

# Adventures in Geochemical Molecular Modeling

**James D. Kubicki**

Department of Geosciences  
The Pennsylvania State University  
University Park, PA 16801  
kubicki@geosc.psu.edu

**Randall T. Cygan**

Geochemistry Department  
Sandia National Laboratories  
Albuquerque, NM 87185-0750  
rtcyan@sandia.gov

## Introduction

Somewhere between the ideas that we can calculate everything we need to know about chemical systems and that everything coming out of a computer is nonsense, lies the truth about molecular modeling. In essence, we might be able to calculate everything if we could solve the Schrödinger equation for macroscopic systems. However, numerous approximations to the full Schrödinger equation are generally made in even the most sophisticated quantum mechanical calculations.

Furthermore, the size of most simulated systems involves dozens to perhaps thousands of atoms and therefore does not approach the molar-sized scale of materials, to say nothing of a geochemical system of macroscopic aqueous solution and minerals. Limitations, such as these, should not discourage the use of molecular modeling techniques in geochemistry. Judicious choice of model systems and techniques allows a great deal of useful information to be extracted from a "computer experiment." In this article, we outline a variety of techniques that can be employed to address different types of geochemical problems.

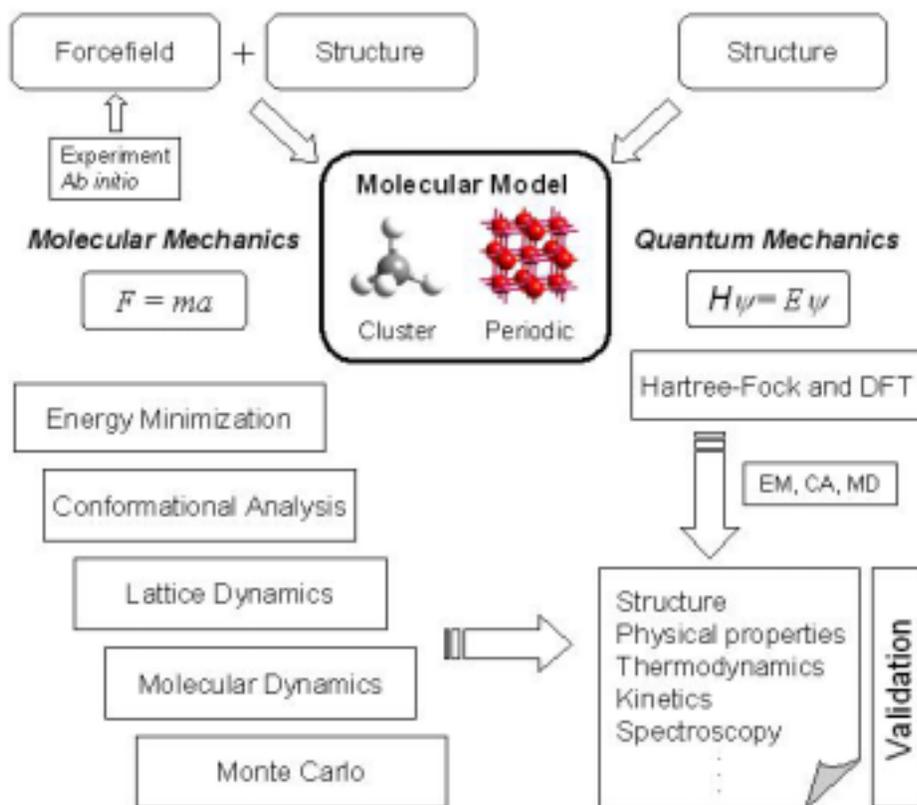


Figure 1. Flow diagram for molecular modeling methods indicating input requirements and potential results; model *M* may be comprised of a molecular cluster or a periodic cell.

A need to explain experimental results motivated the earliest molecular modeling by scientists from the earth science community (Gibbs et al., 1972; Tossell et al., 1972; Woodcock et al., 1976). That motivation remains just as compelling today as it was 30 years ago. Tossell et al. (1972), Tossell (1973), and Collins et al. (1972) focused on the molecular orbital interpretation of X-ray emission and photoelectron spectroscopy of minerals. Though G. V. Gibbs shared an interest in X-ray emission spectroscopy (White and Gibbs, 1967), his more immediate motivation was a desire to explain mineral structure: bond lengths and angles (Brown et al. 1969; Brown and Gibbs, 1969; Gibbs et al., 1972).

An unmistakable trend in recent years finds mineralogists and chemists augmenting traditional phenomenological observations with atomic-scale measurements using physical methods, including spectroscopy plus magnetic and X-ray scattering methods. This shift in experimental approach creates an increasing demand by laboratory experimentalists for

atomic-scale simulations. The link between atomic-scale processes, the point of departure for molecular modeling, and macroscopic observables remains tenuous, a result of the inherent heterogeneity and complexity of earth materials. The connection between molecular modeling and observables is more direct when both are on the same time and length scale.

### Molecular Modeling Methods

A number of molecular modeling techniques have been developed, each with advantages and disadvantages. Practical computational limits generally force trade-offs when selecting a method to study a given chemical system. One can either choose highly accurate quantum mechanical methods that detail the position of every nucleus and the electron density within a system but limit the number of atoms explicitly included, or a molecular mechanics approach can be adopted where up to millions of

are no adjustable or empirically-derived parameters used in calculating the molecular energy. *Ab initio* calculations are preferable because there is less chance for the results to be biased by the researcher or in error due to inadequacies of the interatomic potential. In practice, a number of approximations are made within the molecular orbital calculations that have a direct bearing on the reliability of the results obtained. For example, the Hartree-Fock approximation allows a large saving of computer time and memory, but it neglects the effect of electron correlation on the calculated energy of the system. Methods such as Møller-Plesset second-order perturbation theory (MP2; Møller and Plessett, 1934) and density functional theory (DFT; Kohn and Sham, 1965) can be used to account for electron correlation, but the added computational demands can make including these corrections prohibitive for moderate-size molecules. DFT is much faster than MP2 and is increasingly used in conjunction with molecular orbital calculations. Semi-empirical quantum methods provide a computationally faster but

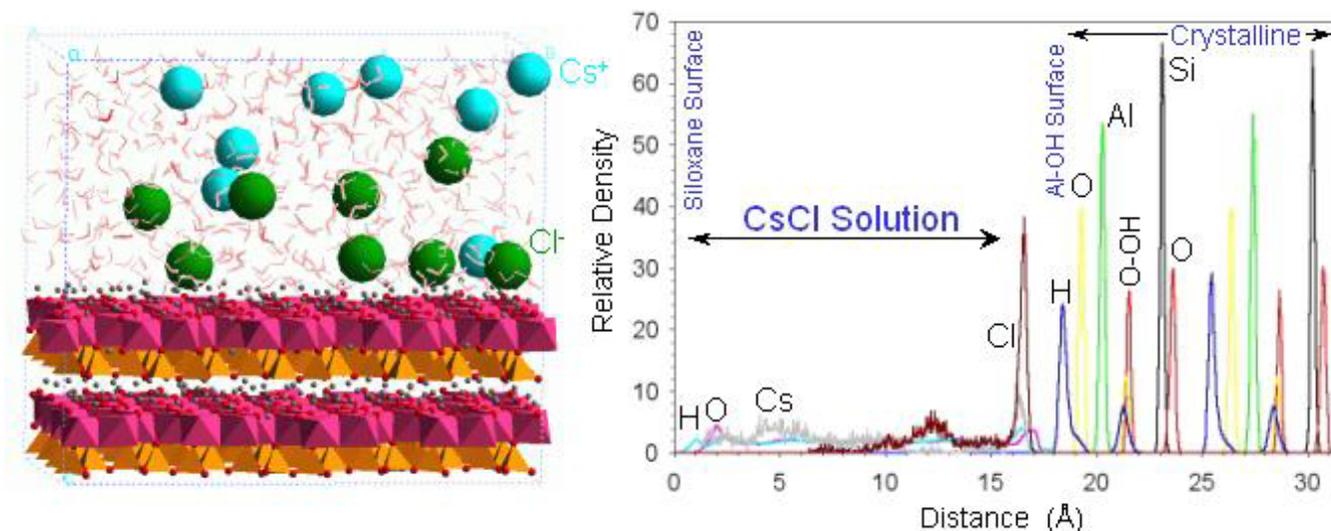


Figure 2. (Left) Snapshot from the equilibrium trajectory of the molecular dynamics simulation of the kaolinite-CsCl solution interface. (Right) Atomic density profile derived from the equilibrium trajectory; aqueous solution on left and kaolinite substrate on right.

atoms may be simulated but atomic interactions are approximated by simple equations. As computational hardware and software become more powerful, the gap between these two end-members diminishes. Quantum mechanical calculations will handle larger numbers of atoms and molecular mechanics simulations will have increasingly accurate interatomic potentials. The two approaches have been linked recently by the development of embedded simulations (Dapprich et al., 1999) in which simple interactions such as bond stretching are handled analytically but more subtle interactions are funneled into subroutines for quantum mechanical calculations. Figure 1 summarizes the various choices the theoretical geochemist has in pursuing accurate molecular simulation.

### Molecular Orbital and Density Functional Theories

*Ab initio* molecular orbital theory is a method used to model the electron density in an atom or molecule with respect to the average electron density. The term "*ab initio*" signifies that there

less theoretically robust alternative method (Pople and Beveridge, 1970).

### Molecular Mechanics

Some molecular modeling simulations avoid calculation of electron densities altogether. Instead, each atom is given a set of parameters that fit into analytical equations used to describe atomic interactions such as Coulombic, van der Waals, bond stretching, etc. The equations and parameters governing these interactions are termed a force field. Neglecting to solve for the electron density of the system saves vast amounts of computer time, so many more atoms—up to millions in some recent cases—may be treated with molecular mechanics. The main problem in utilizing this approach is determining that the force field accurately represents the atomic interactions.

Once an appropriate interatomic potential has been set up to describe the interactions of all the atoms in a system, several

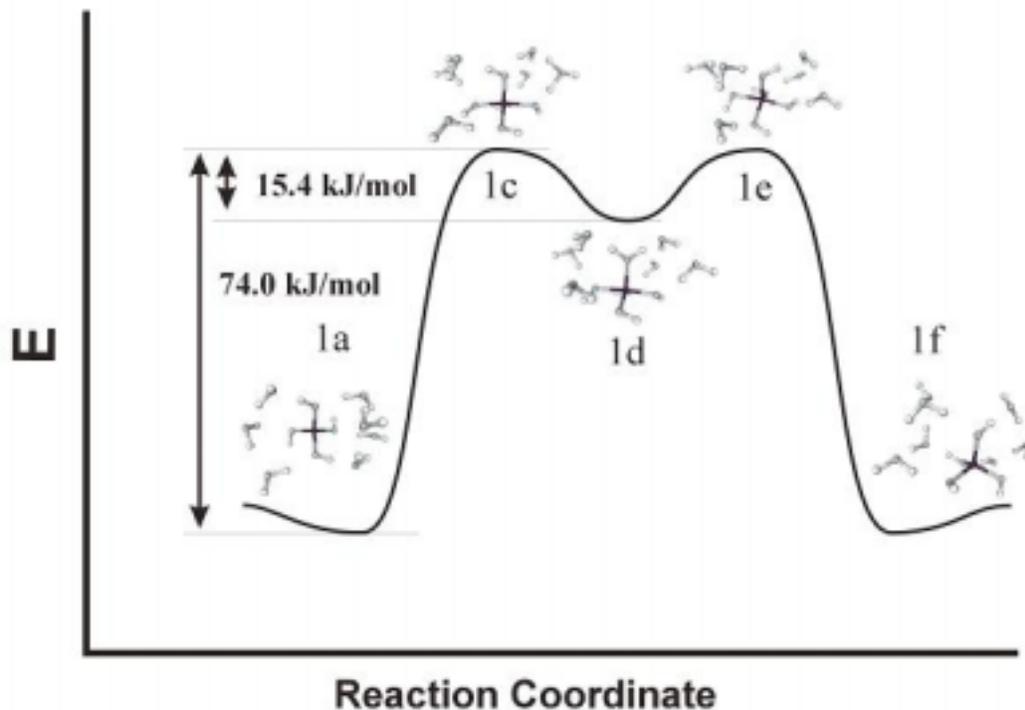


Figure 3. Transition state diagram of the potential energy as a function of reaction pathway for oxygen isotope exchange in a silicic acid solution (from Ph.D. dissertation of M. A. Felipe).

possible methods exist for modeling the system. In order of increasing complexity these are static energy calculations, energy minimizations, lattice dynamics, Monte Carlo calculations, and molecular dynamics simulations. As the name implies, static energy calculations simply take an input configuration, subject the system to the force field equations, and derive an energy for the given state. For example, static energy calculations can be used with molecular orbital theory or molecular mechanics where configurations have been predetermined or obtained by refinement of experimental data.

Monte Carlo calculations are useful for investigating complex systems where several potentially stable structures may exist and where sampling many initial configurations often leads to the equilibrium structures. Molecular dynamics simulations take atomic movement one step further. Time is explicitly included in the calculation and all the atoms move in concert according to classical mechanics equations based on the potential energy surface and the kinetic energies of the atoms for a given temperature. Molecular dynamics is the method of choice for studying dynamical properties of systems, such as diffusion or other time-dependent processes.

Several books on molecular modeling techniques are available (Cygan and Kubicki, 2001; Hehre et al., 1986; Frenkel and Smit, 1996; Young, 2001) and may help in choosing the appropriate method for an application. Commercial software is available for most of the above types of simulations through companies such as Accelrys, Inc. (San Diego, California), Gaussian, Inc. (Carnegie, Pennsylvania), Hypercube Inc. (Waterloo, Ontario), and Tripos Inc. (St. Louis, Missouri). Additionally, clearinghouses such as the Quantum Chemistry Program Exchange (Indiana University) and CCP5 (Daresbury, United Kingdom) make programs and subroutines available at nominal cost.

### Mineralogy and Interfacial Processes

Molecular simulations of mineral structures have been performed since the 1980's using molecular mechanics (e.g., Catlow et al., 1982) and later in the 1990's with periodic quantum methods (e.g., Brodholt et al., 1996). However, only recently have geochemists attempted to simulate the complex chemistry associated with mineral surfaces and mineral-fluid interfaces (e.g., Greathouse et al., 2002). The development of an accurate interfacial model would greatly benefit our understanding of adsorption, dissolution, growth, and reactions that occur at or near the mineral surface, and that control many environmental processes. Clay minerals, in particular, present a noteworthy challenge due to their complex crystallography and composition, and nano-sized grains—all contributing factors to our inability to fully characterize these common soil phases. Cygan and coworkers have recently developed an adsorption model of the kaolinite-solution interface based on the large-scale molecular dynamics simulation of cesium sorption. Figure 2 presents the periodic simulation cell used to evaluate the partitioning of Cs and Cl ions from an aqueous solution to either the siloxane or hydroxyl basal sheets of the kaolinite, and the equilibrium atomic density profile. The simulation results indicate that ion pairing may contribute to adsorption of Cs to the hydroxyl sheet and not just solely to the negatively-charged siloxane surface. Recent synchrotron-based measurements of real systems exhibit evidence of this form of metal-anion complexes on the mineral surface.

### Isotope Geochemistry

A few studies have used molecular modeling techniques in isotopic geochemistry. Casey et al. (1990) performed one of the

first studies of H<sub>2</sub>O-silica dissolution that combined experimental data and ab initio results relevant to isotopic effects. Driesner et al. (2000) performed ab initio calculations to model H and O isotope fractionation by hydrated ions. Schauble et al. (2001) used molecular mechanics methods to estimate Fe-isotope fractionation factors. Because isotope exchange fractionation depends on the zero-point vibrational energies of reactants and products, this type of study is also helpful in estimating fractionation values when they are difficult to obtain experimentally.

Theoretical studies of isotopic exchange have used ab initio calculations to describe the pathway between reactants and products in a step-by-step manner and to find the transition state and activation energy barrier to isotope exchange (Felipe et al., 2003). Although isotopic exchange equilibria have been measured, limited experimental data exist on the rates of isotopic exchange for many important geochemical species. Hence, these modeling studies are used to provide estimates of the rate data and can be tested against the known isotope exchange equilibrium constants. Figure 3 represents such a reaction pathway as the potential energy changes during O-isotopic exchange in aqueous H<sub>4</sub>SiO<sub>4</sub>. The results predict that oxygen exchange between dissolved silicic acid and water should occur on the order of tens of seconds. Consequently, oxygen isotopes in silica precipitated from solution should reflect the isotopic composition of the solution rather than the original source material.

### Organic Geochemistry

As in isotope geochemistry, the number of papers employing molecular modeling techniques in organic geochemistry is disproportionately small. This is surprising given that much of

the development of molecular modeling techniques has focused on organic chemistry. Many of the force fields have been parameterized for organic compounds. In addition, the emphasis on detailed reaction mechanisms found in many organic chemistry textbooks suggests that molecular modeling would be useful for addressing problems in organic geochemistry. The reason for this discrepancy may be that much of naturally occurring organic matter (e.g., humic and fulvic acids, black carbon, etc.) is poorly characterized on a molecular scale. This makes the step of creating a fundamental molecular model difficult.

Despite this difficulty, some recent studies have made excellent use of molecular modeling methods. For example, Cody and Sághi-Szabó (1999) predicted changes in the <sup>13</sup>C NMR chemical shielding in lignin based on ab initio calculations designed to constrain the calculated values as a function of torsion angle. Peters (2000) used molecular mechanics to predict the elution behavior of epimers in order to help interpret chromatographic results on petroleum samples. Kubicki and coworkers have been modeling soot and its interactions with polycyclic aromatic hydrocarbons (PAHs). In this case, MD simulations of hexane soot material have resulted in a structure that is similar to what is known about soot particles. However, the same force field that produced a reasonable soot structure predicts incorrect values for PAH-soot partitioning. Ab initio calculations at the MP2 level of theory were required in order to obtain partitioning energies that would correlate (Figure 4) with observed equilibrium constants (Bucheli and Gustafsson, 2000). This illustrates the point that the correct prediction of one property of a system by molecular modeling does not guarantee the correct prediction of all properties. Verification of modeling results by as many pieces of experimental data as possible is critical in producing a robust methodology.

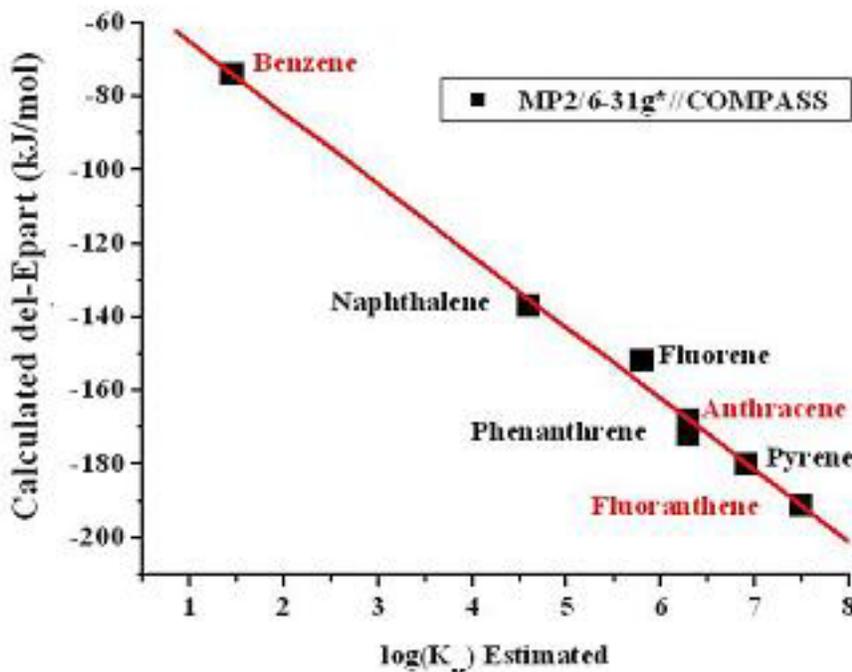


Figure 4. Comparison of ab initio partitioning energies and observed log  $K_{sc}$  values for various polycyclic aromatic hydrocarbons on soot; experimental values are from Bucheli and Gustafsson (2000). The benzene, anthracene, and fluoranthene values were predicted based on the correlation of the ab initio experimental results.

## Biogeochemistry

Biogeochemistry has become a hot area of research and some workers are beginning to bring molecular modeling techniques to bear on these problems. Molecular simulations have been common in biochemistry, especially drug development, for some time. However, the complication of dealing with biological molecules and inorganic geomaterials is a daunting challenge. For example, a great deal of work has been carried out on polysaccharides, a major component of bacterial surfaces, but only a few studies address the chemistry of these biomolecules under geochemical conditions. Notably, the simulation of bacterial membranes by Straatsma and coworkers (e.g., Shroll and Straatsma, 2002) has been advancing this field. Another good example of biogeochemical molecular modeling is provided by Sahai and Tossell (2002) in their calculations relating to the stabilities of biogenic silica species. Also, studies that model geochemical reactions relevant to microbial processes in the environment are beginning to be published (e.g., Duff et al., 2002).

One area where biology and geology meet is in the cycling of silicon. This process and related germanium complexation with natural organic matter can then affect the Ge/Si ratio in river waters that eventually reach the ocean. NMR spectroscopic studies have suggested that Si may form complexes with certain organic molecules in aqueous solution, and XAFS studies have shown that other types of organic molecules preferentially bind to Ge rather than form Si complexes. However, the interpretation of NMR spectra of Si-organic solutions has been controversial. Kubicki and Heaney (2003) demonstrated Si-organic complexation could be responsible for the observed NMR peaks and that other aqueous species did not fit the analytical data. This role for molecular modeling—testing proposed structural models based on spectral interpretations—is an excellent one that allows experiment and theory to complement one another. Such an approach is catching on with spectroscopists in geochemistry as they attempt to decipher the often ambiguous data collected on complex and often disordered geologic materials.

## Future Directions in Geochemical Molecular Modeling

The phenomenal growth in computing power is allowing researchers to rapidly increase the realism and accuracy of their molecular simulations. In addition, more emphasis is being placed on reaction kinetics and understanding fundamental mechanisms in geochemistry. Just about everywhere one looks for equilibrium situations, the effects of kinetics and disequilibrium can be seen. Process-oriented geochemistry is evolving to examine mechanisms to explain reactions. Understanding initial and final states is not enough; the path between the two should also be understood.

Molecular modeling techniques will ultimately become accepted if they can be integrated into coursework at the undergraduate and graduate levels. The beautiful three-dimensional graphics now possible on a desktop personal computer lend themselves to producing visual aids in mineralogy and crystallography. Concepts in geochemical kinetics can also be illustrated clearly by the use of molecular modeling graphics. However, we must always be cautious not to let deceptively accurate graphics make us believe the simulation results are always accurate. Beyond

simple graphics, modeling can be used to examine nucleation, adsorption, complexation, and relationships between structure and thermodynamic stability. Once geochemists begin to see molecular modeling techniques as a tool with weaknesses and strengths—like any other tool in the array of methods for studying geochemical problems—then molecular simulations can be truly integrated into geochemical research rather than being reserved for specialists.

## References

- Brodholt, J., Patel, A., and Refson, K. (1996) An ab initio study of the compressional behavior of forsterite. *American Mineralogist*, 81(1-2), 257-260.
- Brown G.E. and Gibbs, G.V. (1969) Oxygen coordination and the Si-O bond. *American Mineralogist*, 54, 1528-1539.
- Brown G.E., Gibbs, G.V., and Ribbe, P.H. (1969) The nature and the variation in length of the Si-O and Al-O bonds in framework silicates. *American Mineralogist*, 54, 1044-1061.
- Bucheli, T.D. and Gustafsson, O. (2000) Quantification of the soil-water distribution coefficient of PAHs provides mechanistic basis for enhanced sorption observations. *Environmental Science & Technology*, 34(24), 5144-5151.
- Casey, W.H., Lasaga, A.C., and Gibbs, G.V. (1990) Mechanisms of silica dissolution as inferred from the kinetic isotope effect. *Geochimica et Cosmochimica Acta*, 54(12), 3369-3378.
- Catlow, C.R.A., Thomas, J.M., Parker, S.C., and Jefferson, D.A. (1982) Simulating silicate structures and the structural chemistry of pyroxenoids. *Nature*, 295(5851), 658-662.
- Cody, G.D. and Sághi-Szabó, G. (1999) Calculation of the <sup>13</sup>C NMR chemical shift of ether linkages in lignin derived geopolymers: Constraints on the preservation of lignin primary structure with diagenesis. *Geochimica et Cosmochimica Acta*, 63(2), 193-205.
- Collins, G.A.D., Cruickshank, D.W.J., and Breeze, A. (1972) Ab initio calculations on the silicate ion, orthosilicic acid and their L2,3 X-ray spectra. *Journal of the Chemical Society, Faraday Transactions II*, 68, 1189-1195.
- Cygan, R.T. and Kubicki, J.D. (2001) *Molecular Modeling Theory: Applications in the Geosciences*. Reviews in Mineralogy and Geochemistry, Volume 42. 531 p. Geochemical Society, Washington, D.C.
- Dapprich, S., Komaromi, I., Byun, K.S., Morokuma, K., and Frisch, M.J. (1999) A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives. *Journal of Molecular Structure (Theochem)*, 461-462, 1-21.
- Driesner, T., Ha, T.K., and Seward, T.M. (2000) Oxygen and hydrogen isotope fractionation by hydration complexes of Li<sup>+</sup>, Na<sup>+</sup>, K<sup>+</sup>, Mg<sup>2+</sup>, F<sup>-</sup>, Cl<sup>-</sup>, and Br<sup>-</sup>: A theoretical study. *Geochimica et Cosmochimica Acta*, 64(17), 2755-2769.
- Duff, M. C., Coughlin, J. U., and Hunter, D. B. (2002) Uranium co-precipitation with iron oxide minerals. *Geochimica et Cosmochimica Acta*, 66(20), 3533-3547.
- Felipe, M., Kubicki, J.D., and Rye, D.M. (2003) Hydrogen isotope exchange kinetics between H<sub>2</sub>O and H<sub>4</sub>SiO<sub>4</sub> from ab initio calculations. *Geochimica et Cosmochimica Acta*, in press.
- Frenkel, D. and Smit, B. (1996) *Understanding Molecular Simulation*. 443 p. Academic Press, San Diego.
- Gibbs, G.V., Hamil, M.M., Louisnathan, S.J., Bartell, L.S., and Yow, H. (1972) Correlations between Si-O bond length, Si-O-Si angle and bond overlap populations calculated using extended Hückel molecular orbital theory. *American Mineralogist*, 57, 1578-1613.
- Greathouse, J.A., O'Brien, R.A., Bemis, G., and Pabalan, R.T. (2002) Molecular dynamics study of aqueous uranyl interactions with quartz (010). *Journal of Physical Chemistry B*, 106, 1646-1655.
- Hehre, W.J., Radom, L., Schleyer, P.v.R., and Pople, J.A. (1986) *Ab Initio Molecular Orbital Theory*. 548 p. John Wiley and Sons, New York.
- Kohn, W. and Sham, L.J. (1965) Self-consistent equations including exchange and correlation effects. *Physical Review*, 140, 1133-1138.
- Kubicki, J.D. and Heaney, P.J. (2003) Molecular orbital modeling of aqueous organosilicon complexes: Implications for silica biomineralization. *Geochimica et Cosmochimica Acta*, in press.
- Møller, C. and Plessett, M.S. (1934) Note on an approximate treatment for many-electron systems. *Physical Review*, 46, 618-622.
- Peters, K.E. (2000) Petroleum tricyclic terpanes: Predicted physicochemical behavior from molecular mechanics calculations. *Organic Geochemistry*, 31(6), 497-507.
- Pople, J.A. and Beveridge, D.L. (1970) *Approximate Molecular Orbital Theory*. 214 p. McGraw-Hill, New York.
- Sahai, N. and Tossell, J.A. (2002) <sup>29</sup>Si NMR shifts and relative stabilities calculated for hypercoordinated silicon-polyalcohol complexes: Role in sol-gel and biogenic silica synthesis. *Inorganic Chemistry*, 41, 748-756.
- Schauble, E.A., Rossman, G.R., and Taylor, H.P. (2001) Theoretical estimates of equilibrium Fe-isotope fractionations from vibrational spectroscopy. *Geochimica et Cosmochimica Acta*, 65(15), 2487-2497.
- Shroll, R.M. and Straatsma, T.P. (2002) Molecular structure of the outer bacterial membrane of *Pseudomonas aeruginosa* via classical simulation. *Biopolymers*, 65(6), 395-407.
- Tossell J.A. (1973) Interpretation of K X-ray emission spectra and chemical bonding in oxides of Mg, Al and Si using quantitative molecular orbital theory. *Geochimica et Cosmochimica Acta*, 37, 583-594.
- Tossell J.A., Vaughan D.J., and Johnson K.H. (1972) X-ray photoelectron, X-ray emission, and UV spectra of SiO<sub>2</sub> calculated by the SCF-Xa scattered wave method. *Chemical Physics Letters*, 20, 329-334.
- White, E.W. and Gibbs, G.V. (1967) Structural and chemical effects on the Si K<sub>α</sub> X-ray line form silicates. *American Mineralogist*, 52, 985-993.
- Woodcock, L.V., Angell, C.A., and Cheeseman, P. (1976) Molecular dynamics studies of the vitreous state: Simple ionic systems and silica. *Journal of Chemical Physics*, 65, 1565-1577.
- Young, D.C. (2001) *Computational Chemistry: A Practical Guide for Applying Techniques to Real World Problems*. 381 p. John Wiley and Sons, Inc., New York.