J. Phys. Chem. A **1997,** *101,* 227−233

IMOMO(G2MS): A New High-Level G2-Like Method for Large Molecules and Its Applications to Diels−Alder Reactions

Robert D. J. Froese, Stéphane Humbel, Mats Svensson, and Keiji Morokuma*

Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, Atlanta, Georgia 30322

Received: October 1, 1996[⊗]

A new and less expensive G2-type approach, G2MS, which can be used for accurate energy prediction for up to seven to eight atoms has been proposed and tested against the standard G2 data set. The results compare well with other G2 methods. The G2MS method performs an extrapolation of correlation and basis set effects, while the integrated MO+MO (IMOMO) method provides an extrapolation of electronic and steric effects from a small model to a large real system. Thus, using G2MS as the high-level method in IMOMO is a natural approach to accurate energy predictions for large molecular systems. The G2MS method predicts activation energies for ethylene + butadiene and ethylene + cyclopentadiene of 23.9 and 18.5 kcal/mol, respectively. The IMOMO(G2MS:MP2) method has been used to obtain accurate activation barriers for a number of Diels−Alder reactions, including the dimerization of butadiene where the calculated value of 23.5 kcal/mol is within 1 kcal/mol of two experimental values. For the addition of acrylic acid to 2,4-pentadienoic acid, a nearly quantitative agreement in the branching ratio for the product regio- and stereoisomers has been obtained. Calculations of the activation barriers for larger Diels−Alder reaction systems were performed, including the reaction of maleic anhydride with isoprene and 2-*tert*-butyl-1,3-butadiene, where the conformation of the reactant diene is found to be an important factor in determining the activation energy.

I. Introduction

High-level *ab initio* approaches, especially those including electron correlation, are well-known to be computationally very expensive. The high dependence of the cost on the number of atoms and basis functions in the molecule makes it impossible to study most chemical systems at a high and reliable level of theory. Furthermore, to achieve chemical accuracy in the energies, not only do we have to consider the method of electron correlation, but the basis set also is of importance. The G2 method,[1] as well as the CBS[2] and G2M[3] methods, has been developed to extrapolate the correlation and basis set effects to obtain chemical accuracy in the bond dissociation energies and are applicable to systems with up to six to seven heavy (non-hydrogen) atoms.

In studying larger systems, two approaches have been used in the past: either small models of the real system in question can be adopted or the real system is treated at a relatively low level of theory. Both of these approximations have obvious limitations as the former excludes electronic and steric effects from the part of the molecule not present in the small model and the latter requires a low-level description of the correlation and basis set effects resulting in a loss of accuracy. We have recently developed a family of integrated methods, which incorporate both of these approaches into one calculation. Several different schemes have been suggested[4,5] in which a molecule is divided into various parts described at different levels of theory where the total energy is expressed as a sum and difference of their energies. We have developed the integrated MO+MM (IMOMM),[5,6] the integrated MO+MO (IMOMO),[7,8] and the more general ONIOM (our own N-layered integrated mO + mM) methods[9] and tested these for different systems. In the IMOMO scheme, for instance, the total energy is given as

$$E(\text{IMOMO}) = E(\text{high,model}) + [E(\text{low,real}) - E(\text{low,model})]$$

where high and low refer to high and low levels of theory, respectively, and model and real refer to the small model system and the real system, respectively. The term in the square brackets can be viewed as the electronic effect of the real system on the high-level energetics of the model system. These approaches can also be considered as an extrapolation of the level of computation for the increased size of the molecule.

In the present study, we focus on IMOMO and develop an extrapolation scheme which not only corrects for correlation and basis set effects, like G2, but also extrapolates for the size of the molecule, in the spirit of IMOMO. For this purpose, we at first introduce a relatively inexpensive G2-like scheme, G2MS, aiming for systems consisting of up to eight to ten heavy atoms, and test it using the standard G2 set of 32 first row molecules.

This new G2MS method will be used as the model in the IMOMO(G2MS:MO) approach. This IMOMO(G2MS:MO) scheme will be tested on a variety of Diels−Alder reactions,[10] which are very important in synthetic organic chemistry and have been studied in detail both experimentally and theoretically.[11,12] The process is believed to be a concerted one, even though there are many studies suggesting a stepwise biradical mechanism.[13] A variety of semiempirical[13,14] and *ab initio* studies[11,12,15] have been made on Diels−Alder reactions, and, in general, it has been shown that high levels of electron correlation are required to obtain activation energies in quantitative agreement with experiment. Due to strong electronic and correlation effects, the Hartree−Fock method overestimates these activation energies and MP2 often underestimates them. The results are also sensitive to the basis set. Our interest in these types of reactions comes from the difficulty in obtaining good agreement with experiment in these Diels−Alder reactions and the extension to substituted systems where electronic effects are very important.

[⊗] Abstract published in *Advance ACS Abstracts,* December 15, 1996.

S1089-5639(96)03019-8 CCC: $14.00

Our studies on Diels–Alder reactions will look at different systems. First, we will test our new G2MS approach to the butadiene + ethylene and cyclopentadiene + ethylene reactions. Following this test of G2MS, we will study activation energies for the dimerization of butadiene to determine the effectiveness of the IMOMO(G2MS:MP2) approach. Then we will use the IMOMO(G2MS:MO) method to study (1) regio- and stereoselectivity in the addition of acrylic acid to 2,4-pentadienoic acid and (2) the activation barriers in the acrolein + isoprene, acrolein + 2-*tert*-butyl-1,3-butadiene, maleic anhydride + isoprene, and maleic anhydride + 2-*tert*-butyl-1,3-butadiene.

II. New G2MS Scheme and the IMOMO(G2MS:MO) Method

A. Definition and Test of the G2MS Scheme. In order to attain chemical accuracy (1–2 kcal/mol) in bond dissociation and atomization energies, a combination of a high-level method like CCSD(T) and a large basis set like 6-311+G(3df,2p) is required and can be adopted for only the very smallest chemical systems containing up to two to three heavy atoms. As discussed in the previous section, several schemes including G2,[1] CBS,[2] and G2M[3] have been developed to estimate the basis set dependence of the correlation energy at a lower level of theory and have been applied for accurate evaluation of the energetics for systems containing up to six to seven heavy atoms.

The main purpose of this subsection is to develop a less expensive scheme having chemical accuracy and applicable to systems containing as many as eight to ten heavy atoms, as measured by the present capacity of a standard workstation. This criterion is chosen as such since the active centers of various organic and inorganic reactions usually contain less than eight to ten heavy atoms. High-level single-reference *ab initio* methods, such as CCSD(T) have been shown to be reliable even in cases where a near-degeneracy is important.[16] A single reference approach has a large cost advantage over a multireference method, such as MR-CI, and is easier to handle. Therefore, it is natural to pick CCSD(T) as the "method of choice" in describing the high-level correlation effects. The low-level correlation method, MP2, is adopted for incorporating the effect of a large basis set. It is well-known that due to the slow convergence of the correlation energy with the size of the basis set, energetics for chemical reactions in which the number of electron pairs change will be erroneous. As with G2 and G2M methods, an empirically optimized correction term, usually referred to as HLC, will be included. Thus, the new scheme we suggest, called G2MS, is defined as

$$E(\mathrm{G2MS}) = E[\mathrm{CCSD(T)/6\text{-}31G(d)}] + E[\mathrm{MP2/6\text{-}311{+}G(2df,2p)}] - E[\mathrm{MP2/6\text{-}31G(d)}] + \mathrm{HLC_{G2MS}} \quad (1)$$

The CCSD(T)/6-31G(d) calculation forms the basis of the energy expression, and the two MP2 calculations correct for the basis set effects. Equation 1 looks very similar to the G2MP2 and G2(MP2,SVP) schemes.[1b,c] One main difference is in the basis set of the CCSD(T)[17] calculation, which is 6-311G(d,p) in G2MP2, making the G2MS approach substantially less expensive, and 6-31G in G2(MP2,SVP), making the G2MS substantially more reliable. Replacing MP2/6-311+G(3df,2p) in G2MP2 by a MP2/6-311+G(2df,2p) calculation results in an additional computational saving of 1.5–2 times, depending on the size of the system. Spin contamination for open shell systems at the MP2 level can cause difficulties and loss of reliability. In order to avoid these problems, we used the ROMP2 methods as implemented in Gaussian94[18] for the open shell systems and the RMP2 method for the closed shell systems for our test data. The UMP2 method was also used for the open shell systems, and little or no difference was seen as compared to the ROMP2 approach. Thus, we recomend the more standard UMP2 method. Coupled cluster calculations for open shell systems were performed at the unrestricted level. The HLC term is determined to minimize the mean absolute deviation from experiment in the atomization energies for the standard G2 set of 32 molecules and is −3.8 kcal/mol in G2MS for each doubly occupied orbital, vs −3.0 kcal/mol in each of G2 and G2MP2. The geometries and frequencies for the zero-point energy correction (ZPC) are calculated at the B3LYP/6-31G(d) level. B3LYP optimizations, as used in the G2M method, are less expensive and scale better with the system size than MP2 optimizations.[3] In addition, B3LYP geometries and frequencies have been shown to be very reliable[19] and more stable to spin contamination for several different types of systems.[3] Though the G2M method uses B3LYP/6-311G(d,p) optimized geometries and frequencies, this basis set is too large for the size of the systems we are aiming for.[20]

TABLE 1: Deviations (kcal/mol) in the Atomization Energies with Respect to Experiment for the G2MS Approach, As Compared with the G2MP2 and G2 Schemes

	exp	(G2MS)	(G2MP2)[a]	(G2)[b]
H_2	103.3	−1.1	1.0	0.6
LiH	56.0	0.2	0.2	0.6
BeH	46.9	−4.5	−2.2	−1.4
CH	79.9	0.8	0.3	0.6
CH_2 ($^3B_1$)	179.6	−1.3	−1.4	−1.0
CH_2 ($^1A_1$)	170.6	1.5	1.0	1.4
CH_3	289.2	−0.8	−0.5	−0.1
CH_4	392.5	−0.5	0.2	0.7
NH	79.0	−1.2	−1.4	−1.1
NH_2	170.0	−0.2	−0.1	0.1
NH_3	276.7	−0.7	−0.1	−0.2
OH	101.3	0.0	0.5	0.3
OH_2	219.3	0.2	1.2	0.3
FH	135.2	0.9	1.7	1.1
Li_2	24.0	3.3	2.3	1.9
LiF	137.6	−1.0	0.3	−0.1
C_2H_2	388.9	0.8	−2.2	−1.7
C_2H_4	531.9	0.8	−0.7	−0.2
C_2H_6	666.3	−0.7	−0.4	0.3
CN	176.6	−0.6	−0.8	−0.6
HCN	301.8	2.3	0.8	1.0
CO	256.2	1.9	2.8	1.8
HCO	270.3	0.0	1.7	1.1
H_2CO	357.2	1.3	2.7	2.1
H_3COH	480.8	0.1	1.9	1.5
N_2	225.1	0.3	−1.2	−1.3
H_2NNH_2	405.4	−1.6	−0.6	−1.0
NO	150.1	−1.2	1.3	0.5
O_2	118.0	−2.5	−2.1	−2.4
HOOH	252.3	−0.2	1.2	−0.2
F_2	36.9	−0.2	0.7	−0.3
CO_2	381.9	2.0	4.2	2.7
av abs dev		1.1	1.2	1.0
max abs dev		4.5	4.2	2.7

[a] Taken from ref 1a. [b] Taken from ref 1b.

The results using the G2MS scheme of eq 1 are given in Table 1 and compared to the results from the G2[1a] and G2MP2[1b] schemes. The average deviation from experiment is 1.1 kcal/mol for G2MS, compared to 1.0 kcal/mol for G2 and 1.2 kcal/mol for G2MP2. The maximum deviation for G2MS is 4.5 kcal/mol, similar to 4.2 kcal/mol for G2MP2. Overall, the least expensive G2MS method performs as well as G2MP2. The G2MS maximum error occurs for BeH, for the which G2MP2 error is 2.2 kcal/mol. The large G2MS error for BeH is an example of MP2 not being able to account for the basis set effect in CCSD(T). The largest error for G2MP2 is for CO_2,

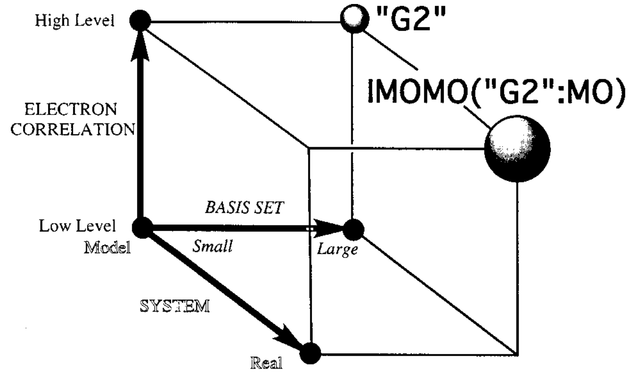


Figure 1. Schematic representation of the simultaneous extrapolation scheme labeled IMOMO("G2":MO). The basis set and the electron correlation are extrapolated as in G2-type schemes and are labeled as "G2", and, for the molecular size, the extrapolation from a small model to the large real system is handled by the IMOMO method.

for which the G2MS error is only 2.0 kcal/mol, an advantage of G2MS in combustion calculations.

We have attempted to further reduce the computational cost by using a smaller basis set, but the average and maximum deviations increased substantially. For example, by removing the *f* function for the heavy atoms in the large MP2 calculation in the G2MS scheme (MP2/6-311+G(2d,2p)), the average deviation increased by 0.5 to 1.6 kcal/mol and the maximum deviation by over 2 kcal/mol to *ca.* 7 kcal/mol. We conclude that 6-311+G(2df,2p) is the smallest possible basis set that could be used in the MP2 calculation if an accuracy of 1−2 kcal/mol is required.

B. IMOMO(G2MS:MO) Method. One of the reasons for developing the G2MS approach is to use it as a high-level method for the model in an integrated MO+MO (IMOMO) approach. Figure 1 illustrates the concept of a three-dimensional extrapolation scheme, connecting the G2-type approaches and IMOMO. The G2-type scheme is a two-dimensional extrapolation of the effect of a large basis set evaluated at a lower level (such as MP2) of correlation and the effect of high-level (such as CCSD(T)) electron correlation evaluated with a small basis set with the goal of estimating the results of high-level electron correlation with a large basis set. Now add to this scheme a third axis representing the size of the system. With IMOMO-(high:low), one can extrapolate the results of a "high"-level calculation for a small model system and of a "low"-level calculation for the large real system and estimate the high-level result for the real system. If one uses a G2-like extrapolation scheme as the high level method for the small model system in the IMOMO scheme, one can estimate the high correlation, large basis set results for the large real system. With the G2MS method, one can include eight to ten heavy atoms in the model system describing the reaction center, which would be sufficient for most reactions including electrocyclic reactions where the reaction directly involves reorganization of many electrons and atoms. The real system can contain 20 or so heavy atoms if the MP2 method is required as the low-level method, and much more if the HF method is satisfactory. One may also include molecular mechanics (MM) as an external layer and construct an ONIOM method such as ONIOM3(G2MS:MP2:MM).

C. Test of the G2MS Scheme for the Diels−Alder Additions of Butadiene + Ethylene and Cyclopentadiene + Ethylene. In order to test the G2MS scheme for a system relevant to the present study, we have chosen to study a few different Diels−Alder reactions, beginning with the smallest possible [4+2] cycloaddition, i.e., *s-trans*-1,3-butadiene + ethylene. The synchronous concerted transition state for this reaction is shown in Figure 2. It is worth noting that the B3LYP structures using the 6-31G and 6-31G(d) basis sets are very similar. The activation barrier differs by less than 0.1 kcal/mol between these two structures as calculated at the B3LYP/6-31G(d) level. On the basis of this result, all Diels−Alder transition states throughout this paper are optimized and the ZPC is calculated at the B3LYP/6-31G level.

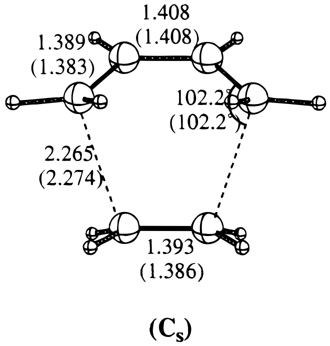


Figure 2. Important geometrical parameters (bond lengths, Å, bond angles, deg) of the B3LYP/6-31G and B3LYP/6-31G(d) (in parentheses) optimized transition state for the ethylene + butadiene reaction.

TABLE 2: Activation Barriers (kcal/mol) with Different Methods and Basis Sets for the Diels−Alder Reaction between Ethylene and *s-trans*-Butadiene and between Ethylene and Cyclopentadiene[a]

	s-trans-butadiene					cyclo-pentadiene
	6-31G (d)	6-31G (d,p)	6-311G (d,p)	6-311+G (2df,2p)	6-311+ (3df,2p)	6-31G(d)
HF	46.3	46.6	49.1	50.9	50.5	41.4
MP2	20.2	19.4	17.5	16.6	16.4	14.1
MP3	29.6	29.0	28.2			
MP4(SDTQ)	25.1	24.5	22.8			
CCSD	31.9	31.4	30.8			
CCSD(T)	27.6	27.0	25.7			22.3
G2MS				23.9		18.5
G2MP2[b]					24.6	

[a] The B3LYP/6-31G optimized geometry and ZPC was used. [b] G2MP2 = CCSD(T)/6-311G(d,p) + MP2/6-311+G(3df,2p) - MP2/6-311G(d,p), using the B3LYP/6-31G geometry and ZPC.

The activation barrier at different levels of theory including the B3LYP/6-31G(d) ZPC of 2.4 kcal/mol is presented in Table 2. The correlation effects (CCSD(T) − HF) are very large for this reaction, in fact more than 20 kcal/mol. The perturbation series does not converge satisfactorily as the MP4 − MP3 difference is −5.4 kcal/mol (in the 6-311G(d,p) basis set) while the MP3 − MP2 difference is +10.7 kcal/mol. The triple effect for CCSD is large, 5.1 kcal/mol, but is moderately stable (within 0.8 kcal/mol) with different basis sets.

Table 2 illustrates the usefulness of the G2MS approach. We cannot afford the very expensive CCSD(T)/6-311+G(2df,2p) calculation, but in G2MS this target is estimated by considering the effect of the very large basis set at the much less expensive MP2 level. The effects of expanding basis sets at the MP2 level as shown in Table 2 show that the activation energy is converged at 6-311+G(2df,2p), as the change in activation energy going to 6-311+G(3df,2p) is very small. The CCSD(T) has not converged at the 6-31G(d) level, as the barrier decreases by 1.9 kcal/mol going to 6-311G(d,p), vs the corresponding MP2 decrease of 2.7 kcal/mol, indicating that the correspondence of CCSD(T) vs MP2 is good and should be better on going to 6-311+G(2df,2p). The G2MS scheme gives a barrier of 23.9 kcal/mol, and the G2MP2 barrier is very close at 24.6 kcal/mol. Two different experimental numbers have been reported in the literature, an old value of 27.5 kcal/mol[21] at very high

TABLE 3: Activation Barriers (kcal/mol) with Different Pure MO and IMOMO Methods for the Diels−Alder Reaction between Two *s-trans*-1,3-Butadiene Molecules and between Maleic Anhydride and Acrolein as Dienophiles and *s-trans* Isoprene (2-methyl) and *s-cis* 2-*tert*-Butyl-1,3-butadiene (2-*tert*-butyl) as Dienes

diene	HF	MP2	CCSD(T): MP2	G2MS: MP2	exp
			Butadiene Dienophile		
butadiene	48.0	18.3	27.2(30.1[b])	23.5	23.7 ± 0.2[c] 24.5[d]
			Maleic Anhydride Dienophile		
2-methyl	37.2	5.5	11.4	9.2	12.2[e]
2-*tert*-butyl	33.7	0.7	6.8	4.6	6.5[e]
			Acrolein Dienophile		
2-methyl	43.3	14.5	21.3	17.6	18.7[f]
2-*tert*-butyl	40.3	11.0	18.0	14.2	
			Rotational Isomerization of Diene[g]		
2-methyl TS	7.6	7.6	7.3	7.9	
2-methyl *cis*	3.6	3.4	3.4	3.7	
2-*tert*-butyl TS	2.6	2.4	3.1	3.7	
2-*tert*-butyl *cis*	−2.0	−1.6	−1.4	−1.1	

[a] The B3LYP/6-31G optimized geometry and ZPC were used. The basis set used for energies is 6-31G(d), except for G2MS for which the basis sets are defined in eq 1. [b] IMOMO(CCSD(T):HF). [c] Reference 24. [d] Reference 25. [e] Reference 29. [f] Reference 28. [g] The rotational transition state and *s-cis* isomer of isoprene (2-methyl) and 2-*tert*-butyl-1,3-butadiene (2-*tert*-butyl), relative to the *s-trans* isomer.

temperatures and a more recent and presumably more reliable estimate of 25.1 kcal/mol.[22] The present G2MS and G2MP2 results both support the latter value.

We have slightly increased the size of our system to study the addition of ethylene to cyclopentadiene, and the results shown in Table 3 were not quite as successful at the G2MS level of theory. The G2MS approach differed from experiment by over 5 kcal/mol as the predicted value of 18.5 kcal/mol differed from the experimental value of 23.7 kcal/mol,[23] but this reference suggests their predicted value of E_a could be too large and a value of 21.9 kcal/mol is more plausible. The experimental data for this reaction were for reactions at temperatures between 521 and 570 K, and the high temperatures may lead to a slight overestimation of the activation barrier.

D. Test of the IMOMO(G2MS:MP2) Scheme for the Diels−Alder Addition of Butadiene + Butadiene. The first system examined for testing the IMOMO(G2MS:MO) method was the Diels−Alder dimerization of butadiene, using ethylene + butadiene as the model system. A few activation energies have been experimentally determined at different temperature ranges, with lower temperature values of 23.7 ± 0.2 and 24.5 kcal/mol.[24,25] The calculated activation barriers are shown in Table 3. As with all Diels−Alder systems studied, the pure HF method gives an activation barrier which was too large, while the pure MP2 overcorrects and leads to a barrier which is too small. The IMOMO(CCSD(T):MP2) scheme with the unextrapolated CCSD(T) gives a barrier too high by 3.5 kcal/mol. HF as the lower level method seems to be unable to properly take account of the electronic effect of butadiene conjugation in the dienophile and introduces an additional error of 2.9 kcal/mol. The IMOMO(G2MS:MP2) value for the activation energy was estimated to be 23.5 kcal/mol, within 0.2 and 1.0 kcal/mol of the two experimental values. The analysis of the G2MS barrier shows that the barrier for the model system is 23.1 kcal/mol, with a ΔE(MP2, model→real) correction of +0.4 kcal/mol. It should be noted that calculations at any G2-type method for the eight carbon real system would be nearly impossible. There would be difficulty doing G2 or CBS calculations with the IMOMO scheme with a six carbon model system, and G2MP2 calculation could be done with some difficulty, but there is no problem with the present G2MS method.

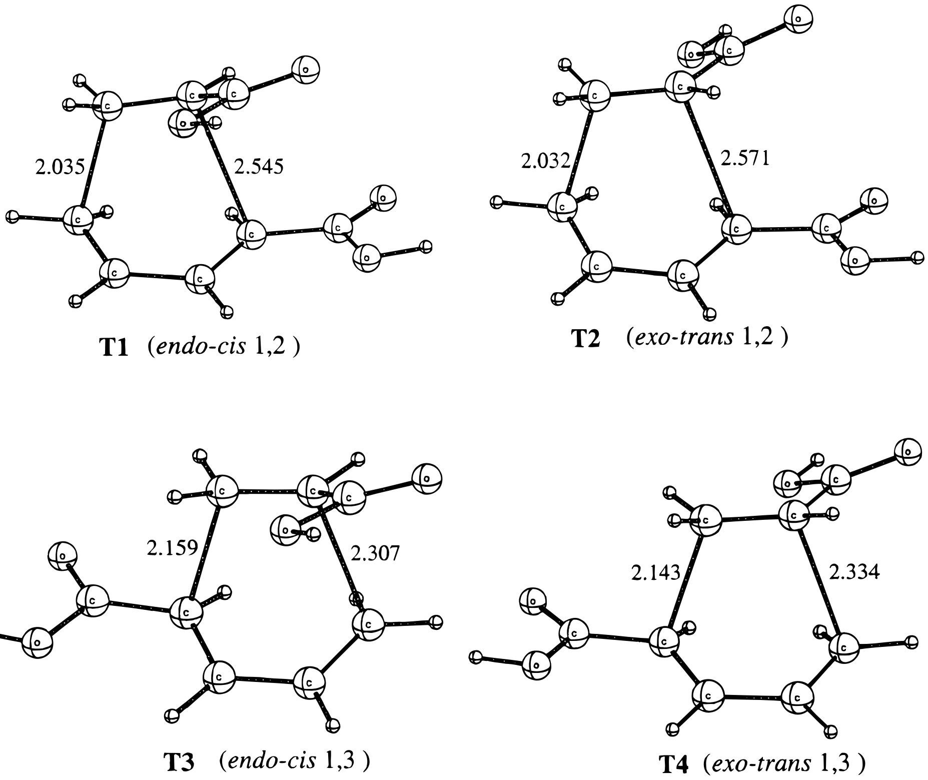
III. Branching Ratio in the Acrylic Acid + 2,4-Pentadienoic Acid Diels−Alder Addition

The reaction of acrylic acid with 2,4-pentadienoic acid was studied with the IMOMO(G2MS:MP2) method using ethylene + butadiene as the model system. All geometries were optimized and ZPC were calculated at the B3LYP/6-31G level. Figure 3 depicts the four transition state structures and the bond distances of the forming bonds, while Table 4 shows the relative energies of four transition states and four products, relative to the reactant.

The transition state structures as depicted in Figure 3 are very asymmetric, and, as a general trend, the more asymmetric the transition state, the lower the barrier. This trend has been seen before,[9,12] and the simplest reaction, ethylene + butadiene with a perfectly symmetric transition state as discussed above, has a high barrier of *ca.* 25 kcal/mol.[22] At the highest level of integrated theory, IMOMO(G2MS:MP2), which was found in the preceding section to be quantitatively accurate, the transition states T1 and T2 for head-to-head (HH or 1,2) addition are favored over those, T3 and T4, for head-to-tail (HT or 1,3) addition by at least 2 kcal/mol, and the *endo* T1 or T3 is slightly favored over the corresponding *exo* T2 or T4, respectively.[26] Thus, the barrier heights increase in the order: T1 < T2 < T3 < T4. An earlier suggestion that at the transition state orbital interactions of the dienophile are more important than the diene[9,12] is consistent with the present finding. Due to their early nature, steric interaction plays little role in these transition states, while on the other hand, steric interactions are critically important in the energetics of the cyclohexadiene products. Of the four products, the two most stable ones are *trans*, with the *trans*-1,2 species being lowest in energy. The relative energies of the products: P2 < P4 < P3 < P1, are nearly independent of the order of the transition states.

The IMOMO(G2MS:MP2) relative free energies, ΔG, at the four transition states at 110 °C give the branching ratio as shown in Table 4. The IMOMO derived branching ratio of 76:23:1:1 agrees well with the experimental ratio of 61:22:9:8 at 110 °C.[24,27] Considering the sensitivity of the branching ratio to slight energy changes, the agreement is excellent.

A systematic comparison can be made among a variety of methods listed in Table 4. Clearly the pure HF method is not suitable, as not only are the activation barriers approximately 25 kcal/mol too high, but also the order of the transition states is incorrect. Any IMOMO method with HF as the low-level method gives an incorrect order of transition states, indicating that the electronic effect of the carboxyl group cannot be accounted for by the HF method. The pure MP2 method predicts the correct ordering of the transition states, though the barrier is underestimated by as much as 30%. All of the IMOMO methods with MP2 as the low level give the correct order of the transition states. Thus, the role of the high-level calculation of the six carbon model system is to give the correct absolute value of the barrier by including high-order electron correlation. Because the model systems for the four transition states are similar in structure, it might be expected that the main energy difference between any pair of transition states in the IMOMO(G2MS:MP2) approach would come from the MP2 calculation for the real system. However, the difference between T1 and T4 at the IMOMO(G2MS:MP2) level is 2.8 kcal/mol while that at the pure MP2 level is 2.1 kcal/mol, indicating a substantial energy correction of 0.7 kcal/mol comes from the



T1 (*endo-cis* 1,2) **T2** (*exo-trans* 1,2)

T3 (*endo-cis* 1,3) **T4** (*exo-trans* 1,3)

Figure 3. Bond distances (Å) of the forming bonds at the four B3LYP/6-31G optimized transition states in the addition of acrylic acid with 2,4-pentadienoic acid.

TABLE 4: Activation Barriers (kcal/mol) for the Four Transition States in Figure 3 and Energies of Reaction (kcal/mol) for Four Products with Different Pure MO and IMOMO Methods for the Diels−Alder Addition of Acrylic Acid to 2,4-Pentadienoic Acid[a]

	T1 (H,H,endo)	T2 (H,H,exo)	T3 (H,T,endo)	T4 (H,T,exo)	P1 (1,2,cis)	P2 (1,2,*trans*)	P3 (1,3,cis)	P4 (1,3,trans)
HF	45.5	46.4	43.6	44.0	−27.8	−30.8	−28.4	−31.3
MP2:HF	18.2	19.2	16.7	17.1	−36.3	−38.7	−36.5	−39.2
CCSD(T):HF	25.3	26.2	24.5	24.8	−30.6	−33.0	−30.8	−33.5
MP2	11.4	12.2	13.3	13.5	−42.4	−44.9	−42.7	−44.7
MP3:MP2	21.0	21.7	23.2	23.4	−40.4	−43.0	−40.8	−42.7
MP4:MP2	23.1	23.8	25.5	25.6	−38.4	−41.0	−38.9	−40.8
CCSD(T):MP2	18.4	19.1	21.1	21.3	−38.4	−41.0	−38.9	−40.8
G2MS:MP2	14.4	15.1	17.1	17.2	−33.2	−35.6	−33.6	−35.4
Δ*G*(G2MS:MP2)	0.00	0.30	2.31	2.36				
branching ratio based on Δ*G*(G2MS:MP2)					75.5	22.9	0.9	0.7
exp branching ratio[b]					61	22	9	8

[a] The B3LYP/6-31G optimized geometry and ZPC was used. The basis set used for energies is 6-31G(d), except for G2MS for which the basis sets are defined in eq 1. Free energy predictions were calculated at the experimental temperature of 110 °C. [b] References 24a and 27.

geometry changes in the model system. The lowering of the barrier going from the unextrapolated IMOMO(CCSD(T):MP2) to the extrapolated IMOMO(G2MS:MP2) method is as large as 4 kcal/mol; estimating the results for a large basis set for the small model system is essential to obtain a correct activation energy.

IV. Diels−Alder Additions of Acrolein and Maleic Anhydride Dienophiles to Isoprene and 2-*tert*-Butyl-1,3-butadiene

With the IMOMO(G2MS:MP2) method, we can now study accurately the potential energy surfaces of Diels−Alder reactions with complicated substituents on both diene and dienophile. In this section, we have examined the activation barriers and compared with experiment the following four reactions: (1) acrolein + isoprene, (2) acrolein + 2-*tert*-butyl-1,3-butadiene, (3) maleic anhydride + isoprene, and (4) maleic anhydride + 2-*tert*-butyl-1,3-butadiene. The last system was the largest and contained 15 heavy atoms, but we used ethylene + butadiene as the small model system in all of these IMOMO predictions. Changing the diene from isoprene to 2-*tert*-butyl-1,3-butadiene, reactions 1 and 2 or reactions 3 and 4 should allow comparisons for the activation energies for different dienes, whereas the comparison of (1) with (3) or (2) with (4) will determine the effect of the dienophile on the reaction. For all of these reactions, the geometries were optimized and the ZPC was calculated and included at the B3LYP/6-31G level. For reactions 1 and 2, only the HH *endo* transition state was studied, as this was shown to have the lowest energy of the four, while for (3) and (4), the *endo* species was the only one examined, without HH/HT distinction because of the symmetry of the dienophile.

One of the issues which need to be clarified in these reactions of substituted 1,3-butadienes is the energy of *s-cis* and *s-trans* isomers of the reactant dienes and the transition state linking the two. As shown at the bottom of Table 3 for isoprene, the *s-cis* isomer lies 3.7 kcal/mol above the *s-trans* and the barrier from *s-trans* is 7.9 kcal/mol. With a barrier for isomerization lower than the barrier for the addition reaction (vide infra), the addition reaction should take place exclusively from the *s-cis* reactant. However, the large *tert*-butyl group at the 2-position of 2-*tert*-butyl-1,3-butadiene destabilizes the *s-trans* reactant, and the intrinsically unstable *s-cis* reactant is now the more stable isomer lying 1.1 kcal/mol below the *s-trans* one. The barrier for isomerization is 3.7 kcal/mol from the *s-trans* and 4.8 kcal/mol from the *s-cis*, and the addition reaction should also take place only from the *s-cis* reactant. The activation energies for Diels−Alder additions quoted in this section and Table 3 are all relative to the lowest energy reactants, *s-cis* for 2-*tert*-butyl-1,3-butadiene and *s-trans* for isoprene.

For the acrolein + isoprene reaction, the IMOMO(G2MS:MP2) barrier of 17.6 kcal/mol shown in Table 3 is only 1.1 kcal/mol lower than experiment,[28] and the prediction for acrolein + 2-*tert*-butyl-1,3-butadiene of 14.2 kcal/mol is expected to be quite reliable. This 3.4 kcal/mol decrease in the activation barrier going from 2-methyl to 2-*tert*-butadiene mainly comes from the fact, as discussed above, that in the latter the intrinsically stable *trans* reactant is sterically congested and destabilized and the reaction takes place from the intrinsically less stable *cis* reactant. The barrier height measured from the *cis* reactant changed little between the two dienes.

Maleic anhydride is considered to be one of the best dienophiles for Diels−Alder reactions due to the strong electron withdrawing substituents leading to an electron-deficient C=C double bond. Experimentally, the activation barriers for maleic anhydride with isoprene and 2-*tert*-butyl-1,3-butadiene are 12.2 and 6.5 kcal/mol,[29,30] respectively, which are very low compared to many other activation energies we have studied. The IMOMO(G2MS:MP2) activation energies are 9.2 and 4.6 kcal/mol, respectively; however, they are about 3 and 2 kcal/mol too low compared with experiment. Possibly the electronic effect of the acid anhydride group is too strong to be handled at the low MP2 level in conjunction with ethylene + butadiene as the model reaction at the high G2MS level. Pure MP2 predictions for Diels−Alder reactions are always too low; thus, possibly the electronic effects of the carbonyl substituents in maleic anhydride are too important to be calculated at the lower level MP2 approach. It is also possible, especially for 2-*tert*-butyl-1,3-butadiene, that the activation barrier for the addition reaction is so low that it competes with the *cis*−*trans* isomerization reaction of the reactant, complicating the kinetics. New experimental studies of the rate are desired.

V. Conclusions

A new G2-type approach, G2MS, has been proposed, tested, and compared with the G2 and G2MP2 approaches against experimental atomization energies for the standard G2 set of first-row compounds. Overall, the average absolute deviation is 1.1 kcal/mol, with the largest deviation of 4.5 kcal/mol for BeH, and the method performs as well as the others. However, this method is considerably less expensive than G2MP2 and G2 and can be used routinely for up to eight non-hydrogen atoms. This method can be integrated effectively into the IMOMO scheme, to extrapolate the correlation and basis set effects with respect to the size of the system. The combination of these two methods, in the form such as IMOMO(G2MS:MP2), is a natural and relatively inexpensive approach to accurate energy predictions for large systems, up to 7−9 heavy atoms at the active center and 20 or more heavy atoms exerting electronic effects on the active center, where no practical method has been available in the past.

The G2MS and IMOMO(G2MS:MP2) approaches have been applied with success to a number of Diels−Alder reactions. The G2MS activation energy for ethylene + butadiene is 23.9 kcal/mol, only 1.2 kcal/mol lower than experiment. Our data for the addition of ethylene with cyclopentadiene is slightly in error, possibly due to the high temperatures at which the gas phase experiments were done. Some of the other experiments like ethylene + butadiene were also examined at very high temperatures, and there could very easily be a 1−3 kcal/mol effect in raising the experimental activation energy on the basis of the higher temperatures.[31] The predictions for the butadiene dimerization were in near exact agreement with experiment.

On the basis of the free energies calculated with the IMOMO(G2MS:MP2) scheme, the branching ratio for the Diels−Alder addition of acrylic acid to 2,4-pentadienoic acid was derived, which is in nearly quantitative agreement with experiment despite the fact that very slight energy differences lead to large errors in the branching ratio. Very low activation energies have been obtained for reactions of maleic anhydride with isoprene and 2-*tert*-butyl-1,3-butadiene. The 4.6 kcal/mol (5.7 kcal/mol experimentally) decrease in the activation energy on going from isoprene to 2-*tert*-butyl-1,3-butadiene has been attributed to the fact that while the reactant isoprene is the *s-trans* isomer, 2-*tert*-butyl-1,3-butadiene is in the intrinsically less stable *s-cis* form, from which the reaction barrier is smaller.

Acknowledgment. The authors are grateful to Drs. Stefan Sieber, Toshiaki Matsubara, and Jamal Musaev and Mr. Qiang Cui for their discussions and assistance and to Prof. James Coxon for comments on the manuscript. The use of the Emerson Center computing facilities is acknowledged. The present research is in part supported by Grants CHE-9409020 and CHE-9627775 from the National Science Foundation and by Grant F49620-95-1-0182 from the Air Force Office of Scientific Research. R.D.J.F. acknowledges a Postdoctoral Fellowship from the Natural Sciences and Engineering Research Council of Canada. M.S. acknowledges a Postdoctoral Fellowship from the Swedish Natural Science Research Council.

References and Notes

(1) (a) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221. (b) Curtiss, L. A.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1993**, *98*, 1293. (c) Curtiss, L. A.; Redfem, P. C.; Smith, B. J.; Radom, L. *J. Chem. Phys.* **1996**, *104*, 5148.

(2) (a) Montgomery, J. A., Jr.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **1994**, *101*, 5900. (b) Nyden, M. R.; Petersson, G. A. *J. Chem. Phys.* **1981**, *75*, 1843. (c) Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Al-Laham, M. A.; Shirley, W. A.; Mantzaris, J. *J. Chem. Phys.* **1988**, *89*, 2193. (d) Petersson, G. A.; Al-Laham, M. A. *J. Chem. Phys.* **1991**, *94*, 6081. (e) Petersson, G. A.; Tensfeldt, T. G.; Montgomery, J. A., Jr.; *J. Chem. Phys.* **1991**, *94*, 6091.

(3) Mebel, A.; Morokuma, K.; Lin, M. C. *J. Chem. Phys.* **1995**, *103*, 7414.

(4) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700.

(5) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170.

(6) Matsubara, T.; Maseras, F.; Koga, N.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 2573.

(7) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959.

(8) Svensson, M.; Humbel, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 3654.

(9) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.*, in press.

(10) Diels, O.; Alder, K. *Justus Liebigs Ann. Chem.* **1928**, *460*, 98.

(11) For example: (a) Brown, F. K.; Houk, K. N. *Tetrahedron Lett.* **1985**, *26*, 2297. (b) Gajewski, J. J.; Peterson, K. B.; Kagel, J. R. *J. Am. Chem. Soc.* **1987**, *109*, 5545. (c) Brown, F. K.; Houk, K. N.; Burnell, D. J.; Valenta, Z. *J. Org. Chem.* **1987**, *52*, 3050. (d) Loncharich, R. J.; Brown, F. K.; Houk, K. N. *J. Org. Chem.* **1989**, *54*, 1129.

(12) For example: (a) Jorgenson, W. L.; Lim, D.; Blake, J. F. *J. Am. Chem. Soc.* **1993**, *115*, 2936. (b) Froese, R. D. J.; Organ, M. G.; Goddard, J. D.; Stack, T. D. P.; Trost, B. M. *J. Am. Chem. Soc.* **1995**, *117*, 10931.

(13) (a) Dewar, M. J. S.; Pierini, A. B. *J. Am. Chem. Soc.* **1984**, *106*, 203. (b) Dewar, M. J. S. *J. Am. Chem. Soc.* **1984**, *106*, 209. (c) Dewar, M. J. S.; Olivella, S.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1986**, *108*, 5771.

(14) For example: (a) Basilevsky, M. V.; Tikhomirov, V. A.; Chlenov, I. E. *Theor. Chim. Acta* **1971**, *27*, 2791. (b) Kikuchi, O. *Tetrahedron* **1971**, *27*, 2791. (c) Marshall, J. A.; Grote, J.; Audia, J. E. *J. Am. Chem. Soc.* **1987**, *109*, 1186.

(15) (a) Burke, L. A.; Leroy, G.; Sana, M. *Theor. Chim. Acta* **1975**, *40*, 313. (b) Ortega, M.; Oliva, A.; Lluch, J. M.; Bertran, J. *Chem. Phys. Lett.* **1983**, *102*, 317. (c) Bernardi, F.; Bottoni, A.; Robb, M. A.; Field, M. J.; Hillier, I. H.; Guest, M. F. *J. Chem. Soc., Chem. Commun.* **1985**, 1051.

(16) Bauschlicher, C. W., Jr.; Partridge, H. *J. Chem. Phys.* **1994**, *100*, 4329.

(17) The higher order correlation calculation in the G2, G2MP2, and G2(MP2,SVP) schemes is QCISD(T).

(18) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*, Revision A.1; Gaussian, Inc.: Pittsburgh, PA, 1995.

(19) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (b) Ricca, A.; Bauschlicher, C. W., Jr. *J. Chem. Phys.* **1994**, *98*, 12899. (c) Ricca, A.; Bauschlicher, C. W., Jr. *J. Chem. Phys.* **1995**, *99*, 5922. (d) Ricca, A.; Bauschlicher, Jr., C. W. *J. Chem. Phys.* **1995**, *99*, 9003.

(20) For the 32 molecules in Table 1, the maximal B3LYP/6-31G(d) energy difference between the B3LYP/6-311G(d,p) and B3LYP/6-31G(d) optimized geometries is less than 0.1 kcal/mol. The computational savings in the optimizations are 4−5 times.

(21) Rowley, D.; Steiner, H. *Discuss. Faraday Soc.* **1951**, *10*, 198.

(22) Li, Y.; Houk, K. N. *J. Am. Chem. Soc.* **1993**, *115*, 7478.

(23) Walsh, R.; Wells, J. M. *J. Chem. Soc., Perkin Trans. 2* **1976**, 52.

(24) (a) Sauer, J.; Sustmann, R. *Angew. Chem., Int. Ed. Engl.* **1980**, *19*, 779. (b) Kistiakowsky, G. B.; Ransom, W. W. *J. Chem. Phys.* **1939**, *7*, 725.

(25) Huybrechts, G.; Luyckx, L.; Vandenboom, Th.; van Mele, W. H. *Int. J. Chem. Kinet.* **1977**, *9*, 283.

(26) For the transition states, we use the terms HH/HT and *endo/exo* for the regio- and stereoselectivity, respectively, and for the products, 1,2/1,3 and *cis/trans*, respectively. We sometimes interchange these terms to avoid confusion.

(27) Fleming, I.; Michael, J. P.; Overman, L. E.; Taylor, G. F. *Tetrahedron Lett.* **1978**, 1313.

(28) Kistiakowsky, G. B.; Lacher, J. R. *J. Am. Chem. Soc.* **1936**, *58*, 123.

(29) Craig, D.; Shipman, J. J.; Fowler, R. B. *J. Am. Chem. Soc.* **1961**, *83*, 2885.

(30) Solvent effects appear to be important in this reaction as a range of values from 11.8−14.2 have been predicted in a separate study (Dewar, M. J. S.; Pyron, R. S. *J. Am. Chem. Soc.* **1970**, *92*, 3098).

(31) For example, in butadiene dimerization, experiments at 446−660 K led to a barrier of 23.7 kcal/mol (ref 21b above), while experiments at 691−923 K led to a barrier of 26.8 kcal/mol (ref 17 above).