

The effective fragment model for solvation: Internal rotation in formamide

Wei Chen and Mark S. Gordon

Citation: *The Journal of Chemical Physics* **105**, 11081 (1996); doi: 10.1063/1.472909

View online: <http://dx.doi.org/10.1063/1.472909>

View Table of Contents: <http://scitation.aip.org/content/aip/journal/jcp/105/24?ver=pdfcov>

Published by the AIP Publishing

Articles you may be interested in

Solvation dynamics study of 3-aminophthalimide in n-butanol solution at different temperatures

J. Chem. Phys. **105**, 9771 (1996); 10.1063/1.472847

A comparative study of first hyperpolarizabilities of the acidic and basic forms of weak organic acids in water

J. Chem. Phys. **105**, 9633 (1996); 10.1063/1.472794

On the origin of size inconsistency of the second-order state-specific effective Hamiltonian method

J. Chem. Phys. **105**, 6887 (1996); 10.1063/1.471982

Solvent effects on the potential energy surface of the 1:1 complex of water and formamide: Application of the polarizable continuum model to the study of nonadditive effects

J. Chem. Phys. **104**, 5539 (1996); 10.1063/1.471793

Theoretical study of the internal rotation in neutral and protonated glyoxal: From gas phase to aqueous solution

AIP Conf. Proc. **330**, 234 (1995); 10.1063/1.47690



AIP | APL Photonics

APL Photonics is pleased to announce
Benjamin Eggleton as its Editor-in-Chief



The effective fragment model for solvation: Internal rotation in formamide

Wei Chen and Mark S. Gordon

Department of Chemistry, Iowa State University, Ames, Iowa 50011

(Received 19 July 1996; accepted 19 September 1996)

The internal rotation of formamide with 0–5 water molecules oriented along the N–C bond has been studied by the full *ab initio* self-consistent field theory and using the effective fragment (EFP) method. For each case, the EFP geometries, harmonic vibrational frequencies, rotational barriers, and intrinsic reaction coordinates for the internal rotation are found to be in excellent agreement with their *ab initio* counterparts. The global energy minimum structures for four and five water complexes are predicted to be formamide bonded to two adjacent waters, with all water molecules in a ring. Probably due to the structural constraints, the complexes containing less than four waters have cyclic structures with the two ends of formamide connected by a sequence of water molecules. The internal rotation barrier of formamide–water complexes increases from 15.3 kcal/mol with no water to 19.0 kcal/mol with four waters and seems to saturate at four to five waters. When electron correlation corrections are added, the estimated internal rotation barrier is ~ 20 kcal/mol, in very good agreement with experimental measurements. © 1996 American Institute of Physics. [S0021-9606(96)00948-8]

I. INTRODUCTION

The development of methods for treating solvation effects is an area of increasing activity. Most of these methods are based on continuum models,¹ but several of the more recent methods include discrete solvent molecules.² Both of these approaches are important. Discrete methods permit the incorporation of individual interactions between solute (broadly defined) and solvent molecules, as well as solvent–solvent interactions. Cleverly designed algorithms^{2–4} facilitate the efficient determination of the arrangements of the solvent molecules and the ability to investigate solvation effects in excited states. Even with very efficient algorithms, however, it is difficult to include a sufficient number of discrete solvent molecules to realistically simulate bulk behavior. Ultimately, one would like to interface a discrete solvation model with a continuum method, in order to provide a broad consistent picture of solvation chemistry and dynamics.^{2(a),2(e),2(f),2(h)}

As a first step in such an endeavor, we have recently introduced the “effective fragment potential” (EFP) method for treating discrete solvent effects.^{5,6} This model starts from the *ab initio* Hamiltonian for the solute, which might be a simple molecule, or a reacting group of molecules, plus some (presumably small) number of solvent molecules. The remaining solvent molecules are then treated by adding their effects as one electron terms in the *ab initio* Hamiltonian. At present, the model has been developed at the Hartree–Fock level of theory. So, these one-electron terms represent the three types of interactions that one expects at this level of theory: Electrostatic, polarization, and exchange repulsion/charge transfer. Subsequent extensions of the model, already under development, will include correlation in the *ab initio* part of the treatment, and therefore, higher order solvent effects (e.g., dispersion, exchange dispersion).

In the present paper, following a summary of the solvation model and the computational methods in Sec. II, we

illustrate the accuracy of the method by applying it to an analysis of the internal rotation of formamide in the presence of 0–4 water molecules in Sec. III. All calculations have been performed using fully *ab initio* methods, and then by systematically replacing *ab initio* water molecules by EFP water molecules. Conclusions are presented in Sec. IV.

II. THEORY AND COMPUTATIONAL METHODS

A. The effective fragment method

As noted above, the effective fragment model treats each solvent molecule explicitly, by adding one-electron terms directly to the *ab initio* Hamiltonian,

$$\mathbf{H} = \mathbf{H}_{\text{AR}} + \mathbf{V}, \quad (1)$$

where \mathbf{H}_{AR} is the *ab initio* Hamiltonian that describes the “action region” (AR) of the system. The three one-electron terms in \mathbf{V} , representing the potential due to the solvent (fragment) molecules, correspond to electrostatic, polarization, and exchange repulsion/charge transfer interactions. For the μ th solvent molecule, the effective fragment interaction Hamiltonian with an electron in the AR is given by⁶

$$V_{\text{el}}(\mu, s) = \sum_{k=1}^K V_k^{\text{Elec}}(\mu, s) + \sum_{l=1}^L V_l^{\text{Pol}}(\mu, s) + \sum_{m=1}^M V_m^{\text{Rep}}(\mu, s), \quad (2)$$

where s represents the electronic coordinates. The three terms on the right-hand side (RHS) of Eq. (2) represent the electrostatic, polarization, and exchange potential/charge transfer interactions, respectively. Similar terms are added to represent the interactions between nuclei in the AR and solvent molecules, as well as solvent–solvent interactions. Of course, there are no exchange repulsion/charge transfer terms in the nuclear–solvent interaction. The solute (including the

desired number of solvent molecules) is explicitly treated with the *ab initio* wave function of choice, while the remainder may be represented by effective fragments.

An accurate, relatively compact, representation of the electrostatic potential in the important interaction regions is achieved using a *distributed* multipolar analysis⁷ (DMA) of the spectator charge distributions. K in the first term of Eq. (2) is the number of expansion points. In the present implementation, each nuclear center and each bond midpoint is chosen to be an expansion point, (e.g., five expansion points for a water molecule), and the expansion is extended through octupoles,

$$V_k^{\text{Elec}}(\mu, s) = \frac{q_k(\mu)q_s}{r_{sk}} - \sum_a^{x,y,z} \mu_a^k(\mu)F_a(\mathbf{r}_{sk}) - \frac{1}{3} \sum_{a,b}^{x,y,z} \Theta_{ab}^k(\mu)F_{ab}(\mathbf{r}_{sk}) - \frac{1}{15} \sum_{a,b,c}^{x,y,z} \Omega_{abc}^k(\mu)F_{abc}(\mathbf{r}_{sk}), \quad (3)$$

where q , μ , Θ , and Ω are the charge, dipole, quadrupole, and octupole, respectively, and F_a , F_{ab} , and F_{abc} are the solute electric field, field gradient, and field hessian. Since the expression in Eq. (3) is a point charge model, it must be modified to account for overlapping electron densities as the molecules (fragment and solute or fragment and fragment) approach each other. This is accomplished by multiplying the expression in Eq. (3) by a distance-dependent cutoff function,

$$V_k^{\text{Elec}}(\mu, s) \rightarrow [1 - \beta_k(\mu)e^{-\alpha_k(\mu)r_{sk}^2}]V_k^{\text{Elec}}(\mu, s). \quad (4)$$

The polarization of the spectator molecules by the electric field of the active (*ab initio*) molecules [second term in Eq. (2)] is treated by a self-consistent perturbation model⁸ employing bond and lone pair localized orbital dipole polarizabilities α_{ab}^l extracted from finite-field perturbed Hartree–Fock calculations on isolated spectator (fragment) molecules and centered at the centroids of the L localized valence orbitals ($L=4$ for water)

$$V_l^{\text{Pol}}(\mu) = - \sum_{a,b}^{x,y,z} F_a(\mathbf{r}_1)\alpha_{ab}^l(\mu)\langle F_b(\mathbf{r}_1) \rangle, \quad (5a)$$

$$\alpha_{xy}^l = \lim_{F_y \rightarrow 0} \frac{\mu_x^l(F_y) - \mu_x^l(0)}{F_y}. \quad (5b)$$

Here, F is the field due to the *ab initio* part of the system, while α_{xy}^l is a component of the polarizability of the fragment molecule.⁹

The exchange-repulsion/charge transfer interaction between the active and spectator molecules is modeled by one-electron terms in the *ab initio* Hamiltonian that have the form of simple Gaussian functions located at the spectator atom centers ($J=3$ for water)

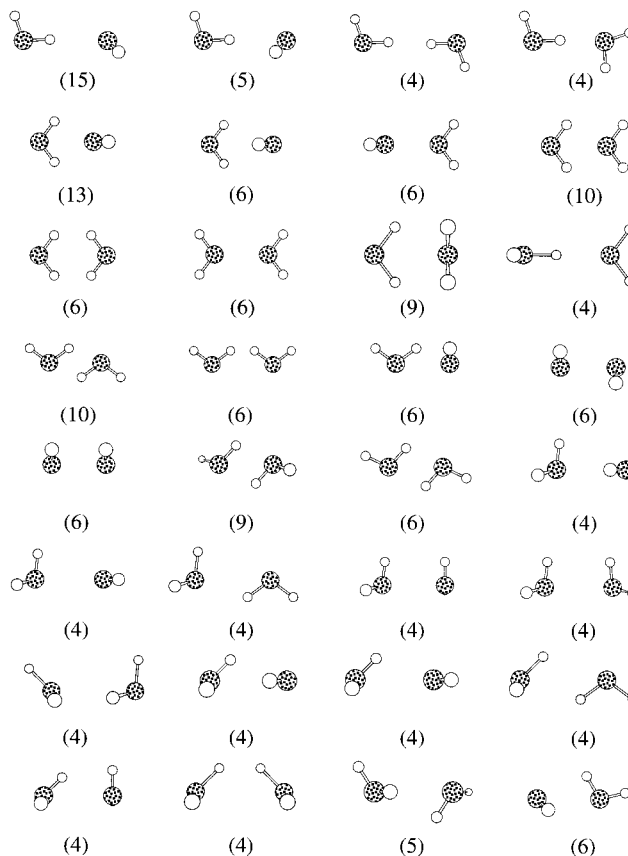


FIG. 1. A schematic of the water dimer orientations used to fit the repulsive potential. See the text for details.

$$V_m^{\text{Rep}}(\mu, s) = \sum_j^J \beta_{m,j}(\mu)e^{-\alpha_{m,j}(\mu)r_{j,s}^2}. \quad (6)$$

The Gaussian functions are optimized by a fitting procedure. *Ab initio* calculations are performed on some number of points (e.g., 192 for water dimer to represent water as the solvent). The *ab initio* exchange repulsion/charge transfer potential is obtained by subtracting the sum of electrostatic plus polarization energies from the total potential to obtain the term $E_{\text{rem}}^{(ab)}$. Then, V_m^{Rep} is fitted to $E_{\text{rem}}^{(ab)}$

$$\Delta = \sum_p^p w_p \left[\left\langle \Psi \left| \sum_m^M V_m^{\text{Rep}} \right| \Psi \right\rangle_p - E_{\text{rem}}^{(ab)}(p) \right]^2, \quad (7)$$

where w_p is a weighting factor that has been set to unity in the fitting process for water. Ψ is the wave function of AR and M is 2. The fragment–fragment interactions are treated in a similar manner, except that exponential, rather than Gaussian functions are used to represent the exchange-repulsion/charge transfer interaction.

A schematic of the water dimer orientations used to fit the repulsive potential is given in Fig. 1. The number in parentheses by each orientation represents the number of O–O distances (ranging from 2 to 5 Å) used for that orientation. For the fragment–solute interaction, a subset of 26 structures were used in the fit, while all 192 points were used to fit the fragment–fragment interaction. The quality of the

TABLE I. Statistics for repulsion parameter fitting.

	Interaction energy (kcal/mol)						
	<-2	<-1	<0	<1	<3	<10	<20
N_{sample}	35	64	93	112	143	168	177
Solute-Fragment							
rms ^a	0.34	0.66	0.57	0.56	0.60	0.91	0.97
err ^b	0.21	0.27	0.25	0.24	0.26	0.38	0.43
Fragment-Fragment							
rms	0.11	0.14	0.13	0.15	0.23	0.36	0.58
err	0.08	0.08	0.08	0.08	0.10	0.17	0.24

$$^{\text{a}}\text{rms} = \sqrt{\sum_i (\Delta E_i)^2 / (N_{\text{sample}} - 1)}.$$

$$^{\text{b}}\text{err} = |\Delta E| / N_{\text{sample}}.$$

fit may be assessed by reference to Table I. For low energy structures (interaction energies <1.0 kcal/mol), the rms error is less than 0.5 kcal/mol for the solute-fragment interaction and less than 0.2 kcal/mol for the fragment-fragment interaction. The agreement tends to deteriorate as the dimer structures move up the repulsive wall. These latter structures were not included in the solute-fragment fit, and they are unlikely to be important in most chemistry of interest.

The necessary equations have been derived⁶ and coded for the analytic gradients of the entire (*ab initio*+fragments) system. The availability of analytic gradients means that one can also perform vibrational analyses using finite differences of these gradients. So, one can determine the manner in which geometries and energetics of minima, transition states, and reaction paths (and therefore, the reaction dynamics) are modified by the presence of the solvent. The entire code described here has been added to the electronic structure code GAMESS.¹⁰

B. Geometry optimizations

The searches for the minimum energy structures (MIN) and the transition states (TS) for the internal rotation of the amino group in the formamide-water complexes were carried out in the laboratory-fixed Cartesian coordinate frame (lab frame). The potential energy surface in the lab frame for the system containing the EFP's can be expressed as a Taylor expansion,

$$E(\xi = \{\mathbf{x}, \mathbf{Q}\}) = E(\xi_0) + \mathbf{g}^T \Delta \xi + \frac{1}{2} \Delta \xi^T \mathbf{H} \Delta \xi + \dots, \quad (8)$$

where \mathbf{x} are the Cartesian coordinates for the atoms in the action, or *ab initio*, region (AR), \mathbf{Q} are the Cartesian coordinates for the effective fragment potentials (EFP's) in the fragment, or spectator region (SR), \mathbf{g} is the gradient, and \mathbf{H} is the Hessian. For the μ th EFP, \mathbf{Q}^μ includes the three Cartesian components (X^μ, Y^μ, Z^μ) for the translations of the center of mass (CM) and the three Cartesian components ($\theta_x^\mu, \theta_y^\mu, \theta_z^\mu$) for the rotations about the CM.⁶ The total number of degrees of freedom for the system is three times the number of atoms in the AR plus six (for three translations +three rotations) times the number of EFP's in the SR. The gradient of the translational (\mathbf{g}_T) and rotational (\mathbf{g}_R) components of the m th EFP is then

$$\mathbf{g}_T^\mu = \sum_{i \in \mu} \nabla_i E, \quad (9a)$$

$$\mathbf{g}_R^\mu = \sum_{i \in \mu} [(\mathbf{r}_i - \mathbf{r}_\mu) \times \nabla_i E + \boldsymbol{\tau}_i], \quad (9b)$$

where the index i runs over all expansion centers in the EFP, \mathbf{r}_i and \mathbf{r}_μ are the Cartesian coordinates of center i and the CM, respectively. $\boldsymbol{\tau}_i$ is the torque acting on the fragment multipole or induced dipole expansion centers due to the electric fields of the AR and the other EFP's.

Two quasi-Newton-based algorithms, the *quadratic approximation* (QA)¹¹ and the *eigenvector following* (EF)¹² methods, were used for the geometry searches, with the necessary modifications for including the EFP's (rigid rotors) in the optimizations, as discussed above. The two methods provide similar efficiency in determining stationary points on the formamide-water potential energy surfaces. Rather than storing the six Cartesian components of each fragment, the Cartesian coordinates of the expansion centers are updated using the following transformations, which are invariant with respect to the lab frame:¹³

$$\Delta \mathbf{r}_i(\Delta \mathbf{X}^\mu, \Delta \boldsymbol{\theta}^\mu) = \Delta \mathbf{X}^\mu + \mathbf{R}(\Delta \boldsymbol{\theta}^\mu)(\mathbf{r}_i - \mathbf{r}_\mu), \quad (10)$$

where \mathbf{X} and $\boldsymbol{\theta}$ are the translational and rotational coordinates, respectively. The rotational matrix \mathbf{R} is

$$\mathbf{R}(\Delta \boldsymbol{\theta}^\mu) \equiv e^{\mathbf{J}(\Delta \boldsymbol{\theta}^\mu)} = \mathbf{1} + \frac{\sin(|\Delta \boldsymbol{\theta}^\mu|)}{|\Delta \boldsymbol{\theta}^\mu|} \mathbf{J}(\Delta \boldsymbol{\theta}^\mu) + \frac{1 - \cos(|\Delta \boldsymbol{\theta}^\mu|)}{(\Delta \boldsymbol{\theta}^\mu)^2} [\mathbf{J}(\Delta \boldsymbol{\theta}^\mu)]^2, \quad (11)$$

$\mathbf{1}$ is a unit matrix and $\mathbf{J}(\Delta \boldsymbol{\theta}^\mu)$ is the first order rotation matrix,

$$\mathbf{J}(\Delta \boldsymbol{\theta}^\mu) = \begin{bmatrix} 0 & -\Delta \theta_z^\mu & \Delta \theta_y^\mu \\ \Delta \theta_z^\mu & 0 & -\Delta \theta_x^\mu \\ -\Delta \theta_y^\mu & \Delta \theta_x^\mu & 0 \end{bmatrix}. \quad (12)$$

The Hessian \mathbf{H} is updated by the BFGS¹⁴ and Powell¹⁵ schemes, in the search for minima and transition states, respectively. The six overall translations and rotations are projected out from the Hessian during the optimization, using the projection matrices given in the next subsection.

The choice of the initial Hessian \mathbf{H}_0 can be critical for effective geometry optimizations, especially for large species with small amplitude motions. For small or medium size systems a common practice is to compute the exact initial Hessian from *ab initio* electron structure theory. However, for large systems or in the absence of an analytic Hessian, computing the exact Hessian may not be practical. Several approaches have been proposed^{16,17} for formulating an initial approximate Hessian. For EFP calculations, the following choice for \mathbf{H}_0 is made:

$$\mathbf{H}_0 = \begin{bmatrix} \mathbf{H}^A(\mathbf{x}) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}^T(\mathbf{X}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{H}^R(\boldsymbol{\theta}) \end{bmatrix}. \quad (13)$$

\mathbf{H}^A is the analytically calculated Hessian for the AR without the EFP, while \mathbf{H}^T and \mathbf{H}^R are the Hessians for translational and rotational EFP coordinates, respectively, and are taken to be diagonal

$$\mathbf{H}_{ij}^T = 0.01 \times \delta_{ij}, \quad (14a)$$

$$\mathbf{H}_{ij}^R = 0.03 \times \delta_{ij}, \quad (14b)$$

where δ is the Kronecker delta. Equation (14) would also be useful for weakly interacting molecules with constrained internal geometries.

C. Intrinsic reaction coordinate

The *intrinsic reaction coordinate* (IRC), according to Fukui,¹⁸ is the steepest descent path from saddle point to minimum in mass-weighted (MW) Cartesian coordinates. The IRC, $\mathbf{q}(s)$, can be expressed in the differential form,

$$\frac{d\mathbf{q}}{ds} = - \frac{\mathbf{g}^M(s)}{|\mathbf{g}^M(s)|}, \quad (15)$$

where $\mathbf{g}^M(s)$ is the gradient at s in MW Cartesian coordinates. Usually, the MW coordinates are obtained by multiplying the mass matrix \mathbf{M} by the Cartesian coordinate vector, where \mathbf{M} is a diagonal matrix containing the square roots of the nuclear masses. For a system including the EFP's, the mass matrix is expressed as

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}^A & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{M}^R \end{bmatrix}, \quad (16)$$

where \mathbf{M}^A is the submatrix for atoms in the AR and \mathbf{M}^T is the submatrix for translational coordinates of the EFP's. Each block of \mathbf{M}^R , which corresponds to the rotational coordinates of the EFP's, is expressed as

$$\mathbf{M}_\mu^R = [\mathbf{I}_\mu]^{1/2}, \quad (17)$$

where \mathbf{I}_μ contains the moments of inertia of the EFP and is normally not diagonal in the lab frame. If the set of principle axes of each EFP is used as an instantaneous body-fixed frame, Eq. (16) will become diagonal. If the coordinates are transformed back to the lab frame, a result identical to Eq. (10) will be obtained. For a linear rotor some special care is

needed to obtain \mathbf{M}_μ^R . The MW gradient and Hessian can then be obtained using the inverse matrix of \mathbf{M} .

The Gonzalez–Schlegel second order algorithm¹⁹ was employed to integrate Eq. (15) to obtain the IRC for the internal rotation of formamide in the presence of 0–4 water molecules. We have extended the existing IRC code in GAMESS to include the EFP part of the system. Since the mass matrix for the EFP's (rigid rotors) varies during the course of integration, the moments of inertia (and thus the mass matrix) of the EFP's are evaluated at each structure for which the gradient is calculated. This guarantees that the MW gradient on the IRC is always tangent to the reaction path. The Cartesian coordinates of the expansion centers are updated via the back transformation of the MW displacements to Cartesian displacements and then following the transformation of Eq. (10).

The translational and rotational components of the Hessian are removed during the constrained optimization, via the projection method,²⁰

$$\mathbf{H}_p^M = [\mathbf{1} - \mathbf{P}] \mathbf{H}^M [\mathbf{1} - \mathbf{P}], \quad (18)$$

where the projection matrix \mathbf{P} is composed of the six unit vectors corresponding to the translations and infinitesimal rotations of the entire system.

$$\mathbf{P}_{i\gamma, i'\gamma'} = \sum_{\lambda=1}^3 \mathbf{L}_{i\gamma, \lambda}^{\text{Tr}} \mathbf{L}_{i'\gamma', \lambda}^{\text{Tr}} + \sum_{\lambda=1}^3 \mathbf{L}_{i\gamma, \lambda}^{\text{Rot}} \mathbf{L}_{i'\gamma', \lambda}^{\text{Rot}}. \quad (19)$$

Indexes i and i' run over all Natm atoms and 2 Nfrg EFP's, and γ and γ' run over Cartesian components (x, y, z). The translational and rotational vectors \mathbf{L} are

$$\mathbf{L}_{i\gamma, \lambda}^{\text{Tr}} = \begin{cases} \sqrt{m_i/M} \delta_{\gamma, \lambda} & (i \leq \text{Natm} + \text{Nfrg}), \\ 0 & (i > \text{Natm} + \text{Nfrg}), \end{cases} \quad (20a)$$

$$\mathbf{L}_{i\gamma, \lambda}^{\text{Rot}} = \begin{cases} \sum_{\alpha, \beta} (\mathbf{I}_0)_{\lambda, \alpha}^{-1/2} \sqrt{m_i} (\mathbf{r}_i - \mathbf{r}_0)_\beta \epsilon_{\alpha, \beta, \gamma} & (i \leq \text{Natm} + \text{Nfrg}) \\ \sum_{\alpha} (\mathbf{I}_0)_{\lambda, \alpha}^{-1/2} (\mathbf{I}_i)_{\alpha, \gamma}^{1/2} & (i > \text{Natm} + \text{Nfrg}) \end{cases} \quad (21b)$$

where M and \mathbf{I}_0 are the mass and moments of inertia of the entire system, m_i is the mass of the i th atom or the EFP, and \mathbf{I}_i are the moments of inertia of the EFP's. $\epsilon_{\alpha, \beta, \gamma}$ is the usual totally asymmetric Cartesian tensor.

D. *Ab initio* method and basis set

Internal rotation in formamide has been examined using both full *ab initio* and EFP calculations for the 0–2 water cases, while only EFP calculations were performed for the 3–5 water cases. For the purposes of this study, we chose the modest Dunning–Hay²¹ double-zeta plus polarization (DZP) basis set to carry out the *ab initio* part of the investigation. This choice is based on a balance between accuracy and computational costs. Jasien *et al.*²² and Boggs and Niu²³ have noted that inclusion of polarization functions reduces

TABLE II. *Ab initio* (HF/DZP) vs effective fragment timings (s).

System	Number of basis functions	Energy				Gradient			
		CPU		Wall clock		CPU		Wall clock	
		Ab	EFP	Ab	EFP	Ab	EFP	Ab	EFP
Formamide	60	35		38		88		93	
Formamide+water	85(60)	139	40	442	46	271	131	575	138
Formamide+two water	110(60)	331	43	1112	51	584	172	1366	181
Formamide+three water	135(60)	596	47	2034	56	982	214	2421	223
Formamide+four water	160(60)	964	51	3254	60	1494	254	3788	273

the formamide internal rotation barrier by a few kcal/mol, while Wang *et al.*²⁴ note that adding a second set of polarization functions or extending the double-zeta basis to the triple-zeta level has little effect on the predicted tautomerization barrier of formamide with one water molecule. As discussed below, the present results for internal rotation in formamide and the formamide–water complex are in good agreement with the experimental measurements and the previous computational results obtained with the Dunning correlation-consistent double-zeta plus polarization (cc-pVDZ) basis sets.²⁵

E. Computational savings

The goal of developing the effective fragment potential model is to carry out calculations that would not be feasible by fully *ab initio* methods. Table II lists the computational costs of a single energy evaluation and energy plus gradient for formamide–water complexes, using full *ab initio* calculations and the EFP method. The calculations were carried

out on an IBM RS6000/370 workstation, using the conventional SCF algorithm. Both CPU and wall clock requirements of the calculations are listed. Direct SCF calculations for formamide plus four waters were also performed. As expected, the CPU requirement for the direct calculation is larger than the wall clock time for the conventional method (3950 vs 3790 s). The number of basis functions for formamide is 60. Each water contributes 25 basis functions to the full *ab initio* calculations. As seen from the table, the time required for an EFP energy calculation behaves linearly with respect to the number of waters, and requires about four more seconds (both CPU and wall clock) for each additional water. For the energy plus gradient calculations, the CPU requirement of an additional water also increases linearly. The time savings for formamide plus up to four waters are factors of 4.1, 7.5, 10.9, and 13.9, respectively. Since SCF calculations scale as the fourth power of the number of basis functions, the time saving should increase dramatically when more EFP's are included in the calculations.

TABLE III. Geometry parameters^a of formamide equilibrium structure and internal rotation transition state.

	Min					TS		
	SCF ^b	MP2 ^c	Expt. ^d			SCF ^b	MP2 ^c	
C–N	1.353	1.354	1.371	1.376	1.352	1.428	1.428	1.447
C–O	1.196	1.190	1.217	1.193	1.219	1.187	1.181	1.210
C–H	1.092	1.101	1.115	1.102	1.098	1.089	1.098	1.111
N–H ₁	0.992	0.997	1.012	1.002	1.002	1.004	1.010	1.027
N–H ₂	0.995	1.000	1.015	1.014	1.002	1.004	1.010	1.027
N–C–O	125.0	125.0	125.1	123.8	124.7	125.0	125.2	125.8
N–C–H	112.8	112.5	111.6	113.2	112.8	113.6	113.4	112.4
C–N–H ₁	121.2	120.2	118.8	120.0	118.5	108.6	107.5	105.4
C–N–H ₂	119.2	118.0	116.5	117.1	118.5	108.6	107.5	105.4
H–C–N–O	180.0					180.0		
H ₁ –N–C–O	–179.8					57.5		
H ₂ –N–C–O	0.4					57.5		

^aBond lengths in Å and angles in degree.

^bThe first column is SCF/DZP and the second column is SCF/cc-DZP from Ref. 24.

^cMP2/cc-DZP from Ref. 24.

^dThe first column from Ref. 26 and the second column Ref. 27.

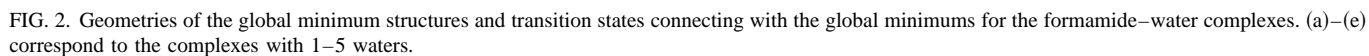


TABLE IV. Internal rotation energetics of formamide with water (kcal/mol).

Water	DZP						CCDZP ^a			
	<i>E</i>				<i>H</i> ₀ ^b		<i>E</i>			<i>H</i> ₀
	SCF	MP2	MP4	EFP	SCF	EFP	SCF	MP2	CISD	SCF
0	15.6	16.4	15.5		15.3		14.9	15.5	14.4	14.5
1	18.1(18.1) ^c	19.7	18.8	17.9	17.1	16.7	17.5	19.0	18.0	16.5
2	19.3(19.2)	21.3	20.4	19.1	17.9	17.5				
3	19.5(19.4)	21.5	20.6	19.3	18.1	17.8				
4	20.5	23.0		20.3	19.0	18.5				
5				20.2		18.4				

^aFrom Ref. 24.^bThe zero-point energies are not scaled.^cData in parentheses are obtained with constrained internal water geometry at the structure of isolated water molecule.

III. RESULTS AND DISCUSSION

A. Molecular structures

The geometries for the formamide equilibrium structure and internal rotation transition state are well known. The MIN structure has C_1 symmetry, while the TS has C_s symmetry. The SCF/DZP geometry parameters are compared with the previous calculations²⁴ and experimental measurements^{26,27} in Table III. The differences between SCF/DZP and MP2²⁸ cc-pVDZ or experimental results are very small, on the order of 0.01 Å for bond lengths and 1° for bond angles. Adding waters to formamide to form formamide–water complexes has only a very small effect on the internal geometries of both Min and TS, as has been seen in many hydrogen-bonding complexes.²⁹ The essential features for the corresponding global minimum structures (GMIN) and transition states connecting with GMIN for the complexes containing 1–5 waters are shown in Fig. 2. Complete geometries in Cartesian coordinates for all complexes with either full *ab initio* or EFP waters are available as supplementary material. For the *ab initio* structures (with

1–3 waters), geometries are given for both full optimizations and for optimizations in which the internal water geometry has been fixed at that of the isolated water molecule, since EFP waters have fixed internal geometries. Freezing the internal water geometries seems to have only a small effect. In general, the EFP geometries are in reasonable agreement with the *ab initio* ones, with bond lengths and bond angles within 0.05–0.08 Å and 5°–10°, respectively. The agreement seems to improve as the number of water molecules in the system increases.

As seen in Fig. 2, the GMIN for the complexes containing up to 3 waters have cyclic structures, with the two ends of the formamide molecule (NH at one end and C=O at the other end) connected by a linear sequence of n waters. For the formamide: (H₂O) _{n} ($n=4, 5$) complexes, the minimum energy complexes (GMIN) are somewhat different [Figs. 2(d)–2(e)]. The water molecules are still on one side of the formamide molecule, but these waters themselves form a cycle, such that only two of the waters are hydrogen bonded to formamide. The formamide complex structures with four

TABLE V. Comparison of vibrational frequencies.

System	Imaginary frequency ^c	ZPE ^{a,b}		rms ($\nu_{ab} - \nu_{EFP}$) ^c	
		Min	TS	Min	TS
Formamide	503i	0.0	0.0		
Formamide+water					
All <i>ab initio</i>	492i	2.6	1.8		
EFP water	496i	2.9	1.9	23	17
Formamide+two water					
All <i>ab initio</i>	482i	5.4	4.3		
EFP water	484i	6.0	4.8	19	16
Formamide+three water					
All <i>ab initio</i>	488i	7.6	6.5		
EFP water	486i	8.6	7.4	18	13
Formamide+four water					
All <i>ab initio</i>	480i	10.6	9.4		
EFP water	481i	12.0	10.6	17	15

^aZero-point energy difference between formamide–water complex and isolated formamide and water.^bIn kcal/mol.^cIn cm^{−1}.

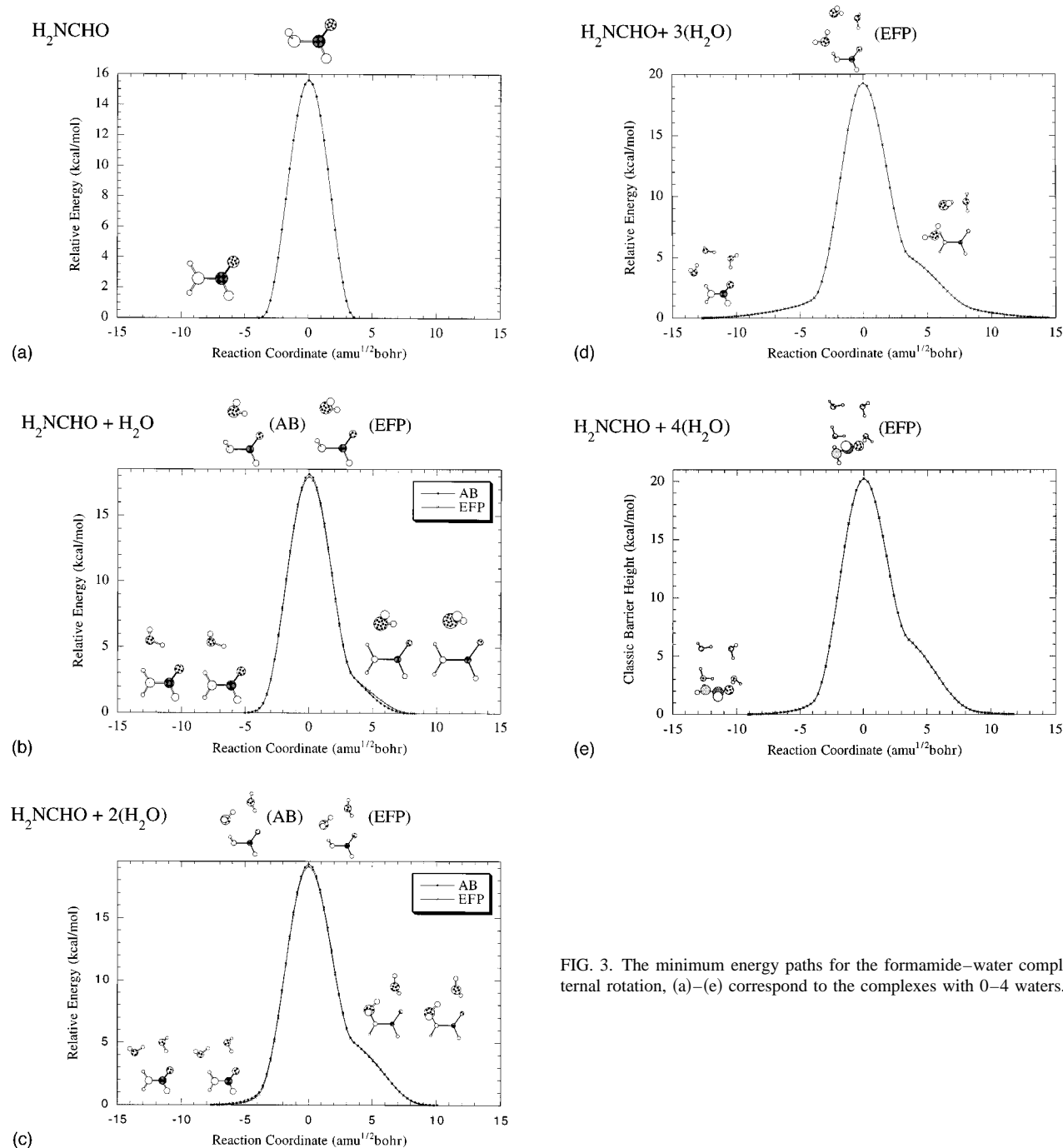


FIG. 3. The minimum energy paths for the formamide–water complex internal rotation, (a)–(e) correspond to the complexes with 0–4 waters.

or five waters that are analogous to those shown in Figs. 2(a)–2(c) for the complexes with 1–3 waters also exist on the potential energy surface, but are 1 kcal/mol higher (for both EFP and *ab initio* waters) than the GMIN shown in Figs. 2(d) and 2(e). The same is true for the transition state structures shown below. Since our main intent here is to test the reliability of the effective fragment method, we focus on lowest energy structures.

B. Internal rotation barriers and vibrational frequencies

The essential features of the internal rotation energetics and vibrational frequencies are summarized in Tables IV and

V, respectively. The EFP barrier heights, both in absolute and relative terms, are in excellent agreement with the full *ab initio* results. Both levels of theory (full *ab initio* and EFP water) predict a significant (>2 kcal/mol) increase in the barrier height upon the introduction of the first water. This is also in agreement with the previously reported *ab initio* calculations.²⁴ Addition of a second water causes a smaller (≈ 1 kcal/mol) increase in the barrier height, and addition of the third water results in a very small (≈ 0.2 kcal/mol) increase in the barrier. However, addition of the fourth water causes approximately a 1 kcal/mol increase in both the EFP and *ab initio* barrier height. This relatively large change may be related to the change in geometry noted above, upon go-

ing from three to four waters. [As noted earlier for four waters there is a geometry that looks like those found for 1–3 waters [Figs. 2(a)–2(c)], but the structure is predicted to be higher in energy by both the *ab initio* and EFP calculations.] For five waters, the EFP method predicts a geometry that is very similar to that for the four water case, and the five water barrier is virtually the same as that for four waters. So, it appears that the effect of solvation on the predicted barrier height for formamide is saturated by 4–5 waters.

It is of interest to explore the effect of correlation on the energy results discussed above. This was explored with the DZP basis set by performing both MP2²⁸ and MP4²⁹ calculations at the SCF geometries. For formamide itself, the MP2 and MP4 barriers are 16.4 and 15.5 kcal/mol, compared with the SCF value of 15.6 kcal/mol. For the formamide–water complex, the MP2 and MP4 barriers are 19.7 and 18.8 kcal/mol, compared with the 18.1 kcal/mol SCF barrier height. Similar results are found as additional waters are added: The more reliable MP4 barrier is approximately 1 kcal/mol higher than the SCF and fragment barriers. So, the addition of correlation corrections has only a small effect, and the discussion of the SCF results appears to be quite reasonable. Note that the experimental value for the formamide internal rotation in aqueous solution is 21.3 ± 1.3 kcal/mol.³¹ Based on the information about the effect of electron correlation, it is likely that the SCF and EFP limit will be about 2 kcal/mol too low. Adding the expected MP4 correction of ~ 1 kcal/mol will bring the calculations into very good agreement with experiment.

Table IV also lists the previous results using the cc-pVDZ basis set.²⁴ The effect of correlation corrections for the barrier heights in formamide and formamide–water complexes is similar for this basis set, although the absolute values are slightly lower.

The label ΔZPE in Table V summarizes the difference in zero-point vibrational energies (ZPE) for the complex, relative to the separated molecules. This provides one measure of the quality of the predicted vibrational frequencies. A second such measure is the rms deviation of the vibrational frequencies, relative to the *ab initio* frequencies. These rms deviations are also summarized in Table V. Both of these comparisons suggest that the EFP and *ab initio* results are in excellent agreement. Since average and rms deviations can be misleading, it is worth noting that for the formamide: (H₂O) complex, nearly all EFP frequencies are within 30–40 cm^{−1} of the corresponding *ab initio* values, the largest frequency error is 80 cm^{−1}, and the imaginary frequency that corresponds to the internal rotation is 492*i* cm^{−1} (*ab initio*) and 496*i* cm^{−1} (EFP). Similar comments apply to the formamide: (H₂O)₂, formamide: (H₂O)₃ and formamide: (H₂O)₄ complexes. In particular, the imaginary frequencies are consistently within 4 cm^{−1} of their *ab initio* counterparts. The complete sets of vibrational frequencies are available as supplementary material.

C. Minimum energy paths

The minimum energy paths (MEP) for the formamide internal rotation are shown in Fig. 3 for formamide: (H₂O)_{*n*},

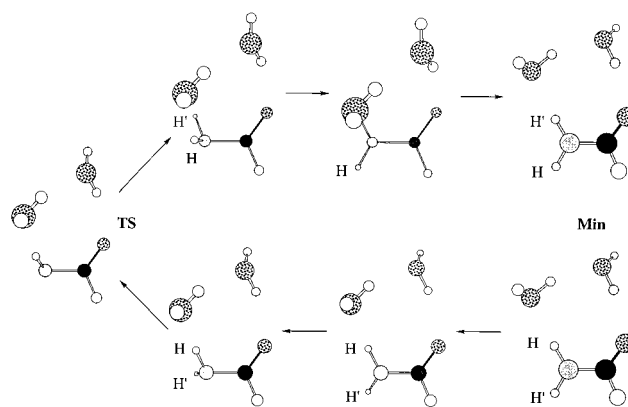


FIG. 4. Rotational isomerization of formamide–water complex.

n = 0–4. First, note that there is excellent agreement between the EFP and *ab initio* reaction paths. This is very encouraging, since it suggests that predictions using the EFP method that are related to dynamics are likely to be reliable. An increasing distortion appears (in both the EFP and *ab initio* paths) on the product side of the MEP, as the number of water molecules is increased. This may be understood by examining the structures along the MEP for the two water case, shown in Fig. 4. In the reaction channel, as the formamide molecule begins to rotate, the essential hydrogen bonding interaction is maintained. However, as the transition state is passed and the system enters the product channel, this hydrogen bonding interaction is temporarily lost. This results in a rather flat portion of the path. One expects that this behavior, especially the apparent asymmetry of the MEP, will disappear as the number of water molecules is increased.

IV. CONCLUSIONS

The internal rotation of formamide with 0–5 water molecules oriented along the N–C bond has been studied by the full *ab initio* self-consistent field theory and using the effective fragment method. For each case, the EFP geometries, harmonic vibrational frequencies, rotational barriers, and intrinsic reaction coordinates for the internal rotation are found to be in excellent agreement with their *ab initio* counterparts. The global energy minimum structures for four and five water complexes are predicted to be formamide bonded to two adjacent waters, with all water molecules in a ring. Probably due to the structural constraints, the complexes containing less than four waters have cyclic structures with the two ends of formamide connected by a sequence of water molecules.

The internal rotation barrier of formamide–water complexes increases from 15.3 kcal/mol with no water to 19.0 kcal/mol with four waters and seems to saturate at four to five waters. When electron correlation corrections are added, the estimated internal rotation barrier is ~ 20 kcal/mol, in very good agreement with experimental measurements.

The effective fragment model appears to be a powerful tool for the study of chemical reactions in solvation, requir-

ing much less computer resources than fully *ab initio* calculations.³²

ACKNOWLEDGMENTS

This work was supported by a grant from the National Institute for Standards and Technology. The authors thank Dr. Jan Jensen for helpful suggestions regarding the manuscript.

- ¹See for example, (a) M. Szafran, M. M. Karelson, A. R. Katritzky, J. Koput, and M. C. Zerner, *J. Comp. Chem.* **14**, 371 (1993); (b) D. J. Giesen, C. J. Cramer, and D. G. Truhlar, *J. Phys. Chem.* **99**, 7137 (1995); (c) R. Bianco and J. T. Hynes, *J. Chem. Phys.* **102**, 7864 (1995); (d) F. R. Tortonda, J.-L. Pascual-Ahuir, E. Silla, and I. Tunon, *J. Phys. Chem.* **99**, 12525 (1995); (e) T. N. Truong and E. V. Stepanovich, *J. Chem. Phys.* **103**, 3709 (1995); (f) D. J. Giesen, J. W. Storer, C. J. Cramer, and D. G. Truhlar, *J. Am. Chem. Soc.* **117**, 1057 (1995).
- ²See for example, (a) A. Warshel, *J. Phys. Chem.* **83**, 1640 (1979); (b) B. T. Thole and P. T. Van Duijnen, *Theor. Chim. Acta* **55**, 307 (1980); (c) B. T. Thole and P. T. Van Duijnen, *Chem. Phys.* **71**, 211 (1982); (d) U. C. Singh and P. A. Kollman, *J. Comp. Chem.* **5**, 129 (1984); (e) A. Warshel and G. King, *Chem. Phys. Lett.* **121**, 124 (1985); (f) G. King and A. Warshel, *J. Chem. Phys.* **91**, 3647 (1989); (g) M. J. Field, P. A. Bash, and M. J. Karplus, *J. Comp. Chem.* **11**, 700 (1990); (h) V. Luzhkov and A. Warshel, *J. Am. Chem. Soc.* **113**, 4491 (1991); (i) J. Gao, *ibid.* **116**, 1563 (1994); (j) M. A. Thompson, E. D. Glendening, and D. F. Feller, *J. Phys. Chem.* **98**, 10465 (1994); (k) F. Maseras and K. Morokuma, *J. Comp. Chem.* **16**, 1170 (1995); (l) J. Gao, *J. Phys. Chem.* **96**, 537 (1992); (m) H. Liu, F. Müller-Plathe, and W. F. van Gunsteren, *J. Chem. Phys.* **101**, 1722 (1994); (n) A. H. de Vries, P. Th. van Duijnen, A. H. Juffer, J. A. C. Rullmann, J. P. Dijkman, H. Merenga, and B. T. Thole, *J. Comput. Chem.* **16**, 37 (1995).
- ³(a) T. N. Truong and E. V. Stefanovich, *Chem. Phys. Lett.* (in press); (b) M. A. Thompson, *J. Phys. Chem.* **100**, 14492 (1996).
- ⁴Z. Luzhkov and A. Warshel, *J. Am. Chem. Soc.* **113**, 4491 (1991).
- ⁵J. H. Jensen, P. N. Day, M. S. Gordon, H. Basch, D. Cohen, D. R. Garmer, M. Krauss, and W. J. Stevens, in *Modeling the Hydrogen Bond*, edited by Douglas A. Smith (ACS Symposium Series 569, 1994), p. 139.
- ⁶P. N. Day, J. H. Jensen, M. S. Gordon, Simon P. Webb, W. J. Stevens, M. Krauss, D. R. Garmer, H. Basch, and D. Cohen, *J. Chem. Phys.* **105**, 1968 (1996).
- ⁷(a) A. J. Stone, *Chem. Phys. Lett.* **83**, 233 (1981); (b) A. J. Stone and M. Alderton, *Mol. Phys.* **56**, 1047 (1985).
- ⁸D. R. Garmer and W. J. Stevens, *J. Phys. Chem.* **93**, 8263 (1989).
- ⁹Note that α_{xy}^I in Eq. (5b) is not symmetric in the general case of multiple fragments. However, it is found via numeric tests that symmetrization of the polarizability tensor by averaging α_{xy}^I and α_{yx}^I has little effect on either the energy or the gradient.
- ¹⁰M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, *J. Comp. Chem.* **14**, 1347 (1993).
- ¹¹P. Culot, G. Dive, V. H. Nguyen, and J. M. Ghuysen, *Theor. Chim. Acta* **82**, 189 (1992); T. Helgaker, *Chem. Phys. Lett.* **182**, 503 (1991).
- ¹²A. Banerjee, N. Adams, J. Simons, and R. Shepard, *J. Phys. Chem.* **89**, 52 (1985).
- ¹³H. F. King, *J. Phys. Chem.* **94**, 5617 (1990).
- ¹⁴(a) C. G. Broyden, *J. Inst. Math. Appl.* **6**, 76 (1970); (b) R. Fletcher, *Comput. J.* **13**, 317 (1970); (c) D. Goldfarb, *Math. Comput.* **24**, 23 (1970); (d) D. F. Shanno, *ibid.* **24**, 647 (1970).
- ¹⁵M. J. D. Powell, *Math. Prog.* **1**, 26 (1971).
- ¹⁶H. B. Schlegel, *Theor. Chim. Acta* **66**, 333 (1984).
- ¹⁷T. H. Fischer and J. Almlöf, *J. Phys. Chem.* **96**, 9768 (1992).
- ¹⁸K. Fukui, *Acc. Chem. Res.* **14**, 363 (1981).
- ¹⁹C. Gonzalez and H. B. Schlegel, *J. Phys. Chem.* **94**, 5523 (1990).
- ²⁰W. H. Miller, N. C. Handy, and J. E. Adams, *J. Chem. Phys.* **72**, 99 (1980).
- ²¹T. H. Dunning and P. J. Hay, in *Modern Theoretical Chemistry, Methods of Electronic Structure Theory*, edited by H. F. Schaefer III (Plenum, New York, 1977), Vol. 3.
- ²²P. G. Jasien, W. J. Stevens, and M. Krauss, *J. Mol. Struct.* **139**, 197 (1986).
- ²³J. E. Boggs and Z. Liu, *J. Comp. Chem.* **6**, 46 (1985).
- ²⁴W.-C. Wang, J. C. Facelli, and J. Simons, *Int. J. Quantum Chem.* **45**, 123 (1993).
- ²⁵T. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).
- ²⁶C. C. Costain and J. M. Dowling, *J. Chem. Phys.* **32**, 158 (1960).
- ²⁷E. Hirota, R. Sugisaki, C. J. Neilsen, and G. O. Sorensen, *J. Mol. Spectrosc.* **49**, 251 (1974).
- ²⁸C. Moller and M. S. Plesset, *Phys. Rev.* **46**, 618 (1934).
- ²⁹See for example, *Modeling the Hydrogen Bond*, edited by Douglas A. Smith (ACS Symposium Series 569, 1994).
- ³⁰R. Krishnan and J. A. Pople, *Int. J. Quantum Chem.* **14**, 91 (1978).
- ³¹H. Kamei, *Bull. Chem. Soc. Jpn.* **41**, 2269 (1968).
- ³²See AIP Document No. PAPS JCPSA-105-11081-10 for 10 pages of Cartesian coordinates and vibrational frequencies. Order by PAPS number and journal reference from American Institute of Physics, Physics Auxiliary Publication Service, Carolyn Gehlbach, 500 Sunnyside Boulevard, Woodbury, New York 11797-2999. Fax: 516-576-2223, e-mail: paps@aip.org. The price is \$1.50 for each microfiche (98 pages) or \$5.00 for photocopies of up to 30 pages, and \$0.15 for each additional page over 30 pages. Airmail additional. Make checks payable to the American Institute of Physics.