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

NATIONAL CENTER FOR RESEARCH RESOURCES
BIOMEDICAL TECHNOLOGY AREA

ANNUAL PROGRESS REPORT (8/1/05 – 7/31/06)

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, 2005
   B. TO (Month, Day, Year): July 31, 2006
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%

Grant Number: P41RR05969
Report PD: (8/1/05 – 7/31/06)

Contents

Grant Number: P41RR05969

Report PD: (8/1/05 – 7/31/06)

We'll have some text here.

The preparation of this report was coordinated by Tim Skirvin

Grant Number: P41RR05969
Report PD: (8/1/05 – 7/31/06)

Highlights

NAMD-G: Grid Technology for NAMD Simulations

When Resource scientists, collaborating with the National Renewable Energy Laboratory to convert the hydrogenase enzyme into an affordable means of large-scale production of hydrogen gas [1], developed a novel algorithm for accurately predicting gas migration pathways inside any protein [2], they were faced with an opportunity and a problem. The opportunity was to characterize gas migration processes in a large set of proteins for which experimental investigation is not practical and/or possible. Each protein studied would require the trajectory from a molecular dynamics simulation of moderate duration that could then be analyzed to determine pathways for any gas of interest (hydrogen, oxygen, nitrogen, methane, etc.). While the computational resources to perform such simulations for several hundred proteins were readily available, the human effort required to copy the input files to remote supercomputers, "baby-sit" the running simulations, and retrieve and analyze the results was daunting—automation was clearly needed.

The Resource immediately turned to the grid computing experts at the National Center for Supercomputing Applications (NCSA, http://www.ncsa.uiuc.edu/), who had years of experience with the tools and services of the TeraGrid (http://teragrid.org/), a national distributed high-performance computing infrastructure on which the Resource's simulation program NAMD (http://www.ks.uiuc.edu/Research/namd/) was already available. NCSA was eager to demonstrate the power of grid technologies for NAMD users, and both parties agreed to construct an integrated system that would provide single sign-on authentication and access to TeraGrid computational resources, a transparent interface to high-speed data management and file transfer utilities available within the TeraGrid software stack, and a workflow model tailored to the way that the biomolecular researcher actually proceeds, thus offloading the burdens of frequent monitoring of simulations and of managing the details of job submission and of restarts.

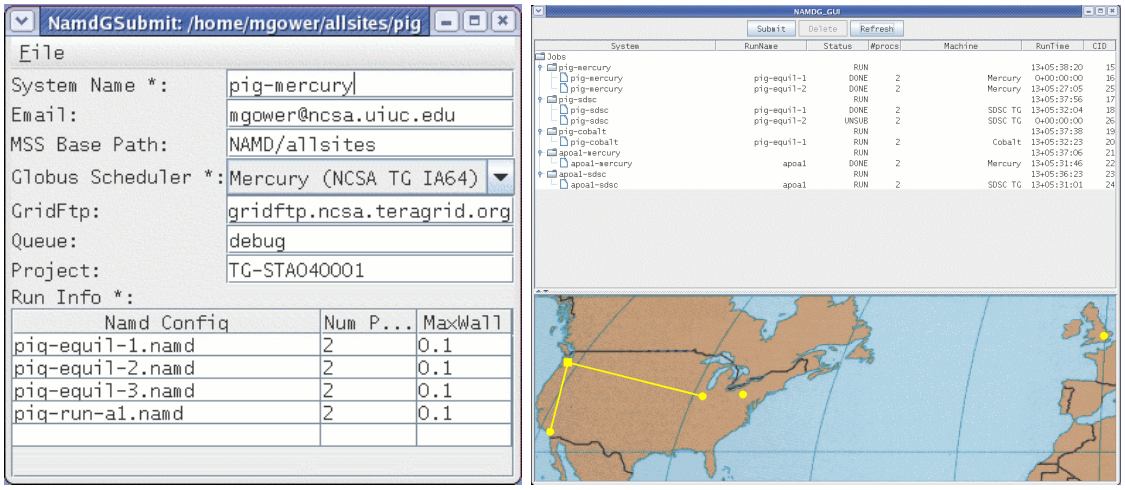

Figure 1: NAMD-G job submission window (left) and graphical monitor (right), showing automated file transfer and job execution, as demonstrated at SC2005.

A prototype of NAMD-G was successfully demonstrated at the SC2005 conference in November, distributing jobs submitted from the show floor to machines both across the US and internationally to the UK. Having reached this milestone, the Resource/NCSA development team began refining NAMD-G into a reliable and transparent tool that could be distributed to the NAMD user community. Resource scientists now use NAMD-G for their daily work, providing essential feedback, feature requests, and bug reports; new users are added as the developers are able to support them. The tight nature of the NCSA collaboration will allow the NAMD-G developers to anticipate and adapt to changes in the TeraGrid software stack and environment, shielding the users from these disruptions.

The workflow capabilities of NAMD-G go beyond simply executing a sequence of runs specified by the researcher. NAMD-G automatically splits simulations that require more computing time than the limit imposed by the remote batch system. While transferring files securely to and from remote supercomputers, NAMD-G also saves both input and output datasets to the NCSA tape archive. Finally, NAMD-G keeps the researcher informed by sending email when each portion of a simulation is completed, including an excerpt of the output for quick visual inspection. For its users, NAMD-G transforms computing from a well, requiring significant labor to extract each bucket of resources, into a free-flowing tap that can be left to run while their minds are more usefully engaged.

Workshops, Tutorials, and Case Studies

Figure 2: Images from the Workshops

Life sciences, perhaps the oldest area of human inquiry, is one that is constantly evolving, and today combines quantitative and computational methods in ways that open new areas of research and enable new discoverlies. Essentially, life sciences is rapidly tranforming into a new discipline, and what is learned in the laboratory today by tomorrow is applied in medicine and industry. What this means for education is that scientists must be trained to deal with both quantitative information and computational methods. In response, the Resource has refined a training approach that emphasizes conceptual understanding, learning by doing, and the opportunity for deeper learning. These goals have been realized in recent Resource-led workshops conducted in nationally and international locations, and through the expansion of tutorial materials and the release of intellectually rich case studies.

Workshop efforts by the Resource generally follow a model of conceptual learning via lectures from Resource faculty or staff in the first part of a workshop day, followed by hands-on learning in the second part of the day. Morning lectures provide a foundation for later hands-on activities, by providing a foundation in theory, and by describing how that theory and related techniques are manifested in computation, simulation, and modeling. Later hands-on sessions give participants the opportunity to put what they've learned to work, whether it be modeling a particular molecule or building a computer framework to support modeling. Recent workshops by the Resource focus on two areas: computational biophysics, and how to build clusters of computers to support scientific work.

Five computational biophysics workshops were held from May 2005 to April 2006: Lake Tahoe, California; Chicago, Illinois; San Franciso, California; Pittsburgh, Pennsylvania; and Frankfurt, Germany. While the Lake Tahoe and Chicago workshops were funded via an NIH grant, and the San Francisco workshop was funded by the National Science foundation, the Pittsburgh and Frankfurt workshops were funded by other organizations, the Pittsburgh Supercomputer Center and the Max Planck Institute for biophysics respectively. A mobile computing laboratory, comprised of 22 Macintosh laptops humming with needed files and computational biology software (e.g., VMD and NAMD), was designed and implemented by the Resource to support the Chicago, Pittsburgh, and San Francisco workshops. By the time of the Frankfurt workshops, workshop materials had evolved so far that participants could use their own laptops. The workshops ran from

three and a half days to a full length in schedule, and typically included 20 participants each, though by Frankfurt that number had nearly doubled, to 37 participants.

Three cluster building workshops were conducted by the Resource from May 2005 to mid-April 2006, all held at the Beckman Institute, where Resource staff can assure the substantial equipment and connectivity needs required by the workshop curriculum. In these workshops, participants - whether they be system administrators or scientists - hear lectures on the concepts underlying the development of PC-based clusters running the Linux operating system, such as why they're useful, how they work, programming techniques, and how to design queus for the linked computers. In the hands-on session of the workshop, participants actually build a functioning cluster, and try out their own applications. A day and half in length, the cluster workshops generally enroll 22-24 participants.

A realization of the intent to teach the marriage of life sciences and computational tools, the Resource tutorials and case studies are available to the biomedical community via the Resource web site, allowing learners to at their own pace develop skills and in-depth understanding. With their origins in a 2003 summer school requiring a Unix-based laboratory, the Resource tutorials have since been expanded to allow their use across the most popular computer platforms - Windows, Macintosh, and Linux/Unix. Further, the use of third-party software - which introduces both financial and licensing complications - was eliminated from the most recent versions of the tutorials, so now any scientist with a laptop and an internet connection can proceed at no cost through the tutorials, getting hands-on experience with application of VMD and NAMD to scientific questions.

Another recent addition to the Resource training trove are eight case studies, each of which exploit VMD's molecular graphics capabilities to teach molecular cell biology. The case studies start out like a conventional textbook chapter, but utilize VMD molecular graphics to offer a much more detailed view of the subjects than commonly possible in textbooks. The case studies induce the reader to inspect by means of VMD the systems introduced. For this purpose the case studies show the molecular systems through graphical images, but provide also the files that permit the reader to regenerate the images with few mouse clicks and then explore in depth the structures shown using the program VMD. Students can rotate the images, enlarge them, alter the views in many ways, and analyze structures and sequences. In this fashion, a more in-depth understanding of the use of visualization tools to explore scientific questions is developed.

Looking forward, the Resource intends to continue workshop activities, with one workshop already planned for November 2006, and more tutorials and case studies are also under development. More on the aforementioned training efforts can be found by exploring the Resource training website: http://www.ks.uiuc.edu/Training/

Simulating an Entire Life Form

Viruses are tiny, parasitic organisms responsible for diseases ranging from AIDS to the common cold. Composed only of a small genome, a protein coat (capsid), and sometimes a few other accessory molecules, there is some debate on whether or not viruses should even be considered alive, yet the damage they cause cannot be questioned. Biophysical research on viruses is currently an area of great interest, both because of the possibility of developing new treatments for viral diseases like bird flue and because of the promise shown by the use of viral particles in nanotechnology [3, 4].

Because of their ability to analyze and manipulate structures at the atomic level, molecular dynamics and other computational techniques will be of great use in analyzing viruses and developing novel antiviral agents. Over the last several years a number of computational studies have been performed to understand the mechanism of antiviral drugs [5, 6], viral stability [7, 8], and important structural transitions in viral life cycles [9–11]. However, although viruses are quite small for biological systems, even the smallest viruses present a formidable challenge for computational study due to their size [8, 12]. Because of this, previous computational studies on viruses have either focused on only a small piece of the virus, attempted to simulate the entire virus by applying periodic boundary conditions, or used coarse grained representations for a part or all of the system. While all of these techniques do provide useful results, they are also limited by their inherent approximations.

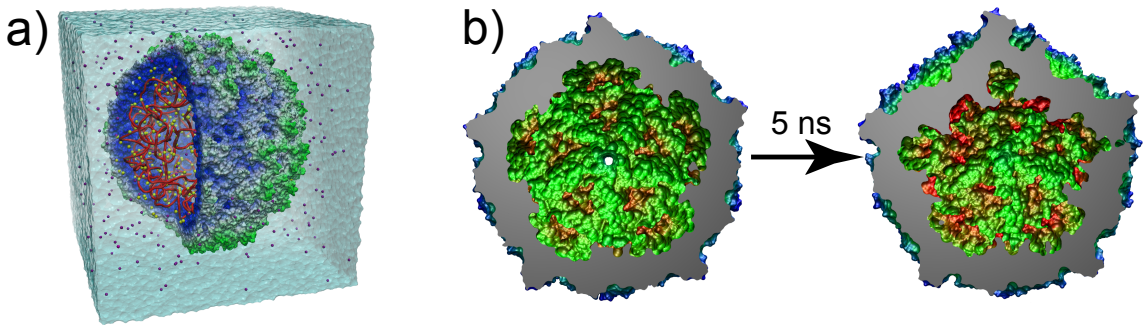

Figure 3: Simulating the all-atom structure of Satellite Tobacco Mosaic Virus (STMV). (**a**) The complete simulated system consists of the viral RNA genome (red) enclosed in the protein capsid, and of the solvation water box with the magnesium (yellow) and chloride (purple) ions. (**b**) The STMV capsid proved to be unstable when simulated without the RNA. The capsid collapses on the time scale of 5-10 ns, which agrees with various experimental observations and implies that the RNA plays an important role in the assembly and stability of STMV.

In order to meet the challenge of studying the behavior of complete viral particles through computational methods, the Resource has begun to perform all-atom molecular dynamics studies on complete viral particles (http://www.ks.uiuc.edu/Research/STMV/). The

first simulation of this nature was carried out on Satellite Tobacco Mosaic Virus (STMV), a small and well studied plant virus. Although STMV has been subjected to numerous experimental investigations, there are still a number of open questions on how it assembles and disassembles (both critical steps in the life cycle of any virus), and regarding the cause of its exceptional stability. The results of our initial simulations on STMV have been carried out and reported in a joint publication with our collaborator Alex McPherson [13]. By studying the stability of the complete virion and its isolated components, simulations by the Resource illustrated that previous speculation that STMV assembly was mostly capsid protein-driven [14] are likely incorrect; in tandem with recent experimental results from the McPherson lab [15] these molecular dynamics studies suggest instead a more concerted assembly mechanism starting with the assembly of RNA that subsequently recruits 60 capsid proteins to build a shell around itself [13]. These studies also allowed the elucidation of a number of other physical properties of the STMV particle that would be difficult or impossible to obtain through experimental means. The Resource has already begun further simulations to study how the STMV particle disassembles when infecting the cell; because of the virion's impressive stability the answer to this question has so far eluded experimentalists.

Gateway to the Nucleus

The nucleus of the eukaryotic cell is central to the cell's ability to function. It houses the genetic information of the cell and is responsible for the protection and proper expression of that information, as the genetic code is translated into useful cellular machinery. Isolation from the rest of the cell is vital in protecting that genetic information, but the nucleus also needs to communicate with the rest of the cell, exchanging proteins and RNA with the cytoplasm. The nuclear pore complex (NPC) is the gatekeeper of the nucleus. It enables the efficient transfer of molecules between the nucleus and cytoplasm, and at the same time it must protect the nuclear components by preventing harmful or unncessary cargo from traversing the nuclear envelope.

It is well known that the NPC uses shuttle proteins (called transport receptors) to move cargo across the nuclear envelope. However, the precise manner in which the NPC recognizes these transport receptors and allows them to cross, while preventing harmful molecules from doing so, remains a mystery. This is largely due to the immense size of the NPC [16], which has made many experimental techniques unviable. The Resource, however, has successfully begun to attack the enigmatic gating mechanism by means of computer modeling.

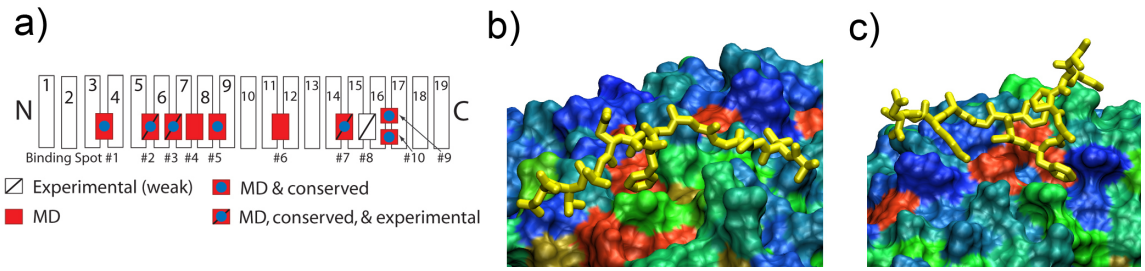

Figure 4: Binding of nuclear pore proteins to the surface of importin-beta. (**a**) A schematic of importin-beta showing all binding spots which were discovered or verified by the Resource in red. (**b**) Nuclear pore protein (yellow) binding to the surface of importin-beta between HEAT repeats 6 and 7. The surface of importin-beta is colored according to residue conservation, with red representing the most conserved residues. (**c**) Nuclear pore protein binding to the surface of importin-beta between HEAT repeats 8 and 9.

In order to gain insight into the mechanism of gating, the Resource has used the computer to investigate properties of transport complexes which are not accessible by experiments (http://www.ks.uiuc.edu/Research/npc/). Our pioneering simulations, a "definitive first in the field" [17], combine molecular dynamics simulations with protein sequence analysis to determine relevant binding spots for NPC proteins on the surface of transport receptors. By performing simulations of the NPC proteins in solution with the transport receptor, the Resource has obtained a detailed all-atom view of the binding of the NPC proteins. By subsequently aligning the sequences of transport receptors from several different species, we are able to determine which of these binding events are likely to

be relevant for NPC recognition. To date, experimental studies have identified binding sites [18–24] via other techniques, but the conditions necessary for the experiments imply that only the strongest binding spots are seen. The Resource is able to probe a greater level of binding sensitivity which reveals that transport receptors interact much more extensively with the NPC than previously realized. Our simulations on the transport receptor importin-beta (and its cargo) [25] have enabled us to verify three out of the four experimentally discovered binding spots [18–20]. They also led to the discovery of five novel binding spots which were hitherto unknown and which are believed to be relevant for in vivo NPC recognition. Furthermore, in a convincing display of the predictive capability of computational simulations for biology, the Resource identified another binding spot, independent of experimental aid, which was discovered concurrently by experimentalists and reported only shortly before the Resource published the discovery [21].

Control of Gene Expression

Each cell of an organism stores its complete genome, and has to continuously regulate the expression of genes based on which proteins are necessary for relevant cellular processes. The regulation of gene expression is carried out at different levels in the cell. One of the principal forms of regulation is via transcriptional control, where, in order to shut down a gene, the RNA polymerase is prevented from binding to the DNA binding site or "promoter", inhibiting the transcription of the DNA into mRNA. A classic example of this process is the *lac* operon in *E. coli*, a set of genes that code for enzymes involved in lactose uptake and catabolism [26]. Without lactose in the environment and thus no requirement to express these genes, the *lac* repressor protein (LacI) binds to distant sites in DNA and forms a loop with the intervening DNA, trapping the promoter of the genes inside the loop and blocking access to RNA polymerase, thereby preventing expression. LacI is the simplest genetic switch known and has been the paradigm of gene control since its discovery fifty years ago [27]. LacI has been extensively studied resulting in a wealth of genetic, structural and biochemical data [28, 29]; and yet the detailed mechanics of transcriptional repression by LacI remain unknown. DNA looping by proteins is a common mechanism for gene repression across all kingdoms of life [30,31]. Understanding the mechanisms used by LacI to arrest DNA in a looped configuration will shed light onto basic mechanisms of gene control.

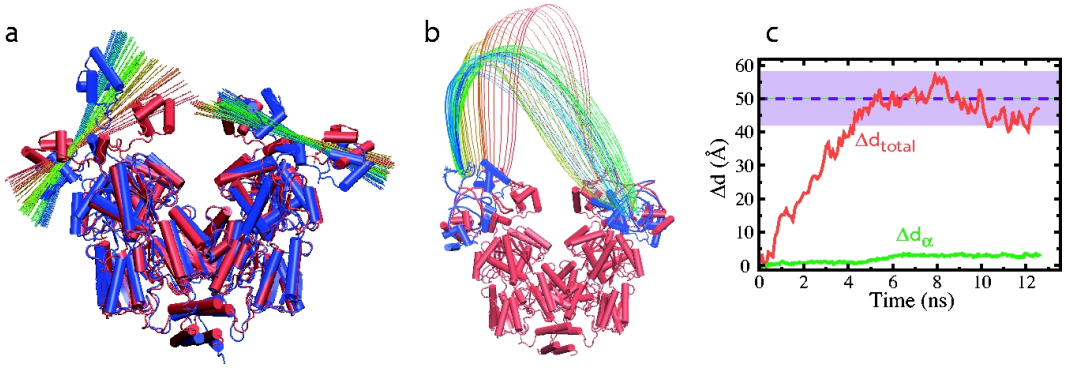

Figure 5: Head group and DNA loop motion during the multiscale simulation. The protein structure is shown before and after the simulation. The lines in (**a**) represent the long principal axis of each head group after every 200 ps of the simulation. The structure of the DNA loop in (**b**) is drawn with the same frequency. (**c**) The distance d between two fluorophores attached to the DNA loop in [32]. Δd_{total} shows the change in d during the simulation; Δd_α shows how d would change if the head groups were immobile with respect to the core domains. The gray band with the dotted line shows the experimental range and the average estimated for d [32].

The study of the LacI-DNA complex required the Resource to first furnish a complete structural model of LacI from disjoint and partially resolved crystallographic and NMR data [33–35]. This has been achieved and the resulting structure has been deposited in the Protein Data Bank (accession code 1Z04) [36, 37]. None of the experimentally

available structures contains information on the DNA loop, because its flexibility makes it difficult to resolve. The inclusion of the loop in the study of the complex is important because forcing the DNA into a looped configuration results in mechanical energy being stored in the DNA and transmitted to the protein. The mechanism by which the protein withstands the strain arising from the DNA and maintains it in a looped configuration is unknown.

The Resource has developed an elastic rod model of DNA that accounts for all the physical properties of DNA: electrostatics, intrinsic twist and bending anisotropy [36, 38–40]; and obtains the structure of the DNA loop based on the protein structure. Even with the structure of the DNA loop known, the size of the resulting protein-DNA complex is too large for molecular dynamics studies. To overcome these difficulties, the Resource has developed a multiscale method to study the interaction between protein and looped DNA (http://www.ks.uiuc.edu/Research/Multiscale/) which combines the developed elastic rod model of DNA with an all-atom description of the protein and protein-bound DNA resolved in the experimental structures [41]. The method builds an equilibrium structure of the DNA loop using the elastic rod model according to boundary conditions provided by the all-atom structure [36, 38–40], obtaining both the geometry of the DNA loop and the forces arising from the protein-DNA interaction. These forces are in turn incorporated into the molecular dynamics all-atom description of the protein. Iterative rounds of the two methods reveal the structural dynamics of the LacI-DNA system [37].

The multiscale simulations revealed the structural dynamics of the LacI-DNA complex [37]. Previously, it was assumed that the protein controlled the DNA by "opening" the cleft formed between the two dimers (Fig. 5). The results of the simulation show that the ability of the protein to maintain its grasp on DNA is due to the extreme flexibility of the DNA-binding domains, or head groups, while the dimers remain in the "V"-configuration observed in the crystal structure (Fig. 5). The behaviour shown in the simulations is also in agreement with experimental data [32], which had determined the distance between two distant base pairs inside the DNA loop. The data has previously been interpreted as an opening of the LacI cleft. The simulation reveals similar distances without opening of the cleft (Fig. 5), which shows that simple reorientation of the head groups can relieve mechanical energy stored in the loop.

The Resource continues to investigate the dynamics of the LacI-DNA complex. A recent study reported the energetics of longer DNA loops formed in the cell [40]. A collaboration with Rob Phillips (Caltech) has been established to further study the opening of the LacI cleft under force; and another one with Francisco Vanzi (Florence University) will explain the role of mutations in the head group linkers, which presumably affect flexibility of the head groups, on the ability of LacI to maintain its grasp on DNA. Finally, the study of other possible binding geometries of looped DNA is under way.

Subprojects

BTA UNIT: C

TITLE: Mechanisms of Gating and Selectivity of Aquaporins

KEYWORDS: water channels, membrane proteins, gating, selectivity, ion conduction

AXIS I: AXIS 1

AXIS II: AXIS 2

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

INVEST2: Fangqiang Zhu
DEGREE2: Ph.D.
DEPT2: Physics
NONHOST2:

INVEST3: Jin Yu
DEGREE3: M.S.
DEPT3: Physics
NONHOST3:

INVEST4: Yi Wang
DEGREE4: B.S.
DEPT4: Biophysics
NONHOST4:

INVEST5: Andrea Yool
DEGREE5: Ph.D.
DEPT5: Physiology and Pharmacology
NONHOST5: University of Arizona

INVEST6: Richard Neutze
DEGREE6: Ph.D.
DEPT6: Chemistry and Bioscience
NONHOST6: Chalmers University of Technology, Sweden

INVEST7: Per Kjellbom

DEGREE7: Ph.D.

DEPT7: Plant Biochemistry

NONHOST7: Lund University, Sweden

INVEST8: Ioan Kosztin

DEGREE8: Ph.D.

DEPT8: Physics

NONHOST8: University of Missouri

INVEST9: Chris Chipot

DEGREE9: Ph.D.

DEPT9: University of Henri Poincare

NONHOST9: France

% 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 [42, 43]. They are widely distributed in all forms of life. Through modulating water permeability of cellular membranes AQPs play an important role in water homeostasis of living cells. MD simulations of membrane-embedded, fully hydrated models of various AQPs with NAMD [44] have revealed novel mechansims for selective function of these channels. 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.

We have investigated the mechanisms that gate the water pores and the central pore of AQPs [**?**, 45]. In close collaboration with A. Yool (U. Arizona) who was the first to propose cGMP-induced ion conductivity of AQPs [**?**, 46], protein conformational changes induced by binding of cGMP were studied. Based on our simulations, a gating mechanism has been proposed: one of the cytoplasmic loops (loop D) plays a critical role in controlling the accessibility of the central pore to water and hydrated ions [?]. The conformation of this flexible loop is highly perturbed by freely diffusing cGMP during the simulations, mainly through strong interactions of the nucleotide with an arginine-rich region of the loop. Retraction of the D loops of the four monomers away from the central pore not only physically unblocks the entrance of the central pore, but also results in conformational changes of a ring of pore-lining, hydrophobic residues that form a gate and block the access of water. The helix bearing these hydrophobic residues is immediately connected to loop D, and it is very likely that the conformational coupling of loop D and this helix is the molecular mechanism of detecting cGMP binding inside the central pore [?].

The involvement of arginines in the gating mechanism was successfully verified by experimental measurements performed by our collaborator on a double mutant species in which two of the arginines had been knocked out [?].

The mechanism of phosphorylation mediated gating of water pores in a plant AQP was investigated through a close collaboration [45] with the group of R. Neutze (Chalmers U. Tech., Sweden). The x-ray structure of the channel in its closed form was used to study conformational changes triggered by phosphorylation. In the closed (unphosphorylated) form, loop D, which is 4-5 residues longer in plant AQPs vs. mammalian ones, is held in its closing position through hydrogen bonds with the N-terminus of the protein. Examination of simulation results using the visualization program VMD [47] shows that upon phosphorylation, the connection between the N-terminus and loop D breaks, and that the latter is free to undergo large conformational changes. These changes, in turn, result in the opening of the water pores through two complementary mechanisms: 1) displacement of loop D from the cytoplasmic mouth of the channel, and, 2) retraction of a hydrophobic residue (Leu197) from the pore [45].

Besides the investigation of gating mechanisms of AQPs, we also tackled the problem of selectivity in these water channels. A comparative study [48] was performed on two bacterial AQPs from the same species, i.e., *E. coli* GlpF and AqpZ, which are structurally highly homologous, but functionally different. While AqpZ is a pure water channel [49, 50], GlpF also conducts glycerol. In an earlier study [51], we calculated the PMF associated with permeation of glycerol through GlpF. Calculation of accurate free energy profiles from MD simulations presents an important challenging problem. The Resource collaborates with several theoretical groups, such as Chris Chipot (France) and Ioan Kosztin (U. Missouri) to overcome technical difficulties associated with such calculations.

Although AqpZ is not a glycerol channel, an artificially induced passage of glycerol through AqpZ was achieved in SMD simulations, from which we constructed the PMF of glycerol conduction along the pore,and identified the barriers that make AqpZ impermeable to glycerol under normal conditions [48]. The results indicate that a high barrier against glycerol permeation arises not only in the selectivity filter region of AqpZ, which is believed to account for the main structural difference between AqpZ and GlpF, but along the entire channel. Furthermore, the difference in substrate selectivities of AqpZ and GlpF may not be simply due to only those residues that directly line the channel [48]. To convert a water channel into a glycerol channel one may have to identify remote residues that do not directly interact with the permeant, but nevertheless control the channel diameter through shifting and tilting helices forming the pore [48].

BTA UNIT: C

TITLE: Gas Migration Pathways in Hydrogenase and Other Proteins

KEYWORDS: gas migration, hydrogen production, sampling and free energy algorithms

AXIS I: 2 9

AXIS II: 74C 74D 89

INVEST1: Jordi Cohen
DEGREE1: M.S.
DEPT1: Physics
NONHOST1:

INVEST2: Michael Seibert
DEGREE2: Ph.D.
DEPT2:
NONHOST2: National Renewable Energy Laboratory, CO

INVEST3: Maria Ghirardi
DEGREE3: Ph.D.
DEPT3:
NONHOST3: National Renewable Energy Laboratory, CO

INVEST4: Paul King
DEGREE4: Ph.D.
DEPT4:
NONHOST4: National Renewable Energy Laboratory, CO

INVEST5: Kwiseon Kim
DEGREE5: Ph.D.
DEPT5:
NONHOST5: National Renewable Energy Laboratory, CO

INVEST6: Anton Arkhipov
DEGREE6: M.S.
DEPT6: Physics
NONHOST6:

INVEST7: Rosemary Braun

DEGREE7: Ph.D.

DEPT7: Physics

NONHOST7:

% BTA $: BTA %

ABSTRACT: Many proteins interact with gas molecules, such as oxygen and carbon monoxide, to perform their function. Unlike in the case of other ligands, the reactive sites for 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 in understanding how proteins such as globins, oxygenases, cytochrome-oxidases and many others work. The direct applications of this knowledge also leads to very relevant work. For example, the knowledge of gas migration pathways inside the hydrogenase enzyme allows scientists to perform targeted mutations that lead towards the goal of a novel enzyme which can be used for affordable large-scale production of H_2 as a source of renewable energy.

The Resource has developed a novel algorithm, called implicit ligand sampling, for accurately predicting gas migration pathways inside any protein, a feat that was not previously possible [2]. This methodology was then applied to the case of myoglobin, a protein whose gas conduction properties have been well characterized experimentally and theoretically over decades. The computation used the NAMD [44] molecular dynamics software, in conjuction with the Resource's newly developed NAMD-G grid submission software, to farm out the computations for myoglobins of four different species over the TeraGrid computational resources. The gas migration pathways and energy barriers found using implicit ligand sampling applied to myoglobin match the known properties very well and furthermore found new pathways in areas of myoglobin that are hard to characterize experimentally [2]. This study now opens the gates for the characterization of gas migration processes in a large set of proteins for which experimental investigation is not practical and/or possible. A set of closely related studies also investigated oxygen and hydrogen gas migration inside hydrogenase from *Clostridium pasteurianum* [1, 52, 53]. 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 human-usable hydrogen-producing enzyme.

BTA UNIT: C

TITLE: Anthrax Toxin

KEYWORDS: anthrax toxin, protective antigen, molecular dynamics, pore

AXIS I: 7a

AXIS II: 66,89

INVEST1: Mu Gao
DEGREE1: Ph.D.
DEPT1: Beckman Institute
NONHOST1:

% BTA $: BTA %

ABSTRACT: Bacillus anthracis, the cause of anthrax, is one of the most lethal bacteria (http://www.ks.uiuc.edu/Research/anthrax/). The bacterium attacks the cells of the host's immune system, the so-called macrophages, as well as many tissue cells. For this purpose the anthrax bacterium releases three types of proteins, or toxins, into the blood stream of the host: protective antigen (PA), lethal factor (LF), and edema factor (EF) [54, 55]. To invade a host cell, the toxins take an intricate entry route that involves binding to cellular receptors and inducing the cell to internalize the toxins in a bubble like endosome. The bubble membrane wall is then punctured by PAs forming a pore upon an acidifying trigger from the host, and finally the LFs and EFs slip into the cell through the pore.

Although the general picture of the anthrax intoxication pathway has been established, many details still remain a mystery. One key question is how the PA-receptor complex initiates formation of the pore for ferrying the toxic cargo. Recently, a high resolution crystal structure of PA in complex with its cellular receptor capillary morphogenesis gene 2 (CMG2) resolved has become available [56], providing an opportunity to study the pore-formation mechanism with MD simulations.

The project required extremely extensive (136ns from 92,000 atom system) MD simulations that became feasible only through optimizing the Resource program NAMD for multi-processor (128) compuation. We have performed simulations of the PA-CMG2 complex with NAMD to study how acidic conditions in the endosome trigger conformational changes of the PA complex necessary for pore formation. We also carried out SMD simulations to understand the role of interactions between PA and CMG2 [57]. Simulations revealed that under neutral pH conditions PA orients one of its loops so that a positively charge amino acid (Arg344PA) interacts

optimally with a negatively charged amino acid (Glu122CMG2) on the receptor. The interaction between the pair, a so-called salt bridge, prevents the toxin from dissociation. Low pH conditions induce the protonation of two amino acids involving or neighboring the salt bridge (His121CMG2 and Glu122CMG2), which triggers rapid dissociation of PA from the receptor for the pore formation purpose. The amino acids, His121 and Glu122 of CMG2, and Arg344 of PA, appear to serve as a pH-sensitive switch that controls the pore formation.

BTA UNIT: C

TITLE: Control of Gene Expression

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

AXIS I: 2, 7a, 9, 28 (Gene control)

AXIS II: 42, 74g, 74h, 77, 89

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: Elizabeth Villa
DEGREE3: B.S.
DEPT3: Center for Biophysics and Computational Biology
NONHOST3:

% BTA $: BTA %

ABSTRACT: The mechanical manipulation of DNA is central to all aspects of genetic regulation in cells [30]. Regulatory proteins mechanically manipulate DNA away from its equilibrium configuration by introducing bends, loops, twists and supercoils [31] against mechanical strain arising from the DNA. In some cases, this is done to prevent other proteins from binding DNA resulting in gene expression. Little is known about how regulatory proteins are able to handle and withstand the forces stemming from DNA. In many cases, changes in DNA structure are known to occur in protein-DNA interaction but the structure cannot be resolved experimentally when the changes are too large or when the DNA becomes disordered, e.g., when regulatory proteins force DNA into loops.

The *lac* repressor (LacI) is a classic example of genetic regulation via DNA looping. It controls the function of the *lac* operon in *E. coli*, a set of genes involved in lactose catabolism [29]. The all-atom structure of LacI bound to its DNA binding sites has been determined, but with the DNA loop missing [33, 34]. 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 has developed a multiscale method [37] for the simulation of protein-DNA complexes involving long segments of flexible DNA. This method combines conventional molecular dynamics simulation using NAMD with a mathematical model of DNA, the so-called elastic rod model. 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 [36, 38–40]. An all-atom molecular dynamics simulation of the protein incorporates these forces using the SMD method [58], revealing the structural dynamics of the protein-DNA interaction. The geometry of the elastic rod model of the DNA loop and the forces arising from it are updated and exchanged every 10 ps of MD simulation. The multiscale method was applied to the LacI in complex with a 76 base pair elastic loop [37]. The MD part of the simulation encompassed 320,000 atoms, simulated with the NAMD program on 256 processors. The results of our simulation corrected a long held view of the mechanism of LacI. While it was assumed previously that LacI opened its cleft in a hinge-like massive motion between the dimers to control the ability of LacI to enforce the DNA loop. The simulations revealed that the ability of LacI to maintain the DNA in a looped configuration is actually due to the extreme flexibility of its DNA-binding head groups connected to the rest of the protein through flexible linkers, while the dimers are essentially immobile with respect to each other. The computationally modeled behavior is in good agreement with experiment [32], and provides a new interpretation of existing data.

BTA UNIT: C

TITLE: Light-induced signaling in LOV domains

KEYWORDS: photoreceptor, phototropin, LOV, QM/MM, molecular dynamics

AXIS I: AXIS 1

AXIS II: AXIS 2

INVEST1: Markus Dittrich
DEGREE1: Ph.D.
DEPT1: Beckman Institute
NONHOST1:

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

% BTA $: BTA %

ABSTRACT: Phototropins are photoreceptors crucial for the regulation of many essential processes (http://www.ks.uiuc.edu/Research/biological_photoreceptors/) in plants, such as phototropism, chloroplast relocation, and stomatal opening and closing [59]. The actual photosensitive elements in phototropins are the Light, Oxygen, and Voltage (LOV) sensitive domains. Absorption of blue light by the LOV domains leads to the formation of a covalent adduct state between the initially non-covalently bound chromophore flavin mono-nucleotide (FMN) and a cysteine residue of the protein. This in turn activates an attached serine-threonine kinase domain, eventually leading to downstream signal transmission by an as yet unknown mechanism. Most plant phototropins contain two distinct LOV domains, LOV1 and LOV2, which are known to exert different levels of influence on the signaling kinase domain of the protein. X-ray crystal structures of LOV domains in their ground and photoactivated states have recently been determined at high resolution [60–62].

To investigate LOV domain function and signaling, the Resource has pursued two complementary approaches. In the first one, combined quantum mechanical/molecular mechanical (QM/MM) simulations were used to examine the photoexcitation event leading to adduct formation at an electronic level [63]. The computational model consisted of approximately 20,000 atoms, 37 of which were

treated quantum mechanically at the HF/6-31G(2d,2p) level of theory. The simulations examined the singlet ground state, several excited triplet state intermediates, the singlet flavin-cysteinyl adduct state, and the pathways connecting them, providing a detailed view of the states' energetics and electronic properties.

Based on experimental data, several proposals regarding the proper adduct formation pathway have been put forward [62, 64, 65]. The conducted QM/MM simulations clearly identified a neutral triplet radical state as the physiologically relevant mechanism. Finally, it was found that the accumulation of spin-density on the cysteine sulfur in the neutral triplet radical state, combined with the large spin orbit coupling constant of the latter, provides an efficient intersystem crossing mechanism, possibly explaining why such a neutral radical species has not yet been observed experimentally.

In the second approach, all-atom molecular dynamics (MD) simulations were applied to investigate how the system evolves after the photoreaction has occurred. For this portion of the study, the MD program NAMD [44] was applied to a set of four systems of approximately 30,000 atoms each: The dark and light states of LOV1 and LOV2. Each system was solvated and ionized using the program VMD [47] and simulated using NAMD in the NPT ensemble. For each of these four systems five molecular dynamics runs of 12 ns each (240 ns total) were performed, starting from an equilibrated form of the crystal structure, to observe the different behavior of the LOV1 and LOV2 domains before and after the photoreaction.

The results of these simulations, described in [66], revealed two distinct mechanisms for photoactivation in LOV1 and LOV2. In simulations of LOV1 it was found that the photoreaction significantly increases the probability for salt bridge formation between two highly conserved residues, greatly stabilizing the surrounding surface of the protein. In the case of LOV2 the photoreaction caused major changes in the dynamics of a loop which interacts with the protein's Ja helix, which was previously shown to be involved in LOV2 activation [67]. Thus, the simulations reveal how photoactivation leads to different structural changes in LOV1 and LOV2, and suggest how these changes could lead to kinase activation.

The project became feasible only through close collaboration with the NCRR Resource at the Pittsburgh Supercomputing Center where an advanced NIH funded computer capable of large scale quantum chemistry calculations was made available.

BTA UNIT: C

TITLE: Mechanosensitive Ion Channels

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

AXIS I: AXIS 1

AXIS II: AXIS 2

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

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

INVEST3: Trudy A. van der Straaten
DEGREE3: Ph.D.
DEPT3: Beckman Institute
NONHOST3:

INVEST4: Umberto Ravaioli
DEGREE4: Ph.D.
DEPT4: Beckman Institute
NONHOST4:

% BTA $: BTA %

ABSTRACT: Perception of sound and regulation of blood pressure or cell volume are archetypical examples of biological processes mediated by mechanosensitive (MS) ion channels. This class of membrane proteins serve as molecular switches, opening and closing in response to stress conveyed through other proteinaceous structures of the cell or the cellular membrane itself [68, 69]. In bacteria, MS channels are thought to act as safety valves preventing cell burst upon osmotic shocks [70, 71].

Crystal structures of two MS channels have been solved: the closed form of the MS channel of large conductance (MscL) from *M. tuberculosis* [72] and the putative open form of the MS channel of small conductance (MscS) from *E. coli* [73]. Both channels

are activated by mechanical stress in the cell membrane, providing a controlled response to the osmotic pressure of the environment.

Despite the detailed view of the molecular architecture of MscS and new experimental results shedding light on its function [74–76], key questions remain unanswered. Is MscS only a safety valve? What residues are relevant for MscS gating? Is the crystal structure conformation really open? How does the closed state of MscS look? What is the role of the large MscS cytoplasmic domain?

Several multi-nanosecond, all-atom molecular dynamics simulations were carried out by the Resource to explore the dynamics of MscS in its native environment (protein, lipid bilayer, water, and ions: 224,000 atom system) as reported in [77]. Electrostatic properties and conduction of ions through MscS were explored by the Resource using a coarse-grained particle-based description based on the Boltzmann transport Monte Carlo method [78]. As recently reported [79], single channel current-voltage curves were computed over time scales of hundreds of nanoseconds for channel conformations derived from all-atom molecular dynamics simulations reaching an unprecedented overall simulation time of 5 microseconds. The coarse-grained simulations revealed that channel conformations similar to those of the crystal structure exhibit low conductance, whereas conformations reached after opening the channel by means of steered molecular dynamics simulations match experimentally determined conductances. However, while experiments indicate a slight preference for anionic currents, the simulated channel strongly selects anions over cations. Further all-atom simulations (`http://www.ks.uiuc.edu/Research/MscSchannel/`) are currently being carried out in which different electrostatic fields are applied to the system and measured using the particle-mesh-Ewald method implemented in NAMD [44] and VMD [47]. In addition, modeling of missing residues in the crystal structure of MscS is being guided by data obtained from electron paramagnetic resonance experiments performed by our collaborator, Eduardo A. Perozo.

BTA UNIT: C

TITLE: Assembly of High-Density Lipoproteins

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

AXIS I: AXIS 1

AXIS II: AXIS 2

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: ABSTRACT

High-density lipoproteins (HDL) (http://www.ks.uiuc.edu/Research/Lipoproteins/) are protein-lipid particles 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. Low levels of HDL have been implicated in an increased risk of coronary heart disease. The production, transformation and degradation of HDL particles 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 to form discoidal lipoprotein particles. Continued efflux of cholesterol and lipids as well as the esterification of cholesterol results in the transformation of the discoidal particles to mature spherical particles, which transport the cholesterol to the liver [80].

Two X-ray crystal structures of lipid-free apo A-I have been determined, one of the entire 243 residue protein [81], and the other of the 200 residue lipid binding domain [82]. 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, exhibiting a variety of shapes and sizes, structural studies have proven difficult [83]. 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 our 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 [84]. We utilize these homogeneous and well-characterized nanodisc particles [85, 86] in our molecular dynamics studies [87, 88].

In order to investigate the assembly of discoidal HDL particles (i.e. nanodiscs), the Resource has developed a coarse-grained (CG) protein-lipid model [88] and implemented it into the molecular dynamics program NAMD [44]. Due to the long timescales required for large molecular movements and assembly processes, a CG model was crucial in allowing one to simulate out to microscecond timescales (all-atom molecular dynamics is limited to nanoseconds). The stability of the CG model was tested by running CG molecular dynamics simulations on pre-assembled discoidal HDL particles (63,000 CG beads simulated for 62.5 to 285 ns) and were found to reproduce overall structural features such as bilayer thickness and disc diameter at various temperatures. The CG model was then used to simulate the assembly of nanodiscs starting from various configurations. During the 1.5 microsecond simulations the lipids quickly aggregated together and the proteins slowly attached themselves to either side of the lipid bilayer [88].

BTA UNIT: C

TITLE: Sequencing DNA with a nanopore device

KEYWORDS:

AXIS I:

AXIS II:

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

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

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

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

% BTA $: BTA %

ABSTRACT: A nano-meter size pore, a so-called nanopore, can be manufactured in thin inorganic membranes. The 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 [89]. 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

(http://www.ks.uiuc.edu/Research/nanopore/). Atomic-scale modeling was carried out in three directions: (i) we determined conditions for permeation and trapping of dsDNA in silicon nitride pores; (ii) investigated the influence of the nanopore confinement on the conformation of ssDNA; (iii) developed an atomic model to be used with the second-generation nanopores, that are integrated with MOS-capacitor membranes, to obtain a better resolution. The novelty of the project has mainly been in the combination of inorganic (silicon nitride and silicon dioxide) and organic materials (physiological solution, DNA) as well as in the enormous computational requirements.

Translocation of dsDNA through silicon nitride pores. Molecular dynamics simulations were carried out to determine the threshold diameter for permeation of dsDNA through silicon nitride (Si3N4) pores. A 1.3 V bias was applied to drive dsDNA through double-conical pores in 10-nm-thick membranes. The translocation of dsDNA was observed to halt when the pores narrowed to 2.5-nm in diameter, which was recognized as a threshold diameter. We have also discovered that the permeation of dsDNA is conditioned not only by the nanopore diameter, but also by the strength of the electric field [90,91]. Indeed, when the force exerted on DNA by the electric field exceeds the threshold for the overstretching transition, the DNA helix deforms such that it can pass through the pore constriction smaller than 2.5-nm in diameter. This discovery opened up a possibility for creating a nanopore trap for a single dsDNA molecule. Simulations assessing the efficacy of such a trap are being carried out.

Conformational dynamics of ssDNA is a 1-nm-diameter pore. Through molecular dynamics simulations we have discovered that when ssDNA is confined to a pore smaller than 1.5 nm in diameter, the DNA bases tilt collectively towards the 5' end [91, 92]. In a 1-nm-diameter pore, the DNA bases aa well as their sequence specific dipole orient along the pore axis and the fluctuations of the DNA conformation are greatly suppressed, both factors contributing to a stronger sequence-depended signal. By selectively pruning interactions between DNA, electrolyte, and the nanopore we have determined that in pores less than 1.5 nm in diameter the counter-ions condense at the DNA backbone, reducing the DNA translocation velocity by two orders of magnitude. We also observed that the translocation of ssDNA through a 1.0-nm-diameter pore creates periodic alternations of the electrostatic potential in the pore that can be used to accurately count the number of nucleotides permeating through the pore. We are investigating the possibility of determining the type of DNA nucleotides confined in a 1.0-nm pore by averaging over their conformational fluctuations.

Atomic-scale model of silica surface. The Resource has been extending its rather

successful computational methodology used in silicon nitride nanopores [90, 91, 93–95] to solid state nanopores in MOS-capacitor membranes. A major challenge for simulating those systems is the lack of a refined mathematical description, so-called forcefield, for the interaction energy between the silica (SiO2) nanopore surface with bio-macromolecules in aqueous environment. Another difficulty is that the properties of silica surface strongly depend on the particular arrangement of the surface atoms. We have developed a silica surface model to use with MD simulations that accurately reproduces atomic-scale roughness of the surface and its hydrophobicity [96]. The silica surfaces were obtained using different annealing cycles on an amorphous silica slab. The interactions between water and silica were calibrated by measuring a contact angle of 1000, 2000 and 4000 water-molecule droplets with different silica surfaces. Our model is a close representation of the real surface, providing a good compromise between the heterogeneity of silica and the magnitude of the intermolecular forces. This model of silica will be deployed in our computational studies of DNA permeation through nanopores in a MOS capacitor.

BTA UNIT: T

TITLE: Empirical Nanotube Model for Biological Applications

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

AXIS I: 9

AXIS II: 39, 77

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

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

INVEST3: Umberto Ravaioli
DEGREE3: Ph.D.
DEPT3: Department of Electrical and Computer Engineering and Beckman Institute
NONHOST3:

INVEST4: Yan Li
DEGREE4: Ph.D.
DEPT4: Physics
NONHOST4:

INVEST5: Slava Rotkin
DEGREE5: Ph.D.
DEPT5: Physics
NONHOST5: Lehigh University

INVEST6: Michael Strano
DEGREE6: Ph.D.
DEPT6: Chemical and Biomolecular Engineering and Beckman Institute
NONHOST6:

% BTA $: BTA %

ABSTRACT: Carbon nanotubes (CNTs) are rolled sheets of graphite discoved in 1991 [97]. With narrow nano-size diameter, CNT can grow as long as several microns in length. CNTs made from a single graphite sheet are referred to as single-walled carbon nanotubes (SWNTs). Due to their superb electronic, thermal, chemical, as well as mechanical properties, SWNTs have shown extraordinary potential in bioengineering, biomedicine and nanotechnological applications, in particular, as drug delivery devices [98–100] and biosensors [101,102]. It has been shown that SWNTs can shuttle various cargo including water [103], protons [104], polymers [105], and nucleic acids [106] across cellular membranes, opening a new path for drug delivery. In addition, the optical signals of SWNTs are sensitive to environment changes. Monitoring the changes in the optical signal allows sensitive detection of target biomolecules, qualifing nanotubes as bio-sensors [101]. Theoretical and computational modeling of the interaction between SWNTs (`http://www.ks.uiuc.edu/Research/nanotube/`) and biomolecules can shed light on the properties of CNT-based nano-biological systems, and establish new concepts for controlling/tuning the performance of such systems to design and optimize these devices.

To characterize the interactions between SWNTs and biomolecules in MD simulations, it is crucial to consider the polarizable nature of SWNTs due to their highly delocalized pi-electrons. To address this issue, the Resource has developed a semi-empirical method [107], namely a tight binding model, that takes into account the polarization effect of SWNTs at low computational cost. In this method, the induced atomic partial charges, i.e, the dielectric response of SWNTs, are predetermined through density functional theory calculations to improve the quality of electrostatics in the existing MD forcefield. Key properties such as electronic energy spectrum and screening constants, computed with the new tight binding approach agree well with results from first principle calculations despite its computational simplicity [108, 109]. The method has been employed to study the systems of water-SWNTs [107–109] and ion-SWNTs [110], both of which are of fundamental importance for biological applications. In the water-SWNT systems, the atomic partial charges on the nanotube edges are found to greatly contribute to the total interaction energy, while the polarization of the SWNT significantly lowers the electrostatic energy in the center of the tube, e.g., for ions [109]. A simulation of a SWNT with a potassium ion revealed a THz oscilation of the ion, demonstrating the accuracy and efficiency of the polarizable SWNT model [96]. The Resource is currently making efforts to implement the semi-empirical method into NAMD [111], so that the electronic response of SWNTs can be evaluated on-the-fly during MD simulations. The application of this method to larger systems such as DNA-SWNT complexes are now being investigated.

BTA UNIT: C

TITLE: Structure and mechanism of photosynthetic unit of Rhodobacter sphaeroides

KEYWORDS: Cell tumography, atomic force microscopy, systems biology, bioenergetics

AXIS I: AXIS 1

AXIS II: AXIS 2

INVEST1: Melih Sener
DEGREE1: Ph.D.
DEPT1: Dept. Physiology and Biophysics
NONHOST1: Cornell U. Weill Medical College, NY

INVEST2: Neil Hunter
DEGREE2: Ph.D.
DEPT2: Dept. Molecular Biology and Biotechnology
NONHOST2: U. of Sheffield, UK

% BTA $: BTA %

ABSTRACT: ABSTRACT

The photosynthetic unit (PSU) of the purple bacterium Rhodobacter sphaeroides is a spherically shaped intrusion of the inner cellular membrane with a diameter of 60 nm. The PSU contains mainly three proteins, light harvesting system I (20) and II (200), as well as the photosynthetic reaction center (20) that are structurally known. Our collaborator in Sheffield succeeded recently to image flattened pieces of the PSU by means of atomic force microscopy, the image permitting the placement of the stated, structurally known proteins in the PSU fragments. With our NY collaborator (until recently a postdoctoral researcher at the Resource) we succeeded to combine the images from multiple fragments to reconstruct an entire spherical PSU, placing all three proteins. The resulting atomic resolution structure contains about 240 proteins with over 6000 chlorophylls, but not yet water and lipids. The structure permitted us to describe the so-called light harvesting process for the entire PSU by which light absorbed by any chlorophyll is funneled towards the reaction center proteins and utilized to charge the PSU membrane. The research may become the first instance in which an entire cellular organelle made of hundreds of proteins is described both structurally and functionally in atomic detail.

BTA UNIT: T, D, S

TITLE: BioCoRE

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

AXIS I: 9

AXIS II: 42, 51, 89

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 [112] (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 handheld PDA) scientists create projects in which all private data is secure and is 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.

Major BioCoRE developments in the past year include a programming interface for supercomputer job submission that allows other programs to easily use BioCoRE to submit computational jobs, and a plugin for VMD that allows users to interact with BioCoRE in a convenient way. In addition, BioCoRE became more "open" in the past year – non-members can now view information about public projects within BioCoRE without registering, and a new "guest user" has been added where users can register for BioCoRE much more quickly than before.

Future BioCoRE efforts will focus on tighter integration with other biomedical packages, such as VMD and NAMD.

BTA UNIT: T

TITLE: VMD

KEYWORDS:

AXIS I:

AXIS II:

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

% BTA $: 10 %

ABSTRACT: VMD [47] 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 extended with new structure building tools that assist biomedical researchers with building all-atom molecular models and preparing simulations. VMD now processes structures with multiple conformations, special bonds, biological unit transformations, and can automatically regenerate some types of missing information. A new volumetric map generation tool synthesizes volumetric density and occupancy maps from molecular dynamics trajectories for use in analysis and visualization. New graphical representations allow researchers to visualize very large molecular models with reduced detail and efficient rendering algorithms. Several new plugins provide additional molecular and volumetric file format support.

More than 10,300 new users registered and downloaded VMD 1.8.3 since the previous progress report. As of April, 2006, over 17,200 unique users had registered and downloaded this version. Three plugin updates were released for VMD 1.8.3 since the previous progress report. Over 11,400 unique users have downloaded developmental versions of VMD since the last progress report. The latest version, VMD 1.8.4, was released on April 16, 2006.

Ongoing VMD developments include additional tools for structure building, multiple sequence alignment, and simulation analysis, new and improved graphical representations, and multiprocessor performance improvements. The next release of VMD is planned for the summer of 2006.

BTA UNIT: T

TITLE: NAMD: Scalable Molecular Dynamics Software

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

AXIS I: 9

AXIS II: 42 89

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

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

INVEST3: David Kunzman
DEGREE3: B.S.
DEPT3: Computer Science
NONHOST3:

INVEST4: Chee Wai Lee
DEGREE4: M.S.
DEPT4: Computer Science
NONHOST4:

INVEST5: Chao Mei
DEGREE5: B.S.
DEPT5: Computer Science
NONHOST5:

% 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 [44, 111]. 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 15,000 registered users as both source code and convenient precompiled binaries. A new publication [44] documents the current capabilities, algorithms, and design of NAMD and provides examples of applications ranging from a tutorial exercise to the large multiscale simulation of [37].

NAMD 2.6b1 was released in July 2005 and has been downloaded by over 3600 users, 650 of whom are NIH-funded. This release added binaries for Linux on Itanium, Altix, and Opteron/Athlon 64/EMT64 and improved serial performance by 30% on IBM AIX and 70% on Mac OS X. Memory usage for simulations of large, highly bonded structures such as covalent crystals and for sparse simulations such as coarse-grained models has been greatly reduced without affecting the performance of typical biopolymer simulations. The new Adaptive Biasing Force method [113], implemented in Tcl, efficiently calculates the potential of mean force along a reaction coordinate by applying adaptive biasing forces to provide uniform sampling. With new Tcl-based boundary potentials, scripted forces may be efficiently applied individually to large numbers of atoms. Finally, the OPLS force field is now supported.

The new 2000-processor Cray XT3 at the Pittsburgh Supercomputing Center (PSC) became available for NAMD porting in August, 2005. As the XT3 employs both a completely new interconnect and a completely new operating system, we anticipated (and encountered) difficulties, but with assistance from Cray and PSC staff NAMD now runs with most features (the XT3 lacks, e.g., the sockets needed for interactive simulations) and with good performance and scalability (tuning continues). Similarly, the single-rack BlueGene/L system at the San Diego Supercomputer Center (SDSC) also became available, allowing the completion of the NAMD port (previously used only for performance tuning) by Resource personnel. It will be impossible to distribute NAMD binaries for the Cray XT3 or BlueGene/L since due to the lack of dynamic linking on this platforms any system upgrade would require recompilation. Finally, NAMD has been ported to the new Mac OS for Intel processors and to the x86-64 version of Windows, both natively for full performance.

BTA UNIT: T

TITLE: NAMD-G: Grid Technology for NAMD Simulations

KEYWORDS: grid computing, molecular dynamics simulation

AXIS I: 9

AXIS II: 42 89

INVEST1: Jordi Cohen
DEGREE1: M.S.
DEPT1: Physics
NONHOST1:

INVEST2: Michelle Gower
DEGREE2: M.S
DEPT2: National Center for Supercomputing Applications
NONHOST2:

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

INVEST4: Richard Kufrin
DEGREE4: M.S
DEPT4: National Center for Supercomputing Applications
NONHOST4:

INVEST5: James Phillips
DEGREE5: Ph.D.
DEPT5: Beckman Institute
NONHOST5:

% BTA $: BTA %

ABSTRACT: NAMD-G is a newly-developed supporting technology for the Resource's popular molecular dynamics program NAMD (http://www.ks.uiuc.edu/Research/namd/). NAMD-G greatly simplifies the work of NAMD users needing to run large numbers of separate simulations, either to study variation in dynamics across a family of proteins or simply for improved sampling. The development of NAMD-G is a close collaboration between Resource scientists and technical staff from the National Center for Supercomputing Applications (NCSA, http://www.ncsa.uiuc.edu/). A prototype of NAMD-G was successfully demonstrated at the SC2005 conference in November and the full version is currently being used for production simulations on NCSA machines by select scientists, who provide substantial feedback to the developers. Wider releases of NAMD-G will occur in stages as bugs and feature requests are resolved, with full public release anticipated by August, 2006.

NAMD-G relies on the services of the TeraGrid (http://teragrid.org/) to automate authentication, data transfer, and job management for NAMD users. Authentication is implemented through Grid Security Infrastructure (GSI), a portion of the Globus Toolkit that uses public key cryptography as the basis for secure authentication and communication. Furthermore, NAMD-G makes use of the NCSA-developed MyProxy credential repository, which provides simplified management of security credentials. For file transfers, NAMD-G uses the NCSA-developed uberFTP client, which can automatically wait for archived files to be read from tape on NCSA's mass storage system. NAMD jobs are managed and submitted to remote supercomputers through Condor-G, a "Globus-aware" subset of the more general Condor system. Another component of Condor called DAGMan (Directed Acyclic Graph Manager) provides the functionality used by NAMD-G to automate a sequence of runs that take a simulation from start to finish.

BTA UNIT: T

TITLE: NAMD-lite

KEYWORDS: molecular dynamics simulation, methods development

AXIS I: 9

AXIS II: 42

INVEST1: David Hardy
DEGREE1: M.S.
DEPT1: Beckman Institute
NONHOST1:

% BTA $: BTA %

ABSTRACT: NAMD-lite (http://www.ks.uiuc.edu/Development/MDTools/namdlite/) is a sequential molecular dynamics program that is compatible with the primary file types of parallel NAMD (http://www.ks.uiuc.edu/Research/namd/) and a subset of its configuration options. The source code of NAMD-lite is organized as a collection of well-documented C libraries to facilitate the development of new methods. The clean, modular design enables scientists to implement and test novel computational techniques that would not be easily done using the parallel NAMD source. To further enhance its appeal as a code development base, NAMD-lite uses source code developed independently from the parallel NAMD source and is licensed under the University of Illinois/NCSA Open Source License (http://www.opensource.org/licenses/UoI-NCSA.php) to provide scientists complete freedom over all of their code modifications. The initial release of the NAMD-lite source code was in September, 2005, and it included web documentation for its computational libraries and an online tutorial. The original release has been followed by several maintenance releases to fix bugs and alpha releases that introduce new features.

The primary purpose of NAMD-lite is to develop a set of reusable code modules towards the implementation of a sequential molecular dynamics program and related tools. The priority is to create molecular dynamics codes that are easy to understand and modify, enabling straightforward design and testing of new methods. Two innovative methods, one related to computation and the other to program infrastructure, are presently available with the current release of NAMD-lite. The multilevel summation method [114] performs the fast evaluation of electrostatic forces. The MDAPI (molecular dynamics application programming interface, http://www.ks.uiuc.edu/Development/MDTools/mdapi/) provides the interfacing

of a front end with a molecular dynamics engine and is intended for use in a future release of parallel NAMD.

NAMD-lite can also be used to facilitate science projects. One such use has been for the simulation of an ion passing through a single-walled carbon nanotube, requiring the implementation of novel force evaluation and time integration techniques, the computational followup to earlier theoretical work [108, 109]. Another ongoing project is the development of polarizable force fields and the implementation of some recently published fast methods for self-consistent polarizability [115].

BTA UNIT: S

TITLE: Computational Facility

KEYWORDS: parallel computing, visualization, network

AXIS I:

AXIS II:

INVEST1: Tim Skirvin

DEGREE1: B.S.

DEPT1: Theoretical and Computational Biophysics

NONHOST1:

% BTA $: 10%

ABSTRACT: Enhancements to the Resource's computational facility (http://www.ks.uiuc.edu/Development/Computers/), have focused in four major areas in the last year: the creation of a formal Visitor Center to support external users of the Resource, upgrading our developer workstations, basic server infrastructure upgrades, and the implementation of a new, easier-to-use web design.

The most important change in the last year has come with the creation of an official Visitor Center. At present, 45 of the 94 users on our systems are visitors; but they have historically used the same office and computational facilities as our regular users, which in many cases has not been appropriate. Now, using new space provided by the Beckman Institute, we can offer visitors an impressive facility to perform short- and long-term collaborations with the Resource. As most visitors prefer to use their own equipment for most work, the facility offers docking stations with 21" LCD monitors, to allow for up to four simultaneous visitors to best utilize their personal laptops. Two graphics workstations installed in the Center offer easy access to our array of computational and visualization resources.

For the second year running, the most obvious change in our environment has been an upgrade to the Resource's desktop workstations, specifically those of the professional software developers, now each equipped with a Sun Ultra 20 workstation. While these desktops are not capable of powerful 3D graphics, they are remarkably quiet, stable, and powerful machines, perfect for the needs of our programmers and administrators. 13 of these systems have been deployed, replacing the 5-year-old Athlon 1333 workstations. These upgrades will ensure that our developers will not be limited by their desktop workstations until the end of our grant period.

Our third major change in the last year has come behind the scenes, as we continued to modernize our existing server infrastructure. We have retired all of our non-rackmount server hardware, increasing our system reliability and freeing valuable space in our server room. To replace our old servers, we have purchased just two additional SunFire V240 servers and an additional 3.5 TB of disk, bringing our total disk capacity to 20 TB.

The final major change in our computational environment has been an upgrade to the look and feel of our main web site. After nearly two years of work, we went live with the new design on 08 Mar 2006. The new design incorporates the most important web technology improvement of the last five years – cascading stylesheets – to offer a web site that is easier to navigate, filled with more useful content, is more responsive, and is even easier to maintain. We look at the web site as our crown jewel, and the new design has given it a much-needed polishing.

Our computational growth has once 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) increased from last year by about 500,000 SUs to a total of 3.6 million. Adjusted for processor speed, this translates to a 30% increase in compute power since last year. This time is supplemented by our local compute clusters, which remain unchanged from last year.

BTA UNIT: C

TITLE: Gateway to the Nucleus

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

AXIS I: AXIS 1

AXIS II: AXIS 2

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

INVEST2: Murray Stewart
DEGREE2: Ph.D.
DEPT2: Laboratory of Molecular Biology
NONHOST2: Univ. of Cambridge

INVEST3: Andrej Sali
DEGREE3: Ph.D.
DEPT3: Biopharmaceutical Sciences
NONHOST3: Univ. of California at San Francisco

% BTA $: BTA %

ABSTRACT: Protection of genetic information is extremely important for the function of the cell. The DNA inside the cell nucleus is the blueprint used by the cell to create the host of molecular machinery that carries out all cellular tasks. The nuclear pore complex (NPC) is the first and foremost guard in that protection. It is a very large macromolecular assembly of proteins that sits in the nuclear envelope and determines which cellular materials may pass into and out of the nucleus. The NPC is perhaps the largest protein structure in eukaryotic cells [116], and as a result, it is difficult to study experimentally. Thus, 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 [116–121]). It is hypothesized that 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 studied the transport receptor importin-beta in the presence of nuclear pore proteins using a combination of molecular dynamics simulations and bioinformatics. NAMD [44] was used to perform molecular dynamics on the importin-beta in a solution with nuclear pore proteins whose concentration ranged from 50-90 mM. Complete systems averaged roughly 150,000 atoms and were simulated between 20-50 ns each, posing a significant challenge and requiring extremely efficient computing on large (128-256 processor) machines. In addition, sequences of importin-beta from eight species were aligned, and conserved residues on the importin-beta surface were thus determined and visualized using VMD [47]. During the course of the simulations, the nuclear pore proteins were found to bind to the surface of importin-beta. Binding of proteins near the conserved residues on the surface indicated an increased likelihood that the observed binding is relevant to in vivo NPC recognition of importin-beta. Using this combination of simulations and sequence information, the Resource was able to confirm three of the four importin-beta binding spots which were experimentally known [18–20]. We also predicted five novel binding sites on the importin-beta surface [25]. Furthermore, in a definitive display of the predictive power of computational science, we identified one binding spot independent of and concurrently with experimental researchers, who reported the binding spot [21] only a few months prior to the Resource. The work has prompted further simulations by the Resource on two other transport receptors, NTF2 and cse1p, which will further help to reveal how interactions between the NPC and transport receptors determine the gating mechanism of nuclear transport.

BTA UNIT: C

TITLE: Multiscale Modeling of Bacterial Ribosome

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

AXIS I: 7 a.

AXIS II: 77

INVEST1: Emma Falck
DEGREE1: Ph.D.
DEPT1: Beckman Institute
NONHOST1:

INVEST2: Leonardo Trabuco
DEGREE2: B.S.
DEPT2: Biophysics
NONHOST2:

INVEST3: Elizabeth Villa
DEGREE3: B.S.
DEPT3: Biophysics
NONHOST3:

INVEST4: Joachim Frank
DEGREE4: Ph.D.
DEPT4: Howard Hughes Medical Institute
NONHOST4: Wadsworth Center, NY

% BTA $: BTA %

ABSTRACT: The ribosome [122] is a cellular machine that synthesizes proteins based on genetic instructions. The genetic code contained in an mRNA molecule is translated into a protein using tRNAs. Each tRNA is charged with an amino acid that matches the tRNA's anticodon – a triplet of bases matching a codon in the mRNA. 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. Because of its fundamental role in constructing cellular structures, the bacterial ribosome is an important target of antibiotics. Hence, 50% of all research on antibiotics is focused on the ribosome.

Currently the most successful approaches to study ribosomes are cryo electron microscopy (cryo-EM) [123] and X-ray crystallography [124]. Cryo-EM offers insights into the function of the ribosome, currently at a resolution of 7 Angstroms or worse. X-ray crystallography yields atomic-scale structural information [124]. 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 the peptidyl bond formation. The whole machinery consists of ribosomal RNAs, about 50 ribosomal proteins, tRNAs, mRNA, ions, and additional protein factors.

Together with Dr. J. Frank, a pioneer in the area of cryo-EM experiments on the ribosome, we have launched a multiscale modeling project to investigate some of the currently unknown details in protein synthesis, e.g., in the recognition step, the exit of the nascent peptide, and the exit of tRNAs from the ribosome. Our first goal is to simulate an all-atom model of the *E. coli* ribosome. This model of a ribosome immersed in a salt buffer is based on a recent X-ray structure [124] and contains approximately 3,000,000 atoms. Modeling such a large structure with flexible parts, catalytic RNAs with modified bases, and structurally and functionally important divalent cations poses a great challenge, which has already resulted in several useful tools for modeling large biomolecular systems. Our further goals include constructing a coarse-grained model, as well as a multi-resolution model, that will allow us to focus on important functional regions at a moderate computational cost. We are also interested in fitting atomic-resolution X-ray structures into cryo-EM maps, which will enable us to study the ribosome during specific stages of protein synthesis captured by cryo-EM.

BTA UNIT: C

TITLE: Molecular Dynamics Simulations of Satellite Tobacco Mosaic Virus

KEYWORDS: Virus, Viral assembly, STMV, RNA, Protein aggregates, Large simulations

AXIS I: AXIS 1

AXIS II: AXIS 2

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: PhD
DEPT3: Molecular Biology and Biochemistry
NONHOST3: University of California, Irvine

% BTA $: BTA %

ABSTRACT: The Satellite tobacco mosaic virus (STMV) is a small satellite virus known to coinfect several species of plants in the presence of tobacco mosaic virus (TMV). Previous experimental studies on this system have provided a high resolution crystal structure for the viral capsid, a somewhat lower resolution structure that allows placement of the backbone of encapsidated RNA [14, 125], and AFM data on the properties of the intact virion and isolated RNA [15,126]. The presence of this data has allowed some hypotheses to be formed on how the virus assembles, but this mechanism is not completely clear, and there is even less information available on how the capsid might *disassemble* upon entering a host cell. The application of all-atom molecular dynamics to this system (`http://www.ks.uiuc.edu/Research/STMV`), while challenging, would both provide specific data on the assembly and disassembly of this virus, and aid in the general development of applying molecular dynamics to viral systems.

A first series of the all-atom MD simulations of the complete STMV particle has been performed by the Resource using the program NAMD [44] and reported in a

joint publication with our collaborator Alex McPherson [13]. In these simulations, which included up to 1,060,000 atoms, the dynamics of the full virion, as well as its components (the capsid and the RNA genome) separately, were observed on a time scale of 10 ns, and analyzed using VMD [47]. Overall, the simulations covered 55 ns of viral dynamics, which required 35 processor-years of computational time at the NCSA Altix supercomputer. The MD study of the complete STMV particle allowed the elucidation of a number of physical properties of the virus, such as the rates of water and ion transport within the particle, distribution of the electric field around the virus, and correlations between the motions of the unit proteins composing the STMV capsid. These characteristics would be difficult or impossible to obtain through experimental means. In the simulations of the complete virion and its isolated components, it was found that the RNA genome of STMV was structurally stable on its own, as was the complete virus, but the capsid without the RNA core was very unstable and collapsed within 5 to 10 ns. These findings agree with recent experimental results from the McPherson lab [15] suggesting that STMV assembly is not mostly capsid protein-driven as assumed previously [14]. Instead, it seems more likely that the protein capsid of STMV arranges itself around a partially preformed RNA core using a concerted assembly mechanism.

The Resource has already begun further simulations to study how the STMV particle disassembles when infecting the cell; because of the virion's impressive stability the answer to this question has so far eluded experimentalists. In addition, future studies in collaboration with Jack Johnson (The Scripps Research Institute) and Xiaowei Zhuang (Harvard U.) will attempt to apply these techniques to somewhat larger, human viruses such as poliovirus.

BTA UNIT: C

TITLE: Transport Mechanism of Lactose Permease

KEYWORDS: co-transporter, membrane protein, salt bridge, proton gradient

AXIS I: AXIS 1

AXIS II: AXIS 2

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: Beckman Institute
NONHOST3:

% BTA $: BTA %

ABSTRACT: Lactose permease (LacY) is an integral membrane protein (`http://www.ks.uiuc.edu/Research/Categories/Membrane/`) that uses the cell's electrochemical proton gradient to actively transport substrates across the cell membrane [127–129]. This protein plays a critical role in transmembrane traffic, and, therefore, is crucial for a healthy metabolism of a wide range of living organisms, including human beings. Malfunction of this transporter is associated with various pathophysiological conditions, such as diabetes and depression in humans [130,131]. Since the first functional characterization of LacY in 1956 [132], various biochemical, biophysical, and structural biological studies [131, 133, 134] have 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 to the protein 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.

Recent crystal structures of LacY from *E. coli* with [130] and without [135] lactose and a large body of experimental data have laid the basis for computational investigation of the lactose/proton co-transport mechanism of LacY. Using NAMD [44] we have initiated molecular dynamics simulations [136] of LacY investigating key steps of the transport scheme. A model of LacY embedded in a fully hydrated lipid bilayer, which provides a native environment for the protein and, thus, allows the system to evolve conformationally, was used for simulations. In order to properly describe the conformational response of lacY, the protein needs to be modeled in its natural environment, i.e., embedded in a sufficiently hydrated lipid bilayer, which results in system sizes of 130,000 atoms or more. The results analyzed by VMD [47] strongly implicate the residue Glu269 as one of the main proton translocation sites, whose protonation state controls several key steps of the transport process. A critical salt bridge between Glu269 and Arg144 was found to keep the cytoplasmic entrance open, however, via a different mechanism from the currently accepted model. After protonation of Glu269, this salt bridge was found to break, and Arg144 to move away from Glu269 establishing a new salt bridge with Glu126. Furthermore, the displacement of Arg144, and consequently of water molecules from the interdomain region, was seen to initiate the closing of the cytoplasmic entrance (reduction of 4 Angstroms in diameter in 10 ns) by allowing hydrophobic surfaces of the N- and C-domains to fuse. Charged Glu269 was found to strongly bind the lactose permeant, indicating that proton transfer from water or another residue to Glu269 is a prerequisite for unbinding of lactose from the binding pocket.

This is one of the first MD studies of a membrane transporter that challenge membrane protein simulations due to the complexity of their function. Revealing the mechanism of co-transport of protons and sugar molecules in LacY requires long simulations of many intermediate states representing different protonation states that are involved in the transport cycle. To study molecular and energetic details of the transport process, the Resource is currently using SMD [58] to induce the lactose permeation through the protein.

BTA UNIT: C

TITLE: The Protein-Conducting Channel

KEYWORDS: translocon, SecY, translocation, protein channel

AXIS I: AXIS 1

AXIS II: AXIS 2

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

INVEST2: Tom Rapoport
DEGREE2: PhD
DEPT2: Cell Biology
NONHOST2: Harvard U.

% 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 serves to help proteins cross or integrate into membranes (depending on whether they are to be soluble or membrane proteins) [137]. It exists in all branches of life and is found in the cytoplasmic membrane in bacteria and archaea and in the membrane of the endoplasmic reticulum in eukaryotes [138–140]. A passive channel by nature, it normally partners with other proteins that drive translocation of an unfolded polypeptide across the channel. For a common mode of translocation, co-translational translocation, this partner is the ribosome which feeds the nascent protein through the channel as it is being made [141, 142]. 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) can allow it to aggregate to lethal levels in the cell [143–145]. However, being able to enhance recognition and passage could benefit artificially created proteins such as insulin [146–148]. In 2004, the first available high resolution structure was released by our collaborator, Tom Rapoport; the protein was resolved at 3.5 Angstroms and obtained from the archaeon *Methanococcus jannaschii* [149]. Based on the new structure, the details of translocation began to come into focus; certain observed 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 the monomer, and not a dimer or tetramer as hypothesized before, serves as the active channel.

The Resource first approached this problem by building a system of the protein, a lipid bilayer large enough to surround it, and water and ions to best represent the channel's native environment, comprising in total a system of greater than 106,000 atoms. Using the program NAMD [44], channel crossing events for small polypeptide helices were simulated, with the system being allowed to relax both before and after [150]. With simulations totaling more than 40 ns, certain structural hypotheses were confirmed as well as new ones made. The pore ring was seen to behave as a tight yet flexible seal before, during, and after translocation, and residues other than the ones originally proposed [149] were suggested to play a role in maintaining the seal. The plug was also seen to behave in a manner expected, leaving the channel under the pressure of the pulled polypeptide but also returning to the pore when allowed to relax. Our combined results further confirmed the idea of the monomer as the active channel. Since this result still leaves open the question of why the protein oligomerizes in nature, we also simulated a back-to-back dimer of the channel, one of two currently proposed dimer forms, which required a noticeably larger system totaling 132,000 atoms. An equilibrium simulation of this structure using NAMD combined with analysis in VMD [47] demonstrated that the plug blocking each channel gained much greater flexibility than in simulations of the monomer. This observed interaction illustrates that two monomers may influence each other in beneficial ways not related to forming a larger channel. With new simulations, we are now examining both dimer forms, accurately constructed based on homology-modelled structures received from structural biologists in the field [151, 152]. Multiple simulation techniques will be used requiring many tens of nanoseconds in the hope of differentiating between the two dimer forms and realizing one as the more likely candidate for the *in vivo* complex.

BTA UNIT: C

TITLE: Alpha Hemolysin

KEYWORDS:

AXIS I:

AXIS II:

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

% BTA $: BTA %

ABSTRACT: Alpha-hemolysin is a self-assembling toxin that, upon binding to the cell membrane, oligomerizes to form a water filled transmembrane channel [153, 154]. Apart from being one of the most widespread toxic agents of bacterial origin encountered by a human organism [153], alpha-hemolysin is the principal component of several biotechnological applications, including an alternative to the sequence gel technology for sequencing DNA [155–158] (http://www.ks.uiuc.edu/Research/hemolysin/). In order to test if a molecular dynamics simulation can accurately predict ionic currents through this membrane channel that has been well characterized by experiment, we built a system comprised of about 290,000 atoms, that included one copy of the protein, a patch of a DPPC lipid bilayer, and 1,000,000 atoms of water solution with KCl. Applying external electrostatic fields produced a transmembrane ionic current; repeating simulations at several voltage biases yielded a current-voltage curve of alpha-hemolysin and a set of electrostatic potential maps, which were found in excellent quantitative agreement with available experimental data [159]. In another simulation, we drove a DNA strand through the pore of alpha-hemolysin by applying an external electric field, and recorded the resulting blockade currents. In accordance with experiments carried out by our collaborator (A. Meller, Harvard U), we observed different kinetics of DNA translocation depending on the global orientation of the DNA strand inside the pore, revealing the molecular mechanism of the experimentally observable differences [92]. These and other results are reported in [92, 159–161].

BTA UNIT: C

TITLE: Titin Z1Z2-Telethonin Complex: A Molecular Glue in Muscle

KEYWORDS: mechanical proteins, titin, Z1Z2, telethonin, muscle, steered molecular dynamics

AXIS I: 13, 20

AXIS II: 39, 58, 60, 74h, 89

INVEST1: Eric Lee
DEGREE1: BS
DEPT1: Medicine and Biophysics
NONHOST1:

INVEST2: Mu Gao
DEGREE2: PhD
DEPT2: Beckman Institute
NONHOST2:

INVEST3: Matthias Wilmanns
DEGREE3: PhD
DEPT3: Cellular & Molecular Imaging
NONHOST3: EMBL Hamburg, Germany

% BTA $: BTA %

ABSTRACT: Titin (http://www.ks.uiuc.edu/Research/telethonin/) is a protein that protects muscle from overstretching, which can occur following a powerful muscle contraction. So important is this protective mechanism that defects in the titin gene have been correlated to the disease muscular distrophy, which is characterized by severe physical weakness. Titin functions as a molecular bungee cord, producing a restoring force when a muscle fiber is extended beyond its normal length. This prevents injury to muscle as a result of overextension.

To perform its function, one end of titin called Z1Z2 must be attached to a rigid structure from which to stretch against. Very recently the atomic structure of the titin Z1Z2 bound to a protein called telethonin became available [162].

Unlike typical ligand-binding in which the ligand is inserted into a receptor pocket, the crystal structure for the titin Z1Z2-telethonin complex revealed that the receptor Z1Z2 domains of two separate antiparallel titin N-termini are joined together by an N-terminal fragment of the telethonin ligand through beta-strand cross-links, a

structural motif that also appears in pathologies like protein fibrillation and plaque formation in Alzheimer's and Parkinson's disease [163–166]. Moreover, it is also known that beta-strand cross-links can form mechanically stable beta-sheets, for example in titin Ig domains and fibronectin type III-like modules. Naturally this raises the question: Does the beta-strand cross-linking between titin and telethonin represent a novel ligand-binding strategy that provides a mechanically stable linkage, and consequently does telethonin function as a rigid anchor for titin?

This question has been addressed by all-atom Steered Molecular Dynamics simulations performed by the Resource using NAMD [44] in order to understand the mechanical design of the Z1Z2-telethonin complex [167]. In these simulations, the complex was stretched with an applied force from multiple directions simultaneously (a specific scheme adapted for this study) in order to measure the strength of the connections between titin and telethonin. The system size involved 140,000 atoms and revealed that the Z1 and Z2 domains are bound strongly to telethonin. Analysis using VMD [47] revealed that the beta-strand cross-links between titin Z1Z2 and telethonin can be thought of as a biological glue at the molecular level. It turns out, in fact, that the Z1Z2-telethonin complex exhibits significantly higher resistance to mechanical stretching forces than titin Z1Z2 alone, suggesting that telethonin plays a role in anchoring one end of titin. Whereas previous studies have revealed that the unraveling of the beta-strands for titin Ig-domains constitutes the dominant unfolding barrier [168], in the case of Z1Z2 in complex with telethonin, the force peak observed corresponded to a detachment of one virtually intact Z2 domain from a beta-strands of telethonin.

The results from these simulations explain how the titin Z1Z2-telethonin complex produces extraordinary resistance to mechanical stress. The major force bearing component of this complex is an extensive intermolecular hydrogen bonding network formed across beta-strands between telethonin and Z1Z2 domains, and not intramolecularly between termini beta-strands of individual Z1 or Z2 domains. This shift to a stronger force bearing interface reduces the chance of unraveling the individual Ig-domains, thus stabilizing the complex. We found that telethonin functions as a strong molecular glue for augmenting the mechanical resilience of muscle proteins in order so that they can carry out their biological functions. A similar mechanism likely exists for the aggregation of pathological fibrils mentioned above.

BTA UNIT: C

TITLE: Voltage-gated Potassium (Kv) Channels

KEYWORDS: K+ channels, Shaker, voltage gating, ion channels, membrane proteins

AXIS I: 21, 13

AXIS II: 74h, 89

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

INVEST2: Emad Tajkhorshid
DEGREE2: Ph.D.
DEPT2: Beckman Institute
NONHOST2:

% BTA $: BTA %

ABSTRACT: Voltage-gated potassium ion channels (Kv channels) are integral membrane proteins present in all major kingdoms of life. In eukaryotic cells, Kv channels work with other cation channels to regulate the electrical activity of the cell. These channels are activated in response to changes in electrical potential, and allow passive conduction of K+ ions, at a high rate, while discriminating against all other cations. In addition to electrical signaling in nervous systems, Kv channels play an important role in regulation of cardiac excitability and the regulation of insulin release [169].

The X-ray crystallographic structure of Kv1.2, a member of the Shaker K+ channel family [170], has provided the first view of the molecular architecture of a Kv channel in a putative open conformation. To study the gating mechanism of Kv1.2, the Resource built a complete model of the Kv1.2 channel, modeling the missing domains of the voltage-sensor helices based on the previously solved structures of KvAP [171], a bacterial Kv channel. The molecular dynamics program, NAMD [44], was used to study the conformational changes of the channel in a lipid bilayer environment in the presence and the absence of an electric field. The studied system, including a patch of POPE lipid bilayer and 0.2 M KCl solution, was comprised of about 200K atoms and was simulated for 14 ns. Investigating the hydration of gating charges in the lipid bilayer lead to a model of gating in which the minor rearrangement of side chains could result in transfer of 4e gating charge per channel.

The ion conduction mechanism through the pore was also studied in a separate system consisting of the pore domain of the Kv1.2 channel in a lipid bilayer environement. The conduction of the K+ ions through the system was reproduced, for the first time, in 30 ns of molecular dynamics simulations with an external biasing potential. The simulations revealed the presence of a three-ion transition state, between two possible two-ion configurations, previously suggested both from X-ray crystallography [172] and MD simulations [173, 174].

BTA UNIT: C

TITLE: Molecular Mechanism of PcrA Helicase

KEYWORDS: helicase, molecular motor, DNA, MD, QM/MM, coarse graining

AXIS I: AXIS 1

AXIS II: AXIS 2

INVEST1: Markus Dittrich
DEGREE1: Ph.D.
DEPT1: Beckman Institute
NONHOST1:

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

INVEST3: Taekjip Ha
DEGREE3: Ph.D.
DEPT3: Physics; Howard Hughes Medical Institute
NONHOST3:

% BTA $: BTA %

ABSTRACT: ABSTRACT

DNA helicases 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'–5' helicase for which several atomic resolution x-ray structures have been reported [175]. The DNA unwinding action of PcrA is caused by helicase translocation along ssDNA, which is driven by the hydrolysis of ATP in a single catalytic site. The translocation proceeds at a rate of approximately 50 nucleotides per second [176].

The Resource has used three complementary approaches to investigate PcrA's overall function, the combination of which is novel and required the use of quantum chemistry, classical molecular modeling, and non-equilibrium statistical mechanics simulations. At the catalytic site 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. Analogous to the Resource's previous studies of ATP hydrolysis in F1-ATPase [177, 178], a proton relay mechanism was identified as the physiologically relevant proton transfer pathway during nucleophilic attack. The "arginine finger" residue R287 from a protein domain neighboring the catalytic site was found to be crucial for transition state stabilization, thereby, adding to the list of similarities between PcrA and F1-ATPase at the catalytic site level. Employing computational mutation studies, it could be shown that the position of Q254 with respect to the terminal phosphate group of ATP greatly influences the energy of the ATP hydrolysis product state, since a slight decrease in side chain length (mutation Q254N) changed the reaction energy profile from endothermic in the wild type to almost equi-energetic. This suggests a mechanism by which protein conformational changes induced by DNA translocation are coupled to the catalytic reaction in the binding site of PcrA.

The ATP hydrolysis powered ssDNA translocation of PcrA was then studied by molecular dynamics (MD) simulations. Using the program NAMD [44], 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 several nanoseconds. On the basis of these simulations, interaction energies between the protein and the ssDNA nucleotides were evaluated, resulting in effective potentials governing the domain movements of PcrA. The simulations revealed that during the ATP hydrolysis cycle the domains 1A and 2A alternatively become weakly and strongly bound to ssDNA in such a manner that unidirectional translocation of PcrA along ssDNA results. The alternating effective potentials were then utilized in a coarse grained Langevin dynamics description that corroborated an inchworm mechanism of PcrA translocation that had previously been proposed based on x-ray crystallography studies [175]. The combined approaches of MD simulation and Langevin dynamics description identified key amino acids affecting the unidirectional translocation of PcrA, suggesting candidate residues for mutations of PcrA that may achieve reverse translocation.

BTA UNIT: T

TITLE: MultiSeq: A VMD Multiple Alignment Plug-in

KEYWORDS: bioinformatics, multiple alignment, phylogenetic tree, evolution

AXIS I:

AXIS II:

INVEST1: Elijah Roberts
DEGREE1: B.S.
DEPT1: Biophysics
NONHOST1:

INVEST2: John Eargle
DEGREE2: B.A.
DEPT2: Biophysics
NONHOST2:

INVEST3: Dan Wright
DEGREE3: B.S.
DEPT3: School of Library and Information Sciences
NONHOST3:

INVEST4: Zaida Luthey-Schulten
DEGREE4: Ph.D.
DEPT4: Beckman Institute
NONHOST4:

% BTA $: BTA %

ABSTRACT: Multiseq (`http://www.ks.uiuc.edu/Research/vmd/plugins/multiseq/`) is a VMD plug-in that provides an environment for the analysis of bioinformatic data, both structure and sequence based. Structures can be loaded into Multiseq using the VMD capabilities, including direct loading of PDB structures over the Internet. Sequence data can be read in by Multiseq in either FASTA or CLUSTALW file formats. It can also use a locally installed version of BLAST [179] to search a local sequence database using a single sequence, a profile of sequences, or a fragment of a sequence or profile, including interative searches using PSI-BLAST. Various sorts of metadata can be automatically imported through the Internet, such as taxonomy information.

Structures are aligned with a version of STAMP [180], which has been modified for use with RNA and to better align end regions, whereas sequence alignments are performed with CLUSTALW [181]. Since structures are generally more conserved than sequences, a particularly useful feature is the possibility of passing a structural alignment to CLUSTALW as a profile to seed the sequence alignment. The sequence alignment can also be manually edited as necessary.

In order to eliminate biases present in various databases, both sequence [182] and structure [183] QR algorithms are implemented in Multiseq, which allow one to easily obtain a non-redundant data set that is suitable for evolutionary analysis. Distance based phylogenetic trees can be generated for structural similarity using QH [184] with unweighted pair group method with arithmetic averages (UPGMA) [185] and for sequence similarity using the Dayhoff PAM matrices [186] with the neighbor-joining method within CLUSTALW [181].

Multiseq is a very powerful environment to perform sequence and structural analysis in an evolutionary framework, providing valuable insights in biomedical research. All of the aforementioned features are implemented in the new version of Multiseq, of which the release is planned for the summer of 2006.

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 board met this year on May 8, 2006, and produced the following report. The Advisory Board was composed of the following members:

- 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.

Organization

Organizational Structure. The Resource web site continues to be a center point of our organization, used internally for administrative, scientific, and computing needs, and externally as a key access point for the biomedical community to review our collaborative work and developmental efforts and to take advantage of our service, training, and dissemination activities. Virtually all of the Resource's operational data (research development, management, and system administration) are stored and distributed internally through locally developed web-based databases. Similarly, the publicly accessible external website represents our extensive effort to communicate our science to the biomedical community, through journal articles and other papers, various media (images, movies, streaming video) capturing our science, updated summaries of Resource research areas, access to all Resource-produced software, a list of our services, educational, and other content reflecting the work of the group.

The Resource's web site represents our way of seeing and doing things both within and beyond the Resource's formal boundaries, and also represents what we view as the mission of the Resource. Recently, the web site was redesigned, both in structure and style, to provide easier access to Resource materials for all users, and to better communicate changes and new events. Additions to the external website include extensive documentation of our workshops, including all lectures slides, links to tutorials, participant descriptions, and evaluation results. Training pages have also been developed that provide links to constantly updated, cross-platform tutorials used in the workshops.

K. Schulten (Professor, Physics, Beckman, Biophysics, Chemistry) is the Principal Investigator and Program Director of the Resource. E. Tajkhorshid (Research Assistant Professor, Biophysics) the Assistant Director for Research of the Resource, assists the Director in all organizational and scientific activities of the Resource. L. Kale (Professor, Computer Science), Z. Luthey-Schulten (Professor, Chemistry), and Aleksei Aksimentiev (Assistant Professor, Physics) are other Co-Principal Investigators of the Resource. D. Brandon is the Manager of the Resource who coordinates all organizational and dissemination activities. Tim Skirvin is the Systems Administrator of the group who also oversees Service activities of the Resource. The Resource is located at the Beckman Institute for Advanced Science and Technology and K. Schulten, the Resource Director, administratively reports to the Institute Director. The Institute Director reports to the University of Illinois Vice Chancellor for Research.

The Resource members come from a spectrum of disciplines, each of which contributes significantly to the intricate fabric of the Resource's goals and activities. Staff and graduate students are affiliated with fields and departments such as Physics, Computer Science, Biophysics, Chemistry, Mathematics, and Electrical and Computer Engineering.

All Resource members participate in the daily operation of the facility. Members attend weekly group and subgroup meetings, are responsible for specific maintenance tasks at the Resource, attend and present talks in group seminars, and keep continuously informed by spending time at the Beckman Institute as well as through email, various BioCoRE project groups, and the Resource's internal web site which lists meetings, seminars, group jobs, and more.

The PIs and affiliated faculty, in consultation with the other Resource members, determine collaborative and service projects. Selection of technological research and development projects at the Resource is determined by the following criteria:

- Relevance of research to the biological and medical sciences
- Quality and originality of research and conceptual approach
- Computational demands of the research project
- Novelty of algorithmic strategies required for the projects

Continuous interactions with the collaborators and ongoing critical evaluation of the projects ensure relevance, progress and adherence to the criteria outlined above. Local and remote computer time is allocated to projects as needed.

The web-based Resource manual, as well as other useful documents available on our internal site serve as guidelines for new members and as reference resources for old members.

The continually evolving internal site reflects short-and long-term objectives and describes the Resource's structure and daily procedures; it specifies policies and guidelines; it contains a job list detailing the maintenance tasks assigned to Resource members; it offers detailed information on reports, proposals, and special events. The internal site has a vital role in streamlining and systematizing the Resource operation via tips and information on the Resource's internal processes, and on Beckman and UIUC facilities and procedures.

How to Acknowledge Resource Support. A prominent link on the front page of the Resource's external site, as well as links at each application website, leads users and beneficiaries to guidelines on how to acknowledge Resource support in several ways, depending on resources used.

Service, Training and Dissemination

Introduction to Service, Training and Dissemination

The research and development outcomes of the Resource are transferred to the biomedical community through a variety of boundary-spanning service, training, and dissemination activities. The core activities of the Resource can be classified into two general, sometimes overlapping, functional areas:

I. Technological development to create research tools and methods

II. Research and collaborative projects that use and benefit from the tools

Each of these activity areas has vast potential and practical implications for the Resource and the biomedical community at large. Assuring that the results of Resource activities are returned to the biomedical community are the goal of the service, training, an dissemination efforts, for example software production and distribution as found in service, numerous workshops as found in training, and lectures as found in dissemination.

Forces such as the huge genomic data revolution and the increasing pace of structure discovery, the explosive progress in hardware development, nanotechnology, and web technology, along with other factors have infused renewed energy and urgency to our activities and are reshaping our scope and practices daily. The growth of the Resource continues, with 50 members (graduate assistants, postdoctoral associates, developers, faculty, administrative and technical staff); the number and size of systems modeled here are unmatched; and, our computational resources are greater than ever before and are effectively utilized.

The Resource's visibility continues to expand, primarily via increased visits to the Resource web site, and with that, the service, training and dissemination opportunities increase. A primary channel to the biomedical community for outreach efforts in all areas, the use of web technologies allows scientists to access all content and materials generated by the Resource. Using the web site as focus for all activities allows the Resource to make its technological developments and collaborative efforts accessible across organizational and other boundaries, though such use can present other challenges, for example maintaining intellectual property and copyright for scientific images posted on the site. Internally, the web site is a fundamental organizing tool for Resource members, be they administrative, staff, or academic.

Efforts over the past year have increased web site options for service, training, and dissemination. At the newly-designed web site, visitors can now access a script library, see a list of recent publications, review research in updated research categories, and see other current information. Training resources have also increased, via publication of case studies, revised tutorials, and content from five recent workshops.

Service

A variety of valuable services are offered to the biomedical community by the Resource, as outlined below. In particular, the Resource is known for its effectiveness in supporting collaborations and for producing and supporting high-quality software, as described both below and in other sections of this report.

A total of 94 researchers have used the Resource's computational facilities (45 local, 50 remote). Since June 2005, the Resource has experienced an increase of 18% in shared file storage space compared to the same period a year ago (from 17.0 to 20.0 TB). Local compute power and visualization capabilities remains strong, and external supercomputer time has again been allocated, raising the Resource's scaled compute power up 45%.

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 2005 to early April 2006 are listed below:

- UIUC, Materials Science and Engineering Department (cluster building)
- Graduate School of Frontier Biosciences Osaka University, Japan (3D facility)
- Yale University (cluster building)
- Northwestern University (evaluation of FileMaker Pro)
- Illinois State Geological Survey (use of SpamAssassin)

Also, Resource members beta tested Apple Computer's MacOS X on the Intel x86 platform prior to its public release. This hardware and software evaluation gave Apple the chance for feedback prior to releasing the new machines to the public in February 2006, and Resource developers the opportunity to have a version of VMD suited to the new machines by the release date.

Resource Software

The Resource is engaged in intensive development efforts and technology transfer. We distribute a number of software packages, particularly VMD, NAMD and BioCoRE, as well as a number of smaller programs. All Resource-developed programs, binaries and source, are freely available on our web site for easy accessibility, employing where needed a unified distribution mechanism*. The VMD, NAMD and BioCoRE packages are developed, maintained, and distributed by Resource staff. The staff also offers extensive

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

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, in the past year.

Software Licensing
Ongoing discussions with the UIUC Office of Technology Managment, industry, and others allow the Resource to develop licenses that allow broad distribution of our software. Licensing activity over the funding period includes the release of two new MDTools[†] under the UIUC Open Source license, and an expansion of the NAMD license at the request of the Argonne National Laboratory.

Use of VMD, NAMD, and BioCoRE.

VMD has 69,347 registered users (an increase of 13,925 or +25% in the least year), with 16,526 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 1.8.3, released in February 2005, has over 17,200 users, with more than 10,300 new users registering for and downloading this version since April 2005. And, over 11,400 users have downloaded development versions of VMD 1.8.4 since April 2005.

NAMD has 15,571 registered users as of April 2006 (an increase of 3,476 or +29% in the last year), of whom 3,064 are repeat users. 2,546 (16%) of NAMD users are NIH funded. NAMD has been downloaded 4,394 times since mid-April 2005; the latest version, NAMD 2.6b1 released in July 2005, has 3,527 users, of whom 635 are NIH funded.

BioCoRE has 1,334 registered users (an increase of 275, or an increase of +26% in the past year), involved in 390 projects (compared to 322 a year ago, an increase of +21%). Fully 145 (37%) of the projects within BioCoRE have been reported as either completely or partially NIH-funded.

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

- VMD: 1.8.4 released April 2006; next release summer 2006
- NAMD: 2.6 expected June 2006
- BioCoRE: Incremental updates every few weeks

Website Popularity
The appeal and usability of the Resource web site continues to bring in growing numbers

[†]URL:`http://www.ks.uiuc.edu/Development/MDTools/`

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 2005 - March 2006) the software sections on our web site had the following volume of visits:

	Total	Month Avg.
VMD	197,855	16,847
NAMD	84,491	7,040
BioCoRE	24,458	2,038

Table 1: Application web site visits

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 2005-2006 include:

- Improvements to molecule data structures allow VMD to correctly load, select, and display Protein Data Bank structures with multiple conformations, special bonds, biological unit reconstruction transformations, and recovery of atomic element and novel residue information when available
- New volumetric map generation commands and interfaces make it possible to synthesize volumetric density and occupancy maps from atomic coordinates or molecular dynamics trajectories which can be used for both analysis and visualization purposes
- New "shaded points" isosurfaces, isosurface step size controls, and "beads" graphical representations allow researchers to visualize very large molecular models and volumetric datasets on computers with modest graphics capabilities
- New plugins have also been developed, which add support for several new molecular and volumetric file formats

Posters/Presentations/Tutorials/Talks (local and remote, by Resource members and others who informed us):

- March 20-23, 2006, Frankfurt, Germany, *Workshop in Computational Biology.* Tutorial: "VMD Molecular Graphics Tutorial" (Resource staff)
- November 28 - December 1, 2005, Pittsburgh, Pennsylvania, *Workshop in Computational Biology.* Tutorial: "VMD Molecular Graphics Tutorial" (Resource staff)
- September 12, 2005, San Francisco, CA, UCSD, San Diego Supercomputer Center, *Visualization of Large Biomolecular Complexes Workshop.* Lecture: "VMD: Algorithms and Methods for Large Scale Biomolecular Visualization" (John Stone)
- June 26-30, 2005, San Francisco, California, *Workshop in Computational Biology.* Tutorial: "VMD Molecular Graphics Tutorial" (Resource staff)
- June 9-13, 2005, Chicago, Illinois, *Workshop in Computational Biology.* Tutorial: "VMD Molecular Graphics Tutorial" (Resource staff)
- May 23-27, 2005, Lake Tahoe, California, *Workshop in Computational Biology.* Tutorial: "VMD Molecular Graphics Tutorial" (Resource staff)

  Scope of VMD User Support:

  - 6,752 e-mail exchanges in response to user inquiries sent to the vmd@ks.uiuc.edu e-mail address
  - 624 subscribers to the VMD-L mailing list, with 6,894 total postings, and 1,961 postings from the end of April 2005 through March 2006
  - Local face-to-face support has been provided

  There are currently 295 non-Resource users with access to the VMD source code repository, with 61 such users added in the last year.

  List of papers citing VMD: A literature search in the ISI Web of Science citation database in April 2006 yielded 1,345 published journal articles, papers, or books citing the VMD origin paper [47]. Below are 50 recent citations:

  1. Yu, Z. Y., M. P. Jacobson and R. A. Friesner (2006). "What role do surfaces play in GB models? A new-generation of surface-generalized Born model based on a novel Gaussian surface for biomolecules." Journal of Computational Chemistry 27(1): 72-89.
  2. Warren, C. L., N. C. S. Kratochvil, K. E. Hauschild, S. Foister, M. L. Brezinski, P. B. Dervan, G. N. Phillips and A. Z. Ansari (2006). "Defining the sequence-recognition profile of DNA-binding molecules." Proceedings of the National Academy of Sciences of the United States of America 103(4): 867-872.

3. Wang, Y. L., K. Arora and T. Schlick (2006). "Subtle but variable conformational rearrangements in the replication cycle of Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4) may accommodate lesion bypass." Protein Science 15(1): 135-151.

4. Wang, T. and R. C. Wade (2006). "Force field effects on a beta-sheet protein domain structure in thermal unfolding simulations." Journal of Chemical Theory and Computation 2(1): 140-148.

5. Vijayakrishnan, S., R. Qamra, C. S. Verma, R. Sen and S. C. Mande (2006). "Cation-mediated interplay of loops in Chaperonin-10." Journal of Biomolecular Structure and Dynamics 23(4): 365-375.

6. Vicens, Q. and T. R. Cech (2006). "Atomic level architecture of group I introns revealed." Trends in Biochemical Sciences 31(1): 41-51.

7. VandeVondele, J., R. Lynden-Bell, E. J. Meijer and M. Sprik (2006). "Density functional theory study of tetrathiafulvalene and thianthrene in acetonitrile: Structure, dynamics, and redox properties." Journal of Physical Chemistry B 110(8): 3614-3623.

8. van der Straaten, T. A., G. Kathawala and U. Ravaioli (2006). "Device engineering approaches to the simulation of charge transport in biological ion channels." Journal of Computational and Theoretical Nanoscience 3(1): 42-62.

9. Tornroth-Horsefield, S., Y. Wang, K. Hedfalk, U. Johanson, M. Karlsson, E. Tajkhorshid, R. Neutze and P. Kjellbom (2006). "Structural mechanism of plant aquaporin gating." Nature 439(7077): 688-694.

10. Tanizaki, S. and M. Feig (2006). "Molecular dynamics simulations of large integral membrane proteins with an implicit membrane model." Journal of Physical Chemistry B 110(1): 548-556.

11. Tan, M. L., C. H. Kang and T. Ichiye (2006). "The role of backbone stability near Ala(44) in the high reduction potential class of rubredoxins." Proteins-Structure Function and Bioinformatics 62(3): 708-714.

12. Sun, Y., S. Bojikova-Fournier and T. H. MacRae (2006). "Structural and functional roles for beta-strand 7 in the alpha-crystallin domain of p26, a polydisperse small heat shock protein from Artemia franciscana." Febs Journal 273(5): 1020-1034.

13. Stoica, I. (2006). "Characterization of protein matrix motions in the Rb. sphaeroides photosynthetic reaction center." Journal of Molecular Modeling 12(4): 468-480.

14. Spackova, N. and J. Sponer (2006). "Molecular dynamics simulations of sarcin-ricin rRNA motif." Nucleic Acids Research 34(2): 697-708.

15. Spaar, A., C. Dammer, R. R. Gabdoulline, R. C. Wade and V. Helms (2006). "Diffusional encounter of barnase and barstar." Biophysical Journal 90(6): 1913-1924.

16. Sieffert, N. and G. Wipff (2006). "Alkali cation extraction by calix 4 crown-6 to room-temperature ionic liquids. The effect of solvent anion and humidity investigated by molecular dynamics simulations." Journal of Physical Chemistry A 110(3): 1106-1117.

17. Sieffert, N. and G. Wipff (2006). "Adsorption at the liquid-liquid interface in the biphasic rhodium catalyzed hydroformylation of olefins promoted by cyclodextrins: A molecular dynamics study." Journal of Physical Chemistry B 110(9): 4125-4134.

18. Shih, A. Y., A. Arkhipov, P. L. Freddolino and K. Schulten (2006). "Coarse grained protein-lipid model with application to lipoprotein particles." Journal of Physical Chemistry B 110(8): 3674-3684.

19. Sheng, Y. B. and W. Wang (2006). "Comparative all-atomic study of unfolding pathways for proteins chymotrypsin inhibitor 2 and barnase." Physical Review E 73(2).

20. Sapay, N., R. Montserret, C. Chipot, V. Brass, D. Moradpour, G. Deleage and F. Penin (2006). "NMR structure and molecular dynamics of the in-plane membrane anchor of nonstructural. protein 5A from bovine viral diarrhea virus." Biochemistry 45(7): 2221-2233.

21. Sands, Z. A., A. Grottesi and M. S. P. Sansom (2006). "The intrinsic flexibility of the Kv voltage sensor and its implications for channel gating." Biophysical Journal 90(5): 1598-1606.

22. Sahoo, N., W. Beatty, J. Heuser, D. Sept and L. D. Sibley (2006). "Unusual kinetic and structural properties control rapid assembly and turnover of actin in the parasite Toxoplasma gondii." Molecular Biology of the Cell 17(2): 895-906.

23. Sacquin-Mora, S. and R. Lavery (2006). "Investigating the local flexibility of functional residues in hemoproteins." Biophysical Journal 90(8): 2706-2717.

24. Rockwell, N. C. and J. C. Lagarias (2006). "The structure of phytochrome: A picture is worth a thousand spectra." Plant Cell 18(1): 4-14.

25. Riposati, A., T. Prieto, C. S. Shida, I. L. Nantes and O. R. Nascimento (2006). "Low spin states of microperoxidases produced by inter- and intra-peptide

chain sixth ligands: Effect of pH and the oligopeptide type." Journal of Inorganic Biochemistry 100(2): 226-238.

26. Richardson, C. M., D. S. Williamson, M. J. Parratt, J. Borgognoni, A. D. Cansfield, P. Dokurno, G. L. Francis, R. Howes, J. D. Moore, J. B. Murray, A. Robertson, A. E. Surgenor and C. J. Torrance (2006). "Triazolo 1,5-a pyrimidines as novel CDK2 inhibitors: Protein structure-guided design and SAR." Bioorganic and Medicinal Chemistry Letters 16(5): 1353-1357.
27. Ramirez, E., A. Santana, A. Cruz, I. Plasencia and G. E. Lopez (2006). "Molecular dynamics of surfactant protein C: From single molecule to heptameric aggregates." Biophysical Journal 90(8): 2698-2705.
28. Priyakumar, U. D. and A. D. MacKerell (2006). "Computational approaches for investigating base flipping in oligonucleotides." Chemical Reviews 106(2): 489-505.
29. Polverini, E., M. Fornabaio, A. Fasano, G. Carlone, P. Riccio and P. Cavatorta (2006). "The pH-dependent unfolding mechanism of P2 myelin protein: An experimental and computational study." Journal of Structural Biology 153(3): 253-263.
30. Petrone, P. and V. S. Pande (2006). "Can conformational change be described by only a few normal modes?" Biophysical Journal 90(5): 1583-1593.
31. Patargias, G., N. Zitzmann, R. Dwek and W. B. Fischer (2006). "Protein-protein interactions: Modeling the hepatitis C virus ion channel p7." Journal of Medicinal Chemistry 49(2): 648-655.
32. Park, S. and J. G. Saven (2006). "Simulation of pH-dependent edge strand rearrangement in human beta-2 microglobulin." Protein Science 15(1): 200-207.
33. Park, J., B. Kahng, R. D. Kamm and W. Hwang (2006). "Atomistic simulation approach to a continuum description of self-assembled beta-sheet filaments." Biophysical Journal 90(7): 2510-2524.
34. Paramore, S., G. S. Ayton and G. A. Voth (2006). "Extending a spectrin repeat unit. II: Rupture behavior." Biophysical Journal 90(1): 101-111.
35. Papaleo, E., P. Fantucci, M. Vai and L. De Gioia (2006). "Three-dimensional structure of the catalytic domain of the yeast beta-(1,3)-glucan transferase Gas1: a molecular modeling investigation." Journal of Molecular Modeling 12(2): 237-248.
36. Pan, X. L., N. H. Tan, G. Z. Zeng, H. J. Han and H. Q. Huang (2006). "3D-QSAR and docking studies of aldehyde inhibitors of human cathepsin K." Bioorganic and Medicinal Chemistry 14(8): 2771-2778.

37. Olkhova, E., C. Hunte, E. Screpanti, E. Padan and H. Michel (2006). "Multiconformation continuum electrostatics analysis of the NhaA Na+/H+ antiporter of Escherichia coli with functional implications." Proceedings of the National Academy of Sciences of the United States of America 103(8): 2629-2634.

38. Oliver, S. L., C. A. Batten, Y. Deng, M. Elschner, P. Otto, A. Charpilienne, I. N. Clarke, J. C. Bridger and P. R. Lambden (2006). "Genotype 1 and genotype 2 bovine noroviruses are antigenically distinct but share a cross-reactive epitope with human noroviruses." Journal of Clinical Microbiology 44(3): 992-998.

39. Noy, D. and P. L. Dutton (2006). "Design of a minimal polypeptide unit for bacteriochlorophyll binding and self-assembly based on photosynthetic bacterial light-harvesting proteins." Biochemistry 45(7): 2103-2113.

40. Noel, N., J. Flanagan, S. G. Kalko, M. J. R. Bajo, M. D. Manu, J. L. G. Fuster, E. Beutler and J. L. V. Corrons (2006). "Two new phosphoglycerate kinase mutations associated with chronic haemolytic anaemia and neurological dysfunction in two patients from Spain." British Journal of Haematology 132(4): 523-529.

41. Morrone, J. A., K. E. Hasllinger and M. E. Tuckerman (2006). "Ab initio molecular dynamics simulation of the structure and proton transport dynamics of methanol-water solutions." Journal of Physical Chemistry B 110(8): 3712-3720.

42. Moore, M. J. B., C. M. Schultes, J. Cuesta, F. Cuenca, M. Gunaratnam, F. A. Tanious, W. D. Wilson and S. Neidle (2006). "Trisubstituted acridines as G-quadruplex telomere targeting agents. Effects of extensions of the 3,6-and 9-side chains on quadruplex binding, telomerase activity, and cell proliferation." Journal of Medicinal Chemistry 49(2): 582-599.

43. Moll, A., A. Hildebrandt, H. P. Lenhof and O. Kohlbacher (2006). "BALLView: a tool for research and education in molecular modeling." Bioinformatics 22(3): 365-366.

44. Milne, J. L. S., X. W. Wu, M. J. Borgnia, J. S. Lengyel, B. R. Brooks, D. Shi, R. N. Perham and S. Subramaniam (2006). "Molecular structure of a 9-MDa icosahedral pyruvate dehydrogenase subcomplex containing the E2 and E3 enzymes using cryoelectron microscopy." Journal of Biological Chemistry 281(7): 4364-4370.

45. Mattice, W. L. and N. Waheed (2006). "An assessment of the role of quenched randomness in the stereochemical sequences of atactic vinyl polymers." Macromolecules 39(6): 2380-2387.

46. Matthews, J. F., C. E. Skopec, P. E. Mason, P. Zuccato, R. W. Torget, J. Sugiyama, M. E. Himmel and J. W. Brady (2006). "Computer simulation studies of microcrystalline cellulose I beta." Carbohydrate Research 341(1): 138-152.

47. Martinez, T. J. (2006). "Insights for light-driven molecular devices from ab initio multiple spawning excited-state dynamics of organic and biological chromophores." Accounts of Chemical Research 39(2): 119-126.

48. Marsella, L. (2006). "Modeling truncated hemoglobin vibrational dynamics." Proteins-Structure Function and Bioinformatics 62(1): 173-182.

49. Mantz, Y. A., B. Chen and G. J. Martyna (2006). "Structural correlations and motifs in liquid water at selected temperatures: Ab initio and empirical model predictions." Journal of Physical Chemistry B 110(8): 3540-3554.

50. Logadottir, A., P. G. Moses, B. Hinnemann, N. Y. Topsoe, K. G. Knudsen, H. Topsoe and J. K. Norskov (2006). "A density functional study of inhibition of the HDS hydrogenation pathway by pyridine, benzene, and H2S on MoS2-based catalysts." Catalysis Today 111(1-2): 44-51.

Sites with Links to the VMD Site (Google, April 2006): 208 domains; 711 pages.

Development, Distribution, and Use of NAMD

During the reported period, NAMD enjoyed significant improvements and a drastic increase in user numbers. 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 2005-2006 include:

- Port to Mac OS X for Intel processors.
- Port to Windows for x86-64 (Opteron/Athlon 64/EMT64) processors.
- Port to 2000-processor PSC Cray XT3, including significant debugging and tuning necessary to achieve acceptable performance on new platform.
- Port of Tcl interpreter to Cray XT3 and IBM BlueGene/L platforms, providing full NAMD functionality on these unique platforms.
- Improved serial performance on POWER, PowerPC, and BlueGene/L.
- Topology-away load balance and work distribution on BlueGene/L.
- Various improvements to load balancer and work distribution.
- Adaptive Biasing Force free energy calculations.

- Support for the OPLS force field.
- Molecular structure generator handles CHARMM stream files and cross-terms, preserves element and chain identifiers, and is more tolerant of atom naming differences between input topology and PDB files.

Posters/Presentations/Tutorials/Talks (local and remote, by Resource members and others who informed us):

- March 20-23, 2006, Frankfurt, Germany, *Workshop in Computational Biology.* Tutorial: "NAMD Tutorial" (Resource staff)
- November 28 - December 1, 2005, Pittsburgh, Pennsylvania, *Workshop in Computational Biology.* Tutorial: "NAMD Tutorial" (Resource staff)
- June 26-30, 2005, San Francisco, California, *Workshop in Computational Biology.* Tutorial: "NAMD Tutorial" (Resource staff)
- June 9-13, 2005, Chicago, Illinois, *Workshop in Computational Biology.* Tutorial: "NAMD Tutorial" (Resource staff)
- May 23-27, 2005, Lake Tahoe, California, *Workshop in Computational Biology.* Tutorial: "NAMD Tutorial" (Resource staff)

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 40 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
- 435 subscribers to the NAMD-L mailing list, with 3,744 total postings, and 1,575 postings from the end of April 2005 - March 2006
- Over 600 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 224 users with access to the NAMD source code repository, with 52 users added in the last year.

List of papers citing NAMD: A literature search in the ISI Web of Science citation database in April 2006 yielded 352 published journal articles, papers, or books citing the NAMD origin paper [111] or the more recent NAMD paper [44]. Below are 50 recent cites:

1. Zhuang, W., D. Abramavicius, T. Hayashi and S. Mukamel (2006). "Simulation protocols for coherent femtosecond vibrational spectra of peptides." Journal of Physical Chemistry B 110(7): 3362-3374.
2. Zhu, Y. M., D. Smith, C. Verma, W. G. Lim, B. J. Tan, J. S. Armstrong, S. F. Zhou, E. Chan, S. L. Tan, Y. Z. Zhu, N. S. Cheung and W. Duan (2006). "The very C-terminus of protein kinase C epsilon is critical for the full catalytic competence but its hydrophobic motif is dispensable for the interaction with 3-phosphoinositide-dependent kinase-1." Cellular Signalling 18(6): 807-818.
3. Yan, L. M., Y. F. Ma and J. M. Seminario (2006). "Encoding information using molecular vibronics." Journal of Nanoscience and Nanotechnology 6(3): 675-684.
4. Wang, Y. L., K. Arora and T. Schlick (2006). "Subtle but variable conformational rearrangements in the replication cycle of Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4) may accommodate lesion bypass." Protein Science 15(1): 135-151.
5. Tornroth-Horsefield, S., Y. Wang, K. Hedfalk, U. Johanson, M. Karlsson, E. Tajkhorshid, R. Neutze and P. Kjellbom (2006). "Structural mechanism of plant aquaporin gating." Nature 439(7077): 688-694.
6. Shih, A. Y., A. Arkhipov, P. L. Freddolino and K. Schulten (2006). "Coarse grained protein-lipid model with application to lipoprotein particles." Journal of Physical Chemistry B 110(8): 3674-3684.
7. Seroka, P., M. Plosinski, J. Czub, P. Sowinski and J. Pawlak (2006). "Monosaccharides as internal probes for the determination of the absolute configuration of 2-butanol." Magnetic Resonance in Chemistry 44(2): 132-138.
8. Sapay, N., R. Montserret, C. Chipot, V. Brass, D. Moradpour, G. Deleage and F. Penin (2006). "NMR structure and molecular dynamics of the in-plane membrane anchor of nonstructural. protein 5A from bovine viral diarrhea virus." Biochemistry 45(7): 2221-2233.
9. Ramirez, E., A. Santana, A. Cruz, I. Plasencia and G. E. Lopez (2006). "Molecular dynamics of surfactant protein C: From single molecule to heptameric aggregates." Biophysical Journal 90(8): 2698-2705.
10. Pophristic, V., S. Vemparala, I. Ivanov, Z. W. Liu, M. L. Klein and W. F. DeGrado (2006). "Controlling the shape and flexibility of arylamides: A com-

bined ab initio, ab initio molecular dynamics, and classical molecular dynamics study." Journal of Physical Chemistry B 110(8): 3517-3526.

11. Park, S. and J. G. Saven (2006). "Simulation of pH-dependent edge strand rearrangement in human beta-2 microglobulin." Protein Science 15(1): 200-207.
12. Meyer, T., C. Ferrer-Costa, A. Perez, M. Rueda, A. Bidon-Chanal, F. J. Luque, C. A. Laughton and M. Orozco (2006). "Essential dynamics: A tool for efficient trajectory compression and management." Journal of Chemical Theory and Computation 2(2): 251-258.
13. Martinez, L., P. Webb, I. Polikarpov and M. S. Skaf (2006). "Molecular dynamics simulations of ligand dissociation from thyroid hormone receptors: Evidence of the likeliest escape pathway and its implications for the design of novel ligands." Journal of Medicinal Chemistry 49(1): 23-26.
14. Li, A. Q. and E. H. Dowell (2006). "Modal reduction of mathematical models of biological molecules." Journal of Computational Physics 211(1): 262-288.
15. Legge, F. S., A. Budi, H. Treutlein and I. Yarovsky (2006). "Protein flexibility: Multiple molecular dynamics simulations of insulin chain B." Biophysical Chemistry 119(2): 146-157.
16. Lee, E. H., M. Gao, N. Pinotsis, M. Wilmanns and K. Schulten (2006). "Mechanical strength of the titin Z1Z2-telethonin complex." Structure 14(3): 497-509.
17. Kosztin, I., B. Barz and L. Janosi (2006). "Calculating potentials of mean force and diffusion coefficients from nonequilibrium processes without Jarzynski's equality." Journal of Chemical Physics 124(6).
18. Kawatsu, T., D. N. Beratan and T. Kakitani (2006). "Conformationally averaged score functions for electronic propagation in proteins." Journal of Physical Chemistry B 110(11): 5747-5757.
19. Kale, L. V., G. B. Zheng, C. W. Lee and S. Kumar (2006). "Scaling applications to massively parallel machines using Projections performance analysis tool." Future Generation Computer Systems-the International Journal of Grid Computing Theory Methods and Applications 22(3): 347-358.
20. Jensen, M. O. and O. G. Mouritsen (2006). "Single-channel water permeabilities of Escherichia coli aquaporins AqpZ and GlpF." Biophysical Journal 90(7): 2270-2284.
21. Ivanov, I., S. Vemparala, V. Pophristic, K. Kuroda, W. F. DeGrado, J. A. McCammon and M. L. Klein (2006). "Characterization of nonbiological antimicrobial polymers in aqueous solution and at water-lipid interfaces from

all-atom molecular dynamics." Journal of the American Chemical Society 128(6): 1778-1779.

22. Henin, J., B. Maigret, M. Tarek, C. Escrieut, D. Fourmy and C. Chipot (2006). "Probing a model of a GPCR/ligand complex in an explicit membrane environment: The human cholecystokinin-1 receptor." Biophysical Journal 90(4): 1232-1240.

23. Heng, J. B., A. Aksimentiev, C. Ho, P. Marks, Y. V. Grinkova, S. Sligar, K. Schulten and G. Timp (2006). "The electromechanics of DNA in a synthetic nanopore." Biophysical Journal 90(3): 1098-1106.

24. Hajjar, E., B. Korkmaz, F. Gauthier, B. O. Brandsdal, V. Witko-Sarsat and N. Reuter (2006). "Inspection of the binding sites of proteinase3 for the design of a highly specific substrate." Journal of Medicinal Chemistry 49(4): 1248-1260.

25. Gumbart, J. and K. Schulten (2006). "Molecular dynamics studies of the archaeal translocon." Biophysical Journal 90(7): 2356-2367.

26. Gullingsrud, J., C. Kim, S. S. Taylor and J. A. McCammon (2006). "Dynamic binding of PKA regulatory subunit RI alpha." Structure 14(1): 141-149.

27. Gracheva, M. E., A. L. Xiong, A. Aksimentiev, K. Schulten, G. Timp and J. P. Leburton (2006). "Simulation of the electric response of DNA translocation through a semiconductor nanopore-capacitor." Nanotechnology 17(3): 622-633.

28. Freddolino, P. L., A. S. Arkhipov, S. B. Larson, A. McPherson and K. Schulten (2006). "Molecular dynamics simulations of the complete satellite tobacco mosaic virus." Structure 14(3): 437-449.

29. Fitch, C. A., S. T. Whitten, V. J. Hilser and B. Garcia-Moreno (2006). "Molecular mechanisms of pH-driven conformational transitions of proteins: Insights from continuum electrostatics calculations of acid unfolding." Proteins-Structure Function and Bioinformatics 63(1): 113-126.

30. Espinoza-Fonseca, L. M. and J. G. Trujillo-Ferrara (2006). "The existence of a second allosteric site on the M-1 muscarinic acetylcholine receptor and its implications for drug design." Bioorganic and Medicinal Chemistry Letters 16(5): 1217-1220.

31. Corry, B. (2006). "An energy-efficient gating mechanism in the acetylcholine receptor channel suggested by molecular and Brownian dynamics." Biophysical Journal 90(3): 799-810.

32. Correia, B. E., N. Loureiro-Ferreira, J. R. Rodrigues and R. M. M. Brito (2006). "A structural model of an amyloid protofilament of transthyretin." Protein Science 15(1): 28-32.

33. Casanovas, J., D. Zanuy and C. Aleman (2006). "Conducting polymer actuator mechanism based on the conformational flexibility of calix 4 arene." Angewandte Chemie-International Edition 45(7): 1103-1105.

34. Cadena, C., Q. Zhao, R. Q. Snurr and E. J. Maginn (2006). "Molecular modeling and experimental studies of the thermodynamic and transport properties of pyridinium-based ionic liquids." Journal of Physical Chemistry B 110(6): 2821-2832.

35. Bronson, J., O. S. Lee and J. G. Saven (2006). "Molecular dynamics simulation of WSK-3, a computationally designed, water-soluble variant of the integral membrane protein KcsA." Biophysical Journal 90(4): 1156-1163.

36. Benitez, B. A. S., K. Arora and T. Schlick (2006). "In silico studies of the African swine fever virus DNA polymerase X support an induced-fit mechanism." Biophysical Journal 90(1): 42-56.

37. Bastug, T., A. Gray-Weale, S. M. Patra and S. Kuyucak (2006). "Role of protein flexibility in ion permeation: A case study in gramicidin A." Biophysical Journal 90(7): 2285-2296.

38. Baer, M., E. Schreiner, A. Kohlmeyer, R. Rousseau and D. Marx (2006). "Inverse temperature transition of a biomimetic elastin model: Reactive flux analysis of folding/unfolding and its coupling to solvent dielectric relaxation." Journal of Physical Chemistry B 110(8): 3576-3587.

39. Anishkin, A., S. Sukharev and M. Colombini (2006). "Searching for the molecular arrangement of transmembrane ceramide channels." Biophysical Journal 90(7): 2414-2426.

40. Zhang, Q. M., J. Jaroniec, G. Lee and P. E. Marszalek (2005). "Direct detection of inter-residue hydrogen bonds in polysaccharides by single-molecule force spectroscopy." Angewandte Chemie-International Edition 44(18): 2723-2727.

41. Zhang, Q. M., G. R. Lee and P. E. Marszalek (2005). "Atomic cranks and levers control sugar ring conformations." Journal of Physics-Condensed Matter 17(18): S1427-S1442.

42. Zhang, D. Q. and J. A. McCammon (2005). "The association of tetrameric acetylcholinesterase with CoIQ tail: A block normal mode analysis." Plos Computational Biology 1(6): 484-491.

43. Yonkunas, M. J., Y. Xu and P. Tang (2005). "Anesthetic interaction with ketosteroid isomerase: Insights from molecular dynamics simulations." Biophysical Journal 89(4): 2350-2356.

44. Woods, C. J., M. H. Ng, S. Johnston, S. E. Murdock, B. Wu, K. Tai, H. Fangohr, P. Jeffreys, S. Cox, J. G. Frey, M. S. P. Sansom and J. W. Essex (2005). "Grid computing and biomolecular simulation." Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences 363(1833): 2017-2035.

45. Witheesuk, E. J. and R. Samudrala (2005). "Prediction of HIV-1 protease inhibitor resistance using a protein-inhibitor flexible docking approach." Antiviral Therapy 10(1): 157-166.

46. Williams, E. T., C. M. Turman and H. W. Strobel (2005). "NAMD-LP: Filtering and translating NAMD-generated files." Molecular Simulation 31(1): 67-68.

47. Wiley, A. P., M. T. Swain, S. C. Phillips, J. W. Essex and C. M. Edge (2005). "Parametrization of reversible digitally filtered molecular dynamics simulations." Journal of Chemical Theory and Computation 1(1): 24-35.

48. Wells, S., S. Menor, B. Hespenheide and M. F. Thorpe (2005). "Constrained geometric simulation of diffusive motion in proteins." Physical Biology 2(4): S127-S136.

49. Weinbach, Y. and R. Elber (2005). "Revisiting and parallelizing SHAKE." Journal of Computational Physics 209(1): 193-206.

50. Wang, W. and R. D. Skeel (2005). "Fast evaluation of polarizable forces." Journal of Chemical Physics 123(16).

Sites with Links to NAMD site (Google, 2005): 107 domains; 384 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 great investment into the US computational grid eminently useful for biomedical research.

BioCoRE 2005-06 updates include:

- Programming interface for supercomputer job submissions
- VMD Plugin for easy interaction with BioCoRE
- Guest User accounts provide access to BioCoRE with less required information.
- Filesystem access for public projects is now truly public. Visitors no longer need to even have BioCoRE accounts to see the public information.

BioCoRE Posters/Presentations/Tutorials/Talks (local and remote, by Resource members and others who informed us):

- November 3, 2005, Urbana, IL, National Center for Supercomputing Applications, Social Networks and Cyberinfrastructure, *BioCoRE Overview* (Kirby Vandivort)

Scope of BioCoRE User Support:

- 172 emails issued to/from `biocore@ks.uiuc.edu` from April 2005 - March 2006
- 1,106 chat messages sent to the BioCoRE public help project from April 2005 - May 2006 within BioCoRE itself.

Sites with Links to BioCoRE Site (Google, April 2006): 26 domains; 211 pages.

Papers citing BioCoRE: A literature search in April 2006 of the ISI Web of Science citation database yielded the following citations of the BioCoRE origin paper [112]:

- 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.

Further Services

Below we report additional services activities by the Resource. The Resource furnished numerous software tools for biomolecular science, trained visiting scientists,

provided user support, and conducted workshops that provided training on Resource software and computational cluster development.

- *MDTools*

  A collection of programs, scripts, and utilities known collectively as 'MDTools' and posted off the main page of the Resource web site[‡] assist researchers in three areas, 1) simulation tools, 2) databases, and 3) web, programming, and administrative tools. Researchers may use these tools, for example, to make various modeling and simulation tasks easier, or build up larger toolsets by combining the code and utilities listed. Since August 2005, MDTools changes include updates to two tools that assist in using output files from molecular dynamics simulations, and to a collections of code libraries to enable development of molecular dynamics methods; two administrative tools were updated as well. A script library, which will hold useful scripts developed by Resource members that others in the biomedical community may find useful, was also added.

- *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 October 2005 to May 2006, the Resource has hosted 11 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 86,000 users across our technology area)[¶], 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.

[‡]URL:`http://www.ks.uiuc.edu/Development/MDTools/`

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

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

– *Workshops*

The Resource presented three 'hands-on' workshops since May 2005, in Lake Tahoe, California (May 23-27, 2005)[†], Chicago, Illinois (June 9-13, 2005)[‡], San Francisco, California (June 26-30, 2005)[§], Pittburgh, Pennsylvania, (November 28 - December 1, 2005)[¶], and Frankfurt, Germany (March 20-23, 2006)[‖]. The Lake Tahoe and Chicago workshops were supported by NIH funding, the San Francisco workshop by the National Science Foundation[**], the Pittsburgh workshop by the Pittsburgh Supercomputing Center[††], and the Frankfurt workshop by the Max Planck Institute of Biophysics[‡‡]. At each workshop, participants over a period of four to five days receive training in both the operation and application of resource software, and are able to ask Resource faculty and graduate students about using Resource tools for their own research.

Three more workshops presented by the Resource (September 22-23, 2005[†], November 10-11, 2005[‡], and March 16-17, 2005[§]) trained participants in the design, construction, and deployment of PC-based computational clusters. Held at the Beckman Institute, the workshops were organized by Resource staff members, who also provided the lectures and materials to allow participants to construct during the class a fully-functional computational cluster. A fourth cluster workshop is planned for April 2006[¶].

Seminars 2005-2006

Between April 2005 and April 2006 the Resource organized and hosted 21 seminars. An established institution on the UIUC campus, Resource seminars benefit students and faculty from the Beckman Institute as well as other departments and

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

[‡]URL:http://www.ks.uiuc.edu/Training/Workshop/Chicago/

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

[¶]URL:http://www.ks.uiuc.edu/Training/Workshop/PSC/

[‖]URL:http://www.ks.uiuc.edu/Training/Workshop/Frankfurt/

[**]URL:http://www.nsf.gov

[††]URL:http://www.psc.edu/

[‡‡]URL:http://www.mpibp-frankfurt.mpg.de/

[†]URL:http://www.ks.uiuc.edu/Training/Workshop/Cluster1/

[‡]URL:http://www.ks.uiuc.edu/Training/Workshop/Cluster/

[§]URL:http://www.ks.uiuc.edu/Training/Workshop/Cluster3/

[¶]URL:http://www.ks.uiuc.edu/Training/Workshop/Cluster2006/

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 our web site[‖] and are also listed on the main page of the Resource website for greater publicity. Below is a list of the Resource seminars in the past year (mid-April 2005 - start of April 2006):

- April 17, 2006, Mu Gao, University of Illinois, Urbana, IL, *How Anthrax Toxins Enter a Host Cell: A Molecular Dynamics Study*
- April 10, 2006, Peter Freddolino and Anton Arkhipov, University of Illinois, Urbana, IL, *All-atom Molecular Dynamics Study of a Complete Virus*
- March 13, 2006, Kevin Y. Sanbonmatsu, Los Alamos National Laboratory, Los Alamos, NM, *Simulating the Conformational Changes of the Ribosome Observed in Cryo-EM Data*
- March 6, 2006, Carola Hunte, Max-Planck-Institute for Biophysics, Frankfurt, Germany, *Structure and Mechanism of the Cytochrome bc1 Complex*
- March 2, 2006, Petros Koumoutsakos, ETH Zurich, Switzerland, *Multiscale Simulations Using Particles*
- February 27, 2006, Joachim Frank, Howard Hughes Medical Institute, Albany, New York, *The Dynamics of Translation as seen by Cryo-electron Microscopy*
- February 3, 2006, Fadel Samatey, Osaka University, Osaka, Japan, *The Bacterial Flagellum: Structure and Mechanism*
- January 30, 2006, Nathan Baker, Washington University , St. Louis, MO, *Multiscale Methods for Biomolecular Systems: Solvation Models and Membrane Electromechanical Coupling*
- December 1, 2005, Janna Maranas, The Pennsylvania State University, University Park, PA, *The Influence of Trehalose on Melting and Dynamics in Dehydrated Phospholipid Bilayers*
- November 21, 2005, Gianni De Fabritiis, University College of London, London, UK, *Hybrid Molecular Dynamics: Modeling the Mesoscale with Molecular Specificity*
- November 14, 2005, Emma Falck, Beckman Postdoctoral Fellow, University of Illinois, Urbana, IL, *Multiple Scales in Computational Modeling of Soft and Biological Matter*

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

- October 4, 2005, Jan Siewert Marrink, University of Groningen, The Netherlands, *Simulating Phase Transformations of Lipid Membranes*
- August 22, 2005, Mark S. P. Sansom, University of Oxford, Oxford, England, *Modelling a Complex Environment Molecular Dynamics Simulations of Membrane Proteins*
- August 19, 2005, Maria Ghirardi, National Renewable Energy Laboratory, Golden, CO, *Studies on the Expression and Activity of the [FeFe]-hydrogenase Enzyme from Chlamydomonas reinhardti*
- June 30, 2005, Goran Neshich, Cidade Universitria, Brazil, *The Identification of the Folding Essential Residues and Active Site Residues by looking at an extensive DB of the Structure Descriptors in Diamond STING*
- June 15, 2005, Robert Fillingame, University of Wisconsin, Medical School, Madison, WI, *H+ Transporting Rotary Motor of ATP Synthase: Revisiting Old Data with New Structural Models*
- June 14, 2005, Paul Grayson, California Institute of Technology, Pasadena, CA, *Genome Length and Ejection Force in Bacteriophage Lambda*
- June 1, 2005, Keiichi Namba, Osaka University, Osaka, Japan, *Molecular Mechanisms of Bacterial Swimming and Tumbling*
- May 18, 2005, Keith Moffat, The University of Chicago, Chicago, IL, *Nanosecond Time-resolved Crystallography of Blue Light Photoreceptors: What Atoms Move, When?*
- May 2, 2005, Eduardo Perozo, University of Virginia Health Sciences Center, Charlottesville, VA, *Structure and Dynamics of Ion Channels*
- April 28, 2005, Aleksei Aksimentiev, University of Illinois, Urbana, IL, *Transistors to DNA - Using Nanoscale Electronics for Sequencing Genomes*

Training

The Resource recognizes the vital importance of training to the education and professional growth of young scientists. In the last year, the Resource has continued and expanded its training efforts, such as conducting five workshops, and expanded its collection of web-accessible tutorials, such as adding a set of case studies that encourage hands-on examination of molecular systems. Such efforts are in addition to more traditional training programs for graduate student and post-doctoral researchers and instructional presentations about Resource software. Training provided by the Resource capitalized on a range of tools and media:

- Workshops at national and international locations
- Tutorial development and distribution (traditional and on-line)
- Development of case studies posted at the Resource web site
- Presentations and lectures about Resource software
- Graduate student and postdoc training
- Visitor program

Hands-On Workshops in Computational Biophysics

The Resource in the last year (March 2005 - April 2006) presented five 'hands-on' workshops in computational biophysics, in Lake Tahoe, Chicago, San Francisco, Pittsburgh, and overseas in Franfurt, Germany*. 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 course utilized case studies including the properties of membranes and membrane proteins, mechanisms of molecular motors, trafficking in the living cell through water and ion channels, and signaling pathways. 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.

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

Supporting the 'hands-on' component of the Lake Tahoe, Chicago, and San Francisco workshops - none of which had computing facilities - were 22 Macintosh Powerbook G4 laptop computers purchased with local funds for the workshops by the Resource. Loaded with tutorial required software, including VMD, NAMD, Mathematica, Matlab, Spartan, and other programs, and with files needed for each tutorial, the laptops provided sufficient computational power, high-quality graphics, and portability to each site. At each of these three sites, workshop teaching assistants set up the laptops for afternoon laboratory sessions, with two of the laptops reserved for instructor presentations and demonstrations. In this fashion, all participants were provided with needed learning resources, without the workshop having to rely on the availability of computer labs in various locations.

Initially, enrollment in the workshops was limited to 20 participants at each site, but with the development of tutorial versions amenable to any laptop platform, at the Franfurt workshop enrollment was expanded to 37 participants. More detail and participant evaluation of the workshops follow below.

Lake Tahoe Workshop. The first workshop was held May 23-27, 2005, in Lake Tahoe, California, and included lectures from three Resource faculty (Klaus Schulten, Zaida Luthey-Schulten, Emad Tajkhorshid) and tutorial support from two graduate student teaching assistants. Funding for the workshop was provided by the National Institutes of Health. Twenty academic participants, almost evenly distributed amongst undegraduate, graduate, and doctoral degree holders in terms of educational background, and from primarly United States institutions, attended morning lectures (e.g., "Molecular Graphics Perspective of Protein Structure and Function") and participated in afternoon labs each day. Evaluation results indicate that all participants felt that the workshop broadened their understanding of theoretical and computational biophysics principles, and that 85% estimated the workshop improved their ability to carry out original research on the topics of the workshop. More detail on the Lake Tahoe workshop is available at the Resource website[†].

Chicago Workshop. The next workshop was held June 9-13, 2005, in Chicago, Illinois, again with funding support from the National Institutes of Health. Three Resource faculty (Klaus Schulten, Zaida Luthey-Schulten, Emad Tajkhorshid) provided lectures, and three graduate student teaching assistants prepared laptops for afternoon tutorials (e.g., "Simulation of Water Permeation through Nanotubes") and helped participants during the tutorials. Of the twenty participants, one-fourth were working on an advanced degree, while about equal numbers of the other participants had achieved a masters degree or a doctoral-level degree. Evaluation results

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

were again positive, with all participants indicating their knowledge of theoretical and computational biophysiscs had broadened, and with 85% indicating the workshop improved their ability to carry out original research in the workshop's content area. A web site dedicated to summarizing the Chicago workshop contains more detail[‡].

San Francisco Workshop. A third workshop was held in San Francisco, California, from June 26-30, 2005, and again supported by a team of three Resource faculty (Klaus Schulten, Zaida Luthey-Schulten, Emad Tajkhorshid) and three graduate student teaching assistants. Participants were the most highly educated group to attend any of the workshops, with 17 out of 20 at the doctoral level, and several had non-academic affiliations (i.e., government, industry). Lecture topics in the morning (e.g., "Transport in Aquaporins") were followed by afternoon computer labs (e.g., "Stretching Deca-alanine") utilizing the portable computer lab. Results from the workshop evaluation found that all participants agreed that the workshop broadened their understanding of concepts and principles in computational and theoretical biophysics, and 84% concluded that the workshop improved their ability to carry original research in that area. Funding for this workshop was provided by the National Science Foundation[§]. More on the San Francisco workshop can be found at the Resource web site on this event[¶].

Pittsburgh Workshop. Held November 28 - December 1, 2005 at the Pittsburgh Supercomputing Center (PSC) in Pittsburgh, Pennsylvania, this workshop was funded and primarily organized (i.e., except for curricular materials) by PSC for 24 participants. Two Resource faculty (Klaus Schulten, Emad Tajkhorshid) provided lectures on topics such as "Equilibrium Properties of Proteins" and "Nanotubes", while two graduate student teaching assistants helped participants through tutorials at a computer laboratory prepared for the workshop. An evaluation conducted by PSC found 87% strongly agreeing with the statement "I gained new knowledge and insights" and 82% indicating they would recommend the workshop to others. Further evaluation and other details can be found at the Resource web site for this workshop[‖] and at the PSC site for the workshop[**].

Frankfurt Workshop. The first overseas workshop since the Perth, Australia workshop in 2004, the March 20-23, 2006 was organized and funded by the Max Planck

[‡]URL:http://www.ks.uiuc.edu/Training/Workshop/Chicago/

[§]URL:http://www.nsf.gov/

[¶]URL:http://www.ks.uiuc.edu/Training/Workshop/SanFrancisco/

[‖]URL:http://www.ks.uiuc.edu/Training/Workshop/PSC/

[**]URL:http://www.psc.edu/biomed/training/workshops/2005/NAMD/

Institute of Biophysics, and held on site at their facility in Frankfurt, Germany. Three Resource faculty (Klaus Schulten, Zaida Luthey-Schulten, Emad Tajkhorshid) provided lectures (e.g., "Introduction to Bioinformatics") and three graduate student teaching assistants helped 37 participants through cross-platform versions of the tutorials, which participants worked through on their own laptops that they had prepared with needed software and files before the event. Ratings of the relevance of the lecture and tutorial topics are quite high, for example 84% found the molecular graphics and molecular dynamics rating relevant, and 92% found the VMD molecular graphics tutorial relevant. The Resource website on the Frankfurt workshop provides more information on this event[††].

Plans for further workshops are currently under development, with one tentative workshop planned for November 2006 in Pittsburgh, and the Resource has also videotaped one lecture, with the goal of streaming the video as part of an online course.

Tutorials

The availability and utility of the tutorials developed and offered by the Resource was greatly expanded during the last funding period by their conversion to cross-platform versions, or versions amenable to use on all of the most popular computer operating system platforms - Windows, Mac OSX, and Linux/Unix. Originally, the tutorials were developed for use at a local UIUC computer laboratory, and were essentially written around that system. Later editing made the tutorials usable on the Mac system, and were supported by the use of Resource purchased, specially-prepared Mac laptops at several workshops. After more testing, feedback, and revision, the cross-platform tutorials listed below, available via the Resource website[‡‡], can now be used on all platforms, further opening up their use to the biomedical community.

- VMD Molecular Graphics
- NAMD Tutorial (two versions, one for Windows, a second for Mac/Linux)
- Topology File Tutorial
- Stretching Deca-alanine
- Simulation of Water Permeation Through Nanotubes
- Evolution of Protein Structure: tRNA Synthetase

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

[‡‡]URL:`http://www.ks.uiuc.edu/Training/TutorialsOverview/index.html`

Case Studies

A new training tool available on the Resource web site[†] are seven case studies that exploit the molecular graphics program VMD[‡] for teaching molecular cell biology. The case studies start out like a conventional textbook chapter, but utilize VMD molecular graphics to offer a much more detailed view of the subjects than commonly possible in textbooks. Each case study shows the molecular systems through graphical images, but also provides the files that permit the reader to regenerate the images with few mouse clicks. Students can rotate the images, enlarge them, alter the views in many ways, and analyze structures and sequences, which encourages the reader to explore in depth the structures shown. A list of the case studies and a short description of each is provided below.

- *Case Study: Aquaporin*
  The organization of water is critical to most biological processes. Although cell membranes are to some extent water-permeable, they cannot facilitate the rapid exchange of large volumes of water, as required by the kidneys in the human body. Aquaporins (AQPs) are a family of specialized integral membrane proteins which function mainly as water channels. These highly efficient water channels can explain how we secrete tears, saliva and sweat, how our kidneys concentrate urine, and how our brains maintain spinal fluids...
- *Case Study: BPTI*
  Bovine pancreatic trypsin inhibitor (BPTI) is one of the smallest and simplest globular proteins that inhibits protein digestive action (digesting protein to peptides) of the enzyme trypsin in cow (bovine) pancreas. BPTI is a member of the family of serine protease inhibitors. These enzymes have many conserved cysteines that form disulfide bonds that stabilize protein three-dimensional structures. BPTI has a relatively broad specificity in that it can inhibit several kinds of digestive enzymes...
- *Case Study: Membranes*
  Membranes are essential to cellular organisms. They are like fortresses in that they provide a barrier between the inside and outside with guarded drawbridges in the form of proteins that regulate the influx and efflux of material. Unlike the rigid walls of a fortress, membranes are fluid and are able to bend and move. The bricks forming a membrane, called lipids, freely move. Plasma membranes enclose and define the boundaries of a cell, maintaining a barrier between the interior of a cell, the cytosol, and the extracellular environment...

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

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

- *Case Study: Titin Ig Domains*
  In the extreme sport of bungee jumping, a daring athlete leaps from a great height and free-falls while a tethered cord tightens and stretches to absorb the energy from the descent. The bungee cord protects the jumper from serious injury, because its elasticity allows it to extend and provide a cushioning force that opposes gravity during the fall. Amazingly, nature also uses elasticity to dampen biological forces at the molecular level, such as during extension of a muscle fiber under stress. The molecular bungee cord that serves this purpose in the human muscle fiber is the protein titin, which functions to protect muscle fibers from damage due to overstretching...
- *Case Study: Ubiquitin*
  Without a doubt, the most organized and coordinated machine known is the biological cell. Inside its micrometer-scale diameter, a wide variety of macromolecules (DNA, proteins, sugars, lipids, etc.) work together in a cooperative way, balancing energy and matter to keep the cell alive. To maintain harmony and efficiency between various functions, most processes have to be turned on or off according to different cellular stages and changes within the environment. To this end, together with the mechanisms to assemble functional proteins and to turn on their functions, there should be counterparts to suppress and disassemble proteins when they are no longer needed. Ubiquitin is a key player in eukaryotic intracellular protein degradation...
- *Case Study: Water and Ice*
  Water is essential for sustaining life on Earth. Almost 75% of the Earth's surface is covered by it. It composes roughly 70% of the human body by mass. It is the medium associated with nearly all microscopic life processes. Much of the reason that water can sustain life is due to its unique properties...

Cluster Building Workshops

The Resource has produced three workshops, each a day and a half long, on the design, construction, and deployment of computational clusters, in September and November of 2005, and in March of 2006. 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, as well as how cluster design interacts with queuing systems. Additionally, participants are given the chance to 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.

- September 22-23, 2005, Cluster Building Workshop[§]. Evaluation results found that a majority of the responding participants indicated that the workshop broadened their understanding of cluster building concepts 89%, and improved their ability to set up a cluster 83%. A total of 22 persons attened the workshop.
- November 10-11, 2005, Cluster Building Workshop[¶]. Again, a majority of the responding participants indicated that the workshop broadened their understanding of cluster building concepts 91%, and improved their ability to set up a cluster 82%. The workshop was fully attended, with 24 participants at the event.
- March 16-17, 2006, Cluster Building Workshop[‖]. The highest ratings were found by the evaluation of this workshop, with 91% of the responding participants indicating the workshop broadened their understanding of cluster building concepts, and all participants indicating the workshop improved their ability to set up a cluster. Twenty-two participants were part of the workshop.

Another Cluster Building Workshop is planned for April 20-21, 2006.

Resource Library

The Resource library has been expanded by the purchase of 49 new books. Further, to supplement the UIUC library's collection of on-line and print journals, the Resource subscribes to the following journals in the sciences and computing:

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

[§]URL:`http://www.ks.uiuc.edu/Training/Workshop/Cluster1/`

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

[‖]URL:`http://www.ks.uiuc.edu/Training/Workshop/Cluster3/`

- Nature
- Chronicle of Higher Education

Software Presentations

In addition to presentations and tutorials about Resource software that are regular features of the hands-on workshops, other presentations about Resource sofware were provided over the past funding period, as listed below**:

- November 2005, Seattle, WA, SC—05 International Conference for High Performance Computing, Networking, Storage and Analysis, *Exploring Biomolecular Machines with Supercomputers* (Jim Phillips, Jordi Cohen)
- November 2005, Urbana, IL, National Center for Supercomputing Applications, Social Networks and Cyberinfrastructure Conference, *BioCoRE Overview* (Kirby Vandivort)
- October 2005, Urbana, IL, University of Illinois, UIUC CITES Computer Consultant Support Program Conference, *Linux Clustering for Scientific Computing*(Jim Phillips, Tim Skirvin)
- September 2005, San Francisco, CA, University of California-San Diego, San Diego Supercomputer Center, Visualization of Large Biomolecular Complexes Workshop, *VMD: Algorithms and Methods for Large Scale Biomolecular Visualization* (John Stone)

Visitors

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

- Amit Roy, Weill Medical College, Cornell University (May, 2006)
- Ilya Solov'yov, Frankfurt Institute for Advanced Studies, Johann Wolfgang Goethe University (April-May 2006)

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

- Lea Thogerson, Bioinformatic Research Center, University of Aarhus (April, 2006)
- Christopher Chipot, Universite Henri Poincare, France (November, 2005 and February-March, 2006)
- Francois Dehez, Universite Henri Poincare, France (February-March, 2006)
- Danilo Gonzalez, University of Talca, Chile (November-December, 2005)
- Ana Verde, Penn State University (November-December, 2005)
- Ana Araujo, Penn State University (November-December, 2005)
- Jerome Henin, Universite Henri Poincare, France (November, 2005)
- Paul McCreary, Tulane University (October-December, 2005)
- Thomas Bishop, Tulane University (October-November, 2005)

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:

- Deyu Lu Ph.D., Physics, University of Illinois, Fall 2005
- David Hardy Ph.D., Computer Science, University of Illinois, Fall 2005

Postdoctoral Associates: Postdocs currently receiving training at the Resource are:

- Dr. Emma Falck
- Dr. David Hardy
- Dr. Mu Gao
- Dr. Markus Dittrich

Dissemination

Broadscale efforts in dissemination and outreach through the last funding period took advantage of a variety of available delivery mechanisms, from distribution of Resource-produced papers and know-how through the web site, through talks, meetings, workshops, and conferences, software distribution, brochure production, news stories and press releases, and use of Resource images in a variety of third-party publications and academic presentations.

The Resource was featured in a variety of printed and online media for its scientific and computational accomplishments. Of particular media interest was Resource research on the complete simulation of the satellite tobacco mosaic virus*, research on the generation of hydrogen fuel†, and methods of visually depicting aquaporin‡.

Media Coverage

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

- Staff. (2006, March 23). Supercomputer Maps One Million Atoms of a Complete Virus in First Simulation of a Life Form. *National Science Foundation, Press Release 06-049*.
  http://www.nsf.gov/news/news_summ.jsp?cntn_id=106791
- Farrell, N. (2006, March 16). Supercomputer Builds a Virus. Now If It Can Convince a Human To Swallow It. *The Inquirer*.
  http://www.theinquirer.net/?article=30330
- Pearson, H. (2006, March 14). Supercomputer Builds a Virus. Vast Simulation Captures Molecules in Motion. *news@nature.com*.
  http://www.theinquirer.net/?article=30330
- Barlow, J. (2006, March 14). Researchers Simulate Complete Structure of Virus-On Computer. *UIUC News Bureau*.
  http://www.news.uiuc.edu/news/06/0314virus.html See also:

*URL:http://www.news.uiuc.edu/news/06/0314virus.html

†URL:http://www.news.uiuc.edu/news/05/1005hydrogen.html

‡URL:http://www.news-gazette.com/news/local/2005/09/04/researchers_art_tells_scientific_story/

§URL:http://www.ks.uiuc.edu/Publications/stories.cgi

  - (2006, March 14). Researchers Simulate Complete Structure of Virus – On a Computer. *EurekAlert!* http://www.eurekalert.org/pub_releases/2006-03/uoia-rsc031006.php
  - (2006, March 14). Computational Science Breakthrough Simulates an Entire Life Form. Beckman Institute Researchers Make History By Simulating The Tiny Satellite Tobacco Mosaic Virus. *Beckman Institute News.* http://www.beckman.uiuc.edu/news/NewsReleases/STMVsimulation.html

- Staff. (2006, March 10). NCSA Users Publish First Simulation of an Entire Life Form. *NCSA Access.* http://access.ncsa.uiuc.edu/Stories/TobaccoMosaic/ See also:
  - Staff. (2006, March 13). NCSA Users Simulate an Entire Life Form. *HPCWire.* http://www.hpcwire.com/hpc/589081.html
- Staff. (2006, February 28). NCSA Users Publish on Nuclear Pore Selectivity. *NCSA Access.*
- Silicon Graphics. (2006, February 7). Bioinformatics Consortium at the University of Missouri adds SGI Technology for Large-Scale Computational Life Sciences Research. *WebWire Internet Press Release Resource.* http://webwire.com/ViewPressRel.asp?SESSIONID=&aId=9074
- Kline, G. (2006, January 15). A Tool To Find Building Blocks: With an Eye on Early Detection of Health Issues, Researchers Developing Ways To Map Genes. *The News-Gazette.*
- Palucka, T. (2006, January 13). All In Your Brain. Progress Toward New Medicinal Drugs For Appetite, Memory, and Mood. *PSC News and Media Online, and PSC's Projects in Scientific Computing, 2005.* http://www.psc.edu/science/2005/brain/
- Kotulak, R. (2006, January 8). The Cosmic Conversation. How Can Lifeless Particles Evolve into Living Things? They Basically Talk Themselves into It, a Group of Scientists Say. *Chicago Tribune Magazine.* http://www.chicagotribune.com/features/magazine/chi-0601080429jan08,1,5317559.story?page=1&ctrack=1&cset=true&coll=chi-leisuremagazine-hed See also:
  - Kotulak, R. (2006, January 20). Lifeless Particles Talk Themselves into Evolving, Scientists Say. *The Kansas City Star.* http://www.kansascity.com/mld/kansascity/news/nation/13670670.htm
  - Kotulak, R. (2006, January 20). Lifeless Particles Talk Themselves into Evolving, Scientists Say. *ContraCostaTimes.com.* http://www.contracostatimes.com/mld/cctimes/news/nation/13670670.htm

- Barlow, J. (2005, December 21). Membrane Research Opens Window To Benefits For Plants, Humans. *UIUC News Bureau*. http://www.news.uiuc.edu/news/05/1221plantprotein.html
- Staff. (2005, December 12). Finding How Water Channels - Aquaporins - Open and Close May Lead To New Drugs. *News-Medical.net*. http://www.news-medical.net/?id=14941
- Salamone, S. (2005, November 23). IT and Informatics: Weekly New Product Roundup. *Bio-IT World*. http://www.bio-itworld.com/newsitems/2005/nov2005/11-23-05-news-products
- Barlow, J. (2005, September 28). U. of I. Researchers To Play Key Roles in Study of How Life Emerged on Earth. *UIUC News Bureau*. http://www.news.uiuc.edu/news/05/0928earlylife.html
- Barlow, J. (2005, October 5). Research Advances Understanding of How Hydrogen Fuel Is Made. *UIUC News Bureau*. http://www.news.uiuc.edu/news/05/1005hydrogen.html See also:
  - Staff. (2005, December 9). Illinois Scientists Closing in on Renewable Energy Source. *NewEnergyReport.org*. http://www.newenergyreport.org/013671.html
  - Staff. (2005, December 7). Research Advances Hydrogen Fuel Production. *RenewableEnergyAccess.com*. http://renewableenergyaccess.com/rea/news/story?id=40207
  - Staff. (2005, October 20). Research Advances Understanding of How Hydrogen Fuel Is Made. *Physicsorg.com*. http://www.physorg.com/news7005.html
  - Staff. (2005, October 6). Research Advances Understanding of How Hydrogen Fuel Is Made. *ScienceDaily.com*. http://www.sciencedaily.com/releases/2005/10/051006082610.htm
- Staff. (2005, September 5). The Novartis and The Daily Telegraph "Visions of Science" Photographic Awards. Winners - 2005. Concepts. "Ion Channel in The Nervous System" by Dr. Oliver Beckstein. (uses VMD to create image). *Visions of Science website*. http://www.visions-of-science.co.uk/f-2005winners.htm
- Kline, G. (2005, September 4). Researcher's Art Tells Scientific Story. *The News-Gazette Online*. http://www.news-gazette.com/localnews/story.cfm?Number=18910 See also:
  - Kline, G. (2005, November 1). It Was Not Meant To Be Art, But It Is Striking. *Quad City Times*.

http://www.qctimes.net/articles/2005/11/01/features/celebrate/doc43670cfeac11a001355112.txt

* Kline, G. (2005, October 26). Illinois Professor Turns Molecules into Artwork. Models Not Meant as Art, But Striking. *Indiana Daily Student*. http://www.idsnews.com/news/story.php?adid=search&id=32050

- Staff. (2005, August 19). Beckman Researchers Get Big Grant For Developing DNA Sequencing Technology. *Beckman in the News*. http://www.beckman.uiuc.edu/news/featured/DNAgrant.html
- Staff. (2005, August 17). Illinois Researchers Net Federal Genome Funding. *Engineering at Illinois News*. http://www.engr.uiuc.edu/news/index.php?xId=066308960700
- Barlow, J. (2005, August 10). Cells Direct Membrane Traffic By Channel Width, Scientists Say. *UIUC News Bureau*. http://www.news.uiuc.edu/news/05/0810channels.html

Publications

In the past year Resource members have published and/or submitted or presented:

- 32 articles in refereed journals, handbooks, and other publications (2 in press)
- 40 talks by PIs
- 36 talks or posters by other Resource members

Published Articles

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

- Alexander Balaeff, L. Mahadevan, and Klaus Schulten. Modeling DNA loops using the theory of elasticity. Physical Review E, 73:031919, 2006.
- James Gumbart and Klaus Schulten. Molecular dynamics studies of the archaeal translocon. Biophysical Journal, 90:2356-2367, 2006.
- Eric H. Lee, Mu Gao, Nikos Pinotsis, Matthias Wilmanns, and Klaus Schulten. Mechanical strength of the titin Z1Z2/telethonin complex. Structure, 14:497-509, 2006.
- Deyu Lu, Aleksei Aksimentiev, Amy Y. Shih, Eduardo Cruz-Chu, Peter L. Freddolino, Anton Arkhipov, and Klaus Schulten. The role of molecular modeling in bionanotechnology. Physical Biology, 3:S40-S53, 2006.

- Maria E. Gracheva, Anlin Xiong, Jean-Pierre Leburton, Aleksei Aksimentiev, Klaus Schulten, and Gregory Timp. Simulation of the electric response of DNA translocation through a semiconductor nanopore-capacitor. Nanotechnology, 17:622-633, 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.

- Amy Y. Shih, Anton Arkhipov, Peter L. Freddolino, and Klaus Schulten. A coarse grained protein-lipid model with application to lipoprotein particles. Journal of Physical Chemistry B, 110:3674-3684, 2006.

- J. B. Heng, A. Aksimentiev, C. Ho, P. Marks, Y. V. Grinkova, S. Sligar, K. Schulten, and G. Timp. The electromechanics of DNA in a synthetic nanopore. Biophysical Journal, 90:1098-1106, 2006.

- Deyu Lu, Yan Li, Umberto Ravaioli, and Klaus Schulten. Ion-nanotube terahertz oscillator. Physical Review Letters, 95:246801, 2005.

- Markus Dittrich and Klaus Schulten. Zooming in on ATP hydrolysis in F1. Journal of Bioenergetics and Biomembranes, 37:441-444, 2005.

- Timothy A. Isgro and Klaus Schulten. Binding dynamics of isolated nucleoporin repeat regions to importin-b. Structure, 13:1869-1879, 2005.

- James C. Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kale, and Klaus Schulten. Scalable molecular dynamics with NAMD. Journal of Computational Chemistry, 26:1781-1802, 2005.

- Laxmikant V. Kale, Klaus Schulten, Robert D. Skeel, Glenn Martyna, Mark Tuckerman, James C. Phillips, Sameer Kumar, and Gengbin Zheng. Biomolecular modeling using parallel supercomputers. In S. Aluru, editor, Handbook of computational molecular biology, pp. 34.1-34.43. Taylor and Francis, 2005.

- Jerome Math, Aleksei Aksimentiev, David R. Nelson, Klaus Schulten, and Amit Meller. Orientation discrimination of single stranded DNA inside the a-hemolysin membrane channel. Proceedings of the National Academy of Sciences, USA, 102: 12377-12382, 2005.

- James Gumbart, Yi Wang, Aleksei Aksimentiev, Emad Tajkhorshid, and Klaus Schulten. Molecular dynamics simulations of proteins in lipid bilayers. Current Opinion in Structural Biology, 15:423-431, 2005.

- Melih K. Sener, Craig Jolley, Adam Ben-Shem, Petra Fromme, Nathan Nelson, Roberta Croce, and Klaus Schulten. Comparison of the light harvesting networks of plant and cyanobacterial photosystem I. Biophysical Journal, 89:1630-1642, 2005.

- J. B. Heng, A. Aksimentiev, C. Ho, V. Dimitrov, T. Sorsch, J. Miner, W. Mansfield, K. Schulten, and G. Timp. Beyond the gene chip. Bell Labs Technical Journal, 10:5-22, 2005.

- Jordi Cohen, Kwiseon Kim, Paul King, Michael Seibert, and Klaus Schulten. Finding gas diffusion pathways in proteins: Application to O2 and H2 transport in CpI [FeFe]-hydrogenase and the role of packing defects. Structure, 13:1321-1329, 2005.

- Markus Dittrich, Peter L. Freddolino, and Klaus Schulten. When light falls in LOV: A quantum mechanical/molecular mechanical study of photoexcitation in Phot-LOV1 of Chlamydomonas reinhardtii. Journal of Physical Chemistry B, 109:13006-13013, 2005.

- Yi Wang, Klaus Schulten, and Emad Tajkhorshid. What makes an aquaporin a glycerol channel: A comparative study of AqpZ and GlpF. Structure, 13:1107-1118, 2005.

- J. B Heng, A. Aksimentiev, C. Ho, P. Marks, Y. V. Grinkova, S. Sligar, K. Schulten, and G. Timp. Stretching DNA using an electric field in a synthetic nanopore. Nano Letters, 5:1883-1888, 2005.

- Deyu Lu, Yan Li, Umberto Ravaioli, and Klaus Schulten. Empirical nanotube model for biological applications. Journal of Physical Chemistry B, 109:11461-11467, 2005.

- Marcos Sotomayor, David P. Corey, and Klaus Schulten. In search of the hair-cell gating spring: Elastic properties of ankyrin and cadherin repeats. Structure, 13:669-682, 2005.

- Elizabeth Villa, Alexander Balaeff, and Klaus Schulten. Structural dynamics of the Lac repressor-DNA complex revealed by a multiscale simulation. Proceedings of the National Academy of Sciences, USA, 102:6783-6788, 2005.

- Aleksei Aksimentiev and Klaus Schulten. Imaging alpha-hemolysin with molecular dynamics: Ionic conductance, osmotic permeability and the electrostatic potential map. Biophysical Journal, 88:3745-3761, 2005.

- Maria L. Ghirardi, Paul W. King, Matthew C. Posewitz, Pin Ching Maness, Alexander Fedorov, Kwiseon Kim, Jordi Cohen, Klaus Schulten, and Michael Seibert. Approaches to developing biological H2-photoproducing organisms and processes. Biochemical Society Transactions, 33:70-72, 2005.

- Jordi Cohen, Kwiseon Kim, Matthew Posewitz, Maria L. Ghirardi, Klaus Schulten, Michael Seibert, and Paul King. Molecular dynamics and experimental investigation of H2 and O2 diffusion in [Fe]-hydrogenase. Biochemical Society Transactions, 33:80-82, 2005.

- Emad Tajkhorshid, Fangqiang Zhu, and Klaus Schulten. Kinetic theory and simulation of single-channel water transport. In S. Yip, editor, Handbook of Materials Modeling, Vol. I: Methods and Models, pp. 1797-1822. Springer, Netherlands, 2005.

- Yan Li, Deyu Lu, Slava V. Rotkin, Klaus Schulten, and Umberto Ravaioli. Screening of water dipoles inside finite-length armchair carbon nanotubes. Journal of Computational Electronics, 4:161-165, 2005.

- Melih Sener and Klaus Schulten. Physical principles of efficient excitation transfer in light harvesting. In David L. Andrews, editor, Energy Harvesting Materials, pp. 1-26. World Scientific, Singapore, 2005.

- Emad Tajkhorshid, Jordi Cohen, Aleksei Aksimentiev, Marcos Sotomayor, and Klaus Schulten. Towards understanding membrane channels. In Boris Martinac and Andrzej Kubalski, editors, Bacterial ion channels and their eukaryotic homologues, pp. 153-190. ASM Press, Washington, DC, 2005.

- Amy Y. Shih, Ilia G. Denisov, James C. Phillips, Stephen G. Sligar, and Klaus Schulten. Molecular dynamics simulations of discoidal bilayers assembled from truncated human lipoproteins. Biophysical Journal, 88:548-556, 2005.

The following two publications are currently in press:

- Mu Gao and Klaus Schulten. Onset of anthrax toxin pore formation. Biophysical Journal, 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, 2006.

Lectures and Talks

The Resource PIs gave the following talks in the last 12 months[¶]:

Klaus Schulten

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

- March 2006, Frankfurt, Germany, Max Planck Institute of Biophysics, Colloquium, *Towards Understanding Membrane Channels*
- March 2006, Frankfurt, Germany, Max-Planck Institute for Biophysics, Hands-On Workshop on Computational Biophysics
  - *Introduction to Protein Structure*
  - *Statistical Mechanics of Proteins*
- January 2006, Osaka, Japan, Symposium on Dynamics of Biological Systems, *Large Scale Biomolecular Simulation of Cellular Processes*
- December 2005, Bonn, Germany, FIAS seminar / DFG's Review on Mathematics and Scientific Computing, *Proteins with Mechanical Functions*
- November-December 2005, Pittsburgh, PA, Pittsburgh Supercomputing Center, Hands-On Workshop on Computational Biophysics
  - *Introduction to Protein Structure*
  - *Statistical Mechanics of Proteins*
- November 2005, Cold Spring Harbor Laboratory, The Intracellular Molecular Environment, *In situ Molecular Modeling of Cellular Processes*
- October 2005, Zurich, Switzerland, ETHZ Workshop on Multiscale Modelling and Simulation
  - *Physics of Photosynthesis*
  - *The Many Faces of Aquaporins*
  - *Single Molecule Electrical Recording with Natural and Synthetic Nanoporins*
  - *Proteins with Mechanical Functions*
  - *Quantum Biology*
- October 2005, ETHZ Workshop on Multiscale Modelling and Simulation, *Molecular Double Motor ATP Synthase*
- October 2005, Tutzing, Germany, 87th International Bunsen Discussion Meeting, *Mechanical Functions of Proteins*
- September 2005, Los Angeles, CA, IPAM / Multiscale Modeling in Soft Matter and Bio-Physics / UCLA, *Multiscale simulations of DNA-protein, protein-lipid compelxes*

- September 2005, Urbana, IL, University of Illinois, Physics Colloquium, *What is Life? An Answer Sought from Photosynthetic Bacteria*
- August 2005, Urbana, IL, University of Illinois, Molecular and Cellular Biology, *Advances in Molecular Cell Biology and Computational Biophysics*
- August 2005, Bad Honnef Germany, Physikzentrum Bad Honnef, WE-Heraeus-Seminar, *Biochemical Mechanisms for Magnetic Orientation in Animals*

Laxmikant Kale

- September 2005, Knoxville, TN, Oak Ridge National Laboratory, Future Technologies Colloquium Series, *Adaptive MPI: Intelligent runtime strategies and performance prediction via simulation*

Emad Tajkhorshid

- March 2006, Frankfurt, Germany, Max-Planck Institute for Biophysics, Hands-On Workshop on Computational Biophysics, *Simulating Membrane Channels*
- 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*
- February 2006, Salt Lake City, UT, 50th Annual Meeting of the Biophysical Society, *Coupling of Proton Translocation and Protein Conformational Changes in Lactose Permease*
- February 2006, Boston, MA, Brandeis University, *Visualizing the Art of Selective Transport in Membrane Channels at Full Atomic Resolution*
- February 2006, Urbana, IL, University of Illinois at Urbana-Champaign, School of Molecular and Cellular Biology, *Visualizing the Art of Selective Transport in Membrane Channels at Full Atomic Resolution*
- November-December 2005, Pittsburgh, PA, Pittsburgh Supercomputing Center, Hands-On Workshop on Computational Biophysics

  - *Parameters for Classical Force Fields*
  - *Simulating Membrane Channels*

- November 2005, New Haven, CT, Yale University, Department of Physiology, Departmental seminar, *Transmembrane Traffic of Materials Through Pure Lipid Bilayers and Membrane Channels*

- October 2005, Philadelphia, PA, Membranes and Ion Channels, e- cheminfo 2005, *Understanding Membrane Transport at Full Atomic Resolution: Molecular dynamics simulations of lipid bilayers and membrane channels*

- October 2005, Columbus, OH, Ohio State University, Biological Membranes: Structure and Function, *Transmembrane Traffic of Materials Through Pure Lipid Bilayers and Membrane Channels*

- October 2005, Chicago, IL, Rush Medical Center, Departmental Seminar, *Visualizing the Art of Selective Transport in Membrane Channels*

- October 2005, Chicago, IL, Rush University, Department of Physiology, Departmental seminar, *Visualizing the Art of Selective Transport in Membrane Channels*

Zaida Luthey-Schulten

- March 2006, Frankfurt, Germany, Max-Planck Institute for Biophysics, Hands-On Workshop on Computational Biophysics, *Introduction to Bioinformatics*

Aleksei Aksimentiev

- March 2006, Baltimore, MD, American Physical Society Meeting, *Microscopic Kinetics of DNA Translocation Through Synthetic and Biological Nanopores*

- March 2006, Ventura, CA, Gordon Research Conference on Protons and Membrane Reactions, *Animating Atomic-Detail Structures of Ion-Motive Atpases with Molecular Dynamics*

- February 2006, Salt Lake City, UT, Biophysical Society Meeting, *Computing The Conductance of A Membrane Channel From All-Atom Molecular Dynamics Simulations*

- September 2005, Urbana, IL, Focused Workshop on Electronic Recognition of Biomolecules, *Crawling with DNA Through a Nanopore: Molecular Dynamics Perspective*

- August 2005, Bremen, Germany, Summer School: Biosensing with channels, *Molecular Dynamics Simulations of DNA Translocation through Synthetic and Biological Nanopores*

Other Resource members gave the following presentations, talks, poster presentations, or attended meetings:

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

- March 2006, Pittsburgh, PA, PSC, Stiles Group Lab, *Insights into the Mechanism of Molecular Motors* (Markus Dittrich)
- March 2006, Albuquerque, NM, Sandia National Labs, *Quantum Mechanical/Molecular Mechanical Simulations of Biomolecular Systems* (Markus Dittrich)
- February, 2006, Salt Lake City, UT, Biophysical Society 50th Annual Meeting
  - *A Coarsed Grained Protein-Lipid Model with Application to High-Density Lipoprotein Particles* (Amy Y. Shih, Anton Arkhipov, Peter L. Freddolino)
  - *Molecular Basis of Gating in Aquaporin Water Channels* (Yi Wang)
  - *Dynamics of the Translocon SecY Investigated Through MD* (James Gumbart)
  - *Channel Mediated Gas Transport Across Lipid Membranes* (Yi Wang)
  - *Tertiary and Secondary Structure Elasticity of Ankyrin Repeats* (Marcos Sotomayor)
  - *The Mechanical Stability of the Titin Z1Z2-Telethonin Complex as Revealed by Steered Molecular Dynamics* (Eric H. Lee, Mu Gao)
  - *A Computational Study of ATP Hydrolysis and Force Generation in PcrA Helicase* (Markus Dittrich)
  - *Simulated Forced Closure of an F1-Atpase Subunit Leads to Rotation of The Central Stalk* (Barry Isralewitz)
  - *Structure-Based Model of a Stepping Motor on ssDNA: PcrA Helicase* (Jin Yu)
  - *Gas Migration Pathways Inside Proteins: Application to CpI Hydrogenase* (Jordi Cohen)
  - *Molecular Dynamics Simulation of a Voltage-Gated K+ Channel* (Fatemeh Khalili-Araghi)

  - *Silica Parameterization Based on Structural Surface Properties and Wetting Behavior for Silica Nanopore Simulation* (Eduardo Cruz-Chu, Aleksei Aksimentiev)
  - *Coupling of Proton Translocation and Protein Conformational Change in E. coli Lactose Permease* (Ying Yin)
  - *Oxygen Pathways in Myoglobin* (Anton Arkhipov, Jordi Cohen, Rosemary Braun)
  - *The Mechanism of Sugar Transport Across E. coli Lactose Permease* (Ying Yin)
  - *Molecular Dynamics Simulations of the Complete Satellite Tobacco Mosaic Virus* (Peter L. Freddolino, Anton S. Arkhipov)
  - *Exploring the Electro-Mechanical Properties of Single DNA Molecules with a Synthetic Nanopore* (Aleksei Aksimentiev)
  - *Conformational Geometries of Holliday Junction in Conformer Transition* (Jin Yu)

- December 2005, San Francisco, CA, University of California at San Francisco, A. Sali Laboratory, *Quantum Mechanical/Molecular Mechanical Simulations of Biomolecular Systems* (Markus Dittrich)

- November 2005, Seattle, WA, SC—05 International Conference for High Performance Computing, Networking, Storage and Analysis, *Exploring Biomolecular Machines with Supercomputers* (Jim Phillips, Jordi Cohen)

- November 2005, Urbana, IL, University of Illinois, Beckman Institute, TCBG Cluster Workshop

  - *Linux Clusters for High-Performance Computing - An Introduction* (Tim Skirvin)
  - *Linux Clusters - Details and Case Studies* (Jim Phillips, Tim Skirvin)

- November 2005, Urbana, IL, University of Illinois, National Center for Supercomputing Applications, Social Networks and Cyberinfrastructure Conference, *BioCoRE Overview* (Kirby Vandivort)

- October 2005, Urbana, IL, University of Illinois, UIUC CITES Computer Consultant Support Program Conference, *Linux Clustering for Scientific Computing* (Jim Phillips, Tim Skirvin)

- October 2005, Seattle, WA, Univ. of Washington, David Baker Laboratory, *Quantum Mechanical/Molecular Mechanical Simulations of Biomolecular Systems* (Markus Dittrich)

- October 2005, Tutzing, Germany, 87th International Bunsen Discussion Meeting on Mechanically Induced Chemistry, *Tertiary and Secondary Structure Elasticity of Repeat Proteins* (Marcos Sotomayor)
- October 2005, Urbana, IL, University of Illinois, The Beckman Institute, The Eighteenth Annual Cell and Molecular Biology/ Molecular Biophysics Training Grant Research Symposium, *Dynamics of the Lac RepressorDNA Complex Revealed by a Multiscale Simulation* (Elizabeth Villa)
- September 2005, Los Angeles, CA, University of California at Los Angeles, Institute for Pure and Applied Mathematics (IPAM), MA2005 Workshop I: Multiscale Modeling in Soft Matter and Bio-Physics, *Mechanical Interactions between Lac Repressor and DNA Loops* (Elizabeth Villa)
- September 2005, Urbana, IL, University of Illinois, Beckman Institute, TCBG Cluster Workshop
  - *Linux Clusters for High-Performance Computing: An Introduction* (Tim Skirvin)
  - *Linux Clusters: Details and Case Studies* (Jim Phillips, Tim Skirvin)
- September 2005, San Francisco, CA, University of California at San Diego, San Diego Supercomputer Center, Visualization of Large Biomolecular Complexes Workshop, *VMD: Algorithms and Methods for Large Scale Biomolecular Visualization* (John Stone)
- August 2005, Montpellier, France, IUPAB/EBSA International Biophysics Congress, *Poster: Tertiary and Secondary Structure Elasticity of Repeat Proteins* (Marcos Sotomayor)

Outreach

The outreach efforts of the Resource broaden beyond media coverage, scholarly articles, lectures, posters, and workshops to other activities such as brochure development and distribution and granting requests from outside authors and groups to use images produced by Resource members in publications and presentations. Outreach activities, and the success of the Resource web site as an outreach mechanism, are listed and detailed below:

- On-site demonstrations
- Others publish our images
- Brochure development and distribution

- License expansions
- Web site design
- Visits and links to the Resource website

Demonstrations

Visitors to the Resource (i.e., seminar speakers, visiting scientists, others) are commonly provided with a demonstration of 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. Or, demonstrations may cover Resource software, such as NAMD or BioCoRE. There were 18 such presentations by Resource members over the last funding period.

Image Requests

The Resource regulary responds to requests for permissions to use Resource images on other sites, in textbooks, papers, talks, and other media produced by others. A standard response letter, written in cooperation with university intellectual property representatives, grants non-exclusive permission to such requests, which protects Resource copyright while at the same time allowing for further image distribution. In the last funding period, 14 requests were processed, with resource images being used in media such as a nanotechnology and biomedicine textbook, a cell and molecular biology textbook, several academic websites, in-class and conference lectures, brochures, and a documentary being developed for the Discovery channel.

Brochures

Three brochure projects were undertaken by the Resource as a means of communicating information about our programs, research, and software. Each brochure is described below; and completed brochures can be found online at the Resource website[‖]:

- *Bringing Computing to Life* - completed during this funding cycle, the brochure describes the challenges and work culture of the developers who produce VMD, NAMD, and BioCoRE, as well as the system that supports their efforts.
- *Theoretical and Computational Biophysics Workshops: Training for the New Discipline* - the brochure describes the workshop program developed by the Resource, including how the workshops combine theory with hands-on experience, tutorial development and utilization, Resource member contributions, participants, with

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

appendices providing workshop details such as a sample agenda and evaluation highlights.

- *Recent Highlights* - every month, one or more highlights (a paragraph of text and associated images) about Resource research or events are presented on the main web page, providing thumbnail views into Resource activities. Currently, the Resource is updating a prior brochure compilation of these highlights, which covered highlights from January 2001 to March 2004; the revised version will include highlights from March 2004 to April 2006.

Licensing and Distribution

Resource software licenses, which already allow for broad distribution, are upon request expanded further to accommodate the needs of external groups. Such expansions are done in consultation and cooperation with the University of Illinois Office of Technology Management, who provide technical and legal expertise. For example, the Argonne National Laboratory requested a change in wording about license transferability, such that any potential successor operators for that organization would be able to continue using NAMD. Resource members, a representative of the Office of Technology Management, and contacts at Argonne National Laboratory were able to resolve the request within a business week.

Web Site Design

Those visiting the Resource web site in recent months will have encountered a new web site design, created in cooperation with a UIUC-based web site design team, that provides easier access and navigability to desired content and information. The main page of the Resource web site**, for example boasts clearer identification of desired content, such as software web sites and download areas, revised and updated research categories, and more information about recent or pending changes at the Resource such as recent publications, pending workshops and seminars, and other announcements.

Website Popularity

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 672,361 unique visitors to the Resource web site, an average of 56,030 per month, and 1,282 gigabytes of data transferred (i.e., from downloaded pages, images, and files within the site), or an average of 107 gigabytes per month, over the April 2005 - March 2006 time period. The most visited sections of the web site are shown in Table 3.

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

	Total Visitors	Visitors per Month
VMD	197,855	16,487
NAMD	84,491	7,040
BioCoRE	24,458	2,038
Other Research	120,673	10,056
Galleries	63,033	5,252
Papers	33,744	2,812
Seminars	4,707	392

Table 2: Numbers from April 2005 – March 2006

A recent Google search (April, 2006) yielded the following statistics regarding external sites that link to the main Resource web page: 1,003 pages link to the main page, registered under 65 different domains. Example education, scientific resources, and programming or computing-related sites with links to the Resource web site are provided below:

Education:

- *Cornell University: Biology Department, Molecular Biology and Genetics Department, Physics Department, Computer Science Department, and Medical Library*
  bio.cornell.edu, www.mbg.cornell.edu, www.physics.cornell.edu, www.cs.cornell.edu, library.med.cornell.edu
- *University of California at San Diego: Physics Department, Chemistry Department, Keck Lab for Integrated Biology, and McCammon Biophysics Group*
  physics.ucsd.edu, www-chem.ucsd.edu, keck2.ucsd.edu, mccammon.ucsd.edu
- *University of California at Berkeley: Electron Microscopy Group, Computer Science Department, Astronomy Department, and Molecular Graphics and Computation Facility*
  cryoem.berkeley.edu, www.cs.berkeley.edu, astron.berkeley.edu, glab.cchem.berkeley.edu
- *Purdue University: Computer Science Department, Chemistry Department, Nanotechnology Simulation Hub, and Instructional Computing Services*
  www.cs.purdue.edu, www.chem.purdue.edu, www.nanohub.purdue.edu, expert.ics.purdue.edu
- *Duke University: Biology Department, Electrical Engineering Department, and Single Molecule Force Spectroscopy Lab*
  www.biology.duke.edu, www.ee.duke.edu, smfs.pratt.duke.edu

- *New York University: Math Department, Computer Science Department, and Computational Biology/Chemistry/Biomathematics Department*
  www.math.nyu.edu, www.cs.nyu.edu, monod.biomath.nyu.edu
- *Harvard University: Instructional Computing Group, Wagner NMR Structural Research Group, and Molecular Biology Core Facilities*
  www.courses.fas.harvard.edu, gwagner.med.harvard.edu
- *Massachusetts Institute of Technology: Open Courseware Project and Computer Graphics Group*
  ocw.mit.edu, graphics.lcs.mit.edu
- *Scripps Research Institute: Amber Molecular Dynamics and Metalloprotein Database*
  amber.scripps.edu, metallo.scripps.edu
- *Stanford University: Medical Informatics and Computer Science*
  www.smi.stanford.edu, www-cs-students.stanford.edu
- *Yale University: Center for Structural Biology and Database of Macromolecular Movements*
  www.csb.yale.edu, molmovdb.mbb.yale.edu
- *University of Pennsylvania: Engineering Department and Center for Molecular Modeling*
  www.seas.upenn.edu, www.cmm.upenn.edu

Scientific Resources:

- *Protein Data Bank*
  www.rcsb.org
- *Biophysical Journal*
  www.biophysj.org
- *Science Magazine*
  www.sciencemag.org
- *Nature Magazine*
  www.nature.com
- *Howard Hughes Medical Institute*
  www.hhmi.org
- *Wikipedia Online Encyclopedia*
  en.wikipedia.org

- *Slashdot*
  science.slashdot.org
- *Chemistry at Harvard Molecular Mechanics*
  www.charmm.org
- *GROMACS Molecular Dynamics*
  www.gromacs.org
- *CPMD consortium*
  www.cpmd.org
- *Prentice Hall*
  wps.prenhall.com
- *The Foresight Institute Nanotechnology*
  www.foresight.org
- *PhysicsWeb, Physics News and Resources*
  physicsweb.org
- *Bioinformatics Open-Access*
  bioinformatics.org
- *Physical Review Focus*
  focus.aps.org
- *Computational Chemistry List*
  www.ccl.net
- *American Scientist Magazine*
  www.americanscientist.org
- *Microbiology Information Portal*
  www.microbes.info
- *Free Science*
  freescience.info
- *Protein Society*
  www.proteinsociety.org
- *Ernest Orlando Lawrence Berkeley National Laboratory*
  www-vis.lbl.gov

- *Center for Molecular Modeling at National Institutes of Health*
  cmm.info.nih.gov

Programming/Computing Related:

- *Apple Computers*
  www.apple.com
- *OpenGL Programming*
  www.opengl.org
- *Silicon Graphics*
  www.sgi.com
- *Linux Online*
  www.linux.org
- *Java 3D Community*
  www.j3d.org
- *GNU Operating System*
  www.gnu.org
- *PovRay Objects Collection*
  objects.povworld.org
- *VersionTracker Software Downloads*
  www.versiontracker.com
- *FreshMeat Software Downloads*
  www.freshmeat.net
- *Codebeach Developer's Guide*
  www.codebeach.com
- *Sourceforge Open Source*
  sourceforge.net
- *StereoGraphics*
  www.stereographics.com
- *FreeBSD Online*
  www.freebsd.org

- *Beowulf Clusters*
  www.beowulf.org
- *Freeware Web*
  www.freewareweb.com

Bibliography

[1] 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.

[2] J. Cohen, A. Arkhipov, R. Braun, and K. Schulten. Imaging the migration pathways for O_2, CO, NO, and Xe inside myoglobin. *Biophys. J.*, 2006. In press.

[3] L. A. Strawn, T. Shen, N. Shulga, D. S. Goldfarb, and S. R. Wente. Minimal nuclear pore complexes define FG repeat domains essential for transport. *Nature Cell Biol.*, 6:197–206, 2004.

[4] A. S. Blum, C. M. Soto, C. D. Wilson, J. D. Cole, M. Kim, B. Gnade, A. Chatterji, W. F. Ochoa, T. Lin, J. E. Johnson, and B. R. Ratna. Cowpea mosaic virus as a scaffold for 3-d patterning of gold nanoparticles. *Nano Letters*, 4:867–870, 2004.

[5] A. L. Perryman, J.-H. Lin, and J. A. McCammon. HIV-1 protease molecular dynamics of a wild-type and of the v82f/i84v mutant: Possible contributions to drug resistance and a potential new target site for drugs. *Protein Science*, 13:1108–1123, 2004.

[6] B. Speelman, B. R. Brooks, and C. B. Post. Molecular dynamics simulations of human rhinovirus and an antiviral compound. *Biophys. J.*, 80:121–129, 2001.

[7] D. Zhang, R. Konecny, N. A. Baker, and J. A. McCammon. Electrostatic interaction between RNA and protein capsid in cowpea chlorotic mottle virus simulated by a coarse-grain RNA model and a monte carlo approach. *Biopolymers*, 75:325–337, 2004.

[8] D. K. Phelps, B. Speelman, and C. B. Post. Theoretical studies of viral capsid proteins. *Current Opinion in Structural Biology*, 10:170–173, 2000.

[9] F. Tama, O. Miyashita, and C. L. Brooks III. Flexible multi-scale fitting of atomic structures into low-resolution electron density maps with elastic network normal mode analysis. *J. Mol. Biol.*, 337:985–999, 2004.

[10] F. Tama and C. L. Brooks III. Diversity and identity of mechanical properties of icosahedral viral capsids studied with elastic network normal mode analysis. *J. Mol. Biol.*, 345:299–314, 2005.

[11] A. J. Rader, D. H. Vlad, and I. Bahar. Maturation dynamics of bacteriophage hk97 capsid. *Structure*, 13:413–421, 2005.

[12] V. S. Reddy, H. A. Giesing, R. T. Morton, A. Kumar, C. B. Post, C. L. B. III, and J. E. Johnson. Energetics of quasiequivalence: Computational analysis of protein-protein interactions in icosahedral viruses. *Biophys. J.*, 74:546–558, 1998.

[13] 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.

[14] S. B. Larson and A. McPherson. Satellite tobacco mosaic virus rna: structure and implications for assembly. *Curr. Op. Struct. Biol.*, 11:59–65, 2001.

[15] Y. G. Kuznetsov, S. Daijogo, J. Zhou, B. L. Semler, and A. McPherson. Atomic force microscopy analysis of icosahedral virus RNA. *J. Mol. Biol.*, 347:41–52, 2005.

[16] M. P. Rout, J. D. Aitchison, A. Suprapto, K. Hjertaas, Y. Zhao, and B. T. Chait. The yeast nuclear pore complex: Composition, architecture, and transport mechanism. *J. Cell Biol.*, 148:635–651, 2000.

[17] R. Y. H. Lim and U. Aebi. In silico access to the nuclear pore complex. *Structure*, 13:1743–1744, 2005.

[18] R. Bayliss, T. Littlewood, L. A. Strawn, S. R. Wente, and M. Stewart. GLFG and FxFG nucleoporins bind to overlapping sites on importin-β. *J. Biol. Chem.*, 277:50597–50606, 2002.

[19] R. Bayliss, T. Littlewood, and M. Stewart. Structural basis for the interaction between FxFG nucleoporin repeats and importin-β in nuclear trafficking. *Cell*, 102:99–108, 2000.

[20] J. Bednenko, G. Cingolani, and L. Gerace. Importin-β contains a COOH-terminal nucleoporin binding region important for nuclear transport. *J. Cell Biol.*, 162:391–401, 2003.

[21] S. M. Liu and M. Stewart. Structural basis for the high-affinity binding of nucleoporin Nup1p to the *Saccharomyces cerevisiae* importin-β homologue, Kap95p. *J. Mol. Biol.*, 349:515–525, 2005.

[22] R. Bayliss, S. W. Leung, R. P. Baker, B. B. Quimby, A. H. Corbett, and M. Stewart. Structural basis for the interaction between NTF2 and nucleoporin FxFG repeats. *EMBO J.*, 21:2843–2853, 2002.

[23] C. Chaillan-Huntington, C. V. Braslavsky, J. Kuhlmann, and M. Stewart. Dissecting the interactions between NTF2, RanGDP, and the nucleoporin XFXFG repeats. *J. Biol. Chem.*, 275:5874–5879, 2000.

[24] R. Bayliss, K. Ribbeck, D. Akin, H. M. Kent, C. M. Feldherr, D. Görlich, and M. Stewart. Interaction between NTF2 and xFxFG-containing nucleoporins is required to mediate nuclear import of RanGDP. *J. Mol. Biol.*, 293:579–593, 1999.

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

[26] M. Ptashne. *A Genetic Switch*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2004.

[27] F. Jacob and J. Monod. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.*, 3:318–356, 1961.

[28] B. Müller-Hill. *The lac operon*. Walter de Gruyter, New York, 1996.

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

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

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

[32] 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.

[33] 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.

[34] 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.

[35] C. E. Bell and M. Lewis. A closer view of the conformation of the Lac repressor bound to operator. *Nature Struct. Biol.*, 7:209–214, 2000.

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

[37] 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.

[38] 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.

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

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

[41] E. Villa, A. Balaeff, L. Mahadevan, and K. Schulten. Multi-scale method for simulating protein-DNA complexes. *Multiscale Model. Simul.*, 2:527–553, 2004.

[42] 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.

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

[44] 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.

[45] S. Törnroth-Horsefield, Y. Wang, K. Hedfalk, U. Johanson, M. Karlsson, E. Tajkhorshid, R. Neutze, and P. Kjellbom. Structural mechanism of plant aquaporin gating. *Nature*, 439:688–694, 2006.

[46] A. J. Yool and A. M. Weinstein. New roles for old holes: Ion channel function in aquaporin-1. *News Physio. Sci.*, 17:68–72, 2002.

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

[48] Y. Wang, K. Schulten, and E. Tajkhorshid. What makes an aquaporin a glycerol channel: A comparative study of AqpZ and GlpF. *Structure*, 13:1107–1118, 2005.

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

[50] D. F. Savage, P. F. Egea, Y. Robles-Colmenares, J. D. O'Connell III, and R. M. Stroud. Architecture and selectivity in aquaporins: 2.5 Å X-ray structure of aquaporin Z. *PLoS Biology*, 1:334–340, 2003.

[51] M. Ø. Jensen, S. Park, E. Tajkhorshid, and K. Schulten. Energetics of glycerol conduction through aquaglyceroporin GlpF. *Proc. Natl. Acad. Sci. USA*, 99:6731–6736, 2002.

[52] 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.

[53] 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.

[54] R. J. Collier and J. A. T. Young. Anthrax toxin. *Annu. Rev. Cell Dev. Biol.*, 19:45–70, 2003.

[55] M. Moayeri and S. H. Leppla. The roles of anthrax toxin in pathogenesis. *Curr. Opin. Microbiol.*, 7:19–24, 2004.

[56] E. Santelli, L. A. Bankston, S. H. Leppla, and R. C. Liddington. Crystal structure of a complex between anthrax toxin and its host cell receptor. *Nature*, 430:905–908, 2004.

[57] M. Gao and K. Schulten. Onset of anthrax toxin pore formation. *Biophys. J.*, 90:3267–3279, 2006.

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

[59] W. R. Briggs and J. M. Christie. Phototropins 1 and 2: versatile plant blue-light receptors. *Trends Plant Sci*, 7:204–10, 2002.

[60] S. Crosson and K. Moffat. Structure of a flavin-binding plant photoreceptor domain: insights into light-mediated signal transduction. *Proc Natl Acad Sci U S A*, 98:2995–3000, 2001.

[61] S. Crosson and K. Moffat. Photoexcited structure of a plant photoreceptor domain reveals a light-driven molecular switch. *Plant Cell*, 14:1067–75, 2002.

[62] R. Fedorov, I. Schlichting, E. Hartmann, T. Domratcheva, M. Fuhrmann, and P. Hegemann. Crystal structures and molecular mechanism of a light-induced signaling switch: The Phot-LOV1 domain from Chlamydomonas reinhardtii. *Biophys J*, 84:2474–82, 2003.

[63] M. Dittrich, P. L. Freddolino, and K. Schulten. When light falls in LOV: A quantum mechanical/molecular mechanical study of photoexcitation in Phot-LOV1 of *Chlamydomonas reinhardtii*. *J. Phys. Chem. B*, 109:13006–13013, 2005.

[64] T. Swartz, S. Corchnoy, J. Christie, J. Lewis, I. Szundi, W. Briggs, and R. Bogomolni. The photocycle of a flavin-binding domain of the blue light photoreceptor phototropin. *J Biol Chem*, 276:36493–500, 2001.

[65] E. Schleicher, R. M. Kowalczyk, C. W. M. Kay, P. Hegemann, A. Bacher, M. Fischer, R. Bittl, G. Richter, and S. Weber. On the reaction mechanism of adduct formation in LOV domains of the plant blue-light receptor phototropin. *J Am Chem Soc*, 126:11067–76, 2004.

[66] P. L. Freddolino, M. Dittrich, and K. Schulten. Dynamic switching mechanisms in lov1 and lov2 domains of plant phototropins. *Biophys. J.*, 2006. Submitted.

[67] S. M. Harper, L. C. Neil, and K. H. Gardner. Structural basis of a phototropin light switch. *Science*, 301:1541–4, 2003.

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

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

[70] 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.

[71] 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.

[72] 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.

[73] 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.

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

[75] W. Grajkowskim, A. Kubalski, and P. Koprowski. Surface changes of the mechanosensitive channel MscS upon its activation, inactivaction and closing. *Biophys. J.*, 88:3050–3059, 2005.

[76] 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.

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

[78] T. A. v. Straaten, G. Kathawala, A. Trellakis, R. S. Eisenberg, and U. Ravaioli. BioMOCA– a Boltzmann transport Monte Carlo model for ion channel simulation. *Mol. Sim.*, 31:151–171, 2005.

[79] 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.

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

[81] A. A. Ajees, G. 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.

[82] 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.

[83] 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.

[84] 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.

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

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

[87] 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.

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

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

[90] J. B. Heng, A. Aksimentiev, C. Ho, P. Marks, Y. V. Grinkova, S. Sligar, K. Schulten, and G. Timp. Stretching DNA using an electric field in a synthetic nanopore. *Nano Lett.*, 5:1883–1888, 2005.

[91] J. B. Heng, A. Aksimentiev, C. Ho, P. Marks, Y. V. Grinkova, S. Sligar, K. Schulten, and G. Timp. The electromechanics of DNA in a synthetic nanopore. *Biophys. J.*, 90:1098–1106, 2006.

[92] J. Mathé, A. Aksimentiev, D. R. Nelson, K. Schulten, and A. Meller. Orientation discrimination of single stranded DNA inside the α-hemolysin membrane channel. *Proc. Natl. Acad. Sci. USA*, 102:12377–12382, 2005.

[93] A. Aksimentiev, J. B. Heng, G. Timp, and K. Schulten. Microscopic kinetics of DNA translocation through synthetic nanopores. *Biophys. J.*, 87:2086–2097, 2004.

[94] J. B. Heng, C. Ho, T. Kim, R. Timp, A. Aksimentiev, Y. V. Grinkova, S. Sligar, K. Schulten, and G. Timp. Sizing DNA using a nanometer-diameter pore. *Biophys. J.*, 87:2905–2911, 2004.

[95] J. B. Heng, A. Aksimentiev, C. Ho, V. Dimitrov, T. Sorsch, J. Miner, W. Mansfield, K. Schulten, and G. Timp. Beyond the gene chip. *Bell Labs Tech. J.*, 10:5–22, 2005.

[96] 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.

[97] S. Iijima. Helical microtubules of graphitic carbon. *Nature*, 354:56–58, 1991.

[98] N. W. S. Kam, T. C. Jessop, P. A. Wender, and H. Dai. Nanotube molecular transporters: Internalization of carbon nanotube-protein conjugates into mammalian cells. *J. Am. Chem. Soc.*, 126:6850–1, 2004.

[99] P. Cherukuri, S. M. Bachilo, S. H. Litovsky, and R. B. Weisman. Near-infrared fluorescence microscopy of single-walled carbon nanotubes in phagocytic cells. *J. Am. Chem. Soc.*, 126:15638–9, 2004.

[100] N. W. S. Kam, M. O'Connell, J. A. Wisdom, and H. Dai. Carbon nanotubes as multifunctional biological transporters and near-infrared agents for selective cancer cell destruction. *Proc. Natl. Acad. Sci. USA*, 102:11600–5, 2005.

[101] P. W. Barone, S. Baik, D. A. Heller, and M. S. Strano. Near-infrared optical sensors based on single-walled carbon nanotubes. *Nat. Mater.*, 4:86–92, 2005.

[102] C. Staii and A. T. Johnson, Jr. DNA-decorated carbon nanotubes for chemical sensing. *Nano Lett.*, 5:1774–8, 2005.

[103] G. Hummer, J. C. Rasaiah, and J. P. Noworyta. Water conduction through the hydrophobic channel of a carbon nanotube. *Nature*, 414:188–190, 2001.

[104] D. J. Mann and M. D. Halls. Water alignment and proton conduction inside carbon nanotubes. *Phys. Rev. Lett.*, 90:195503, 2003.

[105] C. Wei and D. Srivastava. Theory of transport of long polymer molecules through carbon nanotube channel. *Phys. Rev. Lett.*, 91:235901, 2003.

[106] I.-C. Yeh and G. Hummer. Nucleic acid transport through carbon nanotube membranes. *Proc. Natl. Acad. Sci. USA*, 101:12177–12182, 2004.

[107] 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.

[108] 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.

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

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

[111] 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.

[112] 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.

[113] J. Hénin and C. Chipot. Overcoming free energy barriers using unconstrained molecular dynamics simulations. *J. Chem. Phys.*, 121:2904–2914, 2004.

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

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

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

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

[118] B. Fahrenkrog and U. Aebi. The nuclear pore complex: Nucleocytoplasmic transport and beyond. *Nature Reviews Molecular Cell Biology*, 4:757–766, 2003.

[119] S. A. Adam. The nuclear pore complex. *Genome Biology*, 2:reviews0007.1–0007.6, 2001.

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

[121] D. Görlich and U. Kutay. Transport between the cell nucleus and the cytoplasm. *Annual Reviews Cell Development Biology*, 15:607–660, 1999.

[122] A. Liljas. *Structural Aspects of Protein Synthesis*. World Scientific, Singapore, 2004.

[123] J. Frank. Cryo-electron microscopy as an investigative tool: The ribosome as an example. *BioEssays*, 23:725–732, 2001.

[124] 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.

[125] S. B. Larson, J. Day, A. Greenwood, and A. McPherson. Refined structure of satellite tobacco mosaic virus at 1.8 å resolution. *J. Mol. Biol.*, 277:37–59, 1998.

[126] J. Day, Y. G. Kuznetsov, S. B. Larson, A. Greenwood, and A. McPherson. Biophysical studies on the rna cores of satellite tobacco mosaic virus. *Biophys. J.*, 80:2364–2371, 2001.

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

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

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

[130] 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.

[131] H. R. Kaback. Structure and mechanism of the lactose permease. *Comptes Rendus Biologies*, 328:557–567, 2005.

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

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

[134] J. Abramson, S. Iwata, and H. R. Kaback. Lactose permease as a paradigm for membrane transport proteins (Review). *Molecular Membrane Biology*, 21:227–236, 2004.

[135] O. Mirza, L. Guan, G. Verner, S. Iwata, and H. R. Kaback. Structural evidence for induced fit and a mechanism for sugar/h+ symport in lacy. *The EMBO Journal*, 25:1177–1183, 2006.

[136] Y. Yin, M. Ø. Jensen, E. Tajkhorshid, and K. Schulten. Sugar binding and protein conformational changes in lactose permease. *Biophys. J.*, 2006. Submitted.

[137] G. Blobel. Protein targeting. *Biosci. Rep.*, 20:303–344, 2000.

[138] K. E. S. Matlack, W. Mothes, and T. A. Rapoport. Protein translocation: tunnel vision. *Cell*, 92:381–390, 1998.

[139] A. K. J. Veenendaal, C. van der Does, and A. J. M. Driessen. The protein-conducting channel SecYEG. *Biochim. Biophys. Acta*, 1694:81–95, 2004.

[140] H. Mori and K. Ito. The Sec protein-translocation pathway. *Trends Microbiol.*, 9:494–500, 2001.

[141] T. A. Rapoport, B. Jungnickel, and U. Kutay. Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes. *Ann. Rev. Biochem*, 65:271–303, 1996.

[142] I. B. Holland. Translocation of bacterial proteins - an overview. *Biochim. Biophys. Acta*, 1694:5–16, 2004.

[143] R. S. Hegde and N. S. Rane. Prion protein trafficking and the development of neurodegeneration. *Trends in Neurosciences*, 26:337–339, 2003.

[144] 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.

[145] S. J. Kim and R. S. Hegde. Cotranslational partitioning of nascent prion protein into multiple populations at the translocation channel. *Mol. Biol. Cell*, 13:3775–3786, 2002.

[146] A. Kohara, Y. Yamamoto, and M. Kikuchi. Processing and secretion of human growth hormone with an artificial signal sequence. *Biosci. Biotechnol. Biochem.*, 58:779–781, 1994.

[147] M. Hughes, A. Vassilakos, D. W. Andrews, G. Hortelano, J. W. Belmont, and P. L. Chang. Delivery of a secretable adenosine deaminase through microcapsules–a novel approach to somatic gene therapy. *Hum. Gene Ther.*, 5:1445–1455, 1994.

[148] 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.

[149] B. van den Berg, W. M. Clemons, Jr., I. Collinson, Y. Modis, E. Hartmann, S. C. Harrison, and T. A. Rapoport. X-ray structure of a protein-conducting channel. *Nature*, 427:36–44, 2004.

[150] J. Gumbart and K. Schulten. Molecular dynamics studies of the archaeal translocon. *Biophys. J.*, 90:2356–2367, 2006.

[151] M. Bostina, B. Mohsin, W. Kühlbrandt, and I. Collinson. Atomic model of the *E. coli* membrane-bound protein translocation complex SecYEG. *J. Mol. Biol.*, 352:1035–1043, 2005.

[152] K. Mitra, C. Schaffitzel, T. Shaikh, F. Tama, S. Jenni, C. L. Brooks, N. Ban, and J. Frank. Structure of the *E. coli* protein-conducting channel bound to a translating ribosome. *Nature*, 438:318–324, 2005.

[153] S. Bhakdi and J. Tranum-Jensen. Alpha-toxin of *staphylococcus aureus*. *Microbiol. Rev.*, 55:733–751, 1991.

[154] E. Gouaux. α-hemolysin from *staphylococcus aureus* : An archetype of β-barrel, channel-forming toxins. *J. Struct. Biol.*, 121:110–122, 1998.

[155] S. Howorka, S. Cheley, and H. Bayley. Sequence-specific detection of individual dna strands using engineered nanopores. *Nature Biotech.*, 19:636–639, 2001.

[156] D. Deamer and D. Branton. Characterization of nucleic acids by nanopore analysis. *Acc. Chem. Res.*, 35:817–825, 2002.

[157] W. Vercoutere and M. Akeson. Biosensors for DNA sequence detection. *Curr. Op. Chem. Biol.*, 6:816–822, 2002.

[158] N. Ashkenasy, J. Sánchez-Quesada, H. Bayley, and M. R. Ghadiri. Recognizing a single base in an individual DNA strand: A step toward DNA sequencing in nanopores. *Angew. Chem. Int. Ed. Engl.*, 44:1401–1404, 2005.

[159] A. Aksimentiev and K. Schulten. Imaging alpha-hemolysin with molecular dynamics: Ionic conductance, osmotic permeability and the electrostatic potential map. *Biophys. J.*, 88:3745–3761, 2005.

[160] E. Tajkhorshid, J. Cohen, A. Aksimentiev, M. Sotomayor, and K. Schulten. Towards understanding membrane channels. In B. Martinac and A. Kubalski, editors, *Bacterial ion channels and their eukaryotic homologues*, pages 153–190. ASM Press, Washington, DC, 2005.

[161] J. Gumbart, Y. Wang, A. Aksimentiev, E. Tajkhorshid, and K. Schulten. Molecular dynamics simulations of proteins in lipid bilayers. *Curr. Op. Struct. Biol.*, 15:423–431, 2005.

[162] P. J. Zou, N. Pinotsis, M. Marino, S. Lange, A. Popov, I. Mavridis, M. Gautel, O. M. Mayans, and M. Wilmanns. Palindromic assembly of the giant muscle protein in the sarcomeric Z-disk. *Nature*, 439:229–233, 2006.

[163] L. Tjernberg, D. Callaway, A. Tjernberg, S. Hahne, C. Lilliehook, L. Terenius, J. Thyberg, and C. Nordstedt. A molecular model of Alzheimer amyloid beta-peptide fibril formation. *J. Biol. Chem.*, 274:12619–12625, 1999.

[164] D. Walsh, D. Hartley, M. Condron, M. Selkoe, and D. Teplow. In vitro studies of amyloid beta-protein fibril assembly and toxicity provide clues to the aetiology of flemish variant (Ala692 → Gly) Alzheimer's disease. *Biochem. J.*, 355:869–877, 2001.

[165] J. Karanicolas and C. Brooks. The structural basis for biphasic kinetics in the folding of the WW domain from a formin-binding protein: lessons for protein design? *Proc. Natl. Acad. Sci. USA*, 100:3555–3556, 2003.

[166] J. Wang, S. Gülich, C. Bradford, M. Ramirez-Avarado, and L. Regan. A twisted four-sheeted model for an amyloid fibril. *Structure*, 13:1279–1288, 2005.

[167] E. H. Lee, M. Gao, N. Pinotsis, M. Wilmanns, and K. Schulten. Mechanical strength of the titin Z1Z2/telethonin complex. *Structure*, 14:497–509, 2006.

[168] M. Gao, H. Lu, and K. Schulten. Unfolding of titin domains studied by molecular dynamics simulations. *J. Muscle Res. Cell Mot.*, 23:513–521, 2002.

[169] B. Hille. *Ionic channels of excitable membranes*. Sinauer Associates, Sunderland, MA, 3rd edition, 2001.

[170] 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.

[171] Y. Jiang, A. Lee, J. Chen, M. Cadene, B. T. Chait, and R. MacKinnon. X-ray structure of a voltage-dependent K^+ channel. *Nature*, 423:33–41, 2003.

[172] Y. Zhou and R. MacKinnon. The occupancy of ions in the K^+ selectivity filter: Charge balance and coupling of ion binding to a protein conformational change underlie high conduction rates. *J. Mol. Biol.*, 333:965–975, 2003.

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

[174] J. Åqvist and V. Luzhkov. Ion permeation mechanism of the potassium channel. *Nature*, 404:881–884, 2000.

[175] S. Velankar, P. Soultanas, M. S. Dillingham, H.S.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.

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

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

[178] 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.

[179] S. McGinnis and T. L. Madden. BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucl. Acids Res.*, 32:W20–W25, 2004.

[180] R. B. Russel and G. Barton. Multiple protein sequence alignment from tertiary struture comparison: assignment of global and resiude confidence levels. *Proteins: Struct., Func., Gen.*, 14:309–323, 1992.

[181] R. Chenna, H. Sugawara, T. Koike, R. Lopez, T. J. Gibson, D. G. Higgins, and J. D. Thompson. Multiple sequence alignment with the Clustal series of programs. *Nucl. Acids Res.*, 31:3497–3500, 2003.

[182] A. Sethi, P. O'Donoghue, and Z. Luthey-Schulten. Evolutionary profiles from the QR factorization of multiple sequence alignments. *Proc. Natl. Acad. Sci. USA*, 102:4045–4050, 2005.

[183] P. O'Donoghue, A. Sethi, C. R. Woese, and Z. Luthey-Schulten. Evolutionary profiles derived from the QR factorization. *J. Mol. Biol.*, 346:875–894, 2005.

[184] P. O'Donoghue, , and Z. Luthey-Schulten. On the evolution of structure in the aminoacyl-tRNA sythetases. 67:550–573, 2003.

[185] R. R. Sokal and C. D. Michener. A statistical method for evaluating systematic relationships. *Univ. Kansas Sci. Bull.*, 28:1409–1438, 1958.

[186] M. O. Dayhoff, R. M. Schwartz, and B. C. Orcutt. A model of evolutionary change in protein. In M. O. Dayhoff, editor, *Atlas of Protein Sequence and Structure*, volume 5, pages 345–352. National Biomedical Research Foundation, Washington, DC, 1978.