Theor Chim Acta (1988) 74: 1-10

Theoretica Chimica Acta

An *ab initio* pair potential for the interaction between a water molecule and a formate ion

Kersti Hermansson, George C. Lie, and Enrico Clementi

IBM Corporation, Data Systems Division, Dept. 48B/MS 428, Neighborhood Road,
Kingston, NY 12401, USA

(Received May 29, revised October 13/Accepted October 16, 1987)

The potential surface for the interaction between a rigid formate ion and a rigid water molecule has been investigated by *ab initio* methods. An analytical potential expression was derived to fit the 591 calculated SCF energies. The global minimum on the surface is −16.3 kcal/mol and corresponds to a bifurcated bonding situation.

Key words: Potential surface—Water-formation—SCF calculations

1. Introduction

A knowledge of intermolecular potential energy surfaces is of vital interest in chemistry and has many applications in statistical mechanics and physics. Potential energy functions for water interacting with different small inorganic ions have been introduced over the past few years: alkali ions, [1-9], halide ions, [2-7, 9], Mg [2, 10], Ca [2, 11], NH_4^+[12-13], NO_3^-[14], NO_2^-[14]. Self-consistent field (SCF) calculations on parts of the $HCOO^-$-H_2O surface have been performed by Alagona et al. [15] who studied the variation of the interaction energy along three different directions of approach. The present work deals with the full potential energy surface for a water molecule in the field of a formate ion. This potential has subsequently been used in an MD simulation of an aqueous solution of lithium formation, the results of which are reported elsewhere [16].

Some of the ion-water potential energy functions mentioned above are empirical in nature [5-9], while others are partially-or fully-based on quantum mechanical calculations. Among the quantum-mechanical studies only few have attempted to cover the full six-dimensional (in the general case) intermolecular potential surface of the two rigid molecules. This is no doubt due in part to the complexity of the computations; the derivation of a quantum-mechanical intermolecular energy hypersurface is an arduous task even for the smallest molecular systems: intermolecular energies have to be calculated with 'sufficient' accuracy for a large number of geometrical configurations and, in order for the results to be of use in statistical mechanical and other applications, the potential surface has to be represented by some simple analytical function. The resulting analytical expression which models the real chemical system is thus by necessity impaired by approximations made at various stages in the derivation: (i) choice of *quantum-mechanical model*, bais sets, etc.; (ii) selection of *geometrical configurations*; and (iii) the *fitting* procedure. In the present quantum-mechanical calculations on the energy surface for H_2O-$HCOO^-$ the basis sets used were of double-zeta plus polarization (DZP) type. The energies for a total of 591 configurations, carefully selected to provide a representative sample, were fitted to simple analytical potential expressions. Details of the calculations are given in Sect. 2. The results are discussed in Sect. 3 and a summary is given in Sect. 4.

2. Method

2.1. SCF calculations

Ab initio MO-LCAO-SCF calculations were carried out on the $HCOO^-$-H_2O system for 591 different geometrical configurations using the IBMOL program [17]. The internal geometries of the two molecules were kept fixed at the following values. The formate ion was taken to be planar with the C–O distances equal to 1.250 Å, the C–H distance 1.085 Å, and the O–C–O angle 125.0°. For the water molecule the O–H distances were taken to be 0.970 Å and the H–O–H angle 107.8°. These geometries were selected as typical values for bonded water and formate molecules based on available diffraction studies of crystalline formates [18] and hydrates [19]. If instead, the gas-phase experimental water geometry (O–H distance 0.957 Å, H–O–H angle: 104.5°) is used in the quantum-mechanical calculations the resulting interaction energy at the global minimum would only be changed by about 1 kcal/mol (out of −16 kcal/mol). None of the results discussed and conclusions drawn in this paper would be affected by such a small change.

The basis sets consisted of constracted Gaussian type functions of double-zeta-plus-polarization (DZP) quality. Dunning's [20] ⟨4s2p/2s⟩ contractions of Huzinaga's [21] (9s5p/4s) basis sets were used for carbon, oxygen, and hydrogen, augmented with polarization functions according to Roos and Siegbahn [22], i.e., a set of 3d functions with exponent 0.63 on carbon, 1.33 on oxygen and a 2p function with exponent 0.8 on hydrogen. The energy of a free water molecule (at the geometry used in this paper) is −76.045102 a.u.. The energy of a free water molecule at the experimental geometry is −76.046278 a.u. using the same basis set. Optimization of the free water geometry gives an O–H distance of 0.943 Å and an H–O–H angle of 106.6° with a total energy of −76.046784 hartree (h).

The energy of a free formate ion at the geometry used in this work is −188.232067 h. The optimized free formate geometry using this same basis gives a C–O distance of 1.236 Å, a C–H distance of 1.124 Å and an O–C–O angle of 130.6° with a total energy of −188.235404 h.

The *interaction energies*, defined as E(complex)-E(H_2O)-E($HCOO^-$), were in all cases corrected for possible basis set superposition errors (BSSE's) using the counterpoise method [23]. There are controversies in the literature regarding the appropriateness of using the counterpoise method to correct for the basis set superposition errors [24, 25]. In a separate study conducted by Sordo et al. in our laboratory [26], it has been recommended that the counterpoise method be used for *large* systems containing, say, more than 20 electrons, and involved in strong hydrogen bonds. It is also found that with BSSE corrections, basis set of double-zeta quality gives quite good interaction energies (within ~10% compared with that using triple-zeta plus polarization). Thus it is believed that the errors in our basis set (DZ *plus* polarization) with BSSE correction should be around or less than 10%.

Figure 1 shows the variation of the interaction energy with C–OW (OW denoting the oxygen atom in the water molecule) distance for three different directions of approach, which are chosen to represent three rather extreme cases of interaction behavior for the H_2O-$HCOO^-$ system. The corresponding geometrical configurations are displayed in Fig. 2. The curves in Fig. 1 will be discussed in more

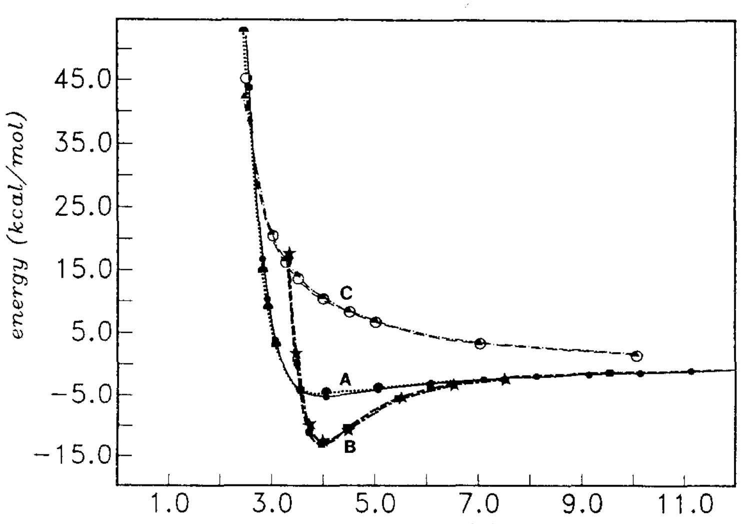

Fig. 1. Potential energy curves from the SCF calculations and from the fitted analytical function [expressions (1) and (2)] for three different geometries of approach. The quantity on the horizontal axis is the C–OW distance. The curves are second-order spline functions through the SCF points and the analytical points, respectively. The conformations *A*, *B*, and *C* are displayed in Fig. 2. —●— SCF for *A* conformation: - -■- - SCF for *B* conformation; —-▲—- SCF for *C* conformation; ·····▲····· fitted for *A* conformation; ———★——— fitted for *B* conformation; - -⊖- - fitted for *C* conformation

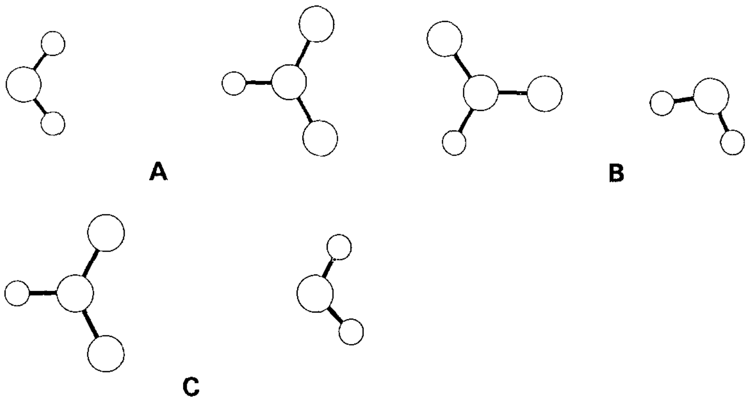

Fig. 2A–C. The geometries corresponding to the curves in Fig. 1, In **A** the C–H vector is co-linear with the HW–OW–HW bisector. In **B** the C–OW vector forms an angle of 8.60° with the "closest" HW–OW vector. In **C** the C–OW vector forms an angle of 8.60° with the HW–OW–HW bisector

detail in the next section; here we just wish to comment on the effect of the BSSE correction on the interaction energy. For the shortest distance of approach (i.e., the most repulsive point) displayed in Fig. 1 the BSSE is 2.6 kcal/mol out of a total interaction energy of +43.6 for curve *A*, 1.5 kcal/mol (out of +17.7) for curve *B* and 1.6 kcal/mol (out of +42.0) for curve *C*. At the minimum in curve *A* the BSSE is 0.7 kcal/mol and 1.6 kcal/mol in curve *B*. Although quite small on a relative scale, the BSSE's are not negligible, and are, in fact, for "medium-range" C–OW distances larger than the errors caused by the fitting procedure, i.e., larger than the difference between the fitted energy and the BSSE-corrected SCF energy.

Finally we would like to make a few comments about the correlation energy in this section. It is well known that calculating the correlation energy for a system like ours is not trivial even with today's supercomputer. Since the system studied here is an ion-dipole system, SCF calculations should be able to pick up the main parts of the interactions. With rather large interaction energy, it is thus believed that the correlation correction will not be large relative to the fitting errors and the binding energies. An indication of that is shown in Sect. 2 where we see that a r^{-4} term representing charge-induced-dipole interaction is required in fitting the calculated energies. As another indication, we note that SCF calculations for water-water interaction (which involve mainly dipole-dipole interaction) were found to account for about 80% of the binding energy [27].

2.2. Geometrical configurations

A proper selection of the configurations is necessary to ensure that the derived potential represents any general configuration of the two molecules. Both the translational and rotational degrees of freedom in the energy hyperspace have to be covered as completely as possible. To that end, the water molecule was translated along "rays" originating at the formate C atom and distributed quite evenly in space, within and out of the formate plane. At each position the water molecule was allowed to rotate in a stepwise fashion. Fortunately, both the formate and water molecules exhibit C_{2v} symmetry, which greatly reduces the space to be covered. If, at the fitting stage, it was found that a certain region of space was under-represented, leading to a bad fit in that region, more SCF points were calculated. As mentioned above, altogether 591 different geometries of the water-formate complex were selected.

Regardless of the care taken in this procedure it is always true that parts of the configuration space remain uncovered. One way to verify that all regions are reasonably represented by the fitted analytical expression is to do "spotchecks" to compare the SCF and the fitted interaction energies for "new" points not included in the fitting procedure. Another way is to study " iso-energy" plots (discussed in the Sect. 2.4) to ensure that no spurious features occur *between* the points included in the fitting. Both these methods were tested with the present potential.

2.3. Fitting procedure

The intermolecular potential function was expressed as a sum of site-site pair potentials representing both the Coulombic and the non-Coulombic interactions. If U is the total potential energy of the system it is thus written as

$$U = \sum_{1 \leq i < j \leq N} U_{ij} \tag{1}$$

where the sum includes all pairs of sites i on the formate ion and sites j on the water molecule. The sites chosen were the atomic sites in both molecules plus an additional site on the water molecule, positioned at 0.2677 Å from the oxygen site towards the hydrogen atoms along the H–O–H bisector. This latter site is only used in the representation of the Coulombic interaction and was chosen to conform with the MCY water-water potential [28], to facilitate the simultaneous use of the present potential and the MCY potential in computer simulation work. For the non-Coulombic interactions only the atomic sites were employed.

Many different forms for the potential were tested. The final expression used is the following

$$U_{ij} = \frac{q_i q_j}{r_{ij}} + A_{ij}\, e^{-B_{ij} r_{ij}} - \frac{C_{ij}}{r_{ij}^4} + \frac{D_{ij}}{r_{ij}^{10}} \tag{2}$$

where r_{ij} is the site-site separation, q_i and q_j are the site charges and the coefficients A_{ij}, B_{ij}, C_{ij} and D_{ij} are the fitting parameters. There is really no deep theoretical reason for choosing that particular form. We had tried to use r^6 instead of r^4 since it is certainly nicer to have something clinging off faster for the sake of cutoff in the liquid simulations to be carried out later, but it gave a considerable worse fit. This seems to conform to the fact that Hartree–Fock calculations are able to pick up charge-induced-dipole interactions which go as r^{-4} (note that since C_{ij}'s are different, there will be a term proportional to R^{-4} left if we do expand r_{ij} in terms of center to center distance R). It was also found that inclusion of the r^{10} term gives a slight better fitting.

Different values for the formate charges were tested, based on the resulting dipole moments and the Mulliken charges for many different formate-water orientations. The final charges and the optimized coefficients are listed in Table 1. We note that the total charge on the formate ion was not allowed to vary as the potential will later be used in molecular dynamics simulations of liquids and solids where charge neutrality of the system is to be maintained.

The medium and long range intermolecular interactions are to a large extent governed by the charge distribution in the formate ion and the water molecule.

Table 1. Parameters for the analytical pair potential of water-formate. See text for the actual expression. Both oxygen atoms in the formate ion and both hydrogen atoms in the water molecule have the same charges and coefficients. The additional "P" site is defined in the text. Distances are given in Å and energies in kcal/mol. Note that the q values in the table are given in a.u. and thus need to be converted to give the energy in kcal/mol

C-OW	A	82 587.335839
	B	3.7811612
	C	784.208846
	D	−12 891.03115596
O-OW	A	388 358.774757
	B	4.62072393
	C	−284.98943821
	D	−19 274.649306
H-OW	A	4 479.90326164
	B	3.33771265
	C	−47.778605
	D	−194.00145699
C-HW	A	789.30057818
	B	50.55008396
	C	−364.9235748
	D	−1 317.19470684
O-HW	A	5 291.139234
	B	3.40207486
	C	163.7323479
	D	31.05383248
H-HW	A	839.16429966
	B	3.44072033
	C	27.50807387
	D	2.35921021
C	q	−0.150
O	q	−0.500
H	q	+0.150
OW	q	0.
HW	q	+0.7150
P	q	−1.4350

The potential expression (2) has to account (as well as possible) not only for the permanent mono- and multi-polar charge interactions but also for induced polarization effects in an effective manner as explained above. However, it should be cautioned in attributing too much physical significance to the effective charges. The non-Coulombic contributions to the interaction energy (i.e., the second, third and fourth terms in (2)) for the three configurations in Figs. 1 and 2 and for a C-OW distance of 3.5 Å are +2.5 kcal/mol for *A*, +27.5 kcal/mol for *B* and −0.1 kcal/mol for *C*. The non-Coulombic contribution to curve *B* at the minimum (C-OW equal to 4.0 Å) is 1.7 kcal/mol.

The standard deviation of the fitting, defined as $[\sum (E_{\mathrm{SCF}} - E_{\mathrm{fit}})^2/N]^{1/2}$, is equal to 1.91 kcal/mol for the 579 points with interaction energies less than +100 kcal/mol, and equal to 0.80 kcal/mol for the 499 points with interaction energies less than +30 kcal/mol. The quality of the fitting is visualized in Figs. 1 and 3, which show that the fit is very satisfactory for all different geometries included in the fitting. The largest deviations between SCF energies and fitted energies occur for the repulsive conformations: they get worse the more repulsive the points are. However, it should be noted that these highly repulsive regions are pretty unimportant in the practical application of the potentials in a simulation, since they are hardly sampled. There are no particular geometries which are especially badly represented: the difference between SCF and fitted energies are all within about ±2 kcal/mol for all the attractive points, and within ±3 or 4 kcal/mol for interaction energies between 0.0 and +20.0 kcal/mol. As a word of caution it should be pointed out, however, that even though the fit is excellent it cannot be excluded that for some other arbitrary configuration the potential expression (2) may deviate more from the SCF energy than the standard deviation would imply.

The fittings were done with a program adapted from the LSQ8A program written by Probst and Clementi [29].

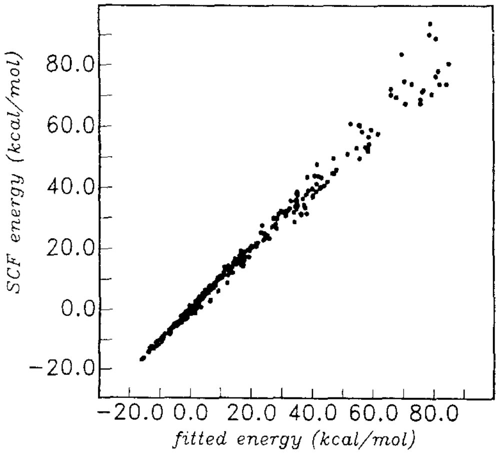

Fig. 3. Correlation between interaction energies from the SCF calculations and those derived from expressions (1) and (2) and the parameters in Table 1

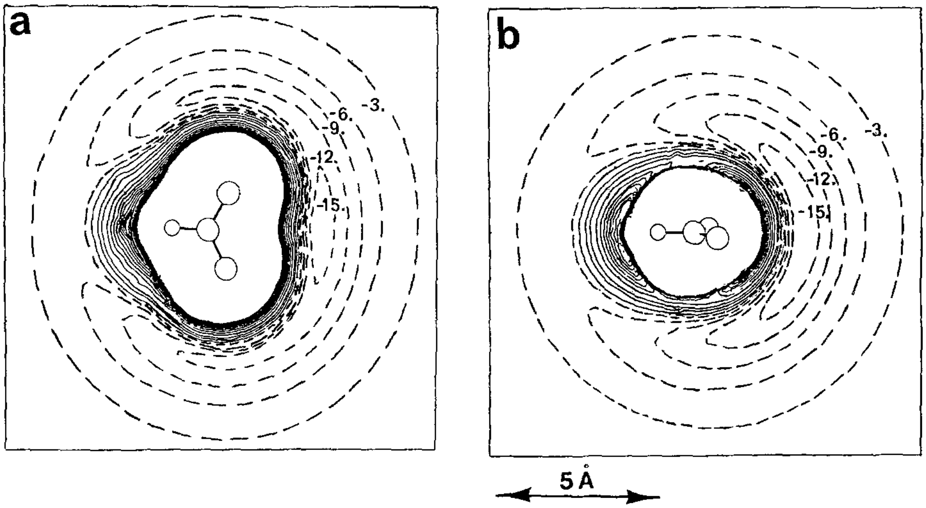

Fig. 4. Iso-energy plots (see text) (**a**) in the formate plane; (**b**) perpendicular to the formate plane, through the C-H bond

2.4. *Iso-energy contour maps*

Figure 4ab shows the iso-energy maps calculated from the potential function (2). The purpose of these maps is to provide a more comprehensive picture of the potential surface than do the one-dimensional plots in Fig. 1. In addition, they serve as a check of the fitted analytical expression in the sense discussed in Sect. 2.2.

The maps were constructed in the following way. The formate ion was positioned at the center of the map and the center of the water molecule was translated to each point in a user-defined grid in the plane studied. At each grid point the water molecule was allowed to rotate in such a way that the interaction energy was minimized. The contours maps were constructed from these final grid values. The program used was a modified version of EMIMAP written by Probst and Clementi [29].

3. Results and discussion

Figures 1 and 4 illustrate the final potential energy surface. The interaction is highly anisotropic. The global minimum is −16.3 kcal/mol and occurs at a C-OW distance of about 3.25 Å (HW···O equal to 1.95 Å), where the water molecule is hydrogen-bonded in a "bifurcated" conformation such that each of its hydrogen atoms binds to an oxygen atom on the formate ion in a symmetrical fashion (i.e., like conformation *C* in Fig. 2 except the water is rotated 180° so that the hydrogen atoms point *towards* instead of away from the ion). When the water molecule binds to the hydrogen side of the formate ion, on the other hand, the minimum energy conformation occurs at −4.7 kcal/mol (curve *A* in Fig. 1) and corresponds to the water molecule lying in the plane of the formate ion with its two hydrogen atoms pointing symmetrically towards the ion.

Other conformations which are very stable (−12 to −15 kcal/mol), although not quite as stable as the bifurcated geometry, occur when the water molecule forms a hydrogen bond to only one of the two formate oxygen atoms (see *B* in Fig. 2). The minimum here occurs at a C–OW distance of approximately 4.0 Å (HW···O approximately equal to 2.0 Å) and the water molecule then prefers to point one OW–HW bond towards the closest formate oxygen atom, while the other OW–HW bond lies in the formate plane pointing as close as it can towards the other formate oxygen atom (i.e. "oposite" its direction in *B* in Fig. 2). As can be seen from Fig. 4, however, the minimum here is very flat and the energy penalty for moving the water molecule out of the formate plane or going from a bifurcated (−16.3 kcal/mol) to a "normal" hydrogen bond is small. The results of Alagona et al. [15] show the same relative difference between the bifurcated and the single-bond systems, although their absolute interaction energies, using a 6-31G** basis set are approximately 4.5 kcal/mol more negative than the present results. The bifurcated configuration was also found to be the most stable geometry for the nitrate-water and nitrite-water complexes studied by Howell et al [14].

As for the variation of the interaction energy with respect to the *orientation of the water molecule* around the formate ion, the following observations can be made. When the water molecule is positioned on the *hydrogen side* of the formate ion the energy is at a minimum when both OW–HW bonds are pointing towards the ion. For the configurations of type *A* in Fig. 2 the interaction is about −5 kcal/mol at a C–OW distance of 4.0 Å (see Fig. 1). If the water molecule here is rotated so that the hydrogen atoms point away from the ion, keeping the C–OW distance fixed, the interaction energy increases to +5 kcal/mol. If, however, the water molecule in configuration *A* is rotated around the C–OW vector (i.e. the HW–OW–HW bisector), the rotation barrier is only about 0.5 kcal/mol, the lowest energy occuring when the two molecules are co-planar.

When the water molecule resides *close to the global minimum* on the oxygen side of the formate ion the energy barrier for rotation around the C–OW vector is much larger. If the water molecule here is rotated 90° so that the hydrogen atoms are still pointing towards the formate ion but one of the formate plane, the minimum moves out to a C–OW distance of about 3.6 Å with an energy minimum of about −11 kcal/mol. If the water molecule is flipped around 180°, so that the oxygen atom is closest to the formate ion, one obtains a highly repulsive potential curve (curve *C* in Fig. 1).

When the water molecule resides *close to one of the two formate oxygen atoms* the minimum energy occurs for a linear hydrogen bond, and corresponds to an interaction energy of about −12 to −15 kcal/mol, as has been discussed above. Here the rotation barrier about the O···HW–OW bond is about 0.8 kcal/mol.

The SCF-calculated energies and the atomic coordinates can be obtained from the authors on request.

4. Summary

The potential surface for the interaction between a rigid formate ion and a rigid water molecule has been derived by *ab initio* calculations. An analytical potential expression was constructed to fit the SCF energies. The global minimum corresponds to an interaction energy of about −16 kcal/mol, while the minimum interaction energy is −5 kcal/mol when the water molecule is located on the hydrogen side of the formation ion. At the global minimum the water molecule is positioned symmetrically on the O–C–O bisector forming a hydrogen bond with each of the two formate oxygen atoms. The energy surface is very flat near the global minimum.

Finally, it should be noted that any realistic simulation of a formate solution should include many-body corrections. This is known to be true even for the case of liquid water [30]. The presence of the formate ions should make the need for many-body corrections even stronger.

References

1. Clementi E, Popkie H (1972) J Chem Phys 57:1077
2. Kistenmacher H, Popkie H, Clementi E (1973) J Chem Phys 58:5627
3. Bounds DG, Bounds PJ (1983) Mol Phys 50:25
4. Dacre PD (1984) Mol Phys 51:633
5. Chandrasekhar J, Spellmeyer DL, Jorgensen WL (1983) J Am Chem Soc 106:903
6. Lybrand TL, Kollman P (1985) J Chem Phys 83:2923
7. Heinzinger K, Vogel PC (1976) Z Naturforschung 31A:463
8. Perez P, Lee WK, Prohofsky EW (1983) J Chem Phys 79:388
9. Thimme Gowda B, Benson SW (1983) J Chem Phys 79:1235
10. Dietz W, Riede WO, Heinzinger K (1982) Z Naturforsch 37a:1038
11. Probst MM, Radnai T, Heinzinger K, Bopp P, Rode BM (1985) J Am Chem Soc 89:753
12. Bohm HJ, McDonald IR (1964) J Chem Soc (Faraday Trans 2) 80:887
13. Welti M, Ha TK, Pretsch E (1985) J Chem Phys 83:2959
14. Howell JM, Sapse AM, Singman E, Synder G (1982) J Phys Chem 86:2345
15. Alagona G, Ghio C, Kollman P (1983) J Am Chem Soc 105:5226
16. Hermansson K, Lie GC, Clementi E (1986) IBM Kingston Tech Rep No 54
17. Ortoleva E, Castiglione G, Clementi E (1980) Comp Phys Comm 19:337
18. Tellgren R, Ramanujam PS, Liminga R (1974) Ferroelectrics 6:191; Fuess H, Burger N, Bats JW (1980) Z Kristallografie 156:219; Bats JW, Fuess H (1980) Acta Cryst B 36:1940; Fuess H, Bats JW, Dannohl H, Meyer H, Schweig A (1982) Acta Cryst B 38:736; Hermansson K, Tellgren R, Lehmann MS (1983) Acta Crystallogr C 39:1507
19. Chiari G, Ferraris G (1982) Acta Crystallogr B Sect B 38:2331
20. Dunning TH (1970) J Chem Phys 53:2823
21. Huzinaga S (1965) J Chem Phys 42:1293
22. Roos B, Siegbahn P (1970) Theor Chim Acta 17:199
23. Boys SF, Bernardi F (1970) Mol Phys 19:553
24. Schwenke DW, Truhlar D (1985) J Chem Phys 82:2418; Frisch MJ, Del Bene JE, Binkley JS, Schaefer HF (1986) J Chem Phys 84:2279
25. Greenwood HH, Plant JS (1986) Int J Quantum Chem 30:127, Bonaccorsi R, Cammi R, Tomasi J (1986) Int J Quantum Chem 29:373; Szczensniak MM, Scheiner S (1986) J Chem Phys 84:6328; Sokalski WA, Lowrey AH, Roszak S, Lewchenko V, Blaisdell J, Hariharan PC, Kaufman JJ (1986) J Comput Chem 7:693 (1986)
26. Sordo JA, Chin S, Sordo TL (1987) IBM Kingston Tech Rep No 130
27. Lie GC, Clementi E, Yoshimine M (1976) J Chem Phys 64:2314
28. Matsouka O, Clementi E, Yoshimine M (1976) J Chem Phys 64:1351
29. Probst M, and Clementi E (1985) IBM Kingston Tech Rep No 33
30. See, for example, Wojcik M, Corongiu G, Detrich J, Mansour MM, Clementi E, Lie GC (1986) In: Dupius M (ed) Super-computer simulations in chemistry. Springer, Berlin Heidelberg New York, p 224