

Electron correlation, basis sets, and the methylene singlet–triplet gap

Emily A. Carter and William A. Goddard III^{a)}

Arthur Amos Noyes Laboratory of Chemical Physics,^{b)} California Institute of Technology, Pasadena, California 91125

(Received 15 September 1986; accepted 9 October 1986)

The effect of basis set and electron correlation on the singlet–triplet splitting (ΔE_{ST}) of CH_2 is examined using the generalized valence bond (GVB) approach. For a standard double zeta plus polarization basis, the GVB based calculation (with only 20–25 spin eigenfunctions) approaches the full CI result ($\sim 220\,000$ spin eigenfunctions) of Bauschlicher and Taylor to within 0.5 kcal/mol for this basis, but both differ substantially from experiment (errors of 2.4 and 2.9 kcal/mol for GVB and full CI, respectively). We have studied the convergence of ΔE_{ST} with basis set and find that an extremely extended basis (triple zeta *sp*, diffuse *sp*, triple zeta *d*, double zeta *f*) for GVB yields $\Delta E_{ST} = 9.03$ kcal/mol, in excellent agreement with the experimental value of 9.09 ± 0.20 kcal/mol.

I. INTRODUCTION

In principle, the methods of quantum chemistry can provide exact answers to quantitative questions of physical and chemical interest such as molecular geometries, vibrational frequencies, excitation energies, bond energies, and even activation energies for chemical reactions. However, even for such small molecules as methylene, calculations must always be restricted in the basis set and level of electron correlation. Hence, it is essential to understand the level of error engendered by restrictions in basis or level of correlation. Because the relative energies of the lowest singlet and triplet states ($\Delta E_{ST} = E_{\text{singlet}} - E_{\text{triplet}}$) of substituted methylenes have been shown to be critical in determining the chemistry of such systems,¹ and because the singlet–triplet gap of methylene is now well established experimentally to be $\Delta E_{ST} = 9.09 \pm 0.20$ kcal/mol,² we have selected CH_2 for a study of the dependence of ΔE_{ST} on basis set and level of correlation.

There are two paradigms for including electron correlation in wave functions. One is to start with the Hartree–Fock (HF) wave function and then to include some level of excitation from occupied to virtual orbitals (commonly single and double excitations denoted HF*S*D). The problem with this approach is that a bias in the HF level of description may well remain upon any fixed level of excitation. Thus, for CH_2 using a standard basis set (valence double zeta plus polarization, VDZ*p*), the $\Delta E_{ST} = 26.1$ kcal/mol for HF and 15.3 kcal/mol for HF*S*D (see Table I). The limit of this approach is to carry out a full CI for the given basis set. Unfortunately such full CI calculations may quickly become impractical ($\sim 220\,000$ spin eigenfunctions for DZ*p* and $\sim 114\,000\,000$ spin eigenfunctions for the TZ3*p2fn* basis discussed below).

The alternative approach, which has proved practical for moderate-sized systems, is to solve self-consistently for the orbitals of the correlated wave function. In general, such wave functions are referred to as MCSCF (multiconfigurational self-consistent field); however, to obtain energy differ-

ences, it is important to have a scheme that specifies which level of MCSCF is consistent for the two states. This is provided by the generalized valence bond (GVB) approach^{3,4} in which there is one valence orbital for each valence electron. Thus, the GVB wave function for both the singlet and triplet states of CH_2 involve six optimized orbitals (in addition to the doubly occupied C 1*s*-like orbital). This wave function is often denoted GVB(3/6) to indicate that three pairs of electrons are described with six orbitals. Advantages of the GVB approach are that (i) the correlated wave function can be interpreted in terms of one-electron orbitals, and (ii) accurate excitation energies and bond energies can be obtained using rather simple wave functions.

Recently⁵ there has been discussion of whether such restricted wave functions may provide energy differences more accurate than is consistent with the basis set being used. In order to understand the magnitude of any such basis set bias, we carried out systematic studies of the convergence of GVB-like wave functions as a function of basis set.

II. GVB CALCULATIONS

The full GVB wave function for $^3\text{CH}_2$ or $^1\text{CH}_2$ would involve six overlapping orbitals optimized simultaneously with the combination of all permissible spin eigenfunctions (five for $^1\text{CH}_2$, nine for $^3\text{CH}_2$).³ Although practical for CH_2 , the $N!$ overlapping terms makes such calculations impractical for large systems, but we have found the following approach to serve quite adequately. Starting with the dominant spin eigenfunction (the valence bond or perfect-pairing spin function), the orbitals are optimized, allowing the two orbitals describing each electron pair to overlap but requiring the orbitals of different pairs to be orthogonal. These GVB orbitals are calculated in terms of the two natural orbitals for each pair to yield the GVB-PP wave function. In terms of these natural orbitals, the GVB-PP wave function of $^1\text{CH}_2$ has $2^3 = 8$ closed-shell determinants. An advantage of the GVB-PP wave function is that all two-electron interactions can be expressed in terms of Coulomb and exchange terms so that it is not necessary to transform the two-electron integrals from the atomic orbital basis to the molecular orbital representation.⁴ For the VDZ*p* basis, GVB-PP

^{a)} Author to whom correspondence should be addressed.

^{b)} Contribution No. 6832.

TABLE I. The singlet-triplet energy gap (ΔE_{ST}) for CH_2 for various levels of calculation within a VDZp basis.^a

Calculation	Total energies (h) ^b		ΔE_{ST} (kcal/mol)	
	1A_1	3B_1	VDZp	DZp
HF	-38.885 87 (1/1)	-38.927 49 (1/1)	26.12	
HF*S*D	-39.016 70 (415/577)	-39.041 07 (479/1485)	15.30	
GVB-PP	-38.938 77 -38.939 40 (8/8)	-38.953 29 -38.953 90 (4/4)	9.11	9.10
GVB-RCI(PP) ^c	-38.941 51 -38.942 12 (18/20)	-38.959 25 -38.959 73 (9/25)	11.13	11.05
GVB-RCI(opt) ^d	-38.941 89 -38.942 50 (18/20)	-38.960 35 -38.960 82 (9/25)	11.59	11.49
GVB-RCI(opt)*S*D ^e	-39.025 49 (7759/16 674)	-39.044 62 (5078/22 230)		12.00
Full CI	-39.027 18 (~44 000/~220 000)	-39.046 26		11.97 ^f
Experiment				9.09 ± 0.20 ^g

^aThe corresponding number of spatial configurations/spin eigenfunctions for each wave function is given below each total energy. Basis sets (DZp is the same as that used for the full CI results) are described in Sec. III.

^b1 h = 1 hartree = 627.5096 kcal/mol = 27.211 617 eV = 219 474.65 cm^{-1} .

^c(PP) = orbitals from the GVB perfect-pairing wave function were used for the CI basis.

^dOrbitals were optimized for the GVB-RCI.

^eAll single and double excitations of the six valence electrons were allowed from the self-consistent GVB-RCI wave function.

^fReference 5.

^gReference 2.

yields $\Delta E_{ST} = 9.11$ kcal/mol, but as the basis is increased, ΔE_{ST} becomes 7.01 kcal/mol (see Table II).

The major discrepancy between GVB-PP and the full GVB wave function is the lack of other spin couplings. These spin couplings are included by carrying out a restricted con-

figuration interaction (RCI) in which the two electrons of each pair are allowed to occupy the two natural orbitals of the pair in all three ways. For $^1\text{CH}_2$, this leads to $3^3 = 27$ configurations (nine of which are not allowed by symmetry). The GVB-RCI wave function can still be interpreted in

TABLE II. ΔE_{ST} (kcal/mol) as a function of basis set.^a

Basis set	HF	GVB-RCI			Full CI
		GVB-PP	PP ^b	Opt ^c	
DZp		9.10 (-0.000 63)	11.05 (-0.003 35)	11.49 (-0.003 73)	11.97 ^d
VDZp	26.11 (0.052 90)	9.11 (0.000 00)	11.13 (-0.002 74)	11.59 (-0.003 12)	
VDZ2p/TZ2p		8.15 (-0.003 79)	9.98 (-0.006 63)	10.42 (-0.007 07)	
VDZ2pn/TZ2p		7.84 (-0.004 50)	9.63 (-0.007 38)	10.06 (-0.007 83)	
TZ2p		8.01 (-0.008 29)	9.69 (-0.011 36)	10.07 (-0.011 85)	
TZ2pf/TZ2p		7.81 (-0.008 83)	9.41 (-0.011 93)	9.84 (-0.012 41)	
TZ3p2fn/TZ2p	24.87 (0.045 14)	7.01 (-0.010 46)	8.60 (-0.013 52)	9.03 (-0.014 02)	(9.09) ^e
Change with basis	1.24	2.09	2.45	2.46	(2.88) ^e

^aNumbers in parentheses are the energies in hartrees of the singlet state relative to the GVB-PP wavefunction within the VDZp basis (total energy = -38.938 77 h).

^bSee Table I, footnote c.

^cSee Table I, footnote d.

^dReference 5.

^eAssuming that the full CI with the most extensive basis leads to the experimental value.

terms of the GVB orbital picture, and the GVB-RCI wave function is generally a good approximation to the full GVB wave function. Using the orbitals from the GVB-PP wave function, the GVB-RCI leads to $\Delta E_{ST} = 11.13$ and 11.05 kcal/mol for VDZ p and DZ p bases, converging to 8.60 kcal/mol for the extended TZ3 $p2fn$ basis. Rather than using the GVB-PP orbitals, we may solve self-consistently for the orbitals of the GVB-RCI wave function,^{4(b)} obtaining thereby an even better approximation to the full GVB wave function. Calculating the orbitals for the GVB-RCI self-consistently, leads to 11.59 and 11.49 kcal/mol for VDZ p and DZ p bases, in good agreement with the full CI result (DZ p) of 11.97 kcal/mol.⁵ For the largest basis set used, this self-consistent GVB-RCI yields 9.03 kcal/mol, in excellent agreement with experiment.

III. CALCULATIONAL DETAILS

A. Basis sets

We used the following basis sets:

VDZ p : The Dunning valence double zeta contractions⁶ of the (9 $s5p$) and the (4 s) Huzinaga Gaussian primitive bases⁷ for carbon and hydrogen (exponents scaled by 1.2) are used with one set of 3 d polarization functions ($\zeta^d = 0.64$) on carbon and one set of 2 p polarization functions ($\zeta^p = 1.0$) on hydrogen. The s combination of the carbon 3 d functions are excluded from all basis sets and all calculations, except for calculations using the DZ p basis set for direct comparison to the full CI result.⁵

DZ p : This basis set is given explicitly in Ref. 5. It differs from the VDZ p basis by having a DZ core on carbon, different 3 d polarization functions for each state, and the s combination of the d functions was included in the calculations.⁸

VDZ2 p /TZ2 p : For carbon, the same (9 $s5p/3s2p$) basis was used as above, with two sets of 3 d functions centered at 0.64 but scaled by 2.3 ($\zeta^d = 0.971$ and 0.422). For hydrogen, the Huzinaga⁷ unscaled (6 s) basis was contracted to triple zeta, with two sets of 2 p functions centered at 0.91 (optimized⁹ for H bonded to C) but scaled by 2.3 ($\zeta^p = 1.38$ and 0.60).

VDZ2 pn /TZ2 p : To the VDZ2 p /TZ2 p basis was added one set of diffuse s ($\zeta^s = 0.045$) and p ($\zeta^p = 0.034$) functions optimized for negative ions of carbon.⁹ This is denoted as "n" for negative ion.

TZ2 p : The Huzinaga (11 $s7p$) basis¹⁰ for carbon was contracted to (6 $s3p$) triple zeta for both core and valence. The carbon 3 d and hydrogen 2 p polarization functions are the same as in VDZ2 p /TZ2 p .

TZ2 pf /TZ2 p : One set of carbon 4 f functions [$\zeta^f = 0.96$, chosen to maximize overlap with the carbon d -function ($\zeta^d = 0.64$)], with the three 4 p combinations removed, was added to the carbon TZ2 p basis.

TZ3 $p2fn$ /TZ2 p : The Huzinaga (11 $s7p$) basis for carbon was contracted triple zeta for both core and valence as before, but diffuse s and p functions were added by scaling out (by 2.5) from the most diffuse exponents of the (11 $s7p$) set, yielding $\zeta^s = 0.0388$ and $\zeta^p = 0.0282$. Three sets of carbon 3 d -polarization functions were added, centered at 0.640 and scaled by 2.5 (leading to exponents $\zeta^d = 1.60$, 0.640 , and 0.256). Two sets of carbon 4 f functions were included, ob-

tained by scaling the previous 4 f exponent of 0.96 by 2.5, yielding $\zeta^f = 1.52$ and 0.607 . The 3 s and 4 p combinations of the d and f functions were removed. For hydrogen, the same Huzinaga triple zeta basis as above was used with two sets of p -polarization functions scaled by 3.0 from $\zeta^p = 0.91$, yielding $\zeta^p = 1.58$ and 0.525 .

B. Geometries

The equilibrium geometries for 1A_1 and 3B_1 CH₂ were taken from the GVB-POL-CI calculations of Harding and Goddard¹¹ who found $\theta_e = 133.2^\circ$ and $R_e = 1.084 \text{ \AA}$ for the 3B_1 state and $\theta_e = 101.8^\circ$ and $R_e = 1.113 \text{ \AA}$ for the 1A_1 state. For calculations involving the DZ p basis set, the geometries were taken from Ref. 5.

IV. RESULTS

The ΔE_{ST} using the VDZ p and DZ p bases for various wave functions are shown in Table I together with experimental (corrected for relativistic effects and zero-point motion)² and full CI⁵ values. (The full CI calculation involved all excitations of the six valence electrons in CH₂, with the C 1 s frozen at the HF level.) The self-consistent GVB-RCI (with 20 to 25 spin eigenfunctions) leads to $\Delta E_{ST} = 11.5$ kcal/mol, only 0.5 kcal/mol below the value $\Delta E_{ST} = 12.0$ kcal/mol obtained from the complete CI for this basis (with $\sim 220\,000$ spin eigenfunctions). Allowing all single and double excitations from the self-consistent GVB-RCI wave function (with one-tenth the number of spin eigenfunctions of the full CI wave function) yields $\Delta E_{ST} = 12.0$ kcal/mol, in complete agreement with the full CI. While a study of basis set convergence at the RCI*SD level is impractical, the small error engendered for ΔE_{ST} (0.5 kcal/mol) with the simpler GVB wave functions suggests that they may prove adequate for studying the convergence of ΔE_{ST} with basis set.

This convergence with basis set is shown in Table II, where we see that for an extremely large basis, the self-consistent GVB-RCI converges to $\Delta E_{ST} = 9.03$ kcal/mol, in excellent agreement with the experimental result. For smaller bases, the values for ΔE_{ST} decrease smoothly as the basis is extended, with the differential effects being quite similar for all three calculational levels. Solving self-consistently for the orbitals of the GVB-RCI wave function contributes ~ 0.43 kcal/mol energy lowering to the excitation energies. This occurs because the RCI wave function includes spin-couplings important for the triplet state which are omitted in the perfect singlet pairing wave function, resulting in orbital shape changes for the self-consistent GVB-RCI. For the uncorrelated HF wave function, the total change in ΔE_{ST} between the VDZ p basis and the full basis set is -1.24 kcal/mol, while it is -2.09 kcal/mol for GVB-PP and -2.46 kcal/mol for GVB-RCI. Similar studies of convergence for full CI as a function of basis set completeness are not available. Indeed, such a full CI test on our extended basis (TZ3 $p2fn$ /TZ2 p) may well be beyond the scope of current computers, since it would require ~ 23 million spatial configurations, ~ 114 million spin eigenfunctions, or ~ 455 million determinants! However, if we assume that the full CI would agree with experiment for the full basis

set we have used, the drop in the excitation energy would have to be 2.88 kcal/mol. This is reasonable since the corresponding quantities for lower level wave functions are 1.24 (HF), 2.09 (GVB-PP), and 2.46 (GVB-RCI) kcal/mol.

V. CONCLUSION

We have shown that the simple GVB description of the two lowest states of CH₂ (¹A₁ and ³B₁) leads to singlet-triplet gaps within 0.5 kcal/mol of the full CI result for a DZp basis set (still off from experiment by 2.4 kcal/mol). However, with GVB it is practical to use extremely extended bases, leading to results within 0.1 kcal/mol of the experimental value.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (Grant No. CHE83-18041). EAC acknowledges a National Science Foundation predoctoral fellowship (1982-1985).

¹E. A. Carter and W. A. Goddard III, *J. Phys. Chem.* **90**, 998 (1986); *J. Am. Chem. Soc.* **108**, 4746 (1986).

² $T_0 = 9.023 \pm 0.014$ kcal/mol is taken from P. R. Bunker, P. Jensen, W. P. Kraemer, and R. Beardsworth, *J. Chem. Phys.* **85**, 3724 (1986), and is corrected for relativistic effects and zero point motion by 0.07 ± 0.20 kcal/mol [I. Shavitt, *Tetrahedron* **41**, 1531 (1985)] to arrive at a nonrelativistic $T_e = 9.09 \pm 0.20$ kcal/mol.

³R. C. Ladner and W. A. Goddard III, *J. Chem. Phys.* **51**, 1073 (1969).

⁴(a) W. J. Hunt, T. H. Dunning, Jr., and W. A. Goddard III, *Chem. Phys. Lett.* **3**, 606 (1969); W. A. Goddard III, T. H. Dunning, Jr., and W. J. Hunt, *ibid.* **4**, 231 (1969); W. J. Hunt, W. A. Goddard III, and T. H. Dunning, Jr., *ibid.* **6**, 147 (1970); W. J. Hunt, P. J. Hay, and W. A. Goddard III, *J. Chem. Phys.* **57**, 738 (1972); F. W. Bobrowicz and W. A. Goddard III, in *Methods of Electronic Structure Theory*, edited by H. F. Schaefer (Plenum, New York, 1977), pp. 79-127; (b) L. G. Yaffe and W. A. Goddard III, *Phys. Rev. A* **13**, 1682 (1976).

⁵C. W. Bauschlicher, Jr. and P. R. Taylor, *J. Chem. Phys.* **85**, 5936 (1986).

⁶T. H. Dunning, Jr., *J. Chem. Phys.* **53**, 2823 (1970).

⁷S. Huzinaga, *J. Chem. Phys.* **42**, 1293 (1965).

⁸Excluding the *s* combination of the 3*d* functions in the DZP basis leads to an increase in ΔE_{ST} (at all levels reported in Table I) of no more than 0.02 kcal/mol and a destabilization of the total energies by no more than 0.000 28 hartrees (0.18 kcal/mol).

⁹R. A. Bair and W. A. Goddard III (submitted).

¹⁰S. Huzinaga and Y. Sakai, *J. Chem. Phys.* **50**, 1371 (1969).

¹¹L. B. Harding and W. A. Goddard III, *Chem. Phys. Lett.* **55**, 217 (1978).