket; the periplasmic side is closed and the

substrate is ready to diffuse into the cell through the opening of the cavity toward the cytoplasmic side of the membrane.

Previous molecular dynamics simulations [100] of the Resource using NAMD2 [28] on systems of LacY embedded in a fully hydrated lipid bilayer demonstrated that a critical buried salt bridge between Glu269 and Arg144 keeps the cytoplasmic half-channel open. After protonation of Glu269, this salt bridge breaks and Arg144 moves toward the surface of the protein, followed by the departure of water molecules from the interdomain space, which then initiates the closing of the cytoplasmic half-channel by allowing the hydrophobic surfaces of the two domains to approach each other. Recently, the molecular and energetic details of the lactose transport across LacY have been probed [101] through steered molecular dynamics simulations [89] by the Resource. Lactose was found to induce a widening of the narrowest parts of the channel during permeation, the widening being largest within the periplasmic half-channel. During permeation, the water-filled pore of LacY only partially hydrated lactose, forcing lactose to interact with channel lining residues. Lactose was seen to form a multitude of direct sugar-channel hydrogen bonds, predominantly with residues of the flexible N-domain, which is known to contribute the major part of LacY's affinity for lactose. Hydrophobic interactions arose predominantly between lactose and hydrophobic residues lining the periplasmic half-channel. The major energy barrier against transport was found within a tight periplasmic half-channel where sugar hydration is minimal and protein-sugar interactions are maximal. Upon unbinding from the binding pocket, lactose underwent a rotation to permeate either half-channel with its long axis aligned parallel to the channel axis. The results hint at the probability of an alternative mechanism for transport, in which lactose permeates LacY without major opening of the periplasmic half-channel.

BTA UNIT: C

TITLE: Mechanosensitive Ion Channels

KEYWORDS: MscS, MscL, Membrane Proteins, Ion Channels, Mechanotransduction

INVEST1: Marcos Sotomayor
DEGREE1: M.S.
DEPT1: Physics
NONHOST1:

INVEST2: Eduardo Perozo
DEGREE2: Ph.D.
DEPT2: Institute of Molecular Pediatric Sciences
NONHOST2: University of Chicago

INVEST3: Valeria Vasquez
DEGREE3: B.S.
DEPT3: Institute of Molecular Pediatric Sciences
NONHOST3: University of Chicago

% BTA $: BTA %

ABSTRACT: The perception of sound and regulation of blood pressure or cell volume are archetypal examples of biological processes mediated by mechanosensitive (MS) ion channels. This class of membrane proteins serves as molecular switches, opening and closing in response to stress conveyed through other proteinaceous structures of the cell or the cellular membrane itself [102, 103]. In bacteria, MS channels are thought to act as safety valves preventing the cell from bursting upon osmotic shock [104, 105].

Crystal structures of two MS channels have been solved: the MS channel of large conductance (MscL) in closed form from *M. tuberculosis* [106] and the MS channel of small conductance (MscS) in a putative open form from *E. coli* [107]. Both channels are activated by mechanical stress in the cell membrane, providing a controlled response to the osmotic pressure of the environment.

Whether the MscS crystal structure represents the fully open conformation of the channel with a conductance of 1 nS (as determined through patch-clamp experiments) remains debated [107–114]. Moreover, two different closed conformations of MscS have been proposed [109, 110] and the voltage dependence of MscS activation and inactivation remains controversial [104, 111, 115–118]. How the channel

dynamically transitions between closed and open conformations is also undetermined [109–111, 117, 119, 120].

Three important studies exploring the transport and gating properties of MscS have already been carried out by Resource scientists and collaborators [109, 114, 118]. In the first study, that followed an approach used earlier for MscL [121–123], the dynamics leading to channel closure of MscS in its native environment with a lipid bilayer, water, and ions (224,000 atom system), was explored [109]. The second study employed a multiscale method to compute conduction properties of the available crystal structure of MscS [114]. The debated voltage dependence of MscS activation and inactivation lead Resource scientists and Resource collaborator E. Perozo, to carry out a third study [118] which also explored the chance of introducing systematic errors in the multiscale description [114]. In this study, further all-atom simulations of MscS and electrophysiological experiments in a clean genetic-background were performed. The simulations utilized biasing electrostatic fields and lasted more than 200 ns in total. The results were in agreement with previous simulations [114] and were compared to electrophysiological measurements that unequivocally determined the conductance of MscS and its slight selectivity for anions. In addition, the experiments revealed that only MscS inactivation is voltage-dependent [116, 118]. Finally, data obtained from electron paramagnetic resonance experiments performed by our collaborator, E. Perozo, is being used to refine a closed conformation of MscS that includes the previously missing N-terminal domain. The EPR-based refinement technique is driving implementation of capabilities for new interaction potentials in NAMD and new tools to handle EPR data in VMD.

BTA UNIT: C

TITLE: Assembly of High-Density Lipoproteins

KEYWORDS: apolipoproteins, Nanodisc, HDL, apo A-I

INVEST1: Amy Shih
DEGREE1: M.S.
DEPT1: Biophysics
NONHOST1:

INVEST2: Peter Freddolino
DEGREE2: B.S.
DEPT2: Biophysics
NONHOST2:

INVEST3: Anton Arkhipov
DEGREE3: M.S.
DEPT3: Physics
NONHOST3:

INVEST4: Stephen Sligar
DEGREE4: Ph.D.
DEPT4: Biochemistry
NONHOST4:

% BTA $: BTA %

ABSTRACT: High-density lipoproteins (HDL) (http://www.ks.uiuc.edu/Research/Lipoproteins/) are protein-lipid assemblies involved in the transport of cholesterol from peripheral tissues to the liver for degradation. HDL is often called "good cholesterol" due to its role in removing excess cholesterol from tissues and blood vessels. Lower than average levels of HDL have been implicated in an increased risk of coronary heart disease. The production, transformation, and degradation of HDL is regulated by the reverse cholesterol transport pathway. Apolipoprotein A-I (apo A-I), the primary protein component of HDL, initially forms lipid-free/poor HDL particles. The incorporation of cholesterol and lipids into lipid-free/poor HDL particles causes a structural change, forming discoidal lipoprotein particles. Continued efflux of cholesterol and lipids as well as the esterification of cholesterol results in the transformation of the discoidal particles into mature spherical particles, which transport the cholesterol to the liver [1].

Two X-ray crystal structures of lipid-free apo A-I have been determined [124, 125]; however, the structure of apo A-I bound to lipid, in either the discoidal or spherical HDL forms, remains unknown. Since natural HDL particles are heterogeneous in size and composition, it has been impossible to obtain consistent structural data on them [7]. However, reconstituted HDL (rHDL), in which purified (and often truncated) apo A-I is used to form HDL particles, can be made into homogeneous particles. Nanodiscs are an engineered rHDL mimic being developed by Resource collaborator S. Sligar (UIUC), which can be self-assembled using a precise set of optimized conditions to form discoidal protein-lipid particles with homogeneous size and composition [9]. The Resource utilizes these homogeneous and well-characterized nanodisc particles [10, 126] in our molecular dynamics studies [14, 127].

Because the assembly and structural transitions of nanodiscs and HDL occur on timescales longer than those accessible using all-atom MD, a coarse-grained molecular dynamics model has been developed and applied to study this system [14]. Recent simulations on timescales of 1-10 microseconds have revealed the full assembly path of nanodiscs starting from randomized lipid-protein-water mixtures, and provided final structures for discoidal HDL particles in agreement with small-angle X-ray scattering (SAXS) results [15,16]. In addition, simulations of nanodiscs in the presence of varying amounts of the detergent cholate illustrated the stages involved in HDL assembly and disassembly in the presence of detergent, an important step during reconstitution of HDL particles [17]; again, close agreement was found with SAXS experiments.

BTA UNIT: C

TITLE: Sequencing DNA with a nanopore device

KEYWORDS: nanopore, DNA sequencing, genotyping, human genome, force spectroscopy, silicon, silica

INVEST1: Aleksei Aksimentiev
DEGREE1: Ph.D.
DEPT1: Beckman Institute
NONHOST1:

INVEST2: Eduardo R. Cruz-Chu
DEGREE2: B.Sc.
DEPT2: Biophysics
NONHOST2:

INVEST3: Gregory Timp
DEGREE3: Ph.D.
DEPT3: Electrical and Computer Engineering
NONHOST3:

INVEST4: Jean-Pierre Leburton
DEGREE4: Ph.D.
DEPT4: Electrical and Computer Engineering
NONHOST4:

% BTA $: BTA %

ABSTRACT: A nanometer-size pore, a so-called nanopore, can be manufactured in thin inorganic membranes. The current most important application for nanopores is DNA sequencing: under the influence of an electric field, DNA translocates through the nanopore, producing electrical signals characteristic of the sequence and length of the DNA strand. Current synthetic nanopores can not reach single-base resolution yet; however, they are among the most promising technologies for cheap DNA sequencing [128]. The Resource has been working in close collaboration with electrical engineers (Gregory Timp and Jean-Pierre Leburton) to understand the physics of synthetic nanopores and improve their resolution. Atomic-scale modeling was carried out in four directions: (i) genotyping with synthetic nanopore; (ii) stretching/unzipping DNA hairpins with a synthetic nanopore; (iii) sensing DNA sequence

with a nanopore capacitor; (iv) modeling of ionic current through silica nanopores (`http://www.ks.uiuc.edu/Research/nanopore/`).

The Resource's collaborator Dr. Timp discovered that proteins called restriction enzymes, which bind selectively to certain DNA sequences, can be used in conjunction with nanopores to find mutations in DNA [129]. The protein-DNA complex is forced by an electric field through a nanopore designed to allow the passage of the DNA but not of the protein. The voltage at which the complex breaks up and DNA freely translocates through the nanopore depends on the sequence. It can be found therefore, using a relatively simple measurement of voltage threshold, whether the DNA sequence contains an error (mutation) in the segment of interest. The Resource has performed molecular dynamics simulations using the program NAMD of a system that consisted of the protein EcoRI (restriction enzyme), DNA, a silicon nitride nanopore, and solvent. The simulations reproduced accurately the experimentally obtained values of the voltage threshold, and allowed one to predict which errors in the DNA sequence most strongly affected the threshold.

By exploiting the peculiar mechanical properties of DNA, it is possible to design a pore geometry to precisely control the conformation of DNA within the nanopore in order to facilitate sequencing. Essential to the mechanics of DNA is the transformation of a DNA double helix into two single strands. As revealed by single molecule spectroscopy, the transition can proceed in two ways, referred to as unzipping and stretching, distinguished by the force required to induce the transition. The Resource has performed atomic-scale simulations to determine the influence of pore geometry on the helix-coil transition of DNA during its passage through the pore. Both modes of the helix-coil transition have been observed, where the pathway taken depended on the geometry of the pore. Understanding the atomic details of the helix-coil transition will allow one to precisely engineer synthetic nanopores for a desired effect on DNA.

Movement of ssDNA driven by an oscillating electric field through a nanopore capacitor fabricated from silicon nitride membrane was simulated using the molecular dynamics program NAMD. It was found that the electrostatic potential of different nucleotides of ssDNA sensed by the capacitor depends on the type of the nucleotide (A, C, G, or T) [130, 131] and movement direction. It was also found that the translocation velocity of ssDNA depends on DNA composition and the movement direction. The preliminary results suggest that the discrimination of nucleotides by their electrostatic potential and velocity of translocation through a nanopore is feasible. Further molecular dynamics studies, in particular simulation of realistic non-uniform DNA sequences, are required to develop a robust measurement and data processing methodology of DNA sequencing based on the above effects. When

implemented in a nanopore sequencing device, this methodology promises to become an affordable high-throughput alternative to the existing sequencing techniques.

The most promising application of solid-state nanopores is to develop an inexpensive, high-throughput genome sequencing technique. However, nanopore measurements are not yet sensitive enough to resolve DNA with single-base precision. A major hurdle to achieve single-base resolution is to relate the measured signal to a particular atomic level event. The Resource has recently developed an amorphous silica model [132] that accurately reproduces atomic-scale roughness of the surface and its wetting properties. Using this novel model, the Resource has performed a systematic study of the ionic conductance through silica nanopores. Macroscopic measurable quantities, namely the ionic current, have been related to molecular details of the nanopore surface.

BTA UNIT: T

TITLE: Empirical Nanotube Model for Biological Applications

KEYWORDS: carbon nanotubes, polarizable model, biosensor, nanotechnology, DNA-CNT complex

INVEST1: Ying Yin
DEGREE1: B.S.
DEPT1: Physics
NONHOST1:

INVEST2: Deyu Lu
DEGREE2: Ph.D
DEPT2: Physics
NONHOST2:

INVEST3: Aruna Arajan
DEGREE3: M.Sc.
DEPT3: Physics
NONHOST3:

INVEST4: Michael S. Strano
DEGREE4: Ph.D.
DEPT4: Chemical and Biomolecular Engineering and Beckman Institute
NONHOST4:

% BTA $: BTA %

ABSTRACT: Carbon nanotubes (CNTs) have shown an extraordinary potential for bionanoengineering applications. A promising application involves SWNT-based optical biosensors. Compared to conventional optical biosensors, carbon nanotubes have a great advantage in that they fluoresce at near-infrared wavelengths [133], where human tissues and biological fluids are nearly transparent [134–137]. Design of SWNT based sensors requires an adequate description of the interactions between nanotubes and their biological environment.

To characterize the interactions between SWNTs and biological environment in MD simulations, it is crucial to consider the polarizable nature of SWNTs due to their highly delocalized pi-electrons. For this purpose, the Resource has developed an electrostatic force field for SWNTs [138–141] that has been reviewed in [142]. The

new force field divides the atomic partial charges for SWNTs into a fixed component and a fluctuating component. The fixed charges are parameterized based on density functional theory calculations to account for the initial charge distribution of isolated SWNT segments [140]. The effect of the environment, *i.e.*, polarization of the SWNTs, is captured by induced charges (the fluctuating component), which are computed from a tight binding Hamiltonian [138–140]. This polarizable force field properly describes the dielectric response of CNTs to the surrounding environment within an MD simulation, but requires on-the-fly quantum chemistry calculations [141]. The scheme was applied to study the oscillation of a potassium ion inside a CNT [141], and the results were tested successfully against more accurate first-principles MD simulations (unpublished results, collaboration with G. Galli). This scheme has also been implemented in NAMD-lite, the Resource's MD code for method prototyping. Using the dynamic force field [141] described above, a SWNT was immersed in a box of water and simulated for 2 ns using NAMD2 [28]. The effect of the dielectric properties of the SWNT on water transport was probed by switching the tight-binding module on and off. The screening effect due to the polarization of the nanotube wall was found to significantly lower the overall dipole moment inside the nanotube, a noticeable response that further demonstrates the value of the tight-binding model.

BTA UNIT: C

TITLE: Protecting the Cell Nucleus

KEYWORDS: nuclear pore complex, nucleoporin, importin-beta, FG-repeat, NTF2, cse1p, nucleus

INVEST1: Timothy Isgro
DEGREE1: M.S.
DEPT1: Physics
NONHOST1:

INVEST2: Ueli Aebi
DEGREE2: Ph.D.
DEPT2: M.E. Müller Institute for Structural Biology, Biozentrum
NONHOST2: University of Basel

% BTA $: BTA %

ABSTRACT: The nucleus of the cell is of central importance to an organism. It serves to store and organize genetic material, while separating and protecting this very important information from the host of other cellular components. While the nucleus requires this protective isolation, it also needs to communicate with the rest of the cell, exchanging proteins and RNA, for a variety of nuclear and cytoplasmic processes which act in concert. The nuclear pore complex (NPC) is the gatekeeper of the nucleus. It is an immense assembly of proteins, perhaps the largest protein structure in eukaryotic cells [143], embedded in the nuclear envelope. Because of its large size, however, pointed experimental study has been difficult, and as a result, the mechanism by which the NPC selectively allows "good" material across the nuclear envelope, while preventing the transit of the "bad", remains unknown. It is known, however, that in order to cross the nuclear envelope, a large molecule must first associate with a transport receptor protein (reviewed in [143–148]). It is hypothesized that so-called "FG-repeat" proteins in the NPC recognize the transport receptor and allow the complex to pass. Understanding precisely how this recognition occurs is vital to determining how the NPC protects the nucleus.

In order to shed light on the gating mechanism of the NPC, the Resource performed two sets of intesive molecular dynamics simulations on complexes formed by the transport receptors NTF2 [149] and Cse1p [150]. The work follows our highly successful simulations on the transport receptor importin-beta [151] (http://www.ks.uiuc.edu/Research/npc/). NAMD [28] was used to perform the simulations on both protein complexes, each in the presence of a solution of nuclear

pore proteins whose concentration ranged from 50-90 mM. Complete systems averaged roughly 60,000 atoms in the case of NTF2 and 250,000 atoms for Cse1p, with total simulation times of 250 ns and 160 ns, respectively. Simulating the dynamics of each atom of these large biological complexes for such long amounts of time was necessary to characterize the binding of FG-repeats accurately. This posed a significant challenge, requiring extremely efficient computing on up to 256 processors on some of the best supercomputers in the world. In addition, sequences of each transport receptor were aligned, and conserved residues on their surfaces were thus determined and visualized using VMD [42]. The simulations were powerful in reproducing previous experimental evidence for FG-repeat binding spots on the surfaces of the transport receptors. Furthermore, the simulations suggest a wide range of new binding spots for each transport receptor. While these results are significant in their own right, the results for the three transport receptors, when taken as a whole, reveal a broad and striking binding pattern for FG-repeats and transport receptors that may suggest how the NPC distinguishes transport receptors from other, inert proteins. By examining the binding spots on the surfaces of importin-beta, NTF2, and Cse1p, and contrasting them with those on the surface of Kap60p, which is inert, it appears that the mere possession of FG-repeat binding spots on the surface is not a sufficient condition for viable nuclear transport. Rather, a dense array of FG-repeat binding spots, spaced appropriately, around 14 Angstroms, is necessary for a transport receptor to be recognized as such by the NPC. This implies a groundbreaking, new hypothesis that the binding of multiple FG-repeats to the transport receptor is necessary for viable nuclear transport and must be a key feature for any future, complete nuclear transport model.

BTA UNIT: C

TITLE: Simulating a Bacterial Organelle: the Photosynthetic Chromatophore

KEYWORDS: photosynthesis, atomic force microscopy, purple bacteria, bioenergetics, energy transfer, quantum biology, light harvesting complex, reaction center, Rhodobacter sphaeroides, spherical membrane, membrane curvature, bacteriochlorophyll, PSU, LH1, LH2, BchI, AFM

INVEST1: Danielle Chandler
DEGREE1: B.Sc.
DEPT1: Physics
NONHOST1:

INVEST2: James Gumbart
DEGREE2: M.S.
DEPT2: Physics
NONHOST2:

INVEST3: Jen Hsin
DEGREE3: B.Sc.
DEPT3: Physics
NONHOST3:

INVEST4: Melih K. Sener
DEGREE4: Ph.D.
DEPT4: Physiology and Biophysics
NONHOST4: Cornell University

INVEST5: C. Neil Hunter
DEGREE5: Ph.D.
DEPT5: Molecular Biology and Biotechnology
NONHOST5: University of Sheffeld

INVEST6: John D. Olsen
DEGREE6: Ph.D.
DEPT6: Molecular Biology and Biotechnology
NONHOST6: University of Sheffeld

% BTA $: BTA %

ABSTRACT: Photosynthesis is the bioenergetic process by which the majority of Earth's biosphere derives energy from sunlight. This efficient sub-cellular process depends on the complex organization of the photosynthetic unit (PSU) at multiple scales spanning from sub-nanoscale atomic processes to macroscale multi-protein and lipid assemblies that are integral components of PSU organelles. One such organelle found in the purple bacterium *Rhodobacter sphaeroides* is the chromatophore, a 70 nm wide bulbous invagination of the inner membrane containing a total of approximately 200 proteins, 5,000 chlorophylls, and 1,700 carotenoids that permit efficient photosynthesis. These 200 chromatophore proteins consist of several multi-protein complexes, including 20 photosynthetic reaction centers (RC) [152, 153], 20 light harvesting complexes 1 (LH1) [154, 155], 150 light harvesting complexes 2 (LH2) [156, 157], five bc1 complexes (bc1) [158] and cytochrome c2s [159], and usually one ATP synthase [160]. Involving 70 million atoms, continued development of coarse-graining tools and the extension of VMD and NAMD capabilities will be necessary to accommodate such a large system computationally. Representing one of the first simulations at the near-cellular level, the chromatophore system will also strenuously test the scalability of VMD and NAMD while providing unique opportunities to investigate one of the most fundamental processes necessary for life on Earth, namely the conversion of light into bioenergy.

Modeling the collective behavior of a PSU requires knowledge of the relative stoichiometry and spatial distribution of its photosynthetic complexes. Collaborators of the Resource have completed atomic force microscopy (AFM) studies of the intracytoplasmic membrane in the purple bacterium that reveal the overall two-dimensional spatial organization of the chromatophore. These AFM data were then used by Resource scientists to construct three-dimensional models of an entire membrane vesicle at the atomic level, using the known structure of the LH2 complex and a structural model of the dimeric RC-LH1 complex. The *in silico* reconstructions permit a detailed description of light absorption and electronic excitation migration, including computation of a 50 ps excitation lifetime and a 95% quantum efficiency [161].

Separate full-atom models of LH1, LH2, and the bc1 complex embedded in membrane patches are currently in development at the Resource. While the structure of the bc1 complex is established and construction of the LH2 model benefits from the availability of similar structures that may be utilized as templates in model development, construction of LH1 utilizes EM data from experimental collaborators. These independent models will investigate the contribution of protein shape to membrane curvature in efforts to understand how the chromatophore membrane initiates and maintains its bulbous curvature and will ultimately assist in correct placement of proteins into the chromatophore sphere. Subsequent lipid placement

and solvation will result in a 90 cubic-nanometer volume system, providing the first functional model of a bacterial organelle for computational investigation at atomic resolution.

BTA UNIT: C

TITLE: Voltage-gating mechanism of Kv1.2

KEYWORDS: K channel, Shaker, voltage-gating, ion channels, membrane proteins

INVEST1: Fatemeh Khalili-Araghi
DEGREE1: B.Sc.
DEPT1: Physics
NONHOST1:

INVEST2: Emad Tajkhorshid
DEGREE2: Ph.D.
DEPT2: Biochemistry
NONHOST2:

INVEST3: Vladimir Yarov-Yarovoy
DEGREE3: Ph.D.
DEPT3: Pharmacology
NONHOST3: University of Washington

INVEST4: Benoit Roux
DEGREE4: Ph.D.
DEPT4: Molecular Pediatric Sciences
NONHOST4: University of Chicago

% BTA $: BTA %

ABSTRACT: Voltage-gated potassium channels (http://www.ks.uiuc.edu/Research/kvchannel) are integral membrane proteins present in all three domains of life. In a specialized class of animal cells, known as excitable cells — including neurons, muscle cells, and endocrine cells — Kv channels work with other cation channels (sodium and calcium channels) to regulate the electrical activity and signaling of the cell [162]. Kv channels activate (open and close) in response to changes in the electrical potential across the cell membrane allowing passive and selective conduction of K+ ions through the channel. Potassium conduction is directed by the electrochemical gradient across the membrane and can achieve very high rates, while still discriminating against all other cations (including the smaller Na+ ions) [162]. In addition to electrical signaling in nervous systems, Kv channels play an important role in the regulation of cardiac excitability and regulation of insulin release. In humans,

malfunction of these channels can result in neurological or cardiovascular diseases such as long QT syndrome or episodic ataxia [163].

The crystal structure of Kv1.2 [164], a member of the Shaker K+ channel family, has provided the first view of the molecular architecture of a mammalian potassium channel in a putative open state at 3.9 Angstrom resolution. In addition to the ion conduction pore, four voltage sensor domains are also partially identified in the structure. The voltage sensors contain several charged residues that respond to the changes in the electric field [165–168],and, thus, control opening and closing of the channel.

To study the gating mechanism of Kv1.2, the Resource's collaborator (Dr. Yarov) built complete models of the Kv1.2 channel in the open state, in which the missing domains of the voltage sensors were modeled according to a database of structurally known membrane proteins. Several possible candidates for the closed state of the channel were also determined using the structure prediction program ROSETTA [169], which was recently adapted to work with membrane proteins [170].

The Resource performed molecular dynamics simulations using the program NAMD [28] to further refine the structures of the closed and open states in the presence of an electric field enforcing spatial restraints obtained from experimental studies (provided by Dr. Roux). The simulated system, consisting of the protein in a patch of DPPC lipid bilayer in a 200mM KCl solution, includes 340,000 atoms and was simulated for about 28 ns. The Resource also implemented an efficient version of the replica exchange method [171] in NAMD, combining temperature as well as voltage tempering, which will be used to further refine the modeled structures. Once the closed and open conformations of the Kv1.2 are obtained, all-atom MD simulations will be performed to determine how the electrical potential can drive the gating of the channel.

In another set of simulations, the voltage-induced conduction of K+ ions through the Kv1.2 channel was studied. The pore domain of Kv1.2 (not including the voltage sensors) was inserted into a lipid bilayer and simulated with and without a voltage bias of 1 V for about 65 ns. Several conduction events, describing a detailed dynamical picture of ion permeation, were recorded [172]. As suggested earlier, the ion movement in the channel takes place in a concerted way, involving 2–3 ions residing mainly at sites identified previously in the crystallographic studies [173] and molecular dynamics simulations [174, 175]. The simulations revealed, however, the jump of ions between these sites and identified the sequence of multi-ion configurations involved in permeation. In addition to the previously suggested mechanism of K+ conduction, other possible scenarios for the ion conduction were also observed, which could be important at higher concentrations of K+.

BTA UNIT: C

TITLE: Multiscale Modeling of a Bacterial Ribosome

KEYWORDS: multiscale, translation, ribosome, RNA, molecular dynamics, cryo-electron mycroscopy

INVEST1: Leonardo Trabuco
DEGREE1: B.S.
DEPT1: Biophysics
NONHOST1:

INVEST2: Elizabeth Villa
DEGREE2: B.S.
DEPT2: Biophysics
NONHOST2:

INVEST3: Joachim Frank
DEGREE3: Ph.D.
DEPT3: Howard Hughes Medical Institute
NONHOST3: Wadsworth Center, NY

% BTA $: BTA %

ABSTRACT: The ribosome [176] is a cellular machine that synthesizes proteins based on genetic instructions. The ribosome moves along the mRNA, catches tRNAs, facilitates the pairing between codons and anticodons, and catalyzes the formation of peptide bonds between amino acids. The bacterial ribosome is an important target of antibiotics; indeed, 50% of all research on antibiotics is focused on the ribosome.

Currently the most successful approaches to image ribosomes are cryo-electron microscopy (cryo-EM) [177] and X-ray crystallography [178]. Cryo-EM offers insights into the function of the ribosome by providing snapshots of different functional states, currently at a resolution of 7 Angstroms. X-ray crystallography yields atomic-scale structural information [178]. These and other experiments show that the ribosome consists of two subunits, the small subunit being responsible for codon-anticodon recognition, and the large subunit for catalyzing peptide bond formation. The whole translation machinery consists of ribosomal RNAs, about 50 ribosomal proteins, tRNAs, mRNA, ions, and additional protein factors.

The Resource has developed an all-atom model of the *Thermus thermophilus* ribosome based on a recent 2.8-Angstrom resolution crystal structure [179] required

completion of missing protein regions; modeling of the L7/12 stalk – a large flexible region – based on different experimentally-determined structures and homology modeling, modeling of the mRNA by combining low-resolution crystallographic data with an analytical model of the polymer chain; iterative ion placement at electrostatic potential minima; and multistep solvation of the system. The model contains 2.8 million atoms and was simulated for 10 ns by means of molecular dynamics (MD) with NAMD2. The average electrostatic potential was calculated from the last 3 ns of this simulation using VMD running on graphics processing units (GPUs), which greatly accelerates the calculations [46]. The average electrostatic potential was incorporated into multiscale simulations where particular regions of the ribosome of around 200,000 atoms are described by all-atom MD, whereas the rest of the system is represented by a mean-field electrostatic potential, allowing for longer times to be reached. These multiscale simulations are underway and will help to address important questions raised by cryo-EM data recently obtained by our collaborator Dr. Frank.

In order to bridge the resolution gap between X-ray crystallography and cryo-EM and obtain "quasi-atomic" models of the ribosome in different functional states, the Resource has developed a new method to perform flexible-fitting of atomic structures into density maps. The method combines an MD simulation with a mass-weighted biasing potential derived from the electron density map. The method has been successfully applied to fit crystal structures into cryo-EM maps of the ribosome from Dr. Frank's lab, rendering "quasi-atomic" models that aid in the interpretation of the experimental results. These models will also be used as starting structures of simulations of the ribosome in different functional states.

BTA UNIT: C

TITLE: The Protein-Conducting Channel

KEYWORDS: translocon, SecY, translocation, protein channel

INVEST1: James Gumbart
DEGREE1: M.S.
DEPT1: Physics
NONHOST1:

INVEST2: Tom Rapoport
DEGREE2: Ph.D.
DEPT2: Cell Biology
NONHOST2: Harvard Univ.

% BTA $: BTA %

ABSTRACT: The protein-conducting channel, more specifically known as the translocon (`http://www.ks.uiuc.edu/Research/translocon/`) or Sec complex, is an evolutionarily ancient protein complex that helps proteins cross or integrate into membranes (depending on whether they are soluble or membrane proteins). Present in all branches of life, the Sec complex is found in the cytoplasmic membrane in bacteria and archaea and in the membrane of the endoplasmic reticulum in eukaryotes. A passive channel, the Sec complex partners with other proteins that drive translocation of an unfolded polypeptide through the channel. In co-translational translocation, a common mode of translocation, this partner is the ribosome which feeds the nascent protein through the channel as it is synthesized. As a key step in protein targeting, translocation can be a deciding factor in the fate of proteins and even the cell as a whole. For example, poor recognition of the prion protein (PrP) leads to its abnormal aggregation and ultimately to lethal levels in the cell [180]. However, being able to enhance recognition and passage across the membrane could increase yields for artificially created proteins such as insulin [181]. In 2004, the Resource's collaborator, Tom Rapoport, released the first high resolution structure of the translocon. Obtained from *Methanococcus jannaschii*, this heterotrimeric membrane protein complex was resolved to 3.5 Angstroms. Based on this structure, specific details of translocation began to emerge. Observed structural elements were proposed to have specific functions, such as a constrictive pore ring and a plug blocking the exit of the channel. It was also proposed that a singular monomer within a dimeric or tetrameric complex serves as the active channel, leaving the

role of oligomerization in question. Two dimeric forms of the channel with different functional behavior have been proposed (a 'back-to-back' and a 'front-to-front' dimer) although which is the *in vivo* state is unknown.

Previous investigations targeted the translocation function of the channel, i.e., moving proteins across the channel [182]. More recently, Resource scientists investigated the ability of the channel to open laterally to the membrane by means of a 'lateral gate', proposed for insertion of membrane proteins [183]. Steered molecular dynamics was used to open the lateral gate in three modeled structures of the channel to determine the relevance of different elements to lateral gating. One surprising result indicated SecE, an accessory protein associated with the main body of the channel, does not function as a clamp as originally presumed. It was also found that only an intermediate amount of gate opening was required to free the channel-blocking plug from its position in the center of the channel. This interaction between the gate and the plug is one possible step in signal sequence gating of the channel. Simulations of the channel with an open gate in both all-atom and a recently developed course-grained representation on timescales of up to one microsecond [14] revealed that lipids do not flood the open channel. Taken together, these results support the back-to-back dimer model over the front-to-front one.

Current efforts include the utilization of experimental ion conductance data through both the native channel and two selected mutants. Simulations reproducing the relative differences in conductance will provide an atomic-scale model of the molecular mechanisms underlying them. Preliminary results in qualitative agreement with experiment have already provided initial validation of the simulations. Understanding these mechanisms will permit the development of a more comprehensive model describing the sequence of events involved in gating the channel. This functional model will then be utilized in efforts to artificially modulate gating of the Sec complex and regulate protein transport across cell membranes.

BTA UNIT: C

TITLE: Structural Dynamics of Viruses

KEYWORDS: Virus, Viral assembly, Viral entry, STMV, Poliovirus, Protein aggregates, Large simulations, Coarse-grained simulations

INVEST1: Peter Freddolino
DEGREE1: B.S.
DEPT1: Center for Biophysics and Computational Biology
NONHOST1:

INVEST2: Anton Arkhipov
DEGREE2: M.S.
DEPT2: Physics
NONHOST2:

INVEST3: Alexander McPherson
DEGREE3: Ph.D.
DEPT3: Molecular Biology and Biochemistry
NONHOST3: University of California, Irvine

INVEST4: Xiaowei Zhuang
DEGREE4: Ph.D.
DEPT4: Chemistry
NONHOST4: Harvard University and Howard Hughes Medical Institute

INVEST5: James Hogle
DEGREE5: Ph.D.
DEPT5: Biological Chemistry and Molecular Pharmacology
NONHOST5: Harvard Medical School

INVEST6: David Belnap
DEGREE6: Ph.D.
DEPT6: Chemistry and Biochemistry
NONHOST6: Brigham Young University

INVEST7: Taekjip Ha
DEGREE7: Ph.D.
DEPT7: Physics

NONHOST7: University of Illinois at Urbana-Champaign and Howard Hughes Medical Institute

% BTA $: BTA %

ABSTRACT: Viruses are parasitic organisms responsible for many diseases. A virus consists of a genome enclosed in a protein coat (capsid), often with a few other accessory molecules. Viral infection depends on the ability of a virion to maintain its stability, disassemble, or assemble depending on the infection stage; these steps are determined by interactions between the viral components, as well as by external factors. Thus, studying the dynamics of viral components is crucial for elucidation of the mechanisms of viral infections. Although a large number of viruses have been structurally resolved (see, e.g., [184]), a static structure does not reveal dynamical properties; also, the spatial and temporal resolution of real-time experiments (e.g. [185, 186]) is usually much lower than that required to distinguish the movements of individual viral proteins. In this situation, MD studies are perfectly suited to complement experiments, as MD simulations are now capable of describing the dynamics of complete viruses at atomic resolution [187] (`http://www.ks.uiuc.edu/Research/STMV`).

The first all-atom simulation of a complete virus, the Satellite Tobacco Mosaic Virus (STMV), has been reported [187] by the Resource (using NAMD [28]) in collaboration with A. McPherson. Based on the simulation, a possible assembly pathway for the virus has been proposed. However, all-atom simulations are currently limited to about 10 nm size and 100 ns duration, while the size of most viruses is much larger and structural transitions of interest are slower. To reach desired length and time scales, the Resource has developed a new coarse-grained (CG) technique for simulations of macromolecular assemblies [64,65], the shape-based CG method, which reproduces the shape of a macromolecule using several point masses distributed over the all-atom structure. The method has been applied to simulate several viral capsids on a microsecond time scale [64]. The results of the CG simulations agreed with experimental observations, where possible, and with all-atom MD simulations [187]. Among the studied capsids, some were found stable while others collapsed rapidly when simulated without the corresponding viral genome, implying different self-assembly routes for different viruses.

In collaboration with J. Hogle, X. Zhuang, and D. Belnap the Resource is now investigating the mechanism of poliovirus infection. The infection starts when the virus binds to protein receptors on the surface of the cell, inducing a structural transition in the capsid, as well as formation of a hole in the membrane; as a result, viral RNA can penetrate into the cell (see, e.g., [188]). The mechanism of this process is unknown. It is also unclear how many bound receptors are needed to start the transition. Using low-resolution cryo-EM structures of the virus-receptor

complex, an X-ray structure of the capsid, and a homology model of the receptor, an all-atom structure of the poliovirus-receptor complex has been built employing the tools developed by the Resource. All-atom simulations of a part of the complex (with a single receptor) have been started, to be followed by simulations of 1/3 of the poliovirus capsid with five, one, or no receptors bound, and by CG simulations of the whole virus with various number of receptors.

Ongoing efforts also focus on the disassembly of STMV, which is known to be mediated by a change in pH (as with many other viruses), although the exact mechanism is unclear. The Resource researchers are performing simulations of STMV at pH 4, 7, and 10; in doing so they determine the protonation state for each titratable group [189] at the beginning, but also frequently during the simulation. In addition, Resource researchers are performing FRET experiments (in Professor Ha's laboratory) on STMV at various pH levels to track the structural changes in the virus and to compare experimental results with computational data.

Resource Summary

BTA unit: (T)

NUMBER PUBLISHED -

Books: **??** Papers: **??** Abstracts: **??**

NUMBER IN PRESS -

Books: **??** Papers: **??** Abstracts: **??**

PUBLISHED:

Books:

Papers:

Abstracts:

IN PRESS:

Books:

Papers:

Abstracts:

BTA unit: (C)

NUMBER PUBLISHED -

Books: **??** Papers: **??** Abstracts: **??**

NUMBER IN PRESS -

Books: **??** Papers: **??** Abstracts: **??**

PUBLISHED:

Books:

Papers:

Abstracts:

IN PRESS:

Books:

Papers:

Abstracts:

BTA unit: (S)

NUMBER PUBLISHED -

Books: **0** Papers: **0** Abstracts: **0**

NUMBER IN PRESS -

Books: **0** Papers: **0** Abstracts: **0**

PUBLISHED:

Books:

None.

Papers:

None.

Abstracts:

None.

IN PRESS:

Books:

None.

Papers:

None.

Abstracts:

Software Releases (2004-2005)

- VMD: 1.8.3 released February 2005
- NAMD: 2.5 in beta; 2.6 expected in June 2005
- BioCoRE: Incremental updates every few weeks*

*URL:`http://www.ks.uiuc.edu/Research/biocore/announce/changeLog.shtml`

- MDTools: released 4 software packages since last report, and updated another 7. New releases:
  - Mail::SpamAssassin::UIUC v3.0.2+tcb3 - spam filtering package (released March 2005)
  - TCB::Backup v0.05 - data backup package (released September 2004; last updated December 2004)
  - TCB::RSS v0.51 - website update announcements (released March 2005)
  - TCB::Webdav v2.04 - network file system mounter (released December 2004; last updated February 2005)

  Updated packages:
  - CGI::SHTML v1.32 (last updated August 2004)
  - DBIx::Frame v1.06 (last updated May 2004)
  - TCB::AddUser v1.06 (last updated June 2004)
  - TCB::Internal v1.04 (last updated August 2004)
  - TCB::Publications v0.99.07 (last updated December 2004)
  - TCB::Seminar 0.99.01 (last updated July 2004)
  - TCB::System 0.99.02 (last updated July 2004)

BTA unit: (D)

NUMBER PUBLISHED -

Books: **??** Papers: **??** Abstracts: **??**

NUMBER IN PRESS -

Books: **??** Papers: **??** Abstracts: **??**

PUBLISHED:

Books:

Papers:

Abstracts:

IN PRESS:

Books:

Papers:

Abstracts:

The Resource advisory committee met last on May 8, 2006. In the months following that meeting, the Resource prepared and submitted its 5-year renewal application, with the required site visit in April 2007. The Resource is working to set the next advisory committee meeting in late July 2007, before the start of the next 5-year funding period.

The May 2006 Advisory Board membership and report is provided below:

- Dr. Angel Garcia, Senior Constellation Chaired Professor in Biocomputation and Bioinformatics, Rensselaer Polytechnic Institute (Chair)
- Dr. Angela Gronenborn, Program Director, Department of Structural Biology, Molecular Biophysics and Structural Biology Graduate Program, University of Pittsburgh School of Medicine
- Dr. Jeffrey Skolnick, Director, Center for the Study of Systems Biology, Georgia Institute of Technology
- Dr. Thom Dunning, Director, National Center for Supercomputing Applications
- Dr. Michael Heath, Director of Computational Science and Engineering, University of Illinois at Urbana-Champaign
- Dr. Benoit Roux, Professor, Pediatrics, Biochemistry and Molecular Biology, University of Chicago

Advisory Board Report, May 8, 2006

NIH Resource FOR Macromolecular Modeling AND Bioinformatics

Summary

The panel is very well impressed by the accomplishments of the Resource. The scientific accomplishments of the faculty and the software developed as part of this resource are state of art. The resource has had the vision to develop tools that have become essential to answer scientific questions. This Resource provides an enabling technology for biomolecular modeling, since software of this quality and capabilities is not available elsewhere. The further development of the tools will be needed to satisfy the community demands and scientific needs of the future.

This project has demanded the use of large computer clusters. Currently the capacity to house the resource clusters is limited. We consider that it is imperative that a suitable computer room is completed before an NIH site visit to increase the likelihood of the successful renewal of this valuable Resource.

In what follows we provide brief descriptions of the various components of this Resource. The resource consists of three main components: NAMD, VMD, and a Modeling suite.

1. NAMD

NAMD is the dominant highly parallel program available today for classical molecular dynamics simulations of large biomolecular systems. NAMD, the Gordon-Bell Award winning program, has a world-wide user base of 15,000 with 3,000 new users added each year. It is heavily used by scientists at NSF's supercomputing centers (in the last several months it consumed 70% of the cycles at PSC and a substantial fraction of the resources at NCSA).

The ability of NAMD to scale to small (1-16 cpus), medium (128 cpus), and large scale clusters (1000s of cpu) have been demonstrated by applications from the user community. Users have been able to take released versions of NAMD and execute it efficiently in the largest computers available, a capability that resulted from a very productive collaboration with computer science faculty at UIUC. This program has enabled the simulation of million atom molecular machines. NAMD provides its users with a complete feature-set that runs on almost every platform available today with excellent parallel scaling.

NAMD developers respond to challenges presented by the Resource's collaborative research by developing general-purpose solutions to address a broad range of current and future needs. New simulation features can be implemented into NAMD by others using Tcl/python. Planned new features will allow NAMD users to employ coarse-grained models, polarizable force fields, implicit solvent models, replica-based methods, and a wider variety of user-written methods in their simulations.

Continuous development of NAMD is required to respond to changing scientific challenges, computer environments, and users demands. For example, the Resource staff are exploring the implementation of NAMD on emerging technologies, specifically multi-core CPUs (the future of all computing systems), graphics processors, the Cell processor, and FPGAs (reconfigurable computing). NAMD already runs on BlueGene/L with excellent scaling and performance, but changes are needed to run million-atom simulations on this unique platform.

2. VMD

VMD has gained over 10,000 new users of VMD 1.8.3 during the past funding period for a total of 17,354 registrants of that version. In the three weeks since the release of VMD 1.8.4 on April 16, 2006, 2,159 users have registered and downloaded the new version. Thus VMD has a very impressive user community. VMD provides both common and specialty structure analysis features, and will be improved by ongoing development of new graphical representations for nucleic acids, carbohydrates, silicon structures involved in bio-nanotechnology, and coarse-grain structures and simulation trajectories. Especially

noteworthy is MULTISEQ that provides a very powerful suite of sequence/structure analysis tools that is extremely useful for biological analysis.

The continued success of VMD depends on its compatibility with key data types and file formats, and interoperability with other software tools. The extensibility of VMD in the form of embedded Tcl/Python scripting languages and plug-in interfaces allows researchers to adapt VMD to new types of data, to interoperate with other research software, and to develop their own research tools as VMD extensions. The Resource has leveraged this extensibility to develop 55 VMD extensions for visualization, analysis, and modeling (36 with graphical interfaces), and to provide support for over 51 molecular file formats. VMD will continue to act as a platform for the development of new research tools, implementing key computational and visualization capabilities needed by tool developers.

The rapidly growing number of structures and their increased sizes challenge VMD in terms of data capacity, computational speed, and interactive display performance. It is (will be) possible to analyze thousands of structures (the entire pdb) thus VMD can be used as a structure mining engine. The use of out-of-core data access and compression algorithms, multi-core processors, programmable shading and offloading of computations to the graphics processing unit will allow VMD to meet these demands. Porting of VMD to Windows XP x64 and the future 64-bit MacOS X will extend support for extremely large datasets to the most commonly available personal computing systems. Plans to exploit the tremendous capabilities of current and future graphics processing units promise an order of magnitude speed-up in floating point processing.

3. Integrated Modeling Suite

Faced with the challenges from the ever-increasing complexity of molecular modeling projects, the process of carrying out computational research on biomolecular systems using the wide variety of available tools needs to be better organized. The Resource is poised to lead the shift towards integrating this large tool chest of techniques through the development of a new Biomolecular Modeling Suite. The Suite is a comprehensive set of tools necessary during the various stages of a modeling-simulation project: setup, production, and analysis, as well as management to allow researchers to tackle this new kind of complex research projects. By leveraging and extending the growing toolset available in VMD, as well as the collaborative features, logging, and file sharing capabilities of BioCoRE, the Resource will create a suite of tools with clear user-friendly interfaces, to enable the seamless execution of increasingly complex modeling projects in biomedical research.

4. Collaborations

The Resource is involved in various collaborations with research groups across the US

and in the world. The demand for tools and expertise of the Resource is high, and numerous research groups approach the Resource every year seeking collaborations. The main criteria for the Resource to engage in collaborations are: the project should be of high quality and pose an exciting biological problem, it has to be of high biomedical relevance, and, most importantly, the methodology required by the project should be challenging and driving technology development in the Resource.

Dr. Emad Tajkhorshid, Assistant Director of Research of the Resource, presented major collaborations of the Resource. The projects cover a wide range of biological and technological applications, all posing new challenges to technology development in all four major core projects, i.e., Structural Dynamics of Macromolecular Assemblies, VMD, NAMD, and the Modeling Suite.

List of presented collaborations was as follows:

- All-atom simulation of a complete virus (1 million atom system)
- Multiscale modeling of protein-DNA complexes
- Multiscale Modeling of Bacterial Ribosome
- Gating a mechanosensitive ion channel
- Stretching nanospring repeat proteins
- Mechanism of Voltage Gating of the Kv1.2 potassium channel
- Engineering hydrogenase as a source of H2 energy
- CG Assembly of Nanodiscs (discoidal HDL)
- Binding Dynamics of Nuclear Pore Proteins
- DNA recording with synthetic nanopores
- Accurate and Efficient Modeling of Carbon Nanotubes in Biological Applications

Significant progress has been made on most of the projects, with published results in high impact journals. The technical and methodological advances needed to accomplish these collaborations are: development of a coarse-grained model for proteins and nucleic acids; the development of multiscale simulation tools (setup and simulation); enhanced sampling with temperature and Hamiltonian replica exchange methodologies; methods to analyze structural dynamics, flexibility, and stress in proteins; incorporation of QM/MM methods; and finally, developments allowing seamless communications between visualization and structural homology modeling tools as well as automated setup, production and analysis

of all-atom simulations. Many projects involved large molecular systems that need to be simulated for long time scales. As such, they posed unprecedented challenges to the size and time scale that are currently achievable in MD simulations. These projects are collaborative in the very best sense, allowing the extensive expertise of the collaborators to bear on the development of cutting-edge simulation methodologies, thus contributing additional insights into the behavior of complex biological systems.

Other activities associated with the Resource are Training, Service, Dissemination and Administration.

5. Training

The training strategies of the Resource concentrate on teaching hands-on sessions of all available tools for different scientific applications. Major successes have been computational biology workshops and, more recently, cluster building workshops. Computational workshops use lectures, tutorials, and case studies to instruct students. Cluster workshops teach how to build computational clusters for long simulations. For those unable to attend workshops, all tutorial material is available from the Resource's web site. Future training avenues include online workshops, creating an online textbook on DVD and in print. Currently the students in the workshops are primarily graduate and post-doctoral researchers. It is planned to modifying the training methodologies to also serve undergraduate and high-school students. This will aid in training the next generation of life scientists. Participants in the workshops are chosen among applicants such that the workshop includes participants from multiple institutions, gender and geographical locations. In some instances NIH and NSF funding has been obtained to help finance the students participation in the workshops. Students evaluate the workshop lectures on a daily basis and provide feedback to the lecturers. In the past, the Resource has been very successful in its training activity. Almost 300 students participated in the computational workshops and 100 in the cluster building ones. This training activity will continue and also will expand to more on-line workshops. Nice new features are the case studies on the web; 9 exist at the present time already. The hands-on workshops are a tremendous achievement and probably the best way to disseminate any of the software tools.

6. Service

The Resource's service efforts are excellent and are focused in three areas: the Resource software suite, the support of external users, and the Resource web site. The software suite is its primary service, providing 60,000+ users (about 20% of which are NIH funded) with powerful tools for their biomedical research. The external users, which make up half of the Resource's system accounts, are supported first through the use of the Resource's flexible and powerful computational facility and second through the brand-new Visitor Center, which offers support and facilities to visitors to the Resource. Finally, the Resource web

site offers access to the combined documentation and knowledge base of the Resource staff, and has served approximately 56,000 users per month over the last year.

A substantial computing resource will be needed by the users of the NIH Resource to realize the benefits of the new modeling capabilities that will be developed during the next funding period. The University will need to provide a facility capable of housing these resources, which are expected to comprise nearly 300 processors plus disk storage. The new computing facility must provide the power and cooling needed by the new computing systems.

7. Dissemination

For an effective dissemination program the Resource uses traditional academic channels as well as electronic channels supported by the Internet for distribution ofresearch results and software. Being familiar with available distribution channels is only one aspect of dissemination; prodedures have to be place a set of dissemination strategies that guide and encourage a Resource in current and future research and software distribution activities. The Resource has developed six strategies that will guide future dissemination activities: Pro-active methods entail getting information out to the biomedical community in a timely fashion, through papers, lectures, and web site postings, new releases etc. Reactive Methods entail responding to external inquiries for materials. Providing Multiple Paths to a unit of information (e.g., a MS) helps with distribution, as does Ease of Access. Easily navigable web design is a must. The notion of 'Maximum Availability' directs that a Resource make all its intellectual products available and maintain their availability, including pursuing open licenses and supporting license revisions, while 'Minimum Time' indicates that any new materials be made available as soon as possible. Utilizing these dissemination strategies has proved effective for the Resource over the last year. Examples of dissemination activities include 36 articles in refereed journals, production of two new brochures (on Resource computing and workshops), 40 talks by PIs and 36 talks or posters by others, 30 stories in media outlets about the Resource, a high volume of requests for publications, and improvements to the Resource web site in both style and usability. Future plans include description of the user experience with the Resource through dedicated web sites and brochure development, updates to web site image and movie galleries, expansion of visitor and undergraduate training programs, participation in local outreach events, and logo development. All of these activities reflect the Resources primary dissemination goal maintaining and increasing the already excellent dissemination goals.

8. Administration

A functioning Resource involves a cluster of activities, including scientific collaborations, service, training, development, and dissemination and the interactions between these activities. Connecting and supporting these activities are administrative structures, both

internal to the Resource and in how the Resource works with external organizations. The Resource exists as a research group within the University of Illinois system, at the Beckman Institute for Advanced Science and Technology. The Primary Investigators and Directors, as well as an Advisory Board provide the leadership of collaborative research, technical and administrative support, and technological developments. The intellectual diversity and senior staff of the Resource, along with a substantial supporting technological infrastructure, greatly contribute to the Resource being able to fulfill its obligations. Administrative features the Resource will be adding in the near future include committees devoted to identifying software features to be developed in the VMD, NAMD, and Modeling Suite applications, and a committee for selecting collaborations to be pursued by the Resource. The Software Features Committees will be constituted of a Primary Investigator and software developer(s) from the Resource, and a rotating external scientist who serves for a one-year period. Quarterly reports for each application on features developed, requested, and pending completion will be review by these committees, who on an annual basis will produce advisory reports on development. The Collaboration Selection Committee will be manned by all Resource PIs, and on quarterly basis will review suggested collaborations collected via the web site and through other contacts. Collaborations will be selected according to criteria including the needed qualities for research (e.g., biomedical relevance, quality and originality of the research) and qualities of the methodology (e.g., project requires technological development, great computational demands). Contested decisions will be brought before the Advisory Board for review and resolution.

One concern of the Advisory Committee is the fragility of the non faculty Resource staff positions. The expertise of the programming staff is essential for the continuing success of this Resource. Measures should be considered to enhance the long term stability of these positions.

Administration

Organization

The organization and operation of the Resource supports development and distribution of software, collaborations, user service, and interactions between researchers and developers. Software development, both of current and planned applications, is the central responsibility of assigned programmers, with input and assistance from other members of the Resource. Software distribution occurs via the Resource web site, with the application web sites managed by the software developers. Server hardware underlying the web site is maintained by the Resource's system administration team. Collaborations with external scientists, where Resource graduate students, postdoctoral associates, and faculty work with outside researchers on projects that require new methodological solutions, benefit from and provide direction to software development.

The Resource's many service, training, and dissemination activities involve all members of the Resource, *e.g.*, hosting external scientists in the visitor center, or providing members of the biomedical community access to Resource computing facilities. Interactions stemming from collaborations, other sources of input from external scientists, and internal contacts between Resource scientists and developers, as supported by administrative structures, produce a dynamic environment that fosters both research and development. Activities of the Resource are supported by both external and internal organizational structures, and by committees.

External Structures. Externally, the Resource resides within the Beckman Institute for Advanced Science and Technology*, at the University of Illinois at Urbana-Champaign (UIUC)†, one of three campuses of the University of Illinois system. The mission of the Beckman Institute is to foster basic, interdisciplinary research as focused around three reseach initiatives: biological intelligence, human-computer intelligent interaction, and molecular and electronic nanostructures. Organizationally, the Resource belongs to the molecular and electronic nanostructures research initiative, where the emphasis is on developing a fundamental understanding of chemical and physical processes involving structures on the nanometer scale. The Resource is involved in close collaborative projects with other groups that are part of this research initiative, mainly in the area of biotechnology.

Administratively, the Director of the Beckman Institute reports to the campus Provost and Vice Chancellor for Academic Affairs. Resource members Drs. Schulten, Luthey-Schulten, Kalé, Tajkhorshid, and Aksimentiev all have faculty appointments at the Beckman Institute. Other contacts with major campus units come through the UIUC faculty

*`http://www.beckman.uiuc.edu/`

†`http://www.uiuc.edu/`

positions of Resource primary investigators. Drs. Schulten and Aksimentiev have appointments at the Department of Physics; Drs. Schulten, Luthey-Schulten, and Tajkhorshid have appointments at the Center for Biophysics and Computational Biology (a unit of the Department of Molecular and Cellular Biology); Drs. Schulten and Luthey-Schulten have appointments at the Department of Chemistry; and, Dr. Kalé has appointments in the Department of Computer Science and in the Department of Electrical and Computer Engineering.

Internal Structures. Internally, the Resource is led by Principal Investigators (PIs) and Directors, with guidance, information, and expertise from the Advisory Committee. Three functional internal subunits - technical and administrative support, technological development, and collaborations - carry out Resource operations. The subunit 'technical and administrative support' includes development and maintenance of computing clusters; maintenance of desktop machines and network connections; and, clerical and administrative support, including interfacing with other campus administrative units. Members of the technological development subunit spend the majority of their time developing software for the Resource. Included under the collaborations subunit is work with external scientists, typically involving one or more Resource graduate students or postdoctoral associates, a faculty member, and a member of the technological development unit.

Any given task carried out by the Resource is likely to involve multiple members of any one of the organizational subunits, for example a collaborative project will typically require support from the other two units. All members participate in the administration of the Resource by taking on tasks related to operation of the Resouce, such as assisting in system administration, or contributing to the web site. Resource members also attend regular all-member and subgroup meetings. A highly-developed internal website tracks meetings and provides information and resources in four main categories: administration and service, events and outreach, databases and records, and a catch-all miscellaneous category. For example, meeting agendas and minutes generated by general meetings of all Resource members are kept on the internal site, providing useful reference documentation for decisions on any number of topics addressed during the meetings.

Committees. Two committees, the *Collaboration Selection Committee* and the *Software Features Committee*, are important administrative tools for the Resource. The purpose of the Collaboration Selection Committee is to decide which collaborations should be pursued by the Resource. The Resource adds about one collaborator a month and on average produces one joint publication per collaborator; successful projects usually take 2-3 years. Suggestions for collaborations come via a number of channels, such as direct requests to Resource PIs, Resource member suggestions, meetings at conferences, and other contacts with Resource members. The PI and Co-PIs of the Resource comprise the membership of the committee, with regular meetings of the committee four times a year.

Selection of collaboration requires careful consideration of both research and methodological criteria, including: biomedical relevance; quality/originality of the research and conceptual approach; experience/quality of the external PI(s); match with Resource member abilities/interests; project challenges to technological/application development; and, the computational demands of the project.

The purpose of the recently dveloped Software Features Committee is to suggest and prioritize new features for each of the major software applications being developed by the Resource. Membership of the committee includes four external scientists who serve for a 12-month period. On an annual basis, committee members review a list of planned features and, as needed, lists of requested and completed features. For each planned feature, committee members are asked to judge if the feature will provide a wide benefit to the biomedical community, if the feature will enhance research by doing something faster or doing something new, and if the feature is consistent with the purpose of its host program. Once collected, Resource faculty and developers review and clarify committee member comments, and their implications for software development.

Allocation of Resource Access

Access to the Resource is provided at three general levels: access to Resource software, to software developers/development, and to Resource expertise. Access to Resource develops software - Visual Molecular Dynamics (VMD)*, Nanoscale Molecular Dynamics (NAMD)†, and Biological Collaborative Environment (BioCoRE)‡ - is provided via the Resource's popular web site§. Information on the number of registered users of each application is provided below, along with statistics on use of the web site and counts of external users accessing the Resource's computational facilities. Users have also access to software support by email. Statistis of this widely-used service are also provided below.

Access to Resource development efforts - the opportunity to interact with software developers - is provided via multiple channels. All major software applications provide e-mail contacts and mailing lists. Further, the VMD application web site provides a Public Project via BioCoRE¶, where the user community can exchange tips and information about VMD, and the NAMD web site provides a wiki‖ of user-modifiable web pages on numerous topics. Information describing exchanges with software developers (e.g., the number of emails with developers) is provided below. The Resource's *Software Features Committee* (described in more detail in the *Organization* section) allows members of the biomedical community to directly voice opinion on the development plans of the Resource's major software applications (results from the committee are pending, and will be summarized in a future report).

Access to Resource expertise is also available via multiple channels. Collaborations, as represented by the subprojects included with this report, represent a long-term access of Resource expertise, and as such are carefully selected by the Resource's *Collaboration Selection Committee* (described in more detail in the *Organization* section). Other accesses of Resource expertise include developer consultations on hardware systems, the Resource's visitor program, and other training efforts as described in the *Training* section. Further, the Resource organizes a seminar program that brings scientists to the University of Illinois at Urbana-Champaign (UIUC) to present their work and to meet with Resource members.

---

*http://www.ks.uiuc.edu/Research/vmd/

†http://www.ks.uiuc.edu/Research/namd/

‡http://www.ks.uiuc.edu/Research/biocore/

§http://www.ks.uiuc.edu/

¶http://biocore.ks.uiuc.edu/biocore/biofs/VMD%20(Public)/index.html

‖http://www.ks.uiuc.edu/Research/namd/wiki/

Access accomplishments by the Resource as related to access/service over the last year include:

- 18,972 additional registered users of the Resource's VMD software
- 4,685 addtional registered users of the Resource's NAMD software
- 605 additional registered users of the Resource's BioCoRE collaboratory
- 7,800 VMD emails, 2,427 NAMD emails, and 1,692 BioCoRE chats and emails were exchanged in user support
- 1,881 citations of the VMD source paper; 542 citations of the NAMD source papers
- 121 external users of the Resource computational facilities
- 5 workshops were presented, on computational biology and cluster building, online and in-person
- 17 visitors to the Resource received training
- 6 consultations to outside groups regarding computer cluster or visualization facility development
- 16 seminars were organized by the Resource

The Resource is engaged in intensive development efforts and technology transfer. A number of software packages, particularly VMD, NAMD and BioCoRE, as well as a number of smaller programs, are freely distributed. All Resource-developed programs, binaries and source, are available on our web site for easy accessibility, employing a unified distribution mechanism**. The VMD, NAMD and BioCoRE packages are developed, maintained, and distributed by Resource staff. The staff also offers extensive user support and has turned the Resource web site into a leading and a widely recognized distribution resource for biomedical software. In this report we are focusing on the distribution and support accomplishments of VMD, NAMD and BioCoRE, over the last year.

Use of VMD, NAMD, and BioCoRE

VMD has 88,319 registered users as of April 2007 (an increase of 18,972 or +27% since March 2006), with 16,579 of those users repeat users (i.e., they have downloaded more than one version of VMD), and 19% of all registrants having NIH funding. VMD 1.8.5, released in August 2006, has over 17,500 registered users; VMD 1.8.6, released in early April, already has 2,523 registered users.

**URL:`http://www.ks.uiuc.edu/Development/Download/download.cgi`

NAMD has 20,256 registered users (as of April 2007) (an increase of 4,685 or +30% since March 2006), of whom 4,282 or 21% are repeat users. 3,392 (17%) of NAMD users are NIH funded. The current version of NAMD, version 2.6 released in August 2006, has 4,497 registered users (814 are NIH funded).

BioCoRE has 1,939 registered users (an increase of 605, or +45% in the past year), involved in 464 projects (compared to 390 a year ago). 176 projects within BioCoRE have been reported as either fully or partially NIH-funded.

The software release schedule of the Resource's lead programs reflects great productivity and lively activity:

- VMD: 1.8.6 released April 2007
- NAMD: 2.6 released August 2006
- BioCoRE: Incremental updates every few weeks

Development, Distribution, and Use of VMD

Below we report service rendered by the Resource through its molecular graphics and structure/dynamics analysis program VMD. The program enjoyed during the reported period significant improvements and a further drastic increase in user numbers.

VMD Enhancements for 2006-2007 include amongst other features:

- a new rendering mode which supports use of Adobe Acrobat3D, allowing researchers to create PDF documents containing 3-D scenes from VMD[††].
- VMD now incorporates expanded support for high-quality molecular scene rendering using Tachyon, NVIDIA Gelato, and PIXAR RenderMan. These renderers support advanced lighting modes such as ambient occlusion lighting, which can be used to create molecular renderings that look more "three dimensional". These renderers also support high dynamic range lighting (HDR) and high precision color image formats.
- VMD includes a new "Field Lines" representation which can be used to display images of electrostatic fields, flow fields, and other volumetric data for which viewing field lines or particle advection traces is helpful.
- VMD 1.8.6 uses just over half as much memory relative to prior versions, allowing structures of up to 72 million atoms to be loaded and displayed on workstations with 16GB of memory, using the 64-bit versions of VMD.

[††]URL:`http://www.adobe.com/products/acrobat3d/`

Scope of VMD User Support:

- 7,800 e-mail exchanges in response to user inquiries sent to the vmd@ks.uiuc.edu e-mail address
- 625 subscribers to the VMD-L mailing list, with 9,445 total postings, and 2,551 postings from the end of April 2006 through March 2007
- Local face-to-face support has been provided

There are currently 377 non-Resource users with access to the VMD source code repository, with 112 such users added in the last year.

Sites with Links to the VMD Site (Google, April 2007): 174 domains; 1,075 pages.

Development, Distribution, and Use of NAMD

During the reported period, NAMD enjoyed significant improvements and continued to increase in its number of registered users. The program is widely considered as uniquely satisfying the demand for an effective program on the new generation of teraflop parallel computers.

NAMD Enhancements for 2006-2007 include among other features:

- Ports to Itanium, Altix, and Opteron/Athlon64/EMT64.
- Port to Mac OS X for Intel processors.
- Ports to Cray XT3 and IBM BlueGene/L (source code only).
- Improved serial performance, especially on POWER and PowerPC.
- Adaptive biasing force free energy calculations.
- Customizable replica exchange simulations.
- Tcl-based boundary potentials.
- Reduced memory usage for unusual simulations.
- Support for CHARMM 31 stream files and CMAP crossterms.
- Support for OPLS force field.

NAMD Availability in Supercomputer Centers:

- Pittsburgh Supercomputing Center

- National Center for Supercomputing Applications
- San Diego Supercomputer Center

Scope of NAMD User Support:

- The NamdWiki user-editable web site contains 45 topical pages, with the ability for users to add their own pages, providing a public whiteboard for sharing NAMD issues, experiences, providing advice, and troubleshooting; sample wiki topics are "NAMD on Blue Gene" and "NAMD at NCSA"
- 535 subscribers to the NAMD-L mailing list, with 5,641 total postings, and 1,897 postings for the April 2006 - March 2007 period
- Over 530 emails exchanged with users via the namd@ks.uiuc.edu e-mail address, a number which excludes questions sent to the Charm++ developers or the NAMD and VMD mailing lists
- Local face-to-face support has been provided

There are currently 260 users with access to the NAMD source code repository, with 36 users added in the last year.

Sites with Links to NAMD site (Google, 2007): 90 domains; 1,055 pages.

Development, Distribution, and Use of BioCoRE

During the reported period, BioCoRE added several significant improvements and has become more widely adopted by the community. BioCoRE continues to be ideally suited for making the investment into the US computational grid useful for biomedical research.

BioCoRE Enhancements for 2006-07 include among other features:

- Significant advances in BioCoRE/VMD interactions (communication via a VMD "control panel", account creation)
- Integration of BioCoRE supercomputer job submission capabilities into VMD Plugins NAMDEnergy and APBSRun.
- BioCoRE Programming Interface now provides access to more BioCoRE functionality

Scope of BioCoRE User Support:

- 143 emails issued to/from `biocore@ks.uiuc.edu` from April 2006 - March 2007

- 1,549 chat messages sent to the BioCoRE public help project from April 2006 - March 2007 within BioCoRE itself.

Sites with Links to BioCoRE Site (Google, April 2007): 18 domains; 971 pages.

Citations of Software Source Papers

All users of Resource software are asked to acknowlege in any journal or other publications the source paper for the software that they used. Searches of online citations databases then provide one means of indicating the use of a software application. Recent citation search results for the VMD, NAMD, and BioCoRE source papers are provided below.

List of papers citing VMD: A literature search in the ISI Web of Science citation database in April 2007 yielded 1,881 published journal articles, papers, or books citing the VMD origin paper [42]. Below are 50 recent citations:

- Humeres, E., C. Mascayano, G. Riadi and F. Gonzalez-Nilo (2006). "Molecular dynamics simulation of the aqueous solvation shell of cellulose and xanthate ester derivatives." Journal of Physical Organic Chemistry 19(12): 896-901.

- Wang, Y., J. Cohen, W. F. Boron, K. Schulten and E. Tajkhorshid (2007). "Exploring gas permeability of cellular membranes and membrane channels with molecular dynamics." Journal of Structural Biology 157(3): 534-544.

- Shih, A. Y., P. L. Freddolino, A. Arkhipov and K. Schulten (2007). "Assembly of lipoprotein particles revealed by coarse-grained molecular dynamics simulations." Journal of Structural Biology 157(3): 579-592.

- Palmer, B., S. Kathmann, M. Krishnan, V. Tipparaju and J. Nieplocha (2007). "The use of processor groups in molecular dynamics simulations to sample free-energy states." Journal of Chemical Theory and Computation 3(2): 583-592.

- Ghosh, P., D. R. Katti and K. S. Katti (2007). "Mineral proximity influences mechanical response of proteins in biological mineral-protein hybrid systems." Biomacromolecules 8(3): 851-856.

- Fowler, P. W., K. Balali-Mood, S. Deol, P. V. Coveney and M. S. P. Sansom (2007). "Monotopic enzymes and lipid bilayers: A comparative study." Biochemistry 46(11): 3108-3115.

- De Fabritiis, G., M. Serrano, R. Delgado-Buscalioni and P. V. Coveney (2007). "Fluctuating hydrodynamic modeling of fluids at the nanoscale." Physical Review E 75(2).

- Solares, S. D. (2007). "Single biomolecule imaging with frequency and force modulation in tapping-mode atomic force microscopy." Journal of Physical Chemistry B 111(9): 2125-2129.

- Calderon, C. P. (2007). "On the use of local diffusion models for path ensemble averaging in potential of mean force computations." Journal of Chemical Physics 126(8).

- Dittrich, M., J. Yu and K. Schulten (2007). "PcrA helicase, a molecular motor studied from the electronic to the functional level. Atomistic Approaches in Modern Biology: From Quantum Chemistry to Molecular Simulations." 268: 319-347.

- Lou, J. Z. and C. Zhu (2007). "A structure-based sliding-rebinding mechanism for catch bonds." Biophysical Journal 92(5): 1471-1485.

- Balog, E., M. Laberge and J. Fidy (2007). "The influence of interdomain interactions on the intradomain motions in yeast phosphoglycerate kinase: A molecular dynamics study." Biophysical Journal 92(5): 1709-1716.

- Wei, K., L. Liu, Y. H. Cheng, Y. Fu and Q. X. Guo (2007). "Theoretical examination of two opposite mechanisms proposed for hepatitis delta virus ribozyme." Journal of Physical Chemistry B 111(7): 1514-1516.

- Amaro, R. E., A. Sethi, R. S. Myers, V. J. Davisson and Z. A. Luthey-Schulten (2007). "A network of conserved interactions regulates the allosteric signal in a glutamine amidotransferase." Biochemistry 46(8): 2156-2173.

- Isgro, T. A. and K. Schulten (2007). "Association of nuclear pore FG-repeat domains to NTF2 import and export complexes." Journal of Molecular Biology 366(1): 330-345.

- Gorfe, A. A., M. Hanzal-Bayer, D. Abankwa, J. F. Hancock and J. A. McCammon (2007). "Structure and dynamics of the full-length lipid-modified H-ras protein in a 1,2-dimyristoylglycero-3-phosphocholine bilayer." Journal of Medicinal Chemistry 50(4): 674-684.

- Rueda, M., C. Ferrer-Costa, T. Meyer, A. Perez, J. Camps, A. Hospital, J. L. Gelpi and M. Orozco (2007). "A consensus view of protein dynamics." Proceedings of the National Academy of Sciences of the United States of America 104(3): 796-801.

- Lippert, R. A., K. J. Bowers, R. O. Dror, M. P. Eastwood, B. A. Gregersen, J. L. Klepeis, I. Kolossvary and D. E. Shaw (2007). "A common, avoidable source of error in molecular dynamics integrators." Journal of Chemical Physics 126(4).

- Stitham, J., E. J. Arehart, S. R. Gleim, K. L. Douville and J. Hwa (2007). "Human prostacyclin receptor structure and function from naturally-occurring and synthetic mutations." Prostaglandins & Other Lipid Mediators 82(1-4): 95-108.

- Pogorelov, T. V., F. Autenrieth, E. Roberts and Z. A. Luthey-Schulten (2007). "Cytochrome c(2) exit strategy: Dissociation studies and evolutionary implications." Journal of Physical Chemistry B 111(3): 618-634.

- Patra, M., M. T. Hyvonen, E. Falck, M. Sabouri-Ghomi, I. Vattulainen and M. Karttunen (2007). "Long-range interactions and parallel scalability in molecular simulations." Computer Physics Communications 176(1): 14-22.

- Thevenard, J., N. Floquet, L. Ramont, E. Prost, J. M. Nuzillard, M. Dauchez, H. Yezid, A. J. P. Alix, F. X. Maquart and S. Brassart-Plasco (2006). "Structural and antitumor properties of the YSNSG cyclopeptide derived from tumstatin." Chemistry & Biology 13(12): 1307-1315.

- Sotomayor, M., V. Vasquez, E. Perozo and K. Schulten (2007). "Ion conduction through MscS as determined by electrophysiology and simulation." Biophysical Journal 92(3): 886-902.

- Pedersen, U. R., G. H. Peters and P. Westh (2007). "Molecular packing in 1-hexanol-DMPC bilayers studied by molecular dynamics simulation." Biophysical Chemistry 125(1): 104-111.

- Kikuzawa, Y., T. Nagata, T. Tahara and K. Ishii (2006). "Photo- and redox-active dendritic molecules with soft, layered nanostructures." Chemistry-an Asian Journal 1(4): 516-528.

- Arkhipov, A., P. L. Freddolino and K. Schulten (2006). "Stability and dynamics of virus capsids described by coarse-grained modeling." Structure 14(12): 1767-1777.

- Muller, M., K. Katsov and M. Schick (2006). "Biological and synthetic membranes: What can be learned from a coarse-grained description?" Physics Reports-Review Section of Physics Letters 434(5-6): 113-176.

- Bellesia, G., M. V. Fedorov and E. G. Timoshenko (2007). "Molecular dynamics study of structural properties of beta-sheet assemblies formed by synthetic de novo oligopeptides." Physica a-Statistical Mechanics and Its Applications 373: 455-476.

- Krystek, S. R., S. R. Kimura and A. J. Tebben (2006). "Modeling and active site refinement for G protein-coupled receptors: application to the beta-2 adrenergic receptor." Journal of Computer-Aided Molecular Design 20(7-8): 463-470.

- Arkhipov, A., P. L. Freddolino, K. Imada, K. Namba and K. Schulten (2006). "Coarse-grained molecular dynamics simulations of a rotating bacterial flagellum." Biophysical Journal 91(12): 4589-4597.

- Stitham, J., S. R. Gleim, K. Douville, E. Arehart and J. Hwa (2006). "Versatility and differential roles of cysteine residues in human prostacyclin receptor structure and function." Journal of Biological Chemistry 281(48): 37227-37236.

- Yin, Y., M. O. Jensen, E. Tajkhorshid and K. Schulten (2006). "Sugar binding and protein conformational changes in lactose permease." Biophysical Journal 91(11): 3972-3985.

- Buehler, M. J. (2006). "Large-scale hierarchical molecular modeling of nanostructured biological materials." Journal of Computational and Theoretical Nanoscience 3(5): 603-623.

- Freddolino, P. L., M. Dittrich and K. Schulten (2006). "Dynamic switching mechanisms in LOV1 and LOV2 domains of plant phototropins." Biophysical Journal 91(10): 3630-3639.

- Cruz-Chu, E. R., A. Aksimentiev and K. Schulten (2006). "Water-silica force field for simulating nanodevices." Journal of Physical Chemistry B 110(43): 21497-21508.

- Puklin-Faucher, E., M. Gao, K. Schulten and V. Vogel (2006). "How the headpiece hinge angle is opened: new insights into the dynamics of integrin activation." Journal of Cell Biology 175(2): 349-360.

- Tuukkanen, A., M. I. Verkhovsky, L. Laakkonen and M. Wikstrom (2006). "The K-pathway revisited: A computational study on cytochrome c oxidase." Biochimica Et Biophysica Acta-Bioenergetics 1757(9-10): 1117-1121.

- Blood, P. D. and G. A. Voth (2006). "Direct observation of Bin/amphiphysin/Rvs (BAR) domain-induced membrane curvature by means of molecular dynamics simulations." Proceedings of the National Academy of Sciences of the United States of America 103(41): 15068-15072.

- Klauda, J. B., B. R. Brooks and R. W. Pastor (2006). "Dynamical motions of lipids and a finite size effect in simulations of bilayers." Journal of Chemical Physics 125(14).

- Foley, M. C., K. Arora and T. Schlick (2006). "Sequential side-chain residue motions transform the binary into the ternary state of DNA polymerase lambda." Biophysical Journal 91(9): 3182-3195.

- Yraola, F., S. Garcia-Vicente, J. Fernandez-Recio, F. Albericio, A. Zorzano, L. Marti and M. Royo (2006). "New efficient substrates for semicarbazide-sensitive amine oxidase/VAP-1 enzyme: Analysis by SARs and computational docking." Journal of Medicinal Chemistry 49(21): 6197-6208.

- Cheng, X. L., H. L. Wang, B. Grant, S. M. Sine and J. A. McCammon (2006). "Targeted molecular dynamics study of C-loop closure and channel gating in nicotinic receptors." Plos Computational Biology 2(9): 1173-1184.

- Lou, J. Z., T. Yago, A. G. Klopocki, P. Mehta, W. Chen, V. I. Zarnitsyna, N. V. Bovin, C. Zhu and R. P. McEver (2006). "Flow-enhanced adhesion regulated by a selectin interdomain hinge." Journal of Cell Biology 174(7): 1107-1117.

- Jang, H., B. Ma, T. B. Woolf and R. Nussinov (2006). "Interaction of protegrin-1 with lipid bilayers: Membrane thinning effect." Biophysical Journal 91(8): 2848-2859.

- Yu, J., A. J. Yool, K. Schulten and E. Tajkhorshid (2006). "Mechanism of gating and ion conductivity of a possible tetrameric pore in aquaporin-1." Structure 14(9): 1411-1423.

- Roh, J. H., J. E. Curtis, S. Azzam, V. N. Novikov, I. Peral, Z. Chowdhuri, R. B. Gregory and A. P. Sokolov (2006). "Influence of hydration on the dynamics of lysozyme." Biophysical Journal 91(7): 2573-2588.

- Khalili-Araghi, F., E. Tajkhorshid and K. Schulten (2006). "Dynamics of K+ ion conduction through Kv1.2." Biophysical Journal 91(6): L2-L4.

- Yu, J., T. Ha and K. Schulten (2006). "Structure-based model of the stepping motor of PcrA helicase." Biophysical Journal 91(6): 2097-2114.

- Henin, J., K. Schulten and C. Chipot (2006). "Conformational equilibrium in alanine-rich peptides probed by reversible stretching simulations." Journal of Physical Chemistry B 110(33): 16718-16723.

- Meyer, G. R., J. Gullingsrud, K. Schulten and B. Martinac (2006). "Molecular dynamics study of MscL interactions with a curved lipid bilayer." Biophysical Journal 91(5): 1630-1637.

List of papers citing NAMD: A literature search in the ISI Web of Science citation database in March 2007 yielded 542 published journal articles, papers, or books citing the current NAMD origin paper [28]. Below are 50 recent cites:

- Sacquin-Mora, S., E. Laforet and R. Lavery (2007). "Locating the active sites of enzymes using mechanical properties." Proteins-Structure Function and Bioinformatics 67(2): 350-359.

- Allen, M. P. (2007). "Educational aspects of molecular simulation." Molecular Physics 105(2-3): 157-166.

- Rinaldo, D., D. M. Philipp, S. J. Lippard and R. A. Friesner (2007). "Intermediates in dioxygen activation by methane monooxygenase: A QM/MM study." Journal of the American Chemical Society 129(11): 3135-3147.

- Voityuk, A. A. and W. B. Davis (2007). "Hole transfer energetics in structurally distorted DNA: The nucleosome core particle." Journal of Physical Chemistry B 111(11): 2976-2985.

- Bastug, T. and S. Kuyucak (2007). "Free energy simulations of single and double ion occupancy in gramicidin A." Journal of Chemical Physics 126(10).

- Sieffert, N. and G. Wipff (2007). "Importance of interfacial adsorption in the biphasic hydroformylation of higher olefins promoted by cyclodextrins: A molecular dynamics study at the decene/water interface." Chemistry-a European Journal 13(7): 1978-1990.

- Haider, S., S. Khalid, S. J. Tucker, F. M. Ashcroft and M. S. P. Sansom (2007). "Molecular dynamics simulations of inwardly rectifying (Kir) potassium channels: A comparative study." Biochemistry 46(12): 3643-3652.

- Iimura, S., T. Umezaki, M. Takeuchi, M. Mizuguchi, H. Yagi, K. Ogasahara, H. Akutsu, Y. Noda, S. Segawa and K. Yutani (2007). "Characterization of the denatured structure of pyrrolidone carboxyl peptidase from a hyperthermophile under nondenaturing conditions: Role of the C-terminal alpha-helix of the protein in folding and stability." Biochemistry 46(12): 3664-3672.

- Borowski, T., S. de Marothy, E. Broclawik, C. J. Schofield and P. E. M. Siegbahn (2007). "Mechanism for cyclization reaction by clavaminic acid synthase. Insights from modeling studies." Biochemistry 46(12): 3682-3691.

- Qiao, R., A. P. Roberts, A. S. Mount, S. J. Klaine and P. C. Ke (2007). "Translocation of C-60 and its derivatives across a lipid bilayer." Nano Letters 7(3): 614-619.

- Kormos, B. L., A. M. Baranger and D. L. Beveridge (2007). "A study of collective atomic fluctuations and cooperativity in the U1A-RNA complex based on molecular dynamics simulations." Journal of Structural Biology 157(3): 500-513.

- Wang, Y., J. Cohen, W. F. Boron, K. Schulten and E. Tajkhorshid (2007). "Exploring gas permeability of cellular membranes and membrane channels with molecular dynamics." Journal of Structural Biology 157(3): 534-544.

- Shih, A. Y., P. L. Freddolino, A. Arkhipov and K. Schulten (2007). "Assembly of lipoprotein particles revealed by coarse-grained molecular dynamics simulations." Journal of Structural Biology 157(3): 579-592.

- Bond, P. J., J. Holyoake, A. Ivetac, S. Khalid and M. S. P. Sansom (2007). "Coarse-grained molecular dynamics simulations of membrane proteins and peptides." Journal of Structural Biology 157(3): 593-605.

- Tozzini, V., J. Trylska, C. E. Chang and J. A. McCammon (2007). "Flap opening dynamics in HIV-1 protease explored with a coarse-grained model." Journal of Structural Biology 157(3): 606-615.

- Printz, M. and C. Richert (2007). "Optimizing the stacking moiety and linker of 2 '-acylamido caps of DNA duplexes with 3 '-terminal adenine residues." Journal of Combinatorial Chemistry 9(2): 306-320.

- Guo, X., A. Y. T. Leung, H. Jiang, X. Q. He and Y. Huang (2007). "Critical strain of carbon nanotubes: An atomic-scale finite element study." Journal of Applied Mechanics-Transactions of the Asme 74(2): 347-351.

- Baker, C. M. and G. H. Grant (2007). "Modeling aromatic liquids: Toluene, phenol, and pyridine." Journal of Chemical Theory and Computation 3(2): 530-548.

- Craig, A. M. and Y. Kang (2007). "Neurexin-neuroligin signaling in synapse development." Current Opinion in Neurobiology 17(1): 43-52.

- Linderoth, L., G. H. Peters, K. Jorgensen, R. Madsen and T. L. Andresen (2007). "Synthesis of sn-1 functionalized phospholipids as substrates for secretory phospholipase A(2)." Chemistry and Physics of Lipids 146(1): 54-66.

- Hang, B. and A. B. Guhaev (2007). "Substrate specificity of human thymine-DNA glycosylase on exocyclic cytosine adducts." Chemico-Biological Interactions 165(3): 230-238.

- Yui, T. and S. Hayashi (2007). "Molecular dynamics simulations of solvated crystal models of cellulose I-alpha and IIII." Biomacromolecules 8(3): 817-824.

- Ghosh, P., D. R. Katti and K. S. Katti (2007). "Mineral proximity influences mechanical response of proteins in biological mineral-protein hybrid systems." Biomacromolecules 8(3): 851-856.

- Maupin, C. M. and G. A. Voth (2007). "Preferred orientations of His64 in human carbonic anhydrase II." Biochemistry 46(11): 2938-2947.
- Fowler, P. W., K. Balali-Mood, S. Deol, P. V. Coveney and M. S. P. Sansom (2007). "Monotopic enzymes and lipid bilayers: A comparative study." Biochemistry 46(11): 3108-3115.
- Brewer, S. H., B. B. Song, D. P. Raleigh and R. B. Dyer (2007). "Residue specific resolution of protein folding dynamics using isotope-edited infrared temperature jump spectroscopy." Biochemistry 46(11): 3279-3285.
- Dal Peraro, M., A. J. Vila, P. Carloni and M. L. Klein (2007). "Role of zinc content on the catalytic efficiency of B1 metallo beta-lactamases." Journal of the American Chemical Society 129(10): 2808-2816.
- Moreira, I. S., P. A. Fernandes and M. J. Ramos (2007). "Hot spot occlusion from bulk water: A comprehensive study of the complex between the lysozyme HEL and the antibody FVD1.3." Journal of Physical Chemistry B 111(10): 2697-2706.
- Chaumont, A. and G. Wipff (2007). "Solvation of "big" spherical solutes in room temperature ionic liquids and at their aqueous interface: A molecular dynamics simulation study." Journal of Molecular Liquids 131: 36-47.
- Mark, P. and L. Nilsson (2007). "A molecular dynamics study of Cyclophilin A free and in complex with the Ala-Pro dipeptide." European Biophysics Journal with Biophysics Letters 36(3): 213-224.
- Ivetac, A., J. D. Campbell and M. S. P. Sansom (2007). "Dynamics and function in a bacterial ABC transporter: Simulation studies of the BtuCDF system and its components." Biochemistry 46(10): 2767-2778.
- Novak, B. R., E. J. Maginn and M. J. McCready (2007). "Comparison of heterogeneous and homogeneous bubble nucleation using molecular simulations." Physical Review B 75(8).
- Shinoda, W., R. Devane and M. L. Klein (2007). "Multi-property fitting and parameterization of a coarse grained model for aqueous surfactants." Molecular Simulation 33(1-2): 27-36.
- Whitford, P. C., O. Miyashita, Y. Levy and J. N. Onuchic (2007). "Conformational transitions of adenylate kinase: Switching by cracking." Journal of Molecular Biology 366(5): 1661-1671.
- Knotts, T. A., N. Rathore, D. C. Schwartz and J. J. de Pablo (2007). "A coarse grain model for DNA." Journal of Chemical Physics 126(8).

- Bachorz, R. A., W. Klopper and M. Gutowski (2007). "Coupled-cluster and explicitly correlated perturbation-theory calculations of the uracil anion." Journal of Chemical Physics 126(8).

- Lasker, K., O. Dror, M. Shatsky, R. Nussinov and H. J. Wolfson (2007). "EMatch: Discovery of high resolution structural homologues of protein domains in intermediate resolution cryo-EM maps." Ieee-Acm Transactions on Computational Biology and Bioinformatics 4(1): 28-39.

- Martinek, V., U. Bren, M. F. Goodman, A. Warshel and J. Florian (2007). "DNA polymerase beta catalytic efficiency mirrors the Asn279-dCTP H-bonding strength." Febs Letters 581(4): 775-780.

- Suresh, A. and C. Verma (2006). "Modelling study of dimerization in mammalian defensins." Bmc Bioinformatics 7.

- Knaggs, M. H., F. R. Salsbury, M. H. Edgell and J. S. Fetrow (2007). "Insights into correlated motions and long-range interactions in CheY derived from molecular dynamics simulations." Biophysical Journal 92(6): 2062-2079.

- Luo, J. and T. C. Bruice (2007). "Low-frequency normal modes in horse liver alcohol dehydrogenase and motions of residues involved in the enzymatic reaction." Biophysical Chemistry 126(1-3): 80-85.

- Dryden, I. L., J. D. Hirst and J. L. Melville (2007). "Statistical analysis of unlabeled point sets: Comparing molecules in chemoinformatics." Biometrics 63(1): 237-251.

- Dittrich, M., J. Yu and K. Schulten (2007). "PcrA helicase, a molecular motor studied from the electronic to the functional level." Atomistic Approaches in Modern Biology: From Quantum Chemistry to Molecular Simulations 268: 319-347.

- Bond, P. J. and M. S. P. Sansom (2007). "Bilayer deformation by the Kv channel voltage sensor domain revealed by self-assembly simulations." Proceedings of the National Academy of Sciences of the United States of America 104(8): 2631-2636.

- Belevich, I., D. A. Bloch, N. Belevich, M. Wikstrom and M. I. Verkhovsky (2007). "Exploring the proton pump mechanism of cytochrome c oxidase in real time." Proceedings of the National Academy of Sciences of the United States of America 104(8): 2685-2690.

- Karain, W. I., B. Ajarmah and N. I. Qaraeen (2007). "The dynamics of inter-residue distances in bovine pancreatic trypsin inhibitor." Physica a-Statistical Mechanics and Its Applications 376: 394-400.

- Inuzuka, T., T. Fujisawa, H. Arimoto and D. Uemura (2007). "Molecular shape of palytoxin in aqueous solution." Organic & Biomolecular Chemistry 5(6): 897-899.
- Bulo, R. E., L. Siggel, F. Molnar and H. Weiss (2007). "Modeling of bovine Type-I collagen fibrils: Interaction with pickling and retanning agents." Macromolecular Bioscience 7(2): 234-240.
- Kim, K., M. E. McCully, N. Bhattacharya, B. Butler, D. Sept and J. A. Cooper (2007). "Structure/function analysis of the interaction of phosphatidylinositol 4,5-bisphosphate with actin-capping protein - Implications for how capping protein binds the actin filament." Journal of Biological Chemistry 282(8): 5871-5879.
- Richards, A. D. and A. Rodger (2007). "Synthetic metallomolecules as agents for the control of DNA structure." Chemical Society Reviews 36(3): 471-483.

Papers citing BioCoRE: A literature search in April 2007 of the ISI Web of Science citation database yielded the following citations of the BioCoRE origin paper [50]:

- Sild S, Maran U, Lomaka A, Karelson M. (2006). "Open computing grid for molecular science and engineering." Journal of Chemical Information and Modeling 46(3): 953-959.
- Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, K. Schulten. (2005). "Scalable molecular dynamics with NAMD." Journal of Computational Chemistry 26(16): 1781-1802.
- Dittrich, M., S. Hayashi, K. Schulten. (2004). "ATP hydrolysis in the beta(TP) and beta(DP) catalytic sites of F-1-ATPase." Biophysical Journal 87(5): 2954-2967.
- Fudos, I. and I. Kyriazis (2004). "Thin client access to a visualization environment." Computational Science - Iccs 2004, Proceedings. 3039: 258-263.
- Dittrich, M., S. Hayashi, et al. (2003). "On the mechanism of ATP hydrolysis in F-1-ATPase." Biophysical Journal 85(4): 2253-2266.
- Phillips, R., M. Dittrich, K. Schulten. (2002). "Quasicontinuum representations of atomic-scale mechanics: From proteins to dislocations." Annual Review of Materials Research 32: 219-233.
- Finholt, T. A. (2002). "Collaboratories." Annual Review of Information Science and Technology 36: 73-107.

Resource Facilities and Computational Expertise

A total of 121 researchers have used the Resource's computational facilities in the last year (57 local, 64 remote). The Resource experienced an increase of 125% in shared file storage space (from 20.0 to 45.0 TB). External supercomputer time has again been allocated, doubling the Resource's scaled compute power from its level in April 2006.

Resource knowledge of visualization solutions, large-memory computers, web utilization, and computational clusters has become an asset to the biomedical community at large via consultations with Resource members. Researchers and organizations have requested and received advice from the Resource for the software and hardware development of their local facilities. Consultations from late April 2006 to early March 2007 are listed below:

- Computational Electronics Group, University of Illinois at Urbana-Champaign (UIUC) (cluster wiring)
- voicesignal.com (Athlon cluster building)
- Library and Information Sciences, (UIUC) (clusters)
- University of Texas at Austin (visualization clusters)
- University of Pittsburgh (clusters)
- Howard Hughes Medical Institute (3D Facility)

Software Application Website Popularity

The appeal and usability of the Resource web site continues to bring in growing numbers of unique visitors. (A visitor is defined as an individual machine accessing a web page on our site; note that this is a much more conservative and accurate method of measuring web traffic than mere web hits.)

In the past year (April 2006 - March 2007) the software sections on our web site showed substantial visitor traffic, as depicted in Table 1.

	Total	Month Avg.
VMD	229,351	19,112
NAMD	104,376	8,698
BioCoRE	29,569	2,464

Table 1: Application web site visits

Further Access

Below we report additional access activities by the Resource. The Resource trained visiting scientists, provided user support, and conducted workshops that provided training on Resource software and computational cluster development.

- *Visitor Program*
  The Resource visitor program invites members of the biomedical community to come to the Resource and get training on Resource software, and the expert analysis of Resource members for scientific research problem of interest to the visitor. From April 2006 to March 2007, the Resource has hosted 17 vistors[†]. Visitors typically fund their visits, while the Resource contributes computing resources, facilities, and local expertise.

- *User Support*
  The Resource strives to release code of high quality, and to distribute bug-free software to the user community. Assisting use in assuring the integrity and reliability of our software is a local prototyping phase, in which Resource members make use of early releases of code and provide feeback to developers before broader release occurs. In terms of providing support to the continually expanding external user community (over 110,000 users)[‡], support is a major undertaking, and taken very seriously by the Resource. Our support guidelines call for the programmers to respond to all support inquiries within 48 hours of receipt or the next business day. Nontrivial inquiries may take longer, though we strive to respond within three business days.

- *Workshops*
  The Resource presented 5 workshops since May 2006, including two on-site and one online 'hands-on' workshops, and two cluster-building workshops, as described below:

  – February 22-27, 2007, Online Workshop on Computational Biophysics: "Computer Simulation of Biological and Artificial Membrane Channels"[†]. An online workshop, hosted from the Resource web site, involving participants in viewing a streaming lecture and completing one or more tutorials on their own, with e-mail help from a teaching assistant and conference calls with the instructor. The workshop provided a brief introduction to modeling and simulation of molecular systems (in particular, molecular dynamics simulation as used

[†]URL:`http://www.ks.uiuc.edu/Overview/People/visitor.cgi`

[‡]Based on total number of dowloads of VMD and NAMD, and registered BioCoRE users

[†]URL:`http://www.ks.uiuc.edu/Training/Workshop/Online/OC1/index.html`

to investigate the molecular mechanism of function and structure-function relationship of a wide range of biological macromolecules), utilizing membrane channels as an example of how theoretical biophysical methods and computer simulation technology can be applied to biological problems.

- November 30 - December 1, 2006, "Cluster Building Workshop"[‡]. Held at the Beckman Institute, UIUC, the workshop helped users and system administrators specify, design, build, and deploy PC Clusters running Linux, and even determine if a cluster is right for a specific application.
- November 14-16, 2006, "Hands-on Workshop on Computational Biophysics,"[§], hosted by the Centro de Bioinformática y Simulación Molecular, in Talca, Chile, at the workshop participants learned how to simulate biological and synthetic membrane channels, stretch proteins, make publication quality images and movies, and study their favorite biomolecules.
- November 6-9, 2006, "Hands-on Workshop on Computational Biophysics,"[¶] hosted by the University of Pittsburgh Department of Structural Biology, in Pittsburgh, PA. Participants learned how to stretch proteins, pull water through molecular channels, mine genomic data, and study their favorite biomolecules.
- April 20-21, 2006, "Cluster Building Workshop"[‖]. Held at the Beckman Institute, UIUC, the workshop helped users and system administrators specify, design, build, and deploy PC Clusters running Linux, and even determine if a cluster is right for a specific application.

Seminars 2006-2007

Between April 2006 and April 2007 the Resource organized and hosted 16 seminars. An established institution on the UIUC campus, Resource seminars benefit students and faculty from the Beckman Institute as well as other departments and institutions. Using financial support from the Beckman Institute and our NIH Resource grant, leading scientists from around the country and around the world are brought to the Beckman Institute to present their work. Resource members also present seminars on occasion. The seminars and their respective abstracts are all posted on the Resource web site[**] and

[‡]URL:http://www.ks.uiuc.edu/Training/Workshop/Cluster4/index.html

[§]URL:http://www.ks.uiuc.edu/Training/Workshop/Talca/index.html

[¶]URL:http://www.ks.uiuc.edu/Training/Workshop/Pittsburgh/index.html

[‖]URL:http://www.ks.uiuc.edu/Training/Workshop/Cluster4/index.html

[**]URL:http://www.ks.uiuc.edu/Services/Seminar/

are also announced on the main page of the Resource website for greater publicity. Below is a list of the Resource seminars in the past year (late April 2006 - April 2007):

- April 23, 2007, Professor Benoit Roux, The University of Chicago, Chicago, IL, *Free Energy Landscape, Conformational Changes, and the Transduction of Biological Signals*
- March 12, 2007, Professor Kenneth W. Olsen, Loyola University Chicago, Chicago, IL, *Simulations of Ligand Binding to Truncated Hemoglobins*
- February 19, 2007, Dr. Margaret Ahmad, Pennsylvania State University, Delaware County Campus, Media, PA, *Magnetic Intensity Affects Cryptochrome-dependent Responses in Arabidopsis Thaliana*
- December 11, 2006, Dr. Kakoli Mitra, HHMI at Wadsworth Center, Albany, NY, *Understanding the Mechanisms of Translation by the Ribosome and Protein Translocation at Membranes*
- December 4, 2006, Professor Yong Duan, University of California, Davis, Davis, CA, *All-atom Modeling of Protein Folding and Aggregation*
- November 28, 2006, Dr. Aaron Oakley, Australian National University, Canberra, Australia, *Michaelis Complex Formation in Enzymes that Hydrolyse Carbohydrate Polymers: Steered MD through Mapping into an Alternate Coordinate Space*
- November 27, 2006, Professor Qiang Cui, University of Wisconsin, Madison, WI, *Mechanochemical Coupling in "Molecular Machines" - Insights from Molecular Simulations at Multiple Scales*
- November 9, 2006, Dr. Lloyd Demetrius, Harvard university, Cambridge, MA, *The Origin of Allometric Scaling Laws in Biology*
- November 1, 2006, Dr. Glenn Martyna, IBM TJ Watson Laboratory, Yorktown Heights, NY, *Treating Manybody Polarization and Manybody Dispersion in Complex Systems: The Quantum Drude Oscillator Formalism*
- September 27, 2006, Professor Jon Widom, Northwestern University, Evanston, IL,. *A Genomic Code for Nucleosome Positioning*
- August 10, 2006, Professor Marcus Elstner, University of Braunschweig, Germany, *The Mechanism of Spectral Shift and Color tuning in Rhodopsins*
- July 31, 2006, Professor Neil Hunter, University of Sheffield, United Kingdom, *The Structure and Organization of Bacterial Photosynthetic Complexes*

- July 28, 2006, Dr. Sameer Kumar, IBM T J Watson Research Center, Yorktown Heights, NY, *Achieving Strong Scaling with NAMD on Bluegene/L*
- June 28, 2006, Professor Willy Wriggers, School of Health Information Sciences and Institute of Molecular Medicine, University of Texas Health Science Center at Houston, TX, *Pleiomorphism of Supramolecular Assemblies*
- June 7, 2006, Mr. Grischa Meyer, School of Biomedical Sciences, University of Queensland, Brisbane, Australia, *Molecular Dynamics Study of MscL Interactions with a Curved Lipid Bilayer*
- April 24, 2006, Joel R. Stiles, Pittsburgh Supercomputing Center, Carnegie Mellon University, Pittsburgh, PA, *Novel Microphysiological Design Principles Predicted from Spatially Realistic Monte Carlo Simulations*

Awards, Honors, and Special Recognitions

There are no items to list for the current year.

Dissemination

Broad-scale efforts in dissemination and outreach through the last year took advantage of a variety of available traditional and electronic delivery mechanisms, including: distribution of Resource-produced papers and know-how via the web site; talks, meetings, workshops, and conferences; software distribution; brochure production; news stories and press releases; and use of Resource images and movies in a variety of third-party publications and academic presentations. Specific accomplishments in dissemination over the last year include:

- 41 articles in refereed journals or other publications
- 808,000 unique visitors to the Resource web site
- 56 talks by Resource Primary Investigators and 32 presentations by other members
- 34 news stories about the Resource in various media outlets
- 96 requests to use Resource images or movies from external publishers or presenters

Following in sections below are details of the Resource's dissemination efforts.

Publications

Below is a list of articles by Resource members and collaborators published over the last year, followed by a list of nine publications currently in press.

- Andrew Aird, Jörg Wrachtrup, Klaus Schulten, and Carsten Tietz. Possible pathway for ubiquinone shuttling in *R. rubrum* revealed by molecular dynamics simulation. *Biophysical Journal*, 92:23-33, 2007.
- Hanning Chen, Boaz Ilan, Yujie Wu, Fangqiang Zhu, Klaus Schulten, and Gregory A. Voth. Charge delocalization in proton channels, I. The aquaporin channels and proton blockage. *Biophysical Journal*, 92:46-60, 2007.
- Timothy A. Isgro and Klaus Schulten. Association of nuclear pore FG-repeat domains to NTF2 import and export complexes. *Journal of Molecular Biology*, 366:330-345, 2007.
- Amy Y. Shih, Peter L. Freddolino, Anton Arkhipov, and Klaus Schulten. Assembly of lipoprotein particles revealed by coarse-grained molecular dynamics simulations. *Journal of Structural Biology*, 157:579-592, 2007.
- Ilia A. Solov'yov, Danielle E. Chandler, and Klaus Schulten. Magnetic field effects in *Arabidopsis thaliana* cryptochrome-1. *Biophysical Journal*, 92:2711-2726, 2007.

- Marcos Sotomayor, Valeria Vasquez, Eduardo Perozo, and Klaus Schulten. Ion conduction through MscS as determined by electrophysiology and simulation. *Biophysical Journal*, 92:886-902, 2007.
- Yi Wang, Jordi Cohen, Walter F. Boron, Klaus Schulten, and Emad Tajkhorshid. Exploring gas permeability of cellular membranes and membrane channels with molecular dynamics. *Journal of Structural Biology*, 157:534-544, 2007.
- Anton Arkhipov, Peter L. Freddolino, Katsumi Imada, Keiichi Namba, and Klaus Schulten. Coarsegrained molecular dynamics simulations of a rotating bacterial flagellum. *Biophysical Journal*, 91:4589-4597, 2006.
- Anton Arkhipov, Peter L. Freddolino, and Klaus Schulten. Stability and dynamics of virus capsids described by coarse-grained modeling. *Structure*, 14:1767-1777, 2006.
- Jordi Cohen, Anton Arkhipov, Rosemary Braun, and Klaus Schulten. Imaging the migration pathways for O_2, CO, NO, and Xe inside myoglobin. *Biophysical Journal*, 91:1844-1857, 2006.
- Eduardo R. Cruz-Chu, Aleksei Aksimentiev, and Klaus Schulten. Water-silica force field for simulating nanodevices. *Journal of Physical Chemistry B*, 110:21497-21508, 2006.
- Markus Dittrich and Klaus Schulten. PcrA helicase, a prototype ATP-driven molecular motor. *Structure*, 14:1345-1353, 2006.
- Markus Dittrich, Jin Yu, and Klaus Schulten. PcrA helicase, a molecular motor studied from the electronic to the functional level. *Topics in Current Chemistry*, 268:319-347, 2006.
- Peter L. Freddolino, Anton S. Arkhipov, Steven B. Larson, Alexander McPherson, and Klaus Schulten. Molecular dynamics simulations of the complete satellite tobacco mosaic virus. *Structure*, 14:437-449, 2006.
- Peter L. Freddolino, Markus Dittrich, and Klaus Schulten. Dynamic switching mechanisms in LOV1 and LOV2 domains of plant phototropins. *Biophysical Journal*, 91:3630-3639, 2006.
- Mu Gao and Klaus Schulten. Onset of anthrax toxin pore formation. *Biophysical Journal*, 90:3267-3279, 2006.
- Mu Gao, Marcos Sotomayor, Elizabeth Villa, Eric Lee, and Klaus Schulten. Molecular mechanisms of cellular mechanics. *Physical Chemistry - Chemical Physics*, 8:3692-3706, 2006.

- Maria L. Ghirardi, Jordi Cohen, Paul King, Klaus Schulten, Kwiseon Kim, and Michael Seibert. [FeFe]-hydrogenases and photobiological hydrogen production. In Lionel Vayssieres, editor, *Solar hydrogen and Nanotechnology, volume 6340 of Proceedings of the Society of Photo-Optical Instrumentation Engineers, pp. 253-258*, 2006.

- Michelle Gower, Jordi Cohen, James Phillips, Richard Kufrin, and Klaus Schulten. Managing biomolecular simulations in a grid environment with NAMD-G. In *Proceedings of the 2006 TeraGrid Conference*, 2006.

- Maria E. Gracheva, Aleksei Aksimentiev, and Jean-Pierre Leburton. Electrical signatures of single-stranded DNA with single base mutations in a nanopore capacitor. *Nanotechnology*, 17:3160-3165, 2006.

- Jerome Henin, Klaus Schulten, and Christophe Chipot. Conformational equilibrium in alanine-rich peptides probed by reversible stretching simulations. *Journal of Physical Chemistry B*, 110:16718-16723, 2006.

- Michael Hoffmann, Marius Wanko, Paul Strodel, Peter H. Koenig, Thomas Frauenheim, Klaus Schulten, Walter Thiel, Emad Tajkhorshid, and Marcus Elstner. Color tuning in rhodopsins: the mechanism for the spectral shift between bacteriorhodopsin and sensory rhodopsin II. *Journal of the American Chemical Society*, 128:10808-10818, 2006.

- Fatemeh Khalili-Araghi, Emad Tajkhorshid, and Klaus Schulten. Dynamics of K+ ion conduction through Kv1.2. *Biophysical Journal*, 91:L72-L74, 2006.

- Paul W. King, Drazenka Svedruzic, Jordi Cohen, Klaus Schulten, Michael Seibert, and Maria L. Ghirardi. Structural and functional investigations of biological catalysts for optimization of solardriven, H2 production systems. In Lionel Vayssieres, editor, *Solar Hydrogen and Nanotechnology, volume 6340 of Proceedings of the Society of Photo-Optical Instrumentation Engineers, pp. 259-267*, 2006.

- Peter H. König, Nilanjan Ghosh, Michael Hoffmann, Marcus Elstner, Emad Tajkhorshid, Thomas Frauenheim, and Qiang Cui. Towards theoretical analysis of long-range proton transfer kinetics in biomolecular pumps. *Journal of Physical Chemistry A*, 110:548-563, 2006.

- Grischa Raphael Meyer, Justin Gullingsrud, Klaus Schulten, and Boris Martinac. Molecular dynamics study of bilayer deformation forces on MscL structure. *Biophysical Journal*, 91:1630-1637, 2006.

- Eileen Puklin-Faucher, Mu Gao, Klaus Schulten, and Viola Vogel. How the headpiece hinge angle is opened: new insights into the dynamics of integrin activation. *Journal of Cell Biology*, 175:349-360, 2006.

- Marcos Sotomayor, Trudy A. van der Straaten, Umberto Ravaioli, and Klaus Schulten. Electrostatic properties of the mechanosensitive channel of small conductance MscS. *Biophysical Journal*, 90:3496-3510, 2006.

- Susanna Törnroth-Horsefield, Yi Wang, Kristina Hedfalk, Urban Johanson, Maria Karlsson, Emad Tajkhorshid, Richard Neutze, and Per Kjellbom. Structural mechanism of plant aquaporin gating. *Nature*, 439:688-694, 2006.

- Ying Yin, Morten Ø. Jensen, Emad Tajkhorshid, and Klaus Schulten. Sugar binding and protein conformational changes in lactose permease. *Biophysical Journal*, 91:3972-3985, 2006.

- Jin Yu, Taekjip Ha, and Klaus Schulten. Structure-based model of the stepping motor of PcrA helicase. *Biophysical Journal*, 91:2097-2114, 2006.

- Jin Yu, Andrea J. Yool, Klaus Schulten, and Emad Tajkhorshid. Mechanism of gating and ion conductivity of a possible tetrameric pore in Aquaporin-1. *Structure*, 14:1411-1423, 2006.

The following nine articles are currently in press:

- James Gumbart, Michael C. Wiener, and Emad Tajkhorshid. Mechanics of force propagation in TonB-dependent outer membrane transport. *Biophysical Journal*, 2007.

- Morten Ø. Jensen, Ying Yin, Emad Tajkhorshid, and Klaus Schulten. Sugar transport across lactose permease probed by steered molecular dynamics. *Biophysical Journal*, 2007.

- Bryan J. Johnson, Jordi Cohen, Richard W. Welford, Arwen R. Pearson, Klaus Schulten, Judith P. Klinman, and Carrie M. Wilmot. Exploring molecular oxygen pathways in Hanseluna Polymorpha copper-containing amine oxidase. *Journal of Biological Chemistry*, 2007.

- Ioan Kosztin and Klaus Schulten. Molecular dynamics methods for bioelectronic systems in photosynthesis. In Thijs Aartsma and Joerg Matysik, editors, *Biophysical Techniques in Photosynthesis*, 2nd Ed. Kluwer Academic Publishers, New York, 2007.

- Eric H. Lee, Jen Hsin, Olga Mayans, and Klaus Schulten. Secondary and tertiary structure elasticity of titin Z1Z2 and a titin chain model. *Biophysical Journal*, 2007.
- Marcos Sotomayor and Klaus Schulten. Single molecule experiments in vitro and in silico. *Science*, 2007.
- Qian Zhao, Grigori Sigalov, Valentin Dimitrov, Brian Dorvel, Utkur Mirsaidov, Steve Sligar, Aleksei Aksimentiev, and Greg Timp. Detecting SNPs using a synthetic nanopore. *Nano Letters*, 2007.
- Amy Y. Shih, Peter L. Freddolino, Stephen G. Sligar, and Klaus Schulten. Disassembly of nanodiscs with cholate. Nano Lett., 2007. In press.
- Amy Y. Shih, Anton Arkhipov, Peter L. Freddolino, Stephen G. Sligar, and Klaus Schulten. Mechanism of discoidal high-density lipoprotein assembly. J. Phys. Chem. B, 2007. In press.

Web Site Design and Popularity

Those visiting the Resource web site starting in early 2007 encountered a new layout of the front page*, based on a design produced and refinded by Resource members, which built upon an earlier design produced by professional web desingers. New 'Group Overview' and 'Community' sections have been added, with the former grouping together information that provides an overview of the Resource (e.g., mission statement, literature reviews, brochures, member information) while the latter section emphasizes changing, current information (news stories, announcements, upcoming seminars).

The amount of traffic to the Resource website, as well as links to the web site from other groups, are telling indicators of the success of Resource outreach efforts. Details on visits and links to the site are provided below.

There have been 807,986 unique visitors to the Resource web site, an average of 67,332 per month during the April 2006 - March 2007 period; visits in March 2007 alone resulted in 195 gigabytes of data transfer (from downloaded pages, images, and files within the site). The most visited sections of the web site are shown in Table 2.

An example service found at the Resource web site is the publications database†, which provides visitors with a searchable database of Resource publications, including searches by title, author(s), journal, subject, year ranges, and fulltext searching. Over 23,600 unique visitors downloaded at least one file copy of an article using the database over the last year.

*URL:`http://www.ks.uiuc.edu/`

†URL:`http://www.ks.uiuc.edu/Publications/Papers/`

	Total Visitors	Visitors per Month
VMD	229,351	19,112
NAMD	104,376	8,698
BioCoRE	29,569	2,464
Other Research	126,189	10,515
Galleries	57,013	4,758
Papers	38,303	3,191
Seminars	5,589	465

Table 2: Web site visitors from April 2006 – March 2007

A recent Google search (April, 2007) yielded the following statistics regarding external sites that link to the main Resource web page: 919 pages link to the main page, registered under 46 different domains. Sample education, scientific resources, and programming or computing-related sites with links to the Resource web site are provided below:

Education:

- *Cornell University: Department of Physiology and Biophysics, and Medical Library*
  physiology.med.cornell.edu, library.med.cornell.edu
- *University of California at San Diego: Keck Laboratory for Integrated Biology, the Electron Microscopy Outreach Program, and Department of Computer Science and Engineering*
  keck2.ucsd.edu, em-outreach.ucsd.edu, www.cse.ucsd.edu
- *University of California at Berkeley: Computer Science Department*
  www.cs.berkeley.edu
- *Purdue University: The Post Lab*
  post.bio.purdue.edu
- *Scripps Research Institute: Amber Molecular Dynamics*
  amber.scripps.edu
- *Stanford University: Helix Group*
  helix-web.stanford.edu
- *University of Pennsylvania: Center for Molecular Modeling*
  www.cmm.upenn.edu

- *University of Michigan: College of Engineering, and Center for Advanced Computing*
  www.engin.umich.edu, cac.engin.umich.edu
- *Carnegie Mellon University: University Libraries*
  www.library.cmu.edu
- *University of Washington: Civil and Environmental Engineering, and Anderson Research Group*
  www.ce.washington.edu, andersenlab.chem.washington.edu
- *Brown University: Center for Computation and Visualization*
  www.ccv.brown.edu
- *Xplora gateway: European gateway to science education*
  www.xplora.org

Scientific Resources:

- *Biophysical Journal*
  www.biophysj.org
- *The Journal of Biological Chemistry*
  www.jbc.org
- *arXiv.org e-Print archive*
  www.arxiv.org
- *Wikipedia Online Encyclopedia*
  en.wikipedia.org
- *The Foresight Institute Nanotechnology*
  www.foresight.org
- *Computational Chemistry List*
  www.ccl.net
- *American Scientist Magazine*
  www.americanscientist.org
- *Biowulf cluster at National Institutes of Health*
  biowulf.nih.gov
- *National Center for Research Resources*
  www.ncrr.nih.gov

- *About: Physics*
  physics.about.com
- *Oak Ridge National Laboratory*
  www.ornl.gov

Programming/Computing Related

- *Apple Computers: Macintosh Products Guide*
  guide.apple.com
- *Linux Online*
  www.linux.org
- *FreeBSD Online*
  www.freebsd.org
- *Adobe Acrobat User Community*
  www.acrobatusers.com

Lectures, Presentations and Posters

The Resource PIs and other members gave the following lectures, presentations, or posters over the last year[‡]:

Klaus Schulten

- April 2007, Urbana, IL, Beckman Institute, Beckman Foundation Board Meeting 2007, *The Next Revolution in Scientific Computing*
- March 2007, Arlington, VA, National Science Foundation, New Frontiers in Dynamic Systems Workshop, *Brain and Brawn in Nanosystems Modeling*
- February 2007, Chicago, IL, University of Chicago, Department of Physics, *Atomic Level Resolution of Cellular Organelles - The Case of Photosynthetic Chromatophores*
- January 2007, Urbana, IL, University of Illinois, Department of Mechanical Science and Engineering, *Molecular Mechanisms Underlying the Mechanics of Living Cells*
- December, 2006, Urbana, IL, Beckman Institute, Presentation to UIUC Provost Linda Katehi
  - *Overview and Computational Microscopy Mission*

[‡]URL:`http://www.ks.uiuc.edu/Publications/Lectures/lectures.cgi`

  - *Training on Computational Biophysics and Summary*

- December, 2006, Urbana, IL, Beckman Institute, Beckman Director's Seminar Series, *Turning Advanced Science into Advanced Teaching*

- December 2006, Urbana, IL, University of Illinois at Urbana-Champaign, National Center for Supercomputing Applications, *Petascale Computing in the Biosciences - Simulating Entire Life Forms*

- December 2006, Atlanta, GA, Georgia Institute of Technology, School of Biology Seminar, *Petascale Computing in the Biosciences - Simulating Entire Life Forms*

- November 2006, Pittsburgh, PA, University of Pittsburgh, Department of Structural Biology, Hands-On Workshop on Computational Biology

  - *Introduction to Protein Structure and Dynamics*
  - *Parameters for Classical Force Fields*

- November 2006, Urbana, IL, Physics Colloquium, University of Illinois at Urbana-Champaign, *From the Atom to the Cell - The Multi-scale Nature of Living Systems*

- November 2006, Urbana, IL, University of Illinois at Urbana-Champaign, Department of Physics, Saturday Physics Honor Program, *Physics of the Body*

- November 2006, Urbana, IL, University of Illinois at Urbana-Champaign, Department of Physics, Physics Colloquium, *Physics - Grasping the Multi-scale Nature of Biological Systems*

- November 2006, Bochum, Germany, Bochum University Conference Center, *PcrA Helicase, a Molecular Motor studied from the Electronic to the Functional Level*

- October 2006, Munich, Germany, Center for Nanoscience, Ludwig Maximilian University, *The Role of Molecular Modeling in Bionanotechnology*

- October 2006, Munich, Germany, Department of Physics, Technical University of Munich, Germany, *Molecular Mechanisms of Cellular Mechanics*

- October 2006, Martinsried, Germany, Max Planck - Institute for Biochemistry, *Structure and Function of Chromatophores of Photosynthetic Bacteria*

- October 2006, Regensburg, Germany, University of Regensberg, Center for Sensory Photoreceptors, *Biochemical Mechanisms for Magnetic Orientation in Animals*

- October 2006, Munich, Germany, Department of Physics, Ludwig-Maximilian-University, Arnold Sommerfeld Lecture, *In Silico Single Molecule Experiments*

- September 2006, Urbana, IL, Beckman Institute, 71st Annual Meeting of the University of Illinois Foundation, Mini-symposium on "Computational Medicine", *Viruses - Seeing the Enemy up Close*
- August 2006, Arlington, VA, National Science Foundation, Workshop on Petascale Computing in the Biological Sciences, *Petascale Computing in the Biosciences - Three Projects Starting Today*
- July 2006, Munich, Germany, Euroscience Open Forum 2006, *Clustering of Function in Biological Cells*
- June 2006, New London, NH, Gordon Conference on Single Molecule Approaches to Biology, *Molecular Mechanisms of Cellular Mechanics*
- May 2006, Norman, OK, University of Oklahoma, Linux Clusters Institute, *Cluster Computing in Everyday Biomedical Research: Past, Present, and Future*
- May 2006, Philadelphia, PA, University of Pennsylvania, Laboratory for Research on the Structure of Matter, *The Role of Molecular Modeling in Bionanotechnology*
- May 2006, Santa Barbara, CA, University of California, Kavli Institute for Theoretical Physics, "New Physical Approaches to Molecular and Cellular Machines" Workshop, *Proteins with Mechanical Functions*
- May 2006, Frankfurt, Germany, University of Frankfurt, Frankfurt Institute of Advanced Studies, *Recent Conceptual and Methodological Developments in Biomolecular Modeling*
- May 2006, Mainz, Germany, Mainz University, *Recent Conceptual and Methodological Developments in Biomolecular Modeling*
- May 2006, Mainz, Germany, Max Planck Institute for Polymer Science, *In Situ Modeling of Biological and Artificial Channels*
- May 2006, Frankfurt, Germany, University of Frankfurt, Deutsche Bank Lectureship, *Inate and Living Matter Biological Physics in Search of the Secret of Life*
- May 2006, Vienna, Austria, University of Vienna, Symposium "Evolution of Biomolecular Structure", dedicated to Professor Peter Schuster *What is Life? An Answer Sought from Photosynthetic Bacteria*
- May 2006, St. Louis, MO, Washington University, Institute for Complex and Adaptive Matter Workshop "Multiscale Interactions and Dynamics in Biological Systems", *Multiscale Simulation of Protein-DNA and Protein-Lipid*

- April 2006, Cambridge, MA, Harvard-MIT Seminar, *Towards Understanding Membrane Channels*

Emad Tajkhorshid

- April 2007, Urbana, IL, Beckman Institute, Beckman Foundation Board Meeting 2007, *Visualizing the Molecular Water Pipes of the Human Body at Work*
- February 2007, Urbana, IL, Beckman Institute, Beckman Open House 2007, Computational Microscope: Visualizing Nanodevices at Work, Presentation: *Molecular Water Pipes in the Human Body*
- December, 2006, Urbana, IL, Beckman Institute, Presentation to UIUC Provost Linda Katehi, *Computational Microscopy of Biological Water Channels*
- November 2006, Pittsburgh, PA, University of Pittsburgh, Department of Structural Biology, Hands-On Workshop on Computational Biology
  - *Parameters for Classical Force Fields*
  - *Simulating Membrane Channels*
- September 2006, Urbana, IL, Beckman Institute, 71st Annual Meeting of the University of Illinois Foundation, Mini-symposium on "Computational Medicine", *Molecular Water Channels Keep You Healthy*
- March 2006, Gomadingen, Germany, German Biophysical Society, International Workshop on "Dynamics of Membranes", *Visualizing the Art of Selective Transport in Membrane Channels at Full Atomic Resolution*
- March 2006, West Lafayette, IN, Purdue University, *Visualizing the Art of Selective Transport in Membrane Channels at Full Atomic Resolution*
- March 2006, Frankfurt, Germany, Max Planck Institute of Biophysics, *Unraveling Molecular Mechanisms of Permeation, Selectivity, and Gating of Membrane Channels at Full Atomic Resolution*

Alek Aksimentiev

- April 2007, Urbana, IL, Beckman Institute, Beckman Foundation Board Meeting 2007, *Silicon Nanopores for Personal Genomics*
- February 2007, Urbana, IL, Beckman Institute, Beckman Open House 2007, Computational Microscope: Visualizing Nanodevices at Work, *Transistors with Holes for Sequencing DNA*

- December 2006, Urbana, IL, Beckman Institute, Presentation to UIUC Provost Linda Katehi, *Computational Microscopy of Synthetic Nanopores for DNA Sequencing*
- November 2006, Talca, Chile, University of Talca, Centro de Bioinformtica y Simulacin Molecular, Hands-On Workshop on Computational Biology
  - *Introduction to Biomolecular Modeling with VMD and NAMD*
  - *All-atom Modeling of Membrane Proteins*
  - *Advanced Topics: Custom Forces and Bionanotechnology*
- November 2006, Talca, Chile, University of Talca, Centro de Bioinformtica y Simulacin Molecular, *Present and Future of Biomolecular Simulations*
- November 2006, Urbana, IL, Department of Physics, Condensed Matter Seminar, *Toward Sequencing DNA with a Synthetic Nanopore*
- October 2006, Urbana, IL, Beckman Institute, Nanohour Lecture Series, *Toward Sequencing DNA with a Synthetic Nanopore*
- October 2006, Fayetteville, AR, University of Arkansas, Department of Physics, Colloquium, *Toward Sequencing DNA with a Synthetic Nanopore*
- September 2006, Urbana, IL, Beckman Institute, 71st Annual Meeting of the University of Illinois Foundation, Mini-symposium on "Computational Medicine", *Tomorrow's DNA Sequencing*
- September 2006, Aussois, France, Gordon Research Conference on Bioelectrochemistry, *Imaging DNA Translocation through Nanopores with Molecular Dynamics*
- September 2006, Liege, Belgium, Third Focused Workshop on Electronic Recognition of Bio-molecules, *Forcing Biomolecules through Synthetic and Biological Nanopores*
- May 2006, Urbana, IL, Center for Nanoscale Science and Technology, *Atomic Resolution Imaging of Nanodevices with Large Scale Molecular Dynamics*

Other TCB members (includes meetings attended and poster sessions):

- April 2007, Urbana, IL, Beckman Institute, Beckman Student Seminar Series, *Stability and Dynamics of Viruses Described by Computer Simulations* (Anton Arkhipov)
- April 2007, Urbana, IL, Beckman Institute, Beckman Foundation Board Meeting 2007

  - *Using Algae to Produce Affordable Hydrogen Energy* (Jordi Cohen)
  - *Visualizing How Viruses Infect Cells* (Peter Freddolino and Anton Arkhipov)

- February 2007, Baltimore, MD, Biophysical Society 51st Annual Meeting (poster sessions)
  - *Tertiary and Secondary Structure Elasticity of Repeat Proteins* (Marcos Sotomayor, Klaus Schulten)
  - *Magnetic Field Effects in Arabidopsis Thaliana Cryptochrome-1* (Ilia A. Solov'yov, Danielle Chandler, Klaus Schulten)
  - *Aquaporin-mediated Gas Conduction across Biological Membranes* (Yi Wang, Jordi Cohen, Walter F. Boron, Klaus Schulten, Emad Tajkhorshid)
  - *Simulation of Protein Translocation in the Bacterial Flagellum* (Zhongzhou Chen, Peter Freddolino, Anton Arkhipov, Klaus Schulten)
  - *Coarsed-grained Molecular Dynamics Simulations of Rotational-induced Structural Transitions in the Bacterial Flagellum* (Peter L. Freddolino, Anton Arkhipov, Klaus Schulten)
  - *Ca+2 Mediated Membrane Association of the GLA Domain to Anionic Lipid Bilayers* (Y. Zenmei Ohkubo, Emad Tajkhorshid)
  - *Molecular Dynamics Studies of Ionic Conductance through Silica Nanopores* (Eduardo R. Cruz-Chu, Aleksei Aksimentiev, Klaus Schulten)
  - *Dynamics of K+ Conduction through Kv1.2* (Fatemeh Khalili-Araghi, Emad Tajkhorshid, Klaus Schulten)
  - *Sugar Transport across Lactose Permease Probed by Molecular Dynamics Simulations* (Ying Yin, Morten Jensen, Emad Tajkhorshid, Klaus Schulten)
  - *Ion Conduction through MscS as Determined by Electrophysiology and Simulation* (Marcos Sotomayor, Valeria Vasquez, Eduardo Perozo, Klaus Schulten)
  - *Assembly of Lipoproteins Revealed by Coarse-grained Molecular Dynamics Simulations* (Amy Y. Shih, Peter Freddolino, Anton Arkhipov, Stephen G Sliger, Klaus Schulten)
  - *Disassembly of Nanodiscs: Using Coarse-grained MD for Insight into Self-assembly* (Amy Y. Shih, Peter Freddolino, Stephen G Sligar, Klaus Schulten)
  - *Substrate Binding and Translocation in the Leucine Transporter* (Leyla Celik, Birgit Schiott, Emad Tajkhorshid)
  - *Modeling and Simulations of a Bacterial Ribosome* (Leonardo Trabuco, Emma Falck, Elizabeth Villa, Klaus Schulten)

  - *Opening of the Lateral Gate of the Translocon SecY* (James Gumbart, Klaus Schulten)
  - *Secondary and Tertiary Structure Elasticity of Titin Z1Z2 and the Titin Chain* (Eric H. Lee, Jen Hsin, Olga Mayans, Klaus Schulten)

- February 2007, Urbana, IL, Beckman Institute, Beckman Open House 2007, Computational Microscope: Visualizing Nanodevices at Work
  - *Good Cholesterol* (Amy Shih)
  - *Views of Viral Infection* (Peter Freddolino and Anton Arkhipov)
  - *Making Renewable Hydrogen Energy* (Jordi Cohen)
  - *Proteins that Wrestle DNA* (Elizabeth Villa)

- January 2007, Urbana, IL, Beckman Institute, Nanohour Lecture Series, *Exploring Gas Permeability and Gating Mechanism of Aquaporins* (Yi Wang)

- January 2007, Urbana, IL, Beckman Institute, Graduate Student Seminar Series, *Computational Studies of Nuclear Pore Complex Transport Dynamics* (Tim Isgro)

- December 2006, Urbana, IL, Beckman Institute, Presentation to UIUC Provost Linda Katehi
  - *Computational Microscopy for Photosynthetic Hydrogen Gas Production* (Jordi Cohen)
  - *Computational Microscopy for Fighting Cardiovascular Diseases* (Amy Shih)
  - *Computational Microscopy for Fighting Viral Infections* (Peter Freddolino)

- November-December 2006, Urbana, IL, Beckman Institute, TCBG Cluster Building Workshop
  - *Linux Clusters for High-Performance Computing: An Introduction* (Tim Skirvin, Jim Phillips)
  - *Linux Clusters: Details and Case Studies* (Jim Phillips, Tim Skirvin)

- April 2006, UIUC, Urbana, IL, Beckman Institute, TCBG Cluster Building Workshop
  - *Linux Clusters for High-Performance Computing: An Introduction* (Tim Skirvin, Jim Phillips)
  - *Linux Clusters: Details and Case Studies* (Jim Phillips, Tim Skirvin)

Media Coverage

The Resource has been featured in a variety of printed and online media for its scientific and computational accomplishments over the last year. Media coverage includes stories on Resource research into hydrogen gas produced by light-absorbing algae[§], how the nuclear pore complex regulates passage of proteins into a cell's nucleus[¶], and the identification of a key gating mechanism in aquaporins[‖].

Stories on the Resource appeared in popular media, scientific journals, online news sources, and more. All news-making stories and their reprints are documented by the Resource at the "In the News" section of the web site[**]:

- Dylewski, A. (April, 2007). "SimCell," Anyone? *The Why Files - Science Behind the News (Whyfiles.org)*.
  http://whyfiles.org/shorties/230simcell/
- Brodie, C. R. (May-June, 2007). Fat Enough for Two Belts. High-Density Lipoprotein Makes Itself Into A Double-Belted Nanodisk. *American Scientist Online - Science Observer*.
  http://www.americanscientist.org/template/AssetDetail/assetid/55155
- Ricker, K., & Morgan, H. (January 10, 2007). Protein Wranglers. *NCSA NEWS*.
  http://access.ncsa.uiuc.edu/Stories/namdg/
- Ricker, K., & Morgan, H. (January 9, 2007). Protein Wranglers. *UIUC Department of Engineering news website*.
  http://www.engr.uiuc.edu/news/index.php?xId=070208160728
- Ricker, K., & Morgan, H. (January 15, 2007). Special Features: Protein Wranglers. *GRIDtoday*.
  http://www.gridtoday.com/grid/1200956.html
- Staff. (November, 2006). Pittsburgh Doubles Capability of BigBen. *PSC News Bureau*.
  http://www.psc.edu/publicinfo/news/2006/2006-11-21-bigbengrows.php
- Staff. (November, 2006). Pittsburgh Doubles Capability of BigBen. *PR Newswire*.
  http://www.prnewswire.com/cgi-bin/stories.pl?ACCT=104&STORY=/www/story/11-21-2006/0004478987&EDATE=

[§]URL:`http://access.ncsa.uiuc.edu/Stories/namdg/`

[¶]URL:`http://www.psc.edu/science/2006/schulten/`

[‖]URL:`http://www.news.uiuc.edu/news/06/0921gating.html`

[**]URL:`http://www.ks.uiuc.edu/Publications/stories.cgi`

- Schneider, M. (November, 2006). Knock, Knock, Who's There? Simulations Reveal New Information about the Gateway to the Cell Nucleus. *PSC News Bureau*. http://www.psc.edu/science/2006/schulten/
- Sidener, J. (September 25, 2006). Games Get Serious. *SignOnSanDiego.com - by the Union Tribune*. http://www.signonsandiego.com/news/tech/personaltech/20060925-9999-mz1b25seriou.html
- Sidener, J. (October 3, 2006). Games Get Serious: Video Games Can Power Molecular Research. *Paramus Post News*. http://www.paramuspost.com/article.php/20061001195653647
- Kloeppel, James E. (September 21, 2006). One Protein, Two Channels: Scientists Explain Mechanism In Aquaporins. *UIUC News Bureau*. http://www.news.uiuc.edu/news/06/0921gating.html
- Staff. (September 21, 2006). One Protein, Two Channels: Scientists Explain Mechanism In Aquaporins. *EurekAlert!.com*. http://www.eurekalert.org/pub_releases/2006-09/uoia-opt092106.php
- Staff. (September 24, 2006). Scientists Explain Mechanism In Aquaporins. *MedicalNewsToday.com*. http://www.medicalnewstoday.com/medicalnews.php?newsid=52504
- Staff. (September 27, 2006). One Protein, Two Channels: Scientists Explain Mechanism In Aquaporins. *Innovations-Report.de*. http://www.innovations-report.de/html/berichte/biowissenschaften_chemie/bericht-71116.html
- Kloeppel, James E. (September 18, 2006). Evolutionary Software to be Released Free of Charge. *UIUC News Bureau*. http://www.news.uiuc.edu/news/06/0918software.html
- Staff. (September 21, 2006). Evolutionary Software to be Released Free of Charge. *EurekAlert!.com*. http://www.eurekalert.org/pub_releases/2006-09/uoia-est091806.php
- Staff. (September, 2006). Evolutionary Software for Free. *Scenta.co.uk*. http://www.scenta.co.uk/scenta/news.cfm?cit_id=1130885&FAArea1=widgets.content_view_1
- McGaughey, S. (Summer, 2006). Aleksei Aksimentiev Finds His Place With DNA Sequencing Challenge. *Synergy*, (2), p. 2-4. http://www.beckman.uiuc.edu/news/synergy/SynergySummer2006.pdf

- Staff. (July, 2006). First Atom-by-Atom Simulation of a Life Form. *Physics Illinois News*.
  http://www.physics.uiuc.edu/research/Highlights/STMV.htm

- Staff. (June 9, 2006). The Cray XT3 - A Unique Resource at PSC. *HPCWire*.
  http://www.hpcwire.com/hpc/686730.html

- Robinson, M. (May 10, 2006). First-Ever Simulation of Functioning Organism Spawned by Ingenuity of Illinois Researchers and Power of SGI Altix. *SGI Press Release at PRNewswire*.
  http://www.prnewswire.com/cgi-bin/stories.pl?ACCT=104&STORY=/www/story/05-10-2006/0004358536&EDATE=

- Staff. (May 10, 2006). First-Ever Simulation of Functioning Organism Spawned by Ingenuity of Illinois Researchers and Power of SGI Altix. *WebWire*.
  http://webwire.com/ViewPressRel.asp?SESSIONID=&aId=13977

- Jaques, R. (May 11, 2006). Scientists Claim Atomic-Level Simulation of Living Organisms. *VNUnet*.
  http://www.vnunet.com/vnunet/news/2155804/boffins-complete-first-atomic

- Staff. (May 8, 2006). Fleeting Moments in the Life of a Virus. *Deccan Herald*.
  http://www.deccanherald.com/deccanherald/may92006/snt171916200658.asp

- Staff. (April, 2006). Harnessing Supercomputers to Model Life. *NSF Current*.
  http://www.nsf.gov/news/newsletter/apr_06/NSF_Current_April_2006.pdf

- Staff. (April 14, 2006). Modeling Enzymes at NCSA. *HPCWire*.
  http://www.hpcwire.com/hpc/620580.html

- Morgan, H., & Ricker, K. (April, 2006). Managing Workflow with NAMD-G. *NCSA Datalink*.
  http://www.ncsa.uiuc.edu/News/datalink/0604/namd-g.html

- Staff: (April 17, 2006). Managing Workflow with NAMD-G. *GRIDtoday*.
  http://www.gridtoday.com/grid/625713.html

- Robinson, M. (April 6, 2006). SGI Ranks Among Bio-IT World's Coveted Bio-IT 50. *SGI Press Release at Yahoo!Finance*.
  http://biz.yahoo.com/prnews/060406/sfth046.html?.v=51

- Staff. (April 6, 2006). SGI Named to Bio-IT 50. *HPCWire*.
  http://www.hpcwire.com/hpc/617323.html

- Davis, R. (March 30, 2006). UI Biologists Construct First Digital Model of Virus. *Daily Illini*. http://www.dailyillini.com/media/storage/paper736/news/2006/03/30/News/Ui.Biologists.Construct.First.Digital.Model.Of.Virus-1765450.shtml?norewrite200603301128&sourcedomain=www.dailyillini.com

- Casey, R. M. (March 28, 2006). Bioinformatics in Structure-Based Drug Design. *B-EYE-Network.com*. http://www.b-eye-network.com/view/2593?jsessionid=d79d5e424b85825b0ffe5b2b144f3d37

- Than, K. (March 27, 2006). Virtual Virus is First Simulation of an Entire Life Form. *LiveScience.com*. http://www.livescience.com/technology/060327_computer_virus.html

- Staff. (March 24, 2006). Partners Map Future of TeraGrid. *HPCWire*. http://www.hpcwire.com/hpc/601499.html

Outreach

Resource efforts to reach the biomedical community in addition to the aforementioned media coverage, scholarly articles, lectures, posters, and workshops are viewed as "outreach" activities, and include items such as the following:

- On-site demonstrations
- Making images and movies available for use by others
- Development and distribution of brochures
- Responding to licensing requests

Demonstrations

Visitors to the Resource (i.e., seminar speakers, visiting scientists, others) are commonly provided with a demonstration using the Resource's visualization facility. Demonstrations typically involve a staff or graduate Resource member loading a VMD-based presentation relevant to the interests of the visitor into the Resource's 3D stereo projection system, and then discussing the science an computation behind the presentation. There were 15 such presentations by Resource members over the April 2006 - March 2007 period.

Image Requests

The Resource regulary responds to requests from external scientists and other for permission to use Resource images and movies in a variety of media, including web sites,

books, papers, press articles, talks and presentations. Recent updates to the web site image gallery[††] and movie gallery[‡‡] include the addition of e-mail links for each image or movie to facilitate user requests. A trend of single requesters asking for use of multiple images and/or movies contributes the count of 96 image requests over the last year. A standard response text, written in cooperation with university intellectual property representatives, grants non-exclusive permission to image and movie requests, which protects Resource copyright while at the same time allowing for image/movie distribution.

Brochures

Resource brochures communicate information about our programs, research, and software, and are available in print and electronic formats, with the electronic versions found online at the Resource website[†]. A brochure completed over the last year is *Highlights 2004-2006*[‡], which compiles the research highlights posted two-three times a month on the Resource main web page, for dates April 2004 to May 2006. The brochure is an addendum to the prior *Recent Highlights* which compiled research highlights from January 2001 - March 2004. A brochure introducing lecture topics and lecturers was also developed as part of a "Mini-Symposium on Computational Medicine" provided for the 71st Annual Meeting of the University of Illinois Foundation, held in September 2006[§].

Licensing and Distribution

Resource software licenses, which already allow for broad use, are upon request reviewed and if needed revised to meet the needs of external groups. Such expansions are done in consultation and cooperation with the University of Illinois Office of Technology Management, who provide needed technical and legal expertise. In April 2006, for example, the Resource is reviewing NAMD licensing requests from a drug company and from a firmware distributor.

[††]URL:`http://www.ks.uiuc.edu/Gallery/Science/Structure/`

[‡‡]URL:`http://www.ks.uiuc.edu/Gallery/Movies/`

[†]URL:`http://www.ks.uiuc.edu/Gallery/Brochure/`

[‡]URL:`http://www.ks.uiuc.edu/Gallery/Brochure/highlights2006.pdf`

[§]URL:`http://www.ks.uiuc.edu/Publications/Brochures/symposium/symposium_brochure.pdf`

Patents, Licenses, Inventions, and Copyrights

No patents, licenses, inventions or copyrights have been granted to Resource Primary Investigators or other members for the current year.

Training

In the last year, the Resource has continued and expanded its training efforts, by conducting five workshops and expanding its collection of web-accessible tutorials. Such efforts are in addition to more traditional training programs for graduate students and postdoctoral researchers and instructional presentations about Resource software. Training provided by the Resource capitalized on a range of tools and media, including:

- Three workshops at national and international locations, including one online
- Five new tutorials
- One new case study and the incorporation of new technologies for case study development
- Graduate student and postdoctoral training
- Visitor program

Hands-On Workshops in Computational Biophysics

The Resource in the last year (April 2006 - April 2007) held three hands-on workshops in computational biophysics. The four to five day long workshops explored physical models and computational approaches used for the simulation of biological systems and the investigation of their function at an atomic level. The courses utilized examples including the properties of membranes and membrane proteins, mechanisms of molecular motors, and conduction through water and ion channels. Relevant physical concepts, mathematical techniques, and computational methods were introduced, including force fields and algorithms used in molecular modeling, molecular dynamics simulations on parallel computers, and steered molecular dynamics simulations. The workshops were designed for graduate students and postdoctoral researchers in computational and/or biophysical fields seeking to extend their research skills to include computational and theoretical expertise, as well as other researchers interested in theoretical and computational biophysics. Theory sessions in the morning were followed by hands-on computer labs in the afternoon in which students were able to set up and run simulations.

Pittsburgh Workshop[*]
The Molecular Modeling and Molecular Dynamics Simulation Workshop was held at the University of Pittsburgh from November 6-9, 2006. The evaluation results showed that an average of 96% of the students who responded agreed that the lectures were relevant

[*]URL:`http://www.ks.uiuc.edu/Training/Workshop/Pittsburgh/`

to their research, and an average of 95% agreed that the tutorials were relevant to their research.

Talca Workshop[†]
The Talca workshop, held November 14-16, 2006 at the Centro de Bioinformática y Simulación Molecular, was the first to utilize the Resource's five newest tutorials. Evaluation results show that 83% of students felt the workshop broadened their understanding of concepts and principles in the field of Computational and Theoretical Biophysics, and 92% said it improved their ability to carry out original research in the field of Theoretical and Computational Biophysics. Twenty-four people responded to the evaluation survey.

Online Workshop[‡]
The Resource held its first online workshop February 22-27, 2007. Twenty-six people participated from around the world. The workshop consisted of an online video of the lectures, corresponding lecture slides, and the Nanotubes tutorial. Students could email a teaching assistant with questions, and for direct communication, two conference calls were held where students could talk to the instructor. Evaluation results show that 100% of students who responded felt the workshop broadened their understanding of concepts and principles in the field of Computational and Theoretical Biophysics, and 83% said it improved their ability to carry out original research in the field of Theoretical and Computational Biophysics. Twenty-one people responded to the evaluation survey.

Cluster Building Workshops

In the last year, the Resource held two cluster-building workshops, each a day and a half long, on the design, construction, and deployment of computational clusters. In the workshops, participants hear lectures from Resource staff (Tim Skirvin, James Phillips) on clustering basics, including why clusters are useful, how they work, when they are needed, basic programming techniques, different design options, and how cluster design interacts with queuing systems. Additionally, participants build their own cluster during the workshop, using racked computers and equipment provided by the Resource. Due to the materials required, all cluster workshops have been held locally, at the Beckman Institute where the Resource is housed. Participants in the cluster workshop have been fairly evenly distributed across levels of education, from undergraduate through and past doctoral education, and the majority have had academic affiliations. Dates and evaluation results for each workshop are provided below.

- April 20-21, 2006, Cluster Building Workshop[§]

[†]URL:`http://www.ks.uiuc.edu/Training/Workshop/Talca/`

[‡]URL:`http://www.ks.uiuc.edu/Training/Workshop/Online/OC1/`

[§]URL:`http://www.ks.uiuc.edu/Training/Workshop/Cluster4/`

Evaluation results found that for a majority of the responding participants, the workshop broadened their understanding of cluster building concepts, (90%), and improved their ability to set up a cluster, (95%). A total of 22 persons attended the workshop.

- November 30-December 1, 2006, Cluster Building Workshop[¶]
  Again, a majority of the responding participants indicated that the workshop broadened their understanding of cluster building concepts, (94%), and improved their ability to set up a cluster, (94%). A total of 21 persons attended the workshop.

Tutorials

Five new tutorials have been written, first used in the Talca workshop. Some are currently being revised, and will be added to the website when completed. The tutorials include:

- *Bionanotechnology Tutorial*
  This tutorial is designed to guide users of VMD and NAMD through all the steps required to set up bionanotechnology molecular dynamics simulations.

- *Ion Conduction, Permeation, Electrostatic Maps*
  This tutorial is designed to guide users of VMD and NAMD through all the steps required to determine permeation and electrostatic properties of membrane channels. It covers the analysis of equilibrium membrane-channel simulations along with the steps needed to induce a voltage across a pore-containing membrane and the techniques used to analyze such non-equilibrium simulations.

- *VMD Images and Movies Tutorial*[‖]
  This tutorial is designed to give users of VMD an introduction to advanced techniques for making custom images and movies. The first section looks at how to use features such as resolution, color, material, depth perception, and volumetric data to produce effects and enhancements for still images. The second part demonstrates how to work with trajectories, by using techniques such as smoothing trajectories, showing multiple frames at once, and making atom selections "follow" a trajectory. It also shows how to create a movie from a trajectory using VMD's Movie Maker plugin.

- *Forces Tutorial*[**]
  This tutorial is designed to guide users of VMD and NAMD in the use of tclForces

[¶]URL:`http://www.ks.uiuc.edu/Training/Workshop/Cluster-2006.11/`

[‖]URL:`http://www.ks.uiuc.edu/Training/Tutorials/vmd-ref/imgmv/imgmv-tutorial.pdf`

[**]URL:`http://www.ks.uiuc.edu/Training/Tutorials/science/forces/forces-tutorial.pdf`

and tclBC scripts. These script-based facilities simplify the process of adding complex forces to systems and implementing boundary conditions.

- *Membrane Proteins Tutorial*
  This tutorial is designed to guide users of VMD and NAMD through all the steps required to set up a membrane protein system for molecular dynamics simulations. It covers steps required to set up a structural model of a membrane protein, then place the protein in its native membrane environment, and minimize and equilibrate the resulting system with NAMD.

Case Studies

A new case study on Ion Channels was recently added.

- *Ion Channels*[†]
  The Ion Channels case study illustrates the molecular mechanisms of ion permeation by focusing on potassium (K+) channel and chloride (Cl-) channels. The functional features of the channels, namely selectivity, permeation, and gating, are connected to their structures and dynamics through the integrated use of VMD. Additional VMD tools are also utilized throughout, such as Multiseq for exploring the structure and sequence conservation.

The Resource strives to update its training materials with any new technologies that may benefit learning. One example is Acrobat3D, a recent development by Adobe which allows interactive 3D visualizations to be placed within PDF files. Those who are in a situation where they cannot use VMD, or are quickly browsing the case study without going through the accompanying exercises can still receive the benefits of an interactive 3D visualization.

The "Water and Ice" case study has been modified to take advantage of this technology[‡], and there are plans to create versions of the other case studies as well. In the original case study, an image was used to show the layout of ice crystals from three different angles. Now, with the 3D enhancement, users can view ice crystals from those angles, and also rotate the ice crystal for a better visualization of the crystal structure.

Resource Library

The Resource library has been expanded by the purchase of 23 new books. Further, to supplement the UIUC library's collection of on-line and print journals, the Resource subscribes to the following journals in science and computing:

[†]URL:`http://www.ks.uiuc.edu/Training/CaseStudies/pdfs/channels.pdf`

[‡]URL:`http://www.ks.uiuc.edu/Research/vmd/minitutorials/acrobat3d/`

- Physics Today
- Science
- Sys Admin
- Journal of NIH Research
- C/C++ Users Journal
- Dr. Dobb's Journal
- Linux Journal
- Nature

Visitors

The Resource visitor program, for ten years, provided scientists with the opportunity to learn how to use Resource-produced software, other software hosted on Resource computers, and to benefit from the knowledge and expertise of Resource members. Visitors provide their own financial support. Resource members spend substantial amounts of time helping visitors achieve their educational and research goals. At the end of their time at the Resource, visitors acquired critical skills and new experiences that they took back to their home laboratories. Visitors during the last funding period include:

- Ilia Solov'yov, Frankfurt Institute for Advanced Studies (April 2006)
- Amitava Roy, Weill Medical College, Cornell (May 2006)
- Lea Thogersen, University of Aarhus (May 2006 and January 2007)
- Ogaga Akoroda, Illinois Wesleyan University (June 2006)
- Axel Kohlmeyer, University of Pennsylvania (July 2006)
- Ken Olsen, Loyola University, Chicago (July 2006)
- Guido Fratesi, Universita degli Studi de Milano (August 2006)
- Brian Dorvel, UIUC (September 2006)
- Ben Warlick, UIUC (September 2006)
- Leyla Celik, University of Aarhus (October 2006)
- Teemu Murtola, Helsinki University of Technology (October 2006)

- Chris Chipot, Universite Henri Poincare (December 2006)
- Francois Dehez, Universite Henri Poincare (December 2006)
- Ramya Gamini, Bradley University (December 2006)
- Jaime Valle, UIUC (December 2006)
- Anahita Tafvizi, Harvard University (January 2007)

Graduates

Recent UIUC graduates and postdoctoral associates who received or are continuing their training at the Resource include:

Ph.D. Recipients: Recent UIUC Ph.D. recipients who received their training at the Resource are:

- Jordi Cohen Ph.D., Physics, University of Illinois, Spring 2007

Postdoctoral Associates: Postdoctoral associates that have recently received or are currently receiving training at the Resource are:

- Emma Falck
- Jordi Cohen
- David Hardy
- Chris Harrison
- Zenmei Ohkubo
- Greg Sigalov
- Markus Dittrich

Res. Macromolecular Modeling & Bioinformatics
Grant Number: P41RR05969 (8/1/06 – 7/31/07)

Bibliography

[1] M. Wang and M. R. Briggs. HDL: The metabolism, function, and therapeutic importance. *Chem. Rev.*, 104:119–137, 2004.

[2] C. E. Young, R. H. Karas, and J. T. Kuvin. High-density lipoprotein cholesterol and coronary heart disease. *Cardiol. Rev.*, 12:107–119, 2004.

[3] S. Lund-Katz, L. Liu, S. T. Thuahnai, and M. C. Phillips. High density lipoprotein structure. *Front. Biosci.*, 8:d1044–1054, 2003.

[4] A. Jonas, K. E. Kézdy, and J. H. Wald. Defined apolipoprotein A-I conformations in reconstituted high density lipoprotein discs. *J. Biol. Chem.*, 264:4818–4824, 1989.

[5] M. A. Tricerri, A. K. Behling Agree, S. A. Sanchez, J. Bronski, and A. Jonas. Arrangement of apolipoprotein A-I in reconstituted high-density lipoprotein disks: an alternative model based on fluorescence resonance energy transfer experiments. *Biochemistry*, 40:5065–5074, 2001.

[6] R. A. G. D. Silva, G. M. Hilliard, L. Li, J. P. Segrest, and W. S. Davidson. A mass spectrometric determination of the conformation of dimeric apolipoprotein A-I in discoidal high density lipoproteins. *Biochemistry*, 44:8600–8607, 2005.

[7] W. S. Davidson and R. A. G. D. Silva. Apolipoprotein structural organization in high density lipoproteins: belts, bundles, hinges and hairpins. *Curr. Opin. Lipidol.*, 16:295–300, 2005.

[8] M. M. Thompson, S. C. Reed, and G. W. Cockerill. Therapeutic approaches to raising plasma hdl-cholesterol levels. *Nat. Clin. Prac. Cardio.*, 1:84–89, 2004.

[9] T. H. Bayburt, Y. V. Grinkova, and S. G. Sligar. Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins. *Nano Letters*, 2:853–856, 2002.

[10] I. G. Denisov, Y. V. Grinkova, A. A. Lazarides, and S. G. Sligar. Directed self-assembly of monodisperse phospholipid bilayer Nanodiscs with controlled size. *J. Am. Chem. Soc.*, 126:3477–3487, 2004.

[11] D. R. Davydov, H. Fernando, B. J. Baas, S. G. Sligar, and J. R. Halpert. Kinetics of dithionite-dependent reduction of cytochrome P450 3A4: heterogeneity of the enzyme caused by its oligomerization. *Biochemistry*, 44:13902–13913, 2005.

[12] T. Boldog, S. Grimme, M. Li, S. G. Sligar, and G. L. Hazelbauer. Nanodiscs separate chemoreceptor oligomeric states and reveal their signaling properties. *Proc. Natl. Acad. Sci. USA*, 103:11509–14, 2006.

[13] T. H. Bayburt, Y. V. Grinkova, and S. G. Sligar. Assembly of single bacteriorhodopsin trimers in bilayer nanodiscs. *Arch. Biochem. Biophys.*, 450:215–222, 2006.

[14] A. Y. Shih, A. Arkhipov, P. L. Freddolino, and K. Schulten. Coarse grained protein-lipid model with application to lipoprotein particles. *J. Phys. Chem. B*, 110:3674–3684, 2006.

[15] A. Y. Shih, P. L. Freddolino, A. Arkhipov, and K. Schulten. Assembly of lipoprotein particles revealed by coarse-grained molecular dynamics simulations. *J. Struct. Biol.*, 157:579–592, 2007.

[16] A. Y. Shih, A. Arkhipov, P. L. Freddolino, S. G. Sligar, and K. Schulten. Mechanism of discoidal high-density lipoprotein assembly. *J. Phys. Chem. B*, 2007. Submitted.

[17] A. Y. Shih, P. L. Freddolino, S. G. Sligar, and K. Schulten. Disassembly of nanodiscs with cholate. *Nano Lett.*, 2007. In Press.

[18] S. Lord. Fibrinogen and fibrin: scaffold proteins in homeostasis. *Curr. Opin. Hematol.*, 14:236–2419, 2007.

[19] J. Collet, H. Shuman, R. Ledger, S. Lee, and J. Weisel. The elasticity of an individual fibrin fiber clot. *Proc. Natl. Acad. Sci. USA*, 102:9133–9137, 2005.

[20] J. Weisel. Fibrinogen and fibrin. *Adv. Prot. Chem.*, 70:247–299, 2005.

[21] W. Liu, L. Jawerth, E. Sparks, M. Falvo, R. Hantgan, R. Superfine, S. Lord, and M. Guthold. Fibrin fibers have extraordinary extensibility and elasticity. *Science*, 313:634, 2006.

[22] Z. Yang, J. M. Kollman, L. Pandi, and R. F. Doolittle. Crystal structure of native chicken fibrinogen at 2.7 å resolution. *Biochemistry*, 40:12515–12523, 2001.

[23] R. Litvinov, O. Gorkun, S. Owen, H. Shuman, and J. Weisel. Polymerization of fibrin: specificity, strength, and stability of knob-hole interactions studied at the single molecule level. *Blood*, 106:2944–2951, 2005.

[24] O. Gorkun, R. Litvinov, Y. Veklich, and J. Weisel. Interactions mediated by the N-terminus of fibrinogen's Bβ chain. *Biochemistry*, 45:14843–14852, 2006.

[25] R. Litvinov, O. Gorkun, D. Galanakis, S. Yakovlev, L. Medved, H. Shuman, and J. Weisel. Polymerization of fibrin: Direct observation and quantification of individual b:b knob-hole interactions. *Blood*, 109:130–138, 2007.

[26] A. Brown, R. Litvinov, D. Discher, and J. Weisel. Forced unfolding of coiled-coils in fibrinogen by single molecule AFM. *Biophys. J.*, 92:39–41, 2007.

[27] M. Sotomayor and K. Schulten. Single-molecule experiments in vitro and in silico. *Science*, 316:1144–1148, 2007.

[28] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten. Scalable molecular dynamics with NAMD. *J. Comp. Chem.*, 26:1781–1802, 2005.

[29] V. A. Boichenko, E. Greenbaum, and M. Seibert. Hydrogen production by photosynthetic microorganisms. In M. D. Archer and J. Barber, editors, *Photoconversion of Solar Energy: Molecular to Global Photosynthesis*, pages 397–452. Imperial College Press, London, 2004.

[30] R. Mertens and A. Liese. Biotechnological applications of hydrogenases. *Curr. Opin. Biotech.*, 15:343–348, 2004.

[31] M. L. Ghirardi, L. Zhang, J. W. Lee, T. Flynn, M. Seibert, E. Greenbaum, and A. Melis. Microalgae: A green source of renewable H_2. *Trends Biotech.*, 18:506–511, 2000.

[32] J. Cohen, K. Kim, M. Posewitz, M. L. Ghirardi, K. Schulten, M. Seibert, and P. King. Molecular dynamics and experimental investigation of H_2 and O_2 diffusion in [Fe]-hydrogenase. *Biochem. Soc. Trans.*, 33:80–82, 2005.

[33] M. L. Ghirardi, P. W. King, M. C. Posewitz, P. C. Maness, A. Fedorov, K. Kim, J. Cohen, K. Schulten, and M. Seibert. Approaches to developing biological H_2-photoproducing organisms and processes. *Biochem. Soc. Trans.*, 33:70–72, 2005.

[34] J. Cohen, K. Kim, P. King, M. Seibert, and K. Schulten. Finding gas diffusion pathways in proteins: Application to O_2 and H_2 transport in CpI [FeFe]-hydrogenase and the role of packing defects. *Structure*, 13:1321–1329, 2005.

[35] P. W. King, D. Svedruzic, J. Cohen, K. Schulten, M. Seibert, and M. L. Ghirardi. Structural and functional investigations of biological catalysts for optimization of solar-driven, H_2 production systems. In L. Vayssieres, editor, *Solar Hydrogen and Nanotechnology*, volume 6340 of *Proc. SPIE*, pages 259–267, 2006.

[36] J. Cohen, A. Arkhipov, R. Braun, and K. Schulten. Imaging the migration pathways for O_2, CO, NO, and Xe inside myoglobin. *Biophys. J.*, 91:1844–1857, 2006.

[37] M. L. Ghirardi, J. Cohen, P. King, K. Schulten, K. Kim, and M. Seibert. [FeFe]-hydrogenases and photobiological hydrogen production. In L. Vayssieres, editor, *Solar hydrogen and Nanotechnology*, volume 6340 of *Proc. SPIE*, pages 253–258, 2006.

[38] Y. Wang, J. Cohen, W. F. Boron, K. Schulten, and E. Tajkhorshid. Exploring gas permeability of cellular membranes and membrane channels with molecular dynamics. *J. Struct. Biol.*, 2007. In press.

[39] B. J. Johnson, J. Cohen, R. W. Welford, A. R. Pearson, K. Schulten, J. P. Klinman, and C. M. Wilmot. Exploring molecular oxygen pathways in *Hanseluna Polymorpha* copper-containing amine oxidase. *J. Biol. Chem.*, 2007. In press.

[40] J. Cohen and K. Schulten. O_2 migration pathways in monomeric globins are determined by residue composition, not tertiary structure. *Biophys. J.*, 2007. Submitted.

[41] L. Kalé, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten. NAMD2: Greater scalability for parallel molecular dynamics. *J. Comp. Phys.*, 151:283–312, 1999.

[42] W. Humphrey, A. Dalke, and K. Schulten. VMD – Visual Molecular Dynamics. *J. Mol. Graphics*, 14:33–38, 1996.

[43] R. D. Skeel, D. J. Hardy, and J. C. Phillips. Correcting mesh-based force calculations to conserve both energy and momentum in molecular dynamics simulations. *J. Comp. Phys.*, 2006. Submitted.

[44] R. D. Skeel, I. Tezcan, and D. J. Hardy. Multiple grid methods for classical molecular dynamics. *J. Comp. Chem.*, 23:673–684, 2002.

[45] D. J. Hardy. *Multilevel Summation for the Fast Evaluation of Forces for the Simulation of Biomolecules*. PhD thesis, Univ. of Illinois at Urbana-Champaign, 2006. Also Department of Computer Science Report No. UIUCDCS-R-2006-2546, May 2006. Available online at `http://www.cs.uiuc.edu/research/techreports.php`.

[46] J. E. Stone, J. C. Phillips, P. L. Freddolino, D. J. Hardy, L. G. Trabuco, and K. Schulten. Accelerating molecular modeling applications with graphics processors. *J. Comp. Chem.*, 2007. Submitted.

[47] W. Wang and R. D. Skeel. Fast evaluation of polarizable forces. *J. Chem. Phys.*, 123:164107 (12 pages), 2005.

[48] G. Lamoureux and B. Roux. Modeling induced polarization with classical Drude oscillators: Theory and molecular dynamics simulation algorithm. *J. Chem. Phys.*, 119:3025–3039, 2003.

[49] G. Lamoureux and B. Roux. Absolute hydration free energy scale for alkali and halide ions established from simulations with a polarizable force field. *J. Phys. Chem. B*, 110:3308–3322, 2006.

[50] M. Bhandarkar, G. Budescu, W. Humphrey, J. A. Izaguirre, S. Izrailev, L. V. Kalé, D. Kosztin, F. Molnar, J. C. Phillips, and K. Schulten. BioCoRE: A collaboratory for structural biology. In A. G. Bruzzone, A. Uchrmacher, and E. H. Page, editors, *Proceedings of the SCS International Conference on Web-Based Modeling and Simulation*, pages 242–251, San Francisco, California, 1999.

[51] G. M. Preston, T. P. Carroll, W. B. Guggino, and P. Agre. Appearance of water channels in Xenopus oocytes expressing red cell CHIP28 protein. *Science*, 256:385–387, 1992.

[52] S. Hohmann, S. Nielsen, and P. Agre. *Aquaporins.* Academic Press, San Diego, 2001.

[53] P. M. T. Deen, M. A. J. Verdijk, N. V. A. M. Knoers, B. Wieringa, L. A. H. Monnens, C. H. van Os, and B. A. van Oost. Requirement of human renal water channel aquaporin-2 for vasopressin-dependent concentration of urine. *Science*, 264:92–94, 1994.

[54] P. M. T. Deen and C. H. van Os. Epithelial aquaporins. *Curr. Opin. Cell Biol.*, 10:435–442, 1998.

[55] P. Agre, M. Bonhivers, and M. J. Borgnia. The aquaporins, blueprints for cellular plumbing systems. *J. Biol. Chem.*, 273:14659–14662, 1998.

[56] M. Borgnia, S. Nielsen, A. Engel, and P. Agre. Cellular and molecular biology of the aquaporin water channels. *Annu. Rev. Biochem.*, 68:425–458, 1999.

[57] J. Yu, A. J. Yool, K. Schulten, and E. Tajkhorshid. Mechanism of gating and ion conductivity of a possible tetrameric pore in Aquaporin-1. *Structure*, 14:1411–1423, 2006.

[58] I. A. Solov'yov, D. E. Chandler, and K. Schulten. Magnetic field effects in Arabidopsis thaliana cryptochrome-1. *Biophys. J.*, 92:2711–2726, 2007.

[59] H. C. Berg. Motile behavior of bacteria. *Physics Today*, 53:24–29, 2000.

[60] F. A. Samatey, K. Imada, S. Nagashima, F. Vonderviszt, T. Kumasaka, M. Yamamoto, and K. Namba. Structure of the bacterial flagellar protofilament and implications for a switch for supercoiling. *Nature*, 410:331–337, 2001.

[61] I. Yamashita, K. Hasegawa, H. Suzuki, F. Vonderviszt, Y. Mimori-Kiyosue, and K. Namba. Structure and switching of bacterial flagellar filaments studied by x-ray fiber diffraction. *Nat. Struct. Biol.*, 5:125–132, 1998.

[62] K. Yonekura, S. Maki-Yonekura, and K. Namba. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature*, 424:643–650, 2003.

[63] A. Kitao, K. Yonekura, S. Maki-Yonekura, F. A. Samatey, K. Imada, K. Namba, and N. Go. Switch interactions control energy frustration and multiple flagellar filament structures. *Proc. Natl. Acad. Sci. USA*, 103:4894–4899, 2006.

[64] A. Arkhipov, P. L. Freddolino, and K. Schulten. Stability and dynamics of virus capsids described by coarse-grained modeling. *Structure*, 14:1767–1777, 2006.

[65] A. Arkhipov, P. L. Freddolino, K. Imada, K. Namba, and K. Schulten. Coarse-grained molecular dynamics simulations of a rotating bacterial flagellum. *Biophys. J.*, 91:4589–4597, 2006.

[66] R. Bonneau, J. Tsai, I. Ruczinski, D. Chivian, C. Rohl, C. E. Strauss, and D. Baker. Rosetta in CASP4: progress in ab initio protein structure prediction. *Proteins Suppl*, 5:119–126, 2001.

[67] R. Bonneau, C. E. Strauss, C. A. Rohl, D. Chivian, P. Bradley, L. Malmstrom, T. Robertson, and D. Baker. De novo prediction of three-dimensional structures for major protein families. *J. Mol. Biol.*, 322:65–78, 2002.

[68] S. Velankar, P. Soultanas, M. Dillingham, H. Subramanya, and D. Wigley. Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell*, 97:75–84, 1999.

[69] M. Dillingham, D. Wigley, and M. Webb. Demonstration of unidirectional single-stranded DNA translocatin by PcrA helicase: measurement of step size and translocation speed. *Biochemistry*, 39:205–212, 2000.

[70] M. Dittrich and K. Schulten. PcrA helicase, a prototype ATP-driven molecular motor. *Structure*, 14:1345–1353, 2006.

[71] J. Yu, T. Ha, and K. Schulten. Structure-based model of the stepping motor of PcrA helicase. *Biophys. J.*, 91:2097–2114, 2006.

[72] M. Dittrich, J. Yu, and K. Schulten. PcrA helicase, a molecular motor studied from the electronic to the functional level. *Topics in Current Chemistry*, 268:319–347, 2006.

[73] J. Yu, T. Ha, and K. Schulten. How directional translocation is regulated in a DNA helicase motor. *Biophys. J.*, 2007. Submitted.

[74] M. Dittrich, S. Hayashi, and K. Schulten. ATP hydrolysis in the β_{TP} and β_{DP} catalytic sites of F_1-ATPase. *Biophys. J.*, 87:2954–2967, 2004.

[75] M. Dittrich, S. Hayashi, and K. Schulten. On the mechanism of ATP hydrolysis in F_1-ATPase. *Biophys. J.*, 85:2253–2266, 2003.

[76] G. M. Suel, S. Lockless, M. A. Wall, and R. Ranganathan. Evolutionarily conserved networks of residues mediate allosteric coupling in proteins. *Nat. Struct. Biol.*, 10:59–69, 2003.

[77] Y. Wang, J. Cohen, W. F. Boron, K. Schulten, and E. Tajkhorshid. Exploring gas permeability of cellular membranes and membrane channels with molecular dynamics. *J. Struct. Biol.*, 157:534–544, 2007.

[78] M. Gower, J. Cohen, J. Phillips, R. Kufrin, and K. Schulten. Managing biomolecular simulations in a grid environment with NAMD-G. In *Proceedings of the 2006 TeraGrid Conference*, 2006. (7 pages).

[79] M. Schlenkrich, J. Alex D. MacKerell, and M. Karplus. In preparation, 1992.

[80] K. S. Matthews. DNA looping. *Microbiological Reviews*, 56:123–136, 1992.

[81] J. H. Miller and W. S. Reznikoff, editors. *The Operon*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New Tork, 1980.

[82] M. Lewis, G. Chang, N. C. Horton, M. A. Kercher, H. C. Pace, M. A. Schumacher, R. G. Brennan, and P. Lu. Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science*, 271:1247–1254, 1996.

[83] A. M. Friedman, T. O. Fischmann, and T. A. Steitz. Crystal structure of *lac* repressor core tetramer and its implications for DNA looping. *Science*, 268:1721–1727, 1995.

[84] E. Villa, A. Balaeff, and K. Schulten. Structural dynamics of the *Lac* repressor-DNA complex revealed by a multiscale simulation. *Proc. Natl. Acad. Sci. USA*, 102:6783–6788, 2005.

[85] A. Balaeff, L. Mahadevan, and K. Schulten. Elastic rod model of a DNA loop in the *lac* operon. *Phys. Rev. Lett.*, 83:4900–4903, 1999.

[86] A. Balaeff, L. Mahadevan, and K. Schulten. Structural basis for cooperative DNA binding by CAP and *Lac* repressor. *Structure*, 12:123–132, 2004.

[87] A. Balaeff, L. Mahadevan, and K. Schulten. Modeling DNA loops using the theory of elasticity. *Phys. Rev. E*, 73:031919, 2006. (23 pages).

[88] A. Balaeff, L. Mahadevan, and K. Schulten. Modeling DNA loops using the theory of elasticity. *Phys. Rev. E*, 2004. Submitted.

[89] B. Isralewitz, M. Gao, and K. Schulten. Steered molecular dynamics and mechanical functions of proteins. *Curr. Opin. Struct. Biol.*, 11:224–230, 2001.

[90] H. Garcia, P. Grayson, L. Han, M. Inamdar, J. Kondev, P. Nelson, R. Phillips, J. Widom, and P. Wiggins. Biological consequences of tightly bent DNA: The other life of a macromolecular celebrity. *Biopolymers*, 85:115–130, 2007.

[91] L. M. Edelman, R. Cheong, and J. D. Kahn. Fluorescence resonance energy transfer over ≈130 basepairs in hyperstable lac repressor-DNA loops. *Biophys. J.*, 84:1131–1145, 2003.

[92] H. R. Kaback. Active transport in *Escherichia coli*: Passage to permease. *Annu. Rev. Biophys. Biophys. Chem*, 15:279–319, 1986.

[93] H. R. Kaback. Molecular biology of active transport: From membranes to molecules to mechanism. *Harvey Lect.*, 83:77–103, 1989.

[94] H. R. Kaback. In and out and up and down with the lactose permease of *Escherichia coli*. *Int. Rev. Cytol.*, 137:97–125, 1992.

[95] J. Abramson, I. Smirnova, V. Kasho, G. Verner, H. R. Kaback, and S. Iwata. Structure and mechanism of the lactose permease of *Escherichia coli*. *Science*, 301:610–615, 2003.

[96] H. R. Kaback. Structure and mechanism of the lactose permease. *C. R. Biol.*, 328:557–567, 2005.

[97] H. V. Rickenberg, G. N. Cohen, G. Buttin, and J. Monod. La galactoside-permease d'*Escherichia coli*. *Ann. Inst. Pasteur*, 91:829–857, 1956.

[98] H. R. Kaback, M. Sahin-Tóth, and A. B. Weinglass. The kamikaze approach to membrane transport. *Nat. Rev. Mol. Cell Biol.*, 2:610–620, 2001.

[99] J. Abramson, S. Iwata, and H. R. Kaback. Lactose permease as a paradigm for membrane transport proteins (Review). *Mol. Memb. Biol.*, 21:227–236, 2004.

[100] Y. Yin, M. Ø. Jensen, E. Tajkhorshid, and K. Schulten. Sugar binding and protein conformational changes in lactose permease. *Biophys. J.*, 91:3972–3985, 2006.

[101] M. Ø. Jensen, Y. Yin, E. Tajkhorshid, and K. Schulten. Sugar transport across lactose permease probed by steered molecular dynamics. *Biophys. J.*, 93:92–102, 2007.

[102] O. P. Hamill and B. Martinac. Molecular basis of mechanotransduction in living cells. *Physiol. Rev.*, 81:685–740, 2001.

[103] S. Sukharev and D. P. Corey. Mechanosensitive channels: Multiplicity of families and gating paradigms. *Science's STKE*, 219, 2004. re4.

[104] B. Martinac, M. Buechner, A. H. Delcour, J. Adler, and C. Kung. Pressure-sensitive ion channel in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA*, 84:2297–2301, 1987.

[105] N. Levina, S. Totemeyer, N. R. Stokes, P. Louis, M. A. Jones, and I. Booth. Protection of *Escherichia coli* cells against extreme turgor by activation of MscS and MscL mechanosensitive channels: identification of genes required for MscS activity. *EMBO J.*, 18:1730–1737, 1999.

[106] G. Chang, R. H. Spencer, A. T. Lee, M. T. Barclay, and D. C. Rees. Structure of the MscL homolog from *Mycobacterium tuberculosis*: A gated mechanosensitive ion channel. *Science*, 282:2220–2226, 1998.

[107] R. B. Bass, P. Strop, M. Barclay, and D. C. Rees. Crystal structure of *Escherichia coli* MscS, a voltage-modulated and mechanosensitive channel. *Science*, 298:1582–1587, 2002.

[108] A. Anishkin and S. Sukharev. Water dynamics and dewetting transitions in the small mechanosensitive channel MscS. *Biophys. J.*, 86:2883–2895, 2004.

[109] M. Sotomayor and K. Schulten. Molecular dynamics study of gating in the mechanosensitive channel of small conductance MscS. *Biophys. J.*, 87:3050–3065, 2004.

[110] M. D. Edwards, Y. Li, S. Miller, W. Barlett, S. Black, S. Dennison, I. Iscla, P. Blount, J. U. Bowie, and I. R. Booth. Pivotal role of the glycine-rich TM3 helix in gating the MscS mechanosensitive channel. *Nat. Struct. Mol. Biol.*, 12:113–119, 2005.

[111] B. Akitake, A. Anishkin, and S. Sukharev. The dashpot mechanism of stretch-dependent gating in MscS. *J. Gen. Physiol.*, 125:143–54, 2005.

[112] A. Anishkin and C. Kung. Microbial mechanosensation. *Curr. Opin. Neur.*, 15:397–405, 2005.

[113] B. Martinac. Structural plasticity in MS channels. *Nat. Struct. Mol. Biol.*, 12:104–105, 2005.

[114] M. Sotomayor, T. A. van der Straaten, U. Ravaioli, and K. Schulten. Electrostatic properties of the mechanosensitive channel of small conductance MscS. *Biophys. J.*, 90:3496–3510, 2006.

[115] P. Koprowski and A. Kubalski. Voltage-independent adaptation of mechanosensitive channels in *Escherichia coli* protoplasts. *J. Membr. Biol.*, 164:253–262, 1998.

[116] V. Vasquez and E. Perozo. Voltage dependent gating in MscS. *Biophys. J.*, 86:545A–545A, 2004.

[117] S. Spronk, D. E. Elmore, and D. A. Dougherty. Voltage-dependent hydration and conduction properties of the hydrophobic pore of the mechanosensitive channel of small conductance. *Biophys. J.*, 90:3555–3569, 2006.

[118] M. Sotomayor, V. Vasquez, E. Perozo, and K. Schulten. Ion conduction through MscS as determined by electrophysiology and simulation. *Biophys. J.*, 92:886–902, 2007.

[119] G. Shapovalov and H. A. Lester. Gating transitions in bacterial ion channels measured at 3 μs resolution. *J. Gen. Physiol.*, 124:151–161, 2004.

[120] T. Nomura, M. Sokabe, and K. Yoshimura. Lipid-protein interaction of the MscS mechanosensitive channel examined by scanning mutagenesis. *Biophys. J.*, 2006. In press.

[121] J. Gullingsrud, D. Kosztin, and K. Schulten. Structural determinants of MscL gating studied by molecular dynamics simulations. *Biophys. J.*, 80:2074–2081, 2001.

[122] J. Gullingsrud and K. Schulten. Gating of MscL studied by steered molecular dynamics. *Biophys. J.*, 85:2087–2099, 2003.

[123] J. Gullingsrud and K. Schulten. Lipid bilayer pressure profiles and mechanosensitive channel gating. *Biophys. J.*, 86:3496–3509, 2004.

[124] A. A. Ajees, G. M. Anantharamaiah, V. K. Mishra, M. M. Hussain, and H. M. K. Murthy. Crystal structure of human apolipoprotein A-I: insights into its protective

effect against cardiovascular diseases. *Proc. Natl. Acad. Sci. USA*, 103:2126–2131, 2006.

[125] D. W. Borhani, D. P. Rogers, J. A. Engler, and C. G. Brouillette. Crystal structure of truncated human apolipoprotein A-I suggests a lipid-bound conformation. *Proc. Natl. Acad. Sci. USA*, 94:12291–12296, 1997.

[126] I. G. Denisov, M. A. McLean, A. W. Shaw, Y. V. Grinkova, and S. G. Sligar. Thermotropic phase transition in soluble nanoscale lipid bilayers. *J. Phys. Chem. B*, 109:15580–15588, 2005.

[127] A. Y. Shih, I. G. Denisov, J. C. Phillips, S. G. Sligar, and K. Schulten. Molecular dynamics simulations of discoidal bilayers assembled from truncated human lipoproteins. *Biophys. J.*, 88:548–556, 2005.

[128] R. F. Service. The race for the $1000 genome. *Science*, 311:1544–1546, 2006.

[129] Q. Zhao, G. Sigalov, V. Dimitrov, B. Dorvel, U. Mirsaidov, S. Sligar, A. Aksimentiev, and G. Timp. Detecting SNPs using a synthetic nanopore. *Nano Lett.*, 2007. In press.

[130] M. E. Gracheva, A. Xiong, J.-P. Leburton, A. Aksimentiev, K. Schulten, and G. Timp. Simulation of the electric response of DNA translocation through a semiconductor nanopore-capacitor. *Nanotechnology*, 17:622–633, 2006.

[131] M. E. Gracheva, A. Aksimentiev, and J.-P. Leburton. Electrical signatures of single-stranded DNA with single base mutations in a nanopore capacitor. *Nanotechnology*, 17:3160–3165, 2006.

[132] E. R. Cruz-Chu, A. Aksimentiev, and K. Schulten. Water-silica force field for simulating nanodevices. *J. Phys. Chem. B*, 110:21497–21508, 2006.

[133] M. J. O'Connell, S. M. Bachilo, C. B. Huffman, V. C. Moore, M. S. Strano, E. H. Haroz, K. L. Rialon, P. J. Boul, W. H. Noon, C. Kittrell, J. Ma, R. H. Hauge, R. B. Weisman, and R. E. Smalley. Band gap fluorescence from individual single-walled carbon nanotubes. *Science*, 297:593–596, 2002.

[134] L. J. McCartney, J. C. Pickup, O. J. Rolinski, and D. J. S. Birch. Near-infrared fluorescence lifetime assay for serum glucose based on allophycocyanin-labeled concanavalin A. *Anal. Biochem.*, 127:216 – 221, 2001.

[135] N. Klonisa, N. H. Quazi, L. W. Deady, A. B. Hughes, and L. Tilley. Characterization of a series of far-red-absorbing thiobarbituric acid oxonol derivatives as fluorescent probes for biological applications. *Anal. Biochem.*, 317:47–58, 2003.

[136] J. V. Frangioni. *In vivo* near-infrared fluorescence imaging. *Curr. Opin. Chem. Biol.*, 7:626–634, 2003.

[137] S. Kim, Y. T. Lim, E. G. Soltesz, A. M. De Grand, J. Lee, A. Nakayama, J. A. Parker, T. Mihaljevic, R. G. Laurence, D. M. Dor, L. H. Cohn, M. G. Bawendi, and J. V. Frangioni. Near-infrared fluorescent type II quantum dots for sentinel lymph node mapping. *Nat. Biotechnol.*, 22:93–97, 2004.

[138] D. Lu, Y. Li, S. V. Rotkin, U. Ravaioli, and K. Schulten. Finite-size effect and wall polarization in a carbon nanotube channel. *Nano Lett.*, 4:2383–2387, 2004.

[139] Y. Li, D. Lu, S. V. Rotkin, K. Schulten, and U. Ravaioli. Screening of water dipoles inside finite-length armchair carbon nanotubes. *J. Comp. Electron.*, 4:161–165, 2005.

[140] D. Lu, Y. Li, U. Ravaioli, and K. Schulten. Empirical nanotube model for biological applications. *J. Phys. Chem. B*, 109:11461–11467, 2005.

[141] D. Lu, Y. Li, U. Ravaioli, and K. Schulten. Ion-nanotube terahertz oscillator. *Phys. Rev. Lett.*, 95:246801, 2005.

[142] D. Lu, A. Aksimentiev, A. Y. Shih, E. Cruz-Chu, P. L. Freddolino, A. Arkhipov, and K. Schulten. The role of molecular modeling in bionanotechnology. *Physical Biology*, 3:S40–S53, 2006.

[143] I. G. Macara. Transport into and out of the nucleus. *Microbiol. Mol. Biol. Rev.*, 65:570–594, 2001.

[144] B. Fahrenkrog, J. Köser, and U. Aebi. The nuclear pore complex: A jack of all trades? *Trends Biochem. Sci.*, 29:175–182, 2004.

[145] B. Fahrenkrog and U. Aebi. The nuclear pore complex: Nucleocytoplasmic transport and beyond. *Nat. Rev. Mol. Cell Biol.*, 4:757–766, 2003.

[146] S. A. Adam. The nuclear pore complex. *Gen. Biol.*, 2:reviews0007.1–0007.6, 2001.

[147] M. P. Rout and J. D. Aitchison. The nuclear pore complex as a transport machine. *J. Biol. Chem.*, 276:16593–16596, 2001.

[148] D. Görlich and U. Kutay. Transport between the cell nucleus and the cytoplasm. *Annu. Rev. Cell. Dev. Biol.*, 15:607–660, 1999.

[149] T. A. Isgro and K. Schulten. Association of nuclear pore FG-repeat domains to NTF2 import and export complexes. *J. Mol. Biol.*, 366:330–345, 2007.

[150] T. A. Isgro and K. Schulten. Cse1p binding dynamics reveal a novel binding pattern for FG-repeat nucleoporins on transport receptors. *Structure*, 2007. Submitted.

[151] T. A. Isgro and K. Schulten. Binding dynamics of isolated nucleoporin repeat regions to importin-β. *Structure*, 13:1869–1879, 2005.

[152] J. Deisenhofer, O. Epp, K. Mikki, R. Huber, and H. Michel. Structure of the protein subunits in the photosynthetic reaction centre of *Rhodopseudomonas viridis* at 3 Å resolution. *Nature*, 318:618–624, 1985.

[153] U. Ermler, G. Fritzsch, S. K. Buchanan, and H. Michel. Structure of the photosynthetic reaction center from *Rhodobacter sphaeroides* at 2.65 Å resolution: Cofactors and protein-cofactor interactions. *Structure*, 2:925–936, 1994.

[154] S. J. Jamieson, P. Wang, P. Qian, J. Y. Kirkland, M. J. Conroy, C. N. Hunter, and P. A. Bullough. Projection structure of the photosynthetic reaction centre-antenna complex of *rhodospirillum rubrum* at 8.5 Å resolution. *J. Mol. Biol.*, 21:3927–3935, 2002.

[155] X. Hu and K. Schulten. A model for the light-harvesting complex I (B875) of *Rhodobacter sphaeroides*. *Biophys. J.*, 75:683–694, 1998.

[156] G. McDermott, S. M. Prince, A. A. Freer, A. M. Hawthornthwaite-Lawless, M. Z. Papiz, R. J. Cogdell, and N. W. Isaacs. Crystal structure of an integral membrane light-harvesting complex from photosynthetic bacteria. *Nature*, 374:517–521, 1995.

[157] J. Koepke, X. Hu, C. Muenke, K. Schulten, and H. Michel. The crystal structure of the light harvesting complex II (B800-850) from *Rhodospirillum molischianum*. *Structure*, 4:581–597, 1996.

[158] A. R. Crofts, E. A. Berry, Z. Zhang, L. Huang, R. Kuras, and M. Guergova-Kuras. On the mechanism of ubiquinol oxidation by the bc_1-complex: The role of the Rieske subunit, and its mobility. *Biochemistry*, 1998. Submitted.

[159] H. L. Axelrod, E. C. Abresch, M. Y. Okamura, A. P. Yeh, D. C. Ress, and G. Feher. X-ray structure determination of the cytochrome c2: Reaction center electron transfer complex from *rhodobacter sphaeroides*. *EMBO J.*, 319:501–515, 2002.

[160] B. A. Feniouk, D. A. Cherepanov, N. E. Voskoboynikova, A. Y. Mulkidjanian, and W. Junge. Chromatophore vesicles of *rhodobacter capsulatus* contain on average one F_oF_1-ATP synthase each. *Biophys. J.*, 82:1115–1122, 2002.

[161] M. K. Sener, J. D. Olsen, C. N. Hunter, and K. Schulten. Atomic level structure and function of an entire biological membrane: the photosynthetic unit of *Rhodobacter sphaeroides*. *Proc. Natl. Acad. Sci. USA*, 2007. Submitted.

[162] B. Hille. *Ionic channels of excitable membranes*. Sinauer Associates, Sunderland, MA, 2nd edition, 1992.

[163] G. J. Siegal, B. W. Agranoff, R. W. Albers, S. K. Fisher, and M. D. Uhler. *Basic neurochemistry, molecular, cellular, and medical aspects*. Lippincott Williams and Wilkins, Philadelphia, 6th edition, 1999.

[164] S. B. Long, E. B. Campbell, and R. MacKinnon. Crystal structure of a mammalian voltage-dependent *Shaker* family K^+ channel. *Science*, 309:897–903, 2005.

[165] S. K. Aggarwal and R. MacKinnon. Contribution of the S4 segment to the gating charge in the *Shaker* K^+ channel. *Neuron*, 16:1169–1177, 1996.

[166] A. Cha, G. E. Snyder, P. R. Selvin, and F. Bezanilla. Atomic scale movement of the voltage-sensing region in a potassium channel measured via spectroscopy. *Nature*, 402:809–813, 1999.

[167] V. Ruta, J. Chen, and R. MacKinnon. Calibrated measurement of gating-charge arginine displacement in the KvAP voltage-dependent K^+ channel. *Cell*, 123:463–475, 2005.

[168] R. Chang, G. S. Ayton, and G. A. Voth. Multiscale coupling of mesoscopic- and atomistic-level lipid bilayer simulations. *J. Chem. Phys.*, 122:244716–244728, 2005.

[169] P. Bradley, K. M. S. Misura, and D. Baker. Toward high-resolution de novo structure prediction for small proteins. *Science*, 309:1868–1871, 2005.

[170] V. Yarov-Yarovoy, D. Baker, and W. A. Catterall. Voltage sensor conformations in the open and closed states in ROSETTA structural models of K^+ channels. *Proc. Natl. Acad. Sci. USA*, 103:7292–7297, 2006.

[171] Y. Sugita and Y. Okamoto. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.*, 314:141–151, 1999.

[172] F. Khalili-Araghi, E. Tajkhorshid, and K. Schulten. Dynamics of K^+ ion conduction through Kv1.2. *Biophys. J.*, 91:L72–L74, 2006.

[173] Y. Zhou, J. H. Morais-Cabral, A. Kaufman, and R. MacKinnon. Chemistry of ion coordination and hydration revealed by a K^+ channel-fab complex at 2.0 Å resolution. *Nature*, 414:43–48, 2001.

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

[175] J. Åqvist and V. Luzhkov. Ion permeation mechanism of the potassium channel. *Nature*, 404:881–884, 2000.

[176] A. Liljas. *Structural Aspects of Protein Synthesis*. World Scientific, Singapore, 2004.

[177] J. Frank. Cryo-electron microscopy as an investigative tool: The ribosome as an example. *BioEssays*, 23:725–732, 2001.

[178] B. S. Schuwirth, M. A. Borovinskaya, C. W. Hau, W. Zhang, A. Vila-Sanjurjo, J. M. Holton, and J. H. D. Cate. Structures of the bacterial ribosome at 3.5 Å resolution. *Science*, 310:827–834, 2005.

[179] M. Selmer, C. M. Dunham, F. V. Murphy IV, A. Weixlbaumer, S. Petry, A. C. Kelley, J. R. Weir, and V. Ramakrishnan. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science*, 313:1935–1942, 2006.

[180] N. S. Rane, J. L. Yonkovich, and R. S. Hegde. Protection from cytosolic prion protein toxicity by modulation of protein translocation. *EMBO J.*, 23:4550–4559, 2004.

[181] S. Barash, W. Wang, and Y. Shi. Human secretory signal peptide description by hidden Markov model and generation of a strong artificial signal peptide for secreted protein expression. *Biochem. Biophys. Res. Commun.*, 294:835–842, 2002.

[182] J. Gumbart and K. Schulten. Molecular dynamics studies of the archaeal translocon. *Biophys. J.*, 90:2356–2367, 2006.

[183] J. Gumbart and K. Schulten. Structural determinants of lateral gate opening in the protein translocon. 2007. Submitted.

[184] C. M. Shepherd, I. A. Borelli, G. Lander, P. Natarajan, V. Siddavanahalli, C. Bajaj, J. E. Johnson, I. C. L. Brooks, and V. S. Reddy. VIPERdb: A relational database for structural virology. *Nucl. Acids Res.*, 34:D386–D389, 2006.

[185] D. E. Smith, S. J. Tans, S. B. Smith, S. Grimes, D. L. Anderson, and C. Bustamante. The bacteriophage phi29 portal motor can package DNA against a large internal force. *Nature*, 413:748–752, 2001.

[186] M. J. Rust, M. Lakadamyali, F. Zhang, and X. Zhuang. Assembly of endocytic machinery around individual influenza viruses during viral entry. *Nat. Struct. Mol. Biol.*, 11:567–573, 2004.

[187] P. L. Freddolino, A. S. Arkhipov, S. B. Larson, A. McPherson, and K. Schulten. Molecular dynamics simulations of the complete satellite tobacco mosaic virus. *Structure*, 14:437–449, 2006.

[188] T. J. Tuthill, D. Bubeck, D. J. Rowlands, and J. M. Hogle. Characterization of early steps in the poliovirus infection process: Receptor decorated liposomes induce conversion of the virus to membrane-anchored entry-intermediate particles. *J. Virol.*, 80:172–180, 2006.

[189] P. Beroza and D. A. Case. Including side chain flexibility in continuum electrostatic calculations of protein titration. *JPC*, 100:20156–20163, 1996.