

American Chemical Society
Division of Computers in Chemistry
ABSTRACTS

224th ACS National Meeting

Boston, MA
August 18-22, 2002

R. A. Wheeler, W. D. Cornell, Program Chairs

SUNDAY MORNING

- **Linking Genomic Information with Drug Discovery**
W. D. Cornell, Organizer Papers 1-5
- **New Developments in Force Fields for Molecular Modeling**
T. A. Halgren, Organizer, Presiding Papers 6-10
- **Classical and Quantum Statistical Mechanics Studies of Solvation**
L. X. Dang, Organizer; P. J. Rossky, Presiding Papers 11-16

SUNDAY AFTERNOON

- **Linking Genomic Information with Drug Discovery**
D. Kallick, Organizer Papers 17-21
- **New Developments in Force Fields for Molecular Modeling**
T. A. Halgren, Organizer, Presiding Papers 22-26
- **Classical and Quantum Statistical Mechanics Studies of Solvation**
L. X. Dang, Organizer; G. A. Voth, Presiding Papers 27-32
- **Chemistry of Computers**
P. Lykos, Organizer Papers 33-39

MONDAY MORNING

- **Linking Genomic Information with Drug Discovery**
D. Kallick, Organizer Papers 40-44
- **New Developments in Force Fields for Molecular Modeling**
T. A. Halgren, Organizer, Presiding Papers 45-49
- **Classical and Quantum Statistical Mechanics Studies of Solvation**
L. X. Dang, Organizer; S. Hammes-Schiffer, Presiding Papers 50-55
- **QSAR in the Brave New World of Structural Biology, Genomics, Combichem, and High-Throughput Screening**
R. D. Clark, Organizer Papers 56-61

MONDAY AFTERNOON

- **Linking Genomic Information with Drug Discovery**
J. Harris, Organizer Papers 62-66
- **New Developments in Force Fields for Molecular Modeling**
P. Norrby, Organizer, Presiding Papers 67-71
- **Classical and Quantum Statistical Mechanics Studies of Solvation**
L. X. Dang, Organizer; B. M. Ladanyi, Presiding Papers 72-77
- **QSAR in the Brave New World of Structural Biology, Genomics, Combichem, and High-Throughput Screening**
R. D. Clark, Organizer Papers 78-83

MONDAY EVENING

- **Sci-Mix**

R. A. Wheeler, Organizer Papers 132, 142-144, 146, 149, 155, 157, 162, 165, 182, 186, 190

TUESDAY MORNING

- **Linking Genomic Information with Drug Discovery**

W. D. Cornell, Organizer Papers 84-88

- **New Developments in Force Fields for Molecular Modeling**

T. A. Halgren, Organizer, Presiding Papers 89-93

- **Computational Chemistry in Chemical Education**

T. J. Zielinski, Organizer; T. J. Zielinski, Presiding Papers 94-99

- **QSAR in the Brave New World of Structural Biology, Genomics, Combichem, and High-Throughput Screening**

R. D. Clark, Organizer Papers 100-104

TUESDAY AFTERNOON

- **Emerging Technologies in Computational Chemistry**

D. B. Boyd, Organizer Papers 105-110

- **Computational Chemistry in Chemical Education**

T. J. Zielinski, Organizer; A. Grushow, Presiding Papers 111-116

- **Classical and Quantum Statistical Mechanics Studies of Solvation**

L. X. Dang, Organizer, Presiding; L. X. Dang, Organizer Papers 117-122

- **QSAR in the Brave New World of Structural Biology, Genomics, Combichem, and High-Throughput Screening**

R. D. Clark, Organizer Papers 123-127

TUESDAY EVENING

- **Computers in Chemistry Posters**

R. A. Wheeler, Organizer Papers 128-203

WEDNESDAY MORNING

- **Semi-Empirical Molecular Orbital Methods**

A. J. Holder, Organizer Papers 204-208

- **Tautomerism and Protonation State in the Context of Virtual Screening**

R. S. Pearlman, Organizer Papers 209-212

- **Classical and Quantum Statistical Mechanics Studies of Solvation**

L. X. Dang, Organizer; G. K. Schenter, Presiding Papers 213-218

- **Computers in Chemistry General Contributions**

R. A. Wheeler, Organizer; J. D. Evanseck, Presiding Papers 219-225

WEDNESDAY AFTERNOON

- **Semi-Empirical Molecular Orbital Methods**

A. J. Holder, Organizer Papers 226-231

- **Tautomerism and Protonation State in the Context of Virtual Screening**

R. S. Pearlman, Organizer Papers 232-236

- **Classical and Quantum Statistical Mechanics Studies of Solvation**

L. X. Dang, Organizer; D. A. Dixon, Presiding Papers 237-242

- **Chemometrics and Chemoinformatics**

B. Lavine, Organizer Papers 243-247

THURSDAY MORNING

- **Semi-Empirical Molecular Orbital Methods**

A. J. Holder, Organizer

Papers 248-253

- **Chemometrics and Chemoinformatics**

B. Lavine, Organizer

Papers 254-258

- **Classical and Quantum Statistical Mechanics Studies of Solvation**

L. X. Dang, Organizer; L. R. Corrales, Presiding

Papers 259-264

- **Computers in Chemistry General Contributions**

R. A. Wheeler, Organizer; C. A. Parish, Presiding

Papers 265-271

THURSDAY AFTERNOON

- **Computers in Chemistry General Contributions**

R. Wheeler, Organizer

Papers 272-277

- **Chemometrics and Chemoinformatics**

B. Lavine, Organizer

Papers 278-282

- **Computers in Chemistry General Contributions**

R. A. Wheeler, Organizer; T. R. Cundari, Presiding

Papers 283-291

DIVISION OF COMPUTERS IN CHEMISTRY

1. TARGET VALIDATION AND DRUG DISCOVERY: AN OVERVIEW. *Deborah Kallick, Target Validation and Drug Discovery, Incyte Genomics, Inc, 3160 Porter Dr, Palo Alto, CA 94304, dkallick@incyte.com*

The new paradigm for target validation and drug discovery employs the wealth of information available from the human genome. The basic steps in the process will be outlined and discussed, with particular attention to issues of interest to computational chemists. The various phases of target validation will be discussed using specific examples from the public domain where possible. Wrap up will illustrate how early stage target validation is linked to downstream drug discovery. This talk will introduce the area of target validation and drug discovery for the uninitiated chemist.

2. DEVELOPMENT AND APPLICATION OF A TECHNOLOGY PLATFORM FOR HIGH THROUGHPUT STRUCTURE-GUIDED DRUG DISCOVERY. *Stephen W. Kaldor, Syrrx, Inc, 10450 Science Center Drive, San Diego, CA 92121, Fax: 858-202-4333, steve.kaldor@syrrx.com*

We have studied the bottlenecks in structure-guided drug discovery and have applied parallel processing, automation, and miniaturization to provide 100- to 1000-fold rate enhancements for a number of the slow steps involved in proceeding from gene to structure (e.g. fermentation, purification, crystallization and data collection). As a result of these streamlining efforts, we are now producing de novo protein structures at a rate of 10-15 per month. Complete structures have been determined in as little as 3 weeks from gene to refined coordinates. Access to such early structural information has allowed us to focus our small molecule lead generation and optimization efforts using techniques such as targeted virtual ligand screening to further enhance our efficiency in identifying potential clinical candidates.

3. DISCOVERY INFORMATICS: INTEGRATING GENOMICS AND PROTEOMICS FOR DRUG DISCOVERY. *Mark D. Adams, Celera Genomics, 45 W. Gude Dr, Rockville, MD 20850, Fax: 240-453-3755, mark.adams@celera.com*

The completion of the human genome, as well as genomes of mouse, Drosophila, and many other model organisms has provided a basic platform for exploring biological problems. Comparative genomics approaches offer a powerful tool for identifying functional elements of the human genome, for defining effective models of human disease, and for choosing the appropriate organisms for efficacy and toxicity studies. Quantitative analysis of protein expression is being used to analyze changes in the protein complement of cells in a variety of conditions, including transformation from normal to tumor phenotype. Correlation of this information with genome information, and the wealth of additional biological data that is accessible via the Internet represents an information integration challenge and opportunity. Methods for prioritizing genes of interest based on genome and proteome analysis will be discussed.

4. GENOMICS-DRIVEN STRUCTURE-BASED DRUG DISCOVERY. *Tod Klingler, Ken Lind, and Vickie Tsui, Computational Chemistry, Structural GenomiX, Inc, 525 Brannan Street, Suite 200, San Francisco, CA 94070, tod@stromix.com*

The drug discovery process has become an information intensive process. With the availability of information (including sequence, expression, structure and function) for every human gene, the challenges for the pharmaceutical industry are no longer with the optimization of a sequential experimental process. Rather, the challenges are the management of a highly parallel, information-rich selection process. Computational methods, including protein structure modeling and virtual screening, are beginning to have a significant impact in this information management process. I will describe the integration of experimental

and computational technologies at SGX and how they are being used to discovery drug leads.

5. KINASES VERSUS PROTEIN-PROTEIN INTERACTIONS: DIFFERENT OPPORTUNITIES AND CHALLENGES IN DRUG DISCOVERY. *Carlos Garcia-Echeverria, Oncology Research, Novartis Pharma AG, WKL-136.4.84, CH-4002 Basel, Switzerland, Fax: 41 61 696 62 46, carlos.garcia-echeverria@pharma.novartis.com*

Phosphorylation and protein-protein interactions constitute essential regulatory steps in basically all biological events. Genomic research is expected to provide a continuous flow of novel kinase and protein-binding targets for therapeutic intervention in many diseases. We are facing a number of challenges to identify inhibitors of protein phosphorylation and signalling. Most of the available kinase inhibitors are believed to act within the ATP binding pocket of the enzyme, a highly conserved region across the whole kinase family. Medicinal chemists have to impart potency and selectivity to a limited number of scaffolds by modulating the interactions between the modified template and the ATP binding site of the selected kinase/s. Protein-protein interactions generally span large surface areas and, unlike enzyme/substrate interactions, they lack deep, well defined binding pockets. Examples highlighting different techniques that may be brought to bear on the discovery of kinase inhibitors and antagonists of protein-protein interactions will be presented.

6. ALL ELEMENT AUGMENTATION TO MM3. *George D. Purvis III, CAChe Group, Fujitsu, 15244 NW Greenbrier Pkwy, Beaverton, OR 97006, Fax: 503-531-9966, gpurvis@CACheSoftware.com*

MM2 and MM3 are extended to all elements by augmenting the force-field with a rule based system in CAChe that estimates missing parameters and adds new hybridizations and bond types. Recently, the rule based system has been refined for bio-molecules. The refinements are described and the results from testing the new bio-MM2 and bio-MM3 against Halgren's conformational and hydrogen bond test set are reported.

7. DEVELOPMENT OF A MORE ACCURATE AMBER UNITED-ATOM FORCE FIELD FOR PROTEIN FOLDING AND LARGE-SCALE BIOMOLECULAR SIMULATIONS. *Ray Luo, Department of Molecular Biology and Biochemistry, Univ. Calif. Irvine, Irvine, CA 92697-3900, Fax: 949-824-8551, rluo@uci.edu, Junmei Wang, Texas Biotechnology Corp, and Peter A. Kollman, Deceased*

Recent advances in both methodology development and computer hardware have opened the door to studies of biochemical systems on a very large scale and biochemical systems of fair complexity on very long time scales. Adapting continuum solvent in these type of studies certainly improves the efficiency of the simulations needed in these systems. With such simplifications in the treatment of solvent, it now makes sense to come back to the united-atom (UA) model for the solute. The previous UA force field for AMBER is rather limited in that the charging scheme is out-dated and the torsional fitting is rather poor, resulting in unreliable structures and dynamic properties for biomolecules. In this study, we have developed a brand-new UA parameter set from the following three aspects: deriving a set of new atomic charges with the DFT/cc-pvtz/HF-631g with the PCM continuum solvent included; refitting the protein main chain torsional parameters by systematic search to optimize the UA force field with respect to potential energies by MP2/cc-pvtz/HF-631g with PCM continuum solvent for the 7 dialanine and 11 tetrapeptide local minima; refitting all side-chain torsional parameters to reproduce the MP2 and/or experimental conformational energies for a series of training molecules. Tests of the new UA parameter set in biomolecular simulations show very encouraging improvement

over the old AMBER UA parameters and a quality similar to that of the AMBER all-atom parameter set.

8. PERFORMANCE OF EMPIRICAL FORCE FIELDS IN THE MODELING OF NUCLEIC ACID STRUCTURE, DYNAMICS AND INTERACTIONS. *Thomas E. Cheatham III, Department of Medicinal Chemistry, University of Utah, 2000 East, 30 South, Skaggs Hall 201, Salt Lake City, UT 82117, Fax: 801-585-5366, tec3@utah.edu*

Empirical force fields for nucleic acids have improved considerably in recent years, such that, in atomistic molecular dynamics simulations (including an explicit representation of the solvent and counterions), accurate modeling of the structure, dynamics and interactions of a wide variety of nucleic acid systems has been performed. With methods that properly represent the long range electrostatic interactions and appropriate boundary conditions, we have been able to reproduce the details of A-tract bending, differences in bendability and twistability of nucleosome positioning sequences, and even the subtle details of DNA—minor groove binding drug interactions. As a part of our efforts, we have also investigated a number of the state-of-the-art additive molecular mechanical force fields for nucleic acids in simulation of a wide variety of DNA duplexes and shown excellent performance in each case, even from force fields that were independently derived. We will highlight recent work as it relates to the performance and evaluation of the current set of well-performing force fields for nucleic acids.

9. OPLS_2002: A NEW VERSION OF THE OPLS-AA FORCE FIELD. *Wolfgang Damm¹, Thomas A. Halgren¹, Robert B. Murphy¹, Alex M. Smondryev¹, Richard A. Friesner², and William L. Jorgensen³. (1) Schrodinger, Inc, 120 W 45th Street, New York, NY 10036, Fax: 646-366-95507, damm@schrodinger.com, (2) Department of Chemistry, Columbia University, (3) Department of Chemistry, Yale University*

A new version of the OPLS-AA force field is presented. The new implementation differs from the previous OPLS-AA versions by using an improved atom-typing procedure, by employing bond-charge increments to construct the partial atomic charges, and by using a different scaling factor for 1,4 electrostatic interactions. The changes improve the representation even of molecules with functionality for which no special parameters have been developed. Extended sets of torsional parameters and bond-charge increments have been fit to reproduce high-level ab initio data. When finished, the 2002 version of the OPLS force field will represent a major expansion of the set of well-parameterized compounds that are of interest for studying pharmaceutical and biomolecular systems.

10. CHARMM BIOMOLECULAR FORCE FIELD: RECENT DEVELOPMENTS AND FUTURE DIRECTIONS. *Alexander D. MacKerell Jr., Department of Pharmaceutical Sciences, University of Maryland, School of Pharmacy, 20 N. Pine St., Baltimore, MD 21201, Fax: 410-706-0346*

The CHARMM Biomolecular Force Field encompasses proteins, nucleic acids, lipids and carbohydrates including non-natural modifications and a variety of small molecules, allowing modeling and MD simulation based studies of heterogeneous systems in the condensed phase. Recent developments and extensions of the force field will be presented, including work on the nonbond parameters, polyunsaturated lipids, the protein backbone and polarizable force fields.

11. QUANTUM MOLECULAR DYNAMICS IN LIQUIDS. *Bruce J. Berne, Department of Chemistry, Columbia University, 3000 Broadway, MC 3103, New York, NY 10027, Fax: 212-932-1289, berne@chem.columbia.edu*

Maximum entropy methods are used in conjunction with path integral Monte Carlo simulations to determine the real time dynamics of time correlation functions, spectral line shapes, tunneling rate constants and self diffusion in condensed systems.

12. SOLVATION OF THE EXCESS PROTON IN COMPLEX ENVIRONMENTS. *Gregory A. Voth, Department of Chemistry & Henry Eyring Center for Theoretical Chemistry, University of Utah, 315 South 1400 East, Room 2020, Salt Lake City, UT 84112, Fax: 801-581-4353, voth@chem.utah.edu*

The solvation states of the excess proton in several important systems will be described using the multi-state empirical valence (MS-EVB) approach combined with Molecular Dynamics simulation. The MS-EVB approach allows for the treatment of explicitly quantum phenomena such as dynamical bond-breaking, which, in turn, strongly influence the solvation structure of the excess proton in various aqueous and biomolecular systems. The differences and similarities between the excess proton in bulk water, water clusters, and several biomolecular systems will be discussed. If time allows, ab initio Molecular Dynamics simulations of several of these systems will be presented based on a new ab initio MD approach.

13. THEORETICAL STUDIES OF PROTON-COUPLED ELECTRON TRANSFER REACTIONS IN SOLUTION. *Sharon Hammes-Schiffer, Department of Chemistry, Pennsylvania State University, 152 Davey Laboratory, University Park, PA 16802, Fax: 814-863-5319, shs@chem.psu.edu*

A theoretical formulation for proton-coupled electron transfer reactions in solution will be presented. The active electrons and protons are treated quantum mechanically, and the free energy surfaces are obtained as functions of collective solvent coordinates corresponding to the proton and electron transfer reactions. This theoretical formulation provides mechanistic information and allows the calculation of rates and kinetic isotope effects. This theory has been applied to experimentally studied proton-coupled electron transfer reactions in iron bi-imidazole, ruthenium polypyridyl, and osmium complexes. The experimentally measured deuterium kinetic isotope effects for these systems range from a moderate value of 2 to an unusually high value of 178. The theoretical calculations are consistent with the experimental data and provide insight into the chemical and physical basis for the observed rates and kinetic isotope effects. These studies also lead to experimentally testable predictions.

14. MODELING NON-ADIABATIC EXCITED STATE CHARGE TRANSFER TO SOLVENT DYNAMICS. *David F. Coker, and Ning Yu, Department of Chemistry, Boston University, 590 Commonwealth Ave., Boston, MA 02215, Fax: (617) 353-6466, coker@bu.edu*

We explore the non-adiabatic excited state charge transfer to solvent dynamics in which molecular iodine is first dissociated producing excited radicals which are then further excited, transferring charge to the surrounding xenon solvent atoms producing possibly both I-Xe⁺ and I-Xe²⁺ ionic complexes. We study these excited state charge transfer to solvent reactions in both liquid and solid xenon solution and study the dynamical excited state rearrangements that lead to these different complexes. Our calculations include effects of solvent polarization and spin orbit coupling using a semi-empirical diatomics in ionic systems approach and use a non-adiabatic molecular dynamics method to incorporate the effects of electronic transitions on the nuclear motion.

15. SIMULATION OF QUANTUM MOLECULAR PROCESSES USING ENTANGLED CLASSICAL TRAJECTORY ENSEMBLES. *Craig C. Martens, Department of Chemistry, University of California, Irvine, Irvine, CA 92697-2025, Fax: 949-824-8571, cmartens@uci.edu*

The time-dependent quantum mechanics of heavy particles moving on a single potential energy surface can often be represented surprisingly well by the evolution of classical trajectory ensembles. However, manifestly quantum mechanical phenomena such as transitions between coupled electronic states, electronic coherence and its decay, or quantum mechanical tunneling require fundamental modification of the purely classical equations of motion. In this talk, I describe a general approach to simulating molecular dynamics with intrinsic quantum effects using classical trajectories and its application to the problems of nonadiabatic dynamics, coherent multistate electronic-nuclear dynamics, and tunneling through potential barriers. When viewed from the trajectory ensemble perspective, quantum effects arise as a breakdown of the

statistical independence of the trajectories in the ensemble and a nonlocal entanglement of their collective evolution.

16.

INVERSION AND ILL-POISED PROBLEM IN COMPUTING NONADIABATIC ELECTRON TRANSFER RATE CONSTANTS. *Jie-Lou Liao, and Gregory A. Voth, The Henry Eyring Center for Theoretical Chemistry, Department of Chemistry, University of Utah, 1400 E., 315 S, Salt Lake City, UT 84112, Fax: 801-581-4353, ljl@hec.utah.edu, voth@chem.utah.edu*

A major difficulty in computing nonadiabatic electron transfer (ET) reaction rate constants is that they, in principle, involve the computation of real-time quantum dynamics. To circumvent this problem, the nonadiabatic ET rate constant formalism can be expressed in terms of an inverse Laplace transform. This inversion is an ill-posed problem. The maximum entropy method (MEM) is explored to solve the problem.

The method is applied to a model problem of a two-state ET system coupled to a dissipative bath. The resulting numerical studies are primarily focused on the effects of the anharmonicity of the intramolecular vibrational modes and the coupling of the motion of these modes to the electronic tunneling in ET reactions. These calculations show how the anharmonicity and the electronic-vibrational coupling can significantly affect the value of an ET rate constant.

17.

INDUSTRIALIZATION OF DRUG DISCOVERY. *Jeffrey S. Handen, PricewaterhouseCoopers Consulting, Philadelphia, PA 19103-4094, jeffrey.s.handen@us.pwcglobal.com*

Recent technological advances in combinatorial chemistry and high-throughput screening combined with the explosion of genomic data and the increased empowerment of the consumer/patient have all converged to start exerting significant pressure on the traditional drug development paradigm of "one size fits all". As industry and society move to mass customization in drug development, or "personalized medicine", there will be a significant challenge posed to the drug discovery industry. That challenge will be how to dramatically increase the speed and volume of discovery in order to both exploit the new technological opportunities afforded by advances in automation, computing, and genomics and serve the growing market expectations for "personalized medicine".

18.

FINDING IMPORTANT DISEASE GENES USING MASS SPECTROMETRY. *Charles R. Cantor, Sequenom, Inc, San Diego, CA 92121, ccantor@sequenom.com*

Sequenom's major internal research program uses high throughput mass spectrometry-based genotyping to discover commercially important disease genes. For a target to be commercially important, it must affect a reasonably large number of people and it must be drugable, either directly or indirectly. Both genomics and genetics have the promise to produce such targets but the pathways and likelihood of success appear to be fairly different. Genomics can filter genomes for molecules known to be in potentially drugable classes, but unless experiments are done on human material it is difficult to decide which targets are large enough to be commercially important. Genetics works directly on humans, but the targets produced may or may not correspond to anything that has been previously drugged successfully. These issues will all be illustrated with the portfolio of targets Sequenom has discovered using genetics, and a fairly detailed look at one target, which we believe may be drugable, will be offered.

19.

USING MAMMALIAN GENETICS TO VALIDATE DRUG DISCOVERY TARGETS FOR SMALL MOLECULE INTERVENTION. *Alan Main, Lexicon Pharmaceuticals, East Windsor, NJ 08520, amain@lexpharma.com*

Lexicon employs a genome-wide pharmaceutical discovery program to rapidly determine the physiologic functions of genes in mammals using its proprietary mouse knockout technologies on an unprecedented scale. Lexicon is focusing its large-scale gene function discovery process on the druggable classes of genes such as receptors, ion channels, key enzymes, and secreted proteins. Through its Genome5000 Program, Lexicon is determining in vivo the function

of 5,000 genes over five years for the discovery of new disease targets. Lexicon uses a battery of tests conducted in vivo to validate targets in an actual mammalian system. Genes have already been identified as drug targets and are rapidly being moved into small molecule drug discovery in the fields of cardiology, immunology and neurology. For example, LG314 knockout mice registered significantly lower triglyceride and cholesterol levels, as well as reduced body fat, with no change in diet. The LG314 target, as well as other examples, will be used to demonstrate the power of Lexicon's technology.

20.

THE PERSONALIZED MEDICINE STRATEGY IN DRUG DISCOVERY. *Geoffrey S. Ginsburg, Molecular Medicine and Research Strategy, Millennium Pharmaceuticals, Inc, 45 Sidney Street, Cambridge, MA 02139*

The post genomic era has brought with it advances in human genome research that opening the door to a new paradigm for practicing medicine that promises to transform healthcare. Personalized medicine, the use of marker-assisted diagnosis and targeted therapies derived from an individual's molecular profile, will impact the way drugs are developed and how medicine is practiced. The pharmaceutical industry has successfully integrated multiple novel genomic technologies into the drug discovery and development process in order to validate and functionally annotate genes leading to thousands of potential new drug targets. Now, knowledge of the molecular basis of disease will lead to additional novel target identification, toxicogenomic markers to screen compounds and improved selection of patients for clinical trials, which will fundamentally change the industry. The traditional linear process of drug discovery and development is already being replaced by an integrated and heuristic approach. In addition patient care will be revolutionized through the use of novel molecular predisposition, screening, diagnostic, prognostic, and pharmacogenomic and monitoring markers. Although numerous challenges will need to be met to make personalized medicine a reality, with time, this approach will replace the traditional trial-and-error practice of medicine. In this session the personalized medicine strategy in drug discovery will be described - a strategy that encompasses the emerging use of genomic information for disease prognosis and individualized therapeutic intervention along the entire drug development process.

21.

A COMPREHENSIVE SYSTEMS BIOLOGY APPROACH TO IDENTIFY KEY NOVEL TARGETS FOR DRUG DEVELOPMENT. *Robert D. Klein, Deltagen, Inc, 1003 Hamilton Ave., Menlo Park, CA 94025, Fax: 650 752 0202, klein@deltagen.com*

Genome sequencing efforts have identified thousands of genes with no known function representing potential new drug targets. The challenge is to identify which targets have therapeutic potential early in the drug discovery process. Using mouse knockout technology, it is now possible to assign in vivo mammalian function to novel members of gene families representing drugable targets, including GPCRs, ion channels, nuclear hormone receptors, kinases, and phosphatases. A comprehensive systems biology approach, integrating extensive phenotypic analysis, microarray expression studies, disease challenge models and pathway analysis has enabled the identification of key novel targets for the treatment of disease. Examples illustrating the discovery of novel drug targets using this systems biology approach will be presented.

22.

CHEMICAL POTENTIAL EQUALIZATION: A MANY-BODY FORCE FIELD FOR MOLECULAR SIMULATIONS. *Darrin M. York, Department of Chemistry, University of Minnesota, 207 Pleasant Street S.E, Minneapolis, MN 55455, york@chem.umn.edu*

A many-body force field for molecular simulations based on the chemical potential equalization (CPE) method is presented. Several extensions of the model are highlighted, including: 1) the coupling of polarization and short-ranged exchange effects, 2) the scaling of the polarizability of polymer systems, and 3) a new method for treating charge transfer at short and long range. These aspects have traditionally been problematic in conventional molecular force fields. Results for several small molecules are compared with ab initio calculations and experiment. Finally, a new coupling scheme between CPE and density-functional theory will be outlined for hybrid quantum mechanical/molecular mechanical simulations.

23.

NEW ELECTROSTATIC AND VALENCE ASPECTS OF A SPECTROSCOPICALLY DETERMINED FORCE FIELD (SDFF) FOR THE PEPTIDE GROUP. *Kim Palmo, Berit Mannfors, Noemi G. Mirkin, and Samuel Krimm, Biophysics Research Division and Department of Physics, University of Michigan, 930 N. University Ave., Ann Arbor, MI 48109, Fax: 734-764-3323, palmo@umich.edu*

It is well known that the electric potential around a molecule, and therefore also the potential derived atomic charges, sometimes depend significantly on structure. We have found, for example, that the electric potential around the N-methylacetamide molecule depends on the NH out-of-plane angle in a way that requires much larger charge variations than can be explained by polarization through the (molecular mechanics) electric field. Geometry dependencies in the electrostatics are therefore explicitly taken into account in our SDFF by using internal coordinate charge-fluxes, originally devised for calculating IR intensities. When the charges are allowed to vary with structure, special attention has to be paid to the way in which the molecular mechanics electrostatic potential energy is calculated.

For a correct description of peptide dynamics the NH out-of-plane motion has to be separated from the peptide CN torsion. In the SDFF this is achieved by defining mutually orthogonal torsion and out-of-plane bend coordinates. The barriers and force constants can then be properly designed to reproduce the potential energy, vibrational frequencies, IR intensities, etc.

24.

TINKER POLARIZABLE ATOMIC MULTIPOLE FORCE FIELD FOR PROTEINS. *Pengyu Ren, and Jay W. Ponder, Department of Biochemistry, Washington University School of Medicine, 660 S. Euclid Ave, St. Louis, MO 63110, Fax: 314-362-7183, pren@dasher.wustl.edu*

The TINKER polarizable force field has been developed for peptide and protein modeling. The permanent electrostatic treatment consists of atom-centered monopoles, dipoles and quadrupoles obtained from ab initio calculations by means of Stone's distributed multipole analysis (DMA). Induced dipoles are computed from Thole's interactive dipole polarizability model in response to the intra- and intermolecular electric fields. Other force field terms include buffered 14-7 vdW, anharmonic bond stretching and angle bending, Fourier torsions, and selected cross terms. Parameters have been derived for a large number of small molecules including water, amides, alcohols, amines, acids, sulfides, aromatics, and other functional classes via gas and liquid simulations. Capped dipeptides have been chosen as model compounds to obtain multipole parameters for larger polypeptides. We utilize a scheme to remove intramolecular polarization from the DMA, exposing the conformation-independent components. Application of the new force field to the structure, energy and thermodynamic properties for small molecules and peptides will be discussed.

25.

POLARIZABLE FORCE FIELDS: NEW PHYSICS, NEW QUESTIONS, NEW PROBLEMS. *Charles L. Brooks III, Department of Molecular Biology, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, CA 92037*

The next key objective in extending the physical nature of empirical force field calculations is the addition of charge polarization. While current force fields incorporate the influences of charge polarization in a mean-field manner and provide quite accurate representations of intermolecular interactions, a physical understanding of how electronic polarizability affects localized interactions will only be achieved through the direct incorporation of this influence. We have developed a first generation polarizable force field based on the CHARMM potential energy functions. This force field utilizes the fluctuating charge model to accomplish charge polarization during the course of dynamics simulations. We will describe the development and testing of this force field with initial applications to a number of systems.

26.

DEVELOPMENT OF A POLARIZABLE FORCE FIELD FOR PROTEINS AND PHARMACEUTICAL COMPOUNDS. *Richard A. Friesner¹, George Kaminski¹, Harry A. Stern¹, and Jon Maple². (1) Department of Chemistry, Columbia University, 3000 Broadway, MC 3110, New York, NY 10027, Fax: 212-854-7454, rich@chem.columbia.edu, (2) Schrodinger*

We describe our most recent efforts to develop a robust and automated methodology for construction of a polarizable force field for an arbitrary organic

molecule. Our model employs fixed and polarizable dipoles as well as atomic point charges, which are fit to highly accurate ab initio quantum chemical data. A key focus of the present talk will be the development of van der Waals parameters which are capable of reproducing both gas phase and liquid state energetics. We have found that transferable dispersive tails for various atoms can be obtained by fitting to liquid state data, which are then complemented by short range parameters adjusted to reproduce the binding energies and structures of molecular pairs. Results for both small molecules and proteins will be presented.

27.

THE ROLE OF SOLVATION DYNAMICS IN HYDROPHOBIC COLLAPSE. *David Chandler, Department of Chemistry, University of California, Berkeley, Berkeley, CA 94720-1460, Fax: 510-642-6262, chandler@gold.cchem.berkeley.edu*

The kinetics and stability of mesoscopic assemblies, including the formation of soft matter structures and the folding of proteins, often involve hydrophobic interactions. This lecture focuses upon the time scales associated with these interactions and demonstrates that solvent degrees of freedom dominate the rate determining steps for the assemblies they produce.

28.

PROTEIN SOLVATION IN MEMBRANES AND AT WATER-MEMBRANE INTERFACES. *Andrew Pohorille¹, Christophe Chipot², and Michael A. Wilson¹. (1) NASA Center for Computational Astrobiology, NASA-Ames Research Center, MS 239-4, Moffett Field, CA 94035, Fax: 65-604-1088, pohorill@raphael.arc.nasa.gov, (2) Laboratoire de Chimie Theorique, Universite Henri Poincare - Nancy I*

Different solvation properties of water and membranes mediate a host of biologically important processes, such as folding, insertion into a lipid bilayer, associations and functions of membrane proteins. These processes will be discussed in several examples involving synthetic and natural peptides. In particular, a mechanism by which a helical peptide becomes inserted into a model membrane will be described. Further, the molecular mechanism of recognition and association of protein helical segments in membranes will be discussed. These processes are crucial for proper functioning of a cell. A membrane-spanning domain of glycoporphin A, which exists as a helical dimer, serves as the model system. For this system, the free energy of dissociation of the helices is being determined for both the wild type and a mutant, in which dimerization is disrupted.

29.

HYDROPHOBIC AND HYDROPHILIC BEHAVIOR IN BULK WATER AND CONFINED SYSTEMS. *Jayendran C. Rasaiah, Department of Chemistry, University of Maine, Orono, ME 04469, Fax: 201-581-1191, rasaiah@maine.edu*

We discuss several unusual examples of hydrophobic and hydrophilic behavior in water and confined systems seen in computer simulations. Removing the charge on an ion makes it move more slowly if it is large and faster if it is small. This is explained by cage formation around large and desolvation of small ions when discharged. The entropy of solvation (R. Lynden-Bell) as a function of the charge has a local minimum near zero charge corresponding to hydrophobic solvation and, maxima on either side associated with cage breaking as the neutral solute forms cations or anions that exhibit hydrophilic solvation. The asymmetry in the solvation entropy as a function of charge is related to the asymmetry in the charge distribution of the water molecule. Recent computer simulations of water between plates in an open system (Vaitheeswaran) show that water can be depleted on applying an electric field between the plates—contrary to expectation from simple theories of electrostriction. Finally we discuss water in open nanotubes (Hummer, Noworyta and Waghe) of varying length and width which exhibit emptying and filling transitions modulated by solvent polarity and the strength of the nanotube-water interactions. The kinetics of these transitions will be discussed.

30.

SYNCHROTRON EXPERIMENTS AND THEORY: PURE WATER AS A FUNCTION OF TEMPERATURE. *Teresa Head-Gordon¹, Greg L.B. Hura², Daniela Russo¹, and Robert M. Glaeser³. (1) Department of Bioengineering, University of California, Berkeley, Berkeley, CA 94720, Fax: 510-486-6488, TLHead-Gordon@lbl.gov, (2) Biophysics Group, University of California, Berkeley, (3) Molecular and Cell Biology, University of California, Berkeley*

We have reported a new, high-quality x-ray scattering experiment on pure ambient water using synchrotron beam line 7.3.3 at the Advanced Light Source

at LBNL. Several factors contribute to the improved quality of our intensity curves for pure water that have been unattainable in the past, including use of a highly monochromatic source, a well characterized polarization correction, a Compton scattering correction that includes electron correlation, and better resolved and more accurate intensities using a modern CCD detector. We also show that in order to extract $g_{OO}(r)$ correctly, the spherical atomic scattering factors used to weight the various molecular structure factors must be modified to account for chemical bonding effects. We will present our newest synchrotron experiments and simulation analysis of liquid water over a temperature range between 2°C and 77°C.

31.

ISSUES OF SOLVATION IN DRUG DESIGN. *Terry R. Stouch, Computer Assisted Drug Design, Bristol-Myers Squibb, H23-07, PO Box 4000, Princeton, NJ 08543-4000, Fax: 609 818 3545, terry.stouch@bms.com*

Issues of solvation in drug design will be discussed including QSAR models for the rapid prediction of solubility, the role of water in protein/ligand binding, and solvation effects on molecular conformation.

32.

PROBING QUANTUM SOLVATION THROUGH INFRARED SPECTROSCOPY: DOPANT-INDUCED INFRARED ACTIVITY IN SOLID HYDROGEN. *Robert J. Hinde, Department of Chemistry, University of Tennessee, 336 Buehler Hall, Knoxville, TN 37996-1600, Fax: 865-974-3454, rhinde@utk.edu*

Atomic and molecular dopants in solid parahydrogen (pH_2) induce infrared activity in the pH_2 host crystal. This IR activity is associated with the pure vibrational $Q_1(0)$ transitions of pH_2 molecules near a dopant, and reflects the fact that dopants break the local symmetry of the solid pH_2 crystal. Analysis of these dopant-induced IR absorption features can thus provide information on the structure of dopant trapping sites in solid pH_2 and on the rovibrational dynamics of both dopant species and the pH_2 "solvent" molecules themselves. Hence these features provide considerable insight into the nature of impurity solvation in a highly quantum solid. We present diffusion quantum Monte Carlo studies of atomic and molecular dopants in solid pH_2 crystals and show how these studies help us unravel the relationship between spectral observables in doped solid pH_2 , such as the lineshapes of dopant-induced absorption features, and the underlying microscopic physics.

33.

NANOWIRES AS BUILDING BLOCKS FOR NANOSCALE SCIENCE AND TECHNOLOGY. *Charles M. Lieber, Department of Chemistry, Harvard University, 12 Oxford Street, Cambridge, MA 02138, Fax: 617-496-5442, cml@cmliris.harvard.edu*

Semiconductor nanowires and carbon nanotubes can function as active device elements and interconnects for the transport of charge carriers and excitons, and thus represent promising building blocks for a myriad of electronic and optoelectronic devices. To realize the potential of such structures requires basic scientific issues, including synthesis, physical properties and hierarchical assembly to be addressed. This presentation will provide examples focused on each of these fundamental scientific questions, and demonstrate how knowledge from such studies can be exploited for creating nanoscale electronic and optoelectronic devices. First, studies focused on the developing a general methodology for predictable and controlled nanowire synthesis will be described. Second, investigations of electrical transport and optical properties of individual nanowires and organized nanowire/nanowire junctions will be presented. The development of a range of specific devices based on the fundamental information obtained in these studies will be outlined. Lastly, efforts directed towards parallel, hierarchical organization of these building blocks into complex structures will be described. Challenges and goals for realizing nanotechnologies in the future will be discussed.

34.

AN INTEGRATED SYSTEMS-ORIENTED APPROACH TO MOLECULAR ELECTRONICS. *J Fraser Stoddart, Department of Chemistry and Biochemistry, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, CA 90095, Fax: 310-206-1843*

The area of molecular electronics has advanced rapidly during the past few years. While there are several reasons for this rapid progress, one enabling

factor has been the development of template-directed methods for synthesizing mechanically interlocked, bistable molecular compounds, which can be rendered amphiphilic and thus susceptible to being self-assembled at air-water interfaces, prior to their being transferred to appropriate substrates as stable monolayers. The time is now ripe to exploit these advances in both noncovalent and supramolecularly assisted covalent synthesis in the context of solid-state device fabrication. During the past three years, we have utilized a number of amphiphilic, bistable catenanes and rotaxanes as the active elements in solid-state switchable tunnel junction devices. In this talk, the template-directed synthesis and characterization of some amphiphilic bistable catenanes and rotaxanes will be discussed as a prelude to discussing the properties of devices fabricated using these mechanically interlocked compounds. An integrated systems-oriented approach toward building a highly efficient computing machine, based on molecular electronics components and self-assembly as a fabrication protocol, will be discussed. An architectural rationale for this approach will be presented. Simple devices, as well as working memory and logic circuits, will be presented. It will be suggested that a systems-oriented approach is likely to be critical for the development of any advanced nanotechnology that emerges out of fundamental nanoscience.

35.

MOLECULAR QUANTUM-DOT CELLULAR AUTOMATA. *Craig S. Lent, Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556, Beth Isaksen, University of Notre Dame, and Marya Lieberman, Department of Chemistry and Biochemistry, University of Notre Dame*

Computers dating back to those composed of electromechanical relay switches have relied on encoding binary information in the on or off state of a current switch. While this has been remarkably successful, it may not be the appropriate paradigm for realizing single-molecule devices. The quantum-dot cellular automata (QCA) approach to molecular electronics involves using charge configurations within molecules to encode bit information. No current flows through the molecule. This QCA approach has two substantial advantages over molecular electronic devices in which molecules are used as miniature current gates or current-carrying wires: power dissipation can be greatly reduced and true power gain is possible. QCA devices using single electron switching have been demonstrated using small metallic dots at low temperatures. We describe how this approach can support general-purpose computation, show the classes of molecules that we are studying as candidate QCA cells, and lay out the "roadmap" towards integration of these molecules in functional devices. We show theoretical results for QCA switching in molecules based on those proposed by Aviram.

36.

MOLECULES AND SUPRAMOLECULAR ARRAYS FOR QUANTUM-DOT CELLULAR AUTOMATA. *Marya Lieberman¹, Sudha Chellamma¹, Yuliang Wang¹, Qingling Hang², Gary Bernstein², and Craig S. Lent². (1) Department of Chemistry and Biochemistry, University of Notre Dame, 251 Nieuwland Science Hall, Notre Dame, IN 46556, Fax: 574-631-6652, mlieberm@nd.edu, (2) Department of Electrical Engineering, University of Notre Dame*

Quantum-dot cellular automata offer a paradigm for molecular computing in which Coulomb interactions between charges in molecular QCA cells are used to transmit and process information. QCA devices consist of arrays of QCA cells—for example, a line of cells forms a wire. QCA wires, AND/OR gates, and other devices have been tested at ultra-low temperature in lithographically fabricated metal-dot cells. For room temperature operation, however, QCA requires molecular-sized cells, which presents some intriguing questions to the chemist. Are there any types of molecules which are suitable as candidates for QCA? If so, how can they be fashioned into the necessary arrays to fabricate QCA devices? Mixed-valence metal complexes appear to meet many of the requirements for molecular QCA, including stability of the mixed-valence oxidation state, size, and potential for self-assembly, but most are encumbered with counterions which may impair the fidelity of QCA operation. We have recently synthesized a neutral mixed-valence ruthenium complex, $[Ru_2(acac)_4B(pz)_4]$, which addresses this counterion problem. We will also discuss the formation of self-assembled monolayers of mixed-valence complexes and describe a novel approach to patterning them to form the types of supramolecular arrays which are necessary for QCA devices.

37. STRUCTURE ANALYSIS OF DISCOTIC LIQUID CRYSTALS CONTAINING 1,3,5-TRI-SUBSTITUTED BENZENE CORE AND CONJUGATED OXADIAZOLE ARMS USING XRD AND COMPUTER SIMULATION METHODS. *C.R. Park, and S.H. Gihm, School of Materials Science and Engineering, Seoul National University, Seoul 151-744, South Korea, Fax: +82-2-885-1748, crpark@snu.ac.kr*

Conjugated and three-armed discotic liquid crystals consisting of a 1,3,5-tri-substituted benzene core, oxadiazole-based conjugated calamitic arms, and one (denoted 3a) or three (denoted 3b) dodecyloxy tails per each arm, were synthesized. From small angle X-ray scattering (SAXS) data, the hyperstructure of 3b could be identified, but, for 3a, it was not straightforward since 3a gave a scattering profile similar to a powder pattern. So, to elucidate the hyperstructure of 3a, Polymorph (that consists of Monte Carlo, energy minimization and cluster analysis) and Rietveld methods were adopted. The starting model structures created by Polymorph were refined using Rietveld method. With the obtained structure in this manner, different mesophase behaviors of 3a and 3b could be explained in terms of different modes of transverse arm-to-arm interactions and space-filling effect.

38. SILICON SEGREGATION FROM SiO_{2-x} SUB-OXIDES: DYNAMICAL MONTE CARLO STUDY. *P. Ballone, Department of Physics, University of Messina, Italy, Contrada Papardo, 98166 Messina, Italy, p.ballone@fz-juelich.de, and R. O. Jones, I F F, Forschungszentrum Juelich, Germany, 52425 Juelich, Germany, Fax: 01149-2461-612850, r.jones@fz-juelich.de*

The Si/SiO₂ system is one of the most important in semiconductor technology, and Si nanocrystals embedded in amorphous SiO₂ are of great interest for advanced FLASH memories and for opto-electronic applications. We simulate the segregation of Si dots from metastable SiO_{2-x} (x=0.2-1) sub-oxides by a nearest neighbor force field model, using Monte Carlo to sample atomic positions and bond interchanges. We follow the motion for times (estimated from the average atomic displacement) approaching mesoscopic scales. On cooling from a random Si-SiO₂ alloy at high temperature (T=2000 K), Si dots form progressively as T decreases to 1000K. We investigate the nucleation process and the interface between Si and SiO₂ matrix. Properties are dominated by potential energy considerations at the lowest temperatures, while entropy contributions from the Si-Si and Si-O bond configuration are important at the temperatures at which segregation takes place. There is an intriguing formal analogy with equilibrium polymerization.

39. MOLECULAR TOOLBOX: A PHYSICAL ORGANIC CHEMISTRY APPROACH TO MOLECULAR ELECTRONICS. *Jun Hu, Chalermchai Khemtong, and Yubiao Liu, Department of Chemistry, The University of Akron, 190 E. Buchtel Comm, Akron, OH 44325-3601, Fax: 330-972-7370, jhu@uakron.edu*

Molecular tunneling junctions based on pi-conjugated organic compounds display nonlinear current/voltage responses such as negative differential resistance (NDR). They are promising components for molecular electronic devices. The I/V characteristics can be considered as the manifestations of the electronic structure features of electrode/molecule interconnections and the "molecular wires" that form the tunneling junctions. To elucidate the mechanisms of the functions of the devices and to establish structure-property relationships in the molecular electronic devices, it is necessary to systematically investigate the I/V behaviors of a set of "molecular wires" with varied functional groups and electronic structure features. We will report our recent progresses in construction of a new set of pi-conjugated organic compounds that form structurally well-defined molecular tunneling junctions by molecular self-assembling on electrode surfaces.

40. ORGANIZING RESEARCH BY GENE FAMILY. *Mark Murcko, Vertex Pharmaceuticals, Cambridge, MA 02139-4242, mark_murcko@vpharm.com*

Abstract not available.

41. IDENTIFYING FUNCTION FIRST TOWARD LEAD DISCOVERY IN THE POST-GENOMIC ERA. *Susan M. Baxter, Peter Domaille, Stephen Betz, Stacy Knutson, Keith Burdick, Brian Hoffman, and Jacquelyn Fetrow, GeneFormatics, Inc, 5830 Oberlin Drive, San Diego, CA 92121, susanbaxter@geneformatics.com*

In the post-genomic era, pharmaceutical researchers must evaluate vast numbers of protein sequences and formulate intelligent strategies for identifying valid targets and discovering leads against them. We have chosen to identify and classify proteins based on their active sites using large-scale computational methods to identify functional families of proteins across the human proteome. Recently we have focused on the serine proteases, proteins that participate in a variety of biological pathways relevant to human disease. Once serine protease panels are identified from genomic or expression data, the three-dimensional structure and the chemical characteristics functional sites are analyzed to identify similarities and differences, recognizing they are key to specific drug design. We use a combination of virtual screening tools and complementary experimental methods, including NMR screening, to identify small molecules that bind generally to several family members or specifically to one family member, reducing the possibility of cross-reactivity across the panel.

42. DRUG DISCOVERY ON A PROTEOMIC SCALE. *Michael V. Milburn, Plexixikon Inc, 91 Bolivar Dr, Suite A, Berkeley, CA 94710, Fax: 510-548-8014, mmilburn@plexixikon.com*

Plexixikon utilizes a proprietary drug discovery platform of compound screening that combines state of the art high-throughput co-crystallography, informatics and parallel biochemical assays to rapidly discover high quality lead compounds for a large set of targets.

Plexixikon's discovery platform is designed to discover low molecular-weight chemical scaffolds that act broadly on protein families sharing common folds, or domains. We will demonstrate how Plexixikon, through a combination of low affinity parallel biochemical screening of the company's scaffold compound collection and high throughput co-crystallographic structure analysis of hits co-complexed with the target, identifies multiple potential scaffolds for drug discovery. Using the available co-structure information, discovery chemistry synthesizes additional scaffold compounds for screening and co-crystallization with the target. Those scaffolds that are novel, chemically tractable, and best suited for drug discovery are further developed to yield proprietary small molecule drug leads for pharmaceutically relevant protein targets.

We will show several examples of chemical scaffolds discovered through Plexixikon's unique discovery platform and their use in rapidly discovering leads for specific new protein targets.

43. HOMOLOGY MODELLING AND SIMULATION STUDIES ON THE ORPHAN NUCLEAR RECEPTOR ERR- γ . *Wendy Cornell¹, Kiyeon Nam², Romain Wolf¹, and Paul Marshall¹. (1) Novartis Pharmaceuticals, Summit, NJ 07901, wendy.cornell@pharma.novartis.com, (2) Novartis Pharmaceuticals and UMDNJ*

Sequencing efforts on the human genome have led to a wealth of data, some of which can be used for protein modelling studies. Here we describe the development and application of a homology model for the estrogen-related receptor- γ (ERR- γ), a member of the steroid hormone binding or "NR3" family. ERR has been classified as an "orphan" nuclear receptor, since a natural ligand is not known. ERR- γ is most closely related to the estrogen receptor ER- α , which was used as a template for homology modelling. Members of the NR3 family have been implicated in a number of diseases such as diabetes, cancer, and reproductive diseases. The ERR- γ homology models and ER- α crystal structures were employed in MM-GB/SA simulation studies aimed at reproducing the opposing effects of the synthetic estrogen DES on the two receptors. These studies provide a structural basis for this behavior and suggests insights into the functional role of ERR- γ .

44. PATENTING 3-D STRUCTURES. *Alicia A. Russo, Baker Botts, LLP, New York, NY 10112-0228, alicia.russo@bakerbotts.com*

Structural genomics is a new field for enhancing all aspects of structural biology and has become a major focus for the development of drugs. New initiatives in structural genomics, both public and private, aim to solve protein structures at

previously unprecedented rates using high-throughput systems. While structural genomics and the information obtained therefrom have become a major focus for drug development efforts, the intellectual property relating to structural genomics has not been readily afforded patent protection. For example, neither the coordinates of protein structures themselves nor the classes of inhibitor molecules identified using the coordinates have been patented. Notwithstanding, the number of granted patents relating to such information has been increasing suggesting a tendency for the United States Patent Office to grant such patents. This may be due to a better understanding of the inventions and/or to recent case law relating to patentable subject. Accordingly, assuming a patent application directed to a structural genomics invention provides an adequate written description and enablement of the invention and that the invention is nonobvious and novel, patent protection should become more feasible for inventions relating to structural information, including the coordinates themselves (as defining a three-dimensional structure of a particular protein); methods of using the coordinates in rational drug design; methods of treatment using a drug having certain structural characteristics; and the drugs themselves that are designed by the methods of rational drug design. This is likely to further foster structural genomics efforts since patent protection is needed to ensure adequate monetary returns for costly and time consuming research and development.

45. EXTENDING THE MM3 FORCE FIELD FOR APPLICATION TO THE STRUCTURE-BASED DESIGN OF METAL ION HOSTS. *Benjamin P. Hay, W. R. Wiley Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, PO BOX 999, Richland, WA 99352, Fax: 509-375-6631, ben.hay@pnl.gov*

One strategy for the design of metal ion hosts is to couple molecule building algorithm with scoring functions that are used to prioritize the candidate structures. Because of their speed, molecular mechanics (MM) models provide a method that is suitable for evaluating large numbers of candidates. Moreover, the process of parameterizing MM models yields structural design criteria that serve as input for molecular builders. This talk will describe how the MM3 force field, originally developed for organic compounds, has been extended to treat metal complexes formed from a variety of ligand and metal types. Examples will illustrate how the extended MM3 models have been used to rank host structures with respect to the degree of binding site organization for targeted metal ions and facilitate the design of more efficient host architectures.

46. INTERPLAY OF BONDING THEORY AND FORCE FIELDS FOR ORGANOTRANSITION METAL COMPLEXES. *Clark R. Landis, Department of Chemistry, University of Wisconsin-Madison, 1101 University Avenue, Madison, WI 53706, Fax: 608-262-6143, landis@chem.wisc.edu*

Fundamental to the application of empirical force field methods to transition metals is the accurate and flexible description of molecular shapes as a function of electron counts, coordination numbers, and bond orders. Simple electron counting schemes based on Lewis-like structures provide a systematic, programmable method for describing the molecular framework. Hybridization and resonance considerations enable derivation of effective potential energy expressions for the description of molecular shapes in a variety of bonding situations. These methods are applied to molecular mechanics description of a large variety of organotransition metal complexes containing metal-ligand single and multiple bonds. This method works best when the metal-ligand bonding environment is largely covalent and its description does not require a large number of resonance structures.

47. DE NOVO STRUCTURAL PREDICTION OF TRANSITION METAL COMPLEXES. *Thomas Cundari, Department of Chemistry, CRONIUM, University of Memphis, Memphis, TN 38152-6060, Fax: 901-678-3447, tcundari@memphis.edu, and Corneliu Buda, Department of Chemistry, University of Memphis*

Transition metal complexes play an important role in applications ranging from catalysis to medicine to waste remediation. In any computer-aided design scenario the prediction of geometry is the first and most important task. Predicting the geometry of transition metal complexes "from scratch" or de novo remains an important, albeit elusive goal. Layers of difficulty beyond the

conformational complexity that is typically observed for organic species are encountered in de novo structural prediction of transition metal complexes. Issues include geometric (e.g., cis versus trans or fac versus mer octahedral complexes), structural (e.g., trigonal bipyramid versus square pyramidal five-coordinate complexes), coordination (e.g., axial versus equatorial in trigonal bipyramid), spin (low versus intermediate versus high spin states), and linkage (e.g., NCS versus SCN, CN versus NC, S-DMSO versus O-DMSO, etc.) isomerism. A combination of molecular mechanics and semiempirical quantum mechanics for de novo structural prediction of transition metal complexes will be discussed.

48. SELECTIVITY IN THE ADDITION OF DIALKYL ZINC TO ALDEHYDES FROM Q2MM FORCE FIELD MODELING. *Per-Ola Norrby, Department of Chemistry, Organic Chemistry, Technical University of Denmark, Building 201, Kemitorvet, DK-2800 Kgs. Lyngby, Denmark, Fax: +45-45933968, pon@kemi.dtu.dk, Torben Rasmussen, Department of Theoretical Chemistry, University of Lund, and Malcolm B. Gillies, Organic Chemistry, Department of Chemistry, Technical University of Denmark*

A new Q2MM force field has been developed for the addition of dialkyl zinc to aldehydes. The Q2MM methodology uses a quantum chemical description of the selectivity-determining reaction step to create a transition state force field, that is, an empirical force field treating the TS as if it were a ground state. Application of the Q2MM force field allows a rapid comparison of transition state conformations and prediction of stereoselectivity in the reaction, but not a determination of absolute reaction rates. The accuracy of and the sources of error in the Q2MM approximation applied to metal-catalyzed reactions will be discussed.

49. IMPORTANCE OF CONFORMATIONAL SEARCHING: UNDERSTANDING PROCHIRAL OLEFIN BINDING TO CHIRAL CYCLOPENTADIENYL COMPLEXES OF RHENIUM AND THE CASE OF THE MISSING CONFORMERS. *David White¹, Aaron M. Gillespie¹, Glenn R. Morello¹, and Michael A. Freeze². (1) Department of Chemistry, University of North Carolina at Wilmington, 601 S. College Rd, Wilmington, NC 28403, whitedp@uncwil.edu, (2) Department of Mathematics and Statistics, University of North Carolina at Wilmington*

Searching the conformational space of a molecule is a particularly difficult task when modeling organometallic catalysts. Most conformational searching algorithms used in organometallic chemistry apply a combination of Monte Carlo methods and molecular dynamics to sample the conformational space of the molecule. However, the consequence of missing important conformers can easily be overlooked. In this talk, the use of molecular mechanics (MM) to demonstrate that steric effects are responsible for the ability of the coordinatively unsaturated $[\eta^5\text{-C}_5\text{H}_5]\text{Re}(\text{PPh}_3)(\text{NO})^+$ fragment to stereospecifically bind prochiral α -olefins will be presented. Geometric refinement of the MM-optimized geometry at the PM3(tm) level of theory in semiempirical quantum mechanics (SEQM) will be discussed as will accurate energy determination by means of DFT methods. "Missing conformers" from the MM search were discovered with a mixed MM-DFT methodology, which will be discussed. Analysis of conformational spaces with fuzzy set theory will be presented. Finally, a computational estimate of diastereoselective excess will be discussed.

50. THE COUPLING OF SOLVENT AND ELECTRONIC DYNAMICS IN EXCITED STATE RELAXATION. *Peter J. Rossky, Institute for Theoretical Chemistry, Dept. of Chemistry & Biochemistry, University of Texas, MC A5300, Austin, TX 78712, Fax: 512-471-1624, rossky@mail.utexas.edu*

The dynamics of electronic excited states play an important role in a variety of scientific and engineering contexts, including biological and technological processes. In this presentation, the strongly coupled evolution of electronic excited states and the nuclear degrees of freedom that constitute their environment in the condensed phase will be discussed. The common themes that emerge will be developed via results obtained from mixed quantum-classical treatments of a variety of specific examples. These include solvated electrons, solvated intramolecular charge transfer species, and solvated proteins involved in biological electron transfer. The results demonstrate that electronic relaxation dynamics and nuclear dynamics are intimately coupled in condensed phases.

51. DYNAMICS IN LIQUID WATER: COMPUTATIONAL APPROACHES TO CALCULATING ULTRAFAST INFRARED SPECTROSCOPY OBSERVABLES. *James L Skinner, Department of Chemistry, University of Wisconsin, 1101 University Avenue, Madison, WI 53706, Fax: 608-262-9918, skinner@chem.wisc.edu*

Very recent ultrafast infrared pump-probe spectroscopy experiments have been used to measure vibrational energy relaxation and structural dynamics in liquid water (actually dilute HOD in D₂O). In this talk I describe our efforts to calculate several relevant spectroscopic observables. Our approach treats the vibrations of the HOD molecule quantum mechanically, and the rotations and translations of all molecules classically. Because of the strong hydrogen-bonding interactions in water, and in order to calculate the experimentally-measured vibrational Stokes shift, we have implemented a novel renormalization scheme whereby the "system" and "bath" Hamiltonians are determined self-consistently for a given vibrational state of the HOD molecule. Our results for the different observables are generally in good agreement with experiment. In particular, we can shed light on the nature and time scales of hydrogen-bond breaking and making dynamics, as measured by transient vibrational hole-burning experiments.

52. SOLVATION DYNAMICS IN MULTICOMPONENT SYSTEMS AND SUPERCRITICAL FLUIDS. *Branka M. Ladanyi, Department of Chemistry, Colorado State University, Fort Collins, CO 80523, Fax: 970-491-3361, bl@lamar.colostate.edu*

Solvation dynamics (SD) is the time-evolution of the change in solute-solvent interaction energy, usually brought about by solute electronic excitation. Analysis of molecular trajectory data obtained from molecular dynamics (MD) computer simulation can provide a wealth of information about the molecular mechanisms of SD. This information includes the types of molecular motions that contribute to different SD time scales, the effects of the location of the solute relative to interfaces in heterogeneous systems, and the influence on SD of the build-up local concentration enhancement in mixtures and of local density enhancement in supercritical fluids. I will discuss how analysis of MD data has helped us understand the molecular mechanisms contributing to SD in reverse micelles, benzene-acetonitrile mixtures, and supercritical CO₂.

53. HOW THE VIBRATIONAL MODES OF A POLYATOMIC MOLECULE SEE A SOLVENT. *Richard M. Stratt, and Yuqing Deng, Department of Chemistry, Brown University, 324 Brook St., Providence, RI 02912, Fax: 401-863-2594*

Vibrationally excited polyatomic molecules can relax in a variety of different ways in solution: the excess energy can be dissipated directly to the solvent, or it can be redistributed between any of a number of different intramolecular modes, with the liquid absorbing (or supplying) just enough energy to make the process work. What we consider is how the solvent participates in these mechanistic choices. Using the prototypical example of a symmetric linear triatomic molecule, we compare the molecular origins of the vibrational friction for the direct vibrational cooling of the symmetric and antisymmetric stretching modes and contrast both of those with intramolecular vibrational energy transfer between these two modes. Instantaneous-normal-mode analysis reveals that a solid-state-like perspective is a plausible starting point for understanding these processes: the solvent does define a band of intermolecular vibrations, and it is only when the energy being transferred falls within that band that the solvent can easily accept energy from a solute. However, it is also possible to discern some more liquid-state-specific details. Despite their different symmetries and different kinematic requirements, all of the different relaxation pathways are apparently driven by the dynamics of the same instantaneously-nearest solvents.

54. ENERGY RELAXATION IN LIQUIDS: HDO IN D₂O. *Philippe A. Bopp, Institut für Physikalische Chemie, Technische Universität Darmstadt, Petersenstrasse 20, D-64287 Darmstadt, Germany, Fax: (+49)(0)6151 16 4298, pab@loriot.ismc.u-bordeaux.fr, and Samuel J.-F. Cousineau, LPCM, Université Bordeaux I*

We report first results on non-equilibrium classical molecular dynamics(NEMD) simulations of vibrationally excited HDO molecules in D₂O. The simulations, run under NVE conditions, attempt to mimic time resolved infrared experiments in

which the OH-oscillator of HDO molecules diluted in D₂O are excited and their relaxation is observed. The simulations use a flexible water model with an effective intramolecular part tuned to reproduce the gas-liquid shifts in the three intramolecular vibrational frequencies. The observed relaxation times are in reasonable agreement with experimental data. We attempt to locate the main paths for the energy flows out of the excited degree of freedom.

55. A NEW THEORY FOR SOLUTION-PHASE VIBRATIONAL ENERGY RELAXATION. *Eitan Geva, and Qiang Shi, Department of Chemistry, University of Michigan, 930 North University, Ann Arbor, MI 48109-1055, Fax: 734-647-4865, eitan@umich.edu*

We present a new class of general quantum-mechanical expressions for the vibrational energy relaxation (VER) rate constant. The new theory builds on the formal similarity between the VER problem and the reactive-flux formalism for barrier crossing. The new expressions are derived from non-linear response theory, and put the VER rate constant in terms of equilibrium correlation functions of vibrational observables. We show that these new expressions reduce to the commonly used Landau-Teller (LT) formula in the limit of weak system-bath coupling. It is argued that the new theory has several important advantages over the LT formula: (1) The new theory avoids the assumption of weak system-bath coupling, which is explicit in the LT theory, and can therefore be used in order to assess the validity of this assumption for a given system; (2) The VER rate constant is expressed in terms of "local" vibrational observables, as opposed to the "global" force that appears in the LT formula; (3) The new theory offers a more direct avenue to the calculation of high-frequency VER rate constants, and avoids the evaluation of the high-frequency tail of the LT power spectrum; (4) Unlike the LT formula, some of the new expressions can be evaluated directly from Centroid Molecular Dynamics (CMD) simulations, as a way for estimating quantum corrections. Results obtained from applying the new theory within the framework of CMD will be reported for several anharmonic model system.

56. VIRTUAL SCREENING WITH TOPOMERIC COMFA. *Richard D. Cramer, Tripos, 1699 South Hanley Road, St. Louis, MO 63144, Fax: 505-995-4439, cramer@tripos.com*

Topomerically aligning the fragments of the input structures has proven to be a remarkably robust starting point for automated CoMFA studies. Acceptable results were obtained in repeating all 15 of 15 published studies, with overall average q² and errors of "true prediction" indistinguishable from the literature summary results. The resulting CoMFA models are perfectly suited for searching the vast ChemSpace database (>10¹³ synthetically accessible drug-like structures), while targeting improved activity as well as similarity, at negligible performance cost. Thus a general "high-throughput-design" methodology exists, capable of shortening high quality information-based complete design-test-synthesis cycles to weeks, in the real world.

57. BUILDING QSAR MODELS FROM LARGE SCREENING SETS. *Paul Blower¹, Kevin Cross¹, Michael Fligner², Roman Khramets¹, Glenn Myatt¹, and Joseph Verducci². (1) LeadScope, Inc, 1245 Kinnear Rd, Columbus, OH 43212, pblower@leadscope.com, (2) Department of Statistics, The Ohio State University*

We have developed novel methods for building preliminary structure-activity relationship (SAR) models from the large, heterogeneous data sets used in primary HTS screens. For these models, each compound is described by a fingerprint consisting of thousands of binary features. We first group active compounds in the training set into structurally meaningful categories, either by simple clustering or stochastic tree classification. The next stage is to identify key features that distinguish each active group from inactive compounds in the neighborhood of the group. These features are likely candidates for pieces of a macrostructure that governs activity. Once we have identified the macrostructures of each group, we use a probability model to infer weightings associated with each such structure appearing in the group. Finally, we build a local prediction model based on the weighted distances from each macrostructure. This SAR model can then be used to efficiently select new compounds for follow-up screening.

58.

4D-QSAR AND COMFA OF LIGAND ACTUATORS FOR THE ECDYSONE RECEPTOR-MEDIATED GENE TRANSCRIPTION SYSTEM. *Robert E. Hormann¹, Orestes Chortyk¹, Dean E. Cress¹, Anton J. Hopfinger², and Christine S. Thompson¹.* (1) RHeoGene, L.L.C, P.O. Box 949, Spring House, PA 19477-0949, Fax: 215-619-1665, RHormann@Rheogene.com, (2) Laboratory of Molecular Modeling and Design, U. Illinois at Chicago

The ecdysone receptor (EcR), a nuclear hormone receptor essential to the gene regulation of insect molting, has been engineered into gene regulation systems potentially useful for biotherapeutic protein production, proteomics, drug discovery, and ultimately, gene therapy. Comparative CoMFA and 4D-QSAR analyses are reported for diacylhydrazine actuation of Bm EcR-mediated b-galactosidase reporter gene transactivation. The relative merits of each QSAR technique are examined, and the resulting QSAR models are discussed in view of an EcR homology model and EcR affinity of the natural ecdysteroid ligands.

59.

MOTION OF MOLECULAR PROBES IN THE FUNCTIONAL SITES OF PROTEINS. *Tamas Kortvelyesi, Sheldon Dennis, and Sandor Vajda, Department of Biomedical Engineering, Boston University, 44 Cummington Street, Boston, MA 02215, kortve@bu.edu*

In computational mapping of a protein small organic molecules are used for the determination of the binding site. The probes cluster in the binding pocket, and the clusters can be further divided into subclusters with different configurations (orientations and conformations). Each subcluster shows a number of hydrophobic and hydrophilic interactions with specific residues, and the analysis provides substantial information on the binding site. Molecular dynamics (MD) calculations were performed to confirm that the subclusters represent the different rotational/translational states of the bound ligand. In particular, we studied the interaction of small organic molecules with hen egg white lysozyme (HEWL) using (i) an explicit water model in a periodic box with the GROMACS program, (ii) an implicit water model with the CHARMM/ACE program. The reaction paths and transition states between the configurations found in the mapping calculations were also calculated using CHARMM/ACE. In the explicit water MD calculations we found that small solvent molecules (e.g. methanol) show a wide range of translational and rotational movement, resulting in more configurations than larger molecules, which move less during the same simulation time. This may occur because of the low free energy barriers between the configurations.

60.

OPTIMIZED LEAD DISCOVERY COMBINING LIGAND-BASED AND STRUCTURE-BASED METHODS. *William Mydlowec, Guido Lanza, and Jessen Yu, Pharmix Corp, 200 Twin Dolphin Drive, Suite F, Redwood Shores, CA 94065, bill@pharmix.com*

We present a method for lead discovery that combines ligand-based and structure-based approaches. In the first phase of this method, an initial compound library is evaluated using molecular docking to simulate activity against a given protein target. A predictive model is then generated to explain this compound-activity data, and a refined screening library is designed to best improve the model by maximizing information (entropy) expected a priori from each new compound. By repeating this process over several iterations, we converge on the most informative model of activity having performed a minimal number of docking simulations. Then in the second phase of this method, a large virtual library is rapidly screened against the learned model to produce an enriched subset of potentially-active compounds. Only this enriched set need be subjected to docking simulation in order to determine the subset of true actives from the entire virtual library. We present applications of this method to several current drug targets, and show that enrichments of one or two orders of magnitude are possible over exhaustive docking.

61.

LIGAND-BASED LEAD DISCOVERY AND APPLICATIONS TO "SCAFFOLD HOPPING". *Guido Lanza, Jessen Yu, and William Mydlowec, Pharmix Corp, 200 Twin Dolphin Drive, Suite F, Redwood Shores, CA 94065, guido@pharmix.com*

We describe a method for ligand-based lead discovery that is capable of finding active compounds in multiple chemical classes. Multiconformer analysis is performed on all ligands, and 3-D molecular descriptors are generated from features and shapes. A machine learning algorithm is then applied to produce a predictive model that explains both active and inactive ligands. We present applications of this method to a wide range of protein targets, including GPCRs, proteases, kinases, isomerases, DNA/RNA-related proteins, and others. Using a training set of several thousand ligands with published activity data, we generated a predictive model for each protein target. We then validated each model by using it to filter a new set of topologically-dissimilar compounds with known activities. The results presented demonstrate the effectiveness of this method in "scaffold hopping" on many current drug targets.

62.

MINING THE NCI'S TUMOR SCREENING AND GENOMIC DATABASES: RELATING MOLECULAR TARGETS TO CANDIDATE LIGANDS. *David G Covell¹, Alfred A. Rabow², Anders Wallqvist², and Robert H. Shoemaker³.* (1) DTP, DCTD, NCI, NCI-Frederick, 430/215, Frederick, MD 21702, Fax: 301-846-5762, covell@helix.nih.gov, (2) NCI-Frederick, (3) STB, DTP

Computational tools are assembled for examinations of bioinformatics questions related to the pharmacology and drug discovery of new anticancer agents. Efforts are currently focused on providing access to the NCI's anti-cancer drug screening and genomic databases. Towards this end a suite of internet accessible (<http://spheroid.ncifcrf.gov>) computational tools is available for database explorations. Initial efforts identified relationships between chemotypes in the NCI's 60 tumor cell screen and their effect on four major classes of cellular activities: mitosis(M), nucleic acid synthesis(S), membrane transport and integrity(N) and phosphatase and kinase mediated cell cycle regulation(P). Detailed analyses of the chemicals within these cellular activity classes finds a strong dependence on chemotype and coherence among cellular response patterns. These results provide a basis for interfacing this screening data with companion mRNA microarray data obtained on the same tumor cells. Attempts are made to link constitutive tumor gene expression levels to specific biological responses. These results isolate general toxic chemotypes from agents with a (putative) cellular chemical activity. Examples are provided for classes of agents known to act in response to induced DNA stress and multidrug resistance. Further analyses identify a number of significantly correlated genes that are involved in constitutive cellular growth, signal transduction and maintenance of ion homeostasis.

63.

NOVEL BIOINFORMATICS TECHNIQUES IN FUNCTIONAL GENOMICS. *Atul J Butte, Division of Endocrinology and Informatics Program, Children's Hospital, Boston and Harvard Medical School, 300 Longwood Avenue, Endocrinology 333LW6, Boston, MA 02130, Fax: 801-729-7231, atul_butte@harvard.edu*

We describe two packages freely available to the academic genomics community. Relevance Networks comprehensively compares all measured genes and phenotypic measurements and builds networks from the highest scoring pairs. Advantages of this method include (1) negative associations are shown (e.g. those from tumor suppressors), (2) disparate data types are included (i.e. clinical, expression, and phenotypic), and (3) multiple connections are allowed (e.g. transcription factors may regulate multiple other genes). The output relevance networks are linked to a database and freely available web-based system (Unchip), taking accession codes and returning the latest information about each probe set. Having analyzed over 600 microarrays, we have examples for and learned (1) not all pathways will be reverse-engineered, (2) looking at simultaneous expression ignores biology taking time, (3) discovered diagnostic models don't imply the underlying molecular physiology, (4) with rapidly changing information on genes already measured, one is never finished analyzing a microarray dataset.

64.

MODEL-CENTRIC DATA INTEGRATION AND THE DEVELOPMENT OF GENOME-SCALE MODELS OF METABOLISM. *Christophe H. Schilling, Genomatica, Inc, 5405 Morehouse Drive, Suite 210, San Diego, CA 92121, Fax: 858-824-1772, cschilling@genomatica.com*

Systems biology is rapidly emerging as a central research paradigm for the integrated analysis of cellular function. At the core of this paradigm is the need to develop computer models of cell function, organ function, and ultimately whole-animal function. These models hold the potential to be catalysts for a new era of discovery, whereby setting the stage for a fundamental change in the way drugs are discovered developed. A critical component to the successful implementation of in silico models is the ability to tightly integrate them with high throughput experimental technologies, such as DNA sequencing, array-based gene and protein expression technologies, and approaches for cellular phenotyping. In this presentation we will discuss the iterative model development process and current efforts in the development and application of genome-scale models of cellular metabolism. This will be followed by examples of how these models are integrated to experimental technologies and used to efficiently guide research with particular emphasis on the integration of gene expression data gathered from recent experiments on common model organisms.

65.

MOLECULAR PROFILING IN DRUG DISCOVERY AND DEVELOPMENT. *Roger G. Ulrich, Rosetta Inpharmatics, Merck & Co, 12040 115th Avenue NE, Kirkland, WA 98034, Fax: 425-821-5354, roger_ulrich@merck.com*

Abstract text not available.

66.

EXPANDING PREDICTIVE TOXICOLOGY APPLICATIONS USING TOXICOGENOMICS. *Donna L. Mendrick, Gene Logic, Inc, Gaithersburg, MD 20878, dmendrick@genelogic.com*

Gene Logic is using a multi-platform gene expression approach to develop the ToxExpress database. We are identifying markers of toxicity in rats and determining how these relate to human responses by testing compounds in vivo in rats and in primary human and rat cells. The samples are profiled on Affymetrix's rat (>26,000 genes and ESTs) or human (>33,000 genes and ESTs) GeneChips® and selected samples are investigated further with READS, our proprietary differential display technology. Selected gene expression changes are confirmed using Q-RT-PCR. Using our proprietary algorithms along with commercially available statistical and clustering packages to mine our database, we are building predictive models using gene expression changes seen with model toxicants and correlating these changes with classical toxicologic parameters such as clinical chemistry, hematology, and histopathological changes. The compounds used to build the database are commercially available, marketed pharmaceuticals and toxicants. To enhance our ability to detect species-specific toxicity, compounds include those that induce toxicity only in rats, only in humans, and in both species.

67.

GROUND STATE NEAR ATTACK CONFORMERS (NACS) AND TRANSITION STATES IN ENZYME REACTION. *Thomas C. Bruice, Department of Chemistry & Biochemistry, University of California, Santa Barbara, Santa Barbara, CA 93106, tcbuice@bioorganic.ucsb.edu*

Among the many conformations of enzyme and substrate present at the active site of E-S are those that closely resemble the stereochemistry of the enzyme and transition state in E-TS. These conformers have been termed Near Attack Conformers or NACs. For an example, in an SN₂ displacement on carbon by a negative oxygen functionality a NAC would have a distance between O and C of ~3.2 Å and an angle of attack within ± 15° of the forming bond in the transition state. In studies of a number of enzymes the NAC populations consist of between 2-70% of the ground state conformers of E-S. Thus NACs may be the main ground state conformer or exist about 2 kcal above the other conformers of E-S. The lowest free energy barrier for an enzymatic reaction is from E-NAC directly to E-TS. The advantage of the near attack conformer is that it forms

spontaneously at ambient temperature (formation is free of strain) and, as a turnstile to enter the transition state, allows this process to take place without more than minor changes in protein conformation. That is to say the preformed structure or the induced fit structure of the enzyme ground state remains much the same when NAC is converted to TS — very little reorganization energy. The advantage of NAC formation in the enzyme as compared to water solvent will be discussed. Examples comparing enzyme binding of NAC and TS will be presented.

68.

COMBINING QUANTUM MECHANICS, MOLECULAR MECHANICS, AND MOLECULAR DYNAMICS TO MODEL ENZYME AND ANTIBODY CATALYSIS. *K. N. Houk, Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, CA 90095-1569, Fax: (310)206-1843, houk@chem.ucla.edu*

The exploration of a variety of enzyme and antibody catalyzed organic reactions will be described in this lecture. QM/MM techniques, combining density functional theory and AMBER molecular mechanics, have been used. QM/MM techniques permit the exploration of bond-making and bond-breaking with quantum mechanics while including non-bonding interactions and dynamics of proteins with force-field techniques. The simultaneous use of QM and MM has clear advantages, but the serial application of QM and MM methods can also provide valuable insights. Density functional theory based quantum mechanical transition structures, MD docking procedures, and force-field-based free energy perturbation or linear interaction energy calculations provide quantitative insights into the origins of catalysis by proteins.

69.

DEVELOPMENT AND APPLICATIONS OF TRANSITION STATE FORCE FIELDS. *Malcolm B. Gillies, Organic Chemistry, Department of Chemistry, Technical University of Denmark, Bldg 201 Kemitorvet, 2800 Kgs. Lyngby, Denmark, Fax: +45-4593-3968, malcolm@kemi.dtu.dk, and Per-Ola Norrby, Department of Chemistry, Organic Chemistry, Technical University of Denmark*

The Q2MM method for the construction of transition state force fields treats transition states as minima on a molecular mechanics potential energy surface. Typical applications have involved automated parameterization using data from quantum chemical calculations, and have allowed accurate prediction of experimental selectivities in asymmetric synthesis. We present further developments using the Q2MM method, including new applications to mechanistic studies, improvements to the parameterization procedure, and enhancements of the software.

70.

EMPIRICAL VALENCE BOND STRATEGIES FOR PHOSPHORYL AND NUCLEOTIDYL TRANSFER REACTIONS. *Jan Florian, Avital Shurki, and Arieh Warshel, Department of Chemistry, University of Southern California, 3620 S McClintock Av., SGM418, Los Angeles, CA 90089-1062, Fax: 213-740-2701, florian@usc.edu*

Reactions that lead to a formation and/or cleavage of PO bonds represent arguably the most important and diverse class of biological reactions. Our strategy of a systematic investigation of these reactions in water and proteins by computer simulations based on empirical valence bond (EVB) methodology is outlined. As a starting point we consider model reactions in water by ab initio methods and evaluate corresponding free-energy surfaces in water. These calculations and key experiments, including linear free-energy relationships, are then used to build the consensus free-energy surface and to refine parameters for the corresponding EVB surface. The EVB surfaces used for solution reactions are also used for studies in proteins, wherein the only modification considered is the change in the environment. Several systems are examined, among them T7 DNA polymerase and GTPase rasp21. The stability of the calculated catalytic enhancement with respect to small changes in EVB and molecular mechanic force fields is demonstrated.

71.

MULTI-CONFIGURATION MOLECULAR MECHANICS. *Donald G. Truhlar, and Titus V. Albu, Department of Chemistry and Supercomputer Institute, University of Minnesota, 207 Pleasant Street Southeast, 139 Smith Hall, Minneapolis, MN 55455-0431, Fax: 612-626-9390, truhlar@umn.edu*

Multi-configuration molecular mechanics (MCM) is a general procedure for fitting potential energy surfaces of reactive systems. It is based on a combination of semiempirical valence bond theory and Shepard interpolation in internal coordinates. This talk will present the method and examples of its use.

72.

FORWARD-BACKWARD SEMICLASSICAL DYNAMICS. *Nancy Makri, Department of Chemistry, University of Illinois, 601 S. Goodwin Avenue, Urbana, IL 61801, Fax: 217-244-3186, nancy@makri.scs.uiuc.edu*

Forward-backward semiclassical dynamics (FBS) provides a practical methodology for including quantum mechanical effects in classical trajectory simulations of polyatomic systems. FBS expressions for time-dependent expectation values or correlation functions take the form of phase space integrals with respect to trajectory initial conditions, weighted by the coherent state transform of a corrected density operator. It is shown that the initial density in finite temperature expressions can be fully quantized by employing the discretized path integral representation of statistical mechanics, thus ensuring a proper treatment of zero point effects and capturing important imaginary components that are absent from purely classical trajectory methods. Optimal sampling is achieved through Monte Carlo or molecular dynamics techniques. Applications to polyatomic clusters and condensed phase processes are presented.

73.

DESOLVATION EFFECTS UPON ION TRANSPORT THROUGH NARROW PROTEIN CHANNELS. *Rob D. Coalson¹, Artem Mamonov¹, Maria Kurnikova², and Abraham Nitzan³. (1) Dept. of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260, Fax: 412-624-8611, coalson@pitt.edu, (2) Dept. of Chemistry, Marquette University, (3) Dept. of Chemistry, Tel Aviv University*

A paradox in ion permeation arises due to the dielectric barrier seen by an ion entering a narrow protein channel. The ion must leave behind the solvation shell that forms around it in bulk water. Inside the channel the ion is surrounded by a low dielectric environment associated with the channel protein and embedding lipid layer. Continuum dielectric calculations based on a rigid protein/channel system show that a dielectric barrier of 10-20 kT arises when a monovalent ion enters a protein channel with radius of 3 Å or less. One would thus predict that no ions could pass through such a channel. In reality, monovalent ions flow easily through narrow ion channels like the K⁺ channel and Gramicidin. We have performed MD simulations on a flexible model of Gramicidin A which reveal that small distortions of the channel structure induced by the presence of an ion in the channel can significantly stabilize the ion, thus compensating for the dielectric barrier. Calculations of ion permeation kinetics based on such relaxed energy barrier profiles give results for the net ion current which are in reasonable agreement with experiment.

74.

MOLECULAR MODELING OF PROTON CONDUCTION IN POLYMER ELECTROLYTE MEMBRANES OF NAFION TYPE. *Lawrence R. Pratt¹, Michael Eikerling¹, Stephen J. Paddison², and Thomas A. Zawodzinski¹. (1) T-12 Theoretical Chemistry, Los Alamos National Laboratory, MS: B268, Los Alamos, NM 85545, lrp@lanl.gov, (2) Computational Nanoscience Group, Motorola Inc*

Molecular mechanisms of proton conduction in polymer electrolyte membranes are addressed by *ab initio* molecular dynamics calculations, utilizing the VASP program, on the model system trifluoromethane sulfonic acid monohydrate solid. After establishing that the experimental crystal is stably simulated, these computations are used to seek crystal defects that might contribute to proton conduction. The most favorable defect found involves formation of a Zundel ion, H₅O₂⁺, and H-bonded, paired sulfonate groups. The energy of this defect is 0.3-0.4eV above the energy of the undefected solid, in suggestive agreement with the observed activation energy for proton transport in minimally hydrated Nafion.

75.

ENSEMBLE-AVERAGED VARIATIONAL TRANSITION STATE THEORY WITH OPTIMIZED MULTIDIMENSIONAL TUNNELING. *Donald G. Truhlar, Jiali Gao, Cristobal Alhambra, Jose Corchado, Mireia Garcia-Viloca, Maria Luz Sánchez, Lakshmi Devi Kesavan, Tina Poulsen, and Jordi Villa, Department of Chemistry and Supercomputer Institute, University of Minnesota, 207 Pleasant Street Southeast, 139 Smith Hall, Minneapolis, MN 55455-0431, Fax: 612-626-9390, truhlar@umn.edu*

Ensemble-averaged variational transition state theory with optimized multidimensional tunneling (EA-VTST/OMT) is a general technique for including quantum mechanical effects in the calculation of rate constants for condensed-phase systems. The first step is the calculation of the classical mechanical potential of mean force along a pre-defined reaction coordinate, for example by umbrella sampling. The second step is adding quantum effects to the free energy of activation of primary-zone vibrations orthogonal to the reaction coordinate. Then one uses variational transition state theory to identify an ensemble of transition states. Several members of this ensemble are sampled in the next stage of calculation. For each ensemble member that is sampled, a minimum energy reaction path of primary-zone atoms is calculated in the static field of the secondary zone, and this reaction path is used to estimate a Keck-Anderson-Bennett-type recrossing coefficient by a VTST calculation with the more optimized reaction coordinate afforded by the reaction-path calculation. Finally tunneling and non-classical reflection are added by optimizing between two multidimensional sets of multidimensional tunneling paths, as given by the small-curvature tunneling approximation and the large-curvature tunneling approximation. Optionally one also adds the free energy change due to equilibrating the secondary zone to each reaction path. This procedure builds on well validated methods developed originally for gas-phase reactions by combining them with an appropriate treatment of condensed-phase systems having multiple reaction paths. This work is supported in part by the NSF and the NIH.

76.

PERTURBED HARD-BODY FLUID THEORETICAL ANALYSIS OF SOLVENT EFFECTS ON CHEMICAL REACTIONS. *Dor Ben-Amotz, Alan Gift, and Brian L. McClain, Department of Chemistry, Purdue University, 1393 Brown Building, West Lafayette, IN 47907-1393, Fax: 765-494-0239, bendor@purdue.edu*

The molecular perturbed hard-body fluid (PHF) model is used to extract repulsive (cavity formation) and attractive (cohesive) contributions to the complete set of solvent excess thermodynamic functions (δG^* , δS^* , δV^* , δH^* , δU^* and δA^*) for chemical reactions in organic liquids. The reactions studied include the formation of a hemiketal from acetone and methanol and conformational equilibria of 1,2-dichloroethane and trans-1,2-dichlorocyclohexane dissolved in diethyl ether and tetrahydrofuran. Pressure and temperature dependent experimental results are compared with PHF predictions. Cavity formation is modeled by treating molecules either as hard spheres or hard anisotropic particles and cohesive interactions are treated using the van der Waals mean field approximation. The results reveal that the excess enthalpy, δH^* , and energy, δU^* , often derive primarily from attractive interactions while repulsive cavity formation more significantly affects the excess entropy, δS^* , and volume, δV^* , in a way that cannot be captured by continuum models for solvation.

77.

AB INITIO MOLECULAR DYNAMICS INVESTIGATION OF PROTON TRANSPORT IN LIQUID METHANOL. *Joseph A. Morrone, and Mark E. Tuckerman, Department of Chemistry, New York University, 100 Washington Square East, New York, NY 10003, Fax: 212-260-7905, jam325@nyu.edu*

Proton transport through aqueous, partially aqueous, or non-aqueous hydrogen-bonded media is a fundamental process in many biologically and technologically interesting systems. Liquid methanol is an example of a hydrogen-bonded liquid that, like water, supports anomalously fast proton transport. Here, we present an *ab initio* molecular dynamics investigation of the microscopic mechanism of proton transport in liquid methanol at 300 K. It is found that the defect structure is a relatively long hydrogen-bonded cationic chain, in which hydrogen bonds in the first and second solvation shells of the excess proton are considerably shorter and stronger than those further away. Along the chain, proton transfer occurs in an essentially random manner. Thus, structural diffusion of the defect chain occurs in a "snake-like" fashion when the excess charge migrates toward

an end of the chain. This causes a weakening and eventual cleavage of the hydrogen bonds at the opposite end. At the protonated end, new hydrogen bonds form due to the strongly associative nature of the excess charge.

78.

QSAR PREDICTION OF DRUG TOXICITY. *John C. Dearden, School of Pharmacy and Chemistry, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, United Kingdom, Fax: +44-151-231-2170, j.c.dearden@livjm.ac.uk*

Until recently, the practice in drug development has been to design in the required activity, but not to test rigorously for toxicity until later in the development process. Nevertheless, it clearly makes sense to eliminate potentially toxic drug candidates as early as possible (fail early, fail cheap!). Toxicity testing itself is, however, not cheap, and increasing interest is focussed on the prediction of toxicity, which additionally has the advantage of speed. This presentation will review what has been achieved in the prediction of drug toxicity using QSAR (quantitative structure-activity relationships), and will relate this to the current use of large, diverse compound libraries. The use of expert systems for toxicity prediction will also be discussed.

79.

ADME PROFILING IN DRUG DISCOVERY: COMPARISON OF RULE-BASED COMPUTATIONAL ALERTS TO MECHANISM-BASED SIMULATION. *Michael B. Bolger¹, Robert Fraczkiewicz², and Boyd Steere². (1) Pharmaceutical Sciences, USC School of Pharmacy and Simulations Plus, Inc, 1985 Zonal Ave. PSC 700, Los Angeles, CA 90089, Fax: 323-442-1390, bolger@usc.edu, (2) Life Sciences Department, Simulations Plus, Inc*

Purpose. To compare a rule-based method (J-Alert) to a mechanistic gastrointestinal simulation for ADME profiling in drug discovery. **Methods.** A new set of rules, called "J-Alert", was developed that contains cutoffs for permeability, solubility, acceptor H-bonding to oxygen, partial charge on the oxygen H-bond donors and acceptors, and a low level cutoff for the Moriguchi log P (I. Moriguchi, 1992). Rank ordering of human intestinal absorption (%HIA) for a library of 138 diverse, well-known drugs by use of the J-alert, was compared to rank ordering by use of a gastrointestinal simulation parameterized with only *in silico* estimates of biopharmaceutical properties. **Results.** Spearman rank correlation coefficients for the rule-based and the mechanistic simulation-based methods were 0.61 ($p < 0.001$) and 0.72 ($p < 0.001$) respectively. **Conclusions.** Application of ultra-high throughput *in silico* estimation of biopharmaceutical properties applied to rule-based and simulation-based methods both produce good predictive ranking when compared to experimental fractions absorbed.

80.

CLASSIFICATION OF CYTOGENETIC TOXICITY OF ORGANIC COMPOUNDS FROM MOLECULAR STRUCTURE. *Nathan R. McElroy, and Peter C. Jurs, Department of Chemistry, Pennsylvania State University, 152 Davey Lab, University Park, PA 16802, nrm126@psu.edu*

Molecular structure features of 297 diverse organic compounds ($42 \leq MW \leq 959$) are used to build predictive models that classify each compound as toxic or non-toxic toward Chinese hamster cells. The literature data used are results of assays that exposed cells to compounds, with and without an enzymatic rat liver homogenate. Each compound is classified as toxic (68) or non-toxic (229), depending on the degree of chromosomal aberration. Topological, geometric, electronic, and polar surface features are calculated to represent each compound. Feature selection is performed using simulated annealing or genetic algorithm routines, combined with classification algorithms including k-nearest neighbor, multiple discriminant analysis, probabilistic neural networks, and support vector machines. Additionally, a compound similarity measure based on atom-pair descriptors was used to remove dissimilar compounds from training, which resulted in improved classification. Training and prediction set classification rates are approximately 85-90%. Methodology and optimal results will be discussed.

81.

QSAR MODELING OF B-LACTAM BINDING TO SERUM PROTEIN. *Lemont B. Kier¹, Lowell H. Hall², and L. Mark Hall². (1) Department of Medicinal Chemistry, Virginia Commonwealth University, Richmond, VA 23298, kier@hsc.vcu.edu, (2) Department of Chemistry, Eastern Nazarene College*

The binding of a drug to proteins in plasma is a powerful determinant on their pharmacodynamic behaviour. The free concentration of the drug, hence the

biological activity is at stake when the drug binds to serum proteins available to it. A goal of these studies is the creation of models that are predictive of the extent of serum protein binding. Of particular interest here is the protein binding of the cephalosporins and the penicillins, chemical classes which are the β -lactam antibiotics. A QSAR model of the binding percent has been created using a combination of topological and E-State indices. Statistical analyses of the model of a database reveals that it is predictive of this important property.

82.

PROPERTY MODELING INDEPENDENT OF 3D INFORMATION: BLOOD-BRAIN BARRIER PENETRATION AND FISH TOXICITY. *Lowell H. Hall, Department of Chemistry, Eastern Nazarene College, 23 East Elm Avenue, Quincy, MA 02170, Fax: 617-745-3905, hall@enc.edu, and Lemont B. Kier, Department of Medicinal Chemistry, Virginia Commonwealth University*

Property modeling is approached through topological representation of molecular structure, independent of 3D structure information. Two properties are modeled: blood-brain barrier partitioning and fish toxicity. Both properties are modeled with a combination of atom type E-State indices and valence molecular connectivity indices. The blood-brain model ($q_2=0.62$) for the set of 106 compounds is validated with an external validation test set and five-fold cross-validation. QSAR models were developed for fish toxicity of a set of 92 compounds including phenols, anilines and substituted aromatic hydrocarbons, yielding excellent statistics ($q_2=0.85$). The model is validated through use of an external validation test set (5 compounds) and ten-fold cross-validation. Detailed structure interpretation is given for the structure indices in each model, providing useful information for further synthesis. Because no 3D structure is required, computations on a virtual library are very fast.

83.

ANALYSIS OF STRUCTURE-ACTIVITY RELATIONSHIPS OF P-GLYCOPROTEIN MEDIATORS FOR DEVELOPMENT OF A NEW P-GLYCOPROTEIN INHIBITOR: PYRONARIDINE. *Chun-zheng Yang, Jing Qi, Shu-bin Wang, and Hui Peng, State Key Laboratory of Experimental Hematology, Peking Medical College, 288 Nanjing Road, 300020 Tianjin, China, czyang@public.tpt.tj.cn*

One of the major reasons for the failure of antineoplastic agents in treatment of solid tumors and hematological malignancies is the presence of drug resistance mechanisms that counter the cytotoxic effects of the antitumor drugs. There were evidences that MDR is induced by a membrane P-glycoprotein (P-gp) which acts as an efflux pump of drugs, resulting in reduction of the intracellular concentration of drugs. Based on analysis of structure-activity relationships, we found that for an effective inhibitors of P-gp, the partition coefficient, logP, of the molecule must be larger than 4 except the other requirements of the structure for the reversal activity of MDR. Based on this finding we selected pyronaridine (PND) as a new Pgp inhibitor. It was found that PND had higher activity for reversal of MDR *in vitro* as well as *in vivo* than verapamil and did not affect the PK of Adriamycin.

84.

STRUCTURAL PROTEOMICS OF EUKARYOTIC PROTEIN FAMILIES. *John F. Hunt, Columbia University, New York, NY 10027, hunt@sid.bio.columbia.edu*

Genome sequencing projects have already determined nearly complete genome sequences of many organisms, including human. The products of these genes are widely recognized as the next generation of therapeutics and targets for the development of pharmaceuticals. While identification of these genes is proceeding quickly, elucidation of their three-dimensional (3D) structures and biochemical functions lags far behind. In some cases, knowledge of 3D structures of proteins can provide important insights into evolutionary relationships that are not easily recognized by sequence alignment comparisons. Thus, structure determination by X-ray crystallography or NMR can sometimes provide key information regarding protein fold class, locations and clustering of conserved residues, and surface electrostatic field distributions that connect a protein sequence with potential biochemical functions. In collaboration with other members of the Northeast Structural Genomics Consortium (www.nesg.org), we are developing technologies that accelerate the process of protein structure determination by X-ray crystallography and NMR. The goal of this pilot project is to develop a "high-throughput" process for structural analysis of proteins on a genomic scale and to apply this in the analysis of protein families containing

one or more members coded by the human and/or eukaryotic model organism genomes.

85.

FUNCTION BASED PROFILING OF PROTEASES. *Jennifer L. Harris, Nicolas Winsinger, Robert Damoiseaux, and Bradley J. Backes, Chemistry, Genomics Institute of the Novartis Research Foundation, 10675 John Jay Hopkins Drive, San Diego, CA 92121, Fax: 858-812-1584, harris@gnf.org*

Complete genome sequences have propelled interest in profiling technologies whereby the impact of a cellular perturbation are investigated on a genome-wide scale. The regulation of protein function through post-translational modification, local environment and protein-protein interaction is critical to cellular function and necessitates the development of technologies aimed at measuring a protein's functional activity rather than a protein's presence. We recently reported a new technology for the preparation of spatially addressable small molecule microarrays from split pool combinatorial libraries. Here we will report the application of such small molecule microarray to an activity-based profile of proteases in complex biological samples.

86.

FROM PROTEINS TO DRUGS - IDENTIFICATION OF REGIONS OF PROTEINS INVOLVED IN INTRAMOLECULAR INTERACTIONS. *Matthew Clark, Scientific Computation, Locus Discovery, Inc, 512 Township Line Road, Blue Bell, PA 19422, Fax: 215 358 2020, mclark@locusdiscovery.com, and Stephan Brunner, Locus Discovery Inc*

The identification of the regions of proteins important for interactions - protein-protein and drug-protein - is a key step in the overall process of going from genomic information to drug discovery. We present and assess novel methods that provide a highly predictive descriptor for binding sites. The results can be viewed graphically to identify regions hospitable for specific binding modes, as well as rigorously analyzed to locate the binding sites and specific binding modes of the residues involved in the binding site.

87.

AUTOMATIC IDENTIFICATION OF PROTEIN FUNCTIONAL SITES USING BIODOCK. *Sabine K. Schlyer, Diana C. Roe, and Malin M. Young, Biosystems Research Department, Sandia National Laboratories, MS 9951, P. O. Box 969, 7011 East Ave., Livermore, CA 94551, Fax: (925) 294-3020, skschly@sandia.gov, mmyoung@sandia.gov*

Automatic identification of protein functional sites using BioDOCK.

Protein active sites and interaction interfaces are often characterized by a conserved spatial arrangement of amino acids that is fold-family specific. The identification of these conserved sites permits focused docking studies that selectively target families of, rather than individual, proteins. To this end, we have developed a program, BioDOCK, that integrates genomic and structural information to automatically select sites that are a) functionally important and b) suitable for docking calculations. We have tested BioDOCK on a diverse fold-family library of 800 protein targets. In our talk, we will summarize the results of this calculation and discuss its implications for drug design.

88.

BINDING MODE OF VEGF RECEPTOR TYROSINE KINASE KDR (VEGFR-2) INHIBITORS. *Georgia McGaughey, Mark Bilodeau, Kathleen Coll, Jay Gibbs, George Hartman, William Huckle, Randall Hungate, Anne Johnson, Richard Kendall, Timothy Koester, Nancy Kohl, Xianzhi Mao, Rosemary McFall, Bo-Sheng Pan, Keith Rickert, Leonard Rodman, Ruth Rutledge, and Kenneth Thomas, Merck Research Laboratories, West Point, PA 19486, georgia_mcgaughey@merck.com*

VEGF induces endothelial cell mitogenic signaling through the receptor tyrosine kinase KDR. The inhibition of this process has been a leading target in the search for anti-angiogenic therapeutics. Although a crystal structure of activated KDR is available, we created a homology model of the unactivated form to reflect our assay conditions. Using a combination of docking studies and molecular mechanical and quantum mechanical calculations, a binding mode of thiazolopyridyl KDR inhibitors is proposed. This binding mode takes advantage of hydrophobic interactions and makes key hydrogen bonds to the hinge region. Binding data of thiazolopyridyl KDR inhibitors in various mutants supports the proposed docking orientations. Finally, sequence alignments of related kinases

reveal amino acid differences within the active site, which could be a means to achieve selectivity.

89.

ACCURATE AND EFFICIENT MANY-BODY POTENTIAL FOR HYDROGEN BONDING - APPLICATION TO WATER. *P. Ballone, Department of Physics, University of Messina, Contrada Papardo, 98166 Messina, Italy, p.ballone@fz-juelich.de, J. Akola, I F F, Forschungszentrum Juelich, Germany, and R. O. Jones, I F F, Forschungszentrum Juelich, 52425 Juelich, Germany, Fax: 01149-2461-612850, r.jones@fz-juelich.de*

An accurate, computationally efficient model for hydrogen bonding is crucial for simulations of water, many organic materials, and most biological systems. While traditional models rely on fitting electrostatic and Lennard-Jones parameters to reproduce the strength, chemical selectivity and harmonic vibrational modes, current applications often demand a more detailed description. We have extended a many-body potential (J. Tersoff) developed to model solids with directional bonds whose strength depend on the local environment. The model discriminates between tetrahedral and trigonal bonding configurations and provides a flexible representation of many-body contributions, whose computational cost scales linearly with the number of particles. This last feature makes it more efficient than many-body potentials that include molecular polarizability. Applications to water clusters, liquid water, and ice gives results comparable to those of the best many-body potentials at a cost comparable to that of traditional models.

90.

DEVELOPING TRANSFERABLE FORCE FIELDS FOR PHASE EQUILIBRIUM CALCULATIONS. *J. Ilja Siepmann¹, Collin D. Wick², John M. Stubbs², Nikolaj Zhuravlev², Ling Zhang², Jun-Seok Lee², Xin Zhao², Marcus G. Martin³, Bin Chen⁴, and Jeffrey J. Potoff⁵. (1) Departments of Chemistry, Chemical Engineering and Materials Science, University of Minnesota, 207 Pleasant St. SE, Minneapolis, MN 55455, Fax: 612-626-7541, siepmann@chem.umn.edu, (2) Department of Chemistry, University of Minnesota, (3) Computational Materials and Molecular Biology, Sandia National Laboratories, (4) CMM & Department of Chemistry, University of Pennsylvania, (5) Department of Chemical Engineering and Materials Science, Wayne State University*

Recent developments of the TraPPE (transferable potentials for phase equilibria) force field will be described. After a brief introduction of the fitting philosophy used for the TraPPE force field, this presentation will focus on the prediction of the thermophysical properties for industrially relevant multicomponent systems.

91.

HARDWARE ACCELERATOR FOR MOLECULAR FORCE FIELDS. *Forrest H. Bennett III, and William Mydlowec, Pharmix Corp, 200 Twin Dolphin Drive, Suite F, Redwood Shores, CA 94065, forrest@pharmix.com*

This talk will present a hardware accelerator that allows us to perform OPLS-AA molecular mechanics computations at 324 times the speed of a 1GHz Pentium III, which represents an 80-fold improvement in price/performance, and a 648-fold improvement in Watts/performance (and hence operating costs). We are able to model systems of up to 223,000 atoms with up to 256 atom types. This is more than sufficient to model large protein-ligand systems with explicit solvent, or even a large GPCR with explicit membrane and water. The single-point energy of a 64,000 atom system is computed by our accelerator in 4.5 seconds. This accelerator is implemented on a Xilinx Virtex-II 6000 field programmable gate array chip (FPGA) running at 160 MHz. This FPGA has 76K logic cells and 2.6 Mbits of on-chip memory. The FPGA is mounted on a standard PCI bus expansion card with a 300 Mbits/s connection to 8 MBytes of memory on the board. Up to 4 of these accelerators can be installed in a PC running Linux or Windows. This is the first system we know of that implements a complete molecular mechanics computation on a single chip.

92.

FORCE FIELD DEPENDENCE OF NMR-BASED, RESTRAINED MOLECULAR DYNAMICS DNA STRUCTURE CALCULATIONS INCLUDING AN ANALYSIS OF THE INFLUENCE OF RESIDUAL DIPOLAR COUPLING RESTRAINTS. *Michael A Kennedy, and Kathleen McAteer, Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, PO Box 999, K8-98, Richland, WA 99352, Fax: 509-376-2303, mkeneddy@pnl.gov*

DNA structure determination by NMR has historically suffered from insufficient distance and dihedral angle restraints to generate highly converged families of

structures using a distance geometry approach. Consequently, restrained molecular dynamics is widely used for calculations since the force field helps drive the structure into a reasonable conformation. Comparison of force fields developed for CHARMM and AMBER in an in-silico experiment show that the force field can be significant compared to NMR restraints in driving final structures to converge. The influence of the force field leads to artificially tight convergence within families of structures and the precision and character of resulting structures depend on the force field used. Incorporation of residual dipolar coupling restraints improves convergence with the target structure and between families of structures, indicating that the force field dependence can be overcome if RDC restraints are employed. A comparative analysis of the force fields will be presented.

93.

A CRITIQUE OF GBSA. *Anthony Nicholls, OpenEye Scientific Software, Inc, 3600, Suite# 1107, Cerrillos Rd, Santa Fe, NM 87507, and Andrew Grant, EST, AstraZeneca, Ltd*

The Generalized Born model of solvation (GBSA) is now widely used as a surrogate for the more computationally expensive solution of the Poisson Equation (PE). Among the attractive attributes of GBSA are an accounting for dielectric screening and desolvation, frame-invariance and access to analytic derivatives. It is known that parameterization is necessary for good agreement with the PE. What is not commonly appreciated is that such parameterization lacks generally applicability, that the pair-wise term lacks any fundamental validity, that errors of greater than 40% in total energy are easy to find, that, when applied to proteins, the long-range form of the model is incorrect and the computational performance lags that of PE solvers, and that calculating the exact GB radius does not actually enhance accuracy. An examination of related models illustrate the fundamental limitations inherent in the loss of directional information in GBSA.

94.

PROTEIN MODELING IN EDUCATION. *George D. Purvis III, CAChe Group, Fujitsu, 15244 NW Greenbrier Pkwy, Beaverton, OR 97006, Fax: 503-531-9966, gpurvis@CACheSoftware.com*

Protein chemistry is becoming increasingly mainstream in chemical education. High-speed desktop and laptop computers and innovative new software are making protein modeling accessible to every chemist. A protein docking experiment designed for teaching students about protein-ligand interactions using Windows(tm) computers is described. In the experiment, students first view the x-ray crystal structure of the tuberculosis PNP protein-inhibitor complex, interpret the binding, then dock the inhibitor in a human protein and suggest ways that the inhibitor could be made more selective.

95.

LUCID: WEB-BASED TEAM LEARNING AND REAL-TIME MULTILEVEL ASSESSMENT FOR INTRODUCTORY CHEMISTRY. *Troy A. Wolfskill, Center for Excellence in Learning and Teaching, University at Stony Brook, Stony Brook, NY 11794-3400, Fax: 631-632-7960, Troy.Wolfskill@sunysb.edu, and David Hanson, Department of Chemistry, University at Stony Brook*

The LUCID Project is developing materials and functional enhancements for web-based learning systems to assist students and teachers in improving student learning outcomes. Team learning activities develop conceptual understanding and problem solving strategies, and include a peer review process for assessing student work. The assessment system provides a rich set of conceptual questions, exercises, and problems, many of which call for open-ended responses that may include such things as chemical reaction equations and molecular structures. These responses are analyzed to identify learning objectives that students have and have not achieved. By tracking student learning in real time by both topic and level of mastery, feedback can be provided to help both students and teachers improve learning.

96.

IMPACT OF ADDING INTERACTIVE PROBLEM-ANALYSIS HELP TO WEB-BASED HOMEWORK ASSIGNMENTS. *Katherine I. Barnhard, and John W. Moore, Department of Chemistry, University of Wisconsin-Madison, 1101 University Ave., Madison, WI 53706, Fax: 608-265-8094, barnhard@chem.wisc.edu*

Most students feel stymied by questions that require them to use problem-solving skills and many just throw up their hands without a clue as to how or

where to start. We have developed an interactive problem-analysis aid for on-line homework assignments. We are using it within the WebCT course management system, but it could just as easily be integrated into any system of web-based problems written in HTML. While working on a problem, the student can press a help button which leads to a series of questions to which the student must respond. Rather than presenting correct answers or entire explanations all at once, this aid attempts to encourage students to employ the thought processes required for higher order problem analysis and problem solving skills. It presents questions that guide the student toward working out the problem. The student response to each question elicits a summary of "what we've learned" and a new question leading to the next step. If there are several possible paths to reach a "correct answer", students are encouraged to choose whichever suits their style best. Student response to these on-line aids will also be discussed.

97.

THE USE OF TECHNOLOGY TO FACILITATE INTERACTIVE LEARNING OF CHEMISTRY AT THE UNITED STATES MILITARY ACADEMY. *Augustus W. Fountain III, Department of Chemistry, United States Military Academy, MADN-CHEM, West Point, NY 10996, Fax: 845-938-3062, augustus-fountain@usma.edu*

Since the late-1980's every cadet has been issued a computer upon entering West Point. Every classroom and cadet room in the barracks are "wired" allowing faculty and cadets access to the Academy LAN and world wide web resources to aid in their learning. Each course in the Chemistry Department sponsors a web site that provides cadets information on their individual course as well as links to other sites that can assist the cadets. Cadets also do web-based homework using Web CT. This system gives cadets immediate feedback to their homework problems. These assets give the instructor an ability to use the many facets of technology in the traditional classroom. Over the years we have used many types of technology to enhance learning. These include interactive CD-ROMS, World Wide Web Resources, the Classroom Performance System (an Infrared response system that supports real-time interaction in the traditional classroom), LabWorks (a computer-based laboratory interface, MathCAD developed problem sets, and various sophisticated molecular modeling environments that unite 3D visualization and animation with quantum chemical calculations, molecular mechanics, and dynamics.

98.

EXAMPLES OF COMPUTER USE IN UNDERGRADUATE PHYSICAL CHEMISTRY. *Danny G. Miles Jr., Department of Science, Mount St. Mary's College, 16300 Old Emmitsburg Road, Emmitsburg, MD 21727, Fax: 301-447-5021, miles@msmary.edu*

Responses to a web-based survey on computer use in undergraduate physical chemistry indicated a need for better implementation of computer-related activities throughout the course. Examples of activities that can be integrated into physical chemistry lecture and laboratory components will be presented. The activities take advantage of computer and computational methods in a variety of ways, including data analysis, simulations, molecular modeling, and real-time data acquisition. Excel, Mathcad, Spartan, and Logger Pro software are used in the examples.

99.

STUDENT MANIPULATION OF MICROSCOPIC REACTION TRAJECTORIES. *Alexander Grushow, Department of Chemistry and Biochemistry, Rider University, Lawrenceville, NJ 08648, Fax: 609-895-5782, grushow@rider.edu*

Given the complex nature of a reaction potential energy surface (PES), it is difficult to fully describe in a lecture setting and have students develop an appreciation for the significance of PES topology and its affect upon reaction cross section. However, this can be overcome by allowing students to discover the effect of manipulation of microscopic variables on an individual reaction trajectory. A computer program, PTRJ, has been written which allows students to run a single reaction trajectory and observe the motions of the atoms as the reaction proceeds. While undergraduates do not understand how a trajectory calculation is preformed, they are able to interpret the motions of the atoms on the PES in a way that introduces them to the concept of a reaction on a microscopic scale. By directing the students' exploration of the different factors which influence the results of a single reaction trajectory, they can begin to

explore the relationship between reactions on the microscopic and macroscopic scales.

100.

PREDICTIVE ADME: WHY MODELS FAIL. *Terry R Stouch, and Stephen Johnson, Computer-Assisted Drug Design, Bristol-Myers Squibb, P.O. Box 4000, Princeton, NJ 08543-4000, Fax: 609-252-6030, Stouch@bms.com*

There is substantial hope that predictive QSAR models of ADME liabilities can help screen prospective drug compounds early in the Drug Discovery cycle. Consequently, the literature is full of models for almost every liability. However, even carefully considered models can fail. We will present case studies detailing issues that derail otherwise reasonable QSAR models and make them useless for routine screening.

101.

ON THE DETECTION OF MULTIPLE BINDING MODES OF LIGANDS TO PROTEINS, FROM BIOLOGICAL, STRUCTURAL, AND MODELLING DATA. *P. Lewi¹, M. de Jonge¹, F. Daeyaert¹, L. Koymans¹, M. Vinkers¹, J. Heeres¹, P.A.J. Janssen¹, E. Arnold², K. Das², A.D. Clark Jr.², S.H. Hughes³, P.L. Boyer³, M.-P. de Béthune⁴, R. Pauwels⁴, and K. Andries⁵.*

(1) Center for Molecular Design, Janssen Pharmaceutica N.V., B-2350 Vosselaar, Belgium, plewi@prdbe.jnj.com, (2) Chemistry and Chemical Biology Department, Center for Advanced Biotechnology and Medicine, Rutgers University, (3) NIH-NCI HIV Drug Resistance Program, (4) Tibotec-Virco, (5) Virology Dept., J&JPRD, Janssen Pharmaceutica N.V

There is evidence that highly potent ligands may bind in different conformational modes to the same binding site of their target protein. Over the past decade we have gathered biological activity data on several hundred non-nucleoside inhibitors of HIV-1 reverse transcriptase (NNRTIs), as obtained from cellular tests of the wild type and a broad panel of mutant viruses. NNRTIs with efficacy against a broad spectrum of clinically relevant mutants were selected for structure analysis by X-ray crystallography. The RT/NNRTI complex structures were used as templates for molecular modelling studies aimed at guiding further synthesis and understanding the inhibition profiles. Binding energies were computed for these inhibitors and for their interactions with the amino acids of the binding site which contribute significantly to their observed biological activity. Comparison of the structures of RT complexed with related NNRTIs revealed that some of the inhibitors had the potential to bind to their target site in multiple conformations. This observation may explain the high potency of some of the NNRTIs against a broad spectrum of HIV mutants resistant to other NNRTIs. It is possible that a single NNRTI could inhibit different HIV-1 RT mutants by binding in alternate conformations. Using a chemometric approach, based on multivariate data analysis, we attempt to detect whether the concept of multiple binding modes of NNRTIs is supported by biological and modelling data.

102.

THEORY AND PRACTICE OF SAFE QSAR. *Alexander Tropsha¹, Yun-De Xiao¹, Min Shen¹, Scott Oloff², and Alexander Golbraikh¹.* *(1) Laboratory for Molecular Modeling, School of Pharmacy, University of North Carolina, CB # 7360, Beard Hall, School of Pharmacy, Chapel Hill, NC 27599-7360, Fax: 919-966-0204, alex_tropsha@unc.edu, (2) Department of Pharmacology, University of North Carolina*

Using k nearest neighbors (kNN) variable selection QSAR method for the analysis of several data sets, we show that many traditional means of QSAR model quality assessment such as leave-one-out (LOO) cross-validated R² (q²) are inadequate criteria to evaluate the predictive ability of the models. We show that there exists no correlation between the values of q² for the training set and the accuracy of prediction (R²) for the test set. Based on extensive validation with both real and simulated data sets, we argue that this observation is a general property of any QSAR model developed with LOO cross-validation. We formulate a set of quantitative criteria for evaluation of predictive ability of QSAR models. We demonstrate that using these criteria for model development results in highly predictive models for a number of applications, including human S9 metabolism turnover rates and inhibition of Cytochrome P3A4.

103.

EFFECT OF TRAINING AND TEST SET DIVERSITY ON PLS STATISTICS. *Robert D. Clark, Tripos, Inc, 1699 S. Hanley Road, St. Louis, MO 63144, bclark@tripos.com*

It is becoming increasingly common in quantitative structure/activity relationship (QSAR) analyses to use external test sets to evaluate the stability and predictivity of the models obtained. In some cases, such as those involving variable selection, an internal test set - i.e., a crossvalidation set - is also used. Care is sometimes taken to ensure that the subsets used exhibit similar response or property distributions, but more often the individual observations are simply assigned "at random." In the special case of MLR without variable selection, it can be analytically demonstrated that this strategy is inferior to others. Most particularly, D-optimal design performs better if the form of the regression equation is known and the variables involved are well behaved. Here we use optimizable k-dissimilarity partitioning to explore the countervailing effects of training and test set size, diversity, and representativeness on PLS model statistics, where no such analytical justification exists.

104.

ARE TOPOLOGICAL PARAMETERS "FUNDAMENTAL PROPERTIES"? EVIDENCE FROM THE SOLUBILITY OF GASES IN LIQUIDS. *Marvin Charton, Chemistry Department, Pratt Institute, DeKalb Ave. and Hall St., Brooklyn, NY 11205, Fax: 718-722-7706mcharton@, mcharton@pratt.edu*

Topological parameters are frequently used in QSAR studies. It has been claimed that they are fundamental molecular properties. If this is indeed the case then they must account for the properties of every type of molecule. We present evidence that the solubility of the rare gases and of diatomic molecules which have constant topology can be modeled by polarizability and in the latter dipole moment as well. As this cannot be done with topological parameters they are not fundamental quantities. They are in fact composite parameters representing a mix of polarizability and steric effects, as we have reported elsewhere.

105.

GEMINALS-BASED MODEL CHEMISTRY. *Vitaly A Rassolov, Department of Chemistry and Biochemistry, University of South Carolina, 631 Sumter St, Columbia, SC 29208, Fax: 803-777-9521, rassolov@mail.chem.sc.edu*

We present a development of a method based on a modification of the Antisymmetrized Product of Strongly orthogonal Geminals (APSG) theory. Our method is size-consistent, well defined, and applicable to large chemical systems. The method is implemented in the atomic orbital basis, with scaling with respect to the system size that is favorable to any correlated method. We demonstrate the formal properties of our method and apply it to studies of geometries and energetics of all diatomic molecules from the G2/97 test set (28 molecules, in both singlet and open shell states).

106.

A LINEAR-SCALING QUANTUM CHEMISTRY METHOD FOR GROUND AND EXCITED STATES: LOCALIZED-DENSITY-MATRIX METHOD. *GuanHua Chen, The University of Hong Kong, Hong Kong, China, Fax: 852-28571586, ghc@everest.hku.hk*

The localized-density-matrix (LDM) method has been developed to calculate the electronic ground and, in particular, excited state properties of very large molecular systems (containing hundreds or more atoms). It is a linear-scaling quantum chemistry method, and has so far been implemented with semiempirical methods and density functional theory (DFT). The key for achieving the linear-scaling of computational costs is the locality of reduced single-electron density matrix. Instead of many-body wave function, the reduced single-electron density matrix is calculated. It is solved in the time domain instead of the usual energy domain. The LDM method has been used to calculate the absorption spectra, circular dichroism spectra and electronic structures of large molecular systems. The largest system whose absorption spectrum has been calculated so far is a polyacetylene oligomer containing 32,000 carbon atoms (PPP Hamiltonian was employed in the calculation). Applications to conjugated polymers, PPV-aggregates, light harvesting system and carbon nanotubes will be given.

107.

MONTE CARLO SIMULATION OF PROTEINS THROUGH A RANDOM WALK IN ENERGY SPACE. *Nitin Rathore*, and *Juan J. de Pablo*, Department of Chemical Engineering, University of Wisconsin-Madison, 1415 Engineering Drive, 1037 Engineering Hall, Madison, WI 53706, Fax: 608-262-5434, rathore@cae.wisc.edu

A new Monte Carlo algorithm that performs a random walk in energy space has been used to study random coil-helix and random coil-beta sheet transitions in model proteins. This method permits estimations of the density of states of a protein via a random walk on the energy surface, thereby allowing the system to escape from local free-energy minima with relative ease. A cubic lattice model and a knowledge based force field are employed for these simulations. It is shown that, for a given amino acid sequence, this new method is able to fold long polypeptides reproducibly. Its results compare favorably with those of annealing and parallel tempering simulations, which have been used before in the same context. This method is used to examine the effect of amino acid sequence and chain length on the folding of several designer polypeptides.

108.

DIRECT CALCULATION OF MOLECULAR FREE ENERGIES. *Chia-en Chang*, Department of Chemistry, University of Maryland, College Park, MD 20742, cchang@wam.umd.edu, and *Michael K Gilson*, University of Maryland Biotechnology Institute

A fast, accurate method has been developed for directly computing the conformational free energy of a molecule in all degrees of freedom. The method, based on internal coordinates, includes an efficient conformational search method and direct computation of the molecular configuration integral in each energy well found during the search. The method uses internal vibrational modes for both the conformational search and integration steps. To speed the calculations, soft degrees of freedom — freely rotatable bonds — are separated from hard degrees of freedom and sampled with Mining Minima, while “hard” degrees of freedom — bond-lengths, bond-angles and bond-rotations within rings — are treated by a modified harmonic approximation which accounts for local anharmonicities. Tests show that the method is faster and more accurate than a previously described method, MINTA. The search method, which is effective for cyclic, macrocyclic, acyclic and mixed molecules, can be used independently of the full-fledged free energy calculations.

109.

HIGH THROUGHPUT PREDICTION OF PASSIVE ADME PROPERTIES FROM FRAGMENTS. *Tudor I. Oprea*¹, *Massimo Baroni*², *Ismael Zamora*³, and *Gabriele Cruciani*². (1) EST Chemical Computing, AstraZeneca, S-43183 Molndal, Sweden, Fax: 46 31 776 3792, tudor.oprea@astrazeneca.com, (2) 4 Chandos Street, Molecular Discovery Ltd, (3) Lead Molecular Design, s.l

Based on the GRID, VolSurf, ChemGPS and GPSVS technologies, PENGUINS (Pharmacokinetics Evaluation and GRID Utilization IN Silico) is designed to manipulate large numbers of compounds starting from 2D structure. In PENGUINS, one can use SMILES or SDF files, in the absence of 3D structures, to predict VolSurf parameters for the water and DRY probes, as well as GPSVS scores. A precomputed database of fragments can be used to recognize input molecules. PENGUINS has the ability to provide fast and accurate predictions for virtual and/or existing chemical libraries, with respect to passive ADME properties such as oral (intestinal) drug absorption and blood-brain barrier drug permeation, as well as water solubility for more than one million compounds per CPU/day.

110.

PRIVILEGED SUBSTRUCTURE SEARCHING FOR FOCUSED SET DESIGN. *Christophe Cleva*, *Daniel Domine*, *Cedric Merlot*, *Jean Bunn*, *Eric Sebille*, *Wolfgang Sauer*, and *Dennis Church*, Scientific Computing Group, Serono Pharmaceutical Research Institute, 14 chemin des Aulx, 1228, Plan-les-Ouates, Geneva, Switzerland, Fax: +41 22-794-6965, christophe.cleva@serono.com

A novel computational method for the rapid identification of privileged substructures associated with one or more biological activities has been developed, implemented and validated on various targets. The technique extracts structural information from data sets which can be used to query compound libraries,

direct sequential HTS and/or combinatorial chemistry campaigns, and predict various pharmacological and toxicological outcomes. Real-life case studies illustrating the use of the methodology in combination with other tools such as diversity analysis and virtual compound libraries in the identification of novel 7-TM receptor ligands, phosphatase inhibitors, ion channel blockers and kinase inhibitors are presented. The presentation will finally touch on the software implementation of the methodology, as well as its integration within the pharminformatics environment.

111.

THEORY AND EXPERIMENT: MOLECULAR MODELING EXERCISES DESIGNED TO INCREASE CONCEPTUAL UNDERSTANDING. *James B. Foresman*, Department of Physical Science, York College of PA, Country Club Road, York, PA 17405, JForesma@YCP.EDU

The partnership between theory and experiment is an important thread to maintain in teaching chemistry. With the increased accuracy and efficiency of today's computational chemistry codes, it is possible to design a useful modeling exercise to accompany almost any traditional topic in the curriculum. This paper describes new examples from a wide variety of contexts (freshman chemistry to graduate level coursework). The examples reinforce concepts in thermochemistry, spectroscopy, solvent effects, intermolecular forces, and reaction mechanisms. Finally, an evaluation of different computational engines is made in the context of student use. Included in the comparison are the many graphical visualization programs that are widely available.

112.

USING WEB-BASED HOMEWORK AND QUIZZES TO REDESIGN A GENERAL CHEMISTRY COURSE. *John W Moore*¹, *Katherine I. Barnhard*¹, *Renee S. Cole*², and *John B. Todd*³. (1) Department of Chemistry, University of Wisconsin-Madison, 1101 University Avenue, Madison, WI 53706, Fax: 608-265-8094, jwmoore@chem.wisc.edu, (2) Department of Chemistry and Physics, Central Missouri State University, (3) Department of Chemistry, Bowling Green State University

With support from the Pew Course Redevelopment project we have created a new learning environment for general chemistry that provides students with a better means of learning chemistry and automates many of the time-consuming duties of teaching assistants. The system includes online homework, online quizzes, an online gradebook, and online tutorials delivered via the WebCT course management system. We have developed 417 homework sets containing 3556 questions, 279 quiz question sets containing 1374 questions, and 37 instructional modules for Web-based delivery. We have incorporated these materials into two different large general chemistry courses, which included developing and implementing procedures for weekly Web-based quizzing of students. We have evaluated online homework through a controlled experiment and found learning equivalent to students who were conventionally taught. We have also taught part of a course using only Web-based tutorials and found that students were able to learn effectively in the absence of lectures and textbook material.

113.

BEYOND QUANTUM CHEMISTRY: A COURSE IN MATHEMATICAL MODELING FOR CHEMISTS. *Michelle M Francl*, Department of Chemistry, Bryn Mawr College, 101 N Merion Ave, Bryn Mawr, PA 19010, Fax: 610 526 5086, mfrancl@brynmaur.edu

Molecular modeling is becoming firmly embedded in the undergraduate chemistry curriculum. Students encounter quantum mechanical modeling as early as their first college chemistry course. Additional exposure to molecular modeling often comes in physical chemistry and biochemistry courses. Students are less frequently challenged to create and work with their own numerical models of chemical phenomena based on the mathematical models developed in physical chemistry. In this course, students designed computer models using Mathematica which included random walk models for polymers, Monte Carlo models for gas deposition and diffusion on surfaces, and kinetic models for oscillating reactions. Students were assessed on the ability of their models to reproduce known values.

114.

ESSENTIAL ROLE OF INFORMATICS IN COMPUTATIONAL BIOCHEMISTRY/CHEMISTRY CURRICULUM. *Masayuki Shibata, and Syed S. Haque, Biomedical Informatics, UMDNJ-SHRP, 65 Bergen Street, Newark, NJ 07107-3001, Fax: 973-972-1054, shibatma@umdnj.edu*

It has been suggested that the 21st century is the Informatics century. Recent advances in technologies have triggered the current explosions of data in many areas of sciences. A large amount of data is generated everyday by high-throughput screenings, various genomics, functional genomics or microarray, and proteomics projects. Conversions of large amount of data into information and knowledge are performed by various informatics tools. However, only a small number of scientists are trained in computational/informatics field on one hand, and in basic scientific field such as chemistry, biochemistry, biotechnology, genetics, molecular biology, and cell biology on the other. This shortage creates an increasing need to provide sound informatics trainings to these scientists. Biomedical Informatics program at UMDNJ-SHRP was developed to satisfy such demands. A detailed description of the Biomedical Informatics curriculum, and its applicability to biochemistry/chemistry curriculum will be discussed. The nature of our curriculum, the characteristics of our student body and the outcome assessments will also be presented.

115.

A RIGOROUS SENIOR-LEVEL MODELING COURSE: HOW MUCH IS TOO MUCH FOR THREE CREDITS? TELL US ABOUT IMPLEMENTATION AND USE OF COMPUTER-SUPPORTED INSTRUCTION AND LEARNING? *Theresa Julia Zielinski, Department of Chemistry, Medical Technology, and Physics, Monmouth University, Edison Science Hall Room 245, West Long Branch, NJ 07764, Fax: 732-263-5213, tzielins@monmouth.edu*

Molecular modeling software is routinely used for drug design and the analysis of chemical reaction mechanisms. A course in molecular modeling and computational chemistry is therefore an important component of the undergraduate BS chemistry curriculum. During this presentation a rigorous course using available software, Spartan, and texts by Wavefunction Inc. will be described. The course consists of a significant number of computational chemistry exercises designed to span a two-hour lecture and three hour laboratory sequence. Substantial homework assignments are included in the syllabus.

116.

WHAT DOES THE EVALUATION OF PHYSICAL CHEMISTRY ON-LINE PROJECTS TELL US ABOUT IMPLEMENTATION AND USE OF COMPUTER-SUPPORTED INSTRUCTION AND LEARNING? *Marcy Towns, Department of Chemistry, Ball State University, Cooper Hall, Muncie, IN 47306, Fax: 765-285-2351, 00mhtowns@bsu.edu, Laura Slocum, University High School of Indiana, Renee S. Cole, Department of Chemistry and Physics, Central Missouri State University, and Theresa Julia Zielinski, Department of Chemistry, Medical Technology, and Physics, Monmouth University*

The Physical Chemistry On-Line (PCOL) consortium makes use of computer supported collaborative learning through on-line modules that engage students at geographically dispersed institutions in activities to facilitate their learning of physical chemistry. The modules use a guided-inquiry approach and rely on collaboration between students in their own class, as well as computer-supported collaboration across institutions via discussion boards, listservs, or WebCT. Our evaluation has used pre- and post-surveys as well as the discussion archives to focus on student computer use, the student perspective of the module, and student interactions on the computer conferencing system used in the projects. This presentation will focus on the development and use of the modules, and what the outcomes of our research tells us about the type of computer conferencing systems used, what type of course content are most suitable to these projects, and role of the faculty as facilitators.

117.

PROPERTIES OF WATER IN MICELLAR SOLUTIONS: RESULTS FROM COMPUTER SIMULATIONS. *Max L. Berkowitz¹, Sanjib Senapati¹, Chrystal D. Bruce¹, and Lalith Perera². (1) Department of Chemistry, University of North Carolina, Chapel Hill, NC 27599, Fax: 919-962-2388, maxb@unc.edu, (2) Academic Technologies and Networking, University of North Carolina*

We present the results on structural and dynamical behavior of water obtained from two computer simulations. In the first simulation we consider a sodium dodecyl sulfate micelle solvated in water. In a second simulation we study a

system containing a reverse micelle based on a phosphate surfactant in a water/carbon dioxide mixture.

118.

SIMULATIONS OF SELECTIVITY AND PERMEATION IN CHANNELS. *Douglas Henderson, Department of Chemistry, Brigham Young University, Provo, UT 84602, Fax: 801-422-0153, doug@huey.byu.edu, Dezso Boda, Department of Physical Chemistry, University of Veszprem, and David Busath, Department of Zoology, Brigham Young University*

Our Monte Carlo simulations of the selectivity of calcium and sodium channels are discussed. Our model is that of Nonner and Eisenberg, where the side chains of the filter are modeled as tethered charged spheres. We find that this model calcium channel is selective for calcium ions and, under physiological conditions, the sodium channel is not selective for calcium ions. The mechanism for selectivity is electrostatic. Also, our molecular dynamics simulations of permeation in some simple model channels are discussed.

119.

ENERGY TRANSFER AND CHEMICAL REACTION DYNAMICS AT INTERFACES.

William L. Hase, Institute for Scientific Computing, Wayne State University, 431 State Hall / Computer Science Department, 5143 Cass Avenue, Detroit, MI 48202, wlh@chem.wayne.edu

Many-atom physical and chemical effects at interfaces may give rise to reaction dynamics different than that observed in the gas phase. These effects include confinement, solvation to adjust reaction energetics, the presence of thermal baths, the co-operativity of many-atom motions, and a broad-range of dynamical time-scales. In this talk, relationships between the dynamics of apparent disparate systems will be discussed. They include: energy transfer in rare gas collisions with hydrogenated surfaces; peptide ion surface-induced dissociation (SID); and energy transfer at the interface of sliding surfaces.

120.

SOLVATION AND IONIZATION STAGES OF HCl ADSORBATE ON ICE NANOCRYSTALS. *V. Buch¹, J.P. Devlin², J. Sadlej³, and N. Uras-Aytemiz². (1) Fritz Haber Institute for Molecular Dynamics, Hebrew University, 91904 Jerusalem, Israel, Fax: 972-2-6513742, viki@fh.huji.ac.il, (2) Department of Chemistry, Oklahoma State University, (3) Department of Chemistry, Warsaw University*

Ionization of hydrochloric acid, $\text{HCl} + \text{H}_2\text{O} \rightarrow \text{H}_3\text{O}^+ + \text{Cl}^-$, is a very basic reaction in chemistry, driven by ion solvation in water. Recently, there has been enhanced interest in the HCl solvation mechanism in the context of the minimal extent of solvation necessary for acid ionization, and the ability of confined-water media such as clusters and ice surfaces to induce the ionization. Here the latter issues are addressed for HCl adsorbate on cold ice particle surfaces. It is shown that these surfaces offer a range of adsorption sites, in which HCl freezes in different recognizable solvation stages progressing towards ionization and ion separation. The observed spectroscopic signatures of these stages are discussed and assigned with the help of Monte Carlo and ab initio calculations.

121.

AN INHOMOGENEOUS MODEL OF PROTEIN DIELECTRIC PROPERTIES: INTRINSIC POLARIZABILITIES OF AMINO ACIDS. *Xueyu Song, Department of Chemistry, Iowa State University, Ames, IA 50011, Fax: 515-294-0105, xsong@iastate.edu*

A simple inhomogeneous model of protein dielectric properties is discussed. A protein in solution is modeled as a collection of polarizable dipoles in a cavity embedded inside a dielectric medium. The intrinsic polarizabilities of 20 amino acids are assumed to be portable to all proteins in nature. A reasonable set of these polarizability values has been obtained by comparing dielectric fluctuations from molecular dynamics simulations with model calculations. The results are consistent within a data set of three small proteins. Applications to solvation dynamics in proteins will be discussed.

122.

VIBRATIONAL RELAXATION AT LIQUID INTERFACES. *Ilan Benjamin,**Department of Chemistry, University of California, Santa Cruz, 1156 High St., Santa Cruz, CA 95064, Fax: 831-459-2935, benjamin@chemistry.ucsc.edu*

The vibrational relaxation of several diatomic and triatomic molecules at the surface of liquid water and other liquids is studied using classical molecular dynamics computer simulations and compared with the same process in the bulk liquids. Both non-equilibrium classical trajectory calculations and equilibrium force autocorrelation functions are used to elucidate the factors that influence vibrational energy relaxation at the liquid surface region. We find that in general vibrational relaxation rates at interfaces are slower than in the bulk due to reduced friction. However, the degree of the slowing-down effect depends on the contribution of electrostatic forces and is correlated with the structure of the first solvation shell.

123.

QSAR STUDIES FOR ESTROGEN RECEPTOR PET TRACERS. *David E. Reichert, and Peter Wolohan, Mallinckrodt Institute of Radiology, Washington University School of Medicine, 510 S. Kingshighway Blvd., Campus Box 8225, St Louis, MO 63110, Fax: 314-362-9940, reichertd@mir.wustl.edu*

As part of our research in developing improved in vivo imaging agents we have developed enhanced Comparative Molecular Field Analysis (CoMFA) models for two isoforms of the estrogen receptor and to a serum protein, sex hormone binding globulin. We sought to predict the selectivity and binding affinities of a series of established and novel estrogenic ligands to both the alpha (ER α) and its newly discovered beta isoform (ER β). A training set of ligands with known binding affinities for both isoforms were built and a traditional CoMFA study was performed. These models were significantly enhanced through the inclusion of solvation terms and a description of the calculated binding energies for the training set. This has led to an accurate model for the prediction of the binding affinity and selectivity of novel estrogen subtype selective ligands has been developed and will hopefully lead to the development of novel subtype specific PET tracers.

124.

CHAIN MELTING TEMPERATURE ESTIMATION FOR PHOSPHATIDYLCHOLINES BY QUANTUM MECHANICALLY DERIVED QSPR. *Andrew J. Holder¹, Derek A. White¹, David M. Yourtee², Alan G. Glaros³, and Robert Smith⁴.*

(1) Department of Chemistry, University of Missouri - Kansas City, UMKC, Flarshheim Hall, Rm 410h, 5110 Rockhill Road, Kansas City, MO 64110, holdera@umkc.edu, whitede@umkc.edu, (2) Department of Pharmacology/School of Pharmacy, University of Missouri - Kansas City, (3) School of Dentistry, University of Missouri - Kansas City, (4) School of Pharmacy, University of Missouri - Kansas City

To further ongoing research into the chemical characteristics of monolayer phospholipid membranes, selected information obtained by the application of the AM1 semiempirical method to twenty-five phosphatidylcholines, each consisting of approximately 150 atoms, was used to develop a multilinear regression for predicting the chain melting temperatures (T_m) of these systems. The resulting model was then applied to twelve similar molecules NOT included in the training set and predictions of T_m to within experimental error were obtained. Prior to this study, models were either based on empirical data or on limited computational methods primarily confined to molecular mechanics methods because the molecules consisted of a prohibitively large number of atoms with respect to ab initio, semiempirical, or density functional quantum mechanical methods. Advances in hardware and software permit us to apply the AM1 semiempirical quantum mechanical method to molecules of this size for the first time.

125.

CLASSIFICATION OF HIV PROTEASE INHIBITORS AND PREDICTION OF THEIR ANTIVIRAL POTENCY FROM MOLECULAR STRUCTURE. *Suhas J. Patankar, and Peter C. Jurs, Department of Chemistry, Penn State University, 152 Davey Laboratory, Penn State University, University Park, PA 16802, sjp9@psu.edu*

Multi-drug-resistant HIV strains are emerging with the initial aggressive multi-drug treatment of HIV patients. This necessitates continued search for

novel inhibitors of viral replication. These protease inhibitors may further be useful as pharmacological agents against other viruses. Classification models of HIV Protease inhibitors are developed using a data set of 123 compounds containing several heterocycles. Their inhibitory concentrations expressed as log (IC₅₀) ranged from -1.52 to 2.12 log units. Initially a two-class problem is explored using several different classification methods including radial basis functions (RBF). A successful classifier was developed with six descriptors that showed predictive ability in the high 80% range for an external prediction set. In addition a three-class problem is also explored.

126.

DEVELOPMENT OF QSAR AND CLASSIFICATION MODELS TO PREDICT INHIBITION AND SELECTIVITY OF DHFR INHIBITORS. *Brian E. Mattioni, and Peter C. Jurs, Department of Chemistry, Pennsylvania State University, 152 Davey Lab, University Park, PA 16802, bem172@psu.edu*

QSAR models are developed to correlate chemical structure and inhibition towards the dihydrofolate reductase (DHFR) enzyme. Predictive models are built using inhibition data from the literature for *pneumocystis carinii* (pc) DHFR, *toxoplasma gondii* (tg) DHFR, and rat liver (rl) DHFR. Each compound is encoded with numerical descriptors, which capture topological, geometric, and electronic molecular features. Subsets of descriptors are analyzed via multiple linear regression, computational neural network, and support vector machine regression analyses to find predictive models. In some cases, support vector machines outperformed neural networks in terms of computational time and the overall root-mean-square errors of the training, cross-validation, and prediction sets. Furthermore, linear discriminant analysis (LDA) and probabilistic neural networks (PNNs) are used to construct models that classify compounds as selective or nonselective inhibitors of pcDHFR and tgDHFR relative to rlDHFR. A modified genetic algorithm was coupled with LDA to find optimal descriptor subsets. Additionally, PNNs afforded comparable results to LDA using smaller descriptor subset sizes in model development.

127.

NEW DEVELOPMENTS IN PEST SHAPE/PROPERTY HYBRID DESCRIPTORS.

Curt M Breneman¹, N Sukumar¹, C. Matthew Sundling¹, Lingling Shen¹, Bo Jiang¹, Bill Katt¹, Minghu Song¹, Hongmei Zhang¹, and Mark J. Embrechts². *(1) Department of Chemistry, Rensselaer Polytechnic Institute, Cogswell Laboratory, 110 8th Street, Troy, NY 12180-3590, Fax: 518-276-4045, brene@rpi.edu, (2) Decision Sciences and Engineering Systems, Rensselaer Polytechnic Institute*

Recent investigations have shown that the inclusion of hybrid shape/property descriptors together with 2D topological descriptors increases the predictive capability of QSAR and QSPR models. At the present time, PEST (Property-Encoded Surface Translator) descriptors may be computed using ab initio or semi-empirical electron density surfaces and/or electronic properties, as well as atomic fragment-based TAE/RECON property-encoded surface reconstructions. The new RECON/PEST algorithm also includes rapid fragment-based wavelet coefficient descriptor (WCD) computation. We now present the use of the RECON/PEST methodology in a virtual high-throughput mode, as well as the use of TAE properties for molecular surface autocorrelation analysis.

128.

SYSTEMATIC COMPARISON OF DENSITY FUNCTIONALS. *Xuelin Wang, Alan Lowe, Qi Wang, Justin Briggie, and Angela K. Wilson, Department of Chemistry, University of North Texas, Box 305070, Denton, TX 76203-5070, Fax: 940-565-4318, akwilson@unt.edu*

The performance of several density functionals including B3LYP, BLYP, B3PW91, BPW91, B3P86, BP86, and MPW1K has been investigated for a series of first- and second-row molecules. The basis sets used in the calculations include the correlation consistent basis sets (cc-pVxZ, where x=D(2), T(3), Q(4), 5), augmented correlation consistent basis sets (aug-cc-pVxZ), and the new ?pc? sets [Jensen, Journal of Chemical Physics, 201, 9113, (2001)]. Atomization energies and molecular structure are reported. A statistical error analysis including mean error, standard deviation, mean absolute deviation, and maximum error has been performed.

129.

AB INITIO STUDIES OF THE STRUCTURE AND BONDING OF SULFUR YLIDES.

Jean M. Standard, Tami K. Johnson, and Beth A. Copack, Department of Chemistry, Illinois State University, Normal, IL 61790-4160, standard@ilstu.edu

Ab initio studies of the structure and bonding of sulfur ylides have been performed at the MP2 level of theory using basis sets ranging from 6-31+G* to 6-311++G(2df,2pd). The sulfur ylides investigated were formed between singlet methylene and hydrogen sulfide or dimethylsulfide. In the optimized H₂S-CH₂ ylide structure, the calculated S-C intermolecular distance ranged from 1.683 Å at the MP2/6-31+G* level to 1.645 Å at the MP2/6-311++G(2df,2pd) level, exhibiting a basis set dependence as a result of the inclusion of f-type polarization functions. The binding energy of each of the complexes, measured relative to the separated monomers, was also determined with vibrational zero-point energy corrections. The calculated binding energy was 44 kcal/mol for the H₂S-CH₂ ylide and 63 kcal/mol for the CH₃SCH₃-CH₂ ylide at the MP2/6-311++G(2df,2pd) level. To further characterize the interaction between the singlet carbene and sulfide monomers, natural population analysis was used to determine the amount of charge transfer from the sulfide monomer to the carbene, and natural bond orbital analyses were performed to examine the bonding interactions in the ylides.

130.

2-AMINOETHOXYDIPHENYL BORATE, A COMPUTATIONAL AND EXPERIMENTAL STUDY.

Marsha A. Collins, and Thomas A. Holme, Department of Chemistry, University of Wisconsin - Milwaukee, 3210 North Cramer Ave., Milwaukee, WI 53211, marshac@uwm.edu

2-Aminoethoxydiphenyl borate (2-APB) is physiologically interesting because of its role as an inhibitor to the release of calcium from the endoplasmic reticulum of various types of cells. The structure of this molecule has been described several ways in literature - as a ring, open, and as a dimer. The studies reported in this poster detail computational and experimental approaches to understand this molecule and what factors contribute to the relative stabilities of the various possible structures.

131.

MECHANISTIC DETAILS OF ALDEHYDE DEHYDROGENASE CHEMISTRY FROM MM AND QM/MM SIMULATIONS.

Troy Wymore¹, Martin J. Field², David Deerfield II¹, Hugh B. Nicholas Jr.¹, and John Hempel³. (1) Biomedical Initiative Group, Pittsburgh Supercomputing Center, 4400 Fifth Avenue, Pittsburgh, PA 15213, Fax: 412-268-8200, wymore@psc.edu, (2) Laboratoire de Dynamique Moleculaire, Institut de Biologie Structurale, (3) Department of Biological Sciences, University of Pittsburgh

Aldehyde Dehydrogenases (ALDH) oxidize aldehydes to their corresponding carboxylic acids and require a cofactor, usually nicotinamide adenine dinucleotide (NAD). Since this catalytic cycle involves several individual reactions (nucleophilic attack, hydride transfer, and (de)acylation) experimental observations are sometimes difficult to interpret since a residue may provide crucial interactions throughout the entire cycle or to just one reaction. Our goal is to use molecular modeling methods to establish the individual roles of active site residues throughout the catalytic cycle. In this presentation, we will detail the results from 1) calculations to assign protonation states of ALDH by solving the appropriate Poisson-Boltzmann equations 2) molecular dynamics (MD) simulations on both the holo and substrate bound (benzaldehyde) form and 3) Quantum Mechanical/Molecular Mechanical (QM/MM) simulations of the initial events in the catalytic cycle. The QM/MM calculations utilize the AM1/OPLS potential to describe the role of proton transfer, proximal water molecules and other stabilizing factors to forming the transition state on the path to the thiohemiacetal intermediate. Cys-243 is shown to have a dual role: 1) forming the thiohemiacetal intermediate with the thiolate sulfur attacking the substrate carbonyl carbon and 2) the backbone amide proton strongly interacting with the substrate carbonyl oxygen.

132.

MODELS OF BIS(OXAZOLINE)-CU(II) COMPLEXES UPON DIELS-ALDER REACTIONS.

Jason DeChancie¹, Orlando Acevedo², and Jeffrey D. Evansek¹.

(1) Department of Chemistry and Biochemistry, Duquesne University, 600 Forbes Ave, Pittsburgh, PA 15282, Fax: 412-396-5683, jdechancie@hotmail.com, (2) Center for Computational Sciences, Duquesne University

The catalytic influence of C_s-symmetric bis(oxazoline)-Cu(II) complexes upon the Diels-Alder reaction have been modeled with the Cu(II) aqua ion. Density functional and Moller-Plesset theory have been used to explore the energetics and structures of the cyclopentadiene and acrylimide Diels-Alder reaction. The importance of the "twisted amide" in acrylimide in both the ground state and transition structure is discussed in terms of the activation barrier lowering and increased stereoselection. The results from the Cu(II) aqua ion will be related to the bis(oxazoline)-Cu(II) catalytic system to clarify how the catalyst enhances the rate and selectivity of the Diels-Alder reaction.

133.

REACTIONS OF METAL ATOMS WITH CARBONYL COMPOUNDS:**COMPARISONS OF THEORY AND EXPERIMENT.**

Craig A. Bayse, Department of Chemistry and Biochemistry, Old Dominion University, Hampton Blvd, Norfolk, VA 23529, Fax: 7576834628, cbayse@odu.edu

Recent experimental studies of reactions of metal atoms in crossed-molecular beams by the Davis group at Cornell University have yielded many interesting results. The reaction of compounds such as HCHO and MeCHO with metal atoms, the initial step is the formation of an η²-carbonyl complex. Subsequent steps involve insertion of the metal into the C-H and C-C bonds of the ligand. Observed final products include hydrogen, CO, MCO, MH₂, M-ketene, MH₂-ketene, and M(H)(Me); but no methane is observed. In this work, we examine the potential energy surfaces of the reactions of yttrium and niobium with the simple aldehydes above using DFT and CCSD(T) methods. The results are in very good agreement with experiment.

134.

AB INITIO DIRECT DYNAMICS SIMULATION OF CH₃CH₂F → CH₂=CH₂ + HF UNIMOLECULAR DISSOCIATION: A CASE STUDY TO COMPARE WIGNER AND QUASICLASSICAL INITIAL CONDITION SAMPLING METHODS.

Lipeng Sun, Department of Chemistry, Wayne State University, Detroit, MI 48202, lpsun@chem.wayne.edu, and William L. Hase, Department of Chemistry, Institute for Scientific Computing, Wayne State University

Solving classical equations of motion and comparing classical trajectory simulation results with experiment require the proper selection of initial conditions. With the development of computer technology and electronic structure calculation methods, it is now possible to perform *ab initio* direct dynamics trajectory simulations. In this work, the Wigner sampling technique, for which the initial normal mode coordinates and momenta are selected according to their quantum mechanical distribution in phase space, was implemented and applied in an *ab initio* direct dynamics study of CH₃CH₂F → CH₂=CH₂ + HF unimolecular dissociation. The trajectories were initialized at the dissociation saddle point and directed towards products. The effect of Wigner sampling of initial conditions on the simulation results was compared with the results obtained using quasiclassical sampling method.

135.

DIRECT DYNAMICS SIMULATION OF PROTONATED DIGLYCINE SURFACE

INDUCED DISSOCIATION. **Yanfei Wang**, and William L. Hase, Department of Chemistry, Wayne State University, chemistry building, Wayne State University, Detroit, MI 48201, yfwang@chem.wayne.edu, wlh@mercury.cs.wayne.edu

Direct dynamics classical trajectory simulations are performed to study fragmentation pathways and energy transfer in collisions of protonated diglycine with a hydrogenated diamond{111} surface, for a collision energy E_i at 30,50,70 and 100 eV and a collision angle of 0 degree. From the fragmentation dynamics, the reaction path which forms the NH₂CH₂⁺ + HCONHCH₂COOH is the most probable pathway among the different reaction types. With increase in the initial translational energy E_i, the percent fragmentation increases. As E_i is increased, energy transfer to the ion E_{int} and the final translational energy E_f decrease, and energy transfer to the diamond surface E_{surf} increases. For reactive trajectories, E_{int} is much higher and E_{surf} much lower than for non-reactive trajectories. E_f is nearly the same for the two sets of trajectories. Shattering fragmentation, in

which as diglycine dissociates as it collides with the surface, becomes important as E_i is increased. Both energy transfers to the ion and to the surface are nearly the same for shattering and non-shattering reaction.

136.

QUANTUM-MECHANICAL REACTION RATE AND VIBRATIONAL ENERGY RELAXATION RATE CONSTANTS FROM CENTROID MOLECULAR DYNAMICS (CMD) SIMULATIONS. *Qiang Shi, and Eitan Geva, Departments of Chemistry, University of Michigan, 930 N University, Ann Arbor, MI 48108, Fax: 734-647-4865, qshi@umich.edu, eitan@umich.edu*

The calculation of quantum real-time correlation functions for many-body systems remains an important challenge in theoretical chemical physics. An approximation for such correlation functions can be obtained from Centroid Molecular Dynamics (CMD) simulations. However, in order for real-time correlation functions obtained from CMD simulations to be directly related to the corresponding quantum correlation functions, one of the operators must be linear in position and/or momentum. This requirement makes it difficult to apply CMD to the calculation of reaction rate and vibrational energy relaxation (VER) rate constants, which are usually expressed in terms of the flux-flux and force-force correlation functions, respectively. This is because the relevant correlation functions involve two nonlinear operators, and therefore cannot be calculated via CMD without further approximations. We present alternative, and completely equivalent, theories of reaction rates and VER, which are based on linear and nonlinear response theory, and that make it possible to avoid this problem. These new theories open the door for a direct application of CMD to the reaction rate and VER problems. For the reaction rate theory, we tested the new method on a system consisting of a double-well potential bi-linearly coupled to a harmonic bath, and found the CMD results to be in good agreement with the exact results for a wide range of frictions and temperatures. For the VER theory, we tested the new method, with relative success, on several model systems, including a Lennard-Jones liquid.

137.

THEORETICAL STUDIES OF THE COLLISION-INDUCED DISSOCIATION OF METAL CLUSTERS. *Pascal Larrégaray, and Gilles Peslherbe, Centre for Research in Molecular Modeling and Department of Chemistry & Biochemistry, Concordia University, 1455, de Maisonneuve Blvd West, Montreal, QC H3G 1M8, Canada, pascal@cermm.concordia.ca*

The unimolecular dissociation of Aluminum clusters (Al_6 and Al_{13}) following a collision with a rare gas atom or a surface are investigated by classical dynamical simulations. Such processes are usually thought to proceed through rapid internal energy redistribution (IVR) followed by statistical dissociation on the ground state energy surface. Assuming a simple kinetic scheme of IVR and statistical unimolecular dissociation, we attempt to characterize the actual rate of IVR, so that error bars can be assigned to binding energies inferred from experimental measurements with statistical models which intrinsically assume an infinite rate of IVR.

138.

STUDY OF THE STRUCTURAL AND DYNAMICAL PROPERTIES OF A BIOMIMETIC COMPOUND OF DIIRON PROTEINS VIA AB INITIO AND HYBRID (QM/MM) MOLECULAR DYNAMICS SIMULATIONS. *Alessandra Magistrato¹, Ursula Röthlisberger², and Michael L. Klein¹. (1) Department of Chemistry, University of Pennsylvania, 231 S 34 Street, Philadelphia, PA 19104, Fax: 215-573-6233, ale@cmm.upenn.edu, (2) Department of Chemistry, ETH Zürich*

A biomimetic four-helix bundle with a binuclear active site (DF1), bearing Zn, Mn or Fe as transition metals, has been synthesized and characterized. The carboxylate bridged binuclear motif of DF1 resembles the active site of numerous binuclear enzymes, such as Manganese Catalase, Methane Monooxygenase (MMO), etc.

Due to the crucial chemical and biological relevance of binuclear enzymes in hydrolytic as well as redox active processes, we have performed a systematic study of the structural properties of the dizinc analogue of DF1 through *ab initio*

and hybrid QM/MM Car-Parrinello molecular dynamics simulations. Different quantum mechanical models of the active site have been chosen in order to gain a qualitative view of how the first and second shell ligands can tune structural and dynamical properties of the active site. In addition, the geometrical restraints and the electrostatic stabilization provided by the whole structure have been explicitly considered, performing QM/MM molecular dynamics simulations.

139.

BRIDGED BINUCLEAR METAL MOTIF: GENERAL FEATURES THROUGH SIMPLE MG(II) AND ZN(II) MODELS. *Petra Munih, Department of Chemistry, The Center for Molecular Modeling, University of Pennsylvania, 231 S. 34th St., Philadelphia, PA 19104-6323, Fax: 215-573-6233, munih@cmm.chem.upenn.edu, Martin Karplus, Dept. of Chem. & Chem. Biol, Harvard University, Gregory A. Petsko, Rosenstiel Basic Medical Sciences Research Center, Brandeis University, and Dagmar Ringe, Departments of Biochemistry and Chemistry, Brandeis University*

The "bridged binuclear metal" motif appears in a number of hydrolytic metallo-enzyme active sites. It consists of two metal cations (usually divalent) bridged by an endogenous ligand (usually a carboxylate, coordinating the metals in a bidentate fashion), an exogenous one (a water molecule, a substrate, hydroxide) or both. In most of these enzymes, the motif has been found to be critical for catalysis. We have investigated the role of the "bridged binuclear metal" motif with Mg^{2+} as the metal, water or hydroxide as the exogenous bridging ligands and water terminal ligands by *ab-initio* DFT. Subsequently, an aminopeptidase substrate analog was added to the motif to describe a prototype amide hydrolysis reaction with Zn^{2+} and Mg^{2+} as the metals. Their respective coordination spheres were completed by either terminal water ligands or aminopeptidase-like ligands. Implications of the results for understanding the role of the motif are discussed.

140.

MOLECULAR DYNAMICS SIMULATIONS OF A SMALL PROTEIN IN AQUEOUS SOLVENT USING A FLUCTUATING CHARGE FORCEFIELD. *Sandeep A. Patel¹, Alexander D. MacKerell Jr.², and Charles L. Brooks III¹. (1) Department of Molecular Biology, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, CA 92037, sandeep@scripps.edu, (2) University of Maryland*

We present results of molecular dynamics simulations of small proteins in aqueous solvent using a fluctuating charge forcefield. We will address issues pertinent to the application of such fluctuating charge models to biomolecular systems as they concern dynamical and equilibrium properties of the systems of interest.

141.

MOLECULAR MODELING OF NAALADASE/PMSA INHIBITORS. *Di-Fei Wang¹, Olaf Wiest¹, Paul Helquist², Marvin Miller¹, and Martin Tenniswood³. (1) Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN 46556, dwang@nd.edu, (2) Walther Cancer Research Center, University of Notre Dame, (3) Department of Biological Sciences, University of Notre Dame*

Prostate specific membrane antigen (PMSA) is expressed exclusively by prostate tumor cells and is a promising target in the development of a new method for fast reliable diagnosis of prostate cancer. In this study, we will present our recent modeling results directed to the design of novel NAALADase/PMSA inhibitors. Homology models of NAALADase/PMSA based on the transferrin receptor and other homologous structures as well as docking of some inhibitors into its active site will be presented.

142.

GTP-BINDING AND ASSOCIATION OF THE *ESCHERICHIA COLI* SIGNAL RECOGNITION PARTICLE PROTEIN, FFH, AND ITS RECEPTOR, FTSY.

Kelly M. Elkins¹, Irmgard Sinning², and Rebecca C. Wade¹. (1) Molecular and Cellular Modeling Group, European Media Laboratory, Schloss-Wolfsbrunnengasse 33, D-69118 Heidelberg, Germany, Fax: +49 6221 533 298, kelly.elkins@eml.villa-bosch.de, (2) Biochemiezentrum, Universität Heidelberg

The signal recognition particle (SRP) is a universally conserved system for protein trafficking. Many SRPs are GTP-binding proteins. Their crucial role in

ensuring that proteins are not misplaced into the wrong cellular location makes them potential targets for drug design. The aim of our work is to derive structural models of the interactions of the SRP and its receptor. This is being done for the SRP system from *Escherichia coli*, which is much simpler than that found in humans and thus provides a good model system. The *E. coli* SRP is a ribonucleoprotein complex composed of a 48 kDa GTPase protein and a 4.5 S RNA. While the crystal structure of the *E. coli* SRP receptor FtsY, also a GTPase, has been solved, the structure of the *E. coli* SRP protein itself has not. Consequently, we have used comparative modelling techniques to build a model of the *E. coli* SRP protein, Ffh, on the basis of SRP structures from other organisms and to dock GTP and Mg²⁺ into their hypothesized sites on both Ffh and FtsY. These modelled structures are being used to build a model of the active SRP:SRP receptor complex. This model should prove useful in understanding how the components of the SRP interact to direct protein traffic within and out of the cell.

143.

IMPACT OF LIPID STRUCTURE AND DYNAMICS UPON LASER-INDUCED

FLUORESCENCE DETECTION AND SEPARATION. *Anne Loccisano¹, Jeffrey D. Evanseck², and Mitchell Johnson².* (1) Center for Computational Sciences, Duquesne University, Department of Chemistry and Biochemistry, 600 Forbes Avenue, Pittsburgh, PA 15282, Fax: (412)396-5683, aloccisano@hotmail.com, (2) Department of Chemistry and Biochemistry, Duquesne University

We are developing capillary electrophoresis with laser-induced fluorescence detection in order to separate lipids, using a positively charged cyanine fluorescent dye. The size, shape, and dynamics of the lipid molecule and the separation medium used are thought to affect the separation and detection of lipids. Molecular dynamics simulations have been carried out to assess the structure and dynamics lipids of different size and shape. A model lipid, deoxylysophosphatidylcholine (LPPC), has been simulated in both water and methanol with box dimensions of 43.6 × 20.4 × 18.3 angstroms with periodic boundary conditions and Ewald electrostatics. It is also of interest to examine the thermodynamic properties of lipids and the differences in their properties due to their size and shapes. The long-term goal is to apply this data to biological systems, such as predicting how lipids bind to substrates depending on shape, size, or degree of unsaturation of the molecule.

144.

STRUCTURE BASED INHA INHIBITOR DESIGN AIDED BY OWFEG. *Guanglei Cui¹, Dawn Marin², Peter Tonge¹, and Carlos Simmerling¹.* (1) Department of Chemistry, State University of New York at Stony Brook, Stony Brook, NY 11790, Fax: 631-632-7960, cui@morita.chem.sunysb.edu, (2) Department of Pharmacology, State University of New York at Stony Brook

InhA, a long-chain enoyl reductase, is essential for mycolic acid biosynthesis in *Mycobacterium tuberculosis*. Inhibition of InhA by the front-line antituberculosis drug isoniazid compromises mycolic acid biosynthesis and consequently interferes with the formation of the bacterial cell wall, however, drug resistance has been developed. The long term goal of this project is to design and synthesize novel InhA inhibitors with the objective of generating novel antituberculosis compounds. The focus of the current study is triclosan, an easily-functionalized, antibacterial additive in many consumer products. While triclosan is a picomolar inhibitor of the homologous enzyme (FabI) from *Escherichia coli*, this compound is only a submicromolar inhibitor of InhA. Using the X-ray structure of triclosan bound to FabI we are exploring modifications to the triclosan core structure that will improve its affinity for InhA. In this study, we employed molecular dynamics (MD) techniques and a one-window free energy grid (OWFEG) method in the design of new triclosan-derived inhibitors. Assuming a similar ligand/enzyme binding motif, we generated the structure of the InhA:NAD⁺:triclosan complex by docking triclosan into InhA using FabI as a template. The resulting structure was then further refined by a 3.5ns explicit solvent MD simulation. OWFEG calculations were performed to identify possible modifications that could enhance the binding. As a control, the same OWFEG calculation was carried out for the FabI system. Five triclosan analogs whose binding affinities have been determined experimentally were evaluated using the OWFEG grid. The results was reasonably consistent with their IC₅₀ measurements. Triclosan analogs suggested by the InhA OWFEG grid are currently being synthesized and their affinity for InhA will be determined.

145.

THE DETERMINATION OF PKA'S IN PROTEINS. *Rasha R. Abd El-Rahman, Jeffrey D. Madura, and Charles T. Dameron, Department of Chemistry and Biochemistry, Duquesne University, 600 Forbes Avenue, 308 Mellon Hall of Science, Pittsburgh, PA 15282, Fax: 412-396-5683, rrrehab@yahoo.com, Madura@duq.edu*

Cells utilize a variety of mechanisms to deal with essential, yet toxic trace elements such as copper. Recent studies have shown that cells use proteins called "metallochaperones" to perform the routing of these metal ions to protect the cell and to ensure delivery to nascent enzymes. The mechanism of binding and release of metals by the metallochaperones is not well understood. The largest of the families of copper chaperones has a common structure known as an "open-faced b-sandwich", consisting of two α -helices overlaying four β -strands. An exposed -Cys-X-X-Cys- metal binding motif is positioned in a loop near one end of the protein. A similar motif is seen in the thioredoxin family of proteins. The cysteinyl thiolates in the thioredoxin, which markedly altered pKa's, serve a catalytic function. We are proposing that, as in the thioredoxins, that the thiolates have perturbed pKa's and furthermore that these differences are important in the release of the metal ions. In order to establish an appropriate procedure to calculate pKa values in the copper chaperones, we are using the program UHBD to calculate the pKa values of cysteines in the thioredoxin proteins. Using a protein dielectric constant ($\epsilon_p=20$) we calculate a pKa value of 7.8 for Cys 32 found in the e-coli thioredoxin this calculated value is consistent with the experimental value of 7.4. The pKa shift for Cys 32 of e-coli thioredoxin has traditionally been rationalized because of its position at the N-terminus end of an α -helix. We are applying this method to the metallochaperones to estimate the thiolate pKa's and investigate their effect on metal binding.

146.

PROBING THE LINK BETWEEN CATALYTIC RATE AND ENZYME DYNAMICS IN THE HYDRIDE TRANSFER REACTION CATALYZED BY DIHYDROFOLATE

REDUCTASE. *Ian F. Thorpe, Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, TPC-06, La Jolla, CA 92037, Fax: 858-784-8688, ithorpe@scripps.edu, and Charles L. Brooks III, Department of Molecular Biology, TPC-6, The Scripps Research Institute*

The enzyme dihydrofolate reductase (DHFR) catalyzes the reduction of 7,8-dihydrofolate (DHF) to 5,6,7,8-tetrahydrofolate (THF) utilizing the nucleotide cofactor NADPH. Mutational studies have identified amino acid residues that, while located far away from the active site (up to 19 Å), have an effect on the catalytic activity of the enzyme when altered. NMR and computational studies have suggested that the presence of these mutations is associated with the absence of certain correlated motions in the protein. We have studied the link between the presence of these mutations and the activation energy for the reaction using combined quantum mechanical/ molecular mechanical (QM/MM) simulations and demonstrate that significant differences in the magnitude and distribution of activation energies exist between wild-type DHFR and certain of these mutants.

147.

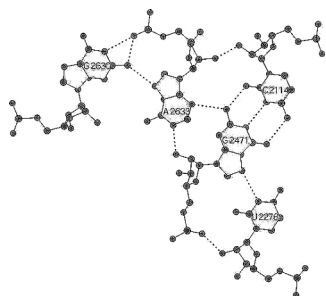
FINDING BASE-PAIRS AND HELICAL REGIONS IN NUCLEIC ACID

STRUCTURES. *Xiang-Jun Lu, Yu-Rong Xin, and Wilma K Olson, Department of Chemistry and Chemical Biology, Rutgers University, 610 Taylor Road, Piscataway, NJ 08854, Fax: 732 445-5958, xiangjun@rutchem.rutgers.edu, xinyu@rutchem.rutgers.edu*

We present a novel algorithm for finding all possible base-pairs (bps), higher order base associations, and double helical regions in nucleic acid structures. Starting from the reference frame defined for an ideal Watson-Crick bp, the method defines a bp using several simple geometrical parameters and works well for canonical as well as uncommon pairs. Horizontally connected bps form a network of higher order base association, and vertically stacked bps form double helical regions. These elements form the basis for identifying RNA structural motifs.

Each bp is uniquely characterized by geometric parameters and hydrogen bonding patterns, which can be used in a searchable database. When applied to the large ribosomal subunit (50S), the method finds 23 out of the 28 possible bps with at least two hydrogen bonds between the bases. Other bps with

hydrogen bonds involving backbone atoms are also identified, along with base triplets, tetrads, and pentads.



148.
LOCATING THE EXPERIMENTAL STRUCTURE OF A NONHELICAL NUCLEIC ACID BY COMBINED LOCALLY ENHANCED SAMPLING AND CONTINUUM ELECTROSTATICS APPROACH. Xiaolin Cheng, and Carlos Simmerling, Department of Chemistry, State University of New York at Stony Brook, Stony Brook, NY 11794, xcheng@ic.sunysb.edu

In this article we show for the first time the combination of Generalized Born (GB) continuum solvation model and Locally Enhanced Sampling (LES). It has been shown that the reduction in barrier heights provided by LES can result in solvated simulations that are more than an order of magnitude more efficient than single copy methods. Continuum solvation methods can also accelerate conformational transitions due to reduction in friction. Since an exact combination of the two algorithms results in a significant increase in computational requirements, we will present approximations that greatly increase efficiency. We have applied the resulting combination to simulation of conformational change in an RNA UUCG tetraloop and show that the combined GB/LES approach is more efficient than use of either GB or LES alone. We believe that this approach is likely to be an important component in all-atom protein structure refinement, where both extensive conformational sampling and the accuracy of the energy function are of critical importance.

149.
MOLECULAR DYNAMICS SIMULATION AND FREE ENERGY ANALYSIS OF POLYCYCLIC AROMATIC CARCINOGEN-DAMAGED DNA: ACCURATE REPRESENTATION AND NEW INSIGHTS. Shixiang Yan¹, Min Wu¹, Nicholas E. Geacintov¹, and Suse Brody². (1) Department of Chemistry, New York University, New York, NY 10003, sy8@scires.nyu.edu, mw402@nyu.edu, (2) Department of Biology, New York University

Benzo[a]pyrene (BP), a prototype polycyclic aromatic hydrocarbon (PAH), widespread in our environment, is present in tobacco smoke, automobile exhaust, and as a contaminant in foods. Upon metabolic activation, BP can be converted to a pair of highly reactive enantiomeric diol epoxides, known as (+)- and (-)-anti-BPDE. Both (+)- and (-)-anti-BPDE react primarily with the guanine residues in DNA by trans epoxide opening, to form the 10S (+)- and 10R (-)-trans-anti-[BP]-N²-dG adducts. The BP moiety resides in the minor groove of the modified DNA duplex in both adducts. However, the 10S (+)-trans adduct has the pyrenyl ring system directed 5' along the modified strand, while in the 10R (-)-trans adduct this ring system is oriented in the 3' direction. This system has been experimentally well characterized both structurally and thermodynamically. We have carried out 3-ns molecular dynamics (MD) simulations, using NMR solution structures (Cosman *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 1914; de los Santos *et al.*, *Biochemistry*, **1992**, *31*, 5245) as the starting models for the 10S (+)- and 10R (-)-trans adducts in a DNA duplex 11-mer. Then we applied the MM-PBSA (molecular mechanics Poisson-Boltzmann surface area) method (Kollman *et al.*, *Acc. Chem. Res.* **2000**, *33*, 889) to compute the free energies of the two adducts. We also calculated the NOE distances from the MD trajectories, and compared them against the experimental NMR NOE distances. Our results showed that the simulated structures are in good agreement with the NMR experimental NOE data, and our complete thermodynamic analysis agrees well with experimental thermodynamic characterization of these adducts (Marky *et al.*, *J. Am. Chem. Soc.* **1996**, *118*, 3804). Moreover, the enhanced exposure to solvent of the BP aromatic ring system in the 10S (+)-trans adduct, deduced through experimental thermodynamic investigations, is structurally rationalized.

150.
PROBING THE MECHANISM OF BASE FLIPPING IN DNA VIA POTENTIAL OF MEAN FORCE CALCULATIONS. Niu Huang¹, Niles K. Banavali², and Alexander D. MacKerell Jr.¹. (1) Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, 20 N. Pine Street, Baltimore, MD 21201, nhuan001@umaryland.edu, (2) Biochemistry and Structural Biology, Weill Medical College of Cornell University

DNA methylation is important for the protection of "self" DNA in bacteria and is known to contribute to the regulation of gene expression in eukaryotes. Essential for the methylation of DNA is the requirement that the target base flip out of the DNA duplex. The exact structural mechanism of base flipping in DNA and DNA-protein complexes is still not clear. In this study we investigated the flipping of cytosine out of duplex DNA in DNA alone and when bound to the (cytosine-C5)-methyltransferase from HhaI (M.HhaI) via potential of mean force calculations. In DNA alone it is shown that spontaneous flipping can occur via both the minor and major grooves. Calculations in the presence of M.HhaI indicate that the enzyme actively facilitates the base pair opening in the ternary complex, with flipping occurring via the major groove, in contrast to the minor groove path suggested by analysis of X-ray crystal structures.

151.
MOLECULAR DYNAMICS SIMULATIONS ON THE WILD-TYPE HUMAN TRANSTHYRETIN AND ITS VARIANTS. Mingfeng Yang, Ming Lei, and Shuanghong Huo, Department of Chemistry and Biochemistry, Clark University, 950 Main Street, Worcester, MA 01610

Transthyretin(TTR) is one of the known human amyloidogenic proteins. Its native state is a homo-tetramer with each monomer having a beta-sandwich structure. Strong experimental evidence suggests that TTR dissociates to a monomeric intermediate state and that the monomers subsequently self-assemble to form amyloid deposits and insoluble fibrils. Our aim is to study the local and global conformational perturbations caused by the single-site mutations, V30M and L55P, in aqueous solution. We begin with the simulations of monomeric wild-type TTR, V30M, and L55P mutants. Nanosecond time scale molecular dynamics simulations at 300K were performed using AMBER. The protein was fully solvated with a rectangular box of TIP3P water molecules. Long-range electrostatic interactions were treated with the particle mesh Ewald summation. The observations of the MD simulation of the wild type TTR is consistent with those from H-exchange experiments. Several residues are identified as sites with intrinsic instability in the wild-type and mutants.

152.
FOLDING AND REFOLDING SIMULATIONS OF BETA-HAIRPIN PEPTIDES IN EXPLICIT SOLVENT. Hongxing Lei, and Paul E. Smith, Department of Biochemistry, Kansas State University, 36 Willard Hall, Manhattan, KS 66506-3702, hle5757@ksu.edu

The folding mechanism of secondary structures is fundamental to an understanding of the protein folding problem. The simplest beta-sheet structure is called a beta-hairpin and contains one turn connecting two strands. Two model peptides (P2: SYINSNGTWT and P5: YITNSNGTWT), with known high populations of a single 3:5 beta-hairpin structure, were chosen for theoretical study. Six molecular dynamics simulations (20 ns each) in explicit solvent were performed to investigate the mechanism of beta-hairpin folding. Two misfolded structures, two extended structures, one correct and one irregular hairpin structure, were used as starting conformations to study the sampling efficiency and folding capability in explicit solvent. Both misfolded structures and one extended structure folded to the correct native structure within 13 ns, according to proton-proton distances corresponding to the experimentally observed Nuclear Overhauser Enhancements. A clustering analysis showed that several turns, involving the residue pairs NS, SN and NG, were highly populated during the simulations. One folding pathway proceeded through an incorrect beta-turn. Another pathway involved an initial loop structure, which then folded according to the zip-down mechanism. Hence, multiple pathways were observed, although no evidence for the zip-up mechanism was found. The formation of cross-strand side chain hydrophobic contacts (residues 1 to 9, 2 to 10 and 3 to 9) were correlated with the folding events, but no apparent correlation between solvent accessible surface area and the folding process was observed. The simulations suggest that the correct register for hairpin formation in this system is deter-

mined by a balance between turn preferences, side chain interactions, and main chain hydrogen bonding.

153.

APPLICATION OF SIMPLICIAL NEIGHBORHOOD ANALYSIS OF PROTEIN PACKING (SNAPP) TO BINDING PROTEINS THAT UNDERGO

CONFORMATIONAL CHANGE. Douglas B. Sherman¹, Shuxing Zhang², J. Bruce Pitner¹, and Alexander Tropsha². (1) BD Technologies, 21 Davis Dr, Research Triangle Park, NC 27709, Fax: 919-597-6400, Douglas_Sherman@bd.com, (2) School of Pharmacy, University of North Carolina at Chapel Hill

Bacterial periplasmic binding proteins (PBPs) transport nutrients into the cytoplasm. PBPs generally consist of two globular domains connected by strands forming a hinge. During ligand binding, hinge motion changes the conformation from the open to closed form. Both forms can be crystallized easily without a ligand, suggesting that the energy difference between the two forms is small. We applied Simplicial Neighborhood Analysis of Protein Packing (SNAPP) to evaluate the relative stability of open and closed forms in PBPs. Using united residue representation of amino acids, SNAPP performs Delaunay tessellation of the protein, producing tetrahedrons with nearest neighbor residues at the vertices. SNAPP calculates log-likelihood scores for all possible quadruplet compositions of amino acids from representative proteins, and the sum of scores for a given protein provides the total SNAPP score. Results of scoring for PBPs will be discussed in terms of the relative energetic stability of the two forms.

154.

A COMPARISON OF THE LOW MODE AND MONTE CARLO CONFORMATIONAL SEARCH METHODS. Jennifer Pratt, and Carol Parish, Department of Chemistry,

Hobart and William Smith Colleges, Box 4178, Geneva, NY 14456, parish@hws.edu

Conformational search methods are used to explore molecular potential energy surfaces and to identify low energy structures on the surface yielding information about molecular flexibility and conformational behavior. This study compares the Low Mode and Monte Carlo conformational search methods on three potent, FDA-approved HIV-1 protease inhibitor drugs; ritonavir, amprenavir and TMC-126. The efficiency of the methods will be discussed. Conformationally accessible structures will be presented along with a cluster analysis performed to identify major conformational families.

155.

COMPARING THE CONFORMATIONAL BEHAVIOR OF A SERIES OF

ENANTIOMERIC CYCLIC UREA HIV-1 INHIBITORS. Kristin Schram, and Carol Parish, Department of Chemistry, Hobart and William Smith Colleges, Box 4178, Geneva, NY 14456, parish@hws.edu

Drug molecules based on the cyclic urea scaffold have been shown to be effective inhibitors of the HIV protease. An enantiomeric series of cyclic ureas that exhibit significantly different binding affinities has been reported in the literature. In this study the conformational flexibility of ten of these cyclic urea enantiomers was determined using the Low Mode and Monte Carlo conformation search strategies with the AMBER^{*} force field and the GB/SA solvent model for water. The efficiency of each method will be compared and the differences in molecular flexibility will be discussed.

156.

MOLECULAR MODELING AND CONFORMATIONAL ANALYSIS OF

POLYAZAMACROLIDES. Emelyn Smith, and Carol Parish, Department of Chemistry, Hobart and William Smith Colleges, Box 4178, Geneva, NY 14456, parish@hws.edu

The structure and conformational flexibility of a series of polyazamacrolides have been studied using the Low Mode – Monte Carlo hybrid conformational search method. We are interested in the conformational behavior of these molecules as they have been reported as crucial constituents in a remarkably complex defense secretion isolated from the ladybird beetle *epilachna borealis*. The secretion contains a combinatorial library of differently sized polyazamacro-

lides built up from three different saturated (w-1)-hydroxyethyl-amino acids. The structural diversity of the arthropod exudate is even further enhanced as the polyazamacrolides have been shown to undergo multiple intramolecular O-to-N acyl migrations. We are interested in how the conformational flexibility and dynamical nature of this system affects its ability to initiate an O-to-N acyl migration via a five membered ring intermediate. Conformationally accessible structures will be presented along with a cluster analysis performed to identify major conformational families.

157.

COMPARING THE CONFORMATIONAL FLEXIBILITY OF HIV-1 INHIBITORS. Julia

James, Jennifer Pratt, Kristin Schram, and Carol Parish, Department of Chemistry, Hobart and William Smith Colleges, Box 4178, Geneva, NY 14456, parish@hws.edu

The human immunodeficiency virus (HIV) encodes an aspartyl protease enzyme (HIV-PR) that cleaves viral polyprotein precursors, allowing the maturation of the HIV virus that causes the autoimmune deficiency disease (AIDS). Protease inhibitors occupy the basket-shaped active site, interfering with the functioning of the enzyme and preventing the maturation of the virus. This study will use molecular modeling methods to study the conformational behavior in water of different FDA-approved HIV protease inhibitor drugs such as indinavir, saquinavir, nelfinavir, amprenavir, lopinavir and ritonavir to determine the similarities and differences in the way these drugs interact with the protease active site. Conformationally accessible structures will be presented along with a cluster analysis performed to identify major conformational families for each inhibitor.

158.

AUTOMATING THE PREDICTION OF ENZYME ACTIVE SITES FROM STRUCTURE ALONE IN THEMATICS. Robert P. Futrelle¹, Ronald J. Williams¹, Wenxu Tong¹,

and Mary Jo Ondrechen². (1) College of Computer Science, Northeastern University, Boston, MA 02115, (2) Department of Chemistry, Northeastern University

THEMATICS is a method that uses theoretical microscopic titration curves to identify active-site residues in proteins of known structure. We have shown that predicted titration curves that do not have the typical Henderson-Hasselbalch shape are significant for proteins; they are markers of active site location and chemical reactivity. We are developing machine-learning methods to perform active site identification automatically. The training data consist of titration curves computed for all of the ionizable residues of each protein where the active site has already been well characterized by biochemists. In the initial studies, titration curve parameters along with relative locations are used as inputs to neural network algorithms. The parameters include effective pKa, maximal slope, and the pH range over which the curve separates from its asymptotes. Estimates are presented of performance on as yet unseen enzymes (cross-validation). Lessons learned from these initial studies are described.

159.

COMPUTATIONAL APPROACHES FOR PREDICTING ANDROGENIC ACTIVITY: HOMOLOGY MODELING OF THE ANDROGEN RECEPTOR AND CALCULATION

OF LIGAND BINDING AFFINITIES. Ai Ni¹, Seong-Jae Yu¹, Robert K. DeLisle², and William J. Welsh¹. (1) UMDNJ, Piscataway, NJ 08854, aini@umdnj.edu, (2) Accelrys

Classically considered the “male” sex hormone, dihydrotestosterone is critically involved in numerous physiological processes. In addition to initiating male sexual differentiation and development of male secondary characteristics, dihydrotestosterone has been implicated in blood pressure regulation, obesity, as well as bone development in conjunction with estrogen. For the past several years, there has been growing concern regarding the impact of synthetic chemicals within the environment upon normal physiological processes of both wildlife and humans. The widespread use of various insecticides, fungicides, and herbicides has resulted in the accumulation of these compounds within the environment. While the impact upon human health status has yet to be fully realized, it is apparent that measures must be taken in order to address these issues. However, standard experimental techniques to evaluate the effects and mechanisms of action of chemical entities are both labor and cost intensive. The

extremely strong conservation of structural characteristics among those nuclear hormone receptors for which structural data exist suggests that this family of receptors is highly amenable to homology modeling. In this study we present the development of a homology model of the androgen receptor ligand-binding domain (LBD) based upon the existing crystal structure of the progesterone receptor, a highly related member of the nuclear hormone receptor superfamily. The recently released x-ray crystal structures of androgen receptor enabled us to evaluate our homology model. Values of the calculated binding energy (BE) to the androgen receptor for a series of 25 structurally diverse compounds showed good correlation with the corresponding experimentally observed values. This research is partially funded by a grant from the U.S. Environmental Protection Agency's Science to Achieve Success (STAR) program.

160. COMPUTATIONAL STUDIES OF RALOXIFENE (EVISTA) DERIVATIVES: 3D-QSAR/COMFA MODELS AND BINDING ENERGY CALCULATION AS A GUIDE FOR SELECTIVE ESTROGEN RECEPTOR MODULATORS (SERMS). *Seong-Jae Yu¹, David Derington², and William Welsh¹. (1) UMDNJ, Piscataway, NJ 08854, yuse@umdnj.edu, (2) Tyco Healthcare*

Raloxifene, marketed by Eli Lilly at Evista[®], was approved by the FDA for the treatment and prevention of osteoporosis in postmenopausal women. We have developed several predictive QSAR models to guide the rational design of novel analogs exhibiting improved binding to and selectivity for the estrogen receptor. These QSAR models were developed based on the relative binding affinity (RBA) data of Grese et al. for 76 raloxifene derivatives (including raloxifene itself). We will present several three-dimensional QSAR (3D-QSAR) models derived from Comparative Molecular Field Analysis (CoMFA) that exhibit excellent self-consistency ($r^2 > 0.90$) and predictive ability ($rCV2 > 0.50$). As a new alignment scheme, receptor-based alignment scheme was applied using Insight II program and compared with Sybyl alignment schemes. Binding energy was calculated for the purposes of predicting binding affinity to estrogen receptor. We will present results of our binding energy calculations based on x-ray crystal structures for estrogen receptor ligand-binding domain complexed with raloxifene. In this study we found a high correlation between the calculated binding energy (kcal/mol) and the corresponding observed relative binding affinity. The CoMFA 3D contour maps, which reveal several strategies for designing novel high-affinity analogs, have been interpreted in terms of the recently published X-ray crystal structure of raloxifene bound to the estrogen ligand-binding domain.

161. COMPUTATIONAL STUDY OF CYCLOOXYGENASE INHIBITORS: 3D-QSAR/COMFA MODELS FOR COX-1 AND COX-2. *Ai Ni, Seong-Jae Yu, and William Welsh, UMDNJ, Piscataway, NJ 08854, aini@umdnj.edu*

Recent research indicate that side effects shown by the classical nonsteroidal antiinflammatory drugs (NSAIDs) is related to their nonselective inhibition of the two isoforms of cyclooxygenase, namely, COX-1 and COX-2. Several 3D-QSAR models were developed using Comparative Molecular Field Analysis (CoMFA) based on the experimental pIC₅₀ data on recombinant human COX-1 and COX-2 for a series of 108 1,5-diarylpyrazole. These 3D-QSAR models yielded excellent self-consistency ($r^2 > 0.90$) and predictive ability ($rCV2 > 0.50$) between the experimental and computed pIC₅₀ values. These statistically significant models can serve as guides for the rational design of highly selective inhibitors of COX-2. Docking of ligands, and calculation of ligand-enzyme binding energies, was performed to elucidate the structural prerequisites for binding affinity and COX-1/COX-2 selectivity. Values of the calculated binding energy correlated strongly with the corresponding experimental pIC₅₀ values.

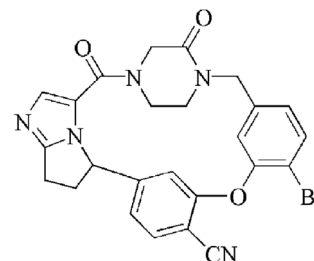
162. DOCKING OF NOVEL MELATONIN LIGANDS TO HOMOLOGY MODELS OF MT₁ AND MT₂. *Emilio Xavier Esposito¹, Troy Wymore², Pui-Kai Li³, and Jeffrey D. Madura¹. (1) Department of Chemistry & Biochemistry and Center for Computational Sciences, Duquesne University, 320 Mellon Hall, 600 Forbes Ave, Pittsburgh, PA 15282, exesposito@kent.net, (2) Pittsburgh Supercomputing Center, (3) Division of Medicinal Chemistry and Pharmacognosy, College of Pharmacy, Ohio State University*

The development of compounds that preferentially bind to specific melatonin receptors is critical in the elucidation of the specific function of the MT₁ and MT₂ receptors *in vivo*. Using homology models of the G-Protein Coupled

Receptors MT₁ and MT₂, docking simulations were performed using known and novel melatonin ligands. We currently have 30 ligands with known bioactivities for specific melatonin receptors. Several of these compounds were selected for the docking study, while others were used in the development of the docking protocol. We will elucidate key chemical characteristics for the binding of compounds to specific melatonin receptors. Conclusions and docking protocols will be discussed with attention given to the novel ligands designed with the aid of QSAR.

163. DOCKING STUDIES OF A CONFORMATIONALLY RIGIDIFIED ANALOG OF L-778,123, A POTENT FARNESYLTRANSFERASE INHIBITOR: INVESTIGATION OF THE EFFECT OF CHARGE MODELS ON DOCKING RESULTS. *Isidro Merino, and Yuan-Ping Pang, Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Foundation, 200 First Street SW, Rochester, MN 55905, Fax: 507-284-9111, merino.isidro@mayo.edu*

Using the RESP charges (Cieplak, et al., J. Comp. Chem., 16, 1357-1377, 1995) derived from quantum mechanics calculations and the CM2 charges (Li, et al., J. Phys. Chem. A, 102, 1820-1831, 1998) obtained from semiempirical calculations, we have performed docking studies of farnesyltransferase complexed with a conformationally rigidified analog of L-778,123 (Williams, T.M., 22nd ACS National Meeting, Chicago, IL, 2001) by employing the EUDOC program (Pang, et al., J. Comp. Chem., 22, 1750-1771, 2001). Both stereoisomers of the analog were used in the docking studies as the absolute stereochemistry of the analog was not known for certain. Herein, we report the computationally predicted stereochemistry of the analog bound in the active site of farnesyltransferase and the effects of charge models on docking results.



Conformationally rigidified analog of L-778,123

164. PFK5 A NOVEL KINASE TARGET FOR ANTIMALARIAL DRUGS: PROTEIN HOMOLOGY MODELING AND IN SILICO DESIGN OF POTENT AND SELECTIVE INHIBITORS. *Susan Keenan, and William Welsh, UMDNJ, Piscataway, NJ 08854, keenansm@umdnj.edu*

Increasing worldwide resistance of Plasmodium falciparum to traditional chemotherapy strategies such as chloroquine and mefloquine demonstrates the urgent need for the discovery of novel chemotherapeutic agents in the fight against malaria. The recent discovery of P. falciparum Protein Kinase 5 (PfPK5) invites the possibility of selectively targeting the life cycle of P. falciparum in order to prevent cerebral malaria. PfPK5 bears a high degree of sequence homology (>65%) to a structurally conserved family of mammalian kinases known as the cyclin-dependent kinases (CDKs). The CDKs are the key regulatory elements governing the ordered progression of the mammalian cell cycle. The x-ray crystal structures of CDK2, CDK6, and of the ligand-binding domain of CDK4 have been solved, thus providing structural templates upon which to base a three-dimensional model of PfPK5. Using computer-based homology modeling strategies we have developed a three-dimensional model of PfPK5, which has provided visual representation of the structural prerequisites for inhibitor binding. Structural information obtained from the homology model and from three-dimensional quantitative structure-activity relationship (3D-QSAR) models for both P. falciparum inhibition and CDK1 inhibition based on analysis of

published biological data on the inhibitory activity of a series of purine derivatives (ATP analogues) have guided the rational design of inhibitors for potential use as novel antimalarial agents. In addition, we have developed an in silico screening method to enable prediction of both the potency and selectivity of a drug candidate for inhibition of PIPK5 activity.

165.

PREDICTING BIOLOGICAL ACTIVITY OF A SET OF COMPOUNDS USING MACHINE LEARNING ALGORITHMS.

Kiyeon Nam¹, Christian Honer², Christoph Schumacher², Carol Berry², Paul Marshall², and Wendy Cornell². (1) Novartis Pharmaceuticals and UMDNJ, Summit, NJ 07901, Kiyeon.Nam@pharma.Novartis.com, (2) Novartis Pharmaceuticals

Determining the biological activity of a large number of compounds is a challenging step in the lead finding process, especially in the era of high throughput screening (HTS) when a number of active compounds may be identified. This step requires significant resources, since the assays are more sophisticated than those employed in HTS and analysis of the results involves subjective decision making which cannot easily be automated. Thus, it will be very advantageous if we can develop biological profiles for compounds, based on existing data from related targets, to prioritize compounds for testing. Here, we present a comparative study using machine learning algorithms to make a binary decision (agonist or antagonist) on nuclear receptor ligands with unknown biological activity. Classification models are developed based on the properties learned from known estrogen receptor α (ER- α) ligands. Then, the models are tested against compounds identified in a glucocorticoid receptor (GR) HTS.

166.

RATIONAL COMPUTER-AIDED DESIGN OF OPIOID ANALOGUES: 3D-QSAR MODELS FOR δ , μ , AND κ OPIOID RECEPTORS.

Seong-Jae Yu¹, Anil Nair², John Duchek³, and William Welsh¹. (1) UMDNJ, Piscataway, NJ 08854, yuse@umdnj.edu, (2) Aventis Pharmaceuticals, (3) Tyco Healthcare

Among the three classes of opioid receptors, designated delta (δ), kappa (κ), and mu (μ), recent evidence suggests that δ -selective opioids could be potentially useful as analgesics devoid of the numerous side effects (e.g., respiratory depression, physical dependence and gastrointestinal effects) associated with narcotics such as morphine. Moreover, selective antagonists of δ receptors have been shown to modulate the development of tolerance and dependence to μ agonists such as morphine, to modulate the behavioral effects of drugs of abuse such as cocaine, and to elicit favorable immunomodulatory effects. The δ -selective opioids thus represent extremely attractive candidates for a broad range of novel pharmaceutical applications including powerful yet safe analgesics, immunomodulatory agents for treating immune disorders, and new treatments for drug addiction. In this study, 3D-QSAR models were developed based on receptor binding affinities (K_i) data of Raynor et al. These 3D-QSAR models yielded strong correlations between the experimental and computed binding affinities (cross validated $r^2 > 0.5$) and also explain receptor selectivity of ligand. These models are useful to guide the rational design and rapid screening of the new opioid analogues for their selectivity toward the delta, μ and kappa receptors. Our basic design strategy is based on the "message-address" concept developed to rationalize the separate pharmacophoric features of opioids that confer affinity versus selectivity.

167.

RATIONAL DESIGN OF BETA-PEPTIDE INHIBITORS OF P53-MDM2 INTERACTION.

Haizhen Zhong, Leah M. Sandvoss, and Heather A Carlson, Department of Medicinal Chemistry, The University of Michigan, 428 Church Street, Ann Arbor, MI 48109, zhongh@umich.edu

The inhibition of the tumor suppressor protein p53 by mdm2 has been observed in a number of human cancers. In normal cells, p53 protects the cell from tumor formation in the mechanism of the cell cycle arrest or apoptosis. However this function is inhibited in some cancer cells by the amplification of the mdm2 oncoprotein. With the goal of restoring p53 function in cancers that contain overexpressed mdm2, we have designed a class of novel p53 mimetic. From the QM/MM calculations and MD simulations, we have obtained a series

of di-substituted beta-proline oligomers with high affinity. The details of these calculations, together with their biological implications as inhibitors of the p53-mdm2 protein-protein interaction, will be presented.

168.

A THEORETICAL INVESTIGATION ON THE BINDING OF HLA-A2-PEPTIDE COMPLEXES.

Tugba Gul Kucukkal¹, Emilio Xavier Esposito¹, Jeffrey R. Kovacs², Sai Prasanth Chamarthy², Wilson S. Meng², and Jeffrey D. Evanseck¹. (1) Department of Chemistry & Biochemistry and Center for Computational Sciences, Duquesne University, 600 Forbes Avenue, Pittsburgh, PA 15282, Fax: 412-396-5683, kucukka671@duq.edu, (2) Division of Pharmaceutical Sciences, Duquesne University

HER2/neu is a transmembrane glycoprotein that expresses in many tumors, including ovarian and breast cancers. The HER2/neu peptide IISAVVGIL (GP2) is recognized by tumor-specific cytotoxic T lymphocytes in the context of the human class I major histocompatibility complex (MHC) molecule HLA-A2. One limiting-factor for using this peptide as a tumor vaccine is its poor affinity for HLA-A2. The purpose of this computational study is to design better MHC binding GP2 analogs. The third and sixth residue positions are of interest due to the potential of increasing hydrogen bonding between the ligand and the protein. Initially, energy minimizations have been performed with Amber 94 without solvation effects, yet water molecules from the crystal structure were retained. With the best set of modified GP2-like structures identified, docking studies have been performed utilizing MOE-Dock to further refine the search for a "better" peptide.

169.

THE DOCKING AND BINDING AFFINITY CALCULATIONS OF TIMBO ANALOGUE INHIBITORS IN HIV-1 RT.

Zhigang Zhou, Chemistry and Biochemistry Department, Duquesne University, 600 Forbes Av, Mellon Hall 320, Pittsburgh, PA 15282, Fax: 412-396-5683, zhou7454@duq.edu, and Jeffrey D. Madura, Department of Chemistry and Biochemistry, Duquesne University

TIMBO analogues have been reported to be potent inhibitors of HIV-1 RT. Their binding structures in NNRTI active site are important to calculate their binding affinities to design new inhibitors. In the work, we use docking and our developed free energy calculation method to explore the binding pattern and affinity of more than 50 TIMBO analogues to elucidate their interaction and inhibition with HIV-1 RT. The free energy calculation method, which was developed and verified from known structures of HIV-1 RT/NNRTI complexes, includes solvation energy from Poisson Boltzmann method. Also we try to develop a QSAR model of these TIMBO inhibitors to help design more potent inhibitors.

170.

USE OF DOCK AND PB IN REPRODUCING THE BINDING MODE AND PREDICTING THE BINDING FREE ENERGIES OF NON-NUCLEOSIDE INHIBITORS TO HIV-1 REVERSE TRANSCRIPTASE.

Zhigang Zhou, Chemistry and Biochemistry Department, Duquesne University, 600 Forbes Av, Mellon Hall 320, Pittsburgh, PA 15282, Fax: 412-396-5683, zhou7454@duq.edu, and Jeffrey Madura, Department of Chemistry and Biochemistry, Duquesne University

Non-nucleoside Reverse Transcriptase inhibitors (NNRTI) are one of major group of inhibitors to treat AIDS as the enzyme key role in the reproduction of Human Immunodeficiency Virus type 1 (HIV-1). NNRTIs have been investigated and several protein/inhibitor complexes have been determined experimentally. Using generated docked complexes, we have applied Poisson-Boltzmann methods into calculation of the relative binding free energy for these inhibitors. The calculated results reproduce crystal binding positions and orientations (RMSD < 1Å). Based on an ensemble of binding structures, the binding free energies were calculated by combining the computed energies from Autodock, which includes vdW energy, electrostatic energy, solvation, and hydrophobic energy from buried area change, with the Poisson-Boltzmann solvation energy. This method gave the best correlation to the experimental activities of these inhibitors.

171.

UNUSUAL BINDING OF OLIGOSACCHARIDES IN THE ACTIVE SITE CLEFT OF FAMILY 18 CHITINASES. *Pranav Dalal*¹, *Nathan N. Aronson Jr.*², *Mikhail F. Alexyev*², *Lauren Amable*², *Brian A. Halloran*², *Patrick VanRoey*³, and *Jeffrey D. Madura*¹. (1) Center of Computational Sciences, Duquesne University, Department of Chemistry and Biochemistry, 600 Forbes Avenue, Pittsburgh, PA 15226, Fax: 412-396-5683, dalal@duq.edu, (2) Department of Biochemistry and Molecular Biology, Univ. of South Alabama, (3) Wadsworth Center, New York State Dept. of Health

Chitinases belong to a class of enzymes commonly known as glycosyl hydrolases which hydrolyze oligosaccharides and their derivatives. Chitinases are specific for the polysaccharide chitin which is a polymer of N-acetyl glucosamine residues linked by $\beta(1\rightarrow4)$ bonds. Experimental data is available on various chitinases and their subtle differences, however the mechanism of action for these enzymes is still not understood at the molecular level. We have carried out molecular dynamics simulations, fluorescence spectroscopy, and isothermal calorimetric experiments on these systems. The results from the computational and experimental work will be presented along with our proposed mechanism of binding and hydrolysis.

172.

NOVEL METHODOLOGIES IN TRAINING SET SELECTION FOR PHARMACOPHORE MODELING: A PHARMACOPHORE MODEL OF P38 MAP KINASE INHIBITORS. *Shikha Varma*, and *Remy Hoffmann*, 9685 Scranton Rd, Accelrys Inc, San Diego, CA 92121-3752, shikha@accelrys.com

Training set selection from a given SAR data is the first step in deriving a predictive QSAR model. The quality of the resultant model is highly dependent upon the compounds that are used to derive the model. A 3D pharmacophore (HypoGen) generation within Catalyst derives information from the training set in two ways, from the activity data and from the structures themselves. Therefore, while selecting a training set for developing a predictive pharmacophore model in Catalyst, it is critical to have diversity both in activity data as well as in the structure of the compounds, thus maximizing the information content of the training set. This process can often be a daunting task if there is a large amount of SAR data and if the compounds are structurally very similar. The objective of the work presented here is to develop an automated method for selecting a training set that can be used for Catalyst hypothesis generation from a large collection of compounds. Functionalities within Cerius and Catalyst are used in conjunction to develop an automated training set selection protocol. This protocol involves selection of the most diverse set of compounds on the basis of their multi-conformer 3D pharmacophoric fingerprints, multi-conformer shape indices, and activity. An application of this methodology is illustrated by selecting a training set and developing a pharmacophore model of p38 MAP Kinase inhibitors from SAR data collected for 130 pyridyl-imidazole containing compounds, and its ability to predict activities of compounds outside the training set. Docking the pharmacophore model into the active site of p38 MAP Kinase crystal structure further validates the methodology by examining the pharmacophoric features and receptor interactions.

173.

QSAR STUDY OF GENOTOXICITY AND CYTOTOXICITY OF A SERIES OF NAPHTHOQUINONES. *Jane L Roberts*¹, *John C Dearden*¹, and *Rodney F Bilton*². (1) School of Pharmacy and Chemistry, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, United Kingdom, Fax: +44-151-231-2170, j.c.dearden@livjm.ac.uk, (2) School of Biomolecular Sciences, Liverpool John Moores University

Naphthoquinones are widely used compounds, and some have pharmaceutical applications. Nonetheless, some of them are known to be genotoxic. We therefore carried out a number of in vitro toxicity tests on a series of 29 substituted naphthoquinones, and subjected the results to QSAR analysis. The tests were: comet tail test for genotoxicity, and the lactate dehydrogenase release test and the trypan blue exclusion test for cell membrane damage. The results showed that the compounds fell into two classes: those that were genotoxic and those that were cytotoxic. QSAR analysis allowed us to determine the molecular characteristics that determine each of the toxic effects.

174.

CONSIDERATION OF MOLECULAR WEIGHT DURING COMPOUND SELECTION IN VIRTUAL DATABASE SCREENING. *Yongping Pan*, and *Alexander D. Mackereil Jr.*, Pharmaceutical Sciences, University of Maryland, 20 N Pine Street, Baltimore, MD 21201, yongping@outerbanks.umaryland.edu

Compound selection in virtual database screening when targeting a biological macromolecule is typically based on the total interaction energy between the chemical compound and the biological macromolecule. However, this approach alone does not consider the contribution of the compound size or molecular weight to the energy score and, therefore, is biased towards the selection of large compounds. To account for molecular weight during energy based screening, we propose a normalization strategy based on the total number of heavy atoms in the chemical compound. This approach is computationally efficient and produces molecular weight distributions of selected compounds similar to that of the original databases used in the virtual screening. The proposed normalization procedure yields a similar set of selected compounds regardless of the cut-off of the number of heavy atoms being used in the database screening protocol, eliminating the artifact introduced by the arbitrary selection of the heavy atom cutoff.

175.

OEChem. *Matthew T. Stahl*, *A. Geoffrey Skillman*, and *Roger Sayle*, OpenEye Scientific Software, 3600 Cerrillos Rd, Suite 1107, Santa Fe, NM 87507

Representations of molecules and chemistry that are used inside computers are usually task specific and are frequently orthogonal to representations used by unrelated tools. Communicating chemical data between disparate applications can be challenging as representations of chemistry in software may not have the ability to store a specific type of data, or they may not be able to deduce chemical properties from the data that has been provided. A programming library called OEChem was designed to address chemical data deduction, representation, and interconversion issues common in chemical software. This paper will present the architecture and functionality of OEChem.

176.

PyOEChem. *Matthew T Stahl*, *Roger A. Sayle*, and *A. Geoffrey Skillman*, OpenEye Scientific Software, 3600 Cerrillos Rd, Suite 1107, Santa Fe, NM 87507

Interpreted languages such as Python facilitate application development by providing rich functionality in the core language and accessibility to a wide variety of modular extensions. Applications written in Python can be prototyped rapidly and modified easily. These features may Python an ideal development platform for chemical data processing and inter-application communication. A number of chemistry specific Python modules have been developed and used in the chemical software community. PyOEChem is the recently developed union of OEChem, a powerful programming library for chemistry, and Python. The architecture of PyOEChem and example applications will be presented in this paper.

177.

Ptree-BASED APPROACH TO MINING GENE EXPRESSION DATA. *fei pan Sr.*¹, *Xin Hu*², and *William William*¹. (1) Department of Computer Science, North Dakota State University, Fargo, ND 58102, fei.pan@ndsu.nodak.edu, (2) Department of Pharmaceutical Science, North Dakota State University

We propose a new "Data Mining Ready" data structure, called Peano count tree (P-tree) and apply for the NCI-60 cancer cell lines data set analysis. The data set (Nature Genetics. 2000;24(3):236-44) comprises 1376 gene expression profiles of 60 cancer cell lines and the growth inhibition of 1400 chemical compounds on the same cell lines, providing an excellent test case in the context of linking genomic data mining with high-throughput drug design. In this paper, we reconstructed the data set using P-tree first, and several Ptree-based clustering approaches have been used to analyze the gene expression profiles and drug activity pattern. Results with traditional clustering methods were also compared. The P-tree based data mining techniques hold a promise to gene expression data analysis.

178.

APPLICATION OF FUZZY METHODS TO THE ANALYSIS OF CONFORMATIONAL SPACES FOR ORGANOMETALLIC MOIETIES. *David White, Michael A. Freeze II, Aaron M. Gillespie, and Glenn R. Morello, Department of Chemistry, University of North Carolina at Wilmington, 601 S. College Rd, Wilmington, NC 28403, Fax: (910) 962-7107, whitedp@uncwil.edu, freezem@uncwil.edu*

Chemical reactions are typically dominated by a large number of conformationally distinct molecules in solution. Computational methods must deal with a more modest number of structures in order for the methods to complete expediently. To compare computed physical properties with experiment, all important conformers that participate in the experimental reaction must be identified. Molecular mechanics is typically used to generate the conformational space and the resulting plot of torsion angle(s) versus energy can be complicated, depending on the number of rotatable bonds and how they interact with one another. The task of identifying the clusters of conformers about a minimum is difficult. In this poster, the application of fuzzy methods to the cluster analysis of the conformational space of prototypical organometallic catalysts will be presented. The effect of coordination geometry on conformational with a variety of monodentate and bidentate ligands will be presented. Different fuzzy clustering algorithms will be discussed.

179.

THEORETICAL STUDY ON THE GAS-CHROMATOGRAPHIC ELUTION OF COMPOUNDS ON COLUMNS WITH DIFFERENT POLARITY. *Tamas Kortvelyesi, Department of Biomedical Engineering, Boston University, 44 Cummington Street, Boston, MA 02215, kortve@bu.edu, Karoly Heberger, Central Research Institute of Chemistry, Hungarian Academy of Sciences, Krisztina Nagy, Central Research Institute for Chemistry, Hungarian Academy of Sciences, and Miklos Gorgenyi, Department of Physical Chemistry, University of Szeged*

The separation of compounds with different polarities can be improved by the correct selection of the stationary phase (SP) of the gas-chromatographic (GC) columns. Different (relative) scales were developed for describing the polarity of the stationary phases on GC columns (e.g. McReynolds, dGE(CH₂), etc.). For the characterization of SPs (phthalates, adipates, sebacates, phosphates, citrates, nitrils, siloxanes), the electronic structures, hydrophobicity and geometries of these compounds were calculated at the level of PM3 semiempirical quantum chemical method and empirical formulas. The calculated chemical descriptors were statistically analyzed and correlated with the McReynolds constants and dGE(CH₂) to find the similarities and differences between the polarity scales of the SPs. On the basis of the statistical analysis the significant physical and chemical descriptors were chosen which are important in the solvent-solute interactions. The type of the significant interactions were also characterized at the columns and at the polarity scales, too.

180.

COMFA STUDY OF PROTONATED METHYLPHENIDATE PHENYL-SUBSTITUTED ANALOGS. *Kathleen M. Gilbert¹, Milind Misra², Ronald A. Buono², Carol A. Venanzi², Margaret M. Schweri³, Qing Shi⁴, and Howard M. Deutsch⁴.* (1) *Department of Chemistry and Environmental Science, New Jersey Institute of Technology, 323 King Boulevard, Newark, NJ 07102, Fax: 973-596-3596, kxg2248@njit.edu,* (2) *Chemistry Division, New Jersey Institute of Technology,* (3) *School of Medicine, Mercer University,* (4) *School of Chemistry and Biochemistry, Georgia Institute of Technology*

Protonated methylphenidate (pMP) and several phenyl-substituted pMP analogs were analyzed using Comparative Molecular Field Analysis (CoMFA) to develop a pharmacophore for dopamine transporter (DAT) binding. This research is part of an interdisciplinary study on using methylphenidate (MP) analogs to block the binding of cocaine to the DAT as a treatment for addiction. A random search conformational analysis using key pMP torsional angles was performed to create conformer families representing possible bioactive conformations. The lowest energy pMP conformer of each family was used as a template to create phenyl-substituted pMP analogs. Partial least squares analysis was used to determine the combination of electrostatic and steric cutoffs that yielded the highest predictability (q^{*2}). q^{*2} values above 0.5 were achieved for all conformer families. The best model was used to propose a pharmacophore to predict DAT binding affinity. The results were compared to a previous CoMFA study on neutral MP.

181.

2-DIMENSIONAL CONFORMATIONALLY SAMPLED PHARMACOPHORE: A NOVEL APPROACH TO LIGAND BASED PHARMACOPHORE DEVELOPMENT APPLIED TO δ OPIOID LIGANDS. *Denzil Bernard, Andrew Coop, and Alexander D. MacKerell Jr., Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, 20 North Pine Street, Baltimore, MD 21201, Fax: 410-706-0346, dbern002@umaryland.edu*

A pharmacophore is a model describing the spatial relationship between atoms and/or functional groups in a ligand that are important for its interaction with a receptor. To date ligand based pharmacophores have typically used low energy conformers of known receptor ligands. However, this method does not take into account the inherent dynamic nature of molecules, which sample a wide range of conformations any of which could be the bound form. In the present study molecular dynamics (MD) simulations were used as a means to sample the conformational space of ligands to include all accessible conformers at room temperature in the development of a pharmacophore. Based on these conformers, probability distributions of selected distances and angles in a series of δ specific opioid ligands were obtained and correlated with agonist versus antagonist activities. A 2-dimensional conformationally sampled pharmacophore for distinguishing δ opioid receptor agonists and antagonists was thus developed.

182.

NEW QUANTUM INDICES FOR MACROMOLECULAR CHARACTERIZATION. *Jana Khandogin, and Darrin York, Department of Chemistry, University of Minnesota, 207 Pleasant St. SE, Minneapolis, MN 55455-0431, Fax: 612-626-7541, khand0115@tc.umn.edu*

The present paper describes a linear-scaling semiempirical-based approach for the characterization of biological macromolecules. Quantum indices such as the electrostatic potential map, local softness, local hardness, proton potential and solvent response properties are presented that can be used to help predict the binding affinity and chemical reactivity of biomolecules. The method is applied to the Hiv-1 nucleocapsid protein complex and the hammerhead ribozyme. The linear-scaling approach allows the calculation of quantum molecular descriptors that have considerable promise in QSAR studies and rational drug design.

183.

NOVEL QSAR/QSPR METHODS FOR PREDICTING THE ACUTE TOXICITY OF ORGANIC COMPOUNDS. *Bo Jiang¹, Curt M Breneman¹, N Sukumar¹, and Mark J Embrechts².* (1) *Department of Chemistry, Rensselaer Polytechnic Institute, Cogswell 306, 110 8th St, Troy, NY 12180, Fax: 518-276-4045, jiangb@rpi.edu,* (2) *Department of Decision Sciences & Engineering Systems, Rensselaer Polytechnic Institute*

Measurement of the acute toxicity of organic compounds is a non-trivial task using traditional experimental methods, and has been subject to numerous computational treatments. We now wish to report a novel QSAR methodology for predicting acute fathead minnow toxicity of organic compounds using novel descriptors and modern machine learning methods. QSAR models for the acute aqueous toxicity of 375 diverse organic compounds from the COMPUTOX toxicity database were developed using sets of RECON/TAE and PEST descriptors developed in the Breneman group. The RECON algorithm provides a means for the rapid reconstruction of molecular charge density-based descriptors using atomic charge density fragments. The PEST (Property-Encoded Surface Translator) descriptor method was developed as a hybrid shape/property adaptation of the Zauhar Shape Signature method. Through the use of RECON and PEST molecular descriptors and modern machine learning regression techniques such as Artificial Neural Networks, Partial Least Squares, Kernel Partial Least Squares and SVMs as implemented in the STRIPMINER/Analyze software package, a set of predictive 3-D QSAR models for toxicity have been created and validated.

184.

NOVEL USE OF SELF-ORGANIZING FEATURE MAPS IN QSAR STUDIES. *Aaron J. Clauset, William S. Caldwell, and Jeffrey D. Schmitt, Molecular Design Group, Targacept Inc, 200 East First Street, Winston-Salem, NC 27101-4165, aaron.clauset@targacept.com*

We report the development of a novel QSAR technique that utilizes self-organizing feature maps (SOMs) for dataset modeling and visualization. In

contrast to linear techniques, which fail to model non-linear relationships within descriptor and target-variable space, SOMs are non-linear and preserve topological mappings. The output of SOM methodology bears some similarity to clustering algorithms. We show that when paired with objective variable selection strategies (stepwise regression and steepest-ascent hillclimbing), the SOM QSAR outperforms linear QSARs in both cross-validated (leave-group-out) and randomization tests for real-world QSAR datasets.

185. PHARMACOPHORE IDENTIFICATION AND BIOACTIVITY PREDICTION BASED ON THE TAE ELECTRONIC CONFORMATION METHOD. *Minghu Song, Curt M Breneman, Dechuan Zhuang, and N. Sukumar, Department of Chemistry, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180, Fax: 518 276-4887, songm@rpi.edu*

The pharmacophore concept is widely used in the field of computer-aided drug design, even though it is an abstract and qualitative concept that accounts for common sets of features shared by a series of active molecules. It has been observed that the presence of a specific pharmacophore does not assure favorable binding affinity. For instance, affinity can be modulated by interactions between portions of the binding site with ligand features outside of the pharmacophore.

The electron-conformation (EC) method is one of methods to quantitatively identify of pharmacophore (Pha). Bersuker has defined an electron-conformational matrix of congruity (ECMC) which encodes the topological, geometrical and electronic features of molecules. The Pha can be identified by a small sub-matrix of ECMC, which is common for active compounds (or within certain tolerances) and absence in inactive compounds.

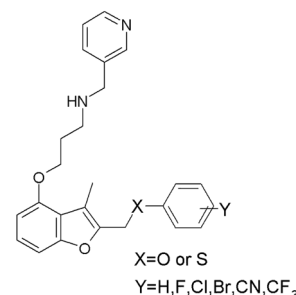
In this study, the TAE Electronic-density-derived atomic properties are utilized to quantitatively represent the atomic nature of electron donor-acceptor strength. The above atomic electronic indices as well as the intramolecular atomic distances are employed to construct an ECMC matrix for a set of Angiotensin Converting Enzyme (ACE) inhibitors. Once the Pha is identified by the usual EC procedure, it can be used as a screening criterion for compounds in a large database. The next step of this project will take into account the modulating influence of neighboring groups by constructing a parameterized activity modulation function. This function will be validated using compounds "held out" from the training set.

186. PREDICTION OF GLASS TRANSITION TEMPERATURES FOR POLYMERS USING THE TAE/RECON METHOD. *Qiong Luo¹, N Sukumar¹, Curt M. Breneman¹, Kristin Bennett², and Mark J. Embrechts³.* (1) Department of Chemistry, Rensselaer Polytechnic Institute, Cogswell 306, 110 8th St, Troy, NY 12180, Fax: 518 276-4045, luoq@rpi.edu, (2) Department of Mathematical Sciences, Rensselaer Polytechnic Institute, (3) Decision Sciences and Engineering Systems, Rensselaer Polytechnic Institute

The Transferable Atom Equivalent (TAE) method has been fruitfully employed in QSAR and ADME property prediction. Recent developments of this method for the prediction of glass transition temperature (Tg) of polymers will be discussed. The repeat unit structures of end-capped monomers were used to construct the appropriate atom type environment for the quantitative structure-property relationships (QSPR) investigation. The RECON algorithm determines atom types and environments from the SMILES strings for the repeat units, assigns the closest match to each atom from a library of atom types, and combines the densities of the atomic fragments to compute a large set of QSPR descriptors. The library of atomic charge density fragments employed allows for rapid retrieval of atomic/residue descriptors and molecular assembly. QSPR descriptors for hundreds of polymers can be computed within a few minutes. Several machine learning methods are used to generate the models after descriptor generation. Optimization routines (genetic algorithm) are utilized to find the information-rich subsets of descriptors for prediction. Partial Least Squares Regression (PLS), Support Vector Machine Regression (SVM) and Artificial Neural Networks (ANN) are utilized to build predictive models.

187. QSAR ANALYSIS OF FUNGAL N-MYRISTOYLTRANSFERASE INHIBITORS: CORRELATION BETWEEN ENZYME INHIBITORY ACTIVITY AND INTERACTION ENERGIES CALCULATED BY AB-INITIO MOLECULAR ORBITAL METHODS. *Kenji Morikami¹, Kiyoshi Hasegawa¹, Satoshi Sogabe¹, Yasuhiko Shiratori¹, Hirotsato Ebiike¹, Ken-ichi Kawasaki¹, Miyako Masubuchi¹, Tatsuo Ohtsuka¹, Kiyooki Sakata², Toshihiko Fujii², Yuko Aoki², and Nobuo Shimma¹.* (1) Department of Chemistry, Nippon Roche Research Center, 200 Kajiwara, Kamakura 247-8530, Japan, Fax: +81-467-45-6824, kenji.morikami@roche.com, (2) Department of Mycology, Nippon Roche Research Center

We have discovered a series of novel benzofuran derivatives as a new antifungal agent targeting fungal N-myristoyltransferase. The enzyme inhibitory activity of the inhibitors showed clear parabolic correlation with electrostatic potential derived atomic charges of X. This correlation was closely related to the hydrogen bonding abilities of the inhibitors, which were controlled by substituents (Y) on the aromatic rings adjacent to X. It was also suggested that the electrostatic interactions between fluorine atoms of the inhibitors and hydrogen atoms of the enzyme played an important role in exhibiting the potent enzyme inhibitory activity. These findings were consistent with the experimental observations by X-ray crystallography.



188. TACS (TARGACEPT ACTIVE CONFORMATION SEARCH): A NEW METHOD FOR PREDICTING THE CONFORMATION OF A LIGAND BOUND TO ITS PROTEIN TARGET. *Josef Klucik, Aaron J. Clauset, Phillip S. Hammond, Rebecca Harris, William S. Caldwell, and Jeffrey D. Schmitt, Molecular Design Group, Targacept Inc, 200 East First Street, Winston-Salem, NC 27101-4165, josef.klucik@targacept.com*

Using a combination of exhaustive conformational analysis, calculated molecular properties, and a variety of regression techniques TACS successfully determines the binding or "active" conformation of protein bound ligands. Unlike other active conformation or pharmacophore search methods, the TACS hierarchical pruning method does not require manual molecular alignments nor does it require subjective judgment on the part of investigator.

189. VIRTUAL HIGH THROUGHPUT SCREENING OF ION RECEPTORS. *Timothy K. Firman, and Benjamin P. Hay, W. R. Wiley Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, PO BOX 999, Richland, WA 99352, Fax: 509-375-6631, Timothy.Firman@pnl.gov*

The ultimate goal of our research program is the effective automated design of ion receptors. This poster describes the development of molecular structure generating and scoring algorithms that describe a first step toward achieving this goal. These algorithms have been implemented in a program, HostDesigner, that runs on personal computers. Applications are presented that demonstrate HostDesigner is capable of rapidly building and evaluating millions of candidate structures, and generates reasonable molecular structures with complementary binding sites for targeted ion guests.

190.

APPLICATION OF THEORETICAL LINEAR SOLVATION ENERGY RELATIONSHIPS TO HUMAN INTESTINAL ABSORPTION. *D.R. Rakijian¹, Tri Duong¹, George R. Famini², Leland Wilson¹, and Marvin Payne¹.* (1) Department of Chemistry and Biochemistry, La Sierra University, 4700 Pierce St., Riverside, CA 92515, Fax: 909-785-2144, richard767@yahoo.com, bbboy38@yahoo.com, mpayne@lasierra.edu, (2) International Programs Office, U.S. Army Soldier Biological and Chemical Command

The theoretical linear solvation energy relationship (TLSE) extends linear free energy (LFER) methods by using theoretical descriptors to correlate various phenomena, rather than empirical descriptors, obviating the need for measured chemical data. The TLSE descriptors are derived from structurally-based molecular orbital calculations and represent chemically intuitive parameters. We studied the correlation of TLSE descriptors with human intestinal absorption (HIA) based on an extensive compilation of absorption data previously analyzed with the Abraham descriptors. Considering the complexity of the system we find a good overall correlation between HIA and the TLSE parameters (R^2 near 0.74). The significant descriptors represent molecular volume, polarizability index, and charge variance terms. The largest contribution is associated with the polarizability index, indicating that a compound that is easily polarized will have high transport efficiency through the intestinal mucosa.

191.

"GLATENT MEMBRANE PERMEABILITY" CONCEPT: QSAR ANALYSIS OF INTER- AND INTRA-LABORATORICALLY VARIABLE CACO-2 CELL PERMEABILITY. *Fumiyoshi Yamashita, Shin-ichi Fujiwara, and Mitsuru Hashida,* Department of Drug Delivery Research, Kyoto University, 46-29 Yoshidashimoadachi-cho, Kyoto 606-8501, Japan, Fax: 81-75-753-4575, yama@pharm.kyoto-u.ac.jp

Caco-2 cell monolayers grown on a filter support are the most widely used systems for predicting intestinal absorption. However, inter- and intra-laboratory variability in Caco-2 permeability makes it difficult to analyze quantitative structure/permeability relationships for a large number of compounds. We proposed the "latent membrane permeability" concept, assuming that all Caco-2 permeability datasets share a hidden, common relationship between their membrane permeability and physicochemical properties. This means that the coefficient vectors in the regression equations derived from different data sets are parallel to one another. An iterative calculation method was developed to deal with this conceptual approach and applied to the analysis of Caco-2 permeability data sets from different sources. Residual sum of squares given by regression of individual data sets was not statistically different from that for the "latent membrane permeability" concept. Thus, an essential structure/permeability relationship could be extracted from plural data sets.

192.

DIFFUSIVE STEP METHOD FOR DETERMINING THE FULLY ANISOTROPIC ROTATIONAL DIFFUSION TENSORS OF SPIN LABELS TETHERED TO MACROMOLECULES. *David E. Budil¹, Maria Micalizzi¹, Lan Jiang¹, Eva Darien², and Peter M. Gannett³.* (1) Department of Chemistry, Northeastern University, 360 Huntington Avenue, Boston, MA 02115, dbudil@neu.edu, (2) Basic Pharmaceutical Sciences, West Virginia University, (3) Department of Basic Pharmaceutical Science, West Virginia University

High field electron paramagnetic resonance is the only experimental method that can measure the fully anisotropic rotational diffusion of a molecule, that is, resolve different rates of rotation around molecular X, Y, and Z axes. This is particularly important for the dynamics of spin-labels, which are tethered to a macromolecule and therefore exhibit full rotational anisotropy and tilting of the principal diffusion axes away from the nitroxide symmetry axes. Existing molecular dynamics methods for measuring the rotational diffusion tensor based on time autocorrelation function (ACF) analysis have not been applied to tethered probes, and such analysis also has very limited ability to resolve the full diffusion tensor. We present a diffusive step method for fully determining the rotational diffusion tensor and diffusion tilt of a reorienting spin label using relatively short trajectories. The approach gives reasonable agreement with ACF methods for freely diffusing nitroxides. For nitroxides tethered to a DNA molecule, labels with a rigid acetylenic link exhibit significantly different

rotational diffusion properties from those with a flexibly attached label. The implications for EPR analysis of spin-labeled biomolecules are briefly discussed.

193.

QUANTUM CHEMICAL CALCULATIONS OF CADMIUM CHEMICAL SHIELDING TENSORS IN BIOLOGICALLY RELEVANT MOLECULES. *Srikanth S. Kidambi,* Department of Chemistry, University of Michigan, 930 N. University Ave., Ann Arbor, MI 48109, skidambi@mspacman.gccc.itd.umich.edu, and A. Ramamoorthy, Biophysics Research Division & Department of Chemistry, University of Michigan

The chemical shift interaction contains valuable information about the local environment of a nucleus and therefore is useful to understand the chemical bonding, conformation, and dynamics of molecules. For example, chemical shifts of metals in inorganic and biological complexes can provide insights into the nature of coordinating ligands, coordination number, and the coordination geometry. The direct detection of the most prevalent metals, such as zinc and calcium, using NMR experiments to determine their chemical shifts is rather difficult. Therefore, it becomes essential to use ^{113}Cd , which has spin $I=1/2$, as a surrogate probe for zinc, calcium containing bio-complexes. Unlike ^1H , ^{13}C and ^{15}N , ^{113}Cd spans a chemical shift range of 900 ppm, which makes it a valuable tool to distinguish different metal coordination geometry and ligand type. On the other hand, there were instances when isotropic chemical shift was insufficient to understand the geometry and the effect of ligands. This can be overcome by determining the magnitude and orientations of the individual components of the chemical shift anisotropy. To further understand the chemical principles underlying the variation of the ^{113}Cd CSA tensor, solid-state NMR experiments can be supplemented with ab initio calculations.

In our earlier study, we determined the power of DFT calculations and bigger basis set in identifying the CSA tensors of cadmium nuclei. As an extension to that work, we looked at several biologically relevant molecules such as Cd-porphyrins and Cd-crownether complexes. These molecules form discrete units, which mimic many of the metalloproteins. Using our established method, we calculated the ^{113}Cd CSA tensors in these molecules and compared with the experimental values. Our results suggest that DFT calculations estimate chemical shielding tensors to a great accuracy in discrete moieties and hence will be a valuable tool in elucidating the structure around metal centers.

194.

MOLECULAR DYNAMICS STUDIES ON POLYMER GOLD CRYSTALLIZATION FROM AN ORIENTED AMORPHOUS STATE. *Akira Koyama,* Graduate School of Human and Environmental Studies, Kyoto University, Yoshida Nihonmatsu Chou, Kyoto 606-8501, Japan, Fax: +81-75-753-6804, koyama@phys.h.kyoto-u.ac.jp, Takashi Yamamoto, Department of Physics, Yamaguchi University, Koji Fukao, Polymer Science and Engineering, Kyoto Institute of Technology, and Yoshihisa Miyamoto, Faculty of Integrated Human Studies, Kyoto University

The molecular process of crystallization from an oriented amorphous state was reproduced by molecular dynamics simulation for a realistic poly(ethylene) model. The initial oriented amorphous state was obtained by uniaxially drawing an isotropic glassy state at 100 K. By the temperature jump from 100 K to 330 K, there occurred crystallization into the fiber structure, during the process of which we observed the developments of various order parameters. The real space image and its Fourier transform revealed that a hexagonally ordered domain was initially formed, and then a highly ordered crystalline state with stacked lamellae developed after further adjustment of the relative heights of the chains along their axes.

195.

MULTICANONICAL MONTE CARLO STUDY OF LIQUID-SOLID PHASE TRANSITION OF ARGON. *Chizuru Muguruma,* Faculty of Liberal Arts, Chukyo University, 101 Tokodachi, Kaizu-cho, Toyota 470-0393, Japan, Fax: +81-565-46-1298, Yuko Okamoto, Department of Theoretical Studies, Institute for Molecular Science, and Masuhiro Mikami, Research Institute for Computational Sciences, National Institute of Advanced Industrial Science

To investigate the liquid-solid phase transition of a bulk system, the multicannonical Monte Carlo Method was applied to a 108 argon system. Though we can observe the border between liquid and solid states around 1.35 kcal/mol during the long projection run, the flat probability distribution was obtained from the simulation. The expectation values calculated by reweighting method are

compared with the averaged values obtained by the conventional canonical Monte Carlo calculation at each temperature. The conventional canonical Monte Carlo Method shows the hysteresis loops for averaged energies, averaged pressures and radial distribution functions at 100-200 K, whereas The multicanonical Monte Carlo calculation does not show the hysteresis. Assuming the multicanonical weight is enough reliable, the melting/freezing point of this system is around 150K from the free-energy curve.

196.
LARGE-SCALE SIMULATIONS OF CLAY-POLYMER NANOCOMPOSITES USING GRAND-CANONICAL MONTE CARLO AND MOLECULAR DYNAMICS APPROACHES. *Pascal Boulet, S. Stackhouse, and P.V. Coveney, Centre for Computational Science, Department of Chemistry, Queen Mary, University of London, Mile End Rd., London, United Kingdom, P.Boulet@qmul.ac.uk*

Development of novel clay-polymer nanocomposites has become a very important field of research for several decades. These new categories of materials have proved to exhibit enhanced properties (such as electrical and mechanical properties) compared with the corresponding bare polymer. Appropriate description of these materials by atomistic simulations (Monte Carlo and molecular dynamics) requires the utilisation of an adequate force field. Furthermore, realistic simulations necessitate the use of large models and consequently of large-scale atomistic calculations. For this purpose we have used the Teppen's force field which has been especially developed to properly describe clays. We have interfaced this force field to the LAMMPS program package (Large-scale Atomic/Molecular Massively Parallel Simulator). The validation of this scheme has been done by comparing the experimental and theoretical swelling curves of the sodium-montmorillonite clay. As to the clay-polymer nanocomposites materials, several monomers have been chosen to be intercalated into the Na-montmorillonite clay galleries. Intercalations have been modelled performing grand-canonical Monte Carlo simulations. The corresponding force field is a modified version of the dreiding force field. Resulting structures served as starting points for molecular dynamics simulations using the Teppen's force field in LAMMPS. d-spacing parameters have been compared to experimental data and are presented in this poster.

197.
AN EFFICIENT METHOD FOR THE DIRECT CALCULATION OF THE EXCESS ENTROPY USING THE AVERAGE OF ACCEPTANCE RATIOS. *Seung Do Hong, and Du-Jeon Jang, School of Chemistry, Seoul National University, Seoul 151-742, South Korea, Fax: +82-2-889-1568, sdhong@snu.ac.kr, djjang@plaza.snu.ac.kr*

We present an efficient method for the direct calculation of the excess entropy from the canonical ensemble average of acceptance ratios during a single Monte Carlo or molecular dynamics simulation. The acceptance ratios are averaged over virtual random configurations generated by separate parallel Monte Carlo procedures. No reference system is needed except the ideal gas and so the absolute excess entropy is obtained. We obtained very stable excess entropies for model systems including the two-center-Lennard-Jones system. This method is very efficient and easy to implement as well.

198.
A NEW HYDROGEN BONDING PARADIGM: DIFFERING ELECTRICAL INTERACTIONS IN C-H...O AND O-H...O HYDROGEN BONDS OF FORMIC ACID. *Weili Qian, and Samuel Krimm, Biophysics Research Division and Department of Physics, University of Michigan, 930 N. University, Ann Arbor, MI 48109, wlqian@umich.edu*

It is now well understood that the blue-shifted CH stretch frequency of many CH...O hydrogen bonds, as compared to "normal" red-shifted OH or NH bands, results from the shortened CH bond length. The explanation of the origin of this different behavior, however, has been controversial. We have studied two examples that shed light on the role of the common electrical interactions in explaining the results: (1) an ab initio study of a formic acid monomer in a

constant Onsager reaction field and (2) a QM/MM study of the seven equilibrium formic acid dimer structures. In the former case, the different bond length changes result from the parallelism or anti-parallelism of the electric field and the bond dipole derivative. In the latter case, involving a non-constant electric field plus exchange repulsion, similar but more general electrical interactions are involved. These results provide a new perspective on hydrogen bonding.

199.
THEORETICAL APPLICATIONS OF CHEMICAL AND PHYSICAL CO₂ SEQUESTRATION. *Thomas J. Dick, Department of Chemistry and Biochemistry, Duquesne University, 308 Mellon Hall, 600 Forbes Ave, Pittsburgh, PA 15282, Fax: 412-396-5683, dick251@duq.edu, Pranav Dalal, Center of Computational Sciences, Duquesne University, and Jeffry D. Madura, Department of Chemistry & Biochemistry and Center for Computational Sciences, Duquesne University*

Many CO₂ sequestration methods have been proposed. Currently, we are investigating two methods; brine sequestration and coal sequestration. Mineral trapping in brine aquifers is the most probable method for CO₂ sequestration. If the model brine system is to be accurate, the solubility of CO₂ needs to be calculated correctly. We have undertaken potential mean force calculations to determine the solubility of CO₂ in bulk water and brine solutions. For the method of CO₂ sequestration by coal, carbon nanotubes are used to model the pores in coal. Molecular dynamics simulations have been used to investigate the flooding effect of water and CO₂ in carbon nanotubes. A comparison of the two CO₂ sequestration methods will be presented.

200.
ELUCIDATING THE FORM OF MgO-CH₄ AND CH₄-CH₄ INTERACTIONS IN THE ADSORBED PHASE. *P.J. Stimac, R.J. Hinde, and J.Z. Larese, Department of Chemistry, University of Tennessee, 336 Buehler Hall, Knoxville, TN 37996-1600, Fax: 974-3454, pstimac@utk.edu*

Larese et al. have obtained high resolution inelastic neutron scattering spectra for one to six layers of methane adsorbed on the (100) face of magnesium oxide at a temperature of 1.5 K. These studies probe the rotational transitions of the adsorbed methane molecules. For monolayer coverage the spectra do not exhibit transitions corresponding to free rotation of methane molecules on the MgO surface. However, as the methane coverage increases the spectra change significantly, a transition corresponding to free methane rotation emerges and the spectra begin to resemble that of bulk methane. The evolution in the structure of these spectra as a function of surface coverage provides significant insight into the nature of the CH₄-CH₄ and MgO-CH₄ interactions. In this work we focus on modeling the CH₄-CH₄ and MgO-CH₄ interactions in an attempt to reproduce the rotational spectra by solving the Schrodinger rigid rotor equation for a spherical top.

201.
QUANTUM CHEMICAL STUDIES OF THE SOLVENT DECOMPOSITION IN LITHIUM-ION BATTERY ELECTROLYTE. *Young-Kyu Han, Sang Uck Lee, and Jaehoon Jung, Analytic R&D center, LG Chem, Ltd. Research Park, 104-1, Moonji-dong, Yuseong-gu, Taejon 305-380, South Korea, Fax: 82-42-863-7466, ykhan@lgchem.co.kr*

Electrochemical reactions involving solvent (or additive) decompositions at the electrode-electrolyte interface play a crucial role that affects capacity, cycle life, and safety in lithium-ion rechargeable battery. It has been proposed that the initial reaction is an electron transfer from the electrode to the solvent (or additive) molecules that are coordinated with lithium ions. DFT and ab initio MP2 calculations were used to investigate the structure and stability of M-Li_n (n=0, 1, and 2) complexes, where the M are the solvents (EC, PC), and the additives (VC, ES, and GS). Although the molecules are geometrically similar, it is found that the reactions with lithium atoms may provide various reaction products depending upon the structures and stabilities. The transition states that have been investigated for the reductive bond-cleavage reactions are also presented and discussed. For propane sultone (PS), we have tried to elucidate the role as a solvent additive in organic polar solutions for lithium-ion batteries, with the cluster models (EC)_nLi⁺(PS) (n=0-3) and the polarized continuum model to account for bulk solvent effect.

202. INTERPRETING BROWN DWARF ATMOSPHERIC ABSORPTION SPECTRA: THE NA — HE DIMER. *T.C. Lillestolen, and R.J. Hinde, Department of Chemistry, University of Tennessee, Knoxville, TN 37996, tlillest@utk.edu*

Since the first confirmed discovery of a brown dwarf in 1995 improved observational techniques have allowed astronomers to record the absorption spectra of brown dwarf atmospheres. Evident in these spectra is the Na D line which is broadened by interactions with other constituents of the atmosphere. In this work we present the ground state and low-lying excited state potential energy surfaces of the Na — He dimer, along with transition dipole moments for electronic excitation of the dimer. This information allows us to investigate the broadening of the Na D line due to molecular collisions in the brown dwarf atmosphere.

203. MONTE CARLO SIMULATIONS FOR ISOTOPIC SUBSTITUTIONS IN NEUTRON STUDIES. *James Wesley Mancillas, Department of Chemistry, Univ. Tenn, Buehler Hall, Knoxville, TN 37996, jmancill@utk.edu*

The determination of partial structure factors using neutron diffraction with isotopic substitution (NDIS) relies on the subtraction of experimental data sets for two samples whose only difference is the isotopic identity at a known atomic site in the sample. Implicit in this method is the underlying assumption that isotopic substitution leaves the atomic-level structure of the sample unchanged. For H/D substitution, which is one of the most common NDIS techniques, this assumption is rendered suspect by the fact that large differences in quantum mechanical zero-point vibrational amplitudes necessarily accompany isotopic substitution.

We investigate these quantum deviations from isostructural behavior by conducting Metropolis Monte Carlo simulations of isolated H₂ and D₂ molecules dissolved in fluid Ar. We treat all degrees of freedom but the solute vibrational mode classically; the solute vibration is modeled as the appropriate molecular $v=0$ quantum state. From these simulations we obtain radial pair distribution functions $G(r)$ centered at the H or D site. Our results help clarify the role of quantum effects in H/D NDIS experiments and contribute to a deeper fundamental understanding of quantum effects in condensed phases.

204. SEMIEMPIRICAL CALCULATION OF CRYSTAL STRUCTURES. *James J. P. Stewart, Stewart Computational Chemistry, 15210 Paddington Circle, Colorado Springs, CO 80921-2512, jstewart@fujitsu.com*

By use of Born Von-Kármán periodic boundary conditions, semiempirical methods can be applied to crystals, including minerals. A brief background to the theory involved will be given. Application to different types of crystal will be presented, along with discussion of the thermochemistry and structural features involved.

205. SEMIEMPIRICAL QUANTUM CHEMISTRY: WHERE WE ARE AND WHERE WE ARE GOING. *Kenneth M. Merz Jr., Department of Chemistry, The Pennsylvania State University, 152 Davey Laboratory, University Park, PA 16802, Fax: 814-863-8403, merz@psu.edu*

Semiempirical SCF-MO theories based on the NDDO approximation have had tremendous impact on our understanding of chemical and biological systems. Thus, it is worthwhile to reflect on the current strengths and weaknesses of available NDDO based semiempirical quantum chemistry methodologies with an eye towards how we can improve the current state-of-the-art. In this presentation, we will briefly review the development and status of semiempirical SCF-MO methods and we will discuss ways in which the current approaches can be improved. In particular, we will discuss the development of semiempirical NDDO-DFT methodologies, which are formalistically better aligned with modern ab initio and DFT methodologies. Finally, we will discuss the integration of our NDDO-DFT approach with traditional ab initio and DFT methods to form a combined QM/QM approach.

206. SEMI-EMPIRICAL PARAMETERS FOR TRANSITION METALS. *Warren J. Hehre, and Jianguo Yu, Wavefunction, Inc, 18401 Von Karman Avenue, Suite 370, Irvine, CA 92612, Fax: 949-955-2118, hehre@wavefun.com*

The PM3 semi-empirical method has been parameterized for 21 transition metals, based solely on reproducing experimental equilibrium geometries for inorganic and organometallic compounds. The resulting method employs an $(n-1)d, ns, np$ valence basis set (n is 4, 5 and 6 for first, second and third row metals, respectively), and is intended to be used in conjunction with previously published PM3 parameters for main-group elements and for Zn, Cd and Hg. It is fully implemented in the Spartan series of programs.

Comparison of PM3 equilibrium and transition-state geometries, equilibrium conformations and reaction and activation energies with experimental data where available, and/or with the results of density functional calculations is provided. An attempt is made to establish a suitable role for semi-empirical methods in descriptions of transition metal inorganic and organometallic compounds.

207. SEMI-EMPIRICAL MODELING OF TRANSITION METAL COMPLEXES. *Corneliu Buda, Department of Chemistry, University of Memphis, J. M. Smith Chemistry Building, 3744 Walker Avenue, Memphis, TN 38152, Fax: 901-678-3447, cbuda@memphis.edu, and Thomas Cundari, Department of Chemistry, CROMIUM, University of Memphis*

Technetium chemistry is important in radiopharmaceuticals and nuclear waste remediation. To achieve the aim of computer-aided design of metal complexes the most important step is to predict an appropriate geometry. The large diversity of ligands, -1 to 7 formal oxidation states and coordination numbers from 3 to 7 for Tc makes it very complicated to choose an appropriate computational method for these complexes. Semiempirical quantum mechanics (SEQM) provides a compromise between accuracy and computational demands for modeling metal complexes. Comparisons between PM3(tm) standard parameters and PM3(tm)-GA (genetic algorithm-optimized parameters for Tc) were made for 197 Tc complexes. Another challenge is to consistently predict the correct spin state for metal complexes. Comparisons between SEQM and ab initio methods were made for 50 Tc complexes whose structures are known. SEQM methods ZINDO/1 and PM3(tm) yield consistent spin predictions as compared with much more expensive density functional and Hartree-Fock methods.

208. PROGRESS TOWARDS A "NEXT GENERATION" NDDO-BASED SEMIEMPIRICAL TECHNIQUE. *Tim Clark, Paul Winget, Cenk Selcuki, Anselm Horn, and Bodo Martin, Friedrich-Alexander-Universität Erlangen-Nürnberg, Computer-Chemie-Centrum, Nögelsbachstrasse 25, D-91052 Erlangen, Germany, Fax: +49-9131-8526565, clark@chemie.uni-erlangen.de*

We have examined the systematic weaknesses and sources of error of current semiempirical techniques and have identified areas in which considerable improvements can be made. In order to provide a systematic, high quality set of parameterization data, we have set up a database that will be freely available online and that contains validated (using G2, G3 or DFT calculations) results for a large set of parameterization data. Results obtained using the prototype of the new technique will be presented.

209. INSIGHT INTO THE DOCKING OF METALLOPROTEINASES. *Xin Hu, and William H. Shelper, Department of Pharmaceutical Sciences, North Dakota State University, Sudro Hall #11, College of Pharmacy, Fargo, ND 58105, Fax: 701-231-8333, xin.hu@ndsu.nodak.edu*

Docking of metalloproteinase remains a challenge due to the polyhedral coordination and lack of well parameterization associated with the metal ion. We investigated the roles of zinc metal in the docking process by docking 20 zinc-dependent matrix metalloproteinases ligands to their crystal structures. Several docking/scoring approaches (DOCK, FlexX, GOLD, and AutoDock) have been used to assess the accuracy and reliability. While most of these docking programs have been found to have a zinc binding problem, correlation between the docking accuracy at the zinc binding site and the reliability of predicted energy suggests one way to improve the performance by dealing with the

zinc-ligand interactions properly. Several factors related to the metal docking were discussed such as the charges model, zinc parameters, tautomerism and protonation state of ligand, and coordination geometry consideration.

210.

PROTONATION-INDUCED STEREOISOMERISM IN TERTIARY AMINES: APPLICATION OF QM-BASED MOLECULAR DYNAMICS METHODS. *Philip S. Hammond¹, Roberto Car², Rebecca Harris¹, William S. Caldwell³, and Jeffrey D. Schmitt¹.* (1) *Molecular Design, Targacept, Inc, 200 East First Street, Suite 300, Winston-Salem, NC 27101-4165, Fax: 336-480-2107, phil.hammond@targacept.com,* (2) *Department of Chemistry, Princeton University,* (3) *Vice President, Drug Discovery, Targacept, Inc*

A variety of biologically active small molecules contain tertiary amines that become chiral centers when protonated. Using a combination of molecular mechanics and conformational analysis, we have demonstrated that this phenomenon has an appreciable effect on a molecule's energy hypersurface, leading to a differentiation in molecular shape, divergent sampling of energetically accessible conformational minima and may have a significant impact on molecular recognition. Proton-induced stereoisomerism may influence binding in a variety of stereoselective receptors. Numerous pharmacologically important molecules exhibiting protonation-induced stereoisomerism are studied using Car-Parinello ab initio molecular dynamics. Standard measures of topological and conformational similarity are used to quantitate this effect.

211.

TAUTOMERISM OF NUCLEOBASE DERIVATIVES AND THEIR SCORE IN VIRTUAL SCREENING TO THYMIDINE KINASE. *Pavel Pospisil, Patrick Ballmer, Gerd Folkers, and Leonardo Scapozza, Department of Applied Biosciences, Pharmaceutical Chemistry, Swiss Federal Institute of Technology, Winterthurerstr. 190, 8057 Zurich, Switzerland, Fax: 0041-1-6356884, pospisil@pharma.ethz.ch*

The therapeutically important Herpes simplex virus type 1 thymidine kinase (HSV1 TK) is able to accommodate several pyrimidine and purine bases in its active site. Tautomers of 30 known nucleobase analogs were docked into the active site using the programs AutoDock and FlexX. Additionally, different rotameric states of amino acid side chains crucial for ligand binding were considered. The docking results were compared with the available structural information and binding affinity data. The results showed the impact of tautomerism in predicting ligand binding orientation. Furthermore, the tautomers of several compounds achieved different scores as compared to earlier screenings in which tautomerism and rotameric states were disregarded. In order to perform further investigations using an extended library of 1200 drug-like purine and pyrimidine analogs from ACD, a program to generate all tautomers of each database entry was developed. Subsequent ranking and scoring revealed new hits.

212.

GENERATING PHARMACOPHORE QUERIES AND SMALL MOLECULE REPRESENTATIONS FOR VIRTUAL SCREENING. *E Keith Davies, Department of Chemistry, Oxford University, Central Chemistry Laboratory, South Parks Road, Oxford, United Kingdom, Fax: +44 1865 275905, Keith.Davies@Chem.ox.ac.uk, and Catherine J Davies, Trewaren Consultants Ltd*

There are many approximations including protein conformation, presence of water molecules, undetermined chirality, tautomerism and protonation state which have the potential to materially affect the results from virtual screening. The approach used by THINK endeavours to represent the possible interactions between a small molecule and a receptor. The resulting pharmacophores are weighted and may be adjusted to accommodate prejudices which may be supported by experimental binding or other data. The implementation used, the consequential small molecule representations, some validation examples and impact on the results will be discussed.

213.

ELECTRONIC STRUCTURE AND CHEMICAL POTENTIALS OF MOLECULES IN SPC/E WATER. *Ruth M. Lynden-Bell, Atomistic Simulation Group, School of Mathematics and Physics, Queen's University Belfast, Belfast BT7 1NN, United Kingdom, Fax: +44-2890-683274, R.Lynden-Bell@qub.ac.uk, and Stuart Murdock, Atomistic Simulation Group, School of Maths and Physics, Queen's University Belfast*

Some molecules change their structure significantly in solution and others do not. We describe some calculations of the electronic structure of triazoles and *n*-methyl piperidinol in aqueous solution using a hybrid method together with classical simulations to explore phase space as suggested by R.H.Wood et al. (J. Chem. Phys. 110 (1999) 1329. We find that Wood's method gives a good classical model in a single iteration. Further simulations using this model are used with the cumulant expansion to find the difference in excess chemical potential of the quantum molecule and the good classical molecule and hence the effect of solvation on equilibrium constants. The molecules are significantly polarised with enhanced dipole moments, but the fluctuations in polarisation are not very large, and the quantum-classical corrections relatively small. The solvation effects are interpreted using information from the probability distribution functions of the oxygen and hydrogen sites relative to the solute.

214.

PREDICTING ABSOLUTE FREE ENERGIES OF SOLVATION FOR THE PROTON AND THE HYDROXIDE ION. *David A. Dixon¹, David F. Feller¹, Chang-Guo Zhan¹, and Joseph S. Francisco².* (1) *W. R. Wiley Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, PO Box 999, Richland, WA 99352, Fax: 509-375-6631, david.dixon@pnl.gov,* (2) *Department of Chemistry, Purdue University*

Ionic dissociation is an important process in aqueous chemical systems. The absolute solvation free energy of a single ion is difficult to measure. We have used high-level, first-principles supermolecule-continuum calculations to predict the absolute hydration free energy of the proton and the hydroxide anion. Part of the solvent surrounding the solute is treated quantum mechanically and the remaining bulk solvent is approximated by a dielectric continuum medium based on a recently developed self-consistent reaction field model known as the surface and volume polarization for electrostatic interaction (SVPE) model. The sum of the absolute hydration free energies of the proton and hydroxide is calculated to be -366.9 kcal/mol, in excellent agreement with the experimental thermodynamic value of -366.6 ± 0.1 kcal/mol. We will also discuss applications of this computational approach to the prediction of the gas phase and solution acidities of HNO, HONO, HOONO, and HONO2.

215.

GENERALIZED BORN CONTINUUM SOLVATION MODELS FOR PROTEIN STRUCTURE PREDICTION. *Richard A. Friesner, Department of Chemistry, Columbia University, 3000 Broadway, MC 3110, New York, NY 10027, Fax: 212-854-7454, rich@chem.columbia.edu*

We have developed an improved version of the generalized Born model which provides good accuracy for modeling biomolecular systems. A surface formulation of the generalized Born model is used, providing a correct description of the solvent accessible region of the system. A novel treatment developed by Levy and coworkers provides qualitatively improved accuracy for the nonpolar solvation term as compared to traditional surface area based approaches. The model is parametrized for protein simulations against a combination of small molecule experimental solvation free energies, explicit solvent simulations, and experimental protein structural data. The method is applied to the prediction of protein loop and side chain conformations

216.

TOWARDS AB INITIO CAVITIES FOR DIELECTRIC CONTINUUM SOLVATION MODELS. *Michel Dupuis¹, Donald M. Camaioni¹, Daniel M Chipman², and John Bentley².* (1) *Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, P.O. Box 999, K8-91, Richland, WA 99352, michel.dupuis@pnl.gov,* (2) *Notre Dame Radiation Laboratory, University of Notre Dame*

Hydration energies of multifunctional charged and zwitterionic organic radicals are not predicted accurately with existing dielectric continuum models. The variability can be traced to the selection of the atomic radii and/or iso-density

contour values that define the solute cavities. We are developing a novel approach to the definition of cavities that embodies the chemical interaction between solute and solvent. It is based on obtaining a surface of minimum interaction energy for a solute-water complex, extracted from ab initio-derived interaction potentials or directly from ab initio calculations. An iso-density surface of the solvent water molecule is subsequently 'rolled' on top of the minimum interaction energy surface. The surface of closest proximity between the rolled water density and the solute serves as the cavity. In this presentation we will describe the approach and illustrate its application to the case of water in water, where the water solute acts as both hydrogen-bond donor and hydrogen-bond acceptor, and for some ions in water.

217.

CHARGE FLOW AND SOLVENT DYNAMICS IN THE PHOTODISSOCIATION OF MOLECULAR ANIONS. *Robert Parson, Nicole Delaney, James Faeder, and Matthew Thompson, JILA and Department of Chemistry and Biochemistry, University of Colorado, 440 UCB, Boulder, CO 80309-0440, Fax: 303-492-5235, rparson@jila.Colorado.edu*

Experimental studies of the photodissociation and recombination of molecular ions embedded in mass-selected clusters, carried out by the Lineberger and Neumark groups, present a major challenge to theory. Since the polarizable solvent strongly perturbs the solute electronic structure, one must deal with multidimensional potential surfaces that explicitly include the solvent degrees of freedom. Our approach is to construct an effective Hamiltonian that describes the electronic structure of the solvated molecule. By representing the solute wave functions in terms of a Distributed Multipole Expansion, the necessary interaction matrix elements can be efficiently calculated. At each step of a molecular dynamics simulation the effective Hamiltonian matrix is constructed and diagonalized, providing the energies, forces, and nonadiabatic transition probabilities required to propagate classical trajectories. Applied to solvated I_2^- and ICl^- , the simulations successfully reproduce the experimental product distributions, pump-probe spectra, and time-resolved photoelectron spectra. A physical picture inspired by the Marcus theory of electron transfer reactions allows us to organize and interpret the computational results.

218.

CONTINUUM AND HYBRID SOLVENT MODELS BASED ON THE FINITE-DIFFERENCE POISSON-BOLTZMANN METHOD. *Ray Luo, Department of Molecular Biology and Biochemistry, Univ. Calif. Irvine, Irvine, CA 92697-3900, Fax: 949-824-8551, rluo@uci.edu*

Continuum solvent models have been widely accepted as an efficient treatment of solvation in computational studies of static and dynamic properties of biochemical systems. However, there has been no satisfactory treatment of solvation in large biochemical systems unless the time-consuming Poisson-Boltzmann approach is used. Recently, we implemented an efficient finite-difference Poisson-Boltzmann treatment (FDPB) of solvation and long-range electrostatic interactions for biomolecular simulations (Luo, David, and Gilson, in press, JCC 2002). We adopted the strategy of Particle-Particle and Particle-Mesh to achieve the desired efficiency and accuracy in computing total electrostatic energy and forces in a dynamics simulation. Molecular dynamics simulation with the Poisson-Boltzmann method was shown to be robust and efficient for many systems tested. The application of FDPB is further extended to the hybrid explicit-implicit solvation models. To test the validation of this idea, we first implemented a hybrid model for equilibrium simulations in a water-droplet setting. High quality simulation results as indicated by small RMS deviation from crystal structures have been observed for three systems studied. The efficiency of the current implementation is similar to that using 15 Angstrom cutoff in those water-droplet simulations. These encouraging preliminary results show promise in developing more general and robust hybrid solvent models based on the similar approach.

219.

EVALUATING PROTEIN-LIGAND INTERACTIONS THROUGH FLEXIBLE DOCKING. *Tad Hurst, ChemNavigator, 6166 Nancy Ridge Drive, San Diego, CA 92121, Fax: 858-625-2377, thurst@chemnavigator.com*

As the global research emphasis shifts from genomics to proteomics, the question of how copious amounts of bioinformation will ultimately be used to accelerate the discovery of therapeutic compounds becomes more prominent. At

the same time, the number of commercially accessible compounds that can be tested for pharmaceutical efficacy has exploded into the millions.

ChemNavigator is addressing this need by offering advanced docking technology that more efficiently evaluates protein-ligand interactions. ChemNavigator has developed ultra-fast 3-D technology that will allow millions of structures to be docked into thousands of protein targets. In addition to rapid analysis, this technology will allow flexible ligand docking against the entire surface of a protein, not requiring specification of an active site.

This presentation details how ChemNavigator's novel 3-D flexible docking technology can assist life science researchers by allowing them to quickly and efficiently filter millions of structures in their search for novel therapeutic compounds.

220.

AFFINITY BASED HIGH THROUGHPUT SCREENING OF ORPHAN TARGETS: PRACTICAL SOLUTIONS FOR REMOVING PROMISCUOUS BINDERS. *Kenneth M. Comess¹, Martin J. Voorbach¹, Michael L. Coen¹, Hua Tang¹, Lan Gao¹, Xueheng Cheng¹, Mark E. Schurdak¹, Bruce A. Beutel², and David J. Burns¹. (1) Department of Biological Screening, Abbott Laboratories, Dept. R4PN, Bldg. AP10-1, 100 Abbott Park Road, Abbott Park, IL 60064-6099, Fax: 847-935-4200, kenneth.m.comess@abbott.com, (2) Dept. 47P, Bldg. AP52, Abbott Laboratories*

In the last year, Abbott Laboratories has fully implemented an affinity selection with mass spectrometric detection (ASMS) technique for high throughput screening of known and orphan targets. Although novel potent inhibitors have been discovered, certain functional groups and specific nuisance compounds also have been observed as promiscuous hits against a variety of targets. Since the assay type as well as readout is a constant in ASMS, it is possible to avoid such undesirable compounds through an empirically derived database of exclusionary criteria. We chose heat inactivated and dialyzed fetal calf serum as an excellent promiscuous target for a comprehensive screen against several hundred thousand compounds. We then analyzed and compiled several million data points to construct the exclusionary database. This strategy has led to a streamlined approach to lead discovery and prioritization, as well as interesting conclusions on nuisance functional groups and pharmacophores.

221.

COMPUTATIONAL SOLVENT MAPPING FOR THE ANALYSIS OF ENZYME ACTIVE SITES. *Sandor Vajda¹, Sheldon Dennis¹, Tamas Kortvelyesi¹, Michael Silberstein², and Karl Clodfelter². (1) Department of Biomedical Engineering, Boston University, 44 Cummington Street, Boston, MA 02215, Fax: 1-617-353-6766, vajda@bu.edu, (2) Program in Bioinformatics, Boston University*

A major challenge in structural genomics is the elucidation of biochemical and biological properties of enzymes. The two main sources of information on specific molecular interactions are the structures of the enzyme (or its homologues), co-crystallized with various ligands, and site directed mutagenesis of the putative binding site residues. Since both approaches are slow and labor-intensive, developing a method for determining the functional site on the basis of protein structure has been the target of intensive research.

We have developed a novel method that can find the residues of enzyme binding sites. The method moves molecular probes - small organic molecules or functional groups - around the protein surface in order to identify their most favorable binding positions. The mapping procedure reproduces the available experimental solvent mapping results, eliminating the problem of spurious local minima associated with previous computational methods. A very important result is that using at least six different probes, the consensus sites found by the mapping are always in the major subsites of the functional site, and the amino acid residues that interact with the probes also bind the specific ligands of the protein. Thus, computational mapping provides detailed information on the functional sites of proteins. The approach is less sensitive to variations in the structure of the protein than docking methods, and it is remarkably robust against changes in the algorithm and energy parameters. Preliminary applications of mapping to docking and drug design will be discussed.

222.

THEMATICS: IDENTIFICATION AND CHARACTERIZATION OF ACTIVE SITES FROM STRUCTURE ALONE. *Mary Jo Ondrechen, Department of Chemistry, Northeastern University, Boston, MA 02115, Fax: 617-373-8795, James G. Clifton, Rosenstiel Basic Medical Sciences Research Center, Brandeis University, and Dagmar Ringe, Departments of Biochemistry and Chemistry, Brandeis University*

Structural genomics efforts are rapidly increasing the pace of discovery of protein structures, many of which are of unknown function. THEMATICS, for Theoretical Microscopic Titration Curves, is a new method (based on established computational tools) for identification and characterization of enzyme active sites from the structure alone. Predicted titration curves for an ionizable residue in a protein express the mean net charge as a function of pH for that residue in an ensemble of protein molecules. A typical curve has sigmoid shape, with a sharp fall-off near the pKa. A small fraction of the ionizable residues in a protein molecule are predicted to deviate markedly from this typical Henderson-Hasselbalch form. These perturbed residues are functionally significant; a cluster of two or more perturbed residues in close physical proximity is a reliable predictor of active site location. Results of predicted active site features are presented for some proteins of pharmaceutical interest.

223.

IDENTIFICATION OF A SECOND BINDING SITE IN THE ESTROGEN RECEPTOR.

Willem P van Hoorn, Molecular Informatics, Structure and Design, Pfizer Global Research and Development, Ramsgate Road, Sandwich CT13 9NJ, United Kingdom, Fax: +44-1304-658422, willem_van_hoorn@sandwich.pfizer.com

Fluorescence spectrometry data by Tylumenkov and Klinge suggest the presence of a second binding site in both subtypes ER α and ER β of the Estrogen Receptor (ER). A cavity previously described as a solvent channel was located in close proximity to the steroid binding site of both ER subtypes. Derivatives of a tetrahydrochrysenone (THC) compound, speculated in literature to bind to a second binding site, were docked successfully in the second sites identified. However, computation of accurate interaction scores indicates preferred binding to the steroid binding site over the second binding site of both ER α and ER β for all THC derivatives. Therefore, binding to this second site is probably not the reason why the THC derivatives are agonists on ER α and antagonists on ER β . Most likely, the smaller steroid binding site of ER β compared to ER α and/or the apparent larger flexibility of helix 12 of ER β , make ER β more readily adopt an antagonist conformation.

224.

SUPRAMOLECULAR TRANSPORT OF RETINOL AND FENRETINIDE:

IMPLICATIONS IN BREAST CANCER. *Amy Marie Waligorski, Center for Computational Sciences, Duquesne University, Department of Chemistry and Biochemistry, 600 Forbes Avenue, Pittsburgh, PA 15282, Fax: 412-396-5683, waligor805@duq.edu, and Jeffrey D. Evanseck, Department of Chemistry & Biochemistry and Center for Computational Sciences, Duquesne University*

Molecular dynamics simulations were carried out for both retinol and fenretinide complexes. The complexes consisted of the specific retinoid (either retinol or fenretinide) bound to retinol binding protein (RBP) and transthyretin (TTR). Our newly developed force field parameters for alternating double and single bonds in the pi-conjugation were used in this simulation work. Fenretinide has been shown to cause apoptosis in certain tumor cells. The transport system involves a tetramer of TTR and two RBP-complexed retinoids. However, one postulate is that the fenretinide-RBP bound to TTR is sufficiently destabilized and degraded by the kidney. The scientific challenge is to derivatize fenretinide or mutate RBP to increase the binding of the TTR complex with RBP while maintaining its chemotherapeutic power. The molecular dynamics simulations provide the necessary microscopic insight about the molecular interactions and forces that are responsible for binding of these complexes and how to improve upon them.

225.

QM AND MIXED QM/MM SIMULATIONS OF MODEL NERVE AGENTS WITH ACETYLCHOLINESTERASE. *J. B. Wright, Natick Soldier Center, Materials Science Team, U. S. Army Soldier and Biological Chemical Command, ATTN: AMSSB-RSS-M, Kansas Street, Natick, MA 01760, Fax: 508-233-5521, jeffery.wright@natick.army.mil, Margaret M. Hurley, Computational and Information Sciences Directorate, U.S. Army Research Laboratory, Gerald Lushington, Molecular Graphics and Modeling Laboratory, University of Kansas, William E. White, Edgewood Chemical and Biological Center, and Jason A. Morrill, Weapons and Materials Research Directorate, United States Army Research Laboratory*

Nerve agents pose a major threat to DoD personnel and operations - not only from overt military attacks but also from terrorist activity and other low intensity conflicts. This project entails characterization of the structural features, which lie at the heart of the interaction between agents (or other more reversible inhibitors) and the enzyme acetylcholinesterase. Large biological systems present inherent difficulties to the modeling community. We are using quantum and mixed quantum/classical techniques, QM/MM, to examine structural effects in model nerve agent reactions in the active site of the enzyme, as well as the role of surrounding residues. Through this we hope to provide a sound theoretical method for prediction of toxicity of novel agents, as well as antidote effectiveness. This work represents part of a DoD HPCMO Challenge project.

226.

VOLUME CHANGE STUDIES USING QUANTUM MECHANICALLY DERIVED

QSPR. *Andrew J. Holder¹, Derek A. White¹, Robert Smith², Cecil C. Chappelow³, Charles Pinzino³, J. D. Eick⁴, and Jason A. Morrill⁵. (1) Department of Chemistry, University of Missouri - Kansas City, UMKC, Flarsheim Hall, Rm 410h, 5110 Rockhill Road, Kansas City, MO 64110, holdera@umkc.edu, whitede@umkc.edu, (2) School of Pharmacy, University of Missouri - Kansas City, (3) Midwest Research Institute, (4) Department of Oral Biology/School of Dentistry, University of Missouri - Kansas City, (5) Weapons and Materials Research Directorate, United States Army Research Laboratory*

This study was conducted in support of ongoing research into a photo-initiated cationically polymerized nonshrinking dental restorative material. The purpose of this study was twofold. First, selected information gleaned from semiempirical calculations using AM1 was employed to develop a multilinear regression to aid in understanding the mechanism or mechanisms of volume change upon polymerization of selected study monomers. Second, a detailed statistical study of the descriptors found by the regression analysis coupled with the chemical implications of those descriptors was used to screen potential expanding monomers as well as to predict structures and/or moieties specifically related to volume change. Regression models were developed for each of the monomer types as well as for a combination of all monomers. Chemical information derived from the models is discussed and evaluated with respect to previously developed theories of volume change upon polymerization.

227.

TOPOLOGICALLY INTERESTING MOLECULES WITH ONLY ONE SIDE. *Michelle M. Francl, and Dipannita Kalyani, Department of Chemistry, Bryn Mawr College, 101 N Merion Ave, Bryn Mawr, PA 19010, Fax: 610-526-5086, mfrancl@brynmawr.edu*

The structures of several classes of large molecules having the same topology as Möbius strips are explored using semi-empirical and ab initio methods. These molecules have potential as chiral hosts in host-guest chemistry. Suitability for host-guest chemistry and the thermodynamic stability of various prototypes will be presented. Some of these systems contain regions that are highly strained. The factors that control the degree of strain have been elucidated.

228.

MODEL FOR SOLUBILITY PARAMETER PREDICTIONS DERIVED BY QSPR.

Andrew J. Holder¹, J. D. Eick², James E. Code¹, and Cecil C. Chappelow³. (1) Department of Chemistry, University of Missouri - Kansas City, UMKC, Flarsheim Hall, Rm 410h, 5110 Rockhill Road, Kansas City, MO 64110, holdera@umkc.edu, codej@umkc.edu, (2) Department of Oral Biology/School of Dentistry, University of Missouri - Kansas City, (3) Midwest Research Institute

This research developed a computational model to predict approximate Hildebrand solubility parameters (δ_1) of solution monomers. A monomer with high

solubility parameter requires more energy for dispersal than is gained by mixing it with a material of low solubility parameter, resulting in immiscibility. Solubility parameter theory is largely based on dipoles. The hypothesis tested was that predicted δ_s should correlate ($R^2 > 0.900$) with literature δ_s . Information from semiempirical AM1 calculations was employed to develop a multilinear regression for predicting δ_s . A detailed statistical study of the descriptors determined by regression analysis coupled with the chemical implications of the descriptors was done. Predicted solubility parameters for chemical structures outside the training set of molecules externally validated the model by a least squares fit ($R^2=0.964$), supporting the proposed hypothesis.

229.

SEMI-EMPIRICAL CALCULATIONS OF ROD-LIKE DINUCLEAR RUTHENIUM COMPLEXES FOR DYE SENSITIZED PHOTOVOLTAIC APPLICATIONS. *John Walker¹, Lynn Samuelson¹, Ravi Mosurkal², and Jayant Kumar².* (1) U.S. Army Natick Soldier Center, Department of the Army, U.S. Soldier/Chemical Biological Command, Natick Soldier Center, Kansas Street, Natick, MA 01760, Fax: 508-233-5521, John.Walker@natick.army.mil, (2) Department of Chemistry, Center for Advance Materials, University of Massachusetts-Lowell

Photosensitization of porous nanocrystalline TiO₂ film electrodes using polypyridine complexes has intensively been investigated for solar cell applications. An important step for the next generation of solar cells based on this technology is to prepare sensitizers that have an intense response in the solar spectrum. In this study we have carried out molecular modeling on rod-like homometallic dinuclear ruthenium complexes with the typical formula [(Cl)(dcbpy)Ru(terpy-ph-terpy)Ru(dcbpy)(Cl)], where dcbpy is 4,4'-dicarboxy-2,2'-bipyridine, and terpy-ph-terpy is 4,4'-(1-4-phenylene)bis-(2,2'-6',2''-terpyridine). Computations were performed using the following programs: Cerius24_6, Dmol3, and Zindo. All codes are commercially available (Accelrys, San Diego, CA). HOMO-LUMO calculations were carried out in support for syntheses of the Ru complexes. Experimentally obtained photovoltaic properties of the synthesized sensitizers are compared to the computational results.

230.

SEMIEMPIRICAL STUDY OF THE MECHANISM OF CATIONIC HOMOPOLYMERIZATION AND COPOLYMERIZATION OF MONOMER SYSTEMS.

Andrew J. Holder¹, Matthew D. Miller¹, J. D. Eick², and Cecil C. Chappelow³. (1) Department of Chemistry, University of Missouri - Kansas City, UMKC, Flarsheim Hall, Rm 410h, 5110 Rockhill Road, Kansas City, MO 64110, holdera@umkc.edu, mdma95@umkc.edu, (2) Department of Oral Biology/School of Dentistry, University of Missouri - Kansas City, (3) Midwest Research Institute

In support of ongoing research into dental restorative materials, a computational study was undertaken to determine the activation energies, enthalpies of reaction, the endo- or exothermicity of, and potential mechanisms for the cationic homopolymerization and copolymerization of monomer systems of interest. Enthalpies of activation and reaction were obtained by modeling various reaction mechanisms using the AM1 semiempirical quantum mechanical method. Assuming similar steric effects for similar systems, these enthalpies can be directly compared and the relative reactivity inferred. The systems studied included 1,5,7,11-tetraoxa-spiro[5.5]undecane, 3,9-bis-(4-methyl-7-oxa-bicyclo[4.1.0]hept-3-ylmethyl)-1,5,7,11-tetraoxa-spiro[5.5]undecane, and bisphenol-A-diglycidyl ether. Following the modeling the calculated results were compared to results obtained experimentally to verify accuracy and identify comparative strengths and weaknesses of the methods employed.

231.

SEMIEMPIRICAL CALCULATIONS OF REACTIVE SURFACES FOR TWISTED INTRAMOLECULAR CHARGE TRANSFER (TICT). *M Maroncelli, and Weiping Song, Department of Chemistry, Penn State University, 152 Davey Laboratory, University Park, PA 16802, Fax: 814-863-5319, mpm@chem.psu.edu*

In this study, Austin Model 1 (AM1/Cl) calculations coupled to Conductor-like Screen Model (COSMO) continuum solvation model were used to calculate 1-dimensional potential energy surfaces of the excited-state twisted intramolecular charge transfer (TICT) reactions in a series of 4-N,N-bridged-aminobenzonitrile compounds in solution. These solvent-equilibrated surfaces were used to develop 2-dimensional (solvent polarization + intramolecular twisting motion) potential surfaces, which are capable of semi-quantitatively reproducing the

steady-state absorption and emission spectra of these species in a range of solvents. The nature of these surfaces, as well as preliminary results concerning reaction dynamics on them will be discussed and compared to experimental data.

232.

PROTOPLEX: USER-CONTROL OVER TAUTOMERIC AND PROTONATION STATE.

R. S. Pearlman¹, R. Khashan¹, D. Wong¹, and R. Balducci². (1) Laboratory for the Development of CADD Software, University of Texas, College of Pharmacy, Austin, TX 78712, pearlman@naphthyl.phr.utexas.edu, (2) CADD Software LLC

The tautomeric state and/or protonation state, hereinafter referred to as the "protomer", which happens to appear in a corporate database or third-party database of compounds may or may not be the protomer appropriate for a given task. Physical property- and ADMET-related software often yields incorrect predictions unless structures are input in completely unionized form or the protonation state "predominant" at a particular pH. Moreover, such software usually considers only the tautomeric state represented on input. Similarly, software used for virtual screening will yield false negatives if the protomer in the virtual screening library is not the protomer preferred by the receptor. ProtoPlex generates all protomers of drug-sized compounds subject to user-specified rules regarding the number of such protomers and the protonation- and tautomerization-sites to be "multiplexed." StereoPlex was developed over 12 years ago to provide the same type of functionality with respect to atom- and bond-centered stereochemistry and the important relationship between ProtoPlex, StereoPlex, and vHTS software will be discussed.

233.

A TAUTOMER AND PROTONATION PRE-PROCESSOR FOR VIRTUAL

SCREENING. *Jens Sadowski, Structural Chemistry Laboratory, AstraZeneca, Mölndal S-43183, Sweden, jens.sadowski@astrazeneca.com*

The tautomeric and protonation state of a substrate or inhibitor molecule binding to its protein receptor is essential for the binding mode. Often, there exist a number of equally reasonable different tautomers and protonation states for drug-like molecules. In these cases, it is difficult to predict the correct state without knowing the receptor environment. On the other hand, this information is needed in order to predict the complex correctly.

In principle, this behaviour—shifting hydrogens over the molecular skeleton—could be seen as a number of additional degrees of freedom which should be considered by, e.g., docking programs. Unfortunately, most of the commonly used docking codes neglect this. Therefore, a method has been developed that is capable to generate a number of reasonable tautomers and protonation states for drug-like molecules before docking.

The method was applied to study the impact of tautomerism and protonation state on receptor binding. A dataset of several thousand experimental complex structures from the Brookhaven Protein Databank was used to study this. Statistics of the effect of tautomerism and protonation on the binding mode are presented.

234.

TO CLEAN OR NOT TO CLEAN? THE QUESTION OF APPROPRIATE STRUCTURAL REPRESENTATION IN CORPORATE DATABASES.

M. J. Polley, J. T. Swanson, W. Homer, and R. D. Clark, Tripos, Inc, 1699 S. Hanley Road, St. Louis, MO 63144, mpolley@tripos.com

Corporate databases contain structures from a myriad of sources, and in a variety of forms, some much better than others. Different techniques may require these structures to be represented in any one of a number of forms. For example, docking programs typically want structures in the protonation or tautomeric state that they will have in the binding site of the target protein. In contrast, a ClogP calculation requires the structures be represented in their neutral form. It is therefore critical to think about how structures are represented in a corporate database and to carefully consider techniques for producing the desired representation. This talk discusses some of the work done at Tripos to provide tools to address the issue of consistent and appropriate structures for cheminformatics applications.

235.

AUTOMATED DATABASE IONIZATION. *Jasna Klicic, and Mark Reboul, Schrodinger, Inc, 120 W 45th St., New York, NY 10036, Fax: 646-366-9550*

We have created a computational method to expand a database of neutral organic molecules by generating a selection of protonation states for each entry. Based on user's preferences, a selection of protonation states can range from generating only ionization states that are predominant in the solution at a given pH to generating all possible states. The latter option has an important application for preparing a database for virtual screening. Docking a variety of protonation states for each molecule will compensate for variability of the effective pH in different parts of a receptor binding site. The program 'ionizer' is directly compatible with the Maestro structural file format.

236.

HOOKED ON PROTONICS. *Roger Sayle, and Geoffrey Skillman, OpenEye Scientific Software, Suite 1107, 3600 Cerrillos Road, Santa Fe, NM 87507, Fax: 505-473-0833, roger@eyesopen.com*

Most modern chemistry, in particular chemoinformatics, is built upon the valence-bond model of a molecule. Unfortunately, the mobility of formal charges and protons, a common feature of many organic molecules under physiological conditions, makes it difficult to discuss the most basic concept of this model: the equality of two compounds. This talk describes the complexity inherent in the interaction between tautomerism, dissociation, resonance and representation. Computational methods are described for handling the resulting equivalence classes and enumerating physiologically relevant forms of drug-like molecules. In addition the importance to docking studies and property prediction will be illustrated.

237.

AB-INITIO SIMULATIONS OF SIMPLE AQUEOUS SOLUTIONS. *Giulia Galli, Eric R. Schwegler, and Jeffrey C. Grossman, Lawrence Livermore National Laboratory, Livermore, CA 94551, galli@lml.gov*

In this talk I plan to describe some recent progress in ab-initio simulations of solvation effects in simple systems, including Na⁺, Mg⁺⁺ and dimethyl phosphate in water, as well as hydrophobic molecules such as methane and silane.

238.

AB INITIO MOLECULAR DYNAMICS STUDIES OF ION-WATER INTERACTIONS. *Douglas J. Tobias¹, I-Feng W. Kuo¹, and Christopher J. Mundy². (1) Department of Chemistry, University of California, Irvine, Irvine, CA 92697-2025, Fax: 949-824-8571, dtobias@uci.edu, (2) Chemistry and Materials Science, Lawrence Livermore National Laboratory*

The interaction between water molecules and ions generally modifies the electronic structure of both species. When water molecules interact with anions, the OH bonds are generally weakened. When the ion solvation shell is non-centrosymmetric, ions can be polarized. This talk reports investigations based on ab initio (Car-Parrinello) molecular dynamics simulations of electronic polarization effects on ion-water interactions in three settings: halide-ion clusters, where ionic polarization is largely responsible for ion adsorption at aqueous interfaces; the structure of the solvation shell of the dimethylphosphate anion (a model for nucleic acid backbone and phospholipid headgroups), where polarization effects are surprisingly small; the vicinity of the retinal Schiff base in the proton pump bacteriorhodopsin, where a key water molecule in the proton transfer pathway is found to be highly polarized by a cluster of surrounding charged residues.

239.

TREATING MANYBODY DISPERSION AND POLARIZATION WITHIN A UNIFIED FORMALISM. *Glenn Martyna, Physical Science Division, IBM Research, TJ Watson Research Center, PO Box 218, Yorktown Heights, NY 10598, Fax: 914-945-4506, martyna@us.ibm.com, and Troy Whitfield, Department of Chemistry, University of Pennsylvania*

Manybody polarization and manybody dispersion are key environmentally dependent effects which are ignored in nearly all atomic level treatments of solvation. Here, a novel formulation of manybody polarization and manybody dispersion, the quantum Drude oscillator model, is presented. It is shown how the model can be solved efficiently and accurately using Feynman Path Integrals to yield the structure and dynamics of atomistic systems. It is also shown how the model can be parameterized for both simple, xenon, and more complex systems, water. Results for both xenon and water demonstrate that both manybody dispersion and polarization change markedly with density and influence the structure and dynamics of these systems.

240.

FIRST PRINCIPLES STUDY OF AQUEOUS HYDROXIDE SOLUTIONS. *Bin Chen¹, Ivaylo Ivanov¹, and Michael L. Klein². (1) CMM & Department of Chemistry, University of Pennsylvania, 231 S. 34 Street, Philadelphia, PA 19104, Fax: 215-573-6233, binchen@cmm.chem.upenn.edu, (2) Department of Chemistry, University of Pennsylvania*

Experimental studies on alkaline solutions have demonstrated that the properties of water change significantly in the presence of hydroxide ions. Also the hydroxide ion has an unusually high mobility in water, comparable to that of the proton. In spite of many experiments, confusion persists regarding the solvation structures and the mobility mechanism of the hydroxide ion. To gain insight into the structure and dynamical behavior of the hydroxide ion in solution, Car-Parrinello molecular dynamics (CPMD) simulations were carried out for aqueous NaOH solutions, under ambient conditions, over a wide range of concentrations. The simulation results reconcile conflicting structural interpretations from previous experimental and theoretical studies on this system. Specifically, a variety of hydroxide complexes are present in solution and the dominant type of complex is dependent upon the concentration. Analysis of the CPMD trajectories supports the view that the transport mechanism of the hydroxide ion is different from that of proton.

241.

RAPID CALCULATION OF FREE ENERGY AND SOLUTION STRUCTURE WITH HIGH LEVEL AB INITIO MODELS OF SOLUTE-WATER INTERACTIONS. *Wenbin Liu, Robert H. Wood, and Douglas J. Doren, Department of Chemistry and Biochemistry, University of Delaware, Newark, DE 19716, Fax: 302-831-6335, wbliu@udel.edu*

We have recently proposed methods to obtain free energies and structures of models with QM/MM potential energies. The hybrid model, V1, consists of First Principles solute-solvent interactions for a cluster containing n water molecules and approximate MM models for all other interactions. Configurations are sampled from simulations with an approximate empirical model, V0. Calculations of V1 are only needed for 50 to 100 statistically independent configurations of a simulation with V0. Free energies of the QM/MM model are calculated by free energy perturbation from the simulation with V0. Non-Boltzmann weighting of the pair correlation function of the simulation with V0 gives the pair correlation function of the QM/MM model. In practice, this approach allows calculations that are several orders of magnitude faster than full First Principles simulations. These new methods have been used to calculate free energies, pair correlation functions, and running coordination numbers for aqueous sodium ion, chloride ion, and the NaCl pair at temperatures from 573 to 973 K and water densities from 0.01 to 0.725 g/cm³.

242.

AB INITIO MOLECULAR DYNAMICS VIA A NOVEL COMBINATION OF DISCRETE VARIABLE REPRESENTATIONS AND PLANE WAVE BASIS SETS. *Yi Liu, Dept. of Chemistry, New York University, 100 Washington Square East, New York, NY 10003, Fax: 212-260-7905, liuy@ariadne.chem.nyu.edu, and Mark E. Tuckerman, Department of Chemistry, New York University*

A simple, rigorous and efficient combined real space/reciprocal space approach for electronic structure calculations within density functional based *ab initio*

molecular dynamics calculations is presented. Specifically, a novel combination of discrete variable representations (DVRs) commonly used in reactive scattering and a plane wave basis is employed. A DVR is a set of approximate position eigenfunctions that are spatially localized. Here, the DVR is employed to treat the noninteracting kinetic energy and short-ranged interactions while a plane wave basis set is used to treat the long-ranged energy terms. By incorporating the screening function methodology recently introduced by Martyna, Tuckerman, and coworkers, cluster, wire and surface boundary conditions can be treated as well as full three-dimensional periodicity. The method is considerably simpler and more flexible than Gaussians and scales inherently as $O(N^2)$. The method is compared to the full plane-wave based approach commonly used in *ab initio* molecular dynamics for periodic and cluster systems and shown to yield much faster convergence with grid size.

243.

CHEMOMETRICS: PAST, PRESENT, AND FUTURE. *Barry K. Lavine, Chemistry, Clarkson University, 8 Clarkson Ave, Potsdam, NY 13699-5810, Fax: 315-268-6610, bklab@clarkson.edu, and Jerome Workman Jr., Analytical & Measurement Technology, Kimberly Clark Corporation*

Chemometrics has enjoyed tremendous success in the areas related to calibration of spectrometers and spectroscopy-based measurements. These chemometric-based spectrometers have been widely applied for process monitoring and quality control. However, chemometrics has the potential to revolutionize the very roots of problem solving. The chemometric approach to scientific problem solving, which attempts to explore the implications of data so that hypotheses, i.e., models of the data, are developed with greater awareness of reality, can be summarized as follows: (1) measure the phenomena or process using chemical instrumentation that generates data inexpensively, (2) analyze the multivariate data, (3) iterate if necessary, (4) create and test the model, and (5) develop fundamental multivariate understanding of the process. The new approach does not involve a thought ritual; rather it is a method involving many inexpensive measurements, possible a few simulations, and chemometric analysis. It is a true paradigm shift since multiple experimentation and chemometrics are used as a vehicle to examine the world from a multivariate perspective. Mathematics is not used for modeling per se but more for discovery and is thus a data microscope to sort, probe, and to look for hidden relationships in data.

244.

TRANSFERRING AND IMPROVING THE ROBUSTNESS OF MULTIVARIATE CALIBRATIONS. *Steven D. Brown, Hu-Wei Tan, Robert Feudale, and Tony Myles, Department of Chemistry & Biochemistry, University of Delaware, Brown Laboratories, Newark, DE 19716, sdb@udel.edu*

One of the limiting factors in extracting quantitative relationships from chemical measurements is the inexact nature of chemical theory. First-principles modeling based on limiting laws like Beer's relation apply in only simple systems, yet rapid, reliable quantitation is often needed in complex mixtures. Quantitative methods based on soft modeling offer a route to calibration of complex mixtures whose spectra may not be well described by first-principles models. In essence, soft modeling uses a correlative relationship to calibrate multivariate variation present in the instrumental signature to variation in the desired property. The ability to model complex systems accurately is a major triumph of the field of chemometrics. Unintended variation in the instrumental response limits the utility of soft modeling in many practical applications, however. Environmental or instrumental effects corrupt the soft modeling and prevent widespread, practical use of multivariate methods based on soft modeling. We have explored several mathematical methods to correct the multivariate calibrations for these unintended effects and have developed several methods that function well with spectroscopic data. One method can be used to transfer NIR spectral data without need for standard spectra measured on the master and the target instruments. Another method requires only one spectrum to permit a transfer, while others based on wavelets offer routes to background removal and correction needed to perform standardization and robust calibrations. In many cases, the standardization methods work well enough to permit transfer of a calibration developed on one instrument to be used on another instrument with different variations due to environmental factors. These and related methods can be used to remove background and other undesired variation to improve the robustness of a calibration.

245.

APPLICATION OF MODIFIED ALTERNATING LEAST-SQUARES (MALS) REGRESSION TO SPECTROSCOPIC IMAGE ANALYSIS. *T. M. Hanczewicz¹, S.L. Zhang¹, Philip Hopke², and J.-H. Wang². (1) Data Analysis and Spectroscopy, Unilever Research, US, 45 River Road, Edgewater, NJ 07020, (2) Department of Chemical Engineering, Clarkson University*

Much work has been published in the literature and presented at conferences that deal with the analysis of spectroscopic imaging data and the difficulty there is in handling such data. This kind of data tend to be problematic due to the large size of the data sets and the rich chemical component content present in spatially overlapped, and/or spectroscopically unresolved form. The main goal of such analysis is to accurately resolve both component spectral response and the associated images so that physically and chemically meaningful information can be extracted from the data. The authors will demonstrate in this paper a new algorithm for solving such problems in a quick and efficient manner that is superior to the methods currently in use. Analysis of synthetically generated and real Raman imaging data sets will be used to show the significance of modified alternating least-squares (MALS) regression as a superior method of analysis compared with several other well-established mathematical algorithms. The performance of MALS will be compared with that of ordinary alternating least-squares regression (ALS) and fast non-negative least-squares (FNNLS) regression as the engine for a self-modeling curve resolution (SMCR) algorithm in applications of spectroscopic image analysis. The authors will show that MALS is superior in terms of computational speed, stability, and component resolution ability in the analysis of both synthetic and real data sets. The results will be discussed in terms of speed and performance of the analysis as compared to that of FNNLS. The authors will also show that MALS is superior to ordinary ALS in all performance aspects.

246.

CHEMOMETRIC TOOLS IN BIOINFORMATICS. *Anders Berglund, and Fredrik Pettersson, Research Group for Chemometrics, Umeå University, 901 87 Umeå, Sweden, Fax: + 46-90-138885, anders.berglund@chem.umu.se*

Multivariate techniques such as Principal Component Analysis (PCA) and Partial Least Squares (PLS) regression are today widely used in the chemometric society. We are applying these techniques on examples from the bioinformatic field. Many of the challenges and/or problems are similar to those in chemical examples, for instance, more variables than there are objects, noisy variables, missing values and so on. This has made us believe that applying these multivariate techniques will, in many examples, be an alternative to more classical methods that are used today. We will show examples where we have analyzed microarray data using a modeling approach based on PLS regression. We will also present other interesting examples from the bioinformatic field.

247.

MAPPING PROTEIN ENERGY LANDSCAPES BY CHEMOMETRIC METHODS. *Jeffrey D. Evanseck, Department of Chemistry and Biochemistry, Duquesne University, 600 Forbes Ave, Pittsburgh, PA 15282, Fax: 412-396-5683, evanseck@duq.edu, and Brita G. Schulze, Analytics and Production, MBT Munich Biotechnology*

The biomolecular motion of carbonmonoxy myoglobin (MbCO) has been investigated by chemometric methods. Principal component analysis was used to analyze ten short (400 ps) and two longer time (1.2 ns) molecular dynamics simulations, starting from five different crystallographic and solution phase structures centered in a 37 Å radius sphere of water. The motion of two amino acid residues, His64 and Arg45, were identified as the main factors for the observed A-states. The chemometric methods used to develop this new A-state model will be described. The model rationalizes the A3 state based upon the fluctuating electrostatic field generated by the gate-like dynamics of His64 and Arg45 is presented, which is consistent with previously reported time scales for substate interconversion and mutation studies.

248.

NATURAL BOND ORBITAL EVALUATION OF AM1-PREDICTED C—H—O HYDROGEN BONDS IN DIMERS OF 1,5,7,11-TETRAOXASPIRO[5.5]UNDECANE.

Andrew J. Holder, and C. David Harris, Department of Chemistry, University of Missouri - Kansas City, UMKC, Flarsheim Hall, Rm 410h, 5110 Rockhill Road, Kansas City, MO 64110, Fax: 816-235-6543, holdera@umkc.edu, harrisc@umkc.edu

Spiroorthocarbonates have potential utility as components of low-shrinkage polymeric materials. Density functional theory, ab-initio Hartree-Fock theory, and NBO analysis has been used to characterize AM1-predicted individual C—H—O intermolecular interactions in calculated gas phase van der Waals dimers of different conformations of 1,5,7,11-tetraoxaspiro[5.5]undecane. Relative contributions to the stabilization energy arising from charge transfer for any C—H—O close contacts predicted in the calculated geometries are approximated and compared. Emphasis is placed on identification of and differentiation between short-range intermolecular internuclear distances that represent hydrogen bonds and those that represent forced close contacts.

249.

THE USE OF SEMIEMPIRICAL METHODS FOR CONFORMATIONAL STUDIES OF NOVEL ANGIOGENESIS INHIBITORS.

J. Phillip Bowen¹, Scott Furness¹, Thomas Phillip Robinson¹, Tedman Ehlers¹, Richard Hubbard IV¹, Corey Fowler¹, Jack L. Arbiser², and David Goldsmith³. (1) Department of Chemistry, University of Georgia, Computational Center for Molecular Structure and Design, Athens, GA 30602, Fax: 706-542-2673, bowen@ccmsd.chem.uga.edu, (2) Department of Dermatology, Emory University, (3) Department of Chemistry, Emory University

The development of angiogenesis inhibitors is a promising approach for anti-tumor agents. Semiempirical methods, along with other computational approaches, have been used to understand and predict conformational preferences for two classes of potentially useful antiangiogenic compounds: decalin dione and curcumin analogs. Based on a pharmacophore hypothesis, a series of curcumin analogs have been designed and synthesized. Subsequent in vitro assays have indicated excellent tumor cell growth inhibition. In particular, semiempirical methods have been used to probe stereoelectronic effects that may be responsible for the observed activity based on the proposed pharmacophore model and available structure-activity data.

250.

QSAR MODEL FROM SEMIEMPIRICAL METHODS TO PREDICT INDUCTION OF THE CYP1A1 GENE.

Andrew J. Holder¹, Kimberly M Andreassen¹, David M. Yourtee², J. D. Eick³, and Cecil C. Chappelow⁴. (1) Department of Chemistry, University of Missouri - Kansas City, UMKC, Flarsheim Hall, Rm 410h, 5110 Rockhill Road, Kansas City, MO 64110, holdera@umkc.edu, kma9bd@umkc.edu, (2) Department of Pharmacology/School of Pharmacy, University of Missouri - Kansas City, (3) Department of Oral Biology/School of Dentistry, University of Missouri - Kansas City, (4) Midwest Research Institute

Induction of cytochrome P450 (CYP) genes is an area of great interest in the development of therapeutics and biomaterials. Cytochrome P450 enzymes, eventual products of the CYP genes, are responsible for the metabolism of endogenous and exogenous chemicals. Alterations in the amount of this enzyme present in the body can lead to deleterious health consequences due to the build up of substances that may be toxic or rapid elimination of beneficial substances. The cytochrome P450 1A1 (CYP1A1) enzyme is responsible for the metabolism of polycyclic aromatic hydrocarbons and nitroarenes. To aid in the development of dental materials, a QSAR model predicting the induction ability of materials was produced using information from the AM1 semiempirical method. A detailed statistical study of the descriptors found by the regression coupled with chemical and biochemical intuition was used to interpret the model.

251.

NEW SEMIEMPIRICAL QUANTUM METHODS FOR BIOLOGICAL REACTIVITY.

Darrin M. York, Department of Chemistry, University of Minnesota, 207 Pleasant Street S.E, Minneapolis, MN 55455, york@chem.umn.edu

A new semiempirical model Hamiltonian is presented for biological reactions. The Hamiltonian is based on the NDDO approximation with several extensions including dispersion and hydrogen bonding terms and pairwise short-ranged core-core interactions. A large quantum database of biological molecules,

hydrogen bond complexes, and reactions involving phosphates and phosphoranes is outlined. A new non-linear optimization suite and strategy for semiempirical parameter development is described. The resulting semiempirical model Hamiltonians provide very high-level accuracy for hydrogen bonding and phosphate reactivity with over three orders of magnitude of computational savings with respect to density-functional calculations. Applications to linear-scaling semiempirical quantum and hybrid quantum mechanical/molecular mechanical calculations are presented for nucleic acids and phosphate hydrolysis reactions.

252.

NEW SEMIEMPIRICAL METHOD FOR FAST AND ACCURATE PROTEIN MODELING.

Nikolai Anikine¹, Vladislav Bugaenko¹, Vladimir Bobrikov¹, Alexey Andreyev¹, and Victor M. Anisimov². (1) Quantum Biochemistry Group, Konstantina Fedina-3/24, Moscow 105215, Russia, aanikin@swf.chem.ac.ru, (2) FQS Poland, Starowislna 13-15, Krakow 31-038, Poland, victor@fqspl.com.pl

New semiempirical method has been developed for a purpose of fast and accurate protein modeling. The method provides more accurate computational results than Density Functional Theory methods (B3LYP/6-31G*) in application to protein modeling. Fast linear scaling SCF engine implemented eliminates a computer resource bottleneck of application of quantum-chemistry methods to biologically important molecular systems demonstrating calculations of million atom bio-molecules on a single CPU based personal computer. The method allows splitting large molecule on smaller fragments and performing SCF calculations on these parts separately. Different level of accuracy can be assigned to the molecular fragments. This way the whole system is considered quantum-mechanically at the same level of theory eliminating thus a problem of transition zone between parts of system treated by different methods common for QM/MM and QM/QM methods. While a scope of the current work is limited to accurate description proteins, DNA/RNA, drugs, carcinogens and their complexes, the method itself is extendable to accurate description of other classes of chemical compounds, for instance, carbon nanotubes, liquid crystals, polymers, and etc.

253.

QUANTUM MECHANICALLY DERIVED QSARS FOR REVERSIBLE INHIBITORS OF ACETYLCHOLINESTERASE.

Jason A. Morrill, Weapons and Materials Research Directorate, United States Army Research Laboratory, AMSRL-WM-BD, Aberdeen Proving Ground, MD 21005-5066, Fax: 410-306-1909, jmmorrill@arl.army.mil, J. B. Wright, Natick Soldier Center, Materials Science Team, U. S. Army Soldier and Biological Chemical Command, Margaret M. Hurley, Computational and Information Sciences Directorate, U.S. Army Research Laboratory, and William E. White, Edgewood Chemical and Biological Center

Investigations involving inhibition of acetylcholinesterase have been of key interest to military and emergency response personnel due to the role that it plays in nerve agent toxicity, which is known to occur through "irreversible" modification of SER200. Acetylcholinesterase inhibition also has implications for the treatment of Alzheimer's disease and has led to a number of therapeutics for reversible inhibition. In order to provide improved prophylactic measures for nerve agent exposure we are investigating reversible inhibitors of acetylcholinesterase through the development of Quantitative Structure-Activity Relationships (QSARs) derived on a semiempirical quantum mechanical basis. We are also using semiempirical methods to investigate the reactions leading to irreversible inhibition with the intent to provide a sound theoretical model of nerve agent toxicity as well as antidote efficacy. Results from the QSAR studies as well as reaction modeling will be discussed. This work represents part of a DoD HPCMO Challenge project.

254.

PREDICTION OF PROTEIN RETENTION TIMES IN ANION EXCHANGE CHROMATOGRAPHY SYSTEMS USING SUPPORT VECTOR MACHINE REGRESSION.

Minghu Song¹, Jinbo Bi², Curt M Breneman¹, Nihal Tugcu³, N Sukumar¹, Kristin P. Bennett², and Steve M. Cramer³. (1) Department of Chemistry, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180, Fax: 518 276-4887, songm@rpi.edu, (2) Department of Mathematics, Rensselaer Polytechnic Institute, (3) Department of Chemical Engineering, Rensselaer Polytechnic Institute

Quantitative Structure-Retention Relationship (QSRR) models are developed for predicting protein retention times in anion exchange chromatography. Constitu-

tional, topological, electrostatic and TAE (Transferable Atom Equivalent) electron-density-based descriptors are computed directly from the protein crystal structures. A novel algorithm involving Support Vector Machine (SVM) regression is used to obtain Quantitative Structure-Retention Relationship (QSRR) models with high predictive accuracy. To accomplish this, a two-step computational strategy was adopted. In the first step, a sparse linear SVM was utilized as a variable selection method in which the relative importance of selected descriptors is analyzed using a bootstrap starplot visualization approach. Subsequently, the selected features are used to produce nonlinear SVM models. In order to ensure robust predictions, a bootstrap aggregate technique - "bagging" - is used. After validation, these predictive models may be used as part of an automated virtual high-throughput screening (VHTS) process.

255.

ON THE EFFECTIVENESS OF BOOTSTRAP AGGREGATION FOR CONSTRUCTING NEURAL NETWORK ENSEMBLES FOR QSAR AND QSPR. *Dimitris K. Agrafiotis, Walter Cedeno, and Victor S. Lobanov, 3-Dimensional Pharmaceuticals, Inc, 665 Stockton Drive, Exton, PA 19341, Fax: 610-458-8249, dimitris@3dp.com*

Despite their growing popularity among neural network practitioners, ensemble methods have not been widely adopted in QSAR and QSPR. Neural networks are inherently unstable, in that small changes in the training set and/or training parameters can lead to large changes in their generalization performance. Recent research has shown that by capitalizing on the diversity of the individual models, ensemble techniques can minimize uncertainty and produce more stable and accurate predictors. In this work, we present a critical assessment of the most common ensemble technique known as bootstrap aggregation, or bagging, as applied to QSAR and QSPR. Although aggregation does offer definitive advantages, we demonstrate that bagging may not be the best possible choice, and that simpler techniques such as retraining with the full sample can often produce superior results. These findings are rationalized by decomposing the generalization error into a term that measures the average generalization performance of the individual networks, and a term that measures the diversity among them. For networks that are designed to resist over-fitting, the benefits of aggregation are clear but not overwhelming.

256.

OPTIMIZATION OF MDL SUBSTRUCTURE SEARCH KEYS FOR THE PREDICTION OF ACTIVITY AND TOXICITY. *Douglas R. Henry, Product Development, MDL Information Systems, Inc, 14600 Catalina St., San Leandro, CA 94577, Fax: 510-614-3616, dough@mdl.com, and Joseph L. Durant Jr., MDL Information Systems*

There has been much recent interest in the use of MDL substructure search keys as descriptors for the unsupervised and supervised prediction of biological activity. Properly designed and optimized keys can perform on a par with most other chemical structure descriptors that have been examined for predicting the type of therapeutic activity a compound possesses. The parallel problem of predicting the toxicity of structures, which is crucial to ADMET studies and lead development, has received much less attention. In this talk we describe the use of several chemometric techniques to address the problem of combined activity and toxicity prediction using keys as descriptors. We apply these techniques to the structural data mining of the MDL MDDR and Toxicity databases. We describe the results of these studies, and provide some recommendations for further research in the field.

257.

ASYMMETRIC CLUSTERING OF CHEMICAL DATASETS: AN INVESTIGATION. *John D. MacCuish, and Norah E. MacCuish, Mesa Analytics & Computing, LLC, 212 Corona St., Santa Fe, NM 87501, Fax: 509-472-8131, john.maccuish@mesaac.com, norah.maccuish@mesaac.com*

We investigate asymmetric clustering of compound data as a viable alternative to more commonly used algorithms in this area such as Wards, exclusion region grouping, and complete link clustering. We show that the asymmetric Tversky measure, more commonly applied to similarity searching in compound databases, can be used in an asymmetric clustering algorithm as an effective means to cluster compounds for template extraction without the size bias usually associated with more common clustering measures and methods. We also show how this measure's asymmetry can be used in a popular exclusion region grouping algorithm. We show the results of the combination of these

measures and algorithms with several chemical datasets. The algorithm development and results are incorporated in the Khoros software development platform.

258.

GENETIC ALGORITHMS FOR DEVELOPING STRUCTURE-ACTIVITY CORRELATIONS IN LARGE PHARMACEUTICAL AND OLFACTORY DATABASES.

Barry K. Lavine¹, Charles E. Davidson¹, Curt Breneman², and William Kaat². (1) Chemistry, Clarkson University, 8 Clarkson Ave., Potsdam, NY 13699-5810, Fax: 315-268-6610, bklab@clarkson.edu, (2) Department of Chemistry, Rensselaer Polytechnic Institute

We have developed and tested a genetic algorithm (GA) for pattern recognition, which identifies molecular descriptors that optimize the separation of the activity classes in a plot of the two or three largest principal components of the data. Furthermore, we have generalized the fitness function of the pattern recognition so that it can tackle problems in multivariate calibration and messy pattern recognition. For example, outliers and nonlinear relationships in data can be handled using a Kohonen neural network implemented in a toroidal configuration in lieu of principal component analysis. Transverse learning has been introduced into the GA by coupling the Hopkins statistic to the original fitness function of the pattern recognition GA. The Hopkins statistic searches for features that increase the clustering of the data whereas the original fitness function of the pattern recognition GA identifies feature subsets that create class separation. Scaling the Hopkins statistics using a sigmoid transfer function and applying an influence function to deweight observations with high leverage and thereby robustify the Hopkins statistic ensures that our modified Hopkins statistic is a meaningful metric to assess clustering. We will be able to explore the structure of a data set, for example, discover new classes, by simply tuning the relative contribution made by the Hopkins statistic and the original fitness function to the overall fitness score. For training sets with small amounts of labeled data and large amounts of unlabeled data, this approach is preferable, as our previous results have shown since the information in the unlabeled data is used by the fitness function to guide feature selection and prevent overfitting. The efficiency and efficacy of the pattern recognition GA has been demonstrated through development of structure-activity correlations in large pharmaceutical and olfactory databases.

259.

SOLVATION IN MICROHETEROGENEOUS FLUIDS. *J. Ilya Siepmann¹, Bin Chen², Collin D. Wick³, John M. Stubbs³, and Li Sun³. (1) Departments of Chemistry, Chemical Engineering and Materials Science, University of Minnesota, 207 Pleasant St. SE, Minneapolis, MN 55455, Fax: 612-626-7541, siepmann@chem.umn.edu, (2) CMM & Department of Chemistry, University of Pennsylvania, (3) Department of Chemistry, University of Minnesota*

Configurational-bias Monte Carlo simulations in the isobaric-isothermal and Gibbs ensembles are used to investigate (i) the partitioning of normal alkanes and primary alcohols between water, wet or dry 1-octanol and between water/methanol mixtures and n-hexadecane, (ii) the retention characteristics of alkane and alcohol solutes in gas-liquid chromatography using squalane, dioctylether, and polyethylene oxide stationary phases and in reversed-phase liquid chromatography, and (iii) the temperature dependence of solvation in water, 1-octanol, or n-octane phases.

260.

GOOD + BAD = BAD: FORWARD AND REVERSE FEP CALCULATIONS SHOULD NOT BE AVERAGED. *David A Kofke, and Nandou Lu, Department of Chemical Engineering, University at Buffalo, The State University of New York, 303 Furnas Hall, Buffalo, NY 14260-4200, Fax: 716-645-3822*

While it is well known that free-energy perturbation (FEP) calculations performed in "forward" and "reverse" directions differ systematically, the usual countermeasure of simply averaging the two results does not have merit. One of the averages is much more reliable than the other, and appreciation of this fact is needed to apply FEP calculations correctly and efficiently. We argue that entropy is a central quantity to consider when applying FEP methods: FEP should be applied only in a direction in which the entropy decreases, and multistage strategies must be formulated to ensure that the configurations important to the target form a subset of configurations sampled by the reference. We apply the analysis to generate heuristics that indicate if a FEP

calculation is being applied correctly and whether the result has converged; we also provide recommendations for selection of multistage intermediates that optimize the accuracy and precision of the overall result.

261.

PARALLEL TEMPERING MONTE CARLO SIMULATIONS OF $(\text{H}_2\text{O})_n$, $(\text{H}_2\text{O})_n^+$, AND $(\text{CO}_2)_n$ CLUSTERS. *K. D. Jordan, Department of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260, jordan@imap.pitt.edu*

The parallel tempering Monte Carlo method is used to study structural transformations in $(\text{H}_2\text{O})_n$, $n \leq 20$, $(\text{H}_2\text{O})_n^+$, $n=6,8$, and $(\text{CO}_2)_n$, $n=6, 8, 12$, clusters. Inherent structure populations are calculated as a function of cluster temperature and used to aid in characterizing the transformations. The $(\text{H}_2\text{O})_8^+$ cluster is found to be particularly intriguing, in that there are two pronounced transitions, one from cubic-like to ring structures and the other from ring to branched-chain structures. The potential energy surface of $(\text{H}_2\text{O})_{20}$ is sufficiently "rugged" that at the lower temperatures the parallel tempering Monte Carlo simulations, even when run for 10^9 Monte Carlo moves, fail to achieve equilibrium.

262.

SOLVATION THERMODYNAMICS OF SMALL IONIC WATER CLUSTERS RELEVANT TO NUCLEATION. *Shawn M. Kathmann, Gregory K. Schenter, and Bruce C. Garrett, Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, PO Box 999 MS:K1-96, Richland, WA 99352, shawn.kathmann@pnl.gov*

Observation has shown that vapor-to-liquid nucleation is extremely sensitive to trace contaminants. It is of fundamental importance to understand the molecular processes involved in nucleation phenomena as they are crucial to many scientific endeavors. Small aqueous atomic or molecular clusters provide the bridge between the vapor and liquid phases. Accurate Helmholtz free energies for small aqueous clusters are crucial to the theoretical description of nucleation kinetics. Any trace species that can reduce the activation barrier to nucleation reduces the degree of metastability the system can reach, thus giving rise to a greater nucleation rate. Recently, we have shown that *tenths of a kcal/mol* differences in the underlying atomic or molecular interactions can have a profound influence on the thermodynamics and kinetics of clusters relevant to nucleation. We will present calculations comparing the free energies of small water clusters both with and without ions and discuss the consequences on the overall nucleation rate.

Acknowledgement: DOE, Division of Environmental Science, Office of Biological and Environmental Research (SMK and BCG), and DOE, Division of Chemical Sciences, Office of Basic Energy Sciences (GKS).

263.

NOVEL PHASE BEHAVIOR OF WATER INSIDE NANOTUBE AND SLIT NANOPORE. *Xiao Cheng Zeng, Department of Chemistry, University of Nebraska-Lincoln, 536 Hamilton Hall, Lincoln, NE 68588, Fax: 402-472-9402, xzeng1@unl.edu, Kenichiro Koga, Department of Chemistry, Cornell University, G.T. Gao, Department of Chemistry, US Naval Academy, and Hideki Tanaka, Department of Chemistry, Okayama University*

Phase behavior of water confined into nanometer-sized hydrophobic pores was studied using molecular dynamics simulation. (1) For cylindrical pores solvated in water, it is found that the "phase behavior" is very sensitive to the diameter of the pore. Using carbon nanotubes with diameters ranging from 1.1 to 1.4 nm, the simulation suggests a first-order freezing-like transition to hexagonal and heptagonal ice nanotubes, and a continuous phase transformation into solid-like square or pentagonal ice nanotubes, as the temperature is lowered slowly. That the isotherms for a given diameter are found to be similar to those near the liquid-gas critical point of bulk fluids suggests possible existence of a phase boundary terminated by a critical point. (2) For slit pore immersed in water under high pressure, both polymorphic and polyamorphic phase transition were observed. The latter transition arises in the confined supercooled water, that is when crystallization to the bilayer ice is prohibited.

264.

HYDRATION FREE ENERGIES AND ENTROPIES OF WATER IN PROTEIN INTERIORS. *Steven Rick, and L. Renee Olano, Department of Chemistry, University of New Orleans, New Orleans, LA 70148, Fax: 504-280-6860, sricks@uno.edu*

In the interiors of nearly all proteins there are cavities large enough to contain one or more water molecules. Water in these positions have important influences on protein structure, function, and dynamics. Most often, these water molecules are detected in polar environments and less polar cavities are commonly taken to not contain water. Alternatively, these non-polar cavities may contain water which is too disordered to be resolved by most X-ray studies and some recent NMR and X-ray experiments have identified water molecules in hydrophobic cavities. We will present calculations of the free energy for the transfer of water molecules from the bulk liquid to both polar and non-polar cavities. Water in the less-polar environments may be stabilized by entropy and we will discuss the entropic contribution to the hydration free energies. In addition, we will describe how the polarizability of the bound water molecule influences its stability.

265.

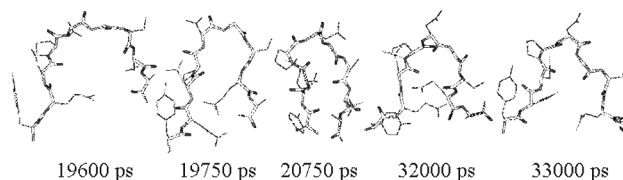
EFFECT OF ION BINDING ON THE CONFORMATIONAL DYNAMICS OF CALMODULIN. *Yuk Y Sham, Ruhong Zhou, Frank Suits, and Robert Germain, Computational Biology Center, IBM, IBM Thomas J Watson Research Center, P O Box 218 & Route 134, Yorktown Heights, NY 10598, Fax: 9149454104, ruhongz@us.ibm.com*

Nanoseconds molecular dynamics simulation were carried out to study the effect of binding on the conformation dynamics of calmodulin. It was found the removal of calcium ions results in the enhance flexibility of the calcium binding region. The flexibility of the commonly known flexible region (res 76-78) within the central helical linker between the N and C domain remains mostly unchanged between the apo and holo form. Removal of calcium ions also results in the collapse of the N-domain to a structure very similar to the experimentally observed "closed" conformation. These observations suggest a) the activation of calmodulin to bind to its target substrates involve stabilization of the "open" conformation and b) calcium binding does not affect the unfolding of the flexible region of the central helical linker, an essential requirement for calmodulin to attain a folded conformation.

266.

DIRECT OBSERVATION OF THE FOLDING AND UNFOLDING OF A B-HAIRPIN IN EXPLICIT WATER THROUGH COMPUTER SIMULATION. *Xiongwu Wu, Laboratory of Biophysical Chemistry, NHLBI, National Institutes of Health, Washington, MD 20892, Fax: 301-402-3404, WuXW@nhlbi.nih.gov, Shaomeng Wang, Department of Internal Medicine & Medicinal Chemistry, University of Michigan, and Bernard R. Brooks, Laboratory of Biophysical Chemistry, National Heart, Lung and Blood Institute, National Institutes of Health*

The cooperative folding and unfolding of a b-hairpin structure are observed in explicit water at native folding conditions through self-guided molecular dynamics simulation. The folded structure agrees excellently with the NMR NOE data. After going through a fully hydrated state, the peptide folds into a b-hairpin structure in a highly cooperative process. During the folding process it is observed that sidechain interaction occurs first, while intra-peptide hydrogen bonds only form at the final stage. On the contrary, the unfolding process starts with the breaking of inter-strand hydrogen bonds. Energetic analysis indicates that the driving force of the folding is the intra-peptide interaction, while the solvent interaction is against the folding.



267.

COMPUTATIONAL STRUCTURAL PROTEOMICS: GUIDING MUTAGENESIS EXPERIMENTS WITH RIGID BODY DYNAMICS AND B-SPLINES. *Richard E. Gillilan, Liru You, and Tim Huffaker, Molecular Biology and Genetics/MacCHESS, Cornell University, Biotechnology Bldg, Ithaca, NY 14853, reg8@cornell.edu*

A wide variety of cellular processes involve the docking of two or more proteins. Both the alpha and beta forms of tubulin interact with multiple protein partners and are part of a larger network of interactions. We examine the interaction of the tubulins with a chaperone protein rbl2 using a combination of quaternion-based rigid body mechanics, B-spline interpolation and the MMFF empirical forcefield. This grid-based approach allows for rapid evaluation of protein-protein forces with full summation of electrostatic terms (no cutoffs) smoothly joined to multipole expansions at extreme long range. Surface contact residues were extracted from docked models and used in two-hybrid mutagenesis experiments. Mutations were also subjected to a functional assay. Docking results show significant overlap with the experimental binding surface. The predicted structure is an unexpected, end-on rbl2 orientation that, in accordance with experimental observation, prevents alpha-tubulin binding due to its overlap with the alpha-beta interface.

268.

HIGH-THROUGHPUT DATA ANALYSIS. *David Rogers, SciTegic, Inc, 9665 Chesapeake Dr, Suite 401, San Diego, CA 92123, Fax: 858-279-8804, drogers@scitegic.com*

As the explosive growth in the amount of cheminformatics data continues, so does the desire for a "magic bullet", some software, method, or methodology that would unlock the value contained in these data streams. This talk will argue against such a solution, and instead propose that the solution will be a confluence of different techniques.

In particular, several new technologies have been appearing that, when tied together, form a framework for addressing the requirements of high-throughput data analysis. They include: novel internet technologies; the breakdown of the metaphor of the database-oriented "table" in favor of distributed, pipelined data; linear-scaling learning methods adapted to noisy assay results; novel structural descriptors that capture the high-dimensional nature of structural information. This talk will tie these threads together as a framework for high-throughput data analysis.

269.

COMPUTATIONAL DESIGN OF BIOMIMETIC MATERIALS. *Bin Chen¹, Robert J. Doerksen¹, and Michael L. Klein². (1) CMM & Department of Chemistry, University of Pennsylvania, 231 S. 34 Street, Philadelphia, PA 19104, Fax: 215-573-6233, binchen@cmm.chem.upenn.edu, (2) Department of Chemistry, University of Pennsylvania*

The design of polymers that mimic the complex structures and functions of bio-molecules is of paramount importance with both fundamental and practical implications. However, this effort is often limited by our knowledge of the microscopic-level driving forces behind certain structures and properties. Computations allow for direct probing of this information. Moreover, given a primary sequence for a polymer, they can be utilized to predict its conformational properties. This presentation will report some recent progress in using computations to guide the design of biomimetic antimicrobial material. Specifically, the conformational properties of proposed polymer backbones were examined with different levels of computational approaches before they were synthesized. First, density functional theory calculations are employed to search for stable conformations of selected polymer fragments and to determine the torsional potentials around various bonds on the polymer backbone. The torsional potentials are then applied to predict the conformational properties of the polymers using classical simulations. These calculations also yield useful insights on how various structural motifs (e.g., conjugation, hydrogen bonding, and side chain substitution) enforce desired conformation. Available experimental results show that the resulting polymers have promising antibacterial activities.

270.

STRUCTURAL BASIS FOR THE RESISTANCE OF OLIGONUCLEOTIDES CONTAINING NNK-INDUCED O⁶-[4-OXO-4-(3-PYRIDYL)BUTYL]GUANINE ADDUCTS TO 3'→5' EXONUCLEASES. *Mahadevan Seetharaman¹, Alexis Ogdie², Soobong Park², David M. Ferguson¹, and Natalia Tretyakova³. (1) Department of Medicinal Chemistry, University of Minnesota, 8-101 WDH, 308, Harvard St SE, Minneapolis, MN 55455, maha@vwl.medc.umn.edu, (2) University of Minnesota Cancer Center, University of Minnesota, (3) Department of Medicinal Chemistry and University of Minnesota Cancer Center, University of Minnesota*

Tobacco-specific nitrosamine, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK), is a chemical carcinogen thought to be involved in the initiation of lung cancer. Previously we had shown O⁶-[4-oxo-4-(3-pyridyl)butyl]guanine (O⁶-POB-dG), a pyridyloxobutylated promutagenic DNA adduct derived from NNK, to block snake venom phosphodiesterase, a 3'-exonuclease. In contrast, the 5'-exonuclease from bovine spleen was capable of digesting past the adduct. So, we had extended our study to two other 3'-exonucleases, one from T4 DNA polymerase and the other from *E. Coli* DNA polymerase. *E. Coli* Pol I is capable of digesting DNA containing O⁶-POB-dG, while the 3'-exonuclease activity of T4 is hindered, but not completely blocked, since DNA cleavage past the adduct can be achieved after prolonged incubation with excess enzyme. Here, we present results of molecular modeling of O⁶-POB-dG containing trinucleotide docked into the 3'-exonuclease domains of *E. Coli* Pol I and T4 DNA polymerase. Our docking studies indicate that the side chains of two amino acids in T4 DNA Polymerase, Phe 120 and Lys 119, may block the entrance of the O⁶-POB-dG-containing DNA into the active site of the enzyme. No such residues are present in the 3'-exonuclease active site of *E. Coli* Pol I, consistent with the ability of this enzyme to digest pyridyloxobutylated DNA. We propose that the observed differential effects of O⁶-POB-dG on various 3'-exonuclease enzymes are the result of the orientation of pyridyloxobutyl group relative to the geometry of the exonuclease active site.

271.

ADAPTIVE WAVELET THRESHOLDING FOR DENOISING DNA MICROARRAYS. *Christian S. Uehara, AnnMarie H. Spexet, and Ioannis A. Kakadiaris, Department of Chemical Engineering, University of Houston, 4800 Calhoun Ave., Houston, TX 77204-4792, csuehara@swbell.net*

DNA microarrays are a common, high-throughput technique for monitoring gene expression in biological systems. With the current state of genomics and proteomics, the study and description of protein interaction maps in organisms has become increasingly feasible. However, analysis of microarray experiments often suffers from significant noise and geometrical problems, due to the laboratory techniques and the imaging methods involved. These problems often curtail the use of fully automated analysis and can hinder the extraction of useful information.

This paper discusses a method for denoising array images that is particularly suited for application to high-throughput methods. Differences in suitability based on the stage of analysis and the noise structure are investigated. Adaptive wavelet thresholding is investigated for a range of acquired images as is the suitability of a new family of analytical functions. Ultimately, the goal is to aid automated analysis while providing image data quality statistics and image data extraction.

272.

SUMMATION APPROXIMANTS: AN EASY WAY TO IMPROVE THE ACCURACY OF AB INITIO QUANTUM CHEMISTRY. *David Z. Goodson, Department of Chemistry and Biochemistry, University of Massachusetts at Dartmouth, 285 Old Westport Rd, North Dartmouth, MA 02747, Fax: 508-999-9167, dgoodson@umassd.edu*

The accuracy of quantum chemistry calculations with high-level treatment of electron correlation, in particular, fourth-order Moeller-Plesset perturbation theory and the CCSD(T) coupled-cluster theory, can be significantly improved with little or no increase in computational cost by using summation approximants that model the way the underlying theory converges toward the full configuration-interaction limit. Recently developed summation methods will be described and examples of calculations of molecular properties will be presented to illustrate the amount of improvement that can be expected.

273.

EFFICIENT PARALLELIZATION OF SCF AND DFT ANALYTIC SECOND DERIVATIVES. *Theresa L. Windus¹, Edoardo Apra¹, Robert J. Harrison², and Benny G. Johnson³.* (1) High Performance Computational Chemistry Group, Pacific Northwest National Laboratory, 902 Battelle Boulevard, P.O. Box 999, MSIN: K1-96, Richland, WA 99352, Fax: 509-375-6631, theresa.windus@pnl.gov, (2) EMSL, PNNL, (3) Quantum Simulations, Inc

An efficient parallelization of the SCF and DFT analytic second derivatives with respect to nuclear coordinates in the massively parallel computational chemistry code, NWChem, will be presented. The presentation will cover implementation issues such as optimization of the multiple Fock builds and distribution of data. Scalability of the implementation will be shown through several examples on several different platforms.

274.

INTERMOLECULAR AND INTRAMOLECULAR ENERGY TRANSFER IN NE-ATOM COLLISIONS WITH A N-HEXYLTHIOLATE SELF-ASSEMBLED MONOLAYER (SAM). *Tianying Yan, and William L. Hase, Department of Chemistry, Institute for Scientific Computing, Wayne State University, Detroit, MI 48202, tyan@chem.wayne.edu*

In computer simulations of Ne-atoms colliding with an n-hexylthiolate self-assembled monolayer (SAM), a Boltzmann component occurs in the final translational energy distribution $P(E_f)$ of the scattered atoms. It has been shown that the vast majority of the Ne + SAM/Au(111) collisions are direct and trapping desorption is unimportant. (Yan, T.-Y.; Hase, W. L.; Barker, J. R. *Chem. Phys. Lett.* **2000**, *329*, 84.) The incident Ne-atom does not have enough time to "feel" the surface temperature and bounces off with a translational energy distribution which depends on a combination of different dynamical processes. Energy transfer occurs by both exciting vibrations of the SAM and inducing conformational changes of the SAM's alkyl chains. The rate of intramolecular vibrational energy redistribution (IVR) among the surface vibrational modes, which may contribute to Boltzmann component in $P(E_f)$, is also investigated.

275.

COMPUTER-AIDED CONSTRUCTION OF PRESSURE-DEPENDENT REACTION NETWORKS FOR LARGE, GAS-PHASE CHEMICAL MECHANISMS. *David M. Matheu, Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave. Rm 66-270, Cambridge, MA 02139, Fax: 617-252-1651, dmmatheu@mit.edu, William H. Green Jr., Department of Chemical Engineering, MIT, and Jeffrey M. Grenda, Corporate Strategic Research, ExxonMobil Research and Engineering*

A host of vital, current, and developing technologies, including new engine designs and partial oxidation processes, involve complex, gas-phase chemical mechanisms, with hundreds of species and thousands of reactions. The size and complexity of the chemical models needed for these processes can make them almost impossible to construct by hand. Chemists and engineers have thus turned to artificial intelligence software tools to build these large mechanisms automatically. Such tools, however, have (until now) been unable to handle the crucial effects of pressure-dependence — pressure-dependent reactions exist in sub-networks that thwart simple inclusion in large chemical mechanisms. We answer this challenge with a software tool which can build and add pressure-dependent reactions "from scratch". Our algorithm constructs truncated pressure-dependent networks, as needed, within a large chemical mechanism. It includes all important pathways, and safely discards unimportant ones, providing the chemist with an appropriate chemical model for the conditions of interest.

276.

FUNCTIONALITY MAPPING AS A METHOD TO IDENTIFY STEREOCONTROL ELEMENTS FOR ASYMMETRIC REACTIONS. *Manoranjan Panda, and Marisa C. Kozlowski, Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, PA 19104, Fax: (215) 573-7165, mpanda@sas.upenn.edu*

A computational method to determine the energetically favorable positions of functional groups with respect to the transition states of stereoselective reactions based on force field energy minimization will be presented. The parameters of this functionality mapping, the characteristics of the target transition states, and the features of the probe structures will be outlined. Our method was found to reproduce the positions of the stereodiscriminating

fragments for a series known chiral ligands including the Masamune dimethylborolane, dimethylborane, the Corey stein reagent, the Roush allylboronate tartrates, and the secondary amine Diels-Alder catalysts described by MacMillan. Functionality mapping can be used to better understand the specific interactions of ligand functional groups in the transition states leading to the products. The method can determine if a chiral ligand imparts the observed selectivity by stabilizing one reaction pathway, by destabilizing a reaction pathway, or by a combination of both. Orientational as well as positional information about potential functional groups is readily obtained. In addition to its utility as an analytical tool, the functionality mapping can be used to explore starting points for the design of new chiral ligands.

277.

AN ALGORITHM FOR CONFORMATIONAL SAMPLING USING DISTANCE CONSTRAINTS. *Huafeng Xu, and Dimitris Agrafiotis, Research Informatics, 3-Dimensional Pharmaceuticals, Inc, 8 Clarke Drive, Cranbury, NJ 08512, Fax: 609-655-6930, hxu@3dp.com*

We present an algorithm for conformational sampling based on distance constraints derived from the connection table of a molecule. Unlike conventional distance geometry approaches, where an expensive triangular smoothing procedure is usually required before the conformations can be generated, ours capitalizes on the fact that the full distance matrix is overdetermined, and a small fraction of the distances, together with reasonable upper and lower bounds, suffice to generate reasonable conformations. Our algorithm can generate very rapidly a large number of distinct conformations. These conformations can be subsequently refined by gradient minimization techniques to select the ones with the lowest energies.

278.

MACHINE LEARNING MODELS FOR HIGHLY-MULTIDIMENSIONAL MOLECULAR DESCRIPTORS. *Jessen Yu, and William Mydlowec, Pharmix Corp, 200 Twin Dolphin Drive, Suite F, Redwood Shores, CA 94065, jyu@pharmix.com*

Highly-multidimensional molecular descriptors can present difficulties in encoding, construction, and performance for machine learning models. We describe several models that are well-suited for multimillion bit molecular descriptors. Support Vector Machines (SVM), originally developed for character and speech recognition, demonstrate excellent generalization performance on nonlinear data, even when the size of the training set is significantly smaller than the dimensionality of the problem. Binary Decision Diagrams (BDD), used in synthesis and verification of sequential logic circuits, can compactly encode huge sets and can be solved for satisfiability in constant time. Maximum-entropy models, originally developed for digital signal processing, can effectively extract information from datasets that are very biased or noisy. We compare the effectiveness of these models using 3-D pharmacophore keys (~10 million bits) on several lead discovery problems involving current drug targets.

279.

SARSTEP: A NOVEL CONCEPTUAL FRAMEWORK FOR QSAR. *Lee W. Herman¹, David Clemens², and Andrei Caracoti¹.* (1) Computational Chemistry, Millennium Pharmaceuticals, Inc, 75 Sidney St., Cambridge, MA 02139, Fax: 617-551-8911, herman@mpi.com, (2) ChemInformatics, Millennium Pharmaceuticals, Inc

SARstep, a novel conceptual framework for QSAR, is described.

For the last forty years, the concept of QSAR has been framed within the concept of *state*. The implication for this premise is that there is some biological *state* (activity) that may be implied from any given chemical *state* (structure) via some kind of mathematical representation (linear equation, neural net, decision tree, KNN model, etc. . . .). An alternate strategy is to represent the structure-activity landscape as the sum of all pair-wise changes in biology with respect to chemistry (*Hsteps*).

We will present an overview of the SARstep concept, report on the implementation, and highlight the application of the methodology to the lead optimization process.

280.

FROM DECISION TREE TO DECISION FOREST – A NOVEL CHEMOMETRICS APPROACH FOR QSAR STUDY. *Weida Tong, Huixiao Hong, Hong Fang, Qian Xie, and Roger Perkins, Bioinformatics and Computational Science, Northrup Grumman, 3900 nctr road, Jefferson, AR 72079, Fax: 870 543 7382, wtong@nctr.fda.gov*

Forecast combination techniques are known to often yield better predictions than any one of many prediction models that might be combined. In earlier chemometric applications, the multiple models to be combined have been developed by altering the training set in ways to maximize the difference between individual models. Multiply sampling the same pool of data to obtain multiple training sets, however, poses the risk of over fitting the noise in the data, possibly resulting in poorer prediction of the composite model than the individual models. We suggest a novel approach we have named Decision Forest that is not subject to the noise over fitting problem. The Decision Forest approach combines several decision tree models that are each developed using a unique set of descriptors (attributes). When models of similar predictive quality are combined using the Decision Forest method, quality compared to the individual models was consistently and significantly improved, as well as prediction of external test sets. An example will be presented for prediction of binding affinity of 230 chemicals to the estrogen nuclear receptor.

281.

EARLY PREDICTION OF BIOAVAILABILITY BY ADAPTIVE FUZZY PARTITIONING (AFP). *Marco Pintore¹, Nadège Piclin¹, Han van de Waterbeemd², and Jacques R. Chretien¹.* (1) *BioChemics Consulting, Centre d'Innovation, 16, rue Leonard de Vinci, Orleans cedex 2 45074, France, Fax: 33-238417221, marco.pintore@univ-orleans.fr, jacques.chretien@univ-orleans.fr,* (2) *PDM, Department of Drug Metabolism, Pfizer Global Research and Development*

Early ADMET is among the up to date challenges. An Adaptive Fuzzy Partition (AFP) algorithm, based on Fuzzy Logic, was applied on a bioavailability data set, including 272 compounds and subdivided into four ranges of activity. The AFP method consists in modeling molecular descriptor-activity relationships by dynamically dividing the descriptor hyperspace into a set of fuzzy subspaces. A large set of molecular descriptors was tested and the most relevant parameters were selected with help of an innovative procedure based on genetic algorithm concepts and stepwise method. After building several AFP models on a training set, the best ones were able to predict correctly 75% of the test set compounds. Furthermore, an improvement of about 10-15% in the validation results was got as regard to other prediction methods from literature. Furthermore, the prediction power was increased up to 25% employing a data set with a better-optimized molecular diversity.

282.

ON THE EFFECTIVENESS OF BOOTSTRAP AGGREGATION FOR CONSTRUCTING NEURAL NETWORK ENSEMBLES FOR QSAR AND QSPR. *Dimitris K. Agrafiotis, Sergei Izrailev, and Walter Cedeno, 3-Dimensional Pharmaceuticals, Inc, 665 Stockton Drive, Exton, PA 19341, Fax: 610-458-8249, dimitris@3dp.com*

Feature selection continues to be the method of choice for minimizing the risk of overfitting in structure-activity and structure-property correlation. Here we present artificial ant colony systems and particle swarms, two novel population-based optimization paradigms inspired by the behavior of insect colonies. Artificial ant colonies mimic the mechanism that allows real ants to find the shortest path between a food source and their nest using deposits of pheromone as a communication agent. Particle swarms explore the search space through a population of individuals, which adapt by returning stochastically towards previously successful regions, influenced by the success of their neighbors. Using several examples from the QSAR literature, we show that these methods compare favorably to more established optimization techniques such as simulated annealing, and are able to identify a better and more diverse set of solutions given the same amount of simulation time.

283.

NOVEL METHOD OF AB INITIO EVALUATION OF THE PSEUDO JAHN-TELLER EFFECT. *Isaac B. Bersuker, Nikolai Balabanov, David Pekker, and James E Boggs, Institute for Theoretical Chemistry, Department of Chemistry and Biochemistry, University of Texas, Austin, TX 78712, Fax: 512-471-8696, bersuker@eeyore.cm.utexas.edu*

Ab initio calculations of the instability of high symmetry molecular configurations which occurs due to (and only to) the negative pseudo Jahn-Teller contribution (K_v) of the excited states to the curvature of the adiabatic potential energy surface encounters significant difficulties because of the lack of sufficiently accurate data on excited states, and the singularities in the positive nonvibronic contribution (K_0) to the curvature $K=K_0+K_v$. We worked out a method that overcomes both these difficulties by (1) reducing the sum of excited state contributions to expressions with wavefunction derivatives, and (2) excluding the singularities in K_0 by means of a transformation that cancels the intraatomic contributions which do not affect the molecular K value. Formulas are derived for K_0 and K_v at different levels of theory. Numerical calculations for several series of compounds allow for explicit evaluation of the origin of instabilities and rationalization of data on molecular geometry.

284.

β -X ELIMINATION IN $Ta(H)_2(OH)_3$: A DENSITY FUNCTIONAL THEORY STUDY. *Christopher D. Taylor, and Thomas R. Cundari, Department of Chemistry, University of Memphis, 3744 Walker, Memphis, TN 38152, Fax: 901-678-3447, cdtaylor2@memphis.edu*

The β -X elimination reaction is a significant hurdle for efficient and direct synthesis of functionalized polymers by Ziegler-Natta catalysis. The $H_2C=CHX$ insertion and β -X elimination (to form ethylene) reactions are studied for the model catalyst $Ta(H)_2(OH)_3$ for a series of vinyl-halides and -ethers using density functional theory. Increasingly realistic models are addressed using QM/MM techniques. The conclusions drawn regarding the relative rates of insertion and elimination, as well as the mechanisms by which these reactions occur are interpreted. Both reactions involve a four-centered transition state, and there is a lower barrier to halide β -X elimination as opposed to the ether moieties. Conversely, there is a higher barrier to vinyl-halide insertion than for vinyl-ethers. Comparison has been made to experiment. While there is close agreement regarding trends, and the calculations add insight not available from experiment, there are some discrepancies regarding olefin-binding and catalyst geometries.

285.

DFT AND ONIOM MODELING OF THE SUZUKI CROSS-COUPLED REACTION, Pd VS. Ni. *Yong Zhao, Department of Chemistry, University of Memphis, Univ. of Memphis, Memphis, TN 38152, yongzhao@memphis.edu, and Thomas Cundari, Department of Chemistry, CROMIUM, University of Memphis*

The Suzuki coupling reaction is one of the fundamental organic synthetic transformations for the formation of carbon-carbon bonds. The process works well in the presence of a wide range of functional groups. This research focuses on modeling of the catalytic cycle, including oxidation addition, transmetalation and reductive elimination. Effects of different aryl (Br, I, Triflate, Cl) and ligand substituents (PH₃, P(t-Bu)₃ and P(o-tolyl)₃) have been explored. Pd has been almost exclusively used in Suzuki reaction. The possibility of using cheaper Ni as a replacement for Pd is considered and computed. In this study, the effective core potentials (ECPs) of Stevens et al. have been employed in conjunction with the B3LYP hybrid functional. For bulky phosphine ligands ONIOM (B3LYP/CEP-31G:UFF) is applied to the model.

286.

COMPUTATIONAL STUDY OF SOLVOLYSIS REACTIONS USING THE EXPLICIT SOLVENT MODEL. *Jiri Mareda, Jean-Francois Fuchs, and Michele Rizzello, Department of Organic Chemistry, University of Geneva, 30, Quai Ernest Ansermet, Geneva 1211, Switzerland, Fax: 41-22-329-6102, jiri.mareda@chiorg.unige.ch*

The reactivity of bridgehead halides exemplified by bicyclo[n.1.1] heptyl substrates has been posing an intriguing problem to chemists. These halides solvolyze much more rapidly than would be expected on the basis of the strain accumulated in the corresponding cations. In particular, 1-bicyclo[3.1.1]heptyl derivatives show extraordinary reactivity even under mild solvolytic conditions,

leading to a mixture of both rearranged and un-rearranged products. In order to investigate the mechanism of these reactions by *ab initio* methods we propose the model where the effects of solvent on cationic rearrangements are modeled by explicit definition of one, three, and four waters. The combined strategy of the discrete-continuum model was also adopted. Density functional theory and MP2 methods, employing several basis sets, were used to investigate the potential energy surfaces for the heterolytic C-O bond cleavage of protonated 1-bicyclo[3.1.1]heptanol. The computations indicate, that the initial reaction intermediate of the studied processes is the water complex of unsymmetrical bicyclobutonium cation for which the integrity of bicyclo[3.1.1]heptyl structure is still preserved. The solvent-cation complexes containing up to four solvent molecules were located on the potential energy surfaces relating to these reactions. The stabilization of the cationic moieties by solvent molecules together with the bonding properties for these complexes will also be discussed. The relative stabilities of the solvent-cation complexes will be compared with results computed on the C7H11+ surface in vacuo.

287.

INTERACTION MODEL FOR FUNCTIONAL GROUPS USING PROPERTIES FOUND IN QUANTUM CALCULATIONS. *Steven G. Arturo, and Dana E. Knox, Department of Chemical Engineering, New Jersey Institute of Technology, University Heights, Newark, NJ 07102-1982, sga5892@njit.edu*

A model for use in group-contribution methods that describes the interactions of functional groups has been developed. Functional groups are investigated for several n-alkanes, alcohols and water. For each molecule, group properties are determined using quantum calculations at the MP2 level and with Bader's Theory of Atoms in Molecules (AIM). Definable properties within the AIM theory that are evaluated include group charge, dipole moment, polarizability and volume. These properties are used in a binary interaction energy model that gives the average interaction energy as a function of the intergroup distance. It is constructed using classical attraction terms and repulsion due to quantum effects evolving from the overlap between the interacting groups' wavefunctions. The different property values for the same group in different molecules translate into differences among the interaction energy functions involving that group. Hydrogen bonding energies and distances are predicted for the hydroxyl/methyl, hydroxyl/methylene, and hydroxyl/water intergroup interactions.

288.

PREDICTIVE MODELING OF CHEMICAL REACTIONS. *L.J. Soltzberg, Nancy E. Lee, Ayako Honda, Michelle Sanford, and Mojisola Sekoni, Department of Chemistry, Simmons College, 300 The Fenway, Boston, MA 02115, lsoltzberg@simmons.edu*

We have used off the shelf software (Spartan Pro) to model the outcomes of a variety of chemical reactions. Geometries of the reactant molecules are separately optimized, and the molecules are given an initial orientation that might correspond to a reactive collision. Geometry optimization is then run on the assembly. Application of this procedure with a low level *ab initio* Hamiltonian has replicated the correct products and relative reactivities for the reduction of hydroquinone by various alkylammonia borane complexes. With a semi-empirical Hamiltonian, application of this method to the oxazaborolidine-catalyzed reduction of diverse ketones gives not only the correct products but reproduces the enantioselectivity of these reductions when carried out with a chiral catalyst.

289.

POLARIZATION AND DYNAMICS OF WATER AND SCHIFF BASE WITHIN BACTERIORHODOPSIN. *I-Feng W. Kuo¹, Christopher J. Mundy², and Douglas J. Tobias¹.* (1) Department of Chemistry, University of California, Irvine, Irvine, CA 92697-2025, ikuo@uci.edu, (2) Chemistry and Materials Science, Lawrence Livermore National Laboratory

Bacteriorhodopsin (bR) is a prototypical light-driven proton pump which transports a proton across the cell membrane through a photocycle with well

determined intermediates. From experimental as well as theoretical studies, the existence of an extensive hydrogen bonded network composed of key polar residues as well as discrete water molecules inside the transmembrane helices are believed to play vital roles in the functionality of the protein. To gain further insight into this unique and polar region within bR, we used *ab initio* (Car-Parrinello) molecular dynamics where the forces are derived from the electronic structures which were solved through Density Functional Theory with a gradient corrected B-LYP exchange-correlation functional. To make the calculation tractable, a cluster model based on crystallographic data was used. This cluster incorporates polar residues and water around the Schiff base which forms the extensive hydrogen bonded network leading to the extracellular domain. With the explicit calculation of the electronic structure, the polarization and dynamics of water molecules as well as the protonated Schiff base were analyzed in detail for the ground state of bR. Furthermore, vibrational modes are computed and compared to available experimental data.

290.

THEORETICAL INVESTIGATION OF THE STRUCTURE AND VIBRATIONAL FREQUENCIES OF HOMOGENEOUS AND HETEROGENEOUS CLUSTERS OF HPLC SOLVENTS: WATER, METHANOL, AND ACETONITRILE. *John M. Craig, Donald D. Shillady, Sarah C. Rutan, Sally S. Hunnicutt, and Ernst Bezemer, Department of Chemistry, Virginia Commonwealth University, P. O. Box 842006, 1001 W. Main Street, Richmond, VA 23284-2006, hiker@cheerful.com*

Water, methanol, and acetonitrile are common solvents used in HPLC separations. It is highly desirable to model the interaction of these solvents in order to better understand retention and the nature of solvation. Therefore, clusters of these solvent molecules have been studied from a theoretical perspective. Specifically, cluster structures have been optimized at the RHF and MP2 levels, and the effects on geometry and vibrational frequency have been evaluated. The total energies and binding energies have been computed, and the trends are reported as a function of cluster composition. Harmonic frequencies have been calculated and corrected for the effects of anharmonicity by approximating the normal potential using fourth-order vibrational displacements in combination with first-order perturbation theory to fit a Morse potential to each normal mode. A novel self-adjusting method is used to account for the depth or shallowness of each potential, resulting in frequency data with improved accuracy.

291.

INTERMOLECULAR POTENTIALS FOR GAS HYDRATES OBTAINED FROM *AB INITIO* QUANTUM MECHANICS. *Jeffery B. Klauda, and Stanley I. Sandler, Department of Chemical Engineering, University of Delaware, Colburn Lab, Newark, DE 19716, Fax: 302-831-1048, klauda@che.udel.edu*

A systematic method for calculating the interaction energies between various guests and hydrate cages is calculated using *ab initio* quantum mechanics (QM). Interactions between a single guest and symmetric pieces of a cavity are calculated at MP2/6-31+g(3d,3p) for various guest positions and orientations. Then a Lennard-Jones (LJ) potential with an electrostatic term is fit to these QM energies to predict guest occupancies in the hydrate cavities. This differs from the commonly used van der Waals and Platteeuw (vdWP) hydrate model in which the intermolecular potential parameters are fit to hydrate equilibrium pressures of both single and multiple guest hydrates. The procedure here is seen to improve upon the vdWP model with Sloan's Kihara potential parameters when compared to Raman and NMR spectroscopy for guest compositions in the hydrate cavities.