

A statistical approach of density functional effects on the vibrational frequencies and infrared intensities of CH₃F

Idelazil C. Nascimento, Anselmo E. de Oliveira, Roy E. Bruns *

Instituto de Química, Universidade Estadual de Campinas, CP 6154, 13081-970, Campinas-SP, Brazil

Received 22 August 1997; received in revised form 4 December 1997; accepted 9 December 1997

Abstract

Fractional factorial design and principal component analyses are applied to CH₃F vibrational harmonic frequencies and infrared intensities from density functional calculations. Five effects for 16 wave functions of a 2⁵⁻¹ fractional factorial design are investigated: the kinds of exchange (S or B) and correlation (VWN or LYP) functional; the valence basis set (6-31G or 6-311G) and the level of diffuse and polarization functions supplementing this basis set. The main effects on the frequencies are related to modifications of both the exchange and correlation functionals. For the intensities main effects are related to the exchange functional and to the level of polarization functions used. Principal component analysis indicates that for wave functions used in the 2⁵⁻¹ factorial design those including the B exchange and VWN correlation functionals are the best ones for estimating the harmonic frequencies of CH₃F. For the infrared intensities a wave function including the B exchange functional and a basis set supplemented by (3df, 3pd) polarization functions provides the most accurate CH₃F results. The calculated frequency results were significantly improved by use of a B3LYP functional with a 6-311 valence, (3df, 3pd) polarization and + + diffuse function basis set. However several calculations with the BLYP functional provided more accurate infrared intensity estimates than those with the B3LYP functional. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Density functional; Vibrational frequencies; Infrared intensities; Factorial design; Principal component analysis

1. Introduction

Quantum chemical applications are often concerned with the calculation of several molecular and spectral properties. Although all ab initio wave functions provide energies satisfying the variational principle optimal ones besides furnish-

ing lower energies also result in accurate quantitative estimates of these properties. The design of optimal wave functions can be characterized as shown in Fig. 1. Several input wave function factors can be modified in such a way so that calculated output properties are in satisfactory agreement with experimental values. Several problems of a multivariate nature arise in this process. How can the effects of various input factor modifications on the molecular properties be effi-

* Corresponding author. Fax.: +55 19 2393805; e-mail: bruns@iqm.unicamp.br

ciently evaluated? Are the results of these modifications independent of one another? How do they vary from one molecule to another? How can the results obtained for the different molecular and spectral properties be efficiently analyzed? How can the wave functions providing the most accurate estimates of these properties be identified?

Recently our laboratory has proposed that multivariate statistical techniques are valuable tools for determining more accurate wave functions [1–3]. Factorial designs [4] (FD) permit the complete evaluation of the effects of wave function modifications on calculated property values while testing a minimum number of molecular wave functions. Possible synergic or antagonistic interaction effects involving several input factors can be determined for each property calculated. Furthermore the effects of all wave function modifications on all calculated properties can be analyzed simultaneously using principal component analysis (PCA) [5–7]. Convenient two-dimensional graphs permit a detailed analysis of how well diverse calculated property values obtained from different molecular wave functions agree with experimental results. Multivariate statistical methods, such as PCA, are capable of treating the statistical correlations involving calculated or experimentally measured properties so that problem dimensionality is reduced and data analysis simplified in comparison with alternative univariate procedures.

In this work factorial design and principal component analysis are applied to density functional calculations of the harmonic vibrational frequencies and infrared intensities of methyl fluoride. Besides having been subject to extensive experimental [8–12] and theoretical investigations [13–15], the CH_3F frequency values have been corrected for anharmonic effects [8] and its intensity values [10] have been used to calculate reference polar tensors for empirical intensity estimates for a variety of fluorine-containing molecules [16–18]. Previous theoretical treatments have employed conventional molecular orbital wave functions and, besides confirming spectroscopic band assignments, have been most useful in determining dipole moment derivative signs. Sosa

and Schlegel [13] have calculated all the fundamental vibrational frequencies and infrared intensities of methyl fluoride using some 30 different wave functions, in an attempt to evaluate the importance of correcting for electron correlation for their accurate calculation, as well as to estimate the effects of including multiple sets of polarization and diffuse functions in the atomic basis set. Quantitative factor effect values for type of basis set, presence or absence of diffuse and polarization functions as well as the level of electron correlation treatment, Hartree–Fock or Møller–Plesset 2, were later determined in our laboratory using a 2^4 factorial design [1] consisting of 16 different wave functions. In addition to important main factor effects, large polarization function–correlation level interaction effects were found for several fundamental frequencies (ν_1 , ν_3 , ν_4 , and ν_6) and for one intensity (A_4).

Recent studies have shown that the density functional method provides accurate estimates of the frequencies [19,20] and polar tensor elements [21,22] of several molecules. Accordingly, it seems appropriate to critically examine the importance of several characteristics of density functional calculations on the fundamental frequencies and intensities of methyl fluoride so that they can be compared with effects determined earlier for modifications of conventional molecular orbital wave functions. This is especially important since density functional theory provides an alternative means of including electron correlation [23]. Here five factors or characteristics of density functional calculations are examined at two levels each: (1) type of exchange functional, Slater (S) or Becke 88 (B); (2) type of correlation functional, Vosko–Wilk–Nussair 1980 (VWN) or Lee–Yang–Parr

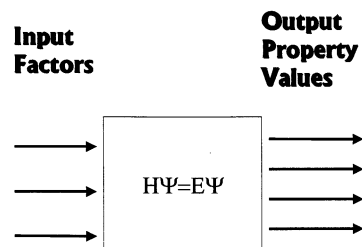


Fig. 1. Multivariate input factors and output property values.

(LYP); (3) basis set; 6-31G or 6-311G; (4) basis set supplemented by polarization functions, one set of d functions on both carbon and fluorine atoms and one set of p functions on the hydrogen atoms (d, p) or three sets of d functions and one set of f functions on both carbon and fluorine atoms and three sets of p functions and one set of d functions on the hydrogen atoms (3df, 3pd); and (5) basis set supplemented by diffuse functions on all atoms or just on the carbon and fluorine ones. A complete factorial design treating these five factors would require $2^5 = 32$ different density functional calculations. Here we show that a fractional factorial design (FFD) consisting of only 16 density functional calculations is capable of furnishing accurate main and two factor interaction effect values of these density functional characteristics. Furthermore the fractional factorial design results for all frequency and intensity values are conveniently analyzed using the principal component transformation. The levels of the factors of the 2^{5-1} fractional factorial design were adjusted following Gaussian 94 [24] options including both the Local Spin Density Approximation (LSDA) and Becke's 1988 functional.

2. Calculations

Density functional calculations of the CH_3F fundamental vibrational frequencies and infrared intensities were performed as prescribed by the 2^{5-1} fractional factorial design in Table 1. These harmonic frequencies were determined by the analytic evaluation of the second derivative of the energy with respect to nuclear displacement and each of the five wave function characteristics (or factors in statistical terminology) were investigated at two levels. The $-$ and $+$ signs indicate whether a factor is in its low or high level. Since the absolute sign definitions are arbitrary (some or all of the above definitions could be reversed) relative signs are important in constructing factorial designs. A full 2^5 factorial design would require the calculation of 32 different wave functions corresponding to all possible combinations of the two levels of the five factors. The design given in Table 1 is a half fraction of the

Table 1

A 2^{5-1} fractional factorial design for the calculation of the CH_3F vibrational frequencies and intensities

Factors	Levels	
	$-$	$+$
1. Exchange functional	S	B
2. Correlation functional	VWN	LYP
3. Basis set function	6-31G	6-311G
4. Polarization function	(d,p)	(3df,3pd)
5. Diffuse function	$+$	$++$
Wave function	Factorial designation	
1. SVWN/6-31 + + G(d, p)	$-----+$	
2. BVWN/6-31 + G(d, p)	$+- ----$	
3. SLYP/6-31 + G(d, p)	$-+ ----$	
4. BLYP/6-31 + + G(d, p)	$++ -- ++$	
5. SVWN/6-311 + G(d, p)	$-- + --$	
6. BVWN/6-311 + + G(d, p)	$+ - + +$	
7. SLYP/6-311 + + G(d, p)	$- + + +$	
8. BLYP/6-311 + G(d, p)	$++ + --$	
9. SVWN/6-31 + G(3df, 3pd)	$--- + -$	
10. BVWN/6-31 + + G(3df, 3pd)	$+ -- + +$	
11. SLYP/6-31 + + G(3df, 3pd)	$- + + +$	
12. BLYP/6-31 + G(3df, 3pd)	$+ + - +$	
13. SVWN/6-311 + + G(3df, 3pd)	$-- + + +$	
14. BVWN/6-311 + G(3df, 3pd)	$+ - + + -$	
15. SLYP/6-311 + G(3df, 3pd)	$- + + + -$	
16. BLYP/6-311 + + G(3df, 3pd)	$++ + + +$	

full factorial design. Only those sign combinations for which their product yields a positive sign ($-----+$, $+-----$, ..., $+++++$) are included in the 2^{5-1} fractional factorial used here. A statistically equivalent 2^{5-1} fractional factorial design for which only combinations whose products yield negative signs could just as well have been chosen. Whereas the full factorial design permits the determination of all interaction effects in an unambiguous way the cost of performing only half of the calculations using the design in Table 1 (or its statistical equivalent) is that principal effect values are confounded with fourth order ones and second order interaction effects are confounded with third order ones. Since full factorial designs previously applied to molecular orbital calculations of the frequencies

and intensities of several molecules, including CH_3F , have shown that third and higher order interaction effects are not important it seems secure to assume that they are also negligible for the CH_3F density functional calculations investigated here.

A main or interaction effect is defined as the change in a calculated parameter (frequency or intensity) as the level is changed from low to high,

$$(\text{Ef}_i) = (R_+)_i - (R_-)_i \quad (1)$$

(Ef_i) is the effect value for the i th factor or combination of factors. $(R_+)_i$ and $(R_-)_i$ are average results of calculated frequencies or intensities at the high (+) and low (−) levels of the factor or combination of factors.

For main effects which describe the change in parameters for isolated changes from the low (−) to high (+) level in only one factor, the signs in Eq. (1) are identical to those in Table 1 for that factor. Interaction effects of two or more factors are calculated using the same equation, except that the (+) and (−) levels are determined by multiplying the signs of the factors involved in the interaction. This simple equation can be used to calculate the main and interaction effects because the fractional factorial design in Table 1 is orthogonal. Effects values can not be calculated using Eq. (1) for other subsets of 16 sign combinations chosen for a half-fraction design, except for the alternative that is statistically equivalent to the one in Table 1, having all of its signs in the last column reversed.

An alternative to calculating individual effect values for each frequency or intensity value in attempts to determine how they are affected by wave function modifications is provided by the multivariate principal component method. Instead of the univariate treatment for which each frequency or intensity parameter is treated separately principal component analysis allows all these parameters to be treated together. In this way their statistical correlations can be studied in a multivariate procedure that permits the identification of wave function modifications that affect the various spectral or molecular parameters. Furthermore the quality of the wave functions used for calculating frequencies and intensities can be

judged more objectively using multivariate methods since an arbitrary choice, made by the investigator, of the most relevant variable to be used in a univariate procedure becomes unnecessary.

Principal components for the frequency and intensity results of the 2^{5-1} fractional design were calculated separately. They are eigenvectors of the covariance or correlation matrix, $\mathbf{X}^t\mathbf{X}$, where \mathbf{X}^t is the transpose of the \mathbf{X} data matrix. The \mathbf{X} matrix has 16 rows, each one corresponding to one of the wave functions in the fractional factorial design in Table 1 and six columns, one for each of the six fundamental frequencies or infrared intensities. The intensity values were centered on their respective means while the frequencies were auto-scaled [1], i.e. after mean centering the frequencies were divided by the standard deviations of each fundamental frequency for the fractional factorial results. As such all transformed frequency values have the same variance and are treated in the multivariate analysis on an equal footing. On the other hand, strong bands are favored in the principal component calculation of the intensity data, since a large experimental intensity is usually accompanied by a large variance for the results of a factorial design.

Bidimensional principal component graphs provide accurate representations of $3N-6$ dimensional frequency or intensity data providing that almost all of the statistical variance is concentrated in the first two principal components [1–3]. In other words, if the 16 points, one for each type of wave function in the factorial design, form an approximately planar array in $3N-6$ dimensional frequency or intensity space (six for our applications on methyl fluoride) a bidimensional principal component graph is capable of accurately describing the geometrical relationships between the points in this array. Using these graphs the $3N-6$ experimental frequency and intensity values can be compared with the theoretical results for the wave functions of the factorial design and wave function modifications providing better overall agreement between theoretical and experimental results can be conveniently determined.

All the harmonic vibrational frequencies and infrared intensities of CH_3F were carried out with the Gaussian 94 program [24], on RISC 6000

Table 2
Calculated and experimental fundamental harmonic vibrational frequencies for CH₃F (cm⁻¹)

Wave function	ν_1	ν_2	ν_3	ν_4	ν_5	ν_6
	A_1 CH ₃	A_1 CH ₃	A_1 CF	E CH ₃	E CH ₃	E CH ₃
	Sym. str.	Sym. bend.	Str.	Antisym. str.	Antisym. def.	Antisym. def
-----+	2969.4	1418.3	1086.4	3066.7	1424.5	1145.1
+-----	3009.7	1464.6	961.7	3098.4	1480.8	1163.6
-+-----	2924.0	1394.1	1094.3	3019.2	1399.3	1129.4
+ + ----+	2966.4	1438.8	972.9	3053.7	1452.9	1146.2
--+----	2957.6	1411.8	1075.0	3048.6	1415.9	1142.0
+ - + - +	2997.0	1459.4	954.3	3079.3	1474.9	1158.0
- + + - +	2910.8	1387.8	1085.4	2999.4	1387.7	1125.1
+ + + - -	2955.2	1435.1	965.4	3036.1	1448.7	1143.6
---+---	2953.8	1404.7	1095.8	3042.6	1411.8	1144.5
+ ---++	2994.0	1459.0	972.2	3074.8	1474.3	1165.4
- + - ++	2912.0	1377.7	1103.7	2999.7	1383.6	1126.7
+ + - + -	2953.4	1432.8	983.3	3032.8	1446.6	1148.8
--+++	2961.1	1412.5	1092.9	3051.0	1417.0	1148.7
+ - + + -	3002.1	1467.1	969.6	3082.4	1482.7	1170.0
- + + + -	2919.7	1385.6	1102.2	3008.5	1389.5	1130.7
+ + + + +	2960.4	1440.7	980.7	3039.6	1453.5	1152.9
B3LYP ^a	3032.3	1484.8	1047.6	3113.8	1493.8	1193.5
Expt ^b	3031.2	1490.2	1059.2	3131.5	1497.8	1206.4

^a B3LYP/6-311++G(3df, 3pd); ^b Harmonic vibrational frequencies as corrected from the observed values in Ref. [9].

IBM workstations at CENAPAD/SP. These calculations were made using the energy-minimum equilibrium geometries for each of the CH₃F wave functions in Table 1.

3. Results

The calculated fundamental vibrational frequencies for the 2⁵⁻¹ fractional factorial design of methyl fluoride are presented in Table 2. Experimental frequencies, corrected for anharmonicity [8] are also included in this table. These values can be conveniently analyzed on a graph having the first two principal components as axes. The first principal component is given in Table 3 as a function of the auto-scaled frequencies and accounts for 88% of the total data variance. All loading coefficients have similar absolute values indicating that all frequencies are important for explaining variations in the values of Table 2. However all these loadings are positive except for

ν_3 . This implies that all the frequency values are correlated with each other except for ν_3 , the CF stretching frequency, which is anticorrelated with the other five frequencies. In other words the wave function modifications, on the whole, affect

Table 3
Principal component equations for the calculated frequencies and intensities of CH₃F^a

	Explained variance (%)
Frequencies	
PC ₁ = 0.42 ν_1 + 0.43 ν_2 - 0.37 ν_3 + 0.40 ν_4 + 0.42 ν_5 + 0.41 ν_6	88
PC ₂ = 0.32 ν_1 - 0.16 ν_2 + 0.71 ν_3 + 0.52 ν_4 - 0.21 ν_5 + 0.19 ν_6	10
Intensities	
PC ₁ = 0.12 A_1 + 0.95 A_4 - 0.26 A_5	87
PC ₂ = -0.92 A_3 - 0.11 A_4 - 0.33 A_5	11

^a The frequency values, ν_p , are auto-scaled whereas the A_i intensity values are mean-centered.

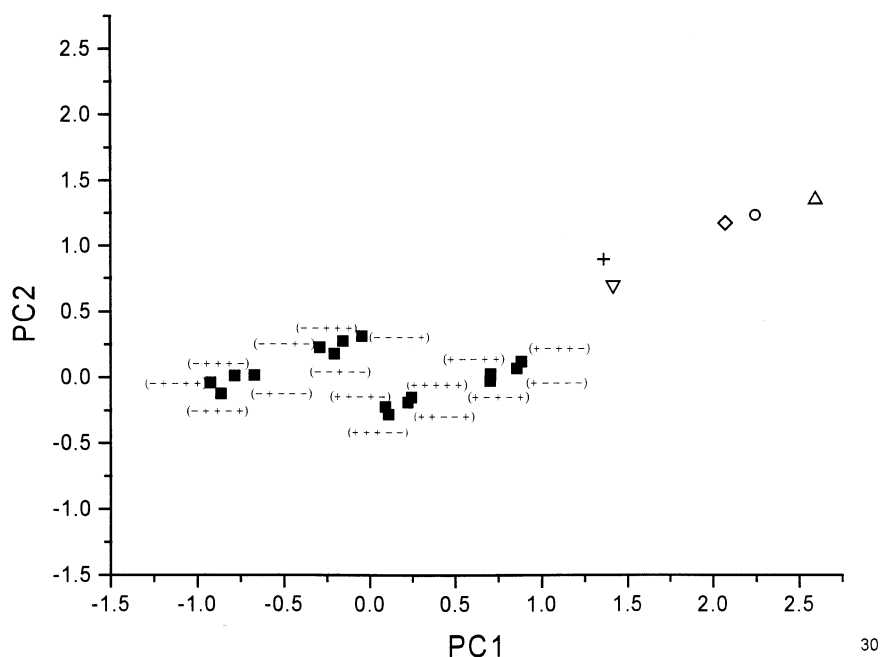


Fig. 2. Principal component plot of calculated and experimental (harmonic) frequencies of CH_3F (cm^{-1}). The CISD/3-21G and MP2/6-311G(d, p) frequencies were taken from Ref. [13] and the MP2/6-311 + G(d, p) ones from Ref. [1]. ■ factorial design; ○ CISD/3-21G; Δ MP2/6-311G(d, p); ▽ exp.; ◇ MP2/6-311 + G(d, p); + B3LYP/6-311 + G(3df, 3pd).

all the CH_3 frequency values in similar ways whereas they have opposite effects on the CF stretching frequency as can be verified in Table 2. The second principal component explains much less variance than the first one and has important positive contributions from the three stretching frequencies, especially ν_3 , the CF stretching frequency.

The bidimensional graph of these two principal components explains 98% of the total variance and is shown in Fig. 2. Since all but 2% of the total variance is explained in this graph it can be considered to be an accurate representation of the six dimensional frequency data. Note that points with mostly positive first principal component scores were calculated using the B exchange functional whereas those corresponding to the S functional have negative score values. Furthermore for points obtained with equivalent exchange functionals those corresponding to the VWN correlation functional tend to have more positive first principal component scores than those calculated using the LYP functional. Considering that all the

frequencies have positive loadings except for the CF stretching one, it can be concluded that changing the S exchange functional for a B one, in the presence of either the VWN or LYP correlation functional, increases all the frequency values except for the CF stretching one which is lowered. Furthermore changing the VWN functional by the LYP one, increases the CF stretching frequency while lowering all the others—for results with a fixed exchange functional. These deductions can be verified upon inspection of the frequency values in Table 2.

The CF stretching frequency dominates the second principal component. Since it has a positive loading score points for results obtained from using an S exchange potential are somewhat more positive than those obtained using a B one. This is consistent with the conclusion drawn about the CF frequency values from the scores of the first PC.

The calculated infrared intensities for the 2^{5-1} fractional factorial design and the experimentally measured intensities are presented in Table 4.

Table 4
Calculated and experimental fundamental infrared intensities for CH₃F (km mol⁻¹)

Wave function	A_1	A_2	A_3	A_4	A_5	A_6
	A_1 CH ₃	A_1 CH ₃	A_1 CF	E CH ₃	E CH ₃	E CH ₃
	Sym. str.	Sym. bend.	Str.	Antisym. str.	Antisym. def.	Antisym. def.
-----+	34.2	0.6	111.02	46.4	14.9	0.0
+-----	37.2	4.3	109.3	74.1	7.0	1.1
-+-----	33.5	0.3	111.8	41.3	18.1	0.0
+ + - - -	36.6	3.0	109.7	69.4	8.1	0.7
-- + - -	31.7	0.2	113.5	44.6	18.0	0.0
+ - + - -	35.5	3.0	110.2	76.8	8.1	1.2
- + + - -	32.0	0.1	115.0	42.5	18.0	0.0
+ + + - -	34.8	1.9	110.2	70.9	10.4	0.8
--- - + -	31.0	0.0	105.6	42.2	15.3	0.4
+ - - + +	35.1	1.3	104.5	71.2	6.9	2.1
- + - + +	30.5	0.0	106.3	41.2	15.3	0.2
+ + - + -	33.7	0.8	104.6	65.2	8.7	1.9
- - + + +	32.0	0.0	104.0	37.8	15.3	0.3
+ - + + -	35.0	1.2	101.9	62.4	8.2	1.9
- + + + -	30.8	0.0	104.4	33.1	17.3	0.1
+ + + + +	34.8	0.9	102.2	60.1	8.5	1.5
B3LYP ^a	33.6	0.9	110.0	57.0	10.2	2.1
Expt ^b	24.7	0.9	95.3	61.0	8.7	2.7

^a B3LYP/6-311 + + G(3df, 3pd); ^b Experimental intensity values of Ref. [11].

These intensity values were not auto-scaled since it seems reasonable that the larger intensity values should have larger weights in the multivariate analysis than those for weak bands. Large measured intensities correspond to large variances in the calculated intensities for a factorial design as seen before [1–3]. The loadings for the first two principal components are given in Table 3 and Fig. 3 shows a clear separation for results obtained using S exchange functional and those corresponding to the B functional. The first principal component explains 87% of the total data variance and has a large positive loading for A_4 . Since the wave functions with the B correlation functional all result in positive first PC scores whereas those with the S one have negative scores it is straightforward to deduce that the A_4 intensities calculated with the B functional are larger than those obtained with the S functional. Inspection of values in Table 4 confirms this. The B functional results in A_4 intensity values 20–30 km mol⁻¹ higher than those calculated with the S functional.

The second principal component explains only 11% of the total variance of the intensity data. Its most important loading is negative and pertains to A_3 . Fig. 3 shows that calculations using (3df, 3pd) polarization functions have larger second PC scores than those obtained with (d, p) functions. This is consistent with the results in Table 4. Calculations using the larger set of polarization functions (last eight rows in Table 4) result in A_3 intensity values 5–10 km mol⁻¹ smaller than those using the simpler polarization set.

The conclusions from the multivariate PC analysis can also be obtained using factorial design models for which main and interaction effect values are calculated using Eq. (1). Table 5 contains the effect values on the fundamental frequencies. Modifications in the exchange and correlation functionals clearly produce the largest changes in the calculated frequencies. The exchange functional effect is the largest one. Substituting the S functional for the B one significantly increases all

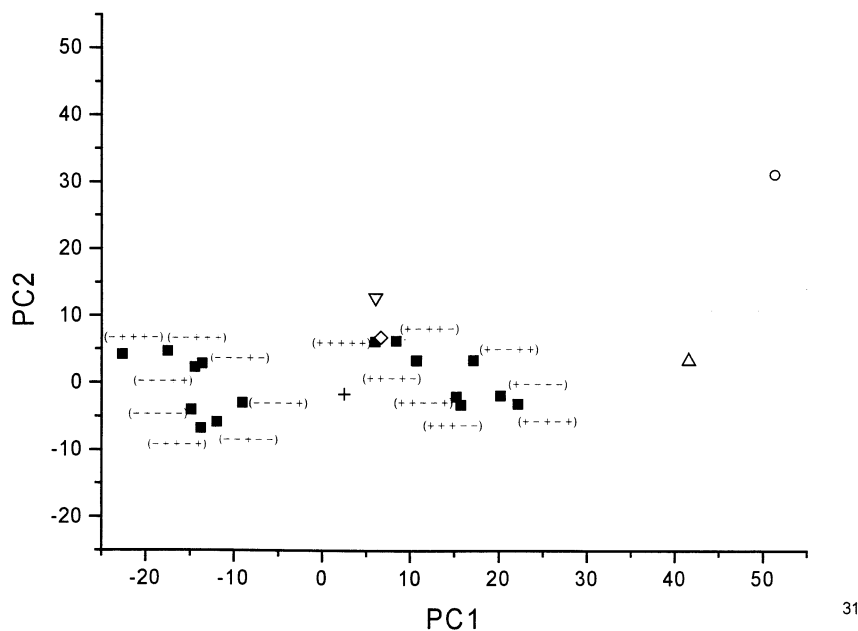


Fig. 3. Principal component plot of calculated and experimental intensities of CH_3F (km mol^{-1}). The CISD/3-21G and MP2/6-311G(d, p) intensities were taken from Ref. [13] and those of the MP2/6-311 + + G(d, p) wave function from Ref. [1]. ■ factorial design; ○ CISD/3-21G; Δ MP2/6-311G(d,p); ▽ exp.; ◇ MP2/6-311 + + G(d, p); + B3LYP/6-311 + + G(3df, 3pd).

the fundamental frequencies by $20\text{--}50\text{ cm}^{-1}$ whereas it decreases the CF stretching frequency by around 120 cm^{-1} . This is consistent with the opposite sign calculated for the loading of the CF stretching frequency relative to those calculated for the other frequencies for the first PC of the frequency data. Substituting the VWN correlation functional by the LYP functional results in a lowering of all the fundamental frequencies by $17\text{--}44\text{ cm}^{-1}$ except for the CF stretching frequency, which increases by 10 cm^{-1} . This is also consistent with the higher first PC scores for points obtained with the VWN correlation functional than for those obtained with the LYP one. Other wave function modifications do not have much effect on the calculated frequencies except for the basis set-polarization function interaction. Simultaneous changes in the basis set from 6-31G to 6-311G and in the polarization functions from (d, p) to (3df, 3pd) increases the CH_3 stretching and bending frequencies by about 10 cm^{-1} more than expected from models considering only main effect values. These second order interaction ef-

fects are substantially smaller than the main effect values for the exchange and correlation functional factors.

Table 6 contains the main and interaction effect values on the methyl fluoride intensities. The largest effect values are found on changes of the S exchange functional for the B one. This modification increases the CH_3 antisymmetric stretching intensity by 27.6 km mol^{-1} . All the calculations with the S functional underestimate this intensity by 15 km mol^{-1} or more. Indeed inspection of the calculated intensity values in Table 6 shows that use of the B exchange functional results in more accurate intensity values for all vibrational modes except for the CH_3 symmetric stretching mode. This is clearly seen in the principal component graph in Fig. 3. The experimental point is closer to the group of theoretical points obtained using the B functional than those originating from calculations with the S functional. The use of appropriate polarization functions is also important for accurate intensity calculations even within the density functional procedure as the values in

Table 5
Main and interaction effects of wave function modifications on the fundamental CH₃F frequencies^a

Effects	ν_1	ν_2	ν_3	ν_4	ν_5	ν_6
	CH ₃ sym. str.	CH ₃ sym. bend.	CF str.	CH ₃ antisym str.	CH ₃ antisym bend	CH ₃ rock
Exchange functional	41.2	50.6	−122.0	32.7	60.6	19.5
Correlation functional	−42.9	−25.6	10.0	−44.4	−27.5	−16.7
Basis set	—	—	−5.6	−5.4	—	—
Basis set/polarization function interaction	9.9	6.7	—	13.3	7.1	—

^a Effects with absolute values less than 5.0 cm^{−1} are not included in this table.

Table 6 indicate. The inclusion of more polarization functions, (d, p)→(3df, 3pd) lowers the CF stretching intensity by about 7 km mol^{−1}. As mentioned earlier in this section the last eight rows in Table 4 have A_3 values 5–10 km mol^{−1} smaller than those of the first eight rows.

4. Discussion

Previous factorial design studies using ab initio molecular orbital wave functions showed that similar effect values for frequencies are obtained for similar types of molecular vibrations. However this is not true for the intensities. The CH₃ symmetric and antisymmetric stretching frequencies, ν_1 and ν_4 , have −38.4 and −41.6 cm^{−1} effect values on changing from 6-31G to a 6-311G basis set and −120.4 and −102.4 cm^{−1} effect values on changing from the HF to the MP2 level. Even the second order correlation level/polarization function interaction effect values are about the same for these two vibrations, 34.2 and 38.0 cm^{−1}. On the other hand effect values for the intensities of these vibrational modes are very different. For example, the inclusion of diffuse orbitals increases the A_1 intensity by 4 km mol^{−1} whereas it causes a 24.2 km mol^{−1} reduction in the A_4 intensity.

This same type of behavior holds for the density functional calculations as well. Changing the exchange functional from S to B increases the ν_1 and ν_4 CH₃ stretching frequencies by 41.2 and 32.7 cm^{−1}. In contrast this change increases the

A_1 intensity by 3.4 km mol^{−1} but results in a much larger 27.6 km mol^{−1} increase in A_4 . The VWN to LYP change decreases ν_1 and ν_4 by 42.9 and 44.4 cm^{−1}, respectively, whereas it hardly affects the A_1 intensity value but decreases the A_4 intensity by 4 km mol^{−1}. As such systematic changes in calculated frequency values upon wave function modifications are more commonly observed than for intensity values, not only for ab initio wave functions but for density functional ones as well.

In general the effects for the density functional calculations are smaller than those observed earlier by us for molecular orbital wave functions. The electron correlation effects for the ab initio wave functions, corresponding to a change from the Hartree–Fock to the Møller–Plesset 2 level, have values between −89 and −120 cm^{−1} for all the CH₃F frequencies. These values are much larger in absolute magnitude than the ones obtained here upon changing from a VWN to a LYP functional. The inclusion of polarization orbitals in the basis set of the ab initio wave function resulted in a 115 cm^{−1} increase in the CF stretching frequency. Here changing the polarization function level from (d, p) to (3df, 3pd) does not have any main effect on the frequencies with absolute value above 5 cm^{−1}. Also second order interaction effects are much smaller (7–13 cm^{−1}) for our density functional calculation than for the ab initio wave functions (12–38 cm^{−1}).

The intensity effect values for the density functional calculations are also much smaller than those for the ab initio wave function. Only one

Table 6

Main and interaction effects of wave function modifications on the fundamental CH₃F intensities^a

Effects	A_1	A_2	A_3	A_4	A_5	A_6
	CH ₃ sym. str.	CH ₃ sym. bend	CF str.	CH ₃ antisym str.	CH ₃ antisym bend	CH ₃ rock
Exchange functional	3.4	1.9	−2.4	27.6	−8.3	1.3
Correlation functional	—	—	—	−4.0	1.3	—
Basis set	—	—	—	−2.9	1.2	—
Polarization functions	−1.6	−1.2	−7.2	−6.6	—	—
Diffuse functions	—	—	—	1.5	−1.0	—
Basis set/polarization function interaction	1.2	—	−1.9	—	—	—

^a Effects with absolute values less than 1 km mol^{−1} are not included in this table.

effect value is larger than 20 km mol^{−1} for the density functional calculations whereas six effects have higher absolute values for the ab initio wave function calculations.

Of the density functional results the (+ − + −) and (+ − − −) ones are in best agreement with the experimental anharmonicity corrected values for the frequencies with root mean square errors of 19.3 and 19.6 cm^{−1}, respectively. All the quantum chemical values reported here under-estimate these observed values. The ab initio wave functions in our previous study [1] using a 2⁴ full factorial design providing frequency values in best agreement with the observed ones were the MP2/6-311 + + G(d, p) and MP2/6-311G(d, p) with 16.5 and 17.7 cm^{−1} errors. Although these errors are comparable to those for the density functional results their calculated frequencies over-estimate the observed harmonic values. This is clearly seen in the principal component graph in Fig. 2. Whereas the density functional PC score points are lower and to the left of the experimental score values the MP2/6-311 + + G(d, p) point obtained from the results of Ref. [1] is above and to the right. For comparison points corresponding to the ab initio CISD/3-21G and MP2/6-311G(d, p) results of Ref. [13] are included in Fig. 2. They are also located to the extreme right and are well above the experimental point.

The two density functional calculations with results in closest agreement with the measured

intensities, (+ + + + +) and (+ − + + −), have smaller root mean square errors (5.0 km mol^{−1} each) than the ab initio wave function of the 2⁴ factorial design reported earlier to be in best agreement, the HF/6-31G wave function with a 6.9 km mol^{−1}. Furthermore the principal component projection for this latter design showed that the MP2/6-311 + + G(d, p) wave function was in second best agreement. In other words, the simplest and most complex wave function of the 2⁴ factorial design for the ab initio calculations resulted in the most accurate intensity estimates. The two density functional calculations with best results are more similar to each other than the ab initio ones, only differing in their correlation functional and in the quantity of diffuse functions used. This might indicate that the density functional 2^{5−1} results are converging to more accurate values whereas this does not appear to be true for the 2⁴ ab initio ones of our previous study.

In an attempt to obtain more accurate theoretical results for both the CH₃F frequencies and intensities a B3LYP functional calculation was performed. Since the most accurate frequencies and intensities obtained with the BLYP functional used 6-311G valence, (3df, 3pd) polarization and (+ +) diffuse functions a B3LYP/6-311 + + G(3df, 3pd) calculation was performed. The results are included in Tables 2 and 4 and in Figs. 2 and 3.

The B3LYP/6-311++G(3df,3pd) frequency results are in excellent agreement with the anharmonic-corrected values (within 0.04–1.10% depending on the fundamental frequency). Its corresponding score point in Fig. 2 is much closer to the experimental score point than those for any other set of theoretical frequencies. These results certainly suggest that density functional calculations can provide very accurate vibrational frequency results compared with conventional *ab initio* results.

However the B3LYP/6-311++G(3df,3pd) results listed in the Table 4 are not in as good agreement with the experimental intensity results as are the (++++) or BLYP/6-311++G(3df,3pd) results or the (+--+ -) or BVWN/6-311+G(3df,3pd) results. The B3LYP results have a root mean square (rms) error of 7.23 km mol⁻¹, significantly larger than the 5.0 km mol⁻¹ rms errors for the calculations indicated above. These conclusions are easily confirmed by inspection of Fig. 3 where the latter calculations are represented by points much closer to the experimental point than the one for B3LYP/6-311++G(3df,3pd).

In summary, based on these results for CH₃F, the density functional method seems to provide an attractive alternative to *ab initio* calculations for calculating vibrational frequencies and infrared intensities. However as results for strictly *ab initio* calculations have shown previously [1–3], the frequency results appear to be converging to accurate values as the wave function factors are systematically changed whereas this does not appear to be true for the corresponding intensity values.

Acknowledgements

The authors thank FAPESP (95-3580-8) and CNPq for partial financial support. ICN thanks FAPESP and AEO CNPq for graduate student fellowships. We would like to thank a referee for suggesting the B3LYP calculations.

References

- [1] A.L.M.S. de Azevedo, B. de Barros Neto, I.S. Scarminio, A.E. Oliveira, R.E. Bruns, *J. Comput. Chem.* 17 (1996) 167.
- [2] J.B.P. da Silva, M.N. Ramos, R.E. Bruns, *Spectrochim. Acta Part A* 53A (1997) 733.
- [3] (a) R.E. Bruns, P.H. Guadagnini, I.S. Scarminio, B. de Barros Neto, *J. Mol. Struct. (THEOCHEM)* 397 (1997) 197; (b) A.E. Oliveira, R.E. Bruns, *J. Mol. Struct. (THEOCHEM)* 397 (1997) 187.
- [4] G.E.P. Box, W.S. Hunter, J.S. Hunter, *Statistics for Experimenters*, Wiley, New York, 1978.
- [5] K.V. Mardia, J.T. Kent, J.M. Bibby, *Multivariate Analysis*, Ch. 8, Academic Press, London, 1979, p. 213.
- [6] E. Suto, M.M.C. Ferreira, R.E. Bruns, *J. Comp. Chem.* 12 (1991) 885.
- [7] F.P.S.C. Gil, A.M. Amorim da Costa, R.E. Bruns, J.J.C. Teixeira-Dias, *J. Phys. Chem.* 99 (1995) 634.
- [8] G.M. Barrow, D.C. McKean, *Proc. R. Soc. London* 213A (1952) 27.
- [9] J.W. Russell, C.D. Needham, J. Overend, *J. Chem. Phys.* 45 (1966) 3383.
- [10] S. Kondo, S. Saeki, *J. Chem. Phys.* 76 (1982) 809.
- [11] C. Dilauro, I.M. Mills, *J. Mol. Spectrosc.* 21 (1966) 386.
- [12] J. Duncan, D.C. McKean, G.K. Spiers, *Mol. Phys.* 24 (1972) 553.
- [13] S. Sosa, H.B. Schlegel, *J. Chem. Phys.* 86 (1987) 6937.
- [14] E. Suto, H.P. Martins Filho, R.E. Bruns, *J. Mol. Struct. (THEOCHEM)* 282 (1993) 81.
- [15] C.F. Blom, A. Muller, *J. Mol. Spectrosc.* 70 (1978) 449.
- [16] W.B. Person, Prediction of infrared intensities by transfer of atomic polar tensors, in: W.B. Person and G. Zerbi (Eds.), *Vibrational Intensities in Infrared and Raman Spectroscopy*, Ch. 14, Elsevier, Amsterdam, 1982, p. 269.
- [17] J.H. Newton, W.B. Person, *J. Chem. Phys.* 64 (1976) 3036.
- [18] W.B. Person, J. Overend, *J. Chem. Phys.* 66 (1977) 1442.
- [19] B.G. Johnson, P.M. Gill, J.A. Pople, *J. Chem. Phys.* 98 (1993) 5612.
- [20] C.H. Choi, M. Kertesz, *J. Phys. Chem.* 100 (1996) 16530.
- [21] D.P. Chong, A.V. Bree, *Chem. Phys. Lett.* 210 (1993) 443.
- [22] D.P. Chong, D. Papousek, *Chem. Phys. Lett.* 193 (1992) 399.
- [23] D.K.W. Hok, R. Neumann, N.C. Handy, *J. Phys. Chem.* 100 (1996) 6225.
- [24] Gaussian 94, Revision D.2, M.J. Frisch, G.W. Trucks, H.B. Schlegel, P.M.W. Gill, B.G. Johnson, M.A. Robb, J.R. Cheeseman, T. Keith, G.A. Petersson, J.A. Montgomery, K. Raghavachari, M.A. Al-Laham, V.G. Zakrzewski, J.V. Orates, J.B. Foresman, J. Cioslowski, B.B. Stefanov, A. Nanayakkara, M. Challacombe, C.Y. Peng, P.Y. Ayala, W. Chen, M.W. Wong, J.L. Andres, E.S. Replogle, R. Gomperts, R.L. Martin, D.J. Fox, J.S. Binkley, D.J. Defrees, J. Baker, J.P. Stewart, M. Head-Gordon, C. Gonzalez, J.A. Pople, Gaussian, Pittsburgh PA, 1995.