Exact geometries from quantum chemical calculations[☆]

Dieter Cremer[*], Elfi Kraka, Yuan He

Department of Theoretical Chemistry, Göteborg University Reutersgatan 2, S-41320 Goteborg, Sweden

Abstract

For seventeen molecules, complete basis set (CBS) geometries are obtained for Møller–Plesset perturbation methods at second (MP2), fourth (MP4), and sixth order (MP6) as well as for the Coupled Cluster methods CCD, CCSD, and CCSD(T). The correlation consistent basis sets cc-pVDZ, cc-pVTZ, and cc-pVQZ were systematically applied and calculated geometries extrapolated to the limit of an infinitely large basis set. MP6 equilibrium geometries are more accurate than MP2 or MP4 geometries at the CBS limit and provide AH bond lengths with an accuracy of 0.001 Å. However, AB bonds are always predicted too long because of the lack of sufficient coupling effects between p-electron correlation at MP6. CCSD(T) provides reasonable AB bond lengths although these are in general too short by 0.003 Å. Due to error cancellation very accurate geometries are obtained at the CCSD(T)/cc-pVTZ and CCSD(T)/cc-pVQZ level of theory. With the help of the accurate equilibrium geometries obtained in this work, several experimentally based geometries could be corrected. The effects of HF-optimized basis sets, diffuse functions or the frozen core approximation on geometry optimizations are discussed. It is emphasized that the use of the cc-pVDZ or any other VDZ + P basis set should be avoided in correlation corrected ab initio calculations. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Molecular equilibrium; Thermodynamics; Quantum chemical methods

1. Introduction

The knowledge of molecular equilibrium geometries is a basic requirement for the understanding of the electronic structure, thermodynamic stability, and chemical reactivity of molecules. Over the last century, experimental work has led to an impressive collection of reliable geometrical data of molecules utilizing spectroscopic (microwave, infrared, Raman, etc.) and diffraction methods (electron diffraction, X-ray diffraction, neutron diffraction, etc.) [1–10]. These techniques yield information on average nuclear positions in ground and excited vibrational states of a molecule, thermally averaged values of distances and angles, or molecular parameters such as effective ground state rotational constants that are related to the geometrical parameters of a molecule[11]. Because of the Heisenberg principle, it is not possible to measure equilibrium geometries r_e directly. These have to be derived from experimentally based molecular structures (r_0, r_g, r_α, r_z, also r_s, r_m, r_a, r_v) using additional information such as the molecular force field, rotation-vibration couplings, coriolis couplings, etc.

The equilibrium geometry r_e of a molecule, although not measurable, is well-defined and directly

---

[☆] Dedicated to Professor Marit Trætteberg on the occasion of her 70th birthday.

[*] Corresponding author. Tel.: +46-31-7735597; fax: +46-31-7735590.

E-mail address: cremer@theoc.gu.se (D. Cremer).

comparable to calculated geometries based on the use of reliable quantum chemical methods [7,11]. Actually, calculated r_e geometries can be used to predict, verify or improve measured geometries in two different ways.

1. Experimentally derived r_e geometries can be used to test the reliability of a given quantum chemical method. If the latter is found to be reliable it can be applied to determine the equilibrium geometry of molecules, for which it is not possible to obtain an experimentally based r_e geometry.
2. Quantum chemical methods can be used to calculate the r_z, r_0, r_g, etc. geometry of a molecule, which can be directly compared with measured geometries. If the method in question is known to lead to reliable geometries (point 1) this approach is the basis for directly checking, complementing or predicting measured geometries.

Clearly, most of the quantum chemical work has focused on task 1 where both the testing of quantum chemical methods and the prediction of unknown r_e-geometries have played a large role [12–21]. In recent times, more and more research has focused on task 2) to support the work of spectroscopists and diffractionists in structure determination. Typical examples are the recent structure determination of cyclopropane [22], dioxirane [23] or benzene [24].

In this work, we will contribute to a better understanding of quantum chemically obtained geometries. While experimental geometries depend on the measuring technique, the environment of the molecule (gas phase, liquid, solution or solid), and the state of the molecule (electronic, vibrational, rotational), calculated geometries of a given molecular state depend on the quantum chemical method applied, in particular the amount of electron correlation covered, and (within any LCAO approach) on the number of basis functions used. Since electron correlation effects and basis set effects both influence the calculated geometry, it is difficult to separate both effects. Thirty years of quantum chemical work have led to considerable knowledge about the accuracy of molecular properties calculated with a given basis set and a given method. Nevertheless, it is still an art in itself to choose method and basis set for the calculation of a molecular geometry in such a way that the

best compromise between computational cost and accuracy of geometrical parameters is obtained.

In this work, we will solve the method/basis set problem in a somewhat different way from what is usually done, namely by eliminating one of the two factors determining the accuracy of a calculated molecular geometry. This can be accomplished by determining that molecular geometry, which will be obtained when the basis set used with a particular quantum chemical method becomes infinitely large, i.e. when the complete basis set (CBS) limit is reached. Geometries obtained at the method/CBS level do no longer depend on deficiencies of a truncated basis set and, therefore, reflect directly the accuracy of the method chosen, which in turn is related to the electron correlation effects covered by it.

Although the advantages of using CBS limit geometries in quantum chemical investigations is obvious, it is not obvious how to determine them. The computational cost of single determinant methods, which cover just pair correlation effects (as described by double (D) excitations) scale with $O(M^5)$ or $O(M^6)$ where M denotes the number of basis functions. The inclusion of three-electron correlation effects (described by triple (T) excitations) leads to $O(M^7)$ and $O(M^8)$ methods while the inclusion of connected four-electron correlations (described by quadruple (Q) excitations) already requires $O(M^9)$ and $O(M^{10})$ methods. Utilizing supercomputers, a pair-correlation method can still be carried out with more than 1000 basis functions, the limit for the simplest T excitation methods is at about 250 basis functions (but extension to 450 basis functions is possible with direct methods, see, for example Ref. [25]), and reduces to just 50–60 basis functions in the case of the Q excitation methods.

Since a hextuple zeta basis augmented with d-, f-, g-, h-, and i-type polarization functions such as [7s6p5d4f3g2h1i/6s5p4d3f2g1h] can be considered to lead to energies and geometries close or equal to the basis set limit of a given ab initio method, a geometry optimization of a small molecule such as methane would include already 504 basis functions, which is already too large even for the simplest T method. Clearly, CBS limit geometries have to be determined in another way than just extending M to large enough values so that the CBS limit of a given property is reached within calculational accuracy.

However, one can make use of the fact that calculated geometrical parameters Q monotonically converge with increasing basis set to the CBS value $Q(CBS) = Q(M = \infty)$ and that this convergence behavior can be modeled with appropriate exponential functions. Work by Dunning and co-workers [26–30] and more recently by Cremer and co-workers [31–33] has confirmed this in many cases and results obtained by these authors constitute the basis for the present investigation. Furthermore, it can be shown that Q values determined with closely related ab initio methods possess the same or similar convergence behavior, which can also be exploited to obtain $Q(CBS)$ values for expensive methods in an economic way. In the case of Q being the molecular energy E, this was observed at an early stage by Pople and led to the Gp ($p = 1, 2, 3$) methods [34–36], which are based on the extrapolation to $E(CBS)$ values of highly correlated methods using the basis set dependence of E calculated for less sophisticated but also less costly methods.

We will present in this work CBS limit geometries as obtained with different correlation corrected methods. The goal of this work is to determine the method that leads to the most accurate r_e-geometries. For this purpose, we select a set of suitable test molecules (Table 1) for which (with a few exceptions) accurate experimentally based equilibrium geometries are known [37–66]. Since all molecules considered represent closed/open shell systems without significant multireference character, we will apply single determinant theory starting from a Hartree–Fock (HF) reference wave function. The correlations corrected methods used are limited to single determinant many body perturbation theory (MBPT) [67] and single determinant coupled cluster theory (CC) [68] because these theories lead to hierarchies of size-extensive methods (contrary to configuration interaction theory), which make it possible to investigate the influence of electron correlation effects of increasing order on the accuracy of calculated geometrical parameters. Such a relationship cannot be deduced from correlation corrected methods based on density functional theory (DFT) (for a description on DFT methods, see Ref. [69]) because this theory includes electron correlation in an unspecified way and does not lead to a hierarchy of methods, which lead with increasing sophistication to increasing accuracy of calculated results [70].

MBPT will be used with the Møller–Plesset (MP) perturbation operator [71] at increasing orders n of perturbation theory (MPn series). The MPn series is known to oscillate at lower orders for many molecular properties such as energy, geometry, dipole moment, etc.[67,72]. The even order methods such as second (MP2), fourth (MP4) (for a recent review, see Ref. [67]), and sixth order MP (MP6) [73–78] are the more attractive MP methods because they include always new correlation effects (pair correlation at MP2; orbital relaxation and three-electron correlation at MP4, connected four-electron correlations, disconnected 5- and 6-electron correlations at MP6), which lead to a better quantum chemical description while the odd order MPn methods introduce coupling corrections between correlation effects newly introduced at order n − 1. Cremer and He [73–78] showed that extrapolation of even order MPn methods and that of odd order MPn methods leads for energies to the same infinite order limit identical (or close to) the full CI (FCI) value. There are not sufficient data to confirm the same convergence behavior for MPn geometries and, therefore, it will not be possible to use MPn/CBS geometries to extrapolate to FCI/CBS geometries.

A better description of correlation effects is accomplished by projected CC methods [67,68] because these contain infinite order correlation effects. For example, the D excitations in CCD [79,80] cover all possible disconnected pair correlation effects for a given molecule while an extension to single (S) excitations as in CCSD [81] includes also all infinite order orbital relaxation effects. The addition of all infinite order three-electron correlation effects as described by the T excitations is obtained in CCSDT [82,83], which is a very accurate but also a very expensive method (cost being proportional to N_{iter} $O(M^8)$ where N_{iter} defines the number of iterations needed in order to obtain self-consistent CC amplitudes). Accordingly, one compromises mostly by adding T excitations in a perturbative way to CCSD thus yielding CCSD(T) [84], which scales with N_{iter} $O(M^7)$ and, therefore, can be applied for larger M (see above). Since CCSD(T) covers about 70% of the T effects [85,86] contained in the more complete CCSDT method, it provides reasonable estimates of CCSDT geometries.

We will present the results of this work by describing in the next chapter the computational

Table 1
Experimental geometries of the molecules investigated. Bond lengths are given in Å and angles in degree. Reference parameters (*value chosen* are exclusively r_e parameters or averages over several published r_e parameters as indicated in the table. For improved values, see text

Molecule	Parameter	Type	Exp. (error)	Year	Value chosen	Comment	Ref
CH_2 ($^3B_1$)							
	$r(CH)$	r_e	1.07530 (11)	1988	1.0753	Empirical fit of spectrscopic data	[37]
	$\theta(HCH)$	r_e	133.9308 (21)	1988	133.93		[37]
CH_2 ($^1A_1$)							
	$r(CH)$	r_e	1.1112	1988	1.1112	Improved to 1.1081 value kept	[37]
	$\theta(HCH)$	r_e	101.954	1988	101.954		[37]
CH_4 ($^1A_1$)							
	r(CH)	r_e	1.086 (2)	1994	1.0856	Average	[38]
		r_e	1.0849 (31)	1986			[39]
		r_e	1.0857 (10)	1979			[40]
HCN ($^1\Sigma$)							
	r(CH)	r_e	1.0655	1997	1.0655	Average	[40]
		r_e	1.06501 (8)	1992			[42]
		r_e	1.065825 (103)	1988			[43]
		r_e	1.06549 (24)	1971			[44]
	r(CN)						
		r_e	1.1534 ± 0.0005	1997	1.1533	Average	[41]
		r_e	1.15324 (2)	1992			[42]
		r_e	1.153193 (16)	1988			[43]
		r_e	1.15321 (5)	1971			[44]
HNC ($^1\Sigma$)							
	r(NH)	r_e	0.9960643 (29)	(1993)	0.9951	Average	[45]
		r_e	0.99696 (145)	1988			[43]
		r_e	0.9940 (8)	1978			[46]
		r_e	0.99357 (117)	1976			[47]
	r(NC)	r_e	1.1683506 (16)	(1993)	1.1685	Average	[45]
		r_e	1.16836 (23)	1988			[43]
		r_e	1.1689 (2)	1977			[46]
		r_e	1.1645 (28)	1976		Not taken for average	[47]
CO ($^1\Sigma^+$)							
	r(CO)	r_e	1.128230 (1)	1993	1.1282		[48]
		r_e	1.1282293 (10)	1993			[49]
		r_e	1.1283226 (1)	1992			[50]
CO_2 ($^1\Sigma_g^+$)							
	r(CO)	r_e	1.15995 (<1)	1986	1.1599		[51]
NH_3 ($^1A_1$)							
	r(NH)	r_0	1.0162	1982	1.0112	Average	[52]
		r_e	1.0118	1957			[53]
	$\theta(HNH)$	r_0	107.47	1982	106.54	Average	[52]
		r_e	106.69	1957			[53]
N_2 ($^1\Sigma_g^+$)							
	r(NN)	r_e	1.0975992 (13)	1997	1.0976		[54]
N_2H_2 (^1A_g)							
	r(NH)	r_e	1.029 (1)	1997	1.029		[55]
		r_0	1.0284 (78)	1974			[56]
	r(NN)	r_e	1.247 (1)	1997	1.247		[55]
		r_0	1.2517 (21)	1974			[56]
	$\theta(NNH)$	r_0	106.3 (1)	1997	106.3		[55]
		r_0	106.9 (8)	1974			[56]
HNO ($^1A'$)							
	r(NH)	r_0	1.0628	1958	1.0584	Improved to 1.0553	[57]

Table 1 (*continued*)

Molecule	Parameter	Type	Exp. (error)	Year	Value chosen	Comment	Ref
	r(NO)	r_0	1.2116		1.2075	Improved to 1.2074	[57]
	θ(HNO)	r_0	108.58		108.58	Improved to 108.00	[57]
H_2O ($^1A_1$)							
	r(HO)	r_e	0.9579205	1996	0.9579	Average	[58]
		r_e	0.957626 (11)	1994			[59]
		r_e	0.958290 (16)	1994			[60]
		r_e	0.95781 (3)	1979			[61]
	θ(HOH)	r_e	104.500	1996	104.50	Average	[58]
		r_e	104.510 (4)	1994			[59]
		r_e	104.522 (3)	1994			[60]
		r_e	104.478 (2)	1979			[61]
H_2O_2 (^1A)							
	r(OH)	r_e	0.9670	1993	0.9670	Improved to 0.9621	[62]
	r(OO)	r_e	1.4556	1993	1.4556	value kept	[62]
	θ(OOH)	r_e	102.32	1993	102.32	Improved to 99.80	[62]
	τ(HOOH)	r_e	113.70	1993	113.70	Improved to 112.00 for average	[62]
O_3 ($^1A_1$)							
	r(OO)	r_e	1.2717 (2)	1970	1.2717		[63]
	θ(OOO)	θ_e	116.7833 (192)	1970	116.7833		[63]
FH ($^1\Sigma^+$)							
	r(FH)	r_e	0.9169	1971	0.9169		[64]
F_2 ($^1\Sigma_g^+$)							
	r(FF)	r_e	1.4119	1976	1.4119		[65]
HOF ($^1A'$)							
	r(OH)	r_e	0.9657	1988	0.9657		[66]
	r(OF)	r_e	1.4350	1988	1.4350		[66]
	θ(HOF)	r_e	97.54	1988	97.54		[66]

procedure applied to obtain CBS limit geometries. In Chapter 3, we will compare calculated geometries of 17 small molecules with the experimentally based reference geometries listed in Table 1. On the basis of this investigation we will answer a number of interesting questions: (1) Does MP6 lead to reliable molecular geometries or do we have to apply even higher orders of MPn theory for this purpose?; (2) Which correlation effects have to be covered to obtain useful geometries?; (3) How do MP and CC methods perform in comparison?; (4) Is CCSD(T) sufficient to determine accurate geometries?; (5) What is the most economic way to calculate reliable geometries?

2. Determination of CBS limit geometries using MPn and CC theory

The geometries discussed in this work were calculated with standard MP2, MP3, and MP4 theory (for a recent review, see Ref. [67]) using an unfrozen core

and analytical energy gradients [14]. In addition, MP6 calculations (both full and frozen core) were carried out with the program of He and Cremer [73–78], which is part of the program package COLOGNE2000 [87]. Since analytical gradients are not available for MP6, geometries were calculated numerically with the Davidon–Fletcher–Powell algorithm [88]. Beside the complete MP6 calculations, which scale with $O(M^9)$, also the less costly MP6(M7) [73–76] calculations (scaling with $O(M^7)$) were carried out to investigate larger molecules. In these cases, differences between MP6 and MP6(M7) geometries obtained for smaller basis sets were utilized to predict MP6 geometries for large basis sets according to the procedure outlined by He and Cremer [31]. All CC calculations were carried out with full core and analytical energy gradients as implemented in the program ACES II [89,90].

In Table 1, the experimental reference geometries of the molecules investigated in this work are listed. For most molecules considered, several experimentally

based r_e geometries were published [37–66] where we preferably used those geometries published in the last decades. Otherwise it was difficult to select a particular r_e geometry out of several published and, therefore, we simply used averaged geometrical parameters as a reasonable choice for the purpose of comparison (see Table 1). In a few cases, we found r_e geometries differing so strongly from other published geometries that we excluded them from the averaging procedure. For $CH_2(^1A_1)$, $HNO(^1A')$, and $H_2O_2(^1A)$, r_e geometries are available, which are derived from ab initio calculations and fitting of spectroscopic data (CH_2 [37]) or from r_0 geometries with the help of appropriate models (HNO [91], H_2O_2 [63]). We used these data in our analysis of ab initio geometries to test their reliability, but based on the results of this work we were able to replace them by more accurate r_e geometries

Uncertainties in experimental r_e bond lengths and bond angles are often smaller than 10^{-3} Å and $0.1°$, respectively (Table 1). Therefore, these parameters were calculated with an accuracy of 10^{-4} Å and $10^{-2}°$. This required more stringent convergence criteria in both the SCF and the geometry optimizations than are normally applied. A convergence criterion of 10^{-8} with regard to changes in the density matrix was used for the SCF iterations and 10^{-5} with regard to changes in the forces (expressed in mdyn/Å) for the geometry optimization.

CBS limit geometries were determined by combining the use of (a) correlation consistent basis sets with (b) a suitable extrapolation formula as will be described in the following.

Basis sets. An investigation of the dependence of calculated geometries on the basis set employed requires the use of basis sets, which systematically and consistently improve the description of the space around the atoms of a molecule. In addition, it requires basis sets, which are designed for correlation corrections rather than HF calculations. The correlation consistent (cc) polarized (p) valence multiple zeta basis sets (cc-p VmZ) of Dunning [92–97] fulfill these requirements and, therefore, these basis sets were used throughout this work. The size of a cc-pVmZ basis is determined by the cardinal number m, which for $M = 2$ (cc-pVDZ), 3 (cc-pVTZ), and 4 (cc-pVQZ) defines basis sets of VDZ + P, VTZ + P, and VQZ + P quality given by (9s4p1d/4s1p) [3s2p1d/2s1p], (10s5p2d1f/

5s2p1d) [4s3p2d1f/3s2p1d], and (12s6p3d2f1g/6s3p2d1f) [5s4p3d2f1g/4s3p2d1f] contractions, respectively, [91]. Dunning and co-workers [26–30,93–97] showed that calculated molecular properties Q obtained with cc-pVmZ basis sets converge monotonically to limit values $Q(\infty)$ for $m \to \infty$, which can be used to represent the corresponding CBS limit for the properties in question.

Since Dunning's cc-pVDZ basis [81] is less used than Pople's 6-31G(d,p) basis, [98] we also tested how geometries change in dependence of the composition of a given VDZ + P basis set. In some cases, aug-cc-pVmZ basis sets [93,94] as well as the larger cc-pV5Z basis set were employed for test calculations.

Calculation of CBS limits geometries. A number of different extrapolation procedures are suggested in the literature, [26,27,99,100] which are used to derive from a limited number of calculations with increasing basis set the CBS limit value of Q. Since changes in the property value decrease exponentially for increasing basis set and become zero for the CBS limit, most authors used exponential fitting functions. Recently, He and Cremer [31] tested several extrapolation equations suggested by Dunning and co-workers [26,27] for extrapolating energies and geometrical parameters (Q = energy E, bond length r, bond angle θ, dihedral angle τ) to the CBS limit and found Eq. (1) to be useful in terms of applicability, accuracy, and interpretability:

$$Q(m) = Q(\infty) + ae^{-m} + be^{-m2} \tag{1}$$

The fitting parameters a and b and the CBS limit $Q(\infty)$ can be determined from three geometry optimizations using a given ab initio method and cc-pVmZ basis sets for $M = 2$, 3, and 4 [31]. Dunning and co-workers [26–30] derived more precise CBS values by applying also the cc-pV5Z and cc-pV6Z basis sets in selected cases and solving for the three unknowns of Eq. (1) in a least squares fashion. However, such a procedure can no longer be applied for expensive methods such as MP4, CCSD(T) or MP6 in the case of larger molecules and, therefore, we preferred the 3-point procedure for all molecules considered thus loosing some accuracy but gaining a common basis for the comparison of CBS limit geometries at different levels of theory.

Despite this simplification, it was still not feasible

to carry out MP6 or CCSD(T) geometry calculations for some of the molecules using cc-pVTZ and cc-pVQZ basis sets. In these cases, we exploited the fact that the fitting parameters a and b of Eq. (1) are closely related for one and the same geometrical parameter when obtained with related MP or CC methods. For example, MP4/CBS limit values determined at the MP4 level with three cc-pVmZ points or just with one point using the curve Q(m) determined at the MP2 level differ only slightly (the mean absolute deviation $|\mu|$ for 20 bond lengths and seven angles is about 0.001 Å and 0.1°, respectively. For more details, see Ref. [31]). Hence, we applied the a, b fitting coefficients determined with MP4 (or CCSD) to calculate Q(CBS) values at the MP6 (or CCSD(T)) level of theory.

In the case of the CCSD(T) geometries, it was possible to compare predicted geometries with recently published CCSD(T) geometries obtained with a *direct* CC method[21,101]. Predicted AH bond lengths are, compared to calculated values, on the average somewhat too short, which leads to an average error of -0.0010 Å for CBS limit values. This means that calculated AH bond lengths are in somewhat better agreement with experimental values. The corresponding errors for predicted AB bond length are smaller (see Table 2) both for a specific basis set and the CBS limit. In the case of bond angles, deviations between predicted and calculated CCSD(T) angles are negligible. Hence, the analysis of calculated CCSD(T) geometries leads to the same general conclusions as found when using predicted CCSD(T) geometries. This provides additional evidence for the reliability of predicted geometries in cases where now and in the near future geometry claculations will not be possible for large basis sets (MP6/cc-pVQZ, etc.).

Error analysis. Deviations of calculated geometrical parameters from the corresponding experimental values were analyzed using mean deviation μ, mean absolute deviation $|\mu|$, and standard deviation σ:

$$\mu = \frac{1}{n} \sum_{i=1}^{n} \Delta_i, \text{ for } \Delta_i = Q_i(\text{cal}) - Q_i(\text{exp}) \qquad (2)$$

$$|\mu| = \frac{1}{n} \sum_{i=1}^{n} |\Delta_i| \qquad (3)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (\Delta_i - \mu)^2}{n - 1}} \qquad (4)$$

where Q_i(cal) denotes a calculated and Q_i(exp) the corresponding experimental value of a geometrical parameter $Q_i = \{r, \theta, \tau\}$.

3. Comparison of calculated and experimental geometries

A complete list of all calculated equilibrium geometries is given in the supporting information to this work. Deviations from experimental geometries are analyzed in Figs. 1–6 where the analysis is done separately for MP (Figs. 1–3) and CC methods (Figs. 4–6). A summary of the error analysis is given in Table 2 using the definitions of Eqs. (2)–(4). Geometrical parameters are separated into AH bond lengths, AB bond lengths (A and B heavy atoms), and bond angles (covering both HAH, HAB, and BAB bond angles as well as the HOOH dihedral angle) in view of the different electronic features influencing these geometrical parameters: There are less electrons in the vicinity of an AH bond (bonding electron pair plus possible electron lone pairs at A) than in the vicinity of an AB bond (one or more bonding electron pairs plus possible electron lone pairs both at A and B) and, therefore, it is much easier to describe an AH bond length accurately than an AB bond. Bond angles XAX at a central atom A depend on the AX bond lengths, the bond polarity of AX, and nonbonding interactions between the terminal atoms X, which should make it more difficult to get accurate bond angles than bond lengths. However, bond angles are normally needed with a lower accuracy (±0.1 rather than $\pm0.01°$) and, therefore, requirements for their calculation are less stringent than those for determining bond lengths.

Inspection of Figs. 1–6 and the data in Table 2 leads to the following conclusions:

1. For an increase of the size of the basis set, bond lengths become shorter while bond angles widen irrespective of the use of a particular MP or CC method. This is a general observation, which was first made at the HF level of theory[12–15,102],

Table 2
Mean deviation μ, mean absolute deviation $|\mu|$, and standard deviation σ of calculated geometrical parameters. μ, $|\mu|$ and σ are defined in Eqs. (2)–(4) and are given in Å or deg. Numbers in parentheses denote error values obtained when replacing estimated CCSD(T) geometries by calculated geometries.

Parameter	cc-pVDZ MP2	cc-pVTZ	cc-pVQZ	CBS	cc-pVDZ MP4	cc-pVTZ	cc-pVQZ	CBS	cc-pVDZ MP6	cc-pVTZ	cc-pVQZ	CBS
AH bonds												
μ	0.0098	−0.0033	−0.0037	−0.0038	0.0129	−0.0008	−0.0009	−0.0009	0.0137	−0.0000	−0.0001	−0.0000
$\|\mu\|$	0.0098	0.0035	0.0037	0.0038	0.0129	0.0016	0.0011	0.0011	0.0137	0.0015	0.0005	0.0006
σ	0.0034	0.0033	0.0029	0.0028	0.0051	0.0020	0.0010	0.0011	0.0058	0.0017	0.0007	0.0009
AB bonds												
μ	0.0170	0.0019	−0.0004	−0.0016	0.0249	0.0098	0.0079	0.0069	0.0258	0.0103	0.0085	0.0069
$\|\mu\|$	0.0170	0.0091	0.0076	0.0067	0.0249	0.0098	0.0079	0.0070	0.0258	0.0103	0.0085	0.0069
σ	0.0093	0.0101	0.0095	0.0091	0.0065	0.0048	0.0041	0.0038	0.0095	0.0060	0.0059	0.0056
Bond angles												
μ	−0.50	−0.06	−0.02	0.01	−0.65	−0.08	−0.06	−0.05	−0.59	−0.02	0.00	0.01
$\|\mu\|$	1.90	0.59	0.36	0.28	2.02	0.66	0.42	0.27	2.03	0.68	0.42	0.29
σ	2.77	0.84	0.47	0.40	2.63	0.92	0.51	0.38	2.61	0.93	0.58	0.49

| Parameter	cc-pVDZ CCD	cc-pVTZ	cc-pVQZ	CBS	cc-pVDZ CCSD	cc-pVTZ	cc-pVQZ	CBS	cc-pVDZ CCSD(T)	cc-pVTZ	cc-pVQZ	CBS
AH bonds												
μ	0.0105	−0.0044	−0.0050	−0.0053	0.0115	−0.0036	−0.0041	−0.0044	0.0132	−0.0019 (13)	−0.0024 (16)	−0.0027 (17)
$\|\mu\|$	0.0105	0.0044	0.0050	0.0053	0.0115	0.0036	0.0041	0.0044	0.0132	0.0019 (15)	0.0024 (16)	0.0027 (19)
σ	0.0054	0.0017	0.0018	0.0023	0.0053	0.0017	0.0016	0.0021	0.0052	0.0015 (17)	0.0011 (11)	0.0017 (19)
AB bonds												
μ	0.0016	−0.0156	−0.0188	−0.0207	0.0075	−0.0108	−0.0141	−0.0160	0.0200	0.0017 (18)	−0.0017 (11)	−0.0036 (30)
$\|\mu\|$	0.0096	0.0156	0.0188	0.0207	0.0104	0.0108	0.0141	0.0160	0.0200	0.0023 (24)	0.0022 (17)	0.0036 (30)
σ	0.0118	0.0115	0.0119	0.0122	0.0095	0.0085	0.0091	0.0095	0.0097	0.0023 (20)	0.0028 (24)	0.0032 (27)
Bond angles												
μ	−0.36	0.34	0.45	0.50	−0.40	0.40	0.36	0.33	−0.60	0.20	0.16	0.14
$\|\mu\|$	1.39	0.50	0.52	0.68	1.47	0.54	0.43	0.68	1.81	0.57	0.33	0.25
σ	1.93	0.54	0.57	0.71	2.10	0.70	0.48	0.77	2.51	1.12	0.44	0.30
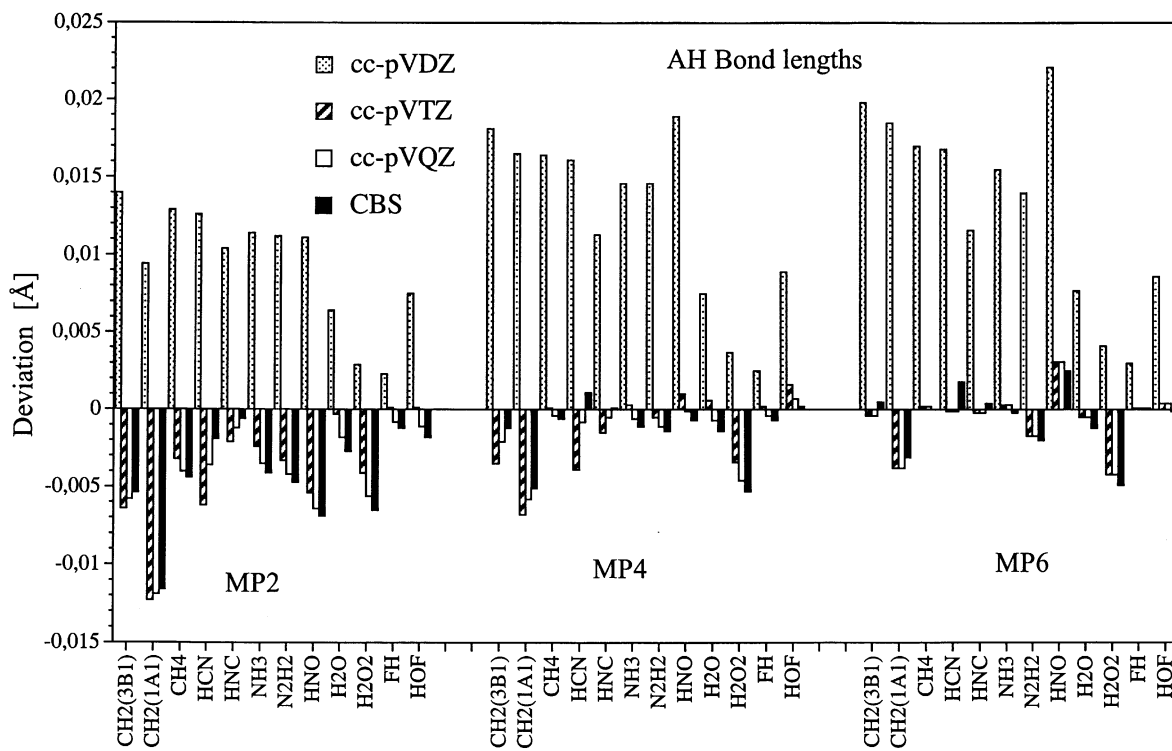
Fig. 1. Errors (relative to the chosen reference values listed in Table 1) of calculated AH bond lengths obtained at the MP2, MP4, and MP6 level of theory for the cc-pVDZ, cc-pVTZ, cc-pVQZ, and cc-pV∞Z (CBS limit) basis sets.

but which also applies at all levels of correlation corrected ab initio theory. Enlarging the basis set leads to an improved description of the valence shell of an atom thus allowing that electron density will be packed more closely around a nucleus in a molecule even if correlation effects hinder clustering directly at the nucleus. Hence, the nuclei are better screened and as a consequence nuclear repulsion is reduced thus leading to smaller bond lengths. In a three-atom situation, smaller bond lengths imply stronger 1,3 electrostatic repulsions and a widening of the bond angle, which explains the reverse relationship between trends in bond lengths and trends in bond angles. Exceptions from the general trend of bond shortening and angle widening with increasing basis set size are only found for some cc-pVDZ results and actually indicate the drawbacks of using VDZ + P basis sets in correlation corrected calculations.

2. Use of VDZ + P basis sets in correlation corrected calculations of geometries (and other properties) is inappropriate and has to be avoided. Although most correlation corrected ab initio calculations were and are still carried out with VDZ + P basis sets, results obtained in this work with the cc-pVDZ basis, clearly indicate that VDZ + P basis sets lead to large errors in calculated geometries where errors increase when higher order correlation errors are included as for example in the series MP2, MP4, MP6 (Figs. 1–3) or CCD, CCSD, CCSD(*T*) (Figs. 4–6). VDZ + P bond lengths become longer and VDZ + P bond angles smaller the more electron correlation effects are covered by a given method. Two important guidelines for future work can be derived from these observations. (a) Correlation corrected ab initio calculations have to be carried out at least with a VTZ + P basis set to obtain reasonable geometries and other molecular
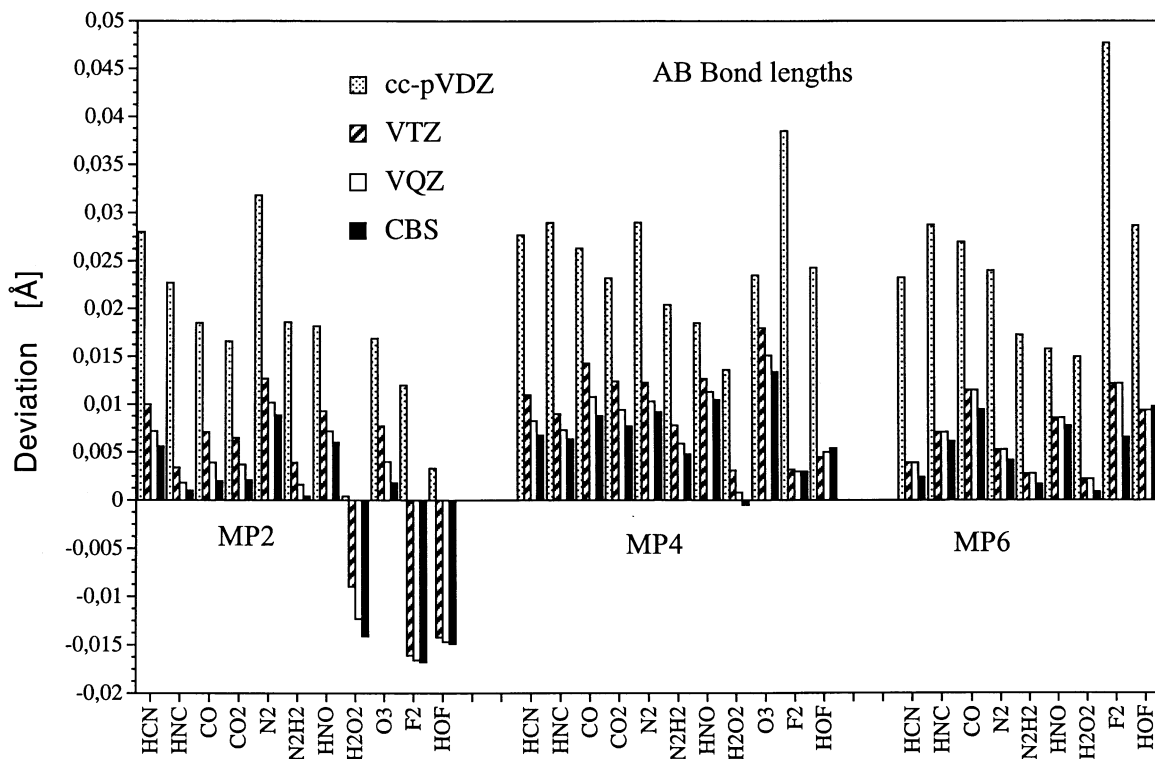
Fig. 2. Errors (relative to the chosen reference values listed in Table 1) of calculated AB bond lengths obtained at the MP2, MP4, and MP6 level of theory for the cc-pVDZ, cc-pVTZ, cc-pVQZ, and cc-pV∞Z (CBS limit) basis sets.

properties. (b) The extrapolation procedure applied in this work would become more accurate if it could be based on a cc-pVTZ, cc-pVQZ, and a cc-pV5Z basis set, however at the moment this is not possible for many molecules because of computational limitations.

3. Bond lengths increase and angles become (slightly) smaller when including higher order correlation effects. However, these trends are somewhat different for MP and CC methods. Electron correlation effects can be best analyzed by using CBS geometries thus largely excluding the influence of basis set truncation. For MP methods the mean deviation (increases for AH bonds from -0.0038, to -0.009 and 0 Å, for AB bonds from -0.0016 to 0.0069 and 0.0069 Å while for CC methods the μ-values are -0.0053, -0.0044, -0.0017 and -0.0207, -0.0160, -0.0030 Å (Table 2). Hence, on the

average CBS bonds are shorter for CC than MP methods that of course reflects the way how electron correlation effects are covered by the different methods.

MP methods introduce with each even order new connected p-particle effects: at MP2 pair correlation ($p = 2$), at MP4 connected three-electron correlation ($p = 3$), at MP6 connected four-electron correlation ($p = 4$), etc. These p-particle effects lead to a stronger separation of electrons in molecular space, an extension of the electron density distribution, deshielding of the nuclei, and hence longer bond lengths. Actually, the connected p-particle effects of MP theory introduced at even orders exaggerate electron correlation, which requires corrections at odd orders that lead to a coupling between p-particle correlations, a reduction of the effects of the latter and, by this, to less deshielded nuclei and shorter bond lengths.
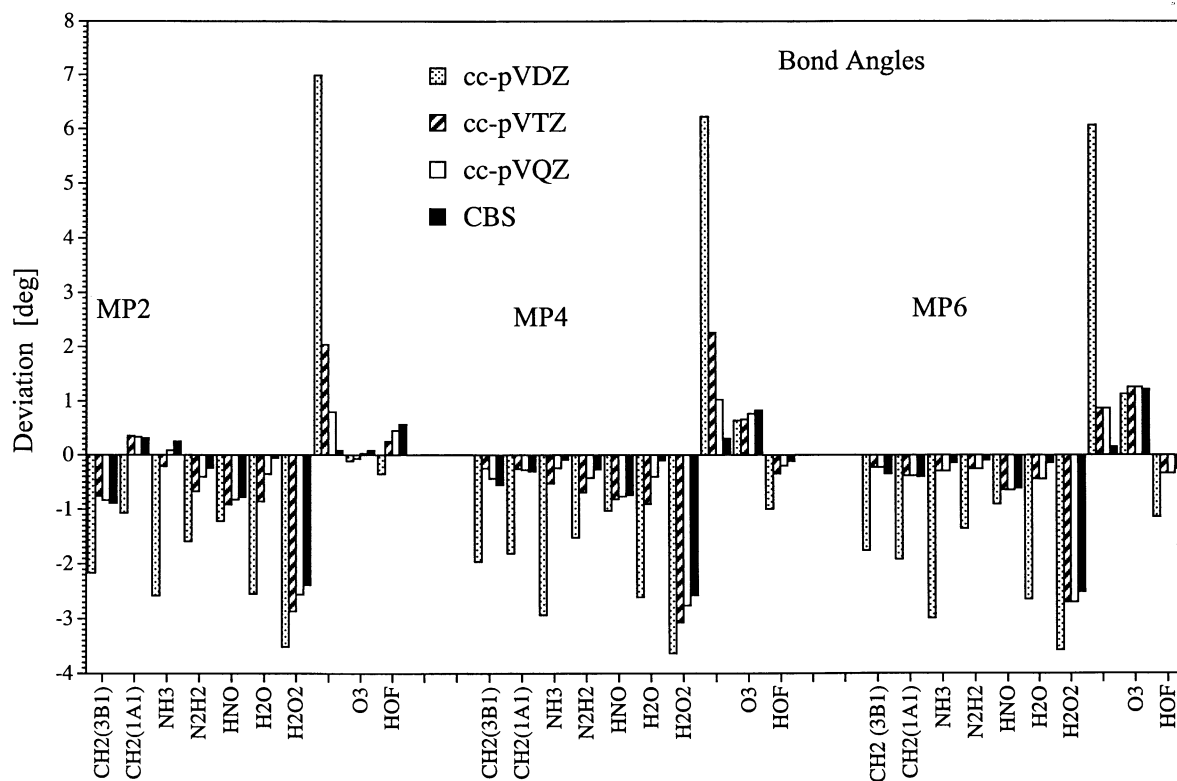
Fig. 3. Errors (relative to the chosen reference values listed in Table 1) of calculated angles obtained at the MP2, MP4, and MP6 level of theory for the cc-pVDZ, cc-pVTZ, cc-pVQZ, and cc-pV∞Z (CBS limit) basis sets.

The coupling corrections require high values of order n before a given type of electron correlation is sufficiently corrected and described in a balanced way. For example, the strong exaggeration of pair correlation at MP2 is first corrected at MP3 by coupling of the D excitations (the pairwise correlated electrons of different pairs start to *see* each other), however this coupling correction is not sufficient and is further corrected at MP4, MP5, MP6, etc. up to a sufficiently high order. CC methods include the coupling corrections in a more economic way by covering a given type up to infinite order because of the exponential form of the wave operator and by this of the CC wave function. For example, CCD includes pair correlation as described by MP2, however covers also all possible coupling corrections such as disconnected Q effects at fourth order, DD effects at fifth order, DDD effects at sixth order, etc. (for a

discussion of how these effects are included into CCD, see Ref. [85]). Hence, an exaggeration of bond lengths for a given p-particle correlation effect as found for even ordered MPn methods is not possible at the CC level, which is the reason why CC/CBS bond lengths are always shorter than MP/CBS bond lengths obtained by including the same p-particle correlation ($p = 2$: CCD and MP2; $p = 1$ and 2: CCSD; $p = 1, 2$ and 3: MP4 and CCSDT or, as an approximation to the latter, CCSD(T); $p = 1, 2, 3, 4$: MP6).

The CC and MP methods used in this work are not directly comparable because they cover different p-particle correlation effects. However, a user of these methods may be more interested in the computer time requirements and on this basis the $O(M^5/M^6)$ methods MP2, CCD, CCSD and the $O(M^7)$ methods MP4, CCSD(T) may be related while MP6 as an $O(M^9)$ method is in an
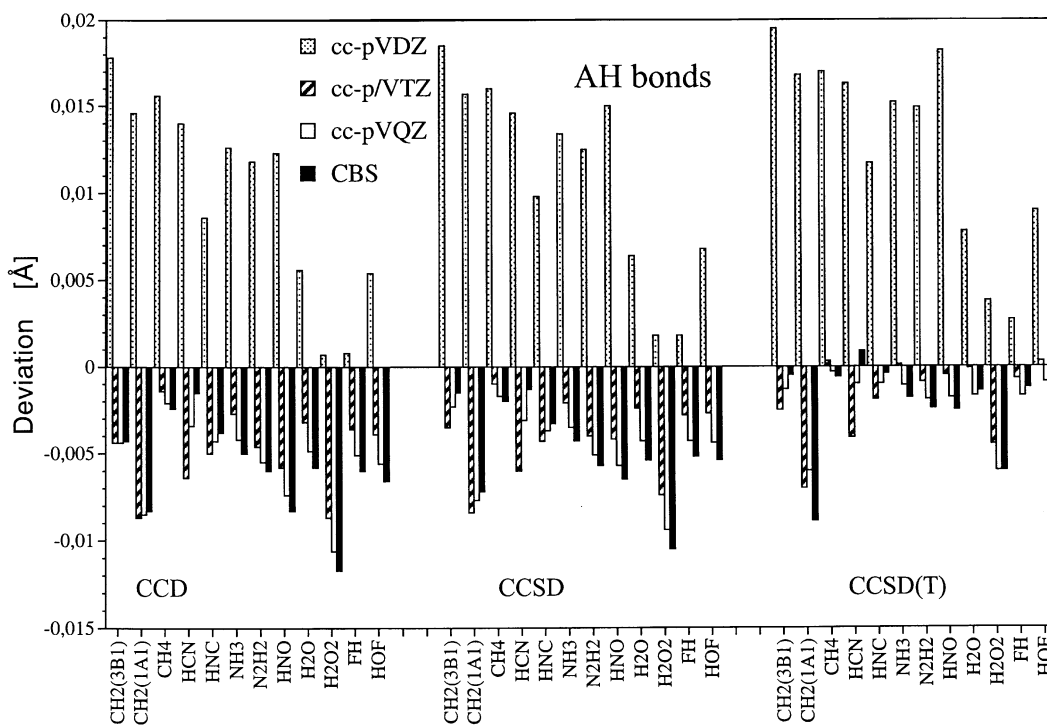
Fig. 4. Errors (relative to the chosen reference values listed in Table 1) of calculated AH bond lengths obtained at the CCD, CCSD, and CCSD(*T*) level of theory for the cc-pVDZ, cc-pVTZ, cc-pVQZ, and cc-pV∞Z (CBS limit) basis sets.

extra class. On the basis of this relation, it can be said that *a CC method always leads to shorter (larger) bond lengths (bond angles) than the corresponding MP method* (see μ values in Table 2) *due to the inclusion of more coupling effects in the former.*

4. The accuracy of calculated geometries increases with increasing order of the electron correlation effects covered by the ab initio method used. MP6 provides the best equilibrium geometries when using MP theory, CCSD(*T*) when using CC theory. The accuracy of calculated geometries is reflected by both mean deviations μ, mean absolute deviations $|\mu|$, and standard deviations σ of the CBS results, which all decrease when including more and more electron correlation effects. In particular, σ, which describes the scattering of errors in calculated geometrical parameters, is a reliable indicator for increased accuracy (leading to less scattering) with an increased coverage of electron correlation effects.

Also, the increase rather than decrease of σ for AB bonds and bond angles when replacing MP4 by MP6 indicates the magnitude of errors caused by the extrapolation procedure applied to predict MP6/large basis set geometries.

5. Individual method-basis set combinations may lead to reasonable geometries due to a cancellation of basis set and correlation errors.

(5a) *MP2:* If one has to use MP2 for geometry calculations, one should apply a VTZ + P basis because MP2/cc-pVTZ leads to reasonable geometries (in particular AH bond lengths, see Table 2 and Figs. 1–3). A cc-pVQZ leads to slightly better results, however these do not compensate for the exceedingly high computational cost. In no case, can one expect useful geometries from MP2/cc-pVDZ.

(5b) *CCD and CCSD:* From the results obtained in this work it is difficult to see why either CCD or CCSD should be applied. The computational accuracy is low and comparable
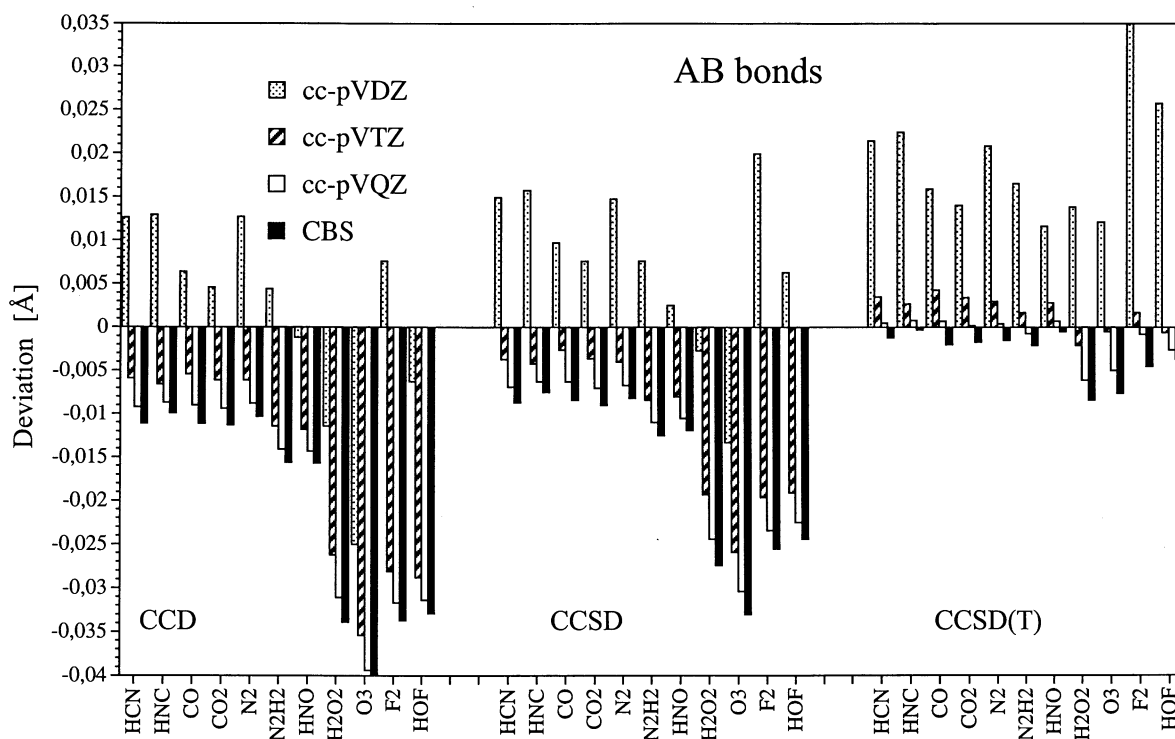
Fig. 5. Errors (relative to the chosen reference values listed in Table 1) of calculated AB bond lengths obtained at the CCD, CCSD, and CCSD(*T*) level of theory for the cc-pVDZ, cc-pVTZ, cc-pVQZ, and cc-pV∞Z (CBS limit) basis sets.

to that of MP2 calculations, however an MP2/ cc-pVTZ calculation is much less costly than a CCSD/cc-pVTZ calculation $(N_{iter}\mathrm{O}(M^6)$ compared to $\mathrm{O}(M^5))$.

(5c) *MP4 and CCSD(T):* Accurate AH bond lengths and reasonable AB bond lengths are obtained with MP4/cc-pVTZ while MP4/cc-pVDZ geometries are poor. Instead of spending the cost for an MP4 calculation $(\mathrm{O}(M^7))$ one should invest directly into a CCSD(*T*)/cc-pVTZ geometry optimization $(N_{iter}\mathrm{O}(M^7))$ because results are excellent with this approach (AH errors: −0.0013, 0.0015, 0.0017; AB errors: 0.0018, 0.0024, 0.0020 Å, Table 2), in particular because deviations are relatively regular and provide the possibility of systematic empirical corrections.

(5d) *MP6:* The most expensive MP approach presently in use does not provide more accurate data than MP4 (apart from a few exceptions) using the cc-pVTZ basis. MP6/cc-pVTZ AH

bond lengths are rather accurate but AB bond lengths are still far too long which is indicative of the lack of important coupling effects. In view of the high computational cost of MP6 $(\mathrm{O}(M^9))$ there is no reason to use this method.

6. *The highest accuracy for AH bonds is obtained at the MP6/CBS level and for AB bonds at the CCSD(T)/CBS level.* The most accurate AH bond lengths ever obtained are calculated at the MP6/CBS limit (errors: −0.001, 0.0006, 0.0009 Å). If atom A possesses two electron lone pairs, then there will be at the most three electron pairs in or close to the AH bonding region. MP6 covers disconnected hextupel (three electron pairs) correlations and, accordingly, it is able to adequately describe AH bonds. MP6/CBS values for AB bond lengths suffer from the fact that this method does not contain enough coupling effects for connected four-electron correlations needed for AB bonds and, consequently, these parameters are too long. However, CCSD(*T*)
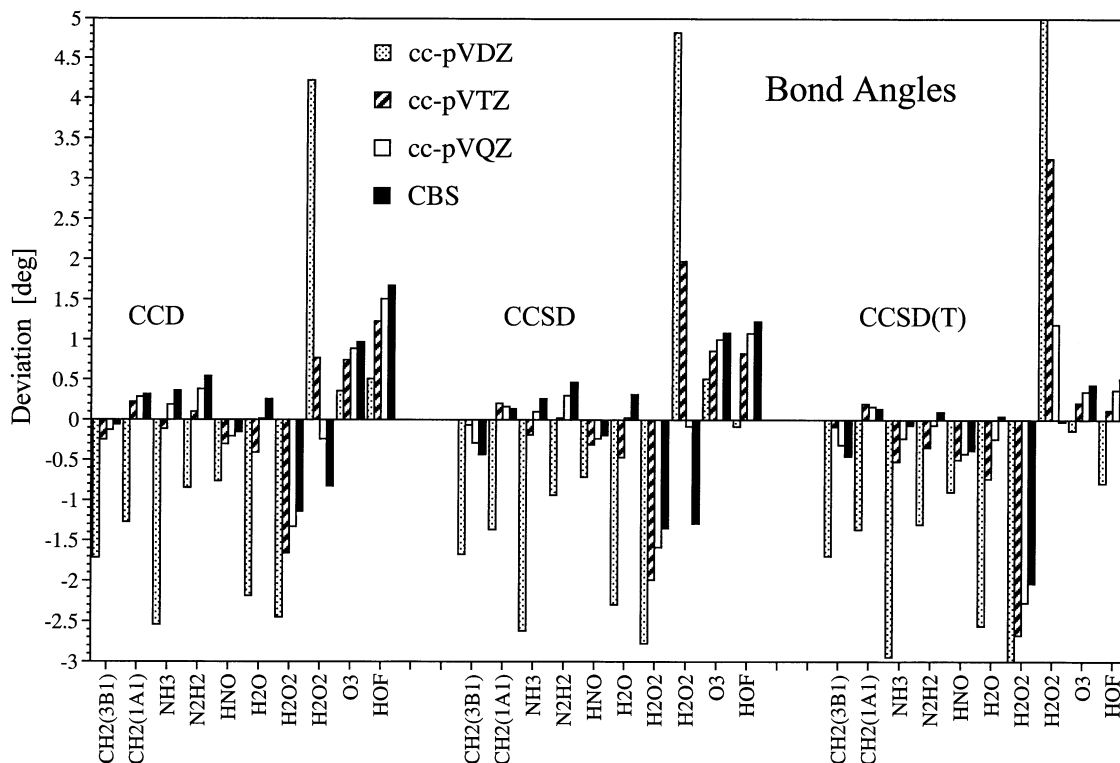
Fig. 6. Errors (relative to the chosen reference values listed in Table 1) of calculated angles obtained at the CCD, CCSD, and CCSD(T) level of theory for the cc-pVDZ, cc-pVTZ, cc-pVQZ, and cc-pV∞Z (CBS limit) basis sets.

has a sufficient number of coupling effects for three-electron correlations (yielding short bond lengths) while it lacks connected four-electron correlation effects not to speak of the 30% of T effects not covered by CCSD(T). Consequently, AB bond lengths are too short. There are situations such as for the OO bond in H_2O_2(^1A) (see Figs. 2 and 5) or F_2O_2(^1A) [103] where MP6 is superior to CCSD(T) in the CBS limit because the former method contains connected four-electron correlations important for a correct description of OO bonding in these molecules.

7. In view of (5) and (6), the best compromise in terms of computational cost and accuracy is obtained at the CCSD(T)/cc-pVTZ or CCSD(T)/cc-pVQZ level of theory. Clearly, this level of accuracy can no longer be reached for medium-sized or larger molecules. On the other hand, the extrapolation procedure described in this work provides a basis for estimating MP6

or CCSD(T)/cc-pVmZ geometries even for larger molecules provided a geometry optimization can be carried out for these methods with a cc-pVDZ basis set (loosing in this way some of the accuracy normally reached when larger basis set calculations are possible).

8. In view of the MP6/CBS and CCSD(T)/CBS results presented in this work, currently used r_e geometries for some of the molecules listed in Table 1 have to be considered to be erroneous. These values (still contained in Figs. 1–6) were corrected as described in the following and used in the error analysis presented in Table 2.

  (8a) CH_2($^1A_1$). Jensen and Bunker [37] fitted measured vibrational energies of this molecule to a model potential function partially based on CI calculations. Although not explicitly discussed in their work, their fit seems to support the equilibrium geometry obtained in the CI calculations. For the CH bond length, a

value of 1.111223 Å was obtained[37], which according to our calculations is too long. At the CCSD(T)/CBS level, we obtained 1.1051, at the MP4/CBS level 1.1061, and at the MP6/CBS level 1.1081 Å where the latter value is preferred because of the lower accuracy of MP4 and CCSD(T) in the case of AH bonds. The HCH angle given by Bunker and Jensen (101.9) agrees well with our CBS values although it may be 0.4° too large.

(8b) HNO($^1A'$). For this molecule, a r_0 geometry (r(NH) = 1.0628, r(N=O) = 1.2116, θ(HNO) = 108.58°) is available [57], from which we estimated r_e values of 1.0584, 1.2075 and 108.58 on the basis of previous calculations [92]. However, the MP6/CBS value for the NH bond (1.0553 Å) is clearly shorter than our previous estimate and, therefore, it is taken as a more reliable r_e(NH) value. In the case of the NO bond we considered that the CCSD(T)/cc-pVTZ and the cc-pVQZ value should bracket the true r_e value because the former approach leads to a too long, the latter to a too short AB bond length (see Table 2 and discussion above). Accordingly, we get r_e(NO) = 1.2074 by averaging these two values. Similarly, θ(HNO) = 108.00 was obtained as the average of the MP6/CBS and CCSD(T)/cc-pVTZ value considering that the NH bond length is better described by MP6/CBS and the NO bond length better by CCSD(T)/cc-pVTZ.

(8c) H$_2$O$_2$(^1A). In the case of hydrogen peroxide, Winnewisser and co-workers [62] investigated the torsional potential of the molecule utilizing measured rotational constants. Their starting geometry was a r_0 geometry published by Flaud and co-workers [104] and the equilibrium geometry was calculated at the minimum of the torsional potential. When using the r_e geometry of Winnewisser and co-workers exceptionally large deviations are obtained for r_e(OH) and θ_e(OOH) at MP4/CBS, CCSD(T)/CBS, and MP6/CBS (Figs. 1, 3, 4 and 6), which clearly indicates errors in the experimentally based r_e geometry. Therefore, we consider the MP6/CBS values r_e(OH) = 0.9621, θ_e(OOH) = 99.80, and τ_e(HOOH) =

112.0 (keeping the r_e(OO) value of Winnewisser and co-workers [62], which agrees well with the MP6/CBS value) to represent a more correct r_e geometry.

9. Limitations of the approach for getting CBS limit geometries are given by the performance of the cc-pVDZ basis at a given method level. On the one hand the cc-pVDZ basis is clearly not suited for getting reasonably accurate geometries with correlation corrected methods while on the other hand this basis set has to be employed for getting CBS limit geometries at reasonable cost (see discussion above). Actually, CBS limit geometries based on cc-pVDZ, cc-pVTZ, and cc-pVQZ geometries are still reasonable as we found when using in addition the cc-pV5Z for some small molecules. Calculations with the cc-pVDZ became however problematic at the MP6 level when describing multiple bonds involving electronegative atoms. For example at the MP6/cc-pVDZ level the CO$_2$ molecule dissociates and the OO bonds of O$_3$ become unreasonably long, which is the reason why these molecules were excluded from the MP6 analysis. Hence, the necessary use of the cc-pVDZ basis limits the accuracy and applicability of the procedure to predict MP6/CBS geometries.

Apart from the general observations (1)–(9) referring to the most likely method/basis set combinations, we tested also some calculational strategies, which are commonly used in quantum chemical studies of equilibrium geometries even though their applicability may be questioned.

10. *Use of HF-optimized rather than correlation-consistent basis sets.*. Although correlation-consistent basis sets have been available for more than a decade, basis sets are frequently employed, which were originally designed for use in HF rather than correlation corrected calculations. HF-optimized basis sets (VTZ + P and better) reflect the tendency of HF to cluster electrons in the average field description as close as possible at the nuclei, which leads to a rather contracted molecular electron density, strong shielding of the nuclei, and too short (large) bond lengths (bond angles). The contraction of the electron density can be anticipated from the exponents of the basis functions and the weight of

the inner shell functions, which are both larger than in a correlation-consistent basis set.

If one applies a HF-optimized basis set in a correlation-corrected ab initio calculation, the tendency of separating electrons by the method is somewhat reduced by the basis set that does not describe adequately regions farther away from the nuclei. Consequently, bond lengths become shorter than those calculated with the corresponding correlation consistent basis set, which reminds of the shortening caused by the coupling effects between p-particle correlations. Hence, a HF-optimized basis set applied with a MPn method can lead to better geometry while in the case of CC methods it should lead to a deteriorated equilibrium geometry (for an exception, see Ref. [103]). However, even in the case of MP methods the use of HF optimized basis sets cannot be recommended as a means to improve r_e geometries because improvements obtained for the wrong reason will spoil generally derived basis set trends as discussed in this work, i.e. results become unpredictable.

11. *Use of Pople's 6-31G(d,p) rather than Dunning's cc-pVDZ basis.* In line with the discussion given under 9) we find that MPn/6-31G(d,p) ($n = 2, 4, 6$) values for AH bonds are much closer to experimental values than MPn/cc-pVDZ values. On the other hand, AB bond lengths obtained with Pople's 6-31G(d,p) are 10–20% longer than the corresponding cc-pVDZ values, i.e. they deviate even stronger from experimental values. This result can be easily explained considering the composition of the two basis sets. Both basis sets contain the same number of basis functions for H and first row atoms, only that the 6-31G(d,p) is based on cartesian d-type polarization functions while the Dunning basis sets exclusively work with spherical polarization functions, which means that there is an extra 3s function in the 6-31G(d,p) basis. The major difference between the two basis sets, however, results a) from the larger number of s-type primitives used in the cc-pVDZ basis and b) the sharing of s,p exponents typical of Pople basis sets. Hence, the cc-pVDZ is much more flexible in the inner shell region and can accommodate more negative charge in this region thus leading to a better

shielding of the heavy atoms. This is confirmed by the fact that AB bonds are shorter at the HF/cc-pVDZ than at the HF/6-31G(d,p) level of theory where this difference is kept at the correlation corrected levels.

Since the H basis sets contain the same number of primitives, the shielding of the H nucleus is described by the cc-pVDZ and the 6-31G(d,p) basis in a comparable way. When describing AH bonds, it plays a role that the exponents of the cc-pVDZ functions are smaller than those of the 6-31G(d,p) functions thus leading to a more expanded charge distribution and a deshielding of the H nucleus. This second effect dominates when describing AH bonds and, therefore, AH bonds calculated at the HF/cc-pVDZ or the corresponding correlation corrected levels are always longer than those obtained at the same level of theory with the 6-31G(d,p) basis. Hence, in some cases (e.g. AH_n molecules) MP2/6-31G(d,p) will lead to surprisingly accurate geometries because of error cancellation, however this advantage is lost as soon as heavy atom bonds have to be described.

12. *Use of cc-pVmZ basis sets augmented by diffuse basis functions.* Often it is recommended to add diffuse functions when describing hetero atoms possessing electron lone pairs. We find in agreement with previous work [31–33,72] that the use of an aug-cc-pVDZ or even an aug-cc-pVTZ basis set cannot be recommended for obtaining better geometries. The changes in the geometry upon the addition of diffuse basis functions are erratic and often in the wrong direction. It is far better to increase the cardinal number m of the basis than adding diffuse functions where computational cost are in both cases comparable or slightly larger for using the next higher cc-pVmZ basis set. Cremer and co-workers have shown that diffuse functions added to a basis set not saturated in the sp space (such as VDZ + P basis) erroneously introduce higher order correlation effects and by this lead to unpredictable results[72]. This effect is not so strong in the case of an aug-cc-pVTZ basis but still detectable. Apart from this, it can be seen that geometries calculated with the aug-cc-pVmZ basis sets converge to the same CBS limit geometries as

those calculated with the normal cc-pVmZ basis sets.

13. *Use of frozen core descriptions*. For highly correlated ab initio methods, one often reduces computational cost by freezing core electrons in the correlation treatment where the common argument is that core electrons do not influence molecular properties such as the geometry. We have investigated this effect at the MP6 level and we find that a frozen core leads to a lengthening of both AH and AB bonds by about 0.001–0.002 Å. This is unexpected because inclusion of the core electrons in correlation corrected descriptions should lead to deshielding of the nucleus and hence to longer rather than shorter bond lengths. On the other hand, correlation of core electrons will always lead to small effects and, therefore, will have hardly any consequence on the shielding/deshielding of a nucleus. It seems to be more important that the number of correlating electrons is increased when unfreezing the core. In this way, the number of possible disconnected four-, five-, and six-electron combinations becomes considerably larger and by this the coupling corrections for two- and three-electron correlation effects. Coupling effects always shorten the bond, which explains the difference between frozen and unfrozen geometry optimizations.

4. Conclusions

Using a generally applicable method for predicting CBS limit geometries (for exceptions, see point 9) at higher orders of MPn and CC theory, results have been obtained for 17 small molecules that shed light on the accuracy obtainable with correlation corrected ab initio methods nowadays available. Results of this investigation are summarized in the headings of discussion points (1)– (13), which clearly show that there exist no generally applicable method that leads to r_e geometries accurate better than ± 0.001 Å or $\pm 0.01°$. However, for AH bonds this accuracy is already reached at the MP6/CBS level.

As for the questions posed in the introduction, we are now able to give some specific answers.

(1) *Does MP6 lead to reliable molecular geo-metries or do we have to apply even higher orders of MPn theory for this purpose?* Clearly, MP6 is not capable of describing AB bonds accurately. In the case of a CO, OF or FF bond, there are more electron pairs (up to 7) close to the bonding region than MP6 can correlate either in the connected (probably not important) or disconnected fashion (probably important). We predict that accurate AB bond lengths will be obtained at MP8 or MP10, however these methods are outside present or future computational possibilities.

(2) *Which correlation effects have to be covered to obtain useful geometries?* It is generally known that three-electron correlation effects are important to get reasonable geometries. Our present and previous work [31–33,105] also shows that connected four-electron correlation effects play a role for bonds involving electronegative atoms or for multiple bonds. The rule of thumb is that all electron pairs in the vicinity of a bond must be correlated in at least a disconnected fashion where sufficient coupling effects have to be included. This requirement can only be fulfilled by CC methods.

(3 + 4) *How do MP and CC methods perform in comparison? Is CCSD(T) sufficient to determine accurate geometries?* All available MP or CC methods have their deficiencies. AB bond lengths come out too long at the MP and too short at the CC level, in particular the often praised CCSD(T) method is less accurate in the CBS limit than previous calculations with truncated basis sets suggested. However, in general it is always advisable to invest in CCSD(T) (definitely not CCD or CCSD) and to obtain a CCSD(T)/cc-pVTZ or CCSD(T)/cc-pVQZ geometry, which both provide accurate geometrical parameters due to error cancellation.

(5) *What is the most economic way to calculate reliable geometries?* We suggest the following sequence of calculations to obtain exceedingly more accurate r_e geometries: (a) MP2/cc-pVTZ; (b) CCSD(T)/cc-pVTZ; (c) extrapolation to CCSD(T)/CBS geometries followed by averaging of CCSD(T)/cc-pVTZ and CCSD(T)/cc-pVQZ geometries.

We are presently investigating how CBS limit geometries obtained in this work can be extrapolated to the CC or full CI limit using methods such as CCSDT or the recently developed QCISDTQ method[105].

Acknowledgements

References

[1] M.D. Harmony, V.W. Laurie, R.L. Kuczkowski, R.H. Schwendeman, D.A. Ramsay, F.J. Lovas, W.J. Lafferty, A.G. Maki, J. Phys. Chem. Ref. Data 8 (1979) 619.

[2] G. Herzberg, Spectra of Diatomic Molecules, Van Nostrand, Princeton, NJ, 1950.

[3] K.P. Huber, G. Herzberg, Molecular Spectra and Molecular Structure, Constants of Diatomic Molecules, Van Nostrand, New York, 1979.

[4] Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology, New Series, vol. II/7: Structure Data of Free Polyatomic Molecules, K.-H. Hellwege, A.M. Hellwege (Eds.); Vol II/15: Supplement, K.-H. Hellwege, A.M. Hellwege (Eds.); vol. II/21: Supplement, K. Kuchitsu (Ed.), Springer, Berlin, Heidelberg, 1976, 1987, 1992.

[5] MOGADOC, Molecular Gasphase Documentation, University Ulm, 1999 described in J. Vogt. Struct. Chem. 3 (1992) 147.

[6] J. Vogt, J. Mol. Spectrosc. 155 (1992) 413.

[7] M.D. Harmony, Acc. Chem. Res. 25 (1992) 321.

[8] A. Domenicano, I. Hargittai (Eds.), Accurate Molecular Structures Oxford University Press, Oxford, 1992.

[9] F.H. Allen, O. Kennard, D.G. Watson, L. Brammer, A.G. Orpen, R. Taylor, J. Chem. Soc. Perkin Trans. II (1997) S1.

[10] F.H. Allen, in Ref. [8], p. 355.

[11] K. Kuchitsu, S.J. Cyvin, in: S.J. Cyvin (Ed.), Molecular Structure and Vibrations, Elsevier, Amsterdam, 1972, pp. 183–211.

[12] W.J. Hehre, L. Radom, P.V.R. Schleyer, J.A. Pople, Ab Initio Molecular Orbital Theory, Wiley, New York, 1986.

[13] D. Cremer, in: P.V.R Schleyer, N.L. Allinger, T. Clark, J. Gasteiger, P.A. Kollman, H.F. Schaefer III, P.R. Schreiner (Eds.), Encyclopedia of Computational Chemistry, vol. 3, Wiley, Chichester, UK, 1998, p. 1706 (For a recent review see).

[14] J. Gauss, D. Cremer, Adv. Quant. Chem. 23 (1992) 205.

[15] P.J. DeFrees, B.A. Levi, S.K. Pollack, W.J. Hehre, J.S. Binkley, J.A. Pople, J. Am. Chem. Soc. 113 (1991) 1507.

[16] B.G. Johnson, P.M. Gill, J.A. Pople, J. Chem. Phys. 97 (1992) 7846.

[17] G. Scuseria, T.L. Lee, in: S.R. Langhoff (Ed.), Quantum Mechanical Electronic Structure Calculations with Chemical Accurary, Kluwer Academic Publishers, Dortrecht, 1995, p. 47.

[18] H.F. Schaefer III, J.R. Thomas, Y. Yamaguchi, B.J. Deleeuw, G. Vacek, in: D.R. Yarkony (Ed.), Modern Electronic Structure Theory, World Scientific, Singapore, 1995.

[19] N. Oliphant, R.J. Bartlett, J. Chem. Phys. 100 (1994) 6550.

[20] T. Helgaker, J. Gauss, P. Jørgenson, J. Olsson, J. Chem. Phys. 106 (1997) 6430.

[21] H. Asger, P. Jørgenson, J. Gauss, T. Helgaker, Chem. Phys. Lett. 274 (1997) 235.

[22] J. Gauss, D. Cremer, J.F. Stanton, J. Phys. Chem. 104 (2000) 1319.

[23] J.F. Stanton, C.L. Lopreore, J. Gauss, J. Chem. Phys. 108 (1998) 7190.

[24] J. Gauss, J.F. Stanton, J. Phys. Chem., A 104 (2000) 2865.

[25] H. Koch, B. Fernandez, O. Christiansen, J. Chem. Phys. 108 (1998) 2784 (and papers cited therein).

[26] K.A. Peterson, D.E. Woon, T.H. Dunning Jr., J. Chem. Phys. 100 (1994) 7410.

[27] T. van Mourik, A.K. Wilson, T.H. Dunning Jr., Mol. Phys. 96 (1999) 529.

[28] K.A. Peterson, T.H. Dunning Jr., J. Mol. Struct. (Theochem) 400 (1997) 93.

[29] T.H. Dunning Jr., K.A. Peterson, J. Chem. Phys. 108 (1998) 4761.

[30] K.A. Peterson, T.H. Dunning Jr., J. Chem. Phys. 106 (1997) 4119.

[31] Y. He, D. Cremer, J. Phys. Chem. A 104 (2000) 7679.

[32] Y. He, D. Cremer, J. Mol. Phys. 98 (2000) 1415.

[33] Y. He, D. Cremer, Theoret. Chem. Accounts. 105 (2000) 110.

[34] J.A. Pople, M. Head-Gordon, D.J. Fox, K. Raghavachari, L.A. Curtiss, J. Chem. Phys. 90 (1989) 5622.

[35] L.A. Curtiss, K. Ragavachari, G.W. Trucks, J.A. Pople, J. Chem. Phys. 94 (1991) 7221.

[36] L.A. Curtiss, K. Raghavachari, P.C. Redfern, V. Rassolov, J.A. Pople, J. Chem. Phys. 109 (1998) 7764.

[37] P. Jensen, P.R. Bunker, J. Chem. Phys 89 (1988) 1327.

[38] H. Hollenstein, R.R. Marquardt, M. Quack, M.A. Suhm, J. Chem. Phys. 101 (1994) 3588.

[39] K. Ohno, H. Matsuura, Y. Endo, E. Hirota, J. Mol. Spectrosc. 118 (1986) 1.

[40] D.L. Gray, A.G. Robiette, Mol. Phys. 37 (1979) 1901.

[41] P. Botschwina, M. Horn, M. Matuschewski, E. Schick, P. Sebald, J. Mol. Struct. 400 (1997) 119.

[42] S. Carter, I. Mills, H.C. Handy, J. Chem. Phys. 97 (1992) 1606.

[43] R.C. Woods, Phil. Trans. Roy. Soc. London, A 324 (1988) 141.

[44] G. Winnewisser, A.G. Maki, D.R. Johnson, J. Mol. Spectrosc. 39 (1971) 149.

[45] T. Okabayashi, M. Tanimoto, J. Chem. Phys. 99 (1993) 3268.

[46] R.A. Creswell, A.G. Robiette, Mol. Phys. 36 (1978) 869.

[47] E.F. Pearson, R.A. Creswell, M. Winnewisser, G. Winnewisser, Z. Naturforsch A 31 (1976) 1394.

[48] N. Authier, N. Bagland, A. Le Floch, J. Mol. Spectrosc. 160 (1993) 590.

[49] J.F. Ogilvie, M.C.C. Ho, Chin. J. Phys. (Taipei) 31 (1993) 721.

[50] J.A. Coxon, P.G. Hajigeorgiou, Can. J. Phys. 70 (1992) 40.

[51] G. Graner, D. Rossetti, Mol. Phys. 58 (1986) 627.

[52] E.A. Cohen, H.M. Pickett, J. Mol. Spectrosc. 93 (1982) 83.

[53] W.S. Benedict, E.K. Plyler, Can. J. Phys. 35 (1957) 1235.

[54] M.L. Orlov, J.F. Ogilvie, J.W. Nibler, J. Mol. Spectrosc. 185 (1997) 128.

[55] J. Demaison, F. Hegelung, H. Buerger, J. Mol. Struct. 413 (1997) 447.

[56] M. Carlotti, J.W.C. Johns, A. Trombetti, Can. J. Phys. 52 (1974) 340.

[57] F.W. Dalby, Can. J. Phys. 36 (1958) 1336.

[58] O.L. Polansky, P. Jensen, J. Tennyson, J. Chem. Phys. 105 (1996) 6490.

[59] O.L. Polansky, P. Jensen, J. Tennyson, J. Chem. Phys. 101 (1994) 7651.

[60] C.D. Paulse, J. Tennyson, J. Mol. Spectrosc. 168 (1994) 313.

[61] A.R. Hoy, P.R. Bunker, J. Mol. Spectrosc. 74 (1979) 1.

[62] G. Pelz, K.M.T. Yamada, G. Winnewisser, J. Mol. Spectrosc. 159 (1993) 507.

[63] T. Tanaka, Y. Morino, J. Mol. Spectrosc. 33 (1970) 538.

[64] F.C. DeLucia, P. Helminger, W. Gordy, Phys. Rev., A 3 (1971) 1849.

[65] A. Edwards, B. Goods, C. Long, J. Chem. Soc. Faraday Trans II 72 (1976) 984.

[66] L. Halonen, T.-K. Ha, J. Chem. Phys. 89 (1988) 4885.

[67] T.D. Crawford, H.F. Schaefer III, in: K.B. Lipkowitz, D.B. Boyd (Eds.), Reviews in Computational Chemistry, vol. 14, VCH, Weinheim, 2000, p. 33.

[68] J. Gauss, in: P.V.R. Schleyer, N.L. Allinger, T. Clark, J. Gasteiger, P.A. Kollman, H.F. Schaefer III, P.R. Schreiner (Eds.), Encyclopedia of Computational Chemistry, vol. 1, Wiley, Chichester, UK, 1998, p. 615.

[69] R.G. Parr, W. Yang, International Series of Monographs on Chemistry 16: Density-Functional Theory of Atoms and Molecules, Oxford University Press, New York, 1989.

[70] Y. He, J. Gräfenstein, E. Kraka, D. Cremer, Mol. Phys 98 (2000) 1639.

[71] C. Møller, M.S. Plesset, Phys. Rev. 46 (1934) 618.

[72] B. Forsberg, Z. He, Y. He, D. Cremer, Int. J. Quant. Chem. 76 (2000) 306.

[73] Z. He, D. Cremer, Int. J. Quant. Chem. 59 (1996) 15.

[74] Z. He, D. Cremer, Int. J. Quant. Chem. 59 (1996) 31.

[75] Z. He, D. Cremer, Int. J. Quant. Chem. 59 (1996) 57.

[76] Z. He, D. Cremer, Int. J. Quant. Chem. 59 (1996) 71.

[77] D. Cremer, Z. He, J. Phys. Chem. 100 (1996) 6173.

[78] D. Cremer, Z. He, J. Mol. Struct. (Theochem) 7 (1997) 398.

[79] J.A. Pople, R. Krishnan, H.B. Schlegel, J.S. Binkley, Int. J. Quant. Chem. 14 (1978) 545.

[80] R.J. Bartlett, G.D. Purvis III, Int. J. Quant. Chem. 14 (1978) 561.

[81] G.D. Purvis III, R.J. Bartlett, J. Chem. Phys. 76 (1982) 1910.

[82] J. Noga, R.J. Bartlett, J. Chem. Phys. 86 (1987) 7041.

[83] J. Noga, R.J. Bartlett, J. Chem. Phys. 89 (1988) 340(E).

[84] K. Raghavachari, G.W. Trucks, J.A. Pople, M. Head-Gordon, Chem. Phys. Lett. 157 (1989) 479.

[85] Z. He, D. Cremer, Int. J. Quant. Chem. Symp. 25 (1991) 43.

[86] Z. He, D. Cremer, Theor. Chim. Acta 85 (1993) 305.

[87] E. Kraka, J. Gräfenstein, Y. He, J. Gauss, F. Reichel, L. Olsson, Z. Konkoli, Z. He, D. Cremer, Program package COLOGNE, Göteborg University, Göteborg, 2000.

[88] R. Fletcher, M.J.D. Powell, Comput. J. 6 (1963) 163 (W. Davidon, Argonne National Lab. Report, ANL-5990).

[89] J.F. Stanton, J. Gauss, J.D. Watts, W.J. Lauderdale, R.J. Bartlett, ACES II, Quantum Theory Project, University of Florida, 1992.

[90] J.F. Stanton, J. Gauss, J.D. Watts, W.J. Lauderdale, R.J. Bartlett, Int. J. Quant. Chem. Symp. 26 (1992) 879.

[91] D. Cremer, unpublished results.

[92] T.H. Dunning Jr., J. Chem. Phys. 99 (1989) 1007.

[93] R.A. Kendall, T.H. Dunning Jr., R.J. Harrison, J. Chem. Phys. 96 (1992) 6796.

[94] K.A. Peterson, R.A. Kendall, T.H. Dunning Jr., J. Chem. Phys. 99 (1993) 1930.

[95] K.A. Peterson, R.A. Kendall, T.H. Dunning Jr., J. Chem. Phys. 99 (1993) 9790.

[96] K.A. Wilson, T.H. Dunning Jr., J. Chem. Phys. 106 (1997) 8718.

[97] A. Wilson, T. van Mourik, T.H. Dunning Jr., J. Mol. Struct. (Theochem) 388 (1996) 339.

[98] P.C. Hariharan, J.A. Pople, Theor. Chim. Acta 28 (1973) 213.

[99] J.R.L. Martin, R.R. Taylor, Chem. Phys. Lett. 225 (1994) 473.

[100] J.R.L. Martin, T.J. Lee, Chem. Phys. Lett. 258 (1996) 136.

[101] K.L. Balk, J. Gauss, P. Jørgensen, J. Olsen, T. Helgaker, J.F. Stanton, J. Chem. Phys. (2001) (in press).

[102] D. Cremer, in: S. Patai (Ed.), The Chemistry of Peroxides, Wiley, New York, 1983.

[103] E. Kraka, Y. He, D. Cremer, J. Phys. Chem. A (2001) (in press).

[104] J.-M. Flaud, C. Camy-Peyret, J.W.C. Johns, B. Carli, J. Chem. Phys. 91 (1989) 1504.

[105] Y. He, Z. He, D. Cremer, Chem. Phys. Lett. 317 (2000) 535.