8430 J. Phys. Chem. A **2005,** *109,* 8430−8437

Uncertainties in Scaling Factors for ab Initio Vibrational Frequencies

Karl K. Irikura,*,† Russell D. Johnson III,† and Raghu N. Kacker‡

Physical and Chemical Properties Division and Mathematical and Computational Sciences Division, National Institute of Standards and Technology, Gaithersburg, Maryland 20899-8380

Received: May 26, 2005; In Final Form: July 25, 2005

Vibrational frequencies determined from ab initio calculations are often scaled by empirical factors. An empirical scaling factor partially compensates for the errors arising from vibrational anharmonicity and incomplete treatment of electron correlation. These errors are not random but are systematic biases. We report scaling factors for 40 combinations of theory and basis set, intended for predicting the fundamental frequencies from computed harmonic frequencies. An empirical scaling factor carries uncertainty. We quantify and report, for the first time, the uncertainties associated with the scaling factors. The uncertainties are larger than generally acknowledged; the scaling factors have only two significant digits. For example, the scaling factor for HF/6-31G(d) is 0.8982 ± 0.0230 (standard uncertainty). The uncertainties in the scaling factors lead to corresponding uncertainties in predicted vibrational frequencies. The proposed method for quantifying the uncertainties associated with scaling factors is based on the *Guide to the Expression of Uncertainty in Measurement*, published by the International Organization for Standardization (ISO). The data used are from the Computational Chemistry Comparison and Benchmark Database (CCCBDB), maintained by the National Institute of Standards and Technology, which includes more than 3939 independent vibrations for 358 molecules.

1. Introduction

One of the most popular uses of computational quantum chemistry models is to predict vibrational frequencies for spectroscopy, thermochemistry, and reaction kinetics. However, vibrational frequencies predicted from quantum chemistry models seldom agree with the corresponding experimental frequencies. A common practice is to apply an empirical scaling factor to the computed harmonic frequency to bring it closer to the "true" fundamental frequency. The empirical scaling factor carries uncertainty. However, to our knowledge, no one has quantified the uncertainties associated with the scaling factor and the corresponding scaled computed frequency. This paper attempts to quantify these uncertainties.

It is well recognized that an experimental measurement is incomplete in the absence of a quantitative and valid expression of its associated uncertainty. Unfortunately, predictions from computational quantum chemistry models are often reported without their associated uncertainties, making them incomplete statements. The term *virtual measurement* refers to a scaled (or otherwise corrected) computed result together with its associated uncertainty. The term *virtual measurement* emphasizes the analogy with experimental (physical) measurement along with its associated uncertainty. Our investigations of the uncertainties associated with predictions from quantum chemistry models are intended to help "virtual measurements" supplant "calculated results."

Scaling of Computed Frequencies. Harmonic vibrational frequencies are computed by mass-weighting and diagonalizing the force constant matrix, which is a matrix of second-order derivatives of the electronic energy with respect to nuclear coordinates. Because the force constants are the underlying quantities, the first approach was to apply empirical scaling to the force constants.[1−4] This remains an effective approach.[5] Another approach is uniform scaling of the harmonic vibrational frequencies. The resulting predictions are less accurate (more uncertain), but the latter approach has the advantage of simplicity. The present study deals with this simpler approach.

The first published survey of uniform scaling focused on HF/3-21G calculations.[6] One of the conclusions of that study was that "the harmonic frequencies calculated at this level exceed the anharmonic observed frequencies by an average of 180 cm^{-1} or 12.3%". It was therefore recommended to multiply the computed frequencies by the factor 0.89 to counter their bias of about 12.3%. A more precise factor, 0.8929 (=1/1.12), was adopted for HF/6-31G(d) calculations as part of the popular "Gaussian-*n*" methods for quantitatively predicting molecular energetics.[7] The reason for increasing the number of seemingly significant digits from two (0.89) to four (0.8929) remains unclear.

It is now a common practice to apply such scaling factors to the computed frequencies obtained from quantum chemistry models. Although the computed quantities are harmonic frequencies, it is implicitly assumed that the scaling factor absorbs most of the anharmonic effects, as well as errors in the force constants due to approximations in the electronic structure calculations. Thus, the scaled computed frequencies serve as predictions of the fundamental frequencies.

As the computation of vibrational frequencies has become routine and the number of computational models, each characterized by a level of theory and a basis set, has proliferated, the demand for scaling factors has increased. Several scaling-factor studies have been published. For example, on the basis of data for 36 small molecules, it was reported in 1982 that HF/6-31G(d) and MP2/6-31G(d) harmonic frequencies are 12.6% and 7.3% higher, respectively, than the corresponding experimental fundamentals.[8] The corresponding scaling factors are 0.8881

---

* To whom correspondence should be addressed.

† Physical and Chemical Properties Division.

‡ Mathematical and Computational Sciences Division.

10.1021/jp052793n

and 0.9320, respectively. In 1993, a study involving 122 molecules (1066 frequencies) yielded a scaling factor of 0.9427 for MP2/6-31G(d)[9] and a scaling factor of 0.8953 for HF/6-31G(d); however, the authors of that study recommended that for HF/6-31G(d) the previously recommended value (0.8929) continue to be used because there was little difference. Two studies were published independently in 1996 using the same set of 122 molecules as that in the 1993 study.[9] One of them focused on five density functionals in conjunction with the 6-31G(d) basis sets.[10] The other, encompassing 19 theoretical models, is the most thorough and most cited such study to date.[11] Because the same experimental data were used, the results for HF/6-31G(d) and MP2/6-31G(d) were the same as those in the 1993 study.[9] In more recent work, additional basis sets have been addressed. For a set of 900 frequencies, scaling factors were reported for HF (0.9066), MP2 (0.9649), and four density functionals in conjunction with Sadlej's pVTZ basis sets.[12] For the frequencies above 1000 cm^{-1} for a set of 41 molecules, scaling factors were recently reported for HF, B3LYP, and MP2 combined with six of Dunning's correlation-consistent basis sets.[13] Most recently, a study of B3LYP concluded by recommending that a single scaling factor be used for several related basis sets.[14] As an alternative to the uniform scaling of computed vibrational frequencies, an improved fit can be obtained by including a second scaling parameter,[12,15,16] but this has not become popular.

Most of the systematic studies report root-mean-square (rms) residuals from least-squares fitting of scaling factors. Some also provide histograms of residuals before and after the scaling. However, no study has attempted to quantify the uncertainties associated with the scaling factors and the scaled computed frequencies. The purpose of the present study is to establish an approach to quantify those uncertainties. This is important because (i) a vibrational frequency predicted from a quantum chemistry model is incomplete without a valid expression of its associated uncertainty, (ii) scaling factors are typically quoted to four digits without consideration of their significance, and (iii) the use of seemingly highly precise scaling factors may be unwarranted.

Computational Comparison and Benchmark Database (CCCBDB). The CCCBDB[17] is a large, Web-accessible database containing calculated results from many quantum chemistry models and the corresponding experimental results. The two goals of the CCCBDB are to provide benchmark, evaluated experimental data for testing new computational methods and to illustrate how well the current computational methods perform, relative to experimental measurements, in predicting the properties of gas-phase molecules. The CCCBDB initially focused on gas-phase thermochemistry (enthalpies of formation, entropies, and heat capacities) and related properties (moments of inertia and vibrational frequencies). It was later expanded to include some other properties that are often predicted by using quantum chemistry models. In particular, the CCCBDB includes experimental vibrational data for 386 molecules, for a total of 4824 distinct vibrational frequencies. The uncertainties associated with the experimental measurements range from less than 1 to 15 cm^{-1}.[18] Despite continued efforts to ensure the integrity of experimental data in the CCCBDB, some incorrect experimental values may remain. Indeed, the computational results in the CCCBDB are useful for identifying questionable experimental vibrational frequencies or mode assignments. The CCCBDB includes tools for comparing the calculated and experimental vibrational frequencies for one or more molecules. The CCCBDB may be used to specify a scaling factor for a given quantum chemistry model and one or more molecules. Given the large number of experimental vibrational frequencies and over a million calculated vibrational frequencies covering more than 140 computational models, the CCCBDB may be used to estimate the uncertainties associated with the scaling factors and the scaled computed vibrational frequencies.

2. Guide to the Expression of Uncertainty in Measurement

The uncertainty associated with a computed vibrational frequency arises primarily, but not exclusively, from its bias (systematic error) with respect to the "true" fundamental frequency. Before publication of the *Guide to the Expression of Uncertainty in Measurement*[19] by the International Organization for Standardization (ISO), there was no generally accepted approach to account for the uncertainty arising from a bias. The approach recommended by the ISO *Guide* is now generally accepted. Furthermore, the ISO *Guide* is the de facto international standard for quantifying the uncertainty in measurement. The ISO *Guide* recommends that all uncertainties be expressed as *standard uncertainties* or as *uncertainty intervals*, defined below. A *measurand* is a property, denoted by Y, which is subject to measurement or prediction. A computed result or an experimental result, denoted by y, is an *estimate* for Y. The *uncertainty* is a parameter associated with the estimate y that characterizes the dispersion of the values that could reasonably be attributed to Y on the basis of all available information. The *standard uncertainty* is uncertainty expressed as a standard deviation, denoted by $u(y)$.

The ISO *Guide* is based on the concept of a *measurement equation*. In its simplest form, a measurement equation is a mathematical function, $Y = f(Q_1, ..., Q_N)$, that represents the process used for determining an estimate y and its associated standard uncertainty $u(y)$ from the estimates and their associated standard uncertainties for various input quantities, $Q_1, ..., Q_N$. Each input and output quantity of a measurement equation is regarded as a variable with a *state-of-knowledge probability distribution* having an expected value and a finite standard deviation. The estimate y is determined by substituting the estimates $q_1, ..., q_N$ for the input variables in the measurement equation $Y = f(Q_1, ..., Q_N)$. That is,

$$y = f(q_1, ..., q_N) \tag{1}$$

The standard uncertainties $u(q_1), ..., u(q_N)$ associated with the input estimates $q_1, ..., q_N$ are components of uncertainty in determining the estimate y. The measurement equation $Y = f(Q_1, ..., Q_N)$ is approximated about y by a Taylor series as

$$Y \approx Y_{\text{linear}} = y + \sum_i c_i (Q_i - q_i) \tag{2}$$

where $c_1, ..., c_N$ are partial derivatives of Y with respect to $Q_1, ..., Q_N$ evaluated at $q_1, ..., q_N$. Treating q_i and $u(q_i)$ as the expected value and standard deviation of Q_i, the variance of Y_{linear} gives the following expression for propagating the uncertainties associated with the input estimates:

$$u^2(y) = \sum_i c_i^2 u^2(q_i) + 2\sum_{(i<j)} c_i c_j u(q_i) u(q_j) r(q_i, q_j) \tag{3}$$

where $r(q_i,q_j)$ is the correlation coefficient between Q_i and Q_j for $i, j = 1, ..., N$ and $i \neq j$. The estimate y and the standard uncertainty $u(y)$ are the expected value and the standard deviation of Y_{linear}. The ISO *Guide* regards y and $u(y)$ as approximating the expected value and the standard deviation of a state-of-knowledge probability distribution for Y.

The uncertainty may alternatively be expressed as an *uncertainty interval* $[y \pm ku(y)]$, for some *coverage factor*, k. By convention, chemists express the uncertainty as an interval, $[y \pm ku(y)]$, with a supposed *coverage probability* of about 95%. The coverage probability of an uncertainty interval can be stated only under special conditions. In this study, we do not discuss specification of uncertainty intervals $[y \pm ku(y)]$ with a stated (supposed) coverage probability such as 95%. This is an interesting topic for investigation in the future.

ISO *Guide's* Approach To Quantify the Uncertainty Arising from Biases. Suppose that x_0 is a preliminary estimate such as the output of a computational model for the value Y_0 of a measurand. We use the symbols $E(\cdot)$, $S(\cdot)$, and $V(\cdot)$ for the expected value, the standard deviation, and the variance, respectively, of the variable indicated in the argument. Suppose the expected value $E(x_0)$ of the *sampling probability distribution* of x_0 is X_0. That is, X_0 is the hypothetical average of infinitely repeated independent evaluations of x_0 under given conditions. The difference $x_0 - X_0$ is the random error, and the ratio x_0/X_0 is the fractional random error in x_0. In quantum chemistry, the random error arises from a variety of small contributions, such as the nonzero convergence thresholds that create dependence upon the choice of the initial geometry and wave function. The random error is generally negligible. Otherwise, the uncertainty arising from the sources of random error must be quantified and incorporated. The difference $X_0 - Y_0$ is the *additive bias*, and the ratio X_0/Y_0 is the *fractional bias* (or multiplicative bias) in x_0. The bias arises from systematic (nonrandom) effects in the process of evaluating x_0. The ISO *Guide* recommends that the estimate x_0 be corrected or scaled to counter its bias, thus providing a corrected or scaled virtual measurement, y_0, for Y_0. From this viewpoint, we refer to x_0 as an uncorrected or unscaled virtual measurement for Y_0 with a small standard uncertainty $u(x_0)$, mainly from random effects. A measurement equation is required to incorporate a correction or scaling factor for the bias. Thus, X_0 and Y_0 are regarded as variables with probability distributions representing the states of knowledge about the expected value X_0 and the value Y_0 of the measurand, respectively. The ISO *Guide* identifies the expected value $E(X_0)$ of a state-of-knowledge distribution for X_0 with the uncorrected or unscaled virtual measurement x_0 and the standard deviation $S(X_0)$ with the standard uncertainty $u(x_0)$.

The measurement equation that corresponds to the additive bias $X_0 - Y_0$ is

$$Y_0 = X_0 + C_0 \tag{4}$$

where C_0 is a variable with a probability distribution representing the state of knowledge about the negative of bias $(Y_0 - X_0)$. Suppose that the expected value $E(C_0)$ and the standard deviation $S(C_0)$ are c_0 and $u(c_0)$, respectively. Then, a corrected estimate, y_0, for Y_0 is determined by substituting x_0 for the variable X_0 and c_0 for the variable C_0 in the measurement equation (eq 4). Thus,

$$y_0 = x_0 + c_0 \tag{5}$$

The combined standard uncertainty $u(y_0)$ is determined by propagating the variances $V(X_0) = u^2(x_0)$ and $V(C_0) = u^2(c_0)$ and the covariance $C(X_0,C_0) = u(x_0,c_0)$. Because the probability distributions for X_0 (random effects) and C_0 (negative of bias) are specified independently, we have $u(x_0,c_0) = 0$. Thus, the expression for propagating uncertainties based on the measurement eq 4 is

$$u^2(y_0) = u^2(x_0) + u^2(c_0) \tag{6}$$

The measurement equation that corresponds to the multiplicative bias X_0/Y_0 is

$$Y_0 = X_0 C_0 \tag{7}$$

where C_0 is a variable with a probability distribution representing the state of knowledge about the reciprocal of bias Y_0/X_0. Thus,

$$y_0 = x_0 c_0 \tag{8}$$

The expected value c_0 is a scaling factor for x_0. A linear Taylor series approximation of the measurement equation (eq 7) about y_0 simplifies to

$$\frac{Y_0 - y_0}{y_0} \approx \frac{X_0 - x_0}{x_0} + \frac{C_0 - c_0}{c_0} \tag{9}$$

The expression for propagating uncertainties based on the linear approximation (eq 9) of the measurement eq 7 is

$$u_r^{\,2}(y_0) \approx u_r^{\,2}(x_0) + u_r^{\,2}(c_0) \tag{10}$$

where $u_r(y_0) = u(y_0)/y_0$, $u_r(x_0) = u(x_0)/x_0$, and $u_r(c_0) = u(c_0)/c_0$ are the *relative standard uncertainties* associated with y_0, x_0, and c_0, respectively.[20] In the case of vibrational frequencies computed from quantum chemistry models, y_0, x_0, and c_0 are positive so the relative standard uncertainties $u_r(y_0)$, $u_r(x_0)$, and $u_r(c_0)$ are well defined. Thus

$$u(y_0) = y_0 u_r(y_0) \approx y_0[u_r^{\,2}(x_0) + u_r^{\,2}(c_0)]^{1/2} \tag{11}$$

where y_0 is defined by eq 8. Thus, the ISO *Guide*'s approach to quantify the uncertainty arising from bias requires one to specify the scaling factor c_0 and its associated standard uncertainty $u(c_0)$ on the basis of all available knowledge and scientific judgment. In the rest of this paper, we propose scaling factors c_0 and their associated standard uncertainties $u(c_0)$ for computational models for vibrational frequencies. The resulting scaled estimate y_0 and its associated standard uncertainty $u(y_0)$ are regarded as approximating the expected value and the standard deviation of a state-of-knowledge probability distribution for Y_0, the "true" fundamental vibrational frequency.

3. Methodology

Additive Bias. In an earlier paper,[21] we showed how the Computational Chemistry Comparison and Benchmark Database (CCCBDB)[17] may be used to specify the correction c_0 and the standard uncertainty $u(c_0)$ associated with the atomization enthalpy of a target molecule, as obtained from a quantum chemistry calculation. Briefly, the principal challenge is to identify a class of molecules in the CCCBDB with the following three characteristics: (i) The bias for the target molecule is believed to be of the same sign and of similar magnitude as the biases for the molecules in the class. (ii) The estimated biases for the molecules in the class appear to have an approximately normal distribution and do not have an excessively large spread. (iii) The number, m, of molecules in the class is sufficiently large. Once the class is selected, the correction c_i for each molecule in the class ($i = 1, ..., m$) is estimated as the difference between the experimental and computed values. The associated standard uncertainty $u(c_i)$ is typically approximated as the uncertainty in the corresponding experimental value. Suppose the mean and standard deviation of $c_1, ..., c_m$ are $\mu = (1/m)\sum_i c_i$ and $\sigma = [(1/m)\sum_i (c_i - \mu)^2]^{1/2}$, respectively. Our recommendation was to use, as the correction and its associated standard uncertainty, $c_0 = \mu$ and $u(c_0) = [(1/m)\sum_i u^2(c_i) + \sigma^2]^{1/2}$. The corresponding prediction y_0 and uncertainty $u(y_0)$ are determined from eqs 5 and 6.

Fractional Bias. The calculated result x_0 is an unscaled vibrational frequency. We will use the CCCBDB to specify a scaling factor, c_0, and its associated standard uncertainty, $u(c_0)$, to counter fractional bias (i.e., multiplicative bias) in x_0. The corresponding scaled result y_0 and the standard uncertainty $u(y_0)$ are determined from eqs 8 and 11. The first task is to identify a class of vibrational frequencies in the CCCBDB that meet the three requirements listed in the previous paragraph. Suppose the "true" fundamental frequencies for the identified class are Y_i ($i = 1, ..., m$), the corresponding unscaled calculated results are x_i with standard uncertainties $u(x_i)$, and the experimental results are z_i with standard uncertainties $u(z_i)$. Suppose the expected value of the sampling distribution for x_i is X_i. Then, the fractional bias in x_i is X_i/Y_i. The unscaled calculated result x_i is an estimate for X_i, and the experimental value z_i is an estimate for Y_i; so $b_i = x_i/z_i$ is an estimate for the fractional bias X_i/Y_i, and the estimated scaling factor is $c_i = 1/b_i = z_i/x_i$.

In accordance with the ISO *Guide*, the unscaled calculated result x_i and the uncertainty $u(x_i)$ are regarded as the expected value and the standard deviation of a state-of-knowledge distribution for X_i. Likewise, the experimental result z_i and the uncertainty $u(z_i)$ are regarded as the expected value and the standard deviation of a state-of-knowledge distribution for Y_i. Let $C_i = Y_i/X_i$ be a variable representing the correction for fractional bias in x_i. In the ISO *Guide*, the expected value and standard deviation of a state-of-knowledge distribution for C_i are determined from a linear approximation of $C_i = Y_i/X_i$ about $c_i = z_i/x_i$. Thus, the expected value of a state-of-knowledge distribution for C_i is identified with $c_i = z_i/x_i$. Because the probability distributions for Y_i and X_i are determined independently, the covariance between Y_i and X_i is zero. So, the relative standard deviation for C_i, denoted by $u_r(c_i)$, is approximated[20] as $u_r(c_i) \approx [u_r^2(z_i) + u_r^2(x_i)]^{1/2}$, where $u_r(c_i) = u(c_i)/c_i$, $u_r(z_i) = u(z_i)/z_i$, and $u_r(x_i) = u(x_i)/x_i$. Thus,

$$u(c_i) = c_i u_r(c_i) \approx c_i[(u(z_i)/z_i)^2 + (u(x_i)/x_i)^2]^{1/2} \tag{12}$$

Following our approach for an additive bias,[21] the first step is to identify a class of vibrational frequencies for which the fractional biases $b_i = x_i/z_i$ are believed to be similar to the unknown fractional bias in the computed result x_0. For example, if x_0 is a C−H stretching frequency, one might choose a class of C−H stretching frequencies. However, nearly all previous studies have treated all available frequencies as a single class. We do the same in the present study. The distribution of the b_i values for the selected class should be examined to verify that it is approximately normal and that the spread is not excessively large.

According to the belief that the fractional bias in x_0 is similar to the biases for the class identified in the CCCBDB, each of the m state-of-knowledge distributions for $C_1, ..., C_m$ may be attributed to C_0. Suppose the probability density function (PDF) for C_i, represented by c_i and $u(c_i)$, is $p_i(\bullet)$. We propose that the PDF $p(\bullet)$, attributed to C_0, be defined as the linear combination $p(y) = \sum_i \kappa_i p_i(\bullet)$ of the PDFs $p_i(\bullet)$, where $\kappa_i = a_i/\sum_i a_i$ and $a_1, ..., a_m$ are nonnegative "weights" attributed to the PDFs $p_1(\bullet), ..., p_m(\bullet)$, respectively. A combined probability distribution with PDF $p(\bullet) = \sum_i \kappa_i p_i(\bullet)$ is referred to as a mixture probability distribution. The expected value and standard deviation of the PDF $p(\bullet)$ are $\sum_i \kappa_i c_i$ and $[\sum_i \kappa_i u^2(c_i) + \sum_i \kappa_i (c_i - \sum_i \kappa_i c_i)^2]^{1/2}$, respectively.[22] Thus, the scaling factor c_0 and its associated standard uncertainty, $u(c_0)$, may be specified as

$$c_0 = \sum_i \kappa_i c_i = \frac{\sum_i a_i c_i}{\sum_i a_i} = \frac{\sum_i a_i (z_i/x_i)}{\sum_i a_i} \tag{13}$$

and

$$u(c_0) = [\sum_i \kappa_i u^2(c_i) + \sum_i \kappa_i (c_i - c_0)^2]^{1/2} = \left[\frac{\sum_i a_i u^2(c_i)}{\sum_i a_i} + \frac{\sum_i a_i (c_i - c_0)^2}{\sum_i a_i}\right]^{1/2} \tag{14}$$

Scaling Factor Based on Least-Squares Theory. A common approach for determining the weights, a_i, and, hence, the scaling factor c_0 and its associated standard uncertainty $u(c_0)$ is to fit the linear model

$$z_i = c_0 x_i + e_i \tag{15}$$

using least-squares theory. Here, $e_i = (z_i - c_0 x_i)$ is the difference between the experimental result z_i and the corresponding scaled computed result $c_0 x_i$. Accordingly, the scaling factor c_0 based on least-squares theory is that value which minimizes the least-squares objective function

$$\Delta^2 = \sum_i e_i^{\,2} = \sum_i (c_0 x_i - z_i)^2 \tag{16}$$

The solution of minimizing the objective function (eq 16) is

$$c_0 = \frac{\sum_i x_i z_i}{\sum_i x_i^{\,2}} \tag{17}$$

Comparison of eqs 13 and 17 shows that c_0 corresponds to the weights $a_i = x_i^2$ or the probabilities $\kappa_i = x_i^2/\sum_i x_i^2$. From eq 14, the corresponding standard uncertainty $u(c_0)$ associated with the scaling factor c_0 of eq 17 is

$$u(c_0) = \left[\frac{\sum_i x_i^{\,2} u^2(c_i)}{\sum_i x_i^{\,2}} + \frac{\sum_i x_i^{\,2} (c_i - c_0)^2}{\sum_i x_i^{\,2}}\right]^{1/2} \tag{18}$$

where $c_i = z_i/x_i$. Then, the scaled computed vibrational frequency is $y_0 = x_0 c_0$ from eq 8, with standard uncertainty $u(y_0)$ determined from eq 11.

Scaling Factor and Its Associated Approximate Standard Uncertainty. As illustrated in the following, the second term of eq 18 is the dominant term. We consider the popular HF/6-31G(d) quantum chemistry model. Our data set for this model contains $m = 3508$ frequencies (each degenerate frequency is counted only once). The histogram of the corresponding estimated biases, $b_i = 1/c_i$, is shown in Figure 1. Because both terms of eq 18 have the same denominator, we need only to compare numerators. The numerator of the second term of eq 18 is

$$\sum_i x_i^{\,2} (c_i - c_0)^2 \tag{19}$$

For our HF/6-31G(d) frequencies, the values of $|x_i(c_i - c_0)|$ range from 0.009 to 332 cm^{-1}, with a median value of 21 cm^{-1}

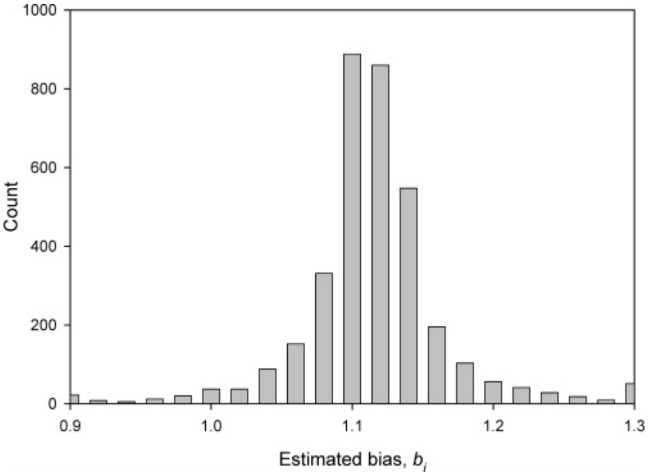

Figure 1. Distribution of estimated biases, b_i, for HF/6-31G(d) calculations of 3508 vibrational frequencies of 312 molecules.

and a mean value of 30 cm^{-1}. Using eq 12, the numerator of the first term of eq 18 is

$$\sum_i x_i^2 u^2(c_i) = \sum_i u^2(z_i) + \sum_i c_i^2 u^2(x_i) \qquad (20)$$

We consider the two pieces of eq 20 separately. The second piece is $\Sigma_i c_i^2 u^2(x_i)$, where $u(x_i)$ is the small standard uncertainty, mainly from random effects, associated with the calculated frequency x_i. Using the HF/6-31G(d) model, we computed the 27 vibrational frequencies of the propane molecule 1000 times, starting from randomized initial coordinates (Gaussian 03 software[23,24]). Because the optimized molecular geometry depends weakly upon the initial coordinates and the force constants depend on the geometry, a distribution of values was obtained for each vibrational frequency. The standard deviation of the distribution for the frequency x_i is an estimate of the standard uncertainty $u(x_i)$. The uncertainties $u(x_i)$ range from 0.05 to 0.7 cm^{-1}, with a median value of 0.2 cm^{-1}. A different quantum chemistry software package (GAMESS[23,25]), with different default convergence criteria, produced uncertainties ranging from 0.03 to 0.4 cm^{-1}, with a median value of 0.1 cm^{-1}. Mean values from the two quantum chemistry packages agreed within the corresponding standard uncertainties and always within 0.1 cm^{-1}. Because c_i is not far from 1, the range of values for $c_i u(x_i)$ is not far from the range of values for $u(x_i)$. The first piece of eq 20 is $\Sigma_i u^2(z_i)$, where $u(z_i)$ is the uncertainty associated with the experimental value z_i of the vibrational frequency. The uncertainties associated with the experimental values are not always reported. However, the compilation by Shimanouchi[18] does include estimated uncertainties. There are 1019 frequencies in the CCCBDB for which we have both HF/6-31G(d) values and the experimental uncertainties estimated by Shimanouchi. These experimental uncertainties range from 1 to 30 cm^{-1}, with a median of 15 cm^{-1} and a mean of 10 cm^{-1}. Thus, we conclude that the first term of eq 18 is smaller than the second. Henceforth, for the approximate standard uncertainty $u(c_0)$ associated with the scaling factor c_0, we will use

$$u(c_0) \approx \left[\frac{\sum_i x_i^2 (c_i - c_0)^2}{\sum_i x_i^2}\right]^{1/2} \qquad (21)$$

To evaluate the usefulness of this approximation, we computed $u(c_0)$ from eq 21 and from eq 18, neglecting the tiny second piece $\Sigma_i c_i^2 u^2(x_i)$ of eq 20. Using the Shimanouchi subset of 1019 frequencies, from eq 17, we obtain $c_0 = 0.8990$. From eq 21, we obtain $u(c_0) = 0.0210$. From eq 18, neglecting $\Sigma_i c_i^2 u^2(x_i)$, we obtain $u(c_0) = 0.0222$. The uncertainty values 0.0210 and 0.0222 are not appreciably different, justifying the use of eq 21.

Some previous studies report the root-mean-square (rms) difference between experimental and scaled frequencies and use rms as a surrogate for uncertainty. The definition is rms $= (\Delta^2/m)^{1/2}$, where Δ^2 is given by eq 16 and m is the number of frequencies in the class. Substituting into the approximation given by eq 21 and using the definition $c_i = z_i/x_i$ leads to the relation

$$u(c_0) \approx \frac{\text{rms}}{\sqrt{\langle x^2 \rangle}} \qquad (22)$$

where $\langle x^2 \rangle = (1/m)\sum_i x_i^2$. We do not use this relation further but include it for comparison with previous studies.

Equation 21 can be used to find a convenient approximation to $u(y_0)$, defined by eq 11, for the uncertainty associated with scaled computed vibrational frequency y_0. In this case, we consider $u_r(x_0)$ and $u_r(c_0)$. For the propane molecule discussed previously, the relative uncertainties $u_r(x_i)$ range from 0.000 03 to 0.003, with a median value of 0.0001 (Gaussian 03). A different software package (GAMESS) produced relative uncertainties ranging from 0.000 02 to 0.002, with a median value of 0.000 05. In contrast, $u_r(c_0) \approx 0.02$, using the uncertainties estimated by Shimanouchi described previously. Assuming that $u_r(x_0)$ will be in the range found for the vibrations of propane, $u_r(x_0)$ is negligible relative to $u_r(c_0)$. Thus, eqs 8 and 11 can be combined to produce the approximate standard uncertainty $u(y_0)$, given by eq 23.

$$u(y_0) \approx y_0 u_r(c_0) = (y_0/c_0) u(c_0) = x_0 u(c_0) \qquad (23)$$

In summary, the scaling factor c_0 is given by eq 17, and its associated approximate standard uncertainty $u(c_0)$ is given by eq 21. The corresponding scaled computed vibrational frequency y_0 is given by eq 8, and its associated approximate standard uncertainty $u(y_0)$ is given by eq 23.

4. Selection of Vibrational Frequencies in the CCCBDB

Molecules Represented in the CCCBDB. The molecules in the CCCBDB were chosen for having well-determined enthalpies of formation and not for well-determined vibrational frequencies. Nonetheless, vibrational frequencies for half of the molecules in the CCCBDB are available in evaluated compilations of experimental data. Experimental frequencies are available for 80 diatomic molecules and 306 polyatomic molecules. Of these 386 molecules, 224 are organic and 162 are inorganic. Also, 70 are free radicals. The experimental vibrational frequencies are taken from the compilations by Shimanouchi,[18] by Huber and Herzberg,[26] by Jacox,[27] and by Sverdlov, Kovner, and Krainov,[28] and also from some individual journal papers.

Integrity of the Data. Before an experimental frequency and a computational frequency are compared, it must be verified that they correspond to the same vibrational mode of the same molecule. In experimental studies, assigning the observed frequencies to specific vibrational modes can be challenging. The difficulty increases with the number of vibrational frequencies in the molecule. For a small molecule with a rotationally resolved spectrum, vibrational frequencies can be measured to a precision better than 1 cm^{-1} and the assignment is unambiguous. Over half of the molecules in the CCCBDB with vibrational frequencies are small (i.e., have five or fewer atoms). For larger

TABLE 1: Scaling Factors, c_0, and Their Associated Standard Uncertainties, $u(c_0)$, Expressed as $c_0 \pm u(c_0)$, for the Prediction of Vibrational Fundamentals, Arranged by Theoretical Method and Basis Set[a]

	3-21G	6-31G(d)	6-31G(d,p)	6-31+G(d,p)	6-311G(d,p)
HF	0.9043 ± 0.0407 [3295] (0.89,[c] 0.9085[d])	0.8982 ± 0.0230 [3508] (0.8929,[e] 0.8881,[f] 0.8953[g])	0.9024 ± 0.0241 [3419] (0.8992[d])	0.9039 ± 0.0236 [3360]	0.9085 ± 0.0246 [3478] (0.9051[d])
MP2	0.9527 ± 0.0448[b] [3316]	0.9411 ± 0.0245 [3360]; 0.9421 ± 0.0254[b] [3252] (0.9320,[f] 0.9427,[g] 0.9427,[d] 0.9434[b,d])	0.9356 ± 0.0255[b] [3407] (0.9370[b,d])	0.9398 ± 0.0285[b] [3399]	0.9502 ± 0.0250[b] [3416] (0.9496[b,d])
QCISD[b]	0.9680 ± 0.0467 [2947]	0.9519 ± 0.0202 [2263] (0.9537[d])	0.9396 ± 0.0205[b] [3409]	0.9433 ± 0.0230 [3448]	0.9541 ± 0.0200 [3263]
BLYP	0.9923 ± 0.0472 [3348]	0.9903 ± 0.0253 [3210] (0.995,[h] 0.9940,[i] 0.9945[d])	0.9910 ± 0.0253 [3447]	0.9933 ± 0.0274 [3398]	0.9962 ± 0.0259 [3298]
B3LYP	0.9627 ± 0.0403 [3325]	0.9594 ± 0.0200 [3310] (0.963,[h] 0.9613[i])	0.9611 ± 0.0201 [3468]	0.9632 ± 0.0211 [3347]	0.9669 ± 0.0205 [3474] (0.9679[j])
B3PW91	0.9595 ± 0.0371 [3387]	0.9557 ± 0.0196 [3276] (0.9573[d])	0.9575 ± 0.0207 [3446]	0.9592 ± 0.0218 [2854]	0.9631 ± 0.0208 [3482]
mPW1PW91	0.9534 ± 0.0391 [3486]	0.9491 ± 0.0194 [3433]	0.9508 ± 0.0204 [3467]	0.9525 ± 0.0213 [3412]	0.9568 ± 0.0205 [3436]
PBEPBE	0.9885 ± 0.0431 [3299]	0.9843 ± 0.0237 [3377]	0.9851 ± 0.0238 [3317]	0.9874 ± 0.0254 [3327]	0.9910 ± 0.0248 [3447]

[a] The number of independent vibrational frequencies in each data set is given between square brackets. Literature values are between parentheses. [b] Frozen-core approximation. [c] Reference 6. [d] Reference 11. [e] Reference 7. [f] Reference 8. [g] Reference 9. [h] References 43 and 44. [i] Reference 10. [j] Reference 14.

molecules, vibrational frequencies may overlap. Moreover, overtones (harmonics) and combination bands (cross-harmonics) may be difficult to distinguish from the fundamental frequencies. Such complications may lead to mistaken assignments. Another common mistake is interchange of two assignments. This mistake is often revealed upon comparison with the frequencies computed using quantum chemistry models. Experimental frequencies may also be assigned to the wrong molecule. This is harder to detect and cannot be corrected by using computational data alone. Finally, the experimental sample may include a mixture of conformations that may lead to incorrect assignments.

Computational mistakes include an inappropriate choice of the quantum chemistry model or incorrect execution of the appropriate calculations. For example, a quantum chemistry model based on Hartree−Fock theory would be inappropriate for weakly bound van der Waals molecules because such molecules are held together by forces that are not included in Hartree−Fock theory. More subtly, the harmonic vibrational model itself is inappropriate for multiple-well potentials such as soft torsions or inversions. This is the reason that the largest values of the correction, c_i, are usually for torsional vibrations. Examples of incorrect execution are use of the wrong conformation and the failure to converge to the electronic ground state of the molecule.

Suitable Class of Vibrational Frequencies. As discussed in section 3, the main task in determining a scaling factor and its associated standard uncertainty is to identify an appropriate class of reference vibrational frequencies within the CCCBDB. This is challenging because the three required characteristics are somewhat vague: (i) the biases within the class are believed to be similar to that for the target frequency, (ii) the estimated biases within the class comprise an approximately normal distribution without an excessively large spread, and (iii) the cardinality of the class is sufficiently large. Classifying vibrational modes as stretch, bend, or torsion is advantageous when individual force constants are scaled[5] and would probably also be helpful for frequency scaling. However, for the present study, we choose to make no distinctions among frequencies; that is, we consider all vibrational frequencies as a single class. This is to conform to the current common practice as discussed in section 1.

5. Results and Discussion

Choice of Quantum Chemistry Models. The CCCBDB contains vibrational frequencies computed using over 140 quantum chemistry models.[17] For this study, we have selected 40 models, listed in Table 1. The selected models are all the combinations of the basis sets 3-21G, 6-31G(d), 6-31G(d,p), 6-31+G(d,p), and 6-311G(d,p) with the theoretical methods HF (Hartree−Fock), MP2 (second-order Møller−Plesset perturbation theory), and QCISD (quadratic configuration interaction including single and double excitations) and the density functionals BLYP (Becke exchange[29] with Lee−Yang−Parr correlation[30]), B3LYP (Becke three-parameter hybrid exchange[31] with LYP correlation), B3PW91 (B3 exchange with Perdew−Wang correlation[32]), mPW1PW91 (modified[33] Perdew−Wang exchange[34] and PW91 correlation), and PBEPBE (Perdew−Burke−Ernzerhof exchange and correlation[35,36]). The principal quantities in Table 1 are the scaling factors c_0, computed using eq 17, and their associated uncertainties $u(c_0)$, computed using eq 21. When available, values of c_0 from the literature are shown in Table 1 between parentheses. Because the CCCBDB remains under continual development, the data sets for the various models include different numbers of vibrational frequencies, shown in Table 1 between square brackets.

Robustness of Scaling Factors and Their Associated Uncertainties. As discussed in section 3, the class used to determine a scaling factor should have estimated biases that display an approximately normal distribution without an excessively large spread. Consider the HF/6-31G(d) model, which is one of the most popular models in current use. The distribution of estimated biases, b_i, in the CCCBDB for this model is shown in Figure 1. It appears approximately normal; however, the spread is rather large. Moreover, it is skewed by several large values of b_i. Such “outliers” are often due to mistakes, as discussed in section 4. The skewness is conveniently characterized by the coefficient of skewness, $\eta_3 = [\sum(b_i - \mu)^3/m]/\sigma^3$, which is zero for a normal distribution.[37] For the data of Figure 1, $\eta_3 = 0.44$. To investigate whether the skewness and large

TABLE 2: Influence of Outliers in the b_i Distribution on Scaling Factor c_0 and Associated Uncertainty $u(c_0)$[a]

truncation	m	σ'	η_3	c_0	$u(c_0)$
none	3508	0.0641	0.44	0.8982	0.0230
$\pm 10\sigma$	3507	0.0629	0.92	0.8982	0.0229
$\pm 8\sigma$	3501	0.0582	0.62	0.8982	0.0228
$\pm 6\sigma$	3492	0.0542	0.46	0.8982	0.0226
$\pm 4\sigma$	3468	0.0480	0.17	0.8983	0.0215
$\pm 3\sigma$	3438	0.0437	0.16	0.8984	0.0206
$\pm 2\sigma$	3359	0.0374	0.07	0.8987	0.0188
$\pm 1\sigma$	3031	0.0256	0.08	0.8990	0.0153

[a] Results are for the HF/6-31G(d) data set. For each truncated distribution, the number of data is m, the standard deviation is σ', and the coefficient of skewness is η_3.

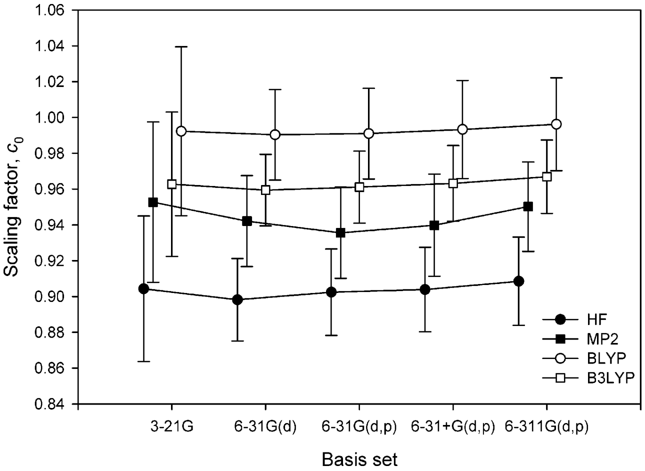

Figure 2. Scaling factors, c_0, for HF, MP2, BLYP, and B3LYP theories in conjunction with five basis sets. Error bars represent standard uncertainties. The small horizontal offsets are for visual clarity.

spread affect the results for c_0 and $u(c_0)$, we truncate the b_i distribution at various multiples of σ. The results are listed in Table 2. We note that even severe truncation at $\pm 3\sigma$ has a negligible effect on the value of the scaling factor c_0 and decreases the standard uncertainty $u(c_0)$ by only 10%.

Trends in Scaling Factors and Their Associated Uncertainties. Table 1 is the largest table of vibrational scaling factors yet assembled, revealing some simple patterns. Some of the results are presented graphically in Figure 2 to show the trends more clearly. For a given theoretical method, the value of the scaling factor, c_0, depends only weakly upon the basis set. This was noted previously for the B3LYP density functional.[14] QCISD (not displayed in Figure 2) and MP2 show the strongest basis set dependence. This is as expected, because only these two methods make use of the virtual orbitals, which are affected more than the occupied orbitals when the basis set is enlarged.

The uncertainties $u(c_0)$ are about twice as large for the 3-21G basis set as for the other basis sets, all of which include polarization functions. This suggests that polarization functions are important for avoiding markedly poor predictions of vibrational frequencies. In support of this suggestion, data from a recent study[14] show that rms drops more than 2-fold when polarization functions are added to 6-311G basis sets. Likewise, for B3LYP/6-31G(d), we have $u(c_0) = 0.02$ (Table 1), but for B3LYP/6-31G, we find[17] $u(c_0) = 0.04$.

The values of $u(c_0)$ are surprisingly insensitive to the theoretical method employed. For example, $u(c_0) = 0.023$ for HF theory and $u(c_0) = 0.020$ for QCISD theory (both with the 6-31G(d) basis set) despite the neglect of electron correlation in HF theory and the high-level treatment of correlation in QCISD theory. This suggests that poor performance for outliers is due to inadequacy of the harmonic oscillator model, which is common to all the computations summarized in Table 1. Likewise, even higher-level calculations have large errors after scaling[38] but are reliable after correcting for anharmonicity.[39]

Application Examples. The scaling factors and their associated uncertainties proposed here are appropriate for predicting the fundamental vibrational frequencies of molecules that are well represented by those in the CCCBDB. They are not intended for predicting zero-point energies or thermodynamic quantities or for predicting frequencies of molecules such as metal clusters or weakly bound complexes.

Consider the molecule BF_2OH, which is not included in the CCCBDB. Its vibrational frequencies have been measured experimentally in a cryogenic neon matrix.[40] The highest frequency in the molecule (ν_1) is the O−H stretching vibration. An HF/6-31G(d) calculation yields the unscaled prediction $x_0 = 4135.0$ cm^{-1}. Using the scaling factor $c_0 = 0.8982$ and the uncertainty $u(c_0) = 0.0230$ from Table 1, we obtain a virtual measurement of $y_0 = 3714$ cm^{-1} with a standard uncertainty of $u(y_0) = 95$ cm^{-1}. For comparison, the corresponding experimental measurement is $z_0 = 3712.5$ cm^{-1} with a standard uncertainty of $u(z_0) = 0.1$ cm^{-1}. The lowest frequency in the molecule (ν_7) is the in-plane BOH bend. The HF/6-31G(d) calculation yields the unscaled prediction $x_0 = 470.7$ cm^{-1}, which becomes the virtual measurement of 423 cm^{-1} with a standard uncertainty of 11 cm^{-1}. This is compared with the experimental measurement of 447.5 cm^{-1} with a standard uncertainty of 0.1 cm^{-1}.

6. Conclusions

The uncertainties presented in Table 1 reveal that the scaling factors are accurate to only two significant figures. The common practice of reporting four significant figures overstates the precision of vibrational scaling factors.

Table 1 and Figure 2 show that the scaling factors depend only weakly upon the basis set. Including d-functions reduces the uncertainty, $u(c_0)$, but the scaling factor, c_0, is essentially unchanged. Thus, these scaling factors may be used for basis sets not included in Table 1. For example, the scaling factor for HF calculations using the cc-pVTZ basis sets[41,42] can be assumed equal to that for the closest calculation in Table 1, HF/6-311G(d,p): $c_0 = 0.91 \pm 0.02$. As a check, explicit HF/cc-pVTZ calculations yield the scaling factor $c_0 = 0.9099$ and the standard uncertainty $u(c_0) = 0.0250$,[17] in agreement with this estimate.

The uncertainties, $u(c_0)$, in scaling factors represent the standard deviations of their respective state-of-knowledge probability distributions. For a predicted vibrational frequency, the estimated uncertainty $u(y_0)$ (eq 23) likewise represents a standard deviation. However, until the underlying probability distributions have been characterized, it is not possible to obtain probabilistic uncertainty intervals. Improved methods for classifying vibrational frequencies will probably lead to distributions that are more nearly normal and to smaller uncertainties.

The thorough study by Scott and Radom included scaling factors not only for vibrational fundamentals but also for molecular properties that depend on the vibrational partition function.[11] The present study of uncertainties is currently being extended to include such properties.

References and Notes

(1) Kosmutza, K.; Pulay, P. *Theor. Chim. Acta* **1975**, *37*, 67−75.
(2) Blom, C. E.; Altona, C. *Mol. Phys.* **1976**, *31*, 1377−1391.

(3) Török, F.; Hegedüs, Á.; Kósa, K.; Pulay, P. *J. Mol. Struct.* **1976**, *32*, 93−99.

(4) Botschwina, P. *Chem. Phys. Lett.* **1974**, *29*, 580−584.

(5) Baker, J.; Jarzecki, A. A.; Pulay, P. *J. Phys. Chem. A* **1998**, *102*, 1412−1424.

(6) Pople, J. A.; Schlegel, H. B.; Krishnan, R.; Defrees, D. J.; Binkley, J. S.; Frisch, M. J.; Whiteside, R. A.; Hout, R. F.; Hehre, W. J. *Int. J. Quantum Chem. Symp.* **1981**, *15*, 269−278.

(7) Pople, J. A.; Head-Gordon, M.; Fox, D. J.; Raghavachari, K.; Curtiss, L. A. *J. Chem. Phys.* **1989**, *90*, 5622−5629.

(8) Hout, R. F.; Levi, B. A.; Hehre, W. J. *J. Comput. Chem.* **1982**, *3*, 234−250.

(9) Pople, J. A.; Scott, A. P.; Wong, M. W.; Radom, L. *Isr. J. Chem.* **1993**, *33*, 345−350.

(10) Wong, M. W. *Chem. Phys. Lett.* **1996**, *256*, 391−399.

(11) Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502−16513.

(12) Halls, M. D.; Velkovski, J.; Schlegel, H. B. *Theor. Chem. Acc.* **2001**, *105*, 413−421.

(13) Sinha, P.; Boesch, S. E.; Gu, C. M.; Wheeler, R. A.; Wilson, A. K. *J. Phys. Chem. A* **2004**, *108*, 9213−9217.

(14) Andersson, M. P.; Uvdal, P. *J. Phys. Chem. A* **2005**, *109*, 2937−2941.

(15) Panchenko, Y. N.; Pupyshev, V. I.; Bock, C. W. *J. Mol. Struct.* **2000**, *550*, 495−504.

(16) Berezin, K. V.; Nechaev, V. V.; Krivokhizhina, T. V. *Opt. Spectrosc.* **2003**, *94*, 357−360.

(17) Johnson, R. D., III. *NIST Computational Chemistry Comparison and Benchmark Database,* version 11; NIST Standard Reference Database Number 101, May 2005; http://srdata.nist.gov/cccbdb/.

(18) Shimanouchi, T. *Tables of Molecular Vibrational Frequencies, Consolidated*; U.S. GPO: Washington, DC, 1972; Vol. I.

(19) *Guide to the Expression of Uncertainty in Measurement*, 2nd ed.; International Organization for Standardization (ISO): Geneva, 1995.

(20) Kacker, R.; Jones, A. *Metrologia* **2003**, *40*, 235−248.

(21) Irikura, K. K.; Johnson, R. D., III; Kacker, R. N. *Metrologia* **2004**, *41*, 369−375.

(22) Stuart, A.; Ord, J. K. *Kendall's Advanced Theory of Statistics: Distribution Theory*, 5th ed.; Oxford University Press: New York, 1987.

(23) Certain commercial materials and equipment are identified in this paper to specify procedures completely. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that the material or equipment identified is necessarily the best available for the purpose.

(24) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; 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*, revision B.02; Gaussian, Inc.: Wallingford, CT, 2004.

(25) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347−1363.

(26) Huber, K. P.; Herzberg, G. *Molecular Spectra and Molecular Structure: IV. Constants of Diatomic Molecules*; van Nostrand Reinhold: New York, 1979.

(27) *Vibrational and Electronic Energy Levels of Polyatomic Transient Molecules*; Jacox, M. E., Ed.; National Institute of Standards and Technology: Gaithersburg, MD, 2003.

(28) Sverdlov, L. M.; Kovner, M. A.; Krainov, E. P. *Vibrational Spectra of Polyatomic Molecules*; Wiley: New York, 1974.

(29) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.

(30) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.

(31) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(32) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244−13249.

(33) Adamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664−675.

(34) Perdew, J. P. In *Proceedings of the 21st Annual International Symposium on the Electronic Structure of Solids*; Ziesche, P., Eschrig, H., Eds.; Akademie: Berlin, 1991.

(35) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.

(36) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396.

(37) Evans, M.; Hastings, N.; Peacock, B. *Statistical Distributions*, 3rd ed.; Wiley: New York, 2000.

(38) Irikura, K. K. *Chem. Phys. Lett.* **2005**, *403*, 275−279.

(39) Helgaker, T.; Ruden, T. A.; Jørgensen, P.; Olsen, J.; Klopper, W. *J. Phys. Org. Chem.* **2004**, *17*, 913−933.

(40) Jacox, M. E.; Irikura, K. K.; Thompson, W. E. *J. Chem. Phys.* **2000**, *113*, 5705−5715.

(41) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007−1023.

(42) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *98*, 1358−1371.

(43) Rauhut, G.; Pulay, P. *J. Phys. Chem.* **1995**, *99*, 3093−3100.

(44) Rauhut, G.; Pulay, P. *J. Phys. Chem.* **1995**, *99*, 14572.