and next through an internal conversion to the a^3A_1 state. The molecule will dissociate along the a^3A_1 potential energy curve to the products H and Mn(CO)$_5$ in their ground state. Irradiation at 229 nm will bring the system into the b^1E state. From there the molecule goes down along the b^1E potential energy curve corresponding to the Mn–CO elongation until it reaches a potential well. At this point the system evolves to the a^1E state through internal conversion, and dissociation will occur along the a^1E potential energy curve with the formation of the products CO and HMn(CO)$_4$, the latter as a square pyramid with H apical.

The nature of the a^3A_1 state that raises the energy of the b^3A_1 $\sigma \rightarrow \sigma^*$ photoactive state for the metal–hydrogen bond homolysis was not clearly identified. More sophisticated calculations, with larger basis sets, similar to those reported for Fe(CO)$_5$[56] are required. It is difficult to interpret the results obtained in low-temperature matrix experiments[28] since the data are largely dependent on the experimental conditions. Gas-phase experiments would be needed in order to ascertain the mechanism of photodissociation of HMn(CO)$_5$ proposed in the present study. The main reasons responsible for the different behavior upon irradiation of HCo(CO)$_4$ and HMn(CO$_5$) were outlined. The homolysis of the metal–hydrogen bond in monohydrides is a rather general reaction, although it may be obscured in some systems by the competition with the photoelimination of other ligands. The general ocurrence of these two photoreactions results from the dissociative character of the potential energy curves associated with a certain type of excited state. From this study and the previous one,[20] it is clear that the homolysis of the metal–hydrogen bond results from the dissociative character of the potential energy curve for the triplet state corresponding to a $\sigma_{M-H} \rightarrow \sigma^*_{M-H}$ excitation. The dissociative character of the curve associated to the $d \rightarrow \sigma^*_{M-H}$ excitation is responsible for the ligand dissociation.

Further work is needed in order to understand the photochemical behavior of other RMn(CO)$_5$ systems as a function of the fragment R. The understanding of the influence of a π acceptor ligand on the photoreactivity of these complexes through the presence of a low-lying metal-to-ligand charge-transfer state is another challenge.

Acknowledgment. The author is grateful to Dr. A. Veillard, Pr. R. L. Sweany, and Pr. M. Poliakoff for helpful discussions. The calculations have been carried out on the CRAY-2 computer of the CCVR (Palaiseau, France) through a grant of computer time from the Conseil Scientifique du Centre de Calcul Vectoriel pour la Recherche.

Registry No. HMn(CO)$_5$, 16972-33-1.

Stabilities of Hydrocarbons and Carbocations. 1. A Comparison of Augmented 6-31G, 6-311G, and Correlation Consistent Basis Sets

Janet E. Del Bene,[*,†] Donald H. Aue,[*,‡] and Isaiah Shavitt[*,§]

Contribution from the Departments of Chemistry, Youngstown State University, Youngstown, Ohio 44555, University of California at Santa Barbara, Santa Barbara, California 93106, and The Ohio State University, Columbus, Ohio 43210. *Received July 1, 1991*

Abstract: Ab initio calculations have been performed at correlated levels of theory on several hydrocarbons and carbocations in order to investigate the basis set dependence of their computed hydrogenolysis energies, which determine the stabilities of these species relative to H$_2$ and CH$_4$, and from which proton affinities, hydride ion affinities, and isomerization energies may be derived. One series of basis sets studied is derived from the 6-31G(d,p) basis by augmentation with diffuse functions on carbon and additional first and second polarization functions on carbon and hydrogen. Also studied, to a more limited extent, is the triple-split 6-311G(2df,2pd) basis. The 6-31G and 6-311G series of basis sets are compared with the correlation-consistent polarized valence double-, triple-, and quadruple-ζ basis sets of Dunning, in which the sp part of the basis is systematically improved as the polarization space is augmented. The hydrogenolysis energies computed with the 6-31G basis sets show poorer convergence patterns upon augmentation, including divergence from experimental values in some cases, and show an excessive dependence on the presence of diffuse functions on carbon and on the choice of five versus six components for d polarization functions. These problems are attributed to an inadequate representation of the sp part in the valence double-split 6-31G basis sets. Better results, and reduced sensitivity with respect to the number of d components and to the addition of diffuse functions, are found with the triple-split 6-311G(2df,2pd) basis set. The correlation-consistent basis sets have been found to show smooth convergence of the computed hydrogenolysis energies, both internally and with respect to experimental data, and no significant dependence on the presence of diffuse functions or on the number of components for the d functions, except for the double-split member of the series. The correlation-consistent polarized valence triple-ζ basis set is recommended for accurate calculations of reaction energies, with 6-311+G(2df,2pd) (with diffuse functions) a reasonable alternative.

Introduction

As part of a continuing investigation of basis set and correlation effects on computed reaction energies,[1-7] we have embarked on a project to investigate reaction energetics of hydrocarbons and carbocations.[8] Our aim is to determine the level of theory which provides agreement with experimental data to within 1–3 kcal/mol, close to the uncertainty limits of many of the experiments, for computed proton affinities of hydrocarbons, hydride ion affinities of carbocations, and isomerization energies of both neutral and charged species. To accomplish this goal, it is necessary to assess the basis set, correlation, and geometry dependence of the computed properties. Such a detailed assessment is particularly important if we want to ensure that the chosen level of theory is

---

[†] Youngstown State University.

[‡] University of California at Santa Barbara.

[§] The Ohio State University.

(1) Del Bene, J. E. *J. Comp. Chem.* **1989**, *5*, 603.

(2) Del Bene, J. E. *Struct. Chem.* **1990**, *1*, 19.

(3) Del Bene, J. E.; Shavitt, I. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1989**, *23*, 445.

(4) Del Bene, J. E.; Shavitt, I. *J. Phys. Chem.* **1990**, *94*, 5514.

(5) Del Bene, J. E.; Shavitt, I. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1990**, *24*, 365.

(6) Del Bene, J. E.; Stahlberg, E. A.; Shavitt, I. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1990**, *24*, 455.

(7) Del Bene, J. E.; Shavitt, I. *THEOCHEM* **1991**, *80*, 499.

(8) Del Bene, J. E.; Aue, D. H.; Shavitt, I. Part 2, to be published.

0002-7863/92/1514-1631$03.00/0

substantially converged with respect to further improvements in the methodology, instead of just giving correct results by accidental cancellation of errors, since such cancellation cannot be relied upon in general.

As part of this project, we have computed the proton affinities, hydride ion affinities, and isomerization energies for a complete series of C_1, C_2, and C_3 neutral hydrocarbons and carbocations for which experimental data are well-known.[8] The basis sets which have been used are the 6-31G(d,p) basis set[9,10] and those derived from it by adding diffuse functions on carbon[11,12] and additional first and second polarization functions on carbon and hydrogen.[13] The results of this study show that the proton affinities and hydride ion affinities computed with this series of basis sets do not always converge with increasing basis set size.[8] In fact, in some cases, energies obtained with smaller basis sets in this series are in better agreement than those from much larger basis sets with what may be regarded as nearly converged computed values. In trying to ascertain the nature of the difficulties, we sought to isolate basis set effects on the neutral hydrocarbon from those on the carbocation by computing the hydrogenolysis energies of each of these species with various basis sets. This approach allows the evaluation of basis set effects on the stability of each hydrocarbon and carbocation independently, relative to H_2 and CH_4. As an example, the hydrogenolysis energy of C_2H_2 is the energy of the reaction

$$C_2H_2 + 3H_2 \rightarrow 2CH_4 \qquad (1)$$

while that of the corresponding cation $C_2H_3^+$ is the reaction energy for

$$C_2H_3^+ + 3H_2 \rightarrow 2CH_4 + H^+ \qquad (2)$$

The proton affinity of C_2H_2 can then be obtained as the difference in energy between reactions 1 and 2. Similarly, the hydride ion affinity of $C_2H_3^+$ is obtained as the energy difference between reaction 2 and the hydrogenolysis of C_2H_4, plus the energy of the reaction $H^+ + H^- \rightarrow H_2$.

Implementation of this approach has, in fact, demonstrated that the hydrogenolysis energies of both the neutral hydrocarbons and the corresponding cations display the same convergence problems as observed for the proton affinities and hydride ion affinities. It has also been found that diffuse functions on carbon are generally required to improve agreement between theory and experiment, independently of the nature of the polarization space. Furthermore, the choice of five or six components for each set of d polarization functions can have a significant effect on the values of the hydrogenolysis energies, particularly for the larger basis sets, as well as on the convergence behavior of these energies with increasing basis set size.

Since these 6-31G basis sets are widely used for quantum chemical studies, it is important to understand this behavior and characterize it in detail. Our hypothesis is that the problems which have been encountered arise from an inadequate description of the sp part of the split-valence 6-31G(d,p) basis set. To test this hypothesis we examine in this paper hydrogenolysis energies computed for a subset of small hydrocarbons and carbocations chosen to represent common structural types. These types include structures with multiple bonds (acetylene and ethylene), a strained ring (cyclopropene), and nonclassical carbocations ($C_2H_3^+$ and $C_2H_5^+$), which represent special challenges to electronic structure theory.

The hydrogenolysis energies calculated with augmented 6-31G(d,p) basis sets are compared here with hydrogenolysis energies computed with the Dunning correlation-consistent polarized double-, triple-, and quadruple-split valence basis sets.[14] The Dunning basis sets have been chosen for comparison because they systematically expand the sp part of the basis set (in terms of both the primitive and contracted representations) as additional polarization functions are added. Clearly, this study is not intended to be an even contest between the two series of basis sets. Traditionally, improvements to the 6-31G basis have left the sp part unchanged as additional polarization functions and diffuse functions have been added. It is precisely this question of the need for improvement of the sp part that is a major focus of this investigation. We have also included limited studies with the triple-split 6-311G basis,[15] the only currently available sp-augmented successor to the 6-31G basis in this series of basis sets, but it should be noted that the 6-311G basis has been characterized as not being a true triple-split valence basis set, because the three-component s function in this basis contributes more to the inner-shell 1s orbital than to the valence shell.[16] Moreover, it has been demonstrated[13,17,18] that the need for the addition of diffuse functions, shown in the current study and elsewhere[13] for the 6-31G basis, persists for the 6-311G basis as well. This deficiency was traced[18] to the fact that the outermost p-type Gaussian in the 6-311G basis (exponent 0.145585 for carbon) is significantly less diffuse than in other basis sets of comparable size.

In this paper we demonstrate that the problems encountered with the augmented 6-31G(d,p) basis set do not occur when the sp part of the basis is systematically improved, as illustrated here by the Dunning polarized triple- and quadruple-split valence basis sets and, to a more limited extent, by the 6-311G basis. The criteria used here in assessing the performance of the various basis sets are the apparent convergence of the calculated total energies and hydrogenolysis energies for a particular wave function model as the basis sets are systematically augmented and improved, as well as the sensitivity of the results to the addition of diffuse functions and to the choice of five or six components for the d polarization functions. The Dunning quadruple-split basis set is considerably larger than any of the augmented 6-31G and 6-311G bases used, and is assumed to be close to the complete basis limit for the corresponding wave function model. The results for the quadruple-split basis serve as an "internal convergence" benchmark that we have used, together with comparisons with experimental data, to assess the accuracy of the various theoretical models. The internal convergence criterion is more useful in assessing basis set performance than comparison with experimental results, since such comparisons may be clouded by experimental uncertainties, by geometry effects, by vibrational energy corrections, and by questions of the convergence of the correlation model used.

A study comparing several basis sets in the 6-31G, 6-311G, and related series for MP2 calculations on hydrocarbons was published recently by Guo and Karplus,[19] but it focused on geometries and vibrational frequencies rather than reaction energies and did not consider diffuse functions or additional polarization functions. Among other recent basis set studies, a comparison of the 6-31G(d,p) basis with other basis sets in MP2 calculations on the hydrogen fluoride dimer was published by Cybulski et al.[20]

Methods

The geometries of the hydrocarbons and carbocations were optimized at the Hartree–Fock level with the 6-31G(d) basis set. Many of these structures had been optimized previously, and are available from the Carnegie-Mellon Quantum Chemistry Archive.[21] Hartree-Fock vibrational frequencies were computed for each structure to ensure that it is an equilibrium structure (no imaginary frequencies) on the potential

---

(9) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.

(10) Dill, J. D.; Pople, J. A. *J. Chem. Phys.* **1975**, *62*, 2921.

(11) Spitznagel, G. W.; Clark, T.; Chandrasekhar, J.; Schleyer, P. v. R. *J. Comp. Chem.* **1982**, *3*, 363.

(12) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comp. Chem.* **1983**, *4*, 294.

(13) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265.

(14) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.

(15) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.

(16) Grev, R. S.; Schaefer, H. F., III *J. Chem. Phys.* **1989**, *91*, 7305.

(17) Pople, J. A.; Head-Gordon, M.; Fox, D. J.; Raghavachari, K.; Curtiss, L. A. *J. Chem. Phys.* **1989**, *90*, 5622.

(18) Scuseria, G. E.; Schaefer, H. F., III *J. Chem. Phys.* **1989**, *90*, 3629.

(19) Guo, H.; Karplus, M. *J. Chem. Phys.* **1989**, *91*, 1719.

(20) Cybulski, S. M.; Chalasiński, G.; Moszyński, R. *J. Chem. Phys.* **1990**, *92*, 4357.

(21) Whiteside, R. A.; Frisch, M. J.; Pople, J. A. *The Carnegie-Mellon Quantum Chemistry Archive*; Carnegie-Mellon University: Pittsburgh, 1983.

Table I. Number of Contracted and Primitive Functions in the Basis Sets[a]

designation	composition	no. of functions per atom[b] C (5d)	C (6d)	H (5d)	H (6d)
6-31G(d,p)	[3s2p1d/2s1p]	14	15	5	5
	(10s4p1d/4s1p)	27	28	7	7
6-31+G(d,p)	[4s3p1d/2s1p]	18	19	5	5
	(11s5p1d/4s1p)	31	32	7	7
6-31G(2d,2p)	[3s2p2d/2s2p]	19	21	8	8
	(10s4p2d/4s2p)	32	34	10	10
6-31+G(2d,2p)	[4s3p2d/2s2p]	23	25	8	8
	(11s5p2d/4s2p)	36	38	10	10
6-31G(2df,2pd)	[3s2p2d1f/2s2p1d]	26	28	13	14
	(10s4p2d1f/4s2p1d)	39	41	15	16
6-31+G(2df,2pd)	[4s3p2d1f/2s2p1d]	30	32	13	14
	(11s5p2d1f/4s2p1d)	43	45	15	16
6-311G(2df,2pd)	[4s3p2d1f/3s2p1d]	30	32	14	15
	(11s5p2d1f/5s2p1d)	43	45	16	17
6-311+G(2df,2pd)	[5s4p2d1f/3s2p1d]	34	36	14	15
	(12s6p2d1f/5s2p1d)	47	49	16	17
cc-pVDZ	[3s2p1d/2s1p]	14	15	5	5
	(9s4p1d/4s1p)	26 (34)	27 (35)	7	7
cc-pVDZ+	[4s3p1d/2s1p]	18	19	5	5
	(10s5p1d/4s1p)	30 (38)	31 (39)	7	7
cc-pVTZ	[4s3p2d1f/3s2p1d]	30	32	14	15
	(10s5p2d1f/5s2p1d)	42 (50)	44 (52)	16	17
cc-pVTZ+	[5s4p2d1f/3s2p1d]	34	36	14	15
	(11s6p2d1f/5s2p1d)	46 (54)	48 (56)	16	17
cc-pVQZ′	[5s4p3d2f/4s3p2d]	46	49	23	25
	(12s6p3d2f/6s3p2d)	59 (68)	62 (71)	25	27
cc-pVQZ′+	[6s5p3d2f/4s3p2d]	50	53	23	25
	(13s7p3d2f/6s3p2d)	63 (72)	66 (75)	25	27
cc-pVQZ	[5s4p3d2f1g/4s3p2d1f]	55	58	30	32
	(12s6p3d2f1g/6s3p2d1f)	68 (77)	71 (80)	32	34
cc-pVQZ+	[6s5p3d2f1g/4s3p2d1f]	59	62	30	32
	(13s7p3d2f1g/6s3p2d1f)	72 (81)	75 (84)	32	34

[a] For each basis set, the composition and number of functions on each atom are given in the first row in terms of contracted functions and in the second row in terms of primitive functions. [b] The (5d) and (6d) headings refer to the number of components used for d functions. The numbers in parentheses take into account the repetition of primitives required for use with integral programs which are not designed for generalized contractions.

energy surface. The structures used for the cations $C_2H_3^+$ and $C_2H_5^+$ are the bridged structures, which are local minima on the Hartree–Fock potential energy surface (but correspond to global minima on the correlated surfaces). The optimized HF/6-31G(d) structures were used in all subsequent calculations. It is not expected that the differences between these structures and structures optimized with better basis sets or at correlated levels would have had an appreciable effect on the comparisons and conclusions presented in this study.

The basis sets investigated in this study include those that can be obtained by augmenting the split-valence plus polarization 6-31G(d,p) basis set[9,10] by adding a set of diffuse s and p functions[11,12] on carbon atoms {6-31+G(d,p)}, by splitting the polarization functions into two sets, with exponents doubled and halved relative to the original functions[13] {6-31G(2d,2p) and 6-31+G(2d,2p)}, and by adding second polarization functions on all atoms[13] {6-31G(2df,2pd) and 6-31+G(2df,2pd)}. These basis sets are balanced, in the sense that the numbers of sets of first and second polarization functions are the same on carbon and hydrogen. We have also included the 6-311G basis set with the largest set of polarization functions, both without {6-311G(2df,2pd)} and with {6-311+G(2df,2pd)} diffuse functions on carbon. We have also investigated the effect on the 6-31+G results of adding a diffuse s function on hydrogen, but found that it has no significant energy consequences. The compositions and sizes of the various basis sets are listed in Table I.

The comparative basis sets used in this study are the correlation-consistent polarized split-valence basis sets of Dunning,[14] denoted by cc-pVXZ, where X = D for valence double-split, T for triple-split, and Q for the quadruple-split basis (see Table I). The cc-pVDZ basis set includes one set of d polarization functions on carbon and p functions on hydrogen. The cc-pVTZ basis contains two sets of d functions and a set of f functions on carbon, and two sets of p functions and a set of d functions on hydrogen. The cc-pVQZ basis has three sets of d functions, two sets of f functions, and a set of g functions on carbon, and three sets of p functions, two sets of d functions, and a set of f functions on hydrogen. We have truncated this last basis set by omitting the g functions on carbon and f functions on hydrogen, since the Gaussian series of computer programs[22] used for all of these calculations cannot handle g functions. This truncated quadruple-split basis is denoted by cc-pVQZ′. A recent highly accurate study of the binding energy of N_2 confirmed the high quality of the Dunning cc basis sets in comparison with other types of bases.[23]

The s exponents of the hydrogen cc-pVDZ basis have been scaled by a factor of $1.2^2 = 1.44$. This scaling is necessary because the original s functions of the hydrogen correlation-consistent basis sets had been taken from atom-optimized calculations, unlike the polarization functions, which had been optimized in configuration interaction calculations on H_2. Scaling lowers the cc-pVDZ total energies of the molecules in this study by 0.6–1.4 mE_h per hydrogen atom at the Hartree–Fock level and by 0.9–1.8 mE_h at correlated levels. (Without scaling, the cc-pVDZ energies are often higher than the corresponding 6-31G(d,p) values, particularly for H_2 and the saturated hydrocarbons.) Scaling affects the cc-pVDZ hydrogenolysis energies by less than 0.1 kcal/mol per hydrogen atom involved in the reaction. Several tests of scaling for the cc-pVTZ hydrogen basis have shown effects smaller by about an order of magnitude, and no scaling has been employed in the results presented here for the cc-pVTZ and cc-pVQZ′ bases.

The outermost exponents for the carbon s and p Gaussians in the 6-31G, 6-311G, and the cc-pVXZ basis sets are compared in Table II. Calculations have also been carried out with cc basis sets augmented with diffuse s and p functions on carbon, using an exponent of 0.04, which is

---

(22) Frisch, M. J.; Head-Gordon, M.; Schlegel, H. B.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; DeFrees, D. J.; Fox, D. J.; Whiteside, R. A.; Seeger, R.; Melius, C. F.; Baker, J.; Martin, R.; Kahn, L. R.; Stewart, J. J. P.; Fluder, E. M.; Topiol, S.; Pople, J. A. *Gaussian 88*; Gaussian, Inc.: Pittsburgh, 1988.

(23) Werner, H.-J.; Knowles, P. J. *J. Chem. Phys.* **1991**, *94*, 1264.

Table II. Exponents of Outer s and p Gaussian Functions in Several Carbon Atom Basis Sets[a]

basis set	smallest s exponent	smallest p exponent
6-31G	0.1687145	0.1687145
6-311G	0.145585	0.145585
cc-pVDZ	0.1596	0.1517
cc-pVTZ	0.1285	0.1209
cc-pVQZ	0.1111	0.1007

[a] When diffuse s and p functions are added, these have an exponent of 0.0438 for 6-31+G and 6-311+G and 0.04 for the cc-pVXZ+ basis sets.

Table III. Comparison of Various Estimates for Zero-Point Energies (kcal/mol)

molecule	0.9-scaled HF/6-31G(d)	MP2/6-31+G(d,p)	0.93-scaled HF/6-31G(d)	av of 0.9 × HF and MP2	exptl[a]
H_2	6.0	6.6	6.1	6.3	6.21[b]
CH_3^+	19.1	20.5	19.7	19.8	
CH_4	27.0	29.1	27.9	28.1	28.0[c]
C_2H_2	16.6	16.5	17.2	16.6	16.5[d]
$C_2H_3^+$	20.7	22.0	21.4	21.4	
C_2H_4	30.9	32.7	32.0	31.8	31.6[e]
$C_2H_5^+$	36.7	39.5	37.9	38.1	
C_2H_6	45.0	48.4	46.5	46.7	46.9[f]
cyclopropene	34.2	35.9	35.3	35.1	
C_3H_8	62.4	66.9	64.5	64.7	64.6[g]

[a] From eq 3. [b] Huber, K. P.; Herzberg, G. *Molecular Spectra and Molecular Structure. Volume IV. Constants of Diatomic Molecules*; Van Nostrand Reinhold: New York, 1979. [c] Gray, D. L.; Robiette, A. G. *Mol. Phys.* **1979**, *37*, 1901. [d] Strey, G.; Mills, I. M. *J. Mol. Spectrosc.* **1976**, *59*, 103. [e] Duncan, J. L.; Hamilton, E. *J. Mol. Struct.* **1981**, *76*, 65. [f] Hansen, G. E.; Dennison, D. M. *J. Chem. Phys.* **1952**, *20*, 313. [g] Gayles, J. N., Jr.; King, W. T.; Schachtschneider, J. H. *Spectrochim. Acta Part A* **1967**, *23*, 703.

about one-third the value of the smallest s and p exponents in the cc-pVTZ basis set, and similar to the diffuse functions (exponent 0.0438) used in the 6-31+G and 6-311+G bases. These augmented sets are denoted cc-pVXZ+ (X = D, T, or Q′).

For all basis sets, separate calculations have been carried out for two modes of use of the d polarization functions. One choice, consistent with the original form in which the polarized 6-31G basis sets had been developed,[9,10] and corresponding to the default setting for these bases in the Gaussian program,[22] uses six components for each d set (containing the factors x^2, y^2, z^2, xy, xz, yz, respectively). The other choice, commonly employed with most other bases, including 6-311G, uses five components ($2z^2 - x^2 - y^2$, $x^2 - y^2$, xy, xz, and yz). The six-components choice provides a 3s-type combination ($x^2 + y^2 + z^2$); if the s basis function space is well-covered in the original basis, this additional 3s function should not have a significant impact on the results.

The correlation energies were evaluated using fourth-order many-body Møller–Plesset perturbation theory (MP4).[24-29] For the MP4 calculations, inner-shell electrons on the carbon atoms were frozen in their Hartree–Fock molecular orbitals. We have also computed correlation energies using the method of quadratic configuration interaction (QCISD) and QCISD with noniterative inclusion of triple excitations [QCISD(T)].[30] Those data will be analyzed in detail in another paper,[8] but a few comparisons between MP4 and QCISD(T) energies are also made in the present study.

While internal convergence, primarily by comparison with the results of the very large cc-pVQZ′ basis, is used as the principal measure of basis set quality in this study, computed hydrogenolysis energies will also be compared here with values derived from experimental heats of formation. To facilitate the comparison, the experimental data at 298 K have been back-corrected to "electronic" hydrogenolysis energies at 0 K by subtracting zero-point vibrational and thermal contributions. The zero-point energies were obtained from spectroscopically-derived data whenever these were available, using the contribution

$$\tfrac{3}{8}\omega_e + \tfrac{1}{8}\nu \qquad (3)$$

for each normal mode (or each component of a degenerate mode). In this formula ω_e is the harmonic frequency, taken from a published force-field expression obtained by fitting spectroscopic energy level data for various isotopomers of the molecule of interest, and ν is the corresponding fundamental frequency. Equation 3 is based on the energy level formula

$$E_\upsilon = \omega_e(\upsilon + \tfrac{1}{2}) - \omega_e x_e(\upsilon + \tfrac{1}{2})^2 \qquad (4)$$

which is exact for a Morse oscillator (see, e.g., Miller and Ganda-Kesuma[31]).

When force field data were not available, zero-point energies were computed from ab initio frequencies. It has been common practice to use scaled Hartree–Fock harmonic frequencies, with a scale factor 0.9 or 0.89 for the 6-31G(d) basis, to estimate vibrational frequencies,[32] but this scaling was chosen to reproduce spectroscopic fundamental frequencies. We have found that a scale factor of 0.93 produces good agreement with experimental zero-point energies obtained from eq 3 for the hydrocarbons in Table III. Even better agreement, however, with a maximum zero-point energy error of 0.2 kcal/mol for all cases tested, was obtained using the average of 0.9-scaled HF/6-31G(d) and unscaled MP2/6-31+G(d,p) harmonic frequencies. The relevant data for the molecules included in this study (plus propane) are presented in Table III, including the spectroscopically-derived values where available.

Thermal vibrational energies for all species were evaluated from the 0.9-scaled HF frequencies. Rotational and translational energy changes and the pV work term were evaluated classically. The electronic hydrogenolysis energies obtained from these precedures and the data used in their derivation are shown in Table IV. These hydrogenolysis energies may then be compared directly with the computed ab initio values.

The MP4 and QCISD calculations reported in this study were performed using the Gaussian 88 system of computer programs[22] on the Cray Y-MP8/864 computer at the Ohio Supercomputer Center and on the AMDAHL 5868 computer at Youngstown State University.

Results

Table V reports the computed HF and MP4 total energies for H_2, the neutral hydrocarbons CH_4, C_2H_2, C_2H_4, C_2H_6, and C_3H_4 (cyclopropene), and the carbocations CH_3^+, $C_2H_3^+$, and $C_2H_5^+$ (the latter two in their nonclassical bridged structures). The hydrogenolysis energies computed from these data are reported in Table VI. Table VII reports selected QCISD(T) total energies. Comparisons of hydrogenolysis energies computed with several basis sets with experimentally derived values are shown in Figure 1. For reasons that are made clear in the next section, the computed results displayed in Figure 1 for the 6-31G basis sets include diffuse functions and six components for the d basis functions, while the results for the 6-311G and correlation-consistent basis sets are without diffuse functions and with five d components.

Discussion

Convergence with Increasing Basis Set Size. The type of basis set convergence problems which motivated this study are apparent from the variations of the MP4 total energies (Table V) and hydrogenolysis energies (Table VI) with the various 6-31G basis sets. These problems are particularly evident in Figure 1, which compares the MP4 and QCISD(T) hydrogenolysis energies computed with the 6-31G bases with six d components and with added diffuse functions with those obtained with the 6-311G and cc basis sets with five d components and no diffuse functions. The dependence of the results on the number of components for the d polarization functions and on the presence of diffuse functions will be discussed in detail below. At this point we focus on the convergence within a consistent set of bases, for example, all with

---

(24) Bartlett, R. J.; Silver, D. M. *J. Chem. Phys.* **1975**, *62*, 3258; **1976**, *64*, 1260, 4578.

(25) Binkley, J. S.; Pople, J. A. *Int. J. Quantum Chem.* **1975**, *9*, 229.

(26) Pople, J. A.; Binkley, J. S.; Seeger, R. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1976**, *10*, 1.

(27) Krishnan, R.; Pople, J. A. *Int. J. Quantum Chem.* **1978**, *14*, 91.

(28) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1978**, *68*, 2114.

(29) Bartlett, R. J.; Purvis, G. D. *Int. J. Quantum Chem.* **1978**, *14*, 561.

(30) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968.

(31) Miller, K. J.; Ganda-Kesuma, F. S. *J. Mol. Spectrosc.* **1991**, *145*, 429.

(32) Pople, J. A.; Schlegel, H. B.; Krishnan, R.; DeFrees, D. J.; Binkley, J. S.; Frisch, M. J.; Whiteside, R. J.; Hout, R. F.; Hehre, W. J. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1981**, *15*, 269. Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986. Pople, J. A.; Head-Gordon, M.; Fox, D. J.; Raghavachari, K.; Curtiss, L. A. *J. Chem. Phys.* **1989**, *90*, 5622.

Table IV. Experimentally Derived Electronic Hydrogenolysis Energies (kcal/mol)

molecule	enthalpy of formation $\Delta H^\circ_{f,298}$	enthalpy of hydrogenolysis[a] ΔH°_{298}	zero-point vibrational contribution[b]	thermal contribution + pV term[c]	electronic hydrogenolysis energy
H^+	365.7[d]				
H_2	0.0				
CH_4	−17.8[e]				
CH_3^+	261.3[d]	86.6	2.0	−0.6	85.2
C_2H_2	54.3[e]	−89.9	20.9	−3.7	−107.1
$C_2H_3^+$	265.9[d]	64.2	16.0	−2.8	51.0
C_2H_4	12.5[e]	−48.1	12.0	−1.9	−58.2
$C_2H_5^+$	215.6[d]	114.5	5.5	−0.6	109.6
C_2H_6	−20.1[e]	−15.5	2.9	−0.1	−18.3
cyclopropene	66.2[e]	−119.6	24.1	−3.8	−139.9

[a] See eqs 1 and 2. [b] Based on the data of Table III; experimental values used where available, average of 0.9 × HF and MP2 otherwise. [c] Vibrational contributions based on 0.9-scaled HF frequencies (Table III); classical rotational and translational contributions. [d] Lias, S. G.; Bartmess, J. E.; Liebman, J. F.; Holmes, J. L.; Levin, R. D.; Mallard, W. G. In Gas-Phase Ion and Neutral Thermochemistry; *J. Phys. Chem. Ref. Data*, **1988**, *17*, Suppl. No. 1. [e] Pedley, J. B.; Naylor, R. D.; Kirby, S. P. Thermochemical Data for Organic Compounds; 2nd ed.; Chapman and Hall: London, 1986.

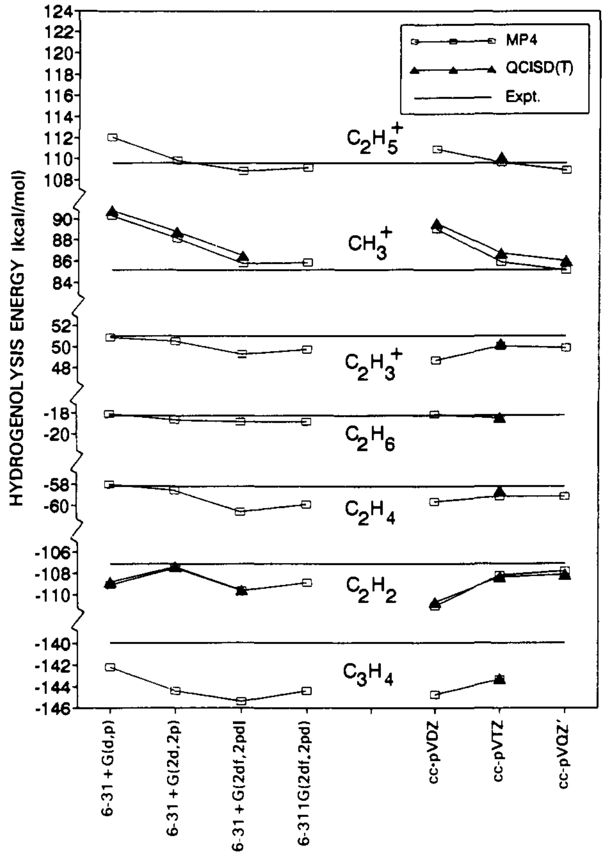

Figure 1. Comparison of computed and experimental hydrogenolysis energies. In this comparison, the 6-31G basis sets include diffuse functions and six components for d basis functions, while the 6-311G and cc basis sets are without diffuse functions and with five d components. The proper comparison for the largest of the 6-31+G bases and for 6-311G is with the comparably sized cc-pVTZ basis. The cc-pVQZ′ results provide an indication of the probable converged limit for the corresponding correlation model. The uncertainty in the experimentally derived values (from Table IV) is probably about 1 kcal/mol, but may be higher for cyclopropene.

five d components and without diffuse functions. We emphasize again that the appropriate comparison for the largest of the 6-31G basis sets and for 6-311G is with the similarly sized cc-pVTZ basis, and that the much larger cc-pVQZ′ basis is used here primarily to provide an indication of the probable convergence limit for each type of wave function (HF, MP4, or QCISD(T)).

We first consider the convergence of the HF and MP4 total energies. As expected, the choice of basis has a much greater effect at the MP4 level than at the Hartree–Fock level. Comparing the smallest bases in the two series of basis sets, we find that the energies obtained with the cc-pVDZ basis are almost always lower than the corresponding 6-31G(d,p) values. (The exceptions are minor differences in the HF energy of H_2 and the MP4 energy of $C_2H_3^+$.) Interestingly, the amount of correlation energy recovered at the MP4 level with the cc-pVDZ basis is less than with 6-31G(d,p) for the three cations and for acetylene and cyclopropene.

Except for the case of H_2, a much greater MP4 energy change for the 6-31G bases occurs in the second basis extension step, going from (2d,2p) to (2df,2pd), than in the first, from (d,p) to (2d,2p). Very significant differences are found between the total energies obtained with the 6-31G(2df,2pd) and cc-pVTZ bases, always in favor of the latter. Using the cc-pVQZ′ results as nearly converged benchmark values in order to gauge the convergence achieved with each of the other two bases, we find that the differences between the 6-31G(2df,2pd) total energies and the corresponding cc-pVQZ′ values are about 3–5 times greater at the Hartree–Fock level, and 2–3 times greater at the MP4 level, than the differences between the cc-pVTZ and cc-pVQZ′ results. This difference persists even if we compare the more comparably sized 6-31+G(2df,2pd) and cc-pVTZ bases; for the hydrocarbons, the 6-31+G(2df,2pd) energies are higher than the corresponding cc-pVTZ values by 8–25 mE_h at the HF level and by 10–40 mE_h at MP4. Although the 6-311G(2df,2pd) basis is only marginally larger (by one hydrogen s function, see Table I) than 6-31+G(2df,2pd) and equal in size to cc-pVTZ, it produces energies which are significantly lower than 6-31+G(2df,2pd), but higher than cc-pVTZ hydrocarbon energies by 0.5–3.6 mE_h at the HF level and by 1–5 mE_h at the MP4 level.

Turning next to the hydrogenolysis energies, and ignoring variations of 0.2 kcal/mol or less, the variations within the 6-31G and cc series of basis sets are mostly monotonic, with the notable exception of the 6-31G series MP4 results for acetylene when six d components or diffuse functions are used and for cyclopropene with five d components. For acetylene, protonated acetylene, and ethylene, the changes in the computed MP4 hydrogenolysis energies with extension of the 6-31G polarization space are substantially greater in the second extension step than in the first when six d components are used. The 6-311G(2df,2pd) basis usually, but not always, continues the trend established in the 6-31G series, the major exceptions being the MP4 hydrogenolysis energies of acetylene and ethylene and the six d component results for cyclopropene. For the cc basis sets, essentially all changes are much smaller in the second step, from TZ to QZ′, than in the first, but it should be remembered that the first extension step in the cc series is equivalent to the entire combined change, from 6-31G(d,p) to 6-311G(2df,2pd), in the G series. Examining the differences of the MP4 hydrogenolysis energies with the various 6-31G basis sets from the nearly converged cc-pVQZ′ values, the closest results for acetylene and ethylene are found not with the largest (2df,2pd) polarization set but with (2d,2p) or even (for ethylene without diffuse functions) with (d,p). Significantly better agreement is generally found for 6-311G(2df,2pd), and particularly for 6-

Table V. Computed Total Energies (E_h)

	five d components		six d components			five d components		six d components	
basis set	HF	MP4	HF	MP4	basis set	HF	MP4	HF	MP4
Molecule: H_2									
6-31G(d,p)	−1.131 329	−1.164 537	−1.131 329	−1.164 537	cc-pVDZ	−1.131 042	−1.166 025	−1.131 042	−1.166 025
6-31+G(d,p)	−1.131 329	−1.164 537	−1.131 329	−1.164 537	cc-pVDZ+	−1.131 042	−1.166 025	−1.131 042	−1.166 025
6-31G(2d,2p)	−1.131 951	−1.168 594	−1.131 951	−1.168 594	cc-pVTZ	−1.132 975	−1.171 643	−1.132 997	−1.171 761
6-31+G(2d,2p)	−1.131 951	−1.168 954	−1.131 951	−1.168 594	cc-pVTZ+	−1.132 975	−1.171 643	−1.132 997	−1.171 761
6-31G(2df,2pd)	−1.131 975	−1.170 168	−1.131 987	−1.170 419	cc-pVQZ′	−1.133 484	−1.172 868	−1.133 494	−1.172 922
6-31+G(2df,2pd)	−1.131 975	−1.170 168	−1.131 987	−1.170 419	cc-pVQZ′+	−1.133 484	−1.172 868	−1.133 494	−1.172 922
6-311G(2df,2pd)	−1.132 991	−1.171 617	−1.133 004	−1.171 721					
6-311+G(2df,2pd)	−1.132 991	−1.171 617	−1.133 004	−1.171 721					
Molecule: CH_3^+									
6-31G(d,p)	−39.236 261	−39.367 330	−39.236 291	−39.368 750	cc-pVDZ	−39.237 414	−39.367 749	−39.237 587	−39.369 348
6-31+G(d,p)	−39.236 670	−39.367 915	−39.236 702	−39.369 331	cc-pVDZ+	−39.238 119	−39.368 564	−39.238 222	−39.370 101
6-31G(2d,2p)	−39.237 638	−39.374 335	−39.238 644	−39.376 704	cc-pVTZ	−39.247 751	−39.402 841	−39.247 805	−39.403 246
6-31+G(2d,2p)	−39.237 964	−39.374 887	−39.238 898	−39.377 219	cc-pVTZ+	−39.247 805	−39.402 928	−39.247 863	−39.403 319
6-31G(2df,2pd)	−39.238 146	−39.388 611	−39.239 162	−39.392 461	cc-pVQZ′	−39.250 187	−39.409 909	−39.250 251	−39.410 261
6-31+G(2df,2pd)	−39.238 466	−39.389 123	−39.239 398	−39.392 788	cc-pVQZ′+	−39.250 199	−39.409 929	−39.250 271	−39.410 288
6-311G(2df,2pd)	−39.246 919	−39.401 646	−39.247 105	−39.402 189					
6-311+G(2df,2pd)	−39.246 932	−39.401 698	−39.247 125	−39.402 235					
Molecule: CH_4									
6-31G(d,p)	−40.201 678	−40.387 396	−40.201 704	−40.388 610	cc-pVDZ	−40.203 262	−40.391 933	−40.203 326	−40.393 204
6-31+G(d,p)	−40.202 104	−40.388 805	−40.202 131	−40.390 034	cc-pVDZ+	−40.203 639	−40.393 065	−40.203 701	−40.394 357
6-31G(2d,2p)	−40.202 498	−40.401 436	−40.203 583	−40.403 762	cc-pVTZ	−40.213 460	−40.437 533	−40.213 497	−40.438 152
6-31+G(2d,2p)	−40.203 128	−40.403 048	−40.204 204	−40.405 392	cc-pVTZ+	−40.213 463	−40.437 613	−40.213 499	−40.438 191
6-31G(2df,2pd)	−40.203 203	−40.420 530	−40.204 493	−40.425 363	cc-pVQZ′	−40.216 229	−40.447 105	−40.216 287	−40.447 543
6-31+G(2df,2pd)	−40.203 821	−40.422 002	−40.204 998	−40.426 468	cc-pVQZ′	−40.216 231	−40.447 118	−40.216 291	−40.447 558
6-311G(2df,2pd)	−40.212 795	−40.436 409	−40.212 939	−40.437 093					
6-311+G(2df,2pd)	−40.212 819	−40.436 578	−40.212 955	−40.437 203					
Molecule: C_2H_2									
6-31G(d,p)	−76.821 700	−77.103 897	−76.821 837	−77.106 707	cc-pVDZ	−76.828 404	−77.108 715	−76.828 792	−77.111 930
6-31+G(d,p)	−76.827 075	−77.109 853	−76.827 195	−77.112 669	cc-pVDZ+	−76.830 223	−77.111 232	−76.830 613	−77.114 534
6-31G(2d,2p)	−76.822 383	−77.121 177	−76.825 192	−77.127 319	cc-pVTZ	−76.850 546	−77.187 722	−76.850 667	−77.188 292
6-31+G(2d,2p)	−76.828 472	−77.127 468	−76.831 235	−77.133 716	cc-pVTZ+	−76.850 925	−77.188 176	−76.851 040	−77.188 730
6-31G(2df,2pd)	−76.827 635	−77.154 445	−76.831 003	−77.161 487	cc-pVQZ′	−76.855 555	−77.203 788	−76.855 606	−77.204 294
6-31+G(2df,2pd)	−76.833 127	−77.159 770	−76.836 438	−77.166 971	cc-pVQZ′+	−76.855 650	−77.203 908	−76.855 700	−77.204 401
6-311G(2df,2pd)	−76.847 999	−77.184 426	−76.848 313	−77.185 862					
6-311+G(2df,2pd)	−76.849 405	−77.185 809	−76.849 692	−77.187 239					
Molecule: $C_2H_3^+$									
6-31G(d,p)	−77.084 311	−77.363 442	−77.084 425	−77.366 247	cc-pVDZ	−77.088 002	−77.363 436	−77.088 479	−77.366 698
6-31+G(d,p)	−77.084 998	−77.364 725	−77.085 084	−77.367 514	cc-pVDZ+	−77.088 637	−77.364 581	−77.088 973	−77.367 758
6-31G(2d,2p)	−77.088 096	−77.378 613	−77.090 513	−77.384 282	cc-pVTZ	−77.111 020	−77.439 831	−77.111 167	−77.440 584
6-31+G(2d,2p)	−77.088 779	−77.379 839	−77.091 035	−77.385 464	cc-pVTZ+	−77.111 125	−77.440 049	−77.111 268	−77.440 769
6-31G(2df,2pd)	−77.093 127	−77.412 202	−77.096 164	−77.419 298	cc-pVQZ′	77.116 016	−77.455 013	−77.116 090	−77.455 577
6-31+G(2df,2pd)	−77.093 717	−77.413 121	−77.096 593	−77.420 199	cc-pVQZ′+	−77.116 050	−77.455 083	−77.116 129	−77.455 644
6-311G(2df,2pd)	−77.109 106	−77.437 151	−77.109 588	−77.438 763					
6-311+G(2df,2pd)	−77.109 195	−77.437 403	−77.109 623	−77.438 935					
Molecule: C_2H_4									
6-31G(d,p)	−78.038 731	−78.350 159	−78.038 839	−78.353 060	cc-pVDZ	−78.044 647	−78.356 796	−78.044 834	−78.360 050
6-31+G(d,p)	−78.042 945	−78.355 705	−78.043 043	−78.358 573	cc-pVDZ+	−78.046 569	−78.359 870	−78.046 740	−78.363 043
6-31G(2d,2p)	−78.039 412	−78.369 491	−78.041 708	−78.374 840	cc-pVTZ	−78.064 393	−78.437 545	−78.064 492	−78.438 427
6-31+G(2d,2p)	−78.044 198	−78.374 952	−78.046 441	−78.380 253	cc-pVTZ+	−78.064 860	−78.438 131	−78.064 958	−78.438 939
6-31G(2df,2pd)	−78.041 914	−78.403 215	−78.044 441	−78.410 727	cc-pVQZ′	−78.069 547	−78.454 263	−78.069 639	−78.454 953
6-31+G(2df,2pd)	−78.046 530	−78.408 207	−78.048 985	−78.415 532	cc-pVQZ′+	−78.069 692	−78.454 418	−78.069 785	−78.455 101
6-311G(2df,2pd)	−78.061 649	−78.434 090	−78.062 077	−78.435 461					
6-311+G(2df,2pd)	−78.063 148	−78.435 590	−78.063 528	−78.436 827					
Molecule: $C_2H_5^+$									
6-31G(d,p)	−78.320 881	−78.625 751	−78.320 938	−78.628 571	cc-pVDZ	−78.324 126	−78.628 527	−78.324 241	−78.631 674
6-31+G(d,p)	−78.321 158	−78.626 726	−78.321 214	−78.629 543	cc-pVDZ+	−78.324 747	−78.629 606	−78.324 913	−78.632 706
6-31G(2d,2p)	−78.323 712	−78.642 806	−78.325 676	−78.647 813	cc-pVTZ	−78.344 431	−78.706 471	−78.344 547	−78.707 376
6-31+G(2d,2p)	−78.323 981	−78.643 696	−78.325 875	−78.648 646	cc-pVTZ+	−78.344 547	−78.706 667	−78.344 651	−78.707 524
6-31G(2df,2pd)	−78.326 056	−78.677 497	−78.328 354	−78.685 032	cc-pVQZ′	−78.349 399	−78.721 957		
6-31+G(2df,2pd)	−78.326 294	−78.678 240	−78.328 516	−78.685 542					
6-311G(2df,2pd)	−78.342 317	−78.703 537	−78.342 950	−78.705 012					
6-311+G(2df,2pd)	−78.342 465	−78.703 796	−78.343 004	−78.705 155					
Molecule: C_2H_6									
6-31G(d,p)	−79.238 174	−79.581 603	−79.238 234	−79.584 023	cc-pVDZ	−79.241 263	−79.588 830	−79.241 476	−79.591 496
6-31+G(d,p)	−79.238 701	−79.584 204	−79.238 761	−79.586 643	cc-pVDZ+	−79.241 861	−79.590 949	−79.242 058	−79.593 637
6-31G(2d,2p)	−79.239 450	−79.605 068	−79.241 443	−79.609 726	cc-pVTZ	−79.260 019	−79.673 780	−79.260 076	−79.674 885
6-31+G(2d,2p)	−79.240 322	−79.607 735	−79.242 287	−79.612 395	cc-pVTZ+	−79.260 032	−79.673 953	−79.260 090	−79.674 989
6-31G(2df,2pd)	−79.240 947	−79.641 774	−79.243 236	−79.650 390					
6-31+G(2df,2pd)	−79.241 805	−79.644 240	−79.243 980	−79.652 384					
6-311G(2df,2pd)	−79.258 279	−79.671 054	−79.258 736	−79.672 506					
6-311+G(2df,2pd)	−79.258 332	−79.671 421	−79.258 775	−79.672 772					
Molecule: C_3H_4 (Cyclopropene)									
6-31G(d,p)	−115.830 395	−116.275 279	−115.830 526	−116.279 264	cc-pVDZ	−115.837 476	−116.281 030	−115.837 943	−116.285 628
6-31+G(d,p)	−115.833 881	−116.281 414	−115.833 992	−116.285 373	cc-pVDZ+	−115.838 920	−116.284 479	−115.839 289	−116.288 954
6-31G(2d,2p)	−115.831 299	−116.298 464	−115.834 304	−116.305 943	cc-pVTZ	−115.866 144	−116.397 684	−115.866 271	−116.398 913
6-31+G(2d,2p)	−115.835 194	−116.304 225	−115.838 141	−116.311 701	cc-pVTZ+	−115.866 444	−116.398 277	−115.866 566	−116.399 425
6-31G(2df,2pd)	−115.837 350	−116.351 326	−115.840 760	−116.361 093					
6-31+G(2df,2pd)	−115.840 951	−116.356 364	−115.844 278	−116.365 974					
6-311G(2df,2pd)	−115.862 555	−116.392 608	−115.863 109	−116.394 754					
6-311+G(2df,2pd)	−115.863 665	−116.394 208	−115.864 135	−116.396 158					

Table VI. Hydrogenolysis Energies (kcal/mol)

	five d components		six d components			five d components		six d components	
basis set	HF	MP4	HF	MP4	basis set	HF	MP4	HF	MP4
$CH_3^+ + H_2 \rightarrow CH_4 + H^+$									
6-31G(d,p)	104.1	90.7	104.1	90.8	cc-pVDZ	103.7	89.0	103.7	89.2
6-31+G(d,p)	104.1	90.1	104.1	90.3	cc-pVDZ+	103.9	88.8	103.9	89.0
6-31G(2d,2p)	104.9	88.8	104.8	88.8	cc-pVTZ	105.0	85.9	105.0	85.9
6-31+G(2d,2p)	104.7	88.1	104.6	88.1	cc-pVTZ+	105.0	85.9	105.0	85.9
6-31G(2df,2pd)	104.7	86.8	104.6	86.3	cc-pVQZ′	105.1	85.1	105.1	85.1
6-31+G(2df,2pd)	104.6	86.2	104.4	85.8	cc-pVQZ′+	105.1	85.1	105.1	85.1
6-311G(2df,2pd)	104.9	85.9	104.9	85.9					
6-311+G(2df,2pd)	104.9	85.8	104.9	85.8					
$C_2H_2 + 3H_2 \rightarrow 2CH_4$									
6-31G(d,p)	−117.8	−111.2	−117.7	−111.0	cc-pVDZ	−116.1	−111.1	−115.9	−110.7
6-31+G(d,p)	−114.9	−109.3	−114.9	−109.1	cc-pVDZ+	−115.4	−111.0	−115.3	−110.5
6-31G(2d,2p)	−117.2	−110.4	−116.8	−109.5	cc-pVTZ	−111.4	−108.2	−111.3	−108.4
6-31+G(2d,2p)	−114.2	−108.5	−113.8	−107.5	cc-pVTZ+	−111.1	−108.0	−111.0	−108.2
6-31G(2df,2pd)	−114.7	−110.5	−114.2	−111.7	cc-pVQZ′	−110.7	−107.8	−110.7	−107.9
6-31+G(2df,2pd)	−112.1	−109.0	−111.4	−109.6	cc-pVQZ′+	−110.7	−107.8	−110.7	−107.9
6-311G(2df,2pd)	−112.1	−108.9	−112.0	−108.7					
6-311+G(2df,2pd)	−111.2	−108.2	−111.2	−107.9					
$C_2H_3^+ + 3H_2 \rightarrow 2CH_4 + H^+$									
6-31G(d,p)	47.0	51.6	47.1	51.9	cc-pVDZ	46.8	48.7	47.0	49.2
6-31+G(d,p)	46.9	50.7	46.9	50.9	cc-pVDZ+	46.7	48.0	46.9	48.4
6-31G(2d,2p)	49.5	51.2	49.7	51.8	cc-pVTZ	52.1	50.0	52.2	49.9
6-31+G(2d,2p)	49.2	49.9	49.2	50.5	cc-pVTZ+	52.2	50.0	52.2	50.0
6-31G(2df,2pd)	51.9	51.2	52.2	50.1	cc-pVQZ′	52.7	49.8	52.7	49.7
6-31+G(2df,2pd)	51.5	50.0	51.8	49.3	cc-pVQZ′+	52.7	49.9	52.7	49.8
6-311G(2df,2pd)	51.8	49.7	51.9	50.0					
6-311+G(2df,2pd)	51.8	49.6	51.9	50.0					
$C_2H_4 + 2H_2 \rightarrow 2CH_4$									
6-31G(d,p)	−64.0	−60.0	−64.0	−59.7	cc-pVDZ	−62.6	−59.6	−62.6	−59.2
6-31+G(d,p)	−61.9	−58.3	−61.8	−58.0	cc-pVDZ+	−61.9	−59.1	−61.9	−58.7
6-31G(2d,2p)	−63.8	−60.4	−63.7	−59.9	cc-pVTZ	−60.6	−59.1	−60.6	−59.2
6-31+G(2d,2p)	−61.6	−59.0	−61.5	−58.6	cc-pVTZ+	−60.3	−58.9	−60.3	−58.9
6-31G(2df,2pd)	−63.1	−61.2	−63.1	−62.2	cc-pVQZ′	−60.2	−59.1	−60.2	−59.2
6-31+G(2df,2pd)	−61.0	−59.9	−60.9	−60.6	cc-pVQZ′+	−60.1	−59.0	−60.1	−59.1
6-311G(2df,2pd)	−61.5	−59.9	−61.4	−59.8					
6-311+G(2df,2pd)	−60.6	−59.2	−60.5	−59.1					
$C_2H_5^+ + 2H_2 \rightarrow 2CH_4 + H^+$									
6-31G(d,p)	113.1	113.0	113.1	113.2	cc-pVDZ	112.8	110.9	112.7	111.3
6-31+G(d,p)	112.7	111.8	112.7	112.0	cc-pVDZ+	112.7	110.1	112.7	110.5
6-31G(2d,2p)	114.6	111.1	114.5	111.4	cc-pVTZ	115.1	109.6	115.2	109.6
6-31+G(2d,2p)	114.0	109.7	113.8	109.8	cc-pVTZ+	115.2	109.6	115.2	109.6
6-31G(2df,2pd)	115.2	110.9	115.0	109.9	cc-pVQZ′	115.4	108.9		
6-31+G(2df,2pd)	114.6	109.5	114.5	108.8					
6-311G(2df,2pd)	114.7	109.2	114.9	109.4					
6-311+G(2df,2pd)	114.7	109.1	114.9	109.3					
$C_2H_6 + H_2 \rightarrow 2CH_4$									
6-31G(d,p)	−21.2	−18.0	−21.2	−18.0	cc-pVDZ	−21.5	−18.2	−21.4	−18.1
6-31+G(d,p)	−21.4	−18.1	−21.4	−18.1	cc-pVDZ+	−21.6	−18.3	−21.5	−18.2
6-31G(2d,2p)	−21.1	−18.3	−21.2	−18.3	cc-pVTZ	−21.3	−18.6	−21.3	−18.6
6-31+G(2d,2p)	−21.3	−18.7	−21.4	−18.7	cc-pVTZ+	−21.3	−18.6	−21.3	−18.6
6-31G(2df,2pd)	−21.0	−18.3	−21.2	−18.8					
6-31+G(2df,2pd)	−21.2	−18.6	−21.4	−18.9					
6-311G(2df,2pd)	−21.5	−18.9	−21.4	−18.8					
6-311+G(2df,2pd)	−21.5	−18.9	−21.4	−18.8					
$C_3H_4 + 4H_2 \rightarrow 3CH_4$ (Cyclopropene)									
6-31G(d,p)	−156.5	−143.5	−156.4	−143.3	cc-pVDZ	−155.7	−144.7	−155.5	−144.3
6-31+G(d,p)	−155.1	−142.4	−155.0	−142.2	cc-pVDZ+	−155.5	−144.7	−155.4	−144.3
6-31G(2d,2p)	−155.9	−145.2	−156.0	−144.9	cc-pVTZ	−152.1	−143.3	−152.0	−143.4
6-31+G(2d,2p)	−154.6	−144.7	−154.8	−144.4	cc-pVTZ+	−151.9	−143.1	−151.8	−143.1
6-31G(2df,2pd)	−153.3	−144.1	−153.6	−146.4					
6-31+G(2df,2pd)	−152.2	−143.7	−152.3	−145.4					
6-311G(2df,2pd)	−153.0	−144.4	−152.9	−144.1					
6-311+G(2df,2pd)	−152.4	−143.7	−152.3	−143.4					

311+G(2df,2pd), with results mostly comparable to those of cc-pVTZ. An illustration of the convergence behavior for acetylene is shown graphically in Figure 2.

Five versus Six d Components. Examination of the data in Tables V and VI shows a significant difference between the 6-31G, 6-311G, and cc series of basis sets in the effect of the number of components used for the d polarization functions. This behavior is illustrated graphically for the case of acetylene in Figure 3.

The effect on the total energies of the hydrocarbons is much less at the Hartree–Fock level than at MP4. In both cases the contribution of the sixth d component to the total energy tends to be somewhat greater in magnitude for the smallest of the cc bases, cc-pVDZ (0.06–0.48 mE_h at the HF level and 1.3–4.6 mE_h at MP4), than for the smallest of the 6-31G bases, 6-31G(d,p) (0.03–0.14 mE_h at HF and 1.2–4.0 mE_h at MP4). However, this contribution increases substantially, to as much as 3.4 mE_h at the

Table VII. Selected QCISD(T) Total Energies[a]

basis set	H_2	CH_3^+	CH_4	C_2H_2
6-31+G(d,p)	−1.165 107	−39.370 536	−40.390 953	−77.113 159
6-31+G(2d,2p)	−1.169 156	−39.378 564	−40.406 216	−77.133 882
6-31+G(2df,2pd)	−1.171 007	−39.394 085	−40.427 114	−77.166 536
cc-pVDZ	−1.166 590	−39.369 043	−40.392 894	−77.109 576
cc-pVTZ	−1.172 228	−39.404 141	−40.438 112	−77.186 823
cc-pVQZ′	−1.173 463	−39.411 204	−40.447 620	−77.202 538

basis set	$C_2H_3^+$	C_2H_4	$C_2H_5^+$	C_2H_6	C_3H_4[b]
cc-pVTZ	−77.439 677	−78.438 402	−78.707 268	−79.674 470	−116.397 125

[a] Total energies computed with six components for the d functions for the 6-31+G basis sets and five for the cc basis sets. [b] Cyclopropene.

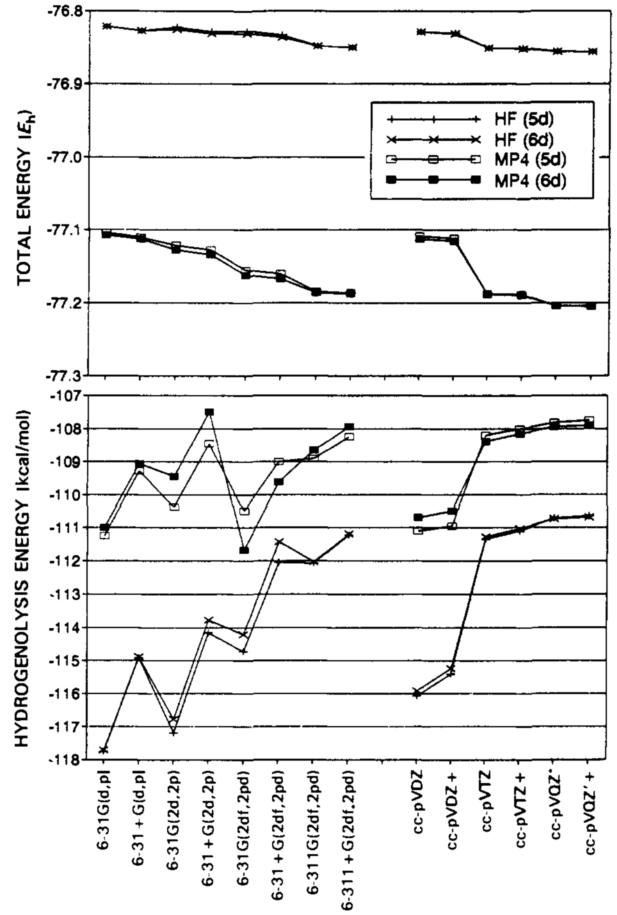

Figure 2. Total and hydrogenolysis energies of acetylene computed with different basis sets. Shown are SCF and MP4 energies computed with both five (5d) and six (6d) components for the d polarization functions.

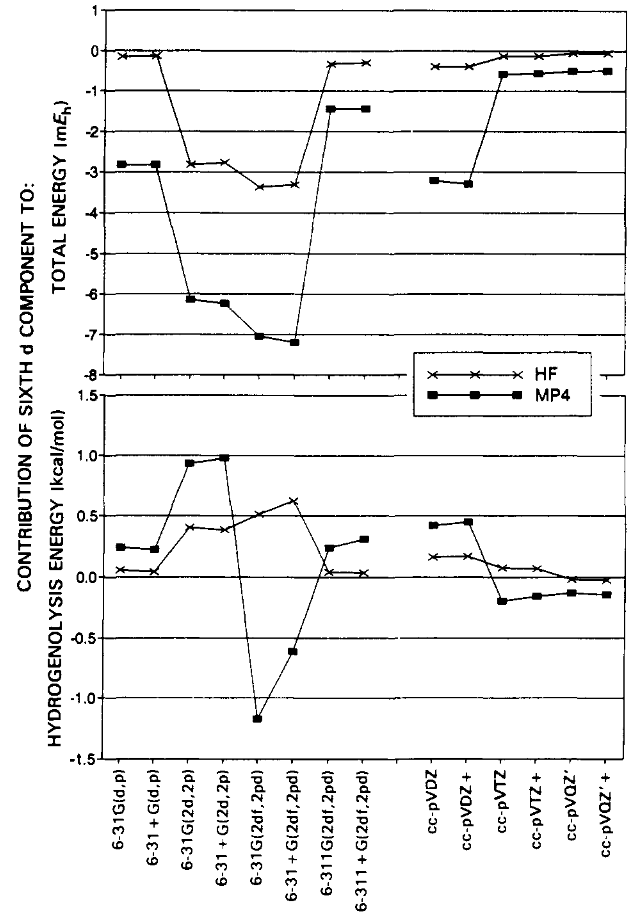

Figure 3. Contributions of the sixth d component to the total and hydrogenolysis energies of acetylene computed with different basis sets. Shown are the contributions to the SCF and MP4 energies.

HF level and 9.8 mE_h at MP4, with the increase of the polarization space from (d,p) to (2d,2p) and to (2df,2pd) for the 6-31G bases. On the other hand, the magnitude of the sixth d component contribution decreases strongly with increase in basis size for the cc series, to no more than 0.15 mE_h (HF) and 1.2 mE_h (MP4) for cc-pVTZ and no more than 0.09 mE_h (HF) and 0.7 mE_h (MP4) for any of the cc-pVQZ′ results. In all likelihood, this difference in behavior between the two series of basis sets reflects the fact that the sp space is systematically increased when more polarization functions are added in the cc series, but not in the 6-31G series, in which case the sixth d component (an s-type function) compensates for some of the sp deficiencies. Indeed, the magnitude of the sixth d component contribution is much smaller for 6-311G(2df,2pd) (0.1–0.6 mE_h at HF and 0.5–2.1 mE_h at MP4) than for 6-31G(2df,2pd), though it is still greater than for cc-pVTZ.

The hydrogenolysis energies display similar trends for the effects of the six d components. At the Hartree–Fock level these effects are negligible for 6-31G(d,p), 6-311G(2df,2pd), and all the cc bases, and reach at most 0.6 kcal/mol in magnitude for the larger 6-31G bases. At MP4 the largest effects for the 6-31G(d,p), 6-311G(2df,2pd), and cc-pVDZ bases are 0.3, 0.4, and 0.5 kcal/mol, respectively, but increase to as much as 2.3 kcal/mol for the larger 6-31G bases, while decreasing to 0.2 kcal/mol or less for the larger cc bases. It is also noteworthy that in the cases of acetylene and ethylene the MP4 hydrogenolysis energies computed with the polarized 6-31G basis sets show poorer convergence behavior when six d components are used rather than five (see Figure 2).

Effect of Diffuse Functions. The addition of diffuse s and p functions on the carbon atoms has a substantial effect on the Hartree–Fock total energies, particularly for the unsaturated hydrocarbons, for all the 6-31+G basis sets, and for cc-pVDZ+, and it also has a significant effect on the MP4 correlation energies for these basis sets. For the unsaturated neutral hydrocarbons, the magnitude of the diffuse functions contribution is greater for the smallest of the 6-31+G basis sets (3.5–5.4 mE_h at the HF level and 5.5–6.1 mE_h at MP4) than for the corresponding cc basis, cc-pVDZ+ (1.4–1.9 mE_h at HF and 2.5–3.4 mE_h at MP4). The effects are smaller, and comparable for the two bases, for the

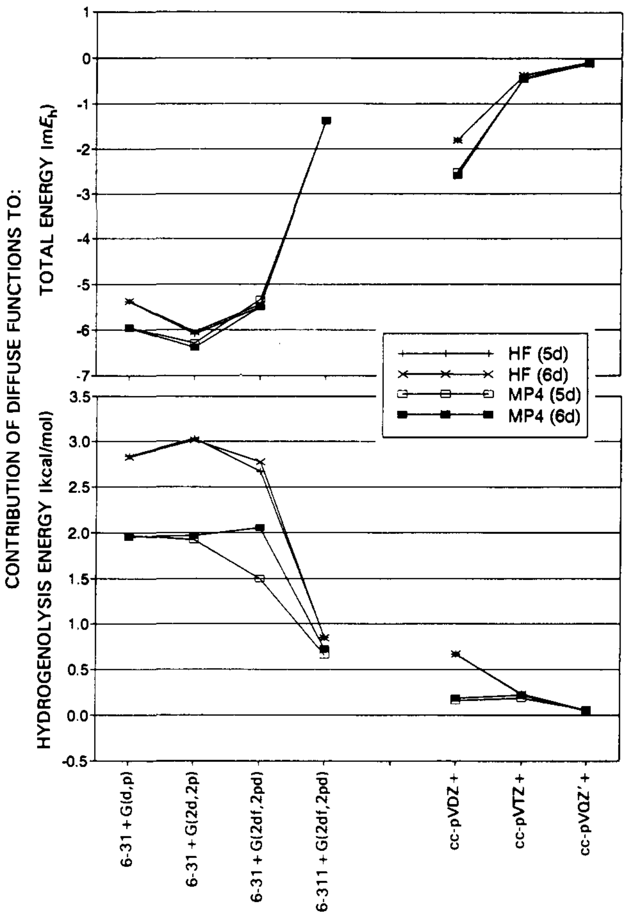

Figure 4. Contributions of diffuse basis functions to the total and hydrogenolysis energies of acetylene computed with different basis sets. Shown are the contributions to the SCF and MP4 energies computed with both five (5d) and six (6d) components for the d polarization functions.

saturated hydrocarbons and the cations. As the 6-31+G(d,p) basis is augmented with additional polarization functions, the contribution of diffuse functions to the HF and MP4 total energies does not vary much, but is substantially reduced when the sp part of the basis is enhanced to 6-311+G. However, for acetylene, ethylene, and cyclopropene, the diffuse functions' contribution to the total energy for the 6-311+G(2df,2pd) basis is still as large as 1.0–1.6 mE_h, compared to a maximum contribution of 0.6 mE_h for cc-pVTZ+. The magnitude of this contribution is decreased substantially with increase in basis size in the cc series of basis sets, to at most 0.5 mE_h (HF) and 0.6 mE_h (MP4) for cc-pVTZ+ and at most 0.15 mE_h (HF and MP4) for cc-pVQZ′+. These effects are illustrated for the case of acetylene in Figure 4.

The effect of diffuse functions on the hydrogenolysis energies is far from negligible for all the 6-31+G basis sets, and is particularly large for the unsaturated molecules, amounting to 2–3 kcal/mol for C_2H_2 and C_2H_4. In contrast, the effect of diffuse functions is never more than 0.7 kcal/mol for 6-311+G(2df,2pd) and 0.8 kcal/mol for cc-pVDZ+, and decreases to at most 0.3 kcal/mol for cc-pVTZ+ and 0.1 kcal/mol for cc-pVQZ′+. For the neutral hydrocarbons, most of the diffuse functions' contribution to the hydrogenolysis energies occurs at the Hartree–Fock level, with relatively small effects on the correlation energy contribution. The opposite is true for the hydrogenolysis energies of the carbocations. In most cases the MP4 hydrogenolysis energies computed with the 6-31+G(2df,2pd) basis set (particularly with five d components) and with 6-311+G(2df,2pd) are better converged, relative to the best available cc basis results, than the corresponding values without diffuse functions. Thus, the diffuse functions are important for all of the 6-31G basis sets and even, to a much lesser extent, for 6-311G, in order to bring their hydrogenolysis energies into better agreement with converged theoretical values or with experiment. Hydrogenolysis energies computed with the cc-pVDZ basis are also improved by added diffuse functions, but less so than for the 6-31G bases. The reasons for these effects may be traced to the magnitudes of the exponents for the outermost Gaussians in the different basis sets (Table II).

The need for diffuse functions in the 6-31G and (to a lesser extent) 6-311G basis sets carries with it some practical disadvantages because of problems associated with slower SCF convergence and difficulties with CI and QCI convergence in some cases.[8]

Comparison with Experimental Hydrogenolysis Energies. As noted earlier, the polarized 6-31G basis sets, unlike the 6-311G and cc bases, were developed and are commonly used with six components for the d polarization functions. Because of this, and based on the analysis of the effects of the diffuse functions, comparisons of the hydrogenolysis energies computed with the 6-31G basis sets with experimental values will be confined to the bases using six d components and including diffuse basis functions, while comparisons for 6-311G and the correlation-consistent basis sets will focus on the cases of five d components and no diffuse functions.

The computed MP4 hydrogenolysis energies are compared with the experimentally derived values (taken from Table IV) in Figure 1. As noted previously, such comparisons are subject to some uncertainties in the experimental heats of formation and zero-point energies, and the basis set effects may be obscured by lack of convergence of the MP4 model. In order to provide an indication of the effect of the limitations of the MP4 model, we have included in Figure 1 some results obtained with the infinite-order QCISD(T) model (Table VII). The differences between the MP4 and QCISD(T) hydrogenolysis energies are all less than 1 kcal/mol, and often less than 0.5 kcal/mol, for the cases studied, indicating that convergence of the MP series for the hydrogenolysis energies has largely been achieved at MP4.

Where significant differences exist among the computed hydrogenolysis energies, cc-pVTZ and cc-pVQZ′ generally give the best agreement with experimental values. The cc-pVTZ and cc-pVQZ′ MP4 and QCISD(T) hydrogenolysis energies reproduce the experimental values to within about 1.5 kcal/mol, close to the probable experimental uncertainty, except for cyclopropene. While the errors of the theoretical models would generally be expected to increase with molecular size, the 3.5 kcal/mol difference from experiment for cyclopropene at the QCISD(T)/cc-pVTZ level appears excessive and, based on the results for the other molecules, is unlikely to change much at the cc-pVQZ′ level. This difference may well be due to a greater experimental error in this case.

Although good agreement with the experimental values is also usually obtained with the 6-311G(2df,2pd) basis, the hydrogenolysis energies of the neutral unsaturated hydrocarbons are lowered by 0.7 kcal/mol, and brought into better agreement with experiment, with the 6-311+G(2df,2pd) basis. Agreement is generally poorer for cc-pVDZ and for the various 6-31+G bases, with deviations as large as 2–5 kcal/mol in some cases. Except for the notable improvement of the agreement with experiment for CH_3^+ and, to a lesser degree, for $C_2H_5^+$, as the polarization space is increased in the 6-31+G basis sets, the hydrogenolysis energies obtained with the largest of these bases, (2df,2pd), are generally in worse agreement with experiment than those obtained with the smaller bases. As in the case of other trends, this behavior is particularly noticeable for acetylene and ethylene.

Overall Comparison of the Basis Sets. The numbers of primitive and contracted functions per carbon and hydrogen atom for each of the basis sets included in this study are listed in Table I. The 6-31G(d,p) and cc-pVDZ basis sets are of comparable size, and generally produce comparable energies. Both are sensitive to the choice of five or six d components and to the inclusion of diffuse functions, indicating significant deficiencies in their sp part. The larger cc basis sets, in contrast to the larger 6-31G bases, are not sensitive to these effects, and therefore the proper comparisons between the two series of basis sets should include diffuse functions for all the 6-31G bases and need not include them for the larger cc bases. Furthermore, only five d components need to be included for the latter, while the use of six components produces significant energy lowerings (but not necessarily better convergence of the hydrogenolysis energies) in the former. The 6-311G(2df,2pd) basis, which is equal in size to cc-pVTZ, is much less sensitive than 6-31G(2df,2pd) to the number of d components and the presence of diffuse functions, though the addition of diffuse functions in 6-311+G(2df,2pd) still produces a quite noticeable effect for the unsaturated neutral hydrocarbons, significantly more so than in the case of cc-pVTZ, and improves agreement with cc-pVTZ and cc-pVQZ′.

The number of contracted basis functions in the 6-31+G-(2df,2pd) basis, 30 per carbon atom and 13 per hydrogen with five d components, and 32 and 14, respectively, with six, is comparable to the numbers for the 6-311G(2df,2pd) and cc-pVTZ bases (30 per carbon atom and 13 per hydrogen, using five d components). As seen above, however, total and hydrogenolysis energies computed with the cc-pVTZ basis and, to a somewhat lesser extent, with 6-311G(2df,2pd) are closer to the nearly converged values, particularly for the unsaturated species. The 6-311+G(2df,2pd) basis produces results which are generally as good as cc-pVTZ, but is somewhat larger (34 contracted functions per carbon atom and 14 per hydrogen). The cc-pVTZ basis with five d components and without diffuse functions is therefore recommended as the preferred basis set for accurate work on hydrocarbon hydrogenolysis energies and, by extension, for other hydrocarbon reaction energies.

When an integral program (such as that in the Gaussian series[22]) which cannot handle generalized contractions is used, there is a penalty associated with the use of the cc basis sets in the required repetition of integral evaluation in terms of the primitive Gaussians which appear in more than one contraction. This disadvantage relates to the integral calculation only, and does not affect the Hartree–Fock iterations or the correlation treatments. Because only Hartree–Fock-occupied orbitals are represented by long contractions in the cc basis sets, the disadvantage is minimized by leaving out of the contractions those primitive functions which also appear as uncontracted basis functions in the cc bases, so that the only required repetition is a single repetition of a subset of the primitive s functions on each carbon atom. (Because of this feature, the Dunning cc basis sets are significantly cheaper than atomic-natural-orbital basis sets[33] for use with programs not designed for generalized contractions.) For the cc-pVTZ basis this repetition of primitive Gaussian basis functions increases the effective number of primitive functions per carbon atom, for the purpose of integral evaluation, to 50, compared to 43 (or 45, with six d components) for 6-31+G(2df,2pd) and 6-311G(2df,2pd) and 47 for 6-311+G(2df,2pd) (Table I). The number of primitive Gaussians per hydrogen atom is the same, 16, for these four bases.

Some timing comparisons have been made for MP4 calculations with the cc-pVTZ, 6-31+G(2df,2pd), 6-311G(2df,2pd), and 6-311+G(2df,2pd) basis sets on $C_2H_3^+$, carried out using Gaussian 88 in an 8 million word memory partition on the Cray Y-MP8/864. The repetition of primitive functions in the integral calculations was minimized as described in the preceding paragraph. As expected, the cc-pVTZ integral time is much longer (504 versus 251 s of CPU time) than the 6-31+G(2df,2pd) time, even though there are about 1 million more contracted integrals with the latter basis set (7 323 084 compared to 6 385 126) because of the use of six d components. (The Gaussian programs do not use integral symmetry in MP4 calculations, hence the relatively large number of integrals in this example.) However, the total times for the full MP4 runs (integrals, SCF, and MP4) differ by much less (1003 s of CPU time for cc-pVTZ, compared to 845 s for 6-31+G(2df,2pd)) because, with fewer integrals and fewer basis functions, the time required to construct and diagonalize the Fock matrix and to transform the integrals from an AO to an MO basis for the MP4 calculation is shorter with the cc-pVTZ basis set. In addition, the SCF calculation converged in one fewer cycle with this basis. The total CPU time for the 6-311G(2df,2pd) and 6-311+G(2df,2pd) MP4 calculations was 795 and 1031 s, respectively.

Summary and Conclusions

The data obtained in this study support the following observations and conclusions:

1. Hydrogenolysis energies computed with the augmented 6-31G basis sets often show poor convergence patterns, particularly when six components are used for the d functions. Moreover, hydrogenolysis energies computed with the largest of these basis sets sometimes do not agree with experimental data as well as energies computed with the smaller basis sets. In contrast, the cc basis sets show monotonic convergence, with the converged hydrogenolysis energies approaching the experimental values.

2. The choice of five versus six d components generally has little effect on total and hydrogenolysis energies computed with the cc basis sets, with the exception of cc-pVDZ, and a small effect for 6-311G(2df,2pd). In contrast, for the augmented 6-31G basis sets, the use of six components leads to significant total energy lowerings, particularly for 6-31G(2df,2pd) and 6-31+G(2df,2pd). Correlated hydrogenolysis energies show poorer convergence patterns for the 6-31G basis sets using six d components rather than five. Since there also is a computational advantage to the use of five d components, these observations suggest that there is little to be gained by using six components for the d functions in the 6-31G basis sets.

3. Diffuse functions are generally required to bring all 6-31G computed hydrogenolysis energies into better agreement with experimenal values. In contrast, diffuse functions have a smaller effect on hydrogenolysis energies computed with the 6-311G and cc-pVDZ bases, and essentially no effect for the larger cc basis sets, and are not needed there.

4. While in many cases either 6-31+G(d,p) or 6-31+G(2d,2p) reaction energies may be in reasonable agreement with experimental data, they cannot generally be considered internally converged with respect to basis completeness. In any particular case, and especially when unsaturated C–C bonds or strained bonds are present, these basis sets may give reaction energies which deviate significantly from converged values. Neither the split-valence 6-31G nor the cc-pVDZ basis sets, with any choice of polarization functions and with or without diffuse functions, can be relied upon for consistently accurate predictive work.

5. The hydrogenolysis energies computed with the cc basis sets appear to be essentially converged, to within about 1 kcal/mol or better, with respect to basis set size at cc-pVTZ. Thus, cc-pVTZ is the basis set of choice for quantitative work at reasonable computational expense. The 6-311+G(2df,2pd) basis is a reasonable alternative.

Acknowledgment. The support of the Ohio Supercomputer Center and the Youngstown State University Computer Center is gratefully acknowledged. The work of I.S. was partially supported by a grant from Cray Research, Inc.

Registry No. H_2, 1333-74-0; CH_3^+, 14531-53-4; CH_4, 74-82-8; C_2H_2, 74-86-2; $C_2H_3^+$, 14604-48-9; C_2H_4, 74-85-1; $C_2H_5^+$, 14936-94-8; C_2H_6, 74-84-0; C_3H_8, 74-98-6; cyclopropene, 2781-85-3.

---

[33] Almlöf, J.; Taylor, P. R. *J. Chem. Phys.* **1987**, *86*, 4070.