The Elusive Equilibrium Bond Length in Organic Polyatomic Molecules: Finally Obtainable from Spectroscopy?

MARLIN D. HARMONY

Department of Chemistry, University of Kansas, Lawrence, Kansas 66045

Received January 22, 1992

While numerous properties contribute to our understanding of molecules, surely the most characteristic attribute of a molecule is its three-dimensional structure, i.e., the geometrical arrangement of its constituent atoms. There can be no doubt that the development of chemistry in the 20th century has been paralleled, if not led, by advancement in our quantitative knowledge of molecular structure. When Lewis[1] published his famous 1923 monograph (*Valence and the Structure of Atoms and Molecules*), the fields of X-ray crystallography and molecular spectroscopy were in their infancy. Among the few interatomic distances quoted by Lewis were the very new gas-phase spectroscopic results for HF, HCl, and HBr, namely, 0.92, 1.26, and 1.40 Å, respectively. By 1939, when Pauling[2] published his classic work *The Nature of the Chemical Bond*, the situation had changed markedly. A large volume of X-ray data was available for crystals, and a smaller but growing body of gas-phase data from molecular spectroscopy and electron diffraction was beginning to appear. By now it was firmly established that carbon–carbon single, double, and triple bonds had characteristic lengths of ca. 1.54, 1.33, and 1.20 Å, respectively, and extensive tables of atomic covalent and ionic radii had been developed. Still, at this time there existed only a relatively small number of *precision* gas-phase structures for polyatomic organic molecules having more than a couple of inequivalent bond lengths. The principal era for the extensive determination of

gas-phase structures of polyatomic organic molecules began in 1945 with the new microwave technology developed in World War II. Since that time, structural results of various degrees of completion and precision have been reported for many hundreds of molecules by means of microwave (pure rotational) spectroscopy.[3] At the same time, theoretical and instrumental advances led to a greatly increased body of gas electron diffraction structures, and a spirited and mostly productive, several decade competition was underway.

Our primary goal in the next section is to summarize and evaluate concisely, but in some detail, the principal methodologies of structure determination from gas-phase spectroscopic data. We shall not discuss the numerous important contributions of gas electron diffraction, but instead refer the reader to an excellent review.[4] Finally, solid-state structures will not be discussed, since our interests lie with isolated molecules essentially free of intermolecular perturbations.

Molecular Structure Theory

The seminal work of Born and Oppenheimer[5] laid the foundation for the quantum mechanical treatment of molecular structure. By methods which are by now well established, the Hamiltonian is separated into two parts, one depending upon electronic variables (with clamped nuclei) and one depending upon nuclear variables (with a "potential" energy obtained from the first problem). Approximate solution of the clamped nuclei

(1) Lewis, G. N. *Valence and the Structure of Atoms and Molecules*; The Chemical Catalog Co.: New York, 1923.
(2) Pauling, L. *The Nature of the Chemical Bond*; Cornell University Press: Ithaca, NY, 1948.
(3) For a modern treatise on microwave spectroscopy, see, for example: Gordy, W.; Cook, R. L. *Microwave Molecular Spectra*; Wiley-Interscience: New York, 1984.
(4) Kuchitsu, K.; Cyvin, S. J. *Molecular Structure and Vibrations*; Cyvin, S. J., Ed.; Elsevier: Amsterdam, 1972; pp 183–211.
(5) Born, M.; Oppenheimer, J. R. *Ann. Phys.* 1927, 84, 457–484.

electronic Hamiltonian by the methods of ab initio quantum theory leads to the Born–Oppenheimer potential energy surface given for a state n by $U^{(n)} = E_{el}^{(n)} + U_{NN}$, where the first term on the right is the purely electronic energy and the second term represents the nuclear repulsion energy. $U^{(n)}$ can be considered to be a function of the $3N - 6$ (or 5) internal structural parameters $[r_i, \theta_j, i = 1 \ldots n_1, j = 1 \ldots n_2, n_1 + n_2 = 3N - 6$ (or 5)] defining the nuclear geometry. The global minimum of $U^{(n)}(r,\theta)$ in $3N - 6$ dimensional space yields the total energy $U_e^{(n)}$ and structure $r_{1e}, r_{2e}, \ldots, \theta_{1e}, \theta_{2e}, \ldots$ in the equilibrium (e) configuration.

The second half of the complete Hamiltonian leads to the energy states associated with the nuclear motions. Subject to appropriate approximations, this Hamiltonian separates into a rotational and a vibrational energy problem, with parameters involving the potential energy surface $U^{(n)}(r,\theta)$.[6] By probing the rotational and vibrational states spectroscopically, one obtains access to the experimental determination of molecular structure and various potential constants (derivatives of $U^{(n)}(r,\theta)$). In order to see how the molecular structure or geometry results, we illustrate briefly the procedure for the diatomic molecule, in which case $U^{(n)} = U^{(n)}(r)$. In this case, the form

$$U^{(n)}(r) = a(r - r_e)^2 + b(r - r_e)^3 + c(r - r_e)^4 + \ldots \quad (1)$$

will be adequate for the low-lying states. It is well known that the resulting vibrational energy states follow the expression (in standard notation)

$$E_v = \omega_e(v + \tfrac{1}{2}) - x_e\omega_e(v + \tfrac{1}{2})^2 + y_e\omega_e(v + \tfrac{1}{2})^3 + \ldots \quad (2)$$

where ω_e involves the second derivative of $U^{(n)}(r)$ at $r = r_e$ (i.e., ω_e depends upon a in eq 1), and y_e depends additionally upon the higher order constants in $U^{(n)}(r)$. The states are approximately equally spaced for low v, but converge as v increases.

The rotational energy wave equation is similarly parametrized to handle the vibrational effects from $U^{(n)}(r,\theta)$ that impact the rotational energies and spectrum. For the $^1\Sigma$ diatomic molecule, the rotational energies are written

$$E_r = B_v J(J + 1) - D_v J^2(J + 1)^2 \quad (3)$$

to a suitable approximation, although more accurate forms are available. B_v and D_v are the parameters that carry the vibrational effects from $U^{(n)}(r)$. The *effective* rotational constant B_v varies from one vibrational state to another according to

$$B_v = B_e - \alpha_e(v + \tfrac{1}{2}) + \gamma_e(v + \tfrac{1}{2})^2 + \ldots \quad (4)$$

where

$$B_e = \frac{h}{8\pi^2 I_e} \qquad I_e = \mu r_e^2 \qquad \mu = \frac{m_1 m_2}{m_1 + m_2} \quad (5)$$

and α_e and γ_e are vibration–rotation constants with γ_e normally much smaller than α_e. r_e will correspond very closely to the minimum of the Born–Oppenheimer potential function $U^{(n)}(r)$ in the absence of relativistic effects or breakdown of the Born–Oppenheimer approximation. Such effects on structure are generally small (<0.0001 Å) for ground-state singlet molecules

Table I
Rotational Constants and Bond Lengths for $^{133}Cs^{35}Cl$ and $H^{35}Cl^a$

v	B_v (MHz)	r_v (Å)
	CsCl	
0	2156.191	2.9097
1	2146.092	2.9165
2	2136.013	2.9234
3	2125.955	2.9303
	$B_e = 2161.189$	$r_e = 2.9063$
	HCl	
0	312990.9	1.2839
1	303876.4	1.3030
2	294835.9	1.3228
3	285847.5	1.3434
	$B_e = 317582.6$	$r_e = 1.2746$

^{a} See ref 9.

containing first or second period elements. It is well known, however, that species with heavy atoms (e.g., Sn, Pb, I) can incur rather large contributions from these neglected terms.[7]

We arrive then at spectroscopy. From the vibrational–rotational band structure of infrared or optical (UV–visible) spectra, from Raman spectra, or from pure rotational (microwave, commonly) spectra, it is essentially straightforward to determine B_v for diatomic molecules in several vibrational states. Thus, several of the rotational constants, $B_0, B_1, B_2, \ldots$, can typically be evaluated, leading to the determination of B_e, α_e, γ_e, $\ldots$ and thus the internuclear distance r_e. By 1950, such measurements had already been extensively reported, so that Herzberg was able to tabulate over 250 experimental r_e values for the ground and numerous excited states of more than 100 diatomic molecules.[8] Thus, the diatomic molecule data bank had grown substantially in the 27 years since Lewis[1] reported the new results for HF, HCl, and HBr. While Herzberg's reported r_e values for these species, 0.9171, 1.2746, and 1.4138 Å, differed from the earlier report by less than 1%, the precision and meaning of the values improved immensely.

Although our interest is ultimately in polyatomic rather than diatomic molecules, it is useful to consider results for two typical diatomics, viz., $^{133}Cs^{35}Cl$ and $^1H^{35}Cl$. From experimental data for B_v in various low vibrational states,[9] it is possible to compute not only B_e and thus r_e but also r_v, the characteristic effective bond length in the state v given by

$$r_v = (I_v/\mu)^{1/2} \quad (6)$$

The results summarized in Table I show that r_e is very nearly the value of r obtained by extrapolation of the r_v values to $v = -1/2$. The increase in r_v as v increases is the natural result of the increased vibrational amplitude in the excited vibrational states. It can be shown, in fact, that to a high approximation

$$r_v = \langle 1/r^2 \rangle_v^{-1/2} \quad (7)$$

that is, r_v, often called the *effective* bond length in the state v, is the inverse square root of the average of $1/r^2$ in the vibrational state v. The data in Table I show that the vibrational effects are substantial, especially for a

(6) For an elementary and clear description of this procedure, see: Wilson, E. B.; Decius, J. C.; Cross, P. C. *Molecular Vibrations*; McGraw-Hill: New York, 1955.

(7) Pyykkö, P. *Chem. Rev.* **1988**, *88*, 563–594.

(8) Herzberg, G. *Spectra of Diatomic Molecules*; Van Nostrand: Princeton, NJ, 1950.

(9) Lovas, F. J.; Tiemann, E. *J. Phys. Chem. Ref. Data* **1974**, *3*, 609–770.

case such as HCl involving the very light hydrogen atom, for which r_v varies 0.02 Å per vibrational quantum. It is also important to point out that r_e values are isotopically invariant to a high approximation, which should be the case if the Born–Oppenheimer approximation is valid. Thus, the experimental r_e values for $D^{35}Cl$, $T^{35}Cl$, and $H^{37}Cl$ are identical to the $H^{35}Cl$ value to within a precision of 1×10^{-4} Å.

On the other hand, r_v values are not generally isotopically invariant; e.g., for DCl and TCl the r_0 values are 1.2812 and 1.2801 Å, respectively, while for HCl Table I shows a value of 1.2839 Å. Qualitatively, the states of the heavier isotopomers sink more deeply into the potential well, thus approaching more closely the vibrationless equilibrium state. We see, then, that even for diatomic molecules, the concept of "structure" needs to be rather carefully defined. For any given molecule, $r_e, r_0, r_1, ...$ have different values. And while one expects and finds that r_e values are isotopically invariant, one also finds that r_0 values are not.

There are still other ways of defining and computing internuclear distances. For a hypothetical vibrationless (and isotopically invariant) diatomic molecule, it is easily shown that the magnitude of the z-coordinate of the ith atom in the center of mass axis system of a "parent" or reference isotopomer can be computed from[10]

$$z_i = (\Delta I / \mu')^{1/2} \qquad (8)$$

Here ΔI is the difference in moments of inertia of the parent isotopomer of mass M and the isotopomer having nucleus i of mass m replaced by an isotope of mass $m + \Delta m$, and $\mu' = M\Delta m/(M + \Delta m)$. Then from two such substitutions, one for each nucleus, one straightforwardly obtains the internuclear distance. Of course, for the hypothetical *rigid* molecule, such a procedure will merely identically reproduce the isotopically invariant distance obtainable from each individual moment of inertia. Similarly, if the method is applied to the *equilibrium* moments of inertia, I_e, of a real molecule such as HCl, one will again reproduce (to within $\sim$0.0001 Å) the isotopically invariant distance r_e. On the other hand, if one were to apply eq 8 to the $v = 0$ moments of inertia, I_0, for $H^{35}Cl$, $D^{35}Cl$, and $H^{37}Cl$, the *substitution*[11] bond length $r_s = 1.2783$ Å would be obtained. Note that the value lies between r_e and r_0, a result shown by Costain to be generally true for diatomic molecules.[11]

Of course, there is little point in carrying out such r_s evaluations for diatomic molecules since the required experiments on three isotopomers involve perhaps more experimental effort than merely evaluating B_e (from a minimum of two vibrational states) and hence r_e for a *single* isotopomer. The method has, however, very practical applications for polyatomic molecules, as we shall now illustrate.

The rotational energy levels of a polyatomic molecule in any vibrational state are now functions of three rotational constants, A_v, B_v, and C_v, referred to the three molecular principal axes a, b, and c. For a molecule with no degenerate vibrational modes, the A rotational constant varies in the vibrational states of the $3N - 6$ normal modes according (in analogy to eq 4) to

(10) Kraitchman, J. *Am. J. Phys.* **1953**, *21*, 17–24.
(11) Costain, C. C. *J. Chem. Phys.* **1958**, *29*, 864–874.

$$A_v = A_{v_1 v_2 ... v_{3N-6}} = A_e - \sum_i^{3N-6} \alpha_i^a(v_i + \tfrac{1}{2}) + ... \qquad (9)$$

with analogous equations for B_v and C_v. To this approximation, it is clear that the determination of the *equilibrium* rotational constants for a given isotopic species requires the measurement of A_v, B_v, and C_v in a minimum of $3N - 6 + 1$ vibrational states. For molecules with sufficient symmetry to permit degenerate vibrational modes, the sums in eq 9 extend only over the distinct modes, and $v_i + 1/2$ is replaced by $v_i + d_i/2$ for a mode of degeneracy d_i.

Thus, for a molecule such as SO_2 (C_{2v} symmetry), it is necessary to determine the rotational constants in *each* of *four* vibrational states such as $(v_1, v_2, v_3) = (000)$, (100), (010), and (001); for CH_4 (T_d symmetry, with four normal modes), rotational constants (actually only B for this spherical rotor) are needed in at least five vibrational states, while for a simple C_s molecule such as formic acid (HCOOH) in one of its two stable conformations, rotational constants must be measured in $3(5) - 6 + 1 = 10$ vibrational states. For SO_2, the resulting A_e, B_e, and C_e values would permit (with redundancy) the evaluation of both the SO bond distance and OSO bond angle, and the single B_e value for CH_4 would permit evaluation of the CH bond distance. Indeed, these experimental results are available for these two molecules. On the other hand, for formic acid, with eight independent structural parameters under C_s symmetry, one set of A_e, B_e, C_e values for the normal isotopic species is clearly not sufficient for the evaluation of eight parameters. In fact, it is necessary to study two additional isotopic species (such as DCOOH and $H^{13}COOH$), which would yield nine rotational constants for determination of the eight parameters. Thus, an equilibrium structure for just one formic acid conformer requires complete spectral studies in a total of 30 vibrational states, 10 for each of the 3 isotopomers. This formidable task has never been accomplished for a species such as formic acid, and it is representative in a small way of the problem faced in the experimental evaluation of equilibrium (r_e) structures of polyatomic molecules.

Thus, except for a goodly number of linear and nonlinear triatomic molecules and a very few simple and highly symmetric polyatomic molecules such as NH_3, CH_4, and CH_3F, spectroscopists have been forced to report generally either r_0 (effective) or r_s (substitution) structures. In the former method, rotational constants are evaluated in the ground vibrational (all $v_i = 0$) state (only) for a number of isotopic species sufficient to yield at least as many independent A_0, B_0, C_0 values as unknown parameters. Such r_0 structures (obtained by nonlinear least-squares fitting) are notoriously unreliable because of the residual zero-point vibrational effects. They exhibit bond distances and angles which may be either larger or smaller (although distances are more often larger) than the r_e values by as much as 1–2% or more. Moreover, the structures are strongly dependent upon the selection of isotopomers, and there is *no* theoretical meaning to the resulting parameters as there is for diatomic molecules (see eq 7).

Because of the inherent problems with r_0 structures, Costain suggested the r_s substitution method outlined earlier for diatomic molecules.[11] The use of moment of inertia *differences* in Kraitchman's[10] general equa-

324 *Acc. Chem. Res., Vol. 25, No. 8, 1992*

Harmony

tions (analogous to eq 8 for diatomics) leads to some cancellation of the zero-point vibrational terms, so it is found generally that the resulting structural parameters are substantially more consistent than those obtained from r_o structures. The method requires that ground-state rotational constants be obtained for a parent (usually the normal) isotopomer as well as for all of the monosubstituted isotopomers. Thus, for the one conformational form of formic acid, rotational spectral analyses are needed for $H^{12}C^{16}O^{16}OH$, $D^{12}C^{16}O^{16}OH$, $H^{12}C^{16}O^{16}OD$, $H^{13}C^{16}O^{16}OH$, $H^{12}C^{17}O^{16}OH$, and $H^{12}C^{16}O^{17}OH$. While some challenging isotopic enrichment chemistry may be involved, the six spectral data sets are substantially easier to obtain than the 30 data sets needed for an r_e structure. For molecules with symmetrically equivalent atoms, such as formaldehyde ($O{=}CH_2$), only one of the equivalent hydrogen atoms need be substituted, and moreover, it is possible to obtain substitution coordinates from multiple isotopic substitution (such as $O{=}CD_2$), which may be chemically simpler. The method still suffers from incompletely cancelled zero-point vibrational effects, especially for hydrogen atom parameters, and as for r_o parameters, there is no strict theoretical significance to the parameters. Nevertheless, for the past 30 years, the r_s structure has been the best commonly available spectroscopic structure for most polyatomic molecules.[12] The results have been coherent overall (with heavy-atom (non-hydrogen) distances often, but not always, within perhaps 0.005 Å of the expected but unattainable r_e values) and have played an important role in shaping our understanding of many of the details of structural chemistry. As ab initio molecular quantum chemistry has matured during this same time span, spectroscopic r_s parameters have provided important benchmarks for the quality of the computations for polyatomic organic molecules.

The situation at the present time is, unfortunately, that the remaining uncertainties in r_s parameters (say, ±0.005 Å or so) represent a noise level of approximately the same size as that which is now achievable from the best ab initio computations on small molecules. Because of this, it has become increasingly difficult to definitively and carefully evaluate and further calibrate the results of ab initio computations for polyatomic molecules. Thus, important remaining questions are what size basis set (6-31G, 6-31G*, 6-31G**, 6-311G**, etc.) and what level of electron correlation (MP1, MP2, CISD, etc.) are needed to achieve a reliably known structural precision for polyatomic organic molecules with several heavy (second or third period) atoms. With small basis sets, and little or no consideration of correlation, ab initio geometries are not highly reliable. On the other hand, complete geometry searches with the largest basis sets and the highest level of electron correlation remain costly and time-consuming computational endeavors in the case of polyatomic molecules with several carbon, nitrogen, or oxygen atoms. Thus, it became clear to us that it would be highly desirable if spectroscopic structures could be obtained with greater accuracy and reliability than are currently available from r_s structures. The following section

Table II
Comparison of Experimental Values of I_o, I_m, and I_s for OCS[a]

species	I_m/I_e	I_o/I_e
16-12-32	0.999 938	1.002 944
16-13-32	0.999 840	1.002 882
18-12-32	0.999 989	1.002 906
16-12-34	0.999 971	1.002 925

[a] See ref 13.

summarizes our efforts to obtain near-equilibrium structures for polyatomic organic molecules.

Near-Equilibrium Structures

Watson's Contribution (r_m Structures). In 1973, Watson described a new structural procedure which appeared to have great potential for obtaining structures very close to equilibrium.[13,14] The method begins with the application of the substitution procedure of Costain. Illustrating the method for a linear molecule, the substitution moment of inertia, I_s, is first computed from substitution coordinates obtained from eq 8:

$$I_s = \sum_i m_i z_i^2 \qquad (10)$$

Then Watson showed with beautiful clarity that, correct to first order in the isotopic mass changes,

$$I_e = I_m - \frac{1}{M}\sum_i m_i \Delta m_i \frac{\partial^2(M\epsilon)}{\partial m_i^2} + \ldots \cong I_m - \delta \quad (11)$$

in which

$$\epsilon = I_o - I_e \qquad (12)$$

is the zero-point vibration–rotation contribution to I_o and

$$I_m = 2I_s - I_o \qquad (13)$$

The m_i and Δm_i are the masses of the parent isotopomer atoms and the changes in mass, respectively, upon isotopic substitution, and M is the mass of the parent isotopomer. Thus, I_m, computed simply from eq 13 from the experimental I_o and computed I_s values of the parent isotopomer, is by eq 11 the zeroth order approximation to I_e. In fact, I_m is very close to I_e, typically within $\sim$0.005% if hydrogen-containing molecules are excluded. This contrasts with I_o, which typically differs from I_e by $\sim$0.5%. Stated differently, δ/ϵ (see eqs 11 and 12) typically has a value of $\sim$0.01. The significant result is that the simple procedure outlined, based only upon measurements of ground-state moments of inertia, I_o, for the parent and all necessary monosubstituted isotopomers, leads to moments of inertia I_m remarkably close to I_e. Table II summarizes experimental values of I_o/I_e and I_m/I_e for several OCS isotopomers, clearly showing the excellent quality of the I_m values.[13] The procedure is easily generalized to any polyatomic molecule, in which case I_o, I_s, and I_m values will be defined for the three principal axes a, b, and c. In addition, for species with two or more equivalent atoms, multiple isotopic substitution equations can be utilized in place of Kraitchman's single substitution equations.[15] An interesting and fundamentally important aspect of the procedure is that *imaginary* substitution coordinates are retained in the

(12) There are still other, less common, types of spectroscopic structures. For example, the r_z structure provides bond lengths which represent the separation of atoms in their average ground-state positions. See: Oka, T. *J. Phys. Soc. Jpn.* 1960, *15*, 2274–2279.

(13) Watson, J. K. G. *J. Mol. Spectrosc.* 1973, *48*, 479–502.
(14) Smith, J. G.; Watson, J. K. G. *J. Mol. Spectrosc.* 1978, *69*, 47–52.
(15) Chutjian, A. *J. Mol. Spectrosc.* 1964, *14*, 361–370.

The Equilibrium Bond Length in Organic Molecules

Acc. Chem. Res., Vol. 25, No. 8, 1992 **325**

Table III
Structures by Various Methods^a

	r_s	r_m	r_m^ρ	r_e
		OCS		
CO	1.1605	1.1587	1.1551	1.1543
CS	1.5596	1.5593	1.5621	1.5628
		N_2O		
NN	1.1466	1.1281	1.1277	1.1284
NO	1.1695	1.1842	1.1846	1.1841
		$COCl_2$		
CO	1.1836	1.1808	1.1767	1.1756
CCl	1.7346	1.7351	1.7371	1.7381
ClCCl	112.24	112.08	111.89	111.79

^{a} Distances in Å, angles in degrees. See ref 16 for original data and details.

calculation of I_s. Such coordinates lead to negative terms in eq 10 and provide meaningful and necessary mass-dependent vibration–rotation contributions to I_s and hence to I_m.

Since the I_m values are very accurate approximations to I_e, one might reasonably expect that they will lead to accurate near-r_e structures. Indeed, for linear molecules involving only one parameter and thus requiring only a single I_m value, such as CO or CO_2, or for C_{2v} nonlinear molecules such as SO_2 which involve two structural parameters but require only the $(I_m)_a$ and $(I_m)_b$ moments for a single parent isotopomer, the (distance) structural parameters are typically within 0.0001 Å (and an analogous angle precision) of the true r_e values. Unfortunately, when the method is extended to molecules requiring I_m values for two or more parent isotopomers, the results deteriorate substantially. Thus, Watson's[13] reported OCS r_m structure in Table III shows errors of +0.4% in C=O and –0.2% in C=S, even though the I_m values are all within ~0.02% of the I_e values. The origin of this distorted structure lies in the neglected term δ in eq 11. The terms in the sum fluctuate in magnitude and more importantly in sign (depending upon the substitution pattern) for the various parent species, thus leading to the possibility of discordant I_m values. (Note in Table II the discordant ratio for the ^{13}C species.) The problem has been discussed extensively by Watson and by our own group in several papers.[16-18] The net result is that r_m structures cannot, in general, be relied upon as good approximations to r_e structures for even small polyatomic molecules, although they may, fortuitously (see N_2O in Table III), be satisfactory. Of course, for large low-symmetry polyatomic molecules, the isotopic data requirements become entirely too unwieldy to be practical in any case.

A Pivotal Observation (r_m^ρ Structures). Having been involved for over 15 years in polyatomic molecular structure determinations at the r_o and r_s levels, the microwave group at Kansas began in the mid 1980s a search for more reliable, precise, and accurate (in the sense of agreement with or closeness to r_e structures) spectroscopic structural methods. We were led rather quickly into a detailed critical study of the application of Watson's theory. Initially, although aware of the

(16) Harmony, M. D.; Taylor, W. H. *J. Mol. Spectrosc.* **1986**, *118*, 163–173.
(17) Harmony, M. D.; Berry, R. J.; Taylor, W. H. *J. Mol. Spectrosc.* **1988**, *127*, 324–336.
(18) Berry, R. J.; Harmony, M. D. *J. Mol. Spectrosc.* **1988**, *128*, 176–194.

Table IV
I_s/I_o **Values for** $COCl_2$^{a}

isotope	$[I_s/I_o]_a$	$[I_s/I_o]_b$
16-12-35	0.998 277 8	0.998 042 2
16-12-37	0.998 262 6	0.998 084 0
18-12-35	0.998 327 8	0.998 064 8
18-12-37	0.998 305 7	0.998 086 6
16-13-35	0.998 221 2	0.997 983 2
16-13-37	0.998 198 6	0.998 011 4
18-13-35	0.998 247 0	0.997 982 0
18-13-37	0.998 222 8	0.997 987 2

^{a} From Taylor, W. H. Ph.D. Dissertation, University of Kansas, Lawrence, KS, 1986.

potential deficiencies outlined in the previous section, we became intrigued with the possibility of applying the method directly, or perhaps in some approximate form, to bona fide polyatomic molecules (more than 2–4 atoms). While we were concerned with the problems associated with the fluctuating δ terms of eq 8, we focused attention upon ways of reducing the isotopic data requirements. Upon investigating extensive data sets for several non-hydrogen-containing molecules (including OCS, N_2O, SO_2, $COCl_2$), we were immediately struck by the fact that the I_m/I_o ratio for any given molecule was independent, to a high approximation, of which isotopomer was chosen as the parent species.[16] From Watson's definition of eq 13, we can write

$$I_m/I_o = 2I_s/I_o - 1 \qquad (14)$$

Thus, the constancy of I_m/I_o traces back to a constancy of I_s/I_o. Table IV shows the results for the extensive $COCl_2$ data, wherein it is observed that, to within a few parts in 10^4, the ratios are constant for each individual axis. (The closeness of the a-axis and b-axis values is fortuitous and not a general expectation or observation.) From this observation, which was repeated numerous times for other molecules, we were led to postulate that I_s/I_o needed to be evaluated only once from a single set of conventional Costain substitution data and that the resulting value would be within ~$1/10^4$ of the true value for any other substitution data set (SDS) with a different parent.[16,17] Moreover, the single experimentally determined I_s/I_o value could then be used to compute approximate Watson I_m values for all of the isotopomers of the SDS.

We define (for a particular principal axis)

$$\rho = [I_s]_1/[I_o]_1 \qquad (15)$$

where $[I_s]_1$ is computed for a particular isotopic SDS, usually with the normal abundant isotopomer as the parent and $[I_o]_1$, the ground-state moment of the parent species. Then scaled moments of inertia are defined by

$$[I_m^\rho]_i = (2\rho - 1)[I_o]_i \qquad i = 1 \ldots L \qquad (16)$$

for all L members of the SDS. Now $[I_m^\rho]_1$ is, in fact, an exact Watson I_m value (see eq 14) and is thus within a few parts in 10^4 of I_e; moreover, since ρ is constant to within a few parts in 10^4 also, we expect that each of the $[I_m^\rho]_i$ values will be an excellent approximation to $[I_e]_i$, in general within a few parts in 10^4! Thus, the scaled moment of inertia procedure of eqs 15 and 16, using only a single conventional SDS, is expected to yield excellent absolute approximations to the I_e values for *all* the members of the SDS. Thus, our initial primary aim of reducing the quantity of necessary data

326 *Acc. Chem. Res., Vol. 25, No. 8, 1992*

Harmony

Table V
Published Heavy-Atom Test Molecule r_m^o Structures^a

molecules: OCS, N_2O, $COCl_2$, OCSe, CO_2, SO_2, S_2O, NOCl
distances: 15, with mean absolute deviation = 0.0007 Å

^{a}See refs 16 and 17.

was clearly satisfied by the procedure, and moreover, the scaled moments were expected to be of essentially the same absolute quality as Watson's I_m moments.

But now, how well do the I_m^o values fare in structural computations? One might suppose that they would perform no better than I_m values and perhaps somewhat poorer since they are merely approximations to I_m. We determine structural parameters generally from the L I_m^o values of a linear molecule SDS, the $2L$ I_m^o values (e.g., $(I_m^o)_a$ and $(I_m^o)_b$) of a planar asymmetric rotor SDS, or the $3L$ I_m^o values of a general asymmetric rotor SDS by a conventional nonlinear least-squares fit. The resulting r_m^o structural parameters for heavy-atom[19] molecules are found to be in remarkably good agreement with known r_e parameters.[16,17] Distances, as illustrated in Table III, appear to be generally within 0.001 Å (perhaps 0.002 Å at the outside) of r_e values, with comparable angle quality. Moreover, the results are generally far more reliable than r_s structural parameters and are always comparable to or better than the Watson r_m values. There appear to be essentially no structural distortions common to the r_s method, such as one long bond and one short bond.

To date, a total of eight heavy-atom molecules (summarized in Table V) for which reliable r_e structures are known have been found, with only one minor exception, to yield distance parameters within 0.001–0.002 Å of r_e. This success has been the cause of some consternation in the structural community, since the scaling lacks a rigorous theoretical foundation. Empirically, the high-quality results occur not merely and not even most importantly because the I_m^o values are within ~0.01% of the true I_e values. Rather, the success arises because the I_m^o scaled moments have, on average, a more nearly correct isotopic mass dependence than the I_m moments. In particular, if the parent member of the SDS is the "normal" isotopic species, the errors in the I_m^o values (i.e., $I_m^o - I_e$) of the substituted isotopomers of the SDS exhibit no unusual fluctuations in sign or magnitude relative to the error $(1/10^4)$ in the parent I_m^o value. Model force field computations[17] for OCS have confirmed these empirical observations and have shown further that a key requirement for the well-behaved, homogeneous behavior is that the "parent" species of the SDS be chosen such that all isotopic substitutions satisfy either $\Delta m_i > 0$ or $\Delta m_i < 0$. For many of the common light elements (H, C, N, O, Cl, S), the normal or most abundant nuclide is at the same time the nuclide of lowest mass, and consequently, the choice of the normal species as the parent of the SDS leads to the appropriate homogeneous set of scaled moments. Finally, for the asymmetric rotor, it is important to note that the scaling parameters ρ_a, ρ_b, and ρ_c (which are in general not numerically the same) appear to yield a molecule of the correct size in all three axis directions. This is a necessary ingredient in eliminating the structural distortions common to the r_0 and r_s methods

Table VI
Published r_m^o Structures of Hydrogen-Containing Test Molecules^a

molecules: HCO^+, HCN, HN_2^+, HNC, HCCH, H_2CO, CH_2Cl_2, C_2H_4
heavy-atom distances: 8, mean absolute deviation = 0.0009 Å
X–H distances: 8, mean absolute deviation = 0.0017 Å

^{a}See refs 20 and 21.

for nonlinear polyatomic molecules.

Extension to General Polyatomic Molecules (r_m^o Structures). It was clear very early that the scaling of eqs 15 and 16 was not satisfactory for molecules involving hydrogen. The problem is that due to the large zero-point vibrational effects involving hydrogen, the quality of the I_s/I_0 values involving H ↔ D substitution deteriorates by approximately a factor of 10 compared to the values involving only heavy-atom substitution. The result is that the scaling procedure of eq 16 leads to I_m^o values for deuterium species which are too small (when $\rho < 1$); that is, the scaling procedure overcorrects in the case of the deuterium species. Structures computed for hydrogen-containing molecules then exhibit distortions in *both* light-atom and heavy-atom parameters because of the discordant I_m^o values of the deuterated species. This matter has been investigated in detail both experimentally and theoretically.[18,20] A general way of handling the difficulty is to correct the $(I_m^o)^D$ values empirically by an amount Δ equivalent to a small extension of the X–D bond by an amount $\delta r_D = 0.0028$ Å. Then the corrected scaled moment for a deuterium species is given by (for the a-axis moment)

$$(I_m^o)_a{}^D{}_{corr} - (I_m^o)_a{}^D = \Delta_a = \pm 2m_D \sum_i^n (b_i \delta b_i + c_i \delta c_i) \quad (17)$$

where b_i, c_i and δb_i, δc_i are the coordinates of the deuterium atoms and the components of $\delta \vec{r}_D$, respectively, and m_D is the mass of a deuterium atom. The equation applies generally for a case in which the isotopic substitution involves replacement of n symmetry-equivalent hydrogen atoms by n deuterium atoms. Similar equations apply for the other axes with obvious permutation of a, b, c, and the plus or minus sign is selected according to whether $\rho_g < 1$ or $\rho_g > 1$, respectively.[20,21] The empirical parameter $\delta r_D = 0.0028$ Å has been selected to provide the best overall fit for five hydrogen-containing test molecules.[20] While one would suppose that the optimum value should depend upon the heavy-atom X involved in the X–H bonds, our work shows that a single global average value is sufficient to provide excellent heavy-atom parameters. Table VI summarizes the results for eight hydrogen-containing molecules, showing that the r_m^o heavy-atom distances are within 0.0009 Å of the true r_e values. Somewhat surprisingly, the X–H parameters for these same species are also very good, with a mean absolute deviation of 0.0017 Å.

These results and subsequent studies[21,22] strongly suggest that heavy-atom r_m^o distances are likely to be within 0.001–0.002 Å of the true r_e values for any hydrogen-containing polyatomic molecule. It appears, however, that the hydrogen atom parameters will not

(19) We use this term to mean basically atoms other than hydrogen, although in practice the heavy atoms include primarily elements from carbon to bromine.

(20) Berry, R. J.; Harmony, M. D. *Struct. Chem.* **1990**, *1*, 49–59.
(21) Harmony, M. D. *J. Chem. Phys.* **1990**, *93*, 7522–7523.
(22) Tam, H. S.; Choe, J.-I.; Harmony, M. D. *J. Phys. Chem.* **1991**, *95*, 9267–9272.

Table VII
Selected Values of Near-Equilibrium Bond Distances (in Å)

	C—C	C=C	C≡C	C—O	C=O	C—Cl
HCCH			1.203			
ClCCH			1.203			1.636
H_2CO					1.203	
CH_2Cl_2						1.764
HCO_2H				1.340	1.196	
H_2CCH_2		1.332				
H_3CCH_3	1.522					
CH_2ClCH_3	1.510					1.789
$CH_3CH_2CH_2$	1.496	1.334				
$CH_3CH_2CH_3$	1.522					

a See refs 21 and 22 for detailed discussion.

generally be reliable estimates of the r_e values,[21,22] although they are in many cases. This suggests that the H → D correction procedure of eq 17, while entirely adequate for the determination of heavy-atom parameters, is not sufficiently accurate to guarantee reliable r_m^ρ values for the hydrogen parameters. It is likely that some of this problem arises from the use of a single X–H bond invariant parameter $\delta r_D = 0.0028$ Å. At this stage, however, it does not seem profitable to complicate the procedure by the inclusion of a bond-dependent value for δr_D, since there is serious doubt as to whether the simple scaled-moment procedure is likely to provide an adequate assessment of the H ↔ D vibration–rotation contributions in any case. Finally, it should be stressed that the H → D correction procedure of eq 17 is absolutely essential to obtaining reliable heavy-atom distances. If the corrections are not made, the $(I_m^\rho)^D$ moments induce errors, which may be substantial, into the best fit least-squares heavy-atom parameters. Thus, for C_2H_2,[20] the uncorrected data set yields a C≡C distance which deviates from r_e by 0.0075 Å, while the data set using the H → D correction produces a distance within 0.0009 Å of r_e.

The procedures outlined in eqs 15–17 thus provide a means of obtaining *heavy-atom* structures of general polyatomic molecules which, by all available detailed evidence and substantial circumstantial evidence, are expected to be very close to the true Born–Oppenheimer r_e values. Still, there are a few caveats. First, as with any structure determination method, the r_m^ρ procedure requires that the raw data (I_0 values) be of high quality. Second, the scaling procedure is known to be potentially deficient in cases involving especially low-frequency anharmonic wagging vibrations involving hydrogen.[18] Third, it must be possible to obtain a complete set of substitution data. Thus, the r_m^ρ proce-

dure makes no provision for handling molecules with atoms such as fluorine, for which no stable isotope exists, or molecules with incomplete isotopic substitution.

Near-r_e Structures of Organic Polyatomic Molecules

The procedures outlined in the previous sections provide a practical, tractable scheme for the determination of near-r_e heavy-atom structures for polyatomic molecules from spectroscopic data. Table VII summarizes some recently obtained results for several molecules and bonds of interest in organic structural chemistry.[20-22] All of the reported distances are known or expected to be within 0.001–0.002 Å of the true equilibrium values. For several of the species, the r_m^ρ results provide the only near-r_e spectroscopic values that are likely to be obtained experimentally. The results provide in a small way, and for the first time, the opportunity for a careful and detailed assessment of environment or substituent effects upon experimental near-r_e bond lengths. Moreover, the results should provide important benchmarks for ab initio computational methods. In this regard, we call attention to the prototype r_e C–C bond length in ethane, 1.522 Å, which is substantially smaller than suggested by earlier non-r_e spectroscopic determinations, but which is actually in very good agreement with the estimated r_e value from electron diffraction.[21,23] It is interesting to note also that, among the characteristic carbon–carbon single, double, and triple bond lengths described by Pauling,[2] only the single bond length changes substantially in the equilibrium configuration.

Summary

The scaled moment of inertia structure procedure (r_m^ρ) has been demonstrated to provide an excellent approximation to the true r_e heavy-atom structure of a molecule. Finally, after more than a half-century of spectroscopic structural studies, it now appears possible to obtain near-r_e bond lengths and angles for organic polyatomic molecules of modest size (6–8 heavy atoms). The procedure is applicable to any molecule for which a *complete* conventional spectroscopic r_s structure determination is possible.

(23) Duncan, J. L.; McKean, D. C.; Bruce, A. J. *J. Mol. Spectrosc.* 1979, 74, 361–374.