Jorgensen Symposium: 30 Years in the Trenches
Department of Chemistry
Yale University
Poster Presentation Abstracts
April 29th 2005

(1) **The influence of inter- and intramolecular hydrogen bonding on Kemp decarboxylations from QM/MM simulations**

Orlando Acevedo and William L. Jorgensen, Department of Chemistry, Yale University, New Haven, CT 06520-8107, orlando.acevedo@yale.edu

The Kemp decarboxylation reaction for benzisoxazole-3-carboxylic acid derivatives has been investigated using QM/MM calculations in protic and dipolar aprotic solvents. Aprotic solvents have been shown to accelerate the rates of reaction by 7-8 orders of magnitude over water; however the inclusion of an internal hydrogen bond effectively inhibits the reaction with near solvent independence. The effects of solvation and intramolecular hydrogen bonding on the reactants, transition structures, and the rate of reaction are elucidated using two-dimensional potentials of mean force (PMF) derived from free energy perturbation calculations in Monte Carlo simulations (MC/FEP). In addition, a discrepancy for the experimental rate in chloroform has been studied in detail with the conclusion that ion-pairing between the reactant anion and TMG^+ counterion is responsible for the anomalously slow reaction rate.

(2) **MC/CRA: Applications to nucleic acids**

Gabriela Barreiro and William L. Jorgensen, Department of Chemistry, Yale University, New Haven, CT 06520-8107, gabriela.barreiro@yale.edu

Monte Carlo Simulations of two RNA molecules, conducted using concerted rotations with flexible bond angles (CRA) for backbone sampling, were performed to study the extension of this method to nucleic acids. The systems under investigation are a RNA hairpin containing a GAAA tetraloop and the Trans-activation response element (TAR) RNA from human immunodeficiency virus type 1 (HIV-1). All simulations were generated using the OPLS-AA force field with the generalized Born (GB/SA) model. Preliminary results are encouraging and indicate that the MC/CRA simulations of nucleic acids can be competitive with molecular dynamics.

(3) **An undergraduate computational chemistry program**

Daqing Gao, Department of Chemistry, Queensborough Community College of The City University of New York, Bayside, NY 11364, dgao@qcc.cuny.edu

A computational chemistry research program is being established at QCC thanks to the support of the CUNY research foundation. This lab aims at introducing undergraduate students to different computer platforms (PC, Linux, and Unix) and the popular computational chemistry programs and graphics, including Gaussian, Spartan, ChemDraw, Chem3D, etc. A representative research project, the calculation of the pK_a's of weak organic carbon acids in aqueous solution with several easily accessible methods will be presented.

(4) **The Curtius rearrangement of benzoyl azide, mechanism and solvent effects**

Kasper P. Jensen and William L. Jorgensen, Department of Chemistry, Yale University, New Haven, CT 06520-8107, kasper.jensen@yale.edu

One of the most important synthetic pathways for generating N-containing heterocyclic compounds is the Curtius rearrangement of "rigid" alkylazides or of "non-rigid" carbonylazides to isocyanates. Whether nitrenes are involved in these reactions is not completely settled yet, and understanding the solvent effects on these reactions remains desirable. We have studied the thermal Curtius reaction of benzoyl azide with two complementary methods: a) B3LYP on cluster models including thermodynamics, continuum solvation, and large basis sets; b) the combined PDDG/OPLS approach including Monte Carlo sampling along the reaction coordinate (QM/MM potential of mean force). Both concerted and non-concerted (involving nitrenes) pathways have been studied and compared for the two approaches.

(5) **QSPR analysis of selectivity of between alkali-metal ions to 18-Crown-6: A Monte Carlo simulation study**

Hag Sung Kim, Department of Environmental Fine Chemistry, Ulsan College, Ulsan, 680-749, S. Korea, hskim@mail.ulsan-c.ac.kr

The solvent effect on differences in stability constant ($\Delta\log K_s$) as well as the relative free energies of binding of alkali-metal ions to 18-crown-6 have been studied by the Monte Carlo simulation of SPT and we have compared differences in stability constant ($\Delta\log K_s$) as well as the relative free energies of binding (selectivity) of alkali-metal ions to 18-crown-6 in this study with those of the published works. There is good agreement among the studies if we consider both methods used to obtain the stability constant ($\Delta\log K_s$) of binding of alkali-metal ions to 18-crown-6 and standard deviations. We have reported the quantitative solvent-polarity relationships (QSPR) studied on the solvent effects the relative free energies of binding of Cs^+ and Rb^+ ions to 18-crown-6. From the calculated coefficients of QSPR, we have noted that DN(Donation number) dominates the differences in relative solvation Gibbs free energies of Cs^+ and Rb^+ ions but Aj (solvent acidity) dominates the differences in the stability constant ($\Delta\log K_s$) as well as the relative free energies of binding of Cs^+ and Rb^+ ions to 18-crown-6. We found that the calculated binding free energies of 18-crown-6 complexes in aqueous solutions are a little underestimate the experimental data; however, they followed the correct sequence, $K^+ > Rb^+ > Cs^+ > Na^+$, our assumed binding free energies of 18-crown-6 complexes in aqueous solutions closely parallel the binding Gibbs free energies of the experimental.

(6) **Selectivity of between alkali-metal ions to 12-Crown-4 in MeOH: A Monte Carlo simulations study**

Hag Sung Kim Department of Environmental Fine Chemistry, Ulsan College, Ulsan, 680-749, S. Korea, hskim@mail.ulsan-c.ac.kr

We have studied the differences in stability constant ($\Delta\log K_s$) as well as the relative free energies of binding (selectivity) of cations to 12-crown-4 i.e. the selectivity of 12-crown-4 to cations using a Monte Carlo simulation of SPT in CH_3OH. The selectivity of 12-crown-4 to cations followed the sequence as $K^+ > Cs^+ > Na^+ > Rb^+ > Li^+$ *i.e.*, the selectivity of 12-crown-4 to K^+ is more favorable than to other cations in CH_3OH. We found 12-crown-4/cation complexes with the 12-crown-4 of almost C_s symmetry that have been never found in any crystal or *ab initio* structure of the cation complexes of 12-crown-4. We also found that 12C4/cation complexes with the 12-crown-4 of almost C_s symmetry are more stable in CH_3OH solutions than 12-crown-4 /cation complexes with the 12-crown-4 of almost C_4 symmetry. ACKNOWLEDGMENT**:** This work was supported by Korea Research Foundation Grant (KRF-2003-015-C00259).

(7) **Molecular modeling calculations on 2-aryl-substituted benzimidazoles: Prediction of activity for novel analogs**

Marilyn Kroeger Smith[1,2], Christopher Michejda[1], David Farnsworth[1], James Lipchock[2,3], Sarah Vannoy Lipchock[2,3] , William L. Jorgensen[3], Julian Tirado-Rives[3], Robert Rizzo[4], and Richard Smith, Jr[1,2] (1) NCI-Frederick Cancer Research and Development Center, Frederick, MD 21702, (2) Department of Chemistry, McDaniel College, Westminster, MD 21157, (3) Department of Chemistry, Yale University, New Haven CT 06520, and (4) Stony Brook University, Stony Brook, NY 11794, kroeger@ncifcrf.gov

In recent years, a number of drugs have been developed to inhibit the HIV-1 virus, including many inhibitors of both the nucleoside and non-nucleoside classes. The non-nucleoside inhibitors are potentially highly effective and non-toxic drugs; however, a great number of these compounds suffer decreased efficacy in the face of viral mutation. Thus, the search for drugs that are resistant to mutations while being potent RT inhibitors continues. With this goal in mind, the 2-aryl-substituted benzimidazole class of compounds was first reported in 1997. The best member of the class at that time, 1-(2,6-difluorobenzyl)-2-(2,6-difluorophenyl)-4-methylbenzimidazole (4-Methyl BPBI), had an IC_{50} = 200nM. A crystal structure of this lead compound has subsequently been solved and has been employed for computer modeling purposes. In the search for more potent analogs of 4-methyl BPBI, we have carried out Monte Carlo modeling calculations using the MCPRO software program. Linear response calculations of this type have previously been validated in studies of a series of RT non-nucleoside inhibitors of HIV-1 RT. In the current study, correlations between calculated energies and physical descriptors that were obtained using a training set of 24 analogs against wild-type and several mutants were used to predict the activity of 29 proposed BPBI analogs.

(8) **A model for inhibition estimation for GSK3 using Monte Carlo simulations along with an extended linear response approach**

Patrick S. Lee and William L. Jorgensen, Department of Chemistry, Yale University, New Haven, CT 06520-8107, patrick.s.lee@yale.edu

Glycogen synthase kinase-3 (GSK3) has been a target for extensive research for its potential role in treating Alzheimer's disease and Type 2 diabetes. In this study, Monte Carlo statistical mechanics simulations in conjunction with the extended linear response approach were used to develop a model to predict the inhibition of GSK3. A model was acquired for 131 inhibitors, which entailed derivatives of indirubins, paullones, and pyrazolo[1,5-b]pyridazines. A model for each core of these derivatives was first generated and then the three sets were combined to generate a generic GSK3 model. A statistically acceptable correlation was achieved, evidenced by the $R^2 > 0.51$ and the $q^2 > 0.61$, using descriptors for EXX-LJ, size of hydrophobic area of ligands upon complexation, internal angle bending energy component of the solute, and solute-solvent LJ energy.

(9) **Development of scoring functions and FEP-based lead optimization protocols with applications to dihydrofolate reductases**

Siegfried Leung and William L. Jorgensen, Department of Chemistry, Yale University, New Haven, CT 06520-8107, siegfried.leung@yale.edu

Protein-ligand interactions of *Plasmodium falciparum* dihydrofolate reductase (pfDHFR) are investigated via a rapid force-field-based scoring approach and MC/FEP calculations. Using BOMB, ligands are constructed from the core structure that is bound to the protein. A conformational search with variable key host's sidechain dihedrals is performed to identify the best structure for each ligand in the binding site. The binding affinity of each ligand is subsequently evaluated by BOMB's scoring functions. In order to model the protein-ligand interactions with high accuracy, MC/FEP calculations are performed to estimate the binding affinity or activity of the ligands. Aided by a better understanding of the actual interactions from the MC/FEP results, scoring functions in BOMB can therefore be improved to describe the protein-ligand interactions more accurately.

(10) **Quest: An intergrated web-based information system for drug discovery**

Dongchul Lim, LG Life Sciences Ltd., Seoul, Korea, cadchem@hotmail.com

We have been developing a platform-independent web-based information system (Quest) which allows researchers to easily access medicinal chemistry, high throughput screening, pharmacokinetics, formulation, and X-ray crystallography data. Data mining and molecular modeling can be performed on this data and the results can be seen to ordinary users through the web. The interface of Quest is designed as easy as possible so that users with little computational experience can use the system without fear. Users with read-only privileges can browse and search data and execute on-the-fly modeling tools, while those with administrative privileges can also insert or update chemical, biological, and modeling data. We expect the Quest system will facilitate an efficient and rapid drug discovery.

(11) **Structural and energetic investigation of the L100I single mutation and L100I K103N double mutation of HIV-1 reverse transcriptase on TMC125 analogues**

Theresa M. Lyons and William L. Jorgensen, Department of Chemistry, Yale University, New Haven, CT 06520-8107, theresa.lyons@yale.edu

The relative fold resistance of a series of compounds active against the HIV-1 Reverse Transcriptase (RT) will be analyzed via the MC/FEP methodology. The influence of a single and double amino acid mutation in the binding site of HIV-1 RT on a series of TMC125 analogues will be examined. Computed relative fold resistance energies will provide structural and energetic insight into the potency of the inhibitors against HIV-1 RT, the single mutant L100I, and the double mutant L100I K103N.

(12) **Biological pathway mapping with genomic structure information**

Fenglou Mao[1], Zhengchang Su[1,2], Victor Olman[1], and Ying Xu[1,2], (1) Computational Systems Biology Lab, Biochemistry & Molecular Biology Department, University of Georgia, Athens, GA 30602 (2) Computational Biology Institute, Oak Ridge National Laboratory, fenglou@csbl.bmb.uga.edu

Mapping known biological pathways across microbial genomes represents a highly important technique in functional studies of biological systems. Existing methods mainly rely on sequence-based orthologous gene mapping, often leading to sub-optimal mapping results. The reason is that sequence similarity information alone does not contain sufficient information for accurate identification of orthology relationship. In this poster, we present a new algorithm for pathway mapping across microbial genomes. The algorithm takes into account both sequence similarity information and genomic structure information such as operons and regulons. One basic premise of our approach is that a microbial pathway could generally be decomposed into a few operons or regulons. We have formulated the pathway mapping problem as to maximize the sequence similarity between mapped gene pairs under the constraint that the mapped genes are grouped into a few operons, preferably co-regulated in the target genome. We have developed an integer programming (IP) algorithm for solving this constrained optimization problem, and implemented the algorithm as a computer software P-MAP. We have tested this software on a number of pathways where the homologous pathways are well characterized in both the template and target genomes (hence we can evaluate the mapping accuracy). We found that using genomic structure information as constraints could greatly improve the pathway-mapping accuracy over methods using sequence similarity information alone.

(13) **Modeling study of several active actinonin analogs in human peptide deformylase**

David Maxwell, Dept. of Experimental Diagnostic Imaging, University of Texas M.D. Anderson Cancer Center, Houston, TX 77054, david.maxwell@di.mdacc.tmc.edu

Peptide deformylase (PDF) is a metalloprotease that catalyzes the cleavage of the formyl group from the methionine residue of newly created polypeptides. PDF activity is essential for bacterial growth and has been known for a long time in prokaryotes, but was only recently discovered in humans. The peptide-like antibiotic Actinonin and several analogs have been identified as potent inhibitors of human PDF. These inhibitors demonstrate clear antiproliferative effects in multiple cancer cell lines. A model of the human form of PDF was constructed from the available crystal structures of the bacterial form. The resultant homology model was studied by molecular dynamics simulation in conjunction with actinonin and several analogs to determine the primary interactions responsible for activity in these inhibitors.

(14) **Structure prediction using Monte Carlo sampling, local backbone moves and implicit GB/SA solvent**

Sara Nichols, Jakob P. Ulmschneider, and William L. Jorgensen, Department of Chemistry, Yale University, New Haven, CT 06520-8107, sara.nichols@yale.edu

Determining the structure of a protein from the primary sequence alone is a challenging task, an ultimate solution of which is likely to involve a variety of disciplines. Folding small systems with molecular mechanics is one way to approach this problem. Current model systems include solution or crystal structures of small peptides and proteins which have secondary structure and no disulfide bonds. Due to the computationally large size of these systems and the complexity of sampling until convergence water is treated with GBSA, a continuum solvent model, in order to decrease the number of pair wise calculations that must be evaluated. In addition, the protein backbone is sampled using concerted rotations (CRA), a procedure which allows local backbone moves while simultaneously biasing against configurations that substantially alter the structure of the protein. Using Monte Carlo sampling and the OPLS-AA force field, different configurations of model systems along the potential energy surface are sampled and compared to the predetermined native structure for analysis of this combination of methods. Previous success with small systems such as the 17 residue trpzip2 (1EOQ) has encouraged investigation into larger systems such as avian pancreatic polypeptide (1PPT) with 36 residues and a ββα-motif forming protein (1FSD) with 28 residues.

(15) **Synthesis and biological evaluation of bis-arylamines as non-nucleoside HIV-1 reverse transcriptase inhibitors based on computer-aided drug design**

Juliana Ruiz-Caro[1], William L. Jorgensen[1], Andrew D. Hamilton[1], Aravind Basavapathruni[2], Christopher Bailey[2] and Karen S. Anderson[2], (1) Department of Chemistry, Yale University, New Haven, CT 06520-8107, (2) School of Medicine, Yale University, New Haven, CT 06520, juliana.ruiz-caro@yale.edu

Targeting HIV-1 reverse transcriptase (RT) with non-nucleoside reverse transcriptase inhibitors (NNRTI) to combat AIDS is complicated, among other factors, by the rapid emergence of drug-resistant strains which are usually cross-resistant to other inhibitors within the class. As a consequence, the search for new classes NNRTI is still being actively pursued. The Jorgensen laboratory has spent many years studying the HIV-RT system and highly effective methods have been developed for computing structures, binding free energies and effects of mutations. Based on these methods, new NNRTIs were designed, synthesized and their biological activity evaluated.

(16) **Reintroduction of the overlap matrix into semiempirical schemes: NO-MNDO and NOPM3**

Kurt W. Sattelmeyer, Ivan Tubert-Brohman, and William L. Jorgensen, Department of Chemistry, Yale University, New Haven, CT 06520, sattlmyr@ramana.chem.yale.edu

NO-MNDO and NOPM3, which differ from MNDO and PM3 by the explicit inclusion of the overlap matrix (S) in the modified NDDO Roothaan-Hall equations, have been implemented and parameterized for hydrocarbons. The mean absolute errors (MAEs) observed on a set of 254 neutral, closed-shell molecules are 3.9 and 3.7 kcal/mol, respectively. NOPM3 is seen to be more accurate (as compared to previous methods) in determining the activation barriers of pericyclic reactions (MAE of 6.9 kcal/mol) while NO-MNDO is not (MAE of 22.6 kcal/mol). These schemes perform comparably well to previously described semiempirical methods for other molecular properties, giving less accurate results for geometrical parameters and, neglecting very small molecules, more accurate results for ionization potentials.

(17) **Dangers in modeling GPCRs – not everything looks like Rhodopsin**

Daniel Severance, Kalypsys, Inc., San Diego, CA 92121, dseverance@kalypsys.com

An analysis of mutagenesis data will be presented which is inconsistent with results of homology modeling via the simple threading approach applied to the Rhodopsin framework. A more critical analysis of the Rhodopsin structure reveals that a critical large scale kink which is often pointed to as a key element in GPCR signaling, is in fact most likely due to an internal interaction in the protein – an interaction which is not conserved in the vast majority of GCPRs. In fact, Rhodopsin itself has other sequence runs with praline which show no such pronounced kink.

(18) **Improved semiempirical methods -- Parameterization of PDDG/PM3 for sulfur**

Ivan Tubert-Brohman, Cristiano R. W. Guimarães, and William L. Jorgensen, Department of Chemistry, Yale University, New Haven, CT 06520-8107, ivan@ramana.chem.yale.edu

The recently developed PDDG/PM3 method has been parameterized for sulfur by using a global optimization that combined genetic approaches with grid searching, simulated annealing, stochastic searching, and Fletcher-Powel local optimizations. For a set of 249 molecules, mean absolute errors (MAE) for heats of formation went from 10.5, 10.6, and 10.0 kcal/mol for the PM3, AM1, and MNDO/d methods, respectively, to 6.6 kcal/mol for PDDG/PM3. Large improvements were observed for problematic functional groups such as sulfones, while also improving simpler functional groups and intermolecular interaction energies. The PDDG/PM3 method gives results that are competitive to those from B3LYP. For molecules in the G2/97 and G3/99 sets, MAEs are 7.4 kcal/mol for B3LYP and 6.6 kcal/mol for PDDG/PM3. If halides are excluded, these numbers change to 7.3 and 3.9 kcal/mol. Other observables such as ionization potentials, dipoles and geometries are similar to those from the original PM3 method.