2047-Pos Board B777
Simultaneous Computation of Dynamical and Equilibrium Information using a Weighted Ensemble of Trajectories
Ernesto Suarez, Steven Lettieri, Metthew C. Zwier, Sundar Raman Subramanian, Lillian T. Chong, Daniel M. Zuckerman.
Computational and Systems Biology, University of Pittsburgh, Pittsburgh, PA, USA.
Equilibrium formally can be represented as an ensemble of uncoupled systems undergoing unbiased dynamics in which detailed balance is maintained on average. Non-equilibrium processes are described by suitable subsets of the equilibrium ensemble. Here, we employ the "weighted ensemble" (WE) simulation protocol to generate equilibrium trajectory ensembles and extract non-equilibrium subsets for computing kinetic quantities. States do not need to be chosen in advance. The procedure formally allows estimation of kinetic rates between arbitrary states chosen after the simulation, along with their equilibrium populations. We also describe a related history-dependent matrix procedure for estimating equilibrium and non-equilibrium observables when phase space has been divided into arbitrary non-Markovian regions, whether in WE or ordinary simulation. The methods are successfully applied and validated on two molecular systems: explicitly solvated methane association and implicitly solvated Ala-4 peptide.

2048-Pos Board B778
Inferring Structural Ensembles from Noisy Experiments and Molecular Dynamics: Correcting Force Field Bias with Bayesian Energy Landscape Tilting
Kyle Beauchamp[1], Vijay Pande[2], Rhiju Das[2].
[1]Memorial Sloan-Kettering Cancer Center, New York, NY, USA, [2]Stanford University, Stanford, CA, USA.
Inferring biomolecular conformation from experiment is a fundamental goal of structural biology. Structure determination often requires the combination of modeling and experiment, but the vast majority of approaches model only a single conformation, provide limited uncertainty information, and inherit biases from assumed force fields when data are limited. Building on recent conceptual advances, we hypothesized that these biases and missing uncertainty estimates could be addressed through Bayesian Energy Landscape Tilting (BELT), a scheme that enables the systematic computation of fully Bayesian "hyperensembles" over conformational ensembles. As a test of BELT's ability to correct force field bias, we show that conformational ensembles of trialanine derived from five different force fields (ff96, ff99, ff99sbnmr-ildn, CHARMM27, and OPLS-AA) and chemical shift and scalar coupling measurements give convergent values of the peptide's α, β, and PPII conformational populations. Furthermore, the ensembles recover set-aside measurements not used in the fitting. BELT's principled combination of simulation and limited experimental data promises rigorous assessment of force field bias and sets a foundation for modeling ensembles and uncertainties in complex biomolecular systems.

2049-Pos Board B779
Optimization of Coarse-Grained Water-Ion Interaction Parameters for Biological Simulation
Joseph Fogarty[1], See-Wing Chiu[2], Eric Jakobsson[2], Sagar Pandit[1].
[1]Univeristy of South Florida, Tampa, FL, USA, [2]University of Illinois, Urbana, IL, USA.
In our previous work, we developed a coarse-grained (CG) (four water molecules per particle) model for polarizable water-water interactions that replicates the bulk properties of water as accurately as many commonly used atomically detailed water models, but with a tenfold reduction in computation time. A key to the accuracy was to use a more complex functional form for the van der Waals interactions, specifically a modified Morse potential in which the short-range repulsion is parameterized independently from the long-range attraction. A major reason for the efficiency increase was the ability to use a longer time step. The parameters of the model were optimized using automated optimization software (ParOpt: Parameter Optimizer) developed in our laboratory. The optimization procedure takes a set of parameters which forms a high dimensional space, a target function defined by the comparison of CG data with either atomistic molecular dynamics data or experimental values, and explores the space to yield the point with best fit. By implementing a simplex-based method we are able to produce optimal points within the high dimensional space without relying on mathematically tidy assumptions about the underlying target function. In this work, we present optimized potential parameters for sodium and chloride ions obtained using the same method as for determination of water-water parameters. Each charged ionic bead represents a partially solvated ion with four water molecules. Our long-term goal is to explore the extent to which it will be possible to apply this approach to more heterogeneous systems, maintaining accuracy and gaining efficiency.

2050-Pos Board B780
The Bio3D Package: New Interactive Tools for Structural Bioinformatics
Xin-Qiu Yao[1], Guido Scarabelli[1], Lars Skjærven[2], **Barry J. Grant**[1].
[1]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA, [2]Structural and Computational Biology Unit, European Molecular Biology Laboratory (EMBL), Heidelberg, Germany.
We present extensive updates to Bio3D, a package for the interactive analysis of biomolecular structure, sequence and molecular simulation data. Features include the ability to read and write biomolecular structure, sequence and dynamic trajectory data, query and search online sequence and structure databases, perform alignment, structural superposition, rigid core and dynamic domain identification, sequence and conformational clustering, distance and correlation matrix analysis, conservation analysis, normal mode analysis, principal component analysis, and many other common structural bioinformatics tasks. Bio3D also leverages the extensive graphical and statistical capabilities of the R environment (http://www.r-project.org) and thus provides a useful framework for the exploratory interactive analysis of biomolecular sequence and structure data. Recent notable additions to the package include consensus dynamical community analysis of network based coupled motions, enhanced coarse grained force fields and methods for normal mode analysis, and multicore support for many time intensive tasks. Here we describe these new capabilities with example applications. The previous version of Bio3D has been downloaded by over 13,700 researchers and cited over 100 times in the last six years. Merging these new methods represents an important advance that we hope will further stimulate biophysicists to use structural bioinformatics methods as an aid in solving their research problems. The Bio3D package is distributed with full source code and extensive documentation as a platform independent R package under a GPL2 license from http://thegrantlab.org/bio3d/.

2051-Pos Board B781
Learning About Transitions: Adaptive Controls for the Molecular Dynamics Database
Sarana Y. Nutanong[1], Yanif Ahmad[1], **Thomas B. Woolf**[2], I-Jeng Wang[3].
[1]Computer Science, Johns Hopkins University, Baltimore, MD, USA, [2]Physiology, Johns Hopkins School of Medicine, Baltimore, MD, USA, [3]Applied Physics Laboratory, Johns Hopkins University, Baltimore, MD, USA.
Improvements in sampling on the events leading to transitions can provide significant insights into what drives biomolecular change. We present additions to our Molecular Dynamics Database (MDDB). We start with the transitions seen in BPTI from DEShaw Research (Science, 2010) using their long-running 1 ms trajectory as the basis for a Markov chain with states and transitions analyzed from the Anton production trajectory. Our algorithm works by resampling snapshot conformations known to be right before transition events, creating an ensemble set of transitions preconditioned on sampling in the space near to the transitions. This enables us to explore the reduced degrees of freedom that drive the transitions and to examine the statistical foundations of the Markov chain description for state transitions in BPTI. To achieve these goals we use an adaptive framework for resampling built around a parallel relational database system and with scripts controlling molecular dynamics codes running on XSEDE sponsored national supercomputers. In addition to BPTI, we will show results that start from other trajectories defined from peptides and from other long-running protein simulations. Our scripts for the initial analysis and the resampling thus readily generalize to both long and short trajectory runs and can be used to increase sampling on a broad range of transition events.

2052-Pos Board B782
Management of Molecular Simulation Database
Anand Kumar, **Yicheng Tu**, Sagar Pandit.
University of South Florida, Tampa, FL, USA.
Large amount of molecular simulations (MS) data produced in scientific studies is stored in computer flat files. Any information requested by the users (or program) is handled by operating system. Often, the same files are accessed multiples times when different information is requested or different analytical query is executed. Multiple users accessing same file would result in multiple access of the file. The I/O bandwidth requirement is huge and multiple accesses delay requests that arrive in sequence. Therefore, there is need for managing the MS data effectively so that access to it is efficient.
In this work, we propose the idea of storing MS data in a database management system (DBMS) and develop novel indexing strategies to help optimize the process of a wide range of queries. The query-plan generation feature of DBMS minimizes number of accesses to the file system. Multiple functions (queries) accessing same simulation file can share single access request made to the file system.