

**AAAS, Pacific Division  
2015 San Francisco  
Symposium Abstracts  
17 Theory, Experiment, and Computations**

*Sodium Diffusion in Type II Silicon Clathrates*, **JASON G. SLINGSBY<sup>1</sup>, NICHOLAS A. RORRER<sup>1</sup>, LAKSHMI KRISHNA<sup>2</sup>, ERIC S. TOBERER<sup>2</sup>, CAROLYN A. KOH<sup>1</sup>, and C. MARK MAUPIN<sup>1</sup>** (<sup>1</sup>Chemical and Biological Engineering Department, Colorado School of Mines, 1500 Illinois Street, Golden, CO 80401; <sup>2</sup>Physics Department, Colorado School of Mines, 1500 Illinois Street, Golden, CO 80401; [cmmaupin@mines.edu](mailto:cmmaupin@mines.edu)).

Earth abundant semiconducting type I and II silicon clathrates have attracted significant attention as photovoltaic materials due to their wide band gaps. To realize the semiconducting properties of these materials guest species, present during the synthesis process, must often times be completely evacuated from the host cage structure post synthesis. A common guest species utilized in the synthesis of silicon clathrates is sodium, which templates the clathrate cage formation. Previous experimental investigations have identified that it is possible to evacuate sodium from type II clathrates to an occupancy of less than 1 sodium per unit cell, but any significant sodium removal from type I has not been achieved. This work investigates the energetics, kinetics, and resulting mechanism of sodium diffusion through type II silicon clathrates by means of biased molecular dynamics and kinetic Monte Carlo simulations. Well-tempered metadynamics has been used to determine the potential of mean force for sodium moving between clathrate cages, from which the thermodynamic preferences and transition barrier heights have been obtained. Kinetic Monte Carlo simulations based on the metadynamics results have identified the mechanism of sodium diffusion and eventual evacuation from type II silicon clathrates. The overall mechanism consists of a coupled diffusive process, where the large occupied hexakaidehedral cages empty their sodium guests to adjacent empty large cages. This migration of sodium guests changes the local environment around the occupied small pentagonal dodecahedral cages increasing the ability of sodium guests to leave the small cages. This coupled process continues through the critical cross-over point that is identified as the point where large and small cages are equally occupied by sodium guests. Further sodium removal results in the majority of sodium guests residing in the large cages as opposed to the small cages, in agreement with experiment, and ultimately a sodium free structure

*Exploring the Limits of Computational Drug Screening*, **OWEN M. MCDOUGAL** (Department of Chemistry and Biochemistry, Boise State University, 1910 University Drive, Boise, ID 83725; [owenmcdougal@boisestate.edu](mailto:owenmcdougal@boisestate.edu)).

$\alpha$  Conotoxins are small, cysteine rich peptides that demonstrate an uncanny ability to differentiate between biomacromolecular receptors in mammals. Alpha-conotoxins are particularly potent and selective for heteropentameric nicotinic acetylcholine receptors in neurons. Understanding how these peptides act is an amazingly complex problem from a molecular biology, chemistry, and computational chemistry perspective. Here we discuss the limitations of sampling the ligand binding space using the computational screening program AutoDock, and propose workflow solutions to overcome common pitfalls. The goal of this research is to understand the role played by the binding of alpha-conotoxin MII on neuronal acetylcholine receptors, and provide insights into how this information may lead to the development of therapeutic drugs for Parkinson's disease.

*Automated Software Development and Optimization*, **VIKRANT SHARMA** (Jesuit High School, 5508 NW 131<sup>st</sup> Ave Portland OR 97229; [sindpower@hotmail.com](mailto:sindpower@hotmail.com)).

Despite significant advances in mathematics, the process of developing and debugging software code has not fundamentally evolved over the same period. This paper proposes, demonstrates and verifies the automated development of software code utilizing a novel approach of modifying and combining genetic algorithms and traditional optimization methods. Beginning with a core randomized seed code, off-springs are generated by a recursive process while satisfying a dynamic objective or forcing function, developing automatically a new set of software code within the subset of a generation. At the same time, the objective function evolves in sync with the offspring code, thus creating the next generation of software code. The process in combination repeats recursively until desired and optimized software is produced. This paper demonstrates the development of the code, discusses the novel method along with its strengths and limitations, and compares this innovative and revolutionary approach to the current art.

**Santrupti Nerli**, *STUDENT* Department of Computer Science, San Jose State University, [santrupti.nerli@sjsu.edu](mailto:santrupti.nerli@sjsu.edu)

*Using Statistical Models to Predict Splice Sites and Cryptic Splice Sites*, **SANTRUPTI NERLI\***, and **SAMI KHURI** (Department of Computer Science, San Jose State University, One Washington Square, San Jose, CA 95192; santrupti.nerli@sjsu.edu).

Alternative splicing events are the most crucial events responsible for the creation of normal proteins. Identification of the splice sites that govern the alternative splicing events has become very critical. In this work, we developed three statistical models to identify splice sites and cryptic splice sites in the human genome. The three models we developed are: (1) Position Weight Matrix (PWM) (2) Maximal Dependence Decomposition (MDD) and (3) Hidden Markov Model (HMM). Our models are evaluated by performing 10-fold cross validation of 5' and 3' authentic splice site data obtained from Homo Sapiens Splice Site Database (HS3D). We also performed classification of 5' and 3' cryptic splice sites for the breast cancer type 1 susceptibility protein (BRCA1) and the beta-globin gene (HBB), which we collected from the literature. We observed that PWM and HMM successfully predict splice sites and cryptic splice sites with good accuracy. However, MDD outperforms them for all our datasets with the prediction accuracy that is represented by the area under Receiver Operating Characteristic (ROC) Curve of 0.96 and 0.93 for 5' and 3' splice site data, 0.99 and 0.95 for 5' and 3' BRCA1 cryptic splice site data and 0.89 and 1.0 for 5' and 3' HBB cryptic splice site data. Thus, our results show that the predictor we developed is pretty accurate. Subsequently, we are working on remodeling the HMM to accommodate variable length sequences since the current model is trained on fixed length splice sites.

*The Impact of Active Site Protonation on Substrate Ring Conformation in Melanocarpus albomyces Cellobiohydrolase Cel7B*, **TIMOTHY C. SCHUTT**, **VIVEK S. BHARADWAJ**, **DAVID M. GRANUM**, and **C. MARK MAUPIN** (Chemical and Biological Engineering Department, Colorado School of Mines, 1500 Illinois Street, Golden, Colorado, 80401, tschutt@mymail.mines.edu).

The ability to utilize biomass as a feedstock for liquid fuel and value-added chemicals is dependent on the efficient and economic utilization of lignin, hemicellulose, and cellulose. In current bioreactors, cellulases are used to convert crystalline and amorphous cellulose to smaller oligomers and eventually glucose by means of cellulase enzymes. A critical component of the enzyme catalyzed hydrolysis reaction is the degree to which the enzyme can facilitate substrate ring deformation from the chair to a more catalytically active conformation (e.g. skewed boat) at the -1 subsite. It is found that the protonation state of critical active site residues (i.e. Glu212, Asp214, Glu217, and His228) in *Melanocarpus Albomyces* (Ma) Cellobiohydrolase Cel7B impacts the substrate's orientation and ring conformation and disrupt of the intra-enzyme hydrogen bonding network and enhance the sampling of various ring puckering conformations for the substrate ring at the +1 and -1 subsites. In particular it is observed that the protonation state of Asp214 dictates the accessibility of the glycosidic bond to the catalytic acid/base Glu217 by influencing the  $\phi/\psi$  dihedral angles and the puckering of the ring structure. The protonation-orientation-conformation analysis has revealed an active site that primarily utilizes two highly coupled protonation schemes; one protonation scheme to orient the substrate and generate catalytically favorable substrate geometries and ring puckering conformations and another protonation scheme to hydrolyze the glycosidic bond. In addition to identifying how enzymes utilize protonation state to manipulate substrate geometry, this study identifies possible directions for improving catalytic activity through protein engineering.