

The accurate determination of molecular equilibrium structures

Keld L. Bak

UNI-C, Olof Palmes Allé 38, DK-8200 Århus N, Denmark

Jürgen Gauss

Institut für Physikalische Chemie, Universität Mainz, D-55099 Mainz, Germany

Poul Jørgensen and Jeppe Olsen

Department of Chemistry, Aarhus University, DK-8000 Aarhus C, Denmark

Trygve Helgaker

Department of Chemistry, University of Oslo, P.O. Box 1033, Blindern, N-0315 Oslo, Norway

John F. Stanton

Institute for Theoretical Chemistry, Department of Chemistry and Biochemistry,
The University of Texas at Austin, Austin, Texas 78712

(Received 8 September 2000; accepted 30 January 2001)

Equilibrium structures have been determined for 19 molecules using least-squares fits involving rotational constants from experiment and vibrational corrections from high-level electronic-structure calculations. Equilibrium structures obtained by this procedure have a uniformly high quality. Indeed, the accuracy of the results reported here likely surpasses that reported in most experimental determinations. In addition, the accuracy of equilibrium structures obtained by energy minimization has been calibrated for the following standard models of *ab initio* theory: Hartree–Fock, MP2, CCSD, and CCSD(T). In accordance with previous observations, CCSD(T) is significantly more accurate than the other models; the mean and maximum absolute errors for bond distances of the 19 molecules are 0.09 and 0.59 pm, respectively, in CCSD(T)/cc-pCVQZ calculations. The maximum error is obtained for R_{OO} in H_2O_2 and is so large compared with the mean absolute error that an experimental reinvestigation of this molecule is warranted. © 2001 American Institute of Physics. [DOI: 10.1063/1.1357225]

I. INTRODUCTION

In the development of *ab initio* electronic structure methods, molecular equilibrium structures have been one of the properties receiving the most attention.^{1–3} The accuracy with which equilibrium structures can be calculated from first-principles quantum mechanics has improved steadily over the years. Indeed, electronic structure calculations now represent a viable alternative to experiment for determining accurate molecular equilibrium structures.

Hierarchies of basis sets and wave-function models have been introduced that give a systematically improved accuracy in calculations of equilibrium structures.⁴ In a previous study, some of us examined the accuracy obtained by the following standard hierarchy of *ab initio* models.² Hartree–Fock theory, second-order Møller–Plesset theory (MP2), coupled-cluster singles-and-doubles (CCSD) theory,⁵ and CCSD theory with perturbative treatment of triples (CCSD(T)).⁶ In accordance with previous investigations, we found that the CCSD(T) method gave structures of high quality.² In particular, for 19 small molecules where the equilibrium structure is reasonably well-known from experiment, CCSD(T) gives a mean intrinsic error of about 0.2 pm. Clearly, these error estimates assume that the experimentally determined equilibrium structures are correct; more reliable calibration of the accuracy of CCSD(T) could be obtained if the accuracy of the experimental structures is also improved.

The present paper reports a detailed investigation of the uncertainties and inconsistencies associated with both the calculated CCSD(T) structures and the experimental structures used as reference data in Ref. 2.

Experimental equilibrium structures are typically obtained from least-squares fits involving experimental rotational constants for the vibrational ground state and the corresponding vibration-rotation interaction constants.⁷ Whereas accurate experimental rotational constants are available for most small molecules, accurate vibration-rotation interaction constants are more scarce. In principle, they can be obtained from rotationally resolved vibrational spectra. However, these constants are difficult to obtain for larger molecules and assumptions are often made about the anharmonicity of the force field, leading to uncertainties in the experimental equilibrium structures. As has been amply documented in the literature, accurate vibrational corrections to rotational constants can be obtained from quantum chemical calculations and subsequently used for the determination of highly accurate structures.^{8–13} In addition, recently developed analytic derivative techniques—in particular, those for CCSD(T) second derivatives¹¹—provide efficient schemes for the precise calculation of vibration-rotation interaction constants, enabling the routine determination of the required vibrational contributions.^{11–13} In this paper, we use calculated sums of vibration-rotation interaction constants together with experimental rotational constants to obtain a set of empirical equi-

librium structures. This set of geometrical parameters has a uniformly high accuracy and is used to judge the quality of previously determined experimental values. This new set of parameters is also used to reassess the accuracy of equilibrium structures obtained from standard *ab initio* calculations.

The calculations previously used to establish an intrinsic CCSD(T) error of about 0.2 pm were carried out in basis sets constructed to describe valence correlation, although all electrons were included in the correlation treatment. Since the valence basis sets are not sufficiently flexible to adequately describe core correlation, the question naturally arises as to what extent this has affected the obtained accuracy. Here we address this question by carrying out all-electron calculations using the corresponding core-valence basis sets.

In the previous investigation, we considered 19 closed-shell systems that comprise a representative set of molecules containing hydrogen and first-row atoms joined by single, double and triple bonds. The systems chosen for that study had a varying degree of single-configurational dominance;⁴ a single molecule (ozone) has relatively significant multiconfiguration character. We are in the process of building a database which serves to calibrate the accuracy of a range of theoretical approaches for various molecular properties that range from equilibrium structures to atomization energies. So that statistics are not skewed by including difficult molecules that exhibit significant multiconfigurational character, ozone has been excluded and H₂ has been included because of its role in a wide variety of important chemical reactions. The resulting set of 19 molecules is precisely that which is involved in the study presented here, and will be used in other calibration work that is currently being carried out in our groups.

In this paper, we report a statistical analysis of the accuracy of calculated equilibrium bond lengths and bond angles for the nineteen molecule test suite.

II. COMPUTATIONAL DETAILS

We consider the determination of equilibrium structures for the nineteen molecules H₂, CH₂, CH₄, H₂O, NH₃, HF, N₂, F₂, CO, HCN, HNC, C₂H₂, C₂H₄, CH₂O, N₂H₂, HOF, H₂O₂, CO₂, and HNO. As mentioned above, they are the same as those of Ref. 2 except that ozone has been replaced by H₂. For these molecules (except for H₂ and the nonrigid H₂O₂ molecule), equilibrium structures have been obtained using experimental rotational constants and sums of calculated vibration-rotation interaction constants. The vibrational corrections to the rotational constants are usually expressed as¹⁴

$$B_v = B_e - \sum_r \alpha_r^B (v_r + \frac{1}{2}) + \dots, \quad (1)$$

i.e.,

$$B_0 = B_e - \frac{1}{2} \sum_r \alpha_r^B, \quad (2)$$

where B_e is inversely proportional to a principal moment of inertia of the equilibrium structure and B_v the rotational constant in the vibrational state v . The sum is taken over the

TABLE I. Wave function/basis set convergence of the vibrational corrections to the rotational constant (in cm⁻¹) and of the empirical structure (in pm) for hydrogen fluoride.

Method/basis	Vibrational corrections	Empirical structure
CCSD(T)/cc-pVDZ	0.4028	91.666
CCSD(T)/cc-pVTZ	0.3857	91.704
CCSD(T)/cc-pVQZ	0.3935	91.687
CCSD(T)/aug-cc-pVDZ	0.4186	92.560
CCSD(T)/aug-cc-pVTZ	0.3876	91.699
CCSD(T)/aug-cc-pVQZ	0.3908	91.693
CCSD(T)/cc-pCVDZ	0.4025	91.667
CCSD(T)/cc-pCVTZ	0.3856	91.704
CCSD(T)/cc-pCVQZ	0.3939	91.686
MP2/cc-pVQZ	0.3875	91.700
CCSD/cc-pVQZ	0.3893	91.698
Experimental value ^a	0.3961	91.680

^aD. U. Webb and K. N. Rao, J. Mol. Spectrosc. **28**, 121 (1968).

normal modes r , α_r^B is the vibration-rotation interaction constant and v_r is the corresponding quantum number. Accurate experimental rotation constants for the vibrational ground state B_0 are available for most small molecules and for many of the corresponding isotopomers. We use these rotational constants together with sums of vibration-rotation interaction constants calculated at the all-electron CCSD(T)/cc-pVQZ (Ref. 14) level [except for C₂H₄, where CCSD(T)/cc-pVTZ is used] to determine B_e and subsequently the equilibrium structure by least-squares fitting. Details of which rotational constants and which isotopomers are used will be given in a future publication.¹⁵ The vibration-rotation interaction constants are obtained using second-order perturbation theory, starting from the rigid-rotor harmonic-oscillator Hamiltonian. Details can be found in Ref. 16. The vibration-rotation interaction constants have been calculated from quadratic and cubic force fields using a local version of the ACESII program.¹⁷ The wave function/basis set convergence of vibrational corrections to the equilibrium structures is illustrated in Table I through calculations on the HF molecule.

Equilibrium structures have also been obtained by direct energy minimization using the HF, MP2, CCSD, and CCSD(T) models and the cc-pCVDZ, cc-pCVTZ, and cc-pCVQZ core-valence sets.¹⁸ Comparisons have also been made with calculations using the corresponding augmented valence sets aug-cc-pVxZ (Ref. 19) and with the previous calculations using the valence sets cc-pVxZ.¹⁴ The calculations reported here are all-electron calculations (no orbitals are excluded from the correlation procedure) and have been carried out using the ACESII program.¹⁷

III. EXPERIMENTAL EQUILIBRIUM STRUCTURES

Experimental equilibrium bond lengths are given in Table II and bond angles in Table III; both sets of values are marked with Exp. Only independent bond angles are listed in Table III; where angles are dependent, we have used the smallest bond angle.

The experimental equilibrium structures have varying accuracy. The diatomic structures are given with three significant digits after the decimal point and the uncertainty is mostly in the last digit. For larger systems, the experimental

TABLE II. Equilibrium bond lengths R_e in pm.

Molecule	Bond	Expt.	Emp.	Emp. - Expt.	CCSD(T)/ cc-pCVQZ
H ₂	R_{HH}	74.144 ^a	74.19
HF	R_{FH}	91.680(8) ^a	91.69	0.01	91.58
H ₂ O	R_{OH}	95.72 ^b	95.75	0.03	95.71
HOF	R_{OH}	96.57(16) ^c	96.78	0.21	96.57
H ₂ O ₂	R_{OH}	96.7 ^d	96.19
HNC	R_{NH}	99.40(8) ^e	99.49	0.09	99.53
NH ₃	R_{NN}	101.1(6) ^f	101.16	0.06	101.12
N ₂ H ₂	R_{NH}	102.9(1) ^g	102.86	0.04	102.84
HNO	R_{NH}	...	105.17	...	105.24
C ₂ H ₂	R_{CH}	106.215(17) ^h	106.13	-0.09	106.21
HCN	R_{CH}	106.501(8) ⁱ	106.53	0.03	106.55
C ₂ H ₄	R_{CH}	108.1(2) ^j	108.07	0.03	108.09
CH ₄	R_{CH}	108.58(10) ^k	108.59 ⁿ	0.01	108.64
N ₂	R_{NN}	109.768(5) ^a	109.77	0.00	109.81
CH ₂ O	R_{CH}	110.05(20) ^j	110.07	0.02	110.08
CH ₂	R_{CH}	110.7(2) ^l	110.63	0.07	110.68
CO	R_{CO}	112.832 ^a	112.84	0.01	112.89
HCN	R_{CN}	115.324(2) ^j	115.34	0.02	115.38
CO ₂	R_{CO}	115.995 ^m	116.01	0.01	116.04
HNC	R_{CN}	116.89(2) ^e	116.87	-0.02	116.93
C ₂ H ₂	R_{CC}	120.257(9) ^h	120.37	0.11	120.37
CH ₂ O	R_{CO}	120.33(10) ^j	120.47	0.14	120.43
HNO	R_{NO}	...	120.86	...	120.85
N ₂ H ₂	R_{NN}	124.7(1) ^g	124.57	-0.13	124.67
C ₂ H ₄	R_{CC}	133.4(2) ^j	133.07	-0.33	133.12
F ₂	R_{FF}	141.193 ^a	141.24	0.05	141.13
HOF	R_{FO}	143.50(31) ^c	143.44	-0.06	143.26
H ₂ O ₂	R_{OO}	145.56 ^d	144.97

^aK. P. Huber and G. H. Herzberg, *Constants of Diatomic Molecules* (Van Nostrand-Reinhold, New York, 1979).

^bA. R. Hoy, I. M. Mills, and G. Strey, *Mol. Phys.* **24**, 1265 (1972); see also A. R. Hoy and P. R. Bunker, *J. Mol. Spectrosc.* **74**, 1 (1979) $R_{OH}=95.78$.

^cL. Halonen and T. K. Ha, *J. Chem. Phys.* **89**, 4885 (1988).

^dG. Pelz, K. M. T. Yamada, and G. Winnewisser, *J. Mol. Spectrosc.* **159**, 507 (1993).

^eR. A. Creswell and A. G. Robiette, *Mol. Phys.* **36**, 869 (1978).

^fJ. L. Duncan and I. M. Mills, *Spectrosc. Chem. Acta* **20**, 523 (1964); see also W. S. Benedict and E. K. Plyler, *Can. J. Phys.* **35**, 1235 (1957). $R_{NH}=101.24$, $\theta_{HNH}=106.7$.

^gJ. Demaison, F. Hegelund, and H. Bürger, *J. Mol. Struct.* **413**, 447 (1997).

^hA. Baldacci, S. Ghersetti, S. C. Hurlock, and K. N. Rao, *J. Mol. Spectrosc.* **59**, 116 (1976).

ⁱS. Carter, I. M. Mills, and N. C. Handy, *J. Chem. Phys.* **97**, 1606 (1992); see also G. Winnewisser, A. G. Maki, and D. R. Johnson, *J. Mol. Spectrosc.* **39**, 149 (1971) $R_{CN}=115.321(5)$, $R_{CH}=106.549(24)$.

^jJ. L. Duncan, *Mol. Phys.* **28**, 1177 (1974); see also K. Yamada, T. Nakagawa, K. Kuchitsu, and Y. Morino, *J. Mol. Spectrosc.* **38**, 70 (1971) $R_{CO}=120.3(3)$, $R_{CH}=109.9(9)$, $\theta_{HCH}=116.5(12)$.

^kD. L. Gray and A. G. Robiette, *Mol. Phys.* **37**, 1901 (1979); see also L. S. Bartell and K. Kuchitsu, *J. Chem. Phys.* **68**, 1213 (1978) $R_{CH}=108.62(24)$; E. Hirota, *J. Mol. Spectrosc.* **77**, 213 (1979) $R_{CH}=108.70(7)$.

^lH. Petek, D. Nesbitt, D. C. Darwin, P. R. Ogilby, C. B. Moore, and D. A. Ramsay, *J. Chem. Phys.* **91**, 6566 (1989).

^mG. Graner, C. Rossetti, and D. Bailly, *Mol. Phys.* **58**, 627 (1986).

ⁿSee, J. F. Stanton, *Mol. Phys.* **97**, 841 (1999).

equilibrium structures are generally given with fewer significant digits and have larger uncertainties. Many of the experimental structures have been obtained from least-squares fits using experimental rotational and vibration-rotation interaction constants. While accurate experimental rotational constants are usually available for the vibrational ground state, the quality of the vibration-rotation interaction constants is more dependent on the considered system, and is frequently complicated by the presence of Coriolis resonances, although the latter problem is avoided when the *sum* of the rotation-vibration constants is explicitly calculated.²⁰ The difficulties involved in obtaining experimental vibration-rotation inter-

TABLE III. Equilibrium bond angles (in degs). See Table I for footnotes.

Molecule	Bond angle	Expt.	Emp.	CCSD(T)/cc-pCVQZ
H ₂ O	HOH	104.52 ^b	104.51	104.22
HOF	HOF	97.54(50) ^c	97.94	97.85
HNO	HNO	...	108.27	108.08
H ₂ O ₂	HOO	102.32 ^d	...	99.99
NH ₃	HNH	106.7 ^f	107.25	106.36
N ₂ H ₂	HNN	106.3(2) ^g	106.36	106.17
C ₂ H ₄	HCH	117.37(33) ^j	117.14	117.10
CH ₂ O	HCH	116.30(25) ^j	116.74	116.44
CH ₂	HCN	102.4(4) ^l	102.44	102.03

TABLE IV. Statistical measures of errors in calculated bond lengths.

	cc-pVxZ			aug-cc-pVxZ			cc-pCVxZ		
	D	T	Q	D	T	Q	D	T	Q
HF:									
$\bar{\Delta}$	-1.65	-2.49	-2.59	-1.79	-2.49	-2.59	-1.70	-2.52	-2.60
$\bar{\Delta}_{\text{abs}}$	1.80	2.49	2.59	1.88	2.49	2.59	1.82	2.52	2.60
Δ_{max}	6.48	8.33	8.49	7.42	8.42	8.57	6.47	8.40	8.51
Δ_{std}	1.95	1.98	2.03	2.04	2.00	2.05	1.93	2.00	2.03
MP2:									
$\bar{\Delta}$	1.34	-0.13	-0.23	1.49	-0.08	-0.18	1.28	0.00	-0.18
$\bar{\Delta}_{\text{abs}}$	1.34	0.55	0.51	1.49	0.51	0.48	1.28	0.46	0.46
Δ_{max}	3.16	1.66	1.71	3.34	1.47	1.76	3.07	1.68	1.70
Δ_{std}	0.74	0.67	0.61	0.73	0.61	0.61	0.70	0.64	0.59
CCSD:									
$\bar{\Delta}$	1.14	-0.56	-0.72	1.13	-0.56	-0.69	1.06	-0.43	-0.67
$\bar{\Delta}_{\text{abs}}$	1.16	0.57	0.72	1.13	0.58	0.69	1.08	0.45	0.67
Δ_{max}	2.05	2.01	2.44	2.02	2.06	2.50	2.02	2.09	2.45
Δ_{std}	0.63	0.52	0.63	0.48	0.54	0.66	0.61	0.57	0.66
CCSD(T):									
$\bar{\Delta}$	1.68	0.02	-0.10	1.73	0.05	-0.06	1.61	0.17	-0.04
$\bar{\Delta}_{\text{abs}}$	1.68	0.20	0.13	1.73	0.19	0.10	1.61	0.22	0.09
Δ_{max}	4.51	0.45	0.61	3.74	0.48	0.61	4.42	0.49	0.59
Δ_{std}	0.80	0.23	0.17	0.71	0.22	0.16	0.78	0.18	0.16

action constants generally increase with the size of the system and, for the larger systems, experimental information is often supplemented with empirical assumptions or with data obtained from *ab initio* electronic structure calculations. This approach leads to less accurate experimental structures, as is reflected in the absolute accuracies and quoted uncertainties associated with the experimental equilibrium bond lengths and bond angles given in Tables II and III for increasing size of the molecules.

We have calculated sums of the vibration-rotation interaction constants at the all-electron CCSD(T)/cc-pVQZ level and used these together with experimental rotational constants for the vibrational ground state to obtain the empirical (Emp.) bond lengths and bond angles in Tables II and III. Detailed information for the experimental rotational constants and the isotopomers that have been used to obtain the empirical structures will be given in a future publication.¹⁵ In Table II, we have also given the difference between the empirical and experimental (Emp. - Expt.) bond distances. In general, the empirical bond distances are in close agreement with the experimental ones—the differences are mostly of the same magnitude or smaller than the uncertainty in the experimental structures, in particular for the smaller molecules and for the diatomics, where the experimental bond distances have been determined with high precision.

The largest difference between empirical and experimental bond distances is -0.33 pm and occurs for the CC distance in C_2H_4 . In view of the assumptions made about the force field to obtain the experimental structure of this molecule,²¹ we believe that the empirical bond distance is the most trustworthy, noting that the discrepancy between theory and experiment for C_2H_4 has already been pointed out by Martin *et al.*²² Since the experimental rotational constants

have been accurately determined for all molecules considered here and since the calculated vibration-rotation interaction constants should have comparable accuracy for all molecules (given that all are relatively well-described by a single configuration), the empirical equilibrium structures should have a relatively uniform level of accuracy. In general, therefore, for the larger molecules, the empirical equilibrium structures should be more accurate than the experimental ones. Consequently, it is more appropriate to use this set than the experimental one for investigating the accuracy of equilibrium structures obtained by energy minimisations using standard *ab initio* models.

For H_2O_2 where experimental rotational constants are not yet available, we have not determined an empirical equilibrium structure, but will address this in future work.²³

IV. STATISTICAL ANALYSIS OF CALCULATED EQUILIBRIUM STRUCTURES

We now report a statistical analysis of the accuracy obtained for bond lengths and bond angles in calculations for the nineteen molecule test suite using core-valence basis sets. We consider the Hartree–Fock, MP2, CCSD, and CCSD(T) models with the cc-pCVDZ, cc-pCVTZ, and cc-pCVQZ basis sets. The calculated bond lengths and bond angles are compared with the empirical ones of Tables II and III, supplemented by accepted equilibrium structures for H_2 and H_2O_2 . In Tables IV and V, we give the mean error $\bar{\Delta}$, the standard deviation in the error Δ_{std} , the mean absolute error $\bar{\Delta}_{\text{abs}}$, and the maximum absolute error Δ_{max} for the calculated bond lengths and bond angles. In Fig. 1, the errors are plotted for the bond lengths; in Fig. 2, normal distributions of the errors are displayed for the bond lengths.

TABLE V. Statistical measures of errors in calculated bond angles using the cc-pCVxZ basis sets.

	Molecules from Table II			H ₂ O ₂ and NH ₃ excluded		
	D	T	Q	D	T	Q
HF:						
$\bar{\Delta}$	0.15	0.85	0.99	0.37	0.95	1.07
$\bar{\Delta}_{\text{abs}}$	0.78	1.13	1.25	0.81	1.32	1.41
Δ_{max}	2.00	2.76	2.84	2.00	2.76	2.84
Δ_{std}	1.03	1.09	1.11	1.01	1.23	1.27
MP2:						
$\bar{\Delta}$	-1.76	-0.83	-0.53	-1.29	-0.49	-0.24
$\bar{\Delta}_{\text{abs}}$	1.79	0.86	0.59	1.32	0.53	0.30
Δ_{max}	3.56	2.90	2.56	2.60	0.88	0.52
Δ_{std}	1.20	0.87	0.80	0.85	0.38	0.27
CCSD:						
$\bar{\Delta}$	-1.53	-0.53	-0.24	-1.10	-0.24	0.01
$\bar{\Delta}_{\text{abs}}$	1.53	0.64	0.45	1.10	0.38	0.27
Δ_{max}	3.33	1.94	1.58	2.31	0.73	0.69
Δ_{std}	1.12	0.70	0.63	0.80	0.41	0.37
CCSD(T):						
$\bar{\Delta}$	-1.81	-0.83	-0.52	-1.31	-0.48	-0.21
$\bar{\Delta}_{\text{abs}}$	1.81	0.83	0.52	1.31	0.48	0.21
Δ_{max}	3.66	2.68	2.33	2.58	0.83	0.41
Δ_{std}	1.22	0.80	0.72	0.81	0.26	0.13

In Table IV, we have also listed the errors obtained for bond lengths using the valence cc-pVDZ, cc-pVTZ, and cc-pVQZ basis sets and, for comparison, the errors obtained with the corresponding augmented aug-cc-pVxZ basis sets. In Ref. 2, a statistical investigation of bond lengths similar to that described above was performed, the differences being

(1) H₂ has replaced O₃ in the set of test molecules, and (2) only cc-pVxZ basis sets were used in the previous study, and (3) experimental rather than empirical results of Table II were used for reference. For the Hartree–Fock, MP2, and CCSD models, the statistics are only slightly affected by these differences, the changes being small compared with the intrinsic errors of the models. By contrast, for the most accurate correlated model CCSD(T), it turns out to be important to use reference data of high and uniform quality as discussed in the following. For the effect of using valence rather than core-valence basis sets in all-electron calculations, see Sec. V.

Whereas the DZ basis is too small to be useful for correlated studies [particularly at the CCSD(T) level], the TZ and QZ basis sets both give results of good quality (see Table IV). In the previous study, no significant improvement was found at the CCSD(T) level in going from the TZ basis to the QZ basis, the mean absolute errors being 0.23 pm and 0.22 pm, respectively.² By contrast, the analysis reported here with improved standards of calibration shows that there is indeed a significant difference in accuracy between these basis sets at the CCSD(T) level. More specifically, the mean absolute errors are reduced from 0.22 to 0.09 pm in going from cc-pCVTZ to cc-pCVQZ; 0.20 to 0.13 from cc-pVTZ to cc-pVQZ; and 0.19 to 0.10 from aug-cc-pVTZ to aug-cc-pVQZ. Collectively, these results suggest that considerable improvement in accuracy is obtained in QZ calculations relative to those carried out at the TZ level. In conclusion, the DZ sets are only useful for exploratory calculations at the Hartree–Fock level, while correlated calculations require at least TZ basis sets; for the highly accurate CCSD(T) model, QZ sets are recommended.

The trends discussed above for bond lengths are also

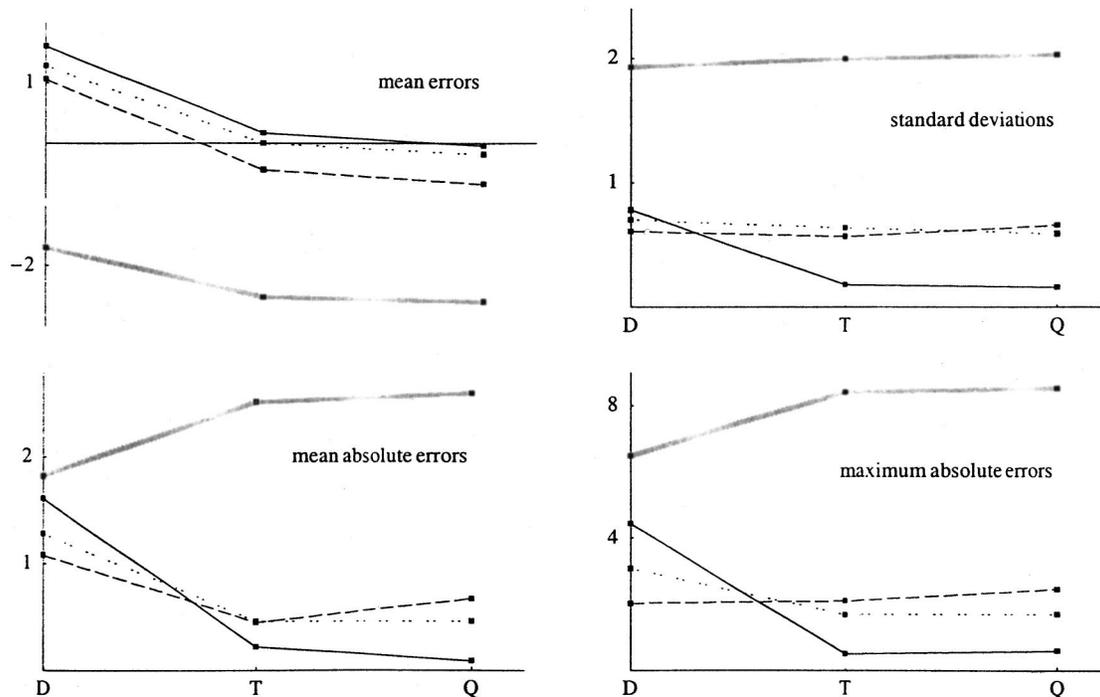


FIG. 1. Errors relative to the empirical bond lengths (pm). We have used gray lines for Hartree–Fock, dashed lines for CCSD, dotted lines for MP2, and solid lines for CCSD(T).

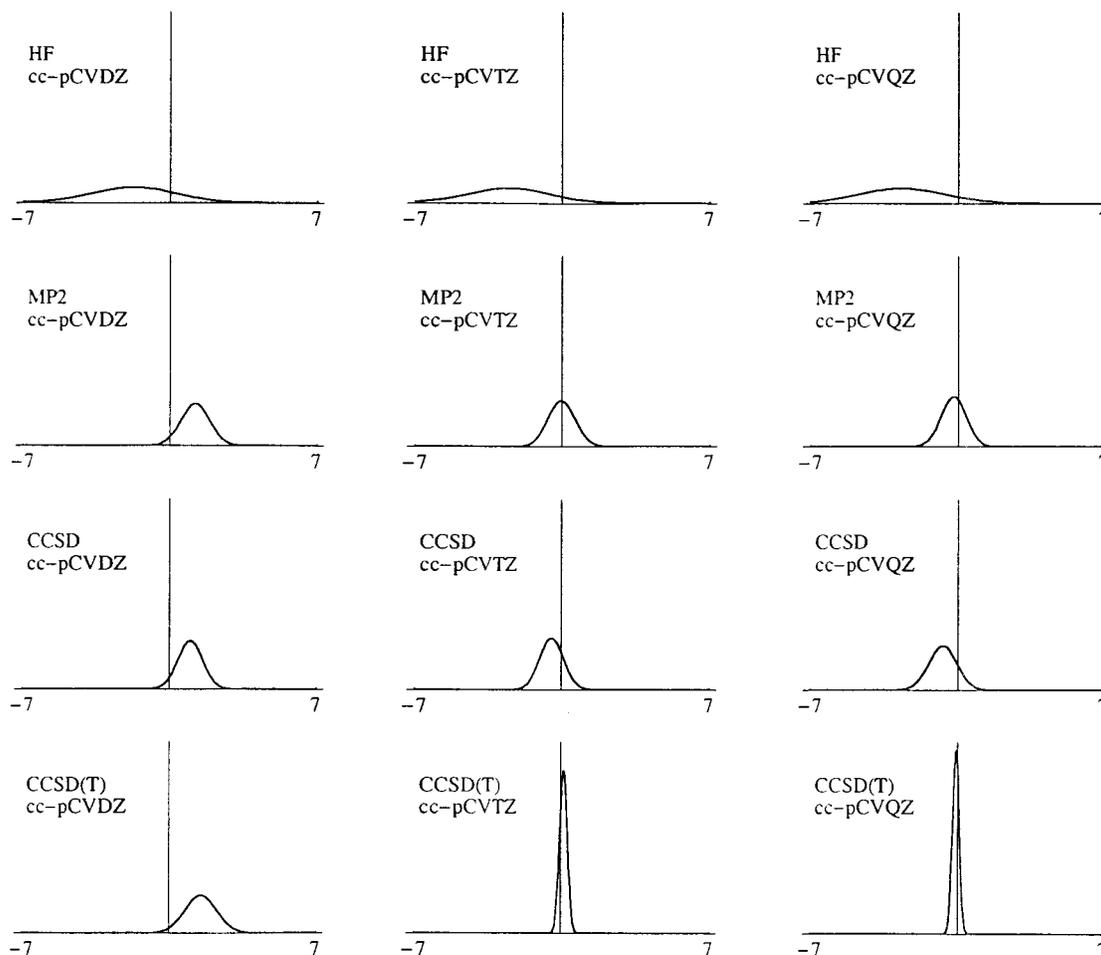


FIG. 2. Normal distributions of errors in the calculated bond distances (pm). For ease of comparison, all distributions have been normalized to one and plotted against the same horizontal and vertical scales.

found for bond angles (see Table V). However, only nine angles are considered in the statistical analysis. This is a rather small sample, especially considering the large uncertainties associated with the experimental bond angles of the nonrigid molecules H_2O_2 and NH_3 . For H_2O_2 and NH_3 , the differences between the experimental and calculated CCSD(T)/cc-pCVQZ bond angles in Table IV are 2.33° and 0.89° , respectively—significantly larger than for the remaining molecules. If H_2O_2 and NH_3 are excluded from the sample, it transpires that bond lengths and bond angles have much the same convergence characteristics with respect to both the one- and the N -electron spaces (see Table V). Note also that CCSD(T) gives bond angles that are always too large, whereas bond angles determined with MP2 and CCSD exhibit less systematic behavior.

Comparing the CCSD(T)/cc-pCVQZ and reference equilibrium structures, we find that the maximum error occurs for H_2O_2 , where R_{OH} and R_{OO} are more than 0.5 pm shorter than the experimental estimates. Considering the accuracy with which the remaining CCSD(T)/cc-pCVQZ structures are determined, it is unlikely that the calculated bond lengths are that much in error; especially since the wave function of H_2O_2 is dominated by a single configuration. It is rather likely that the source of these discrepancies is associated with the experimentally inferred equilibrium structure. We

suspect that an empirical structure (which cannot be determined for H_2O_2 due to a lack of reliable microwave data) would be in much better agreement with that determined at the CCSD(T) level using the cc-pCVQZ basis set. In Table VI, statistical measures of errors are given for the CCSD(T) bond lengths for all molecules except H_2O_2 . Comparing Tables IV and VI, it is clear that the CCSD(T)/cc-pCVQZ errors are due in large part to H_2O_2 and that a much improved convergence is obtained when H_2O_2 is excluded. This is substantiated looking at Fig. 3 where the normal distribution of errors is given for the CCSD(T) calculation with and without the H_2O_2 molecule included in the sample. Note also that the experimental OOH bond angle differs from the CCSD(T)/cc-pCVQZ one by more than 2.3° —an order of magnitude more than for the rigid molecules. For the other nonrigid molecule (NH_3), the cc-pCVQZ empirical and calculated bond lengths are in reasonable agreement, whereas the bond-angle difference is 0.9° and thus substantially larger than for the rigid molecules.

We also note that similarly excellent agreement between the calculated CCSD(T)/cc-pCVQZ and empirical structures (obtained by analysis of experimental rotational constants corrected by computed vibrational corrections) has been recently obtained for a larger molecule, namely, propadienylidene ($\text{H}_2\text{C}=\text{C}=\text{C}:$).¹³

TABLE VI. Statistical measures of errors in the CCSD(T) bond lengths with H₂O₂ excluded from the sample molecules.

	cc-pVxZ			aug-cc-pVxZ			cc-pCVxZ		
	D	T	Q	D	T	Q	D	T	Q
$\bar{\Delta}$	1.74	0.05	-0.06	1.76	0.07	-0.02	1.67	0.20	0.00
$\bar{\Delta}_{\text{abs}}$	1.74	0.19	0.09	1.76	0.19	0.07	1.67	0.21	0.06
Δ_{max}	4.51	0.41	0.30	3.74	0.48	0.22	4.42	0.49	0.21
Δ_{std}	0.79	0.22	0.09	0.68	0.21	0.08	0.77	0.14	0.08

It should be realized that the high accuracy of CCSD(T)/cc-pCVQZ equilibrium structures is partly fortuitous. Thus, in CCSD(T) calculations on six selected molecules (N₂, CO, F₂, HF, H₂O, CH₂), basis-set extension beyond cc-pCVQZ causes a bond shortening of up to 0.3 pm; moreover, a full treatment of the triples at the CCSDT rather than CCSD(T) level gives a further shortening up to 0.05 pm.²⁴ A correlation treatment beyond CCSDT—most important being the inclusion of connected quadruple excitations—is expected to increase bond lengths. Assuming that the shifts in the bond lengths from Hartree–Fock to CCSD and from CCSD to CCSDT progress as a geometric series, the bond lengthening due to the connected quadruples will to a large extent counterbalance the shortening from basis-set extension and the full treatment of triples. The high accuracy of the CCSD(T)/cc-pCVQZ bond lengths therefore depends in part upon a delicate balance between systematic deficiencies that give errors of different sign and comparable magnitude.

V. COMPARISON OF ALL-ELECTRON CCSD(T) CALCULATIONS USING VALENCE AND CORE-VALENCE BASIS SETS

Finally, we compare the accuracy obtained in all-electron CCSD(T) calculations using valence and core-valence basis sets. In Table VII, the empirical bond distances are listed together with the difference between empirical and CCSD(T) bond distances obtained using valence and core-valence basis sets. The differences are only given for the TZ and QZ basis sets. With QZ the differences between all-electron valence and core-valence bond lengths are rather small, the TZ differences are larger, especially for bonds to hydrogen. Thus, the largest cc-pVXZ/cc-pCVXZ difference for X=T is 0.54 pm (R_{CH} in HCN), while the largest for X=Q is only 0.16 pm (R_{CH} in CH₂).²⁵

In general, the use of valence rather than core-valence basis sets shortens bond lengths. As the cc-pCVTZ bonds are usually too long, the reduction due to the use of the cc-pVTZ basis does not necessarily impair the overall accuracy, since it benefits from fortuitous error cancellation. Although the bond lengths change significantly from cc-pCVTZ to cc-pVTZ, the overall accuracy is nearly the same for the two sets; the mean absolute errors, for example, are 0.22 pm and 0.20 pm, respectively. At the QZ level, the core-valence basis sets give slightly more accurate results. The differences are mostly negligible, however. Hence the significantly smaller basis set dimension associated with cc-pVQZ relative to cc-pCVQZ calculations is a compelling reason to prefer the former in calculations that correlate all electrons. *It is important to mention that cc-pVQZ calculations with all electrons correlated are significantly more accurate than those which completely exclude core correlation effects.*

VI. SUMMARY

The focus of this paper has been on accurate determinations of equilibrium structures for a newly-developed nineteenth molecule test suite. The equilibrium structures have been determined using two different approaches:

- (1) Least-squares fits to experimental rotational constants corrected by vibration-rotation interaction constants obtained from high-level electronic-structure calculations;
- (2) Direct energy minimization with standard wavefunction/basis-set calculations at a high level of theory.

The first approach is expected to give equilibrium structures (denoted empirical) of uniformly high quality. Except for very small molecules such as diatomics, the accuracy obtained by this approach likely surpasses that of experimen-

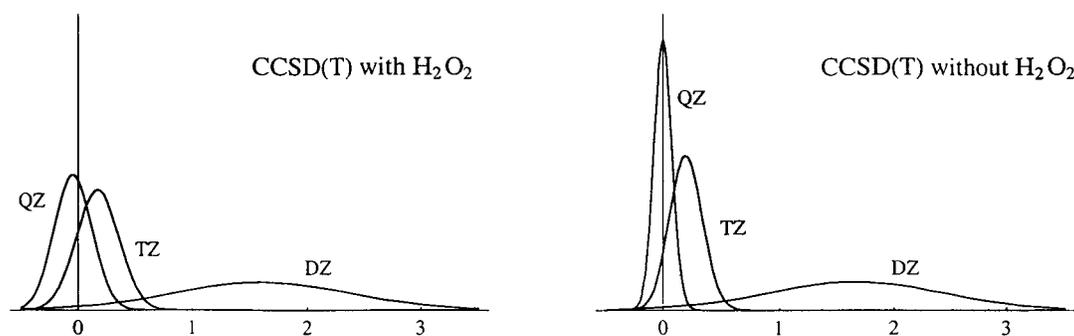


FIG. 3. Normal distributions of errors in the bond distances (pm) in CCSD(T) calculations with and without H₂O₂ included in the sample of molecules.

TABLE VII. Comparison of the empirical bond lengths and bond lengths obtained in all-electron CCSD(T) calculations using valence and core-valence basis sets at the TZ and QZ levels.

Molecule	Bond	Emp.	cc-pVQZ -Emp.	cc-pCVQZ -Emp.	cc-pCVQZ -cc-pVQZ	cc-pVTZ -Emp.	cc-pCVTZ -Emp.	cc-pCVTZ -cc-pVTZ
H ₂	R _{HH}	74.14	0.04	0.04	0.00	0.12	0.12	0.00
HF	R _{FH}	91.69	-0.17	-0.11	0.06	-0.07	-0.02	0.05
H ₂ O	R _{OH}	95.75	-0.13	-0.04	0.09	0.03	0.09	0.06
H ₂ O ₂	R _{OH}	96.70	-0.60	-0.51	0.10	-0.45	-0.40	0.05
HOF	R _{OH}	96.78	-0.30	-0.21	0.10	-0.18	-0.12	0.06
HNC	R _{NH}	99.49	-0.08	0.04	0.12	-0.17	0.08	0.25
NH ₃	R _{NH}	101.16	-0.15	-0.04	0.10	-0.03	0.10	0.13
N ₂ H ₂	R _{NH}	102.86	-0.15	-0.02	0.13	-0.07	0.12	0.19
HNO	R _{NH}	105.17	-0.07	0.07	0.14	0.06	0.25	0.19
C ₂ H ₂	R _{CH}	106.18	-0.05	0.03	0.08	-0.29	0.15	0.44
HCN	R _{CH}	106.52	-0.07	0.03	0.11	-0.38	0.16	0.54
C ₂ H ₄	R _{CH}	108.07	-0.09	0.02	0.11	-0.19	0.13	0.32
CH ₄	R _{CH}	108.59	-0.06	0.05	0.11	0.00	0.16	0.17
N ₂	R _{NN}	109.77	0.04	0.04	0.00	0.29	0.29	0.00
CH ₂ O	R _{CH}	110.07	-0.11	0.01	0.12	-0.21	0.15	0.36
CH ₂	R _{CH}	110.63	-0.11	0.05	0.16	-0.21	0.26	0.47
CO	R _{CO}	112.84	0.05	0.05	0.00	0.41	0.42	0.01
HCN	R _{CN}	115.34	0.04	0.04	0.01	0.34	0.35	0.01
CO ₂	R _{CO}	116.01	0.00	0.03	0.03	0.32	0.36	0.04
HNC	R _{CN}	116.87	0.06	0.06	-0.01	0.25	0.36	0.11
C ₂ H ₂	R _{CC}	120.35	0.06	0.02	-0.04	0.23	0.32	0.09
CH ₂ O	R _{CO}	120.47	-0.05	-0.04	0.01	0.18	0.24	0.06
HNO	R _{NO}	120.86	-0.04	-0.01	0.03	0.17	0.32	0.15
N ₂ H ₂	R _{NN}	124.57	0.06	0.10	0.04	0.30	0.49	0.19
C ₂ H ₄	R _{CC}	123.07	0.02	0.05	0.03	0.23	0.32	0.10
F ₂	R _{FF}	141.24	-0.13	-0.11	0.02	0.12	0.09	0.03
HOF	R _{FO}	143.44	-0.20	-0.18	0.02	0.00	0.09	0.08
H ₂ O ₂	R _{OO}	145.56	-0.61	-0.59	0.02	-0.21	-0.04	0.17

tally derived equilibrium structures. For example, for ethylene our empirical equilibrium structure is significantly more accurate than a previous experimental structure, the CC bond distance being about 0.3 pm shorter.

In the second approach, equilibrium structures have been determined in Hartree–Fock and all-electron MP2, CCSD, and CCSD(T) calculations using three hierarchies of correlation-consistent basis sets, cc-pVXZ, cc-pCVXZ, and aug-cc-pVXZ. The CCSD(T) model when combined with X=Q basis sets gives very accurate results. A careful calibration of CCSD(T) equilibrium structures reveals that significantly improved accuracy is obtained as one goes from TZ to QZ basis sets. At the QZ level, the mean absolute error compared relative to the reference geometries is 0.09 pm. The maximum absolute errors are obtained for H₂O₂, where R_{OH} and R_{OO} are 0.51 pm and 0.59 pm shorter, respectively, than the experimental bond lengths. This is about a factor 6 larger than the mean absolute error, suggesting that a reinvestigation of the equilibrium structure of this nonrigid molecule is warranted.

The accuracy obtained for equilibrium structures by least-squares fitting and by minimization of the CCSD(T) energy suggests that the determination of equilibrium structures based on experimental information alone is not the most accurate method available, except perhaps for very small molecules. Although experimental rotational constants are needed for the most accurate determination of equilib-

rium structures, these can be supplemented by vibrational corrections from high-level electronic structure calculations to obtain accurate “experimental” equilibrium structures.

It is significant that the accuracy of equilibrium structures obtained by energy minimization using high-level *ab initio* electronic-structure theory is sufficient for most purposes. Thus, mean absolute errors at the all-electron CCSD(T)/cc-pCVTZ and CCSD(T)/cc-pCVQZ levels are 0.2 and 0.1 pm, respectively. It is also worth pointing out that the small error at the CCSD(T)/cc-pCVQZ level in part relies on an intricate balance between errors due to the truncation in the atomic basis and those due to neglect of higher-order correlation effects, which tend to have opposite signs. To summarize, the determination of highly accurate equilibrium structures has moved from an experimental towards a combined experimental, theoretical discipline, where experimental information is restricted to the rotational constants for the vibrational ground state.

ACKNOWLEDGMENTS

This work has been supported by the Danish Natural Research Council (Grant No. 9901973). J.G. acknowledges support from the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie. J.F.S. acknowledges the support of the National Science Foundation and the Robert A. Welch Foundation.

- ¹H. F. Schaefer III, J. Russel Thomas, Y. Yamaguchi, B. J. De Leeuw, and G. Vacik, in *Modern Electronic Structure Theory*, edited by D. R. Yarkony (World Scientific, Singapore, 1995), p. 3.
- ²T. Helgaker, J. Gauss, P. Jørgensen, and J. Olsen, *J. Chem. Phys.* **106**, 6430 (1997).
- ³D. Feller and K. A. Peterson, *J. Chem. Phys.* **108**, 154 (1998).
- ⁴T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electron-Structure Theory* (Wiley, Chichester, 2000).
- ⁵G. D. Purvis and R. J. Bartlett, *J. Chem. Phys.* **76**, 1910 (1982).
- ⁶K. Raghavachari, G. W. Trucks, M. A. Pople, and M. Head-Gordon, *Chem. Phys. Lett.* **157**, 1007 (1989).
- ⁷B. P. Van Eijck, in *Accurate Molecular Structures*, edited by A. Domenicano and I. Hargittai (Oxford University Press, Oxford, 1992), p. 47.
- ⁸P. Pulay, W. Meyer, and J. E. Boggs, *J. Chem. Phys.* **68**, 5077 (1978).
- ⁹A. L. L. East, C. S. Johnson, and W. D. Allen, *J. Chem. Phys.* **98**, 1299 (1993); A. L. L. East, W. D. Allen, and S. J. Klippenstein, *ibid.* **102**, 8506 (1995); W. D. Allen, A. L. L. East, and A. G. Csaszar, in *Structures and Conformations of No-Rigid Molecules*, edited by J. Laane, M. Dakkouri, B. Van der Veken, and H. Oberhammer (Kluwer, Dordrecht, 1993), p. 343.
- ¹⁰M. C. McCarthy, C. A. Gottlieb, P. Thaddeus, M. Horn, and P. Botschwina, *J. Chem. Phys.* **103**, 7820 (1995); C. Bartel, P. Botschwina, H. Bürger, A. Guarnieri, A. Heyl, A. Huckauf, D. Lentz, T. Merzliak, and E. B. Mkdami, *Angew. Chem.* **110**, 3036 (1998).
- ¹¹J. Gauss and J. F. Stanton, *Chem. Phys. Lett.* **276**, 70 (1997).
- ¹²J. F. Stanton, C. Lopreore, and J. Gauss, *J. Chem. Phys.* **108**, 7190 (1998); J. Gauss and J. F. Stanton, *J. Phys. Chem. A* **104**, 1319 (2000); *ibid.* **104**, 2865 (2000); *Int. J. Quantum Chem.* **77**, 305 (2000); J. F. Stanton, *Mol. Phys.* **97**, 841 (1999).
- ¹³J. Gauss and J. F. Stanton, *J. Mol. Struct.* **485–486**, 43 (1999).
- ¹⁴T. H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).
- ¹⁵K. L. Bak, J. Gauss, P. Jørgensen, J. Olsen, T. Helgaker, and J. F. Stanton (to be published).
- ¹⁶I. M. Mills, in *Molecular Spectroscopy: Modern Research*, edited by K. N. Rao and C. W. Matthews (Academic, New York, 1972), p. 115.
- ¹⁷J. F. Stanton, J. Gauss, J. D. Watts, W. Lauderdale, and R. J. Bartlett, *Int. J. Quantum Chem., Symp.* **26S**, 879 (1992).
- ¹⁸D. E. Woon and T. H. Dunning, *J. Chem. Phys.* **103**, 4572 (1995).
- ¹⁹R. A. Kendall, T. H. Dunning, and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
- ²⁰J. F. Stanton and J. Gauss, *Int. Rev. Phys. Chem.* **19**, 61 (2000). A. L. L. East, C. S. Johnson, and W. D. Allen, *J. Chem. Phys.* **98**, 1299 (1993).
- ²¹J. L. Duncan, *Mol. Phys.* **28**, 1177 (1974).
- ²²J. M. L. Martin, T. J. Lee, P. R. Taylor, and J.-P. Francois, *J. Chem. Phys.* **103**, 2589 (1993).
- ²³J. F. Stanton and J. Gauss (in progress).
- ²⁴A. Halkier, P. Jørgensen, J. Gauss, and T. Helgaker, *Chem. Phys. Lett.* **274**, 235 (1997).
- ²⁵The use of cc-pVTZ for carbon in all-electron calculations sometimes gives unexpectedly large changes, perhaps because of a nonsmooth behavior of the *s* exponents in the sequence cc-pVDZ–cc-pV5Z (see, e.g., Ref. 4, p. 313).