

Article

Application of Factorial Design for the *ab initio* Study of *cis*- and *trans*-1,2 Dihaloehtylene Stabilities

João Bosco P. da Silva*

Departamento de Química Fundamental, Universidade Federal de Pernambuco, 50740-250, Recife - PE, Brazil

A técnica quimiométrica Planejamento Fatorial de Dois Níveis foi usada para avaliar os efeitos principais e de interação de modificações da função de onda no cálculo da diferença de energia (ΔE) entre os isômeros *cis* e *trans* 1,2-dihaloetileno, $C_2H_2X_2$ ($X=F$ e Cl). O perfil fatorial é bastante diferente para esses dois sistemas. Os maiores efeitos no conjunto de base para a descrição do *cis*-1,2-difluoroetileno como o isômero mais estável são funções de polarização e funções difusas. Em contraste, correlação eletrônica, a nível MP2, e a inclusão de funções de polarização no conjunto de base são os maiores efeitos descrevendo o *cis*-1,2-dicloroetileno como o isômero mais estável. Em ambos os sistemas, os efeitos de interação são menores que os efeitos principais.

The chemometric technique Two Level Factorial Design was used to evaluate the principal and interaction effects of wave function modifications on calculated energy differences (ΔE) between the *cis* and *trans* isomers of 1,2-dihaloethylenes, $C_2H_2X_2$ ($X=F$ and Cl). The factorial profile of ΔE is quite different for these systems. The addition of polarization and diffuse functions to the basis set have the largest effects in describing *cis* 1,2-difluoroethylene as the most stable isomer. In contrast electronic correlation, at the MP2 level, and the inclusion of polarization functions in the basis set have the largest effects in describing the *cis* form as the most stable one in the 1,2-dichloroethylenes. In both systems the interaction effects are smaller than the principal effects.

Keywords: factorial design, *ab initio*, isomers stability, 1,2-dihaloethylenes

Introduction

When one compares rotational or configurational isomers, the question which naturally arises is: what is the most stable conformer? However even for simple systems, predictions based on electrostatic and steric considerations can go against normal chemical intuition. A classical example is the greater stability of the *cis* form of the 1,2-dihaloethylenes, $C_2H_2X_2$ ($X=F$ and Cl), relative to their *trans* forms^{1,2}. In fact, since the seventies theoretical papers have appeared in the literature trying to explain these experimental results³⁻¹⁰. *Ab initio* calculations using a broad spectrum of DZ (including diffuse and polarization functions) and TZ basis sets were employed at both RHF and MP2 levels in order to evaluate the energy differences between the *cis* and *trans* forms of 1,2-difluoroethylene⁴⁻¹⁰. For these calculations both optimized and experimental geometries have been used. With so many different computational results, a detailed analysis has not been reported in the literature. To our knowledge, there is no report in the literature offering a global vision such as provided by de-

tailed statistical analysis of the importance of the relevant factors (basis sets, level of calculation and molecular geometries) on the calculated relative stabilities of the dihaloethylenes.

Objectively, to study how the factors influence a response, it is crucial to know which of these factors are larger, how they contribute to the response investigated, and furthermore, whether they interact with one another, i.e., whether the effects of one factor depend on the others. In such situations, experimental design (in this case the design of theoretical experiments) and multivariate analysis of the results, represent fundamental steps throughout the investigation. A minimum amount of calculation is desirable but significant results must be obtained¹¹.

In order to analyze all these aspects, *ab initio* calculations of the energy difference between the *cis* and *trans* forms of 1,2 difluoroethylene were performed using factorial design. Both experimental and optimized geometries were used. The multivariate exploratory technique, Two Level Factorial Design (TLFD)¹²⁻¹⁷ was applied to identify and quantify the most important effects. Finally, the analysis was extended to the energy difference between *cis* and *trans* forms of 1,2-dichloroethylene in order to verify if the conclusions obtained for difluoroethylene

e-mail: paraiso@npd.ufpe.br

also hold for its chlorine analogue. Except for a single previous RHF calculation developed by Schäfer *et al.*¹⁸ using 4-21G and 5-31G** basis sets (which points out the *trans* form as the most stable form), to our knowledge this is the first systematic *ab initio* study using several basis sets and different levels of calculation for dichloroethylenes.

Calculations and computational methods

Using the TLFD technique one initially defines two levels for each factor that might influence the response under study. After that, it is necessary to perform experiments for all 2ⁿ possible combinations of n factors at each of the two levels. For our purpose four factors and two levels are chosen: (1) the use of 6-31G or 6-311G basis sets, (2) the presence or absence of diffuse functions, (3) the presence or absence of polarization functions and (4) the use or not of Moller-Plesset perturbative corrections to second order (MP2) over Restricted Hartree-Fock (RHF) calculations. *Ab initio* MO calculations were carried out on all 16 possible combinations of factors for both experimental and completely optimized geometries. For the latter, no imaginary frequencies were observed in the calculations. The MP2 calculations were performed using the frozen-core option that excludes the inner-shells from the correlation energy calculations. The TLFD analysis was performed using the FATORIAL program¹³, while for *ab initio* calculations the GAUSSIAN 92 program¹⁹ was used. The MO calculations were executed on a Risc 6000 IBM workstation and the TLFD ones on a personal microcomputer.

Results and Discussions

The first three numerical columns of Table 1 show the total energy difference (ΔE) between the *cis* and *trans* forms

of 1,2-C₂H₂F₂ for all 16 calculations, whereas the last three columns are related to ΔE values between the *cis* and *trans* forms of 1,2-C₂H₂Cl₂. For each molecular system, the first and second columns of numerical data show ΔE values with and without the zero point energy correction, respectively. On the other hand, in the last column for each molecule ΔE values appear corresponding to the experimental geometries. For the difluoroethylenes the geometry reported by Laurie and Pence²⁰ for the *cis* structure was used, whereas for the *trans* structure the original electron diffraction geometry given by Carlos *et al.* was used²¹. For the dichloroethylenes the experimental geometries of Schäfer *et al.* were used¹⁸.

In Table 2 the codified design matrix corresponding to the data of Table 1 is presented. The notation is as follows: Val indicates whether the 6-31G or 6-311G basis set was used; Dif and Pol whether diffuse or polarization functions were used and Corr whether the calculations were performed at the RHF or MP2 level.

In order to evaluate the principal and interaction effects, the ΔE values in Table 1 are applied to the signs of the corresponding effect column in Table 2; then the average value of the high level (+) minus the average value from the low level (-) is taken. For example, the principal effect owing to the use of polarization functions is given by the difference between the average value of the eight energy differences calculated with polarization functions and the average value of the eight energy differences calculated without these functions. Note that the signs for the interaction effects are obtained from the product of the signs of the principal effect as illustrated in Table 2.

C₂H₂F₂

In Figure 1 the principal and interaction effects obtained from the TLFD analysis for the energy differences of the 1,2-difluoroethylenes are shown.

Table 1. Energy difference, in kcal mol⁻¹, for the *cis* e *trans* forms of 1,2-C₂H₂X₂ (X = F and Cl), 1 cal = 4.18 J.

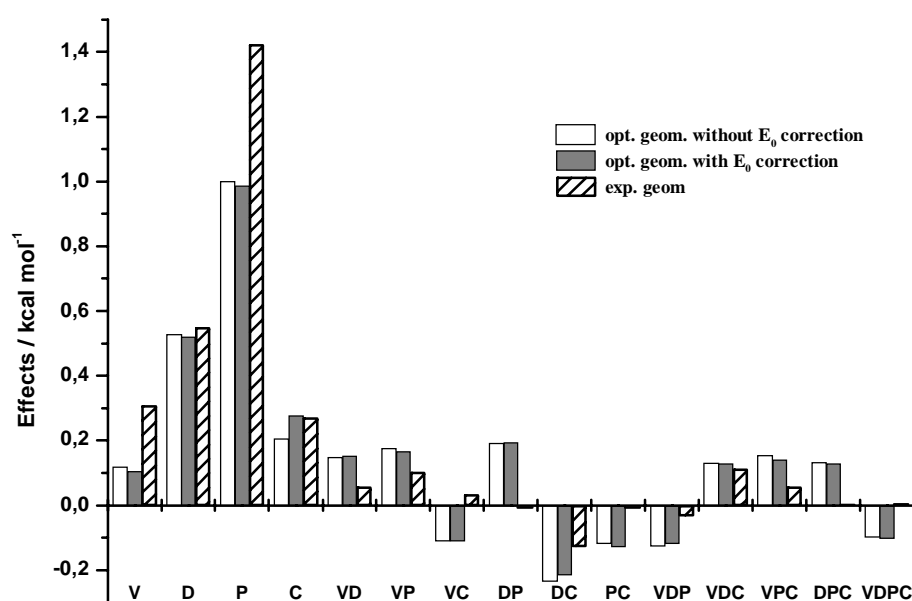
Wave Function	C ₂ H ₂ F ₂			C ₂ H ₂ Cl ₂		
	OPT Geom.	EXP Geom.	EXP Geom.	OPT Geom.	EXP Geom.	EXP Geom.
	ΔE^a	ΔE_0^b	ΔE^c	ΔE^a	ΔE_0^b	ΔE^c
6-31G	-1.27	-1.52	-1.05	-0.63	-0.89	-0.97
6-311G	-1.11	-1.37	-0.90	-0.71	-0.95	-1.08
6-31++G	-0.61	-0.89	-0.44	-0.58	-0.86	-0.85
6-311++G	-0.36	-0.66	-0.13	-0.66	-0.93	-1.03
6-31G**	-0.26	-0.51	0.32	-0.23	-0.46	-0.61
6-311G**	0.00	-0.28	0.63	-0.27	-0.49	-0.63
6-31++G**	0.57	0.28	0.97	-0.09	-0.33	-0.44
6-311++G**	0.81	0.53	1.30	-0.22	-0.44	-0.56
MP2/6-31G	-0.09	-0.29	-0.61	-0.06	-0.30	-0.13
MP2/6-311G	-0.91	-1.10	-0.52	-0.17	-0.40	-0.38
MP2/6-31++G	-0.62	-0.81	-0.27	-0.11	-0.33	-0.14
MP2/6-311++G	-0.44	-0.62	0.01	-0.14	-0.32	-0.33
MP2/6-31G**	-0.08	-0.28	0.63	0.53	0.31	0.38
MP2/6-311G**	0.20	-0.04	1.09	0.44	0.64	0.32
MP2/6-31++G**	0.48	0.29	1.00	0.54	0.40	0.42
MP2/6-311++G**	0.87	0.65	1.54	0.48	0.38	0.39
Exp.		1.08 ^e			0.48 ^f	

^aCalculated as $\Delta E = E_{trans} - E_{cis}$; ^bCalculated as $\Delta E_0 = E_{trans} + E_{0,trans} - (E_{cis} + E_{0,cis})$; ^cCis-C₂H₂F₂, Ref. 20 and Trans-C₂H₂F₂, Ref. 21; ^eRef. 18.; ^fRef. 1.; ^gRef. 2.

Table 2. Codified designing matrix for 2⁴ TLF analysis ^a.

Wave function	Main				Effects					Interaction					
	Val	Dif	Pol	Corr	Val-Dif	Val-Pol	Val-Corr	Dif-Pol	Dif-Corr	Pol-Corr	Val-Dif-Pol	Val-Dif-Corr	Val-Pol-Corr	Dif-Pol-Corr	Val-Dif-Pol-Corr
6-31G	-	-	-	-	+	+	+	+	+	+	-	-	-	-	+
6-311G	+	-	-	-	-	-	-	+	+	+	+	+	+	+	-
6-31++G	-	+	-	-	-	+	+	-	-	+	+	+	-	+	-
6-311++G	+	+	-	-	+	-	-	-	-	+	-	-	+	+	+
6-31G**	-	-	+	-	+	-	+	-	+	-	+	-	+	+	-
6-311G**	+	-	+	-	-	+	-	+	+	-	-	+	-	+	+
6-31++G**	-	+	+	-	-	-	+	+	-	-	-	+	+	-	+
6-311++G**	+	+	+	-	+	+	-	+	-	-	+	-	-	-	-
MP2/6-31G	-	-	-	+	+	+	-	+	-	-	-	+	+	+	-
MP2/6-311G	+	-	-	+	-	-	+	+	+	-	+	-	-	+	+
MP2/6-31++G	-	+	-	+	-	+	-	-	+	-	+	-	+	-	+
MP2/6-311++G	+	+	-	+	+	-	+	-	+	-	-	+	-	-	-
MP2/6-31G**	-	-	+	+	+	-	-	-	-	+	+	+	-	-	+
MP2/6-311G**	+	-	+	+	-	+	+	-	-	+	-	-	+	-	-
MP2/6-31++G**	-	+	+	+	-	-	-	+	+	+	-	-	-	+	-
MP2/6-311++G**	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

^aVal: valence description (- → 6-31G, + → 6-311G); Dif: diffuse functions (- → absent, + → present); Pol: polarization function (- → absent, + → present); Corr: electron correlation (- → RHF calculations, + → MP2 calculations).

**Figure 1.** Principal and interaction effects on the wave function modifications for the energy difference calculations between the *cis* and *trans* forms of 1,2-difluoroethylene (1 cal = 4.18 J).

In this figure one may observe that only two effects are really very important, i.e., the diffuse and polarization function principal effects, where the latter is about twice as large. For both effects their interpretations are straightforward. The inclusion of diffuse or polarization functions tends to increase positively the calculated energy difference between the *trans* and *cis* forms of 1,2-difluoroethylene. For example, although ΔE is -1.52 kcal mol⁻¹ using the 6-31G basis set (therefore, pointing out the *trans* form as the most stable conformer) its value

changes to -0.89 kcal mol⁻¹ when diffuse functions are used in the 6-31++G basis set. A similar trend occurs for the use of polarization functions and thus the ΔE values are -1.52 and -0.51 kcal mol⁻¹ for the 6-311G and 6-311G** basis sets, respectively. Since, in general, it is not a simple task to use a complete enough wave function to achieve the true experimental potential energy surface, one may assume that a ΔE value, such as in Table 1, for a given wave function, is only meaningful with the optimized structure obtained with the same wave function.

Therefore, the effect of using the experimental geometry with *ab initio* wave functions might artificially overestimate the principal Pol effect. Another interesting result is the small effect of electronic correlation (at MP2 level), for both optimized and experimental geometries. This conclusion is not in agreement with Cremer's results⁵ which point to electronic correlation as an important effect. Cremer's results correctly suggest the *cis* form to be more stable since polarized basis sets produce the larger effect in stabilizing the *cis* form. It is important to stress that the factorial design allows a more complete understanding of these apparently contradictory results. In the work of Gandhi *et al.*⁶, all wave functions describing the *cis* form as having the lower energy also contain polarization functions. They found that the use of diffuse functions also leads to a stabilization of the *cis* form (ΔE (DZ+P)=0.86 and ΔE (DZ++P)=1.19 kcal mol⁻¹). This is consistent with the positive and larger diffuse function effect observed in Figure 1. It is interesting to point out, that similarly for our results Gandhi *et al.* also did not observe large correlation effects for the energy difference using experimental or optimized geometries.

The interaction effects, on the whole, are lower than the principal effects. However, their values are 10-24% of the larger principal Pol effect. When one uses the experimental geometry these interaction effects are even lower. In Table 1 one can observe that the zero point energy correction leads to a decrease in the relative stability of the *cis* form by 0.22 kcal mol⁻¹. This being a systematic difference, explains

the similar profiles in Figure 1 for the effects calculated with or without these corrections.

C₂H₂Cl₂

The behavior of the principal and interaction effects obtained from the TLF_D for the energy difference of the *cis* and *trans* forms of 1,2-dichloroethylene is quite different from the one for the fluorine analogue as can be seen below.

Figure 2 shows that the most important effects are the polarization and correlation ones. Now, the correlation effect is the larger one. While the Dif effect tends to increase ΔE in favor of the *cis* form for the difluoroethylenes it has a non-significant effect in the dichloroethylenes. Except for the Pol-Corr interaction effect which is 21% of the larger Corr effect, the other interaction effects are very small, only 1-9% of the Corr effect value. In fact it is necessary to take into account the Pol-Corr interaction effect in order to understand the larger stabilization of *cis*-dichloroethylene when one compares the use of polarization functions in MP2 and RHF calculation as can be confirmed in Table 1. For example the ΔE values for the 6-31G and 6-31G** basis sets are -0.89 and -0.46 kcal mol⁻¹, respectively, a variation of +0.43 kcal mol⁻¹. On the other hand for the MP2/6-31G and MP2/6-31G** calculations the values are -0.30 and 0.31 kcal mol⁻¹, the variation being +0.61 kcal mol⁻¹.

As for 1,2-difluoroethylene, the use of the experimental geometry tends to overestimate the larger main effects and also to decrease the interaction effects. However for 1,2-dichloroethylene these effects are more pronounced

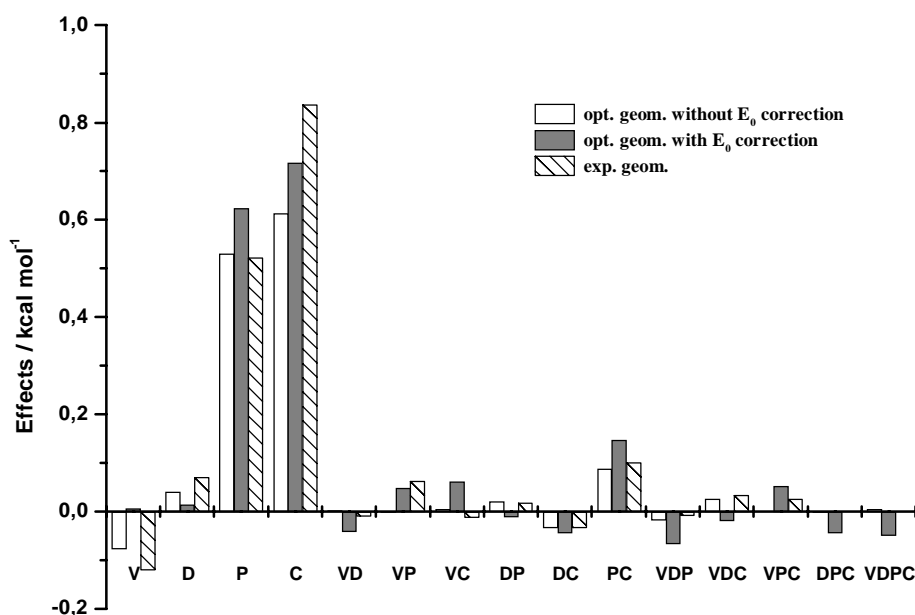


Figure 2. Principal and interaction effects on the wave function modifications for the energy difference calculations between the *cis* and *trans* forms of 1,2-dichloroethylene (1 cal = 4.18 J).

than in the 1,2-difluoroethylene case. The zero point energy contribution for decreasing the *cis* stability is, on the average, the same as the one for 1,2-difluoroethylene system, i.e., 0.22 kcal mol⁻¹.

It is interesting to point out that a similar TLFDD performed for the energy difference between the *cis* and *trans* forms of methyl nitrite (CH₃ONO) showed a different profile compared to the effect values presented here. However, in all three cases the introduction of polarization functions was observed to be fundamental for the correct description of the *cis* conformer as the most stable form ²².

Conclusions

Use of the Two Level Factorial Design exploratory technique shows how one can obtain a complete visualization of the individual and interaction effects of wave function modifications on the *cis* and *trans* 1,2-difluoro and -dichloroethylene relative stabilities. Although the principal effects can be visualized directly from Table 1, the interaction effects can only be quantified using such approaches. For the intervals defined by the low and high levels of each factor defined in this work, one may conclude that: (i) many results previously pointed out in the literature may be understood and confirmed with our TLFDD analysis, (ii) the most important effects for the 1,2-difluoroethylene energy difference are the polarization and diffuse principal effects whereas for 1,2-dichloroethylene the correlation and polarization principal effects are predominant, (iii) in both systems the interaction effects are lower than the main effects, (iv) the zero point energy correction tends to decrease the relative stability of the *cis* forms by 0.22 kcal mol⁻¹ for both systems, (v) the use of experimental geometry tends to overestimate the larger effects and to underestimate the lower interaction effects and (vi) the simple change from fluorine to chlorine in these configurational isomers leads to different requirements on the wave function for an adequate description of the most stable form. This is a good example of the risk of extending conclusions from one system to another even when they are similar, such as here, where atoms in different rows of the periodic table are involved.

Acknowledgments

I would like to thank Professors J. Demnitz, A. Simas, A. Gama (UFPE) and R. Bruns (Unicamp) for valuable comments and suggestions on the final form of this text and the Brazilian agency CNPq agency for a graduate student fellowship (1993 - 1998). The final form of this text also benefited from the referees' comments, which are gratefully acknowledged.

References

1. (a) Craig, N. C.; Entemann, E. A. *J. Am. Chem. Soc.* **1961**, *83*, 3047. (b) Craig, N. C.; Overend, J. *J. Chem. Phys.* **1967**, *51*, 1127.
2. (a) Ambartzmuian, R. V.; Chekalin, N. V.; Doljikov, V. S.; Letokhov, V. S.; Lokhman, V. N. *Optics. Com.* **1976**, *18*, 400. (b) Craig, N. C.; Piper, L. G.; Wheeler, V. L. *J. Phys. Chem.* **1971**, *75*, 1453.
3. Pitzer, K. S.; Hollenberg, J. L. *J. Am. Chem. Soc.* **1954**, *76*, 1493.
4. Binkley, J. S.; Pople, J. A. *Chem. Phys. Lett.* **1977**, *45*, 197.
5. Cremer, D. *Chem. Phys. Lett.* **1981**, *81*, 481.
6. Gandhi, S. R.; Benzel, M. A.; Dykstra, C. E.; Fukunaga, T. *J. Phys. Chem.* **1982**, *86*, 3121.
7. Dixon, D. A.; Smart, B. E.; Fukunaga, T. *Chem. Phys. Lett.* **1986**, *125*, 447.
8. Saebo, S.; Sellers, H. *J. Phys. Chem.* **1988**, *92*, 4266.
9. Muir, M.; Baker, J. *Mol. Phys.* **1996**, *89*, 211.
10. Yamamoto, T.; Tomoda, S. *Chem. Lett.* **1997**, *10*, 1069.
11. Wold, S. *Chemom. Intell. Lab. Syst.* **1995**, *30*, 109.
12. Box, G. E. P.; Hunter, W. G.; Hunter, J. S. *Statistics for Experimenters*; Ed. Wiley, New York, 1978.
13. Neto, B. B.; Scarminio, I. S.; Bruns, R. E. *Planejamento e Otimização de Experimentos*; Editora da Unicamp, Campinas, 1995.
14. de Azevedo, A. L. M. S.; Neto, B. B.; Scarminio, I. S.; de Oliveira, A. E.; Bruns, R. E. *J. Comput. Chem.* **1996**, *17*, 167.
15. Bruns, R. E.; Guadagnini, P. H.; Scarminio, I. S.; Neto, B. B. *Theochem* **1997**, *394*, 197.
16. da Silva, J. B. P.; Ramos, M. N.; Bruns, R. E. *Spectrochimica Acta, Part A* **1997**, *53*, 733.
17. Ramos, M. N.; da Silva, J. B. P.; Bruns, R. E. *Spectrochimica Acta, Part A* **1997**, *53*, 1563.
18. Schäfer, L.; Ewbank, J. D.; Siam, K.; Paul, D. W.; Monts, D. L. *J. Mol. Struct.* **1986**, *145*, 135.
19. Gaussian 92 (Revision C) Frisch, M. J.; Binkley, J. S.; Schlegel, H. B.; Raghavachari, K.; Melius, C. F.; Martin, R. L.; Stewart, J. J. P.; Bobrowicz, F. W.; Rohlfing, C. M.; Kahn, L. R.; Defrees, D. J.; Seeger, R.; Whiteside, R. A.; Fox, D. J.; Fleuder, E. M.; Pople, J. A. *Gaussian, Inc.*, Pittsburg, PA, 1992.
20. Laurie, V. W.; Pence, D. T. *J. Chem. Phys.* **1963**, *38*, 2693.
21. Carlos, J.L.; Karl, R. R.; Bauer, S. H. *J. Chem. Soc., Faraday Trans. 2* **1974**, *2*, 177.
22. da Silva, J. B. P.; da Costa Jr., N. B.; Ramos, M. N.; Fausto, R. *J. Mol. Struct.* **1996**, *375*, 153.

Received: July 12, 1999