<< Back to previous view

[ALLOC-16] Understanding the structure, dynamics, and interactions of biomolecules through simulation

Created: Mar/06/2009 Updated: Mar/06/2009

Status:	Application Submitted
Project:	Allocation Requests
Component/s:	None
Fix Version/s:	None
Security Level:	Principle Investigator and CHPC Staff

Type:	Allocation Request	**Priority:**	standard helpdesk call (5 days)
Reporter:	Thomas Cheatham III	**Assignee:**	Julia D. Harrison
Resolution:	Submitted	**Votes:**	0

File Attachments:	jp8001614.pdf jp801245h.pdf published-full.pdf
Participants of an issue:	Julia D. Harrison and Thomas Cheatham III
Project URL (optional):	http://www.ambermd.org
Department:	Medicinal Chemistry / Pharmaceutics and Pharmaceutical Chemistry
College:	Pharmacy
Phone:	x7-9652
Allocation Sharing Among Users:	All Authorized Users Have Access to Full Award
Arches Cluster:	
Arches Cluster	
Spring 2009 :	50,000
Summer 2009:	50,000
Fall 2009:	50,000
Winter 2010:	50,000
Updraft Cluster:	
Updraft Cluster	
Spring 2009:	250,000
Summer 2009:	250,000
Fall 2009:	250,000
Winter 2010:	250,000
Proposal Submission Instructions:	

Please (in brief) address the following questions/topics regarding your proposal:

- General description of the project.
- Significance of the project and expected impact in the field.
- Describe the numerical techniques or third party software that will be used in the project.
- Describe the characteristics of your computational results and the methodology for analyzing and presenting the results. If applicable, describe the visualization resources necessary for presenting your results.
- Describe the parallelism of your code - particularly if requesting allocation on the updraft system. Include what you know about scaling and optimal job sizes of the codes you expect to run.
- Describe the computer resources available to you in your research group, department, or college, list allocations you already have elsewhere (SDSC, NCSA, Teragrid etc.). Explain the distinction between your proposed work and work already covered by your existing allocations and explain why CHPC resources are sought.
- Resources required. Please indicate CPU, memory, disk space, etc. and explain how you arrived at your estimated figures.
- Indicate sources of funding for the project, including off campus computer time. Describe efforts, if any, to obtain further support.
- If this project has been previously supported by CHPC, please describe the results obtained, including: student thesis, conference

	papers, and journal papers (please include preprints or reprints for CHPC files).
General Description of Project:	Computational resources are requested to augment the large Cheatham allocation at the NSF centers of 10M hours (MCA01S027) for a series of collaborative projects on campus and for analysis of the results. This past year the Cheatham lab has used the local time for extensive QM calculations for parameter development and to support the work of five students and a postdoc. Cheatham himself slowed down in the last two quarters but will be ramping back up soon. A general description is to use MD simulation, docking, and force field development to better understand biomolecular structure. Key projects underway are: (1) defining failures in nucleic acid force fields and overcoming them [Henriksen, Cang, Cheatham, Joung, Hud (Georgia Tech), Santa Lucia (Wayne State), Davis (Med Chem)], (2) RNA-based therapeutics [Cheatham, Davis (Med Chem)], (3) structural remodeling of DHFR upon ligand binding [Cheatham], (4) study of kinases with references to SAXS data [Cheatham, Blumenthal (Pharm/Tox], (5) prediction of substrate selectivity and chemistry in cytochrome P450s [Shahroks, Cheatham, Yost (Pharm/Tox), and (6) designed mutations of coiled-coil proteins [Pendley, Lim (PPC)].
Significance and Expected Impact:	The AMBER force fields are widely applied throughout the world in projects ranging from computer aided drug design to structure prediction. Over 600 sites licenses have been obtained for AMBER 10 and the force fields are applied in all of the major biomolecular simulation codes. If the force fields are wrong, a lot of people are wasting their time. We are often the first group to expose serious deficiencies in the nucleic acid force fields, some of which have been fixed and others which are still outstanding. Moreover, our simulations probe the dynamics and influence of ligand binding on protein and nucleic acid structure which is critical for future drug design. As one example of impact, the Davis group has solved the structure of a small-molecule RNA complex [outside the context of the ribozome!] that is serving as a platform for medicinal chemistry to improve binding from a lead compound to ideally a nM inhibitor. This involves computational docking, NMR refinement, molecular dynamics coupled with experimental assays (Davis), high throughput NMR estimates of binding (Davis) and synthesis (Conboy). A seed grant was funded to Davis and we are now seeking NIH challenge funds.
Numerical Techniques / 3rd Party Software:	AMBER, Gaussian, DOCK, Autodock, ... Cheatham is an AMBER developer and has access to the latest code modifications and improvements.
Computational Results and Analysis/Presentation Methodology:	The "results" are trajectories of energy and position as a function of time that can be micro-analyzed to provide insight into the underlying structure, dynamics and interactions of biomolecules in their native environments of salt, water, and potentially other molecules. The results can be visualized as "movies" but more often hypotheses about the data are tested through analysis of derived data from the simulation. This is largely facilitated by the "ptraj" program (GPL, free) that is written primarily by Cheatham. Required visualization resources are at the desktop level however we should be making use of the visualization wall at CHPC. "ptraj" is being radically modified for improved performance, remote data analysis, and parallel data analysis as a means to overcome the trajectory size bottleneck that looms even though we have > ~75 TB of Cheatham lab disk space locally at CHPC.
Parallelism of Code:	AMBER is highly parallelized and efficient. This is best described at http://ambermd.org. We also take advantage of CHARMM and NAMD for simulation; these are also well parallelized. Given our considerable experience on HPC computers at NSF and beyond, we well understand the limitations of the code. With the emerging multicore architectures this leads to problems due to bandwidth problems that limit full efficiency with high node counts. Therefore we tend to limit our runs to 32-128 processors (or more if we drop cores); however, many of our current runs involve ensembles of loosely coupled jobs. These can make use of many thousands of processors efficiently.
Existing / Outside Computer Resources:	The Cheatham lab has an active TRAC allocation from NSF (MCA01S027) for ~10,000,000 hours for Jan 1 - Dec 31, 2009 on a variety of machines within the TeraGrid (abe, kraken, ranger, pople). Cheatham is also supported by the TeraGrid ASTA program for computational/software support and is a member of the TeraGrid Science Advisory Board. Other resources are local workstations for the lab and the previously mentioned ~75 TB of disk space (we still need more!).
Resources Required:	As the committee knows, we can suck up all available resources. However, we do not want to do this and want to be team players. So, I have requested significant time on updraft (~250K) as this machine appears less crowded, and an additional 50K hours on sanddunearch. The later is particularly important for large scale QM calculations that cannot easily fit in the updraft 1-day queue limit and are too large for the older arches cluster. Please choose what "fits", however we really need time this quarter for sure on sanddunearch for long QM jobs on porphyrins for parameter development.
Sources of Funding:	Awarded: - National Institutes of Health, R01 GM081411 (2/01/08-1/31/13).

	Biomolecular simulation for the end-stage refinement of nucleic acid structure. PI: Cheatham. ~$950,000 direct (Priority score 165, 16.5%). - National Institutes of Health, R01 GM079383 (9/28/07-8/31/12). AMBER force field consortium: A coherent biomolecular simulation platform. PI: Yong Duan (UC Davis), Co-PIs: Cheatham, Carlos Simmerling (Stony Brook), Ray Luo (UC Irvine), Piotr Cieplak (Burnham Inst), Junmei Wang (Incisive, Inc.). ~$3,000,000; ~$50K direct to Cheatham/year (Priority score 125, 2.5%). - Office of Naval Research, N00014-05-1-0457 (4/01/05-9/30/09). A new research tool for the computer simulation of chemical dynamics in complex systems. PI: Greg Voth (Chemistry), Co-PIs: Cheatham (10%), Dave Case (Scripps), Bill Miller (Berkeley) and Bernie Schlegel (Wayne State). ~$3,200,000; 1 mo. Salary only/year to Cheatham. - NSF Cyberinfrastructure Partnership/TeraGrid (4/01/07-3/31/08). LRAC MCA01S027: Insight into biomolecular structure, dynamics, interactions and energetics from simulation. 7th renewal (8th year) Computer time award: ~9.5M node hours awarded in 2009. Pending: - NIH R01 (7/01/09-6/30/14) "Dynamics of large-scale domain motions in cAMP sensor proteins" PI: Blumenthal, Co-PIs: Cheatham, Herron. - NIH R01 (1/01/10-12/31/14) "P450 dehydrogenation mechanisms" PI: Gary Yost, Co-PI: Cheatham - NIH R01 (12/01/09-11/30/12) "Modeling of structure and energetics of co-crystals for pharmaceutical information" PI: Facelli, Co-PI: Cheatham. - NIH Challenge (9/01/09-8/31/11) [drug design targeting the hepatitis delta virus] PI: Davis, Co-PI: Cheatham, Conboy - NSF MRI for the capacity cluster (Co-PI: Cheatham).
Previous CHPC Work Results:	2008: (50) In Suk Joung, and T.E. Cheatham, III. "Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations" J. Phys. Chem. B 112, 9020-9041 (2008). (51) D. Svozil, J.E. Sponer, I. Marchan, A. Perez, T.E. Cheatham, III, J. Luque, M. Orozco, and J. Sponer. "Geometrical and electronic structure variability of the sugar-phosphate backbone in nucleic acids." J. Phys. Chem. B 112, 8188-8197 (2008). (52) S. S. Pendley, Y. B. Yu, and T.E. Cheatham, III. "Molecular dynamics guided study of salt bridge length dependence in both fluorinated and non-fluorinated parallel dimeric coiled-coil proteins." Proteins 74, 612-629 (2009) [DOI: 10.1002/prot.22177]. (53) H. Wang, T.E. Cheatham, III, P.M. Gannett, and J. Lewis. "Differential electronic states observed during the A-B DNA duplex conformational transitions." Soft Matter 5, 685-690 (2009). (54) T. S. Han, M.-M. Zhang, A. Walewska, P. Gruszczynski, T.E. Cheatham, III, D. Yoshikami, B. M. Olivera, and G. Bulaj. "Structurally-minimized disulfide-deficient μ-conotoxin analogs as sodium channel blockers: Implications for designing conopeptide-based therapeutics." Chem. Phys. Chem. (2009) [in press] (55) In Suk Joung, Ö Persil Çetinkol, N. V. Hud, and T.E. Cheatham, III. "Molecular dynamics simulations and coupled nucleotide substitution experiments indicate the nature of A•A base pairing and a putative structure of the coralyne-induced homo-adenine duplex." J. Amer. Chem. Soc. (2008) [submitted]