J. C. Corchado and D. G. Truhlar, in *Combined Quantum Mechanical and Molecular Mechanical Methods*, J. Gao and M. A. Thompson, Eds.; ACS Symposium Series, volume 712; American Chemical Society: Washington, DC, 1998; pp. 106-127.

Chapter 7

Dual-Level Methods for Electronic Structure Calculations of Potential Energy Functions That Use Quantum Mechanics as the Lower Level

José C. Corchado[1,2] and Donald G. Truhlar[1]

[1]Department of Chemistry and Supercomputer Institute, University of Minnesota, Minneapolis, MN 55455–0431
[2]Departamento de Química Física, Universidad de Extremadura, Badajoz 06071, Spain

This chapter overviews the status of three dual-level approaches for potential surface calculations that use quantum mechanics for both the upper and lower level. Three types of approach are singled out for discussion: SEC and SAC calculations, IMOMO calculations with harmonically capped correlated small systems, and dual-level direct dynamics. The scaling external correlation (SEC) and scaling all correlation (SAC) methods are semi-*ab initio* approaches to the calculation of bond energies and barrier heights for chemical reactions. The IMOMO calculations are very similar in spirit to QM/MM methods, but the lower level is quantum mechanical. Dual-level direct dynamics is a general technique for combining levels in dynamics calculations that include quantum mechanical tunneling contributions.

Dual-level methods have become very popular in modern quantum chemistry. The recent surge of interest in methods that combine quantum mechanical solutions for a small part of a system with molecular mechanics for the rest of the system (*1–6*) prompts one to place these methods in the perspective of a larger set of dual-level and multi-level approaches. These approaches have certain elements in common, but additionally they have interesting differences. The purpose of the present overview is to discuss some dual-level methods from this viewpoint.

The diversity of dual-level methods is such that even classifying them creates a stimulating challenge (and, like all review tasks, is doomed to incompleteness). Nevertheless, here is an attempt to enumerate some of the main varieties:

1. Double slash (//) methods for energies of stationary points. This is perhaps the most ubiquitous dual-level approach. A common example would be an MP2/6-311G(*d,p*)//HF/6-31G(*d*) calculation in which a stationary-point geometry is optimized at the HF/6-31G(*d*) level, and then a more accurate energy is calculated by the MP2/6-311//G(*d,p*) method (*7*). This method can be used for both minima and saddle points. One can denote such a theory L2//L1 where L2 and L1 denote the higher and lower

"levels." A variant is L3//L2[L1] where the geometry is optimized at level L1 and vibrational frequencies are also calculated at this level; then the geometry is optimized at a higher level L2, but without a vibrational frequency calculation. The L2 geometry is then used for a single-point calculation at level L3.

2. Gaussian-2 theory. The Gaussian-2 (often abbreviated G2) approach to estimating thermodynamic quantities (*8*) has been extremely successful, and has spawned several variants. G2 theory and its variants are actually carefully layered multi-level theories. The theory assumes that various effects of improving the lower level electronic structure calculation (i.e. by using a larger one-electron basis set or a more accurate treatment of the electron correlation) are additive. Thus, instead of carrying out a full higher level calculation, a sequence of "intermediate" level calculations directed to estimating the effects of several enhancements on the lower level is carried out.

3. CBS. The truncation of the one-electron basis sets used in *ab initio* calculations is frequently the major source of error in the results. The increase in the basis set size involves an increase in the required computational resources that makes impossible to converge a calculation with respect to the size of the basis set even in most small systems. When the interest is mainly focused on differences in the calculated energy for a series of systems or a series of points on a potential energy surface, one can expect some compensation between the errors introduced in the various energy calculations, but even for relative energies, the incompleteness of the one-electron basis set may be severe. In order to minimize the effects of the basis set truncation, an extrapolatory scheme derived by means of perturbation theory and based on the asymptotic basis-set dependence of the pair natural orbital energies has been proposed (*9*). The goal is to obtain an energy comparable to the one that would be obtained with a complete basis set (CBS). Several variants have been proposed (*10*), accounting for a variety of choices of basis sets and methods for including correlation energy.

4. SEC and SAC. The calculation of bond energies is a critical step in the calculation of heats and free energies of reaction, and the calculation of barrier heights is a first step in the calculation of enthalpies and free energies of activation. Unfortunately, although there has been great progress in *ab initio* electronic structure methods, including the G2 and CBS approaches mentioned above, the calculations can be very expensive. Thus semiempirical dual-level approaches can also prove useful, especially for calculating barrier heights, and in this paper we discuss two such methods, namely the scaling-external-correlation (SEC) (*11*) and scaling-all-correlation (SAC) (*12–15*) methods. These methods are designed to circumvent the difficulty that the differential electron correlation energies between species with made and broken bonds or partial broken bonds (as at transition states) are very slowly convergent with respect to the mixing of electronic configurations and the completeness of the one-electron basis sets for electronic structure calculations.

Section 2 of the present chapter provides a review of the SEC and SAC methods.

5. QM/MM methods. The most rapidly growing area of activity in dual-level methods is in the development of algorithms that combine quantum mechanics (QM) with molecular mechanics (MM), where "molecular mechanics" refers to any "classical" force field. Since this subject is covered extensively by the other chapters in the present volume, we will not review it here, but we simply mention that there are two main approaches. The first approach, which is more common (*1–5*), involves applying QM to the primary subsystem and MM to the rest. The second approach, called IMOMM (*6*), involves applying QM to the primary subsystem and applying MM to both the primary subsystems and the entire system—then combining all three calculations. Both methods involve technological challenges in how to treat the boundary, where one locates so called "link atoms" and/or "capping atoms."

6. IMOMO methods. An appealing feature of the IMOMM method mentioned above is that it is very general as to the kinds of levels that are combined, and in particular one can equally well choose a QM level as an MM level for treating the entire system. This is called

IMOMO (*16*). For moderate-size systems, the QM level can be as high as one can afford (*17*), although for very large systems, economic considerations may limit one to semiempirical molecular orbital theory. An alternative approach for very large systems is to use a triple-level scheme, and IMOMO is easily generalized to accommodate this (*18*).

Section 3 of the present chapter reviews two aspects of IMOMO. The first is its ability to include electronic substituent effects ("inductive effects") of the rest of the system on the primary subsystem (*17,19,20*). Second is a new approach we have developed, called a harmonic cap, for optimizing geometries with IMOMO methods (*21*). The approach is actually equally applicable to QM/MM methods, and we believe that it eliminates the "link atom problem," and so it is liable to be particularly interesting to readers of this volume.

7. SCRF/SASA models. Another dual-level approach, albeit of quite a different nature, has been employed for the problem of modeling solvation effects. In this approach, one combines a quantum mechanical method for electrostatics with a classical mechanical method for first-solvation shell effects. In particular one uses the self-consistent reaction field (SCRF) molecular orbital method (*22*) for the former with a treatment based on solvent-accessible surface area (SASA) (*23,24*) and atomic surface tensions (*25,26*) for the latter. The most completely developed set of models of this type is the SM*x* series (*27*) of "solvation models." These have culminated in a suite of models called SM5 models, and these are overviewed in another chapter in this volume (*28*).

8. Dual-level direct dynamics. Direct dynamics is the calculation of dynamical quantities using electronic structure calculations for all required energies, forces, and Hessians, without the intermediacy of an analytic potential energy function (*29–31*). The treatment of atomic motion in direct dynamics calculations can be either classical (*29,32*) or quantal, although quantal effects on nuclear dynamics are included most efficiently by semiclassical (*30,31,33*) methods. (The distinction should not be confused with the use of quantum mechanics or classical molecular mechanics for the potential energy function, which is an entirely separate issue.) Dual-level direct dynamics has been developed so far for calculations based on quasiclassical variational transition state with semiclassical multidimensional tunneling contributions, usually abbreviated VTST/MT for variational transition state theory with multidimensional tunneling (*30,31,33-38*). In practice we use up to four levels, e.g., for OH + NH_3 (*37*) we used QCISD(T)//MP2/aug-cc-pVTZ[MP2/aug-cc-pVDZ]///MP2/6-31G* and ...///PM3-SRP, where L3//L2[L1] is explained above and /// separates the method for high-level calculations at stationary points (before the triple slash) and the method used (prior to interpolation corrections) for everywhere else (after the triple slash). The notation PM3-SRP denotes the use of specific reaction parameters (*30*) starting from the PM3 general parameters (*39*) at the neglect of diatomic difference overlap (NDDO) level of semiempirical molecular orbital theory.

Section 4 of this chapter presents a brief review of dual-level direct dynamics and SRP methods.

2. SEC and SAC

SEC. Correlation energy is defined as the difference between the Hartree-Fock energy and the true energy (*40*). This is usually calculated by optimizing the orbitals of a restricted form of wave function, called the reference state, and adding electron correlation by mixing in additional configurations, either by a variational configuration interaction (CI) method or by perturbation theory or a coupled cluster approach. For some molecules, primarily open-shell systems, transition states, systems with partially broken bonds, and systems with low-lying excited states, the Hartree-Fock self-consistent-field (HF-SCF) method based on a single configuration does not provide a good zero-order description. In such cases, we need a multi-configuration SCF (MCSCF) reference state for a qualitatively correct description (*41*), and when correlation energy is added to such a state, the resulting

calculation is called a multi-reference one. The correlation energy may be decomposed into two parts: internal correlation, also called static correlation, which is the part included in a minimal MCSCF wave function, and external correlation (also called dynamical correlation), which is the remainder even after a good zero-order description has been achieved. The remainder would typically be calculated by a multi-reference CI calculation (MRCI) (*42*). One could equally well include such configuration mixing effects by multi-reference perturbation theory (MRPT), e.g., CASPT2 (*43*). Computationally we have found (*11*) that the fraction of external correlation energy is more a function of the level of theory and the one-electron basis set than of the geometry. By taking advantage of this for a given level, one can scale the external correction energy by a factor that is independent of geometry. In particular, if one assumes that one recovers a fraction F of the external correlation energy, independent of the geometry of the system, one can write (*11*)

$$E(SEC) = E(MCSCF) + \frac{E(MRCI) - E(MCSCF)}{F}. \qquad (1)$$

The SEC method has proven particularly useful for the calculation of barrier heights for chemical reactions. In this case one determines the value of F such that the reactant and product bond energies are correct or the overall exoergicity is correct. Then with F determined on the basis of reactants and products, it can be used at transition state geometries to calculate barrier heights. Some examples (*44–47*) of SEC calculations of barrier heights are provided in Table 1. In all cases, the barrier heights calculated by the SEC method or potential energy surfaces (*47–50*) adjusted to have a barrier height equal to the SEC barrier heights are consistent with experiment.

SAC. Scaling the entire correlation energy, even the internal part, is expected to be less reliable than scaling the external correlation energy, but it is computationally much more tractable since MRCI and MRPT calculations are much more difficult than single-reference ones based on a single-configuration reference state. (A single-reference correlated calculation is one in which there is only a single configuration in the zero-order wave function that serves a starting point for estimating the effect of configuration mixing.) Thus scaling all correlation (SAC) can be very useful. In this approach we write

$$E(SAC) = E(HF) + \frac{E(SRPT) - E(HF)}{F} \qquad (2)$$

Table 1. Computation of barrier heights by MRCI-SEC calculations

Reaction	Barrier (kcal)		Ref.
	MRCI	SEC	
$F + H_2$ collinear	3.7	1.6	(*44*)
bent	...	1.3	(*45*)
$H + DF \rightarrow HCl + D$	43.2	38.4	(*11*)
$H + DCl \rightarrow HCl + D$	20.0	18.1	(*46*)
$Br + H_2$	2.7	1.9	(*47,48*)
$H + DBr \rightarrow HBr + D$	13.4	11.9	(*47,48*)

where *HF* denotes Hartree-Fock and *SRPT* denotes single-reference perturbation theory. One would probably not want to use single-reference CI (SRCI) in eq. (2) because it has severe size-consistency failures, whereas SRPT and single-reference coupled-cluster theory are size consistent. (We use "size consistency" and "size extensivity" as synonymous here. Size consistency (*51*) is not strictly required for using eq. (2), but a method that is not at least approximately size extensive would not be expected to have a geometry-independent F value. MRCI is not strictly size extensive, but it is better than SRCI.)

In early work with the method (*12–14*) we developed general scale factors for various types of bonds with both Møller-Plesset second-order (MP2) and Møller-Plesset fourth-order (MP4) perturbation theory, which are single-reference methods. These general scale factors were determined for several common basis sets. In later work (*15*), we considered other single-reference correlated methods, in particular the coupled-clusters method with double excitations (CCD) (*52,53*) and the quadratic configuration interaction method with single and double excitations (QCISD) (*54*). (The latter method, although named as a CI method, is actually size extensive.)

The basic conclusions of the early systematic studies of the SAC method (*12–14*) are that F is indeed reasonably consistent across systems (for a given basis set and order of perturbation theory), and furthermore there does not appear to be much advantage in MP4-SAC as compared to MP2-SAC. It is instructive to summarize some of the recent results (*15*) we obtain with the SAC method for bond energies. In order to test the SAC method with general scale factors, we considered 13 molecules with experimentally known atomization energies. This test suite consists of H_2, CH_4, NH_3, H_2CO, H_2O, HF, C_2H_2, HCN, F_2, CO, CO_2, N_2, and N_2O, and we will consider the results for the cc-pVDZ (*55*) basis set. For the MP2-SAC calculations, we found that we could not get particularly accurate results with a constant value of F (constant across systems), so we chose $F = 0.918 - 0.143x - 0.176x$, where x is the fraction of atoms that are H. (Note: that F is still independent of geometry, e.g., we use the same value for $CH_3 + H$ as for CH_4). We found a mean |error| per bond of 1.3 kcal per mol. For QCISD-SAC, we used a completely constant value of $F = 0.729$, and we found an average |error| per bond of 1.5 kcal. For CCD-SAC we used $F = 0.729$, and found an average |error| per bond of 1.4 kcal. Note that the average |error|s per bond without scaling the correlation energy are respectively 8.4, 14.5, and 15.6 kcal/mol. The results of the scaling are amazingly good considering none of the three correlation levels considered involves explicit consideration of triple excitations.

Recently Siegbahn *et al.* (*56,57*) proposed a closely related method, which they call PCI-X. In actual applications, PCI-X is basically the same as SAC except that these authors prefer to use a single universal F value.

Correlation Balance. Clearly the success of the SEC and SAC approaches is dependent on the quality of the basis set and in particular on its ability to treat the correlation energy of various geometries and bonds in a consistent way. The idea that a basis set should be electrostatically balanced so that it yields accurate dipole moments and bond polarities is a familiar one, but correlation balance has been less widely studied. Once one attempts to design extrapolation methods (*8–15*) for electronic structure calculations, the question of correlation-balanced basis sets becomes paramount, not only for basis sets but also for approaches to configuration mixing. For example, it is clear from the necessity (see above) to use a variable F in MP2-SAC that coupled-cluster theory gives a more uniform treatment of correlation in bonds to hydrogen and bonds between heavy atoms than MP2 theory does. Some attempts to study the question of correlation balance have been presented (*13,58–61*) but much more work needs to be done.

SAC calculations for kinetics. The area where the SAC method has proved most useful so far (*12,14,58–70*) is for kinetics calculations. This subsection reviews this body of work.

The first application of a SAC method to a reaction was an application of the MP2-SAC scheme for the calculation of the properties of the stationary points along the

$CH_4 + Cl \rightarrow CH_3 + HCl$ reaction-path (*58*). A 6-311G(*2d,d,p*) basis set (*71*) was found to provide a balanced treatment of the effects of electronic correlation on the making and breaking bonds. Therefore, the average of the *F* values obtained for each bond using the MP2/6-311G(*2d,d,p*) was used for the scaling. The MP2-SAC/6-311G(*2d,d,p*) level was chosen for the optimization of the geometries and frequency calculations at the reactants, products, and saddle point for this reaction. The properties obtained in this work for the saddle point were employed as reference data for the calibration of an analytical potential energy surface (*72*), which was used in a VTST/MT calculation of the rate constants and kinetic isotope effects for this reaction.

The first application of MP2-SAC to the calculation of a rate constant was the study of the hydrogen abstraction from CH_4 by the OH radical (*62*). In this study, the basis set employed for SAC calculations was the 6-311G(*2d,p*) basis set, since this is the basis set that provided the most balanced treatment of the correlation energy for the O–H and C–H bonds. In particular, the values of the factor *F* estimated from the H–OH and H–CH_3 bond dissociation energies were 0.875 and 0.864 respectively. The properties of reactants, products, and saddle point were therefore evaluated at the MP2-SAC/6-311G(*2d,p*) level and used in a calculation of the rate constant by means of the zero-order interpolated variational transition-state theory (IVTST–0) (*73*), giving rate constants in good agreement with the experimental results.

The availability of a set of average values for the *F* factors (*12*) allows us to apply the SAC corrections in a more automatic way, without a check of the balance in the basis set for the treatment of the correlation energy. Thus, the MP2-SAC method was also used for calculating single-point energies in a study on the effects of hydration and dimerization of the formamidine rearrangement (*63*). In this study, where the MP2-SAC/6-31G(*d,p*) method was used for calculating only single-point energies, the value of the *F* factor utilized was the average between two previously determined values of the *F* factor, one for the N–H bond and one for the O–H bond. In a study on hydrogen abstraction reactions of ammoniacal compounds (*64*) involving systems for which experimental dissociation energies are not available, the MP4-SAC method was used with *F* equal to the average of the factors proposed in the original SAC paper for the basis sets and bonds involved in the calculations.

The same procedure was used in the calculation of the barrier height for the $NH_3 + H \rightarrow NH_2 + H_2$ reaction (*65*). Table 2 shows a summary of the results obtained for this reaction, in which an MP4-SAC/6-311+G(*d,p*) single point calculation is compared to other schemes for extrapolating the Møller-Plesset series (*74*, *75*). The MP4-SAC results are close to the more expensive QCISD(T)/6-311+G(*d,p*) results, and the use of an average *F* factor leads to a reasonable reaction endoergicity. We note that since the endoergicity is the difference between two bond energies, and since *F* is calculated using experimental bond energies, the use of an *F* factor obtained from one bond for calculations involving a different bond will give energies for this second bond closer to the experimental results when the *F* values for both bonds are more similar. Therefore, the accuracy in the endoergicity is a reflection of the adequacy of an average *F* factor for different reactions. The rate constants obtained using the MP4-SAC results (*65*, *66*) are in good agreement with the experimental results. The same procedure was used for a VTST/MT study of the reaction $OH + NH_3 \rightarrow NH_2 + H_2O$ (*67*, *68*). The MP4-SAC method with a standard *F* factor (*14*) has also recently been used for calibrating an analytical potential energy surface for the $Cl + H_2 \rightarrow ClH + H$ reaction (*69*).

A series of papers have been addressed to understanding the kinetics of the reactions between hydrocarbons and the OH radical. These studies involve the reactions of OH with methane (*59*, *70*), ethane (*60*), and propane (*61*). In all the cases the MP2-SAC method was employed. A modified triple-ζ basis set was obtained in order to get a balanced

Table 2. Barrier height and endoergicity for the NH_3 + H reaction calculated at different levels of calculation[a]

Method	ΔE[b]	$\Delta E^{\neq}$[c]
UHF	-0.2	23.7
PUHF	-2.0	18.6
UMP2	10.8	23.4
PUMP2	9.7	20.0
UMP4	5.3	18.2
PUMP4	4.6	16.2
MP4-SAC	5.7	15.7
PMP∞[d]	4.4	16.1
Feenberg[e]	4.1	15.5
QCISD(T)	5.3	16.8
Exp[f]	6.9	

[a]Results taken from Ref. (*65*). [b]ΔE is the Born-Oppenheimer classical energy of reaction, i. e., it excludes zero point energy. [c]$\Delta E^{\neq}$ is the Born-Oppenheimer classical energy of activation. [d]Extrapolation of the Moller-Plesset series by the methods in Ref. (*74*). [e]Extrapolation of the Moller-Plesset series by the methods in Ref. (*75*). [f]Ref. (*76*).

treatment of the electron correlation for the C–H and O–H dissociation energies, followed by a MP2-SAC single-point energy calculation, except in the propane reaction, for which a geometry optimization was also carried out at the MP2-SAC level. This was part of the input necessary for carrying out calculations of the rate constant and kinetic isotope effects. In the methane and ethane reactions, this information was completed with 2 or 1 extra points on the reaction path, respectively. Then a dynamical calculation was carried out by IVTST methods. The reaction OH + propane was simulated using a dual-level direct dynamics technique (see Section 4) that allows a more reliable calculation of the tunneling contribution.

In conclusion we can say that the use of SAC methods with economical dynamical methods has been a powerful tool for the accurate study of chemical reactions. Nevertheless the SAC methods, although being a good approximation to more accurate, much more expensive calculations, still require a large amount of computer resources for systems with many atoms. Thus, while the treatment of the correlation energies can be limited to a very affordable MP2 calculation, a high quality basis set is still required. The advances in computer hardware and software are opening doors for the study of medium-size systems, such as propane, but we are still far from being able to perform this kind of calculations in the large organic or biological systems, as well as condensed phase systems. In this sense, only the IMOMO (or related) methods seem to be able to propel us ahead.

3. IMOMO

CCSS. A special case of IMOMO is the use of a correlated capped small system (CCSS). Consider, for example, a calculation of the bond energy for the C-H bond in $H\text{-}CH_2CH_2NH_2$. One can take the small system as $H\text{-}CH_2$ and the capped small system as $H\text{-}CH_3$. One could calculate the entire system ($H\text{-}CH_2CH_2NH_2$) only at the lower level, e.g., AM1, HF/3-21G, HF/MIDI!, HF/6-31G*, or B3PW91/6-31G* (or—in the case of IMOMM—the lower level would be MM). One could calculate the capped small system at both lower and higher levels, where the higher (correlated) level might be, e.g., CCSD(T)/6-31G(*d,p*), QCISD(T)/cc-pVTZ, MP2/6-31G(*d,p*), or HF/cc-pVTZ. [The methods and basis sets are explained elsewhere: Ref. (*7*) for MP2, HF, 3-21G, 6-31G*, and 6-31G(*d,p*), Ref. (*77*) for AM1, Ref. (*78,79*) for B3PW91, Ref. (*80*) for CCSD(T), Ref. (*54*) for QCISD(T), Ref. (*81*) for MIDI!, and Ref. (*55*) for cc-pVTZ.] Then the dual-level energy, called the integrated (I) energy, is

$$E^I_{\text{entire}} = E^{\text{low}}_{\text{entire}} + (E^{\text{high}}_{\text{small}} - E^{\text{low}}_{\text{small}}) \tag{3}$$

$$= E^{\text{high}}_{\text{small}} + (E^{\text{low}}_{\text{entire}} - E^{\text{low}}_{\text{small}})$$

Table 3 gives examples of how well this approach works for the molecule given as an example above. Table 4 gives surveys of the performance for several small systems of this type (*17,20*). In these tables, the entire molecule is of the form $H\text{-}CHXCH_2Y$, and the capped small system is $H\text{-}CH_2X$, where X and Y are H, CH_3, NH_2, OH, and F. In both Table 3 and Table 4, standard geometries were used for all calculations—the goal was to test how well electronic effects are included when one uses the same geometry for the integrated calculation as for a more accurate calculation. Notice that in these examples the system is capped very close (geminal) to the bond that is broken, thus providing a challenging test. The goal is to combine a high-level calculation on the capped small system with a low-level calculation on the entire system such that the integrated result is more accurate than either single level. Accuracy is measured by deviation from a full high-level calculation on the entire system (which is feasible for these test cases but which would presumably be unaffordable for the entire system in real applications to large systems). The table shows that the integration of levels is successful, in fact quite dramatically successful. For example, in Table 3 we see that a dual-level calculation is much better than either a higher-level calculation on the capped subsystem or a lower-level calculation on the entire system.

We note specifically the especially good results obtained with the MIDI! basis set in Tables 3 and 4. This is very encouraging because the MIDI! basis set is the first example of a new breed—a basis set optimized specifically to serve as the lower level in dual-level methods (*81*). In particular the MIDI! basis is optimized to give good geometries so it can serve as L2 in L1//L2 calculations and to give balanced charge distributions so it can serve as a starting point for multi-level calculations in the condensed phase, where electrostatics is often of crucial importance.

IMOHC. The second topic we consider in IMOMO theory is the integrated molecular orbital harmonic cap (IMOHC) strategy (*21*) for geometry optimization in IMOMO and IMOMM calculations. One of the most challenging problems that one faces in most QM/MM applications is the treatment of the link atom. (Note: The "link atom" is the atom that becomes the capping atom in the capped small system.) In the case of a geometry optimization the problem of the link atom is very delicate, since usually the link atom has a different nature in the entire system and the capped small system. If the optimization is done in such a way that this atom is constrained to have identical geometrical parameters (e.g., bond lengths) in both systems, the optimization may be unphysical, or it may fail.

Table 3. Bond energy for H-$CH_2CH_2NH_2$ by IMOMO[a]

System	Method	D_e (kcal/mol)	error (kcal/mol)
entire system	QCISD(T)/cc-pVTZ	106.4[b]	0.0
capped small system[c]	QCISD(T)/cc-pVTZ	111.3	4.9
entire system	AM1	84.4	22.0
integrated	QCISD(T)/cc-pVTZ: AM1	105.6	0.8
entire system	HF/MIDI!	79.7	26.7
integrated	QCISD(T)/cc-pVTZ:HF/MIDI!	106.6	0.2

[a]From Ref. (*17*) [b]Presumed accurate. All other calculations are compared to this value using standard geometries. [c]H–CH_3

Table 4. Mean unsigned errors (kcal/mol) in bond energies for H – $CHXCH_2Y \rightarrow H + CHXCH_2Y$

higher level	lower level	< |error| >		
		higher level capped small system	lower level entire system	integrated
	test set 1[a] (nine cases; X,Y = H, CH_3, NH_3, OH, F)			
QCISD(T)/cc-pVTZ	HF/6-31G*	2.1	23.9	0.4
"	HF/MIDI!	2.1	26.1	0.3
"	HF/3-21G	2.1	23.3	0.6
"	AM1	2.1	23.4	1.3
	test set 2[b] (ten cases; X, Y + H, NH_2, OH, F, Cl)			
MP2/6-31G(d,p)	HF/3-21G	1.6	21.4	0.9
CCSD(T)/6-31G(d,p)	HF/3-21G	1.9	21.3	0.9
"	HF/6-31G*	1.9	22.1	0.3

[a]From Ref. (*17*). [b]From Ref. (*20*)

The IMOHC strategy is to circumvent this difficulty by allowing an independent optimization of the link atom in the entire and capped small systems. Thus, if the entire system is formed by a number of atoms N_{entire}, the geometry optimization will involve $3(N_{\text{entire}} + 1)$ coordinates, instead of the usual $3N_{\text{entire}}$. The extra coordinates occur because the link atom occurs both as the link atom itself (denoted L) in the entire system and also, with different coordinates, as the capping atom (denoted M) in the capped small system. All the other atoms in the capped small system have the same coordinates as in the entire system.

The geometry optimization is finished when the $3(N_{\text{entire}} + 1)$ components of the gradient are zero. By applying the distributive law of differentiation, we can write an equation similar to the first half of the equation 3 for the gradients:

$$\nabla E^{I}_{\text{entire}} = \nabla E^{\text{low}}_{\text{entire}} + \left(\nabla E^{\text{high}}_{\text{small}} - \nabla E^{\text{low}}_{\text{small}}\right). \tag{4}$$

An examination of this equation leads us to the conclusion that it will not generally succeed in giving a physical location for the link atom. The behavior we should expect from the optimization algorithm is to find the geometrical conformation that gives nonzero values that cancel out each other for each component of the three terms on the right-hand side of equation 4, thereby giving a zero value for the gradient. However, in practice, the optimization may lead to a geometry with the link atom infinitely separated from the rest of the system, which gives zero components of the gradient for the individual terms for the gradient components related to the link atom. In practice, equation 4 can also lead to unphysically distorted geometries for which $E^{\text{high}}_{\text{small}} - E^{\text{low}}_{\text{small}}$ is negative although $E^{\text{high}}_{\text{small}}$ and $E^{\text{low}}_{\text{small}}$ are both high and positive. This is obviously an undesirable situation, since the link atom of the capped small system (i.e., the capping atom) needs to be in a physical location in order to carry out its capping role in the electronic structure calculations on the small system.

The solution we devised for this problem is to include additive harmonic terms that keep the capping atom at a physically meaningful location in the small system. Thus, the first harmonic cap term we introduce is a correction that prevents the optimization algorithm from locating the capping atom unphysically far from the atom, A, to which it is bonded in the small system. This is given by

$$T_R = \tfrac{1}{2} k_R (R_{\text{A-M}} - R_{\text{eq}})^2. \tag{5}$$

The parameters k_R and R_{eq} are calculated by means of three single-point energy calculations for the capped small system, so that the computational cost is kept to a minimum and accurate results are obtained (*21*).

Similarly, in order to avoid unphysical values for the angles in the capped small system, a second harmonic correction can be introduced,

$$T_\theta = \tfrac{1}{2} k_\theta (\Delta\theta)^2. \tag{6}$$

We define the angle $\Delta\theta$ by

$$\Delta\theta = \arccos(\hat{R}_{\text{A-M}} \cdot \hat{R}_{\text{A-L}}), \tag{7}$$

so that the link atom L and capping atom M have the same orientation. The calculation of k_θ is carried out by means of two or three single-point energy calculations for the capped small system.

Finally, a third term can be defined in order to restrict the possible values of torsion angles in the capped system,

$$T_\phi = \tfrac{1}{2} k_\phi (\Delta\phi)^2, \tag{8}$$

where ϕ indicates the deviation of an L–A–B–C dihedral angle in the entire system from the M–A–B–C dihedral angle in the capped subsystem, where B is an atom bonded to A, and C is an atom bonded to B. Once again, k_ϕ will be estimated by means of a few single-point energy calculations. (Neither of the systems considered in this paper has an L–A–B–C dihedral angle because the capped small system is so small, but we mention the torsion for completeness.)

The total energy for the system will therefore be given by:

$$\begin{aligned} E^I_{\text{entire}} = E^{\text{low}}_{\text{entire}} + (E^{\text{high}}_{\text{small}} - E^{\text{low}}_{\text{small}}) + \\ + \tfrac{1}{2} k_R (R_{\text{A–M}} - R_{\text{eq}})^2 + \tfrac{1}{2} k_\theta (\Delta\theta)^2 + \tfrac{1}{2} k_\phi (\Delta\phi)^2 \end{aligned} \tag{9}$$

The IMOHC method has been tested for the optimization of the geometry of ethane and also for calculating its vibrational frequencies and the C–H bond energy. Some results are shown in Tables 5 and 6. For the results in these tables, k_R and R_{eq} were calculated from three higher-level single-point calculations on methane, with one at the lower-level minimum-energy geometry and the other two having slightly shorter and larger C–H bonds. For this system we found that the only capping term needed is the bond-length controlling term; the parameters k_θ and k_ϕ were therefore set to zero.

Since the most one can expect of any integrated method is to reproduce the results obtained from a complete higher-level calculation, the IMOHC results are compared to those obtained by means of the higher level used in the calculation, MP2/6-31G, as well as to those for the small system (methane and methyl radical) at the higher level of calculation and those for the entire system as the lower level (HF/3-21G). The optimized geometry is much closer to the higher-level optimized geometry than to either the higher-level capped small-system or the lower-level entire-system calculation. The results for the C–H bond energy (Table 6) are also excellent, the error being less than 0.1 kcal/mol, undoubtedly preferable to the 2.6 and 13.7 kcal/mol errors of the single-level calculations.

In a second test, the geometry of the ethylamine molecule (the same molecule as studied in Table 3) was optimized, using the same levels as in the ethane test just described. Once again, the capped system was the methane molecule, with the capping atom being a hydrogen atom and the link atom being a carbon atom. The k_R and R_{eq} parameters were taken from the previous example. In this case, it was necessary to also include the harmonic term for the bend. The bending force constant k_θ was calculated by means of a single-point MP2/6-31G//HF-3-21G calculation at the HF/3-21G optimum geometry for methane, and an MP2/6-31G calculation on a geometry for which one of the C–H bonds in methane deviated by 0.5 degrees from its optimal position at the lower level. (In a less symmetric system, we could need three points.) We obtained $k_\theta = 2.033\text{x}10^{-2}$ kcal mol^{-1} deg^{-2}. This calculation is somewhat arbitrary, since the resulting force constant is dependent on the direction in which the distortion took place. However, we expect that the final results will generally show little dependence on the precise value of k_θ.

Table 5. C–H bond distance (in Å) as predicted by single-level and IMOHC calculations[a]

Level	H–CH_3	H–CH_2	H–CH_2CH_3	H–$CHCH_3$
HF/3-21G	1.0830	1.0717	1.0841	1.0734
MP2/6-31G	1.0959	1.0830	1.0988	1.0867
MP2/6-31G:HF/3-21G		1.0973	1.0854	

[a]All the results in this table are taken from Ref. (*21*).

Table 6. Energy of the C–H bond (kcal/mol) as predicted by single-level and IMOHC calculations[a]

Molecule	Level	ΔE^b	ΔH_0^0
CH_4	MP2/6-31G	99.66	89.90
C_2H_6	HF/3-21G	84.15	73.58
C_2H_6	MP2/6-31G	97.02	87.25
C_2H_6	MP2/6-31G:HF/3-21G	97.09	87.27

[a]All the results in this table are taken from Ref. (*21*). [b]ΔE is the Born-Oppenheimer classical energy of dissociation, i. e., it excludes zero point energy; ΔH_0^0 is the standard-state enthalpy of dissociation at 0 K, including zero point energy.

Table 7. Average unsigned error in the three Cβ–H bond distances in ethylamine (in Å) as predicted by single-level and IMOHC calculations. The reference value is the MP2/6-31G average Cβ–H bond length in ethylamine

Level	H–CH_3	H–$CH_2CH_3NH_2$
HF/3-21G	0.0157	0.0149
MP2/6-31G	0.0028	0.0000[a]
MP2/6-31G:HF/3-21G		0.0018

[a]Zero by definition

Table 8. Average unsigned error in bond angles in ethylamine (in degrees) as predicted by single-level and IMOHC calculations. The reference value is the MP2/6-31G average C_α–C_β–H bond angle in ethylamine[a]

Level	H–CH_3	H–$CH_2CH_3NH_2$
HF/3-21G	0.985	0.187
MP2/6-31G	0.985	0.000[b]
MP2/6-31G:HF/3-21G		0.078

[a]If the hydrogens are labelled D–J as follows: $DENC_\alpha FGC_\beta HIJ$, then the quantity tabulated is the average of the unsigned errors in these three bond angles: $HC_\beta I$, $HC_\beta J$, and $IC_\beta J$. [b]Zero by definition

Table 7 shows the average unsigned error in the C–H bond distance for the β-carbon (where we label the molecule $H_2NC_\alpha H_2C_\beta H_3$), and Table 8 gives the average error in three H–C_β–H bond angles of the small system computed at each single level and also by the dual-level calculations. The "errors" are actually deviations from the optimized MP2/6-31G geometry. Once again, the results are encouraging. The dual-level calculation gives bond distances and bonds angles in the small system much closer to the high-level entire-system calculation than are obtained from either a low-level calculation on the entire system or a high-level calculation on the capped small system.

4. Dual-level direct dynamics

The dual-level direct dynamics approach (*36-38*) is an analog for reaction dynamics of the // approximation in electronic structure calculations on bound-state properties. In the dual-level approach a reaction path is constructed and a complete variational transition state theory calculation including tunneling is carried out at a lower level; the results are then improved by carrying out a reduced number of electronic structure calculations at a higher level without recalculating the reaction path. If HL and LL denote the higher and lower levels, then HL///LL denotes the result of using them together in a dynamics calculation; this is a direct generalization of the popular // notation. Thus, HL///LL indicates a dual-level dynamics calculation based on a potential energy surface calculated at the LL, improved by a small number of HL calculations. One critical distinction between HL//LL and HL///LL, however, is that HL//LL does not involve any geometry optimizations at the higher-level, but HL///LL allows for the possibility that the saddle point and reagent geometries be re-optimized at the higher-level since it is usually dangerous not to do so.

The information from the higher-level is introduced into the lower-level surface by means of interpolated corrections (IC) (*36*, *38*). The method in its simplest form consists of defining an error function that is calculated by means of the HL and LL values for a property and interpolating that function all along the reaction path. For example, let $\omega_m^{HL}(s)$ be the value of the vibrational frequency of the mode m at a distance along the reaction path given by the reaction coordinate, s, as calculated at the higher level. (In the formulation used so far, the only points along the reaction path calculated at the higher level are stationary points, i.e., reactants, products, saddle point, and/or wells on the reaction

path.) Let $\omega_m^{LL}(s)$ be the value of the vibrational frequency of the same mode at the same value of s as calculated at the lower level. We define a function of s that measures how far the lower-level frequency for that mode is from the higher-level one; for example,

$$f(s) = \ln\left[\omega_m^{HL}(s)/\omega_m^{LL}(s)\right]. \tag{10}$$

The value of this function will be available only for those values of s for which the HL calculation has been performed; nevertheless, by using the appropriate functions (*36*) we can interpolate the values of $f(s)$ for any point on the reaction path. Thus, $\omega_m^{DL}(s)$, the dual-level estimated frequency for the mode m at s, will be given by

$$\omega_m^{DL}(s) = \omega_m^{LL} \exp\left[f^{ICL}(s)\right], \tag{11}$$

where

$$f^{ICL}(s) = \text{interpolant}\left\{\ln\left[\omega_m^{HL}(s)/\omega_m^{LL}(s)\right]\right\}, \tag{12}$$

and where ICL indicates "interpolated corrections based on the logarithm" (*38*). Following this idea, the method corrects the energy, frequencies, moments of inertia, and reduced moments of inertia for hindered rotors along the reaction path, although geometries and eigenvectors cannot be corrected by this method. For a detailed discussion on the interpolatory functions and the different ways of defining the corrections the user is referred to the original IC methodology papers (*36*, *38*).

As the higher level the best general options are to take values obtained from *ab initio* calculations at a level as high as possible or from SAC calculations, although some other options are in principle applicable, for example experimentally deduced properties of the saddle point.

As a lower level, the range of options is even broader, since the accuracy in the energies is not as important as in the higher-level calculation. Thus, an analytical potential energy surface for the reaction between atomic oxygen and methane has been used as lower level in a recent application of the dual-level methods (*82*), although this is not expected to be the most widely used choice, since the construction of an analytical potential energy surface can be an extremely time consuming (and perhaps tedious) task. For CH_4 + O the analytical potential energy surface employed was the J1 surface (*83*) for the CH_4 + H reaction slightly modified in order to reproduce the theoretical and experimental results for the CH_4 + O reaction. But the lower-level in most applications to date has been a direct dynamics calculation. In a direct dynamics calculation, whenever an energy, gradient, or Hessian is required in the dynamical calculation, it is computed "on the fly" by means of an electronic structure calculation.

One possibility for the electronic structure level for the lower level in a direct dynamics calculation is to use a density functional theory (DFT) calculation (*84*). This kind of calculation is much more affordable for large systems than *ab initio* calculations of comparable accuracy, and for some types of reactions it is more accurate than semiempirical calculations. DFT is an especially appropriate choice for many systems containing metal atoms, where semiempirical theory appears to be less reliable than for purely organic systems without metals. Thus, we modeled the C–H bond activation involving rhodium complexes by means of DFT. In particular, for the rearrangement of *trans*-$Rh(PH_3)_2Cl(\eta^2\text{-}CH_4)$ to $Rh(PH_3)_2\text{-}Cl(H)(CH_3)$ (*33*), we used the B3LYP DFT level (*85*), with the LANL2DZ basis set (*86*) as the lower level. A dual-level technique

was employed for the correction of only the energy along the reaction path, by using energies for reactants, products, and saddle point computed at the B3LYP level, with a larger basis set.

Some studies have been carried out using as a lower level an *ab initio* calculation using small basis sets and/or inaccurate treatment of the electron correlation. Thus, the first application of the IC method with an *ab initio* lower level was the study of the rate constant of the OH + NH_3 hydrogen abstraction reaction (*37*) using MP2/6-31G(*d,p*) as the lower level. Further comment on the results for this reaction will be provided below.

In calculations carried out so far, the most widely used choice for the lower-level direct dynamics calculation has been the semiempirical neglect-of-diatomic differential-overlap (NDDO) (*39,77,87,88*) method. The main advantage of this way of calculating the lower level is its economy; semiempirical methods are fast and affordable even for large systems, allowing us to calculate larger regions of the potential energy surface. Nevertheless, the lack of quantitative accuracy of these methods is a concern. Any of the usual general parametrizations that make use of NDDO approximations, e.g., MNDO (*88*), AM1 (*77*), and PM3 (*39*), are based on a set of parameters fitted in order to minimize in an average way the discrepancies between calculated and observed properties for a broad training set of test molecules, usually organic molecules. Thus, although these methods behave qualitatively correctly for most reactions, predictive quantitative accuracy is almost never achieved, especially for saddle points (as well as any other nonclassical structure, since the parametrization of the methods did not include such structures). Nevertheless the general parametrizations are sometimes useful with no change in parameters. For example the kinetic isotope effect for the [1,5] sigmatropic rearrangement of *cis*-1,3-pentadiene (*89*) was calculated by direct dynamics based on the MINDO/3 (*90*), AM1, and PM3 semiempirical surfaces. The results are in excellent agreement with the experimental results (within 13%), reflecting the suitability of these semiempirical methods in the study this reaction. Furthermore, since the IC method corrects the deficiencies of the lower level surface, with results usually showing weak dependence on the lower level of calculation, the standard parametrizations of the semiempirical methods are even more useful as the lower-level of a dual-level calculation. However, one can generally do better, as discussed next.

Since the parameters of semiempirical calculations are, in some sense, arbitrary parameters fitted in order to minimize an average error, a solution to the lack of accuracy in the description of a particular reaction is the reparametrization of the semiempirical Hamiltonian in order to improve the description of a particular reaction. Thus, the NDDO methods with specific reaction parameters (SRP) (*30*) can provide us with a economic but accurate description of the potential energy surface suitable for direct dynamics calculations.

The reparametrization of the semiempirical method might seem to be a task as time consuming as the construction and parametrization of an analytical expression for the potential energy surface; but this is not the case. Several aspects of the task make it more affordable:

- The problem of finding a physically meaningful functional form for the mathematical expression of the energy is eliminated. All reactions, independent of the number of atoms involved and the type of reaction can be treated using the same mathematical tools, which have a clear physical meaning.
- The original parametrization constitutes a good starting point for the reparametrization. In fact, we have usually tried to modify the original parameters by no more than a certain percentage (usually 10% or 20%), and one need not change a large number of parameters.
- The parameters to be fitted have a more clear physical meaning than the parameters usually involved in analytical expressions of the potential energy. The

selection of the parameters to be optimized and their fitting can, in principle, be guided by chemical intuition.

- A set of parameters optimized for a reaction can be expected to be at least partially transferable to other similar reactions. For this reason we sometimes call SRP parameters "specific range parameters" when we try to make them useful for some range of systems.
- When the NDDO-SRP surfaces are corrected by means of dual-level techniques (which is the usual procedure), the lower-level description of the reaction does not need to be energetically accurate. Obviously, the final results are expected to be more reliable when the lower level reproduces well the high-quality *ab initio* or experimental data, but when the NDDO-SRP surface is reasonable, the dual-level methods depend only weakly on the lower-level results.
- Automatic or semiautomatic optimization procedures can be used for the fitting of the SRP parameters. This is accomplished by defining an error function to be minimized, where the error function may contain the average deviation of the NDDO-SRP predictions from a few accurately known energies, stationary-point geometries, and/or frequencies. Taking into account the nature of the problem (multidimensional minimization of an error function dependent on a large number of variables, without analytic derivatives and the possibility that the error surface presents a rugged landscape), a particularly useful way of finding the optimum parameters is by using genetic algorithms (*91-93*). Thus, we can define an error function dependent on the set of parameters that we want to fit (*91*), and we can use genetic algorithms to look for the global minimum of that function. The main decisions to be made are the construction of the error function (weighting the errors in those properties that we want to describe more accurately), the choice of starting parameters (usually AM1 or PM3), the choice of which parameters to vary, and the limits (e.g., ±5%, ±10%) over which are allow those parameters to vary from their standard values.

The first applications of NDDO-SRP methods were in single-level calculations. The first work using this approach was the study of the rate constants and primary and secondary kinetic isotope effects for the microsolvated S_N2 reaction $Cl^-(H_2O)_n + CH_3Cl$, with $n = 0, 1, 2$ (*30*). The parameters were fitted to reproduce the experimental value of the electron affinity on Cl (a very important parameter for describing solvation effects) and the classical barrier height inferred (*94*) from a modeling of the $n = 0$ case. The results agree very well with those obtained from an analytic surface (*94*, *95*), which had been much harder to construct. Thus, the same set of parameters was used for a detailed description of the Cl^- solvation (*96*). In later work single-level NDDO calculations with SRP parameters fitted in order to describe experimental properties have been applied to modeling several other kinetic isotope effects (*97-100*).

Nevertheless, a more promising way of using NDDO-SRP methods for VTST/MT calculations is with IC methods. In particular we favor a bootstrap approach in which the higher-level information is used not only in the IC corrections but also for calibrating the SRP parameters (*31*, *36-38*, *61*, *82*, *95b*, *101–103*). An important advantage that we have pointed out previously is the relative independence of the final results on the lower level calculation. As an example, in Table 9 we examine the HL///LL rate constants for the $N_2H_2 + H \rightarrow N_2H + H_2$ reaction calculated using 4 different low-level surfaces, and we compare the results to a full HL calculation (*38*). At low temperatures the effects of a different lower level are more noticeable, but they all are within a reasonable range of the full HL results.

An important advantage of the IVTST-IC approach, especially for reactions with high curvature of the reaction path, is the possibility of including tunneling contributions from regions of the potential energy surface that are far from the reaction-path (the broad region of configuration space traversed by significant tunneling paths is called the reaction

Table 9. Rate constants (10^{-11} cm^3 $molecule^{-1}$ s^{-1}) for the H + N_2H_2 reaction calculated using single-level and dual-level methods[a]

Method	T (K)			
	300	600	1500	3000
HL///MNDO	0.065	0.39	2.7	18
HL///AM1	0.088	0.35	3.5	17
HL///MNDO-SRP	0.120	0.41	3.5	17
HL///AM1-SRP	0.050	0.32	3.8	18
HL	0.068	0.35	3.4	20

[a]In this table, HL denotes MRCI55/cc-pVTZ//CASSCF/cc-pVDZ. the results in this table are all taken from Ref. (*38*).

swath). The portion of the reaction swath that is far removed from the minimum energy path on its concave side is very important for reactions with large reaction-path curvature in the region of the barrier, since most of the tunneling may take place in this region of the surface. The reliable calculation of tunneling effects in reactions with large curvature therefore requires information not only about the reaction path and the harmonic valley surrounding it but about energies in the wider reaction swath. The high cost of the kind of *ab initio* calculations we want to use makes this information unavailable in most cases, but when using semiempirical NDDO or NDDO-SRP methods the use of large-curvature approximations for tunneling is possible (*98*), and the IC methods can also correct the energy in this region of the surface (*36*) on the basis of limited stationary-point data at a higher level. As an example, we note that the reaction-path for the reaction OH + NH_3 reaction (*37*) has an important curvature. As a consequence, tunneling methods that don't include the tunneling probability through the farther out regions of the reaction swath give rate constants that are too low at low temperatures, where tunneling is more important. The use of a semiempirical lower-level surface allows an inexpensive calculation of tunneling along the reaction path and through the reaction swath, leading to low-temperature rate constants in agreement with experimental results (*37*).

5. Summary

Dual-level methods in which the lower level is quantum mechanical can be very useful for bond energies, barrier heights, reaction-path dynamics, and electronic substituent effects on a subsystem of a larger system. The methods may be *ab initio* or partly semiempirical. Basis sets (MIDI!) and semiempirical parameters (SAC, NDDO-SRP) may be optimized specifically for use in dual-level calculations. The following methods were illustrated:

applicable primarily to small systems

- Scaling external correlation (SEC)

applicable to medium-sized systems

- Scaling all correlation (SAC)

applicable to all-sized systems

- NDDO-SRP as lower level
- Variational transition state theory with interpolated corrections
- IMOMO and IMOHC for electronic substituent effects on energies and geometries with DFT, HF, or semiempirical molecular orbital theory as lower level.

6. Acknowledgments

We acknowledgment many contributors to the methods discussed here (see the references) and would especially like to single out the following contributions: SEC: Frank Brown and David Schwenke; SAC: Mark Gordon and Ivan Rossi; IMOMO: E. Laura Coitiño and Molli Noland; dual-level direct dynamics: John Chuang, E. Laura Coitiño, Joaquin Espinosa-García, Patton Fast, Wei-Ping Hu, Yi-Ping Liu, Kiet Nguyen, and Orlando Roberto-Neto. Our work on electronic structure theory is supported in part by the National Science Foundation, and our work on gas-phase variational transition state theory is supported in part by the U.S. Department of Energy, Office of Basic Energy Sciences. JCC aknowledges a Fulbright Scholarship.

Literature Cited

(1) (a) Kao, J. *J. Amer. Chem. Soc.* **1987**, *109*, 3817. (b) Luzhkov, V.; Warshel, A. *J. Comput. Chem.* **1992**, *13*, 199.

(2) (a) Weiner, S. J.; Singh, U. C.; Kollman, P. A. *J. Amer. Chem. Soc.* **1985**, *107*, 2219. (b) Singh, U. C.; Kollman, P. A. *J. Comp. Chem.* **1986**, *7*, 718.

(3) (a) Bash, P. A.; Field, M. J.; Karplus, M. *J. Amer. Chem. Soc.* **1987**, *109*, 8092. (b) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700. (c) Field, M. J. In *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*, Vol. 2; van Gunsteren, W. F., Weiner, P. K., Wilkinson, A. J., Eds.; ESCOM: Leiden, 1993; p. 102.

(4) (a) Gao, J. *J. Phys. Chem.* **1992**, *96*, 537. (b) Gao, J. *Rev. Comp. Chem.* **1996**, *7*, 119.

(5) Cummins, P. L.; Gready, J. E. *J. Comput. Chem.* **1997**, *18*, 1496.

(6) (a) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170. (b) Matsubara, T.; Maseras, F.; Koga, N.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 2573. (c) Deng, L.; Woo, T. K.; Cavallo, L.; Margl, P. M.; Ziegler, T. *J. Amer. Chem. Soc.* **1997**, *119*, 6177.

(7) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.

(8) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221.

(9) Nyden, M. R.; Petersson, G. A. *J. Chem. Phys.* **1981**, *75*, 1843.

(10) (a) Petersson, G. A.; Tensfeldt, T. G.; Montgomery, J. A. *J. Chem. Phys.* **1991**, *94*, 6091. (b) Ochterski, J. W.; Petersson, G. A.; Montgomery, J. A. Jr. *J. Chem. Phys.* **1996**, *104*, 2598.

(11) Brown, F. B.; Truhlar, D. G. *Chem. Phys. Lett.* **1985**, *117*, 307.

(12) Gordon, M. S.; Truhlar, D. G. *J. Amer. Chem. Soc.* **1986**, *108*, 5412.

(13) Gordon, M. S.; Truhlar, D. G. *Int. J. Quantum Chem.* **1987**, *31*, 81

(14) Gordon, M. S.; Nguyen, K. A.; Truhlar, D. G. *J. Phys. Chem.* **1989**, *93*, 7356.

(15) Rossi, I.; Truhlar, D. G. *Chem. Phys. Lett.* **1995**, *234*, 64.

(16) (a) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959. (b) Svensson, M.; Humbel, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 3654.

(17) Noland, M.; Coitiño, E. L.; Truhlar, D. G. *J. Phys. Chem. A* **1997**, *101*, 1193.

(18) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19375.

(19) Coitiño, E. L.; Truhlar, D. G.; Morokuma, K. *Chem. Phys. Lett.* **1996**, *259*, 159.
(20) Coitiño, E. L.; Truhlar, D. G. *J. Phys. Chem. A* **1997**, *101*, 4641.
(21) Corchado, J. C.; Truhlar, D. G. submitted for publication
(22) Tapia, O. In *Quantum Theory of Chemical Reactions*; Daudel, R., Pullman, A., Salem, L., Veillard, A., Eds.; Reidel: Dordrecht, 1980; Vol. 2, p. 25.
(23) Lee, B.; Richards, F. M. *J. Mol. Biol.* **1971**, *55*, 379.
(24) Hermann, R. B. *J. Phys. Chem.* **1972**, *76*, 2754.
(25) Eisenberg, D.; McLachlan, A. D. *Nature* **1986**, *319*, 199.
(26) Cramer, C. J.; Truhlar, D. G. *J. Amer. Chem. Soc.* **1991**, *113*, 8305. Erratum: **1991**, *113*, 9901.
(27) Storer, J. W.; Giesen, D. J.; Hawkins, G. D.; Lynch, G. C.; Cramer, C. J.; Truhlar, D. G.; Liotard, D. A. *ACS Symp. Ser.* **1994**, *568*. 24.
(28) Hawkins, G. D.; Zhu, T.; Li, J.; Chambers, C. C.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G., chapter elsewhere in this volume.
(29) Parrinello, M. In *MOTEC: Modern Techniques in Computational Chemistry*; Clementi, E., Ed.; ESCOM: Leiden, 1991; p. 833.
(30) Gonzàlez-Lafont, A.; Truong, T. N.; Truhlar, D. G. *J. Phys. Chem.* **1991**, *95*, 4618.
(31) Truhlar, D. G. In *The Reaction Path in Chemistry;* Heidrich, D., Ed.; Kluwer: Dordrecht, 1995; p. 229.
(32) (a) Wang, I.; Karplus, M. *J. Amer. Chem. Soc.* **1973**, *95*, 8160. (b) Warshel, A.; Karplus, M. *Chem. Phys. Lett.* **1975**, *32*, 11. (c) Malcome-Lawes, D. J. *J. Chem. Soc. Faraday Trans. 2* **1975**, *71*, 1183. (d) Leforestier, C. *J. Chem. Phys.* **1978**, *68*, 4406. (e) Truhlar, D. G.; Duff, J. W.; Blais, N. C.; Tully, J. C.; Garrett, B. C. *J. Chem. Phys.* **1982**, *77*, 764. (f) Margl, P.; Ziegler, T.; Blöchl, P. E. *J. Am. Chem. Soc.* **1995**, *117*, 12625.
(33) Espinosa-García, J.; Corchado, J. C.; Truhlar, D. G. *J. Amer. Chem. Soc.* **1997**, *119*, 9891.
(34) Truhlar, D. G.; Garrett, B. C. *J. Chim. Phys.* **1987**, *84*, 365.
(35) Truong, T. N.; Lu, D.-h.; Lynch, G. C.; Liu, Y.-P.; Melissas, V. S.; Stewart, J. J. P.; Steckler, R.; Garrett, B. C.; Isaacson, A. D.; Gonzàlez-Lafont, A.; Rai, S. N.; Hancock, G. C.; Joseph, T.; Truhlar, D. G. *Comput. Phys. Commun.* **1993**, *75*, 143.
(36) Hu, W.-P.; Liu, Y.-P.; Truhlar, D. G. *J. Chem. Soc. Faraday Trans.* **1994**, *90*, 1715.
(37) Corchado, J. C.; Espinosa-García, J.; Hu, W.-P.; Rossi, I.; Truhlar, D. G. *J. Phys. Chem.* **1995**, *99*, 687.
(38) Chuang, Y.-Y.; Truhlar, D. G. *J. Phys. Chem. A* **1997**, *101*, 3808.
(39) Stewart, J. J. P. *J. Comp. Chem.* **1989**, *10*, 221.
(40) Löwdin, P.-O. *Adv. Chem. Phys.* **1959**, *2*, 207.
(41) (a) Werner, H.-J. *Adv. Chem. Phys.* **1987**, *69*, 1. (b) Shepard, R. *Adv. Chem. Phys.* **1987**, *69*, 63. (b) Roos, B. O. *Adv. Chem. Phys.* **1987**, *69*, 399.
(42) (a) Shavitt, I. In *Advanced Theories and Computational Approaches to the Electronic Structure of Molecules*; Dykstra, C. E., Ed.; Reidel: Dordrecht, 1984; p. 185. (b) Werner, H.-J. In *Domain-Based Parallelism and Problem Decomposition Methods in Computational Science and Engineering*; Keyes, D. E., Saad, Y., Truhlar, D. G., Eds.; SIAM: Philadelphia, 1995; p. 239.
(43) (a) Andersson, K.; Malmqvist, P.-A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483. (b) Andersson, K.; Malmqvist, P.-A.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.
(44) Steckler, R.; Schwenke, D. W.; Brown, F. B.; Truhlar, D. G. *Chem. Phys. Lett.* **1985**, *121*, 475.

(45) Schwenke, D. W.; Steckler, R.; Brown, F. B.; Truhlar, D. G. *J. Chem. Phys.* **1986**, *84*, 5706.
(46) Schwenke, D. W.; Tucker, S. C.; Steckler, R.; Brown, F. B.; Lynch, G. C.; Truhlar, D. G.; Garrett, B. C. *J. Chem. Phys.* **1989**, *90*, 3110.
(47) Lynch, G. C.; Truhlar, D. G.; Brown, F. B.; Zhao, J.-g. *Hyperfine Interactions* **1994**, *87*, 885.
(48) Lynch, G. C.; Truhlar, D. G.; Brown, F. B.; Zhao, J.-g. *J. Phys. Chem.* **1995**, *99*, 207.
(49) Lynch, G. C.; Steckler, R.; Schwenke, D. W.; Varandas, A. J. C.; Truhlar, D. G.; Garrett, B. C. *J. Chem. Phys.* **1991**, *94*, 7136.
(50) Mielke, S. L.; Lynch, G. C.; Truhlar, D. G.; Schwenke, D. W. *Chem. Phys. Lett.* **1993**, *213*, 10. Erratum: **1994**, *217*, 173.
(51) Pople, J. A.; Binkley, J. S.; Seeger, R. *Int. J. Quantum Chem. Symp.* **1976**, *10*, 1.
(52) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *Int. J. Quantum. Chem.* **1978**, *14*, 545.
(53) Bartlett, R. J. *Annu. Rev. Phys. Chem.* **1981**, *32*, 359.
(54) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968.
(55) Dunning, T. H. Jr. *J. Chem. Phys.* **1989**, *90*, 1007.
(56) (a) Siegbahn, P. E. M.; Blomberg, M. R. A.; Svensson, M. *Chem. Phys. Letters* **1994**, *223*, 35. (b) Siegbahn, P. E. M.; Svensson, M.; Boussard, P. J. E. *J. Chem. Phys.* **1995**, *102*, 5377.
(57) Siegbahn, P. E. M.; Crabtree, R. H. *J. Amer. Chem. Soc.* **1996**, *118*, 4442.
(58) Truong, T. N.; Truhlar, D. G.; Baldridge, K. K.; Gordon, M. S.; Steckler, R. *J. Chem. Phys.* **1989**, *90*, 7137.
(59) Melissas, V. S.; Truhlar, D. G. *J. Chem. Phys.* **1993**, *99*, 1013.
(60) Melissas, V. S.; Truhlar, D. G. *J. Phys. Chem.* **1994**, *98*, 875.
(61) Hu, W.-P.; Rossi, I.; Corchado, J. C.; Truhlar, D. G. *J. Phys. Chem. A*, **1997**, *101*, 6911.
(62) Truong, T. N.; Truhlar, D. G. *J. Chem. Phys.* **1990**, *93*, 1761. Erratum: **1992**, 97, 8820.
(63) Nguyen, K. A.; Gordon, M. S.; Truhlar, D. G. *J. Amer. Chem. Soc.* **1991**, *113*, 1596.
(64) Espinosa-García, J.; Corchado, J. C.; Sana, M. *J. Chim. Phys.* **1993**, *90*, 1181.
(65) Espinosa-García, J.; Tolosa, S.; Corchado, J. C. *J. Phys. Chem.* **1994**, *98*, 2337.
(66) Espinosa-García, J.; Corchado, J. C. *J. Chem. Phys.*, **1994**, *101*, 1333.
(67) Corchado, J. C.; Olivares del Valle, F. J.; Espinosa-García, J. *J. Phys. Chem.* **1993**, *97*, 9129. Erratum: **1994**, *98*, 5796.
(68) Espinosa-García, J.; Corchado, J. C. *J. Chem. Phys.*, **1994**, *101*, 8700.
(69) Allison, T.; Lynch, G. C.; Truhlar, D. G.; Gordon, M. S. *J. Phys. Chem.* **1996**, *100*, 13575.
(70) Melissas, V. S.; Truhlar, D. G. *J. Chem. Phys.* **1993**, *99*, 3542.
(71) (a) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650. (b) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639.
(72) Espinosa-García, J.; Corchado, J. C. *J. Chem. Phys.*, **1996**, *105*, 3517.
(73) Gonzàlez-Lafont, A.; Truong, T. N.; Truhlar, D. G. *J. Chem. Phys.* **1991**, *95*, 8875.
(74) (a) Pople, J. A.; Frisch, M. J.; Luke, B. T.; Binkley, J. S. *Int. J. Quantum Chem. Symp.* **1983**, *17*, 307. (b) Handy, N. C.; Knowles, P. J.; Somasundram, K. *Theor. Chim. Acta* **1985**, *68*, 87.
(75) (a) Feenberg, E. *Ann. Phys. (N.Y.)* **1958**, *3*, 292. (b) Wilson, S. *Int. J. Quantum Chem.* **1980**, *18*, 905.

(76) Chase, M. W. Jr.; Davies, C. A.; Downey, J. R. Jr.; Frurip, D. J.; McDonald, R. A.; Syverud, A. N. "JANAF Thermochemical Tables." *J. Phys. Chem. Ref. Data, Suppl.* **1985**, 14.
(77) (a) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F. *J. Amer. Chem. Soc.*, **1985**, *107*, 3902. (b) Dewar, M. J. S.; Zoebisch, E. G. *J. Mol. Struct. (Theochem)*, **1988**, *180*, 1.
(78) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (b) Perdew, J. P. In *Electronic Structure of Solids '91*; Ziesche, P., Eschrig, H., Eds.; Akademie Verlag: Berlin, 1991; p 11. (c) Perdew, J. P.; Chevary, J. A.; Vosko, S. H.; Jackson, K. A.; Pederson, M. R.; Singh, D. J.; Foilhaus, C. *Phys. Rev. B* **1992**, *46*, 6671. (d) Perdew, J. P.; Burke, K.; Ernzerhof, M. *ACS Symp. Ser.* **1996**, *629*, 453. (e) Perdew, J. P.; Burke, K.; Wang, Y. *Phys. Rev. B* **1996**, *54*, 16533.
(79) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.
(80) (a) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910. (b) Bartlett, R. J.; Watts, J. D.; Kucharski, S. A.; Noga, J. *Chem. Phys. Letters* **1990**, *165*, 513.
(81) Easton, R. E.; Giesen, D. J.; Welch, A.; Cramer, C. J.; Truhlar, D. G. *Theor. Chim. Acta* **1996**, *93*, 28.
(82) Corchado, J. C.; Espinosa-García, J.; Roberto-Neto, O.; Chuang, Y. Y.; Truhlar, D. G. Unpublished results.
(83) Joseph, T.; Steckler, R.; Truhlar, D. G. *J. Chem. Phys.* **1987**, *87*, 7036.
(84) Laird, B. B., Ross, R. B., Ziegler, T., Eds. *Density Functional Theory*; American Chemical Society: Washington, 1996.
(85) Stephens, P. J.; Devlin, F. J.; Ashuar, C. S.; Bak, K. L.; Taylor, P. R.; Frisch, M. J. In *Density Functional Theory*; Laird, B. B., Ross, R. B., Ziegler, T., Eds. Chemical Society: Washington, 1996; p. 105.
(86) (a) Dunning, T. H.; Hay, P. J. In *Modern Theoretical Chemistry*; Schaefer, H. F., Ed.; Plenum: New York, 1977; Vol. 2, pp 1-28. (b) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 270. (c) Wadt, W. R.; Hay, P. J. *J. Chem. Phys.* **1985**, *82*, 284.
(87) (a) Pople, J. A.; Santry, D.; Segal, G. *J. Chem. Phys* **1965**, *43*, S129. (b) Pople, J. A.; Beveridge, D. J. *Approximate Molecular Orbital Theory*; McGraw-Hill: New York, 1970.
(88) Dewar, M. J. S.; Thiel, W. *J. Amer. Chem. Soc.* **1977**, *99*, 4899, 4907.
(89) Liu, Y.-P.; Lynch, G. C.; Truong, T. N.; Lu, D.-h.; Truhlar, D. G; Garrett, B. C. *J. Amer. Chem. Soc.* **1993**, *115*, 2408.
(90) Bingham, R. C.; Dewar, M. J. S.; Lo, D. H. *J. Amer. Chem. Soc.* **1975**, *97*, 1285.
(91) Rossi, I.; Truhlar, D. G. *Chem. Phys. Letters* **1995**, *233*, 231.
(92) Goldberg, D. E. *Genetic Algorithms in Search, Optimization and Machine Learning*; Addison-Wesley: Reading, 1989.
(93) Bash, P. A.; Ho, L.L.; Mackerell, A.D.; Levine, D.; Hallstrom, P. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 3698.
(94) Tucker, S. C.; Truhlar, D. G. *J. Amer. Chem. Soc.* **1990**, *112*, 3338.
(95) (a) Tucker, S. C.; Truhlar, D. G. *J. Amer. Chem. Soc.* **1990**, *112*, 3347. (b) Truhlar, D. G.; Lu, D.-h.; Tucker, S. C.; Zhao, X. G.; Gonzàlez-Lafont, A.; Truong, T. N.; Maurice, D.; Liu, Y.-P.; Lynch, G. C. In *Isotope Effects in Chemical Reactions and Photodissociation Processes*; Kaye, J. A., Ed.; American Chemical Society Symposium Series 502: Washington, DC, 1992; pp. 16-36.
(96) Zhao, X. G.; Gonzàlez-Lafont, A.; Truhlar, D. G. *J. Chem. Phys.* **1991**, *94*, 5544.
(97) Viggiano, A. A.; Paschkewitz, J. S.; Morris, R. A.; Paulson, J. F.; Gonzàlez-Lafont, A.; Truhlar, D. G. *J. Amer. Chem. Soc.* **1991**, *113*, 9404.

(98) Liu, Y.-P.; Lu, D.-h.; Gonzàlez-Lafont, A.; Truhlar, D. G.; Garrett, B. C. *J. Amer. Chem. Soc.* **1993**, *115*, 7806.
(99) Corchado, J. C.; Espinosa-García, J. *J. Chem. Phys.* **1996**, *105*, 3160.
(100) Gonzàlez-Lafont, A.; Truhlar, D. G. In *Chemical Reactions in Clusters*; Bernstein, E. R., Ed.; Oxford University Press: New York, 1996; pp. 3-39.
(101) Robert-Neto, O.; Coitiño, E. L.; Truhlar, D. G. Unpublished results.
(102) Hu, W.-P.; Truhlar, D. G. *J. Amer. Chem. Soc.* **1995**, *117*, 10726.
(103) Hu, W.-P.; Truhlar, D. G. *J. Amer. Chem. Soc.* **1996**, *118*, 860.