Theor Chem Acc (2000) 105:68–76
DOI 10.1007/s002140000187

Theoretical
Chemistry Accounts

Regular article

Prediction of the Raman spectrum of the aqueous formate anion by a combined density functional theory and self-consistent-reaction-field study

A. L. Magalhães, S. R. R. S. Madaíl, M. J. Ramos

CEQUP/Departamento de Química, Faculdade de Ciências, Universidade do Porto, 687 R. campo Alegre, 4169–007 Porto, Portugal

Received: 15 May 2000 / Accepted: 19 June 2000 / Published online: 11 September 2000

Abstract. Formate, $HCOO^-$, is a small ion that can be used as a convenient model to represent general carboxylic acids as well as the terminal side chains of glutamic and aspartic amino acids in biological systems. Several ab initio techniques (Hartree–Fock and density functional theory methods) were used to study the formate anion in aqueous solution, where the solvent was treated as an electrostatic influence by means of a self-consistent-reaction-field (SCRF) method. The comparison of calculated frequencies with Raman experimental data constitutes a good test for the solvation model employed in this work. The results showed that the application of a SCRF method not only produced smaller errors but also improved the linear correlation between predicted and experimental values. This latter characteristic enables a more efficient application of a linear scaling factor, λ, which corrects the calculated frequencies. In fact, the application of λ to the values calculated for the continuum, when compared to those obtained for a vacuum, resulted in smaller root-mean-square errors of the deviations from the experimental data.

Key words: Formate anion – Density functional theory – Solvent effects – Continuum – Raman spectra

1 Introduction

In the recent past, density functional theory (DFT) became a really powerful theoretical tool for a wide range of chemical interests. Nowadays, DFT associates accuracy with low computational cost in such a way that it constitutes an attractive alternative to other post-Hartree–Fock (HF) procedures, such as the second-order Møller–Plesset (MP2) method [1].

Correspondence to: A. L. Magalhães,
e-mail: almagalh@fc.up.pt.

In DFT, the total electronic energy of a molecular system, E_{DFT}, can be written as a functional of the electron density, ρ:

$$E_{\mathrm{DFT}}[\rho] = T[\rho] + E_{\mathrm{ne}}[\rho] + J[\rho] + E_{\mathrm{xc}}[\rho], \quad (1)$$

where $T[\rho]$ is the kinetic energy, $E_{\mathrm{ne}}[\rho]$ the attractive energy between nuclei and electrons, $J[\rho]$ is the Coulomb term of the electron–electron interaction and $E_{xc}[\rho]$ is the exchange–correlation functional, which includes an exchange energy, the largest contribution to $E_{xc}[\rho]$ by far, and a correlation energy. It is usual to separate $E_{xc}[\rho]$ into those two parts, a pure exchange and a correlation part:

$$E_{\mathrm{xc}}[\rho] = E_{\mathrm{x}}[\rho] + E_{\mathrm{c}}[\rho] \quad (2)$$

Several functionals have been proposed for these energy terms, which can be classified roughly into three groups: the local density approximation (LDA) assumes that the density can be treated locally as a uniform electron gas, showing, therefore, a slowly varying function; the gradient-corrected or nonlocal energy functionals which benefit from the inclusion of density derivatives, allowing, in principle, a more correct description of atoms, molecules and their bonds; the hybrid functionals, which include combinations of exact exchange terms as given by the HF theory and DFT exchange–correlation functionals, the coefficients of combination being determined by appropriate fits to experimental data. Nowadays, a wide variety of functionals proposed for $E_{\mathrm{xc}}[\rho]$ may be found in the literature but well-established rules are still missing, which often makes the choice of a particular functional difficult.

Predictions of vibrational spectra are amongst the most important and current applications of DFT techniques since they have shown a good performance in the definition of molecular force fields. The implementation of analytical first and second derivatives has played a crucial role in such success [2]; however, the predicted frequencies have remained affected by errors owing, to a great extent, to anharmonicity effects. The observation that such errors occur approximately in a systematic way led several authors to propose different scaling procedures at semiempirical, HF and DFT levels of quantum

calculation [3, 4, 5, 6, 7, 8, 9]. These methodologies involving scaling have proved to be successful in reproducing vibrational spectra, especially for the selected set of molecules for which they have been calibrated. Transferability of the scaling factors is always a major problem, even for molecules exhibiting similar structural features. Besides, there is another source of errors – the inadequacy of the theoretical methods to include environmental effects such as the natural packing in the crystal structure, the presence of counterions in ionic substances and solvation effects when experimental data are obtained with solution samples. In fact, the great majority of chemical mechanisms, and, in particular, the processes with biological interest, do take place within a surrounding medium; therefore, its influence should not be neglected in a theoretical treatment if accurate results are desired. Among several theoretical methodologies proposed in the last two decades to describe the solvent effect at a quantum mechanical level, continuum models have shown flexibility and enough accuracy to become a popular tool [10, 11, 12, 13]. In these self-consistent-reaction-field (SCRF) solvation methods the solute is placed inside a cavity with appropriate shape, made in a continuous medium characterized by a dielectric constant. The electronic distribution of the solute induces a charge density at the surface of the cavity, which creates a field that modifies the energy and properties of the solute. The effect of the reaction field is solved iteratively in the SCRF method by the inclusion of a supplementary potential term in the solute Hamiltonian.

One of the most important agents in biological processes is the carboxylate group, which rules the interactions established by the side chains of glutamic and aspartic acids. The formate anion has been the subject of many experimental and theoretical studies, and it has also been used as a very convenient model for those terminal side chains, enabling, therefore, the theoretical study of their chemical properties[14].

As mentioned previously, the solvent affects the solute properties to a variable extent, which is not easy to quantify without the use of accurate calculations. In the present work, the solvation models employed only take into account electrostatic effects of the solvent, which are the most important interactions to be considered for a charged molecular system such as the formate anion.

The performance of several DFT functionals in reproducing Raman vibrational spectra of the formate anion in aqueous solution was evaluated in this work. The solvent effect was modeled by two SCRF solvation methods and the influence of the basis set was also considered.

2 Methods

All the calculations were based on HF and DFT methods as implemented in Gaussian98 [15]. The post-HF Møller–Plesset(MP*n*) methods were used also at the second, third and fourth order of perturbation, as a reference to evaluate the ability of DFT functionals to include electronic correlation effects. As far as DFT methods are concerned, the exchange functionals, $\boldsymbol{E}_x[\rho]$, used in this study include Slater's LDA [16] (abbreviated S) with an exchange scaling factor of 2/3, the X_α functional, which is an older version developed by Slater (XA) [16, 17] with a different value for the exchange scale factor, Becke's functional (B) [18], which includes a gradient correction, the functional proposed by Perdew and Wang in 1991 (PW91) [19] and the corresponding version of Adamo and Barone (MPW) [20], and finally the recent functional of Gill (G96) [21]. The correlation functionals, $E_c[\rho]$, considered here are the local spin density form of Vosko, Wilk and Nusair (VWN) [22], the gradient-corrected functional of Lee, Yang and Parr (LYP) [23], the local (non-gradient-corrected) functional of Perdew (PL) [24] and Perdew's functionals with gradient corrections – versions from 1986 (P86) [25] and from 1991 with Wang (PW91) [19]. These functionals were combined to give 33 different methods: S—null, B-null and XA–null, respectively, Slater, Becke and X_α, which include only exchange functional terms to $E_{xc}[\rho]$; and 30 exchange-plus-correlation functionals, $E_{xc}[\rho] = E_x[\rho] + E_c[\rho]$,which were the result of the association of six exchange components S, B, XA, PW91, MPW and G96 with five correlation terms VWN, LYP, PL, P86 and PW91. Additionally, seven hybrid methods were also tested: one using equal contributions of HF exchange and Slater exchange functional without correlation (BHandH) [26]; the same as before but with LYP correlation added (BHandHLYP) [23]; Becke's three-parameter exchange–correlation functional with nonlocal correlation corrections provided by three different terms: Perdew's expression (B3P86) [27], the LYP expression (B3LYP) [23, 27] and the Perdew–Wang expression (B3PW91) [25, 27]; the version proposed by Adamo and Barone of Becke's one-parameter functional with the LYP correlation term (B1LYP) [20, 28] and their version of Becke's one-parameter functional with modified Perdew–Wang exchange and correlation terms (MPW91PW91) [20].

The SCRF methods employed here to model the electrostatic influence of the solvent were the conductor-like solvation model approach of the polarizable continuum model (COSMO–PCM [29]) and the integral equation formulation of the PCM model (IEF–PCM [11, 13]). Therefore, the formate anion was inserted in a cavity with a shape based on interlocking spheres centered on heavy atoms C and O, as defined by the united atom model for HF [30]. This cavity was formed in a continuous medium characterized by a dielectric constant of 78.39 in order to simulate the electrostatic influence of water bulk. The SCRF methods were used at the same HF and DFT levels of those calculations previously performed for a vacuum.

The energy minimizations, both for a vacuum and for a continuum, were performed under the C_{2v} symmetry constraints. The geometrical parameters which were allowed to be optimized are shown in Fig. 1.

The basis sets of atomic functions employed throughout this work were the split-valence 3-21G and the expanded basis set 6-311++G(d,p), where valence orbitals are split into three functions and a complete set of diffusion and polarization functions are added to hydrogens (s and p types) and also to other atoms (p and d types). Previous calculations made by Pople et al. [31] at the HF level have shown that good predictions of vibrational frequencies were obtained with the smaller basis set, 3-21G, for a reasonable set of neutral molecules. The authors concluded that expansion of the basis set to 6-31G(d) did not show an improvement in the calcu-

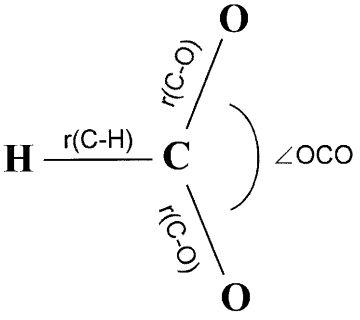

Fig. 1. Geometrical parameters of the formate anion which were allowed to be optimized

lated values when compared with experimental data. This fact, which we have concluded to be still valid for the present charged molecular system, is very helpful because it will enable the study of bigger molecular aggregates. In this work, the calculations were always carried out with the 3-21G and 6-311 + + G(d,p) basis sets and, when thought necessary, intermediate basis sets were also used for comparison.

The statistical analysis of the errors in the predicted harmonic frequencies was based on the root-mean-square (rms) error:

$$\mathrm{rms} = \sqrt{\frac{\sum_i^N \varepsilon_i^2}{N}}, \tag{3}$$

where the summation is extended to all N vibrations and ε_i stands for the relative deviation from the experimental value as defined by

$$\varepsilon_i = \frac{\omega_i^{\mathrm{calc}} - \omega_i^{\mathrm{expt}}}{\omega_i^{\mathrm{expt}}}, \tag{4}$$

in which ω_i^{calc} and ω_i^{expt} are the calculated harmonic frequencies and the experimental fundamentals, respectively. The use of adimensional relative errors seems more suitable for comparisons between different methods and molecular systems.

For each theoretical method, the optimized scaling factor, λ_k, which scales the six harmonic frequencies as

$$\omega_i^{\mathrm{scl}} = \lambda_k \omega_i^{\mathrm{calc}} \tag{5}$$

was obtained by minimizing the relative errors between ω_i^{expt} and ω_i^{scl}. A least-squares fitting procedure was used to minimize such error functions and resulted in the following expression:

$$\lambda_k = \frac{\sum_i^N \frac{\omega_i^{\mathrm{calc}}}{\omega_i^{\mathrm{expt}}}}{\sum_i^N \left(\frac{\omega_i^{\mathrm{calc}}}{\omega_i^{\mathrm{expt}}}\right)^2}. \tag{6}$$

3 Discussion

3.1 Optimized structure in a vacuum

The geometry of the formate anion was optimized for a vacuum at the HF and MP*n* levels, under C_{2v} symmetry constriction. The geometrical parameters calculated with several basis sets and, for comparison, a range of experimental values obtained by X-ray and neutron diffraction techniques with sodium and lithium salts [32, 33, 34] are presented in Table 1. The variety of basis sets employed allows some conclusions to be drawn with respect to the effect of the size of the basis sets on the calculated parameters.

In general, for all the methods presented in Table 1, the calculated values of the C–O and C–H bond lengths decrease with the complexity of the basis set. Compared with the poorest double-ζ basis set, 3-21G, the greatest variation in this series of calculations occurs when polarization functions are added. Further improvements, namely the addition of a third valence function and a set of s- and p-type diffusion functions, produce only smaller changes in these parameters. In particular, the addition of a diffusion function on the hydrogen atom does not bring any significant change in the geometry. HF methods are well known to underestimate bond lengths [31] as happens with the C–O bond; however the C–H bond turns out to be an exception, with its value being overestimated. These two opposite tendencies are partially corrected when some electronic correlation effects are considered with the use of the MP*n* methods. Concerning the bond angle of the carboxylate group, some consistency is observed throughout this table.

The effect of the electronic correlation on the properties of a charged system, such as the formate anion, was also analyzed by a systematic application of several exchange and correlation functionals within the context of DFT methods. The size of the formate anion enabled the computation with the larger basis sets to be carried out in a reasonable time; however, if a bigger molecular system is to be considered in the future, there will be the necessity of choosing a basis set which should present a good compromise between quality and computational effort. Therefore, calculations were performed with 40 different DFT functionals and three different levels of basis sets, namely the smallest double-ζ split valence, 3-21G, the double-ζ with polarization functions, 6-31G(d,p), and the triple-ζ with polarization and dif-

Table 1. Optimized geometry of the formate anion in a vacuum obtained with Hartree–Fock (*HF*) and post-HF methods. The notation is based on Fig. 1. The bond lengths are in angstroms and the bond angles in degrees

Method	Parameter	Basis set of atomic functions				
		3-21G	6-31G(d,p)	6-311G(d,p)	6-311 + G(d,p)	6-311 + + G(d,p)
HF	*r*(C–O)	1.249	1.231	1.227	1.230	1.230
	r(C–H)	1.126	1.130	1.132	1.122	1.220
	∠OCO	130.76	130.96	131.05	130.54	130.54
MP2	*r*(C–O)	1.284	1.261	1.252	1.258	1.258
	r(C–H)	1.152	1.139	1.149	1.133	1.134
	∠OCO	131.76	131.05	131.39	130.43	130.44
MP3	*r*(C–O)	1.271	1.251	1.242	1.246	1.246
	r(C–H)	1.149	1.135	1.146	1.134	1.134
	∠OCO	130.68	130.96	131.30	130.52	130.54
MP4	*r*(C–O)	1.291	1.265	1.258	1.264	1.264
	r(C–H)	1.164	1.144	1.156	1.139	1.140
	∠OCO	130.86	131.07	131.50	130.42	130.48
Expt. [32–34]	*r*(C–O)			1.242–1.252		
	r(C–H)			1.087–1.106		
	∠OCO			125.53–126.40		

fusion functions, 6-311++G(d,p). The huge amount of data obtained is not presented here for simplicity, but it is available upon request. A general analysis of these results shows that the dependence of the calculated bond lengths on the size of the basis set is again observed. The ∠OCO parameter is shown to be insensitive to the nature of the functional with a range of values similar to that of Table 1. Concerning the C–O bond length, the performance of the Slater local spin density exchange functional, regardless of the correlation functional associated with it, is slightly better when compared with other combinations. The good quality of the hybrid functionals must also be mentioned (the BHandH method constitutes the only exception with an underestimated value of 1.231 Å). The calculations for the C–H bond length present some dispersion (the amplitude of the values is 0.048 Å for the biggest basis set), but several functionals show a performance similar to that of the MP*n* methods; this is in fact the case for all the hybrids. The reduction in computation time observed with DFT methods when compared with MP*n*, about 50%, confirms our expectations that they may constitute a good alternative when one needs some electronic correlation effects to be added.

When a direct comparison is made with the available experimental geometries, the deviation from the calculated values for a vacuum is evident; however, some care must be taken with this conclusion because the disagreement may result from a distortion of the internal geometry of the formate, as a consequence of the packing with neighboring ions in the crystal structure.

3.2 Calculated frequencies in a vacuum

For the reasons exposed earlier, the experimental geometry may not be the best molecular characteristic to be used as a reference in the study of the reliability of theoretical methods; therefore, a more objective criterion should be applied, such as the frequencies of molecular vibrations obtained by IR and Raman spectroscopic experiments. However, as for the structural packing in solid crystals, the presence of a solvent may definitely influence the vibrational spectra of the molecular systems. This work compares the frequencies calculated at several theoretical levels with experimental Raman fundamentals obtained with aqueous solutions of formate salts [35, 37]. A first step in the overall strategy of this study was the comparison of experimental data with theoretical values predicted for the ion in a vacuum. At the HF and MP*n* theoretical levels the calculations were carried out with five different basis sets. The results are depicted in Table 2 and Figs. 2 and 3.

For all these methods, the assignments of the frequencies to the C_{2v} normal modes of vibration agrees with previous studies [37, 38] and are indicated at the top of Table 2. The effect of the basis set of atomic functions on the results is evident. In general, when the size of the basis set increases, the majority of the predicted values become smaller, which implies a decrease in the overall rms error of the relative deviations. These results also show that when diffusion functions are added, the values of the rms error suffer a substantial decrease, mainly owing to variations in the predictions of the high-frequency C–O asymmetric stretch and C–H stretch modes;

Table 2. Unscaled harmonic frequencies (cm^{-1}) and root-mean square (*rms*) errors for the formate anion in a vacuum obtained with HF and post-HF methods

Method	Basis set of atomic functions	v_3	v_6	v_2	v_5	v_4	v_1	rms (%)
		(A_1) C–O scissor	(B_1) C–H out-of-plane bend	(A_1) C–O symmetric stretch	(B_2) C–H in-plane bend	(B_2) C–O asymmetric stretch	(A_1) C–H stretch	
HF	3-21G	829.3	1208.2	1433.6	1585.9	1879.3	2686.6	11.8
	6-31G(d,p)	836.4	1215.0	1513.4	1545.2	1918.8	2745.2	12.8
	6-311G(d,p)	839.6	1207.2	1490.1	1524.5	1879.3	2697.0	11.6
	6-311+G(d,p)	824.9	1197.6	1482.3	1514.4	1817.8	2797.3	9.9
	6-311++G(d,p)	824.8	1197.4	1481.9	1514.7	1817.7	2795.1	9.9
MP2	3-21G	769.5	1081.9	1268.6	1501.1	1733.3	2423.3	8.2
	6-31G(d,p)	762.7	1086.0	1368.6	1452.8	1794.1	2633.6	6.5
	6-311G(d,p)	768.6	1069.4	1349.6	1416.1	1770.6	2511.7	6.7
	6-311+G(d,p)	742.8	1051.5	1335.9	1403.2	1670.8	2672.1	3.7
	6-311++G(d,p)	742.8	1051.3	1335.3	1404	1670.8	2667.7	3.7
MP3	3-21G	788.7	1101.1	1333.4	1506.3	1775.6	2467.3	8.1
	6-31G(d,p)	784.8	1117.0	1427.8	1475.5	1837.2	2680.8	7.9
	6-311G(d,p)	792.5	1105.5	1412.6	1443.4	1817.7	2558.8	7.8
	6-311+G(d,p)	773.0	1093.9	1405.2	1434.9	1741.4	2686.6	5.1
	6-311++G(d,p)	773.0	1093.6	1404.8	1435.5	1741.4	2683.9	5.1
MP4	3-21G	747.5	1048.7	1220.1	1449.0	1676.7	2289.7	9.3
	6-31G(d,p)	750.4	1067.3	1334.5	1432.8	1756.2	2567.8	6.1
	6-311G(d,p)	754.0	1046.3	1310.3	1391.3	1728.3	2434.5	7.0
	6-311+G(d,p)	726.2	1025.4	1294.1	1377.7	1625.0	2605.4	4.8
	6-311++G(d,p)	726.2	1025.9	1293.5	1379.3	1625.8	2597.0	4.9
Expt. [35–37]		773	1069	1352	1386	1584	2825	–

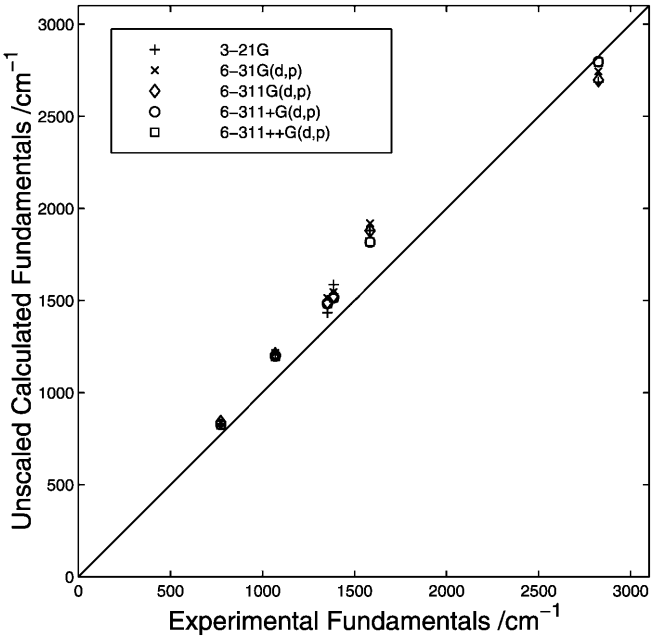

Fig. 2. Comparison between experimental data and Hartree–Fock (*HF*) calculations carried out, for a vacuum, with five different basis sets

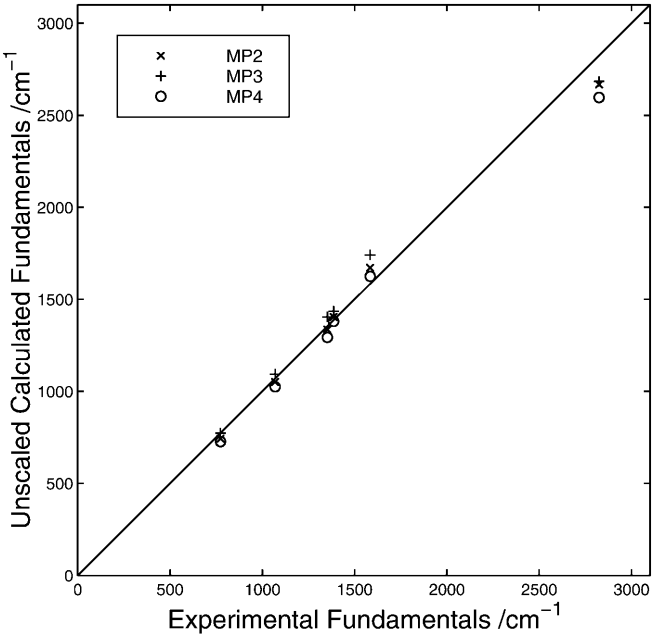

Fig. 3. Comparison between experimental data and MP*n*/6-311++G(d,p) calculations carried out for a vacuum

obviously, the addition of diffusion functions to the hydrogen atoms affects only the C–H stretch mode, but the extent of this effect is small.

The results obtained with the MP*n* methods show smaller rms errors when compared with pure HF predictions. The harmonic frequencies calculated at the HF level are well known to be systematically overestimated by about 12% when compared with the experimental fundamentals [31]. Several explanations may be put forward, for example, the lack of anharmonicity contributions in the theoretical potential of the vibration, the inappropriate size of the basis sets of atomic functions, the absence of electron correlation effects and the improper description of the real ambience felt by the molecular system under study. The observation of a general and relative uniform overestimation has led some authors to propose different scaling factors, depending on the class of molecules, the theoretical method employed and the basis set of atomic functions, which have revealed the ability to reproduce experimental results with high accuracy [1, 5, 39]. However, the application of an empirical factor that scales the harmonic frequencies uniformly must be undertaken with care, because neither all the normal modes follow those statistics nor will all molecular systems be sensitive to the scaling procedure. In particular, a previous study on a positively charged molecular system, namely the guanidinium ion, showed that there is a random dispersion of the calculated harmonic frequencies [40]. For the special case of the formate anion, it is clear from Fig. 2 that there is such a dispersion in the predicted values for a vacuum for all the basis sets considered. In particular, it is interesting to note that the highest frequency (C–H stretch) is even underestimated.

Figure3 shows a similar dispersion for MP*n*/6-311++G(d,p) methods, although a general better agreement with experimental data is achieved. The three smallest frequencies are now slightly underestimated but the C–H stretch mode is predicted with a larger error.

The extrapolation of these observations to the DFT methods may not be straightforward. A systematic study with all 40 DFT methods was carried out with three different basis sets, namely 3-21G, 6-31G(d,p) and 6-311++G(d,p). In order to simplify the comparison with the previous methods, only the results obtained with the biggest basis set are summarized in Fig. 4, and they are grouped in three distinct classes according to the nature of the functional. The rms error of the deviations decreased with the size of the basis set in a similar way to that observed with MP*n* methods. The pure exchange functionals presented the worst predictions, with a slightly better performance being shown by the Xα functional, which shows that the inclusion of correlation terms is important. In general, and concerning the largest basis set, the combination of exchange and correlation functionals improved the quality of the results, decreasing the rms error to about 7%; the XA–VWN combined functional showed here a value of 5.9% for the rms error. The class of the hybrid functionals showed an improved performance, with an rms error around 5%, which compares well with the MP*n* values. The best prediction (4.4%) was obtained here with the B1LYP functional, followed by B3LYP (4.7%). This agrees with previous calculations with 31 organic molecules at the DFT/6-31G* level [4]. The authors concluded there that the B3LYP hybrid functional systematically overestimated the vibrational frequencies by about 5%, due partly to anharmonicity effects and also to a general

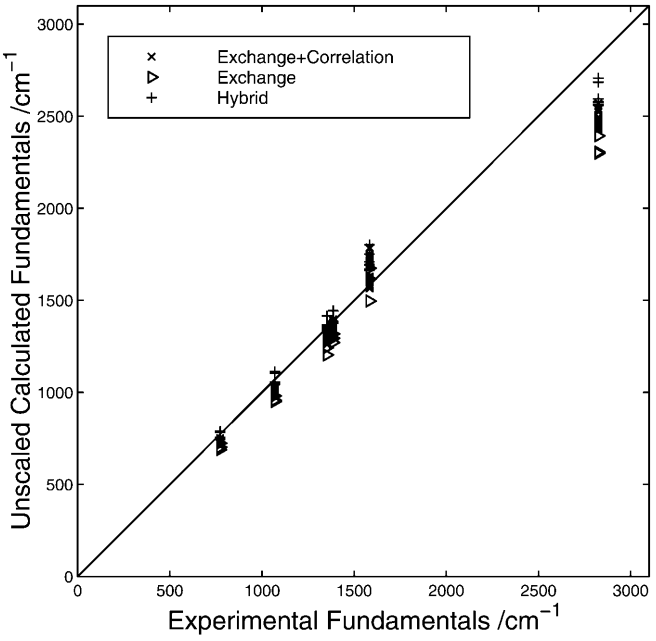

Fig. 4. Comparison between experimental data and the results obtained, for a vacuum, with three different classes of density functional (*DF*) theory functionals – basis set 6-311 + + G(d,p)

overestimation of force constants at the exact equilibrium geometries. In the particular case of the formate anion, the main contribution to the errors is the inaccurate prediction of the two highest frequencies (C–O asymmetric stretch and C–H stretch), since the agreement with the experimental data for the remainding four values is quite good.

Figure 4 shows in a clearer way that, in general, the reliability of the DFT methods is fairly good. The high frequency corresponding to the C–H stretch mode constitutes the worst prediction for all these methods, though some improvement is observed when hybrid functionals are considered.

3.3 Optimized structure in aqueous solution

The same set of three different basis sets was used here within the two SCRF methods employed. The effect of the basis set on the geometry was similar to that found for a vacuum. The IEF–PCM had only a small effect on the geometries previously optimized for a vacuum. In fact, for all methods, the ⟨OCO bond angle suffered a decrease of only 1–3° and the C–O bond length predictions increased by only about 0.3%. The calculated C–H bond length suffered a bigger change under the influence of the reaction field; however, the extent of that perturbation is still small except for the case of the HF/6-311 + + G(d,p) method, where the bond length changed from 1.220–1.099 Å; for all the DFT functionals the decrease was around 3%. The SCRF models only take into account electrostatic effects of the solvent, here shown to have little influence on the equilibrium geometries of the isolated ion.

3.4 Calculated frequencies in aqueous solution

The unscaled harmonic frequencies were calculated at the HF and DFT levels after the insertion of the ion in a cavity formed in a continuous medium characterized by a dielectric constant of 78.39. Both SCRF schemes were used with the same three basis sets: 3-21G, 6-31G(d,p) and 6-311 + + G(d,p); however, for simplicity, only the results obtained at the IEF–PCM/6-311 + + G(d,p) level are presented because they showed better agreement with the experimental data. Nevertheless, all the other results are available upon request.

The data show that, in general, the size of the basis set reduces the overall rms error value. When polarization functions are introduced, there is a significant improvement in the calculated values. The splitting of the valence shell into three functions and the addition of diffusion functions lowered the rms error to a value of about half of that found with the 3-21G basis set. In particular, the predictions of functionals such as Xa–null or those including the Slater local spin density exchange functional can reach 3% rms error. The predictability of B3LYP, with an rms error of 2.0%, is noteworthy. When comparison is made with the results obtained for a vacuum, the conclusion is that for all the methods and basis sets employed the use of either solvation model lowered the rms errors.

For some selected classes of theoretical treatment, the experimental data and the two types of calculated results (for a vacuum and for a continuum) can be easily compared in the "spider plots" presented in Figs. 5–8. The plane is partitioned into three equal sectors, bisected by a broken line. The three axes represent the type of values analyzed (experimental, calculated for a vacuum and calculated for a continuum) and each pair of values is compared in the appropriated sector. These spider

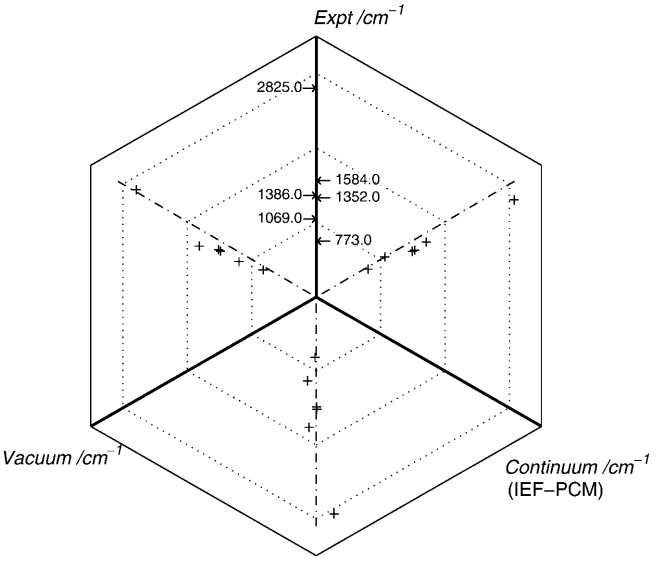

Fig. 5. Comparison between experimental fundamentals and the unscaled harmonic frequencies calculated with the HF/6-311 + + G(d,p) method for a vacuum and within the integral equation formulation of the polarizable continuum model (*IEF–PCM*)

plots may be visualized as simple 3D plots where projections are made in their orthogonal planes.

The results obtained from pure HF calculations are presented in Fig. 5. The rms error value changed from 9.9% (for a vacuum) to 7.8%, which means that the inclusion of the reaction field does not introduce a great improvement in the HF results. In addition, this figure shows that the effect of the continuum is not the same for all six frequencies. There are slight corrections in the lower frequencies contributing to the overall decrease in the rms error despite the worst prediction in the C–H stretch vibration, which changed from 2795.1 cm^{-1} (for a vacuum) to 3073.2 cm^{-1}.

Concerning DFT methods, the results were again collected into three different groups depending on the nature of the functionals used. The results obtained with pure exchange functionals, hybrid functionals and exchange-plus-correlation functionals are summarized in Figs. 6–8, respectively.

An inspection of these three plots reveals that for all classes of functionals the dispersion of values for each calculated frequency is almost conserved, though a general better agreement with experimental data is observed for the continuum. The C–H stretch frequency is the calculation which seems to be more affected by the inclusion of the solvation model. The considerable decrease observed in the overall rms error for the pure exchange and exchange-plus-correlation functionals (Figs. 6, 8, respectively) is due mainly to the correction introduced in that prediction. For all functionals, that calculation suffered a deviation to larger values; in the particular case of the hybrid functionals it even changes from an underestimated value for a vacuum to an overestimation for the continuum, though similar absolute shifts are observed. As the plots in the lower sectors show, the calculations are not all corrected in the same way by the inclusion of the solvation model. This property of the calculated values resulted in an interesting effect, observed in the sectors on the right-hand side (experimental versus continuum), which is an increase in linear correlation between those two variables when compared with the experimental versus vacuum plots (left-hand-side sectors). This improving linearity between the experimental data and the calculations is important because the values become more suitable for the application of an overall linear scaling factor.

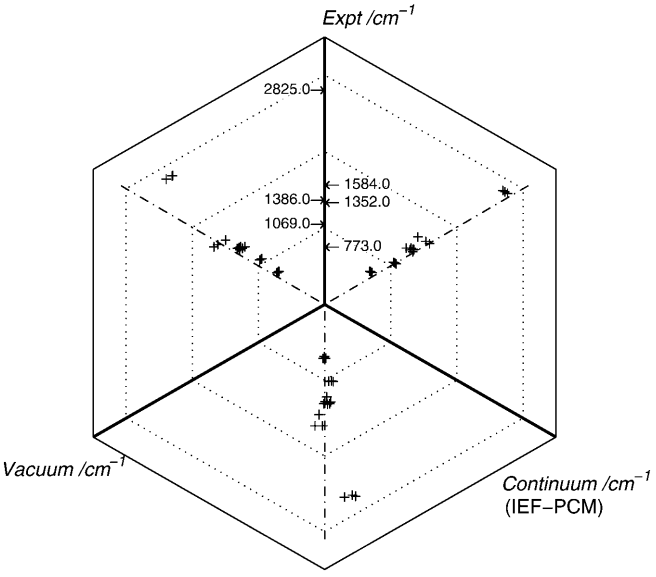

Fig. 6. Comparison between experimental fundamentals and the unscaled harmonic frequencies calculated with the DF/6-311++G(d,p) method (pure exchange functionals) for a vacuum and within the IEF-PCM

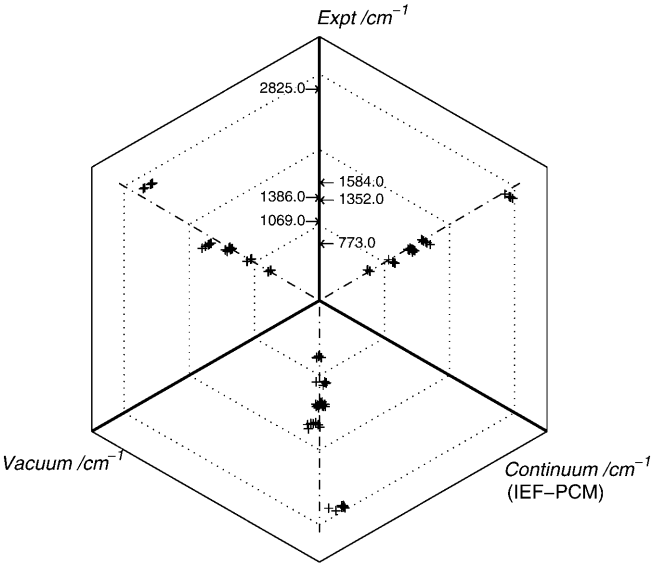

Fig. 7. Comparison between experimental fundamentals and the unscaled harmonic frequencies calculated with the DF/6-311++G(d,p) method (hybrid functionals) for a vacuum and within the IEF–PCM

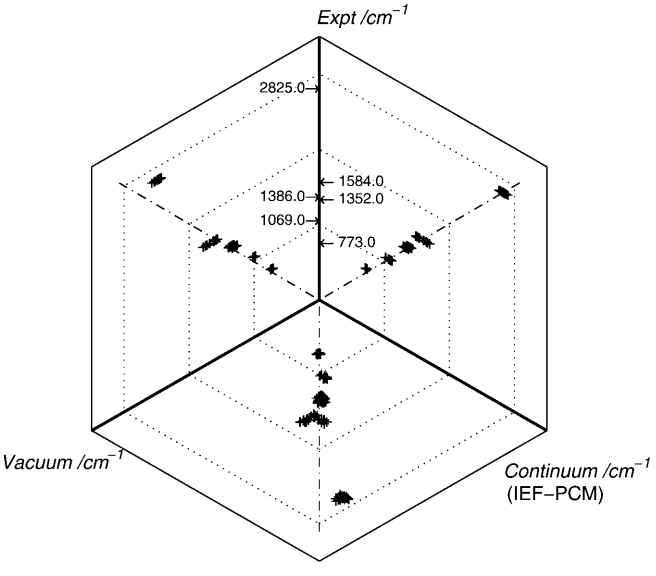

Fig. 8. Comparison between experimental fundamentals and the unscaled harmonic frequencies calculated with the DF/6-311++G(d,p) method (exchange-plus-correlation functionals) for a vacuum and within the IEF–PCM

Table 3. Scaling factors, linear correlation coefficients and rms errors obtained with all the theoretical methods and the 6-311++G(d,p) basis set

Method	Vacuum				IEF–PCM			
	Linear correlation coefficient (r)	rms without correction (%)	Scaling factor (λ_k)	rms with correction (%)	Linear correlation coefficient (r)	rms without correction (%)	Scaling factor (λ_k)	rms with correction (%)
Hartree–Fock	0.99315	9.9	0.91931	4.6	0.99867	7.8	0.93504	3.5
Pure exchange functionals								
DF/B – null	0.99251	11.2	1.11458	4.6	0.99686	7.6	1.06628	4.4
DF/S – null	0.98119	10.7	1.09236	6.6	0.99947	5.6	1.05309	2.4
DF/XA – null	0.98188	8.4	1.05611	6.6	0.99922	3.3	1.01926	2.7
Exchange-plus-correlation functionals								
DF/B–VWN	0.99455	6.9	1.06019	3.9	0.99770	4.7	1.02174	4.1
DF/B–LYP	0.99338	7.8	1.07117	4.2	0.99847	4.8	1.03219	3.6
DF/B–PL	0.99439	7.4	1.06704	3.9	0.99765	4.9	1.02796	4.1
DF/B–P86	0.99073	7.7	1.06394	4.8	0.99908	4.2	1.02709	3.3
DF/B–PW91	0.99078	7.3	1.05856	4.8	0.99904	4.0	1.02240	3.3
DF/XA–VWN	0.98615	5.9	1.00880	5.9	0.99916	3.6	0.98131	3.1
DF/XA–LYP	0.98259	6.8	1.01942	6.6	0.99880	3.3	0.99137	3.2
DF/XA–PL	0.98532	6.2	1.01490	6.0	0.99919	3.2	0.98653	3.0
DF/XA–P86	0.97851	7.4	1.01361	7.3	0.99790	3.9	0.98778	3.7
DF/XA–PW91	0.97899	7.3	1.00886	7.2	0.99791	4.0	0.98361	3.7
DF/S–VWN	0.98559	7.1	1.04086	6.0	0.99928	3.0	1.00681	2.9
DF/S–LYP	0.98191	8.3	1.05225	6.6	0.99913	3.3	1.01757	2.8
DF/S–PL	0.98515	7.5	1.04754	6.0	0.99930	3.1	1.01280	2.8
DF/S–P86	0.97753	8.6	1.04590	7.4	0.99825	3.6	1.01641	3.2
DF/S–PW91	0.97807	8.3	1.04071	7.3	0.99824	3.5	1.01132	3.3
DF/G96–VWN	0.99383	6.6	1.05492	4.1	0.99777	4.4	1.01714	4.1
DF/G96–LYP	0.99243	7.6	1.06560	4.4	0.99852	4.5	1.02748	3.6
DF/G96–PL	0.99354	7.2	1.06168	4.2	0.99773	4.7	1.02331	4.1
DF/G96–P86	0.98968	7.5	1.05865	5.1	0.99907	3.9	1.02250	3.2
DF/G96–PW91	0.98979	7.2	1.05342	5.0	0.99903	3.7	1.01768	3.3
DF/MPW–VWN	0.99469	6.8	1.05978	3.8	0.99945	4.7	1.04169	2.4
DF/MPW–LYP	0.99343	7.8	1.07070	4.2	0.99859	4.7	1.03196	3.6
DF/MPW–PL	0.99448	7.4	1.06656	3.9	0.99779	4.9	1.02772	4.1
DF/MPW–P86	0.99077	7.7	1.06364	4.8	0.99915	4.2	1.02685	3.2
DF/MPW–PW91	0.99086	7.3	1.05824	4.8	0.99922	3.3	1.01750	2.8
DF/PW91–VWN	0.99472	6.8	1.05941	3.8	0.99799	4.5	1.02101	4.0
DF/PW91–LYP	0.99334	7.8	1.07034	4.2	0.99872	4.7	1.03153	3.5
DF/PW91–PL	0.99450	7.3	1.06609	3.9	0.99795	4.8	1.02727	4.0
DF/PW91–P86	0.99065	7.7	1.06345	4.9	0.99922	4.1	1.02639	3.2
DF/PW91–PW91	0.99078	7.3	1.05803	4.8	0.99990	4.5	1.04226	2.0
Hybrid functionals								
DF/B3LYP	0.99273	4.7	1.01939	4.3	0.99983	2.0	1.00596	1.9
DF/B3P86	0.99114	4.8	1.01007	4.8	0.99884	4.2	0.97786	3.5
DF/B3PW91	0.99095	4.9	1.01217	4.8	0.99881	4.1	0.97975	3.5
DF/B1LYP	0.99306	4.4	1.01197	4.3	0.99833	4.4	0.97854	3.9
DF/MPW1PW91	0.99098	4.8	1.00217	4.8	0.99891	4.7	0.96966	3.5
DF/BhandH	0.98886	6.7	0.96066	5.3	0.99944	5.2	0.95447	2.0
DF/BhandHLYP	0.99326	5.6	0.96511	4.2	0.99972	4.8	0.95777	2.0

The statistical parameters obtained with the calculations carried out for a vacuum and within the IEF–PCM are compared in Table 3. The calculations performed for a vacuum showed poor linearity with the experimental data, which is evident from the small values of the linear correlation coefficient depicted in the first column of that table. The rms error values (second column) have been lowered (fourth column) with the correction introduced by the application of the scaling factor λ_k; however, the effect of λ_k was not the same for all methods and, as expected, those with better linear correlation decrease the rms error value more efficiently after correction. The table shows, quantitatively, the important effects caused by the inclusion of the IEF–PCM. First, a decrease in the overall rms error was observed for all functionals, which shows that the IEF–PCM really improves, in this particular molecular system, the prediction of the Raman fundamentals in solution. The other important effect is that the linear correlation coefficient, r, increased for all the theoretical methods. As a consequence, the optimum scaling factor, λ_k, obtained for the continuum now has a higher correcting power, as is shown by the rms error values of the last column. For the HF method a value of 3.5% was obtained, which is far from the

typical 10% of the calculations for a vacuum. Concerning DFT methods, the performances are in general good with a value of about 3% for the rms error after linear scaling. The behavior of some functionals is noteworthy, such as PW91–PW91 and the three types of hybrid functionals proposed by Becke, B3LYP, BHandH and BHandHLYP; these functionals all presented an excellent corrected rms error value of 2%.

4 Conclusions

Although a wide range of DFT applications may be found in the literature, the choice of functionals for a given calculation is not simple because general rules have not been established yet. Therefore, a systematic study was performed here, using several DFT functionals and basis sets for the prediction of the vibrational spectrum of a small charged molecular system, namely the formate anion. The results were compared with Raman spectroscopy data obtained with aqueous solutions of formate salts. When the calculations were performed for a vacuum, the results obtained with some DFT functionals showed, as expected, a quality similar to that of MP*n* methods. However, in order to take into account some solvation effects, an electrostatic term was included in the solute Hamiltonian by means of two SCRF methods. In general, this procedure improved the agreement with experimental data for all the functionals used. Although all the vibrational modes were shown to be sensitive to the influence of the polarizable medium, the largest correction observed in the frequency of C–H stretch mode was by far the most important factor for the excellent results obtained. The ability to increase the linear correlation between calculated harmonic frequencies and the experimental fundamentals constitutes a promising aspect of the SCRF methods, in particular the IEF–PCM model used with the 6-311 + + G(d,p) basis set. This important feature improved the correcting power of the linear scaling factor over the set of six calculated frequencies. In particular, the results obtained with the DF/PW91–PW91 exchange–correlation functional and the three hybrid functionals B3LYP, BHandH and BHandHLYP showed extraordinary agreement with experimental data after the application of the appropriate linear scaling factor. Although other systems need to be tested, these encouraging results indicate that the IEF–PCM can be used to correctly predict Raman spectra of species in solution.

Acknowledgements The authors thank the FCT-Fundação para a Ciência e Tecnologia for financial support under project PRAXIS/P/QUI/11302/98.

References

1. Pople JA, Scott AP, Wong MW, Radom L (1993) Isr J Chem 33: 345–350
2. Pople JA, Krishnan R, Schlegel HB, Binkley JS (1979) Int J Quantum Chem Quantum Chem Symp 13: 225–241
3. Pulay P, Fogarasi G, Pongor, Boggs JE, Vargha AJ (1983) J Am Chem Soc 105: 7037–7047
4. Rauhut G, Pulay P (1995) J Phys Chem 99: 3093–3100
5. Wong MW (1996) Chem Phys Lett 256: 391–399
6. El-Azhary AA, Suter HU (1996) J Phys Chem 100: 15056–15063
7. Manogaran S, Chakraborty D (1998) J Mol Struct (THEOCHEM) 432: 139–151
8. Magdo I, Nemeth K, Mark F, Hildebrandt P, Schaffner K (1999) J Phys Chem A 103: 289–303
9. Grunenberg J, Herges R (1997) J Comput Chem 18: 2050–2059
10. Tomasi J, Persico M (1994) Chem Rev 94: 2027–2094
11. Cancès E, Mennucci B, Tomasi J (1997) J Chem Phys 107: 3032–3041
12. Cossi M, Barone V, Mennucci B, Tomasi J (1998) Chem Phys Lett 286: 253–260
13. Mennucci B, Cancès E, Tomasi J (1997) J Phys Chem B 101: 10506–10517
14. (a) Zheng Y, Ornstein RL (1996) J Am Chem Soc 118: 11237–11243; (b) Sapse AM, Russel CS (1984) Int J Quantum Chem 26: 91–99; (c) Alagona G, Ghio C, Kollman P (1983) J Am Chem Soc 105: 5226–5230
15. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA, Startmann RE, Burant JC, Dapprich S, Millan JM, Daniels AD, Kudin KN, Strain MC, Frakas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gozalez C, Challacombe M, Gill PMW, Johnson BG, Chen W, Wong MW, Andres JL, Head-Gordon M, Replogle ES, Pople JA (1998) Gaussian 98, revision A.1. Gaussian, Pittsburgh, Pa
16. Slater JC (1974) Quantum theory of molecules and solids, vol 4. The self-consistent field for molecular and solids. McGraw-Hill, New York
17. Hohenberg P, Kohn W (1964) Phys Rev B 136: 864–871
18. Becke AD (1988) Phys Rev A 38: 3098–3100
19. Perdew JP, Burke K, Wang Y (1996) Phys Rev B 54: 16533–16539
20. Adamo C, Barone V (1997) Chem Phys Lett 274: 242–250
21. Gill PMW (1996) Mol Phys 89: 433–445
22. Vosko SH, Wilk L, Nusair M (1980) Can J Phys 58: 1200–1211
23. Lee C, Yang W, Parr RG (1988) Phys Rev B 37: 785–789
24. Perdew JP, Zunger A (1981) Phys Rev B 23: 5048–5079
25. Perdew JP (1986) Phys Rev B 33: 8822–8824
26. Becke AD (1993) J Chem Phys 98: 1372–1377
27. Becke AD (1993) J Chem Phys 98: 5648–5652
28. Becke AD (1996) J Chem Phys 104: 1040–1046
29. Barone V, Cossi M (1998) J Phys Chem A 102: 1995–2001
30. Barone V, Cossi M, Tomasi J (1997) J Chem Phys 107: 3210–3221
31. Pople JA, Schlegel MB, Krishnan R, Defrees DJ, Bimkley JS, Frish MJ, Whiteride RA, Mont RF, Mehre WJ (1981) Int J Quantum Chem Quantum Chem Symp 15: 269–278
32. Kidd KG, Mantsch HH (1981) J Mol Spectrosc 85: 375–389
33. Thomas JO, Tellgren R, Almlöf J (1975) J Acta Crystallogr B 31: 1946–1955
34. McKean DC, Duncan JL, Batt L (1962) Spectrochim Acta A 29: 1037–1049
35. Newman R (1952) J Chem Phys 20: 1663–1664
36. Spinner E (1967) J Chem Soc B 879–885
37. Bartholomew RJ, Irish DE (1993) Can J Chem 71: 1728–1733
38. Ito K, Bernstein HJ (1956) Can J Chem 34: 170–174
39. Scott AP, Radom L (1996) J Chem Phys 100: 16502–16513, and references therein
40. Magalhães AL, Gomes JANF (1997) Int J Quantum Chem 61: 725–739