High Accuracy Molecular Heats of Formation and Reaction Barriers: Essential Role of Electron Correlation

Jeffrey C. Grossman[1,2] and Lubos Mitas[2]
[1]*Department of Physics, University of California, Berkeley, California 94720-7300, and Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720*
[2]*National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801*
(Received 20 August 1997)

We demonstrate that the quantum Monte Carlo (QMC) methodology (i) provides barrier heights and heats of formation within ~0.05 eV of experimental values, (ii) confirms recent measurements for the ring inversion of cyclo-octatetraene, and (iii) enables us to predict quantities not yet measured. Density functional methods show a mixed performance in achieving the accuracy required for predictive calculations. Further comparisons show that QMC is competitive in accuracy with the best correlated wave function methods while being applicable to much larger systems because of more favorable scaling. [S0031-9007(97)04708-X]

PACS numbers: 31.15.Ar, 31.25.–v

The most important properties of reactions are geometries, transition state barrier heights, and heats of formation. The fact that these properties might be difficult to estimate by experiments clearly demonstrates the importance of predictive computational approaches. Calculations of reactions by *ab initio* methods remain one of the most challenging problems in computational physics and chemistry, since the accuracy of commonly used approximate methods is often inadequate, while the system sizes which can be studied by higher accuracy approaches is limited to very few atoms. State of the art experiments can now measure transition states to within 0.1 eV for systems as large as 16 atoms [1], which is extremely difficult to achieve by the commonly used correlated wave function methods.

In this Letter we show that barrier heights and heats of formation can be obtained to within 0.05 eV of experiment with the recently developed quantum Monte Carlo (QMC) approaches. In order to gain insight into the performance of different theoretical methods and further understanding of the strengths and weaknesses of these methods, we also employed seven different state of the art computational approaches in condensed matter physics and quantum chemistry. Several examples of molecular reactions which proved to be difficult for commonly used methods [2] are studied. In addition, we present the first accurate calculations of a recently measured reaction, the ring inversion of cyclo-octatetraene, the largest system for which the transition state has been experimentally observed and characterized [1]. Our calculations show that QMC combines the necessary accuracy with the unique ability to perform correlated calculations for large systems. Moreover, the consistency of the QMC results enabled us to make genuine predictions for quantities not yet measured.

In our QMC approach [3–8], we use variational Monte Carlo to find an optimized correlated many-body trial function, $\psi_T(R)$. $\psi_T(R)$ is a product of Slater determinants, D_n, and a correlation factor [9],

$$\psi_T = \sum_n d_n D_n^{\uparrow} D_n^{\downarrow} \exp\left[\sum_{I,i<j} u(r_{iI}, r_{jI}, r_{ij})\right], \quad (1)$$

where I corresponds to the ions, i, j to the electrons, and r_{iI}, r_{jI}, r_{ij} to the distances. Parametrization and optimization of $u(r_{iI}, r_{jI}, r_{ij})$, which represents the electron-electron and electron-electron-ion correlations, is described in Ref. [10]. In the Slater determinant part, we employ natural rather than Hartree-Fock orbitals [6].

To eliminate most of the remaining variational bias we use diffusion Monte Carlo, which is based on the property that the operator $\exp(-\tau H)$, where H is the Hamiltonian, projects out the ground state of any trial function with the same symmetry and nonzero overlap. All QMC results presented here are from diffusion Monte Carlo.

In addition to QMC, we have also calculated barriers and heats of formation within the local density approximation (LDA) [11], the generalized gradient approximation (GGA) with four different functionals [12–15], and the coupled cluster method with singles, doubles, and perturbationally triples [CCSD(T)] [16] using Hartree-Fock (HF) as a reference. Extensive basis sets were employed within each method for all species studied [17].

All geometries were optimized within each method, including the computationally demanding CCSD(T) approach. Because efficient structural optimization has yet to be developed within QMC, it is important to choose carefully the best possible geometries. As is often the case, the LDA and GGA geometries compare very well with experiment for the equilibrium structures. Overall, we find the GGA functional B3-PW91 to give the best match (differing from available experiments by a maximum of 0.02 a.u. and 0.7°, and we therefore use B3-PW91 optimized geometries for our QMC calculations.

In the following, we describe three different organic reactions and compare the results of various theoretical approaches: (i) The simple radical hydroxyl reaction, $OH + H_2 \rightarrow H_2O + H$, involves the exchange of one

hydrogen atom and has proved particularly problematic for the density functional approaches. (ii) The tetrazine reaction, $H_2C_2N_4 \rightarrow 2HCN + N_2$, is interesting because of the possibility of a single-step triple dissociation [18] and has been investigated theoretically by several groups [2,18–20]. (iii) In the vinyl alcohol reaction, $C_2OH_3 \rightarrow$ acetaldehyde, a recent theoretical study has shown that the available experimental data may be suspect [21].

Table I lists our calculated barrier heights and heats of formation for eight different methods plus experiment for each of these reactions. Our QMC results are in agreement to within 0.05 eV of the experimental values for the cases where experimental data are available. For the tetrazine reaction, both the barrier height and the heat of formation are not well established experimentally. The barrier height must fall at or below the threshold for photodissociation (2.25 eV), and the heat of formation has not yet been accurately measured (e.g., two different experiments have given values for the heat of formation -0.04 and -2.21 eV [22]). For vinyl alcohol, our results support a value around 2.47 eV which is different from experimental estimates. As far as is known to us, the experimental values are not available for these quantities and therefore we predict the theoretical values ahead of experiment. Our QMC results are confirmed by the CCSD(T) method which, however, requires enormous computational resources to accommodate the large basis sets needed to achieve such an accuracy.

The Hartree-Fock barriers and heats of formation are typically off by 0.5–1.0 eV. In LDA, one finds a wide range of errors, the worst being the negative barrier for the hydroxyl reaction, thus predicting the transition state to have a lower energy than both the reactants and products. A detailed investigation of the failure of density functional methods to describe the potential energy surface of this reaction can be found in Ref. [23].

Given the small sizes of the molecules in the hydroxyl reaction, it may be somewhat surprising that the various methods underperform so dramatically. The breaking of the hydrogen bond, however, is difficult for many standard methods which rely on a cancellation of systematic errors (i.e., basis set incompleteness, density dependent errors, etc.). In contrast, QMC is less sensitive to these problems and rather relies on a cancellation of the fixed-node error which is usually much smaller, typically 5% of the valence correlation energy. Fixed-node QMC calculations of the barrier height for the hydrogen exchange reaction $H + H_2 \rightarrow H_2 + H$ [24] were in agreement to within $\sim$0.005 eV with the exact calculations [25].

In many cases correct energy differences are found to be "bracketed" by LDA and HF estimates. This observation can be qualitatively understood from the fact that the exchange correlation can be expressed as an integral over the strength of the electron-electron interaction, where one limit in the integrand represents Hartree-Fock while the other LDA. Our calculations show that the effort to "correct" LDA by GGA's produces mixed results and the predictions of various functionals are inconsistent. For example, in the tetrazine reaction, the BLYP result which is the closest to the correct value for the heat of formation fails by about 0.56 eV for the barrier height, and the predictions of the GGA's are on average not much better than LDA. In fact, *none* of the gradient corrected functionals is able to repair completely the LDA error and the GGA's are off by at least 0.4 eV or more—certainly an unsatisfactory result.

As a further test of the strengths of the different theoretical approaches, we also compare atomization energies for the molecules involved in the tetrazine reaction (Table II). Note that LDA exhibits the usual 20%–30% overbinding. Although the GGA results are much improved over LDA, there is still a varying degree of systematic errors depending on the functional used. Quantum Monte Carlo clearly gives values which are very close to experiment, with differences from experiment of less than 2%. On the other hand, CCSD(T) with the largest basis set we could computationally afford shows errors on the order of 6% to 7%. Thus, we are able to make the most accurate prediction to date for the binding energy of the tetrazine molecule, which, despite a number of theoretical studies done in recent years [19,27] has never been explicitly reported, to our best knowledge. The QMC estimate is about 34.36(5) eV which, assuming $\approx$1.5% underbinding, gives a best value so far of about 34.9(1) eV.

TABLE I. Barrier heights (BH) and heats of formation (ΔH_f) in eV for the hydroxyl, tetrazine, and vinyl alcohol reactions calculated with different methods (see text) and compared with available experiments. Theoretical values include their respective zero point energies. The QMC error bars are in parentheses.

	Hydroxyl		Tetrazine		Vinyl alcohol	
Method	BH	ΔH_f	BH	ΔH_f	BH	ΔH_f
HF	1.05	−0.07	3.01	−3.41	3.05	−0.54
LDA	−0.91	−0.69	2.07	−1.99	1.91	−0.34
B-LYP	−0.10	−0.49	1.20	−2.35	2.22	−0.46
B3-LYP	0.03	−0.44	1.74	−2.15	2.41	−0.45
B-PW91	−0.03	−0.64	1.49	−1.73	2.12	−0.45
B3-PW91	0.04	−0.56	1.98	−1.68	2.32	−0.44
CCSD(T)	0.25	−0.57	1.79	−2.53	2.45	−0.46
QMC	0.22(4)	−0.65(4)	1.73(7)	−2.65(7)	2.47(4)	−0.43(4)
EXP	0.17	−0.63	<2.25	···	···	−0.42

TABLE II. Calculated atomization energies (eV) for the tetrazine reaction systems and available experimental data.

	N_2	HCN	Tetrazine	Transition state
LDA	11.03	15.23	42.45	39.98
B-LYP	10.43	13.92	36.29	34.88
B3-LYP	9.97	13.61	35.38	33.42
B-PW91	10.29	13.87	36.64	34.93
B3-PW91	9.80	13.53	35.51	33.31
CCSD(T)[a]	9.37	13.05	33.57	31.37
QMC	9.75(5)	13.46(5)	34.36(5)	32.40(5)
EXP	9.91	13.52	...	...

[a]With the cc-pVTZ [26] basis set.

The largest reaction considered in this study is the ring inversion of cyclo-octatetraene (C_8H_8) which has a transition state (see Fig. 1) that was recently shown to be an open-shell singlet [1], thus violating Hund's rule. These experiments, which mark a significant advance in transition state spectroscopy, place the D_{8h} open-shell singlet $\approx$0.61 eV above the ground state and the D_{8h} triplet roughly 0.35 eV higher in energy than the singlet [1].

Previous theoretical work on the cyclo-octatetraene reaction [28] has shown the tub (D_{2d}) geometry to be the ground state structure, in agreement with experiment. Figure 1 shows the highest relevant single-particle orbitals for the ground and transition state geometries. The low symmetry nondegenerate highest occupied and lowest unoccupied molecular orbitals of the ground state are indicative of the nonequivalent bonding in the molecule. As the reaction proceeds, these orbitals are transformed into two degenerate E_g states which correspond to the equal bonds along the transition state D_{8h} ring. The calculations in the original paper [1] claimed to reproduce the experimental barrier heights using a multiconfiguration HF with a small basis set. However, such a calculation provided only a small amount of correlation energy. In general, for predictions in the 0.05–0.1 eV accuracy range much larger basis sets have to be employed.

Since Hartree-Fock is inappropriate for an open-shell singlet state, we use a two-configuration generalized valence bond (GVB) wave function. This gives an energy for the singlet which is too far (1.39 eV) above the ground state (see Table III). In LDA, the energies are too small for the barrier to the ring triplet and too large for the barrier to the ring singlet, thus reversing the triplet/singlet ordering as in the case of GVB. The GGA results show almost no change for the singlet energy difference and tend to worsen the triplet. The challenges associated with correlating a multireference wave function (i.e., an open-shell singlet) are especially apparent in the density functional approaches, which by construction rely on single-particle densities (or spin densities) and often have difficulties with inherently multireference states such as biradical singlets. Currently, the coupled cluster approach is difficult to apply here with the same quality basis set as in the smaller reactions.

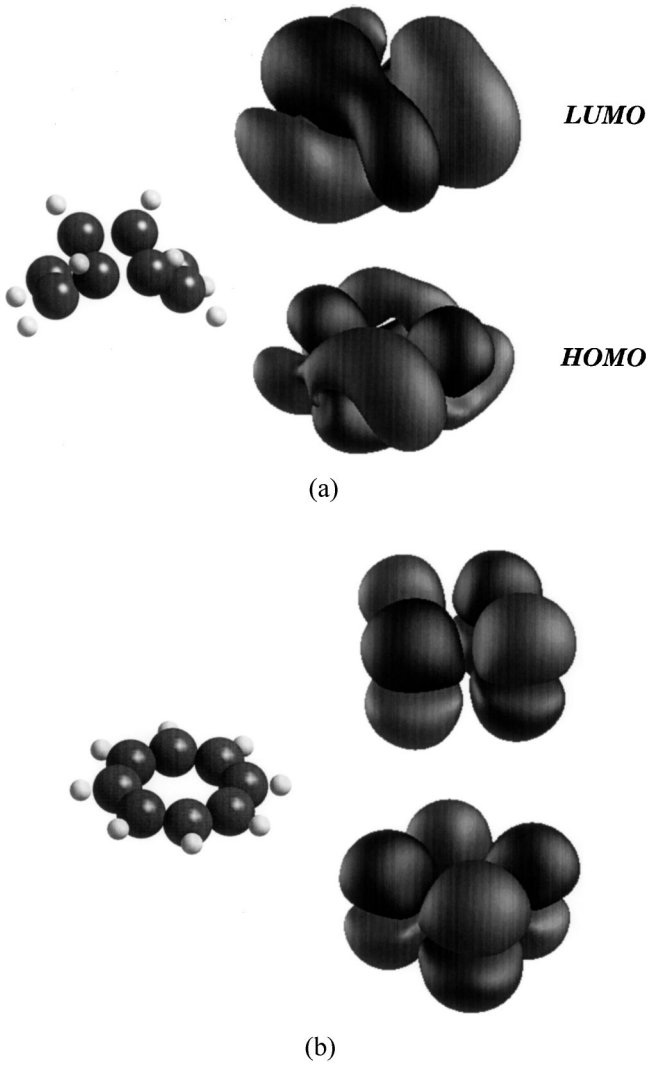

FIG. 1. (a) Geometry and isosurfaces of highest occupied and lowest unoccupied molecular orbitals for the ground D_{2d} state of cyclo-octatetraene; (b) geometry and isosurfaces of the two degenerate orbitals for the singlet and triplet states of the transition D_{8h} state of cyclo-octatetraene.

In our QMC calculations, correlating multideterminantal wave functions is straightforward and, as is clear from Table III, excellent agreement with experiment was obtained using GVB for building the trial function. Thus, for this cyclo-octatetraene reaction, QMC is the *only* method which is capable of correctly predicting both the barrier heights and the energetic ordering of the transition states while, at the same time, providing a high level of correlation by obtaining $\approx$95% of the valence correlation energy.

In summary, we have calculated several organic reactions which revealed a crucial role of electron correlation and its many-body treatment for reliable predictions of barrier heights and heats of formations. The results and comparisons with other methods and experiments show the high accuracy of quantum Monte Carlo which enabled us to predict several quantities not yet measured. The

TABLE III. Energy differences (eV) from ground state (D_{2d}) cyclo-octatetraene to the D_{8h} singlet and triplet ring transition states.

Method	D_{8h} singlet	D_{8h} triplet
GVB	1.40	1.19
LDA	0.95	0.67
B-LYP	0.85	0.52
B-PW91	0.88	0.50
QMC	0.61(7)	0.87(7)
EXP	0.61(7)	0.95(7)

accurate quantum chemistry CCSD(T) approach agrees very well with QMC calculations; however, because of the rapid growth of computer demands with the number of electrons (scaling as $\sim N^7$, where N is the number of correlated electrons), the coupled cluster approach becomes impractical for larger systems. For the largest reaction in this study, the ring inversion of cyclo-octatetraene, QMC convincingly confirms recent experiments on the transition state barrier and the triplet-singlet splitting for the ring transition state. The inherent fixed-node error in QMC, which is roughly 5% of the correlation energy, can be reduced further by ongoing method developments. We believe that the results presented show that advances in both computational power and algorithms have brought a new tool in computational physics to the forefront of simulation technologies. Quantum Monte Carlo offers an accurate and efficient alternative to the traditional correlated approaches and provides new opportunities for studies of electronic structure. The ability of QMC to maintain an accuracy of 0.05 eV for large *sp* systems combined with a favorable scaling (N^3) with number of particles opens possibilities for accurate calculations of reactions involving even larger molecules.

We thank E. Wimmer for suggestions, A. J. Freeman and W. A. Lester for useful discussions, V. H. Crespi for reading the manuscript, and H. Mitasova for help with the figures. This research was supported by NSF Grant No. DMR-9422496, by DARPA Grant No. DA-BT63-95, by U.S. Department of Energy under Contract No. DE-AC03-76SF00098, by NCSA, and by the State of Illinois.

[1] P. G. Wenthold, D. A. Hrovat, W. T. Borden, and W. C. Lineberger, Science **272**, 1456 (1996).
[2] J. Baker, J. Andzelm, and M. Muir, J. Chem. Phys. **102**, 2063 (1995).
[3] L. Mitas, E. L. Shirley, and D. M. Ceperley, J. Chem. Phys. **95**, 3467 (1991).
[4] L. Mitas, Comput. Phys. Commun. **96**, 107 (1996).
[5] L. Mitas and J. C. Grossman, in *Recent Advances in Quantum Monte Carlo Methods,* edited by W. A. Lester, Recent Advances in Computational Chemistry (World Scientific, Singapore, 1997).
[6] J. C. Grossman and L. Mitas, Phys. Rev. Lett. **74**, 1323 (1995).
[7] J. C. Grossman, L. Mitas, and K. Raghavachari, Phys. Rev. Lett. **75**, 3870 (1995).
[8] B. L. Hammond, W. A. Lester, Jr., and P. J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry* (World Scientific, Singapore, 1994).
[9] K. E. Schmidt and J. W. Moskowitz, J. Chem. Phys. **93**, 4172 (1990).
[10] L. Mitas, in *Computer Simulation Studies in Condensed-Matter Physics V,* edited by D. Landau, K. Mon, and H.-B. Schuttler, Springer Proceedings in Physics (Springer-Verlag, Berlin, 1993).
[11] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[12] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
[13] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
[14] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
[15] J. P. Perdew, in *Electronic Structure of Solids '91,* edited by P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991), p. 11.
[16] R. J. Bartlett and J. F. Stanton, in *Reviews in Computational Chemistry,* edited by K. B. Lipkowitz and D. B. Boyd, Reviews in Computational Chemistry Vol. 5 (VCH, New York, 1994).
[17] Surprisingly, up to four basis functions (i.e., quadruple-zeta) in the *s* and *p* symmetry channels are necessary for convergence with GGA functionals; cc-pVQZ [26] basis sets ($2f1g$ functions on C, N, and O, and $2d1f$ functions on H) were used for the CCSD(T) method.
[18] A. C. Scheiner, G. E. Scuseria, and H. F. Schaefer, J. Am. Chem. Soc. **108**, 8160 (1986).
[19] M. Schutz, J. Hutter, and H. P. Lüthi, J. Chem. Phys. **103**, 7048 (1995).
[20] G. E. Scuseria and H. F. Schaefer, J. Phys. Chem. **94**, 5552 (1990).
[21] B. J. Smith, M. T. Nguyen, W. J. Bouma, and L. Radom, J. Am. Chem. Soc. **113**, 6452 (1991).
[22] R. M. Hochstrasser and D. S. King, J. Am. Chem. Soc. **98**, 5443 (1976); D. Coulter, D. Dows, H. Reisler, and C. Wittig, Chem. Phys. **32**, 429 (1978).
[23] J. Baker, J. Andzelm, M. Muir, and P. R. Taylor, Chem. Phys. Lett. **237**, 53 (1995).
[24] D. M. Ceperley and B. J. Alder, J. Chem. Phys. **81**, 5833 (1984).
[25] D. L. Diedrich and J. B. Anderson, Science **258**, 786 (1992).
[26] T. H. Dunning, Jr. J. Chem. Phys. **90**, 1007 (1989).
[27] R. J. Doerksen and A. J. Thakkar, Int. J. Quantum. Chem. **60**, 421 (1996); J. V. Ortiz and V. G. Zakrzewski, J. Chem. Phys. **105**, 2762 (1996); J. Martin and C. Vanalsenoy, J. Phys. Chem. **100**, 6973 (1996).
[28] D. A. Hrovat and W. T. Borden, J. Am. Chem. Soc. **114**, 5879 (1992); C. Trindle and T. Wolfskill, J. Org. Chem. **56**, 5426 (1991); M. J. Dewar and K. M. Merz, Jr., J. Phys. Chem. **89**, 4739 (1985).