

## Final Progress Report for DOE Grant

The functions of proteins can be directly related to specific atomic interactions within and between proteins and other molecular components of cells, which are the basic organizational unit of living systems. A detailed description of these atomic interactions is essential for a comprehensive understanding of biomolecular function. Many of the outstanding issues related to the function of biological systems are difficult to address by using experimental methods alone. Computer simulations based on physical and chemical principles can complement laboratory experiments and provide novel insights into the atomic basis of molecular behavior.

### a) Photodissociation dynamics

The distribution of carbon monoxide after photodissociation in the myoglobin haem pocket was investigated using molecular dynamics simulations at 300 K [1]. The results show that both intermediates (one close to the haem iron and one further away) observed in low temperature x-ray studies of photodissociated CO have a high probability of occurrence, even at ambient temperatures. The fact that the O of CO is oriented toward the haem iron in the closer intermediate provides an explanation for the slow rate of CO geminate rebinding. A refinement against x-ray data generated from the molecular dynamics simulations indicated that the CO has a broader distribution in the haem pocket than is apparent from the experimental electron density. This effect is likely to be general for systems containing highly mobile groups.

### b) Extensions of QM/MM methodology

Analytical second derivatives for combined QM/MM calculations were formulated and implemented in the CHARMM program interfaced with the *ab initio* quantum mechanical GAMESS and CADPAC programs [2]. This makes possible evaluation of vibrational frequencies and infrared intensities in large systems that cannot be treated effectively by QM or MM alone; examples are polarizable molecules in solution and substrates or transition states in enzymes. Test calculations on a number of systems, including formamide in water, butanol, a model transition state structure for triosephosphate isomerase and the active site model of myoglobin, showed that the

DOE Patent Clearance Granted

Mark P Dvorscak  
(630) 252-2393

E-mail: mark.dvorscak@ch.doe.gov  
Office of Intellectual Property Law  
DOE Chicago Operations Office

Oct 30, 2001  
Date

## **DISCLAIMER**

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## **DISCLAIMER**

**Portions of this document may be illegible  
in electronic image products. Images are  
produced from the best available original  
document.**

MM description of the environment can capture much of its polarization effects on the QM region. Thus the implementation of analytical second derivatives within the QM/MM framework has considerable potential for the study of large systems.

A method for calculating the chemical shielding tensor of any atom with the QM/MM approach was developed [3]. The method was described and applied to a number of model systems including the water dimer, NMA-water complexes, cytosine monophosphate, paired and stacked nucleic acid bases, imidazole-metal complexes, and 1'-deoxyribose-metal ion complexes. The results demonstrated that with an appropriate QM/MM partition, good descriptions of the environmental effects on chemical shift tensors are obtained. The typical error compared to full QM calculations was 1-2 ppm for heavy atoms. At distances below 2.5 Å, such as occur in hydrogen bond, larger errors arise due to the lack of Pauli repulsion and magnetic susceptibility of the nearby groups in the current QM/MM model; including the hydrogen bonded molecules as part of the QM region as a way of solving this problem.

### c) Enzyme mechanisms

Three mechanisms proposed for the triosephosphate isomerase (TIM) catalyzed reactions were studied with the QM/MM approach using B3LYP/6-31+G(d,p) as the QM method [4]. The two pathways that involve an enediol species were found to give similar values for the barriers and the calculated rates are in satisfactory agreement with experiment. By contrast, the mechanism that involves intramolecular proton transfer in the enediolate was found to be energetically unfavorable due to electrostatic interactions with His 95, a conserved residue in TIM from different organisms. A perturbation analysis was used to determine the residues that make the major contribution to catalysis.

Chorismate mutase acts at the first branch-point of aromatic amino acid biosynthesis and catalyzes the conversion of chorismate to prephenate [5]. The results of molecular dynamics simulations of the substrate in solution and in the active site of chorismate mutase were reported. Two nonreactive conformers of chorismate were found to be more stable than the reactive pseudodiaxial chair conformer in solution. It was shown by QM/MM molecular dynamics

simulations, which take into account the motions of the enzyme, that when these inactive conformers are bound to the active site, they are rapidly converted to the reactive chair conformer. This result suggested that one contribution of the enzyme is to bind the more prevalent nonreactive conformers and transform them into the active form in a step before the chemical reaction. The motion of the reactive chair conformer in the active site calculated by using the QM/MM potential generated transient structures that are closer to the transition state than is the stable CHAIR conformer.

#### d) RNA recognition

The diversity of RNA tertiary structures provides the basis for specific recognition by proteins or small molecules [6]. To investigate the structural basis and the energetics which control RNA-ligand interactions, favorable RNA binding sites were identified using the MCSS method, which has been employed previously only for protein receptors. Two different RNAs for which the structures have been determined by NMR spectroscopy were examined: two structures of the TAR RNA which contains an arginine binding site, and the structure of the 16S rRNA which contains an amino-glycoside binding site (paramomycin). In accord with the MCSS methodology, the functional groups representing the entire ligand or only part of it (one residue in the case of the aminoglycosides) were first replicated and distributed with random positions and orientations around the target and then energy minimized in the force field of the target RNA. The Coulombic term and the dielectric constant of the force field were adjusted to approximate the effects of solvent-screening and counterions. Optimal force field parameters were determined to RNA. The more favorable binding sites for each residue of the aminoglycoside ligands were then calculated and compared with the binding sites observed experimentally. The predictability of the method was evaluated and refinements were proposed to improve its accuracy.

## References

- [1] A Comparison between Molecular Dynamics and X-ray Results for Dissociated CO in Myoglobin, *Nature Struct. Biol.* **4**, 202-208 (1997), by D. Vitkup, P. A. Petsko, and M. Karplus.
- [2] Molecular Properties from Combined QM/MM Methods. I. Analytical Second Derivative and Vibrational Calculations, *J. Chem. Phys.* **112**, 1133-1149 (2000), by Q. Cui and M. Karplus.
- [3] Molecular Properties from Combined QM/MM Methods. II. Chemical Shifts in Large Molecules, *J. Phys. Chem. B* **104**, 3721-3743 (2000), by Q. Cui and M. Karplus.
- [4] Triosephosphate Isomerase (TIM): A Theoretical Comparison of Alternative Pathways, *J. Am. Chem. Soc.* **123**, 2284-2290 (2001), by Q. Cui, M. Karplus.
- [5] Substrate Conformational Transitions in the Active Site of Yeast Chorismate Mutase: Their Role in the Catalytic Mechanism, *Proc. Natl. Acad. Sci. USA* **98**, 9032-9037 (2001), by H. Guo, Q. Cui, W. N. Lipscomb, and M. Karplus.
- [6] MCSS-Based Predictions of RNA Binding Sites, *Theor. Chem. Acc.* **101**, 131-137 (1999), by F. LeClerc and M. Karplus.