This paper reviews recent advances in computational thermochemistry, solvation modeling, and the calculation of chemical reaction rates in the gas phase and in solution. Recent advances in computational thermochemistry include integrated molecular orbital methods, scaling correlation energy, extrapolation to infinite basis sets, and multi-coefficient correlation methods. Recent advances in solvation modeling include Charge Model 2 (a class IV charge model) and the SM5.42R and SM5.42 solvation models; the solvation models are based on semiempirical molecular orbital theory, the \textit{ab initio} Hartree-Fock method, or density functional calculations in the presence of a reaction field and on atomic surface tensions representing first-solvation-shell effects of water or an organic solvent. Our reaction rate calculations are based on variational transition state theory with multidimensional semiclassical tunneling approximations; in liquid solution we may add either equilibrium or nonequilibrium solvation effects.

INTRODUCTION

Much of our recent work is directed to one or another of two complementary goals: (1) to develop improved practical methods for electronic structure calculations in the gas-phase and in liquid-phase solutions and (2) to develop improved methods for interfacing these electronic structure calculations with dynamical methods for the prediction of chemical reactions rates, especially in systems with hydrogenic motion in the reaction coordinate where it is necessary to take account of quantum effects on the nuclear motion. The present paper provides an overview of some areas where progress has been achieved.

GAS-PHASE ELECTRONIC STRUCTURE

Substituent effects on bond energies are very important for both thermochemistry and kinetics. We have recently shown that integrated molecular orbital methods may be used to calculate accurate substituent effects on bond energies by treating a small capped subsystem at a high level of quantum mechanical electronic structure and the rest of a large system at a much lower level.\textsuperscript{1-3} We have also developed a method for geometry optimization with such a dual-level scheme.\textsuperscript{4,5} In the same spirit one may use molecular mechanics for the low level on the entire system; we have recently applied this
technique to treat an enzyme reaction where the subsystem had 25 atoms and the entire system had 8888 atoms.6

Another type of dual-level or multi-level strategy involves applying two or more levels of electronic structure calculation to the entire system. One strategy we have used is to combine the results at the various levels linearly in order to extrapolate to the limit of full configuration interaction (FCI) or an infinite one-electron basis set (IB) or both (full configuration interaction for an infinite basis is denoted complete configuration interaction or CCI). We have developed several variations on this scheme:

<table>
<thead>
<tr>
<th>Description</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAC</td>
<td>scaling all correlation: combine an uncorrelated and a correlated calculation with a given basis to extrapolate to CCI597</td>
</tr>
<tr>
<td>MC-SAC</td>
<td>multicoefficient SAC: combine an uncorrelated and two or more correlated calculations with a single basis to extrapolate to CCI8</td>
</tr>
<tr>
<td>IB</td>
<td>combine calculations with two basis sets to extrapolate to the infinite-basis-set limit for a given level of electron correlation9-11</td>
</tr>
<tr>
<td>MCCM</td>
<td>multicoefficient correlation method: simultaneous application of MC-SAC and IB to extrapolate to CCI.8</td>
</tr>
</tbody>
</table>

We have organized our MC-CM calculations by the shape of a polygon enclosing the methods in the level-basis plane, and we have proposed Colorado, Utah, and New Mexico methods, where the polygons have the shapes of those states.8 In addition we have proposed multicoefficient Gaussian-212 and multicoefficient Gaussian-313 methods that outperform the original Gaussian-2 and Gaussian-3 methods.

For dynamics calculations we sometimes use a bootstrap technique in which high-level electronic structure calculations are carried out at stationary points, and then the parameters of semi-empirical molecular orbital theory, for example, Austin Model 1 (AM1), are adjusted to reproduce these properties as well as possible; the adjusted parameters are called specific reaction parameters or specific range parameters (SRP), and the modified AM1 is called AM1-SRP.14-20 The advantage is that now the AM1-SRP calculation provides a smooth interpolation of the original high-level calculations between the stationary points. We have also used specific reaction parameters in Becke's 3-parameter hybrid Hartree-Fock-density-functional theory that he derived by adiabatic connection; we call this adiabatic connection-SRP (AC-SRP).21

A more powerful version of the AM1-SRP approach is a dual-level scheme in which the energy is given by a linear combination of ab initio Hartree-Fock calculations and AM1-SRP calculations. This is called HF||AM1-SRP.21 We are currently working on another global interpolation scheme which we hope will accomplish many of the goals of SRP methods but in a more systematic way. The new method is called multiconfiguration molecular mechanics (MCMM), and it involves combining high-level ab initio calculations with the valence bond formalism and molecular mechanics calculations.

There are situations in which each of these many strategies may be the best way to achieve an objective. We believe that such multi-level strategies will become increasingly important in the coming years.

Many one-electron basis sets are available for electronic structure calculations, and in essentially all cases they have been developed on the basis of variational energy calculations. But this strategy ignores the fact that many times the molecular energy is not the goal of a calculation. For example, modest basis sets are often used to obtain molecular geometries or reaction paths, at which or along which higher-level calculations are often used to calculate energies. In other applications, modest-sized basis sets are used to calculate partial atomic charges for large molecules for use with
molecular mechanics calculations of the energy. In order to allow more accurate applications of this type, we developed the MIDI! basis set, which is a heteroatom-polarized split valence basis with polarization functions optimized entirely on the basis of geometries and balanced one-electron charge distributions.\textsuperscript{22,23}

Electronically excited states of very large molecules are of great interest but can typically be modeled only by semiempirical molecular orbital theory such as intermediate-neglect-of-differential-overlap for spectroscopy (INDO/S). We have reparameterized this theory for carbonyl compounds leading to a new version called INDO/S2.\textsuperscript{24}

Electronic structure calculations yield not only energies but also molecular properties. We have developed the concept of class IV charges,\textsuperscript{25} which (like some of our methods for energies) are an attempt to transcend the limitations of truncated CI and finite basis sets at as low a cost as possible. Our second class IV charge model, called Charge Model 2\textsuperscript{26} (CM2), has been parameterized for ground states with neglect-of-diatomic-differential-overlap (NDDO) semiempirical molecular theory,\textsuperscript{26} INDO/S and INDO/S2 semiempirical molecular orbital theory,\textsuperscript{24} \textit{ab initio} Hartree-Fock (HF)\textsuperscript{26,27} theory, density functional theory (DFT),\textsuperscript{26} and adiabatic-connection hybrid HF-DFT (AC).\textsuperscript{26} It has been parameterized for electronically excited states with INDO/S and INDO/S2.\textsuperscript{24} These parameterizations have been shown to yield accurate partial atomic charges. The combination of CM2 with HF/MIDI! is a particularly powerful combination.\textsuperscript{26}

LIQUID-PHASE ELECTRONIC STRUCTURE

Whereas a goal of gas-phase electronic structure calculations is to produce potential energies for atomic motions, the corresponding goal of liquid-phase electronic structure calculations is to produce free energies or potentials of mean force. These can be used to calculate free energies of solvation, partition coefficients, solvatochromic shifts, and solvation effects on reaction rates. We have developed Solvation Model 5 as a fifth-generation approach (following SM1, SM1a, SM2, and SM4—SM3 involved the same approach as SM2) to the calculation of electronic wave functions and free energies in liquids (neat and solutions). SM5 is actually a suite of models, each of which may have more than one parameterization. In general, SM5 models approximate the standard-state free energy of solvation as

$$\Delta G_E^S = \Delta G_{\text{ENF}} + G_{\text{COSMO}}$$

where $\Delta G_{\text{ENF}}$ is a bulk electrostatic term includes solute electronic (E) and nuclear (N) energies and solvent electric polarization (P) free energy, and $G_{\text{COSMO}}$ is a semiempirical first-solvation shell term including cavitation (C), dispersion (D), solvent structure (S), and the breakdown of the bulk electrostatic model in the first solvation shell. The models may be applied with gas-phase geometries (in which case we append R to the model name), or geometries may be optimized in liquid solution using analytic gradients (in which case the model name does not end in R). The SM5 models are:

- SM5.0R\textsuperscript{28,29}
- SM5.2R\textsuperscript{30}
- SM5.4\textsuperscript{31–34}
- SM5.42\textsuperscript{21,35–39}
- SM5CR\textsuperscript{40}

SM5 model in which electrostatics are implicit
SM5 model in which electrostatics are based on class II charges
SM5 model in which electrostatics are based on class IV charges calculated by Charge Model 1
SM5 models in which electrostatics are based on class IV charges calculated by CM2
SM5 model in which electrostatics are treated by the Conductor-like Screening Model (COSMO).
All of these models have been parameterized for aqueous solution and for general organic solvents. The parameterizations are based on over 200 free energies of solvation in 91 solvents (and in some cases on additional free energies of transfer as well). The SM5.0R model is designed for use with molecular mechanics calculations. The SM5.2R model is parameterized for AM1, PM3, MNDO, and MNDO/d. The SM5.4 model is parameterized for AM1 and PM3. The SM5.42R model is parameterized for AM1, PM3, HF/MIDI!, HF/6-31G*, HF/6-31+G*, HF/cc-pVDZ, BPW91/MIDI!, BPW91/6-31G*, BPW91/DZVP, B3LYP/MIDI!, INDO/S, and INDO/S2. The SM5.42 model uses the same parameters as SM5.42R. The SM5CR model is parameterized for AM1, PM3, and MNDO/d. In addition to the original papers, some overviews discussing the SM5 models are available.42-44

The SM5.42R/HF/6-31G*, SM5.42R/BPW91/MIDI!, and SM5.42R/AM1 all yield excellent results for the free energy of transfer of nucleic acid bases from water to chloroform.45

GAS-PHASE DYNAMICS

Variational transition state theory with optimized multidimensional tunneling contributions (VTST/OMT) provides an accurate yet practical method for calculating chemical rate constants.46-48 The optimized multidimensional tunneling calculation allows for corner cutting tunneling paths of two kinds: small-curvature tunneling paths, which are localized in the zero-point-amplitude hypertube enclosing the reaction path, and large-curvature tunneling paths, which are straight lines in isoinertial coordinates. The calculations may be carried out in redundant curvilinear coordinates, which are more physical than rectilinear coordinates.49,50

Our recent emphasis has been on devising algorithms for interfacing VTST/OMT calculations with electronic structure calculations in convenient and efficient ways, and we have especially concentrated on minimizing the amount of electronic structure data required because that would make it more affordable to calculate the required data at high levels.

For example, we have recently shown how the orientation of generalized transition state dividing surfaces may be optimized, and how such optimization allows one to increase the step size used to trace the reaction path or even to use local optimization at nonstationary points to obviate the need for a continuous reaction path. The resulting algorithms are called re-orientation of the dividing surface (RODS)51-53 and variational reaction path (VRP).54

Another way to increase the amount of electronic structure data needed is to use sophisticated interpolation schemes. We have developed a method called interpolated variational transition state theory by mapping (IVTST-M) that allows one to greatly reduce the number of energies, gradients, and Hessians required for VTST calculations with small-curvature tunneling contributions.55 Another approach is to use two levels, including geometry optimization at both levels but reaction-path following only at the lower level; by interpolating the difference between the lower and higher levels one can greatly reduce the amount of data needed at the higher level. This approach, called variational transition state theory with interpolated corrections (VTST-IC), can be used for small-curvature, large-curvature, or optimized multidimensional tunneling contributions.54-57 We have shown that the VTST-IC method is much more accurate than simply adding single-point energy corrections at geometries optimized by the lower level.56 The VTST-IC approach is particularly powerful when the high-level
calculations are used not only as corrections to the lower level but also to optimize SRP
parameters for the lower level.

A recent set of applications\textsuperscript{57-59} of these techniques involved calculating the rate
constant for $\text{H} + \text{C}_2\text{H}_4 \rightarrow \text{C}_2\text{H}_5$ and several isotopically substituted versions of these
reactions. This reaction involves small-curvature tunneling, and we treated it using a
dual-level scheme involving variable scaling of external correlation energy, the RODS
algorithm, redundant curvilinear coordinates, and the IVTST-M algorithm. We obtained
good agreement with experiment both for absolute rate constants and for kinetic isotope
effects.

Other recent applications include $\text{H} + \text{N}_2\text{H}_2 \rightarrow \text{H}_2 + \text{N}_2\text{H}_3$,\textsuperscript{17,49} $\text{OH} + \text{C}_3\text{H}_8 \rightarrow \text{H}_2\text{O} +$\textsuperscript{18} $\text{RCI}({\text{PH}_3})_2(\text{CH}_3) \rightarrow \text{RCI}({\text{PH}_3})_2(\text{CH}_3)\text{H}$,\textsuperscript{60} $\text{O} + \text{HCl} \rightarrow \text{OH} + \text{Cl}$,\textsuperscript{61} $\text{Cl} +$\textsuperscript{20} $\text{H}_2 \rightarrow \text{HCl} + \text{H}$.\textsuperscript{62}

\section*{SOLUTION-PHASE DYNAMICS}

We have extended the gas-phase variational transition state theory formalism to treat
reaction rates in solution, including curvilinear coordinates, small- and large-curvature
tunneling, optimized multidimensional tunneling, and equilibrium and nonequilibrium
solvation effects. The methods are designed to take full advantage of the advances
discussed in all three sections above. We distinguish three general levels of liquid-phase
dynamics calculations:

\begin{itemize}
  \item SES separable equilibrium solvation\textsuperscript{63}
  \item ESP equilibrium solvation path\textsuperscript{21,63,64}
  \item NES nonequilibrium solvation\textsuperscript{65,66}
\end{itemize}

In SES calculations, one calculates stationary point geometries and/or reaction paths in
the gas-phase and then adds the free energy of solvation to each point. In ESP
calculations one first creates a potential of mean force surface by adding free energies of
solvation to the gas-phase potential energy surface; then stationary point geometries
and/or reaction paths are computed using the potential of mean force. Tunneling
calculations involve a function of the potential mean force called the canonical mean
shape potential, which reduces to the potential of mean force in zero order. Nonequilibrium effects may be incorporated via coupling of the solute to collective solvent coordinates.

We have recently demonstrated all three levels of theory for the reaction $\text{H} +$\textsuperscript{21,66} $\text{CH}_3\text{CH}_2\text{OH} \rightarrow \text{H}_2 + \text{CH}_3\text{CHOH}$ in aqueous solution.

\section*{ACKNOWLEDGMENTS}

This research is supported by the National Science Foundation under grant no. CHE97-
25965 (liquid-phase electronic structure and dynamics research and part of the gas-phase
electronic structure research) and by the U.S. Department of Energy, Office of Basic
Energy Sciences under grant no. DOE-DE-FG02-85ER13579 (gas-phase kinetics
research and part of the gas-phase electronic structure research).

\section*{REFERENCES}

457
66. Chuang, Y.-Y.; Truhlar, D. G., to be published.