New Scale Factors for Harmonic Vibrational Frequencies Using the B3LYP Density Functional Method with the Triple-ζ Basis Set 6-311+G(d,p)

M. P. Andersson and P. Uvdal*

Chemical Physics, Department of Chemistry, Lund University, P.O. Box 124, S-22100 Lund, Sweden

Received: September 20, 2004; In Final Form: January 24, 2005

We have calculated optimal frequency scaling factors for the B3LYP/ 6-311+G(d,p) method for fundamental vibrational frequencies on the basis of a set of 125 molecules. Using the new scaling factor, the vibrational frequencies calculated with the triple-ζ basis set 6-311+G(d,p) give significantly better accuracy than those calculated with the double-ζ 6-31G(d) basis set. Scale factors were also determined for low-frequency vibrations using the molecular set of 125 molecules and for zero-point energies using a smaller set of 40 molecules. We have studied the effect on the calculated vibrational frequencies for various combinations of diffuse and polarization functions added to the triple-ζ 6-311G basis set. The 6-311+G(d,p) basis set is found to give almost converged frequencies for most molecules, and we conclude that our optimum scaling factors are valid for the basis sets 6-311G(d,p) to 6-311++G(3df,3pd). The new scale factors are 0.9679 for vibrational frequencies, 1.0100 for low-frequency vibrations, and 0.9877 for zero-point vibrational energies.

Introduction

With the current speed of standard desktop computers and the advent of commercial quantum-chemical programs, it is a rather straightforward procedure even for experimentalists to calculate, for example, vibrational frequencies and use them to aid in interpretation of experimental results. A theoretical quantum-chemical method most widely used today is density functional theory (DFT). For investigations of ground-state properties in strongly bound systems, DFT combines accuracy with computational speed and ease of use. This is particularly true for hybrid DFT methods, which consistently have been shown to be highly reliable. Of all hybrid DFT methods, the B3LYP functional is the most widely used.[1] It is most often combined with the basis set 6-31G(d), and it has been used for almost a decade with consistently good results.[1] Since the mid-nineties, however, computer power has increased by more than an order of magnitude, making the use of a more accurate but calculationally more expensive method feasible. We will demonstrate that increasing the basis set from the double-ζ basis set 6-31G(d) to the triple-ζ basis set 6-311+G(d,p) results in a higher accuracy for the calculated vibrational frequencies.

While being successful for many problems, the absence of long-range correlation in density functionals makes DFT methods inadequate to describe weakly bound systems, see, for example, ref 2 and references therein, for a discussion. Attempts to combine multiconfigurational wave functions with density functionals are also less successful.[3] Finally, for electronically excited states, DFT-based studies are still not straightforward, see, for example, ref 4, for a discussion.

The anharmonicity of the fundamental frequencies is most often taken into account by scaling the calculated harmonic frequencies, and this procedure has been found to work well since the overestimation of vibrational frequencies is fairly uniform. Although it is now possible to calculate the anharmonicities through second-order perturbation theory with commercial programs such as GAUSSIAN 03,[5] this is still prohibitive for larger molecules and the accuracy is similar to that achieved through scaling. A more detailed description is generally needed only when anharmonic resonances are important, see, for example, ref 6. We have, however, showed that alternatively these couplings can be included phenomenologically.[7,8] In the present work we calculate optimum scale factors to be used for scaling the harmonic frequencies for the B3LYP/ 6-311+G(d,p) method. We determine three different scale factors to be used for vibrational frequencies, low-frequency vibrations, and zero-point vibrational energies, respectively.

Calculations

Optimizing the Scale Factors.

The scale factor λ to use was determined by making a least-squares fit of the scaled harmonic frequencies to the experimental fundamentals, i.e., minimizing the following sum:

$$\Delta = \sum_{i}^{N} (\lambda \omega_i^{\text{theor}} - \nu_i^{\text{exp}})^2 \qquad (1)$$

where N is the number of frequencies included in the optimization. The root-mean-square (RMS) error is hence

$$\text{RMS} = \sqrt{\left(\frac{\Delta}{N}\right)} \qquad (2)$$

In this minimization procedure, the high-frequency vibrations tend to dominate the RMS and consequently the scale factor. For some applications, such as statistical thermodynamics and calculations of enthalpies and entropies, the low-frequency vibrations are much more important and the use of another scale factor is appropriate. Following Scott and Radom,[1] we also made a least-squares fit of the scaled inverse harmonic frequencies to the inverse fundamental frequencies, that is, minimized the following sum:

$$\Delta = \sum_{i}^{\text{all}} \left(\frac{1}{\lambda \omega_i^{\text{theor}}} - \frac{1}{\nu_i^{\text{exp}}} \right)^2 \qquad (3)$$

---

* Author to whom correspondence should be addressed. E-mail: per.uvdal@chemphys.lu.se.

10.1021/jp045733a CCC: $30.25

TABLE 1: Optimized Scale Factors for the 6-311G Basis Set with Various Polarization and Diffuse Functions Added[a]

diffuse and polarization functions added to the 6-311G basis set	optimized scale factor for the reduced molecular set of ref 9	RMS error
	0.9737	95
+	0.9738	99
++	0.9742	98
d	0.9623	40
d,p	0.9619	33
+ d,p	0.9613	33
+ df,p	0.9608	35
+ 2d,p	0.9618	31
+ 3d,p	0.9612	32
+ 3df,p	0.9609	32
++ 3df,3pd	0.9604	32

[a] The reduced set of Johnson, Gill and Pople from ref 9 was used.

The RMS is again calculated using eq 2 with Δ taken from eq 3 above. With this minimization scheme, the sum and RMS error are dominated by the lowest frequencies.

When determining the frequency scale factor to use for zero-point energies, we have made a least-squares fit of scaled calculated zero-point energies to experimental zero-point energies analogously to eq 1. For a more detailed description of the procedures, see ref 1.

Quantum-Chemical Calculations.

We have used the Gaussian 03W[5] and the method B3LYP in order to perform geometry optimizations and frequency calculations on a large set of molecules. The default pruned grid was used in all calculations since the effect of increasing the integration grid has been shown to be very small for vibrational frequecies.[1] All optimizations were performed with default cutoffs.

To investigate the effect of diffuse and polarization functions on the vibrational frequencies, we made frequency calculations for a small set of 29 molecules (taken from ref 9, excluding LiH and LiF). This was made for the 6-311G basis set with various key combinations of diffuse and polarization functions added, and we then determined the optimum scale factor and root-mean-square (RMS) error for each of the basis sets investigated. The scale factors and RMS errors are collected in Table 1, and frequency variations for some of the molecules in the set as compared to the “full” basis set 6-311++G(3df,3pd) are shown in Figures 1−3. Since these scaling factors and RMS errors are only to be compared internally, a smaller set of molecules is chosen in order to reduce the number of computations. This strategy is based on the assumption that the trends are well described by this smaller set of molecules, which should be true considering that it contains several different functional groups.

A larger set of molecules consisting of 125 molecules (950 frequencies), taken from Shimanouchi[10,11] and Johnson et al.,[9] was used for determining the optimum scale factors for the 6-311+G(d,p) basis set for frequencies and inverse frequencies. The molecules chosen have a maximum of four heavy atoms (non-H) and no more than 10 atoms in total. No molecules with atoms from the third row or higher in the periodic table were included, and only one isotopomer of each molecule has been included.

When determining the scaling factor for zero-point energies we used a set of 40 molecules which consisted of diatomics from the atomization energies of Table 3 in ref 12 and the molecules in Table 1 in the work of Grev and Janssen[13] (again excluding LiH and LiF). Forty was the number of molecules for which experimentally determined zero-point vibrational energies were available. Experimental zero-point vibrational energies for the diatomics were taken from Herzberg,[14] using standard equations for calculating zero-point energy including anharmonicity, see, for example, ref 1. All other data are taken from Table 1 in Grev and Janssen.[13] These data come from the following sources: singlet and triplet CH_2 from Jensen and Bunker,[15] CH_4 from Gray and Robiette,[16] CH_3Cl from Duncan and Law,[17] H_2O, H_2S, HCO, C_2H_4, and H_2CO from Clabo et al.,[18] and HCN, CO_2, and C_2H_2 from Allen et al.[19]

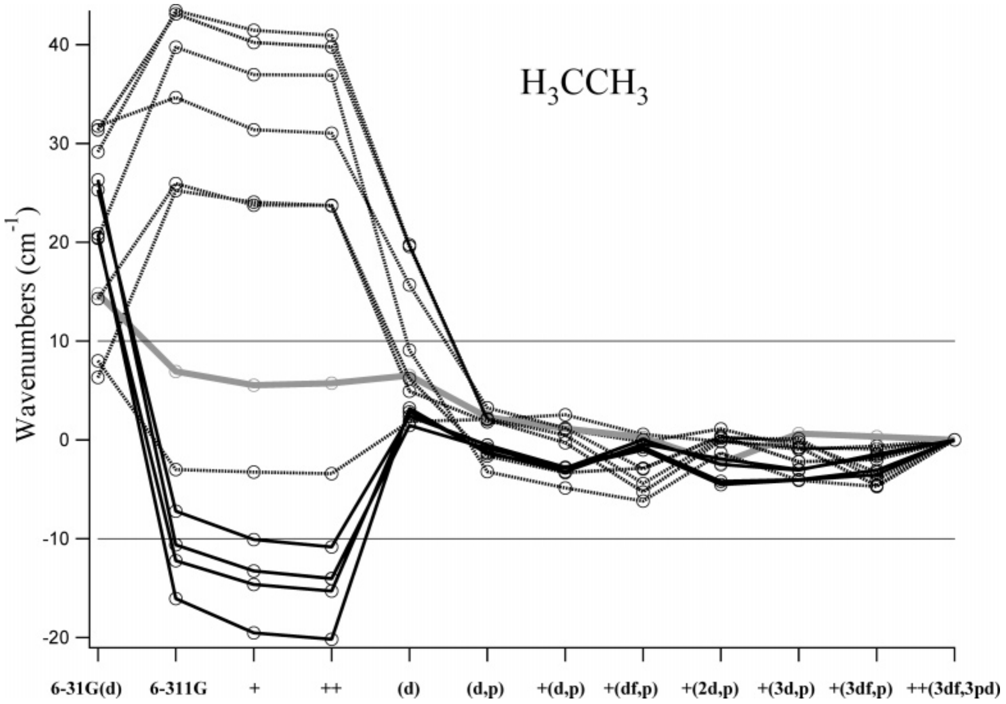

Figure 1. Variation of calculated vibrational frequencies with respect to addition of diffuse functions and polarization functions for H_3CCH_3. All frequencies are compared to the “full” basis set 6-311++G(3df,3pd). CH deformations are drawn as dotted lines, CH stretch modes are drawn as full lines, and the CC stretch is drawn as a full, bold gray line.

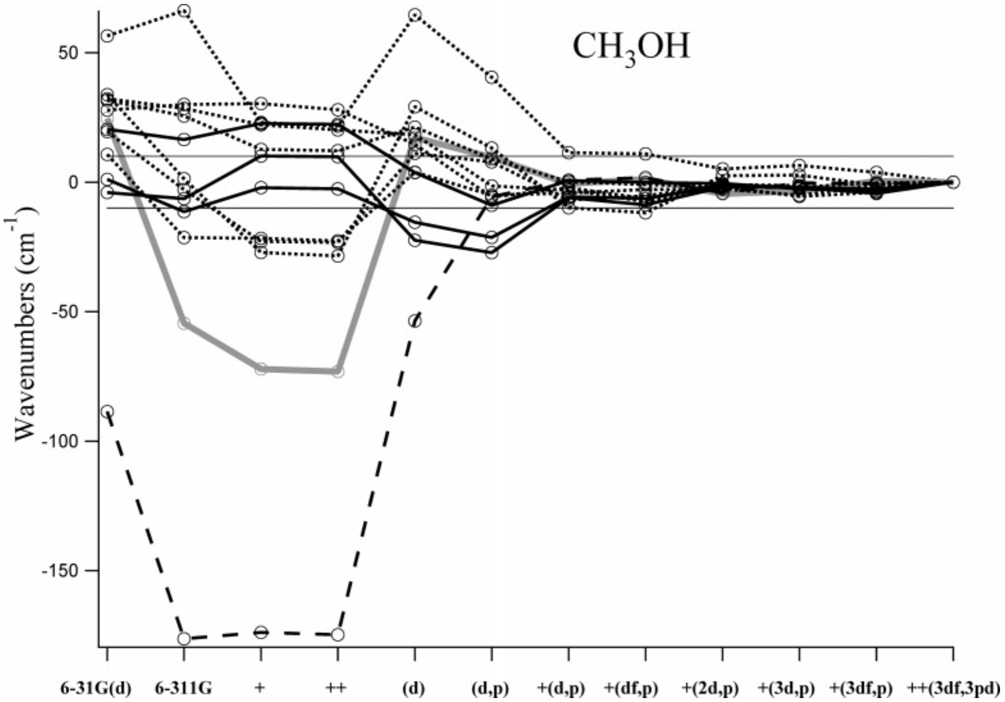

Figure 2. Variation of calculated vibrational frequencies with respect to addition of diffuse functions and polarization functions for CH_3OH. All frequencies are compared to the "full" basis set 6-311++G(3df,3pd). CH deformations and the OH deformation are drawn as dotted lines, CH stretch modes are drawn as full lines, the CO stretch is drawn as a full, bold gray line, and the OH stretch is drawn as a long-dashed line.

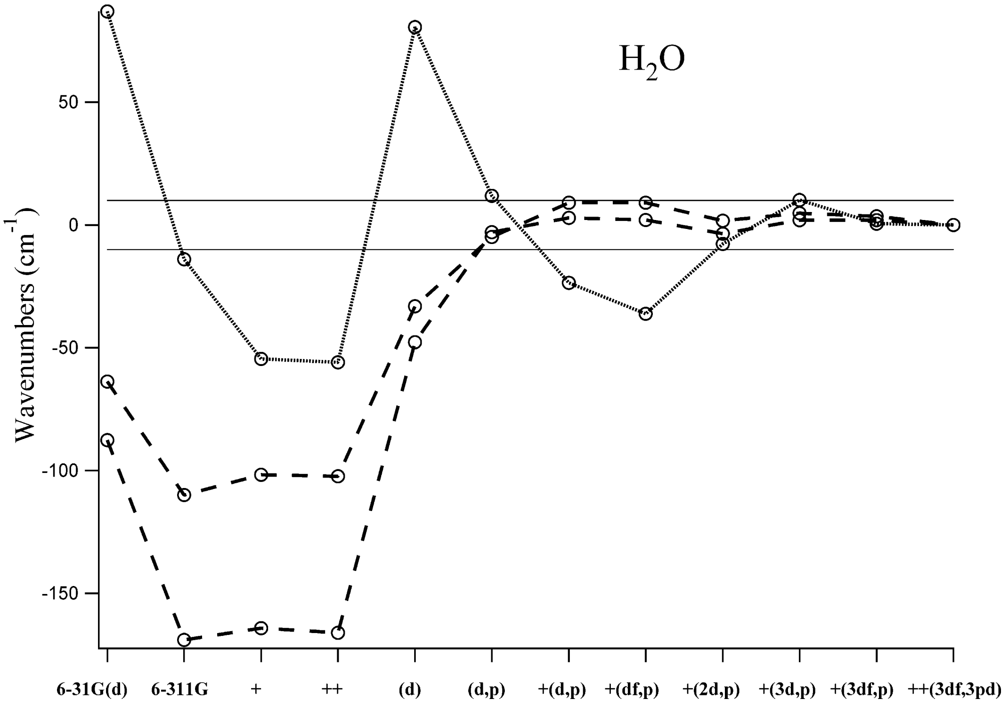

Figure 3. Variation of calculated vibrational frequencies with respect to addition of diffuse functions and polarization functions for H_2O. All frequencies are compared to the "full" basis set 6-311++G(3df,3pd). The OH deformation is drawn as a dotted line, and the OH stretch modes are drawn as long-dashed lines.

Results and Discussion

The accuracy of the calculated frequencies depends on the size and quality of the chosen basis set. This is important since, in general, quantum-chemical calculations are a compromise between accuracy and computational time. Figures 1−3 show how the vibrational frequencies vary with basis set for the molecules H_3CCH_3, CH_3OH, and H_2O. The data indicate that convergence of the vibrational frequencies with respect to addition of diffuse and polarization functions is generally met already at the 6-311G(d,p) level (change $\leq \sim 10$ cm^{-1}). An example can be seen for CH_3CH_3 in Figure 1. For molecules with hydrogen atoms and significant negative charge (e.g., H_2CO, HCN, HCO, and CH_3OH, Figure 2), a diffuse function on heavy atoms (basis set 6-311+G(d,p)) is required in order to reach convergence; for some more difficult cases (e.g., H_2O, Figure 3, and H_2NNH_2), a second polarization function is also needed (6-311+G(2d,p)).

We have also investigated how different types of modes behave differently when the basis set is changed. This was done

TABLE 2: Percentage of Scaled Theoretical Frequencies That Lie within Specified Error Ranges When Compared with Experimental Fundamentals

% error	B3LYP 6-311+G(d,p)	B3LYP 6-31G(d)[a]	B3PW91 6-31G(d)[a]
0−2	93.4	57.5	69.6
2−4	4.0	18.1	16.8
4−6	1.3	10.2	4.0
6−8	0.3	5.3	4.2
8−10	0.2	3.2	1.7
10−12	0.2	2.3	0.6
12−14	0.2	1.1	1.0
14−16	0.0	0.8	0.2
16−18	0.1	0.5	0.0
18−20	0.1	0.1	0.2
20−22	0.0	0.0	0.4
22−24	0.0	0.1	0.2
24−26	0.0	0.1	0.2
26−28	0.1	0.1	0.0
28−30	0.0	0.0	0.0
>30	0.1	0.6	0.8

[a] From ref 9.

for the 12 molecules: CN, CO, H_2O, CH_3CH_3, CH_2CH_2, CHCH, CH_3OH, NH_2NH_2, CH_2O, HCO, HCN, and H_2CO. Figures 1−3 show CH_3CH_3, CH_3OH, and H_2O, respectively. The others are available as supplemental information. Without any polarization functions, CH deformation vibrations are too high. CH stretches are slightly too low for X−CH molecules when X = C or H and too high when X = O or N. CO and CN stretch vibrations are too low, whereas CC and NN stretches are about right. OH stretch vibrations are much too low, whereas NH stretches are too high. There is much less dependence on diffuse functions for all the different modes. One should keep in mind that this is only a very limited set of molecules and that the trends might not be quite as general for a larger set of molecules.

That convergence is generally met at the 6-311G(d,p) level is also evident from the scale factors and RMS values in Table 1, which are essentially the same for 6-311G(d,p) and all larger basis sets investigated. Due to the almost equal scale factors, we have chosen to calculate a single scaling factor to be used for all those basis sets. Importantly the scale factors in Table 1 are evaluated for a limited data set only and calculated in order to be compared to each other. We have chosen to calculate a single accurate scale factor for the 6-311+G(d,p) basis set due to its slightly higher accuracy than 6-311G(d,p) when investigating molecules with significant negative charge.

Another reason for choosing this basis set is that its scaling factor lies between the maximum (6-311G(d,p)) and minimum (6-311++G(3df,3pd)) scaling factors in Table 1 for the relevant basis sets. By using the accurate scale factor determined for 6-311+G(d,p) for all basis sets 6−311(d,p) and larger, we introduce a systematic error. We can estimate the error of the approximation from the spread in scaling factors for these basis sets from Table 1. The total spread of scaling factors is 0.0015 with the scaling factor for the 6-311+G(d,p) in the middle of the interval resulting in an error of about ±3 cm^{-1} for a CH stretch mode. The deviation resulting from this approximation is, as mentioned, systematic but very small.

The scale factor for the 6-311+G(d,p) basis set when optimized with eq 1 for the large set of 125 molecules (950 frequencies) was 0.9679, and the RMS error was 37 cm^{-1}. The old scale factor 0.9614, as determined for the 6-31G(d) basis set, lies outside the 99% confidence interval for our new scaling factor. The new scaling factor 0.9679 should therefore be used for basis sets 6-311G(d,p) or larger. The RMS error is very similar to what Scott and Radom[1] achieved, even though our accuracy is much higher (Table 2). This is as expected since RMS errors are dominated by a few large deviations.

TABLE 3: Recommended Scale Factors To Use with the B3LYP Method for the Basis Sets 6-311G(d,p) up to 6-311++G(3df,3pd) for Calculation of Fundamental Frequencies, Low-Frequency Vibrations, and Zero-Point Vibrational Energies (ZPVE)

level of theory	ω	$1/\omega$	ZPVE
B3LYP/6-311+G(d,p)	0.9679	1.0100	0.9877

In Table 2, the spread of relative errors is shown for the B3LYP/6-311+G(d,p) method and the results are compared to the most widely used method B3LYP/ 6-31G(d) and the best method from Scott and Radom[1] B3PW91/6-31G(d). There is a dramatic increase in accuracy of the calculations upon using a large-enough triple-ζ basis set. Of the scaled calculated frequencies, 93.4% are within 2% of the experimental fundamental. This is to be compared to 57.5% for B3LYP/6-31G(d) and 69.6% for B3PW91/6-31G(d). The much higher accuracy clearly shows that whenever computer power is available the 6-311G(d,p) (or higher) basis set should be used instead of the standard 6-31G(d), when calculating vibrational properties.

The scale factor for low-frequency vibrations was also calculated, and a few methyl rotations dominated the RMS value, as was also observed by Scott and Radom.[1] Following the scheme by Scott and Radom,[1] three such frequencies were removed (resulting in a set of 947 frequencies). Upon reoptimization the RMS error dropped by more than 30%. The scale factor obtained without these frequencies is 1.0100, and the RMS error is 11×10^{-5} cm^{-1}. Again, the old scale factor 1.0013 for the 6-31G(d) basis set lies outside the 99% confidence interval for our new scaling factor. Again, we recommend a change of scale factor for use with basis sets 6-311G(d,p) and higher. The RMS error is again in the same range as that of Scott and Radom.[1]

The optimum scale factor for zero-point vibrational energies was found to be 0.9877 and the RMS error 0.36 kJ/mol. The scaling factor 0.9806 determined for the 6-31G(d) basis set lies outside an 80% confidence interval for our new scale factor. The difference is significant enough that our new scaling factor should be used for 6-311G(d,p) and larger basis sets. In particular, this is true since both scale factors have been determined using almost the same set of molecules. Again, the RMS error is similar to that of Scott and Radom.[1] All optimum scale factors are summarized in Table 3 for convenience.

Summary

We have determined new frequency scaling factors for fundamental frequencies calculated with the method B3LYP/6-311+G(d,p) using least-squares fitting to a large molecular set of 125 molecules (950 fundamental frequencies). Investigations for a limited set of molecules show that the same scaling factors can give a very good approximation to be used for the basis sets 6-311G(d,p) up to 6-311++G(3df,3pd), since most vibrational frequencies change very little upon adding diffuse functions or more polarization functions to the 6-311G(d,p) basis set. We also show that the 6-311G(d,p) basis set (or larger) is superior to the double-ζ basis set 6-31G(d) when calculating vibrational properties. The accuracy of the scaled vibrational frequencies is much higher for the former (93.4% of the calculated frequencies have errors <2% as compared to 57.5% for 6-31G(d)). Using the large molecular set of 125 molecules we calculated a scale factor for low-frequency vibrations, while a set of 40 molecules was used to calculated a zero-point vibrational energy scale factor.

References and Notes

(1) Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502.

(2) Wu, X.; Vargas, M. C.; Nayak, S.; Lotrich, V.; Scoles, G. *J. Chem. Phys* **2001**, *115*, 8748.

(3) McDouall, J. J. W. *Mol. Phys.* **2003**, *101*, 361.

(4) Tozer, D. J.; Amos, R. D.; Handy, N. C.; Roos, B.; Serrano-Andrés, L. *Mol. Phys.* **1999**, *97*, 859.

(5) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*; Gaussian, Inc.: Pittsburgh, PA, 2003.

(6) Andersson, M. P.; Uvdal, P. *Phys. Rev. Lett.* **2003**, *90*, 076103.

(7) Ásmundsson, R.; Uvdal, P. *J. Chem. Phys.* **2000**, *112*, 366.

(8) Andersson, M. P.; Ásmundsson, R.; Uvdal, P. *Surf. Sci.* **2003**, *532–535*, 221.

(9) Johnson, B. G.; Gill, P. M. W.; Pople, J. A. *J. Chem. Phys.* **1993**, *98*, 5612.

(10) Shimanouchi, T. *Tables of molecular vibrational frequencies I*; NSRDS-NBS 39, Washington, DC, 1972.

(11) Shimanouchi, T. *J. Phys. Chem. Ref. Data* **1977**, 993.

(12) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221.

(13) Grev, R. S.; Janssen, C. L.; Schaefer, H. F. *J. Chem. Phys.* **1991**, *95*, 5128.

(14) Herzberg, G. *Molecular Spectra and Molecular Structure*; Van Nostrand: New York, 1945; Vol. I & II.

(15) Jensen, P.; Bunker, P. R. *J. Chem. Phys.* **1988**, *89*, 1327.

(16) Gray, D. L.; Robiette, A. G. *Mol. Phys.* **1979**, *37*, 1901.

(17) Duncan, J. L.; Law, M. M. *J. Mol. Spectrosc.* **1990**, *140*, 13.

(18) Clabo, D. A.; Allen, W. D.; Remington, R. B.; Yamaguchi, Y.; Schaefer, H. F. *Chem. Phys.* **1988**, *123*, 187.

(19) Allen, W. D.; Yamaguchi, Y.; Csaszar, A. G.; Clabo, D. A.; Remington, R. B.; Schaefer, H. F. *Chem. Phys.* **1990**, *145*, 427.