

We developed a linear-scaling semiempirical quantum mechanical (QM) program (DivCon). Using DivCon we can now routinely carry out calculations at the fully QM level on systems containing up to about 15 thousand atoms. We also implemented a Poisson-Boltzmann (PM) method into DivCon in order to compute solvation free energies and electrostatic properties of macromolecules in solution. This new suite of programs has allowed us to bring the power of quantum mechanics to bear on important biological problems associated with protein folding, drug design and enzyme catalysis. Hence, we have garnered insights into biological systems that have been heretofore impossible to obtain using classical simulation techniques.

DOE Patent Clearance Granted  
MPDvorscak 7-19-04  
Date  
Mark P Dvorscak  
(630) 252-2393  
E-mail mark.dvorscak@ch.doe.gov  
Office of Intellectual Property Law  
DOE Chicago Operations Office

## **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.**

## Structure/Function Studies of Proteins Using Linear Scaling Quantum Mechanical Methodologies

Kenneth M. Merz Jr.

September 15, 1996-December 31, 2001

152 Davey Laboratory; Department of Chemistry; The Pennsylvania State University;  
University Park, Pennsylvania 16802  
DE-FG02-96ER62270

We have developed and reported on a linear-scaling semiempirical quantum mechanical (QM) program (DivCon) and we have also parallelized this program in order to leverage more performance out of this program by running it in parallel. We can now carry out calculations at the fully QM level on systems containing up to about 1-15 thousand atoms in a matter of seconds to a few hours. We have implemented a Poisson-Boltzmann (PM) method into DivCon in order to compute solvation free energies and electrostatic properties of proteins. We are in the process of evaluating the performance of this method and we are using it to garner insights into how the electronic structure of a protein is affected by the presence of an implicit solvent. We have also developed a PB approach that incorporates charge transfer at the protein/water interface. We have coded up a semiempirical parameterization program in order to derive new parameter sets (Hamiltonians) that better model protein structure and energetics at the semiempirical level. In conjunction with this project we are also developing a semiempirical water model that is compatible with existing semiempirical parameterizations (e.g., PM3 and AM1). Since the inception of this grant we have begun to pursue other ideas that were not as obvious at the time we wrote the proposal. In particular, we are examining the utility of the combined DivCon/PB approach to determine NMR chemical shifts and coupling constants for biomolecules. Using these tools we have begun development of a "Quantum Bioinformatics" database to store quantum derived quantities (e.g., charges, bond orders, *etc.*) of biomolecules that can then be mined by interested users via a web based interface. We also developed a novel combined QM/QM method. In this approach we use the divide and conquer strategy to integrate semiempirical quantum mechanics with high-level density functional theory (DFT) or *ab initio* quantum mechanics. This allows us to carry out highly accurate QM calculations on the key region of an enzyme using, for example, DFT (e.g., enzyme active site, *etc.*), while treating the rest of the system with semiempirical QM methods. This is a novel approach, that will change the way in which we study complex biological systems.

### Publications:

- 1) "Fast, Accurate Semiempirical Molecular Orbital Calculations for Macromolecules" S. L. Dixon; K. M. Merz, Jr. *J. Chem. Phys.* **1997**, *107*, 879-893.
- 2) "Charge-Transfer Interactions in Macromolecular Systems: A New View of the Protein/Water Interface" G. Nádig; L. C. van Zant; S. L. Dixon; K. M. Merz, Jr. *J. Am. Chem. Soc.* **1998**, *120*, 5593-5594.
- 3) "Parallel Implementation of the Divide and Conquer Methodology" J. Vincent; S. L. Dixon; K. M. Merz, Jr. *Theor. Chem. Acc.* **1998**, *99*, 220-223.
- 4) "Divide and Conquer Approaches for Linear-Scaling Semiempirical SCF-MO Calculations" K. M. Merz, Jr. *Encyclopedia of Comput. Chem.* **1998**, *1*, 761-776.
- 5) "Implementation and Testing of a Frozen Density Matrix-Divide and Conquer Algorithm" M. Ermolaeva; A. van der Vaart; K. M. Merz, Jr. *J. Phys. Chem.* **1999**, *103*, 1868-1875.

- 6) "Divide and Conquer Interaction Energy " A. van der Vaart; K. M. Merz, Jr. *J. Phys. Chem. A* **1999**, *103*, 3321-3329.
- 7) "Fully Quantum Mechanical Description of Proteins in Solution. Combining Linear Scaling Quantum Mechanical Methodologies with the Poisson-Boltzmann Equation" V. Gogonea; K. M. Merz, Jr. *J. Phys. Chem. A* **1999**, *103*, 5171-5188.
- 8) "The Role of Polarization and Charge Transfer in the Solvation of Biomolecules" A. van der Vaart; K. M. Merz, Jr. *J. Am. Chem. Soc.* **1999**, *121*, 9182-9190
- 9) " Charge Transfer Interactions in Biology: A New View of the Protein/Water Interface" ACS *Sym. Ser.* **1999** , *721*, 439-447.
- 10) "Towards All-Electron Modeling of Biomolecular Systems" G. Monard; K. M. Merz, Jr. *Acc. Chem. Res.* **1999**, *32*, 904-911.
- 11) "Charge Transfer in Biologically Important Biomolecules: A Comparison of High Level *ab initio* and Semiempirical Models" a. van der Vaart *Int. J. Quantum Chem.* **2000**, *77*, 27-43.
- 12) "A Quantum Mechanical-Poisson-Boltzmann Equation Approach for Studying Charge Flow Between Ions and a Continuum Dielectric" V. Gogonea; K. M. Merz, Jr. *J. Chem. Phys.* **2000**, *104*, 2117-2122.
- 13) "Charge Flow Between Ions and a Dielectric Continuum: 2. Variational Method for Distributing Charge into the Dielectric" V. Gogonea; K. M. Merz, Jr. *J. Phys. Chem* **2000**, *112*, 3227-3235.
- (14) "Are Many-Body Effects Important in Protein Folding" A. van der Vaart; B. D. Bursulaya; C. L. Brooks, III, K. M. Merz, Jr. *J. Phys. Chem.* **2000** *104*, 9554-9563.
- 15) "QM/QM Methods: 1. A Divide and Conquer Strategy for Solving the Schrödinger Equation for Large Molecular Systems using a Composite Density Functional - Semiempirical Hamiltonian" *J. Chem. Phys.* **2000** *113*, 5604-5613.
- 16) "Critical Assessment of the Performance of the Semiempirical Divide and Conquer Method for Single Point Calculations and Geometry Optimizations of Large Chemical Systems" A. van der Vaart; D. Suarez; K. M. Merz, Jr. *J. Chem. Phys.* **2000**, *113*, 10512-10523
- 17) "Sodium Parameters for AM1 and PM3 Optimized Using a Modified Genetic Algorithm" E. Brothers; K. M. Merz, Jr. *J. Phys. Chem. B* **2001**, *106*, 2779-2785.
- 18) "Charge Transfer in Small Hydrogen Bonded Clusters" A. van der Vaart; K. M. Merz, Jr. *J. Chem. Phys.* **2002**, *116*, 7380-7388.