COMP 1

Probing the α-helical structural stability of stapled p53 peptides: Molecular dynamics simulations and analysis

Goran Krilov(1), gkrilov87@gmail.com, 120 West 45th Street, Tower 45, New York NY 10036, United States; Zuojun Guo(2); Udayan Mohanty(2); Justin Noehre(3); Tomi Sawyer(3); Woody Sherman(1). (1) Schrodinger, Inc, New York NY 10036, United States (2) Department of Chemistry, Boston College, Chestnut Hill MA 02467, United States (3) Aileron Therapeutics, Inc., Cambridge MA 02139, United States

Reactivation of the p53 apoptosis pathway through inhibition of the p53-hDM2 interaction suppresses tumor growth in many human cancers. Stabilization of the helical structure of synthetic p53 analogs via a hydrocarbon cross-link (staple) has been found to lead to increased potency and inhibition of protein-protein binding. However, details of the structure and dynamic stability of the stapled peptides and peptide-protein complexes are not well understood. Extensive replica-exchange molecular dynamics simulations are used to study a series of stapled α-helical peptides in solution, and bound to the protein target. The peptides are found to exhibit substantial variations in predicted α-helical propensities in good agreement with the experimental observations. In addition, we find significant variation in local structural flexibility of the peptides, which appears to be closely related to biological activity. These simulations provide new insights into the design of α-helical stapled peptides and the development of potent inhibitors of protein-protein interfaces.

COMP 2

Density functional theory calculations on matrix metalloproteinases and inhibitors

Mario A Ferron(1), maferron@broncs.utpa.edu, 1201 W. University Drive, Edinburg Texas 78539, United States; Nikolaos Dimakis(1), dimakis@utpa.edu, 1201 W. University Drive, Edinburg Texas 78539, United States; Randy Castillo(2), 1106 N. Shary Road, Mission Texas 78572, United States; Illiana Leal(3); Ricardo Espinoza(3); Nestor Navarro(3). (1) Department of Physics and Geology, The University of Texas-Pan American, Edinburg Texas 78539, United States (2) Science Department, Sharyland High School, Mission Texas 78572, United States (3) Dual Language Program, Pharr-San Juan-Alamo High School, San Juan Texas 78589, United States

Cancer is hardest to control once it has spread throughout the body. The purpose of this research is to determine properties of molecules that are involved in metastasis, a process by which cancer cells spread throughout the human body. We used Density Functional Theory (DFT) to analyze Matrix
Metalloproteinases (MMPs), molecules which are believed to be linked to medical conditions such as arthritis, cirrhosis, and metastasis of malignant cancer cells. In this study, we analyzed structural properties of MMP-9 and MMP-13. MMPs are metalloendopeptidases, metal-based enzymes that degrade extracellular matrix structural components. In the case of MMP-9, it degenerates type IV collagen, which is a major matrix component of the basal lamina. MMP active sites, located across their catalytic domains, consist of a central zinc (II) ion bound to three histidine residues. DFT calculations in the active sites of these molecules may help us understand more their mechanical functions and would provide basis for further research to develop ways for their control. Part of our research includes testing the affinity of new anticancer medications to biological molecules such as hemoglobin. The procedures will be performed using Schrödinger's MAESTRO and JAGUAR, software which is efficient for quantum mechanical \textit{ab initio} phase calculations of metal based systems.

COMP 3

Efficient data mining of the human kinome

\textit{Johannes Maier}\(^{(1)}\), jmaier@chemcomp.com, 1010 Sherbrooke Street West, Suite 910, Montreal Quebec H3A 2R7, Canada \(^{(1)}\) Chemical Computing Group, Montreal Quebec H3A 2R7, Canada

The eukaryotic protein kinases (ePKs) constitute one of the largest protein families in the human genome. EPKs are ATP-dependend phosphotransferase enzymes and key players in a broad variety of integral cellular processes including cell cycle, cell growth and death, metabolism and differentiation. The finding that many viral oncogenes encode kinases with constitutive enzyme activity raised interest in ePKs as pharmacological targets. Understanding and applying the relationship between the binding of inhibitors and the correlated mechanism of inhibition is the basis for targeted drug design. Here, we present a kinase analysis tool that aims to integrate the structural knowledge base available from public sources with practical chemogenomics. The application is based on flexible components that allow a variety of entry points into the kinase knowledge pool as guide to address issues like target selection, profiling panel selection and further exploration of existing structure/inhibitor projects.

COMP 4

Methemoglobinemia caused by 8-aminoquinoline drugs: DFT calculations suggest an analogy to H4B’s role in nitric oxide synthase

\textit{Haining Liu}\(^{(1)}\), hliu@olemiss.edu; \textit{Larry A. Walker}\(^{(2)}\); \textit{N. P. Dhammika Nanayakara}\(^{(3)}\); \textit{Robert J. Doerksen}\(^{(1)}\). \(^{(1)}\) Department of Medicinal Chemistry, University of Mississippi, University MS 38677, United States \(^{(2)}\) Department of
The 8-aminoquinoline class of antimalarial drugs is effective for radical cure against the liver stages of malarial parasite. However, they are toxic to erythrocytes of glucose-6-phosphate dehydrogenase-deficient subjects by a process involving conversion of hemoglobin to methemoglobin and generation of reactive oxygen species. The toxicity is believed to be mediated by one or more oxidative metabolites, such as 5-hydroxyprimaquine. Unfortunately, the chemical mechanism involved remains unclear despite six decades of extensive studies. In this work, we performed density functional theory calculations on a 5-hydroxyprimaquine-hemoglobin model. We found that when $\text{O}_2$ is singly (OOH) and doubly (HOOH) protonated, 5-hydroxyprimaquine donates half an electron and one electron, respectively, to $\text{O}_2$. This thus assists the formation of hydrogen peroxide and may contribute to the toxicity of 8-aminoquinoline drugs. Such an electron-donating ability suggests that 5-hydroxyprimaquine plays an analogous role to that of the H4B cofactor in nitric oxide synthase.

**COMP 5**

**Evaluation and ranking of enzyme designs**

Gert Kiss\(^{(1)}\), gkiss@chem.ucla.edu, 607 Charles E. Young Drive East, Los Angeles CA 90095, United States; Kendall N Houk\(^{(1)}\); Daniela Rothlisberger\(^{(2)}\); David Baker\(^{(2)}\). (1) Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles CA 90095, United States (2) Department of Biochemistry, University of Washington, Seattle WA 98195, United States

In 2008, a successful computational design procedure was reported that yielded active enzyme catalysts for the Kemp elimination. Here, we studied these proteins together with a set of previously unpublished inactive designs in order to determine the sources of activity or lack thereof, and to predict which of the designed structures are most likely to be catalytic. Methods that range from quantum mechanics (QM) on truncated model systems to the treatment of the full protein with ONIOM QM/MM and AMBER molecular dynamics (MD) were explored. The most effective procedure involved molecular dynamics, and a general MD protocol was established. Substantial deviations from the ideal catalytic geometries were observed for a number of designs. Penetration of water into the catalytic site and insufficient residue-packing around the active site are the main factors that can cause enzyme designs to be inactive. Where in the past, computational evaluations of designed enzymes were too time-extensive for practical considerations, it
has now become feasible to rank candidates computationally prior to and in conjunction with experimentation, thus markedly increasing the efficiency of the enzyme design process.

COMP 6

Thermodynamic contributions to PDZ binding affinities: A theoretical and experimental case study

Leslie Vogt\(^{(1)}\), lvogt@fas.harvard.edu, 12 Oxford St, Box 489, Cambridge MA 02138, United States; Ethan Karp\(^{(1)}\); Alan Aspuru-Guzik\(^{(1)}\); Gavin MacBeath\(^{(1)}\).

(1) Department of Chemistry and Chemical Biology, Harvard University, Cambridge MA 02138, United States

Characterizing the thermodynamic contributions to binding in protein-protein interaction domains, such as PDZ domains, can lead to a better understanding of binding selectivity. We performed molecular dynamics simulations of 26 related peptides bound to TIP1 to probe binding of peptides with mutations in the -1, -3, and -4 position from the C-terminus. Bound peptide structures were determined via homology modeling. The binding free energy for each peptide was obtained using the MM/PBSA method with entropic contributions evaluated from normal mode analysis on trajectory snapshots. These values were compared to experimental data obtained via fluorescence polarization and isothermal calorimetry (ITC) to validate the method, particularly for a single mutation series for position -3. Enthalpic and entropic contributions to binding were compared to experimental data using ITC to gauge the accuracy of the molecular dynamics simulations and to derive microscopic insight into the contribution each residue plays in the overall binding affinity.

COMP 7

Aggregation dynamics of partially sticky and structured colloidal spheres under shear stresses

Rigoberto Hernandez\(^{(1)}\), hernandez@gatech.edu, School of Chemistry & Biochemistry, Georgia Institute of Technology, Atlanta GA 30332-0400, United States; Matthew C. Hagy\(^{(1)}\); Alex V. Popov\(^{(1)}\).

(1) School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta GA 30332-0400, United States

Bruce Berne studied the phase diagram and nonequilibrium properties of partially sticky and structured colloidal spheres in the mid-seventies. Mismatches in the positions of the sticky ends with given structures gives rise to frustration and richer morphologies. Meanwhile, we have also found that barriers in the interaction potential leading to novel structuring of the colloids is highly sensitive
to shear. The computational modeling of particles with asymmetric and reactive interactions required specialized simulation codes using Brownian dynamics as the propagator and specialized boundary conditions. The degree of surface roughness of the colloids has been seen to affect the aggregation dynamics and morphology. Additionally, it has been found that shearing of colloid suspension during aggregation can further influence the path of aggregation. On removal of this nonequilibrium perturbation, the shear-driven aggregation structure may remain due to colloidal nonergodicity, and is thereby highly depending on the time-dependent shear profile.

COMP 8

Molecular simulations of electric field effects in condensed phase systems

Alenka Luzar\(^{(1)}\), aluzar@vcu.edu, 1001 West Main St., Richmond VA 23284-2006, United States; Christopher D. Daub\(^{(1)}\); Dusan Bratko\(^{(1)}\). (1) Department of Chemistry, Virginia Commonwealth University, Richmond VA 23284-2006, United States

As accessible experimental length scales become shorter, the modification of interfacial properties of water using electric field (electrowetting), and nanopatterned surface techniques must come to grips with novel effects existing at the nanoscale. We will briefly survey some of our recent progress we have made in understanding the effects of electric field and surface corrugation at the nanoscale on water interfacial tension and on water-mediated interactions by using molecular simulations. We devote particular care to proper control of the pressure affected by a spatially discontinuous electric field. The work gives basic understanding of applied and internal field effects that can operate in condensed phase systems, from modulating local hydrophilicity/hydrophobicity of engineered and biological surfaces, to surface manipulation in nanofluidic devices.

COMP 9

Theories of molecular fluids confined in disordered porous materials

Lev Sarkisov\(^{(1)}\), Lev.Sarkisov@ed.ac.uk, Sanderson Building, Kings Buildings, Edinburgh Scotland EH9 3JL, United Kingdom. (1) Institute for Materials and Processes, University of Edinburgh, Edinburgh Scotland EH9 3JL, United Kingdom

A fundamental description of fluids under confinement is important to our efforts to design next-generation porous materials with specified characteristics. Computer simulation has been a great aid in this process, particularly for crystalline materials and materials with regular porous morphology. The situation
is more complex for disordered porous materials, where a molecular model must capture the essential features of structural disorder on the appropriate length scales. This presents a serious challenge even with modern computer power. Theoretical methods can offer a more efficient alternative while providing broad, general insight to systems of interest. In this talk, we review recent advances in theoretical integral equation approaches to molecular fluids under confinement in disordered media. We focus on replica Ornstein–Zernike-based approaches, and emphasize interaction site fluid and associating fluid applications.

**COMP 10**

**Self-consistent treatment of polarization in condensed-phase QM/MM calculations**

*John M. Herbert*\(^{(1)}\), *herbert@chemistry.ohio-state.edu*, 100 W. 18th Ave., Columbus OH 43210, United States; *Leif D. Jacobson*\(^{(1)}\); *Christopher F. Williams*\(^{(1)}\). (1) Chemistry, Ohio State University, Columbus OH 43210, United States

We describe several hybrid quantum mechanics/polarizable molecular mechanics ("QM/polMM") models, where the MM force field is polarizable and mutual QM-MM polarization effects are treated in a fully self-consistent fashion. This methodology was originally developed to model aqueous electrons, where we find that a self-consistent treatment of electron-water polarization is necessary in order to explain the gross features of the optical spectrum in bulk solution. The new model predicts the existence of polarization-bound quasi-continuum states that are absent in non-polarizable models. Recently, we have extended these methods to many-electron quantum mechanics, in particular, time-dependent density functional theory. We present a variational, Lagrangian-based formulation of TD-DFT/polMM, along with preliminary applications to intramolecular charge-transfer excited states in donor-\(\pi\)-acceptor chromophores. TD-DFT/polMM calculations reproduce the 1 eV shift between absorption and emission spectra in polar solvents.

**COMP 11**

**Developed the potential function of electrostatic interactions for quadrupolar fluid on the basis of consider the atomic charges as variables values for calculation of structure and thermodynamic properties by RISM**

*Svetlana Vladimirovna Keshtova*\(^{(1)}\), sk911y.sv@gmail.com, Vinokurovo 16, 43, Moscow Russia 117447, Russian Federation; *Fatima Keshtova*\(^{(2)}\). (1) Phys Chem, Private researcher, Moscow Russia 117447, Russian Federation (2) Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow Russia 119991, Russian Federation
For investigation of quadrupolar fluid we used modified potential interactions between atoms for obtaining correlation functions. For modeling liquid phase has been taken the theory of integral reference interaction sites model. The molecules of these compounds have three atoms and they are symmetry and linear with large quadrupole moment and lone pair electrons.

COMP 12

Brownian motion in nonequilibrium colloidal suspensions: Dynamics of a stationary solution and swelling colloids

Alexander V Popov(1), popov@chemistry.gatech.edu, 901 Atlantic Drive, Atlanta Georgia 30332-0400, United States; Rigoberto Hernandez(1). (1) School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta Georgia 30332-0400, United States

The dynamics of colloidal particles have been seen by several experimental groups to be affected by their swelling or shrinking due to the changes in the solution temperature or pH. When such changes are fast, the colloids move within a nonequilibrium environment and may access otherwise inaccessible morphologies. In order to better understand this behavior, we have developed a series of reduced-dimensional nonequilibrium models based on the Langevin equation (LE) formalism. The colloidal system can be treated as if each particle is immersed in two subenvironments: a stationary solution and a nonstationary "gas" of identical colloids. To account for nonsynchronous changes of different parts of the environment, a special irreversible form of LE has been derived (J. Chem. Phys. 126(2007)244506). Using numerical simulations of both the reduced-dimensional models and the full-dimensional particle dynamics, we have confirmed the applicability of this approach. It offers a testbed for nonequilibrium theories.

COMP 13

Tautomerism in chemical information management systems

Wendy A. Warr(1), wendy@warr.com, 6 Berwick Court, Holmes Chapel Cheshire CW4 7HZ, United Kingdom. (1) Wendy Warr & Associates, Holmes Chapel Cheshire CW4 7HZ, United Kingdom

Tautomerism has an impact on many of the processes in a chemical information management system including novelty checking during registration into chemical structure databases; storage of structures; exact and substructure searching in chemical structure databases; and depiction of structures retrieved by a search. For this talk the approaches taken by a great many different software vendors and database producers have been compared. Since it is important to take account of the nature of the database and the process for which it is designed,
and the user requirements vary, it is dangerous to lay down the law about what is right and wrong. The comparison is nevertheless of considerable interest.

COMP 14

Tautomerism in large databases

Markus Sitzmann(1), sitzmann@helix.nih.gov, 376 Boyles Str., Frederick MD 21702, United States; Wolf-Dietrich Ihlenfeldt(2), Marc C Nicklaus(1). (1) Chemical Biology Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, DHHS, NCI-Frederick, Frederick MD 21702, United States (2) Xemistry GmbH, Königstein 61462, Germany

We are reporting on a comprehensive tautomerism analysis of one of the largest currently existing sets of real (i.e. not computer-generated) compounds. We used the Chemical Structure DataBase (CSDB) of the NCI CADD Group, an aggregated collection of over 150 small-molecule databases totaling 103.5 million structure records. Tautomerism was found to be possible for more than 2/3 of the unique structures in CSDB. A total of 680 million tautomers were calculated from the original structure records. Tautomerism overlap within the same individual database (i.e. at least one other entry was present that was really only a different tautomeric representation of the same compound) was found at an average rate of 0.3% of the original structure records, with values as high as nearly 2% for some of the databases in CSDB. Tautomeric overlap across all constituent databases in CSDB was found for nearly 10% of the records in the collection.

COMP 15

Tautomerism in drug discovery

Bahaa El-Dien M. El-Gendy(1)(2), bahaa@ufl.edu, University of Florida, Department of Chemistry, Gainesville, FL, Gainesville FL 32611-7200, United States; Alan R. Katritzky(1), C. Dennis Hall(1), Bogdan Draghici(1). (1) Department of Chemistry, University of Florida, Gainesville Florida 32611-7200, United States (2) Department of Chemistry, Benha University, Gainesville Florida 32611-7200, United States

The influence of tautomerism on the precise structure of drugs and thus of their potential to interact in biological systems is discussed from thermodynamic and kinetic aspects. The types of tautomerism encountered in the structure of drugs in current use are surveyed together with the effect of pH, solvent polarity, and temperature.

COMP 16
Quantitative forecasts of biological potency of molecules that can tautomerize

Yvonne C Martin\(^{(1)}\), yvonnecmartin@comcast.net, 2230 Chestnut St, Waukegan IL 60087, United States. (1) Martin Consulting, Waukegan IL 60087, United States

Whether one is using ligand-based 2D or 3D QSAR or structure-based estimates of potency of molecules, tautomerism needs to be addressed. This talk will highlight insights as to when one needs to consider tautomerism and how it can be included in potency forecasts.

COMP 17

New questions about tautomerism in cytosine: Quantum chemical and matrix isolation spectroscopic studies

Geza Fogarasi\(^{(1)}\), fg@chem.elte.hu, Pf. 32., Budapest H-1518, Pf. 32., Hungary; Gabor Bazso\(^{(2)}\); Peter G Szalay\(^{(1)}\); György Tarczay\(^{(2)}\). (1) Laboratory of Theoretical Chemistry, Institute of Chemistry, Eotvos University, Budapest Budapest H-1518, Pf. 32., Hungary (2) Laboratory of Molecular Spectroscopy, Institute of Chemistry, Eotvos University, Budapest Budapest H-1518, Pf. 32., Hungary

In spite of numerous studies, there is much uncertainty about tautomerism in nucleic acids and specifically cytosine. In the gas phase, form 2 dominates but \(\Delta G\) maybe about 1 kcal/mol for both 1 and the “rare” form 3. Spectroscopic studies “see” them but in much smaller abundance. The UV spectrum is normally assigned to 1. Dimerization may also influence tautomerization.

![Fig. 1. Selected isomers and a dimer of cytosine](image)

We present infrared and UV spectroscopic measurements in Ar matrix and discuss them by MP2 and CC quantum chemical calculations, including
electronic excitations. Contributions from isomers/tautomers and/or dimers to the spectra are discussed.

COMP 18

Muonic molecules control in intense laser field

Quan-Fang Wang(1), quanfangwang@yahoo.co.jp, Shatin N.T., Hong Kong, Hong Kong, China. (1) Mechanics and Automation Engineering, The Chinese University of Hong Kong, Hong Kong Hong Kong 100000, China

Recent progresses on laser physical experiments of muonic particles have been reported for ionization of muonic atoms and muonic molecular ions (ddμ). Especially, those particles in strong laser field are well appeared in literatures.

This presentation is aiming at quantum control for the dynamics of muonic molecular described by time depended Schrodinger equation. By controlling through intense laser field (electronic field), we want to find the theoretical and computational approach for matching the existing experimental results, and provide the guidance to real laboratory experiments. More precisely, we simulate laser-controlled nuclear experiments in muonic molecules. Furthermore, other possibility of laser-induced nuclear processes will also be highlighted.

Future perspectives would be focus on applying laser in muon-catalyzed fusion, ionization and dissociation of mesic system, particularly in the extension of current investigations to manipulate electron dynamics.

COMP 19

Exploring the free energy landscapes of chemical reactions

Bernd Ensing(1), b.ensing@uva.nl, Nieuwe Achtergracht 166, Amsterdam NH 1018WV, The Netherlands. (1) Department of Chemistry, University of Amsterdam, Amsterdam NH 1018WV, The Netherlands

Concerted chemical reactions are usually well described by a small set of order parameters, or collective variables, but a good reaction coordinate is not known a priory. When the set of collective variables, for example the lengths of the bonds that are being broken and formed in the reaction, is small enough, the multidimensional free energy surface (FES) of the reaction can be computed, for instance using the metadynamics simulation technique. The lowest free energy pathway, connecting the stable reaction and product states, is a good estimate for the reaction coordinate. However, even converging a 3-dimensional FES can be very computationally demanding. Here, we introduce a new flavor of metadynamics using ideas from path finding methods, such that the computational cost is no longer dependent on the number
of collective variables. Using the improved method, we efficiently obtain the reaction free energy and the reaction pathway in a single simulation.

COMP 20

Three-body interactions in condensed phases of helium

Robert J. Hinde(1), rhinde@utk.edu, Dabney Hall, Knoxville TN 37996, United States. (1) Department of Chemistry, University of Tennessee, Knoxville TN 37996, United States

Three-body interactions play an important role in the properties of high-density condensed phases of helium. We have computed an ab initio three-body interaction energy surface for He$_3$ trimers using coupled cluster quantum chemical techniques that include fully connected triple excitations and a perturbative treatment of quadruple excitations. We evaluate the quality of our three-body surface by using it in quantum Monte Carlo calculations of both the zero-temperature $p$–$V$ curve of solid He and the phonon spectrum of the zero-temperature solid. Our three-body surface appears to describe reliably the He$_3$ three-body interaction for compact He$_3$ geometries that are important in high-density condensed phases of helium. We therefore use it as a benchmark for testing the performance of density functional methods in computing three-body interactions in He$_3$.

COMP 21

Activation of the CO bond in CO$_2$ by Ta and Nb mono and dications

Emilly Hull(1), ehull@iastate.edu, 122 Spedding, Ames IA 50011-3020, United States; Theresa Windus(1). (1) Chemistry, Iowa State University, Ames IA 50010, United States

CO$_2$ is a product of fossil fuel combustion and major greenhouse gas. In order to develop better methods of CO$_2$ sequestration and emissions reduction, the possibility of CO bond activation in CO$_2$ by transition metals deserves consideration. Experimental and theoretical data is presented for the gas phase reactions of Ta and Nb mono and dications with CO$_2$. The potential energy surfaces of these reactions were studied at the DFT and CAS-SCF levels using the LANL2DZ and LANL2TZ ECP for the metals and triplet $\zeta$ basis sets were used for all other atoms. The spin crossings of the reactions (quintet to triplet or quartet to doublet) and the energetics of these reactions will be presented.

COMP 22

Chemical potential from particle removal and insertion: Upper and lower bounds and accurate estimates from computer simulations
Particle insertion and removal provide upper and lower bounds respectively for the excess chemical potential from the corresponding insertion and removal energy distributions determined by computer simulation. We discuss how the chemical potential can be obtained accurately by combining the energy distributions over the whole range of energies beyond the overlap region. Our calculations are compared with the overlap method and Bennet's acceptance ratio method.

COMP 23

DOCK Blaster: Automatic docking with self-assessment

John J Irwin\(^{(1)}\), jji@cgl.ucsf.edu, 1700 4th St, Box 2550, Room BH-501D, San Francisco CA 94158-2330; Brian K Shoichet\(^{(1)}\); Kong T Nguyen\(^{(1)}\); Michael M Mysinger\(^{(1)}\). (1) Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco CA 94158-2330, United States

Structure-based docking is the most practical approach to leverage protein structure for ligand discovery, but the technique retains important liabilities that make it challenging to deploy on a large scale. We have therefore created an expert system, DOCK Blaster, based on UCSF DOCK 3.5.54, to investigate the feasibility of full automation. The method requires a PDB code, sometimes with a ligand structure, and from that alone can launch a full screen of large libraries. A critical feature is self-assessment, which estimates the anticipated reliability of the automated screening results using both pose fidelity and enrichment. To test DOCK Blaster, we have assembled a database of 7408 protein targets from the PDB having a non-covalent ligand bound for which five or more ligands are also available in ChEMBL. We have assessed automated docking performance based on enrichment of experimentally known ligands compared to either property-matched decoys, or the entire ChEMBL database as a proxy for a screening library. For over 55% targets, a logAUC above random of 10 or better was achieved, suggesting that expert-free docking is viable for ligand discovery.

COMP 24

Key directions of mastering docking and scoring approaches

Ghermes Chilov\(^{(1)}\), ghermes@moltech.ru, Leninskie gory, 1/75A, Moscow not applicable 119992, Russian Federation; Oleg Stroganov\(^{(1)}\); Fedor Novikov\(^{(1)}\);
Our experience of developing Lead Finder docking software suggests the following directions of improving docking and scoring. First, increase of docking success rate can be achieved by intelligent reduction of the phase space available to ligands. This can be realized in a number of ways, for example by introducing geometric constraints or energetic preferences for the particular pairs of protein and ligand atoms that are applied during the docking process, and by 3D shape filtration of ligand poses that favors conformations of known active ligands. Second, scoring approaches should concentrate on free energy rather than single point calculations. Our new graph-theory algorithm TSAR (thermodynamic sampling of amino acid residues) explicitly accounts for multi-conformational states of protein side chains, ligand and bound water to estimate the thermodynamic averages of interaction energies. Introduction of protein flexibility would also contribute to the docking accuracy achieved by using the single protein model.

COMP 25

How eHiTS solves the docking and scoring problems

Zsolt Zsoldos\(^{(1)}\), zsolt@simbiosys.com, 135 Queen[apos]s Plate Drive, Suite 520, Toronto Ontario M9W6V1, Canada ; Orr Ravitz\(^{(1)}\). (1) SimBioSys Inc., Toronto Ontario M9W6V1, Canada

Ten lessons learned during ten years of eHiTS docking and scoring development:
1. It is not sufficient to sample a few dozen low energy conformers
2. The search space is vast: brute force systematic or blind random searches fail
3. Need to hit all the good interactions while avoiding bad ones, BUT it is not always clear what is good
4. Location, location, location: buried versus exposed surfaces, pocket depth
5. A weak interaction is better than none: Pi-cation, C-H...O and others
6. Not all H-bonds are created equal: the use of functional group knowledge base
7. The importance of protonation states, induced changes upon binding
8. Beware of data errors: curation of PDB structures, binding data inconsistencies
9. Diversify the 3 aspects of scoring: pose ranking, enrichment and binding energy estimation
10. Different protein families need different weighting schemes

COMP 26
Combining atom-based and residue-based scoring functions in protein-protein docking

Thom Vreven\(^{(1)}\), Thom.Vreven@umassmed.edu, 364 Plantation Street, LRB Room 1006, Worcester MA 01605, United States; Howook Hwang\(^{(1)}\); Zhiping Weng\(^{(1)(2)(3)}\). (1) Program in Bioinformatics and Integrative Biology, University of Massachusetts Medical School, Worcester MA 01605, United States (2) Bioinformatics Program, Boston University, Worcester MA 01605, United States (3) Department of Biomedical Engineering, Boston University, Worcester MA 01605, United States

Most scoring functions for protein-protein docking algorithms are either atom-based or residue-based, but not both. Here we combine atom-based potentials with residue-based potentials. We used our ZRANK program for re-ranking docking predictions, which originally includes only atom-based terms. We incorporated several published interface and pair-wise propensities, and optimize the weights simultaneously. ZDOCK3.0 decoys were used for training/testing, using complexes from our benchmark where ZDOCK finds correct predictions. Using the success rate of the top 2000 predictions (out of 54,000), our atom-based IFACE pair-wise potential reduces the number of cases without hits by 31\% (success rate increases from 67\% to 77\%). Including the best residue-based interface propensities and best pair-wise propensities reduces the number of non-hit cases by 37\% and 44\%, respectively. Adding IFACE to the residue-based potentials, however, only improves the accuracy slightly. We conclude that residue-based and atom-based potentials capture similar information, and can be combined in scoring functions.

COMP 27

SHAPEFIT: Pose prediction with shape and chemistry

Brian P Kelley\(^{(1)}\), kelley@eyesopen.com, 9 Bisbee Court, Santa Fe NM 87508, United States; Scott P Brown\(^{(2)}\). (1) OpenEye Scientific Inc, Santa Fe NM 87508, United States (2) Abbott Laboratories, Abbott Park IL 60064-3500, United States

SHAPEFIT is a method for predicting poses when one possesses knowledge of one or more existing bound conformations. SHAPEFIT aligns a given molecule using a combined shape and MMFF94 forcefield which optimizes shape overlap to the bound ligand while simultaneously limiting conformational strain. SHAPEFIT compares favorably against traditional methods such as rigid overlays and docking into a protein.
To describe protein-ligand interactions realistically, the structural changes in both the receptor and the ligand should be considered during complex formation. We developed a docking approach, DynaDock, which combines various sampling techniques with a new molecular dynamics based refinement method, OPMD (Optimized Potential Molecular Dynamics) and allows for flexible treatment of the whole system. The method works especially well for large, flexible ligands (peptides) and open, surface exposed binding sites.

The approach was tested for 15 protein-peptide complexes (2 to 16 residues). For all peptides docking poses with RMSD values < 2.1Å could be obtained (best scoring RMSD < 3.4Å). (Antes, Proteins, 2010). In addition, the approach was successfully evaluated for docking into the apo-structures of 10 protein-ligand complexes with flexible binding sites and backbone movement upon ligand binding (all best RMSD values < 2.0Å). In all cases full flexibility was allowed for the ligand as well as the receptor.

**COMP 29**

*(kSP) fluctuating charge method based on the consider polarization effects and anisotropy steric van der Waals radiuses of atoms in molecule*

The new scheme to predicting behavior of atomic charges in molecules, which depend on environment described here. The method combines some advantages of the previous techniques (partial equalization of orbital electronegativities; fluctuating charge) and new approaches (two systems of $\sigma$ and $\pi$ electrons have been considered separately; new parameters were used). The method has been avoided mistakes the concept of dot charges on atoms. New technique used Thole's polarizabilities and van der Waals steric radiuses of atoms in molecules, where last obtained from the Natural bond orbital steric analyses. In method suggested new parameters as a tool for the estimation of quantum effects in the frame of classical computer simulation. These new parameters can characterize the degree of perturbation of valence electrons of atoms and bonds in process of creation of chemical bonds. The new method can produce results show good agreement with results of ab initio calculations.
COMP 30

Acidity constants and solvation structures of amino acids via DFTMD

Martina Mangold(1), mm695@cam.ac.uk, Lensfield Road, Cambridge Cambridgeshire CB2 1EW, United Kingdom; Leslie Rolland(2); Francesca Costanzo(3); Michiel Sprik(1); Marialore Sulpizi(1); Jochen Blumberger(4). (1) Department of Chemistry, Cambridge University, Cambridge, United Kingdom (2) Departement de Chimie, Ecole Normale Superieure, Paris, France (3) Dipartimento di Chimica Fisica e Inorganica, Universita di Bologna, Bologna, Italy (4) Department of Physics and Astronomy, University College London, London, United Kingdom

We report on the calculation of absolute acidity constants of amino acid side chain and backbone groups using a recently introduced density functional based molecular dynamics (DFMD) simulation technique [Sulpizi and Sprik, Phys. Chem. Chem. Phys. 10, 2008]. Free energy differences are obtained from the vertical energy gaps between protonated and deprotonated states through removal or insertion of a proton. The proton position for insertion is determined by a dummy proton kept in place by a harmonic potential. The accuracy of our approach is ± 2 pKa units, which is the accuracy expected for DFT. This study is an important first step in the development of a more rigorous pKa value calculation of amino acids in proteins. Moreover, the solvation structures obtained from DFMD may serve as a benchmark for force field or continuum model development.

COMP 31

Variational formulation of the polarizable continuum model

Giovanni Scalmani(1), giovanni@gaussian.com, 340 Quinnipiack St Bldg 40, Wallingford CT 06492, United States; Filippo Lippiari(2)(3); Benedetta Mennucci(3); Eric Cances(4); Marco Caricato(1); Michael J Frisch(1). (1) Gaussian, Inc., Wallingford CT 06492, United States (2) Scuola Normale Superiore, 56126 Pisa, Italy (3) Dipartimento di Chimica e Chimica Industriale, Universita[apos] di
We describe a new formulation of the Polarizable Continuum Model (PCM) of solvation in terms of a free energy functional which depends on both the solute's properties and a set of degrees of freedom representing the polarization of the dielectric continuum. The variational minimization of such functional with respect to all its parameters is equivalent to the solution of the Poisson's equation underlying the PCM model. The application of the recently introduced Continuous Surface Charge (CSC) formalism leads to a mathematically smooth functional which is free of discontinuities and singularities.

The PCM free energy functional is suitable for a number of interesting applications like: i) the simultaneous optimization of the molecular geometry and/or electronic density and solvent polarization, ii) the proper description of the reaction field self-consistent with post-SCF or excited states densities, and iii) the formulation of an extended Lagrangian for molecular dynamics simulations.

COMP 32

Going beyond the Frozen Core Approximation with a coordinate-dependent pseudopotential

Argyris Kahros\(^{(1)}\), argyris@chem.ucla.edu, 607 Charles E. Young Drive East, Los Angeles California 90095-1569, United States; Benjamin J Schwartz\(^{(1)}\). (1) Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles California 90095-1569, United States

Pseudopotentials are necessary to reduce the number of degrees of freedom when performing quantum calculations on systems with large numbers of electrons. Although many pseudopotentials are developed empirically, pseudopotentials can be rigorously calculated using a formalism due to Phillips and Kleinman. All pseudopotentials, however, are subject to the Frozen Core Approximation, which neglects changes in the orbitals of the implicitly-treated electrons with changes in the chemical bonding. In this talk, we show how to go beyond the Frozen Core Approximation and develop pseudopotentials that can respond to changes in the positions of the nuclei in a molecule. We demonstrate this idea for the sodium dimer, where changes in the orbitals of the implicitly-treated electrons are accounted for from the point of bonding to the dissociation limit. We plan to use such coordinate-dependent potentials in mixed quantum/classical MD simulations to rigorously investigate the solvent-induced dissociation of this and related molecules.

COMP 33

Coupled cluster calculations in solution with the polarizable continuum model of solvation
Marco Caricato(1), marco@gaussian.com, 340 Quinnipiac St, Bldg 40, Wallingford CT 06492, United States; Gary W Trucks(1); Michael J. Frisch(1). (1) Gaussian, Inc., Wallingford CT 06492, United States

Coupled cluster theory (CC) provides very accurate estimates of energies and molecular properties. However, such calculations are often limited to gas phase species due to the large computational cost of this level of theory, whereas most of the chemical phenomena take place in solution. We propose an efficient implementation of the polarizable continuum model of solvation (PCM) with the coupled cluster singles and doubles method (CCSD) to take into account the solvent effects on energy and geometry. The PCM approach does not require conformational sampling as for atomistic representations of the solvent molecules and automatically describes mutual polarization effects between solute and solvent. Applications of the CCSD/PCM method to representative molecules are presented.

COMP 34

Modeling molecular charge density with Gaussian multipoles

Dennis M Elking(1), delking@email.unc.edu, Campus Box 3290, Caudill and Kenan Laboratories, Chapel Hill NC 27599-3290, United States; Lalith Perera(2); Lee G Pedersen(1)(2); Robert Duke(1)(2). (1) Department of Chemistry, University of North Carolina-Chapel Hill, Chapel Hill NC 27599, United States (2) Laboratory of Structural Biology, National Institute of Environmental Health Sciences, Research Triangle Park NC 27709, United States

A model for molecular charge density based on atom centered Gaussian multipoles is described. In contrast to atomic point multipole models, Gaussian multipole models are able to account for penetration effects at short range interaction distances. The Gaussian multipoles are fit to the ab initio electrostatic potential surrounding a molecule and tested by comparing electrostatic dimer energies, inter-molecular overlap integrals, and permanent molecular multipole moments with their respective ab initio values. In addition, a method of calculating atomic forces for geometry dependent multipoles is described. It is shown that geometry dependent multipoles are able to reproduce ab initio atomic forces, while static multipole models are able to reproduce total molecular forces and torques.

COMP 35

Dispersion in combined ab initio/effective fragment potential systems

Tony Smith(1), toni@si.msg.chem.iastate.edu, 201 Spedding Hall, Ames Iowa 50011, United States; Lyudmila Slipchenko(2); Mark S. Gordon(1). (1) Department of Chemistry, Iowa State University, Ames Iowa 50011, United States (2)
Dispersion, an attractive intermolecular force arising from a temporary dipole-induced multipole interaction, plays a major role in many systems of chemical and biological interest. An accurate description of dispersion is essential for modeling benzene dimers, DNA base stacking, and other systems involving pi-pi interaction. The general effective fragment potential (EFP2) method, a discrete solvation method in GAMESS, already accounts for fragment-fragment dispersion interaction; an EFP2/QM method is presently in development. A means of modeling the dispersion interaction between an EFP fragment and an ab initio molecule is presented, and results from a selection of dimer systems are demonstrated.

**COMP 36**

**Statistical mechanical basis of coarse graining**

**Hans C. Andersen**(1), hca@stanford.edu, Department of Chemistry, Stanford CA 94305, United States; Avisek Das(1). (1) Department of Chemistry, Stanford University, Stanford CA 94305, United States

Coarse graining is the process of constructing a simplified mechanical model for a molecular material (e.g. an aqueous solution of one or more biomolecules) for which a detailed mechanical model (defined by a Hamiltonian) exists. In comparison to the detailed model, the simplified mechanical model has either fewer degrees of freedom, or simpler interactions, or both, but it is desired that the simplified model correctly describe some specific structural properties of the detailed model. We focus on the multiscale-coarse graining (MS-CG) method of Voth and coworkers. A statistical mechanical theory of coarse graining will be presented. This leads to a variational formulation of the problem of determining the Hamiltonian for a coarse grained model for use in canonical or isothermal-isobaric ensemble simulations. Methods and algorithms for constructing Hamiltonians for coarse grained models will be discussed, as well as numerical tests of these methods and algorithms.

**COMP 37**

**Well tempered-ensemble**

**Michele Parrinello**(1), parrinello@phys.chem.ethz.ch, Via Giuseppe Buffi 13, USI Campus, Lugano Ticino 6900, Switzerland. (1) Department of Chemistry and Applied Biosciences, ETH Zurich, Lugano 6900, Switzerland

In order to alleviate the sampling problem in complex systems characterized by metastable states separated by large barriers, we introduce the well-tempered
ensemble defined by the partition function \( Z_\gamma = \int dU [e^{-\beta U} N(U)]^{1/\gamma} \) where \( U \) is the potential energy and \( N(U) \) the density of states. As \( \gamma \) is varied, \( Z_\gamma \) spans the range from the canonical (\( \gamma = 1 \)) to the multicanonical (\( \gamma = \infty \)) ensemble. We show, that to a first approximation, at intermediate values of \( \gamma \) the average energy is close to its canonical value while its fluctuations are strongly enhanced. These properties can be used to enhance sampling especially in combination with parallel tempering. We show that \( Z_\gamma \) is the ensemble sampled in a well-tempered metadynamics run [A. Barducci, G. Bussi and M. Parrinello, Phys. Rev. Lett. 100, 020603 (2008)] that uses the energy as a collective variable. Canonical ensemble averages are then recovered by applying a recently developed reweighing scheme (M. Bonomi, A. Barducci, M. Parrinello, J. Comput. Chem., 30, 1615 (2009)]. In a series of applications as varied as the Ising model, a Go-like model for HIV protease and the freezing of a Lennard Jones liquid we demonstrate orders of magnitude gain in sampling efficiency.

COMP 38

Molecular simulation: Five decades and counting

Michael L. Klein\(^{(1)}\), mlklein@temple.edu, 1900 North 12th Street, Bio-Life Building 113, Philadelphia PA 19122, United States . (1) Institute for Computational Molecular Science, Temple University, Philadelphia PA 10122, United States

Molecular simulation is now a respected complement to experiment. It was not always like this! Bruce Berne played a key role in enabling the field to take its rightful place in the toolbox of physical scientists. The talk will give a personal overview of the past five decades of molecular simulation.

COMP 39

Chemical dynamics at metal surfaces: The role of electronic excitations

John Tully\(^{(1)}\), john.tully@yale.edu, PO Box 208107, New Haven Connecticut 06520-8107, United States ; Neil Shenvi\(^{(1)}\), Sharani Roy\(^{(2)}\). (1) Department of Chemistry, Yale University, New Haven Connecticut 06520-8107, United States (2) Department of Chemistry, Northwestern University, Evanston Illinois 60208-3113, United States

The adiabatic (Born-Oppenheimer) approximation underlies most of our understanding of chemical reaction dynamics. It has become increasingly apparent, however, that nonadiabatic electronic transitions can sometimes play important roles, particularly in photo-initiated and highly energetic reactions. It is less widely recognized that for chemical reactions at metal surfaces, even at thermal energies, nonadiabatic behavior is the rule rather than the exception. Electron-hole pair creation, electron transfer and hot electron induced motion can
be dominant mechanisms for energy flow and can drastically alter reaction pathways. Recent experiments have demonstrated that reaction exothermicity and molecular vibrational energy can generate highly excited electrons, even resulting in electron emission from the surface. This talk will present progress toward a unified, mixed quantum-classical, picture of nonadiabatic dynamics at metal surfaces, with application to multiquantum vibrational-to-electronic energy transfer in the scattering of nitric oxide from a gold surface.

COMP 40

Coupling Quantum Monte Carlo for electrons with Monte Carlo for Ions

David Ceperley(1), ceperley@uiuc.edu, 1110 W Green St, Urbana IL 61801, United States. (1) Department of Physics, University of Illinois Urbana-Champaign, Urbana IL 61801, United States

In the past decade, we have developed a new method, Coupled Electron-Ion Monte Carlo, to couple simulations of electrons using quantum Monte Carlo (specifically reptation) with Path Integral or classical simulation of ionic degrees of freedom. The goal is to obtain a more accurate ionic energy surface while still obtaining a much lower temperature than with Path Integral simulations of the combined system. We have applied the method to simulations of warm dense hydrogen and hydrogen-helium mixtures. We also discuss recent attempts to simulate a cluster of water molecules with this approach.

COMP 41

Novel molecular dynamics based techniques for enhancing conformational sampling in systems described by rough energy landscapes

Mark E. Tuckerman(1), mark.tuckerman@nyu.edu, 100 Washington Square East, New York New York 10003, United States. (1) Department of Chemistry and Courant Institute of Mathematical Sciences, New York University, New York New York 10003, United States

One of the computational grand challenge problems is the development of methodology capable of sampling conformational equilibria in systems characterized by rough energy landscapes. If met, many important problems, most notably protein structure prediction and enumeration of crystalline polymorphs, could be significantly impacted. In this talk, I will present several new approaches for enhancing conformational sampling in molecular dynamics calculations. These approaches range from straightforward but judicious mass scaling protocols to driven adiabatic dynamics and spatial warping transformations. It will be shown that, in some cases, combining these techniques with replica exchange leads to further enhancements in conformational sampling efficiency. The new techniques will be illustrated on
small polypeptide in solution, long alkane chains, and polymorphic molecular crystals.

COMP 42

Hard and soft interfaces of water

David Chandler(1), chandler@cchem.berkeley.edu, Gilman Hall, Berkeley CA 94720, United States. (1) Department of Chemistry, University of California, Berkeley, Berkeley CA 94720, United States

The formation and response liquid water interfaces have varied effects on the stability and dynamics of nano-scale assemblies in water. Bruce Berne has been a leading contributor to our understanding this facet of nature. This lecture connects with these contributions and describes some of my group’s most recent work aimed at predicting and understanding these effects.

COMP 43

Effects of molecular-scale density fluctuations and surface roughness on hydrophobic hydration

Gerhard Hummer(1), hummer@helix.nih.gov; Jeetain Mittal(2). (1) Laboratory of Chemical Physics, NIDDK, National Institutes of Health, Bethesda Maryland 20892-0520, United States (2) Department of Chemical Engineering, Lehigh University, Bethlehem Pennsylvania 18015, United States

The structure and energetics of water near nonpolar solutes and in nonpolar confinement are central to biomolecular self-assembly and protein function. Molecular-scale fluctuations in the water structure and density at nonpolar interfaces are key determinants of the “hydrophobic effect.” With the help of molecular dynamics simulations, we explore the static and dynamic properties of the water-density fluctuations in the interface of large nonpolar solutes. We show that the water interface is flickering, broadened by capillary-wave fluctuations. These fluctuations result in slow transitions between locally wet and dry regions. We also use simulations to explore the effect of surface roughness on the interfacial free energy. We find that molecular-scale roughness significantly increases the solid-liquid interfacial tension, an effect that can turn a hydrophilic smooth surface into a hydrophobic rough surface. The microscopic origins of these effects will be discussed.

COMP 44

Excited state dynamics of conjugated polymers
Understanding how the molecular-level nuclear structure and dynamics of conjugated polymers and organic donor/acceptor blends are linked to the critical elements of excitation migration, radiative processes, and charge separation appears to be imperative if we are to develop a working chemical intuition about organic electronic devices. In this talk, I will describe progress in using a mixed quantum/classical non-adiabatic molecular dynamics approach that employs an all-atom description of the intermolecular interactions coupled to a semi-empirical (PPP) electronic Hamiltonian. Results exploring several systems at ambient temperature will be discussed, including phenylene-vinylene oligomers (OPV), their intermolecular coupling, and their response to external fields, as well as the coupling of an OPV with an aggregated fullerene cluster.

COMP 45

Local molecular field analysis of hydrogen bonding patterns for water at hydrophilic and hydrophobic silica surfaces

John D. Weeks\(^{(1)}\), jdw@umd.edu, 1109 IPST Building, College Park MD 20742, United States; Richard C. Remsing\(^{(1)}\). (1) Department of Chemistry and Biochemistry, University of Maryland, College Park MD 20742, United States

The structure and dynamics of a slab of SPC/E water in simultaneous equilibrium with its vapor and with either a hydrophobic or hydrophilic (fully hydroxylated) silica surface are studied by molecular dynamics simulations.

Singlet and correlated structural properties at both surfaces are analyzed in detail and compared to those of the liquid-vapor interface. We further examine the static and dynamic hydrogen bonding behavior at the silica surfaces, finding some similarities in the structure, but vast differences in dynamics. This analysis is aided by use of Gaussian-smoothed charged distributions as suggested our recently-developed Local Molecular Field theory.

COMP 46

From water structure and fluctuations to hydrophobicity of proteins and interfaces

Shekhar Garde\(^{(1)}\). (1) Department of Chemical and Biological Engineering, Rensselaer Polytechnic Institute, Troy NY 12180, United States

Water-mediated interactions (e.g., hydrophobic interactions) govern a host of biological and colloidal self-assembly phenomena from protein folding, and
micelle and membrane formation, to molecular recognition. Macroscopically, hydrophobicity is often characterized by measuring droplet contact angles. Such measurements are not feasible for nanoscale surfaces of proteins or nanoparticles. How does one then characterize hydrophobicity/philicity of such interfaces? We present results from theory and simulations of hydration of a variety of surfaces to connect the behavior of water at the nanoscale interfaces and their hydrophobicity. Specifically, we show that water density fluctuations (and not the average local density) provide a quantitative characterization of the interface hydrophobicity. Density fluctuations are enhanced at hydrophobic interfaces and suppressed near hydrophilic ones. Simulations also show how properties of water at interfaces influence solute binding, folding, and dynamics of flexible molecules at interfaces. I will demonstrate that this new perspective on hydrophobicity provides a tool for characterization of hydrophobicity patterns on protein surfaces, which are relevant for binding, recognition, and aggregation.

COMP 47

Statistical thermodynamics of real liquids: Molecular quasi-chemical theory applied to liquid water

Lawrence R. Pratt\(^{(1)}\), lpratt@tulane.edu, 300 Lindy Boggs Center, Department of Chemical and Biomolecular Engineering, 333 Lindy Boggs Center, New Orleans LA 70118, United States. (1) Chemical & Biomolecular Engineering, Tulane University, New Orleans LA 70118, United States

We implement and test a molecular quasi-chemical theory of liquid water. The quasi-chemical perspective is that several physically transparent contributions are involved in a sensitive balance. Network-liquid contributions are precisely defined, and long-ranged interactions are characterized by a conditional distribution of binding energies. For simulated liquid water network-liquid contributions are small (even zero), and the free energy is dominated by long-ranged interactions. The binding-energy distributions for liquid water are observed to be unimodal, and a Gaussian model can be accurate. A broad distribution of interactions remains as the essential difficulty of the molecular theory liquid water.

COMP 48

Investigating molecular mechanisms of specificity in regulation of the HER2 and HER4 tyrosine kinases

Shannon E. Telesco\(^{(1)}\), shannone@seas.upenn.edu, 1509 Pine St, Philadelphia PA 19102, United States; Ravi Radhakrishnan\(^{(1)}\); Rajanikanth Vadigepalli\(^{(2)}\). (1) Department of Bioengineering, University of Pennsylvania, Philadelphia PA 19104, United States (2) Department of Pathology, Anatomy and Cell Biology, Thomas Jefferson University, Philadelphia PA 19107, United States
Design of protein kinase inhibitors that maintain specificity for their targets remains a challenge in drug development. One such target is the ErbB family of tyrosine kinases, which includes EGFR, HER2, HER3 and HER4. Although the ErbB kinases share similar structural features, HER2 is implicated in 20-30% of human breast cancers whereas HER4 is associated with an anticarcinogenic function and a favorable clinical outcome. We perform molecular dynamics and free energy simulations of the HER2 and HER4 kinase structures to investigate the molecular basis for their divergent effects in the mammary gland [Telesco SE and R. Radhakrishnan. 2009. Biophys J 96(6)]. An extensive hydrogen bonding network between the A-loop and C-loop correlates with their concerted motions as revealed by principal component analysis, which may promote alignment of the catalytic residues. Despite these shared dynamical patterns, we postulate key differences in the dimer-mediated activation mechanism of the kinases. Hence we perform umbrella sampling simulations of HER2 and HER4 dimers to compare the conformational changes that occur along the activation pathways. Since there is a growing effort to develop pharmacological compounds that target HER2 or exploit the anticarcinogenic function of HER4, it is crucial to elucidate the molecular properties that confer distinctiveness in kinase regulation.

**COMP 49**

**Riboswitches in motion: Computational and experimental approaches**

*Paula M Petrone*\(^{(1)}\), paula.petrone@novartis.com, 100 Technology Square, Cambridge MA 02139, United States; *Andrea Pomerantz*\(^{(1)}\); *Janetta Dewhurst*\(^{(1)}\); *Ruben Tommasi*\(^{(1)}\); *Lewis Whitehead*\(^{(1)}\). (1) Novartis Institutes for Biomedical Research, Cambridge MA 02139, United States

Riboswitches are ‘intelligent’ elements of mRNA capable of adopting mutually exclusive conformations triggered by an environmental signal in the form of temperature changes or the binding of a small molecule. Conformational
changes in the riboswitch interfere with the structure of downstream mRNA or ribosomal translation of the genetic message, leading to either premature transcription termination or inhibition of initiation, effects which can then be propagated down the corresponding metabolic pathway. By binding its cognate ligand, a riboswitch functions as a key component of a delicate feedback mechanism that kinetically controls the correct concentration of that specific ligand in the cellular environment.

Using a combination of computational and experimental techniques, we elucidate the mechanisms leading to ligand-driven conformational change, and we present a characterization of the unbound state of the riboswitch PreQ1. We also address whether there is a ligand-binding pocket in the apo-structure or, alternatively, the riboswitch responds to an induced-fit model. With molecular dynamics simulations, we focus on the unbound riboswitch and the role of ions and water in the mechanism of unbinding and stabilization of the apo-state. We compare computational predictions to fluorescence spectroscopy measurements of 2-aminopurine-labeled PreQ1 riboswitch sequences to gain a more detailed understanding of ligand-driven changes in the riboswitch structure.

COMP 50

Differential dynamics coupling of plexin GTPase complexes in MD Simulations

Liqun Zhang(1), lxz79@case.edu, 10900 Euclid Ave., Cleveland OH 44106, United States ; Matthias Buck(1). (1) Physiology and Biophysics, Case Western Reserve University, Cleveland OH 44106, United States

Plexin receptors interact directly with Ras/Rho family small GTPases which are involved in cell adhesion and motility. Using NAMD molecular dynamics simulations, the binding of plexin and Rac1/Rnd1 GTPase complexes is investigated. We have found that different protonation states of key Histidine residues in
plexin lead to two different modes of binding with Rac1 using HADDOCK docking simulations. One structure is similar to the homologous RBD:Rnd1 structure solved by the X-ray crystallography, while the other structure has plexin rotated by 60 degrees relative to the first one. Structures from both complexes are stable over 30 ns unrestrained molecular dynamics simulations. Comparison of the internal dynamics of the RBD-GTPase complexes reveals distinct signatures, suggesting a specific dynamical coupling between RBD and different GTPases. The extent of the cross-correlations is more marked than those seen in the GTPases alone, suggesting an allosteric signature, as proposed by a statistical coupling analysis.

COMP 51

Computational study of the ammonium transport mechanism in AmtB: NH₃ permeation vs. NH₃/H⁺ Co-transport

Shihao Wang(1), shihao@gmail.com, 7141 Sherbrooke Street West, SP 265.25, Montreal Quebec H4B 1R6, Canada; Guillaume Lamoureux(1). (1) Department of Chemistry and Biochemistry, Concordia University, Montreal Quebec H4B 1R6, Canada

The transmembrane protein AmtB has an important role in ammonium transport, especially at low external ammonium concentrations. However, whether AmtB is just a channel that permeates NH₃ or a co-transporter that transports NH₃ and H⁺ at the same time is still a question. To study the mechanism of ammonium transport through AmtB, an extensive series of hybrid Quantum Mechanical(QM)/Molecular Mechanical(MM) simulations has been performed. Emphasis has been placed on the deprotonation mechanism of ammonium. Results of Constraint Dynamics simulations have been combined to obtain the Potentials of Mean Force for two possible deprotonation paths: One through protein side chains and the other through solvents. A comparison of the energy barriers reveals the more possible pathway. Statistical results on the solvent and ammonium distribution inside the pore is also shown and the possible mechanisms of ammonia re-protonation and how side chains return back to original state are also presented.

COMP 52

Atomistic perspective of metabolite specificity in the SAM-II riboswitch

Jennifer M Kelley(1), jkelley11@student.gsu.edu, P.O. Box 4098, Atlanta Ga 30302, United States; Donald Hamelberg(1). (1) Department of Chemistry, Georgia State University, Atlanta Georgia 30302, United States

Riboswitches are metabolite sensing motifs usually located in the 5' un-translated region of bacterial mRNAs. Metabolites bind to the aptamer domain of
riboswitches with amazing specificity, modulating gene regulation in a feedback loop as a result of induced conformational changes in the expression platform. Using all-atom molecular dynamics simulations, we show that the ensemble of conformations of the unbound form of the SAM-II riboswitch is a loose pseudoknot structure that periodically visits conformations similar to the bound form, and the pseudoknot structure is only fully formed upon binding the metabolite, S-adenosylmethionine (SAM). These conformational changes are triggered only by SAM, and not by a very similar metabolite, SAH, which differs from SAM by a methyl group and a single positive charge. Unlike the interaction of SAH with the SAM-II riboswitch, the binding of SAM is enthalpically driven and alters the curvature and base-pairing of the expression platform.

COMP 53

Linear interaction energy method for the prediction of protein stability changes upon mutation

Lauren Wickstrom(1), lauren.wickstrom@gmail.com, 610 Taylor Rd, Piscataway NJ 08854, United States; Emilio Gallicchio(1); Ronald Levy(1). (1) Department of Chemistry and Chemical Biology, and BioMaPs Institute for Quantitative Biology, Rutgers University, Piscataway NJ 08854, United States

Computational mutagenesis methods, ranging from statistical, empirical, and physical approaches, have been useful for the understanding and prediction of protein stabilities. Most existing methods are limited because they are unable to reproduce correctly the magnitude of the free energy change (ddG), the difference in free energy of unfolding between wild-type and mutant proteins (> 1.0 kcal/mol deviation from experimental ddG values). In order to overcome this problem, we are developing a computational approach which uses the protein local optimization program (PLOP) [1,2] and the linear interaction energy (LIE) method [3] to predict the changes in the free energy of the native state induced by a single mutation. Initial tests of the PLOP-LIE method have shown an average unsigned error between calculated and experimental ddG values of less than 1.0 kcal/mol. With further development of the LIE functional form and PLOP’s structural prediction capabilities, the PLOP-LIE method should be able to surpass most predictive programs available today.

COMP 54

Molecular mechanics and the evolution of enzyme activity

Brian N. Dominy(1), dominy@clemson.edu, 219 Hunter Laboratories, P.O. Box 340973, Clemson SC 29634-0973, United States; Manoj K. Singh(1); Allyn R. Brice(1); Craig Stanton(1). (1) Department of Chemistry, Clemson University, Clemson SC 29634, United States
Molecular mechanics-based techniques have a long history of successes in characterizing the physical properties of biomolecules that are responsible for their activity. Recently, molecular mechanics approaches have been extended to address questions involving the evolution of biomolecular function, particularly in relation to the development of drug resistance. We have recently applied numerous molecular mechanics-based techniques, using the CHARMM and AMBER packages, to study the effects of mutations on protein activity. The studies examined detailed thermodynamic properties of substrate and transition state binding, as well as the role of dynamics and correlated motions in protein activity. The systems we studied were varied and included HIV protease, beta-lactamase, monofunctional uracil glycosylase, and gastrodianin (an antifungal lectin). This presentation will attempt to both illustrate the variety of molecular mechanisms associated with mutated function in proteins, as well as describe the common mechanisms that emerged from the study of these diverse proteins.

COMP 55

FPGA implementation of cheminformatics and computational chemistry algorithms and its cost/performance comparison with GPGPU, cloud computing and SIMD implementations

Attila Berces(1), attila.berces@chemistrylogic.com, Szuk utca 9, Nyul Hungary 9082, Hungary; Bela Feher(2); Peter Szanto(2); Imre Pechan(2)(3); Laszlo Lajko(3); Zoltan Runyo(3); Peter Laczko(3); Janos Lazanyi(2). (1) Chemistry Logic Kft, Budapest 1116, Hungary (2) Dept. of Measurement and Information Systems, Budapest University of Technology and Economics, Budapest 1117, Hungary (3) evopro Kft, Budapest 1117, Hungary

We have developed binary fingerprint based similarity searching, topological torsional fingerprint based similarity searching, chemical library to library comparison, sphere exclusion and Jarvis Patrick clustering, peptide mass spectrometry fingerprinting, BLAST prefiltering, short read mapping in color space on Silicon Graphics RC100 FPGA card. In addition, we implemented the Autodock docking software on FPGA. We reached 5 to 500 folds acceleration compared to CPU in these implementations. In this presentation the audience will learn what characteristics an algorithm should have to make it worthwhile to implement it on FPGA. We shall also compare the cost/performance characteristics to other alternatives such as cloud computing, GPGPU, and single-instruction-multiple-data (SIMD) optimization.

COMP 56

Technologies for desktop HPC: Application developer's perspective
In the last few years we have witnessed the emergence of a new computing paradigm: computational accelerators. Most prominent examples of such accelerators include FPGAs, Cell/B.E., and most recently GPUs. While these technologies bring unprecedented computing capabilities to the desktop users at a fraction of the cost of a traditional HPC system, their use comes with substantial difficulties due to the need for software reengineering. We survey the landscape of application accelerators for desktop systems and discuss the challenges of re-implementing computational chemistry applications on some of these systems using Hartree-Fock method and molecular dynamics codes as examples.

COMP 57

Faster, cheaper, and better science: Molecular modeling on GPUs

John E. Stone(1), johns@ks.uiuc.edu, 405 N. Mathews Ave., Urbana IL 61801, United States. (1) Beckman Institute, University of Illinois at Urbana-Champaign, Urbana IL 61801, United States

Over the past ten years graphics processing units (GPUs) have evolved from fixed-function single-purpose devices into highly programmable massively parallel co-processors. State-of-the-art GPUs support double-precision floating point arithmetic and achieve performance levels approaching one trillion floating point arithmetic operations per second. Modern GPUs enable software development in dialects of familiar C, C++, and Fortran languages, and GPU acceleration extensions exist for Python, Matlab, and other popular languages and computing tools. The high performance of GPUs has created opportunities for acceleration of many computationally demanding molecular modeling algorithms that contain significant parallelism.

We will describe how GPUs are currently employed to accelerate some of the most computationally demanding tasks involved in molecular dynamics simulation, visualization, and analysis in our NAMD and VMD software, and give an overview of how GPUs are expected to evolve in the next few years.

COMP 58

Folding@home: Petaflops on the cheap today, exaflops soon?
Over the last 10 years, Folding@home has emerged as a very powerful resource. Today, it has multi-petaflop performance, making it the most powerful supercluster in the world. I will talk about how Folding@home works, both in terms of infrastructure and algorithms, and how one can easily reproduce these sorts of approaches in your own lab. I will also very briefly touch on recent results from Folding@home to highlight what petascale power can do to dramatically change the nature of what simulations can inform us about systems of interest.

COMP 59

Protein-ligand docking on the Cell/BE processor with eHiTS Lightning

Zsolt Zsoldos(1), zsolt@simbiosys.com, 135 Queen[sapos]s Plate Drive, Suite 520, Toronto Ontario M9W6V1, Canada ; Orr Ravitz(1). (1) SimBioSys Inc., Toronto Ontario M9W6V1, Canada

The eHiTS flexible docking has proven to be among the most accurate pose prediction tools (http://www.simbiosys.ca/ehits/ehits_validation.html) providing one of the highest enrichment factors based on comparative evaluation studies (http://www.simbiosys.ca/ehits/ehits_enrichment.html). The accurate results of eHiTS have been achieved at the price of longer CPU times in the past, but that has changed with the recent port of the algorithm to the Cell/BE processor (http://www.bio-itworld.com/issues/2008/july-august/simbiosys.html). The revolutionary hardware that powers RoadRunner (the world's current fastest supercomputer) and also available in the low cost SONY PS3 game console, gives eHiTS 30-50 fold speedup compared to a single core Intel/AMD processor. The advantages of the Cell/BE platform over other acceleration techniques (FPGA,GPGPU) will be described, along with the challenges faced during the porting effort. A new proximity data structure is introduced that is optimized for SIMD architectures. It allows efficient evaluation of short range pairwise interactions with optimum cache locality.

COMP 60

Fragment-based druggable hot spot identification in proteins and protein-protein interactions using HPC

Dima Kozakov(1), midas@bu.edu, 44 Cummington Street, Boston MA 02215, United States ; Sandor Vajda(1); Gwo Yu Chuang(1); Ryan Brenke(1); David R Hall(1); Dmitri Beglov(1). (1) Department of Biomedical Engineering, Boston University, Boston MA 02215, United States
Here we present a highly parallel FFT-based method FTMAP for performing computational fragment mapping. Mapping methods place molecular probes on the surface of proteins in order to identify the most favorable binding positions. Since regions of the protein surface that are major contributors to the binding free energy in drug-protein interactions also bind a variety of small organic molecules, mapping can identify such “hot-spots” and the number of probe molecules bound is a good predictor of druggability. The highly parallel nature of our FFT-based approach allows it to be fully scalable, running efficiently on everything from desktop machines with CUDA enabled graphics adapters to an IBM Blue Gene. The method has been applied to both canonical and protein-protein interaction drug targets, successfully predicting binding hot-spots and target druggability. Our public web server is gaining popularity among academic users and generating significant interest from industry.

COMP 61

GPUs: What is all the fuss about?

Brian Cole(1), coleb@eyesopen.com, 9 Bisbee Court, Suite D, Santa Fe NM 87508, United States; Bob Tolbert(1); Anthony Nicholls(1). (1) OpenEye Scientific Software, Santa Fe NM 87508, United States

High performance computing hardware is undergoing a revolution. The best way to achieve increasing performance is through highly parallelized architectures like the graphics processing unit. However, the GPU requires a new assessment of algorithm design based on different memory versus time tradeoffs. Good performance is no longer gained by simply reducing the number of operations, but by organizing the interaction of those operations with a complex hierarchy of memory with varying latencies. Understanding the changing programming paradigm is critical both to selecting which algorithms will benefit from the GPU and how to achieve optimal performance. We will discuss design principles used when porting ROCS to the GPU. We will compare performance of a GPU implementation of ROCS to the highly-tuned production CPU implementation. We will show that higher performance can be achieved on the GPU at a significantly reduced cost compared to CPU clusters.

COMP 62

wwPDB: An organization that provides macromolecular structural data to the global community

Helen M Berman(1), berman@rcsb.rutgers.edu, 610 Taylor Road, Piscatway NJ 08854-8087; Gerard J Kleywegt(2), Haruki Nakamura(3), John L Markley(4). (1) Department of Chemistry & Chemical Biology, Rutgers The State University of New Jersey, RCSB PDB, Piscatway NJ 08854, United States (2) European Bioinformatics Institute, PDBe, Hinxton, United Kingdom (3) Institute for Protein
The Protein Data Bank (PDB) is the worldwide repository for experimentally determined 3D biological macromolecular structures; it currently contains approximately 65,000 entries. The Worldwide Protein Data Bank (wwPDB) was established in 2003 and consists of organizations that act as data deposition, processing, and distribution centers for PDB data.

wwPDB members have published policies and procedures for data processing with input from advisory committees, standing task forces, experimental method developers, and community experts. An overview of the projects undertaken by the wwPDB and how they impact our understanding of biomolecular structure will be presented.

wwPDB members are: RCSB PDB (supported by NSF, NIGMS, DOE, NLM, NCI, NINDS and NIDDK), PDBe (Wellcome Trust, EU, BBSRC, NIH and EMBL), PDBj (BIRD-JST) and BMRB (NLM).

COMP 63

wwPDB common annotation and deposition tool project

Martha Quesada(1), mquesada@rcsb.rutgers.edu, 610 Taylor Road, Piscatway NJ 08854, United States; John Westbrook(1); Tom Oldfield(2); Jasmine Young(1); Jawahar Swaminathan(2); Zukang Feng(2); Sameer Velankar(2); Marina Zhuravleva(1); Gaurav Sahni(2); Chenghua Shao(1); Takanori Matsuura(3); Eldon Ulrich(4); Steve Madding(4). (1) Department of Chemistry & Chemical Biology, Rutgers The State University of New Jersey, RCSB PDB, Piscatway NJ 08854, United States (2) European Bioinformatics Institute, PDBe, Hinxton, United Kingdom (3) Institute for Protein Research, Osaka University, PDBj, Osaka, Japan (4) University of Wisconsin, Madison, BMRB, Madison WI 53706, United States

Experimentally determined 3D structures of biological macromolecules play critical roles in understanding biological processes and in pharmacologically active molecule design and discovery. The wwPDB Common Deposition and Annotation Tool project will produce a set of common deposition and annotation processes and tools to support the goals of data quality and dependability for the next 10 years. In creating a single common process and tool set for use worldwide, the wwPDB is adopting best of breed components of the current systems and combining resources for the development of enhanced tools. The new system makes use of interactive interfaces and visualization tools to facilitate data curation and communication with the originating scientists. New modules supporting sequence alignment and ligand processing will be presented.
Small molecule resources and search tools at the PDB

Dimitris Dimitropoulos\(^ (1) \), ddimitrop@sdsc.edu, Pharmaceutical Sciences Building, 9500 Gilman Drive, Mailcode 0743, San Diego CA 92093, United States; John Westbrook\(^ (2) \), jwest@rcsb.rutgers.edu, 610 Taylor Road, Piscatway NJ 08854, United States. (1) SDSC & SSPPS, University of California, San Diego, RCSB PDB, San Diego CA 92093, United States (2) Department of Chemistry & Chemical Biology, Rutgers The State University of New Jersey, RCSB PDB, Piscatway NJ 08854, United States.

The Protein Data Bank archive contains over 12000 unique chemical components. These chemical components are described in a data dictionary maintained by the Worldwide Protein Data Bank (wwPDB; wwpdb.org). The wwPDB Chemical Component Dictionary, which includes IUPAC atom nomenclature for standard amino acids and nucleotides, stereochemical assignments, aromatic bond assignments, experimental model and computed ideal coordinates, systematic names and chemical descriptors, will be described.

Each of the wwPDB members provides web tools which leverage the content in the Chemical Component Dictionary. We present the tools provided by RCSB PDB for finding small molecules in the PDB and comparing their interactions within the macromolecular environment.

The RCSB PDB is managed by two members of the RCSB: Rutgers and UCSD, and is funded by NSF (DBI-0829586), NIGMS, DOE, NLM, NCI, NINDS, and NIDDK.

Remediation of PDB entries containing molecules with antibiotic and inhibitory pharmacological properties

Miriam (Miri) Hirshberg\(^ (1) \), miri@ebi.ac.uk, Wellcome Trust Genome Campus, Hinxton Cambridge CB10 1SD, United Kingdom; Shuchismita Dutta\(^ (2) \); Kim Henrick\(^ (1) \). (1) PDBe - Protein Data Bank Europe, EMBL-European Bioinformatics Institute, Hinxton Cambridge CB10 1SD, United Kingdom (2) RCSB - Protein Data Bank.
Micro-organisms produce bioactive secondary metabolites with an array of biological activities including antibiotics, antivirals, antitumours and toxins. These post-translationally modified gene products and non-ribosomal products display remarkable chemical diversity. The pharmaceutical industry has designed various synthetic peptides and peptide analogues to mimic natural substrates and inhibit specific targets. There are ~1000 Protein Data Bank (PDB) entries that contain such pharmacologically important molecules. Due to their size and chemical complexity, the PDB historically represented some of these molecules as small proteins, whereas others were treated as small ligands. The wwPDB has undertaken a project to address the inconsistencies in chemical representation, while including additional annotation and references to relevant external databases. A new dictionary has been created to hold data on sequence, chemistry, source, function and a complete SMILES string per entity. A system is being put into place to ensure all new depositions are also consistently annotated.

**COMP 66**

**Synergies of the PDB and the CSD**

*Tjelvar S. G. Olsson*(1), olsson@ccdc.cam.ac.uk, 12 Union Road, Cambridge Cambridgeshire CB2 1EZ, United Kingdom ; *Jason C. Cole*(1); *Simon Bowden*(1); *Colin R. Groom*(1). (1) Cambridge Crystallographic Data Centre, United Kingdom

The Cambridge Structural Database (CSD) contains the crystal structures of over 500,000 organic molecules. These provide a wealth of information on molecular geometry. In addition both intramolecular and intermolecular interactions can be searched and studied. The tremendous chemical diversity and high accuracy of these structures makes an ideal complement to the PDB. This is most evident when studying protein-ligand interactions, where the combination of knowledge embedded in both the PDB and the CSD transcends that of either resource on its own.

**COMP 67**

**Iridium: Prepping PDB data for use in computation chemistry software development**

*Gregory L Warren*(1), greg@eyesopen.com, 9 Bisbee Ct Suite D, Santa Fe NM 87508, United States ; *Thanh Do*(2); *Stephen D Warren*(2). (1) OpenEye Scientific
Computational chemistry, like any other predictive science, uses models of physical phenomenon to make predictions. The Achilles heel of every model is the quality of the experimental data upon which it is built or validated. Historically, in protein-ligand modeling too little time and attention has been paid to this important detail. We present here a protein-ligand structure database called Iridium, so called to reflect the scarcity of reliably curated information. The data in Iridium is from published sources, e.g. structures and electron density can be found in the RCSB; however, enormous time and effort have been taken in annotation and curation. For instance, all structures have been re-refined with MMFF94s used as the force field for the small molecule ligand. Structures are annotated with regard to the quality of the data and for approximately 90% of the structures binding affinity has been obtained from primary literature. We found that 19% of the ligands in this data set contained errors, e.g. incorrect interpretation of bond orders, element types and/or stereochemistry as well as a few instances of missing functional groups. The resultant database divides into three subsets: structure models that should never be used for software development or validation, structure models of questionable reliability or utility and structure models of high reliability. The database will be made freely available with the hope it will reduce the prevalence software with poor predictive capability.

COMP 68

Challenges in positioning, validation, annotation and visualization of chemical compounds bound to proteins

Ruben Abagyan(1). (1) Skaggs School of Pharmacy & PharmSciences, UC San Diego, La Jolla CA 92093, United States

Protein crystallography is a valuable source of information about small molecule ligands of biomacromolecules, experimental and approved drugs, and metabolites. However, atomic models of protein-ligand complexes derived from the crystallographic electron density often suffer from inevitable problems due to intrinsic limitations of the method, limited quality of the underlying data, inevitable ambiguities of data interpretation. Furthermore, the physical understanding of ligand-protein interactions including hydrogen bonding and electrostatic interactions requires the knowledge of protonation states and proton positions. Methods to overcome the obstacles and develop productive models predicting binding other ligands are presented. We also propose a platform for efficient visualization and dissemination of the ligand binding information to non-structural biologists developed in collaboration with the SGC Oxford. The author thank Irina Kufareva, Max Totrov, Eugene Raush, Lee Wen Hwa and Brian Marsden for their contributions.
Understanding of drug-target interactions and substrate binding to neuraminidase of influenza A virus subtypes H5N1 and H1N1-2009

Maturos Malaisree\(^{(1)}\), maturos_ros@hotmail.com, Chulalongkorn University, Phatumwan, Bangkok Bangkok 10330, Thailand; Thanyada Rungrotmongkol\(^{(1)}\); Nadtanet Nantaboot\(^{(2)}\); Adrian Mulholland\(^{(3)}\); Supot Hannongbua\(^{(1)}\). \(^{(1)}\) Department of Chemistry, Chulalongkorn University, Bangkok 10330, Thailand \(^{(2)}\) Department of Chemistry, Mahasarakham University, Mahasarakham 44150, Thailand \(^{(3)}\) School of Chemistry, University of Bristol, Bristol BS8 1TS, United Kingdom

The emergence of influenza A virus subtypes H5N1 and H1N1-2009 has raised concerns of the global flu pandemic. Understanding of the specific binding is the key to success in discovering the new potent neuraminidase (NA) inhibitors. To understand detailed information in the molecular level, molecular dynamics (MD) and QM/MM MD simulations were carried out. The dynamic nature of the three inhibitors, oseltamivir, zanamivir and peramivir, embedded in N1 has been observed. The peramivir system shows the lowest predicted binding affinity, in comparison with oseltamivir and zanamivir. In H274Y mutation strain, the reduction of the pocket size was found to be the source of the oseltamivir resistance. To understand the NA function, the modelled systems of H1N1-1918, H5N1-2005, H1N1-2009 and H2N2-1967 NA bound with SA-\(\alpha\)-2,6-Gal have been revealed the substrate binding in NA active site prior to the cleavage mechanism in the viral propagation step.

Structure-based lead optimization of small molecule Aurora A and CDK1 inhibitors

Deping Wang\(^{(1)}\), Deping.Wang@biogenidec.com, 14 Cambridge Center, Cambridge MA 02142, United States; Lin Zhang\(^{(2)}\); Jean-Yves le Brazidec\(^{(2)}\); Laura Silvian\(^{(1)}\); Doug Marcotte\(^{(1)}\); Junhua Fan\(^{(2)}\); Jer-Hong Chong\(^{(2)}\); Min Teng\(^{(2)}\); Shuo Zhao\(^{(2)}\); Khanh Nguyen\(^{(2)}\); Angela Pasis\(^{(2)}\); Betty Tam\(^{(2)}\); Christina Boykin\(^{(2)}\); Gisela Claassen\(^{(2)}\); Karen Lundgren\(^{(2)}\); Jianhua Chao\(^{(2)}\); Alexey Lugovskoy\(^{(1)}\). \(^{(1)}\) Drug Discovery, Biogen Idec, Cambridge MA 02142, United States \(^{(2)}\) Drug Discovery, Biogen Idec, San Diego CA 92122, United States

Mitotic kinases are the ultimate targets of pathways sensing genotoxic damage and affecting the cell cycle machinery. Many mitotic regulators are aberrantly expressed in tumor cells. These proteins could, therefore, make useful therapeutic targets. The kinases Aurora-A, -B and CDK1 represent such targets, and several Aurora kinases and CDK1 kinases have emerged as attractive targets for the design of anticancer drugs. In this work, we employed a variety of
computational methods, including molecular docking, quantum mechanical calculation and pharmacophore modeling, for structure-based lead optimization to identify novel potent dual Aurora A/CDK1 inhibitors. In addition, we elucidated a structural basis for designing selective Aurora A and B inhibitors.

COMP 71

Structure-based drug design of inhibitors of TNF-α converting enzyme (TACE)

Zhuyan Guo(1), zhuyan.guo@spcorp.com, 2015 Galloping Hill Road, Kenilworth NJ 07033, United States . (1) Department of Medicinal Chemistry, Merck Research Laboratories, Kenilworth NJ 07076, United States

TACE is a multi-domain zinc metalloprotease responsible for the shedding of membrane-bound TNF-a to its soluble active form. Overproduction of TNF-a has been implicated in a variety of inflammatory diseases such as rheumatoid arthritis. Current treatments include anti-TNF biologics such as Enbrel® and Remicade,® However, the advantages of an oral drug such as a potentially improved side effect profile and cost reduction make a small molecule drug highly sought after.

Through structure-based design, we have discovered novel TACE inhibitors of different chemotypes with hydantoin, carboxylate and hydroxamate as zinc binding groups. The compounds have sub-nanomolar potency, excellent selectivity against MMPs, and reasonable PK profiles. The design ideas and molecular modeling techniques that led to the discovery of these compounds will be discussed. Crystal structures of compounds were solved and confirmed the binding of these compounds to TACE.

COMP 72

Evaluation of DOCK6 for pose prediction and enrichment

Robert C Rizzo(1)(2), rizzorc@gmail, Math Tower Room 1-101, Stony Brook USA, United States ; Sudipto Mukherjee(1); Trent E Balius(1). (1) Department of Applied Mathematics and Statistics, Stony Brook University, United States (2) Institute of Chemical Biology & Drug Discovery, Stony Brook University, United States

In an effort to improve virtual screening methods, a dataset consisting of > 700 small molecule ligand-receptor complexes has been constructed from the Protein Databank for use in docking experiments. Statistics for reproducing crystallographically observed binding poses, using a newly modified version of the program DOCK6, will be presented which employed rigid (RGD), fixed anchor (FAD), or flexible (FLX) ligand docking protocols. The current study has focused on using on-the-fly flexible growth protocols in contrast to docking precomputed
conformers (flexibase approach). Total success rates and failures (sampling vs scoring) were also examined for individual protein families and for ligand flexibility (less than 7, 8 to 15, and 15+ rotatable bonds). Application to systems contained within the Astex 85 and UCSF DUD databases will also be presented.

COMP 73

Binding of the macrocyclic noncovalent inhibitor TMC435 to its HCV NS3/NS4A protease target

Maxwell D Cummings\(^{(1)}\), mcummin1@its.jnj.com, Gen De Wittelaan L 11B 3, Mechelen 2800, Belgium; Jimmy Lindberg\(^{(2)}\); Tse-I Lin\(^{(1)}\); Herman de Kock\(^{(1)}\); Oliver Lenz\(^{(1)}\); Elisabet Lilja\(^{(2)}\); Sara Fellander\(^{(2)}\); Vera Baraznenok\(^{(2)}\); Susanne Nystrom\(^{(2)}\); Magnus Nilsson\(^{(2)}\); Lotta Vrang\(^{(2)}\); Michael Edlund\(^{(2)}\); Asa Rosenquist\(^{(2)}\); Bertil Samuelsson\(^{(2)}\); Pierre Raboisson\(^{(1)}\); Kenneth Simmen\(^{(1)}\). (1) Tibotec BVBA, Mechelen 2800, Belgium (2) Medivir AB, Huddinge, Sweden

We are approaching a new era in the treatment of Hepatitis C infection, with the impending arrival of specifically targeted inhibitors of HCV NS3/4A protease and NS5B polymerase (STAT-Cs). At the leading edge of this paradigm shift are several peptidomimetic inhibitors of NS3/4A, currently in late stage clinical trials.

We have recently reported the crystal structure of the HCV protease inhibitor TMC435, currently in Phase 2b clinical trials, in complex with NS3/4A protease. The observed inhibitor binding mode involves an extensive network of intermolecular hydrogen bonds in the catalytic region, as well as an induced fit in an extended S2 subsite of the protease active site. These and other aspects of the new structure will be discussed. Comprehensive understanding of NS3/4A inhibitor binding and structure-based inhibitor design have been facilitated by a wealth of publicly available complex structures, and structural comparison of different NS3/4A inhibitors will also be presented.

COMP 74

Molecular modeling of the aryl hydrocarbon receptor and its interaction with ligands

Ashutosh S Jogalekar\(^{(1)}\), Ashutosh.Jogalekar@sanofi-aventis.com, 1041 Rt. 202/206, Bridgewater NJ 08807, United States; Stephan Reiling\(^{(1)}\); Roy J Vaz\(^{(1)}\). (1) Department of Structure, Design and Informatics, Sanofi-Aventis, Bridgewater NJ 08807, United States

The Aryl Hydrocarbon Receptor (AHR) is one of the principal xenobiotic receptors in living organisms that is responsible for interacting with several drugs and environmental toxins, most notably TCDD (tetrachlorodibenzodioxin). Binding of diverse ligands to AHR initiates an extensive set of downstream gene
expression responses and thus identifies AHR among a key set of proteins responsible for mediating interactions between living organisms and foreign molecules. While extensive biochemical investigations on the interaction of AHR with ligands have been carried out, modeling studies have been much more limited and studies validating specific computational algorithms on this protein are particularly absent. In this study we use molecular dynamics simulations to identify a physically realistic conformation of AHR that is relevant to ligand binding. We then use several docking programs and scoring functions combined with a post-docking MM-GBSA protocol to validate AHR binding with several independent sets of publicly available AHR ligands. The results identify an optimum set of protocols that could prove useful in future AHR ligand discovery and design. Exploration of the details of these protocols sheds light on factors operating in modeling AHR ligand binding.

COMP 75

Methodological advancements to improve the accuracy of induced-fit docking

Woody Sherman\(^{(1)}\), woody.sherman@schrodinger.com, 120 West 45th Street, 17th Floor, New York NY 10036, United States; Zhiyong Zhou\(^{(1)}\); Ramy Farid\(^{(1)}\). (1) Schrodinger, New York NY 10036, United States

Significant advancements in the treatment of protein flexibility in docking have been made in recent years. However, the work to date has typically been on small datasets that are not sufficient to show the robustness of the methodology across broad ligand and target classes. In this work, we describe our efforts to significantly expand our dataset used for induced-fit docking validation and improve the overall robustness and accuracy of the method. We describe the primary challenges and what we have done to address them. Specifically, we discuss the generation of reasonable initial poses, accurate refinement of the protein around those poses, and final scoring to choose the best ligand-receptor complex.

COMP 76

Water in protein binding sites: Consequences for ligand optimization

William L Jorgensen\(^{(1)}\), william.jorgensen@yale.edu, 225 Prospect Street, New Haven CT, United States; Julien Michel\(^{(1)}\); Julian Tirado-Rives\(^{(1)}\); James Luccarelli\(^{(1)}\). (1) Department of Chemistry, Yale University, New Haven CT 06520-8107, United States

An efficient molecular simulation methodology, JAWS, has been developed to determine the positioning of water molecules in the binding site of a protein or
protein-ligand complex. Occupancies and absolute binding free energies of water molecules are computed using a statistical thermodynamics approach. The importance of determining proper water occupancies is illustrated in Monte Carlo/free energy perturbation calculations for ligand series that feature displacement of ordered water molecules in the binding sites of scytalone dehydratase, p38-aMAP kinase, and EGFR kinase. The change in affinity for a ligand modification is found to correlate with the ease of displacement of the ordered water molecule. For accurate results, a complete thermodynamic analysis is needed. It requires identification of the location of water molecules in the protein-ligand interface and evaluation of the free energy changes associated with their removal and with the introduction of the ligand modification. Direct modification of the ligand in free-energy calculations is likely to trap the ordered molecule and provide misleading guidance for lead optimization.

COMP 77

Efficient method for computing the free energies of active site waters: Application to drug discovery

Charles H. Reynolds\(^{(1)}\), creynolds@ansar Isisbio.com, Four Valley Square, 512 East Township Line Rd., Blue Bell Pennsylvania 19422, United States; Jinming Zou\(^{(1)}\); Sia Meshkat\(^{(1)}\); Zenon Konteatis\(^{(1)}\); Anthony Klon\(^{(1)}\). \(^{(1)}\) Ansaris, Blue Bell Pennsylvania 19422, United States

Grand canonical Monte Carlo and systematic free energy methods have been reported previously that allow us to rapidly compute protein-fragment interaction energies. The same methodologies can be employed to compute free energies of binding for water. We have used this approach to identify critical waters in a number of therapeutically interesting protein active sites. Knowledge of the location and affinities of these waters can be useful for designing ligands with improved potency.

COMP 78

Using explicit solvent implicitly

Christopher J Fennell\(^{(1)}\), cfennell@maxwell.compbio.ucsf.edu, UCSF MC 2240, 600 16th St., San Francisco CA 94158-2517, United States; Charles W. Kehoe\(^{(2)}\); Ken A. Dill\(^{(1)}\). \(^{(1)}\) Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco CA 94158-2517, United States \(^{(2)}\) Graduate Group in Bioinformatics, University of California, San Francisco, San Francisco CA 94158-2517, United States

Solvent plays a critical role in biomolecular simulations. It mediates the transfer of small molecules, it bridges interactions between ligands and binding sites, it stabilizes protein structures with external hydrophilic groups and buried
hydrophobic cores, among others. When solvent is modeled explicitly in simulations, the microscopic interactions can be handled rigorously, but obtaining converged solvation energetics can be time–consuming. Here we describe a process, called Semi–Explicit Assembly, where we precompute the solvation response in simple systems and apply it in complex systems. We show that it is possible to have a detailed/explicit–like treatment of solvation at a computational cost similar to the fastest of implicit solvents.

COMP 79

Role of solvent in protein-ligand binding

Robert Abel(1), robert.abel@schrodinger.com, 120 West 45th Street, 17th Floor, New York NY 10036, United States ; Noeris Salam(1); Thijs Beuming(1); Woody Sherman(1); Ramy Farid(1). (1) Schrodinger Inc., New York NY 10036, United States

Calculation of protein-ligand binding affinities continues to be an active area of research. Although many techniques for computing protein-ligand binding affinities have been introduced, ranging from computationally very expensive approaches, such as free energy perturbation (FEP) theory to more approximate techniques, such as empirically derived scoring functions, which, although computationally efficient, lack a clear theoretical basis – their remains pressing need for more robust approaches. The recently introduced WaterMap technology, which calculates the locations and displacement free energies of hydration sites in proteins, was developed to bridge the gap between the accuracy of FEP and the computational efficiency of empirically derived scoring functions. In the present work, we apply WaterMap to a number of pharmaceutically relevant targets, and present a generalized approach for accurate predication of binding affinities that combines solvation terms from WaterMap with other important thermodynamic terms.

COMP 80

Compute the contribution of protein-pocket solvation to ligand-binding affinity by explicit water simulations

Ming-Hong Hao(1), ming-hong.hao@boehringer-ingelheim.com, 900 Old Ridgebury Rd., Ridgefield CT 06877, United States ; Ingo Muegge(1). (1) Department of Medicinal Chemistry, Boehringer Ingelheim Pharmaceuticals, Inc, Ridgefield CT 06877, United States

A significant fraction of ligand-binding free energy in proteins arises from the replacement of water molecules by the ligand in the binding site of proteins. Continuum solvation models based on surface areas do not treat the short-range correlations of water molecules well in the highly irregular and heterogeneous
protein-binding pocket. We have developed a computational procedure to simulate the density distribution and free energy of water molecules in the ligand-binding pocket of proteins using a molecular dynamics procedure (NAMD) with explicit water model (TIP3P). Our results are comparable with literature works (e.g. WaterMap software from Schrodinger Inc.) and show good agreement with crystallized water molecules observed in the X-ray structures of proteins. In our procedure, the distribution of water molecules in the protein-binding pocket is presented as water density on a 3-dimensional grid which we find to provide an intuitive way for visualizing the hydrophobic or polar characteristics of a binding site. The contribution of solvation to ligand-binding free energy is estimated by the difference of free energy of the pocket of water replaced by the ligand in the protein binding site and in the bulk solvent. This contribution is added to the direct ligand-protein interactions in scoring the binding affinity of ligands. We investigated the effects of residue mutations in protein binding-site on ligand binding affinity, including the Tryptophan mutations (W79F, W92F, W108A and W120A) in the high-affinity Streptavidin-Biotin complex and the drug-resistant mutants of HIV protease in complex with the inhibitor U-89360E. In these systems, X-ray crystallography showed no significant differences in the given protein-ligand complex structures between the wild type and mutant proteins. Intermolecular interactions between protein and ligand alone do not fully account for the changes in ligand-binding affinity. The free energy change of solvation in the binding site between wild type and mutants provides a good explanation for the shift in ligand-binding affinity. We also applied the procedure to study the structure-activity relationship of congeneric series of ligands. Our results suggest that binding-pocket solvation is an important factor in understanding the binding affinity of ligands to proteins.

COMP 81

All-atom explicit-solvent fragment-based drug discovery: SILCS ("Site Identification by Ligand Competitive Saturation") molecular dynamics simulations applied to IL-2

Olgun Guvench(1), oguvench@une.edu, 716 Stevens Ave., Portland ME 04103, United States . (1) Department of Pharmaceutical Sciences, University of New England College of Pharmacy, Portland ME 04103, United States

Two challenges in computer-aided drug discovery are incorporation of protein flexibility and an accurate description of solvation effects. Fast in silico screening methods typically employ rigid or near-rigid protein conformations and continuum descriptions of solvation, while more physical and accurate explicit-solvent all-atom molecular dynamics or Monte Carlo methods are very computationally demanding. Site Identification by Ligand Competitive Saturation (SILCS) is a recently-developed computationally-efficient fragment-based drug discovery method that employs all-atom
explicit-solvent molecular dynamics simulations, essentially soaking the target in a 1 molar bath of hydrophobic fragments to compute 3-D probability maps of hot-spots on the protein surface that preferentially bind hydrophobic fragments or water molecules. Applied to the apo crystal structure of IL-2, SILCS identifies two hydrophobic pockets not present in the apo crystal, but later discovered to exist in complexes with small molecule inhibitors and to bind hydrophobic moieties on these molecules.

COMP 82

Past, present, and future of semiempirical methods

James J. P. Stewart(1), MrMOPAC@OpenMOPAC.net, 15210 Paddington Circle, Colorado Springs CO 80921, United States. (1) Stewart Computational Chemistry, Colorado Springs CO 80921, United States

A highly personal overview of NDDO type methods will be given. These methods exemplify semiempirical quantum chemistry in that they show that a combination of theory, reference data, and parameter optimization can result in theoretical structures of useful accuracy. When incorporated as computer programs, the result has been a set of tools that have become widely used. The focus of this talk will be on the philosophical aspects, covering the history of these methods, their steadily-increasing accuracy and range of applicability, current problems and limitations, and speculations regarding the possible futures of such methods.

COMP 83

Monte Carlo study of the axial next-nearest-neighbor Ising model and other microphase forming systems

Patrick Charbonneau(1), patrick.charbonneau@duke.edu, 124 Science Drive, Durham NC 27708, United States; Kai Zhang(1). (1) Department of Chemistry, Duke University, Durham NC 27708, United States

Microphases self-assemble in systems with competing short-range attractive and long-range repulsive interactions, irrespective of the physical and chemical nature of these interactions. Microphases are the frustrated equivalent of gas-liquid coexistence for purely attracting particles. Periodic lamellae, cylinders, clusters, etc. are thus observed in a variety of systems, such as multiblock copolymers, oil-water surfactant mixtures, charged colloidal suspensions, and magnetic materials. Yet controlling the microphase morphology is notoriously difficult. Understanding the equilibrium behavior of model systems would help tune the modulated phases. But even for simple model systems, reliable results for the microphase regime are notoriously difficult to obtain, because of the presence of long-lived metastable states. We present a Monte Carlo simulation method based on thermodynamic integration that avoids this problem and with
which we obtain the microphase behavior of the canonical three-dimensional axial next-nearest-neighbor Ising model. Extensions to particle-based models are discussed.

**COMP 84**

**B2GP-PLYP and friends: Robust, generally applicable, double-hybrid functionals for thermochemistry, thermochemical kinetics, and weak interactions**

*Jan M.L. Martin*(1), gershom@weizmann.ac.il, Kimmelman Bldg., Rm. 252, Rechovot Central Region 76100, Israel. (1) Department of Organic Chemistry, Weizmann Institute of Science, Rechovot 76100, Israel

We report on the development and validation of fifth-rung "double-hybrid" exchange-correlation functionals, which combine a hybrid of (meta-)GGA and HF-type exchange with a mixture of (meta-)GGA and MP2-type correlation. By judicious parametrization, functionals can be found that render superior performance (comparable to composite ab initio methods like G3 theory) for both atomization energies and barrier heights, and that do not share the deficiencies of other "kinetics-friendly" functionals for late transition metal reactions. Performance for weak interactions, while already good, can be further improved by adding an empirical dispersion correction. Performance for vibrational frequencies is intermediate between conventional DFT methods and CCSD(T). Our B2GP-PLYP functional [1] is available in several popular quantum chemistry programs without recoding. Prospects for further improvement will be discussed.


**COMP 85**

**Exploring formation, spectroscopy, and outcome of disulfide radical anions by QM/MM approaches: Insights into a molecular rheostat**

*Elise Dumont*(1), elise.dumont@ens-lyon.fr, 46 allee d'Italie, Lyon Cedex 07 69364, France; Adele D. Laurent(2); Pierre-Francois Loos(2); Xavier Assfeld(2). (1) Department of Chemistry, Ecole Normale Superieure de Lyon, Lyon 69364, France (2) Equipe de Chimie et Biochimie Theoriques, Universite Nancy I, Nancy 54506, France

Disulfide linkages versatility is beautifully exemplified by anti-oxidant enzymes of the Thioredoxin superfamily, embracing a large range of redox potentials.
Radical anions have been identified as key intermediates in the disulfide:dithiol redox mechanism.

We have built up a hybrid methodology specifically tailored to describe electron attachment on biomolecules. Key factors tuning the one-electron addition are highlighted, through inspection of disulfide-bridged systems of increasing complexity [figure 1]. In our scheme, inner competition can be treated, revealing differences between structural vs. redox-active linkages: an analogy between electron affinity and redox potential is drawn. Finally, UV-Vis signature and photodissociation are also considered, in regard to possible biomimetic applications.

Illustration: Huge modulation arising from electrostatic and geometric contributions...

COMP 86

Accelerating self-consistent field convergence with the augmented Roothaan–Hall energy function

Xiangqian Hu\(^{(1)}\), xqhu@duke.edu, 124 Science Dr., Durham NC 27708, United States; Weitao Yang\(^{(1)}\). (1) Department of Chemistry, Duke University, Durham NC 27708, United States

Based on Pulay's direct inversion iterative subspace (DIIS) approach, we present a method to accelerate self-consistent field (SCF) convergence. In this method, the quadratic augmented Roothaan-Hall (ARH) energy function, proposed recently by Host and coworkers [J. Chem. Phys. 129, 124106 (2008)], is used as the object of minimization for obtaining the linear coefficients of Fock matrices within DIIS. This differs from the traditional DIIS of Pulay, which uses an object function derived from the commutator of the density and Fock matrices. Our results show that the present algorithm, abbreviated ADIIS, is more robust and
efficient than the energy-DIIS (EDIIS) approach. In particular, several examples demonstrate that the combination of ADIIS and DIIS ("ADIIS+DIIS") is highly reliable and efficient in accelerating SCF convergence.

**COMP 87**

**Computer algorithm development for modeling of biological macromolecules**

*Dusanka Janezic(1)(2)*  
dusa@cmm.ki.si, Hajdrihova 19, Ljubljana 1000, Slovenia.  
(1) Laboratory for Molecular Modeling, National Institute of Chemistry, Ljubljana 1000, Slovenia  
(2) FAMNIT, University of Primorska, Ljubljana 1000, Slovenia

This presentation focuses on our recent efforts to develop new methods for modeling of biological macromolecules.

In particular, new **Split Integration Symplectic Method** for the numerical solution of molecular dynamics equations, new methods for the determination of vibrational frequencies and normal modes of large systems, and the new **Distributed-Diagonal Force Decomposition Method**, a parallel method for molecular dynamics simulation will be described.

We propose an analytical treatment of the internal high-frequency molecular vibrations in the molecular dynamics simulations using a new form of the classical Liouville propagator. The essence of the work lays in the construction of the second-order integrating algorithm – the **Split Integration Symplectic Method** that is useful for all-atom molecular dynamics simulations of molecular systems described by flexible models. We have developed a computer program for molecular dynamics simulation that implements the **Split Integration Symplectic Method** and is designed to run on specialized parallel computers. We have also developed a parallel method for molecular dynamics simulation, the **Distributed-Diagonal Force Decomposition Method**. Compared to other methods its
communication requirements are lower and it features dynamic load balancing, which increase the parallel efficiency.

The proposed methodological improvements significantly extend the scope of presently used algorithms in terms of length- and time-scales and thus contribute to the general applicability of molecular dynamics simulation algorithms. The simulation results of selected examples will be presented.

By using the theory of protein graphs, we have developed a new approach to finding the protein areas that are functionally important. The approach is ideally suited to finding protein-protein binding sites. The newly-developed method **ProBiS (Protein Binding Sites)** enables the detection of structurally similar binding sites on the query protein. The developed algorithm compares the query protein structure sequentially with all the protein structures in a database of about 24,000 protein structures and identifies, within a few minutes to about an hour, all the query's locally similar structures in this database. On the basis of these aligned structures the algorithm then calculates the conservation scores for all amino acids of the query protein and projects them as colours (or discrete values) onto the surface of the examined protein structure. Due to the challenging computation involved, the comparison of such a number of proteins at a local level and in such a short time has until now been impossible even by using higher performance computer systems. The ProBiS algorithm for detection of structurally similar binding sites in proteins is freely available as a web-tool. Detailed explanation and instructions to users of ProBiS can be found at http://probis.cmm.ki.si.

**COMP 88**

**Common catalytic features among cofactor independent racemases and epimerases: A computational perspective**

**Marco Stenta**(1), marco.stenta@epfl.ch, EPFL SV IBI1 UPDALPE, Office AAB019 (Bâtiment AAB) Station 15, Lausanne Vaud 1015, Switzerland ; **Piero Altoè**(2); **Matteo Dal Peraro**(1); **Marco Garavelli**(2); **Andrea Bottoni**(2). (1) Department of Bioengineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne Vaus 1015, Switzerland (2) Dipartimento di Chimica, Università degli Studi di Bologna, Bollgona 40100, Italy

D-aminoacids play a key role in several biological processes as intermediates, energy source or structural elements. Many pathogenic organisms use specific racemases and epimerases to obtain D-aminoacids from natural L-isomers. Several of those enzymes constitute potential targets for rational drug design, as they are not shared with mammalians. Two enzymes, belonging to the cofactor-independent family, have been studied to understand the common catalytic features of the whole class. Advanced computational techniques based on an accurate QM/MM potential have been employed to understand the reaction mechanism and catalytic proficiency of both *Trypanosoma cruzi* proline racemases and *Haemophilus influenzae* diaminopimelate epimerase. The role of
each residue in the catalytic process has been individuated. Despite consistent
differences in the active site geometry, both enzymes share the same transition
state stabilization mechanism.


COMP 89

Signature of quantum coherent effects

Jianshu Cao(1), jianshu@mit.edu, 77 Massachusetts Ave., Cambridge MA 02139, United States. (1) Department of Chemistry, MIT, Cambridge MA 02139, United States

This talk discusses coherent effects in energy transfer efficiency in light-
harvesting systems and vibrational response of anharmonic systems.

(1) To clarify the optimal conditions in photosynthetic systems and explore the
optimal design principles for artificial energy transfer devices, we calculate the
exciton dynamics in model systems and in FMO systems. These calculations
suggest that noise-enhanced energy transfer is an intrinsically quantum coherent
effects and that spatial-temporal correlations of environmental fluctuations can
help maintain coherence in transport processes.

(2) The simple classical limit of quantum response functions, often used in MD
simulations of large systems, leads to divergence. We discuss the origin of this
divergence and establish a correspondence between quantum transitions and
classical trajectories by means of phase space quantization. We introduce the
concept of crossover time to characterize quantum recurrence and to examine
the validity of quasi-classical dynamics.

COMP 90

Quantum dynamics of excitation energy transfer in nanostructured
environments

David F. Coker(1)(2), david.coker@ucd.ie, University College Dublin, Belfield
Dublin Dublin 4, Ireland; Pengfei Huo(1); Tina Rivera(2); Daniel Montemayor(2). (1)
Department of Chemistry, Boston University, Boston MA 02215, United States
(2) School of Physics, University College Dublin, Boston MA 02215, United States

Quantum dynamical calculations are used to explore excitation energy transfer in
model photosynthetic antenna complexes. Specifically, the role of correlated
environmental motions on long lived excited state coherent dynamics will be
explored. The models are parameterized by large scale classical simulations (e.g. see figure) that include environmental effects. We study the effect of hybridizing natural light harvesting systems with semiconductor solar cell materials.

![Image of a molecular structure](image)

**COMP 91**

**Can impact excitation explain carrier multiplication in carbon nanotube photodiodes?**

*Eran Rabani*\(^{(1)}\), rabani@tau.ac.il, Ramat Aviv, Tel Aviv 69978, Israel. \( (1) \) School of Chemistry, Tel Aviv University, Tel Aviv 69978, Israel

Carrier multiplication is a process where several charge carriers are generated upon the absorption of a single photon in semiconductors. This process is of great technological ramifications for solar cells and other light harvesting technologies. For example, it is expected that when more charge carriers created shortly after the photon is absorbed, the larger fraction of the photon energy can successfully be converted into electricity, thus increasing the device efficiency.

In this talk we will address recent experiments reporting extremely efficient carrier multiplication in carbon nanotube photodiodes at photon energies near the carrier multiplication threshold (twice the quasi-particle band gap). This result is surprising in light of recent experimental and theoretical work on multiexciton generation in other confined materials, such as semiconducting nanocrystals. We propose a possible mechanism based on carrier dynamics leading to impact excitation. We discuss the important time-, energy-, and length- scales of the problem and provide analysis of temperature and gate voltage effects.

**COMP 92**
Electronically non-adiabatic dynamics via semiclassical initial value methods

William H Miller\(^{(1)}\), millerwh@berkeley.edu, Department of Chemistry, Berkeley California 94720, United States Virgin Islands. (1) Department of Chemistry, University of California, Berkeley, Berkeley California 94720, United States Virgin Islands

In the late 1970's Meyer and Miller (MM) [J. Chem. Phys. 70, 3214 (1979)] presented a classical Hamiltonian corresponding to a finite set of electronic states of a molecular system (i.e., the various potential energy surfaces and their couplings), so that classical trajectory simulations could be carried out treating the nuclear and electronic degrees of freedom (DOF) in an equivalent dynamical framework (i.e., by classical mechanics), thereby describing non-adiabatic dynamics in a more unified manner. Much later Stock and Thoss (ST) [Phys. Rev. Lett. 78, 578 (1997)] showed that the MM model is actually not a 'model', but rather a 'representation' of the nuclear-electronic system; i.e., were the MMST nuclear-electronic Hamiltonian taken as a Hamiltonian operator and used in the Schrödinger equation, the exact (quantum) nuclear-electronic dynamics would be obtained. In recent years various initial value representations (IVRs) of semiclassical (SC) theory have been used with the MMST Hamiltonian to describe electronically non-adiabatic processes. Of special interest is the fact that though the classical trajectories generated by the MMST Hamiltonian (and which are the 'input' for an SC-IVR treatment) are 'Ehrenfest trajectories', when they are used within the SC-IVR framework the nuclear motion emerges from regions of non-adiabaticity on one potential energy surface (PES) or another, and not on an average PES as in the traditional Ehrenfest model. Examples are presented to illustrate and (hopefully) illuminate this behavior.

COMP 93

Very accurate semiclassical surface hopping calculations

Michael F. Herman\(^{(1)}\), mherman@tulane.edu, New Orleans LA 70118, United States. (1) Department of Chemistry, Tulane University, New Orleans LA 70118, United States

Many problems in chemistry involve processes in which atoms or molecules undergo transitions between electronic quantum states. A semiclassical surface hopping expansion of the multistate wave function and propagator is discussed.
The surface hopping expansion is capable of obtaining very accurate results even for cases in which interference between significantly different paths plays an important role. It is shown that hops in the classically forbidden region must be included at low energies to maintain the high level of accuracy. High accuracy is obtained even in cases where the transition probability is very small because the transition is strongly classically forbidden. A relatively simple one dimensional approximation, which uses only information obtained near the classical turning point in the motion, is also shown to provide accurate results for these forbidden transitions. A form of the surface hopping expansion, which is free of semiclassical singularities, will also be discussed.

COMP 94

Chemical reaction rates from ring polymer molecular dynamics

David E Manolopoulos(1), david.manolopoulos@chem.ox.ac.uk, PTCL, South Parks Road, Oxford Oxfordshire OX1 3QZ, United Kingdom . (1) Department of Chemistry, Oxford University, Oxford Oxfordshire OX1 3QZ, United Kingdom

In this talk, I will first review the ring polymer molecular dynamics (RPMD) theory of chemical reaction rates, and then discuss some recent developments which enable us to improve on the RPMD rate coefficient in the deep quantum tunnelling regime.

COMP 95

Multiscale models from molecules to life

Yiannis N. Kaznessis(1), yiannis@cems.umn.edu, 421 Washington Ave SE, Minneapolis MN 55455, United States . (1) Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis MN 55455, United States

A vexing question in the biological sciences is the following: how does life emerge from a soup of chemicals? At the crux of the matter lies the doubt that humans can develop faithful mathematical representations of biological dynamics.

We believe that a synthetic biology research programme may liberate empiricism beyond the unaided human brain. Humans can now construct and piece together DNA sequences in order to design new biological systems. Synthetic biology is the discipline that focuses on the construction of these novel biological systems. These systems confer advantages that may indeed help us make a plausible case for so ardent a vision, as to describe biology with mathematics. In the presentation we will explore the tractability of models of synthetic biological systems that are founded on universal laws of chemical thermodynamics and
principles of molecular biology.

COMP 96

Many-scale and multi-scale modeling in computational nanoscience

Peter T Cummings(1)(2), peter.cummings@vanderbilt.edu, VU Station B 351604, 2301 Vanderbilt Place, Nashville TN 37235-1604, United States . (1) Department of Chemical Engineering, Vanderbilt University, Nashville TN 37235-1604, United States (2) Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Nashville TN 37235-1604, United States

Theory, modeling and simulation (TMS) tools constitute key enabling technologies for making fundamental advances in nanoscience and for making nanotechnology a practical reality. Many of the problems encountered in this field are inherently multiscale. In this talk, we provide an overview of the role of TMS in nanoscience, as well as an overview of our many-scale and multiscale TMS research in nanotribology, molecular electronics, and hybrid organic-inorganic nanocomposites.

COMP 97

Coarse-grained and multiscale models of bulk liquids and macromolecules

Teresa Head-Gordon(1), TLHead-Gordon@lbl.gov, 274 Stanley Hall, Berkeley California 94720, United States . (1) Department of Bioengineering, University of California, Berkeley, Berkeley California 94720, United States

I will describe two coarse-grained models and a multiscale model of bulk liquids, aqueous macromolecular solutions and membrane lipid environments. We have recently achieved a fundamental result in deriving an analytical solution for computing the screened electrostatic interaction between arbitrary numbers of proteins of arbitrarily complex charge distributions, assuming they are well described by spherical low dielectric cavities in a higher dielectric salty medium. Smooth and systematic increase or decrease in spatial resolution back and forth between simple dielectric cavities and atomic level descriptions is the centerpiece of a multiscale scheme. I will also describe a coarse-grained model of water to investigate thermodynamic-dynamic relationships as well as a coarse-grained lipid and protein model relevant for lengthscales and timescales relevant for disease aggregation.

COMP 98

Examining methods for general multi-scale modeling using CHARMM
Bernard R Brooks\(^{(1)}\), brb@nih.gov, 5635 Fishers Lane, MSC 9314, Rockville MD 20892-9314, United States . (1) Laboratory of Computational Biology, NHLBI/NIH, Rockville MD 20892-9314, United States

This presentation focuses on our recent efforts to develop general multi-scale macromolecular modeling methods and to apply them to problems of examining protein dynamics and function. One objective in developing multi-scale modeling techniques is to be able to include multiple scale representations within a single study. By combining scales within a single calculation, one can examine properties that would be difficult or too costly to examine with a single model. Examples of modeling scales that can be connected and combined with the “multiscale” command within CHARMM include:

- Grid based map objects (EMAP)
- Coarse-graining using elastic network models (ENM) and Langevin network models (LNM)
- Coarse-grained models using Hydrophobic/Hydrophobic/Neutral models (BLN)
- Atomic models using a classical force field (CHARMM, Amber,…)
- Models employing a quantum mechanical subsystem (CHARMM/Q-Chem,…)

Experiences and examples of newer methods building and expanding upon the multiscale command to be presented and discussed include:

- Structural analysis via single particle electron tomography
- Vibrational subsystem analysis and other newer Hessian based methods
- Examining reaction pathways and free energies
- Rapid exploration of local conformational space using self guided Langevin dynamics
- Newer replica-exchange techniques for enhanced sampling and pKa prediction
- Using CSA (Conformational Simulated Annealing) methods with multiscale modeling

COMP 99

ErbB receptor mediated oncogenic signaling: Molecular systems biology through multi-scale modeling and high-performance computing

Ravi Radhakrishnan\(^{(1)}\), rradhak@seas.upenn.edu, 210 S 33 Street, 240 Skirkanich Hall, Philadelphia PA 19104, United States . (1) Department of Bioengineering, University of Pennsylvania, Philadelphia PA 19104, United States

We focus on an important cell-signaling pathway whereby the activation of the ErbB family growth factor receptors on the cell membrane lead to important cellular decisions such as cell-proliferation, cell-death, or cell-migration. We
employ multiscale simulation approaches to incorporate a molecular context to the signaling pathway involving this receptor to capture how subtle differences in the molecular context of the intracellular environment (e.g., differences in phosphorylation states or closely related mutant proteins can nevertheless translate into crucial differences in the manifestations of the emergent signaling and trafficking responses and cell decisions. Specifically, (1) we employ atomistic models to gain insight into the activation mechanism of the growth factor receptor and how it initiates signaling. (2) We also employ a recently developed a coarse-grained methodology for combining membrane mechanics and dynamics using the time-dependent Ginzburg-Landau (TDGL) formalism together with the stochastic kinetic Monte Carlo (or KMC) based dynamics of curvature inducing proteins for modeling how the receptor endocytosis. Mechanisms (1) and (2) represent activating and deactivating steps in the signaling pathway and a balance between the two is essential for cellular homeostasis. We show how point mutations in the receptor that perturbs a single amino acid position can cause cascading effects of fragility both at the molecular level and at a signaling network level and discuss the mutations in the context of a particular cancer cell line.


COMP 100

Coarse grained molecular dynamics of biological and soft matter systems

Russell H. DeVane(1), devane@temple.edu, 1900 N. 12th Street, Bio-Life Suite 113, Philadelphia PA 19122, United States ; Wataru Shinoda(2); Michael L. Klein(1). (1) Institute for Computational Molecular Science, Temple University, Philadelphia PA 19122, United States (2) RICS, National Institute of Advanced Industrial Science & Technology, Tsukuba Ibaraki 305-8568, Japan

Computational resources continue to rapidly increase allowing theoretical investigation of larger systems on longer timescales. While the capacity of all-atom (AA) molecular dynamics (MD) has reached a substantial level, the spatial and temporal scales of many soft matter and biological systems of interest are still well beyond reach. To address this issue, coarse grain (CG) models have received much attention. The determining factor in the sucess of a CG model is the parametrization methodology. We have recently developed a novel parametrization approach that relies heavily on experimental data and reduces the dependence on AA MD simulations. The resulting CG potential is based upon rather standard functional forms facilitating implementation in conventional MD codes. This approach has been
applied to the study of self-assembly of surfactant and lipid systems as well as structure prediction of proteins. The results demonstrate the ability to make modular transferable CG sites that are capable of accurately modeling systems of interest.

COMP 101

Predicting tautomer preference: Simple rules and unforeseen complexities

Peter J Taylor(1), pjtaylor.cheadle@btinternet.com, 7 Radnормere Drive, Cheadle SK8 5JX, United Kingdom ; Peter W. Kenny(1). (1) AstraZeneca (retired), Cheadle SK8 5JX, United Kingdom

Tautomer ratio depends on phase, so for coherent analysis this must be chosen first. We settle for water as the biological medium, and show inter alia that the gas phase is still more removed from water than even the least polar of organic solvents. We also point out that, while minor tautomers may bind to receptors, this must entail an energetic penalty.

The 'basicity method' is the main source of quantitative data in water but suffers from systematic errors through its inevitable reliance on model compounds. Elimination of these using correction factors not only improves accuracy but has demonstrated structural regularities that have gone unsuspected till now. Their extrapolation leads to plausible predictions amenable to experiment. The effects of benzofusion, and of intramolecular lone pair and dipolar repulsion, exemplify these regularities and will be discussed.

Central to our approach is the realisation that tautomerism takes two forms, 'C-type' and 'N-type,' which depend on different electronic factors. The apparent inconsistencies that result may have helped to inhibit the comprehensive approach to tautomer ratio that is needed, and hopefully their rationalisation will help in its renewal.

COMP 102

Methods for robust and efficient tautomer enumeration, tautomer searching and tautomer duplicate filtering

Szabolcs Csepregi(1), scsepregi@chemaxon.com, Máramaros köz 3/a, Budapest - 1037, Hungary ; Ferenc Csizmadia(1); József Szegedi(1); Zsolt Mohácsy(1); Tamás Csizmazia(1); Szilárd Dóránt(1); Ákos Papp(1); György Pirok(1). (1) ChemAxon Ltd., Budapest 1037, Hungary

Tautomerism is an important and difficult problem in cheminformatics, and has gained much attention recently. [1] The presentation will focus on ChemAxon's approaches and algorithms for handling tautomerism.
There are four main topics to cover:

1. The tautomerization calculator plugin [2] is the basis of most methods. It can identify tautomerizable regions, enumerate all or dominant tautomers and predict the distribution of dominant tautomers. Furthermore, it can provide generic and canonical tautomers that are used by the methods discussed. It first identifies possible proton donors and acceptors and finds the tautomerization paths between them. Depending on the desired operation, it then combines the paths into regions (generic tautomer), combinatorially enumerates all possible tautomeric forms (all tautomers), filters and ranks enumerated structures based on pKa and other criteria (dominant tautomers) or canonicalizes using empirical rules (canonical tautomer).

The tautomerization plugin is also used to improve results of other calculations, such as macro pKa and logP.

2. Tautomer duplicate search uses generic tautomers combined with a hash key. This method also allows fast filtering of tautomers in chemical database tables. It will be shown how this method is able to handle tautomeric migration of H isotopes and interactions with stereochemistry.

3. Tautomer substructure search enumerates tautomers of the query, and searches each of them separately. In case of query H constraints (explicit H), the constraint is enforced on the tautomeric region to retrieve only true tautomers.

4. Standardizer is a tool for performing custom and built-in transformations on molecules. It is integrated with the JChem chemical database system, so that database and query structures are automatically transformed by the specified transformations [3]. It will be shown how the canonical tautomer and custom transformations can be used to handle tautomerism. Custom transformations also allow handling of ring-chain tautomerism.

References:


COMP 103
Tautomerization approach for drug-like molecules

John C. Shelley\(^{(1)}\), John.Shelley@schrodinger.com, 101 SW Main St., Suite 1300, Portland Oregon 97204, United States; Arron P. Sullivan\(^{(1)}\), David Calkins\(^{(1)}\), Jeremy R. Greenwood\(^{(2)}\). (1) Schrodinger, Inc., Portland Oregon 97204, United States (2) Schrodinger, Inc., New York New York 10036, United States

We outline a pragmatic approach for generating the important protonation states, including tautomers, for drug-like molecules in the context of ligand and structure based virtual screening. The emphasis is on generating those states that have significant populations (which we define to be 0.01 mole fraction or more) in solution. These states also encompass the vast majority of those intuited from the examination of more than 2,500 protein-ligand complexes. The overall technology combines the use of many pre-parameterized tautomeric equilibria with Hammett and Taft calculation estimates of pKa values, which in turn can also be used to generate variations in both protonation states and tautomeric states. The overall approach permits the calculation of the mole fractions for the states generated along with their relative free energies. These free energy estimates have been shown to improve the performance of subsequent studies such as docking with Glide.

COMP 104

Acid/base ionization vs. prototropic tautomerism

Robert Fraczkiewicz\(^{(1)}\), robert@simulations-plus.com, 42505 10th Street West, Lancaster CA 93534, United States; Marvin Waldman\(^{(1)}\); Robert D. Clark\(^{(1)}\); Walter S. Woltosz\(^{(1)}\); Michael B. Bolger\(^{(1)}\). (1) Life Sciences, Simulations Plus, Inc., Lancaster CA 93534, United States

The most serious difficulty in computational predictive modeling of tautomerism is the lack of a sufficiently comprehensive database of tautomeric constants. [1] Published data on aqueous protonic ionization is, on the other hand, quite abundant to build successful QSPR models. Moreover, prototropic tautomerism is intimately tied to ionization in more than one way. We present compelling examples of how these ties can be explored to make both qualitative and quantitative predictions regarding tautomers using a truly predictive model of ionization constants. We show a very surprising case where the model refuted the widely accepted tautomeric form of one of the most successful drugs on the market today and how all of these predictions were confirmed beyond any doubt, both experimentally and theoretically. We demonstrate how the complex tautomerism of another very well known drug could be explained and quantified from its predicted ionization patterns. A general theoretical treatment of tautomer and ionization equilibria will be presented as well.

COMP 105

Combinatorial-computational-chemoinformatics approach to finding and analyzing low-energy tautomers

Maciej Haranczyk\(^{(1)}\), mharanczyk@lbl.gov, One Cyclotron Rd, MS 50F-1650, Berkeley CA 94720, United States; Maciej Gutowski\(^{(2)}\). (1) Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley CA 94720, United States (2) Chemistry-School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom

Enumeration of low-energy tautomers of neutral molecules in the gas-phase or typical solvents can be performed by applying available organic chemistry knowledge. However, in esoteric cases such as charged molecules in uncommon, non-aqueous solvents there is simply not enough available knowledge to make reliable predictions of low energy tautomers. We have been developing an approach to address the latter problem and we successfully applied it to discover the most stable anionic tautomers of nucleic acid bases that might be involved in the process of DNA damage by low-energy electrons. The approach involves three steps: (i) combinatorial generation of a library of tautomers, (ii) energy-based screening of the library using electronic structure methods, and (iii) analysis of the information generated in step (ii). In steps i-iii we employ combinatorial, computational and chemoinformatics techniques, respectively. This presentation summarizes our developments and most interesting methodological aspects of our approach.

COMP 106

Comparison of pattern-based and algorithm-based approaches to tautomer informatics

Ben Ellingson\(^{(1)}\), ben@eyesopen.com; Robert Tolbert\(^{(1)}\); A. Geoffrey Skillman\(^{(1)}\). (1) OpenEye Scientific Software, Inc, Santa Fe NM 85708, United States

Tautomers are an important consideration for cheminformatics and molecular modeling. In cheminformatics, a unique tautomer is stored as the singular registration key where it is vital that the unique key can be generated from any tautomer as well as that all tautomers can be generated from the unique key. The stored tautomer is often chosen for aesthetics or computational ease, but chemical implications such as the loss or gain of aromaticity or stereochemistry...
through tautomerization must also be addressed. Molecular modelers are often concerned with small ensembles of low energy tautomers. Unfortunately, determining the low energy tautomers is a complex task, for which sub-kcal/mol accuracy remains computationally intensive [1]. Thus, tautomer prediction for large-scale modeling or cheminformatics remains the domain of approximate. We will discuss two such approximate methods, pattern-based tautomer recognition and atom-type tautomer recognition. The advantages and disadvantages of these approaches will be examined.


COMP 107

Sticking to kinases: Scaffold selectivity and implications for focused library design

Shana L Posy(1), shana.posy@bms.com, POB 4000, Princeton NJ 08543-4000, United States; Stephen R Johnson(1); Deborah Loughney(1); Mark Hermsemier(2); Wayne Vaccaro(3); George Trainor(3); Karl-Heinz Ot(4). (1) Department of Applied Biotechnology, CADD, Bristol-Myers Squibb Co., Princeton NJ 08543, United States (2) Department of Chemistry Informatics, Bristol-Myers Squibb Co., Princeton NJ 08543, United States (3) Department of Discovery Biology, Bristol-Myers Squibb Co., Princeton NJ 08543, United States (4) Department of Applied Biotechnology, Bioinformatics, Bristol-Myers Squibb Co., Pennington NJ 08534, United States

We report the results of a kinome-wide selectivity screen of >20,000 compounds. Analysis of the selectivity patterns for each structural class shows that a broad spectrum of structural scaffolds can achieve specificity for most kinases. Although selective and nonselective compounds are mostly similar in their physicochemical characteristics, we identify specific features that are present more frequently in compounds that bind to many kinases. Our results support a scaffold-oriented approach for building compound collections to screen kinase targets.

COMP 108

Novel classification of proteases in the peptide substrate space reveals cross-family substrate similarities with applications to lead discovery

Sai Chetan K. Sukuru(1), chetan.sukuru@novartis.com, 250 Massachusetts Ave, Cambridge MA 02139, United States; Florian Nigsch(1); Jean Quancard(2); Martin Renatus(2); Rajiv Chopra(1); Natasja Brooijmans(3); Dmitri Mikhailov(1); Zhan Deng(1); Allen Cornett(1); Jeremy L. Jenkins(1); Ulrich Hommel(2); John W.
We present here a novel in silico approach to classify proteases in their peptide substrate space. Intuitively, one would expect that proteases from the same family, based on the traditional classification (aspartic, cysteine, serine and metallo), would be similar in the substrate space. However, our approach reveals that proteases from different families could also have substrates that differ at the cleavage site but are similar away from it. Caspase-3 (cysteine protease) and granzyme B (serine protease) are examples of cross-family neighbors identified by this method and supported by the presence of compounds that inhibit both. In addition, a strategy to assess whether peptide substrate similarity between unrelated proteases could reliably translate into small-molecule or drug like inhibitor similarity was tested on two other cross-family neighbors with no known shared inhibitors. The results indicate that this approach could be prospectively applied to lead discovery for a novel protease target.

COMP 109

Profile-QSAR and surrogate autoshim: Utilizing chemogenomic relationships for drug discovery applications

Prasenjit Mukherjee(1), prasenjit.mukherjee@novartis.com, 4560 Horton Street, Emeryville CA 94608, United States; Eric Martin(1). (1) Department of Computer Aided Drug Discovery, Global Discovery Chemistry, Novartis Institutes for Biomedical Research, Emeryville CA 94608, United States

Hit discovery by high-throughput screening (HTS) carries huge burdens of time and resources. 2D Profile-QSAR and 3D Surrogate Autoshim are two experimentally parameterized, protein family-based, virtual screening methods, which harness the power of chemogenomics to create more efficient and predictive in silico tools. The methods have been applied in multiple real life kinase projects, and have shown consistent prediction accuracy and selectivity, resulting in high hit-rates even while exploring novel chemical space. The methods have been extended beyond what HTS can address to kinase selectivity predictions for hit-to-lead optimization and library design. 2D Profile-QSAR has also been extended to cellular activity predictions for kinase inhibitors. It has proven particularly useful where disconnects exist between cellular and biochemical activity. Over 80 Profile-QSAR models covering most of the human kinome have now been generated on experimental assay data. A kernel-type method interpolates from these models to generate a priori activity predictions for novel kinases with no experimental training data. The sequential floating search technique was implemented to identify a “minimal ensemble” of kinase structures for Surrogate Autoshim that halves the time for pre-docking large commercial databases. The two methods have been successfully extended to serine
proteases, and evaluations against ADME targets such as hERG are now underway.

COMP 110

Introduction of an interpretation layer to link unsupervised and supervised methods for classification of chemicals in biological effects space

James Rathman\(^{(1)}\), rathman.1@osu.edu, 140 W. 19th Ave., Columbus OH 43210, United States ; Chihae Yang\(^{(2)}\). (1) Department of Chemical and Biomolecular Engineering, The Ohio State University, Columbus OH 43210, United States (2) Altamira, LLC, Columbus OH 43235, United States

Current HTS technologies and databases of in vivo testing results provide sufficiently large datasets for the development of new informatics approaches for screening and prioritizing chemicals for further testing. Chemoinformaticians face the significant challenge of developing predictive associations across these highly diverse and noisy data-rich landscapes. One of the largest challenges is relating the high dimensional chemical descriptor space to compound-level biological data. Unsupervised approaches based on chemical descriptors are useful for clustering, but may be inadequate for finding biologically meaningful neighbors. Supervised methods based on biological activity data are used to derive structural alerts or privileged substructures relevant to the endpoint of interest; however, these substructures alone provide limited insight into underlying mechanisms. An “interpretative layer” of analysis is introduced to connect unsupervised and supervised methods through chemical and biological reactivities. Inclusion of reactivity classifiers in this manner augments the descriptor space to incorporate metabolic knowledge and biological endpoints.

COMP 111

Creating reaction schemes to generate an arbitrary number of linked and conjugated polymers

Roberto Olivares-Amaya\(^{(1)}\), olivares@fas.harvard.edu, 12 Oxford St., Cambridge MA 02138, United States ; Sule Atahan\(^{(1)}\); Alan Aspuru-Guzik\(^{(1)}\); Carlos A. Bedolla\(^{(1)}\)\(^{(2)}\). (1) Department of Chemistry and Chemical Biology, Harvard University, Cambridge MA 02138, United States (2) Depto. de Quimica Teorica, Facultad de Quimica, Universidad Nacional Autónoma de Mexico, Cambridge MA 02138, United States

Advances in computational chemistry bring us closer to predicting molecular photovoltaic properties. This will provide experimentalists with valuable data to study a specific set of molecules. In collaboration with IBM, we developed a screensaver (http://cleanenergy.harvard.edu), which allows users to contribute
their computer time to perform electronic structure calculations on a combinatorial molecular library. We generated such library by utilizing a reaction-like approach to form the molecules. We start with a set of 'building blocks' whose reactivity will be given by a specific label (i.e. Mg). The first set of reactions were produced by a facile Grignard-like reaction. The second reaction scheme included a more complex, albeit completely exhaustive fused-ring generation scheme. These two reactions form linked and fused n-oligomers of arbitrary size. Future work will focus on algorithms to pick-and-choose interesting molecular subspaces after an initial set of these has run, as well as a fused-linked scheme for molecule generation.

COMP 112

Quantitative rate calculations from simulations: The Ru$^{2+}$ - Ru$^{3+}$ electron self-exchange reaction in water

**Harald Oberhofer**$^{(1)}$, ho246@cam.ac.uk, Lensfield road, Cambridge Cambridgeshire CB2 1EW, United Kingdom ; **Jochen Blumberger**$^{(2)}$. (1) Department of Chemistry, University of Cambridge, Cambridge Cambridgeshire CB2 1EW, United Kingdom (2) Department of Physics and Astronomy, University College London, London WC1E 6BT, United Kingdom

We show the feasibility of calculating electron exchange reaction rates in condensed systems from scratch by determining the dependence of all parameters occurring in the Marcus theory rate expression on the distance \( r \) of the reactants (Donor and Acceptor). Using both, classical and density functional theory (DFT) simulation techniques we calculate the solvent reorganisation energy \( \lambda(r) \) and the electronic transition matrix element \( H_{ab}(r) \). Together with the potential of mean force \( G(r) \) as a function of the Donor-Acceptor separation this allows us to calculate the overall reaction rate. We also take into account quantum corrections due to the classical nature of the vibrational modes in our model. We illustrate this approach with our results for the Ru$^{2+}$ - Ru$^{3+}$ electron self-exchange reaction in water, which are in good agreement with experiment.

COMP 113

“On water” reactivity and solvent effects for Claisen rearrangements from QM/MM simulations

**Orlando Acevedo**$^{(1)}$, orlando.acevedo@auburn.edu, 179 Chemistry Building, Auburn AL 36849, United States ; **Kira Armacost**$^{(1)}$. (1) Department of Chemistry and Biochemistry, Auburn University, Auburn AL 36849, United States

An “on water” environment, defined by the absence of water solubility of the reactants, has been reported to provide increased rate accelerations, yields, and specificity for several types of organic reaction classes compared to organic
solvents. The aromatic Claisen rearrangements of allyl \( p \)-R-phenyl ethers (R = CH\(_3\), Br, and OCH\(_3\)) and allyl naphthyl ether have been investigated to determine the origin of the “on water” effect using QM/MM Monte Carlo calculations and free-energy perturbation (FEP) theory. The presentation will focus on elucidating the origins of the enhanced reactivity “on water” and in 16 different solvents by exploring the position and orientation of the aromatic ethers at the aqueous interface, hydrophobic effects, solvent polarizability via a polarizable force field, and computed solute-solvent energy pair distributions and radial distribution functions. (J. Am. Chem. Soc. 2010, 132, 1966-1975.)

COMP 114

Effect of hydrogen bonding on the dynamical stability of water clusters

Jason R. Green\(^{(1)}\), jg525@cam.ac.uk, Lensfield Road, Cambridge N/A CB2 1EW, United Kingdom ; Thomas S. Hofer\(^{(2)}\); David J. Wales\(^{(1)}\). (1) Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom (2) Theoretical Chemistry Division, Leopold-Franzens-University of Innsbruck, Institute of General, Inorganic and Theoretical Chemistry, Innsbruck Innrain 52, Austria

Measurements of chaos are capable of characterizing molecular motions through vast, high dimensional potential energy landscapes. The role of water in biological processes as well as its complex hydrogen bond networks, which influence dynamical stability, make it a particularly attractive system to study. Here, we characterize the possible motions of small water clusters and their respective time scales with Lyapunov exponents. As a water model, we use a reactive all-atom forcefield that despite being classical, enables investigations of proton transfer. The utility of measuring chaos in chemical transformations will be discussed and further illustrated with these molecular dynamics simulations of water.

COMP 115

DFT geometry optimization and vibrational frequencies of the quaternary ammonium cation-water [N\((\text{CH}_3)_4^+\)OH\(-/\text{(H}_2\text{O})_n\)] clusters

Dunesh Kumari\(^{(1)}\), deepikashouran@gmail.com, 360 huntington avenue, 414 hurtig hall, Boston MA 02171, United States ; Eugene S. Smotkin\(^{(1)}\); Nicholas Dimakis\(^{(2)}\). (1) Chemistry and Chemical Biology, Northeastern University, Boston MA 02171, United States (2) Department of Physics and Geology, University of Texas-Pan American, Edinburg TX 78541, United States
Alkaline membrane fuel cells are attracting attention due to their potential to replace the acidic proton exchange membrane fuel cells (PEMFCs) because of improved kinetics at the cathode side. Infrared spectroscopy and ab initio calculations are used to study the structure and hydration properties of [N (CH₃)₄]⁺OH⁻/(H₂O)ₙ⁻ at different degrees of hydration (n=1 to 10). The dissociation of the OH⁻ ions results in a change in the point group symmetry of the cluster. This transition in point group symmetry is manifested in the FTIR spectra. The optimized geometry, harmonic vibrational frequencies and infrared intensities of the modeled compound, as it transitions through point group symmetries will be discussed.

**COMP 116**

**Computational study of vibrational spectra of some water containing complexes of atmospheric interest**

*Lauri Halonen*(1), lauri.halonen@helsinki.fi, PO Box 55 (A.I. Virtasen aukio 1), Helsinki FIN-00014, Finland ; Elina Säll(1); Teemu Salmi(1). (1) Laboratory of Physical Chemistry, Department of Chemistry, University of Helsinki, Finland, Helsinki FIN-00014, Finland

We have computed high-frequency stretching and bending vibrational overtone spectra of the water dimer, water trimer and water ammonia complex. These species have been approximated as individually vibrating monomer units. Bond stretch and bond angle coordinate Hamiltonians are constructed for each monomer unit using exact kinetic energy operators within the Born-Oppenheimer approximation. The potential energy surfaces are calculated using the coupled cluster method with correlation consistent basis sets. The explicitly correlated F12 theory, which includes the interelectronic coordinate in the electronic wavefunction, has been adopted for the water ammonia complex. The dipole moment surfaces are obtained with the finite difference method. Eigenvalues of the Hamiltonians are computed variationally. Eigenvectors with the dipole moment surfaces have been used to calculate infrared absorption intensities. Finally, the results have been compared with experimental gas phase infrared and matrix isolation spectroscopic data.

**COMP 117**

**Using a vibrational averaging technique to understand the role of water dimers in atmospheric absorption**

*Ross E. A. Kelly*(1), ross.kelly@ucl.ac.uk, Gower Street, London WC1E 6BT, United Kingdom ; M. J. Barber*(1); J. Tennyson*(1); G C. Groenenboom*(2); A. van der Avoird(2). (1) Department of Physics & Astronomy, University College London, London WC1E 6BT, United Kingdom (2) Theoretical Chemistry, Institute for Molecules and Materials, Radboud University, Nijmegen, The Netherlands
Water is the most important absorber of both long and short range radiation in the atmosphere [1]. Therefore, a proper characterization of all absorption by water vapor is of upmost importance in climate prediction. The spectrum of water monomer has been well studied. In comparison, the role of the water dimer in atmospheric absorption is poorly understood. From a computational viewpoint, simulating this spectrum requires the solution to a 12D problem. However, one may adiabatic separate the monomer and dimer modes [2,3]. This allows the use of rigid 6D monomer approaches which normally involve evaluating a potential energy surface on a specified intermolecular grid with rigid monomers, followed by solution to the nuclear motion problem.

We relax the rigid monomer constraint by creating water monomer wavefunctions using DVR3D [4] and then vibrationally average the potential at each intermolecular grid point [2]. We perform the average on the best available water pair potential [5] with a monomer correction term [6] to create a spectroscopically transferable water pair potential [2]. Unfortunately, this leads to 1,306,613,597,184 potential calls which would take around 7000 days on a single CPU to compute. However, the averages at each point on the dimer grid are completely independent from one another: an ideal case for grid computing. We use the UCL Condor grid which is a distributed high-throughput infrastructure of ~1400 Windows PCs using Condor resource management software. Furthermore, storing the potential points would occupy about 10TB of disk/memory, so that the averaging must be done on the fly.

We first discuss how these calculations have increased the accuracy of existing VRT calculations in comparison with low temperature experiments [2]. Additionally, we report water dimer VRT levels calculated with excited monomers which allow us to give water dimer spectra in the infrared and visible regions of the atmosphere. The agreement with experimental [7] and lower-level theoretical calculations suggests a small but significant contribution of the water dimer to the water continuum problem in the near IR and visible regions in the atmosphere. This is part of the large consortium, Continuum Absorption at Visible and Infrared wavelengths and its Atmospheric Relevance, or simply CAVIAR [8].

8. http://www.met.reading.ac.uk/caviar/

COMP 118

Modeling for the masses: The philosophers' stone of CAMD software

Andrew C Good(1), 153 Second Avenue, Waltham Massachusetts 02451, United States. (1) Drug Discovery and Development, Genzyme Corporation, Waltham Massachusetts 02451, United States

Most companies engaged in the business of writing CAMD software eventually buy into the myth of modeling for the masses. Tempted by theoretical market size, software executives believe they can move beyond their core constituents to the medicinal chemistry community at large through modest simplifications of their interface. This conviction is undone by a core misunderstanding of the factions that exist within a typical large medicinal chemistry community. The dynamics of said factions splinter the theoretical market, both shrinking its size and increasing the complexity of its exploitation. In this talk these dynamics and their ramifications are explained in depth, as are potential methods for overcoming the complexities they engender.

COMP 119

Lightning-fast 3D-shape and feature-based virtual screening

Christian Lemmen(1), CLemmen@BioSolveIT.de, An der Ziegelei 79, 53757 Sankt Augustin, Germany. (1) BioSolveIT GmbH, Germany

Descriptor-based similarity searches are known to be extremely fast and suitable for high throughput virtual screening. Whereas shape-based methods are considered to be more accurate but significantly slower. We have combined these two approaches to gain the better of both worlds, the speed of the descriptor-based search and the accuracy of the shape matching. Application examples and results of benchmark studies will be presented.

COMP 120

Community structure-activity resource: Collecting, curating, and generating protein-ligand data to improve docking and scoring

James B. Dunbar(1), jbdunbar@umich.edu, 428 Church St., Ann Arbor MI 48109-1065, United States; Heather A. Carlson(1). (1) Department of Medicinal Chemistry, University of Michigan, Ann Arbor, Ann Arbor MI 48109-1065, United States
The Community Structure-Activity Resource (CSAR) is a center at the University of Michigan funded by the National Institute of General Medical Sciences. The function of this center is to collect, curate, and disseminate protein-ligand data sets of crystal structures, biological binding affinities, and thermodynamic data to aid in the refinement of docking and scoring methodologies. These data sets are to come from in-house projects at the University of Michigan, other academic labs, and most importantly from industrial, pharma sources. Part of our remit is to augment the deposited data with synthesis, crystallography, and assays to expand the range of properties, binding affinities, and other relevant characteristics involved in docking and scoring. Here, we present CSAR’s capabilities and summarize our current in-house project and potential future targets. We also outline the creation of a dataset (based on the PDB, Binding MOAD, and PDBbind) used in our first community-wide benchmark exercise.

COMP 121

Results of CSAR’s 2010 Benchmark Exercise

Heather A. Carlson(1), carlsonh@umich.edu, 428 Church St., Ann Arbor MI 48109-1065, United States; James B. Dunbar(1); Richard D. Smith(1). (1) Department of Medicinal Chemistry, University of Michigan, Ann Arbor, Ann Arbor MI 48109-1065, United States

The goal of CSAR’s Benchmark Exercises is not to declare winners and losers! Instead, we combine the results of all participants to provide a wider assessment of the field. Here, we present an analysis of which protein-ligand complexes score poorly across the majority of submissions (“globally bad” complexes) and compare their properties to the set of complexes that score well across the majority of methods (“globally good”). It may be tempting to draw conclusions by simply examining the characteristics of the globally bad set, but those characteristics must be rarely observed in the globally good set to gain true insight. Lastly, each participant was asked to submit a standard method and an alternative approach. Several groups showed that the correlation to experiment was the same for vdw/fit-based scores as for full scoring functions that included electrostatics and hydrogen bonding. To help the field overcome this limitation, CSAR will focus on creating datasets that provide a range of hydrogen-bonding characteristics. The overarching goal of our benchmark exercises is to provide insight into what data is most needed to move our field ahead.

COMP 122

Scoring performance of eHiTS on the CSAR dataset

Zsolt Zsoldos(1), zsolt@simbiosys.com, 135 Queen[apos]s Plate Drive, Toronto Ontario M9W6V1, Canada; Orr Ravitz(1). (1) SimBioSys Inc., Toronto M9W6V1, Canada
Numerous studies have pointed out at the inability of scoring functions to perform uniformly well across all biological systems of interest. Some studies suggest guidelines for choosing the best method for a specific problem, others advocate consensus techniques.

An alternative solution is to tailor the scoring function for the system of interest. eHiTS uses a novel scoring method consisting of statistical knowledge focused on interacting surface points and physical terms combined with an adaptive parameter scheme. During the automated tuning of eHiTS-score, receptor targets are clustered according to the chemical and shape similarity of the active site, and weight sets are optimized for each family.

The performance of eHiTS on the CSAR dataset was evaluated using the default parameters (pre-tuned on other data). In addition, the automatic tuning utility was run on one subset of the CSAR data and tested on the other. Results will be presented from both studies.

**COMP 123**

Hydrophobic complementarity: A dominant term in affinity and binding mode prediction

*Leslie A. Kuhn*(1), KuhnL@msu.edu, 502C Biochemistry Building, East Lansing MI 48824, United States ; Matthew E. Tonero*(1). (1) Biochemistry & Molecular Biology, Michigan State University, East Lansing MI 48824, United States

Empirical scoring functions designed for high-throughput docking, containing linear combinations of terms measuring protein-ligand interactions, were tested for affinity prediction. Scoring functions that best predicted affinity were dominated by hydrophobic or shape complementarity terms. Similarly, a scoring function containing only polar terms compensated for the absence of a hydrophobic term by heavily weighting the polar term that correlated most with hydrophobic complementarity. These results are consistent with Eisenberg & McLachlan's observation that the solvation component of the change in Gibbs free energy upon binding is proportional to the surface area and degree of hydrophobicity of atoms buried in the interface. Scoring functions that perform best at affinity prediction are not necessarily optimal for binding mode prediction, though hydrophobic burial is important in both. In other words, tuning scoring functions only to predict the affinity of good ligands in the correct binding mode can limit their applicability, suggesting a broader approach.

**COMP 124**

Docking and scoring for 2010 CSAR benchmark using an improved iterative knowledge-based scoring function with MDock
Based on a physics-based iterative method (Huang & Zou, J. Comput. Chem., 2006, 27, 1865-75; 1876-82), we have extracted a set of distance-dependent all-atom potentials for protein-ligand interactions (ITScore2.0) using a large training set of 1300 protein-ligand complexes. The iterative method circumvents the long-standing reference state problem in traditional knowledge-based scoring functions. ITScore2.0 has been tested with the 2010 CSAR dataset of 345 diverse protein-ligand complexes, and achieved a correlation coefficient of 0.73 between the calculated binding scores and experimental affinity data, compared to 0.58 for the van der Waals (VDW) scoring function and 0.32 for the force field (FF) scoring function consisting of VDW and electrostatic terms. For rigid-ligand docking, ITScore2.0 achieved a success rate of 86.7% in identifying native binding modes, compared to 80.0% and 64.1% for FF and VDW. For flexible-ligand docking, ITScore2.0 yielded a success rate of 79.7%, compared to 71.0% and 52.8% for FF and VDW. The moderate performance of VDW suggests that VDW alone may serve as a benchmark for evaluation of scoring functions. What we have learned through participating in CSAR scoring will be shared.

COMP 125

Quantum glass

David Reichman\(^{(1)}\), drr21032@columbia.edu, 3000 Broadway, New York New York 10027, United States . (1) Chemistry, Columbia University, New York New York 10027, United States

While the origins and causes of classical glassy behavior in liquids is still hotly debated, almost no work has been done to characterize transformations to non-ergodic phases when quantum fluctuations are important. I will discuss several recent experiments to motivate the systematic study of the interplay between quantum fluctuations and glass transitions, and then examine theoretically two interesting examples drawn from recent work in our group.

COMP 126

Energy transfer, coherence, and decoherence in molecular aggregates

Robert J Silbey\(^{(1)}\), silbey@mit.edu, 77 Massachusetts Avenue, Cambridge MA 02139, United States . (1) Department of Chemistry, MIT, Cambridge MA 02139, United States
The excited states of molecular aggregates, such as the photosynthetic light harvesting systems, are often coherent superpositions of localized molecular excitations. Coherence has been experimentally seen in recent femtosecond studies of these systems (and in conjugated polymers as well).

Coherence can enhance the rate of energy transfer among the molecules forming the aggregate. Dynamic and static fluctuations in the environment can cause the coherence to decay; however, it has recently been argued that this decoherence can increase the efficiency of transfer to a trap in the system.

In this talk, we discuss the effect of decoherence on energy transfer and trapping, correlations of the fluctuations on different sites and their effects, and a new theoretical model for understanding these effects from a classical electromagnetism viewpoint and derive a formula that leads to Forster transfer in the incoherent limit and mixed coherent-incoherent energy transfer in general.

**COMP 127**

**Local solvent environment dependent electron transfer reactions in room-temperature ionic-liquids**

Claudio J Margulis(1), claudio-margulis@uiowa.edu, 118 IATL, Iowa City Iowa 52242, United States. (1) Chemistry, University of Iowa, Iowa City Iowa 52242, United States

Maroncelli and co-workers (J. Phys. Chem. B 2007, 111, 13473) have recently shown that the relative population of electron transfer products after photo-excitation of crystal violet lactone (CVL) in novel room-temperature ionic-liquids is absorption wavelength dependent. This is consistent with previous findings by Samanta (J. of Photochemistry and photobiology A-Chemistry, 182, 2, 113-120) and others who studied red-edge excitation phenomena in similar systems. We have recently studied fluorescence as well as excited state intramolecular electron transfer reactions in room temperature ionic liquids. Our simulations show that the kinetics of the intramolecular ET between S-1 and S-2 states of CVL in [Pr-31(+)][Tf2N-] is local solvent-environment-dependent. The same phenomenon which occurs in the case of fluorescence of different probes is due to the fact that emission time
scales are short when compared to solvent relaxation time scales. This behavior which is characteristic of RTILs is not common in conventional solvents at room temperature.

**COMP 128**

**Crystallization of hard aspherical particles**

*Angelo Cacciuto*(1), ac2822@columbia.edu, 3000 Broadway, MC 3123, New York NY 10027, United States . (1) Chemistry, Columbia University, New York NY 10027, United States

We use numerical simulations to study the crystallization of monodisperse systems of hard aspherical particles. We find that particle shape and crystallizability can be easily related to each other when particles are characterized in terms of two simple and experimentally accessible order parameters: one based on the particle surface-to-volume ratio, and the other on the angular distribution of the perturbations away from the ideal spherical shape. We present a phase diagram obtained by exploring the crystallizability of 487 different particle shapes across the two-order-parameter spectrum. Finally, we consider the physical properties of the crystalline structures accessible to aspherical particles, and discuss limits and relevance of our results.

**COMP 129**

**Fluctuations and the second law of thermodynamics: What do we know that Maxwell, Boltzmann and Gibbs did not know?**

*Christopher Jarzynski*(1), cjarzyns@umd.edu, 2203 Chemistry Building, College Park MD 20742, United States . (1) Department of Chemistry and Biochemistry, and Institute for Physical Science and Technology, University of Maryland, College Park MD 20742, United States

The second law of thermodynamics, by outlawing processes that decrease entropy, is intimately related to the "arrow of time" - a term introduced by Eddington to describe the evident directionality of the flow of events. Since the days of Maxwell and Boltzmann it has been appreciated that the second law must be interpreted statistically: in sufficiently small systems, thermal fluctuations blur the boundary between what is thermodynamically allowed and what is forbidden. Recent progress in statistical mechanics has revealed that, away from equilibrium, these fluctuations satisfy rather strong and unexpected laws: equilibrium information is subtly encoded in far-from-equilibrium fluctuations; statistical distributions of work and entropy production exhibit surprising symmetry properties; and our ability to distinguish the direction of time's arrow obeys universal, system-independent identities. I will summarize and illustrate
these results, and will argue that they have refined our understanding of the second law and the nature of irreversibility at the nanoscale.

COMP 130

Microscopic kinetic model exhibiting chiral symmetry breaking

Frank H. Stillinger\textsuperscript{(1)}, fhs@princeton.edu, Princeton University, Princeton NJ 08544, United States. \textsuperscript{(1)} Chemistry Department, Princeton University, Princeton NJ 08544, United States

The emergence scenario of a dominant chirality (handedness) for biomolecules from a presumably non-chiral pre-biotic earth continues to present a major scientific mystery. In order to investigate possible mechanisms a simple lattice model has been constructed whose kinetics displays spontaneous chiral symmetry breaking in chemical formation of enantiomers from non-chiral reactants. This model includes autocatalysis, inhibition, and diffusion of reactants and products. Monte Carlo simulation demonstrates that starting at a precisely symmetrical initial condition with only achiral reactants and solvent present, chiral symmetry breaking in product distributions spontaneously occurs below a positive threshold temperature.

COMP 131

Getting the right answer for the right reason: Accurate and transferable coarse-grained models from structural information

William G Noid\textsuperscript{(1)}, wnoid@chem.psu.edu, 104 Chemistry Building, University Park PA 16802, United States; Wayne Mullinax\textsuperscript{(1)}. \textsuperscript{(1)} Department of Chemistry, The Pennsylvania State University, University Park PA 16802, United States

The present talk will discuss the challenge of systematically determining coarse-grained models that are consistent with the structural properties of atomistic models for multiple systems. In particular, we will discuss our progress in generalizing liquid state theories for extended ensembles. The long term goal of this work is to bridge the gap between physics-based and knowledge-based approaches and determine accurate and transferable coarse-grained models from structural information.

COMP 132

Improved lattice models of the hydrophobic effect

Patrick Varilly\textsuperscript{(1)(2)}, pvarilly@berkeley.edu, 42 Gilman, Pitzer Center for Theoretical Chemistry, Berkeley CA 94720, United States; Amish J Patel\textsuperscript{(1)}; David Chandler\textsuperscript{(1)}. \textsuperscript{(1)} Department of Chemistry, UC Berkeley, Berkeley CA
We present a lattice model of solvation by water and the hydrophobic effect. We carefully implement the ideas of ten Wolde, Sun and Chandler, fixing the salient deficiencies of their original approach at small and large length scales. By construction, our model correctly reproduces the well-known length-scale dependence of solvation free energies, in contrast to GBSA-type models. Using our model, we calculate the water number distribution, $P_V(N)$, in various volumes (in bulk, next to large solutes and in confinement), and show that our results compare favorably with the results of SPC/E calculations, while being several orders of magnitude faster to compute. We report on preliminary work on applying this model to self-assembly of realistic nanoscale systems, where the solvent plays an essential role. Our model can serve as a basis for an accurate description of solvents, without the cost of an explicit representation.

**COMP 133**

**Dissecting force interactions between cellulose and ionic liquids by using the iterative-YGB method**

*Hyung Min Cho* (1), chohm520@gmail.com, 201 Gilman Hall, Berkeley California 94720, United States; *Jhih-Wei Chu* (1). (1) Department of Chemical Engineering, University of California, Berkeley California 94720, United States

Dissolving cellulose using ionic liquids has drawn much attention recently as a promising 'green technology' to make the most of biomass resources without causing serious environmental problems. To characterize the governing mechanisms, understanding how the interactions between cellulose and ionic liquids affect the dissolution process is of fundamental importance. In this study, we employed the iterative-YBG method\(^{(1)}\) to quantify the force interactions between different chemical groups of cellulose and ionic liquids from all-atom MD simulations. The iterative-YBG method can be used to invert the observed structural distribution functions in all-atom MD or in experiments into the interaction forces between chemical groups. The iterative-YBG method is applied to different stages along the pathways of peeling of cellulose chains from a microfibril to elucidate solvent effects on decomposition. As a basis of comparison, cellulose-IL interactions are compared with cellulose-water interactions during the peeling processes.

Fluctuations, slow dynamics, and ordering across length and time scales in nanoparticle self-assembly

**Phillip L Geissler**\(^{(1)(2)}\), geissler@berkeley.edu, Berkeley CA 94720, United States; **Michael Gruenwald**\(^{(1)}\); **Asaph Widmer-Cooper**\(^{(2)}\). (1) Department of Chemistry, U.C. Berkeley, Berkeley CA 94720, United States (2) Chemical and Materials Sciences Divisions, Lawrence Berkeley National Lab, Berkeley CA 94720, United States

Interactions among aggregating nanoparticles are often mediated by complex fluctuating environments, from the malleable hydrogen bonding networks of aqueous solution to the chain molecules that passivate most inorganic nanocrystals. In the latter context we have performed computer simulations that evince collective and frustrated molecular rearrangements, echoing features of pattern formation at much larger scales. In the case of semiconducting nanorods, dense shells of organic ligands are poised near a monolayer ordering transition that can strongly and dynamically shape rod-rod association. For colloidal silver polyhedra, the statistical mechanics of polymer adsorption exerts a striking bias on the geometry of dense packings. The interplay between these fluctuating microscopic forces and the rich kinetics of spontaneous nanoparticle organization poses substantial computational challenges, but also offers unique opportunities to explore how variability at disparate scales can be communicated in both directions.

**COMP 135**

**Generalized mean field theory of coarse-graining**

**Vinod Krishna**\(^{(1)(2)}\), vkrishna@hec.utah.edu, Rm. 2020, 315 South 1400 East, Salt Lake City Utah 84112, United States; **Luca Larini**\(^{(1)}\). (1) Department of Chemistry, University of Utah, Salt Lake City Utah 84112, United States (2) School of Natural Sciences, Institute of Advanced Study, Salt Lake City Utah 84112, United States

A general mean field theory is presented for the construction of equilibrium coarse grained models. It is shown that inverse methods which reconstruct microscopic models from available low resolution data from experiment or by other means can be derived as particular implementations of this theory. Further, it is shown that the theory is also applicable to solving the opposite problem of reduction, where a large amount of high resolution, equilibrium ensemble data for the system is available, and the extraction of relevant information from this data is required. This problem is central to the construction of coarse grained representations of complex statistical mechanical systems, and it is shown that commonly used coarse graining methods are particular cases of the general theory, thus enabling the adaptation of commonly available statistical methods to the study of these problems.
COMP 136

Statistics in multiresolution modeling: An application of Bayesian interference in multiscale coarse-graining

Pu Liu\(^{(1)}\), P.Liu24@its.jnj.com, Welsh & McKean Roads, Spring House Pennsylvania 19477, United States. (1) Research & Development, Johnson & Johnson Pharmaceutical, Spring House Pennsylvania 19477, United States

In multi-resolution modeling, it is often required to derive parameters from different sources like high-resolution simulations and experimental results. Statistics provides us with a powerful toolset which can filter out experimental noise and alleviate the numerical instability in the parameterizations. In the talk, I will show how Bayesian interference, an advanced statistical tool widely used in signal processing and pattern recognition, will be used to regularize the linear equation from multiscale coarse-graining problem, give us reliable parameters and the extent the parameters can be trusted.

COMP 137

Risk assessment of chemicals and prediction of metabolism

Johann Gasteiger\(^{(1)}\), schwab@molecular-networks.com, Henkestrasse 91, Erlangen Bavaria 91052, Germany; Thomas Kleinoeder\(^{(1)}\); Joerg Maruczyk\(^{(1)}\); Christof H. Schwab\(^{(1)}\); Aleksey Tarkhov\(^{(1)}\); Lothar Terfloth\(^{(1)}\). (1) Molecular Networks GmbH, Erlangen 91052, Germany

Legislation within the European Union such as the REACH initiative and the 7th amendment for the "Cosmetic Directive" as well as the Canadian Environmental Protection Act, 1999, heavily suggest non-testing methods as alternatives to in vivo experiments on animals for the hazard and risk assessment of chemicals. Computational methods are therefore becoming of increasing importance for profiling the metabolic and environmental fate and the evaluation of the toxicity of chemical compounds in the risk assessment workflow.

We have developed the MOSES suite of programs by building on our elaborate structure encoding methods [1] and our long-standing experience in the modeling of chemical reactions [2]. The system MOSES.RiskAssessment incorporates an entire workflow for the evaluation of the toxicity and risk posed by a chemical compound. By integrating MOSES.Metabolism, a system for predicting human metabolic and biodegradation reactions, not only the parent compound but also its metabolites can be included in the risk assessment process.

Considering protein structural variations in molecular docking: Ensemble docking with GOLD

Jason Cole(1), admin@ccdc.cam.ac.uk, 12 Union Road, Cambridge CB2 1EZ, United Kingdom; Colin Groom(1). (1) Cambridge Crystallographic Data Centre, Cambridge CB21EZ, United Kingdom

Ligand-induced receptor conformational changes are common in ligand binding, ranging from local rearrangements of side chains to large domain movements. In some cases, even a small change in the receptor conformation can have a remarkable effect on the ligand binding affinity, leading to failure of molecular docking in both binding mode and affinity predictions if this conformational change is not incorporated in docking calculations.

This talk will demonstrate how receptor flexibility is accounted for in GOLD through use of soft potentials, side-chain flexibility and most importantly ensemble docking. In particular a novel methodology for ensemble docking, which avoids the computationally expensive sequential docking of ligands into a multiple protein structures will be described. Not only does this help remove the pain of choosing particular structural representations, but the results suggest we might like to re-evaluate how good we are at docking and virtual screening.

Configurational entropy and mechanical stress in molecular recognition

Michael K. Gilson(1), mgilson@ucsd.edu, 9600 Gilman Drive, MC 0736, La Jolla CA 92093, United States. (1) School of Pharmacy, University of California, San Diego, La Jolla CA 92093, United States

I will present molecular dynamics simulations consistent with long-ranged entropy effects throughout a protein upon binding a peptide. The results are somewhat preliminary, given the challenge of generating converged simulation results, but are qualitatively consistent with the long-ranged changes in orientational order parameters due to binding, which have been observed in NMR studies of binding.

These apparent long-ranged effects raise questions regarding the mechanisms by which binding affects remote parts of the protein. I will explain why the concept of mechanical stress may be useful in thinking about such long-ranged consequences, and will describe our initial computational studies of stress at the molecular level.
One of the most dangerous bioterror agents is the rod-shaped, spore-forming bacterium *Bacillus anthracis*, which is the causative agent of anthrax. Concentrated anthrax spores have been deployed as biological weapons in the United States and elsewhere, resulting in high mortality rates among those exposed. The lethal factor (LF) enzyme is secreted by the bacillus as part of the anthrax lethal toxin, and is mainly responsible for anthrax-related cytotoxicity. As LF can remain in the system long after antibiotics have eradicated the bacilli, the preferred therapeutic modality would be the administration of antibiotics together with an effective LF inhibitor. To date, however, no LF inhibitor is available as a therapeutic or preventive agent. Here we present an original high-throughput computational protocol that successfully identified five promising novel LF inhibitor scaffolds with low micromolar inhibition against that target, demonstrating a 12.8% experimental hit rate. This protocol incorporated topomeric shape-based searching techniques that were particularly effective in identifying potential new leads. Three of the five new hits exhibited experimental IC$_{50}$ values less than 100 μM and may potentially serve as scaffolds for lead optimization. Virtual screening simulations predicted that these preliminary hits
are likely to engage in critical ligand-receptor interactions with nearby residues in at least two of the three (S1', S1-S2, and S2') subsites in the LF binding area. Notably, it was found that micromolar-level LF inhibition can be attained by compounds with non-hydroxamate zinc-binding groups that exhibit monodentate zinc chelation as long as key hydrophobic interactions with at least two LF subsites are retained.

COMP 141

Model-free drug-like filters

Tudor I. Oprea(1), toprea@salud.unm.edu, MSC11 6145, Albuquerque NM 87131, United States; Cristian G. Bologa(1); Oleg Ursu(1). (1) Department of Biochemistry and Molecular Biology, Division of Biocomputing, University of New Mexico School of Medicine, Albuquerque NM 87131, United States

Extended connectivity descriptors computed by the Morgan algorithm have been used for the classification of various molecular properties. The information content encoded by such descriptors can be used to compute any 2D descriptors [1]. As these atom environments are canonical, we extracted them as molecular substructures (SMARTS) queries. Rooted in the information gain concept, already applied to derive selection rules in decision trees [2], we aimed at a better separation between classes of chemicals such as “drugs” and “non-drugs”. The most discriminating atom environments (having the highest information gain) were selected as model-free drug-like filters. These can be used to evaluate third party chemical libraries to assess drug-likeness.


COMP 142

Chemocentric informatics: Enabling bioactive compound discovery through structural hypothesis fusion

Alexander Tropsha(1), alex_tropsha@unc.edu. (1) School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill NC 27599, United States

Historically, computational drug discovery studies have relied on limited sources of data such as biological assays of compound libraries tested against single targets with results published in print. Nowadays, the information resources have broadened dramatically including large chemical genomics databases (e.g., ChEMBL, PubChem, PDSP, ToxCast), digital libraries (e.g., PubMed), gene expression profiles (e.g., cmap), and others. I shall describe a chemocentric informatics strategy integrating different information resources and diverse
computational methodologies towards discovering novel bioactive compounds. I shall describe the use of digital libraries for establishing new datasets to analyze the relationships between chemical structure and biological activity; highlight the importance of chemical data curation; and illustrate how computational models help spotting and correcting erroneous data. I will describe a study combining Quantitative Structure Activity Relationship (QSAR) modeling, virtual screening (VS), text mining, and gene expression profiling of chemicals for identifying novel experimentally confirmed high-affinity GPCR ligands as potential anti-Alzheimer drug candidates.

COMP 143

Computers and drug discovery: From duds to $5B drugs

Robert C Glen(1), rcg28@cam.ac.uk, Lensfield Road, Cambridge Cambridgeshire CB21EW, United Kingdom . (1) Department of Chemistry, University of Cambridge, Cambridge Cambridgeshire CB21EW, United Kingdom

Despite what you may think, given the investment in industrial scale pharmacology and chemistry, drug discovery is still a cottage industry. Small focussed groups of scientists combine diverse expertise from pharmacology and biology to synthesis and design, wrestling with complex and uncertain data. It is a poorly defined science, with undefined outcomes, often guided by rule-of-thumb, intuition and sheer luck. Bringing the logic of computation to the chaos of biology is very difficult, but every so often we succeed beyond our wildest dreams. Since this is the 50th anniversary of The Journal of Chemical Information and Modeling, I would like to review some of our work on novel algorithms and drug discovery, focussing on GPCR's, over the past twenty years and in particular identify some things that worked, some that didn't and also challenge some views of where modelling and computation should be applied, and where it shouldn't (yet).

COMP 144

Weighting and fusion methods for similarity-based virtual screening

Peter Willett(1), p.willett@sheffield.ac.uk, Western Bank, Sheffield South Yorkshire S10 2TN, United Kingdom ; Shereen Arif(1); John Holliday(1); Nurul Malim(1); Christoph Mueller(1). (1) Information School, University of Sheffield, Sheffield South Yorkshire S10 2TN, United Kingdom

Recent work in Sheffield on similarity searching has focussed on the use of data fusion and fragment weighting methods to search the MDDR, WOMBAT and MUV databases. Data fusion involves the combination of multiple similarity searches. The overlap between multiple searches is shown to follow a Zipf-like, power law distribution, with very few molecules (or active molecules) common to multiple searches; and a comparison of a large number of different group-fusion
algorithms shows that one based on molecules' inverse rank positions is the most effective of those tested. Information about the frequencies with which fragments occur in molecules can be used in two ways to increase search effectiveness (when compared with using just the presence or absence of fragments in molecules): using functions of the frequencies of fragment occurrences in individual molecules, and using inverse functions of the frequency of fragment occurrences in the database as a whole.

COMP 145

Lead Finder in the CSAR scoring challenge

Victor Stroylov\(^{(1)}\), rainman@moltech.ru, Leninskie gory, 1/75A, Moscow not applicable 119992, Russian Federation; Ghermes Chilov\(^{(1)}\); Oleg Stroganov\(^{(1)}\); Fedor Novikov\(^{(1)}\); Val Kulkov\(^{(2)}\). (1) "Molecular Technologies", Ltd, Moscow 119992, Russian Federation (2) BioMolTech, Corp., Toronto Ontario M2L 1L1, Canada

Lead Finder is a specialized software package for ligand docking, binding energy evaluation and virtual screening. The standard approach in estimation of binding affinities of protein-ligand complexes of the CSAR test set was the use of Lead Finder v.1.1.14 scoring mode that estimates free energy of protein-ligand binding for the fixed ligand coordinates for each protein-ligand complex. No pre-optimization of either protein or ligand structures were performed. The improvements in the scoring protocol included corrections of protein's and ligand's protonation states, positions of functional hydrogen atoms (for proteins only), and local geometry of nitrogen atoms (for ligands only). No other improvements of Lead Finder's the standard scoring function have been performed. The RMSD of estimated vs experimentally obtained protein-ligand binding energies was found to be equal to 2.07 kcal/mol and 1.98 kcal/mol for the standard and improved protocols correspondingly.

COMP 146

Benchmark of solvated interaction energy (SIE) scoring function on the CSAR-2010 dataset

Traian Sulea\(^{(1)}\), traian.sulea@nrc-cnrc.gc.ca; Qizhi Cui\(^{(1)}\); Herve Hogues\(^{(1)}\); Christopher R Corbeil\(^{(1)}\); Enrico O Purisima\(^{(1)}\). (1) Biotechnology Research Institute, National Research Council Canada, Montreal QC H4P 2R2, Canada

Solvated interaction energy (SIE) is a first-principle function for predicting absolute binding affinities from force-field non-bonded terms, continuum solvation, and scaling for configurational entropy. Standard SIE parametrization applied to the CSAR dataset with binding interfaces refined by constrained
minimization predicted absolute affinities with 2.5 kcal/mol mean-unsigned-error, but with correlation outperformed by buried surface or van der Waals interaction alone. Re-training SIE on CSAR subsets led to increased solute dielectric and reduced electrostatic interactions, stressing the weak signal carried by calculated electrostatics in this heterogeneous dataset. Overestimated complexes implicate highly negatively-charged ligands interacting via metals. Underestimated outliers reveal alternate protonation states that significantly improve SIE predictions. In an upgraded version of the CSAR dataset with reassigned protonation states, 10% of ligands and 20% of proteins are affected. Among other investigated aspects are the sensitivity to polar hydrogens orientation, incorporation of MD-generated ensembles, different solvent models and entropy estimates, and ligand strain.

COMP 147

Protonation states and scoring receptor-ligand poses: It's always the details

Emilio Xavier Esposito\(^{(1)}\), emilio@exeResearch.com, 32 University Drive, East Lansing Michigan 48823, United States . (1) exeResearch LLC, East Lansing Michigan 48823, United States

The protonation state of the receptor – ligand complex has a large influence over the correct approximation of the binding interactions. Using the CSAR dataset, various methods of assigning the complex's protonation state are used to explore the abilities of several scoring functions with respect to protonation state. In conjunction with the complex's protonation state, the 'standard' protocols employed to prepare a receptor for a docking simulation, along with the post-dock refinement of poses, are explored.

COMP 148

Role of active-site solvent in protein-ligand binding affinity calculations

Ye Che\(^{(1)}\), ye.che@pfizer.com, Eastern Point Road, Groton CT 06333, United States ; Veerabahu Shanmugasundaram\(^{(1)}\), (1) Groton Structural Biology, Antibacterials Chemistry/Discovery Technologies, Pfizer PharmaTherapeutics Research & Development, Groton CT 06333, United States

Accurate methods for computing binding affinities of a small molecule to a protein are needed to speed the discovery and optimization of new medicines. An assessment of six scoring functions commonly applied at Pfizer using the CSAR (Community Structure-Activity Resource) set of protein-ligand complexes will be presented. A current weakness amongst these various scoring functions is the treatment of active-site water molecules. Here, we quantitatively estimate the thermodynamic properties of active-site water molecules and capture the effects
of solvent displacement from the protein active site. Water inclusion shows promise in improving current scoring functions and we propose that this could be used more extensively in virtual screening and lead optimization applications.

**COMP 149**

**Flexible docking using a stochastic rotamer library of ligands**

*Nikolay V. Dokholyan*(1), dokh@med.unc.edu, Genetics Medicine, 120 Mason Farm Rd Ste 3097, Chapel Hill NC 27510, United States; Feng Ding*(1); Shuangye Yin*(1). (1) Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill NC 27599-7260, United States

Uncovering structures of molecular complexes via computational docking is at the heart of many structural modeling efforts and virtual drug screening. Modeling both receptor and ligand flexibility is important in order to capture receptor conformation changes induced by ligand binding, but is a major challenge in computational drug discovery. Many flexible docking approaches model the ligand and receptor flexibility either separately or in a loosely-coupled manner, which captures the conformational changes inefficiently. Here, we propose a truly flexible docking approach, MedusaDock, which models both ligand and receptor flexibility simultaneously using sets of discrete rotamers. We developed an algorithm which allows for the building of the ligand rotamer library “on the fly” during docking simulations. MedusaDock benchmarks demonstrate a rapid sampling efficiency and high prediction accuracy in both self-docking (to the co-crystallized state) and cross-docking (to a state co-crystallized with a different ligand), the latter of which mimics the virtual screening procedure in computational drug discovery. We also perform a virtual-screening test for a flexible protein target, cyclin-dependent kinase 2. We find a significant improvement in virtual screening enrichment when compared to rigid-receptor methods. The high predictive power of MedusaDock comes from several innovations, including the generation of a stochastic rotamer library of ligands, the efficient docking protocol, and the novel ligand pose-ranking method. We expect a broad adaption of these methodologies and the application of MedusaDock in ligand-receptor interaction predictions and drug discovery.

**COMP 150**

**Cheminformatics meets molecular mechanics: A combined application of knowledge based pose scoring and physical force field-based hit scoring functions improves the accuracy of virtual screening**
Many scoring functions fail to discriminate between true binders and non-binders (binding decoys), leading to a large number of false positive hits in virtual screening (VS) studies. We have developed a novel binary QSAR-like approach that discriminates geometrical pose decoys from native-like poses for each ligand. We have applied it for filtering (presumed) decoy poses from a library of docked ligand conformations followed by scoring the remaining poses with the MedusaScore physical force field-based scoring. We have demonstrated that this pre-filtering affords a significant improvement of hit rates in virtual screening studies for 5 of the 6 benchmark sets from the Database of Useful Decoys (DUD). Moreover, the top 10 hits in these 5 sets were found to include chemically diverse ligands while yielding high true positive rates (60-100%). We will discuss the methodology as well as the results of applying this approach to CSAR datasets.

COMP 151

Application of free energy methods to water molecules in protein binding sites

Jonathan W. Essex(1), jwe1@soton.ac.uk, Highfield, Southampton Hampshire SO17 1BJ, United Kingdom; Caterina Barillari(1); Russell Viner(2); Michael Bodnarchuk(1). (1) School of Chemistry, University of Southampton, Southampton Hampshire SO17 1BJ, United Kingdom (2) Jealott’s Hill International Research Centre, Syngenta, Bracknell RG42 6EY, United Kingdom

Water molecules play a crucial role in mediating the interaction between a ligand and a macromolecular receptor. An understanding of the nature and role of each water molecule in the active site of a protein could greatly increase the efficiency of rational drug design approaches. In this presentation, a range of different simulation methods, including double decoupling with replica exchange thermodynamic integration, Grand-Canonical Monte Carlo, and JAWS, are used to calculate the absolute binding free energies of a number of water molecules in protein-ligand complexes. The relative merits of each of these methods are discussed. In addition, the development of a number of descriptor-based QSAR models for calculating water binding free energies is described, with a view to reducing the need for expensive free energy simulations.

COMP 152
Which waters are important and how do we identify them?

**Simon Bowden**(1); **Jason C Cole**(1); **Tjelvar Olsson**(1), olsson@ccdc.cam.ac.uk, 12 Union Rd, Cambridge CB2 1EZ, United Kingdom; **Colin Groom**(1); **John Liebescheutz**(1); **Oliver Korb**(1). (1) Cambridge Crystallographic Data Centre, Cambridge CB2 1EZ, United Kingdom

The important role waters play in ligand binding both in terms of thermodynamics and selectivity is well known but identifying which waters are important for the success of a docking experiment is still difficult. Given that consideration of waters involved in primary and secondary mediated protein-ligand contacts has been shown to improve success rates in both native docking and virtual screening, experimenters need tools to help them decide which waters are important and which are not even real.

In this talk we will describe tools which may be of use to identify important waters and to highlight dubious waters. Conserved water structures can also be identified which may have an important influence on ligand binding. The effect of this information when applied to molecular docking will be demonstrated.

**COMP 153**

**Free energies and entropies of water molecules at protein-ligand interfaces**

**Steve W Rick**(1), srick@uno.edu; **Hongtao Yu**(1). (1) Chemistry, University of New Orleans, New Orleans LA 70148, United States

Water molecules are commonly found at the protein-ligand interface. The thermodynamics of these water molecules plays an important role in ligand affinity. In particular, the entropic cost of localizing a water molecule at the binding site can be significant. From the database of crystal structures, it is evident that the local environments of water molecules at the protein-ligand interface can vary considerably. We use molecule dynamics simulations and thermodynamic integration to calculate the free energy, enthalpy, and entropy changes associated with localizing a water molecule at a wide variety of sites at protein-ligand interfaces. Results analyzing how the free energies, enthalpies, and entropies depend on the details of the local environment, including the number of hydrogen bonds and the cavity size, will be presented.

**COMP 154**

**Role of water molecules in docking studies of Cytochromes P450**

**Chris Oostenbrink**(1)(2), chris.oostenbrink@boku.ac.at, Muthgasse 18, Vienna 1190. (1) Institute of Molecular Modeling and Simulation, BOKU University,
Active-site water molecules form an important component in biological systems facilitating promiscuous binding, or an increase in specificity and affinity. Taking water molecules into account in computational approaches to drug design or site-of-metabolism prediction is far from straightforward. The effect of including water molecules in molecular docking simulations of metabolic Cytochrome P450 enzymes is investigated, focusing on pose prediction, virtual screening and free energy estimates. The structure and dynamics of water molecules that are present in the active site simultaneously with selected ligands are described. The transferability of hydration sites between different ligands is investigated. The role of water molecules appears to be very dependent on the protein conformation and the substrate, further enhancing the versatility of these metabolic enzymes.

COMP 155

Modeling explicit waters in docking and scoring

Niu Huang\(^{(1)}\), huangniu@nibs.ac.cn, No. 7 Science Park Road, Zhongguancun Life Science Park, Changping District Beijing 102206, China . (1) National Institute of Biological Sciences, Beijing, Beijing 102206, China

Water molecules play an important role in protein-ligand recognition. However, incorporating explicit waters during docking is challenging in both the sampling and scoring aspects. We explored a method to switch ordered water molecules “on” (retained) and “off” (displaced) during docking screens. This method assumes additivity and scales linearly with the number of waters sampled despite the exponential growth in configurations. We tested this approach for ligand enrichment in screens of a large compound database against 24 DUD targets, exploring up to 8 waters in 256 configurations. Compared to calculations where the water positions were not sampled, enrichment factors increase substantially for 12 of the targets and are largely unaffected for most others. However, in our previous study, the positions of the water molecules were obtained from the x-ray structures, and all waters were treated as equally displaceable without the consideration of the differential energy of water binding. Our recent work in improving the treatment of waters during docking and scoring will be presented.

COMP 156

Desolvation/resolvation: A revolving door that controls the rates of association/dissociation of protein-ligand complexes? Analysis of PCSK9-EGF-A binding kinetics using WaterMap
We hypothesize that desolvation and resolvation processes can constitute rate-determining steps for protein-ligand association and dissociation, respectively. We tested this hypothesis using proprotein convertase subtilisin-kexin type 9 (PCSK9) bound to the epidermal growth factor-like repeat A (EGF-A) of low density lipoprotein cholesterol receptor (LDL-R). We analyzed and compared predicted desolvation properties of wild-type vs. gain-of-function mutant Asp374Tyr PCSK9 using WaterMap, a new method for calculating preferred locations and thermodynamic properties of water solvating proteins (“hydration sites”). We propose that fast $k_{on}$ and entropically driven thermodynamics observed for PCSK9-EGF-A binding is due to functional replacement of water occupying stable PCSK9 hydration sites (exchange of water for polar EGF-A groups). We further propose that relatively fast $k_{off}$ observed for EGF-A unbinding results from limited displacement of unstable water. Slower $k_{off}$ observed for EGF-A and LDL-R unbinding from Asp374Tyr PCSK9 may be due to destabilizing effects of this mutation on PCSK9 hydration sites.

**COMP 157**

**Biophysics-based library design: Discovery of “non-acid” inhibitors of S1 DHFR**

Veerabahu Shanmugasundaram$^{(1)}$, Veerabahu.Shanmugasundaram@pfizer.com, 300 Eastern Point Road, Groton CT 06333, United States; Kris Borzilleri$^{(1)}$; Jeanne Chang$^{(1)}$; Boris Chrunyk$^{(1)}$; Mark E Flanagan$^{(2)}$; Seungil Han$^{(1)}$; Melissa Harris$^{(1)}$; Brian Lacey$^{(3)}$; Richard Miller$^{(3)}$; Parag Sahasrabudhe$^{(1)}$; Ron Sarver$^{(1)}$; Holly Soutter$^{(1)}$; Jane Withka$^{(1)}$. (1) Groton Structural Biology, Antibacterials Chemistry/Discovery Technologies, Pfizer PharmaTherapeutics Research & Development, Groton CT 06333, United States (2) AntiBacterials Chemistry, Pfizer PharmaTherapeutics Research & Development, Groton CT 06333, United States (3) AntiBacterials Research Unit, Pfizer PharmaTherapeutics Research & Development, Groton CT 06333, United States

Methicillin-resistant Staphylococcus aureus (MRSA), the causative agent of many serious nosocomial and community acquired infections, and other gram-positive organisms can show resistance to trimethoprim (TMP) through mutation of the chromosomal gene or acquisition of an alternative DHFR termed "S1 DHFR". To develop new therapies for health threats such as MRSA, it is important to understand the molecular basis of TMP resistance and use that
knowledge to design and develop novel inhibitors that are effective against S1 DHFR. This presentation will highlight and illustrate an effort using a multi-pronged biophysics based strategy that utilizes NMR, thermodynamic, kinetic, structural, computational and medicinal chemistry information in developing an understanding of the mechanism of resistance in S1 DHFR as well as using this prospectively in drug discovery. Specifically this presentation will illustrate computational studies using WaterMap (WM) that developed an understanding of a key element of the mechanism of resistance that was supported by a variety of biophysical experiments and use of these WM calculations in a prospective fashion in library design.

**COMP 158**

**WaterMap: A new approach to mapping water structure, energetics, and free energies of displacement by ligands in protein active sites**

*Richard A. Friesner*\(^{(1)}\), rich@chem.columbia.edu, 3000 Broadway, New York NY 10027, United States. (1) Department of Chemistry, Columbia University, New York NY 10027, United States

We will discuss the use of molecular dynamics simulations, coupled with analytical approximations based on inhomogeneous solvation theory, to determine the location, structure, and approximate excess enthalpy and entropy of water clusters in protein active sites. The methodology, which we refer to as WaterMap, creates a map of water clusters in the active site, each of which is assigned an excess enthalpy and entropy as compared to bulk solution. In situations where water displacement dominates relative binding free energies (as is often the case), the approach can be used to estimate relative binding free energies of two or more ligands. Examples based on pharmaceutically interesting protein targets will be presented. The underlying theoretical rationale for the surprisingly good accuracy obtained, at a computational cost dramatically lower than simulation based alternatives such as free energy perturbation theory, will be discussed.

**COMP 159**

**QM/MM simulations of organic and enzymatic reactions**

*William L Jorgensen*\(^{(1)}\), william.jorgensen@yale.edu, 225 Prospect Street, New Haven CT 06520-8107, United States. (1) Department of Chemistry, Yale University, New Haven CT 06520-8107, United States

Quantum mechanics (QM) and Monte Carlo statistical mechanics (MC) simulations have been used since the early 1980s to study reaction mechanisms and the origin of solvent effects on reaction rates. By 2002 the QM and MC/MM calculations were fully integrated to obtain free-energy surfaces in solution with
no geometrical restrictions. Speed and accuracy demands also led to
development of the improved semiempirical QM method, PDDG-PM3. The
combined PDDG-PM3/MC/FEP methodology has provided excellent results for
free energies of activation for many reactions in numerous solvents. Examples
include Cope, Kemp and E1cb eliminations, S\textsubscript{N}2 and Diels-Alder reactions, as
well as enzymatic reactions catalyzed by the putative Diels-Alderase,
macrophomate synthase, and fatty-acid amide hydrolase [Acc. Chem. Res. 2010,
43, 142-151]. Recent results address Henry and intramolecular S\textsubscript{N}2 reactions,
Diels-Alder reactions on a water surface, and the tautomerase MIF. The
presentation will focus on the accuracy and mechanistic insights that can be
obtained in such QM/MM studies.

COMP 160

Recent advances in the development of a polarizable force field for
simulations of complex biomolecular systems

Benoit Roux\textsuperscript{(1)}, roux@uchicago.edu; Wei Jiang\textsuperscript{(2)}; David Hardy\textsuperscript{(3)}; Jim Phillips\textsuperscript{(3)};
Alex D. MacKerell\textsuperscript{(4)}; Klaus Schulten\textsuperscript{(3)}; Ed Harde\textsuperscript{(1)}; Janamejaya Chowdhary\textsuperscript{(1)};
Xiao Zhu\textsuperscript{(4)}; Pedro Lopes\textsuperscript{(4)}; Jihyun Shim\textsuperscript{(4)}; Chris Baker\textsuperscript{(4)}; Haibo Yu\textsuperscript{(1)}; Albert
Lau\textsuperscript{(1)}. (1) Biochemistry and Molecular Biology, University of Chicago, Chicago IL
60637, United States (2) Advance Leadership Computational Facility, Argonne
National Laboratory, Argonne IL 60439, United States (3) Physics, University of
Illinois in Urbanna-Champaign, Urbanna IL, United States (4) Pharmaceutical
Sciences, University of Maryland, Baltimore MD, United States

Classical simulations based on atomic models play an increasingly important role
in a wide range of applications in physics, biology and chemistry. They are
particularly valuable for the study of soft matter systems involving liquids,
polymers, membranes, microemulsions and surfactants, as well as complex
biomolecules like proteins and nucleic acids. A central issue in classical
simulations concerns the accuracy and predictive value of the force field used.
The latter are mathematical objects constructed from analytical functions, which
are parameterized to approximate the Born-Oppenheimer potential energy
surface and reproduce known experimental results. In practice the neglect of
induced polarization has severely limited the usefulness of MD simulations,
particularly when charged or highly polar species are involved. Furthermore,
induced electronic polarization is also critical to describe accurately the
electrostatic properties of the hydrophobic core of proteins and the hydrocarbon
region of membranes. Here, a simple and efficient implementation of induced
polarizability is developed and optimized for molecular dynamics simulations of
complex biomolecular systems. Electronic induction is represented by introducing
a classical charged Drude particles attached to the polarizable nuclei by a
harmonic spring. The nuclei carry an equal and opposite charge. The model can
be efficiently simulated by considering the dynamics of an extended Lagrangian
in which a small mass is attributed to the Drude particles. The approach for the
treatment of induced polarizability is implemented into the CHARMM program as well as in the highly parallelized code NAMD.

COMP 161

Estimating protein-ligand binding affinities using a fluctuating charge model

Harry A Stern(1), hstern@chem.rochester.edu, RC Box 270216, Rochester NY 14627, United States. (1) Department of Chemistry, University of Rochester, Rochester NY 14627, United States

Electrostatic polarization is a significant factor contributing to protein-small molecule binding. In this work, we develop and parameterize a fluctuating charge model based on density-functional theory calculations, and use it to estimate binding free energies for a set of proteins and small molecules with measured binding affinities and structures determined from crystallography. It is expected that such a model will be a useful tool for more accurate calculations of affinities, as well as providing insight into the forces driving protein-small molecule association.

COMP 162

Applications of variable timestep integrators and polarizable potentials in condensed phase simulations

Steven J Stuart(1), ss@clemson.edu, 369 Hunter Laboratory, Department of Chemistry, Clemson SC 29634, United States. (1) Department of Chemistry, Clemson University, Clemson SC 29634, United States

Accurate simulations of large-scale condensed phase systems require careful attention to both the potential model and the simulation techniques employed. Several modern simulation methods have their roots in approaches pioneered by Bruce Berne. For example, simulation efficiency can be greatly enhanced by taking advantage of heterogeneities in the timescales of the dynamics, with multiple-timestep methods such as rRESPA making use of spatial or physical inhomogeneities. Variable-timestep methods take advantage of temporal heterogeneities in non-equilibrium systems; applications to chemical sputtering and impacts with highly charged ions will be discussed. Heterogeneity in the solvation environment for biomolecular systems is another important challenge, which can be addressed with polarizable force fields. Results using fluctuating-charge water and peptide models demonstrate that incorporation of polarization affects solvation structure and solvent dynamics in the neighborhood of hydrophobic as well as hydrophilic interfaces, ranging in size from small dipeptides to large material surfaces.
Efficient methods for replica exchange and charge transfer models for liquid simulation

Steven W Rick(1), srick@uno.edu, 2000 Lakeshore Drive, New Orleans LA 70148, United States; Alexis J Lee(1). (1) Chemistry, University of New Orleans, New Orleans LA 70148, United States

Two recent projects are presented, both of which are influenced by the work of Bruce Berne. The first is a method for replica exchange. Replica exchange is a method that uses replicas across a range of temperatures (or Hamiltonians) which link the fast sampling replicas (at high temperatures, for example) to the replicas at the data collecting temperatures. Our method combines two lessons learned from Bruce: a particular scaling of the Hamiltonian to span temperature space and the extended Lagrangian technique for propagating extra system variables, in this case related to the Hamiltonian scaling. Applications of the method to protein conformational change will be presented. The second project is the development of a new model to treat charge transfer, beginning with the fluctuating charge approach developed in the Berne group over a decade ago. Charge transfer can be introduced by lifting the neutrality constraint on each molecule, but this creates a new set of difficulties. Methods to surmount these problems, involving the amount of charge transfer at large distances and the implementation using molecular dynamics, will be presented, along with results assessing the importance of charge transfer for liquid water and aqueous solutions.

Modeling co-translational folding of ribosome-nascent chain using AFM probe and computer simulations

Xiancheng Zeng(1), xiancheng.zeng@duke.edu, PO Box 90348, Durham NC 27708, United States; Whasil Lee(1); Huan-Xiang Zhou(2); Weitao Yang(1); Piotr E Marszalek(1); Vann Bennett(1). (1) Duke University, Durham NC, United States; (2) Florida State University, Tallahassee FL, United States

Polypeptide chains are synthesized on ribosomes in cells. The ribosome imposes a vectorial constraint on the co-translational folding of nascent chains that emerge from the ribosomal exit tunnel. To model the vectorial folding, we mechanically steered the relaxation of unraveled consensus ankyrin repeats. Excellent agreement is obtained in the force-extension spectra from AFM and steered molecular dynamics (SMD) with a structure-based coarse-grained model. Based on the AFM experiments and simulations, we found that the partially unraveled proteins refold from the N- to C-terminal sequentially by adding the
newly folding repeats to the folded stack. However, when relaxed from a fully stretched structure, the ankyrin repeats start folding by the nucleation of three N-terminal repeats after a significant delay. We identified the nucleation as the rate limiting step in the folding reaction. The nucleation is then followed by the sequential addition of the remaining repeats, which uses the nucleated core as the folding template.

**COMP 165**

**Multi-scaled explorations of binding induced folding of intrinsically disordered protein inhibitor IA3 to its target enzyme**

*Jin Wang*(1), jin.wang.1@stonybrook.edu, Stony Brook University, Stony Brook New York 11794-3400, United States. (1) Chemistry, Stony Brook University, Stony Brook New York 11794-3400, United States

Biomolecular function is realized by recognition. More and more evidences show function is not only determined by structure but also by dynamics. We explored biomolecular recognition involving large conformational changes-protein folding. In particular, we studied IA3, an intrinsically disordered protein capable of blocking the active site cleft of the protein aspartic proteinase saccharopepsin (YPrA) while its N-terminal residues fold into an amphipathic alpha helix. Combining structure based molecular dynamics simulations at the residue level and a stochastic path method at the atomic level, we developed a multi-scaled approach to explore the underlying mechanisms of binding induced folding of IA3. Both the free energy profile and the associated kinetic paths reveal a common scheme whereby IA3 binding to its target enzyme is prior to its own folding into helix. This theoretical result is consistent with the kinetic experiment. Furthermore, by exploring the landscape and detailed trajectories, we uncovered the important roles of non-native interactions in the initial binding prior to folding of IA3. Unlike the common thinking that non-native interactions contribute only to the roughness of landscapes and serve as negative factors for binding, the non-native interactions here contribute to non-specific binding, reducing significantly the entropy of searching space of the landscape and facilitating the binding process. The multi-scaled simulations of the process of intrinsically disordered proteins becoming structured in the presence of its target binding partner can help to provide information for the design of novel inhibitors of aspartic proteinases.

**COMP 166**
Ligand binding, hinge strain, and local unfolding: Driving forces of conformational change in adenylate kinase

Jason Brokaw(1), jbbrokaw@berkeley.edu, 02A Gilman Hall, Berkeley CA 94720, United States. (1) Department of Chemistry, University of California at Berkeley, Berkeley California 94720, United States

E. coli adenylate kinase (AK) rearranges its domains over large distances in its open-to-closed transition, making it a suitable model system to study the effects of ligand binding and local unfolding on protein conformational changes. Extensive all-atom molecular dynamics simulations from two different initial conformations both in the presence and absence of ligands simulate both partial closing and complete opening transitions without any applied external potential. These simulations provide evidence to characterize the sequential order of the motions two independent domains along two different transition pathways, the effects of local unfolding of an important hinge region on the closed-to-open transition of AK, and the changes in solvation during the transitions of AK. A coarse-grained model based on atomistic data identifies important interresidue coupling during the transition. Local unfolding and alternative interresidue contacts must be taken into account to explain protein conformational behavior.

COMP 167

Systematic identification of order parameters in biophysical systems

Athanassios Z. Panagiotopoulos(1), azp@princeton.edu, A317 Engineering Quad., Princeton NJ 08544, United States; Andrew L. Ferguson(1); Pablo G. Debenedetti(2); Ioannis G. Kevrekidis(2). (1) Department of Chemical Engineering and Institute for the Science and Technology of Materials, Princeton Univ., Princeton NJ 08544, United States (2) Department of Chemical Engineering, Princeton Univ., Princeton NJ 08544, United States

Cooperative couplings between degrees of freedom in biophysical systems lead to effective dimensionalities far less than the 3N-dimensional coordinate space of the constituent atoms, suggesting that the underlying dynamics of the system may be characterized by a relatively small number of order parameters. This effective reduction in dimensionality has been framed as a separation of timescales, whereby the fundamental dynamics reside in a "slow subspace" to which the other degrees of freedom are slaved. Geometrically, the slow subspace may be considered a - possibly highly convoluted - low-dimensional hypersurface, termed the "intrinsic manifold", to which the dynamics are effectively constrained. In this work, we apply a nonlinear dimensionality reduction technique known as the diffusion map (Coifman, R. R., et al. PNAS 102:7426, 2005) to systematically extract the intrinsic manifolds, and good order parameters with which to parameterize them, for a variety of systems of biophysical relevance. In particular, we study alkane chains, dialanine and an
antimicrobial lasso peptide in water. In all cases, novel dynamical features of the free energy landscapes are revealed.

COMP 168

Ab initio
QM/MM minimum free energy path for simulating chemical reaction and redox processes in solution and in enzymes

Weitao Yang(1), weitao.yang@duke.edu, 5310 French Family Science Center, Durham NC 27708, United States. (1) Department of Chemistry, Duke University, Durham NC 27708, United States

Combined QM/MM methods provide an accurate and efficient energetic description of complex chemical and biological systems, leading to significant advances in the understanding of chemical reactions in solution and in enzymes. Ab initio QM/MM methods capitalize on the accuracy and reliability of the associated quantum mechanical approaches, however at a much higher computational cost compared with semiempirical quantum mechanical approaches. Thus reaction path and activation free energy calculations encounter unique challenges in simulation timescales and phase space sampling. Recent developments of the QM/MM minimum free energy path method overcome these challenges and enable accurate free energy determination for reaction redox processes in solution and enzymes. Applications to several solution and enzyme reactions and redox processes will be presented.

COMP 169

Development and application of the string method to characterize conformational transition pathways in complex biomolecular systems

Benoit Roux(1), roux@uchicago.edu; Wenxun Gan(1); Luca Maragliano(1). (1) Biochemistry and Molecular Biology, University of Chicago, Chicago IL 60637, United States

The string method represents the transition pathway between two stable end-point conformations as an ordered sequence of replica of the systems, each parameterized by the value of a large set of order parameters (collective variables). The algorithm based on the swarms-of-trajectories evolves an initial guess for the path toward the most probable pathway by monitoring the average dynamical evolution of the order parameters for each replica along the path. Once the string has converged, it can be used as a basis for computing the free energy and the rate for the transition using a Markovian Milestoning with a Voronoi tessellation to obtain accurate estimates of the free energy profile and
transition rates. The computational framework combining the string method with the milestoning calculation makes it possible to identify the transition state and intermediates along the reaction using all-atom MD simulations. The framework is used to study two processes linked to the activation of Src tyrosine kinase, which are large signaling enzymes. Those are the opening of the activation loop and rotation of alpha-C helix near the catalytic site, and the flipping of the highly conserved and catalytically important DFG motif. In addition, the framework is used to characterize the large conformational transition of the voltage-sensing domain of the Kv1.2 potassium channel, simulated with an explicit lipid membrane.

**COMP 170**

**Computational evaluation of tautomers and zwitterions of D-amino acid oxidase (DAAO) inhibitors**

_Scot Mente_ (1), scot.mente@pfizer.com, 1 Eastern Point Rd., Groton CT 06340, United States. (1) Neuroscience Chemistry, Pfizer Global Research and Development, Groton CT 06340, United States

Quantum mechanical calculations and molecular docking were used in to design novel inhibitors of D-amino acid oxidase (DAAO). Using available x-ray structural information and simple tautomer enumeration tools, reasonable docked poses of a set of small ligands have been obtained. Use of these tools have helped lead to the optimization of the novel non-acidic 3-hydroxyquinolin-2(1H)-one Series (I), as well as the identification of structurally similar 3-hydroxyquinoline (II) and benzotriazole (III). Despite their small sizes, all three of these molecular scaffolds are capable of adopting multiple tautomer or zwitterionic states. The ability to accurately predict these states with quantum mechanical methods will be discussed.

---

**COMP 171**

**Defining states of ionization and tautomerization of thiamin diphosphate at individual reaction intermediates on enzymes: Enzymes that use a rare tautomeric form**
Frank Jordan, frjordan@rutgers.edu, 73 Warren Street, Newark NJ 07102, United States ; Natalia S. Nemeria; Anand Balakrishnan; Siakumar Paramasivam; Tatyana Polenova. (1) Chemistry, Rutgers University, Newark NJ 07102, United States (2) Chemistry and Biochemistry, University of Delaware, Newark DE 19716, United States

The author and coworkers demonstrated on several thiamin diphosphate (ThDP) enzymes that the 1',4'-iminopyrimidine tautomer of ThDP participates at several reaction steps. Hence, ThDP has dual function: an electrophilic covalent catalyst - a function long accepted- and an acid-base catalyst facilitating the ionization of the weak carbon acid to generate the C2 ylide.

It is proposed that ThDP exists in these forms on enzymes: the N1'-protonated 4-aminopyrimidinium (APH+) in protolytic equilibrium with its three conjugate bases, the canonical 4-aminopyrimidine (AP), its 1',4'-iminopyrimidine (IP) tautomeric form, and the C2 carbanion or ylide (Yl). The first three forms have been observed on multiple enzymes in the absence of substrate. In the presence of substrate and analogs, the IP form has been seen on several enzymes along with the APH+ state. Circular dichroism and solid-state NMR methods are being used for the first time to characterize different species. Supported by NIH-GM-050380 and 5P20RR017716.

COMP 172

Do tautomers matter in calculating molecular similarity?

Steven W Muchmore; Isabella Haight; Scott Brown. (1) Cheminformatics, Abbott Laboratories, Abbott Park IL 60048, United States

Compounds that have multiple tautomeric forms, which typically account for about 25% of pharmaceutical company corporate collections, present a challenge in cheminformatic analysis. While widely recognized, their manipulations are often ignored in database registration, substructure searching and similarity searching due to incremental increases in computation time and data management. However, clustering and diversity selection, which are based on similarity calculations, could yield erratic results if they include or exclude molecules that happen to be encoded as different tautomers. We enumerated tautomers for a data set of more than 66,000 compound pairs with associated activity against protein targets used in the assessment of similarity programs (Muchmore et al. J. Chem. Inf. Model. 2008, 48, 941). The similarity value for the highest scoring tautomer pair was compared to the original data to determine if its similarity score increased.

These tautomer similarity values were also applied to single representation results to determine if tautomer enumeration would yield a better estimate of the probability that two compounds will be equipotent.
Automated prediction of tautomeric states in protein-ligand complexes

Sascha Urbaczek(1), urbaczek@zbh.uni-hamburg.de, Bundesstrasse 43, Hamburg Hamburg 20146, Germany ; Stefan Bietz(1) ; Mathias Rarey(1). (1) Center for Bioinformatics, University of Hamburg, Hamburg Hamburg 20146, Germany

Hydrogen bonding plays a major role in the stabilization of protein-ligand complexes. Unfortunately, the positions of hydrogen atoms are not resolved in most structures present in the PDB. This makes it particularly hard to predict adequate tautomeric and protonation states for the atoms and groups involved in the binding. To overcome this difficulty many approaches have been developed to predict the correct protonation of either the ligand or the protein separately using a variety of different methodologies. We present a new method that predicts the tautomeric and protonation states as well as the resulting hydrogen atom positions of both the protein and the ligand simultaneously. The optimization of these states is based on an empirical scoring scheme used also in docking methods. Assuming an optimal hydrogen bonding network, the obtained results indicate that the most stable tautomeric forms in solution do not always correspond to those found in binding modes.

Predicting relative binding affinities in the CSAR Scoring Challenge

Matthew P Jacobson(1), matt.jacobson@ucsf.edu, 600 16th St., Box 2240, Genentech Hall, Room N472C, San Francisco CA 94158-2517, United States ; Chakrapani Kalyanaraman(1). (1) Department of Pharmaceutical Chemistry, UCSF, San Francisco CA 94158-2517, United States

We have been interested in evaluating whether all-atom force fields combined with implicit solvent models can be also used as a docking scoring function. Our prior experience has suggested that such energy functions can be used for, at best, predicting relative binding affinities to a particular binding site, with the best results being achieved for chemically related compounds, such as congeneric series generated in lead optimization. Thus, although predicting absolute binding affinities is a noble challenge, we have not attempted to do so in the CSAR exercise. Instead, with the assistance of the organizers, we focused on series of compounds bound to the same target. The results using the protein-ligand structures as provided
showed essentially no ability to rank order compounds by binding affinity. However, complete energy minimization, and in some cases correcting protonation states, significantly improved the results, to the point where there was some ability to distinguish more potent from less potent compounds, as we have also shown in other work on congeneric series. I will also discuss our attempts to characterize and correct some of the many limitations of this simple scoring scheme.

COMP 175

Surflex: Docking and scoring on CSAR

Ajay N Jain\(^{(1)}\), ajain@jainlab.org, 1450 3rd Street, MC 0128, San Francisco CA 94158, United States. (1) Bioengineering and Therapeutic Sciences, UCSF, San Francisco CA 94158, United States

One of the most challenging aspects of structure-based drug design is binding affinity prediction, since it embeds both the pose determination problem as well as requiring accuracy in estimation of energetic contributions where differences on the order of 1 kcal are large enough to matter. Even in the artificial case where a bound ligand/target structure is known, this remains a challenging problem. We present results for the Surflex family of methods for making predictions on the CSAR 2010 benchmark data set. Results will include straight docking-based pose prediction and scoring, tuned scoring approaches through scoring function optimization and protein structure optimization, and ligand-based approaches.

COMP 176

What we can learn from very large panel docking screens

John J Irwin\(^{(1)}\), jji@cgl.ucsf.edu, 1700 4th St, Box 2550, Room BH-501D, San Francisco CA 94158-2330, United States; Brian K Shoichet\(^{(1)}\); Michael M Mysinger\(^{(1)}\); Kong T Nguyen\(^{(1)}\). (1) Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco CA 94158-2330, United States

Whereas molecular docking is the most practical way to leverage structure for ligand discovery, the method retains important weaknesses. Among the more confounding problems is that docking can work well one target yet fail completely on the next, yet predicting in advance which will succeed or fail is challenging. To
investigate the strengths and weaknesses of docking we have assembled a very large panel of experimental information with which to test it. We have used our automated docking program, DOCK Blaster\(^1\), to study the performance of DOCK 3.5.54 against many protein targets for which experimental control information is available\(^2\). We have focused on two of the seven stated goals of the 2010 CSAR Workshop: to provide a baseline assessment of current scoring functions and to document which targets are most difficult. This approach has enabled us to comprehensively test the effect of changes in sampling, scoring and library composition.

References


**COMP 177**

Docking and scoring of fragments

*Marcel L Verdonk\(^{(1)}\), m.verdonk@astex-therapeutics.com, 436 Cambridge Science Park, Milton Road, Cambridge Cambridgeshire CB4 0QA, United Kingdom . (1) Astex Therapeutics Ltd, Cambridge CB4 0QA, United Kingdom*

Through the application of fragment-based drug discovery, Astex have produced >1,400 in-house X-ray crystal structures of fragments and >2,500 structures of lead-like compounds against a range of drug targets. From this wealth of structural data, we have constructed two test sets, each containing \(~100\) complexes, representing \(~10\) drug targets. In the first test set the ligands are fragments, whereas in the second test set the ligands are lead-like compounds. By applying docking and virtual screening on these sets, we will discuss whether fragments are harder to dock and score than larger compounds, and present our latest experiences on docking and scoring fragments. In addition, we will show how structural data on fragments obtained early on in drug discovery projects can be used to improve docking and scoring during the hit-to-lead phases. Finally, we will show examples of the application of docking and scoring of fragments on actual drug discovery programs.

**COMP 178**

Dynamical investigation of autoinhibitory mechanism of AMP-activated protein kinase
Cheng Peng(1)(2), chelanpeng@gmail.com, 260 Stanley Hall, Berkeley California 94720, United States; Teresa Head-Gordon(1)(3). (1) Bioengineering, University of California, Berkeley, Berkeley California 94720, United States (2) Computer Science and Engineering, Shanghai Jiao Tong University, Berkeley California 94720, United States (3) Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley California 94720, United States

It is structurally shown that autoinhibitory domain (AID) in the α-subunit of AMP-activated protein kinase (AMPK) plays a central role in the allosteric control of AMPK. However, the dynamical mechanism of AID is still unclear. Here we explore the dynamical role of AID in autoinhibition mechanism by using molecular dynamics (MD) simulation. It is found that AID interaction has significant impact on the conformational change of kinase domain (KD). The AID stabilizes the conformation of KD while the lost of AID interaction makes KD easier to undergo conformational change, which may be crucial to the function of AMPK.

COMP 179

Accounting for ligand conformational restriction in calculations of protein-ligand binding affinities

Harry A. Stern(1), hstern@chem.rochester.edu, RC Box 270216, Rochester NY 14627, United States. (1) Department of Chemistry, University of Rochester, Rochester NY 14607, United States

The conformation adopted by a ligand on binding to a receptor may differ from its lowest-energy conformation in solution. In addition, the bound ligand is more conformationally restricted, which is associated with a configurational entropy loss. The free energy change due to these effects is often neglected or treated crudely in current models for predicting binding affinity. We present a method for estimating this contribution, based on perturbation theory using the quasi-harmonic model of Karplus and Kushick as a reference system. The consistency of the method is checked for small model systems. Subsequently we use the method, along with an estimate for the enthalpic contribution due to ligand-receptor interactions, to calculate relative binding affinities. The AMBER force field and generalized Born implicit solvent model is used. Binding affinities were estimated for a test set of 233 protein-ligand complexes for which crystal structures and measured binding affinities are available. In most cases, the ligand conformation in the bound state was significantly different from the most favorable conformation in solution. In general, the correlation between measured and calculated ligand binding affinities including the free energy change due to ligand conformational change is comparable to or slightly better than that obtained by using an empirically-trained docking score. Both entropic and enthalpic contributions to this free energy change are significant.

COMP 180
Assessment of factors determining successful potency prediction in lead optimization: A case study with dual Abl/Src kinase inhibitors

Anna Kohlmann(1), anna.kohlmann@yahoo.com, 26 Landsdowne St., Cambridge MA 02139, United States; Xiaotian Zhu(1); Tianjun Zhou(1); Lois Commodore(1); Ling Xue(1); Yihan Wang(1); Wei-sheng Huang(1); Shuangying Liu(1); Mathew Thomas(1); Qihong Xu(1); Jeff Keats(1); Victor M. Rivera(1); Chester A. Metcalf(1); Tomi K. Sawyer(1); Tim Clackson(1); William C. Shakespeare(1); David C. Dalgarno(1). (1) ARIAD Pharmaceuticals, Inc., Cambridge MA 02139, United States

Computational potency prediction is integrated into the lead optimization cycle at many pharmaceutical companies. The promise of robust scoring methods and QSAR models is to maximize lead potency while minimizing the number of molecules synthesized in the optimization process. However, obtaining a predictive model even for series with information-rich SAR can be challenging.

To understand factors that influence the predictive ability of scoring methods in lead optimization, MM-GBSA, Watermap and a combination of physics-based interaction terms are applied to series of in-house Src and Abl kinase inhibitors. This data set contains over 200 molecules with high-quality enzyme and cellular assay data. We demonstrate successful potency prediction on the set of Src inhibitors bound to a rigid receptor, and discuss the difficulties for predicting the impact of subtle modifications. Conformational changes in the protein that impede the accurate prediction of affinity are also addressed in the example of Abl inhibitors. Finally we explore the utility of incorporating induced-fit calculations into the potency prediction workflow.

COMP 181

Accelerated MD simulation study of the maltose-binding protein

Denis Bucher(1), bucher.denis@gmail.com, Gilman Drive, La Jolla California, United States; Andrew J McCammon(1); Phineus Markwick(1); Barry Grant(1). (1) Chemistry and Biochemistry, University of California, San Diego, La Jolla, United States

Large-scale domain rearrangements in proteins have long been recognized to have a critical function in ligand binding and recognition, catalysis and regulation. Recently the open-to-closed transition in apo maltose-binding protein has been observed by paramagnetic NMR [1]. The timescale of the conformation change (20 ns - 20 ms) is at the limit of what can be observed with conventional MD simulations. Here, we have used accelerated MD simulations, a method recently developed in the McCammon group [2], to enhance the sampling of this slow conformational transition.
Changes of human receptor binding affinity of H1N1 hemagglutinins: Insights from molecular dynamics simulation

Nadtanet Nunthaboot\textsuperscript{(1)}, nadtanet@gmail.com, Faculty of Science, Kham-Riang, Kantarawichai Mahasarakham 44150, Thailand; Thanyada Rungrotmongkol\textsuperscript{(2)}; Maturos Malaisree\textsuperscript{(2)}; Nopphorn Kaiyawet\textsuperscript{(2)}; Panita Decha\textsuperscript{(3)}; Pornthep Sompornpisut\textsuperscript{(2)}; Yong Poovorawan\textsuperscript{(4)}; Supot Hannongbua\textsuperscript{(2)}. (1) Department of Chemistry, Mahasarakham University, Kantarawichai Mahasarakham 44150, Thailand (2) Department of Chemistry, Chulalongkorn University, Bangkok 11330, Thailand (3) Department of Chemistry, Thaksin University, Phattalung, Thailand (4) Center of Excellence in Clinical Virology, Chulalongkorn University, Bangkok 11330, Thailand

The binding of the viral glycoprotein, hemagglutinin (HA), and the human-α2,6-linked sialopentasaccharide host cell receptor (hHAR) is a critical step in the viral replication cycle. Dynamical and structural properties of the four different HAs of Spanish 1918 (H1-1918), swine 1930 (H1-1930), seasonal 2005 (H1-2005) and a novel 2009 (H1-2009) H1N1 bound to the hHAR, were investigated by means of molecular dynamics simulations. In all systems, major interactions between HA residues and hHAR were obtained from Y95 and the conserved residues of the 130-loop, 190-helix and 220-loop. Compared to the three previously recognized H1N1 strains, introductions of K145 and E227, charged residues, increased the HA-hHAR binding efficiency of 2009 HA H1N1. In addition, changing of G225, the non-charged residue, to D225, a negatively charged residue, provides a larger number of hydrogen bonding interactions. The obtained information could help the understanding of how different HAs effectively attach and bind with the hHAR.

Benchmark set for validating and testing free energy methods in molecular design

Michael R. Shirts\textsuperscript{(1)}, michael.shirts@virginia.edu, P.O. Box 400471, Charlottesville VA 22904, United States; Himanshu Paliwal\textsuperscript{(1)}. (1) Department of
There has been intense interest in the simulation and theory communities in determining the most efficient and reliable methods to perform free energy calculations. I will present our group's development of the first version of a benchmark set for calculating free energies of molecular transformations in solution. Included are tests such as methane solvation, dipole inversion, hydrogen bonding networks, and disappearance of large aromatic groups. We also discuss our use of this set to compare the efficiency of a range of equilibrium published free energy methods, including thermodynamic integration, free energy perturbation, the weighted histogram analysis method, the Bennett acceptance method, transition matrix approaches, and the Wang-Landau based methods. We compare methods for estimating the variance such as the bootstrap method and analytical error estimates with direct evaluation of the statistical uncertainty over many copies. An important part of the test set is distribution of files for a number of different computational platforms.

**COMP 184**

**Reaction network analysis and the application of structure-based screening methods for the butanoate metabolic pathways**

**Qin Wang**¹, wangqn@bc.edu, 2609 Beacon St, Merkert Chemistry Center, Chestnut Hill Massachusetts 02467, United States; **Di Wu**², **Rajeev Surendran Assary**², **Linda J Broadbelt**², **Goran Krilov**³. (1) Department of Chemistry, Boston College, Chestnut Hill Massachusetts 02467, United States (2) Department of Chemical and Biological Engineering, Northwestern University, Evanston Illinois 60208, United States (3) Schrodinger, Inc., New York New York 10036, United States

Generation of biochemical reaction pathways consisting of both known and novel reactions to synthesis of carbon-based compounds achieved tremendous interest in this bio-fuels era. Biomass can be converted to energy-rich alcohols using process such as bacterial fermentation involves a series of enzymatic reactions. The concept of generating network of compounds or candidate fuels from an affluent biological compound involve the structural based understanding the chemical transformation. Here we present such as study, where the reaction network generated using generalized enzyme reactions rule from pyruvic acid to 1-butanol and employed a hierarchical computational screening strategy of, network generation, screening based on reaction pathway length and thermodynamics based on group contribution method and various structural based screening tools.

**COMP 185**
Importance of solvent and dynamics on the dehydrogenation of ammonia borane

Paul M Zimmerman(1), zimmerman@stanford.edu, 380 Roth Way, Stanford CA 94305, United States ; Zhiyong Zhang(2); Charles B Musgrave(3). (1) Department of Chemical Engineering, Stanford University, Stanford CA 94305, United States (2) Stanford Nanofabrication Facility, Stanford University, Stanford CA 94305, United States (3) Department of Chemical and Biological Engineering, University of Colorado at Boulder, Boulder CO 80301, United States

The reactivity of ammonia-borane (AB) dimers that leads to dehydrogenation is investigated using ab initio simulations to answer the following three questions: Do the favorable reaction paths differ from gas phase to solvent? Are the products of specific elementary steps affected by nuclear dynamics or do they follow zero-Kelvin minimum energy pathways? When accounting for solvent, what intermediates best account for the dehydrogenation of AB? We show that in ethereal solvent the hydrogen-bonded dimer of AB reacts via two pathways, one leading to its ion-pair isomer, the diammoniate of diborane (DADB) and the other to direct decomposition to \( \text{NH}_2\text{BH}_2 \), \( \text{H}_2 \), \( \text{NH}_3 \) and \( \text{BH}_3 \). Notably, in the gas phase the transition state for DADB formation proceeds through molecular dynamics directly to DADB decomposition products. In addition, the decomposition of DADB in gas and solvent phases dynamically yields products not predicted by the lowest energy 0K energy pathway. We examine how \( \text{NH}_3 \) or \( \text{BH}_3 \) are potential catalysts for AB dehydrogenation and demonstrate the conditions in which these species are likely to catalyze significant AB dehydrogenation. The effects of solvent and dynamics are shown to have fundamental importance in the understanding uncatalyzed AB dehydrogenation. Furthermore, the deviation from minimum energy 0K gas phase reaction pathways is highlighted as a potential problem in all studies that do not account for solvation and dynamics in the theoretical investigations.

COMP 186

pKa predictions of some organic amine solvents

Mohamed Elanany(1), mohamed.elanany@aramco.com, 2296 Bld, Dhahran Eastern Province, Saudi Arabia . (1) R&D Center, Saudi Aramco, Dhahran, Saudi Arabia

Alkanolamines are used in industry to capture CO2. For this application pKa is an important parameter to be measured experimentally. Now Computational modeling is used to calculate pka values of different amine solvents. In addition it
allows solvent screening and designing more efficient solvents. Here, Numerical basis set in Materials studio software is used to predict the pKa values of organic amine solvents. Results from different GGA functional are compared. Preliminary results indicate a better prediction in the case of branched amines as compared to cyclic ones.

COMP 187

Simulated back proton transfer in 3-Hydroxyflavone

Yann Tambouret(1), yannpaul@bu.edu, 725 Commonwealth Avenue, Boston MA 02215, United States ; David Coker(2), (1) Center for Space Physics, Boston University, Boston MA 02215, United States (2) Department of Chemistry, Boston University, Boston MA 02215, United States

3-Hydroxyflavone is the prototypical example of a system that undergoes an excited state, intramolecular proton transfer. Upon fluorescence, the non-equilibrium state relaxes through a back proton transfer. A model is presented for the ground state potential energy surface based on the Cartesian Reaction Plane Hamiltonian, in which both reactive, anharmonic and backbone, harmonic vibrations are characterized. The reaction is simulated using the Multi-Configurational Time-Dependant Hartree method. The dynamics on the surface qualitatively reproduces published emission spectra, but the B3LYP/DFT-parametrized surface has a smaller tunneling barrier than that found in the physical system. This results in a calculated reaction rate that is faster than that measured experimentally.

COMP 188

Ab initio study of hydrogen bonding cooperativity in arylamide

Jessica Amber Geer(1), jgeer@usp.edu, 600 S. 43rd St., Chemistry Department, Philadelphia PA 19104, United States ; Jhenny Galan(1), Zhiwei Liu(1), Vojislava Pophristic(1). (1) Chemistry & Biochemistry, University of the Sciences in Philadelphia, Philadelphia PA 19104, United States

Aromatic oligoamide foldamers are synthetic oligomers that adopt stable secondary structures in solution and can be designed to have therapeutic applications. The design process draws on the principles that govern biopolymer shapes and thus then leads to structures that mimic motifs found in biochemical systems. Among many strategies used in foldamer design, hydrogen bonding (H-bonding) has served as a very effective designing tool due to the strength and directional characteristics found in hydrogen
bonding interactions. We investigate various intramolecular H-bond patterns, with a focus on the influence that one H-bond has on the strength of another, shared one. A comprehensive ab initio study followed by a Natural Bond Orbital (NBO) analysis has been performed on diarylamide model compounds. Our analysis demonstrates to what extent cooperativity between shared H-bonds exist in this type of foldamer units. Using our torsional profile and NBO analysis, we will discuss cooperativity of the shared H-bonds.

COMP 189

Drug discovery technology: The past, the present, and the future

Chris Lipinski(1), 10 Connshire Drive, Waterford Connecticut 06385, United States . (1) Connecticut 06385, United States

Where is early drug discovery technology heading? In the early 1970's we had phenotypic in vivo screening with little or no mechanistic screening and compounds entered the clinic and succeeded at rates not too different from today. Fast forward to the early 1990's and the combinatorial chemistry disaster where drug discovery progress was slowed by very poor chemistry quality. Concomitant to early combichem was the growth of HTS and the gradual evolution and improvement in the quality of screening. By year 2000 we had the genomics explosion. This was great for biology in general but arguably in the short term had little or maybe even a negative effect on drug discovery productivity. Also around 2000 we had the beginnings of fragment screening. Unlike many new technologies this actually turned out to be a big success with its impact mostly limited by the specialized nature of the required screening technology. Concomitant to all the experimental advances (and setbacks) we saw marked improvements on the computational side with the widespread ability to use target structural information in drug design. These computational advances were paralleled by experimental advances in the ease of getting target structural information. The trend to broader access to drug discovery technology across all users continued in the current decade with the explosion of drug discovery screening efforts in academia. In the USA, drivers for this change were the NIH Roadmap screening efforts and the realization by academics that university biology intellectual property is greatly enhanced when there is screening data on chemistry subject matter. Looking into the future I think I can see where chemistry might be heading. We are beginning to see the exploitation of compounds (eg. MWT =10K) that are in between the size of small molecules and biologicals. The impact of conformationally restrained peptides and the various flavors of RNA will depend on progress in tissue delivery. We already know that structurally complex natural products are particularly rich in biology and that the information content of cyclic compounds is much richer than in acyclic compounds. How to cost effectively exploit this knowledge is the issue. Finally, it is really gratifying to me to see that drug discovery for rare, orphan and
neglected disease is finally getting the attention of all the components in the drug discovery process.

COMP 190

Generating novel compounds via rule-based molecular transformations

Joe Leonard(1), jle@chemcomp.com, 1010 Sherbrooke Street West, Suite 910, Montreal Quebec H3A 2R7, Canada; Alex Clark(1). (1) Chemical Computing Group, Montreal Quebec H3A 2R7, Canada

Small molecule scaffold replacement techniques are an important part of drug discovery because of the need to find rapid follow-on compounds or alternate chemical series. Previous approaches used random structural exchange, pharmacophore-based searching or even full de novo methods to generate new structures. We present a method that combines the cheminformatics approach of Drug Guru (Stewart et al) with techniques from ligand/receptor docking. Transformations are drawn from an extensible library of medicinal chemistry design rules of thumb, the resulting compounds can be filtered by various 2D and/or 3D criteria and ligand refinement and scoring can be done. Example of rule creation are discussed, along with examples of transformations relating known drugs.

COMP 191

Vibrational spectroscopy of water and ice

James Skinner(1), skinner@chem.wisc.edu, 74 Cambridge Rd, Madison WI, United States. (1) University of Wisconsin, United States

Vibrational spectroscopy (IR, Raman, 2DIR, pump-pump, SFG, etc.) of water provides useful information about structure and dynamics. Spectra for HOD in H2O or D2O are particularly simple and informative, since the effects of intramolecular and intermolecular vibrational coupling are minimized. On the other hand, spectra for neat H2O or D2O show the fascinating and important effects of vibrational coupling and population transfer. I will focus on the neat systems in this talk, considering the IR and Raman spectra of water and ice, the SFG spectrum of the liquid/vapor interface, and 2DIR spectra and pump-probe rotational anisotropy decay for liquid water.

COMP 192

Rates of reactions activated by non-Gaussian noise

Hermann Grabert(1)(2), grabert@physik.uni-freiburg.de, Hermann Herder Strasse 3, Freiburg D-79104, Germany; Daniel F Urban(2). (1) Freiburg Institute for
While the standard approach to fluctuations in mesoscopic and molecular systems assumes Gaussian fluctuations of the state variables, more recently deviations from Gaussian noise have attracted widespread interest. This development is motivated by the fact that deviations from Gaussian noise – encoded in higher order noise cumulants – provide valuable information about the microscopic processes causing the fluctuations. An important question in this context is how the standard theory of thermally activated escape from a metastable potential well is modified when the system is driven by non-Gaussian noise. The talk will address this issue and present results on escape rates in presence of a Poissonian noise component in addition to standard Gaussian noise. In particular, the modification of the rate of escape by a finite third noise cumulant will be clarified.

**COMP 193**

**Dynamics of photoinduced proton-coupled electron transfer reactions**

*Sharon Hammes-Schiffer*(1), shs@chem.psu.edu, 104 Chemistry Building, University Park PA 16802, United States. (1) Department of Chemistry, Pennsylvania State University, University Park PA 16802, United States

Theoretical studies of photoinduced proton-coupled electron transfer (PCET) reactions in solution and at molecule-semiconductor interfaces will be presented. The quantum mechanical effects of the active electrons and transferring protons, as well as the dynamics of the solvent environment and relevant solute modes, are included in these calculations. The relaxation process following photoexcitation is described in terms of nonadiabatic dynamics on electron-proton vibronic surfaces. This approach provides detailed mechanistic information and enables the investigation of different reaction pathways for processes involving the coupling of electrons and protons in condensed phases. The hydrogen/deuterium isotope effects on the ultrafast dynamics can also be examined with this approach. Applications to experimentally studied PCET reactions will be discussed.

**COMP 194**

**Theory of rainbow scattering from surfaces**

*Eli Pollak*(1), eli.pollak@weizmann.ac.il, Herzl St., Rehovoth n/a 76100, Israel. (1) Chemical Physics Department, Weizmann Institute of Science, Rehovoth 76100, Israel
For many years, the theory of rainbow scattering from surfaces was limited to hard wall potentials. Following work of Prof. Berne, we employ a model in which the scattered particle interacts with the surface continuously. Surface phonons are modeled as a linearly coupled harmonic bath. A classical perturbation theory leads to explicit expressions for the deflection functions and rainbow angles and their dependence on incident scattering angle, energy and surface temperature. The theory was applied to experimentally studied systems, including Ar and Ne on Cu; Ar on Ag(111), LiF(100) and 2H-W(100); Xe on Ge(100); and Kr on Ag(100). It is also developed for 3D scattering and rotational rainbows. The theory led to the discovery of energy loss rainbows, friction-induced rainbow scattering and angular peaks induced by rotational rainbows. In this talk we discuss the theory and new developments and give an experimental demonstration of rainbow scattering from corrugated surfaces.

COMP 195

Heat transfer, heating and cooling in molecular conduction junctions

Abraham Nitzan(1), nitzan@post.tau.ac.il, School of Chemistry, Tel Aviv 69978, Israel. (1) Chemistry, Tel Aviv University, Tel Aviv 69978, Israel

Heating in molecular conduction junction depends on the balance between the rate of heat deposit by the electronic current and the efficiency of heat conduction away from the junction. I will review our recent work on such processes and focus on models for current induced cooling in such systems.

COMP 196

Sorting out electrostatic polarization, dispersion, and hydrogen bonding in solvatochromic shifts

Donald G. Truhlar(1), truhlar@umn.edu, 207 Pleasant ST. SE, 139 Smith Hall, Minneapolis MN 55455-0431, United States; Aleksandr V. Marenich(1); Christopher J. Cramer(1). (1) Department of Chemistry, University of Minnesota, Minneapolis MN 55455-0431, United States

Polarized continuum model (PCM) calculations of solvatochromic shifts on electronic excitation energies include the noninertial and inertial bulk-solvent polarization, which will be called electrostatics, but not dispersion interactions and specific effects like solute-solvent hydrogen bonding. For the \(n\) to \(\pi^*\) excitation of acetone in a diverse set of solvents, we added the nonelectrostatic contributions by dispersion and hydrogen-bonding descriptors, as in J. Li et al., Int. J. Quantum Chem. 2000, 77, 264. The electrostatics include both the fast and slow polarization as well as their interaction, and the solute is treated by time-dependent density
functional theory (TD-DFT). For aqueous solution and for solvation in methanol and heptane, we also estimated the solvatochromic shift by ensemble averaging over supermolecule calculations with up to twelve explicit solvent molecules selected from a molecular dynamics run, with the explicit solvent surrounded by a continuum solvent. The TD-DFT calculations are carried out with the M06 density functional, which has been validated for the treatment of attractive noncovalent interactions. One possible choice of atomic radii parameters are the SMD atomic radii, which were adjusted in previous work to give good free energies of solvation for ions. For 23 solvents, this gives a dispersion contribution to the red of 261–356 cm\(^{-1}\) and a specific hydrogen bonding contribution to the blue of up to 289 cm\(^{-1}\). A larger hydrogen bonding contribution must be invoked if one uses the default radii of Gaussian 09. Other choices of atomic radii and electrostatic treatments will also be considered, and the size of nonelectrostatic effects and the accuracy attainable for predicting solvatochromic shifts and vertical excitation energies in solution will be discussed. This work was supported in part by the Army Research Office.

COMP 197

Multiscale simulation of membrane remodeling by proteins

Gregory A Voth\(^{(1)}\), gavoth@uchicago.edu, 5735 S. Ellis Avenue, Chicago IL 60637, United States. \(^{(1)}\) Department of Chemistry, James Franck Institute, and Computation Institute, University of Chicago, Chicago IL 60637, United States

Many cellular processes occur on mesoscopic time and length scales, relative to atomic dimensions. However, the molecular-level interactions are still very important in defining the nature of the cellular processes. Membrane remodelling by proteins (e.g., the BAR domain, a long banana-shaped homodimer) is one such example. The amphiphysin and endophilin N-BAR domains (the BAR domain plus an N-terminal amphipathic helix) are thought to play a key role in clathrin-mediated endocytosis. We have developed and applied a "bottom-up" multiscale approach consisting of atomistic molecular dynamics (MD), coarse-grained (CG), and mesoscopic simulations to explore membrane remodelling at various length and timescales, in connection with experimental results such as from electron microscopy obtained by the Unger group at Yale. The BAR domain is shown to be a protein module whose shape and electrostatic properties result in liposome tubulation and vesiculation at mesoscopic levels, in agreement with various experimental results. The atomistic-level MD simulations demonstrate the local membrane bending ability of both amphiphysin and endophilin N-BAR domains, while the CG simulations, that are systematically constructed from the atomistic MD data, allow much large scale phenomena to be studied. The end point mesoscopic simulations reveal collective interactions that lead to complex membrane remodelling.
phenomena. Preliminary results from large-scale atomistic MD and CG simulations for the membrane remodelling behavior of the epsin N-terminal homology (ENTH) domain will also be presented.

COMP 198

Self-limited self-assembly of chiral filaments

Yasheng Yang(1), yasheng@gmail.com, 415 South Street, Waltham MA 02453, United States; Robert Meyer(1); Michael Hagan(1). (1) Department of Physics, Brandeis University, Waltham MA 02453, United States

The assembly of filamentous bundles with controlled diameters is common in biological systems and desirable for the development of nanomaterials. We discuss dynamical simulations and free energy calculations on patchy spheres with chiral pair interactions that spontaneously assemble into filamentous bundles. The chirality frustrates long-range crystal order by introducing twist between interacting subunits. For some ranges of system parameters this constraint leads to bundles with a finite diameter as the equilibrium state, and in other cases frustration is relieved by the formation of defects. While some self-limited structures can be modeled as twisted filaments arranged with local hexagonal symmetry, other structures are surprising in their complexity. We discuss the relation between model structures and finite bundles in biological or biomaterials systems, and implications for the design of nanostructured materials with controlled sizes.

COMP 199

Combining theory, computation, and experiment for studying protein aggregation

Rohit V Pappu(1), pappu@wustl.edu. (1) Department of Biomedical Engineering, Washington University in St. Louis, St. Louis MO 63130, United States

This talk will focus on a series of new results that demonstrate the effects of sequence context and chain length on the conformational equilibria and aggregation of proteins containing polyglutamine expansions. Atomistic and mesoscopic simulations are being combined with theoretical tools from polymer physics and in vitro experiments to develop an integrated mechanistic picture of polyglutamine-mediated aggregation.

COMP 200

Understanding the effect of cholesterol on the properties of membranes
Berend Smit\(^{(1)}\), Berend-Smit@Berkeley.edu, 101B Gilman #1462, Berkeley CA 94720-1462, United States; Frederick de Meyer\(^{(1)}\); Jocelyn M Rodgers\(^{(1)}\); Ayelet Benjamini\(^{(1)}\). (1) Department of Chemical Engineering, University of California Berkeley, Berkeley CA 94720-1462, United States

We present a coarse-grained model of a hydrated saturated phospholipid bilayer (dimyristoylphosphatidylcholine, DMPC) containing cholesterol that we study using a hybrid dissipative particle dynamics - Monte Carlo method. This approach allows us to reach the time and length scales necessary to study structural and mechanical changes of the bilayer at various temperatures and cholesterol concentrations. The properties studied are the area per lipid, condensation, bilayer thickness, tail order parameters, bending modulus and area compressibility. Our model quantitatively reproduces most of the experimental effects of cholesterol on these properties, and reproduces the main features of the experimental phase diagrams. We also present all-atom simulation results of the system and use these results to further validate the structure of our coarse-grained bilayer. Based on the changes in structural properties we propose a temperature-composition structure diagram, which we compare with the experimental phase and structure diagrams. Attention is paid to the reliability and interpretation of the model and simulation method and of the different experimental techniques.

In addition, we investigate the effect of cholesterol on the interactions of membrane proteins.

COMP 201

Capsid assembly around RNA and other polyelectrolytes: Multiple roles of the polymeric cargo

Michael F Hagan\(^{(1)}\), hagan@brandeis.edu, 415 South St., MS 057, Waltham MA 02454, United States; Oren M. Elrad\(^{(1)}\); Matthew Perkett\(^{(1)}\). (1) Department of Physics, Brandeis University, Waltham MA 02454, United States

During the replication of many viruses, hundreds to thousands of protein subunits assemble around the viral nucleic acid to form a protein shell called a capsid. Most viruses must form a particular structure to be infectious and do so with astonishing fidelity. Furthermore, recent experiments demonstrate that capsids can assemble around synthetic cargoes such as nanoparticles and inorganic polyelectrolytes. The assembly mechanisms that enable this combination of adaptability and precision are poorly understood.

This presentation will explore two aspects of capsid assembly around polymeric cargoes. First, we use Brownian dynamics with a coarse-grained model to describe the dynamics of capsid assembly around a flexible polyelectrolyte. We identify several mechanisms by which the polymer plays an active role in its
encapsulation, including cooperative polymer-protein motions. The importance of these mechanisms is related to experimentally controllable parameters such as polymer length, protein concentration, and solution conditions. Furthermore, assembly mechanisms are correlated to the efficiency of encapsulation, and we present a "phase diagram" that predicts assembly outcomes as a function of experimental parameters.

For most viruses, assembly into the infectious morphology requires capsid proteins to adopt different conformations, which are arranged in a particular geometry during assembly. Recent experiments suggest that binding to the viral RNA directs these conformational changes in MS2 proteins. We explore this possibility using atomic-resolution molecular dynamics simulations combined with enhanced sampling techniques. We will discuss how these models at different levels of resolution can be integrated to give an overall picture of how the capsid assembly process is directed toward a particular morphology.

COMP 202

Intrinsic bending states of tubulin dimers and protofilaments

Andrea Grafmueller(1), agrafmueller@uchicago.edu, James Franck Institute, and Computation Institute, 5735 S Ellis Ave, Chicago IL, United States ; Gregory A. Voth(1). (1) Department of Chemistry, James Franck Institute, and Computation Institute, University of Chicago, Salt Lake City IL 60637, United States

Microtubules are long, stiff polymers of ab-tubulin heterodimers that exhibit a “dynamic instability”, in which its ends switch stochastically between growth and shrinking. This complex dynamic behavior is linked to the hydrolysis of the guanosine triphosphate (GTP) nucleotide bound to the b-tubulin. The underlying structural mechanism, however, is not well understood.

Due to their size and the long timescale dynamics of these assemblies, simulations of these protein assemblies inherently require a multiscale approach. We use large scale atomistic simulations to explore the mechanical properties of tubulin dimers and short protofilaments for different structural states and nucleotide content. To propagate the effects of these observations to the dynamics of whole microtubules, coarse grained models of the different states of the tubulin dimer are systematically developed based on the dynamics of the underlying atomistic system.
COMP 203

Are the laws of chemistry and physics sufficient for the discovery of new drugs?

**Gerry Maggiora**\(^{(1)(2)(3)}\), gerry.maggiora@gmail.com, 1703 E. Mabel Street, Tucson Arizona 85721-0207, United States. (1) Department of Pharmacology & Toxicology, University of Arizona College of Pharmacy, Tucson Arizona 85721-0207, United States (2) BIOS Institute, Tucson Arizona 85721-0207, United States (3) Translational Genomics Research Institute, Tucson Arizona 85721-0207, United States

Reductionism is alive and well in drug-discovery research. In that tradition, we continually improve experimental and computational methods for studying smaller and smaller aspects of biological systems. Although significant improvements continue to be made, are our efforts too narrowly focused? Suppose all error could be removed from these methods, would we then understand biological systems sufficiently well to design effective drugs? Currently, almost all drug research focuses on single targets. Should the process be expanded to include multiple targets? Recent efforts in this direction have lead to the emerging field of polypharmacology. This appears to be a move in the right direction, but how much polypharmacology is enough? As the complexity of the processes underlying polypharmacology increase will we be able to understand them and their inter-relationships? A number of these questions will be addressed in the talk, which focuses on issues and questions not answers to the drug-discovery conundrum.

COMP 204

How should the world of molecular modeling be organized? Thoughts and speculations

**Anthony Nicholls**\(^{(1)}\), anthony@eyesopen.com, 9 Bisbee Court, Suite D, Santa Fe New Mexico 87508, United States. (1) OpenEye Scientific Software, Sante Fe New Mexico 87508, United States

The world of academia and the world of industry exist at varying distances in different scientific disciplines. The argument will be made that this gap is particularly great in molecular modeling and that this distance hinders development of the field. Some suggestions will be made as to how matters might be improved and how, in an ideal world, the field might look as a result.

COMP 205

Peptide partitioning properties from direct water-to-membrane insertion studies
Peptide partitioning properties are at the heart of biological membrane phenomena, and their precise quantification is vital for \textit{ab-initio} structure prediction. Recently the cellular translocon machinery has been employed to determine the water-to-membrane transfer properties and energetics for a series of polyoleucine segments. We show here that the insertion propensity, pathway and transfer energetics into synthetic POPC bilayers can be fully described by direct atomistic peptide partitioning simulations. The insertion probability as a function of peptide length follows two-state Boltzmann statistics, in excellent agreement with the experiments. The simulations expose a systematic offset between translocon mediated and direct insertion free energies. Compared to the experiment the insertion threshold is shifted towards shorter peptides by \(~2\) leucine residues. The simulations reveal many hitherto unknown atomic-resolution details about the partitioning process and promise to provide a powerful tool for urgently needed calibration of lipid parameters to match experimentally observed transfer energies.

\textbf{COMP 206}

\textbf{Understanding interactions of propylene glycol and skin lipids using molecular dynamics simulation}

\textbf{Jayeeta Ghosh}\textsuperscript{(1)}, jayeeta.ghosh@gmail.com, 50 R Reading Road, Edison NJ 08817, United States; \textbf{Bozena Michniak-Kohn}\textsuperscript{(1)}. (1) New Jersey Center for Biomaterials, Rutgers -The State University of New Jersey, Piscataway NJ 08817, United States

Stratum corneum (SC), the uppermost layer of skin, provides the main barrier to transdermal permeation of drug molecules. In order to effectively modify penetration rates of molecules through skin, it is important to understand the detailed structure of SC lipids from molecular level. In this work we analyzed effects of penetration modifiers, in particular propylene glycol on the structure of lipid membranes using molecular dynamics simulation. United atom models of ceramides 2 and 6 that have similar but distinct molecular structures were simulated. We observed a slightly higher gel to liquid-crystalline transition temperature for ceramide 6 compared to ceramide 2. Structural properties including area per head group, lipid tail order parameter, and hydrogen bonding propensity have been compared between the two lipids and also with experimental data. The effect of varying concentration of propylene glycol was also studied. Correlations were made between experimental permeability data and molecular simulation results.
Comp 207

Conformational prediction of the HIV-1 gp41 membrane-spanning domain

Vamshi K Gangupomu(1), gangupomu@gmail.com, 3141 Chestnut Street, Philadelphia PA 19104, United States; Cameron F Abrams(1). (1) Department of Chemical and Biological Engineering, Drexel University, Philadelphia PA 19104, United States

The HIV-1 envelope glycoprotein (Env) is a trimer of heterodimers of receptor-targeting gp120 and membrane anchoring gp41 fragments that mediates infection by recognizing and binding to cell-surface CD4 and CCR5/CXCR4 receptors and undergoing conformational changes that drive fusion of viral and cell compartments. We use a combination of all-atom molecular dynamics and metadynamics to probe the free energy of membrane spanning domain (MSD) of gp41 in both water and in a viral lipid bilayer to provide a basis for hypothesizing how the Env-mediated fusion reaction occurs in molecular detail. In water, we find that MSD shuttles between two kinked partially helical conformational states. Preliminary metadynamics calculations of MSD's spanning cholesterol-containing lipid bilayers suggest that a single uninterrupted α-helix is not a preferred conformation for bilayer-embedded MSD. We observe several conformations in which the midspan charged residue R694 "snorkels" among either bilayer head-groups, leading to significant water penetration and likely weakening the membrane.

Comp 208

Computing dynamics, electrostatics and spectroscopy at membrane-water interfaces

Jonathan D Hirst(1), jonathan.hirst@nottingham.ac.uk, University Park, Nottingham Nottinghamshire NG7 2RD, United Kingdom; David Robinson(1); Nicholas A Besley(1); Paul O'Shea(2). (1) School of Chemistry, University of Nottingham, Nottingham Nottinghamshire NG7 2RD, United Kingdom (2) School of Biology, University of Nottingham, Nottingham Nottinghamshire NG7 2RD, United Kingdom

Using a combination of classical molecular dynamics simulation and time-dependent density functional theory (TDDFT), we have characterised the spectroscopy of a fluorescent probe, di-8-ANEPPS in several membrane environments. The excited state geometries and atomic charges were obtained from complete active space self-consistent field calculations. These parameters were used in a subsequent decoupled quantum mechanics / molecular mechanics simulation, employing TDDFT (with an asymptotically corrected functional) so that large portions of the explicit immediate surrounding
environment can be included. We have investigated the influence of the nature of the lipids constituting the membrane and the presence of cholesterol, with a particular emphasis on the interfacial region with water. The ordering of water close to the membrane has a pronounced effect on the 'local' dielectric constant and significant implications for the membrane potential, whose molecular origins we are able elucidate.

COMP 209

Lipid10: A comprehensive phospholipid force field for AMBER

Ross C Walker\(^{(1)}\), rcw@sdsc.edu, 9500 Gilman Drive MC0505, La Jolla CA 92093-0505, United States; Åge Skjevik\(^{(2)}\); Knut Teigen\(^{(2)}\). (1) San Diego Supercomputer Center, University of California, San Diego, La Jolla CA 92093-0505, United States (2) Department of Biomedicine, University of Bergen, Bergen 5009, Norway

Current force fields for phospholipids only include a few select head and tail combinations making it difficult to build complex membrane structures. In this talk we provide an overview of a newly developed extensible force field for simulating complex phospholipids. The force field supports all major phospholipid head groups and tails as well as cholesterol. The design approach is such that one can mix and match head groups and tails as required allowing all major phospholipids to be simulated. The lipid10 force field has been designed to be compatible with all other AMBER force fields allowing membrane bound proteins to be effectively simulated. We provide details of the parameterization approach, including the methods used to create independent head and tail units as well as an overview of the validation approaches used.

COMP 210

Adaptive multiscale modeling of bilayer systems

Preston B Moore\(^{(1)}\), p.moore@usp.edu, 600 S 43rd St, Philadelphia PA 19104, United States; Vladimir Perez\(^{(1)}\); Steven O Nielsen\(^{(2)}\); Bernd Ensing\(^{(3)}\). (1) Department of Chemistry & Biochemistry, University of the Sciences in Philadelphia (USP), Philadelphia PA 19104, United States (2) Department of Chemistry, University of Texas at Dallas, Richardson Texas 75080, United States (3) Van 't Hoff Institute for Molecular Sciences, Amsterdam Amsterdam 1018 WV, The Netherlands

Here we report our simulations on ion-channel simulations where we model the ion-channel in all-atom resolution and the surrounding water and lipids at the CG level. Bridging the time and length scale separation between molecular at the mesoscopic phenomena is a severe
challenge for molecular simulation of soft matter and biomolecular systems, particularly membrane systems. Currently, large-scale molecular dynamics (MD) simulations are performed using so-called coarse grain (CG) models. CG models are composed of a reduced number of interaction sites that represent groups of atoms in order to drastically lower the computational demand at the price of losing the atomistic detail and chemical specificity. A promising avenue of research is multiscale modeling through coupling of an atomistic model to a CG model, in analogy to the successful examples found for hybrid quantum/classical approaches. We report on a system of water, lipid, protien which can change representations represents on the fly as the system evolves.

COMP 211

Effect of spherical fullerenes on pulmonary lipid monolayers: Molecular dynamics simulations

Steven O Nielsen(1), steven.nielsen@utdallas.edu, 800 West Campbell Road, BE26, Richardson TX 75080, United States ; Chi-cheng Chiu(1); Preston B Moore(2); Russell DeVane(3); Michael L Klein(3); Wataru Shinoda(4). (1) Department of Chemistry, University of Texas at Dallas, Richardson TX 75080, United States (2) Department of Chemistry and Biochemistry, University of the Sciences in Philadelphia, Philadelphia PA 19104, United States (3) Department of Chemistry, Temple University, Philadelphia PA 19122, United States (4) Research Institute for Computational Sciences, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba Ibaraki, Japan

Airborn spherical fullerenes (C60, C180, C540) can adsorb to the pulmonary lipid monolayer upon inhalation. We use MD simulations to study the effect of these particles on the lipid monolayer to bilayer transition which occurs during respiration. Large effects are observed for clusters of small (C60) particles and for large fullerene particles (C540).

COMP 212

Fragment docking in BACE: Retrospective survey of loop sampling and docking

Johannes H Voigt(1), johannes.voigt@merck.com, 2015 Galloping Hill Rd., K-15-L-0300, Kenilworth NJ 07033, United States ; Peter Orth(1); Yu Sen Wang(2); Daniel Wyss(2); Suresh Babu(3); John Caldwell(1); Sam Chackalamannil(1); Xia Che(4); Jared Cumming(1); Michael Czarniecki(1); Jinqi Duo(1); James Durkin(1); Helen Gu(1); Tao Guo(3); Rachel Hunter(3); Ulrich Iserloh(1); Matthew E Kennedy(4); Reshma Kuvelkar(4); Le Thuy(3); Vincent Madison(1); Robert D Mazzola(1); Brian A
Inhibition of BACE1 is a promising therapeutic target for treatment of Alzheimer's disease.

The series of our current BACE program in Kenilworth originated from an NMR screen of more than 10K fragments. BACE1 like other aspartic proteases has a flexible loop – “flap” – which gives rise to a multitude of possible and observed protein conformations. Therefore correct prediction of the binding pose of the fragment hits and subsequently optimized leads was challenging and elucidation by in house crystallography was crucial. At the onset of the program one published structure – 1fkn – was available; today we have obtained coordinates of more than 1,000 BACE-ligand complexes in house and more than 120 BACE structures have been deposited at the PDB.

In this study we survey how much conformational space can be sampled starting from a diverse set of protein conformation. Various implementations of low mode conformational search algorithms as implemented in MOE (Chemical Computing Group) and Macromodel are compared to loop sampling in Prime (Schroedinger) and molecular dynamics simulations. Pose prediction accuracy by cross-docking and docking to computationally generated protein conformations is assed.

COMP 213

Accurate pose prediction and improved structure-based virtual screening for fragments

Woody Sherman(1), woody.sherman@schrodinger.com, 120 West 45th St., 17th Floor, New York NY 10036, United States. (1) Schrodinger, New York NY 10036, United States

Fragment-based drug design continues to be a growing area of interest in the pharmaceutical industry. Recent advances in experimental technologies have led to an increased number of available crystal structures and screening results. In this work, we describe advances in docking and scoring of fragments. Specifically, we first discuss the importance of accurate tautomer and ionization state prediction in fragment docking. We then show that Glide docking can accurately place fragments when induced-fit of the protein is not significant.
Finally, we demonstrate that Glide can enrich databases for active fragments. However, the default sampling technology and scoring function that is typically used for drug-like molecules may not be optimal when studying fragments and we show how screening results can be improved.

COMP 214

Fragment screening paradigm for drug discovery

Diane Joseph-McCarthy\(^{(1)}\), diane.joseph-mccarthy@astrazeneca.com, 35 Gatehouse Dr., Waltham MA 02451, United States. (1) Infection Computational Science, AstraZeneca, R&D Boston, Waltham MA 02451, United States

The expanding role of fragment screening in the drug discovery process will be discussed. In particular, the combined use of docking and biophysical-based fragment screening will be highlighted. The development of fragment libraries screened at AstraZeneca will also be presented. Finally, examples of computational approaches for both direct and indirect uses of fragment positions and the challenges that lie ahead will be described.

COMP 215

Fragment-based druggability measures in drug targets

David R Hall\(^{(1)}\), drhall@bu.edu, 44 Cummington Street, Boston MA 02215, United States; Chi Ho Ngan\(^{(1)}\); Brandon Zerbe\(^{(1)}\); Dima Kozakov\(^{(1)}\); Sandor Vajda\(^{(1)}\). (1) Department of Biomedical Engineering, Boston University, Boston MA 02215, United States

We describe the FTMAP method based on computational fragment mapping to identify “hot-spot” regions in protein drug targets that bind small molecules. The method places molecular probes on the protein surface, and finds and clusters the energetically favorable sites across probes. We apply our algorithm to a set of 12 drug targets where experimental fragment based approaches were used, and fragment size core and effective inhibitors are known. We demonstrate that mapping applied to unliganded structures of proteins identifies the location of the core as the highest ranked "hot-spot" within the binding site, with other parts of inhibitors delineated by lower ranked "hot-spots". In many cases secondary hot-spots identified by mapping explore different parts of the binding site than natural substrates, in good agreement with high affinity inhibitors. We also demonstrate that consensus clustering of functional groups carried by different probes can provide binding specificity information.
FragVLib: Fragment-based virtual screening library using geometric and chemical patterns of interactions at interface of ligand-receptor complex crystal structures

Raed S. Khashan\textsuperscript{(1)(2)}, raed@email.unc.edu, 301 Beard Hall, CB# 7360, Chapel Hill NC 27599, United States ; Weifan Zheng\textsuperscript{(2)}. (1) Division of Medicinal Chemistry & Natural Products, University of North Carolina at Chapel Hill, Chapel Hill NC 27599, United States (2) Department of Pharmaceutical Sciences, North Carolina Central University, BRITE Institute, Chapel Hill NC 27599, United States

We present a virtual screening library to identify binding fragments. The idea is based on identifying patterns of interactions between ligand and receptor atoms for a database of complexes. This can be done by, first, applying Almost Delaunay tessellation to define interfacial atoms between ligands and their receptors. These interfacial atoms form the pattern of interaction for which both geometric and chemical features are memorized. Thus, given a particular binding pocket, we search the memorized patterns for ones have their receptor atoms matching (geometrically and chemically) any part of the particular binding pocket's atoms. For each match found, the ligand atoms corresponding to the matched receptor atoms in that pattern are identified. These ligand atoms form the binding fragments, which can then be used to search for molecules containing them. Or, they can be augmented using de novo drug design techniques to form new molecular entity. (Check http://www.unc.edu/~raed for Software).

COMP 217

Molecular dynamics studies of water-protein interactions
We use molecular dynamics simulations to study the interaction of water with proteins. With the help of a semi-grand canonical formalism, we determine the structure, dynamics, and thermodynamics of water in the protein interior and at buried sites. We find that water filling of weakly polar protein cavities from the solvent is governed by a subtle balance between the loss in bulk hydrogen bond interactions, the gain in strong hydrogen-bond interactions between confined water molecules, weakly attractive interactions between water and the cavity, and the entropic gain from filling a void space. The simulation results will be compared to X-ray crystallography and NMR experiments. The effects of interfacial and cavity water on protein function and ligand binding will be discussed.

COMP 218

Addressing limitations with the MM-GB/SA scoring procedure using the WaterMap method and free-energy perturbation calculations

Cristiano R. W. Guimaraes(1), cristiano.guimaraes@pfizer.com, 558 Eastern Point Road, Groton Connecticut 06340, United States . (1) CVMD Chemistry, PharmaTherapeutics Research and Development, Pfizer, Inc., Groton Connecticut 06340, United States

The MM-GB/SA scoring technique has become an important computational approach in lead optimization. Despite showing good accuracy, much work is necessary before the method can be applied to rank multiple chemical series. Here, we investigate the poor estimation of protein desolvation provided by GB/SA and the large dynamic range in the MM-GB/SA scoring compared to that of the experimental data. In the former, replacing the GB/SA protein desolvation by the WaterMap free energy liberation of binding-site waters provides the best results. However, the improvement is modest over results obtained with the MM-GB/SA and WaterMap methods individually, apparently due to the high correlation between the free energy liberation and protein-ligand van der Waals interactions. As for the large dynamic range, comparisons between MM-GB/SA and FEP calculations indicate that it has its origin in the lack of dynamical screening of protein-ligand electrostatic interactions and the incomplete description of enthalpy–entropy compensation effects.
Prediction of potency of protease inhibitors by GBSA simulations with polarizable quantum mechanics-based ligand charges and a hybrid water model

Debananda Das\(^{(1)}\), dasd@mail.nih.gov, 10 Center Drive, MSC 1868, Bethesda MD 20892-1868, United States; Hiroaki Mitsuya\(^{(1)(2)}\); Yasuhiro Koh\(^{(2)}\); Yasushi Tojo\(^{(2)}\); Arun Ghosh\(^{(3)}\). (1) HIV and AIDS Malignancy Branch, National Cancer Institute, Bethesda MD 20892-1868, United States (2) Departments of Hematology and Infectious Diseases, Kumamoto University Graduate School of Medical and Pharmaceutical Sciences, Bethesda MD 20892-1868, United States (3) Departments of Chemistry and Medicinal Chemistry, Purdue University, West Lafayette Indiana 47907, United States

Reliable and robust prediction of binding affinity for drug molecules continues to be a daunting challenge. We have simulated the binding interactions and free energy of binding of several protease inhibitors (PIs) with wild-type and various mutant proteases by performing GBSA simulations, in which each PI's partial charge was determined by quantum mechanics and the partial charge accounts for the polarization induced by the protease environment. We employed a hybrid solvation model that retains selected explicit water molecules in the protein with surface generalized Born implicit solvent. We examined the correlation of the free energy with antiviral potency of PIs. The free energy showed a strong correlation with experimentally determined anti-HIV-1 potency. The present data suggest that the presence of selected explicit water in protein, and protein polarization-induced quantum charges for the inhibitor, compared to lack of explicit water and a static force field-based charge model, can serve as an improved lead optimization tool, and warrants further exploration.

COMP 220

Continuum theory and the analysis of active sites

Anthony Nicholls\(^{(1)}\), anthony@eyesopen.com, 9d Bisbee Crt, Santa Fe NM 87508; Mike Word\(^{(1)}\). (1) Department of Research and Development, OpenEye Scientific Software, Inc, Santa Fe NM 87508, United States

Continuum theory for electrostatics free energies at the molecular level was never supposed to work- water is discrete and the very idea of treating its properties as a mean field was considered inappropriate. Yet Poisson-Boltzmann (PB) theory continues to perform as well as, if not better than, explicit water treatments in the estimation of small molecule solvation or macromolecular biophysics. However, it is still assumed PB will fail to correctly describe the physics of the active sites of proteins. As this remains a focus for predictive drug discovery, is this assumption correct? And if it is, can we improve continuum theory by going beyond the mean field limit, i.e. producing a 'virial' expansion of PB? This talk will cover our attempts to date and the physical insight gained.
Prediction of consistent water networks in uncomplexed protein binding sites based on knowledge-based potentials

Michael Betz(1), michael.betz@staff.uni-marburg.de, Marbacher Weg 6, Marburg D 35032, Germany; Gerd Neudert(1); Gerhard Klebe(1). (1) Institute of Pharmaceutical Chemistry, Philipps-University Marburg, Marburg D 35032, Germany

Within the active site of a protein water fulfills a variety of different roles. Solvation of hydrophilic parts stabilizes a distinct protein conformation, whereas desolvation upon ligand binding may lead to a gain of entropy. In an overwhelming number of cases, water molecules mediate interactions between protein and the bound ligand. Therefore, a reliable prediction of water molecules participating in ligand binding is essential for docking and scoring, and is necessary to develop strategies in ligand design. We require some reasonable estimates about the free energy contributions of water to binding.

Useful parameters for such estimations are the total number of displaceable water molecules and the probabilities for their displacement upon ligand binding. These parameters depend on specific interactions with the protein and other water molecules, and thus the positions of individual water molecules.

The high flexibility of water networks makes it difficult to observe distinct water molecules at well defined positions in structure determinations. Thus, experimentally observed positions of water molecules have to be assessed critically, bearing in mind that they represent an average picture of a highly dynamic equilibrium ensemble. Moreover, there are many structures with inconsistent and incomplete water networks.

To address these deficiencies we developed a tool that predicts possible configurations of complete water networks in binding pockets in a consistent way. It is based on the well established knowledge-based potentials implemented into DrugScore, which also allow for a reasonable differentiation between "conserved" and "displaceable" water molecules. The potentials used were derived specifically for water positions as observed in small molecule crystal structures in the CSD.

To account for the flexibility and high intercorrelation we apply a clique-based approach, resulting in water networks maximizing the total DrugScore.

To incorporate as much known information as possible about a given target, we also allow to include constraints defined by experimentally observed water positions.
Our tool provides a useful starting point whenever a possible configuration of water molecules need to be estimated in an uncomplexed protein, and suggests their spatial positions and their classification with respect to some kind of affinity prediction.

In first tests we were able to get classifications and positional predictions which are in good agreement with crystallographically observed water molecules with remarkably small deviations.

COMP 222

Explicit-water modeling of a model protein-ligand binding site predicts the non-classical hydrophobic effect

Demetri T. Moustakas(1), demetri.moustakas@astrazeneca.com, 35 Gatehouse Drive, Waltham MA 02451, United States; Phil W Snyder(2); Woody Sherman(3); George M Whitesides(2). (1) Department of Infection, Computational Sciences, AstraZeneca R&D Boston, Waltham MA 02451, United States (2) Department of Chemistry and Chemical Biology, Harvard University, Cambridge MA 02138, United States (3) Schrödinger, Inc., New York NY 10036, United States

This work reports a study of the thermodynamics of hydrophobic interactions between human carbonic anhydrase II and a series of structurally analogous heteroaromatic sulfonamides. Isothermal titration calorimetry (ITC) established that increasing the non-polar surface area of the ligands resulted in a large enthalpy-dominated increase the binding affinity – the so-called non-classical hydrophobic affect. Subsequent X-ray crystallography studies reveal no significant changes in protein-ligand interactions as a function of increasing the ligand non-polar surface area, suggesting that solute-solvent interactions are responsible for the observed thermodynamic effects. Modeling studies using explicit solvent models suggest that the larger ligands alter both the structure and thermodynamic characteristics of water molecules in the binding site, which contributes significantly to the observed non-classical hydrophobic effect.

COMP 223

New coarse-grained model for water: The importance of electrostatic interactions

Zhe Wu(1), zwu7@wisc.edu, 1101 Univ. Ave., Madison WI 53706, United States; Qiang Cui(1); Arun Yethiraj(1). (1) Department of Chemistry, UW Madison, Madison WI 53706, United States

A new coarse-grained (CG) model for water is developed based on the properties of clusters of four water molecules in atomistic simulations. CG units interact via a soft non-electrostatic interaction. Electrostatic interactions are
incorporated via three charged sites with the charges and model topology chosen to reproduce the dipole moment and quadrupole moment tensor of 4-water clusters. The parameters in the model are optimized to reproduce experimental data for the compressibility, density, and permittivity of bulk water, and the surface tension and interface potential for the air-water interface. This big multipole water (BMW) model represents a qualitative improvement over existing CG water models, e.g., it reproduces the dipole potential in membrane-water interface when compared to experiment, with modest additional computational cost.

**COMP 224**

**Virtual screening targeting the PhoP response regulator to inhibit bacterial virulence**

*Yat T Tang* (1), yattang@wustl.edu, 700 S. Euclid Ave., St. Louis MO 63110, United States; *Eduardo A Groisman* (2)(3), *James J Havranek* (4); *Garland R Marshall* (1). (1) Department of Biochemistry and Molecular Biophysics, Washington University in St. Louis, St. Louis MO 63110, United States (2) Department of Molecular Microbiology, Washington University in St. Louis, St. Louis MO 63110, United States (3) Howard Hughes Medical Institute, St. Louis MO 63110, United States (4) Department of Genetics, Washington University in St. Louis, St. Louis MO 63110, United States

Two-component signal transduction (TCST) is the predominant signaling system used in bacteria to sense and respond to environmental changes in order to survive and thrive under various conditions. TCST is composed of a sensor histidine kinase and a response regulatory receiver. The PhoQ/PhoP two-component signal transduction system in *Salmonella typhimurium* senses and responds to extracellular Mg$^{2+}$ levels by controlling the transcription level of key virulence genes.

A computational approach using virtual screening combined with consensus scoring revealed drug-like compounds that inhibit the function of the PhoP response regulator. Electrophoretic mobility shift assays and fluorescence anisotropy assays were used to test the predictability of the computational
approach. Eight compounds have been identified which disrupts the formation of the protein-DNA complex necessary for regulating transcription.

Results from this study serves as a proof-of-principle for targeting the two-component signal transduction response regulator to inhibit its function as a transcription factor and modulate gene expression and regulation of bacterial virulence. Results using this method suggests targeting TCST response regulators as a promising strategy for the development of novel antimicrobial therapeutics.

COMP 225

Computational investigation of the transport mechanism for monoamine transporters

Bonnie A Merchant\(^{(1)}\), merchantb@duq.edu, 320 Mellon Hall, 600 Forbes Ave, Pittsburgh PA 15282, United States; Jeffry D. Madura\(^{(1)}\). (1) Department of Chemistry & Biochemistry, Duquesne University, Pittsburgh PA 15282, United States

Neurotransmitter sodium symporters are targets for various medicinal and illegal drugs that affect mood and behavior. Of particular interest are the dopamine (DAT) and serotonin (SERT) transporters of which the three-dimensional structures are unknown. A three-dimensional structure homologous to DAT and SERT, both in sequence and in function, is the leucine transporter (LeuT\(_{Aa}\)). Therefore the LeuT\(_{Aa}\) is an ideal system to test new computational tools that can be used to study DAT and SERT function and dynamics, once their structures are determined. The Multi-Configurational Thermodynamic Integration method as implemented in NAMD, was used to study the transport of various chemically relevant substrates in LeuT\(_{Aa}\) and homology models of DAT and SERT. Several sites of low free energy score have been indentified, which correspond to primary and secondary substrate pockets of the transporters. Detailed free energy and structural results of the transport mechanisms will be presented.
COMP 226

Quantum chemistry calculations for arbitrarily complex electrostatic environments: Benzene anion stabilization

Roberto Olivares-Amaya(1), olivares@fas.harvard.edu, 12 Oxford St., Cambridge MA 02138, United States; Michael Stopa(2); Alan Aspuru-Guzik(1). (1) Department of Chemistry and Chemical Biology, Harvard University, Cambridge MA 02138, United States (2) Center for Nanoscale Systems, Harvard University, Cambridge MA 02138, United States

For many nanotechnology applications, it is necessary to include the complex environments in which molecules reside. For instance, in the case of surface-enhanced spectroscopic techniques, molecules are bound to metal nanoparticles. A proper description of the latter is fundamental to obtain more accurate molecular properties. We present a unified finite-difference quantum mechanics in a electrostatic environment approach for determining electronic structure properties. A key feature of our method is the use of arbitrary electrostatic kernels that include non-trivial effects such as screening and gate potentials. Results show that the benzene anion is stable compared to its neutral counterpart at 2.1 nm; when it lies between two metallic plates. A voltage bias sweep shows the electron's tendency to become unbound and be stabilized at the region near the gate. These studies shine light on the pre-charged and charged electronic structure stage under complex boundary conditions. Future work will focus on the stabilization of higher order acenes.

COMP 227

Biomolecular dynamics simulations based on the finite-difference Poisson-Boltzmann method

Qin Cai(1), qcai@uci.edu, 3144 Natural Science I, University of California, Irvine, Irvine CA 92697, United States; Xiang Ye(1); Jun Wang(1); Mengjueh Hsieh(1); Ray Luo(1). (1) University of California, Irvine, United States

Poisson-Boltzmann implicit solvent model has been widely adopted in biomolecular applications, which, however, are mostly used in static conformations. To extend its application to molecular dynamics, we have explored several strategies to enhance the accuracy of the force calculation under the finite-difference scheme. After combining these methods with our previous efforts to improve the finite-difference Poisson-Boltzmann method, including removal of charge singularity in the Poisson-Boltzmann equation (PBE), fast and efficient solutions of the PBE, and accurate interface treatment by the
immersed interface method, we are able to achieve stable molecular dynamics simulations with our tested molecules.

**COMP 228**

**Filtering and ranking enzyme designs using EDGE**

*Scott A. Johnson*(1), sajohn2@chem.ucla.edu, 607 Charles E. Young Drive E., Los Angeles CA 90095-1569, United States; Kendall N. Houk*(1). (1) Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles CA 90095, United States

The "inside-out" protocol for rational enzyme design is the process by which an enzyme structure is gradually built around a quantum mechanical (QM) transition state (TS) for a desired chemical reaction. The protocol hinges on developing a QM "theozyme" (theoretical enzyme) by adding model protein residues to stabilize the TS and lower the activation barrier of the chemical reaction. Although the inside-out protocol has met with success in generating functional enzymes for the Kemp elimination and retro-aldol reactions, the process is limited by the combinatorially explosive generation of enzyme designs by Rosetta. We have written the program EDGE (Enzyme Design Geometry Evaluation) to filter and rank enzyme designs based on theozyme-similarity at key intervals during the inside-out protocol, as well as to prioritize enzyme designs for experimental testing. We couple EDGE with a grid-based workflow that allows for the massive parallelization of the enzyme design process using Rosetta software.

**COMP 229**

**Quantum information and quantum computation for chemistry**

*Alan Aspuru-Guzik*(1), aspuru@chemistry.harvard.edu, 12 Oxford St Room M113, Cambridge MA 02139, United States. (1) Department of Chemistry and Chemical Biology, Harvard University, Cambridge MA 02138, United States

Chemical processes can be formally described using first-principles methods by solving the time-dependent Schrodinger equation. Analytical and numerical solutions of the equations that describe chemical processes are amongst the fundamental activities of theoretical chemistry. The field of quantum information theory emerges from the realization that information is not independent of the fundamental physical laws for processing and storing it. If the system that processes information is quantum mechanical, in certain cases, it can solve algorithmic problems more efficiently than if this information was processed classically. Algorithm
development is the subject of the sub-field of quantum information theory known as quantum computing. Therefore, quantum mechanical processes, such as molecules undergoing chemical reactions, can be thought of as quantum information processing events. My research lies at the intersection of these fields. I seek to employ the formalism, methods and applications of both quantum information and theoretical chemistry to study problems of chemical relevance.

In particular, I have worked on the simulation of chemical systems using quantum computers (theoretically and collaborating with experimentalists), and the development of time-dependent density functional theory for open quantum systems, among other developments summarized here.

COMP 230

Excited-state quantum chemistry for macromolecules and condensed phases

John M. Herbert\(^{(1)}\), herbert@chemistry.ohio-state.edu, 100 W. 18th Ave., Columbus OH 43210, United States. (1) Chemistry, Ohio State University, Columbus OH 43210, United States

A variety of new methods are described for calculation of electronic excitation energies in large systems. The quantum-chemical model of choice for these applications is time-dependent density functional theory (TD-DFT); however, naive application of TD-DFT results in a plethora of spurious charge-transfer excited states. We have developed "long-range corrected" density functionals that avoid this problem, and which provide a balanced description of both localized excitations and charge-transfer excited states. To model the response of the environment to electronic excitation of a chromophore, we have developed a general formalism for coupling quantum chemistry to a polarizable force field, to perform "QM/polMM" calculations. We have also developed a class of continuum solvation models that (unlike previous models) afford smooth potential energy surfaces, and are efficient enough for use in large QM/MM simulations. Representative applications include excited states of DNA oligomers, aqueous electrons, and push-pull chromophores in liquid solution.

COMP 231

Mimicking coarse-grained simulations without coarse-graining: Enhanced sampling by damping short-range interactions

Feng Wang\(^{(1)}\), fengwang@bu.edu, 590 Commonwealth Ave, Boston MA 02215, United States. (1) Department of Chemistry, Boston University, Boston MA 02215, United States
The damped short-range interaction (DSRI) method recently developed in our group is designed to mimic coarse-grained simulations by propagating an atomistic scale system on a smoothed potential energy surface while preserving the overall feature of the free energy landscape. The DSRI method has the benefit of enhanced sampling similar to a coarse grained simulation but does not require coarse-graining. Coupled with a scheme where the mass of various atoms are optimized to accelerate slow events, our method was used to simulate liquid water, alanine dipeptide, alanine penta-peptide in explicit water solvent, and the self-assembly of dimyristoylphosphatidylcholine lipid. In each case, our simulations indicate this method is capable of appreciably accelerating the underlying dynamics without significantly changing the free energy surface. Additional insights from DSRI simulation and the promise of coupling our DSRI method with replica-exchange molecular dynamics will be discussed.

COMP 232

Information flow in biomolecules

Arjan van der Vaart*, vandervaart@cas.usf.edu, 4202 E. Fowler Ave. CHE 205, Tampa Florida 33620, United States. (1) Department of Chemistry, University of South Florida, Tampa Florida 33620, United States

How is information transferred from the binding site of a protein to the site where the action takes place? To address this question, we have introduced the information theory measure of transfer entropy for the analysis of correlated motions in molecular dynamics simulations. The method detects the direction of information flow, and can be used to analyze how information is transferred through the system. Applications of the method to the Ets-1 transcription factor and the ERK2 kinase will be discussed.

COMP 233

Predicting the structure and the binding site of rimonabant for the cannabinoid CB1 receptor

Caitlin E. Scott*, cescott@caltech.edu, 1200 East California Blvd., Beckman Institute (Mail Code: 139-74) Rm. 322, Pasadena CA 91125, United States; Ravinder Abrol; William A. Goddard*. (1) Materials Process and Simulation Center, California Institute of Technology, Pasadena CA 91125, United States

Rimonabant (SR141716A), a selective blocker of the cannabinoid 1 (CB1) G-protein coupled receptor (GPCR), is an anti-obesity drug that causes serious side effects including depression and increased risk of suicide. In order to develop receptor specific drugs to minimize such consequences, we aim to predict the CB1 receptor structure from first principles using a GEnSeMBLE (GPCR Ensemble of Structures in
Membrane BiLayer Environment) method. The docking results from a hierarchical docking program (DarwinDock) operating on the GEnSeMBLE output show that rimonabant is anchored by W5.43 and K3.28 in the CB1 protein, which is consistent with site-directed mutagenesis data. The phenyl rings of rimonabant sandwich W5.43. The ligand's carbonyl forms a hydrogen bond with K3.28, which is common in other analog docking to the CB1 receptor. We expect this study will lead to more effective drug candidates for obesity diseases with reduced side effects.

COMP 233

Predicting the structure and the binding site of rimonabant for the cannabinoid CB1 receptor

Caitlin E. Scott(1), cescott@caltech.edu, 1200 East California Blvd., Beckman Institute (Mail Code: 139-74) Rm. 322, Pasadena CA 91125, United States; Ravinder Abrol(1); William A. Goddard(1). (1) Materials Process and Simulation Center, California Institute of Technology, Pasadena CA 91125, United States

Rimonabant (SR141716A), a selective blocker of the cannabinoid 1 (CB1) G-protein coupled receptor (GPCR), is an anti-obesity drug that causes serious side effects including depression and increased risk of suicide. In order to develop receptor specific drugs to minimize such consequences, we aim to predict the CB1 receptor structure from first principles using a GEnSeMBLE (GPCR Ensemble of Structures in Membrane BiLayer Environment) method. The docking results from a hierarchical docking program (DarwinDock) operating on the GEnSeMBLE output show that rimonabant is anchored by W5.43 and K3.28 in the CB1 protein, which is consistent with site-directed mutagenesis data. The phenyl rings of rimonabant sandwich W5.43. The ligand's carbonyl forms a hydrogen bond with K3.28, which is common in other analog docking to the CB1 receptor. We expect this study will lead to more effective drug candidates for obesity diseases with reduced side effects.

COMP 234

Theoretical studies of hydrogen spillover mechanism for developing next generation automobile catalyst: A quantum chemical molecular dynamics study

Farouq Ahmed(1), farouq80@yahoo.com, Tohoku University, Graduate School of Engineering, 6-6-10-205 Aoba, Aramaki, Aoba, Sendai Miyagi 980-8579, Japan; Ryuji Muira(1); Ai Suzuki(1); Hideyuki Tsuboi(1); Nozomu Hatakeyama(1); Akira Endou(1); Hiromitsu Takaba(1); Momoji Kubo(1); Akira Miyamoto(1). (1) Department of Chemical Engineering, Department of Chemical Engineering,
Hydrogen spillover mechanism has earned intensive interest in the past decades because it plays a vital role in emerging technologies for the reduction of NOx in automobile exhausts. Hydrogen spillover arises in hydrogen-catalyzed reactions on supported metal catalyst. In the present study, we applied quantum chemical molecular dynamics (QCMD) to investigate the mechanism of the hydrogen spillover process on Pt/g-Al2O3 and on Pt/CeO2 catalyst surfaces for the first time. The direct observation of dissociative adsorption of hydrogen and diffusion of hydrogen on Pt/g-Al2O3 and Pt/CeO2 catalyst surfaces were successfully investigated. The diffusion of the hydrogen atom in the gas phase explains high reactivity observed in hydrogen spillover mechanism. The present study also indicates that the CeO2 catalyst support has the strong metal support interaction on Pt, which may enhance the metal catalyst dispersions and rate of hydrogen spillover reaction.

Figure1: Comparison of metal support interaction and structural change during the simulation of spillover hydrogen on Pt/CeO2, Pt/g-Al2O3 catalyst. Snapshots of molecular dynamics calculation at different simulation steps were shown in Figure 1(A) (0 fs to 853 fs) for Pt/CeO2 and Figure1 (B) (0 fs to 155 fs) for Pt/g-Al2O3 catalyst surface, respectively.
Crystallization of polydisperse systems of aspherical particles

William L Miller(1), wlmiller@gmail.com, 3000 Broadway MC 3137, New York NY 10027, United States ; Angelo Cacciuto(1). (1) Department of Chemistry, Columbia University, New York NY 10027, United States

We use numerical simulations to study the crystallization of polydisperse systems of hard aspherical particles. We perform direct simulation of a solid-liquid interface to examine the dependence of the phase transition pressure on the asphericity and shape polydispersity of particles comprising a given system. We present these dependences for several different systems.

COMP 236

Dependent of structural and thermodynamic characteristics of supercritical carbon dioxide fluid from atomic charges in molecule in region of critical point calculated by RISM theory

Fatima Keshtova(1), pantera@realision.ru, Leninskie Gory, GSP-2, Moscow Russia 119991, Russian Federation ; Svetlana Vladimirovna Keshtova(2). (1) Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow Russia 119991, Russian Federation (2) Phys Chem, Private researcher, Moscow Russia 117447, Russian Federation

By the theory of reference interaction site model with Percus-Yevick closure equation has been calculated the structure and thermodynamic properties of supercritical carbon dioxide in region of critical point. In work display that value of partial atomic charges in molecule influence on phase behavior of system near of critical point.

COMP 237

Application of structure and ligand based methods for the investigation of the flexibility of loop regions within a cyclin-dependent kinase (CDK) and the prediction of a possible binding mode of highly selective inhibitors

Ulrich Heiser(1), ulrich.heiser@probiodrug.de, Weinbergweg 22, Halle/Saale 6120, Germany ; Christian Jäger(1); Mirko Buchholz(1); Wolfgang Brandt(2); Lutz Zeitlmann(3); Andre Niestroj(1); Hans-Ulrich Demuth(4). (1) Dept. of Medicinal Chemistry, Probiodrug AG, Halle/Saale 06120, Germany (2) Leibniz Institute of Plant Biochemistry, Halle/Saale 06120, Germany (3) Ingenium Pharmaceuticals GmbH., Martinsried 82152, Germany (4) Probiodrug AG, Halle/Saale 06120, Germany
Cyclin-Dependent Kinase 9 (CDK9) is an ubiquitously expressed serine/threonine kinase. In complex with Cyclin-T, it is an essential part of the Positive Transcription Elongation Factor β (P-TEFβ). P-TEFβ is recruited to promote RNA-Polymerase II processivity by serine 2 phosphorylation in its C-terminal domain. In addition, it was shown, that NF-κB dependent expression of target genes are critically dependent on CDK9 activity. Due to the multiple roles of NF-κB in inflammatory processes an inhibition of CDK9 by small molecules could lead to new therapeutic approaches. Recently the crystal structures of CDK9 in complex with ATP and the potent inhibitor Flavopiridol were published. Flavopiridol is of medium selectivity in the CDK family. In contrast, a medicinal chemistry approach at probiodrug led to highly selective inhibitors of CDK9, whereas crystal structures for these inhibitors are not available as yet. The published crystal structures of CDK9 reveal differences in the glycine rich loop (G-loop) between binding of the co-factor and the inhibitor. An edge to face interaction of the phenylalanine side chain of Phe30 with the inhibitor is not present in the complex with ATP, whereas the side chain of this amino acid has to cover a distance of about 14 Å, when the inhibitor binds to the active site. Because Phe30 is only present in CDK9 and CDK7 within the family of CDKs, the possible role of interactions of Phe30 with inhibitors may play a crucial role for the development of selective CDK9 inhibitors.

In this paper we describe several structure and ligand based methods to investigate the flexibility of the G-loop as an induced fit caused by the binding of Flavopiridol. Further, first insight into the possible binding modes of the selective inhibitors was gained. In doing so, techniques like molecular dynamic simulations, extended docking investigations, small molecule alignments and the development of homology models were applied.

**COMP 237**

**Application of structure and ligand based methods for the investigation of the flexibility of loop regions within a cyclin-dependent kinase (CDK) and the prediction of a possible binding mode of highly selective inhibitors**

**Ulrich Heiser**(1), ulrich.heiser@probiodrug.de, Weinbergweg 22, Halle/Saale 6120, Germany ; Christian Jäger(1); Mirko Buchholz(1); Wolfgang Brandt(2); Lutz Zeitlmann(3); Andre Niestroy(1); Hans-Ulrich Demuth(4). (1) Dept. of Medicinal Chemistry, Probiodrug AG, Halle/Saale 06120, Germany (2) Leibniz Institute of Plant Biochemistry, Halle/Saale 06120, Germany (3) Ingenium Pharmaceuticals GmbH., Martinsried 82152, Germany (4) Probiodrug AG, Halle/Saale 06120, Germany

Cyclin-Dependent Kinase 9 (CDK9) is an ubiquitously expressed serine/threonine kinase. In complex with Cyclin-T, it is an essential part of the Positive Transcription Elongation Factor β (P-TEFβ). P-TEFβ is recruited to promote RNA-Polymerase II processivity by serine 2
phosphorylation in its C-terminal domain. In addition, it was shown that NF-κB dependent expression of target genes are critically dependent on CDK9 activity. Due to the multiple roles of NF-κB in inflammatory processes an inhibition of CDK9 by small molecules could lead to new therapeutic approaches. Recently the crystal structures of CDK9 in complex with ATP and the potent inhibitor Flavopiridol were published. Flavopiridol is of medium selectivity in the CDK family. In contrast, a medicinal chemistry approach at probiodrug led to highly selective inhibitors of CDK9, whereas crystal structures for these inhibitors are not available as yet. The published crystal structures of CDK9 reveal differences in the glycine-rich loop (G-loop) between binding of the co-factor and the inhibitor. An edge to face interaction of the phenylalanine side chain of Phe30 with the inhibitor is not present in the complex with ATP, whereas the side chain of this amino acid has to cover a distance of about 14 Å, when the inhibitor binds to the active site. Because Phe30 is only present in CDK9 and CDK7 within the family of CDKs, the possible role of interactions of Phe30 with inhibitors may play a crucial role for the development of selective CDK9 inhibitors.

In this paper we describe several structure and ligand-based methods to investigate the flexibility of the G-loop as an induced fit caused by the binding of Flavopiridol. Further, first insight into the possible binding modes of the selective inhibitors was gained. In doing so, techniques like molecular dynamic simulations, extended docking investigations, small molecule alignments and the development of homology models were applied.

COMP 238

Molecular simulation of the nucleation and growth of C60 nanoparticles from the supersaturated vapor and from the undercooled liquid

Kenneth Ngale(1), KENNETH.NDUMBENGAL@UND.EDU, 151 Cornell Street Stop 9024, Grand Forks ND 58202, United States; Caroline Desgranges(1); Jerome Delhommelle(1). (1) Chemistry, University of North Dakota, Grand Forks ND 58201, United States

Using molecular simulation, we study the molecular mechanisms underlying the nucleation and growth of C60 nanoparticles from the supersaturated vapor and from the undercooled liquid [1]. We show that in both cases, nucleation proceeds through the formation of small clusters composed of the metastable hexagonal close-packed (HCP) polymorph. This is unlike what is usually found on simple systems, for which nucleation proceeds through the metastable body-centered cubic (BCC) polymorph. We rationalize this observation in terms of the steepness of the repulsion of the potential model for C60. We also observe two different types of growth leading either to crystallites dominated by the metastable HCP polymorph or to crystallites whose structure is predominantly that of the stable face-centered cubic (FCC) polymorph, in excellent agreement with experiments on the crystallization of C60 from the vapor.
COMP 239

Phase equilibria of polyaromatic hydrocarbons by hybrid Monte Carlo Wang-Landau simulations

Andy Magness\textsuperscript{(1)}, magn6494@hotmail.com, 151 Cornell Street Stop 9024, Grand Forks ND 58202, United States; Caroline Desgranges\textsuperscript{(1)}; Jerome Delhommelle\textsuperscript{(1)}; Jason Hicks\textsuperscript{(1)}. (1) Department of Chemistry, University of North Dakota, Grand Forks ND 58202, United States

Using a combination of Hybrid Monte Carlo simulations in the isothermal-isobaric ensemble with a Wang-Landau sampling \cite{1}, we parametrize a force field for polyaromatic hydrocarbons (PAHs). The proposed force field gives an accurate description of the vapour-liquid equilibria of naphthalene, phenanthrene and anthracene. In particular, the model yields a better account of the dependence of the vapour pressure on temperature than existing models. The strategy adopted in this work is markedly different from that followed in previous work, which relied on Gibbs Ensemble Monte Carlo (GEMC) simulations combined with configurational bias (CB) moves. The accuracy of the GEMC-CB method hinges on a reasonably high acceptance rate of the transfer of molecules from one phase to another, a condition difficult to achieve for large molecules like PAHs. Our approach avoids these transfers, allows for a direct determination of the coexistence data in the isothermal-isobaric ensemble and provides a promising alternative to the GEMC-CB approach.

\cite{1} C. Desgranges, J. Hicks, A. Magness and J. Delhommelle, Mol. Phys. 108, 151 (2010).

COMP 240

Molecular simulation of the formation of Fe-Pt nanomaterials for ultrahigh-density magnetic data storage

John Persson\textsuperscript{(1)}, john.persson@und.edu, 151 Cornell Street Stop 9024, Grand Forks ND 58202, United States; Caroline Desgranges\textsuperscript{(1)}; Jerome Delhommelle\textsuperscript{(1)}. (1) Department of Chemistry, University of North Dakota, Grand Forks ND 58202, United States

In recent years, nanoscale materials have started to play a major role in a large number of areas. However, in many cases, predicting and controlling the properties of nanoparticles has remained extremely challenging. In this project, we focus on magnetic nanoparticles, which have potential biomedical
applications (such as e.g. carriers for targeted drug delivery within the human body or as heating agents for local hypothermic cancer treatments) as well as applications in information technology as building blocks for ultra-high-density magnetic data storage. Building on our prior work on gold nanoparticles, we use molecular simulations to elucidate the mechanisms of formation of Fe-Pt nanoparticles. We determine the nucleation pathway and carefully examine how their structure evolves with the size of the nanoparticle. By modifying the conditions of crystallization, we shed light on the best strategies to prevent the formation of structural defects.

COMP 241

Enhanced sampling in folding simulations of Ala5 peptide by DSRI method

Dongshan Wei, dswei@bu.edu, 590 Commonwealth Ave Room 299, Boston MA 02215, United States; Feng Wang. (1) Department of Chemistry, Boston University, Boston MA 02215, United States

The folding dynamics of Ala5 (ACE-ALA5-NME) peptide was studied in explicit water by the damped short-range interaction (DSRI) method recently developed in our group. The DSRI method is designed to mimic coarse-grained simulations by propagating the system along a smoothed potential energy surface while preserving the overall feature of the free energy surface as much as possible. Coupled with a scheme where the mass of hydrogen atoms was set to 2 amu and the masses of all heavier atoms were set to 4 amu, the DSRI method appreciably accelerates the folding of Ala5 in water. Comparing with the standard molecular dynamics simulation, the DSRI method gives a 5- to 10-fold reduction in folding time. Meanwhile, the free energy surface of Ala5 determined by the Ramachandran plot is not significantly changed by the DSRI method.

COMP 242

Monte Carlo statistical analysis of the uncertainty of enrichment metrics for ensemble docking

Ivan Tubert-Brohman, ivan.tubert-brohman@schrodinger.com, 120 West 45th Street, New York NY 10036, United States. (1) Schrodinger, Inc., New York NY 10036, United States

In virtual screening, a common approach to the cross-docking problem is to use an ensemble of receptor conformations, dock every ligand into each conformation, and pick the conformation resulting in the best docking score for
each ligand. However, this method raises a number of questions such as: how many conformations are enough? How should the best conformation be chosen, and what errors are introduced by picking the wrong one? We studied these questions by performing a post hoc analysis after exhaustive cross-docking of known actives plus a 1000-ligand decoy set for over 10 different protein targets. We employed every receptor structure available in our dataset—an average of ~13 per target. The results provide a measure of the sensitivity of various enrichment metrics to receptor selection, as well as the minimum covering sets of receptor conformations that are required to dock every active molecule correctly.

COMP 242

Monte Carlo statistical analysis of the uncertainty of enrichment metrics for ensemble docking

Ivan Tubert-Brohman(1), ivan.tubert-brohman@schrodinger.com, 120 West 45th Street, New York NY 10036, United States . (1) Schrodinger, Inc., New York NY 10036, United States

In virtual screening, a common approach to the cross-docking problem is to use an ensemble of receptor conformations, dock every ligand into each conformation, and pick the conformation resulting in the best docking score for each ligand. However, this method raises a number of questions such as: how many conformations are enough? How should the best conformation be chosen, and what errors are introduced by picking the wrong one? We studied these questions by performing a post hoc analysis after exhaustive cross-docking of known actives plus a 1000-ligand decoy set for over 10 different protein targets. We employed every receptor structure available in our dataset—an average of ~13 per target. The results provide a measure of the sensitivity of various enrichment metrics to receptor selection, as well as the minimum covering sets of receptor conformations that are required to dock every active molecule correctly.

COMP 243

Docking studies of propafenone derivatives into a homology model of the hERG channel in the open state

Andrea Schiesaro(1), andrea.schiesaro@univie.ac.at, Althanstraße 14, Vienna 1090, Austria ; Andreas Windisch(2); Eugen Eugen Timin(2); Steffen Hering(2); Gerhard F. Ecker(1). (1) Department of Medicinal Chemistry, University of Vienna, Vienna 1090, Austria (2) Department of Pharmacology and Toxicology, University of Vienna, Vienna 1090, Austria

Blockade of the hERG channel is correlated with Long QT syndrome, a disorder of cardiac repolarization, which might lead to fatal arrhythmias. In light of this correlation it is important to understand the molecular determinants of this
undesirable interaction. A series of 6 propafenone derivatives were docked into an open state model of the hERG channel. Through cluster analysis of the consensus RMSD matrix considering the common scaffold of the ligands, three clusters containing all compounds were retrieved. The main interactions identified in the three clusters are with Phe656, Thr623 and Ser624. Information from mutagenesis studies, which indicate that only Phe656 is essential for binding of propafenone, allowed to prioritise one cluster. The poses of the selected cluster show that Phe656 is involved in a π-cation interaction with the charged nitrogen of the ligands.

Financial support provided by the University of Vienna under the PhD program “Molecular Drug Targets”.

COMP 243

Docking studies of propafenone derivatives into a homology model of the hERG channel in the open state

Andrea Schiesaro(1), andrea.schiesaro@univie.ac.at, Althanstraße 14, Vienna 1090, Austria ; Andreas Windisch(2); Eugen Eugen Timin(3); Steffen Hering(2); Gerhard F. Ecker(1). (1) Department of Medicinal Chemistry, University of Vienna, Vienna 1090, Austria (2) Department of Pharmacology and Toxicology, University of Vienna, Vienna 1090, Austria

Blockade of the hERG channel is correlated with Long QT syndrome, a disorder of cardiac repolarization, which might lead to fatal arrhythmias. In light of this correlation it is important to understand the molecular determinants of this undesirable interaction. A series of 6 propafenone derivatives were docked into an open state model of the hERG channel. Through cluster analysis of the consensus RMSD matrix considering the common scaffold of the ligands, three clusters containing all compounds were retrieved. The main interactions identified in the three clusters are with Phe656, Thr623 and Ser624. Information from mutagenesis studies, which indicate that only Phe656 is essential for binding of propafenone, allowed to prioritise one cluster. The poses of the selected cluster show that Phe656 is involved in a π-cation interaction with the charged nitrogen of the ligands.

Financial support provided by the University of Vienna under the PhD program “Molecular Drug Targets”.

COMP 244

QSAR, ligand efficiency (LE) and lipophilic efficiency (LipE) studies of a series of benzophenone-type inhibitors of the multidrug transporter P-glycoprotein

Ishrat Jabeen(1), ishrat.jabeen@univie.ac.at, Althanstraße14, Vienna 1090, Austria ; Zahida Parveen(2); Uwe Rinner(3); Peter Chiba(2); Gerhard F. Ecker(1). (1)
P-glycoprotein (P-gp) is an ATP dependent efflux transporter often linked to multidrug resistance. A better understanding of the structural requirements and lipophilic behavior of P-gp inhibitors will aid in the understanding of the molecular basis of ligand recognition. QSAR studies using a data set of benzophenones highlight the importance of partial charge and hydrophobic polar surface area descriptors for high activity. In addition, binding energy of the ligand per atom (ligand efficiency), and lipophilic efficiency were compared with those of P-gp inhibitors which entered clinical studies. For benzophenones, smaller compounds generally exhibited higher LE values. Interestingly, although P-gp inhibitors are highly lipophilic, they showed LipE values below the threshold considered to be necessary for promising drug candidates. This might be due to the fact that the ligand-protein interaction takes place directly in the membrane bilayer.

This work was supported by the Austrian Science Fund (grant F03502) and HEC Pakistan.

COMP 245

Molecular insight into FabH-inhibitor interactions using molecular dynamics simulations and free energy calculations to identify novel scaffolds for inhibition by pharmacophore-based virtual screening

Anuradha Mittal(1), anuradha.mittal@gmail.com, 900 So. Ashland St. - m/c 870, Chicago IL 60607-7173, United States ; Michael E Johnson(1). (1) Center for Pharmaceutical Biotechnology, University of Illinois at Chicago, Chicago IL 60607, United States

The increasing bacterial resistance to available antibiotics has motivated the need for the discovery of novel effective antibacterial agents. The condensing enzyme β-Ketoacyl-acyl carrier protein synthase III (FabH) serves as a potential target for novel antibacterial drug design. The analysis of molecular dynamics trajectories and interaction energy using Molecular Mechanics Poisson-Boltzmann/surface area (MM-PBSA) calculations reveal non-polar van der Waals interactions as the main driving force for the binding of known inhibitors. The various rotamer conformations exhibited by substrate binding residues (as determined from crystal structures and MD simulations) affects the shape of the binding pocket, posing a challenge in drug design. We propose novel inhibitors by pharmacophore-based virtual screening. The binding free
energy analysis of the most promising hits suggest residues in the binding pocket essential for enhanced binding affinity and selectivity of inhibitors. These results provide potentially valuable information to aid the lead optimization process.

COMP 246

Molecular origin of the interaction between PB1 and PB2 subunit in RNA polymerase of influenza virus

**Jin Woo Park**(1), jinu1115@snu.ac.kr, 31-420, Seoul kwanak-gu, Republic of Korea; Won Ho Jo(1). (1) Department of Materials Science and Engineering, Seoul National University, Seoul, Republic of Korea

RNA-dependent RNA polymerase of influenza virus is a multi-functional heterotrimer composed of three subunits (PA, PB1, and PB2). The PB1-PB2 interface is essential for transcription initiation, and thus can be considered as a novel potential drug target site. In this study, to identify the binding hot spots at the interface of the PB1-PB2 complex and to understand the nature of the interactions, a 20-nanosecond unrestrained molecular dynamics simulation with explicit solvent was performed. Pair interaction energies were then calculated to identify the binding hot spots at the PB1-PB2 interface. These analyses indicated that Leu7, Leu20, and Val25 residues of PB2 have strong hydrophobic interactions with PB1, and some charged residues, such as Arg3 and Glu6, have a large electrostatic interaction with PB1 at the interface. These results can provide some insights for designing inhibitors of RNA-dependent RNA polymerase.

COMP 247

Development of minimal pharmacophore hypotheses by molecular dynamics simulation and spatial restriction for screening N5-CAIR mutase fragment-like inhibitors

**Tian Zhu**(1), tzhu6@uic.edu, 900 South Ashland Avenue, Room 3100, Chicago IL 60607, United States; Michael E Johnson(1); Debbie Mulhearn(1). (1) Center for Pharmaceutical Biotechnology, University of Illinois at Chicago, Chicago IL 60607, United States

We used molecular dynamics simulation and spatial restriction for developing minimal pharmacophore hypotheses to identify novel fragment-like inhibitors of N5-CAIR Mutase, a key enzyme in bacterial purine synthesis. A pharmacophore model was generated based on the N5-CAIR Mutase X-ray crystal structure in complex with nitro-AIR. Crucial features were selected based on the flexibility of contributing residues from molecular dynamics simulation and the distance between different feature points, considering that the pharmacophore hypotheses
were intended to be applied over a fragment library. A four-point model was selected and subsequently applied to a database of 554 fragments, and 81 fragments completely or partially match this model without compromising the chemical diversity. We find that combining molecular dynamics simulation results and spacial restriction is useful for developing pharmacophore hypotheses with a rational number of features as a useful tool for reducing the number of compounds to be tested in fragment-based screening.

COMP 248

Computational modeling study of anthranilate synthase from Streptomyces venezuelae for the developing new antimicrobial agents

Jack Z Hu(1), zhu@howard.edu, 520 W Street, NW, Room 324, Washington DC 20059, United States; Meseret Ashenafi(1); Walton M Byrnes(1); William W Southerland(1). (1) Department of Biochemistry and Molecular Biology, Howard University, Washington DC 20059, United States

Anthranilate synthase (AnthS) is an attractive target for the discovery and development of new antimicrobial agents since it is a key regulatory enzyme in the primary metabolic tryptophan biosynthesis pathway in microorganisms. The regulation of tryptophan biosynthesis in the chloramphenicol-producing bacterium Streptomyces venezuelae is of interest because primary and secondary metabolic pathways are closely integrated in the organism. Since the structure of S. venezuelae AnthS is not available, computational modeling was used to build a protein homology model to gain understanding on the interactions between the enzyme and its inhibitors. The crystal structure of the allosteric AnthS from Salmonella typhimurium was selected as the template for modeling. The 3D model of the enzyme was built using the programs MODELLER and Chimera. Energy minimization was performed using CHARMM. The modeled structure of the enzyme provides a molecular target for the structure-based discovery of new antimicrobial agents.

COMP 249

Machine learning application for the high-performance prediction of functional sites in proteins of unknown function

Srinivas Somarowthu(1), somarowthu.s@husky.neu.edu, 360 Huntington Ave, Boston MA 02115, United States; Yujing Wang(2); Andrew Michaelson(3); Huyuan Yang(1); Mary Jo Ondrechen(1)(4). (1) Department of Chemistry and Chemical Biology, Northeastern University, Boston MA 02115, United States (2) College of Professional Studies, Northeastern University, Boston MA 02115, United States (3) Department of Biology, Northeastern University, Boston MA 02115, United States
States (4) Institute for Complex Scientific Software, Northeastern University, Boston MA 02115, United States

One of the major challenges in genomics is to understand the function of gene products, proteins, from their 3D structures. One viable path for acquiring information about protein function is through the identification of the catalytic and binding sites. While many methods are available for this task, the precision tends to be low. What is needed is a powerful, precise and accurate computational method for identifying important residues. Here, we utilize a new machine learning method, Partial Order Optimum Likelihood (POOL), to predict functionally important residues. One advantage of this method is its ability to exploit both the structure and sequence of a protein in order to find the catalytic sites. From the integration of THEMATICS structure based predictions with those of the sequence-based phylogenetic trees of INTREPID, each of which perform admirably on their own, significantly better performance is achieved than that of either of these methods alone.

COMP 249

Machine learning application for the high-performance prediction of functional sites in proteins of unknown function

Srinivas Somarowthu(1), somarowthu.s@husky.neu.edu, 360 Huntington Ave, Boston MA 02115, United States; Yujing Wang(2); Andrew Michaelson(3); Huyuan Yang(1); Mary Jo Ondrechen(1)(4). (1) Department of Chemistry and Chemical Biology, Northeastern University, Boston MA 02115, United States (2) College of Professional Studies, Northeastern University, Boston MA 02115, United States (3) Department of Biology, Northeastern University, Boston MA 02115, United States (4) Institute for Complex Scientific Software, Northeastern University, Boston MA 02115, United States

One of the major challenges in genomics is to understand the function of gene products, proteins, from their 3D structures. One viable path for acquiring information about protein function is through the identification of the catalytic and binding sites. While many methods are available for this task, the precision tends to be low. What is needed is a powerful, precise and accurate computational method for identifying important residues. Here, we utilize a new machine learning method, Partial Order Optimum Likelihood (POOL), to predict functionally important residues. One advantage of this method is its ability to exploit both the structure and sequence of a protein in order to find the catalytic sites. From the integration of THEMATICS structure based predictions with those of the sequence-based phylogenetic trees of INTREPID, each of which perform admirably on their own, significantly better performance is achieved than that of either of these methods alone.

COMP 250
Identification of functional subclasses in the ribulose-phosphate binding barrel superfamily using computational tools

Joslynn S. Lee(1), lee.jos@husky.neu.edu, 125 Hurtig Hall, Boston MA 02115, United States; Srinivas Somarowthu(1); Mary Jo Ondrechen(1). (1) Department of Chemistry and Chemical Biology, Northeastern University, Boston MA 02115, United States

The ability to predict the function of a protein from its three-dimensional structure is an important problem in the post-genomic era. A computational method for the identification of function of proteins within a superfamily is described. Superfamily members have similar 3D structures but can have enormous functional diversity. The ribulose-phosphate binding barrel (RPBB) superfamily consists of six different functional subclasses. The present approach is based on a local structural analysis at the interaction sites of the individual proteins. Functional site predictions using Theoretical Microscopic Titration Curves (THEMATICS), INTREPID and Partial Order Optimum Likelihood (POOL) are combined with 3D structural alignment. The patterns of structurally aligned residues at the local sites of interaction are characteristic of function, and can be used to annotate function. Applications to Structural Genomics (SG) proteins of unknown or uncertain function are reported. Two SG proteins are probably correctly annotated, while others are probably incorrectly annotated.

COMP 251

Comparison of the performance of different scoring functions in docking of antidepressants to serotonin transporter

Sankar G Manepalli(1), manepallisankar@gmail.com, 600 forbes avenue, pittsburgh pa 15282, United States; Jeffry D Madura(1); Chris Surratt(2). (1) chemistry and biochemistry, Duquesne university, pittsburgh pa 15282, United States (2) Mylan school of pharmacy, Duquesne university, pittsburgh pa 15282, United States

The serotonin transporter (SERT), which belongs to the Neurotransmitter Sodium Symporter (NSS) family is responsible for the reuptake of serotonin from the synaptic cleft and has been proposed as an important target in the treatment of anxiety and depression. A homology model of SERT is built based upon LeuT_Aa, which belongs to the same family. A number of antidepressants are docked against SERT using different docking methods. The performance of different scoring functions in predicting experimental trend in activities is compared.

COMP 252
Novel scaffold replacement methodology applied to the discovery of P38 MAP kinase inhibitors

**Joe Leonard**(1), jle@chemcomp.com, 1010 Sherbrooke Street West, Suite 910, Montreal Quebec H3A 2R7, Canada ; Nels Thorsteinson(1). (1) Chemical Computing Group, Montreal Quebec H3A 2R7, Canada

A new application for performing scaffold replacements, fragment linking, and R-group optimization is presented. The application is applied to the discovery of novel P38 MAP kinase inhibitors where the diaryl urea core of doramapimod, a known P38 inhibitor, is replaced to produce promising novel candidate inhibitors.

**COMP 253**

Insights into the cross-reaction between *trypanosoma cruzi* antibodies and human beta-1 adrenergic receptor

**Cesar F de Oliveira**(1), cesar@mccammon.ucsd.edu, 9500 Gilman Drive, La Jolla CA, United States ; Andrew J McCammon(1). (1) Department of Chemistry/Biochemistry, UCSD- HHMI, San Diego CA 92093, United States

Chagas' disease, leishmaniasis and African sleeping sickness are caused by parasites of the trypanosomatidae family. Chagas' disease or American trypanosomiasis is endemic to Central and South America and is caused by the parasite of species Trypanosoma cruzi (T.cruzi). Autoantibodies against the β1-adrenergic receptor were detected in sera from patients with cardiomyopathy and the target epitope was localized on the second extracellular loop of this receptor. Evidence of cross-reactivity between antibodies generated against a moiety present in the extracellular loop of the β1-adrenergic receptor, with epitopes in the C termini of the T. cruzi ribosomal proteins has also been found. The epitope on these ribosomal P proteins is represented by a stretch of acidic amino acids. This sequence is particularly conserved in evolution and a single amino acid differentiates the C terminus of the low molecular weight T. cruzi P2β (TcP2β) protein (peptide R13) from that of its human P counterparts (peptide H13). In rabbit heart preparations, electrical disturbances can be alleviated if chagas patients' sera are incubated with small peptides corresponding the C-terminal region of the P ribosomal proteins from humans and T. cruzi, but interestingly, not from Leishmania braziliensis (peptide A13). Primary sequence comparison shows that the three peptides differ in only one amino acid. In order to shed some light in the molecular basis for cross-reaction between anti-P antibodies and β1 adrenergic receptor, long (1.05 μs) accelerated MD (aMD) simulations of peptides R13, H13 and A13 were performed. For each peptide, structures from high populated conformational states sampled by the aMD simulations were extracted and compared with the structure of the second extracellular loop of our b1-adrenergic receptor model. Our aim is to identify
relevant differences in the dynamical behavior of the peptides, and characterize structural factors that can be playing a important role in the antibody recognition.

COMP 254

Molecular orbital study of keto-enol tautomerism of salvinorin A models

Irina Vinar(1)(2), Irina.vinar@sanofi-aventis.com, 1041 Rt. 202-206, Bridgewater NJ 08770, United States; William J. Skawinski(1); Carol A. Venanzi(1). (1) Department of Chemistry and Environmental Science, New Jersey Institute of Technology, Newark NJ 07102, United States (2) DSAR, Sanofi-Aventis U.S., Newark NJ 07102, United States

The importance of considering tautomeric forms in the design of chemical libraries for drug discovery has been recently underscored. Yet most 3D-QSAR studies assume that only one tautomer is responsible for the observed binding affinity. The present work is the first step in investigating the relative energy of the keto and enol tautomers of salvinorin A analogs in order to evaluate which forms should be included in 3D-QSAR modeling studies of their binding affinity at the kappa and mu opioid receptors. Hartree-Fock and density functional theory calculations are combined with the Polarizable Continuum Solvent Model to predict the solvent effect on the keto-enol tautomeric equilibrium of small models of salvinorin A analogs. The calculations investigate whether the electron-withdrawing nature of the 2-position substituent affects the relative energy of the keto and enol tautomers.

COMP 255

Ligand-steered modeling of the cannabinoid receptor 2: Successful applications to rationalize SAR data of selective CB2 inverse agonist and agonist compounds

Sharangdhar S Phatak(1), Sharangdhar.S.Phatak@uth.tmc.edu, 7000 Fannin Ste. 860B, Houston TX 77030-5400, United States; Claudio N Cavasotto(1), claudio.n.cavasotto@uth.tmc.edu, 7000 Fannin, Ste. 860B, Houston Tx 77030, United States; Philippe Diaz(2); Fanny Astruc-Diaz(2); Jijun Xu(3); Mohammed Naguib(3). (1) School of Health Information Sciences, University of Texas Health Science Center at Houston, Houston TX 77030-5400, United States (2) Department of Biomedical and Pharmaceutical Sciences, University of Montana, Missoula MT 59812-1552, United States (3) Department of Anesthesiology and Pain Medicine, University of Texas MD Anderson Cancer Center, Houston TX 77030, United States

Cannabinoid 2 (CB2) receptor, a member of the therapeutically important Class-A G-Protein Coupled Receptors (GPCRs), is an emerging target for the treatment of neuropathic pain and inflammation. However, the lack of a crystal structure or
a reasonable model for CB2, limits any structure-based drug discovery efforts, e.g. to rationalize and improve the functional activities of compounds based on structural information. There is a need to develop reasonable models for CB2 and other members of the GPCR family in order to drive rational drug discovery. In this study, the ligand-steered modeling method (Cavasotto, C.N. et al. J. Med. Chem. 2008, 51(3), 581-588), was adapted and applied to model the 3-dimensional structures of the inactive and active physiological states of the CB2 receptors. These models were successfully applied to rationalize the structure-activity relationship (SAR) data of a series of potent and selective CB2 inverse agonist and agonist compounds. Thus, the ligand-steered modeling approach provided a good platform for structure-based drug discovery efforts targeting CB2 receptors and class-A GPCRs in general.

In the ligand-steered modeling method, known ligands are explicitly used to shape and optimize the binding site through a docking-based stochastic global energy minimization procedure. The ligand and the receptor are considered flexible throughout the modeling process using a Monte-Carlo-based approach, which ensures proper coverage of the energy landscape. This is important in the case of CB2, as there is limited and inconclusive structural information available about protein-ligand interactions. We have used one representative inverse agonist and agonist compound in the modeling process to build:

i. A ligand-steered homology model of CB2 receptor in an inactive state, using the recently crystallized Beta2-adrenergic crystal structure as a template.
ii. A putative agonist-state ligand-steered homology model of CB2 using a multi-template approach, using the Beta2-adrenergic and ligand-free opsin crystal structures as templates.

Given the limited experimental evidence available, we evaluated the accuracy of our models by their ability to explain the structure-activity relationship data of the other compounds in the series. In both cases, our models were not only able to explain the SAR data, but were also able to reproduce known experimental protein:ligand interaction data, such as hydrogen bonding interaction patterns between the ligands and the lysine or the serine residues in the third transmembrane helix for the inverse agonist and agonist compounds, respectively.

COMP 255

Ligand-steered modeling of the cannabinoid receptor 2: Successful applications to rationalize SAR data of selective CB2 inverse agonist and agonist compounds

Sharangdhar S Phatak\(^{(1)}\), Sharangdhar.S.Phatak@uth.tmc.edu, 7000 Fannin Ste. 860B, Houston TX 77030-5400, United States; Claudio N Cavasotto\(^{(1)}\), claudio.n.cavasotto@uth.tmc.edu, 7000 Fannin, Ste. 860B, Houston Tx 77030, United States; Philippe Diaz\(^{(2)}\); Fanny Astruc-Diaz\(^{(2)}\); Jijun Xu\(^{(3)}\); Mohammed Naguib\(^{(3)}\). (1) School of Health Information Sciences, University of Texas Health Science Center at Houston, Houston TX 77030-5400, United States (2)
Cannabinoid 2 (CB2) receptor, a member of the therapeutically important Class-A G-Protein Coupled Receptors (GPCRs), is an emerging target for the treatment of neuropathic pain and inflammation. However, the lack of a crystal structure or a reasonable model for CB2, limits any structure-based drug discovery efforts, e.g. to rationalize and improve the functional activities of compounds based on structural information. There is a need to develop reasonable models for CB2 and other members of the GPCR family in order to drive rational drug discovery. In this study, the ligand-steered modeling method (Cavasotto, C.N. et al. J. Med. Chem. 2008, 51(3), 581-588), was adapted and applied to model the 3-dimensional structures of the inactive and active physiological states of the CB2 receptors. These models were successfully applied to rationalize the structure-activity relationship (SAR) data of a series of potent and selective CB2 inverse agonist and agonist compounds. Thus, the ligand-steered modeling approach provided a good platform for structure-based drug discovery efforts targeting CB2 receptors and class-A GPCRs in general.

In the ligand-steered modeling method, known ligands are explicitly used to shape and optimize the binding site through a docking-based stochastic global energy minimization procedure. The ligand and the receptor are considered flexible throughout the modeling process using a Monte-Carlo-based approach, which ensures proper coverage of the energy landscape. This is important in the case of CB2, as there is limited and inconclusive structural information available about protein-ligand interactions. We have used one representative inverse agonist and agonist compound in the modeling process to build:

i. A ligand-steered homology model of CB2 receptor in an inactive state, using the recently crystallized Beta2-adrenergic crystal structure as a template.

ii. A putative agonist-state ligand-steered homology model of CB2 using a multi-template approach, using the Beta2-adrenergic and ligand-free opsin crystal structures as templates.

Given the limited experimental evidence available, we evaluated the accuracy of our models by their ability to explain the structure-activity relationship data of the other compounds in the series. In both cases, our models were not only able to explain the SAR data, but were also able to reproduce known experimental protein:ligand interaction data, such as hydrogen bonding interaction patterns between the ligands and the lysine or the serine residues in the third transmembrane helix for the inverse agonist and agonist compounds, respectively.

COMP 256

Development of property-encoded shape distribution descriptors for the robust prediction of polymer glass transition temperatures
Prospective prediction of materials properties is one of the challenges within the emerging field of materials informatics. Through the lens of cheminformatics, solutions can be approached through the creation of robust materials quantitative structure-property relationships (MQSPRs). We discuss novel chemical descriptors derived from MOE EP and Active LP (ALP) property-encoded molecular surfaces that are able to capture information relevant to the glass transition temperature of a diverse set of polymers. These descriptors augment traditional atom count and fragment-based descriptors in QSPR modeling, are quick to compute, and given that they are surface-based, are thus amenable to describing polymer-polymer interfaces. As they are modifications to existing Property Encoded Shape Distribution (PESD) descriptors, performance benchmarks will be demonstrated. Validated QSPR methodologies for these descriptors, in addition to modeling results are presented alongside the descriptors.

COMP 256

Development of property-encoded shape distribution descriptors for the robust prediction of polymer glass transition temperatures

Prospective prediction of materials properties is one of the challenges within the emerging field of materials informatics. Through the lens of cheminformatics, solutions can be approached through the creation of robust materials quantitative structure-property relationships (MQSPRs). We discuss novel chemical descriptors derived from MOE EP and Active LP (ALP) property-encoded molecular surfaces that are able to capture information relevant to the glass transition temperature of a diverse set of polymers. These descriptors augment traditional atom count and fragment-based descriptors in QSPR modeling, are quick to compute, and given that they are surface-based, are thus amenable to describing polymer-polymer interfaces. As they are modifications to existing
Property Encoded Shape Distribution (PESD) descriptors, performance benchmarks will be demonstrated. Validated QSPR methodologies for these descriptors, in addition to modeling results are presented alongside the descriptors.

COMP 257

Molecular dynamics studies of amphiphathic peptides embedded within a lipid bilayer

Thuy Hien T Nguyen\(^{(1)}\), tnguyen513@mail.usp.edu, 600 South 43rd Street, Philadelphia PA 19104, United States; Zhiwei Liu\(^{(1)}\); Preston B Moore\(^{(1)}\). \(^{(1)}\)Department of Chemistry and Biochemistry, University of the Sciences in Philadelphia, Philadelphia PA 19104, United States

Ion-channels embedded within the membrane play a crucial role in numerous cell processes such as signaling, energy conversion, and ion conductance. Homooligomeric ion-channels are a common class of ion-channels and are the main focus of many protein-membrane structure studies. The overall goal of this research is to quantify the structural, dynamical, and energetic properties of transmembrane amphipathic peptides which form homooligomeric ion-channels. In order to achieve this goal, we performed both all-atom (AA) and coarse grain (CG) molecular dynamics (MD) simulations on systems with different oligomeric states, specifically in the tetramer, pentamer and hexamer bundle. These simulations allowed us to determine which oligomeric state is more stable. In addition, we also analyzed the tilt angle, radial distribution, water hydration, area per lipid, and bilayer thickness from both AA and CG simulation data. In summary, our simulations can lead us to understand the assembly of peptides in membranes better, which is of great importance when designing antimicrobial, antiviral, and other pharmaceutical agents that will target ion-channels.

COMP 257

Molecular dynamics studies of amphiphathic peptides embedded within a lipid bilayer

Thuy Hien T Nguyen\(^{(1)}\), tnguyen513@mail.usp.edu, 600 South 43rd Street, Philadelphia PA 19104, United States; Zhiwei Liu\(^{(1)}\); Preston B Moore\(^{(1)}\). \(^{(1)}\)Department of Chemistry and Biochemistry, University of the Sciences in Philadelphia, Philadelphia PA 19104, United States

Ion-channels embedded within the membrane play a crucial role in numerous cell processes such as signaling, energy conversion, and ion conductance. Homooligomeric ion-channels are a common class of ion-channels and are the main focus of many protein-membrane structure studies. The overall goal of this research is to quantify the structural, dynamical, and energetic properties of
transmembrane amphipathic peptides which form homo-oligomeric ion-channels. In order to achieve this goal, we performed both all-atom (AA) and coarse grain (CG) molecular dynamics (MD) simulations on systems with different oligomeric states, specifically in the tetramer, pentamer and hexamer bundle. These simulations allowed us to determine which oligomeric state is more stable. In addition, we also analyzed the tilt angle, radial distribution, water hydration, area per lipid, and bilayer thickness from both AA and CG simulation data. In summary, our simulations can lead us to understand the assembly of peptides in membranes better, which is of great importance when designing antimicrobial, antiviral, and other pharmaceutical agents that will target ion-channels.

COMP 258

Structural basis of pregnane X receptor binding promiscuity

Chi Ho Ngan(1), chihongan@gmail.com, 44 Cummingston St, Boston MA 02215, United States; Dmitri Beglov(1); Aleksandra N. Rudnitskaya(2); Dima Kozakov(1); David J. Waxman(3); Sandor Vajda(1)(4). (1) Department of Biomedical Engineering, Boston University, Boston MA 02215, United States (2) Department of Chemistry, University of Massachusetts, Boston MA 02125, United States (3) Department of Biology, Boston University, Boston MA 02215, United States (4) Department of Chemistry, Boston University, Boston MA 02215, United States

The human pregnane X receptor (PXR) binds a broad range of structurally diverse compounds. Unlike most promiscuous proteins, which adapt their shapes to ligands, the structures of apo and ligand-bound PXR forms are very similar. Using computational solvent mapping hot spots (i.e. regions of the protein surface that are major contributors to the binding free energy) are identified and characterized. The PXR binding site has a well-defined hot-spot structure, four of which reside on the periphery and one near the center. Three of these hot spots are already present in the ligand-free protein. The most important hot spot is defined by W299, F288, and Y306, which are structurally and sequentially conserved. All known PXR ligands interact with this weakly specific region. The ligands can extend into multiple hot spot regions, and the multiplicity of potential binding arrangements explains why PXR can accommodate structurally diverse compounds.

COMP 259

Coarse-grained modeling of antibody self-association

Anuj Chaudhri(1), chaudhri@uchicago.edu, 5735 S. Ellis Ave., Chicago IL, United States; Tim Kamerzell(2); Gregory A. Voth(1). (1) Department of
Therapeutic antibodies have become increasingly important for the treatment of cancer and other diseases. The fastest process of delivering these formulations has been intravenous administration. However frequent administration requires subcutaneous delivery, which requires high concentration formulations. Developing these formulations has been a challenge due to stability and manufacturing issues. High concentration formulations also lead to very high viscosities of these solutions. One of the major reasons for these high viscosities has been related to the reversible self-association network that the antibodies can form. Studying high concentrations of these antibodies using atomistic methods is highly intractable due to the extremely large system sizes and long time scales. In this work, we study this problem of self-association using a coarse-grained model of the Genentech Immunoglobulin (IgG1) molecule. The parameters in the model are calculated using atomistic trajectory data. The model can be used to make predictions on the possible network topologies that can be generated based on experimental observations of the most important interactions in the system.

COMP 260

An N log N generalize Born approximation

Ramu Anandakrishnan(1), ramu@vt.edu, 131 Spring Dr, Newport VA 24128, United States; Alexey V Onufriev(1). (1) Department of Computer Science, Virginia Tech, Blacksburg Virginia 24060, United States

Most practical implicit solvent models, including the popular generalized Born (GB) approximation, have so far had a serious drawback: poor scaling - $O(N^2)$ - with number of atoms. As a consequence applying the GB model to large scale simulations can be computationally prohibitive. We previously developed the hierarchical charge partitioning (HCP) approximation for speeding up computation of pairwise electrostatic interactions in biomolecular systems. The approximation is based on multiple levels of natural partitioning of biomolecular structures into a hierarchical set of its constituent structural components. For a structure consisting of N charges, the computational cost of computing the pairwise interactions via the HCP scales as $O(N \log N)$, under assumptions about the structural organization of biomolecular structures generally consistent with reality. Our previous study showed that the HCP approximation, for simple Coulomb interactions with fixed dielectric constants, is comparable in accuracy to the industry standard particle mesh Ewald (PME) method and significantly more accurate than the spherical cutoff method. A critical benefit of the HCP approximation is that it is algorithmically very simple, and unlike the PME, the HCP is straightforward to use with implicit
solvent models. The HCP approximation for the GB implicit solvent model has now been implemented in the freely available nucleic acids builder (NAB) molecular dynamics package in Amber tools. Simulations of a representative set of structures, ranging in size from approximately 1000 to 3 million atoms, demonstrates that simulations using the HCP approximation for GB can be stable, and multiple orders of magnitude faster than GB simulations without the HCP approximation, consistent with N log N scaling. This study also examines the consequence of violating Newton’s 3rd law by multiscale methods.

COMP 261

Development of internal potentials for a protein coarse-grain forcefield

Kenny Nguyen\(^{(1)}\), nguyen@hydrogen.usp.edu, 600 South 43rd Street, Philadelphia Pennsylvania 19104, United States; Jhenny Galan\(^{(1)}\); Zhiwei Liu\(^{(1)}\); Russell DeVane\(^{(2)}\); Preston Moore\(^{(1)}\). (1) Department of Chemistry and Biochemistry, University of the Sciences in Philadelphia, Philadelphia Pennsylvania 19104, United States (2) Department of Chemistry, Temple University, Philadelphia Pennsylvania 19122, United States

We introduce a method to create internal potentials for a coarse-grain forcefield, which will be used to simulate proteins. Coarse-graining a reduced model method, which allows spatial and temporal scales to exceed all-atom molecular dynamics by three orders of magnitude. Nonbonded forcefield parameters were already developed by DeVane and co-workers for amino acids, and will be used for our study (DeVane et al., J. Chem. Theory Comput. 2009, 5, 2115). Internal potentials, which consist of equilibrium bond distances, bond angles, dihedral angles, and their respective force constants, will be paramertized using all-atom simulation data. Currently, we are using the following proteins: aquaporin, barnase-barstar, chymotrypsin, myoglobin, photosynthetic reaction center, and potassium ion channel to paramertize, test, and validate the coarse-grain forcefield. The goal is to determine a set of internal parameters to simulate proteins using coarse-grain molecular dynamics.

COMP 262

Interaction of 12c4 with alkali metal cation in aqueous solution: Theoretical investigation using polarized continuum model

Hideya Koizumi\(^{(1)}\), hkoizumi@astate.edu, P.O. Box 419, State University AR 72467, United States; Hashim Ali\(^{(1)}\); Tanika Arora\(^{(1)}\); William Andres Burns\(^{(1)}\). (1) Chemistry and Physics, Arkansas State University, State University AR 72467, United States
Interaction between crown ether and alkali metal cation has been studied extensively both gas and aqueous phase. We investigate the interaction of 12c4 and alkali metal ion in aqueous solution using polarized continuum model. We have located multiple geometrical conformers in aqueous solution at several different levels of theories including several types of density functional. Resulting solution phase bond dissociation energies at sufficiently high level of theory for all statistically relevant conformers are compared with those in gas phase by utilizing similar technique used in Hong et al. The relative energies of the geometrical excited states are similar to those found in gas phase. However, much weaker interactions between cation and 12c4 are found in aqueous solution. Although PM6 relative conformational energies are, in general, consistent with DFT result, overestimations of solvent interference are responsible for the faulty prediction of the stability of the cation complexes.

COMP 263

LowModeMD: Conformational search of small molecules, macrocycles, and protein loops

Elizabeth Sourial(1), esourial@chemcomp.com, 1010 Sherbrooke Street West, Suite 910, Montreal Quebec H3A 2R7, Canada; Paul Labute(1). (1) Chemical Computing Group, Montreal Quebec H3A 2R7, Canada

We present a method for conformational search of complex molecular systems including macrocycles and protein loops. The method is based upon perturbing an existing conformation along a molecular dynamics trajectory using atomic velocities with kinetic energy concentrated on the low-frequency vibrational modes followed by energy minimization. A novel Chebyshev polynomial filter is used to heavily damp the high-frequency components of a randomly generated Maxwell-Boltzman velocity vector. The method is efficient, straightforward to implement, and requires only standard forcefield energy and gradient evaluations. The results of several computational experiments suggest that the method is capable of efficiently sampling low strain energy conformations of complex systems with non-trivial non-bonded interaction networks.

COMP 264

Parameterization of small drug-Like ligands using CHARMM General Force Field

Kalyan C Imadisetty(1), imadisettyk@duq.edu, 600 forbes avenue, Pittsburgh PA 15219, United States; Jonathan Gibbons(1); James Branco(1); Jeffry D Madura(1). (1) Chemistry and Biochemistry, Duquesne university, Pittsburgh PA 15219, United States
Monoamine transporters (MAT) like dopamine transporter, serotonin transporter, and norepinephrine transporter are the targets for several antidepressants and addictive psycho stimulants. Ligands such as clomipramine, desipramine, imipramine, dopamine, serotonin, fluoxetine, cocaine, sertraline, benztrapine, citalopram, bupropriion, and methylphenidate act on these MAT's. In computational drug design, accurate prediction of binding free energies requires ligand force field parameters to be well characterized. Comprehensive development of parameters for the above ligands has not been done so far. The above ligands have been parameterized using CHARMM General Force field (CGenFF) method. The parameters developed by this method were validated by calculating hydration energies, binding energies, and pKa's using the free energy perturbation method (FEP) and continuum electrostatic method and comparing the calculated results against the experimental values. The calculated values are in good agreement with the experimental values.

COMP 265

Glycogen phosphorylase inhibitor identification targeting the inhibitor binding site using virtual screening and experimental validation

Shijun Zhong(1), sjzhong@gmail.com, 20 Penn Street, Baltimore MD 21201, United States ; Kyra-Melinda Alexacou(2); Marianna A Charavgi(2); Costas Tiraidis(2); Spyros E Zographos(2); Demetres D Leonidas(2); Evangelia D Chrysina(2); Nikos G Oikonomakos(2); Alexander D MacKerell, Jr.(1). (1) Department of Pharmaceutical Science, University of Maryland Baltimore, Baltimore MD 21201, United States (2) Institute of Organic & Pharmaceutical Chemistry, the National Hellenic Research Foundation, Athens 116 35, Greece

Virtual screening of one million drug-like compounds was performed to identify novel inhibitors of glycogen phosphorylase (GP). Given the function of GP to degrade glycogen leading to the release of glucose such inhibitors have the potential to be used to control the amount of glucose and, therefore, may be of utility for the treatment of type 2 diabetes. The screening targeted the inhibitor binding site of GP. The approach involved a series of steps including two rounds of docking using different scoring approaches, chemical clustering for picking up diverse hits, and filtering with empirical rules for selecting compounds with lead like physical properties. From 175 compounds selected from the screen, 41 were obtained and subjected to kinetic and crystallographic analysis. This process yielded inhibitors with micromolar affinities that bind in the targeted inhibitor site of the protein.
NMR spectroscopy provides structural information of biomolecules. The measured values are averages over time and an ensemble of structures. Therefore, knowledge of distributions and dynamics of the system are required. Molecular Dynamics (MD) simulation generates an ensemble of structures but may suffer from limited sampling in conformational or time space. To enhance conformational sampling we proposed an adaptive restraining method (Christen, M. (2007) J. Biomol. NMR 39 265-273) for structure refinement using a local elevation potential energy function. Experimental values are used as restraints to bias the simulation towards relevant structures. In this study a MD simulation was biased to stereospecifically assigned $^3J_{\alpha\beta}$-coupling constants of Plastocyanin (Moore, J. M. (1991) J. Mol. Biol. 221, 533-555) using this method. The simulated ensemble of structures showed improved agreement to experimental values as J-coupling constants and nuclear Overhauser effect intensities compared to structures derived using single-structure refinement.
J-coupling constants and nuclear Overhauser effect intensities compared to structures derived using single-structure refinement.

COMP 267

First-principles design of conductance switching in functionalized carbon nanotubes

Elise Y Li(1), elise_li@mit.edu, Rm. 13-4066, 77 Mass Ave, Cambridge MA 02139, United States; Nicolas Poilvert(2); Nicola Marzari(2). (1) Department of Chemistry, Massachusetts Institute of Technology, Cambridge MA 02139, United States (2) Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge MA 02139, United States

Functionalization of SWNT through addition reactions represents an effective method to engineer or manipulate carbon nanotubes. For armchair CNTs, the conductivity is often decreased by orders of magnitude by the introduction of monovalent functional groups which disrupt the conjugated \( \pi \) network, whereas in [1+2] cycloadditions of carbenes or nitrenes, the \( sp^2 \) environment and therefore CNT metallicity can be recovered due to the sidewall bond breakage induced by the cyclopropane strain. In real systems, this bond cleavage depends heavily on the chirality and curvature of the tube, and the chemical nature of the addends. Here we explore the underlying mechanism of bond-cleavage chemistry in [1+2] cycloadditions on armchair carbon nanotubes using first-principles calculations. We find the high strain energy in cyclopropane moiety can be compensated by a through space \( \pi \) orbital interaction between the addend and the CNT which lowers the HOMO energy significantly in closed-bond configuration. A bond opening or closing switch marked by large conductance change can therefore be devised by modulating the proximity of the addend \( \pi \) system and the tube surface via optical or electrochemical control, which potentially has extensive applications in nanoscale devices.

COMP 268

Binding free energy of cocaine and citalopram to the serotonin transporter using molecular dynamics

James J Brancho(1), Pittsburgh Pennsylvania 15219; Kalyan Immadisetty(1); Jonathon Gibbons(1); Jeffrey D. Madura(1). (1) Department of Chemistry and Biochemistry, Duquesne University, Pittsburgh Pennsylvania 15219, United States
The specifics of psychoactive molecules interacting with plasmemmal transporter proteins are of particular interest to pharmacologists and medicinal chemists working in the drug design area; computational models of these interactions are therefore very important to drug discovery. Two important compounds in this field are cocaine and citalopram. Topology and parameter files for cocaine and citalopram were developed using the CHARMM General Force Field (CGENFF) and were verified and refined using crystal lattice simulations in CHARMM.

These molecules were then docked to the serotonin transporter homology model (SERT) and the binding free energy was calculated using free energy perturbation in NAMD. The results of free energy perturbation calculations will be presented.

COMP 268

Binding free energy of cocaine and citalopram to the serotonin transporter using molecular dynamics

James J Brancho(1), Pittsburgh Pennsylvania 15219; Kalyan Immadisetty(1); Jonathon Gibbons(1); Jeffrey D. Madura(1). (1) Department of Chemistry and Biochemistry, Duquesne University, Pittsburgh Pennsylvania 15219, United States

The specifics of psychoactive molecules interacting with plasmemmal transporter proteins are of particular interest to pharmacologists and medicinal chemists working in the drug design area; computational models of these interactions are therefore very important to drug discovery. Two important compounds in this field are cocaine and citalopram. Topology and parameter files for cocaine and citalopram were developed using the CHARMM General Force Field (CGENFF) and were verified and refined using crystal lattice simulations in CHARMM.

These molecules were then docked to the serotonin transporter homology model (SERT) and the binding free energy was calculated using free energy perturbation in NAMD. The results of free energy perturbation calculations will be presented.

COMP 269

Prediction of Lapatinib selectivity with kinases in the ErbB family
The ErbB family of tyrosine kinases is an important target for development of anti-cancer drugs in the class known as "molecular targeted therapeutics." In the targeted therapeutics class, one small molecule inhibitor (lapatinib) has been approved for treatment of breast cancer. Unlike many compounds, lapatinib is a dual kinase inhibitor which binds both EGFR and HER2. Interestingly, lapatinib binds much less tightly with ErbB4, another member in the ErbB family, which shares almost 80% sequence homology with EGFR and HER2. A detailed and precise understanding of the key determinants which drive lapatinib specificity and affinity for these targets will guide rational design for the next generation breast cancer drugs. Results from molecular dynamics, free energy calculations and energy component analysis will be presented from simulations of lapatinib with ErbB family members.

COMP 270

Correcting for errors in one-body and two-body interactions of small molecules in DFT with empirical potential forms parameterized through force matching

Yang Song(1), ysong@bu.edu, 590 Commonwealth Ave., Chemistry Department, Boston University, Boston MA 02215, United States; Feng Wang(1). (1) Chemistry, Boston University, Boston MA 02215, United States

Empirical corrections terms are parameterized following the force matching approach to provide corrections to density functional theory. In our approach, the empirical correction terms include one-body terms to capture errors in intramolecular interactions and two-body terms to capture errors in hydrogen bond and in van der Waals interactions. Our approach is tested for correcting the BLYP exchange-correlation functional on water and methane water mixture to the quality of QCISD. Other than dispersion, it is found that both one-body terms and the two-body hydrogen bond correction term significantly improve the potential energy surface described by the BLYP functional.

COMP 271

Identification of novel non-hydroxamate anthrax toxin lethal factor inhibitor scaffolds using in silico and in vitro high-throughput screening methodologies

Ting-Lan Chiu(1), tlchiu@umn.edu, 717 Delaware Street SE, Minneapolis Minnesota 55414-2959, United States; Jonathan Solberg(2); Satish Patil(3); Todd
Anthrax is an infectious disease caused by the gram-positive, rod-shaped bacterium *Bacillus anthracis*. The bacilli secrete a tripartite exotoxin, of which the lethal factor (LF) is the primary agent of cell death. LF has therefore been an attractive target for small-molecule drug design; however, no LF inhibitor has reached the market as a drug. Here we outline a novel computational protocol which screened approximately thirty-five million nonredundant compounds using shape-based searching, docking and scoring, and drug-like filtering, for potential activity against LF and possible use as adjunct therapeutics for those exposed to anthrax in an emergency situation. This *in silico* protocol identified 301 preliminary hits, among which thirty-nine were commercially available and were subjected to *in vitro* screening using high-throughput fluorescence resonance energy transfer (FRET) assays. Five compounds demonstrated low micromolar inhibition against LF; three of these exhibited IC\(_{50}\) values less than 100 μM. Our docking simulations predicted that these three preliminary hits may interact with key residues in at least two of the three LF active-site regions. Our virtual screening results also indicate that the introduction of non-hydroxamate and even monodentate zinc-binding groups is not likely to negatively impact compound activity, as long as critical hydrophobic interactions with at least two of the LF subsites are preserved.
lethal factor (LF) is the primary agent of cell death. LF has therefore been an attractive target for small-molecule drug design; however, no LF inhibitor has reached the market as a drug. Here we outline a novel computational protocol which screened approximately thirty-five million nonredundant compounds using shape-based searching, docking and scoring, and drug-like filtering, for potential activity against LF and possible use as adjunct therapeutics for those exposed to anthrax in an emergency situation. This in silico protocol identified 301 preliminary hits, among which thirty-nine were commercially available and were subjected to in vitro screening using high-throughput fluorescence resonance energy transfer (FRET) assays. Five compounds demonstrated low micromolar inhibition against LF; three of these exhibited IC$_{50}$ values less than 100 μM. Our docking simulations predicted that these three preliminary hits may interact with key residues in at least two of the three LF active-site regions. Our virtual screening results also indicate that the introduction of non-hydroxamate and even monodentate zinc-binding groups is not likely to negatively impact compound activity, as long as critical hydrophobic interactions with at least two of the LF subsites are preserved.

COMP 272

HS-PHARM-GEN: An automated tool for generating receptor-based pharmacophores from ligand-free protein structures

Gwo-Yu Chuang$^{(1)}$, chuang@unistra.fr, 74 route du Rhin, Illkirch Bas-Rhin 67401, France; Caterina Barillari$^{(1)}$; Didier Rognan$^{(1)}$. (1) Department of Pharmacy, University of Strasbourg, Illkirch 67401, France

We have developed HS-PHARM-GEN, a fully automated tool to generate 3-D pharmacophore models using 3-D structure of binding cavities as input, to generate pharmacophore models for apo-proteins where protein-ligand binding mode is not available. HS-PHARM-GEN first determines the hotspot atoms in the binding cavity which are more likely to interact with ligands using a knowledge-based model (HS-PHARM) trained on atom-based cavity fingerprints of protein-ligand complexes [1]. Then it places pharmacophore features near the hotspot atoms using topological rules, and finally the number of features is reduced by removing features with steric clashes to the protein or in unfavorable physiochemical environment, and by clustering. HS-PHARM-GEN has been tested on several different types of protein targets (enzymes, nuclear receptors, GPCRs), and the pharmacophore models produced by HS-PHARM-GEN are able to select the true actives from the decoy sets with high sensitivity, selectivity, and enrichment factors.


COMP 273
Molecular interaction footprints: A docking rescoring method

Trent E. Balius(1), tbalius@aol.com, Math Tower, Stony Brook NY 11794-3600, United States; Robert C. Rizzo(1); Sudipto Mukherjee(1). (1) Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook NY 11794-3600, United States

Our laboratory has previously used "molecular footprints", in the context of molecular dynamics simulations, to study the origins of ligand binding and drug resistance for a variety of targets. Footprints are defined as the per-residue decomposition of intermolecular van der Waals (VDW), Coulombic (ES), or hydrogen bonding (HB), and may be thought of as a protein-ligand interaction signature (or pharmacophore). In this study, we explore the utility of using footprints to compare binding geometries (poses) between two different ligands in the context of docking. The hypothesis is the new scoring function will be useful to identify molecules making similar interaction signatures as a known reference molecule (i.e. inhibitor or native substrate). Results using footprint rescoring, as implemented into the program DOCK, for both pose identification (Stony Brook test set) and enrichment (UCSF DUD database) will be presented. Example applications to virtual screening will also be discussed.

COMP 274

Structural characterization of layered perovskite niobates using projector augmented wave (PAW) method of density functional theory (DFT)

Jhashanath Adhikari(1), jadhikari@clarku.edu, 950 Main Street, Worcester MA 01610, United States; Luis J Smith(1). (1) Carlson School of Chemistry and Biochemistry, Clark University, Worcester MA 01610, United States

Electric Field Gradient (EFG) tensor at the nuclear positions has been calculated using DFT based PAW methods for different layered perovskite niobates. This method has been used for the full structural relaxation/ geometry optimization with and without symmetry restraints. EFG is a sensitive NMR parameter for local structure determination and provides information about the symmetry. The structure of distinct $^{93}$Nb and $^2$H sites observed via solid state NMR in KNb$_3$O$_8$, K$_4$Nb$_6$O$_{17}$, KCa$_2$Nb$_3$O$_{10}$, KSr$_2$Nb$_3$O$_{10}$ and their acid exchanged forms has been identified. The changes in the local structure resulting from compositional changes have been correlated using EFG tensor information. In addition, taking into account the total energy, forces on atoms and XRD patterns of various, slightly modified structures that lie within the uncertainty region, possible different space groups have been proposed. All-electron full-potential LAPW calculations have been performed for a confirmation of the results.

COMP 275
Fragment based design and biophores using geometric and chemical patterns of interactions at interface of ligand-receptor complex crystal structures

**Raed Khashan**(1)(2), raed@email.unc.edu, 301 Beard Hall, CB# 7360, Chapel Hill NC 27599, United States; **Weifan Zheng**(2); a Tropsha(1). (1) Division of Medicinal Chemistry & Natural Products, University of North Carolina at Chapel Hill, Chapel Hill NC 27599, United States (2) Department of Pharmaceutical Sciences, North Carolina Central University, BRITE Institute, Chapel Hill NC 27599, United States

We present a novel technique to mine ligand-receptor complexes and derive meaningful information that can be used to do fragment-based virtual screening, as well as, providing structure-based bioisosteric replacements. The idea is based on identifying patterns of interactions between ligand and receptor atoms for a database of complexes. This can be done by, first, applying Almost Delauney tesselation to define interfacial atoms between ligand and receptor. These interfacial atoms form the pattern of interaction for which both geometric and chemical atom types are memorized. Thus given a binding pocket, we can fill it with fragments using the memorized patterns of interactions (FragVLib). On the other hand, given a particular chemical group, we can identify other groups in the memorized patterns that share the same binding site with that particular chemical group (Geolsosteres). (Check http://www.unc.edu/~raed for Software).
Cryptic adaptive pockets: What can we learn from protein sequence and structure regarding new pocket formation?

Vasu Chandrasekaran\textsuperscript{(1)}, vasu.chandrasekaran@novartis.com, 250 Mass Ave., Cambridge MA 02139, United States; C. Gregory Paris\textsuperscript{(1)}; Amy Keating\textsuperscript{(2)}. (1) GDC-Oncology, Novartis Institutes for Biomedical Research, Cambridge MA 02139, United States (2) Department of Biology, Massachusetts Institute of Technology, Cambridge MA 02139, United States

Methods for computational assessment of protein druggability typically involve, as a first step, a comprehensive detection of all surface pockets, followed by a prediction and identification of a possible ligand binding site through analysis of the pocket characteristics. However, most of the pocket finding programs do not take protein conformational flexibility into account and treat them as static, rigid structures. The drawback of this approach is that it might miss pockets on protein surfaces that form only occasionally, and are not often seen in the unbound experimental structures. This problem is amplified further when studying protein-protein complexes where the interaction surfaces are often flat and apparently lack pockets for small molecules to bind. But, it is known from previous examples of ligand-bound protein structures that some protein-protein interaction (PPI) surfaces are capable of adopting conformations where a pocket is present.

Here, we study some known examples of proteins known to form such “cryptic adaptive pockets” to understand the factors that play a role in the pocket formation, in an attempt to extract descriptors that can predict new pocket formation or pocket expansion when applied to other new protein targets. We have studied various surface characteristics like amino acid composition, sequence conservation and energetic hotspots from multiple available experimental structures for these proteins, and have also investigated their global dynamic properties to understand the types, range and origin of flexibilities associated with the cryptic pocket formation. Results will be summarized, including the possibility that such cryptic pockets are present on many non-PPI surfaces.

COMP 278

Folding mechanism of BBL

Jue Fan\textsuperscript{(1)}, jfan@clarku.edu, 950 Main Street, Worcester MA 01610, United States; Mojie Duan\textsuperscript{(1)}; Li Han\textsuperscript{(2)}; Shuanghong Huo\textsuperscript{(1)}. (1) Department of Chemistry and Biochemistry, Clark University, Worcester MA 01610, United States (2) Department of Mathematics and Computer Science, Clark University, Worcester MA 01610, United States

The folding mechanism of BBL has been investigated under acidic pH without salt as well as under neutral pH with salt in literature. “Downhill” folding
mechanism was proposed based on a combination of spectroscopic techniques and calorimetry and was further supported by dispersion of melting temperatures for individual residues at pH 5.3. Two-state folding behavior was detected at pH 7 with salt. The discrepancy in folding mechanism was previously attributed to the different length of the protein and the presence of fluorescence tag. It was shown recently that the native states were different at different conditions. We have studied the folding pathways of BBL using a graph-based reaction path algorithm. Our results have revealed the complexity of BBL energy landscape. This algorithm is proved to be a new tool to study protein folding and RNA folding.

COMP 279

Engineering intein-modified cell wall-degrading enzymes: Prediction of non-native intein insertion sites with native-like properties

James R Apgar(1), james.apgar@agrivida.com, 200 Boston Ave, Suite 3100, Medford MA 02144, United States; Binzhang Shen(1); Xavier Zuo(1); Xueguang Sun(1); Philip Lessard(1); Sarah Dohle(1); Jonas Emery(1); Gabor Lazar(1); Humberto de la Vega(1); Mary Ross(1); Derek Sturtevant(1); Taran Shilling(1); R. Michael Raab(1). (1) Agrivida Inc., Medford MA 02144, United States

Inteins are self-excising protein domains that modulate the activity of their host enzyme. Agrivida utilizes this switching mechanism to control the activity of cell wall degrading enzymes (CWDE), allowing them to be expressed in plants without detrimental impacts and reducing costs for downstream cellulosic biofuels production. Significant features of native insertion sites were identified and used to predict non-native insertion sites. An SVM model of the insertion site cassettes was generated and utilized to rank order the native sites in the top 25% of all relevant sites. Additionally, ~250 homology models of native exteins were generated to determine local secondary structure and proximity of these insertion sites to the active site residues. Both showed significant correlation with native intein insertion sites. These properties were used to predict suitable insertion sites and design intein-modified CWDE's. Several of these were developed and showed native-like splicing and switchable activity.

COMP 280

SZMAP: Mapping solvent thermodynamics in binding sites

J. Michael Word(1), word@eyesopen.com, 9 Bisbee Court, Suite D, Santa Fe NM 87508, United States; Matthew T Geballe(1); Anthony Nicholls(1). (1) OpenEye Scientific Software, Inc., Santa Fe NM 87508, United States

Continuous solvent models, such as Poisson-Boltzmann theory, can be effective for a wide range of problems yet do not capture discrete solvent effects that can
be important in enclosed spaces such as binding cavities. SZMAP is a hybrid method that employs one explicit water probe in a continuum solvent to map thermodynamic quantities of a water molecule near protein surfaces. This approach offers an explanation for over-estimation by continuum methods of the cost of removing water from sites adjacent to metals and can identify regions where solvent molecules may be stabilized or destabilized by the nature of the binding site environment. This information could be used as a correction factor for continuum solvent calculations or to guide the design of ligand analogs to optimize binding affinity. Examples of SZMAP calculations on congeneric series illustrate the relationship between solvent thermodynamics and binding affinity.

COMP 281

Improvement of an implicit solvent model

Hai Nguyen\(^{(1)}\), hainm.comp@gmail.com, SUNY, Stony Brook NY 11794, United States; Daniel R. Roe\(^{(1)}\); Lauren Wickstrom\(^{(2)}\); Carlos Simmerling\(^{(1)(2)}\). (1) Department of Chemistry, Stony Brook University, Stony Brook NY 11794, United States (2) Graduate Program in Biochemistry and Structural Biology, Stony Brook University, Stony Brook NY 11794, United States

Solvent in Molecular Dynamics (MD) Simulations can be represented by individual solvent molecules (explicit solvent) or by a continuous medium (implicit solvent). Implicit solvent is of great interest due to significantly reducing degrees of freedom and in some cases it is also more computationally efficient than explicit solvent. Generalized Born (GB) model is one of the fastest implicit solvent models and it is widely used in MD simulation. Different levels of approximation make the GB model have noticeable limitations. Because the quality of a GB model is heavily affected by its empirical parameters used in calculating solvation energy, in this study, we have refitted these parameters for GBNNeck, one of the fastest GB models. Solvation energy and effective radii from Poisson-Boltzmann (PB) calculation are used as benchmarks for our work. Comparing to other GB models like GBOBC and GBNNeck, our newly optimized GB model (GBNeck2) has better agreement to PB method in terms of reproducing solvation energies for a variety of systems ranging from peptides to proteins. Secondary structure preferences are also in much better agreement with simulations in explicit water as compared to the original model.

COMP 282

Hydrated or not? The factors affecting the hydration of a protein cavity

Hongtao Yu\(^{(1)}\), htaoyu1@gmail.com; Steve W Rick\(^{(1)}\). (1) Chemistry, University of New Orleans, New Orleans LA 70148, United States
Proteins are well packed biopolymers, but packing defects exist ubiquitously in proteins of all sizes. Cavities large enough to accommodate at least one water molecule account for one percent of total protein volumes, yet only about one out of four of these cavities are observed to be occupied by NMR, X-ray crystallography, or neutron diffraction. Generally, whether a cavity is hydrated by a water molecule or not depends on the environment the water molecule may undergo inside the cavity. In this study, we have calculated the free energies, entropies, and enthalpies for transferring single water molecules into various spherical cavities with molecular dynamics simulation, and the factors that may affect the water occupation are interpreted in terms of cavity sizes and hydrophobicities.

COMP 283

Multi-scale modeling of coarse grained protein interactions: A CHARMMing implementation

Frank C. Pickard(1)(2), frank.pickard@nih.gov, 9000 Rockville Pike, Bethesda MD 20892; Benjamin T. Miller(2); Henry L. Woodcock(3); Henry F. Schaefer(1); Bernard R. Brooks(2). (1) Center of Computational Quantum Chemistry, University of Georgia, Athens GA 30602, United States (2) National Heart, Lung and Blood Institute, National Institutes of Health, Athens GA 30602, United States (3) Department of Chemistry, University of South Florida, Tampa FL 33620, United States

Coarse grained (CG) models are a class of computational tools that efficiently simulate proteins and lipids. In the Klimov-Thirumalai (KT) implementation, each residue is represented by two centers: the Cα and the side-chain center of mass. Structure is biased towards the native state using Leonard-Jones potentials that mimic hydrogen bonding and side-chain interactions. This simplistic model allows for computational speedups, while still retaining semi-quantitative agreement with experiment. The KT model was implemented in CHARMMing, a web-based user interface for the CHARMM molecular simulation package, using Python. To validate the KT model, and its implementation, a mutant of the zipper domain of the yeast protein GCN4 was simulated at both all-atom and CG resolutions. Preliminary results are promising, as computed properties such as melting point and α-helix orientation were consistent. This model will be extended to incorporate transferable parameters to inter-chain dynamics, while retaining the structure based approach for intra-chain dynamics.

COMP 284

Charge transfer model for simulation of water and aqueous solutions
A new model which allows the charge transfer between molecules is presented. This model is based on existing fluctuating charge models, with a different charge neutrality constraint. We will address methods to treat new difficulties associated with charge transfer, involving the amount of charge transfer at large separation and the treatment using molecule dynamics. Results assessing the importance of charge transfer for liquids will be presented.

COMP 285

Statistical analysis of scoring relative to experimental affinities (CSAR 2010 Benchmark Exercise)

Richard D. Smith(1), dicksmit@umich.edu, 428 Church St., Ann Arbor MI 48109-1065, United States; James B. Dunbar(1); Heather A. Carlson(1). (1) Department of Medicinal Chemistry, University of Michigan, Ann Arbor, Ann Arbor MI 48109-1065, United States

As part of the Community Structure-Activity Resource (CSAR) center, a set of 345 high-quality protein-ligand crystal structures were assembled with experimentally determined $K_d$ or $K_i$ information from the literature. We encouraged the community to score the crystallographic poses of the complexes by any method of their choice. The goal of the exercise was to 1) evaluate the current ability of the field to predict activity from structure and 2) investigate the properties of the complexes and methods that appear to hinder scoring. A total of 17 different methods were submitted with numerous parameter variations for a total of 51 sets of scores from 12 participating groups. Spearman rho, Kendall Tau, and linear regression was used to correlate scores to the experimental values. In order to compare differences in $R^2$ between the methods (i.e., is the difference between 0.4 and 0.35 statistically significant?), we used Levene's F-test to analyze the variance in the residuals from each fit. The data across all the participants was combined to identify complexes that were poorly scored across the majority of the scoring methods vs those that were scored well across the majority. The two sets were compared using the Wilcoxon rank-sum test to assess any significant differences in the distributions of 319 ligand properties and 55 quality parameters of the crystal structures. We also investigate differences in the composition of the binding site, including bridging water molecules, to provide insight into limitations of scoring methods.

COMP 286

Hydration and confinement effects on helix formation in a 23-residue polypeptide
Several experiments have suggested that an α-helix forms in polypeptides inside the ribosome exit tunnel. In this work we use molecular dynamics simulation to explore confinement effects on helix stability of a 23-residue polyalanine peptide confined to carbon nanotubes (CNT) of varying diameter open to a water reservoir. The results are different from those reported earlier by Sorin and Pande (JACS. 128, 6316-17 (2006) for the same polypeptide in a periodically replicated nanotube containing water, showing that the influence of solvent on helix formation is determined by whether the confined system is open or closed to an external reservoir.

COMP 287

Assessment and optimization of docking-based virtual screening for GPCR ligands: Not only crystal structures but also homology models

Using the β2-adrenergic receptor as a case study, we have investigated the applicability of crystal structures and homology models to the identification of GPCR ligands through docking-based virtual screening, and have defined methods intended to improve their performances.

Our controlled in silico screenings performed at the receptor crystallized in complex with the inverse agonist carazolol yielded excellent results, with a clearly delineated prioritization of ligands over decoys. Blockers generally were preferred over agonists; however agonists were also well distinguished from decoys. Notably, this trend could be reversed by optimizing the receptor around a bound known agonist prior to the screen. Moreover, we devised a method to improve the general yields of the screen by generating an ensemble of alternative conformations of the receptor that accounts for its flexibility.

Finally, we proved the applicability of docking-based virtual screenings also to homology models endowed with different of accuracy. This last point is of utmost importance, since crystal structures are available only for a limited number of GPCRs, and extends our conclusions to the entire superfamily.
Identification of binding hot spots in protein-macrocycle interactions

Elizabeth A. Villar\(^{(1)}\), eavillar@bu.edu, 590 Commonwealth Ave, Boston MA 02215, United States; Dima Kozakov\(^{(2)}\); Adrian Whitty\(^{(1)}\); Sandor Vajda\(^{(1)(2)}\). (1) Department of Chemistry, Boston University, Boston MA 02215, United States (2) Department of Biomedical Engineering, Boston University, Boston MA 02215, United States

Macrocyclic compounds are common to larger natural products and serve as an important structural class due to their balance between pre-organization and flexibility. However, the use of macrocyclic compounds as therapeutics is mainly through molecules derived from natural products, and they are poorly explored for the development of novel drug molecules. We have investigated the use of FTMAP, our computational mapping algorithm for protein hot spot determination, in identifying macrocycle interaction hot spots. A recent expansion of the fragment probe library used by FTMAP has allowed for increased chemical specificity in our results, and here we present the application of this expanded set mapping to a variety of macrocycle binding protein targets.

COMP 289

Molecular modeling and simulation of the HER3/ErbB3 pseudo-kinase domain
The ErbB family of receptor tyrosine kinases, which includes EGFR (HER1), HER2, HER3, and HER4, is an important target of small-molecule kinase inhibitors in the treatment of human cancers. However, point mutations in cancer cells have led to the emergence of resistance to these drugs and have most recently been shown to be mediated by HER3. Efforts to inhibit HER3 have proven to be difficult because it is believed to be catalytically inactive (a 'pseudokinase') due to amino acid substitutions in the conserved kinase domain. Specifically, it lacks a catalytic base aspartate, which deters from phosphorylating protein substrates. Thus there is a growing interest in characterizing the molecular features of HER3 that, despite seeming to lack catalytic activity, enable it to mediate drug resistance. Here we construct a model of the ternary complex of HER3 kinase bound to ATP and a tyrosine-containing substrate peptide, based on a HER3 crystal structure that has recently been solved in collaborative experiments with the Mark Lemmon lab at UPenn [Shi F. et al. 2010. *PNAS*]. We compare the possible catalytic mechanisms in EGFR and HER3 through free energy calculations and mixed quantum mechanics molecular mechanics (QM/MM) simulations. Our QM/MM simulations reveal that phosphotransfer in HER3 may proceed along an alternate pathway which does not require the conserved catalytic aspartate. Based on the QM/MM activation energies, the HER3 pathway is much slower than EGFR, which correlates with the low level of autocatalytic activity observed in our collaborative experiments. Thus HER3 may represent a novel therapeutic target of small-molecule kinase inhibitors.

**COMP 290**

**Can vibrational spectra distinguish anionic from neutral radicals derived from ubiquinone cofactors in electron transfer proteins?**

*Ralph A. Wheeler*(1), rawheeler@ou.edu, Norman OK 73019, United States; *Scott E. Boesch*(1). (1) Department of Chemistry & Biochemistry, University of Oklahoma, Norman OK 73019, United States

Ubiquinones are important electron transfer (ET) cofactors and their radicals are intermediates implicated in the ET reactions of photosynthesis and respiration. Conflicting experimental results have fueled controversy regarding the identity of radical intermediates in biological ET reactions. Vibrational spectra of anionic and neutral radicals calculated by using the B3LYP method will be presented and compared to find ways to distinguish the radicals from each other experimentally.
Computational study of the effects of water molecules on sulfur oxide reactions

Jean M. Standard\(^{(1)}\), standard@ilstu.edu, Campus Box 4160, Normal IL 61790-4160, United States; Matthew C. Cafarelli\(^{(1)}\); Piotr Gorczynski\(^{(1)}\); Rachel A. Craigmile\(^{(1)}\), (1) Department of Chemistry, Illinois State University, Normal IL 61790-4160, United States

Sulfur-containing compounds such as sulfur dioxide and sulfur trioxide are important species in atmospheric chemistry. A number of previous computational studies have demonstrated the importance of catalytic water molecules in reactions between sulfur trioxide and water to produce sulfuric acid. For example, one additional catalytic water molecule reduces the activation barrier for formation of sulfuric acid from 30 to 12 kcal/mol at the MP2 level. In this work, the effects of catalytic water molecules on other reactions involving sulfur oxides have been investigated. The reactions studied include the reactions of sulfur trioxide with hydrogen halides, as well as reactions of sulfur dioxide and sulfur trioxide with hydroxyl and hydroperoxy radicals. Ab initio (MP2) methods have been employed in the studies along with basis sets up to aug-cc-pV(T+d)Z. The results obtained for structures, charge distributions, and activation energies have been compared with those obtained for the same reactions determined in the absence of catalytic water molecules. The results confirm that catalytic water molecules have a significant impact on most of the sulfur oxide reactions studied.

Eyring: A computer program for computing gas-phase bimolecular chemical reaction rates of polyatomic species

Claudia Zavala-Oseguera\(^{(1)}\), czoseguera@gmail.com, Noria Alta s/n, Guanajuato Guanajuato 36050, Mexico; Aninia Galano\(^{(2)}\); Alonso Ramirez-Manzanares\(^{(3)}\); Gabriel Merino\(^{(1)}\), (1) Departamento de Quimica, Universidad de Guanajuato, Guanajuato Guanajuato 36050, Mexico (2) Departamento de Quimica, Universidad Autónoma Metropolitana-Iztapalapa, Mexico D.F. 09340, Mexico (3) Departamento de Matemáticas, Universidad de Guanajuato, Guanajuato Guanajuato 36240, Mexico

In this word, we introduce the Eyring program, a computer application with a user friendly interface for the computation of chemical rate constant of polyatomic species from first principles. In this first version, the implemented method is based on the conventional transition state theory with two different semiclassical methods for computing the transmission coefficients; namely the one-dimensional Wigner and Eckart methods. In addition, the correct calculation of internal rotations is critical for obtaining reliable values of the rate constant; for
that reason, we include the approximation of Ayala-Schlegel that provides an analytical form for the hindered rotor partition function when the torsional barrier is comparable with $k_B T$.

COMP 293

Conformational sampling of macrocycles using Monte Carlo search techniques

Carolyn Weigelt(1), carolyn.weigelt@bms.com, P.O. Box 4000, Princeton NJ, United States; Roy Kimura(1); Andrew J Tebben(1). (1) Department of Computer Aided Drug Design, Bristol-Myers Squibb, Princeton NJ 08543-4000, United States

Macrocycles are large cyclic molecules which, because of their ring structure, are structurally pre-organized and therefore, have reduced conformational flexibility compared to acyclic molecules. As a result, macrocycles can bind to protein targets with reduced entropic loss and can therefore be highly potent and selective drug discovery targets. Although macrocycles have reduced flexibility, they are not structurally rigid. Effective conformational sampling of macrocyclic space is essential to the successful application of molecular modeling techniques. In this study, we compare different Monte Carlo techniques for the sampling macrocyclic conformational space.

COMP 294

Computational studies on interaction of antioxidant dendrimers with plasmid DNA, pBR322

Leela Rakesh(1), lrakesh@aol.com, Central Michigan University, Pearce 214, Mount Pleasant MI 48859, United States; Choon Y Lee(2), lee1cy@cmich.edu, Central Michigan University, Dow Science 268, Mount Pleasant MI 48859, United States; Pratik Chhetri(1), chhet1p@cmich.edu, Central Michigan University, Pearce 214, Mount Pleasant MI 48859, United States. (1) Department of Mathematics, Central Michigan University, Mount Pleasant MI 48859, United States (2) Department of Chemistry, Central Michigan University, Mount Pleasant MI 48859, United States

Antioxidants have been known for their anti-inflammatory effects for a long time. The purpose of this study was to study the interactions of dendrimers (derived from syringaldehyde, and its analogs derived from vanillin and iodovanillin) with plasmid DNA pBR322. Several computational methods were applied and based on the calculated energy components for the interaction between the dendrimers and the DNA, syringaldehyde based zero generation dendrimer was noticed to have the most efficient antioxidant properties. These results will further be used
to reduce additional laboratory costs for synthesizing generations of those dendrimers.

COMP 295

Designing enzymes with ROSETTA: Challenges and solutions

Seonah Kim(1), seonah@chem.ucla.edu, 607 Charles E. Young drive East, Los Angeles CA 90095, United States; Gert Kiss(1); Robert S Paton(1); Scott A Johnson(1); Geoffrey Nosrati(1); K N Houk(1). (1) Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles CA 90095, United States

De-novo enzyme design is a non-trivial process and is faced with many challenges. Characterization of the background reaction, adequate treatment of flexible substrates/transition states, and correct placement of multiple catalytic residues are essential elements and require special attention. QM theozymes are incorporated into existing stable protein scaffolds using the RosettaMatch program of Baker et al. Multiple transition state conformations must be considered at this stage. The matched geometries are compared to those of the theozyme using the in-house EDGE utility (Enzyme Design Geometry Evaluation). Residues in the vicinity of the active site are then altered and optimized with RosettaDesign, such as to maximize the packing around the ideal QM geometry. To maximize computational efficiency, we utilize the Kepler workflow environment (collaboration with developers at San Diego Super Computing Center) for massive parallelization of RosettaMatch, RosettaDesign, and EDGE. The workflow standardizes the enzyme design process through automation and eliminates the need for unnecessary human intervention.

COMP 296

QSAR modeling of GPCR receptor families, model application for virtual screening, and experimental validation of computational hits

Rima Hajjo(1), hajjo@email.unc.edu, 120 Mason Farm Road, 2069 Genetic Medicine Building, Chapel Hill NC 27599, United States; Simon Wang(1); Bryan L Roth(2); Alexander Tropsha(1). (1) Department of Medicinal Chemistry and Natural Products, University of North Carolina at Chapel Hill, Chapel Hill NC 27599, United States (2) Departments of Pharmacology, Medicinal Chemistry, and Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill NC 27514, United States

Compounds with polypharmacological mechanism of action (i.e., targeting multiple receptors) are considered as promising drug
candidates to treat complex neurological diseases. We have developed QSAR models for each of the serotonin, adrenergic, dopaminergic, muscarinic, and sigma receptor families. The models achieved external classification accuracies of 82 to 94%. These family based QSAR models are attractive because of increased size and chemical diversity of underlying datasets and broader applicability domains of the developed models. They were used for virtual screening (VS) of the World Drug Index database to identify putative ligands for each family of receptors. Twenty hits were tested in binding assays against a panel of receptor subtypes of all studied families. All compounds were found to bind to at least one receptor subtype among the predicted families with binding affinities between 0.0006 – 9.6 µM.

COMP 296

QSAR modeling of GPCR receptor families, model application for virtual screening, and experimental validation of computational hits

Rima Hajjo\(^{(1)}\), hajjo@email.unc.edu, 120 Mason Farm Road, 2069 Genetic Medicine Building, Chapel Hill NC 27599, United States; Simon Wang\(^{(1)}\); Bryan L Roth\(^{(2)}\); Alexander Tropsha\(^{(1)}\). (1) Department of Medicinal Chemistry and Natural Products, University of North Carolina at Chapel Hill, Chapel Hill NC 27599, United States (2) Departments of Pharmacology, Medicinal Chemistry, and Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill NC 27514, United States

Compounds with polypharmacological mechanism of action (i.e., targeting multiple receptors) are considered as promising drug candidates to treat complex neurological diseases. We have developed QSAR models for each of the serotonin, adrenergic, dopaminergic, muscarinic, and sigma receptor families. The models achieved external classification accuracies of 82 to 94%. These family based QSAR models are attractive because of increased size and chemical diversity of underlying datasets and broader applicability domains of the developed models. They were used for virtual screening (VS) of the World Drug Index database to identify putative ligands for each family of receptors. Twenty hits were tested in binding assays against a panel of receptor subtypes of all studied families. All compounds were found to bind to at least one receptor subtype among the predicted families with binding affinities between 0.0006 – 9.6 µM.
**COMP 297**

**Parallel virtual screening for novel antimigraine therapeutics acting on 5-HT1B, 5-HT1D and 5-HT1F serotonin receptors**

*Xiang S. Wang*(1), xswang@email.unc.edu, *Genetic Medicine Building # 2078, Chapel Hill North Carolina, United States*; *Adriano D. Andricopulo*(2); *Tudor I. Oprea*(3); *Alexander Tropsha*(1); *Tiago L. Moda*(1)(2). (1) University of North Carolina at Chapel Hill, United States (2) University of Sao Paulo, Brazil (3) University of New Mexico, United States

Migraine is a recurring, episodic neurological disorder characterized as unilateral, throbbing headache. The only class of marketed drugs to treat migraine specifically is a group of triptans, e.g., sumatriptan and naratriptan, but they are not effective in all patients. Recent findings indicate that compounds that can simultaneously target a triplet of serotonin 5-HT receptors, i.e., 5-HT1B, 5-HT1D, and 5HT1F, may be very effective and safe medications. We have collected three data sets of structurally diverse molecules with known affinity for serotonin 5-HT1B/1D/1F receptors. Multiple externally predictive QSAR models for the three receptor subtypes were obtained using k Nearest Neighbor (kNN) and Support Vector Machine (SVM) methods and validated using five-fold external validation. The models were then used to mine in parallel the Prestwick and Maybridge screening libraries totaling over 55,000 compounds. Finally 49 common hits have been identified and submitted to the Psychactive Drug Screening Program for experimental validation.

**COMP 298**

**Development of semi-empirical models for zinc metalloenzymes**

*Shihao Wang*(1), shihaow@gmail.com, *7141 Sherbrooke Street West, SP 265.25, Montreal Quebec H4B 1R6, Canada*; *Guillaume Lamoureux*(1). (1) Department of Chemistry and Biochemistry, Concordia University, Montreal Quebec H4B 1R6, Canada

Semi-empirical models based on the neglect of diatomic differential overlap (AM1, RM1, PM3) have been widely used in mechanistic studies of organic reactions. However, those models were generally parameterized from small organic molecules, and are not accurate enough for zinc-catalyzed reactions in enzymes. A new methodology is developed to parameterize the models for hydrolysis reactions in zinc metalloenzymes. The training set is prepared by high level *ab initio* data, which includes fragments of the enzyme active sites. A scoring function is defined by the most relevant properties, such as bond lengths, angles, proton affinities, reaction energies and barriers. The parameterization process contains random searching in the parameter space and minimization of the scoring function. This process is repeated for each element until the
parameters are converged. The final parameter set performs very well in the testing set.

COMP 299

Allosterism in MutS proteins during DNA mismatch recognition and repair signaling

Susan N. Pieniazek(1), spieniazek@wesleyan.edu, Hall-Atwater Laboratories-66, 237 Church Street, Middletown Connecticut 06459-0180, United States; Manju M. Hingorani(1); David L. Beveridge(1). (1) Department of Chemistry, Wesleyan University, Middletown Connecticut 06459-0180, United States

Allosteric communication in multi-domain protein architectures is crucial in complex biological processes such as DNA mismatch repair (MMR). MutS proteins initiate MMR through recognition of mismatch DNA and signaling downstream repair. [i] Mismatch recognition by MutS is followed by a marked decrease in the rate of ATP hydrolysis and DNA binding affinity. [ii] Our all-atom molecular dynamics simulations (150 ns) on ATP-bound and ATP-free MutS complexes with mismatch DNA reveal important structural information about the elusive ATP-bound complex, and the dynamical nature of the coupling between the ATPase sites and DNA-binding site ~70 Å away. Overall, ATP-binding reduces coupling by inducing subtle conformational changes and a significant reduction in long-range collective atomic fluctuations.


COMP 300

Computational investigation of the Nitrogen-Boron interaction in o-(N,N-dialkylaminomethyl)arylboronate systems

Joseph D. Larkin(1), josephdlarkin@gmail.com, Bldg. 50, Bethesda MD 20851, United States; John S. Fossey(2); Tony D. James(3); Bernard R. Brooks(1); Charles W. Bock(4)(5). (1) National Heart, Lung, and Blood Institute, The National Institutes of Health, Bethesda MD 20851, United States (2) School of Chemistry, University of Birmingham, Edgbaston Birmingham B15 2TT, United Kingdom (3) Department of Chemistry, University of Bath, Bath BA2 7AY, United Kingdom (4) Department of Chemistry and Biochemistry, School of Science and Health, Philadelphia University, Philadelphia Pennsylvania 19144, United States (5) The Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia Pennsylvania 19144, United States
o-(N,N-dialkylaminomethyl)arylboronate systems are an important class of compounds in diol-sensor development. We report results from a computational investigation of fourteen o-(N,N-dialkylaminomethyl)arylboronates using second-order Møller-Plesset (MP2) perturbation theory. Geometry optimizations were performed at the MP2/cc-pVDZ level and followed by single-point calculations at the MP2/aug-cc-pVDZ(cc-pVTZ) levels. These results are compared to those from density functional theory (DFT) at the PBE1PBE(PBE1PBE-D)/6-311++G(d,p)(aug-cc-pVDZ) levels, as well as to experiment. Results from continuum PCM and CPCM solvation models were employed to assess the effects of a bulk aqueous environment.

COMP 301

Analysis of amino acid sidechain Chi1, Chi2 conformational properties using quantum mechanical and experimental data

Xiao Zhu(1), zhu@outerbanks.umaryland.edu, 20 Penn Street, Baltimore MD 21201, United States ; Pedro Lopes(1); Jihyun Shim(1); Alexander D. MacKerell, Jr. (1). (1) Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, Baltimore MD 21201, United States

Amino acid side chain flexibility is an important property that influences the side-chain interactions in proteins. In molecular mechanics, the conformational properties of sidechains can be modulated by torsional parameters. In this study, we analyze the conformational properties of sidechains via quantum mechanical calculations. One and two-dimensional chi energy surfaces were performed on dipeptides representative of the amino acids. Analysis was performed for relevant peptide backbone conformations corresponding to the alpha helical (alpha R), beta stranded (extended) and alpha L conformations. Calculated QM energy surfaces are indicative of the conformational properties of the different amino acid sidechains and were used as target data for force field optimization as well as a basis for explaining experimental observations.

COMP 302

Structure-based fragment hopping ligand design using predocked fragment database

Fang-yu Lin(1), yjtseng@csie.ntu.edu.tw, No. 1 Roosevelt Rd. Sec. 4, Taipei 106, Taiwan Republic of China ; Y. Jane Tseng(1)(2). (1) Graduate Institute of Biomedical Electronic and Bioinformatics, National Taiwan University, Taipei, Taiwan Republic of China (2) Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan Republic of China

Our method aims to substitute fragments of the inhibitor while leaving the rest of the inhibitor as its original conformation as design template. This approach is
based on a set of fragment scores of docked inhibitors. To find the potential active compounds, known inhibitors are generated through virtual screening and further docked into the receptor active site. By applying our defined fragmentation schemes, each inhibitor is regarded as a composition of several fragments. With binding information, multi-objective scoring function is applied for each fragment. The fragment score is calculated through force-field energy interaction in the binding site and fingerprints of fragment properties. Fragment hopping can be promoted by tuning the relative weights of these terms in the scoring function. The undesirable scores of fragments from the inhibitor serve as query fragments and undergo fragment hopping with a ranked database of predocked fragments. The predocked fragments are a collection of synthesizable docked molecular building blocks. To consider the geometry conformation, the connection rule for the new upcoming fragment is applied during fragment hopping. Finally, new compounds are energy minimized and confirmed with docking method. To validate our approach, a case study with the Human 5-LOX enzyme active site is investigated and its interaction with the suggested new inhibitors is analyzed.

COMP 302

Structure-based fragment hopping ligand design using predocked fragment database

Fang-yu Lin(1), yjtseng@csie.ntu.edu.tw, No. 1 Roosevelt Rd. Sec. 4, Taipei 106, Taiwan Republic of China ; Y. Jane Tseng(1)(2). (1) Graduate Institute of Biomedical Electronic and Bioinformatics, National Taiwan University, Taipei, Taiwan Republic of China (2) Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan Republic of China

Our method aims to substitute fragments of the inhibitor while leaving the rest of the inhibitor as its original conformation as design template. This approach is based on a set of fragment scores of docked inhibitors. To find the potential active compounds, known inhibitors are generated through virtual screening and further docked into the receptor active site. By applying our defined fragmentation schemes, each inhibitor is regarded as a composition of several fragments. With binding information, multi-objective scoring function is applied for each fragment. The fragment score is calculated through force-field energy interaction in the binding site and fingerprints of fragment properties. Fragment hopping can be promoted by tuning the relative weights of these terms in the scoring function. The undesirable scores of fragments from the inhibitor serve as query fragments and undergo fragment hopping with a ranked database of predocked fragments. The predocked fragments are a collection of synthesizable docked molecular building blocks. To consider the geometry conformation, the connection rule for the new upcoming fragment is applied during fragment hopping. Finally, new compounds are energy minimized and confirmed with docking method. To validate our approach, a case study with the Human 5-LOX enzyme active site is investigated and its interaction with the suggested new inhibitors is analyzed.
enzyme active site is investigated and its interaction with the suggested new inhibitors is analyzed.

**COMP 303**

**Fragment-based de-novo design for VEGFR2/3 inhibitors**

*Yi-Syuan Huang*(1), rabbita1101@gmail.com, No. 1 Roosevelt Rd. Sec. 4, Taipei 106, Taiwan Republic of China ; *Y. Jane Tseng*(2), yjtseng@csie.ntu.edu.tw, No. 1 Roosevelt Rd. Sec. 4, Taipei 106, Taiwan Republic of China ; Bo-Han Su*(1). (1) Department of Computer Science and Information Engineering, National Taiwan University, Taipei 106, Taiwan Republic of China (2) Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei 106, Taiwan Republic of China

Vascular endothelial growth factor receptor 2 and 3 (VEGFR-2/3) are members of the protein kinase families and play a pivotal role in angiogenesis and lymphangiogenesis therefore important for cancer treatments. Most kinase inhibitors, “Type I kinase inhibitors”, are ATP competitive. “Type II kinase inhibitors” target the inactive form of kinases which force the DFG domain to stretch out and provide an additional hydrophobic allosteric site. In this study, we established a fragment-based de-novo design strategy to design novel type II VEGFR-2/3 inhibitors. Here, based on the binding mode of sorafenib (Nexavar), we allocate the binding pocket of sorafenib into three functional boxes in 3D into allosteric site, hydrophobic spacer, and ATP binding site. Three steps were employed to design the new VEGFR-2/3 inhibitors. First, fragments from an in house library were docked into the three boxes respectively. Each fragments were evaluated with the efficiency score, GE (group efficiency, $GE = \frac{G}{\text{Heteroatom numbers}}$). The fragments with higher GE were selective with priority. Finally, a series of new structures were generated by adding linker fragments located in the spacer box. Also, we adopted the idea similar to SIFt(Structural interaction fingerprint) strategies to validate the binding mode of these new generating compounds. In our experiment, known VEGFR 2/3 inhibitors, for example, sorafenib can be successfully re-generated and therefore, further validated our approach.

**COMP 304**

**4D-Fingerprints clustering based scaffold hopping**

*Yi-Shu Tu*(1), yjtseng@csie.ntu.edu.tw, No. 1 Roosevelt Rd. Sec. 4, Taipei 106, Taiwan Republic of China ; Bo-Han Su*(2); Y. Jane Tseng*(1)(2). (1) Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei 106, Taiwan Republic of China (2) Department of Computer Science and Information Engineering, National Taiwan University, Taipei 106, Taiwan Republic of China
Most ligand-based scaffold hopping approaches use one or few active compounds as query. However, many active compounds present more than one binding mode. With ligand-based classification on active compounds, more novel active compounds might be able to be discovered with classifying different binding mode. We applied 4D-Fingerprints (4DFP) to scaffold hopping includes more possible conformation in additional to the general 3D-based scaffold hopping. In this work, we applied an additional layer of clustering to our algorithm. Our first step includes a hierarchical clustering of 4DFPs of active compounds and random-selected inactive representatives. The clusters with most active compounds were identified as the "Active Clusters". The 4DFP similarity between the screening compound and known active compounds were calculated to locate the cluster for the screening compound compared with known active compounds. In the hsp90 dataset from MUV, despite the diversity in 2D topologies, the active compounds and inactive compounds were separated with respect to their 4DFPs. Each cluster is considered with different binding mode. By applying 4DFP clustering to scaffold hopping, more novel active compounds with different binding mode might be recovered by comparing with different active clusters.

COMP 305

Simulation-informed predictions of interfacial phenomena: A generalized model for the electric double layer

Alberto Striolo\(^{(1)}\). (1) School of Chemical, Biological and Materials Engineering, The University of Oklahoma, Norman OK 73019, United States

Evidence suggests that the structure of water within \(~1\text{-}2\) nm from a surface dictates a number of properties ranging from molecular-scale (e.g., solute adsorption) to macroscopic phenomena (e.g., drag reduction). We will present evidence for these observations taken both from literature and from our own work. The results will be used to demonstrate that the local properties of water dictate the adsorption of simple electrolytes, the adsorption of surfactants, and also drag reduction.

We will then discuss our attempts to relate the structure of water at solid-liquid interfaces to macroscopic phenomena such as the adsorption of simple electrolytes and that of surfactants. Our work will yield a generalized model to describe the electric double layer and to explain salt-specific Hofmeister effects [F. Hofmeister, Zur lehre von der wirkung der salze, Arch. Exp. Path. Pharm. 25 (1888) 247]. The successful completion of our project is based on the detailed simulation of pure water at interfaces, and the development of accurate models to relate the local properties of interfacial water (e.g., local density, local mobility, and 'local' dielectric constant), to the macroscopic properties of interest. This approach will be extremely beneficial for, e.g.: (1) Quantitatively predict how specific ions partition within pores found in rocks and/or porous membranes, as
required for designing backfiller materials to prevent environmental contaminations due to radionuclide leaks; (2) Quantitatively understand how a solid/nanoparticle surface needs to be modified to control reactant adsorption, and consequently catalytic conversions; (3) Quantitatively predict the structure of surfactant aggregates on, e.g. carbon nanotubes to optimize ultra-centrifugation techniques to prepare carbon nanotube samples mono-dispersed in diameter and chirality.

COMP 306

Pose accuracy using DOCK: Database construction and protocol evaluation

Sudipto Mukherjee\(^{(1)}\), sudipto.mukherjee@gmail.com, Stony Brook, Stony Brook NY 11794-3600, United States; Trent E Balius\(^{(1)}\); Robert C Rizzo\(^{(1)}\). (1) Department of Applied Mathematics & Statistics, Stony Brook University, Stony Brook NY 11794-3600, United States

Docking studies using a newly constructed database (N=780) from the protein data bank will be presented using a modified version of the DOCK 6 program. Partitioning of the database into different protein families (N=25) and ligand flexibility (less than 7, 8 to 15, and 15+ rotatable bonds) facilitated isolation of particularly well-behaved systems (i.e. neuraminidase, HIV reverse transcriptase) and those for which improvements to docking would be desirable (HIV protease, carbonic anhydrase). The database also allowed for comprehensive evaluation of the effectiveness of particular docking protocols (rigid, fixed anchor, flexible ligand docking) and parameters (partial atomic charge models, grid spacing, vdw repulsive exponents). A new tool for tracking ligand sampling during growth illuminates the behavior of the core DOCK anchor-and-grow algorithm and helps identify pathological cases of incorrect sampling, facilitating code debugging and optimization of input parameters. Results from cross-docking for protein family members aligned to a common reference frame will also be discussed.

COMP 307

WITHDRAWN

COMP 307

WITHDRAWN

COMP 308

Conceptual QSAR models for cell-based assays combining QM/MM linear response approach and disposition function
Cell-based assays are used in primary screening in drug development to establish drug efficacy, specificity as well as toxicity. The measured biological activity depends not only on the binding affinity of a ligand to the receptor but also on the ability of the ligand to be transported through the membranes and reach the receptor site. Ligand transport, elimination, and binding to non-receptor cell constituents, affecting the concentration in the receptor surroundings, can be accounted for by the disposition function (DF). DF describes the relationship between the dose and the ligand concentration in the receptor surroundings as a model-based, nonlinear function of ligand's lipophilicity, acidity, and other properties. The present approach combines the DF with QM/MM linear response method for binding affinity prediction. The resulting model helps us to identify structural features essential for good disposition and specific binding and also enable to weed out false positives (good binders but poor disposition or vice versa) in the early stage of drug discovery, saving immense resources.

COMP 309

Charge transfer excited states for large donor-acceptor molecular systems

Tunna Baruah\(^{(1)}\), tbaruah@utep.edu, 500 W. University Avenue, El Paso TX 79968, United States; Marco Olguin\(^{(1)}\); Rajendra R Zope\(^{(1)}\). (1) Physics, University of Texas at El Paso, El Paso TX 79968, United States

The charge-transfer (CT) excitations play an important role in many chemical interactions, such as metal surface-adsorbate catalysis or light-harvesting processes. The widely used time-dependent density functional formalism shows limitations in describing CT excited-states. We present a constrained excited-state method to estimate the energies of singly excited particle-hole states. The applied constraint in this approach is the orthogonality between the ground and excited-states. The expense involved in the calculation of each excited state is comparable to that for the ground state. A comparison of the gas-phase CT excitation energies for a set of aromatic donor with TCNE acceptor shows excellent agreement with experimental values. Its application to the CT excited states for a light harvesting C60-porphyrin-beta-carotene triad demonstrates the applicability of this approach to large systems. Performance appraisal of this method for a large number of excited states including core excitation will be presented and further improvements will be discussed.

COMP 310
WITHDRAWN

COMP 311

Quantitative conformational sampling of glucokinase using second-order orthogonal space random walk

*William Harris*\(^{(1)}\); *Wei Yang*\(^{(1)}\). \((1)\) Department of Chemistry and Biochemistry, Florida State University, Tallahassee Florida 32308, United States

Long time-scale biomolecular events are challenging targets in computational biophysics. To overcome such a bottleneck issue, recently, a novel second-order generalized ensemble free energy simulation method, the orthogonal space random walk (OSRW) algorithm, was developed. This technique allows explicit free energy barriers and the associated “hidden” free energy barriers to be synchronously crossed. In this work, we employed the OSRW algorithm to quantitatively understand a large conformational change of Glucokinase, a pivotal enzyme in carbohydrate metabolism and regulation. In this paper,

<br>

<!--
@page { margin: 0.79in }
P { margin-bottom: 0.08in }
-the role of
the substrate glucose as a conformational selector is clearly revealed in these simulations; more interestingly, an unexpected structural relocation mechanism of the 23-residue helix is also revealed.

COMP 312

Interplay of AAA+ molecular machines, DNA repair enzymes and sliding clamps at the replication fork: A multiscale approach to modeling replisome assembly and function

*Ivaylo Ivanov*\(^{(1)}\), iivanov@gsu.edu, P.O. Box 4098, Atlanta Georgia 30302-4098, United States; *J. Andrew McCammon*\(^{(2)}\); *John A Tainer*\(^{(3)}\). \((1)\) Department of Chemistry, Georgia State University, Atlanta Georgia 30302-4098, United States \((2)\) Department of Chemistry & Biochemistry, University of California – San Diego, United States \((3)\) Department of Molecular Biology, The Scripps Research Institute, United States

<br>
PCNA (sliding clamp) is a toroidal-shaped protein that encircles DNA and plays a pivotal role in replication. During replication sliding clamps have to be opened and resealed at primer-template junctions by a clamp loader ATPase - replication factor C (RFC). To examine this process we employed a combination of steered molecular dynamics and the novel adaptive biasing force (ABF) method. We show that upon clamp opening the RFC/PCNA complex undergoes a large conformational rearrangement, leading to the formation of an extended interface between the clamp and RFC. Binding of ring-open PCNA to all five RFC subunits transformed the free energy landscape underlying the closed- to open state transition, trapping PCNA in an open conformation. Comparison of free energy profiles for clamp opening in the presence and absence of RFC allowed us to substantiate the role of RFC in the initial stage of the clamp-loading cycle.

COMP 313

High-quality ligand-steered modeling and docking evaluation of class A G protein-coupled receptors

Sharangdhar S Phatak(1), Sharangdhar.S.Phatak@uth.tmc.edu, 7000 Fannin Ste. 860B, Houston TX 77030-5400, United States; Claudio N Cavasotto(1), claudio.n.cavasotto@uth.tmc.edu, 7000 Fannin, Ste. 860B, Houston TX 77030, United States. (1) School of Health Information Sciences, University of Texas Health Science Center at Houston, Houston TX 77030-5400, United States

Using three GPCR structures as templates (β2-adrenergic, A2A-adenosine, and rhodopsin receptors), we enhanced and validated the ligand-steered modeling method through cross-modeling of, and high-throughput docking on four target/template cases. We raised and answered three questions:

1. Can the ligand-steered modeling generate a small set of near-native models by using one structural template? In all cases, the ligand-steered modeling method identified a small-set of models with at least one near-native model.

2. How well do the ligand-steered models perform in docking? The enrichment factor obtained by the best ligand-steered models was superior to those obtained by the crude models and random selection in all cases.

3. How well do the ligand-steered models perform in assessing cross-selectivity between β2 and A2A ligands? In all cases, the ligand-steered models either identified a higher percentage of target specific actives, and/or discarded higher percentage of “other” actives, as compared to the crude models.

COMP 314

Multiple land simultaneous docking (MLSD): Orchestrated dancing of ligands in binding sites of protein
Present docking methodologies simulate only one single ligand at a time during docking process. In reality, the molecular recognition process always involves multiple molecular species. Typical protein-ligand interactions are, for example, substrate and cofactor in catalytic cycle; metal ion coordination together with ligand(s); and ligand binding with water molecules. In order to simulate the real molecular binding processes, we propose a novel multiple ligand simultaneous docking (MLSD) strategy which can deal with all the above processes, vastly improving docking sampling and binding free energy scoring. The work also compares two search strategies: Lamarckian Genetic Algorithm (LGA) and Particle Swarm Optimization (PSO), which have respective advantages depending on the specific systems. The methodology proves robust through systematic testing against several diverse model systems: E. coli PNP complex with two substrates (figure 1), SHP2NSH2 complex with two peptides (figure 2) and Bcl-xL complex with ABT-737 fragments (figure 3). In all cases, the final correct docking poses and relative binding free energies were obtained. In PNP case, the simulations also capture the binding intermediates and reveal the binding dynamics during the recognition processes, which is consistent with the proposed enzymatic mechanism. In the other two cases, conventional single ligand docking fails due to energetic and dynamic coupling among ligands, whereas MLSD results in the correct binding modes. These three cases also represent potential applications in the three areas: exploring enzymatic mechanism; interpreting noisy X-ray crystallographic map; and aiding fragment-based drug design; respectively.

COMP 315

Ligand effects on geometric and electronic structure of the Au_{25}(SR)_{18}^{-} nanoparticle

Christine M. Aikens(1), cmaikens@ksu.edu, 213 CBC Building, Manhattan Kansas 66506, United States. (1) Department of Chemistry, Kansas State University, Manhattan Kansas 66506, United States

Density functional theory (DFT) and time-dependent DFT (TDDFT) have recently been employed to examine the geometric and electronic structure, optical properties, chiroptical properties, and other physical properties of gold and silver nanoparticles including Au_{25}(SR)_{18}^{-} (R = H, CH_{3}, CH_{2}CH_{3}, CH_{2}CH_{2}Ph) and Au_{11}(P_{2}C_{4}H_{8})_{4}Cl_{2}^{+}. In many cases, DFT provides a qualitative or even quantitative agreement with experiment.

In this presentation, the electronic structure of the Au_{25}(SPh)_{18}^{-} nanoparticle is
contrasted with the recently established structure of Au$_{25}$(SH)$_{18}^-$. The geometry of the nanoparticle is calculated at the Xa/DZ level of theory. Various orientations of the phenyl ligands have been treated in this work, and the lowest energy structure is found to have significant pi-stacking. This structure differs from the ligand orientation in the crystal structure of Au$_{25}$(SCH$_2$CH$_2$Ph)$_{18}^-$. The optical absorption spectra for Au$_{25}$(SPh)$_{18}^-$ and para-substituted Au$_{25}$(SPhX)$_{18}^-$ nanoparticles (X = F, Cl, Br, CH$_3$, OCH$_3$) differ substantially from Au$_{25}$(SH)$_{18}^-$ due to splitting of the superatom orbitals.

COMP 316

Flipping process of 8oxoG in Fpg studied by mutagenesis in silico and umbrella sampling simulation

Fu Lin$^{(1)}$, linfu3200@gmail.com, CMM building G90, Stony Brook New York, United States; Arthur John Campbell$^{(1)}$; Christina Bergonzo$^{(1)}$; Author Grollman$^{(1)}$; Carlos Simmerling$^{(1)}$. (1) Chemistry, State University of New York at Stony Brook, Stony Brook New York 11794, United States

Damaged DNA, resulted from the attack of reactive species on the genome spontaneously, are an essential source of carcinogenic mutations. Formamidopyrimidine-DNA glycosylase (Fpg) is a primary participant in the repair of 8-oxoguanine (8oxoG), an abundant oxidative DNA lesion. To structurally and kinetically characterize the eversion of 8-oxoG from intrahelical site to the active site via exo-site, disulfide crosslinking (DXL) and crystallographic techniques have been widely introduced to this area. Even though some unprecedented clue to the flipping process of 8oxoG hinted by novel crystal structures obtained with these techniques was disclosed, however, it also may bring some artifacts such as structure distortion caused by DXL, which are implicit in these experimental techniques. Therefore nudged elastic band (NEB), mutagensis in Silico and umbrella sampling simulation technique were adopted in our study. From our preliminary study of the mutation of important amino acids in the Fpg, it indicates that these residues may be involved in this 8oxoG flipping process. We hope these simulation techniques could help us unveil some new detail about this flipping process and could also provide some guidance of the design of our experiment.

COMP 317

Protein folding/unfolding under force

Julio M Fernandez$^{(1)}$, jfernandez@columbia.edu, 1011A Fairchild Center, 1212 Amsterdam Avenue, New York NY 10027, United States; Ronen Berkovich$^{(2)}$; Tzu-Ling Kuo$^{(3)}$; Sergi Garcia-Manyes$^{(1)}$; Jingyuan Li$^{(4)}$; Itay Barel$^{(2)}$; Hui Lu$^{(5)}$; Bruce J Berne$^{(4)}$; Michael Urbakh$^{(2)}$; Joseph Klafter$^{(2)}$. (1) Department of
Force-clamp spectroscopy has opened up new approaches to study protein folding/unfolding at the single molecule level. Under constant force single ubiquitin proteins show non-exponential unfolding kinetics indicating the presence of static disorder in the activation energy barrier for unfolding. The measured variance in $\Delta E$ shows both force dependent and independent components, where the force dependent component scales with $F^2$, in excellent agreement with static disorder theory. Extended proteins folding after a quench in the pulling force follow one of a multitude of possible trajectories, each reflecting a distinct conformational ensemble. We model the free energy of the protein as a combination of an entropic elasticity term together with a short range potential representing enthalpic hydrophobic interactions. By solving the Langevin equation under conditions of a force-quench we generate folding trajectories that reproduce, uncannily, the collapse trajectories observed experimentally. Our studies provide novel benchmarks to uncover the physics of protein dynamics.

COMP 318

Multiscale experimental and computational studies of cadherin-mediated cell adhesion: From molecule to cell

Barry Honig(1), bh6@columbia.edu, 1130 St. Nicholas Ave, New York New York 10032, United States . (1) Center for Computational Biology and Bioinformatics and Department of Biochemistry and Molecular Biophysics, Columbia University and Howard Hughes Medical Institute, New York New York 10032, United States

A variety of cell-cell adhesion processes are mediated by the (trans) dimerization of cadherin proteins presented from apposing cell surfaces. A multiscale combined experimental/theoretical approach to this general problem will be described with particular focus on the development of a correlation between measured binding affinities at the molecular level and cell-cell adhesive specificity. Following cell-cell contact, cadherin dimers cluster together to form cell-cell junctions. Since such clustering does not occur in the absence of cell-cell contact, there is a necessary coupling between trans and lateral (cis) interactions. The molecular mechanism that underlies this coupling will be discussed and is characterized in detail through combination of x-ray crystallography and coarse grained simulations. In addition to the insights that will be discussed in the context of cadherin function, general principles that may underlie the molecular basis of many cell-cell adhesion processes will be proposed.
**COMP 319**

Link between folding landscapes of riboswitches and control of gene expression

*Dave Thirumalai*(1), thirum@umd.edu, College of Chemical & Life Sciences, College Park MD 20742, United States. (1) Department of Chemistry, University of Maryland, College Park MD 20742, United States

I will describe how quantitative descriptions of the folding landscapes of riboswitches can be used to understand their role in regulating gene expressions.

**COMP 320**

Urea's action on hydrophobic interactions in both physical and biological systems

*Ruhong Zhou*(1), ruhongz@us.ibm.com, 1101 Kitchawan Road, Yorktown Heights NY 10598, United States. (1) Computational Biology Center, IBM Thomas J Watson Research Center, Yorktown Heights NY 10598, United States

Urea has been widely used as a chemical denaturant for proteins for more than a century; however, its molecular mechanism is still not well understood. We use extensive microseconds molecular dynamics simulations of several proteins (such as lysozyme, gamma-D crystallin, and CI2), as well as model systems (such as polymers and carbon nanotubes), in 8 M urea to study the protein chemical denaturing mechanism. We observe a 2-stage penetration of proteins in general, with urea penetrating the hydrophobic core before water, forming a "dry globule." The direct dispersion interaction between urea and protein backbones and side chains is stronger than for water, which gives rise to the intrusion of urea into the protein interior and to urea's preferential binding to all regions of the protein. Meanwhile, we find little change in the structure of water on addition of 8 M urea except that a very small fraction of water engaged in two simultaneous H-bonds with the same urea molecule slow down the reorientational dynamics. The simulations with the model polymer systems also show that the preferential binding and the consequent weakened hydrophobic interactions are driven by enthalpy and are related to the difference in the strength of the attractive dispersion interactions of urea and water with the polymer chain. Our study supports the "direct interaction mechanism" whereby urea has a stronger dispersion interaction with protein than water.

**COMP 321**

Simulations of protein aggregation in the cellular milieu
A number of diseases, known as amyloid diseases, are associated with pathological protein folding. Incorrectly or partially folded peptides or proteins can self-assemble into a variety of neurotoxic aggregate species, ranging from small soluble oligomers to amyloid fibrils. I will introduce a novel off-lattice coarse-grained peptide model that can be used to simulate the aggregation process from monomers to fibrils. The effects of beta-sheet propensity and of surfaces on the morphology of the aggregates will be discussed.

COMP 322

Exploring the earliest stages of the amyloid β protein aggregation pathway

John E. Straub(1), straub@bu.edu, 590 Commonwealth Avenue, Boston University, Boston MA 02215, United States. (1) Department of Chemistry, Boston University, Boston MA 02215, United States

The aggregation Amyloid β (Aβ) peptide has been linked to the neurodegenerative Alzheimer's Disease and implicated in other amyloid diseases including cerebral amyloid angiopathy. An essential step in the generation of Aβ peptide is the cleavage of the amyloid precursor protein (APP) by a transmembrane protease. Efforts to determine the structures of β-amyloid peptides and APP in a membrane environment are essential to the aim of providing a molecular basis for the cleavage mechanism. We report the structures of amyloid β peptide isoforms as well as a model of APP in a membrane environment determined by replica-exchange molecular dynamics simulation. The simulated structures provide insights into APP structure and the cleavage mechanism that are essential to a complete understanding of the Aβ peptide aggregation pathways.

COMP 323

Reconstructing the energy landscape of single molecules and complexes from force measurements

Gaurav Arya(1), garya@ucsd.edu, 9500 Gilman Drive., MC 0448, La Jolla CA 92093-0448, United States. (1) Department of Nanoengineering, University of California, San Diego, La Jolla CA 92131, United States
I will describe in this talk our group’s recent efforts in developing theoretical/computational models to reconstruct the free energy landscape of single molecules being stretched through pulling devices such as optical tweezers and AFM cantilevers. Specifically, I will describe the development of theoretical models to properly capture the effects of the pulling device stiffness and connecting linkers in extracting the height and location of free energy barriers and the associated rates from constant-force and constant-velocity measurements. I will end the talk by discussing new challenges for multi-scale modeling in the field of single-molecule force spectroscopy.


COMP 324

Protein-ligand binding free energy calculations for pharmaceutical targets with the binding energy distribution analysis method (BEDAM)

LAPYLOSA MAURO(1), maurolap@eden.rutgers.edu, 605, Taylor Rd, PISCATAWAY NJ 08854, United States; EMILIO GALICCHIO(1); RONALD M LEVY(1). (1) BIOMAPS, RUTGERS UNIVERSITY, PISCATAWAY NJ 08854, United States

The prediction of ligand binding free energies is helpful for the development of novel drugs. We present a formulation for the absolute free energy for protein-ligand binding with implicit solvation. An efficient computational method called binding energy distribution analysis method (BEDAM) is developed based on the calculation, by advanced molecular dynamics sampling with implicit solvation (AGBNP2) and reweighting techniques, of the distribution of binding energies between the receptor and the ligand. We present applications of the BEDAM method to a series of pharmaceutical targets. We showed that BEDAM accurately predicts trends in the experimental data thanks in part to the extensive conformational sampling achieved in the simulations. The tests we performed indicate that BEDAM is a promising new tool for in-silico drug design.

COMP 325

New algorithms for enhanced sampling and applications to protein folding in explicit solvent

Jianpeng Ma(1), jpma@bcm.tmc.edu, One Baylor Plaza, BCM-125, Houston TX 77030, United States. (1) Biochemistry and Molecular Biology, Baylor College of Medicine, Houston TX 77030, United States

We present enhanced sampling methods for simulating complex molecular systems. In a generalized ensemble, several approaches are developed to maximize the sampling efficiency and accuracy for systems
such as proteins in explicit solvent. The method addresses questions in molecular dynamics simulations such as slow protein motion due to interactions between protein and surrounding water molecules. Results are shown in terms of accelerating protein motion and converging thermodynamic quantities.

COMP 326

High-order generalized ensemble theory to achieve long timescale sampling of complex biomolecules

Wei Yang(1). (1) Chemistry and Biochemistry, Florida State University, Tallahassee FL 32306, United States

Starting with our second-order generalized ensemble development, the orthogonal space random walk algorithm, we formulate a general high-order generalized ensemble theory, which allows us to achieve very efficient sampling of long timescale biomolecular events. The essential idea of this theory is to synchronously accelerate the order parameter move, strongly coupled protein motions (motion 1), further protein motions that strongly couple with motion 1, and so on. The theory will be discussed and several challenging applications will be presented.

COMP 327

Target-blind conformational sampling in all-atom protein simulations with temperature-accelerated molecular dynamics

Cameron F. Abrams(1); Harish Vashisth(1). (1) Department of Chemical and Biological Engineering, Drexel University, Philadelphia PA 19130, United States

We describe recent advances and applications of the method of temperature-accelerated molecular dynamics (TAMD; CFA and E. Vanden-Eijnden, PNAS 2010; 107:4961) to enhance the rates of large-scale conformational sampling in all-atom, explicit-solvent protein simulations. TAMD is a collective-variable-based method in which an MD system is augmented with slow variables, each of which is tethered to one predefined collective variable and allowed to evolve slowly but hyperthermally under diffusive dynamics such that the forces they experience approximate negative gradients in the physical-temperature free energy surface associated with those collective variables. We show in this talk how TAMD predicts a novel mechanism for executing the so-called “DFG-flip” in the activation of insulin-receptor tyrosine kinase involving a transient two-turn alpha-helix. We also discuss how to sample the conformational repertoire of fluorophore labels for predicting fluorescence anisotropy decays, with particular application toward developing a conformational assay for HIV-1 gp120.
Q-dynamics: A transition state theory based strategy for modeling rare event dynamics

Chun-Yaung Lu(1), nordicalu@mail.utexas.edu, 1 University Station A5300, Austin Texas 78712-0165, United States; Graeme Henkelman(1). (1) Department of Chemistry and Biochemistry, The University of Texas at Austin, Austin Texas 78712-0165, United States

Simulation of rare events has long been considered a challenging task from either the theoretical or computational point of view. For decades, many methods have been developed to calculate reaction rates. Among them the transition state theory (TST) method, which combines the concept of kinetics and statistical mechanics, is probably the most long-lasting and widely accepted one. It is well-understood now that the exact rate can be obtained from the TST rate together with a dynamical correction factor. Here, we propose a new strategy (Q-dynamics) using both quantities of a close-shaped dividing surface to calculate the escape rate. The reaction path is determined on the fly and need not to be known a priori. The method is highly parallelizable due to the thermodynamic nature of TST. Long time dynamics trajectories are achieved by iterating the simulations. Q-dynamics are shown to be accurate in the study of surface diffusion on the Al (100) surface. Several other examples of atomic scale dynamics are also presented in this work.

Computer-guided discovery of autotaxin inhibitors

Abby L. Parrill(1), aparrill@memphis.edu, 3744 Walker Avenue, Memphis TN 38152, United States; E. Jeffrey North(1); Adrienne Hoeglund(1); Heidi Bostic(2); Michael D Best(2); Daniel L Baker(1). (1) Department of Chemistry, The University of Memphis, Memphis TN 38152, United States (2) Department of Chemistry, The University of Tennessee, Knoxville TN 37996, United States

Autotaxin (ATX) is a secreted glycoprotein with lysophospholipase D (LPLD) activity that generates the bioactive lipid lysophosphatidic acid (LPA) from lysophosphatidylcholine (LPC). Both ATX and LPA have been linked to the promotion and progression of cancer, as well as cardiovascular disease, neuropathic pain, and obesity. Despite the fact that ATX inhibitors have the potential to be useful chemotherapeutics for multiple indications, few examples of potent, drug-like ATX inhibitors are described in the current literature. Structure-based screening against a comparative model produced the first published non-lipid ATX inhibitors. These allowed the development of a binary QSAR model to rapidly prioritize additional candidate inhibitor selection. ATX inhibitors identified using these early computational tools fueled database mining using
pharmacophore models. These methods together have successfully contributed to the identification of small, drug-like, and structurally diverse ATX inhibitors with sub-micromolar $K_i$ values.

COMP 330

Probing hydrophilic domains and interdomain bridges in selected morphological models of hydrated Nafion using large-scale molecular dynamics simulations

Craig K Knox$^{(1)(2)}$, craig.knox@us.army.mil, RDRL-WMM-G, Bldg. 4600, Rm. C237, 4600 Deer Creek Loop, Aberdeen Proving Ground MD 21005-5069, United States; Gregory A Voth$^{(1)}$. (1) Department of Chemistry and Center for Biophysical Modeling and Simulation, University of Utah, Salt Lake City UT 84112, United States (2) Army Research Laboratory, Salt Lake City UT 84112, United States

Atomistic molecular dynamics simulations were performed to study hydrated Nafion systems large enough (~ 2 million atoms, ~ 30 nm box length) to directly observe several hydrophilic domains at the molecular level. These systems consisted of six of the most significant and relevant morphological models of Nafion to-date: (1) the cluster-channel model, (2) the parallel cylinder model, (3) the local-order model, (4) the lamellar model, (5) the rod network model, and (6) a 'random' model that does not directly assume any particular geometry, distribution, or morphology. Each system was initially built to closely approximate the proposed hydrophilic cluster structure in a given model. Molecular dynamics simulations were then used to observe resulting changes from and behavior of the assumed initial configurations. These simulations revealed fast intercluster 'bridge' formation and network percolation in all models. Sulfonate groups were found inside these bridges and played a significant role in percolation. Sulfonates also strongly aggregated around and inside clusters. Cluster surfaces were analyzed to study the hydrophilic-hydrophobic interface. Interfacial area and cluster volume significantly increased in all models during the simulations, suggesting the need for model refinement to account for these observations. Radial distribution functions and structure factors were also calculated. All nonrandom models clearly exhibited the characteristic experimental scattering peak, underscoring the insensitivity of this measurement to hydrophilic domain structure and highlighting the need for future work to clearly distinguish morphological models of Nafion.

COMP 331

Computational design of organometallic photochromic systems with chelating, bifunctional tethered side-chains
Studies of photochromic systems based on reversible photo-substitution reactions have shown promise for use in molecular devices. Complexes of the type CpMn(CO)$_2$L (where L is a chelating, bifunctional tethered side-chain) are being designed as bistable, photochromic systems. Our group has recently reported experimental and computational results in support of a linkage isomerization of ($\eta^1$:$\eta^5$-C$_5$H$_4$R)Mn(CO)$_2$, where R = CH$_2$C(O)pyridyl. We have shown that the more stable pyridyl-coordinated isomer can be converted to the keto-coordinated isomer by irradiation with visible light. The higher-energy keto-coordinated isomer thermally isomerizes to the pyridyl-coordinated isomer via a three-step non-dissociative mechanism involving $\pi$-bound intermediates. We have utilized density functional theory calculations during the design of new systems of this type. Computationally-derived stationary points on the potential energy surfaces are presented for thio/pyridyl, imine/pyridyl, and related systems. Time-dependent density functional theory is employed in the computation of the UV-Visible spectra of proposed synthetic targets.

COMP 332

Modeling detailed chemical kinetics of JP-10 combustion using automated reaction mechanism generation

Gregory R. Magoon\(^{(1)}\), gmagoon@mit.edu, 77 Massachusetts Avenue, Room 66-264, Cambridge MA 02139, United States; William H. Green\(^{(1)}\); Oluwayemisi O. Oluwole\(^{(2)}\); Hsi-Wu Wong\(^{(2)}\); Simon E. Albo\(^{(2)}\); David K. Lewis\(^{(2)}\). (1) Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge MA 02139, United States (2) Aerodyne Research, Inc., Billerica MA 01821, United States

JP-10 is a high-energy-density jet fuel composed of a polycyclic hydrocarbon. We have used RMG (Reaction Mechanism Generator), an open-source reaction mechanism generation program, to develop the first highly-detailed chemical kinetic model for JP-10 combustion. Our study incorporates recent refinements to RMG allowing on-the-fly calculation of thermodynamic quantities using semi-empirical methods, eliminating the need for additive ring corrections for cyclic species. Previous manually-constructed JP-10 combustion models were limited to a few hundred reactions. Using RMG, we were able to generate a highly-detailed mechanism with over 300 species and over 7,000 reactions. During the course of mechanism generation, the program considered over 20,000 species and over 1 million reactions for inclusion in the final model. The RMG-generated model accurately reproduces experimental ignition delay measurements.
Ongoing efforts toward mechanism refinement, including the use of *ab initio* quantum chemistry calculations to improve the accuracy of important parameters, will also be described.

COMP 333

**Computational (re)design and engineering of lipases: Toward a "Diels-Alderase"**

*Mats Linder*(1), linder@physchem.kth.se, Teknikringen 30/36, Stockholm Stockholm 10044, Sweden; Tore Brinck*(1). (1) Department of Physical Chemistry, Royal Institute of Technology, Stockholm, Sweden

The objective of this project is to find and optimize an enzyme that efficiently catalyzes the Diels-Alder reaction. This versatile [2+4]-cycloaddition has received much interest over the years, and emerging understanding of enzyme promiscuity provides a promising outlook for efficient, green catalysis.

By exploiting enzymatic promiscuity, numerous researchers have performed new, synthetically useful reactions, using wild type and/or mutated variants.¹ More than any other enzyme family, hydrolases (EC 3.X.X.X) have been utilized, and especially lipases (EC 3.1.1.3).² Apart from using the reversible ester hydrolysis for kinetic resolution, the catalytic Ser-His-Asp/Glu triad, together with the oxyanion hole (OxHole), has been used to manage e.g. aldol addition, epoxidation and Michael additions in lipases.³ A common denominator of these promiscuous reactions is that they use at least two elements of the catalytic machinery; for example, while the nucleophilic Ser is often suppressed, the basic His and the OxHole are often central in the new mechanism.

Lipases have been the focus of this study, due to their impressive record of catalytic promiscuity in *vitro*. The hypothetic mechanism of the Diels-Alder reaction in a hydrolase only involves the OxHole, which is supposed to stabilize the transition state (TS). By combined quantum chemical and molecular dynamics studies on structures available in the Protein Data Bank (PDB), we predict the catalytic effect to be favorable but modest for some engineered variants, when only OxHole participates. Apart from stabilization of the TS, favorable binding of the reacting diene and dienophile into the active site is the main issue to resolve, and it turns out that many candidates are too small for harboring them in a TS-like geometry.
In our presentation, we show how the geometrical aspects can be controlled using rational point mutations in the active site (fig.1), and how the catalytic effect of an OxHole can be improved by adding activating groups on both the diene and dienophile. We also compare the 'direct' mechanism (only involving the OxHole) with possible pathways that utilize more of the catalytic machinery.

References


COMP 334

Dynamic modeling and optimization of alcoholic fermentation of a grape

Hassan Ait Haddou(1), haithaddou@yahoo.com; haddou@irit.fr, 6 allée Emile Monso, BP 34038, Toulouse cedex 4 Toulouse 31029, France . (1) Laboratoire de Génie Chimique, Unstitut National Polytechnique de Toulouse, Toulouse France 31400, France

We propose a dynamic model to predicate the evolution of the constituents of the alcoholic fermentation. The parameters of this model are concentrations in sugars, ethanol, biomass, aminoacids, co2, and azote. These elements are determined by a study of sensibility. This mathematical model gives a good results, by using a dynamic programming algorithm, for considered constituents and it can be easily extended to other constituents.
Recent advances in computational fragment-based lead discovery

Carsten Detering(1), detering@biosolveit.de, An der Ziegelei 79, Sankt Augustin NRW 53757, Germany; Holger Claussen(1); Marcus Gastreich(1); Markus Lilienthal(1); Christian Lemmen(1). (1) BioSolveIT, Sankt Augustin 53757, Germany

It has frequently been shown in the course of the last years that using fragments as a starting point for buildup is a very sensible approach to finding promising new lead structures. Fragment Growing, linking and merging have been employed to successfully improve binding affinity of new chemical entities. Moreover, searching fragment spaces for novel entities that meet a certain pharmacophore or synthetic criteria is a very powerful means of quickly ascertain new lead compounds with different scaffolds and improved binding motifs.

In this contribution we bring an overview of 2D and 3D methods capable of using fragments to find, change or improve new chemical entities, scaffold hop across compound classes, and the sensible design of the underlying fragment spaces. We show some improvements we recently introduced to excel these approaches.

Some recent applications in the industry that verify these methods will be shown.

Let Gibbs be your guide: Free energy simulations in fragment-based drug design

Charles H. Reynolds(1), creynolds@ansarisbio.com, Four Valley Square, 512 East Township Line Rd., Blue Bell Pennsylvania 19422, United States; Zenon Konteatis(1); Anthony Klon(1); Sia Meshkat(1); Jinming Zou(1). (1) Anaris, Blue Bell Pennsylvania 19422, United States

We have developed a fast free energy method for simulating the interaction of small molecule fragments and their target proteins. This approach provides an accurate and efficient means for assessing the potential binding locations and affinities for thousands of molecular fragments. As such it is far more comprehensive than most other fragment-based approaches, especially those that rely on relatively low-throughput biophysical methods. Molecular fragments identified in this process can, in collaboration with medicinal chemistry, be used to design potent ligands with diverse molecular structures. We will describe the
application of these fragment free energy calculations to drug discovery, including the design of novel p38 kinase inhibitors.

COMP 337

Fragment-to-lead and peptide mimetic design using Fragment Molecular Orbital QM calculations

Richard J Law(1), richard.law@evotec.com, 114 Milton Park, Abingdon Oxfordshire OX144SA, United Kingdom; Osamu Ichihara(1); Michael Mazanetz(1); Michelle Szeto(1); Sandeep Pal(1); Mark Whittaker(1); David Hallett(1). (1) Computational Chemistry, Evotec, Abingdon Oxfordshire OX144SA, United Kingdom

Screening of low molecular weight weak binders, “fragments”, and obtaining hits is a well understood process that can be achieved by many different assay techniques. Less well defined is how to proceed once a hit is obtained. Computational chemistry, and the application of multiple techniques, plays a vital role in understanding and ranking the many potential routes for fragment expansion design. Protein-ligand interactions are routinely investigated by docking and the results are often ranked using molecular mechanics (MM) based scoring functions. MM scoring functions have many limitations and as a consequence scoring functions do not adequately predict ligand binding affinity nor do they describe the interactions in sufficient detail as to accurately and illustratively guide medicinal chemistry. To rationalize binding at a quantum level, we demonstrate the application of the fragment molecular orbital (FMO) method as a novel computational drug design methodology. As well as using FMO to rank docking results, it can also be used to achieve an atomistic level understanding of fragment binding interactions. Further to this it can be used to perform virtual fragment expansion to help guide subsequent rounds of fragment-to-lead chemistry. Examples using CDK2 and Hsp90 will be shown. Understanding of the important interactions involved in peptide/protein-protein interactions can aid in the design of mimetic inhibitors; examples using Renin and APRIL will illustrate this.

COMP 338

Fragment based screening by free energy computer simulations: A critical assessment

Jonathan W. Essex(1), jwe1@soton.ac.uk, Highfield, Southampton Hampshire SO17 1BJ, United Kingdom; Michael Bodnarchuk(1); Russell Viner(2). (1) School of Chemistry, University of Southampton, Southampton Hampshire SO17 1BJ, United Kingdom (2) Jealott’s Hill International Research Centre, Syngenta, Bracknell RG42 6EY, United Kingdom
Fragment based screening is becoming an increasingly common approach for the identification of new protein inhibitors. In this presentation, the relative merits of two free energy based simulation methods in this context will be assessed: Grand Canonical Monte Carlo, in which fragments are inserted and deleted from the protein binding site using a Metropolis test, and the new JAWS method, where fragments compete for binding locations through the sampling of a reaction coordinate which determines to what extent a particular fragment is present in the binding pocket. While Grand Canonical Monte Carlo has been used in the context of fragment based drug discovery, the JAWS method has, to our knowledge, not. This presentation will compare the performance, both in terms of accuracy and efficiency, of these two methods in identifying fragment binding locations in a number of protein-ligand systems.

COMP 339

Quantum mechanical pair-wise decomposition analysis of protein kinase B inhibitors: Validating a new tool for fragment-based drug design

Alan C Gibbs\(^{(1)}\), AGIBBS@its.jnj.com, P.O. Box 776, Spring House Pennsylvania 19477, United States; Charles H. Reynolds\(^{(2)}\); Lance M. Westerhoff\(^{(3)}\); Martin B Peters\(^{(3)}\); Xiaohua Zhang\(^{(3)}\). \(^{(1)}\) Research and Development, Johnson & Johnson Pharmaceutical, Spring House PA 19477, United States \(^{(2)}\) Research and Development, Ansaris Biopharma, Blue Bell PA 19422, United States \(^{(3)}\) Research and Development, QuantumBio Inc, State College PA 16803, United States

Quantum mechanical semiempirical comparative binding energy analysis calculations have been carried out for a series of protein kinase B (PKB) inhibitors derived from fragment- and structure-based drug design. Seven scoring functions were evaluated based on both the PM3 and the AM1 Hamiltonians. The optimal models obtained by partial least-squares analysis of the aligned poses are predictive as measured by a number of standard statistical criteria and by validation with an external data set. The interaction energy map makes it easy to identify the residues that have the largest absolute effect on ligand binding, and the structure-activity relationship (SAR) map highlights residues that are most critical to discriminating between more and less potent ligands. Taken together, the interaction energy and the SAR maps provide useful insights into drug design that would be difficult to garner in any other way.

COMP 340

Fragment based drug design using differential chemical shift perturbation induced upon binding

Lance Westerhoff\(^{(1)}\), lance@quantumbioinc.com, 200 Innovation Blvd., State College Pennsylvania 16803, United States; Jianzhong Liu\(^{(1)}\); Scott Brown\(^{(2)}\).
Andrew M. Petros(2); Kenneth M Merz(3); Bing Wang(3). (1) Quantumbio Inc, State college PA 16803, United States (2) Abbott Laboratories, Abbott Park IL 60064, United States (3) Department of Chemistry, University of Florida, Gainesville FL 32611, United States

Recently, we married semiempirical, linear scaling, quantum mechanics NMR chemical shift perturbation (CSP) prediction with experimental CSP determination in order to characterize biological systems and determine which docked binding mode is most similar to the experimental binding mode. Herein, we describe a fast, optimized method to score pharmaceutically relevant protein/ligand complexes and show how well it chooses the correct ligand binding mode from a set of docked poses. This method is applied to Bcl-xL protein and three different sets of ligand fragments – each with significantly different binding modes with significantly varied experimental CSPs. It is demonstrated that the overall predictability of the method is dependent upon the magnitude of the experimental CSP. If the experimental atomic CSPs of crucial ligand or fragment atoms are above 0.6ppm, the resulting NMRScore is shown to be highly predictive; however, if these crucial atomic CSPs drop below 0.6ppm, the predictability may be affected.

COMP 341

Flexible alignment in 3D and its applications

Adrian Kalaszi(1), akalaszi@chemaxon.com, Maramaros koz 3/a, Budapest Hungary 1037, Hungary; Miklos Vargyas(1), mvargyas@chemaxon.com, Maramaros koz 3/a, Budapest Hungary 1037, Hungary; Ferenc Csizmadia(1); Gabor Imre(1); Odon Farkas(2). (1) ChemAxon Ltd., Budapest, Hungary (2) Lorand Eotvos University, Budapest, Hungary

Fully flexible 3D alignment method without the need for sampling the conformational space is presented. This approach offers advantages over methods that rely on multiple conformations. (1) Higher accuracy is achieved, since unlike discrete sampling it is not prone to missing optimal conformation. (2) Smaller database size is needed, pre-calculated conformers are not stored.

The alignment procedure maximizes the intersection of the volume of the molecules being aligned. These volumes can be colored by atomic properties enabling the use of various similarity scores, like shape-, chemical- and biological similarity.

Similarity scores serve as the basis for 3D virtual screening. However, the calculation of 3D similarities is computationally intensive. In order to achieve high performance fast pre-filtering by 3-dimensional molecular descriptors has been introduced. These shape descriptors are pre-calculated by the application of the
alignment machinery itself in an initialization stage.

Shape descriptors include intra-molecular distance ranges, molecular volume, planar projection area ranges.

The presentation overviews the mathematical apparatus introduced, elaborates on the implemented methods and presents results.

COMP 342

Template-constrained fragment alignment (TCFA)

Richard D Cramer(1), cramer@tripos.com, St. Louis MO 63144, United States. (1) Chief Scientific Officer, Tripos Intl, St. Louis, MO MO 63144, United States

TCFA, a third approach to generating structural alignments for 3D-QSAR, combines the speed and convenience of topomers with some of the greater control that manual approaches provide. To perform TCFA, the user supplies one or more template conformations of either R-groups or entire structures. Wherever there is an appropriate and rooted match in atom and bond types between a candidate structure to be aligned and the template structure, the coordinates of the template atoms are simply transferred. To position the remaining atoms in the candidate, the topomer methodologies are used. TCFA can also provide greater focus in ligand similarity searching.

COMP 343

Chemical fragments that hydrogen bond to Asp, Glu, Arg, and His side chains in protein binding sites

Edith Chan(1), edith.chan@ucl.ac.uk, Gower Street, London London WC1E 6BT, United Kingdom; Roman Laskowski(2); David Selwood(1). (1) Wolfson Institute for Biomedical Research, University College London, London London WC1E 6BT, United Kingdom (2) European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton Cambridge CB10 1SD, United Kingdom

Fragment-based drug discovery is an established and successful paradigm. We present an analysis of the chemical fragments from lead-like ligands in the Protein Data Bank that form hydrogen bonds to the side chains of Asp, Glu, Arg, and His, which are the most common residues found in ligand binding sites. A fragment is defined as the largest ring assembly containing the atoms involved in hydrogen bonding. In total 462 fragments were found in 2,038 ligands from over 8,000 protein ligand structures in the PDB. The results show which fragments have a higher propensity for interaction with specific side chains. Some fragments interact with Asp but not with Glu, and vice versa, despite these side chains sharing the same chemical moiety. Arg side chains form hydrogen bonds
almost exclusively with O-mediated ligands and the fragments are the most diverse. Hydrogen bond distances from the imidazole of His showed a wider range than the other three amino acids.

COMP 344

Designing FBS libraries for drug discovery: Capturing fragment potential

**Johanna M Jansen**(1), johanna.jansen@novartis.com, 4560 Horton Street, Emeryville CA 94608, United States ; **Prasenjit Mukherjee**(1); **Mika K Lindvall**(1); **Yongjin Xu**(1); **Thomas J Hoeffel**(1); **Eric J Martin**(1); **Roger W Thompson**(1); **Jason Kondracki**(1). (1) Global Discovery Chemistry/Oncology & Exploratory Chemistry, Novartis Institutes for Biomedical Research, Emeryville CA 94608, United States

For a balanced design of experimental FBS libraries, it is desirable to characterize fragments using descriptors that capture the potential of these fragments in the context of expanded molecules. An approach is presented for enumerating kinase hinge-binding fragments with a standard R-group set. These enumerated libraries are then used to explore kinase selectivity potential using the Novartis proprietary 2D Profile-QSAR method. Validation of this approach is presented using data from a test set of several combinatorial libraries and predictions for companion libraries enumerated with the standard R-groups around the same core. An analysis of the selectivity potential in a virtual compound space of roughly 100 hinge binders, enumerated to 2 million compounds is performed to explore the feasibility of this approach for the design of FBS libraries.

COMP 345

CADD methodologies for lead optimization: Recent advances

**Richard D Cramer**(1), cramer@tripos.com, 1699 S. Hanley Road, St. Louis MO 63144, United States . (1) Chief Scientific Officer, Tripos Intl, St. Louis MO 63144, United States

During lead optimization, the candidate structures are very similar, often differing only by an R-group, and therefore exhibit similar potencies, with a pIC50 variance little greater than a log unit. Although "locally-derived" 3D-QSAR models are the only CADD approach than can meaningfully rank such similar candidates, 3D-QSAR has required tedious and somewhat subjective manual structural alignments. However, emerging results from make-and-test applications of the new topomer CoMFA approach, combined with exceptional eas in use, strongly encourage its general use. An unprecedented "RGVS" capability to identify the most promising and novel R-groups from among 10E6 candidates could even be project critical.
Other CADD advances for lead optimization include a docking capability that can also be "locally optimized" when multiple experimental structures with binding energies are known, and a multi-criteria optimization function adaptable to individual project goals coupled with a project-proven de novo engine.

**COMP 346**

Software for drug discovery: Are we solving the most relevant problems?

*W. Patrick Walters*\(^{(1)}\), Pat_Walters@vrtx.com, 130 Waverly Street, Cambridge Massachusetts 02139, United States. \(^{(1)}\) Vertex Pharmaceuticals Inc, Cambridge Massachusetts 02139, United States

Drug discovery is a complex process which requires teams to simultaneously optimize many parameters including target activity, selectivity, physical properties, and pharmacokinetics. Traditionally, molecular modeling software has focused on a relatively narrow subset of the drug discovery process. While some of the available computational techniques have an impact on projects, many others are of limited practical utility. In order to maximize the impact of modeling tools, software developers need to focus on the most relevant problems. This presentation will focus on the relevance of current molecular modeling software, and highlight unmet needs.

**COMP 347**

Base-by-base ratcheting of single stranded DNA through a solid-state nanopore: The DNA transistor

*Gustavo Stolovitzky*\(^{(1)}\), gustavo@us.ibm.com, P.O.Box 218, Yorktown Heights NY 10598, United States; *Hongbo Peng*\(^{(1)}\); *Binquan Luan*\(^{(1)}\); *Stefan Harrer*\(^{(1)}\); *Stas Polonsky*\(^{(1)}\); *Stephen Rossnagel*\(^{(1)}\); *Glenn Martyna*\(^{(1)}\). \(^{(1)}\) T.J. Watson Research Center, IBM Research, Yorktown Heights NY 10598, United States

A key challenge to realize nanopore-based DNA sequencing is to control the translocation of the DNA through the nanopore, providing enough dwelling time of each DNA base at a sensor location or moving each base back and forth through the sensor, to allow for repeating measurements. We recently proposed a device, the DNA transistor, designed to control the translocation of a single-stranded DNA (ssDNA) with single-base resolution. Here we report on the present status of our efforts to develop the DNA transistor. In terms of fabrication we assembled stacks of nanometer scale dielectric/metal multilayer structures. Preliminary characterization of the fabricated structures show that the resulting device is electrically sound and that it can detect the passage of DNA through the nanopore. All-atom molecular dynamics simulations results demonstrate that when pulled by an optical tweezer as in a single-molecule experiment or driven by a biasing electric field as in a high-throughput sequencing mode, the DNA
transistor forces ssDNA to slowly transit a nanopore in a ratchet-like fashion, providing a platform for nanoelectronic DNA sensing technologies.

**COMP 348**

**Exploring landscapes for protein allostery, binding, and folding: Effective potentials, replica exchange dynamics, and kinetic network models**

*Ronald Levy*\(^{(1)}\), ronlevy@lutece.rutgers.edu, 610 Taylor Road, Piscataway New Jersey 08854; *Michael Andrec*\(^{(1)}\); *Emilio Gallicchio*\(^{(1)}\); *Weihua Zheng*\(^{(1)}\). (1) Department of Chemistry & BioMaPS Institute, Rutgers University, Piscataway New Jersey 08854, United States

Advances in computational biophysics depend critically on the development of accurate effective potentials and powerful sampling methods to traverse the rugged energy landscapes that govern protein folding, binding, and fitness. I will review work in my lab over the last several years concerning the construction of all-atom effective potentials [1] for proteins and multi-scale methods for simulating their allostery, binding, and folding. Replica exchange (RE) is a generalized ensemble simulation method for accelerating the exploration of free-energy landscapes which define many challenging problems in computational biophysics, including protein allostery, binding, and folding. We have clarified some of the obstacles to obtaining converged thermodynamic information from RE simulations [2,3]. I will discuss these issues and new multi-scale approaches to recover rates and pathways for protein allostery, binding, and folding using the combined power of replica exchange, a kinetic network model with flux analysis, and effective stochastic dynamics [4, 5].

**References**

Aging: Rejuvenation and the ultimate fate of supercooled liquids

Peter G. Wolynes\(^{(1)}\), pwoynes@ucsd.edu, 9500 Gilman Drive, MS 0371, La Jolla California 92093, United States . (1) Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla California 92093, United States

I will discuss recent advances in the random first order transition theory of glasses, focusing on how spatiotemporal patterns develop in aging and rejuvenating glasses and also the problem of the ultimate fate of liquids at deep undercooling.

COMP 350

Millisecond-long molecular dynamics simulations of proteins on a special-purpose machine

David E. Shaw\(^{(1)(2)}\), David.Shaw@DEShawResearch.com, 120 W. 45th St., 39th Floor, New York NY 10036, United States . (1) D. E. Shaw Research, New York NY 10036, United States (2) Center for Computational Biology and Bioinformatics, Columbia University, New York NY 10036, United States

Molecular dynamics (MD) simulation has long been recognized as a potentially powerful tool for understanding the structural, dynamic, and functional characteristics of proteins at an atomic level of detail. Many biologically important phenomena, however, occur over timescales that have previously fallen far outside the reach of MD technology. We have constructed a specialized, massively parallel machine, called Anton, that is capable of performing all-atom simulations of proteins in an explicitly represented solvent environment at a speed roughly two orders of magnitude beyond that of the previous state of the art. Using novel algorithms developed within our lab, the machine has now simulated the behavior of a number of proteins for periods as long as a millisecond -- approximately 100 times the length of the longest such simulation previously published -- revealing aspects of protein dynamics that were previously inaccessible to both computational and experimental study.

COMP 351

Application of Markov State Models to dramatically enhance sampling of complex biomolecular simulations

Vijay Pande\(^{(1)}\), pande@stanford.edu, Stanford University, Stanford CA 94305, United States . (1) Department of Chemistry, Stanford University, Stanford CA 94305, United States
One of the major challenges of connecting simulation to many application areas, especially biological applications such as protein folding or lipid vesicle fusion, is the relatively long time scales found experimentally (milliseconds to seconds) compared to what is typically possible computationally (nanoseconds to microseconds). I will describe our efforts to break past these time scale barriers with Markov State Models (MSM's). In our MSM formulation, we have means to automatically identify relevant states, very efficiently sample transitions from these states using adaptive methods, and then to finally statistically test and compare models. I will demonstrate this method using a diverse set of biological examples, including protein folding, misfolding of Alzheimer's peptides, and lipid vesicle fusion.

COMP 352

Multiscale theory and simulation of biomolecular systems

Gregory A Voth\(^{(1)}\), gavoth@uchicago.edu, 5735 South Ellis Avenue, Chicago IL 60637, United States. (1) Department of Chemistry, James Franck Institute, and Computation Institute, University of Chicago, Chicago IL 60637, United States

A multiscale theoretical and computational methodology will be presented for studying biomolecular systems across multiple length and time scales. The approach provides a systematic connection between all-atom molecular dynamics, coarse-grained modeling, and mesoscopic phenomena. At the heart of the approach is the multiscale coarse-graining method for rigorously deriving coarse-grained models from the underlying molecular-scale interactions. Applications of the multiscale approach will be given for membranes and proteins, although the overall methodology is applicable to other complex condensed matter systems. Recent applications to protein-mediated membrane remodeling and to large biomolecular complexes such as the HIV-1 virion will be described. The theoretical/conceptual challenges and opportunities for this area of multiscale simulation will be emphasized.

COMP 353

Multi-scale simulation methods for modeling two-step ligand-protein binding processes

Myungshim Kang\(^{(1)}\), myungk@ucr.edu, 0426 Chemical Sciences, Riverside CA 92521, United States; Chia-en Angelina Chang\(^{(1)}\). (1) Department of Chemistry, University of California Riverside, Riverside CA 92521, United States
Ligand-protein association is a fundamental step in many biological processes. Some ligand-protein association rates are controlled by diffusion-control one-step association processes. On the other hand, some ligand-protein systems employ more complicated two-step association processes that involve short-range interactions and conformational changes of ligands and proteins. However, studying the association pathways and predicting rate constants for two-step associations is challenging, but it is of interest and has practical applications such as drug design. Here we use HIV-1 protease as one of our study cases. The protein has flaps over its active site, whose opening and closing can “gate” ligand binding. Multiple ligands are introduced to the systems to reflect the biological concentration of ligands. The work uses multi-scale simulation methods, combining Brownian dynamics with a coarse-grained model, atomistic simulations and continuum models to study the molecular encounter process.

COMP 354

Multiscale simulations of F1-ATPase

Jingzhi Pu(1)(2), pu@tammy.harvard.edu, 402 N. Blackford Street, LD 326, Indianapolis IN 46202, United States; Martin Karplus(2)(3). (1) Dept. of Chem. and Chem. Biol., Indiana Univ. - Purdue Univ. Indianapolis, Indianapolis IN 46202, United States (2) Dept. of Chem. and Chem. Biol., Harvard University, Indianapolis IN 46202, United States (3) Laboratoire de Chimie Biophysique (ISIS), Universite de Strasbourg, Cambridge MA 02138, United States

The realization that many essential functions of living cells are performed by nanoscale motors consisting of protein complexes has given rise to an intense effort to understand their mechanisms. The fundamental question concerning such systems is how to explain the macroscopic phenomena in terms of the atomic structures and microscopic forces involved. Multiscale modeling using coarse-grained as well as all-atom simulations based on molecular mechanics or combined quantum mechanics/molecular mechanics, provide the essential link between the static structures, as determined by x-ray crystallography, and function. Complementary to single molecule experiments, computer simulations may also be useful to obtain information that is not available from experiment. The presentation will be focused on understanding the working mechanisms of the rotary motor F1-ATPase by use of multiscale modeling.

COMP 355

Constructing a mechanistic model for a processive cellulose-degrading enzyme from molecular simulation

Gregg T. Beckham(1), Gregg.Beckham@nrel.gov, 1617 Cole Blvd., Golden CO 80401, United States; James F. Matthews(1); Yannick J. Bomble(1); Lintao Bu(1); William S. Adney(1); Michael E. Himmel(1); Mark R. Nimlos(2); Michael F.
The Family 7 processive cellulases are the basis of many enzyme cocktails for cellulosic biomass conversion. Despite their importance, a detailed mechanistic understanding of the Family 7 enzymes remains elusive. Here we present recent results illustrating our ongoing efforts to construct a molecular-level model of the action of a Family 7 cellulase, using the Trichoderma reesei Cel7A enzyme as the basis for study. The three sub-domains of Cel7A, namely the Family 1 carbohydrate binding module, the O-glycosylated linker peptide, and the catalytic domain are studied independently to ascertain isolated structure-function relationships. The knowledge gained from studies of the individual domains is vital in developing a quantitative, mechanistic model of the processive action of Cel7A. Additionally, the free energy cost to decrystallize cellulose is an essential part of this model; thus, we also present a complementary study in which we measure the molecular-level basis for cellulose recalcitrance, which Cel7A (and other cellulase enzymes) must overcome to degrade the plant cell wall.

COMP 356

Allostery of protein as determined by dynamics fluctuations and correlations

Ming S. Liu\textsuperscript{(1)}, ming.liu@csiro.au, Private Bag 33, Clayton South VIC 3169, Australia. (1) Division of Mathematics, Informatics and Statistics, CSIRO, Clayton South 3169, Australia

Dynamics coupling, correlated motion and allosteric cooperativity appear to be conserved in the long range communication and conformational transitions of many proteins. For example, a simple mutation can produce marked effects at distant sites via undefined pathways for a conventionally non-allosteric protein. There is reconciling evidence on allostery mechanisms for the 'induced-fit' scheme and the 'population-shift' theory, where dynamics plays an essential role in allosteric regulations. Using model systems of kinase AdK and signal regulator NtrC, we derived a dynamics criterion to determine possible allostery in globular proteins: Given two distinctive conformational states, dynamical fluctuations and correlations, either amongst the distant functional motifs or different subunits, can be accounted for by the conformational transitions between them. If the dynamics correlations result in both correlated and anti-correlated modes of motions (Figure 1), allosteric cooperativity will occur simultaneously.
Figure 1. Difference matrix of dynamic correlations in protein – the Aquifex AdK case: red, a correlated motion; blue, an anti-correlated motion; and red (blue) regions correspond to same (opposite) direction distortions. The presence of both positive and negative correlations indicates the existence of an allosteric cooperativity during conformational changes, and this was proved by NMR experiments.

COMP 357

Investigating conformational changes of biological macromolecules using multi-resolution Markov State Models

Xuhui Huang(1), xuhuihuang@ust.hk, Clear Water Bay, Kowloon, Hong Kong Special Administrative Region of China. (1) Department of Chemistry, Hong Kong University of Science and Technology, Hong Kong Special Administrative Region of China

Simulating biologically relevant timescales at atomic resolution is a challenging task since typical atomistic simulations are at least two orders of magnitude shorter. Markov State Models (MSMs) provide one means of overcoming this gap without sacrificing atomic resolution by extracting long time dynamics from short simulations. MSMs coarse grain space by dividing conformational space into long-lived, or metastable, states. This is equivalent to coarse graining time by integrating out fast motions within metastable states. By varying the degree of coarse graining one can vary the resolution of an MSM; therefore, MSMs are inherently multi-resolution. In the talk, I will introduce a new algorithm Super-level-set Hierarchical Clustering (SHC), to our knowledge, the first algorithm
focused on constructing MSMs at multiple resolutions. The key insight of this algorithm is to generate a set of super levels covering different density regions of phase space, then cluster each super level separately, and finally recombine this information into a single MSM. SHC is able to produce MSMs at different resolutions using different super density level sets. To demonstrate the power of this algorithm we first apply it to a RNA hairpin, generating MSMs at different resolutions. We validate these MSMs by showing that they are able to reproduce the original simulation data. Furthermore, long time folding dynamics are extracted from these models. We also apply SHC to study the role of water in the hydrophobic collapse of the Mellitin tetramer. Using MSMs, we are able to identify pathways adapting a drying induced hydrophobic mechanisms, and metastable intermediate states on these pathways.

COMP 358

Discovering conformational sub-states essential to protein function: A multi-scale approach

Arvind Ramanathan\(^{(1)}\), aramanat@andrew.cmu.edu, 5000 Forbes Ave, Pittsburgh PA 15213, United States; Christopher J. Langmead\(^{(2)}\); Chakra S. Chennubhotla\(^{(3)}\); Pratul K. Agarwal\(^{(4)}\). (1) Lane Center for Computational Biology, Carnegie Mellon University, Pittsburgh PA 15213, United States (2) Department of Computer Science, Carnegie Mellon University, Pittsburgh PA 15213, United States (3) Department of Computational Biology, University of Pittsburgh, Pittsburgh PA 15213, United States (4) Computational Biology Institute and Computer Science & Mathematics Division, Oak Ridge National Lab, Oak Ridge TN 37831, United States

Conformational transitions are essential for protein function. These transitions are governed by internal motions in the protein, which are hierarchical, involving a multitude of spatial and temporal scales. However, characterizing the hierarchy of functionally relevant motions using traditional approaches including quasi-harmonic analysis has been challenging. To elucidate the multi-scale nature of the conformational landscape, a general approach, Quasi-Anharmonic Analysis (QAA) is described, which highlights the statistical regularities in the conformational and energetic landscape. By effectively exploiting the higher-order correlations, in addition to the variance observed from long time-scale simulations, QAA can provide mechanistic insights into biologically relevant conformational transitions.

QAA, when used to describe a microsecond ensemble of ubiquitin conformations, elucidates a conformational hierarchy that allows its binding region to be modulated to accommodate a variety of substrates. These sub-states, while sharing significant conformational and energetic homogeneity,
can be used in a self-consistent way to reveal novel insights into the binding motions of ubiquitin (Fig. 1). Enzyme motions along a reaction in CypA are also analyzed to reveal conformational sub-states and hierarchy of motions that may impact the catalytic turnover.

COMP 359

Introduction to cross pharma high performance computing forum

John C Morris(1), john.morris@pfizer.com, 35 Cambridge Park Drive, Cambridge MA 02140, United States; Zheng Yang(2). (1) Massachusetts Research Business Technology, Pfizer, Cambridge MA 02140, United States (2) Department of Computational and Structural Chemistry, GlaxoSmithKline Pharmaceuticals, Collegeville PA 19426, United States

High Performance Computing (HPC) within the pharmaceutical industry is a growing and critical component of research due to the large scale analytical demands driven by modern research methods and advancements in computational chemistry and bioinformatics methods to model biological systems. HPC has become a necessary capability to facilitate the analysis of the terabytes of scientific data being generated from technologies such as Next Generation Sequencing, modeling complex drug-target interaction, and statistical analysis. To support the industrialization of scientific research, integrated and coordinated HPC information technology tools, methods, and capabilities are needed. The Cross Pharma HPC forum is a group of scientists, engineers, and key stakeholders within the pharmaceutical industry working together to promote
best practices, coordinate activities, optimize methods, and leverage experience in the non-competitive areas within HPC. In this talk, the history, current status, and future directions of HPC in the pharmaceutical industry will be discussed.

COMP 360

Applications and use of cloud computing in the pharmaceutical industry

Michael D Miller(1), michael.d.miller@pfizer.com, Eastern Point RD, Groton CT 06349, United States; David M Powers(2); Gregory Stiegler(3); Jeremy Martin(4). (1) Research Business Technology, Pfizer, Groton CT 06340, United States (2) Research and Development IT, Eli Lilly, Indianapolis Indiana, United States (3) Scientific Computing, Bristol-Myers Squibb, Princeton New Jersey, United States (4) System Support Department, Information Technology, GlaxoSmithKline R&D Ltd, Harlow, Essex CM19 5AW, United Kingdom

Technological advances across the sciences have enabled basic drug research with an unprecedented amount of data. As a result, the application of computational methods are becoming an increasingly important approach in drug discovery and development. The need for increased computing capacity has reached the point where, today it can become rate limiting. As a result Pharmaceutical companies have begun exploring the use of cloud computing to address these needs. We will present on some of the challenges Pharmaceutical companies have faced in using cloud resources and the different approaches that have been taken to address them.

COMP 361

Current trends of high performance computing in Pharma

Stephen Litster(1), stephen.litster@novartis.com, 200 Technology Square 8th Floor, Cambridge MA 02139, United States; Jeremy Martin(2). (1) NITAS Scientific Computing, Novartis Institutes of BioMedical Research, Cambridge MA 02139, United States (2) Department of System Support, Information Technology, GlaxoSmithKline R&D Ltd, Harlow Essex CM19 5AW, United Kingdom

The world of high performance computing (HPC) has evolved quickly, as exemplified by recent developments in hardware (e.g. Intel Nehalem multi-core CPUs with integrated memory controller), software (e.g. NAMD, a highly scalable molecular dynamics program), computing services (e.g. cloud computing), and storage (TB+ scale file systems). Given these recent developments and much lower cost of entry into HPC, Pharma based Scientific Computing groups are beginning to apply traditional HPC techniques to “non-traditional” (e.g. High Content Screening) and emerging areas of research (e.g. Next Generation Sequencing).
We present here a number of case studies highlighting the current trends of HPC in the pharmaceutical industry and its to impact scientific workflows.

COMP 362

Challenges of HPC and collaboration opportunities in Pharma

Robert Stansfield\(^{(1)}\), rob.stansfield@sanofi-aventis.com, MailCode M-203A, PO Box 6800, Bridgewater NJ 08807, United States; Michael D Miller\(^{(2)}\), michael.d.miller@pfizer.com, Eastern Point Road, Groton CT 06340, United States. \(^{(1)}\) R&D Information Solutions, sanofi-aventis U.S., Bridgewater NJ 08807, United States (2) Research Business Technologies, Pfizer, Groton CT 06340, United States

High Performance Computing (HPC) in Pharmaceutical R&D is well established in computational chemistry and computational biology for drug discovery, but is increasingly seeing broader application across research and development. In addition, internal capacity is being supplemented by external "cloud computing". In consequence, the issues around providing HPC services to in-house scientists in an optimal way for the entire company become more visible and critical. From a technical perspective, HPC requires a holistic view across compute, network, and storage capabilities. From an organizational perspective, effective governance – roles, responsibilities, prioritization and decision making across multiple different groups, operations, and support to end-user scientists - makes all the difference. For these reasons at least, HPC deserves a place in strategic planning. These issues will be explored, as well as the opportunities afforded by pre-competitive collaboration in the Cross-Pharma HPC Forum for identifying best practices.

COMP 363

Importance of pharmaceutically relevant benchmark data

Christian Lemmen\(^{(1)}\); Marcus Gastreich\(^{(1)}\); Carsten Detering\(^{(1)}\); Peter Oledzki\(^{(1)}\). \(^{(1)}\) BioSolveIT, Germany

As a methods developer it is highly desirable to have high-quality benchmark data available for testing. Unfortunately the number and quality of such data sets that are published and freely available is quite limited. Also if such data is available, methods developers may use the data for training and testing. It was therefore quite valuable for the community to have a data-set that was not only well prepared and standardized such as to be applicable for various computer programs, but also to be able to preform blind studies which prevented the developers from overfitting their methods in an effort to improve performance. We will share our experience from application studies based on such data and
explain in which way this guided the further development of our docking technology.

**COMP 364**

**Insights from “real world” docking with RosettaLigand**

_Ian W Davis_\(^{(1)}\), ian.w.davis@gmail.com, 302 E. Pettigrew St., Durham NC 27701, United States. \( (1) \) Department of Biochemistry, University of Washington, Seattle WA, United States

Predicting the 3-D interactions of a drug-like molecule with a protein receptor remains a difficult problem. However, retrospective tests often unintentionally make it seem easier than it is. First, the method or parameters may be over-fit to the test data. Second, the test cases may not be representative of future cases to be predicted. Finally, fair comparison with methods that were evaluated on different test cases may be difficult. In protein structure prediction, progress has been driven by making available previously unseen data for blind predictions. Now, the CSAR project promises to do the same for small molecule docking and scoring. I experienced these benefits firsthand by running the RosettaLigand docking program against an extensive blind test set offered by GlaxoSmithKline. I'll share insights into search and scoring problems uncovered by this and other blind tests of RosettaLigand, and discuss ideas for overcoming them.

**COMP 365**

**Recent experiences with molecular docking in MOE**

_Chris Williams_\(^{(1)}\), cw@chemcomp.com, 1010 Sherbrooke Street West, Suite 910, Montreal Quebec H3A 2R7, Canada; _Paul Labute_\(^{(1)}\). \( (1) \) Chemical Computing Group, Montreal Quebec H3A 2R7, Canada

We summarize CCG’s experiences in the last several years with protein:ligand docking, binding affinity prediction, validation collections and experiments and how these experiences have affected MOE’s docking methodology. The results of docking experiments are presented along with a discussion of issues relevant to the reliable application of docking methodology.

**COMP 366**

**Advances in induced-fit docking: What we have learned (and not learned) from the GSK dataset**

_Woody Sherman_\(^{(1)}\), woody.sherman@schrodinger.com, 120 West 45th Street, 17th Floor, New York NY 10036, United States; _Devleena Shivakumar_\(^{(1)}\). \( (1) \) Schrodinger, New York NY 10036, United States
The ability to accurately predict the structure of ligand-receptor complexes when protein flexibility is required is an important and challenging part of structure-based drug design that has made significant advances in recent years. Since our initial publication on induced-fit docking (Sherman et al., J. Med. Chem., 2006, 49, 534–553), we have been working on developing the next generation protocol to improve the accuracy of ligand pose and protein structure predictions. Here, we present the new methodological advancements and improved results. We also discuss how the GSK dataset has (or has not) helped us in our development efforts.

COMP 367

Surflex-Dock and the GSK Dataset: A retrospective

Essam Metwally(1), essam.metwally@certara.com. (1) Research and Development, Tripos International, St. Louis MO 63146, United States

In, "A Critical Assessment of Docking Programs and Scoring Functions," by Warren et al, several docking programs were evaluated on a pharmaceutically relevant dataset where the results were unknown to the dockers a priori. Surflex-Dock was later evaluated in near identical conditions. While Surflex-Dock was among the top docking programs when used for virtual screening, and the top program when used in "Virtual Crystallography", the knowledge garnered from the design and setup of the experiment was invaluable. Much of what was learned has since been incorporated into Surflex-Dock and what was learned will be discussed.

COMP 368

How to build a better mousetrap: Experiences with docking/scoring benchmarks – including the GlaxoSmithKline set – and what has been learned

Lance M Westerhoff(1), lance@quantumbioinc.com, 200 Innovation Blvd, State College PA 16803, United States. (1) QuantumBio inc., State College PA 16803, United States

Gaining actionable intelligence from protein/ligand interaction characterization and scoring has been largely dependent upon the experience of the practitioner. Vendors and academicians alike have been hard at work developing software to better capture these important interactions and provide more trustworthy results that are able to treat structures of varying complexity and diversity. In
order to measure the success of these new tools, various benchmarks have been developed. With exposure to several major benchmark sets QuantumBio has experienced the gambit of issues that can arise when treating benchmarks with highly variable motifs and with experimental data of unknown quality. A comparison of the current state of available benchmarks is discussed along with rules of thumb that allow one to treat them. Significant focus is paid to benchmarks, such as the GlaxoSmithKline set, that are pharmaceutically relevant and the lessons that have been learned and applied to newer software versions is explored.

COMP 369

Pocket similarity: Are alpha carbons enough?

Howard Feldman(1), hfeldman@chemcomp.com, 1010 Sherbrooke Street West, Suite 910, Montreal Quebec H3A 2R7, Canada. (1) Chemical Computing Group, Montreal Quebec H3A 2R7, Canada

A novel method for measuring protein pocket similarity was devised, using only the alpha carbon positions of the pocket residues. Pockets were compared pairwise using an exhaustive 3D C-alpha common subset search and grouping residues by physicochemical properties. At least five C-alpha matches were required for each hit, and distances between corresponding points were fit to an Extreme Value Distribution (EVD) resulting in a probabilistic score or likelihood for any given superposition. To test the utility of this score, it was successfully used to cluster a set of 85 structures from 13 diverse protein families based on binding sites alone, as well as to cluster 25 kinases into a number of subfamilies. Using a test kinase query to retrieve other kinase pockets, it was found that a specificity of 99.2% and sensitivity of 97.5% could be achieved using an appropriate cutoff score. The search itself took from 2-15 minutes on a single 3GHz CPU to search the entire PDB (130,000 pockets), depending on the number of hits returned.

COMP 370

Chemically meaningful net atomic charges and atomic multipole moments that reproduce the electrostatic potential in periodic and nonperiodic materials

Thomas A. Manz(1), thomasamanz@gmail.com, 311 Ferst Drive, N.W., Atlanta GA 30332-0100, United States. David S. Sholl(1). (1) School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta GA 30332-0100, United States

Net atomic charges (NACs) are widely used to understand the chemical states of atoms in a material and to represent the electrostatic potential, V, of a material
outside its electron distribution. An ideal method for calculating NACs should fulfill this dual purpose and be applicable to both periodic and nonperiodic materials, but existing NAC definitions fall short of this goal. We present a new approach, Density Derived Electrostatic and Chemical (DDEC) charges, that overcomes these limitations. DDEC is an atoms-in-molecules method that uses a distributed multipole expansion to formally reproduce V exactly outside the electron distribution. We compare different charge methods for a variety of materials periodic in 0, 1, 2, and 3 dimensions. The DDEC method performed consistently well for molecules, nonporous solids, solid surfaces, and porous solids like metal organic frameworks. We anticipate this method will find widespread use for the calculation of forcefield charges in complex materials.

COMP 371

Ultrafast electrostatic similarity: The sound of inevitability?

Paul C Hawkins(1), phawkins@eyesopen.com, 9 Bisbee Court, Santa Fe NM 87508, United States ; Anthony Nicholls(1). (1) Applications Science, OpenEye Scientific Software, Santa Fe NM 87508, United States

A number of recent publications have shown that the electrostatic similarity tool EON has significant utility both in guiding compound selection for low throughput screening and in prospective lead hopping experiments. While a single EON comparison is rapid (a few hundredths of a second), repeated sampling of terminal rotors can become expensive. Yet, in order to reliably capture the maximal electrostatic similarity, such an evaluation is often necessary, e.g. if such rotors have a significant dipole component orthogonal to their axis of rotation. As such, we have developed a faster but approximate algorithm that improves comparison by 3-4 orders of magnitude. We will present on its efficacy and application.

COMP 372

Bringing GPUs to mainstream molecular dynamics packages, acceleration of AMBER molecular dynamics simulations using NVIDIA GPUs: Achieving high performance without sacrificing accuracy

Ross C Walker(1), rcw@sdsc.edu, 9500 Gilman Drive MC0505, La Jolla CA 92093-0505, United States . (1) San Diego Supercomputer Center, University of California, San Diego, La Jolla CA 92093-0505, United States

Recent work in close collaboration with NVIDIA has produced a GPU accelerated version of the AMBER Molecular Dynamics Code PMEMD that runs between 20 and 130 times the speed of a single 2.8GHz Intel Nehalem Processor, with even higher performance on multiple GPUs, but which does not make sacrifices in the accuracy or validity of such calculations to achieve this. The GPU accelerated
version supports both explicit solvent particle mesh ewald (PME) and implicit solvent simulations and is available as part of the new AMBER 11 package. This talk will provide an overview of the AMBER software, background behind this GPU work, benchmarks, the impact that GPU accelerated MD can have on the field, the techniques used to achieve the performance seen without sacrificing accuracy and finally the validation methods used to ensure simulations are directly equivalent to CPU based calculations. Ensuring that a GPU implementation of a MD package provides results that are indistinguishable from the CPU code is extremely tricky and often the desire to take shortcuts to boost performance can affect accuracy with unpredictable results. We have developed a comprehensive validation suite that can be used to perform the detailed testing that is required to ensure the approximations necessary for GPU performance do not impact the scientific results. Additionally we will discuss how we have made careful use of mixed single and double precision arithmetic in the AMBER implementation to achieve equivalence in the results without excessively compromising performance. Finally we provide examples of recent breakthrough simulations conducted using GPU enabled AMBER 11.

COMP 373

Quantum polarized fluctuating charge (QPFC) model: A practical method to include quantum mechanical ligand polarizability in force-field-based molecular simulation

S. Roy Kimura(1), roy.kimura@bms.com, 5 Research Parkway, Wallingford CT 06492, United States; David R. Langley(1); Ramkumar Rajamani(1). (1) Department of Computer Assisted Drug Design, Bristol Myers Squibb, Wallingford CT 06492, United States

We present a tractable method to include ligand polarization in molecular dynamics (MD) simulation of biomolecular systems. The method involves periodically saving coordinate snapshots to disk during the MD simulation, and evaluating the electrostatic potential (ESP) around the ligand using Quantum Mechanical (QM) calculations run in parallel on an extra set of processors. The QM ESPs for the ligand are calculated in the presence of the electric field set up by the protein and solvent partial charges. Ligand partial charges are assigned to atom centers using the multi-conformer Restrained Electrostatic Potential (RESP) fit method using several successive coordinate snapshots from the MD simulation. The updated charges are introduced back into the simulation in real-time when the next snapshot is saved. The result is a simulation whose ligand partial charges respond to the electrostatic field of its evolving environment. Test simulations on the HIV Protease-Atazanavir system is discussed.

COMP 374

Status of the ReaxFF reactive forcefield: Development and applications
The ReaxFF method enables large-scale (>>1000 atoms) molecular dynamics simulations on chemically reactive systems. The method combines a bond order/bond distance concept with a polarizable charge method and employs relatively long-range bond orders, enabling a accurate description of transition state energies. While initially developed for hydrocarbons and first-row element chemistry, the method has currently been applied to a significant section of the periodic system, including covalent, metallic, ionic and ceramic materials. Furthermore, the ReaxFF program has been distributed to over 100 academic groups.

In this presentation we will provide a status-report on the currently available ReaxFF force fields, the program environments that currently support ReaxFF and the plans for further method development. This includes recent extensions of ReaxFF to aqueous-phase chemistry, enabling applications to biochemical reactions, extensions and applications to combustion chemistry, applications to material failure and simulations on catalytic materials.

COMP 375

Parametrization of site-specific force fields for computational enzymology

Guillaume Lamoureux(1), glamoure@alcor.concordia.ca, 7141 Sherbrooke Street West, Montreal Quebec H4B 1R6, Canada. (1) Department of Chemistry and Biochemistry, Concordia University, Montreal Quebec H4B 1R6, Canada

We have developed a framework that allows the easy parametrization of semi-empirical force fields for enzyme simulations based on ab initio energy surfaces of the fragments forming the active site. To compensate for the inherent lack a transferability of force fields, the framework emphasises quick parametrization of site-specific models and quick retrieval of existing models from any similar active site. The ab initio surfaces are explored using simple coordinates transformations based on the modes of vibration of the fragments, or using "scans" of internal coordinates. They are generated using a web interface and stored as a database. The force field parameters are adjusted to minimize a chi-square function assembled from the chi-square functions of the different fragments. The approach will be illustrated with zinc-containing metalloenzymes, which have a clearly identifiable active site and a finite catalytic "repertoire". Results for QM/MM simulations using site-specific, AM1-like models as "QM" model will be
presented. Compared to conventional, DFT-based QM/MM simulations of enzymatic reactions, those reaction-specific force fields lead to significant speedup.

COMP 376

Approaches to the treatment of multidrug resistant gram negative infections

Mark C Noe\(^{(1)}\), mark.c.noe@pfizer.com, Eastern Point Road, Groton CT 06340, United States ; Steven J Brickner\(^{(1)}\); Thomas Gootz\(^{(1)}\); Michael Huband\(^{(1)}\); Mark E Flanagan\(^{(1)}\); John Mueller\(^{(1)}\). (1) Department of Antibacterials Research, Pfizer Global Research and Development, Groton CT 06340, United States

Each year, over 4.3 million people worldwide contract hospital-based bacterial infections, approximately half of which are caused by Gram negative organisms. The widespread emergence of genes that confer multidrug resistance in these pathogens threatens to undermine the clinical utility of several antibiotic classes, including the fluoroquinolones, cephalosporins, carbapenems and aminoglycosides. Particularly concerning are the extended spectrum beta lactamases, including carbapenemases, which are advancing at an alarming rate and compromise the effectiveness of the most widely used classes to treat Gram negative infections. This talk will review the medical need for new antibacterial agents, some of the challenges associated with discovering new antibiotics, examples of potentially enabling technologies and recent advances in our understanding of privileged targets for antibacterial therapy. An example of one antibacterial drug discovery program will be presented.

COMP 377

Physicochemical property space of antibiotics

Heinz E Moser\(^{(1)}\), hmoser@achaogen.com, 7000 Shoreline Court, South San Francisco CA 94080, United States . (1) Department of Chemistry, Achaogen, South San Francisco California 94080, United States

While there have been enormous discovery efforts during the past decades to identify novel classes of antibacterials with clinical utility against Gram-negative pathogens, no first-in-class compounds have been successfully developed to use in humans for roughly half a century, and none is currently in clinical evaluation. Predictably, this lack of success has been met by an increasing prevalence of Gram-negative pathogens causing serious infections in hospitals and critical care settings. Recent outbreaks caused by multi-drug resistant (MDR) or pan-resistant organisms such as \textit{K. pneumoniae} have been reported recently and leave physicians with few to no treatment options. This presentation focuses on the physico-chemical property space of antibacterial drugs and how an
understanding of this property space can assist in the discovery and lead optimization of antibiotics, in particular that of antibacterial drugs active against Gram-negative bacteria. Specific examples will be presented and discussed in detail.

COMP 378

Physicochemical properties correlated with Gram-negative antibacterial activity of compounds in the Pfizer corporate library

Jeremy T Starr\textsuperscript{(1)}, jeremy.starr@pfizer.com, 445 Eastern Point Road, Groton CT, United States; Rishi Gupta\textsuperscript{(1)}; Veerabahu Shanmugasundaram\textsuperscript{(1)}. (1) Department of Antibacterials and Discovery Technologies, Pfizer Pharmatherapeutics Research and Development, Groton CT 06340, United States

Correlation of computed physicochemical properties of Pfizer proprietary compounds with their respective \textit{E. coli} or \textit{P. aeruginosa} MICs has led to the identification of a physicochemical fingerprint associated with higher probability of whole cell activity with a cytosolic target and presumed passive cell penetration. A computational tool has been designed to calculate a desirability quotient based on these parameters which demonstrates positive differentiation of higher scoring compound classes.

COMP 379

Combining lessons from computational design of gram positive antibacterials with datamining to aid the design of novel gram negative antibacterials

Charles J. Eyermann\textsuperscript{(1)}, joe.eyermann@astrazeneca.com, 35 Gatehouse Drive, Waltham MA 01776, United States. (1) Infection Discovery, AstraZeneca, Waltham MA 02451, United States

Our approach to address the emergence of resistant bacterial strains has been to identify new chemotypes with a novel mode of action. A significant effort has been made to develop novel inhibitors against gram positive strains like Methicillin-resistant \textit{Staphylococcus aureus} (MRSA) These efforts have provided a number of key lessons related to target isozyme specificity and drug safety margins. Work to identify novel MruI inhibitors of \textit{H. pylori} has also provided insights into the physiochemical properties that impact gram negative antibacterial activity. Combining the lessons learned from the above research efforts with datamining of existing gram negative agents provides a framework to aid in the optimization of novel leads for gram negative antibacterials.

COMP 380
Targeting gram-negative pathogens: Drug design to improve antibiotics permeation?

Matteo Ceccarelli (1), mceccare@katamail.com, Laboratory for Computational Materials Science, Monserrato I-09042, Italy; Eric Hajjar (1); Amit Kumar (1); Paolo Ruggerone (1). (1) Department of Physics, Universita degli Studi di Cagliari and Sardinian, Monserrato I-09042, Italy

Gram-negative bacteria are protected by an outer membrane and to function, antibiotics have to diffuse passively through outer membrane channels, known as porins, such as OmpF in E.coli (Pages, J. M. et al. Nat. Rev. Microbiol. 2008, 6, 893). Bacterial strains can modulate their susceptibility to antibiotics by under-expressing or mutating the structures of porins, becoming resistant, in the worst case, to different antibiotics families. These multidrug resistant bacteria are now ubiquitous in both hospitals and the larger community and the resurrection of tuberculosis provides one ominous example highlighting the risk associated with evolved drug resistance (Cars, O. et al. Brit. Med. J. 2008, 337, 726). Moreover, many pharmaceutical companies abandoned this field and no truly novel active antibacterial compounds are currently in clinical trials. A major current dilemma for the pharmaceutical industry is whether to develop drugs for new targets or promote those drugs presently on the market (Weiss, D. et al. Nat. Rev. Drug. Discov. 2009, 8, 533.), identifying bottlenecks of existing antibiotics to suggest chemical modifications. Following such a strategy, we revealed the complete permeation pathways of β-lactams and fluoroquinolones antibiotics through porins using metadynamics simulations and found that experimental results remarkably confirmed the computational predictions. Further, simulations revealed its potentiality to overcome experimental limitations and provide microscopic details on the permeation process (Hajjar, E. et al. Biophys. J. 2010, 98, 569; Mahendran K. et al. J. Phys. Chem. B, IN PRESS).

Here we follow the paradigm for selecting antibiotics with better permeation properties using computer simulations only. Taking advantage of the atomic level of detail that the simulations provide we find that the diffusion of ampicillin through OmpF is governed by a subtle balance of interactions with partners in the porin channel: we draw, for the first time, the complete inventory of the rate-limiting interactions and map them on both the porin and antibiotics structure. Our methodology, which can be conveniently employed to study other porins/antibiotics, allows identifying the functional groups that govern optimal translocation. Such findings will directly benefit rational antibiotics design, by defining for example, some appropriate pharmacophores within high throughput screening strategies.

COMP 381

Structure-based lead optimization of novel bacterial type II topoisomerase inhibitors
Neil D Pearson(1), Neil.Pearson@gsk.com, 1250 South Collegeville Road, Collegeville Pennsylvania 19426, United States; Zheng Yang(2); Benjamin D Bax(3); Michael N Gwynn(4). (1) Department of Antibacterial Chemistry, Infectious Diseases Center of Excellence in Drug Discovery, GlaxoSmithKline Pharmaceuticals, Collegeville Pennsylvania 19426, United States (2) Department of Computational and Structural Chemistry, GlaxoSmithKline Pharmaceuticals, Collegeville Pennsylvania 19426, United States (3) Department of Computational and Structural Chemistry, GlaxoSmithKline Pharmaceuticals, Stevenage Hertfordshire SG1 2NY, United Kingdom (4) Department of Antibacterial Microbiology, Infectious Diseases Center of Excellence in Drug Discovery, GlaxoSmithKline Pharmaceuticals, Collegeville Pennsylvania 19426, United States

The emergence of multi drug resistant Gram negative pathogens is a major concern given the paucity of new therapies in clinical development. GSK has discovered a novel series of inhibitors of both DNA gyrase and topoisomerase IV (NBTIs) with a unique mechanism and no target based cross resistance to established classes of antibacterials including the fluoroquinolones. Optimisation of the Gram positive selective early leads led to new series which afforded good activity versus Gram negative pathogens. GSK subsequently solved the first X-ray structure of a NBTI inhibitor in complex with S.aureus DNA gyrase and DNA providing unprecedented knowledge for lead optimization and the design of novel inhibitors. This talk will discuss how the structural information enabled the medicinal chemistry team to design new subunits as well as illustrating when optimization of interactions with the binding site have been well served by traditional medicinal chemistry.

COMP 382

Fragment-based development of tetrazole inhibitors against class A beta-lactamase

Yu Chen(1), ychen1@health.usf.edu, 12901 Bruce B. Downs Blvd, MDC07, Tampa FL 33612, United States. (1) Department of Molecular Medicine, University of South Florida, Tampa FL 33612, United States

The production of beta-lactamases is the predominant cause of resistance to beta-lactam antibiotics, such as penicillins, in Gram-negative bacteria. Whereas high throughput screening has appeared insufficient for the development of new beta-lactamase inhibitors, fragment-based methods provide an effective approach in sampling novel chemical space in antibiotics discovery. We have previously used fragment-based molecular docking to identify mM range tetrazole inhibitors against CTX-M Class A beta-lactamase and to subsequently evolve their affinities to ~10 micromolar. New compounds have now been synthesized using the micromolar-affinity tetrazole scaffold, based on some
similarities between this scaffold and beta-lactam antibiotics or on X-ray crystal structures of the inhibitor-bound complexes. Other fragment compounds have also been tested to probe regions of the active site not sampled by existing inhibitors. Combining the fragment-based approach with molecular docking, X-ray crystallography and chemical synthesis, we hope to eventually develop these tetrazole compounds into nM inhibitors.

COMP 383

Elucidating the molecular mechanisms underlying the nucleation and growth of nanoparticles

Jerome Delhommelle(1), jdelhommelle@chem.und.edu, 151 Cornell Street Stop 9024, Grand Forks ND 58202, United States; Kyle D Watson(1); John Persson(1); Caroline Desgranges(1). (1) Chemistry, University of North Dakota, Grand Forks ND 58202, United States

We review our recent work on the molecular simulation of the crystallization of nanoparticles. The aim of this work is to obtain a complete understanding of the molecular mechanisms underlying crystal nucleation and growth, and, in particular, to shed light on the polymorph selection process. For this purpose, we carry out three different types of molecular simulation: (i) to determine the phase diagram of the simulated system, (ii) to simulate the crystal nucleation event and (iii) to gain a direct access to the crystal growth mechanism. We present results obtained on a variety of systems, ranging from model systems [1,2] to metal nanoparticles/nanoalloys [3] and to semiconductor nanoparticles.


COMP 384

Molecular and mesoscale mechanisms of Brittle Bone Disease (Osteogenesis imperfecta): A computational materiomics study

Markus J. Buehler(1), mbuehler@MIT.EDU; Markus Buehler(1). (1) Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge MA 02139, United States

Collagen is a crucial structural protein material, formed through a hierarchical assembly of tropocollagen molecules, arranged in collagen fibrils that constitute the basis for larger-scale fibrils and fibers. Osteogenesis imperfecta is a genetic disorder in collagen characterized by mechanically weakened tendon, fragile bones, skeletal deformities and in severe cases prenatal death. Even though many studies have attempted to associate specific
mutated types with phenotypic severity, the mechanisms by which a single point mutation influences the mechanical behavior of tissues at multiple length-scales remain unknown. Here we show by a hierarchy of full atomistic and mesoscale simulation that osteogenesis imperfecta mutations severely compromise the mechanical properties of collagenous tissues at multiple scales, from single molecules to collagen fibrils. Mutations that lead to the most severe osteogenesis imperfecta phenotype correlate with the strongest effects, leading to weakened intermolecular adhesion, increased intermolecular spacing, reduced stiffness, as well as a reduced failure strength of collagen fibrils (Gautieri et al., Biophysical J., 2009). Our findings provide insight into the microscopic mechanisms of this disease and lead to explanations of characteristic osteogenesis imperfecta tissue features such as reduced mechanical strength and lower cross-link density. Our study explains how single point mutations can lead to catastrophic tissue failure at much larger length-scales. The use of a materiomics approach to understand mechanisms of disease at multiple scales provides a powerful new avenue to develop fundamental insight into disease etiology and progression, and has the potential to impact our ability to treat severe genetic, infectious and neurodegenerative diseases.

COMP 385

Mimicking coarse-grained simulations without coarse-graining: Enhanced sampling by damping short-range interactions

Feng Wang(1); Dongshan Wei(1). (1) Department of Chemistry, Boston University, Boston MA 02215, United States

The damped short-range interaction (DSRI) method recently developed in our group is designed to mimic coarse-grained simulations by propagating an atomistic scale system on a smoothed potential energy surface while preserving the overall feature of the free energy landscape. The DSRI method has the benefit of enhanced sampling similar to a coarse grained simulation but does not require coarse-graining. Coupled with a scheme where the mass of various atoms are optimized to accelerate slow events, our method was used to simulate liquid water, alanine dipeptide, alanine penta-peptide in explicit water solvent, and the self-assembly of dimyristoylphosphatidylcholine lipid. In each case, our simulations indicate this method is capable of appreciably accelerating the underlying dynamics without significantly changing the free energy surface. Additional insights from DSRI simulation and the promise of coupling our DSRI method with replica-exchange molecular dynamics will be discussed.

COMP 386
Mechanism of sliding clamp opening by the clamp loader RFC: Insight from atomistic simulations

Ivaylo Ivanov(1), iivanov@gsu.edu, P.O. Box 4098, Atlanta Georgia 30302-4098, United States; J. Andrew McCammon(2); John A. Tainer(3). (1) Department of Chemistry, Georgia State University, Atlanta Georgia 30302-4098, United States (2) Department of Chemistry & Biochemistry, University of California, San Diego, United States (3) Department of Molecular Biology, The Scripps Research Institute, United States

PCNA is a ring-shaped protein (sliding clamp) that can encircle and slide along DNA. Recently, groundbreaking structural biology work has provided a glimpse into the workings of the cellular machinery responsible for opening and loading sliding clamps onto DNA and opened up the possibility of investigating clamp-loader assemblies in atomic detail. Herein we present results from atomistic simulations of the initial steps in the clamp-loading cycle. We have used molecular dynamics to obtain an atomistic model of the complex of the clamp loader RFC with an open clamp. The out-of-plane twisting of the open PCNA led to a right-handed helical conformation and the formation of an extended interface with RFC. We characterized this interface in terms of residue contacts, electrostatic and shape complementarity. Additionally, low-resolution elastic network modeling revealed the global motions of the PCNA/RFC assembly and provided insight into the camp-opening process.

COMP 387

Spatial heterogeneity in the temperature of a rectangular nanorod dragged across a surface

Alexander V Popov(1), popov@chemistry.gatech.edu, 901 Atlantic Drive, Atlanta Georgia 30332-0400, United States; Rigoberto Hernandez(1). (1) School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta Georgia 30332-0400, United States

The atomic oscillations of a nanoscale α-alumina rectangular rod dragged across a much larger surface made of the same material are investigated using reduced-dimensional models. The nonequilibrium distribution of the temperature among the atomic sublayers of this system has recently been reported by Hase and coworkers (J.Chem.Phys.122(2005)094713) using MD simulations. We perform a stochastic analysis of this phenomenon. The sublayers of the nanorod are treated as elements of a pseudo one-dimensional vibrational chain. Every such element is propagated by the temperature-ramped irreversible Langevin equation. A key concept of our model lies in the separation of the environment into two baths: the thermostated phonon gas which is responsible for the global energy relaxation and the local nonequilibrium bath which leads to the energy
dissipation within the chosen layer. This correctly captures the dynamics of energy transfer across the layers.

**COMP 388**

**Molecular dynamics simulations of nanoparticles and surfactants at oil/water interfaces: Chemisorption vs. physisorption**

*Steven O Nielsen*(1), steven.nielsen@utdallas.edu, 800 West Campbell Road, BE26, Richardson TX 75080, United States; *Chuong T Nguyen*(1); *Udayana Ranatunga*(1). (1) Department of Chemistry, University of Texas at Dallas, Richardson TX 75080, United States

In this study we use MD simulations to give physical insight into the interfacial properties of oil/water interfaces functionalized with nanoparticles (NPs) and chemisorbed or physisorbed amphiphilic ligands. We explore the synergistic action of NPs and surfactants to modify the interface properties. Implications as to the mechanical strength of colloidosomes are discussed.

**COMP 389**

**Aggregation dynamics of spherically-symmetric reactive colloidal particles**

*Matthew C Hagy*(1), hagy@gatech.edu, 901 Atlantic Drive, NW, Atlanta Georgia 30332-0400, United States; *Susan J Fitzwater*(2); *Thomas H Pierce*(2); *Mark Schure*(2); *Alexander V Popov*(1); *Rigoberto Hernandez*(1). (1) Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta Georgia 30332-0400, United States (2) The Dow Chemical Company, Spring House Pennsylvania 19477, United States

Despite the importance of colloidal aggregation in many industrial materials and biological systems, its dynamics remains poorly understood. The problem is particularly challenging because the dynamics span over 10 orders of magnitude in the time scale, raging from picoseconds for particle motion up to the macroscopic time scales of minutes for bulk aggregation. Their treatment therefore required specialized simulation codes using Brownian dynamics as the propagator and treating structure at multiple length scales. The results indicate an influence on the aggregation kinetics due to colloid volume fraction and energetic barrier strength. The early stages of aggregation are primarily determined by activated capture processes. These rates can be obtained directly from Kramers' rate theory. They scale with volume fraction and barrier strength. Later stages of aggregation involve reorganization into highly ordered domains. In this case, the kinetics are path dependent and cannot be modeled analytically.
COMP 390

Coarse-grained models for processes that involve large-scale membrane deformation

Qiang Cui(1), cui@chem.wisc.edu, Dept. of Chem., Univ. of Wisc. Madison, 1101 University Ave, Madison WI 53706, United States. (1) Chemistry, UW-Madison, Madison WI 53706, United States

I'll discuss our progress towards developing effective coarse-grained models for studying processes that involve large-scale membrane deformations, such as mechanosensation, action of antimicrobial materials and membrane fusion. We will discuss two classes of models: particle based models and continuum mechanics models. The former highlights the importance of properly treating electrostatics at the water/membrane interface, even at the coarse-grained level, while the latter helps emphasize the importance of considering boundary condition for studying membrane deformation in response to external chemical or mechanical perturbations.

COMP 391

Coarse-graining electrostatics in multiscale molecular simulations

Matteo Dal Peraro(1), matteo.dalperaro@epfl.ch, EPFL SV IBI1 UPDALPE, AAB 017, Station 15, Lausanne 1015, Switzerland; Michele Cascella(2); Davide Alemani(1); Enrico Spiga(1); Francesca Collu(2). (1) Institute of Bioengineering, Swiss Federal Institute of Technology, EPFL Lausanne, Lausanne 1015, Switzerland (2) Department of Chemistry, University of Bern, Bern, Switzerland

All-atom (AA) molecular dynamics (MD) is a powerful tool to investigate the structure and function of biomolecular systems, but it remains still computationally unaffordable to thoroughly sample size and time scales that are relevant to most of the biological processes. Coarse-grained (CG) schemes have been introduced to overcome these limits; nonetheless, the lack of universality and transferability still afflicts CG models and limits their general applicability to study the mechanism of relevant biological systems.

We have recently introduced a robust scheme to account for the intrinsic non-radial nature of backbone-backbone interactions in CG MD simulations of proteins [1]. A potential term mimics the backbone dipole-dipole interactions, and is able to naturally stabilize elementary secondary structure motifs, such as α-helices and β-sheets, and to modulate transitions to super-secondary structure assemblies. Moreover, the scheme, extended to side-chains and solvent, can be used to model electrostatic contributions in a multiscale (AA/CG)

Thus, this new scheme represents a promising step towards the development of CG force fields for proteins without additional biases on the secondary structure. This feature may contribute to advance our mechanistic understanding of the function of large protein assemblies and networks.


COMP 392

Understanding the cellulosome and its assembly with coarse-grain modeling: Toward improving the CBP process

Michael Crowley\(^{(1)}\), Michael.Crowley@nrel.gov, 1617 Cole Blvd., Golden CO 80401, United States; Yannick J. Bomble\(^{(1)}\), Mark R. Nimlos\(^{(2)}\), Michael E. Himmel\(^{(1)}\). (1) Chemical and Biosciences Center, National Renewable Energy Laboratory, Golden CO 80401, United States (2) National Bioenergy Center, National Renewable Energy Laboratory, Golden CO 80401, United States

The mechanism of assembly for the enzyme subunits of the natural scaffoldin is not currently known. In our study we focused on the cellulosome-integrating protein CipA of C. thermocellum and cellulosomal enzymes from families 5, 9 and 48. These three enzymes are representative of the variety (mass, volume, modularity) of enzymes secreted by C. thermocellum. The first coarse-grained model to study the formation and function of the cellulosome assembly was developed within CHARMM. This work aims at understanding the mechanisms involved in the sequential binding of the cellulosomal enzymes to the CipA scaffold of C. thermocellum. A large study of the effect of several key physical properties on binding to the scaffoldin protein is conducted. The modularity of the enzymes was found to be one of the main influences on the cellulosome assembly process.

COMP 393

Developing multiscale simulation models to study aggregation in peptide-based materials

Christine Peter\(^{(1)}\), peter@mpip-mainz.mpg.de, Ackermannweg 10, Mainz 55128, Germany. (1) Theory Group, Max Planck Institute for Polymer Research, Mainz 55128, Germany

Hierarchical simulation approaches systematically link simulation methods at several levels of resolution. One example are multiscale simulations that combine an all-atom and a coarse-grained description. They allow to address problems that require both accounting for chemically specific interactions and
processes on the high resolution level as well as the time and length scales that can only be achieved on the coarse-grained (mesoscopic) level. These hierarchical methods are ideally suited to study biomolecular aggregation, for example of small peptides - processes that play an important role both in biomedical as well as biomaterials applications. By systematically linking the simulation levels we can consistently switch back and forth between them. The latter can be achieved using an efficient backmapping procedure through which we obtain well equilibrated high-resolution structures of the aggregates formed. In order to keep the models on the different levels properly linked, it is of essential importance that they are both thermodynamically as well as structurally consistent, sample the same conformation space, etc. We will address these requirements and the challenges they pose to the development of reduced-resolution models.


[4] Conformational sampling in atomistic and coarse grained peptide models, O. Bezkorovaynaya, C. Peter, manuscript in preparation

COMP 394

Recent progress in multiscale molecular dynamics simulation of soft matter

Bernd Ensing(1), b.ensing@uva.nl, Nieuwe Achtergracht 166, Amsterdam NH 1018WV, The Netherlands; Preston B. Moore(2); Steven O Nielsen(3). (1) Department of Chemistry, University of Amsterdam, Amsterdam 1018WV, The Netherlands (2) Department of Chemistry and Biochemistry, University of the Sciences in Philadelphia, Philadelphia PA 19104, United States (3) Department of Chemistry, University of Texas at Dallas, Dallas TX 75083, United States

We present a new hybrid all-atom/coarse-grain (AA/CG) method, in which an active part of the system is treated in atomistic detail while the environment is included at a coarse-grain level. An intermediate healing region couples the two representations and allows particles that travel between them to adapt their AA/CG resolution on the fly. Although such coupling may work for simple liquids,
a second advancement is needed for hybrid simulation of more complex systems. This advancement is a reverse mapping technique that uses rigid body rotation dynamics to reintroduce the chemical details into a coarse-grain trajectory. The reverse mapping technique can be used in a hierarchical multiscale modeling approach, but here we incorporate it into the hybrid AA/CG method to precondition the atomic details of the CG representation and construct a robust hybrid multiscale method to model complex soft matter systems. Swelling of polyethylene in theta solvent serves as an illustration.

COMP 395

Monte Carlo simulations of biophysical systems:
The importance of quality of random number generators

George Kaminski(1), gkaminski@wpi.edu, 100 Institute Rd, GP 4019, Worcester MA 01609, United States. (1) Department of Chemistry and Biochemistry, Worcester Polytechnic Institute, Worcester Massachusetts 01609, United States

Monte Carlo simulations of molecular systems require applications of random number generators (RNG). We have used three RNGs in Monte Carlo simulations of pure liquid butane and methanol and hydrated tridecaalanine. A significant effect (such as, for example, a 24% or even 100% increase in the molecular volume) can occur if a wrong choice of the random number generator is made. At the same time, simple statistical tests do not necessarily lead to any warnings. We conclude that the optimal way of testing a new random number generator to be employed in Monte Carlo simulations is with molecular simulations of a well-studied system, such as pure liquid butane of methanol. It is also important to note that results of different Monte Carlo simulations (or Monte Carlo and molecular dynamics simulations) may contain non-trivial differences simply because of random number generator choices.

COMP 396

Developing multisite \(\lambda\)-dynamics for structure-based drug design

Jennifer L. Knight(1), jeknight@umich.edu, 910 N. University Ave., Ann Arbor MI 48109, United States; Charles L. Brooks(1)(2). (1) Department of Chemistry, University of Michigan, United States (2) Department of Biophysics, University of Michigan, United States

Free energy calculations are fundamental to obtaining accurate theoretical estimates of many important biological phenomena including hydration energies, protein-ligand binding affinities and energetics of conformational changes. We have been developing Multi-Site \(\lambda\)-dynamics simulations in which the
conventional "$\lambda$" is treated as a dynamic variable throughout the simulations and free energy differences are simultaneously evaluated for chemical variations at multiple modification sites on an identified molecular framework. Here, we present recent developments in Multi-Site $\lambda$-dynamics and discuss its application in computing relative solvation free energies and relative binding free energies for diverse series of inhibitors targeting HIV-1 reverse transcriptase, dihydrofolate reductase and $\alpha$-mannosidase.

COMP 397

A new computational method for modeling the solvation at interfaces: Bridging fluctuating hydrodynamics and molecular dynamics simulations

Jhih-Wei Chu$^{(1)}$, jwchu@berkeley.edu, 101A Gilman, UC Berkeley, Berkeley CA 94720, United States; Nikolaos Voulgarakis$^{(1)}$, Barry Shang$^{(1)}$. (1) Chemical Engineering, UC Berkeley, Berkeley CA 94720, United States

In this work, we demonstrate that the framework of fluctuating hydrodynamics can be applied to model the solvation at interfaces. First, we present a scheme that couples field variables and particle variables and how the coupled scheme can be used to simulate the dynamics of a system with mixed physics. The calculation of the solvation free energy of hydrophobic solutes with different sizes is used as an example. Second, we demonstrate that the fluctuating hydrodynamics equations can be directly applied to simulate liquid-vapor interfaces and the results predicted by gradient theories and the capillary wave theory can be reproduced. These results illustrate that the combined molecular dynamics and fluctuating hydrodynamics methodology can be used as a general framework for multiphysics simulations. Perspectives of applying this method to model the solvation and dynamics in complex molecular systems will also be discussed.

References:
Voulgarakis and Chu, JCP, 130, 2009, 134111
Voulgarakis, Satish, and Chu, JCP, 131, 2009, 234115

COMP 398

Novel method to control pressure in molecular dynamics simulations

D. Vladimir Perez$^{(1)}$, dperezveloz@mail.usp.edu, 600 South 43rd Street, Philadelphia Pennsylvania 19104, United States; Preston B Moore$^{(1)}$; Steven O Nielsen$^{(2)}$. (1) West Center for Computational Chemistry and Drug Design, Department of Chemistry & Biochemistry, University of the Sciences in Philadelphia, Philadelphia Pennsylvania 19104, United States (2) Department of Chemistry, University of Texas at Dallas, Richardson Texas 75083, United States
Pressure is a fundamental thermodynamic quantity and most processes occur at constant pressure. Many Molecular Dynamics (MD) simulations algorithms use the NPT ensemble (Constant Number of particles, Pressure, and temperature) for this reason. The usual computational way to control pressure in MD simulations is to use the virial expressions which involve modifications of the distances between all the particles in the simulation cell in order to emulate the volume fluctuation of real systems. For many systems this method is satisfactory, but in certain systems this method may be inadequate, e.g. the simulation of adsorption energy on a crystal surface where the lattice parameters must be kept constant. We present a new way to control the pressure by the use of a wall or piston that would compress or dilate the system (depending on the desired pressure) and keep the lattice spacing constant. We present results for confined Lennard-Jones systems and compare and contrast the current barostat methods.

COMP 399

Hydration effects on protein dynamics and function: A microscopic perspective

Donald Hamelberg\(^{(1)}\), dhamelberg@gsu.edu, P. O. Box 4098, Atlanta GA 30302-4098, United States; Quentin Johnson\(^{(1)}\); Urmi Doshi\(^{(1)}\). \(1\) Department of Chemistry, Georgia State University, Atlanta Georgia 30302-4098, United States

Water molecules make up an integral part of protein structures, providing stability, controlling the plasticity of binding interfaces, and assisting in catalysis. Protein motions are accompanied by the formation and breakage of hydrogen-bonding network of the surrounding water molecules. This ordering and reordering of water also adds to the underlying roughness of the energy landscape of proteins and thereby alters their dynamics. We have developed a model to extract the contribution of water to the dynamics and function of proteins from molecular dynamics simulations. We show that water contributes an additional roughness to the energy landscape of proteins. At lower temperatures this roughness, which becomes comparable to \(k_B T\), can considerably slow down protein dynamics. The broader implications of this effect for the function of some classes of enzymes will also be discussed, since the landscape topology of their substrates would change upon moving from an aqueous environment into the binding site.

COMP 400

Lambda hopping: An efficient replica exchange-based sampling method for free energy perturbation calculations
Yujie Wu(1), yujie.wu@schrodinger.com, 120 West 45th St., 17th Floor, New York NY 10036, United States; Teng Lin(1); John C Shelley(1); Woody Sherman(1). (1) Schrodinger, New York NY 10036, United States

Free energy perturbation (FEP) is an attractive method for calculating binding free energies. However, the accuracy depends strongly on the sampling convergence of the system, which is notoriously difficult for complex systems that have a rough free energy landscape (e.g., most receptor-ligand systems). To enhance sampling, and thus convergence, in FEP calculations we have developed a method, called lambda hopping, as an extension of that from Essex et al. (J. Phys. Chem. B 2003, 107, 13703) in which simulations for different lambda windows are switched by a Hamiltonian-mutation replica exchange. In particular, we show that the lambda schedule and the interaction mapping can be systematically devised such that major energy barriers are smoothed out for the intermediate lambda windows. We demonstrate that highly efficient sampling can be achieved with lambda hopping for both ligand-in-water and ligand-in-complex alchemical mutations.

COMP 401

Improving umbrella sampling with knowledge

Daniel J Sindhikara(1), sindhikara@gmail.com, Furo-cho, Nagoya Aichi 464-8602, Japan; Yuko Okamoto(1). (1) Department of Physics, Nagoya University, Nagoya Aichi 464-8602, Japan

Molecular simulations are significantly hindered by rough energy landscapes. Often, sampling is desired along a specific physical reaction coordinate (RC). Umbrella sampling can be used to enforce sampling along a reaction coordinate, but requires many simulations to sample the entire span of the RC. Here, we present a method to reduce the required number of simulations in umbrella sampling by using knowledge of the PMF landscape.

COMP 402

Balancing search, scoring, speed, and flexibility: AutoDock and AutoDock Vina

Arthur J. Olson(1), olson@scripps.edu, Mail Stop 5, 10550 N. Torrey Pines Rd., La Jolla Ca 92037, United States; Oleg Trott(1); Ruth Huey(1); Michael Pique(1); Stefano Fori(1); David S Goodsell(1). (1) Department of Molecular Biology, The Scripps Research Institute, La Jolla CA 92037, United States
This past year we released a new version of AutoDock, and introduced a new
docking program AutoDock Vina. Our AutoDock4.2 release highlights receptor
flexibility and methods for covalent docking. AutoDock's configuration of input
parameters for search and scoring enables a large degree of choice in
customizing docking.

AutoDock Vina was developed with the goal of simplified input and rapid
throughput. The
scoring function was generated using a machine-learning approach, training the
scoring function to reproduce both the binding modes and affinities of PDBBind.
Vina's search uses a Metropolis criterion to accept random mutations and a
quasi-Newton local optimization
using gradient information. Vina uses the same molecular file format as
AutoDock. However in Vina, unlike AutoDock, specification of search
parameters, clustering, and explicit calculation of grid maps are not exposed to
the user.

We have run re-docking, cross-docking and timing tests on both AutoDock and
AutoDock Vina. Results will be discussed.

COMP 403

Incremental gains and quantum leaps: A perspective on the development
of eHiTS

Orr Ravitz(1), ravitz@simbiosys.ca, 135 Queen[apos]s Plate Dr., Unit 520,
Toronto Ontario M9W 6V1, Canada ; Aniko Simon(1); Zsolt Zsoldos(1). (1)
SimBioSys Inc., Toronto Ontario M9W 6V1, Canada

The development of the docking program, eHiTS, and its native scoring function,
has been aimed at addressing the three fundamental challenges of the docking
paradigm: binding mode prediction, identification of active compounds, and
binding affinity evaluation. The chosen fragment- and statistically-based
approach has proven to be a fruitful framework, allowing great flexibility and
versatility in algorithm development, data handling, and implementation of
advanced computer science techniques. eHiTS' underlying principles of
exhaustiveness, accuracy, speed, simplicity, comprehensive experimental data
utilization, and adjustable scoring brought eHiTS to the forefront of the field, and
continue to guide the development efforts to this day. We give a retrospective
account of the evolution of eHiTS and explore the changes that gave rise to
major advances in accuracy and performance and other modifications that
yielded moderate improvements. We provide the overview through the prism of
the Astex and DUD benchmarks as well as other test cases.

COMP 404
Improvement of protein/ligand complexes using a new geometry optimization force field: A prelude to docking and scoring

Colin McMartin\(^{(1)}\), cmcma@ix.netcom.com, 603 Colebrook Rd., Colebrook CT 06021, United States. (1) Research, Thistle Soft, Colebrook CT 06021, United States

Accurate structure based potency prediction depends on accurate modeling of the interactions between ligand and protein.

This talk presents results obtained with tools developed using accurate data from high resolution PDB and small molecule crystal databases:

1. A force field (Crystal-Geo) for geometry optimization.


3. An accurate method for addition of explicit water to models.

The methods have been applied to PDB files (150 high resolution and 85 from the Aztex list plus 110 HIV-protease complexes).

After optimization many high resolutions structures show:

1. Geometry statistics conforming to HighRes-DB values

2. Reduction of ligand strain energy to low values

These changes only require small atomic movements (rmsd 0.1Å).

After optimization significant numbers of low resolution PDB files acquire geometry and energy characteristics of high quality PDB complexes.

COMP 405

Hybrid docking with FRED

Mark McGann\(^{(1)}\), mcgann@eyesopen.com, 222 3rd Street Suite 3120, Cambridge MA, United States. (1) OpenEye Scientific Software, Santa Fe NM 87508, United States

Structure-based virtual screening methods, i.e. docking programs, make use of the structure of a target protein to screen active molecules from a virtual database. Even though protein structures are commonly crystallized in the presence of a ligand, the ligand information is typically discarded by docking
programs. Similarly ligand-based virtual screening methods typically do not make use of protein structure information even when that information is readily available. FRED is a molecular docking program that is capable of using the information present in the structure of a bound ligand (if present), in addition to the structure of the protein, to enhance the virtual screening performance. This hybrid approach is shown to be superior to either pure ligand or pure structure based approaches.

COMP 406

Development of the Glide docking program and scoring function and evaluation of enrichment performance using the DUD data set

Richard A. Friesner(1), rich@chem.columbia.edu, 3000 Broadway, New York NY 10027, United States. (1) Department of Chemistry, Columbia University, New York NY 10027, United States

We will discuss the development of the Glide docking program over the past five years, and present an evaluation of the current performance of the program using the DUD data set. A significant number of new terms have been incorporated into the scoring function, including models for evaluating strain energy induced by rings in the ligand, improved evaluation of salt bridge interactions, and incorporating of active site water locations and free energies as compared to bulk solution as determined by the WaterMap program. We will additionally consider induced fit effects as such effects are essential in obtaining even reasonably accurate assessment of ligand potency in many if not most cases.

COMP 407

Docking and scoring in discovery studio

Allister J Maynards(1), amaynard@accelrys.com, 10188 Telesis Ct, Suite 100, San Diego CA 92121, United States; Jurgen Koska(1); Tedman Ehlers(1). (1) Accelrys, Inc., San Diego CA 92121, United States

This presentation covers a review of the docking tools available from Discovery Studio, including comparative results for both binding mode prediction and virtual screening. In particular CDOCKER is a CHARMM based docking routine for high-accuracy docking. Enhancements and automation of the CDOCKER will be discussed and comparisons made to other docking routines.

COMP 408
Application of free energy perturbation to drug discovery research: What have we learnt so far, and what's next?

Devleena Shivakumar(1), devleena.shivakumar@schrodinger.com, 120 West 45th St 17th floor, New York City NY 10036, United States; Woody Sherman(1).
(1) Schrodinger, Inc., New York City New York 10045, United States

The accurate prediction of relative binding free energy differences has been one of the primary objectives for computational methods since the inception of molecular modeling and computer-aided drug design. While progress has been made using empirical models, there are still substantial limitations. Free energy perturbation (FEP) is a method based on sampling and statistical mechanics to compute free energy differences between two molecules. While FEP was first described over 50 years ago, there have been limited applications in drug discovery and lead optimization due to the formidable computational costs, and difficulty in setting up the simulations. With recent advances in the efficiency of molecular dynamics programs, automation of the FEP-based calculations, and the continuing increase in available compute power it is now becoming feasible to run FEP in real drug discovery projects. In this work we describe a large-scale FEP project designed specifically to make FEP practical in drug discovery. We first describe validation of the underlying methodology and force field by computing absolute and relative solvation free energies of small diverse molecules, and comparing its performance with other conventional methods. We then present a large set of pharmaceutically relevant targets and compounds that are being used for the prediction of relative binding free energies. Finally, we present preliminary results and discuss the implications for FEP in drug discovery.

COMP 409

Binding site water in relative free energy calculations for drug-like compounds

John C. Shelley(1), John.Shelley@schrodinger.com, 101 SW Main St., Suite 1300, Portland Oregon 97204, United States; David Calkins(1); Arron P. Sullivan(1); Yujie Wu(2); Teng Lin(2); Byungchan Kim(2); Mee Y. Shelley(1); Robert J. Abel(2). (1) Schrodinger, Inc., Portland Oregon 97204, United States (2) Schrodinger, Inc., New York New York 10036, United States

Accurate prediction of the relative binding free energies for drug-like ligands in protein binding sites would have a significant impact on drug development. While such calculations should in principle give accurate predictions, many technical hurdles remain before this technology can enter routine use in predicting structure-activity relationships (SAR) in lead optimization projects. One primary
The challenge is the role of binding site water molecules, particularly those trapped in "dead-end" pockets into which ligand variations are to be explored. We will discuss our experiences with addressing this issue including the use of Grand Canonical Monte Carlo calculations to sample such water molecules in conjunction with FEP calculations.

COMP 410

Optimal π-stacking interaction energies in parallel-displaced aryl/aryl dimers are predicted by the dimer heavy atom count

John M Sanders\(^{1}\), john_sanders@merck.com, 770 Sumneytown Pike, PO Box 4; WP53F-301, West Point PA 19486, United States. (1) Department of Chemistry, Modeling & Informatics, Merck & Co., Inc., West Point PA 19486, United States

It is generally accepted that large basis sets are required for the accurate calculation of interaction energies for weakly bound aryl dimers, but it has recently been reported that MP2(full)/6-31G* energies, though inaccurate in absolute terms, are well-correlated with estimated CCSD(T)/CBS results. It is now shown that this correlation holds for MP2/aug-cc-pvdz and SCS-MP2/aug-cc-pvdz values. Linear regression of published CCSD(T)/CBS results with MP2 or SCS-MP2 results has been used to correct systematic errors observed with both MP2 theories, and these corrections are applied to 27 parallel-displaced aromatic dimers of interest in medicinal chemistry. The optimal computed interaction energies are found to be strongly correlated with the heavy atom counts of the aryl/aryl dimers. This relationship between heavy atom count and interaction energy also applies to a series of 14 aryl/non-aryl dimers such that a single linear regression equation accounts for all of the dimers studied.

COMP 411

Computational alanine scanning with linear scaling semi-empirical quantum mechanical methods

David J Diller\(^{1}\), djrdiller@gmail.com, 176 Hickory Corner Road, East Windsor NJ 08520, United States; Xiaohua Zhang\(^{2}\); Lance M Westerhoff\(^{2}\). (1) Consultant, East Windsor NJ 08520, United States (2) QuantumBio Inc, State College PA 16802, United States

Alanine scanning is a powerful experimental tool for understanding the key interactions in protein-protein interfaces. Linear scaling semi-empirical quantum mechanical calculations are now sufficiently fast and robust to allow meaningful calculations on large systems such as proteins, RNA and DNA. In particular, they have proven useful in understanding protein-ligand interactions. Here we ask the question: can these methods developed for protein-ligand scoring be useful for
computational alanine scanning? To answer this question, we assembled 15 protein protein complexes with available crystal structures and sufficient alanine scanning data. We show that with only one adjusted parameter the quantum mechanics based methods out perform both buried accessible surface area and a potential of mean force and compare favorably to a variety of published empirical methods. Finally, we closely examined the outliers in the data set and discuss some of the challenges that arise from this examination.

COMP 412

Computational insight into the specific inhibition of cyclophilins A and B

Orlando Acevedo(1), orlando.acevedo@auburn.edu, 179 Chemistry Building, Auburn AL 36849, United States; Somisetti V Sambasivarao(1). (1) Department of Chemistry and Biochemistry, Auburn University, Auburn AL 36849, United States

Cyclophilins (Cyp) are a family of cellular enzymes possessing peptidyl-prolyl isomerase activity which catalyze the cis-trans interconversion of peptide bonds amino-terminal to proline residues. Of specific interest are human CypA and CypB which have been identified as valid drug targets for hepatitis C virus (HCV), inflammatory diseases, and multiple cancers. However, current treatments rely on global cyclophilin inhibitors, such as cyclosporin A (CsA), that elicit large side effects. Compounds that selectively bind with CypA or CypB are urgently needed. Computer-aided drug development is a viable method for delivering such clinical candidates. Atomic-level computer models of the proteins have been constructed from high-resolution crystal structures. Free energy perturbation (FEP) calculations in conjunction with Monte Carlo (MC) statistical mechanics simulations have been utilized to reproduce the experimental binding affinities of potent acylurea- and aryl 1-indanylketone-based inhibitors and provide an origin for the basic differences in binding and potency of the compounds.

COMP 413

Structural promiscuity of human cytochromes P450: Understanding drug-drug interactions and metabolism via in silico fragment mapping

Raeanne L. Napoleon(1), raeanne@bu.edu, 590 Commonwealth Ave, Boston MA 02215, United States; Chi Ho Ngan(2); Dima Kozakov(2); Sandor Vajda(1)(2). (1) Department of Chemistry, Boston University, Boston MA 02215, United States (2) Department of Biomedical Engineering, Boston University, Boston MA 02215, United States

The interplay between structure and function in proteins, and in protein-ligand complexes, is essential for biological activity. As such, we
utilize robust computational techniques to probe the molecular surface of human cytochromes P450 (CYP) in combination with fragment based docking methods to identify druggable hot-spots. Identification of these hot-spots is critical for the structure-based design methods of pharmaceuticals. Here we present findings of the application of our new algorithms for the CYP 3A4 and 2A6. These particular isoforms catalyze the metabolic clearance of a large number of clinically used drugs, it is the most abundant protein used in the metabolism of xenobiotics, and are the most abundant protein found in the liver. We employ a computational fragment mapping program (FTMap) that identifies fragment binding hot-spots. This method includes a rigid body docking algorithm that is based on the Fast Fourier Transform correlation approach and has been extended to use pairwise interaction potentials. Our results show that these techniques provide metabolic and drug-drug interaction information that would be advantageous for selecting potential drug candidates and for insight into the protein structure-function relationship.

COMP 414

Geolsosteres: Structure-based approach to finding bioisosters using geometric and chemical patterns of interacting atoms at receptor-ligand interfaces

Raed S. Khashan(1)(2), raed@email.unc.edu, 301 Beard Hall, CB# 7360, Chapel Hill NC 27599, United States; Weifan Zheng(2); Alexander Tropsha(1). (1) Division of Medicinal Chemistry & Natural Products, University of North Carolina at Chapel Hill, Chapel Hill NC 27599, United States (2) Department of Pharmaceutical Sciences, North Carolina Central University, BRITE Institute, Chapel Hill NC 27599, United States

Bioisosteric compounds have different chemical functional groups but similar bioactivity, i.e., bioisosters is typically defined via comparing chemical structures of ligands. We introduce a novel, structure based approach to defining bioisosters as chemical fragments of ligands that interact with binding sites with similar geometrical and chemical features. We have developed a library of bioisosters based on x-ray characterized protein-ligand complexes. The interacting atoms at the interface of receptor-ligand complexes were identified with computational geometry approach termed Almost Delaunay tessellation. Interacting atomic patterns at the protein-ligand interface were stored in a special database where both ligand and receptor components of each pattern were annotated. Bioisosters for a particular chemical group can be found by searching this database for other chemical groups interacting with similar patterns of receptor atoms. We will present examples of unusual bioisosters and discuss potential
application of our approach. (Check http://www.unc.edu/~raed for Software).

COMP 415

Energy triplets for writing epigenetic marks: QM/MM free energy simulations of protein lysine methyltransferases

Hong Guo\(^{(1)}\), hguo1@utk.edu, 1414 Cumberland Ave, Knoxville TN 37996, United States. (1) Biochemistry & Cellular & Molecular Biology, University of Tennessee, Knoxville TN 37996, United States

The free energy profiles are obtained from quantum mechanical/molecular mechanical (QM/MM) free energy simulations for the first, second and third methyl transfers in different protein lysine methyltransferases (PKMTs) and their mutants. The results of the simulations suggest that the relative efficiencies of the chemical steps involving the three methyl transfers in PKMTs from S-adenosyl-L-methionine to the ε-amino group of the target lysine may determine how the epigenetic marks of lysine methylation are written. Two different energy triplets are proposed as important parameters for the prediction of product specificity.

COMP 416

Multiscale quantum modeling of mechanisms of ribozyme catalysis

Darrin M. York\(^{(1)}\), york@umn.edu, 207 Pleasant St. S.E., Minneapolis MN 55455, United States. (1) Chemistry, University of Minnesota, Minneapolis MN 55455, United States

Multiscale quantum models are used to study mechanisms of catalysis for several prototype ribozymes, including the hairpin and hammerhead ribozyme. These system are compared in terms of their metal ion dependence, conformational variations, pKa values, free energy profiles and chemical mechanisms. Simulation results are compared with crystallographic and biochemical experiments in order to provide a more clear mechanistic interpretation. New semiempirical and density-functional quantum models are discussed, as well as novel methods for calculation of free energy profiles. These
methods extend the arsenal of accurate theoretical tools that can be applied to study complex biocatalysis problems.

COMP 417

Nanoscale effects on heterojunction electron gases in core/shell nanowires

Bryan M Wong\(^{(1)}\), bmwong@sandia.gov, MS 9403, PO Box 969, Livermore CA 94551, United States; Francois Leonard\(^{(2)}\). (1) Materials Chemistry Department, Sandia National Laboratories, Livermore CA 94551, United States (2) Materials Physics Department, Sandia National Laboratories, Livermore CA 94551, United States

The unique properties of semiconducting heterostructure nanowires hold great promise for their incorporation in next-generation transistors, circuits, and nanoscale devices. The reduction in dimensionality produced by confining electrons in these heterostructure nanowires results in a dramatic change in their electronic structure, leading to novel properties such as ballistic transport and conductance quantization. One area of particular interest is in the formation of heterojunction electron gases in III-nitride core/shell nanowires which may provide a route towards quasi-one-dimensional electron gases.

In order to tailor these nanostructures with the desired physical properties, we must first understand their electronic properties as a function of size and material composition. To this end, we developed a self-consistent Poisson-Schrodinger approach to calculate the properties of heterojunction electron gases in polar and non-polar AlGaN/GaN core-shell nanowires. We find that the nanoscale size of these wires leads to the appearance of quasi-one-dimensional electron gases at the corners of the hexagonal and triangular cross-sections, in contrast to what would be expected from analogy with bulk heterojunctions. Our results allow a guided understanding of low-dimensional electron gas formation in freestanding semiconductor heterostructure nanowires.
COMP 418

Simulations of highly reactive complex chemical processes via an adaptive, multilevel QM/MM technique

Michael R. Salazar(1), msalazar@uu.edu, 1050 Union University Dr., Jackson TN 38350, United States. (1) Department of Chemistry, Union University, Jackson TN 38350, United States

The complexity associated with highly reactive chemical systems at high temperatures and pressures is particularly well suited for multiscale modeling techniques. In these systems, there may indeed be 100s-1000s of elementary steps connecting the starting fuels with the final products with an exponential growth in the intermediate species formed with the number of elementary steps. In addition, the whole reactive process may take nano- to microsecond time scales for completion with elementary steps on the picosecond order. Thus, these complex reactive processes involve multiple multiscale problems: system size, reaction times, and formed chemical species, not to mention the diversity in the collisional processes and the resulting diversity in the levels of quantum mechanical/molecular mechanical (QM/MM) theory necessary to accurately and rapidly calculate such diverse processes.

An adaptive, multilevel QM/MM methodology has been formulated to perform simulations of such complex reactive processes and recent progress will be reported. The method breaks the total potential into a multibody expansion of spatially-resolved time-dependent groups over which multiple levels of QM theory may be employed. Linear computational scaling with system size has been demonstrated for the method, while employing high levels of QM theory. Such a method, when applied to the complexity of highly reactive systems, is capable of elucidating the intricacies of the chemistry and the overall kinetics from a “bottom-up” dynamics point of view that employs high levels of QM theory.

COMP 419

Utilizing organic syntheses and microbial iron assimilation processes for the development of new antibiotics
Pathogenic microbes have rapidly developed resistance to all known antibiotics. To keep ahead in the "microbial war," extensive interdisciplinary effort is needed. Resistance develops primarily to overuse of antibiotics that can result in alteration of microbial permeability, alteration of drug target binding sites, induction of enzymes that destroy antibiotics (ie, beta-lactamases) and even cause efflux of antibiotics. A combination of chemical syntheses, microbiological and biochemical studies will demonstrate that the known critical dependence of iron assimilation by microbes for growth and virulence can be exploited for the development of new approaches to antibiotic therapy. Iron recognition and active transport relies on the biosyntheses and use of microbe-selective iron chelating compounds called siderophores. Our studies demonstrate that siderophores and analogs can be used for

- Iron transport-mediated drug delivery ("Trojan Horse").
- Induction of iron limitation (Development of new agents to block microbial iron assimilation).
- Converting microbe-induced chemistry of iron into a process that is lethal to microbes.

COMP 420

Utilization of bacterial iron transport systems for drug delivery

Ute Moellmann(1), ute.moellmann@googlemail.com, Beutenbergstrasse 11a, Jena Thuringia D-07745, Germany ; Lothar Heinisch(1). (1) Department of Molecular and Applied Microbiology, Leibniz Institute for Natural Product Research and Infection Biology - Hans Knoell Institute, Jena D-07745, Germany

The outer membrane permeability barrier is an important resistance factor of bacterial pathogens. In combination with other factors like drug inactivating enzymes, target alteration and efflux, it can increase resistance dramatically. A strategy to overcome this membrane mediated resistance is the misuse of bacterial transport systems. Most promising systems are those for iron transport. They are
vital for virulence and survival of bacteria in the infected host, where iron depletes is a defense mechanism against invading pathogens. We synthesized biomimetic siderophores as shuttle vectors for active transport of antibiotics through the bacterial membrane. Structure activity relationship studies resulted in ampicillin siderophore conjugates highly active against *Pseudomonas aeruginosa* and other Gram-negative pathogens, which play a crucial role in destructive lung infections in cystic fibrosis patients and in severe nosocomial infections. The mechanism of action, *in vitro* and *in vivo* efficacy were demonstrated.

**COMP 421**

**Activity of BAL30072, a novel siderophore sulfactam**

*Malcolm G P Page*(1), malcolm.page@basilea.com, PO Box, Basel BS CH-4005, Switzerland . (1) Basilea Pharmaceutica International Ltd, Basel CH-4005, Switzerland

BAL30072 is a monocyclic β-lactam antibiotic belonging to the sulfactams. BAL30072 showed potent activity against multidrug-resistant (MDR) *Pseudomonas aeruginosa* and *Acinetobacter* spp., including many carbapenem-resistant strains. BAL30072 was bactericidal against both *Acinetobacter* spp. and *P. aeruginosa*, even against strains that produced metallo-β-lactamases that conferred resistance to all other β-lactams tested, including aztreonam. It was also active against many species of MDR Enterobacteriaceae, including isolates that had a class A carbapenemase or a metallo-β-lactamase. Unlike other monocyclic β-lactams, BAL30072 was found to trigger spheroplasting and lysis of *E. coli*, rather than the formation of extensive filaments. The basis for this unusual property is its inhibition of the bifunctional penicillin-binding proteins PBP 1a and PBP 1b in addition to its high affinity for PBP 3, which is the target of monobactams such as aztreonam.

**COMP 422**

**Targeting bacterial multidrug efflux pumps**

*Olga Lomovskaya*(1), olomovskaya@mpexpharma.com, 11535 Sorrento Valleeu Road, San Diego California 92037, United States ; Scott Hecker*(1). (1) Mpex Pharmaceuticals, San Diego California 92037, United States

Powerful techniques of modern drug discovery such as comparative genomics, ultra-high-throughput screening, structure-guided drug design and combinatorial chemistry have been used to identify novel targets and optimize novel, preferentially broadspectrum antibiotics to combat antibiotic resistance. However, despite the fact that these employed targets are broadly conserved in bacteria, no drug candidate advanced using these methods has demonstrated relevant activity against most gram-negative bacteria. Thus, the outlook for new
antibiotics appears unchanged from present in that of all approved classes of antibiotics, representatives of only three classes (fluoroquinolones, β-lactams and aminoglycosides) have clinical utility for the treatment of gram-negative bacteria such as *Pseudomonas aeruginosa*.

Multidrug resistance (MDR) efflux pumps play a prominent and proven role in gram-negative intrinsic resistance. Moreover, these pumps also play a significant role in acquired clinical resistance. Together, these considerations make efflux pumps attractive targets for inhibition in that the resultant efflux pump inhibitor (EPI)/antibiotic combination drug should exhibit increased potency, enhanced spectrum of activity and reduced propensity for acquired resistance. To date, at least one class of broad-spectrum EPI has been extensively characterized. While these efforts indicated a significant potential for developing small molecule inhibitors against efflux pumps, they did not result in a clinically useful compound. Stemming from the continued clinical pressure for novel approaches to combat drug resistant bacterial infections, a second-generation programs have been initiated based on a number of recent developments in the field, including structural elucidation of all three individual components of MDR efflux pumps and ligand-based insights into the mechanism-of-action of drug transporters. Building upon previous efforts, these new approaches show early promise to significantly improve the clinical usefulness of currently available and future antibiotics against otherwise recalcitrant gram-negative infections.

**COMP 423**

**Interaction of β-peptides with membranes**

*Jagannath Mondal*(1), *jmondal@wisc.edu*, 1101 University Avenue, Madison WI 53706, United States; *Xiao Zhu*(1); *Qiang Cui*(1); *Arun Yethiraj*(1). (1) Department of Chemistry, UW Madison, Madison WI 53706, United States

A new class of anti-microbial agents named β-peptides have recently been reported that show interesting sequence dependent activity and selectivity. In this work we investigate the interaction of these molecules with a model membrane in an effort to obtain physical insight into the mechanism of anti-microbial activity. We investigate the effect of sequence on the adsorption of these β-peptides to a membrane using computer simulations with both implicit and explicit solvent and membrane. Two classes of molecules are investigated: 10-residue oligomers of 14-helical sequences, and four sequences of random co-polymeric β-peptides. The oligomers of interest are two isomers, globally amphiphilic (GA) and non-GA, of two 10-residue 14-helical sequences. The penetration of the molecules into the membrane and the orientation of the molecules at the interface depend strongly on the sequence. We attribute this to the propensity of the β-phenylalanine (βF) residues for membrane penetration. The membrane adsorption studies are consistent with potential of mean force calculations using the same model. Results are similar when the membrane and solvent are treated
in an implicit or explicit fashion. For the four sequences of random-co-polymeric β-peptides, the extent of stabilization of free-energy correlates with their efficiency to segregate the hydrophobic and cationic residues. The simulations are in qualitative accord with experiments on the minimum inhibitory concentration, and suggest simple strategies for the design of candidates for anti-microbial beta-peptides.

COMP 424

Molecular modeling of beta-lactamase inhibitors

Sookhee Nicole Ha(1), sookhee_ha@merck.com, One Merck Drive, Whitehouse Station New Jersey 08889, United States; T. Blizzard(1); H. Chen(1); S. Kim(1); J. Wu(1); K. Young(1); Y. Park(1); A. Ogawa(1); S. Raghoobhar(1); R. Painter(1); N. Hairston(1); S. Lee(1); A. Misura(1); T. Felcetto(1); P. Fitzgerald(1); N. Sharma(1); Jun Lu(1); E. Hickey(1); J. Hermes(1); M. Hammond(1). (1) Merck & Co., Inc, Whitehouse Station New Jersey 08889, United States

Resistance against new antibiotics usually appears within few years after their marketing. Expression of the beta-Lactamase is the most common mechanism of resistance to the beta-Lactam antibiotics in Gram-negative bacteria. To maximize delaying the drug resistance, we have developed a beta-Lactamase inhibitor for combination therapy. We report our efforts on optimization of bridged mono-bactam analogs.

COMP 425

Assembly and function of large Gram-negative bacterial machines studied by molecular simulation integrated with experimental data

Matteo Dal Peraro(1), matteo.dalperaro@epfl.ch, EPFL SV IBI1 UPDALPE, AAB 017, Station 15, Lausanne 1015, Switzerland. (1) Institute of Bioengineering, Swiss Federal Institute of Technology, EPFL Lausanne, Lausanne 1015, Switzerland

Gram-negative bacteria have evolved several means to attack their hosts and defend themselves from external attacks. Here, we use molecular simulations closely integrated with new experimental data to dissect the structural and dynamic features of the assembly mechanism of three large bacterial machines.

(i) We propose a four-helix model of E.coli PhoQ two-component system transmembrane domain, which is consistent with new experimental cross-linking data, and can explain the bacterial response to divalent cations and antimicrobial peptides. (ii) We study, with the aid of site-directed mutagenesis, the role of the pore-forming loop and the C-terminal pro-peptide for the heptamerization of pore-forming toxin aerolysin from A.hydrophila. Finally, (iii) we model the
needle formation and regulation for the type III secretion system from Y. enterocolitica (injectisome) based on fresh genetic and mutagenesis results.

The full comprehension of the structural assembly of these bacterial machines can contribute, on one side, to unveil their fundamental biological function, and, on the other, will permit to develop rational strategies to specifically interfere with them for therapeutic intervention.

COMP 426

Design of potent, broad-spectrum AccC inhibitors

Li Xiao\(^{(1)}\), li.xiao@sproc.com, 2015 Galloping Hill Road, Kenilworth New Jersey NJ 07033, United States; Cliff Cheng\(^{(2)}\); Gerald W Shipps\(^{(2)}\); Aileen Sonano\(^{(1)}\); Peter Orth\(^{(1)}\); Todd Black\(^{(1)}\). \(^{(1)}\) Merck Research Laboratory, Kenilworth New Jersey NJ 07033, United States \(^{(2)}\) Merck Research Laboratory, Cambridge Massachusetts MA 02139, United States

The biotin carboxylase (AccC) is part of the multi-component bacterial acetyl coenzyme-A carboxylase (ACCase) and is essential for pathogen survival. We identified and validated AccC as an antibacterial drug target for our in-house AS/MS screen. An initial hit, 2-(2-chlorobenzylamino)-1-(cyclohexylmethyl)-1H-benzo[\(d\)]imidazole-5-carboxamide \((1)\), was identified, and x-ray crystallography and computer modeling were utilized in its optimization. In this presentation we report our biology, chemistry and structure based drug design efforts in discovering a novel series of AccC inhibitors, exemplified by \((R)-2-(2\text{-chlorobenzylamino})-1-(2,3\text{-dihydro-1H-inden-1-yl})-1H\text{-imidazo[4,5-}b\text{]pyridine-5-carboxamide (2). These inhibitors are potent and selective for bacterial AccC with good cell-based activity against a sensitized strain of \(E. coli\) (HS294 \(E. coli\)).

COMP 427

Application of the Nudged Elastic Band method to the Su-Schrieffer-Heeger model reveals a new conduction mechanism in trans-polyacetylene

Andre L. Botelho\(^{(1)}\), abotelho@alum.bu.edu, 15 St Mary\’s St, Brookline MA 02446, United States; Xi Lin\(^{(1)}\); Minghai Li\(^{(1)}\). \(^{(1)}\) Division of Materials Science and Engineering, Boston University, Boston MA 02215, United States

Charges in the backbone of conductive polymers are trapped in self-localized defects known as solitons, polaron, and bipolarons. To study the mechanisms of electrical conduction, various studies have calculated the energy barriers of soliton migration by assuming the transition state is when the soliton is halfway between two dopant ions. Without the need for this
assumption, a Nudged Elastic Band calculation, using the Su-Schrieffer-Heeger model for trans-polyacetylene, confirms this migration mechanism when a single soliton is considered. Expanding the system to include the interaction between two neighboring solitons, we find a new migration mechanism in which the two solitons go through a locally stable bipolaron state. This mechanism has a lower activation energy than the single soliton motion at the light doping levels studied.

COMP 428

Computational high-throughput screening of sorbents for reversible CO$_2$ capture at high temperature

Ujjal Das$^{(1)}$, ujjal@mit.edu, 77 Massachusetts Avenue, 66-264, Cambridge, Massachusetts 02139, United States; William H Green$^{(1)}$. (1) Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States

Development of efficient computer screening methods has led to the discovery of many new materials. In this work, we present a density functional theory-based high-throughput screening scheme to identify sorbents for carbon dioxide capture from syngas at high temperature. High temperature CO$_2$ capture is particularly important because this process is more energy efficient. We have adopted multi-step screening method. At first we have computed reaction free energy of CO$_2$ in more than 300 reactions. While a large majority of these reactions are not suitable for reversible CO$_2$ capture, we have also identified multiple reactions with favorable reaction thermodynamics. Further screening based on CO$_2$ adsorption kinetics and sorbent capacity sorts out a few promising candidates that can be tested in experimental set up. The CO$_2$ capture ability of one such sorbent is currently under study in our laboratory.

COMP 429

Excited-state properties of dye-sensitized solar cells and light-harvesting molecules

Bryan M Wong$^{(1)}$, bmwong@sandia.gov, MS 9403, PO Box 959, Livermore CA 94551, United States; Joseph G Cordaro$^{(1)}$. (1) Materials Chemistry Department, Sandia National Laboratories, Livermore CA 94551, United States

Dye-sensitized solar cells (DSCs) have gained immense interest in the last few years due to their potential for converting clean solar energy to electricity at low cost. In particular, current research is now directed towards organic dye sensitizers which are less expensive and easier to synthesize. In order to develop these highly-efficient sensitizers, it would be extremely useful and cost-
COMP 430

Computational study of electronic states of monomer units of epoxy polymers

Uros T. Novakovski\textsuperscript{1}, utasic@gmail.com, 2300 Sixth Street, N.W, Washington DC 20059, United States; Steven L. Richardson\textsuperscript{1}; Stephen Christensen\textsuperscript{2}. (1) Department of Electrical and Computer Engineering, Howard University, Washington DC 20059, United States (2) The Boeing Company, Seattle WA 98124, United States

Potential for chemical degradation of polymers within carbon composites, is of critical importance to the aviation safety. Significant risk of exposure to ultraviolet radiation of composites that are a major structural component of the new generation planes, warrants our investigation of excited electronic states of these materials. This work focuses on small monomer units of amine-epoxy-based polymers, and on their ultraviolet light absorption as a possible initiation step in the chemical degradation. We model a range of ethanolamine-derived molecules, and probe the effects of chemical substitution and geometry conformation on the electron density distribution in the ground versus excited electronic states. Our goal is to map the energy profiles, the electronic and vibrational characteristics, and optical transition probabilities of excited electronic states, primarily the two
lowest lying ones. We also explore changes in the molecular geometry following electronic excitation, in the attempt to elucidate likely dissociation pathways.

COMP 431

Migration mechanism of self-localized defects in conductive polymers

Yongwoo Shin(1), yshin@bu.edu, 15 saint Mary's street, Brookline MA 02446, United States; Xi Lin(1). (1) Division of Materials Science and Engineering, Boston university, Brookline MA 02446, United States

Conductive polymers have been widely studied in the last couple of decades. The fundamental factor of charge transport is known to be the migration of self-localized defects such as the soliton, polaron, and exciton. In our previous work, we searched for the activation energy barrier of the motion of these defects in trans-polyacetylene (t-PA), which we found to be on the order of 0.01 eV for a soliton and 0.001 eV for a polaron using ab initio Hartree-Fock (HF) calculations. Nevertheless, the mechanism of migration of defects has yet to be clearly explained theoretically. In this project we present the migration mechanism of defects through the Goldstone mode, which is a break in the translational symmetry, as well as the other localized modes. Finally, the energy landscape of conductive polymers, restricted by the localized modes, is presented by the tight-binding method as well as HF calculations. Our systems included not only t-PA but also poly-phenylacetylene.

COMP 432

Asymmetric deposition of Pt electrodes on (001) ferroelectric surfaces

Miguel Angel Méndez Polanco(1), mpolanco@sas.upenn.edu, The Makineni Theoretical Laboratories, Philadelphia Pennsylvania 19104-6323, United States; Ilya Grinberg(1); Andrew M. Rappe(1). (1) Department of Chemistry, University of Pennsylvania, Philadelphia Pennsylvania 19104-6323, United States

Ferroelectrics materials posses a spontaneous bulk polarization even in the absence of an external field and its magnitude and direction can be varied by applying an external electric field. This possibility for polarization-switching can be utilized in various applications but it has been of particular interest in non-volatile memory applications of FE thin films. The basic setup for such devices consists of a ferroelectric oxide sandwiched between metal electrodes and consequently, two interfaces are formed that influence their overall performance. Due to their relevance in technological applications one important subject of study is that of understanding the existence of a critical size for the stabilization of Ferroelectricity in ultra-thin sized films of perovskite oxides to foster the
development of devices that satisfy the actual miniaturization requirements without sacrificing optimal performance.

Using Density Functional Theory we have studied the effect of symmetric and asymmetric metal film deposition on the Ferroelectric stability of (001) PbO-terminated PbTiO$_3$ thin films. We found that the asymmetric deposition of the Pt layers can stabilize a ferroelectric phase in theoretically unstrained ultra-thin films. Also, the Pt electrode deposition is influenced by the Polarization of the oxide surface such that for a Pt Multilayer capacitor: P$^-$ favors the Top arrangement of the interfacial Pt atoms whereas for the P$^+$ surface the preference is switched to Hollow deposition thus, evidencing the dependence of the Oxide/Metal interactions at the interface on the chemical environment.

This preferential deposition can be understood by means of the Pt-O bonding that is favored in the P$^-$ surface by placing Pt on top of O and Pb and, on the other hand the stabilization of the ferroelectricity is enhanced by the possibility for further off-center ionic displacements (increasing rumpling values) when the Pt layers occupy hollow sites on the P$^+$ surface. These two effects participate cooperatively rendering this particular capacitor arrangement the most stable.

We will discuss the effect of the interfacial environment by contrasting these results with other ABO$_3$ perovskite thin films as well as the effect of the in-plane strain on the stability of the asymmetric capacitors.

**COMP 433**

**Exploring protein conformational changes with accelerated molecular dynamics in NAMD**

**Yi Wang**(1), yiwang@ucsd.edu, 9500 Gilman Dr, La Jolla CA 92093-0365, United States; J. Andrew McCammon(1). (1) Chemistry and Biochemistry, Howard Hughes Medical Institute, University of California, San Diego, La Jolla CA 92093, United States

Accelerated molecular dynamics (aMD) enhances conformational space sampling by reducing energy barriers separating different states of a system. Here we present the implementation of aMD in the highly efficient parallel molecular dynamics program NAMD and offer exemplary applications performed on systems up to 60,000 atoms. Our results indicate that while providing significantly enhanced sampling, aMD simulations have only a small overhead in comparison to classical MD simulations. A 10-ns aMD simulation performed on the bacterial enzyme RmlC successfully revealed its transition from apo- to holo-state, which is not observed in a 50-ns classical MD simulation. We demonstrate that aMD can be applied efficiently to explore the conformational changes of complex biomolecules, especially when little is known about their alternative structures and transition reaction coordinates.
COMP 434

Pseudo-chair conformation of carboxyphosphate

Venkata S Pakkala\textsuperscript{(1)}, pakkalav@duq.edu, 600 Forbes Avenue, Pittsburgh p 15282, United States; Steven M Firestine\textsuperscript{(2)}; Jeffrey D Evanseck\textsuperscript{(1)}. (1) Department of Chemistry and Biochemistry, Duquesne University, Pittsburgh Pennsylvania 15282, United States (2) Eugene Applebaum College of Pharmacy and Health Sciences, Wayne State University, Detroit Michigan 48201, United States

For over 40 years, carboxyphosphate has been postulated as a key intermediate in several carboxylase enzymes. Unfortunately, this compound is extremely unstable (\(t_{1/2}\) of 70 ms), thus precluding direct experimental studies. Therefore, we have utilized high level ab initio (MP2 and CCSD(T)), DFT (B3LYP, BB1K, M05-2X, M06-2X and MPW1K) and ONIOM(DFT:AMBER) methods to investigate the structure and energetics of carboxyphosphate in vacuum, in a PCM continuum solvation model and in the active site of N5-CAIR synthetase, an enzyme shown to proceed via the formation of carboxyphosphate. We report here, for the first time, that carboxyphosphate adopts a “pseudo-chair” conformation and calculations reveal that this conformation is found to be the most stable in vacuum, solvent and the active site. This study has implications in the development of the carboxyphosphate analogs as potential inhibitors, in understanding the instability of the compound, and in elucidating the mechanisms of enzymes utilizing this compound.

COMP 435

Analysis of vibrational spectra of polypeptides in terms of localized vibrations

Christoph R Jacob\textsuperscript{(1)(2)}, christoph.jacob@kit.edu, Wolfgang-Gaede-Str. 1a, Karlsruhe, Germany; Markus Reiher\textsuperscript{(2)}. (1) Center for Functional Nanostructures, Karlsruhe Institute of Technology (KIT), Karlsruhe 76131, Germany (2) Laboratorium für Physikalische Chemie, ETH Zurich, Karlsruhe 76131, Germany

While nowadays efficient quantum chemical methods allow for the calculation of vibrational spectra of large (bio-)molecules, such calculations also provide a large amount of data. In particular for the vibrational spectra of polypeptides, a large number of close-lying normal modes contribute to each of the experimentally observed bands, which hampers the analysis of the calculated spectra considerably.

Here, we discuss how vibrational spectra obtained from quantum chemical calculations can be analyzed by transforming the calculated normal modes contributing to a certain band in the vibrational spectrum to a set of localized
modes [1]. We demonstrate that these localized modes are more appropriate for the analysis of calculated vibrational spectra of polypeptides and proteins than the delocalized normal modes.

We apply this methodology to investigate the influence of the secondary structure on infrared and Raman spectra of polypeptides [2]. As a model system, a polypeptide consisting of twenty (S)-alanine residues in the conformation of an α-helix and of a 3_10-helix is considered. In particular, we show how the use of localized modes facilitates the analysis of the positions and of the total intensities of the bands in the vibrational spectra, and how the couplings between localized modes determine the observed band shapes. Finally, this analysis is applied to analyze the Raman optical activity (ROA) spectra of these helical polypeptides, which provides a detailed picture of the generation of ROA bands in proteins [3].


COMP 436

Conformational coupling between LOV and kinase domains in phototropins: A computational perspective

Marco Stenta(1), marco.stenta@epfl.ch, EPFL SV IBI1 UPDALPE, Office AAB019 (Bâtiment AAB) Station 15, Lausanne Vaud 1015, Switzerland ; Matteo Dal Peraro(1). (1) Department of Bioengineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne Vaud 1015, Switzerland

Phototropins constitute an important class of plant photoreceptors playing key roles in many physiological responses to light, including phototropism, chloroplast movement and stomata opening. Phototropins feature, along with a serine-threonine kinase domain, two LOV (light-, oxygen- or voltage-regulated) domains, each binding a FMN (flavin mononucleotide).

Blue light affects the kinase domain by triggering, in the LOV domain, the formation of a covalent intermediate between the FMN cofactor and a nearby cysteine residue. Despite X-ray structures provided solid ground for mechanistic hypothesis, the molecular details of the inter-domain communication process are still unknown. By using accurate QM/MM (quantum mechanics/molecular mechanics) calculations we investigated the formation/breaking of the FMN/Cys covalent intermediated. We investigated the coupling between the LOV and kinase domains
by means of long MD (molecular dynamics) simulations and detailed PES (potential energy surface) explorations (MM level).


**COMP 437**

**Conformational sampling of macrocycles through accelerated molecular dynamics simulation**

*S. Roy Kimura* (1), roy.kimura@bms.com, 5 Research Parkway, Wallingford CT 06492, United States. (1) Department of Computer Assisted Drug Design, Bristol Myers Squibb, Wallingford CT 06492, United States

Macrocyclization is a strategy used in medicinal chemistry to lock a molecule in its bioactive conformation. The resulting decrease in conformational flexibility often leads to higher potencies due to the reduced entropy loss upon binding, and sometimes improved physical chemical properties such as bioavailability. Conformational searches of macrocycles are usually performed by temporary ring opening and Monte Carlo (MC) sampling to overcome the energy barriers between low energy states. However, widely available MC algorithms can only be used in conjunction with simplified continuum solvents such as dielectrics or Generalized Born-related models. In this study, we assess the use of molecular dynamics simulation in explicit solvent with periodic high-temperature pulsing as a method to overcome the characteristic energy barriers of macrocycles. The pros and cons of this methodology versus MC sampling are discussed.

**COMP 438**

WITHDRAWN

**COMP 439**

**Docking with GOLD: An overview of recent developments and docking performance**

*Jason C Cole* (1), cole@ccdc.cam.ac.uk, 12 Union Road, Cambridge Cambridgeshire CB2 1EZ, United Kingdom; John W Liebescheutz (1); Tjelvar S. G. Olsson (1); Colin Groom (1). (1) Cambridge Crystallographic Data Centre (CCDC), Cambridge CB2 1EZ, United Kingdom

The GOLD docking program has now been in academic and commercial research for over 10 years, and is routinely used for both lead discovery and optimisation. Many changes and improvements to the algorithms forming part of
GOLD have been made in this time period, such as the inclusion of additional scoring functions, knowledge-based constraints, handling of protein flexibility and key waters.

This talk will review some of these changes and present some of the current ongoing developments in GOLD. Results will be presented that illustrate the algorithmic efficacy based on public domain data sets.

COMP 440

Hyde: Scoring for lead optimization

Christian Lemmen\(^{(1)}\); Marcus Gastreich\(^{(1)}\); Carsten Detering\(^{(1)}\); Peter Oledzki\(^{(1)}\).

\(^{(1)}\) BioSolveIT, Germany

It is highly desirable to have a scoring function that provides guidance for the design of compounds with optimized bioactivity. Hyde [1] is such a scoring function. Its basic principle is a balanced assessment of the energetics of desolvation. Only three major factors are taken into consideration: (a) local hydrophobicity, (b) solvent accessible surface, and (c) contact surface area. Based on these, energetically favorable and unfavorable contributions to the binding affinity can be assessed on an atomic level.

It has been demonstrated previously that Hyde is able to distinguish between strong binders, weak binders, and non-binders [2]. However, systematically missing are terms regarding repulsion and strain which rendered Hyde not entirely applicable to conformationally strained or clashing poses. We have therefore further improved the scoring function and now consider the respective terms in an optimization phase prior to the actual score assessment. Further, we coupled it to a graphical interface. Hyde has never been calibrated for an improved correlation with measured binding affinities.

Atomic contributions can now be visualized, which turns out to be particularly helpful in a lead-optimization setup. One may immediately identify energetically unfavorable arrangements, like an H-bonding group without a counter-part in an otherwise hydrophobic pocket. Medicinal chemists will immediately have ideas how to alter a given structure in order to gain activity. The interface allows these changes to be tried out in an interactive manner like on a virtual workbench. We will demonstrate the performance of Hyde based on benchmark datasets as well as on published data of congeneric compound series.


COMP 441
Accurate cross-docking using internal coordinate mechanics (ICM), grids and multiple conformers

Ruben Abagyan\(^{(1)}\); Maxim Totrov\(^{(2)}\). (1) Skaggs School of Pharmacy & PharmSciences, UC San Diego, La Jolla CA 92093, United States (2) Molsoft, La Jolla CA 92037, United States

The internal coordinate docking protocol in which an explicit full atom and flexible ligand is sampled and scored in the field of four groups of grid potentials derived from a protein structure has been gradually improved and refined since it was first published in 1997. Currently, as the self-docking success rates approach 100%, the main limitation to successful docking appears to be the correct protonation and tautomerization state of both ligand and the pocket atoms, as well as the conformational flexibility and rearrangements of the binding site groups. The relative scoring of the same compound bound different proteins is further complicated by variability of protein entropic contributions to binding. We show that the efficient four dimensional docking protocol including a small number of pocket conformers allows to achieve the success rate of 80 to 90% in a cross-docking experiment. The ROC curves and the score-based discrimination values are presented for nuclear receptors and kinases. We thank Giovanni Bottegoni, Irina Kufareva, Manuel Rueda, Sojung Park and Seva Katritch.

COMP 442

Lead Finder docking and scoring evaluation on public data sets

Oleg Stroganov\(^{(1)}\), ostroganov@moltech.ru, Leninskie gory, 1/75A, Moscow not applicable 119992, Russian Federation ; Ghermes Chilov\(^{(1)}\); Fedor Novikov\(^{(1)}\); Victor Stroylov\(^{(1)}\); Val Kulkov\(^{(2)}\). (1) [quot]Molecular Technologies[quot], Ltd, Moscow 119992, Russian Federation (2) BioMolTech, Corp., Toronto Ontario M2L 1L1, Canada

Lead Finder is a software for ligand docking, binding energy evaluation and virtual screening. Its algorithm for ligand docking combines the classical genetic algorithm with various optimization procedures and resourceful exploitation of the knowledge generated during docking run. Lead Finder uses three distinct molecular mechanics functionals tailored for: ligand-docking, binding energy predictions; virtual screening. All scoring functions can be calculated at a number of levels of accuracy ranging from less accurate and very fast, to more accurate and resource-intensive. The docking success rate and virtual screening efficiency of Lead Finder have been evaluated on the well-known Astex and DUD test sets correspondingly. The docking
success rate on the Astex diverse test set was 90.6%. The average area under the ROC curves (AUC) obtained during the virtual screening using the DUD set was 0.70. The latter result can be improved if a careful preparation of protein and ligand structures was performed.

COMP 443

Results of docking experiments in MOE using the Astext 85 collection

Chris Williams\(^{(1)}\), cw@chemcomp.com, 1010 Sherbrooke Street West, Suite 910, Montreal Quebec H3A 2R7, Canada; Paul Labute\(^{(1)}\). (1) Chemical Computing Group, Montreal Quebec H3A 2R7, Canada

The results of computational experiments on the use of MOE docking functionality on the Astex 85 collection of protein:ligand complexes. We focus on self-docking experiments to predict ligand poses. The effects of several algorithmic parameters are presented along with the details of successful and unsuccessful cases.

COMP 444

RosettaLigand: Flexible docking into comparative models

Jens Meiler\(^{(1)}\), jens.meiler@vanderbilt.edu, 7330 Stevenson Center, Station B 351822, Nashville TN 37235, United States. (1) Chemistry, Vanderbilt University, Nashville TN 37235, United States

For many therapeutic targets no experimental structures is available. We benchmark ligand docking into comparative models using RosettaLigand. Full ligand and protein flexibility is required to account for inaccuracies of the comparative model and absence of the ligand in the experimental structure. RosettaLigand identified the binding mode within 2 Å RMSD for over 70% of the test cases; including nine protein/ligand complexes where comparative models had been submitted in the 8\(^{th}\) CASP experiment. RosettaLigand's ability to predict and rank the binding free energy of protein/ligand complexes is assessed on a dataset of HIV protease inhibitors. Extensions for RosettaLigand to utilize pharmacophore maps for scoring and perform drug design will be introduced. (grant support: NIH-NIMH-082254, NIH-NIGMS-080403).
Comp 445

Surflex-Dock: Progress since 2004

Ajay N Jain(1), ajain@jainlab.org, 1450 3rd Street, MC 0128, San Francisco CA 94158, United States. (1) Department of Bioengineering and Therapeutic Sciences, UCSF, San Francisco CA 94158, United States

Operational use of molecular docking methods involves making predictions of ligand pose, ligand database rank, or ligand affinity. All of these predictions occur, in practice, using protein structures that are bound to a different ligand than the one(s) being docked. Frequently, predictions for ligands are sought that are desired (or known) to be very different from any existing true binders. However, benchmarks for evaluating docking method tend toward reproduction of cognate ligand poses for testing geometric accuracy and toward virtual screening assessments where analog bias corrupts the interpretation of results. We present results for Surflex-Dock showing how the use of protein pocket adaptation is key to accurate pose prediction when docking novel ligands, and how scoring function customization using small amounts of ligand binding data yields significant improvements in virtual screening for novel ligands.

Comp 446

Tackling difficult protein targets for novel anti-cancer treatments
Protein–protein interactions (PPIs) are involved in many biological processes from cellular signal transduction to programmed cell death. Therefore, they are viable and important targets for therapeutic intervention. However, designing small-molecule inhibitors for PPIs is generally recognized to be harder than for other targets because of the flat, featureless and the high plasticity of PPIs binding surface. In this presentation, we discuss our experience in working with PPIs using a combination of experimental and computational techniques to address different phases of drug discovery, from target assessment, hit identification and ligand optimization. We will demonstrate the importance of considering protein flexibility early on during ligand optimization. We will also highlight the challenges posed by PPIs and propose general strategies forward.

COMP 447

Pose grading for docking applications

Scott A Wildman(1), wildman@biochem.wustl.edu, 660 S. Euclid Ave, Box 8231, St. Louis MO 63110, United States; Julie Penzotti(1). (1) Center for Computational Biology, Department of Biochemistry and Molecular Biophysics, Washington University in St. Louis, St. Louis MO 63110, United States

Proper sampling in docking applications generates more poses than can reasonably be evaluated by physics-based methods, and empirical scoring methods often fail to identify the best pose. We apply an intermediate grading step to evaluate the large number of initial poses. This eliminates infeasible poses, allowing detailed characterization to focus on only the best poses. This is accomplished with a fast calculation of pose grades based on the interacting protein surfaces and the properties of the atoms that comprise those surfaces. We demonstrate the approach using available docking test sets.

COMP 448

Alignment of protein surfaces using shape and chemical features: Application to detect local similarity among protein-ligand binding sites

Brajesh K Rai(1), brajesh.rai@pfizer.com, Eastern Point Road, Groton CT 06340, United States; Gregory A Bakken(1); W Jeffrey Howe(1). (1) Computational Sciences CoE, Pfizer PharmaTherapeutics Research & Development, Groton CT 06340, United States

We have developed a novel computational method that seeks to establish similarity between a pair of arbitrary-size protein surfaces based on the alignment...
of their shape and chemical features. The algorithm represents the molecular surfaces by a collection of (pseudo)atom-centered Gaussians. These pseudo-atoms are colored based on the physico-chemical properties of the underlying surface. Alignment between a pair of protein surface shapes is obtained by maximizing their volume of intersection as well as their color overlap. Maximal superposition between a pair of surfaces is obtained by generating multiple initial alignments each of which are locally optimized. The method is capable of identifying and correctly overlaying locally similar surface patches even if the overall similarity between the two binding sites is weak. This has potential application in drug discovery as small molecule fragments retrieved from related binding sites can be used as a starting point for further ligand idea generation.

COMP 449

Ligand-steered homology modeling and high-throughput docking:
Successful evaluation in aminergic GPCRs

Claudio N Cavasotto(1), claudio.n.cavasotto@uth.tmc.edu, 7000 Fannin, Ste. 860B, Houston TX 77030, United States; Sharangdhar S Phatak(1), Sharangdhar.S.Phatak@uth.tmc.edu, 7000 Fannin Ste. 860B, Houston TX 77030-5400, United States. (1) School of Health Information Sciences, University of Texas Health Science Center at Houston, Houston TX 77030-5400, United States

Using three GPCR structures as templates (β2-adrenergic, A2A-adenosine, and rhodopsin receptors), we enhanced and validated the ligand-steered modeling method through cross-modeling of, and high-throughput docking on four target/template cases. We raised and answered three questions:

1. **Can the ligand-steered modeling generate a small set of near-native models by using one structural template?** In all cases, the ligand-steered modeling method identified a small-set of models with at least one near-native model.

2. **How well do the ligand-steered models perform in docking?** The enrichment factor obtained by the best ligand-steered models was superior to those obtained by the crude models and random selection in all cases.

3. **How well do the ligand-steered models perform in assessing cross-selectivity between β2 and A2A ligands?** In all cases, the ligand-steered models either identified a higher percentage of target specific actives, and/or discarded higher percentage of “other” actives, as compared to the crude models.

COMP 450

More better charges for metalloproteins
Modeling of binding between metalloproteins and small molecules is generally regarded as a difficult problem. In particular, most of the current docking programs are ineffective in finding the correct binding poses when metalloproteins are involved. It was found in our previous research that the difficulty stems from inadequate partial charges used for docking with force field based energy scoring. We had devised a method in which partial charges on metal ions along with surrounding protein atoms are rescaled after QM/MM energy calculations before docking with the new charges. The method proved to be quite successful in describing correct binding modes of ligands in binding sites with metal atoms but rather cumbersome to perform and costly to be practical. In an attempt to conceive a docking protocol that retains the same spirit as our previous method and yet can be practical, we tested a few different implementations of combining QM/MM calculations with force field docking. The results show that a proper pre-docking charge rescaling through QM/MM calculations combined with QM/MM docking with minimum QM region can yield compatible accuracy as our previous method.