Scientific paper

QTAIM Densities and *Ab Initio* Basis Sets: A Chemometrical Analysis of the Intermolecular Electronic Densities of the Hydrogen-Bonded Complexes $C_2H_4O\cdots HX$ (X = F, CN, NC, and CCH)

Boaz Galdino de Oliveira,[1],* Regiane de Cássia Maritan Ugulino de Araújo,[1]
Antônio Bezerra de Carvalho[1] and Mozart Neves Ramos[2]

[1] *Departamento de Ciências Farmacêuticas, Universidade Federal de Pernambuco*
50740-521, Recife, Pernambuco – Brazil

[2] *Departamento de Química Fundamental, Universidade Federal de Pernambuco*
50739-901, Recife, Pernambuco – Brazil

** Corresponding author: E-mail: boazgaldino@gmail.com*

Received: 22-10-2008

Abstract

The intermolecular electronic densities of the hydrogen-bonded complexes $C_2H_4O\cdots HX$ (X = F, CN, NC, and CCH) were determined by calculations performed at the level of Density Functional Theory (DFT). The B3LYP functional and the Pople's 6-31G split-valence basis set were combined with (11) valence, (++) diffuse and (d,p) polarization functions with the purpose to determine the optimized geometries of the $C_2H_4O\cdots HX$ complexes. Through these geometries, their intermolecular electronic densities $\rho(r,x)^{QTAIM}$ calculated from Quantum Theory of Atoms in Molecules (QTAIM) were examined through the Hierarchical Cluster Analysis (HCA) and Two-Level Factorial Designs (TLFD) statistical techniques. Although it is well-known that the QTAIM approach was projected with a slight dependence from *ab initio* basis sets, our results revealed a good relationship between $\bar{\rho}(r,x)^{QTAIM}$ and the values estimated by the chemometric analysis ($\rho(r,x)^{EST}$) computed essentially by valence (11) and diffuse (++) functions.

Keywords: QTAIM; basis sets; B3LYP; hydrogen complexes

1. Introduction

1939 was considered a remarkable year for the chemistry with the publication of "Nature of the Chemical Bond" of Linus Pauling.[1] Several concepts about chemical bonds are reported in this book, whose reading has aided theoreticians and experimentalists to interpret many questions in chemistry and physics, such as the interaction energy measurements and related phenomena.[2–4] According to Pauling's concepts, it was established that if the distance between two interactive atoms is shorter or at least coequal than sum of the van der Waals radii, we can affirm that these atoms are chemically bound one with another.[5–6] Particularly, it is important to stress that some intermolecular systems are directly involved with the van der Waals radii, for instance the hydrogen-bonded complexes.[7] As it is widely known that hydrogen-bonded complexes are formed when proton donors ($H^{+\delta}$) interact with negative electronic centers (Y), which can be lone electron pairs (O, N, S and F atoms) or π and pseudo-π bonds,[8] it is evident that the hydrogen bond distance ($H^{+\delta}\cdots Y$) must also be shorter or at least coequal than the van der Waals radii of the Y and $H^{+\delta}$ species.[9] However, if this intermolecular criterion is not obeyed, the hydrogen-bonded complex is not formed because the orbital overlapping between HOMO (Y, π and pseudo-π bonds) and LUMO (proton donors) is hindered. On the other hand, some decades after Pauling has divulged his ideas about chemical bonds, Pople *et al.*[10] have started the development of *ab initio* basis sets with the purpose to describe the molecular orbital through implementation of mathematical formulations into software packages.[11] Nowadays, albeit many computational schemes are used routinely by many theoretical groups, many researchers continue the development of more sophisticated *ab initio* basis sets.[12–14]

Historically, in midst of many computational algorithms developed by Pople and others,[15] the 6-31G split-valence is a useful basis sets[16] because its algorithm is composed by 6 contracted Gaussian-type functions which describe the inner shell orbital, whereas the valence shell is represented by (1) valence functions. Moreover, 6-31G can be combined with other wave functions, such as the (++) diffuses and (d,p) polarizations,[17–18] by which the complete basis set 6-311++G(d,p) is composed. In terms of hydrogen-bonded complexes, it is well established that this basis set yields a good description of polarizability and intermolecular strength, although the intermolecular interaction is often well described through the application of (11) valence functions because the chemical bond is formed by outer charge density shells. Nevertheless, it is established that (++) diffuse functions are essential to describe the overlap orbital, and finally, the (d,p) polarization functions extends the angular momentum of the electronic density cloud by addition of "d" and "p" orbitals on "heavy" and hydrogen atoms, respectively. So, surely that the application of a complete basis set such as 6-311++G(d,p) is recommended if the main goal is the characterization of the electronic structure in a minimum of the potential energy surface. However, some decades ago Richard Bader also developed another useful theoretical tool to study the electronic structure and chemical bonds through the quantum-mechanical theorems. This method was named Quantum Theory of Atoms in Molecules (QTAIM)[19] and in its formulation Bader admitted that the fragmentation of a molecular system furnishes conditions for modeling depletion and concentration of charge density in specific chemical sites.[20] Therefore, it becomes established that a high value of charge density suggests typical features for covalent and unsaturated bonds, whereas low values of charge density denotes an accumulation of electronic density localized on disconnected nuclei, then characterizing the existence of intermolecular and intramolecular contacts, *e.g.* hydrogen bonds.[21] Undoubtedly, we should say that the van der Waals radii are important for QTAIM theory because the bonded and non-bonded atomic radii are taken into account to determine Bond Critical Points (BCP)[22] between acceptor/donor of protons.[23–25] By definition, BCP is a source used to calculate topological parameters such as electronic density and Laplacian field, which give us an idea how is the atomic behavior into a chemical bond.[26] In according with the gradient of the electronic density $\nabla\rho(r,x)$, at any $n(\vec{r})$ point with x coordinates on molecular surface, the QTAIM protocol rationalizes the molecular structure into sub-systems ruled by the zero-flux of the charge density (see equation 1).

$$\nabla\rho(r,x).\ n(\vec{r}) = 0 \qquad \forall\, \vec{r} \in \partial\Omega \qquad (1)$$

So, it is known that the electronic density is the quantum mechanical observable adopted by the QTAIM, although it is admitted that QTAIM results depends slightly from *ab initio* basis sets, *i.e.* there is no criteria to affirm which basis set is ideal to perform QTAIM calculations. Nevertheless, some studies have been published where small basis sets have been used successfully for determination of optimized geometries.[27–30] In practice, the coordinates of these optimized geometries were taken as entrance data ("in puts") and the QTAIM calculations were developed at this same theoretical level, where no restriction to small basis sets was accused, *i.e.* the 6-31G(d), 6-31+G(d,p) or 6-31G(d,p).[31–32] Consequently, it is natural that in these cases the applicability of the QTAIM theory confronts with the concepts of basis sets proposed by Pople and co-workers[15,18] by which it is admitted that an accurate description of the chemical bond is directly related to a complete basis set. With the purpose to solve this confrontation of theories, two questions must be answered:

i) Is there dependence of the QTAIM electronic density from *ab initio* basis sets?

ii) Or else, this dependence really affects the description of the molecular system?

To answer these two questions, it would be interesting to evaluate the behavior of the intermolecular electronic density derived from QTAIM theory when several theoretical levels are used. As such, firstly we have elected the $C_2H_4O\cdots HX$ hydrogen-bonded complexes with X = F, CN, NC, and CCH as our investigation systems. After that, we have admitted the importance of a chemometric analysis of the QTAIM intermolecular electronic density by considering the combination of the (11) valence, (++) diffuse and (d,p) polarization functions into the 6-31G split-valence basis set. To execute this chemometric analysis, two techniques were applied: Hierarchical Cluster Analysis (HCA)[33] and the Two-Level Factorial Design (TLFD).[34]

2. Computational Methods

The calculations to determine the optimized geometries of the $C_2H_4O\cdots HX$ hydrogen-bonded complexes with X = F, CN, NC or CCH were carried out in GAUSSIAN 98W software.[35] Density Functional Theory (DFT)[36–37] was selected as our standard level of calculation because this method takes into account electronic correlation effects, which are essential in studies of hydrogen complexes.[38–43] Moreover, it is known that DFT is efficient not only to study molecular parameters of hydrogen complexes, but it performs calculations at a lower computational cost in comparison with other approaches, *e.g.* MP2 perturbation theory.[44] In terms of Bader's molecular topology, the QTAIM calculations were performed in GAUSSIAN 98W program.

3. Chemometrical Procedures

The results of the HCA and TLFD calculations were obtained through the application of the STATISTICA 5.0[45] and MATLAB 7.0.1[46] programs, respectively.

4. Results

4. 1. Optimized Geometries of the $C_2H_4O{\cdots}HX$ Hydrogen-bonded Complexes

Recently, the optimized geometries of the hydrogen-bonded complexes $C_2H_4O{\cdots}HF$, $C_2H_4O{\cdots}HCN$, $C_2H_4O{\cdots}HNC$, and $C_2H_4O{\cdots}C_2H_2$ were examined in detail by our research group.[47–48] At the B3LYP/6-311++G(d,p) level of theory, the optimized geometries of these complexes are presented in Fig. 1.

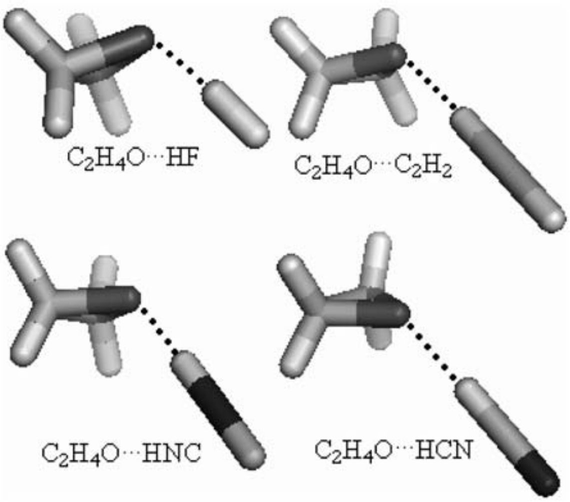

Figure 1. Optimized geometries of the $C_2H_4O{\cdots}HF$, $C_2H_4O{\cdots}C_2H_2$, $C_2H_4O{\cdots}HNC$ and $C_2H_4O{\cdots}HCN$ using B3LYP/6-311++G(d,p) calculations

Note that the complete basis set formed by (11) valence, (++) diffuse and (d,p) polarization functions was used. Besides computation of the attraction energies (both BSSE and ZPE corrections were included)[49–50] and characterization of vibrational red-shifts on proton donors,[51] structurally these hydrogen-bonded complexes were fully examined in terms of their (O···H) hydrogen bond distances.[1] However, by using the small basis set 6-31G(d,p), other theoretical study about the structures of heterocyclic hydrogen-bonded complexes was also documented[52] and, only to mention, we have obtained unusual results for the (O···H) hydrogen bond distances. In other words, we have concluded that the theoretical characterization of hydrogen bonds in heterocyclic hydrogen-bonded complexes seems to be decisively affected when different basis sets are used. Even though the geometries of the $C_2H_4O{\cdots}HF$ and $C_2H_4O{\cdots}C_2H_2$ hydrogen-bonded complexes have been determined experimentally through the Fourier-Transform Microwave Spectroscopy (FTMS),[53] we believe that a characterization of the (O···HX) hydrogen bonds must still be investigated with more accuracy. Besides DFT and/or *ab initio* calculations, theoretically we have assumed that the (O···HX) hydrogen bond and its properties can be studied on the basis of the QTAIM approach.

4. 2. QTAIM Density: Characterization of the (O···HX) Hydrogen Bonds

The identification of hydrogen bonds by means of the QTAIM theory is ruled by a *virial* theorem, wherein high and low centers of charge density are identified in agreement with the kinetic ($K(\vec{r},x)$) and potential ($U(\vec{r},x)$) operators, respectively. The predominance of $K(\vec{r},x)$ indicates a low concentration of electronic density in (O···HX) hydrogen bonds, *i.e.* the interpretation of $K(\vec{r},x)$ is related to the positive Laplacian field, which in this situation says that the electronic density is concentrated on separated nuclei.

Table 1. Values of the Laplacians for the (O···HX) hydrogen bonds of the $C_2H_4O{\cdots}HX$ hydrogen-bonded complexes calculated from QTAIM calculations.

Theoretical levels	$C_2H_4O{\cdots}HX$ hydrogen-bonded complexes			
	HF	**HNC**	**HCN**	**HCCH**
B3LYP/6-31G	0.179	0.144	0.096	0.070
B3LYP/6-311G	0.164	0.141	0.103	0.077
B3LYP/6-31++G	0.202	0.150	0.107	0.074
B3LYP/6-311++G	0.172	0.142	0.104	0.076
B3LYP/6-31G(d,p)	0.138	0.111	0.067	0.052
B3LYP/6-311G(d,p)	0.146	0.123	0.082	0.062
B3LYP/6-31++G(d,p)	0.148	0.112	0.063	0.048
B3LYP/6-311++G(d,p)	0.149	0.119	0.078	0.055

* All values are given in electronic units (e.u.).

In this insight, the intermolecular interaction is called of closed-shell interaction and in some cases, simply as hydrogen bonds. It is by the QTAIM postulates and in

Table 2. Values of the electronic densities for the (O···HX) hydrogen bonds of the $C_2H_4O{\cdots}HX$ hydrogen-bonded complexes calculated from QTAIM calculations.

Theoretical levels	$C_2H_4O{\cdots}HX$ hydrogen-bonded complexes			
	HF	**HNC**	**HCN**	**HCCH**
B3LYP/6-31G	0.0554	0.0443	0.0299	0.0226
B3LYP/6-311G	0.0573	0.0420	0.0271	0.0210
B3LYP/6-31++G	0.0619	0.0424	0.0270	0.0208
B3LYP/6-311++G	0.0597	0.0400	0.0259	0.0192
B3LYP/6-31G(d,p)	0.0457	0.0392	0.0258	0.0196
B3LYP/6-311G(d,p)	0.0465	0.0378	0.0228	0.0178
B3LYP/6-31++G(d,p)	0.0502	0.0379	0.0218	0.0173
B3LYP/6-311++G(d,p)	0.0467	0.0352	0.0206	0.0152

* All values are given in electronic units (e.u.).

agreement with the positive $\nabla^2\rho(r,x)$ values and low $\rho(r,x)$ amounts (van der Waals complexes have smaller $\rho(r,x)$ values, ~10^{-3} e/a_0^3)[8] presented in Tables 1 and 2 respectively, we would like to say that all $C_2H_4O\cdots HX$ systems can be considered as typical hydrogen-bonded complexes.[54]

4. 3. HCA: Clusters of the Electronic Densities of the (O···HX) Hydrogen Bonds

The study of molecular similarity based on electronic density concepts is a problem that must be treated carefully. It was by this insight that Mezey[55–56] has divulged some works in which the electronic density has been used as a parameter for predicting molecular similarity. Nevertheless, we do not have desire to propose a molecular similarity among the $C_2H_4O\cdots HX$ complexes, in fact we are focusing our study in a similarity on 6-31G split-valence by taking into account the values of the intermolecular electronic densities calculated using the QTAIM protocol. In a first step, to adjust the HCA technique at the study of intermolecular properties, $\rho(r,x)^{QTAIM}$ in this current work, we have consulted a report by Popelier and Aicken[57] where a study of molecular similarity among amino acids and their derivatives was proposed. By this context, the Fig. 2 illustrates the dendrogram of the $\rho(r,x)^{QTAIM}$ values (see Table 2) of the (O···HX) hydrogen bonds of the $C_2H_4O\cdots HX$ complexes.

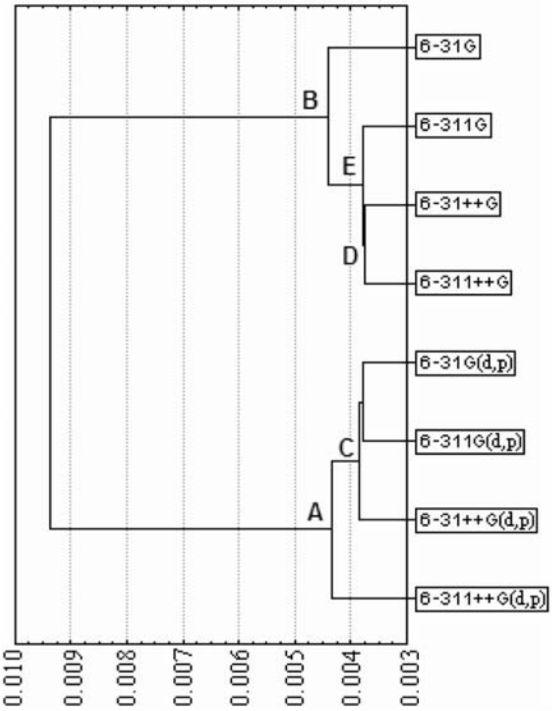

Figure 2. HCA dendogram of the electronic densities of the (O···H) hydrogen bonds of the $C_2H_4O\cdots HX$ hydrogen-bonded complexes.

From HCA protocol, the classification of the basis sets was obtained through a Euclidian relationship, which produces an S_{ab} similarity as follows:

$$S_{ab} = 1 - \frac{d_{ab}}{d_{max}} \tag{2}$$

According to the equation (2), we can understand S_{ab} molecular similarity as the distance (d_{ab}) between two samples "*a*" and "*b*" related to the maximum distance (d_{max}) for any other samples. Analyzing the cluster graph depicted in the Fig. 2, we can observe that the A and B clusters were formed at a S_{ab} distance of ~0.0042. The cluster A is characterized by (d,p) polarization functions, although it is also formed a sub-cluster C which describes other basis sets, such as 6-311G(d,p) and 6-31++G(d,p). In spite of this, based on the results published by our research group,[58] we should mention that a recent chemometric analysis revealed the great importance of (d,p) polarization functions for calculating binding energies, intermolecular distances and new vibrational modes.[59] In this work, we can perceive that (d,p) polarizations functions are also important, mainly for describing intermolecular electronic density.[60] On the other hand, the cluster B and sub-clusters D and E are formed basically by (11) valence and (++) diffuse functions, but it is very important to emphasize that D and E characterize (++) and (11) at a quite similar S_{ab} distance to that computed for (d,p) in sub-cluster C.

4. 4. TLFD: Effect of the *Ab Initio* Basis Sets on QTAIM Electronic Density

For the Two-Level Factorial Designs (TLFD), the modifications on Pople's split-valence 6-3*ij*G*k*[15–16] are treated at two levels by combining the (11) valence = *i*, (++) diffuse = *j* and (d,p) polarization = *k* functions as follows:

i) *i* = Using triple-zeta instead of (11) valence functions;

ii) *j* = Include or not (++) diffuse functions;

iii) *k* = Include or not (d,p) polarization functions.

All these factors were used to generate an algebraic model which describes all main effects on the 6-3*ij*G*k* basis sets for the calculations of the intermolecular electronic densities of the (O···HX) hydrogen bonds of the $C_2H_4O\cdots HX$ hydrogen-bonded complexes. In an overview, this model can be symbolized according to equation (3)

$$\rho(r,x)^{EST} = \rho(r,x)^{QTAIM} + \sum_{m=1}^{n} Y\rho(r,x), X^{Ef} \tag{3}$$

where $\rho(r,x)^{EST}$ and $\rho(r,x)^{QTAIM}$ correspond to values of the electronic densities of the (O···HX) hydrogen bonds estimated by the chemometric model (EST) and calculated by QTAIM, as well as X^{Ef} representing the main

Table 3. Main effects obtained from 2^3 TLFD for the modifications on B3LYP/6-3*ij*G*k* calculations for description of the electronic densities of the (O···H) hydrogen bonds of the $C_2H_4O\cdots HX$ hydrogen-bonded complexes using QTAIM approach.

Effects	**$C_2H_4O\cdots HX$ hydrogen-bonded complexes**				
	HF	**HNC**	**HCN**	**HCCH**	**$\bar{\rho}(r, x)^{AIM}$**
(11)	–	–0.0011	–0.0010	–	–0.00053
(++)	0.0017	–0.0010	–0.0013	–0.0011	–0.00043
[(11)-(++)]	–0.0057	–0.0023	–0.0024	–0.0017	–0.00300
(d,p)	–0.0010	–	–	–	–
[(11)-(d,p)]	–	–	–	–	–
[(11)-(++)–(d,p)]	–	–	–	–	–

TLFD effects of the values obtained from B3LYP/6-3*ij*G*k* calculations. In this equation it is stipulated that the Y coefficient is equal to +1 and –1 when the corresponding factor is present or absent, respectively. By considering the three factors ascribed as (11) valence, (++) diffuse and (d,p) polarization functions, we have applied the TLFD in 2^3 order, whose final results are gathered in Table 3.

We can observe that (d,p) polarization functions are not the most important effects, this being a great surprise because this function is important for description of electronic properties of strained and polarized systems, such as the C_2H_4O molecule[29–30] studied in this work. By the contrary, even though (d,p) seems to be very important to quantify other similar electronic properties, *e.g.* binding energies,[61] the greatest effects obtained in this work are related to (11) and (++) functions, and mainly by the combination between them, *i.e.* 6-311++G. From these two main effects, the TLFD algebraic model was built, which is mathematically represented by the equation (4) and graphically illustrated in Fig. 3. The relationship between $\bar{\rho}(r,x)^{QTAIM}$ and $\rho(r,x)^{EST}$ expressed by equation (4) yielded a R^2 linear coefficient of 0.99.

$$\rho(r,x)^{EST} = \bar{\rho}(r,x)^{QTAIM} - 0.00053(11) - \\ - 0.00043(++) - 0.003[(11)-(++)] \tag{4}$$

Nevertheless, it is well established in the specialized literature that several works related to the QTAIM theory were performed by using functions with higher angular momentum, *i.e.* (d,p) single[62] and (2d,2p)[63] double, or multiple polarization functions such as (3d,3p)[64-65] or (3df,2pd).[66] As it was not necessary to include "d" and "p" polarization functions in our TLFD model (see equation 4), it is very interesting to observe that the intermolecular electronic densities $\rho(r,x)^{QTAIM}$ of the $C_2H_4O\cdots HX$ complexes are described only by (11) valence and (++) diffuse functions.

However, it is widely known that the determination of hydrogen bond energy is one of the most important investigation points in intermolecular systems.[67] According to Duijneveldt and Murrell,[3] as well as Morokuma and Umeyama,[4] the quantification of the hydrogen bond energy (δE_{A-B}) can be physically measured as follows:

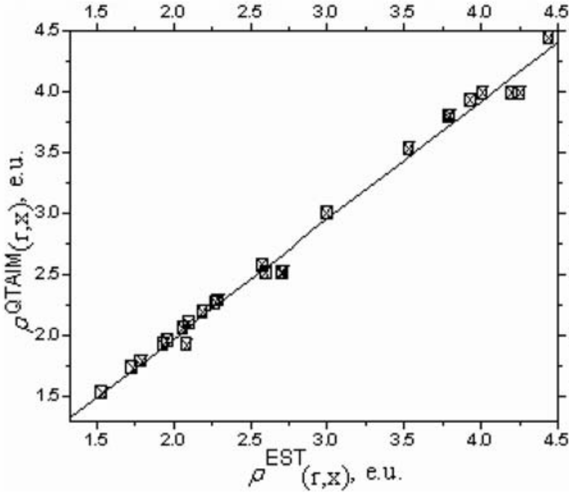

Figure 3. Relationship between $\rho(r,x)^{QTAIM}$ and $\rho(r,x)^{EST}$ values.

$$\Delta E_{A-B} = E_{A-B} - (E_A + E_B) \tag{5}$$

Known as supermolecule approach, this is an indispensable argument in studies of hydrogen-bonded complexes.[68–69] However, there is some kind of system wherein this energy difference cannot be applied, *i.e.* intramolecular interactions. To solve this problem, some time ago Wojtulewski and Grabowski[70] documented an interesting study where the intramolecular hydrogen bond energy was successfully predicted without taking into account the supermolecule approach.[71] So, it has been demonstrated that the hydrogen bond energy can be estimated through the analysis of QTAIM topological parameters,[72] where a relationship between hydrogen bond energies ΔE and electronic density amounts $\rho(r,x)^{QTAIM}$ was obtained. In this work, however, our chemometric analysis has shown that the (++) diffuse functions are the most important effects for calculating the electronic densities of the hydrogen bonds of the $C_2H_4O\cdots HX$ hydrogen-bonded complexes. As such, it is expected that a complete basis set such as 6-311++G(d,p) could yield a good prediction model for the hydrogen bond energies. By analyzing the graph illustrated in the Fig. 4 and taking into account the R^2 linear coefficient of 0.98, a good relationship between the hydrogen bond energy and $\rho(\mathrm{r,x})^{\mathrm{QTAIM}}$ is demonstrated. In this sense a good relationship between hydrogen bond energies and their intermolecular electronic densities was obtained just by applying 6-31++G functions.

$$\Delta E^{C} = 1011.1\ \rho(\mathrm{r,x})^{\mathrm{QTAIM}} - 4.54\ ,\quad R^2 = 0.98 \qquad (6)$$

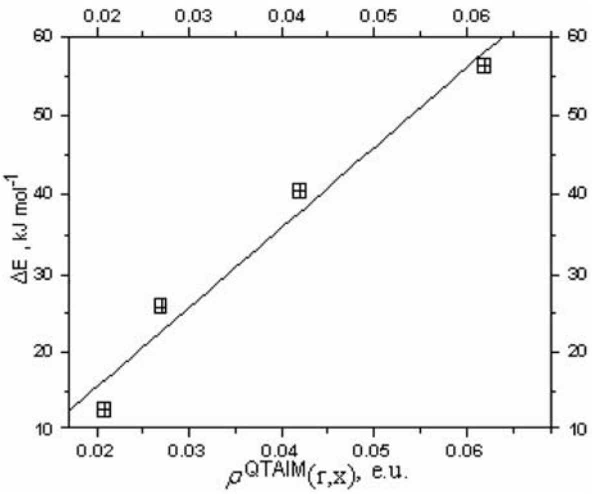

Figure 4. Relationship between $\rho(\mathrm{r,x})^{\mathrm{AIM}}$ values and hydrogen bond energies of the $C_2H_4O\cdots HX$ complexes using the B3LYP/6-31++G calculations.

As polarization functions (d,p) have provided the lowermost contributions to the TLFD model, even so we have tested the 6-31(d,p) and 6-311++(d,p) calculations. Unfortunately, 6-31(d,p) and 6-311++(d,p) basis sets have not provided efficient relationships between ΔE and $\rho(\mathrm{r,x})^{\mathrm{QTAIM}}$, by which their R^2 linear coefficients are 0.85 (see Figure 5) and 0.96 (see Figure 6), respectively. Over all these years a large number of studies wherein the QTAIM theory has been successfully applied was published,[18–29, 73–76] but in general these studies were effectuated by using complete *ab initio* basis sets. Besides quantification of ΔE, we would like to say that other parameters can be used to predict the hydrogen bond strength of the $C_2H_4O\cdots HX$ hydrogen-bonded complexes, for instance the vibrational red-shifts of the HX proton donors and their relationships with $\rho(\mathrm{r,x})^{\mathrm{QTAIM}}$ amounts.[77]

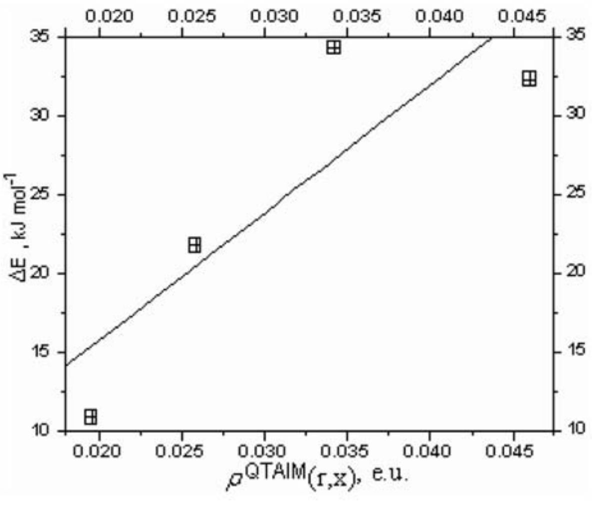

Figure 5. Relationship between $\rho(\mathrm{r,x})^{\mathrm{QTAIM}}$ values and hydrogen bond energies of the $C_2H_4O\cdots HX$ complexes using the B3LYP/6-31G(d,p) calculations.

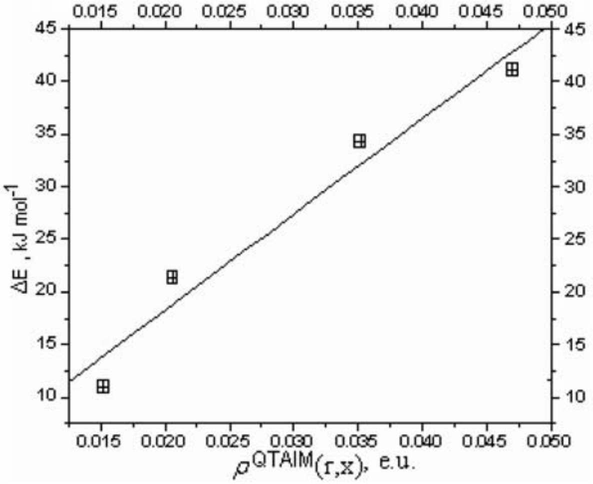

Figure 6. Relationship between $\rho(\mathrm{r,x})^{\mathrm{QTAIM}}$ values and hydrogen bond energies of the $C_2H_4O\cdots HX$ complexes using the B3LYP/6-31++G(d,p) calculations.

5. Discussion

It was demonstrated that a complete basis set cannot always provide a better reproduction and/or prediction of a molecular property. Even though the vibrational parameters of the $C_2H_4O\cdots HX$ hydrogen-bonded complexes have been recently published,[47–48] the chemometric study presented in this work seems to be very promising because there are some limitations in QTAIM calculations, as exposed by Bader,[78] Mohallem,[79] Site,[80] and Kryachko.[81] However, it is well established that QTAIM calculations strongly depend on the atomic size, specially in the electronegativity.[82] Thus, it is known that the description of electronegativity is well correlated with *ab initio* basis sets. Because the basis sets can be grouped into two basic classes known as Slater Type Orbital (STO) and Gaussian Type Orbital (GTO), see equations (7) and (8), respectively, it is also prudent to remember that STO have a nuclear "cusp"[83] Thus, the combination of GTO to form STO yields the split-valence basis set, and thereby, the electronic density is automatically considered.

$$\mathrm{STO} = \frac{\zeta^3}{\pi^{0,5}}\, e^{\left(-\zeta\, \vec{r}\right)} \qquad (7)$$

$$\mathrm{GTO} = \frac{2\zeta}{\pi^{0,75}} e^{\left(-\zeta\, \vec{r}^{2}\right)} \tag{8}$$

According to Cassam-Chenaï and Jayatilaka[84] some limitations of the QTAIM theory and its relationship with *ab initio* wave functions were debated. However, we are not proposing a direct connection between QTAIM densities and *ab initio* basis sets, by contrary, we are presenting chemometric results by which the (11) valence and (++) diffuse functions are important for prediction of the intermolecular electronic density and hence the hydrogen bond strength.

6. Conclusions

In this chemometrical work the relationship of the QTAIM intermolecular electronic density and *ab initio* basis sets was studied. Although a slight connection between QTAIM topological parameters and *ab initio* basis sets is widely known, here we have demonstrated that $\rho(\mathrm{r,x})^{\mathrm{QTAIM}}$ values of the (O···HX) hydrogen bonds of the C_2H_4O···HX complexes are essentially dependent on (11) valence and (++) diffuse functions. Considering that the C_2H_4O molecule has its electronic structure well-founded in terms of polarization effects, even so the (d,p) functions have contributed insignificantly for the chemometric analysis. Because the QTAIM intermolecular electronic densities are calculated from BCP localized between the O and H atoms, the prominent contribution of the (11) valence and mainly (++) diffuse functions must be related to the quantification of the charge density in the (O···HX) hydrogen bonds. Once again, we have no desire to propose a new insight about basis set quality, but our chemometric analysis suggests that the most appropriated *ab initio* basis set to be used for development of QTAIM calculations must be carefully selected.

7. Acknowledgments

The author would like to gratefully thank the CAPES and CNPq Brazilian funding agencies.

8. References

1. L. Pauling, The Nature of the Chemical Bond, Cornell University Press, 1939.
2. R. S. Mulliken, *J. Chem. Phys.* **1955**, *23*, 1833–1840.
3. F. B. van Duijneveldt, J. N. Murrell, *J. Chem. Phys.* **1967**, *46*, 1759–1767.
4. H. Umeyama, K. Morokuma, *J. Am. Chem. Soc.* **1977**, *99*, 1316–1332.
5. R. Taylor, O. Kennard, *J. Am. Chem. Soc.* **1982**, *104*, 5063–5070.
6. N. Balakrishman, *J. Chem. Phys.* **2004**, *121*, 5563–5566.
7. P. Hobza, Z. Havlas, *Chem. Rev.* **2004**, *100*, 4253–4264.
8. I. Rozas, I. Alkorta, J. Elguero, *J. Phys. Chem. A* **1997**, *191*, 9457–9463.
9. G. R. Desiraju, The Weak Hydrogen Bond: In Structural Chemistry and Biology, Thomas Striner Science, 2001.
10. W. J. Hehre, R. F. Stewart, J. A. Pople, *J. Chem. Phys.* **1969**, *51*, 2657–2664.
11. T. Baba, M. Takeuchi, H. Nakai, *Chem. Phys. Lett.* **2006**, *424*, 193–198.
12. D. E. Woon, T. H. Dunning Jr., *J. Chem. Phys.* **1993**, *98*, 1358–1371.
13. R. A. Kendall, T. H. Dunning Jr., J. Harrison, *J. Chem. Phys.* **1992**, *96*, 6796–6806.
14. T. H. Dunning Jr., *J. Chem. Phys.* **1989**, *90*, 1007–1023.
15. J. S. Binkley, J. A. Pople, W. J. Hehre, *J. Am. Chem. Soc.* **1980**, *102*, 939–947.
16. B. F. Shchegolev, *J. Struct. Chem.* **1989**, *29*, 933–936.
17. M. J. Frisch, J. A. Pople, J. S. Binkley, *J. Chem. Phys.* **1984**, *80*, 3265–3269.
18. T. Clark, J. Chandrasekhar, G. W. Spitznagel, P. v. R. Schleyer, *J. Comp. Chem.* **1983**, *4*, 294–301.
19. R. F. W. Bader, Atoms in Molecules. A Quantum Theory, Clarendon, Oxford, U.K., 1990.
20. R. F. W. Bader, *J. Chem. Phys.* **1980**, *73*, 2871–2883.
21. D. Cremer, E. Kraka, *Croat. Chem. Acta.* **1984**, *57*, 1259–1281.
22. R. F. W. Bader, *Theor. Chem. Acc.* **2001**, *105*, 276–283.
23. S. J. Grabowski, *J. Phys. Chem. A* **2000**, *104*, 5551–5557.
24. R. F. W. Bader, P. M. Beddall, P. E. Cade, *J. Am. Chem. Soc.* **1971**, *93*, 3095–3100.
25. R. F. W. Bader, M. T. Carroll, J. R. Cheeseman, C. Chang, *J. Am. Chem. Soc.* **1987**, *109*, 7968–7979.
26. P. L. A. Popelier, *J. Phys. Chem. A* **1998**, *102*, 873–1878.
27. P. L. A. Popelier, *J. Phys. Chem. A* **1999**, *103*, 2883–2890.
28. U. Koch, P. L. A. Popelier, *J. Phys. Chem. A* **1995**, *99*, 9747–9754.
29. D. Cremer, E. Kraka, *J. Am. Chem. Soc.* **1985**, *107*, 3800–3810.
30. D. Cremer, E. Kraka, *J. Am. Chem. Soc.* **1985**, *107*, 3811–3819.
31. N. Trendafilova, G. Bauer, T. Mihaylov, *Chem. Phys.* **2004**, *302*, 95–104.
32. X. Li, G. Wu, Y. A. Abramov, A. V. Volkov, P. Coppens, *Proc. Natl. Aca. Sci. USA*, **2002**, *99*, 12132–12137.
33. B. S. Everitt, S. Landau, M. Leese, Cluster Analysis, 4th Ed., Oxford University Press, 2001.
34. R. Mukerjee, D. F. J. Wu, A Modern Theory of Factorial Design, Springer Series in Statistics, 2006.
35. M. J. Frisch *et al*, Gaussian 98W, Revision A.11.2, Gaussian, Inc., Pittsburgh PA, 1998.
36. P. Hohenberg, W. Kohn, *Phys. Rev. B* **1964**, *136*, 864–871.
37. W. Kohn, L. J. Sham, *Phys. Rev. A* **1965**, *140*, 1133–1138.
38. E. S. Kryachko, M. T. Nguyen, *J. Chem. Phys.* **2004**, *115*, 833–841.
39. A. Vila, A. M. Graña, R. A. Mosquera, *Chem. Phys.* **2002**, *281*, 11–22.

40. O. Lamarche, J. A. Platts, *Chem. Eur. J.* **2002**, *8*, 457–466.
41. T. S. van Erp, E. J. Meijer, *J. Chem. Phys.* **2003**, *118*, 8831–8840.
42. Y. Zhao, O. Tishchenko, D. G. Truhlar, *J. Phys. Chem. B* **2005**, *109*, 19046–19051.
43. J. A. Alonso, A. Mañanes, *Theo. Chem. Acc.* **2007**, *117*, 467–472.
44. C. Møller, M. Plesset, *Phys. Rev.* **1934**, *56*, 618–622.
45. Statistica for Windows, StatSoft, Inc., 2325 East 13th Street, 74104, Tulsa, OK, 1995.
46. The MATLAB v 5.1, The Language of Technical Computing, The MathWorks, 1997.
47. B. G. Oliveira, E. C. S. Santos, E. M. Duarte, R. C. M. U. Araújo, M. N. Ramos, A. B. Carvalho, *Spectrochim. Acta A* **2004**, *60*, 883–1887.
48. B. G. Oliveira, E. M. Duarte, R. C. M. U. Araújo, M. N. Ramos, A. B. Carvalho, *Spectrochim. Acta A* **2005**, *61*, 491–494.
49. S. B. Boys, F. Bernardi, *Mol. Phys.* **1970**, *19*, 553–557.
50. D. A. McQuarrie, Statistical Thermodynamics, Harper and Row, New York, 1973.
51. K. Hermansson, *J. Phys. Chem. A* **2002**, *106*, 4695–4702.
52. B. G. Oliveira, R. C. M. U. Araújo, A. B. Carvalho, M. N. Ramos, *Quím. Nova* **2007**, *30*, 1167–1170.
53. T. J. Balle, W. H. Flygare, *Rev. Sci. Instrum.* **1981**, *52*, 333–345.
54. B. G. Oliveira, R. C. M. U. Araújo, A. B. Carvalho, M. N. Ramos, *J. Mol. Model.* **2009**, *15*, 123–131.
55. P. G. Mezey, *Mol. Phys.* **1999**, *96*, 169–178.
56. P. G. Mezey, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 224–230.
57. P. L. A. Popelier, F. M. Aicken, *J. Am. Chem. Soc.* **2003**, *125*, 1284–1292.
58. B. G. Oliveira, R. C. M. U. Araújo, F. F. Chagas, A. B. Carvalho, M. N. Ramos, *J. Mol. Model.* **2008**, *14*, 949–955.
59. B. G. Oliveira, R. C. M. U. Araújo, A. B. Carvalho, M. N. Ramos, *J. Mol. Model.* doi: 10.1007/s00894-008-0422-9.
60. A. Y. Li, *J. Phys. Chem. A* **2006**, *110*, 10805–10816.
61. B. G. Oliveira, R. C. M. U. Araújo, A. B. Carvalho, M. N. Ramos, *J. Theor. Comp. Chem.* **2007**, *6*, 647–660.
62. B. G. Oliveira, R. C. M. U. Araújo, A. B. Carvalho, M. N. Ramos, *Spectrochim. Acta A* **2007**, *68*, 626–631.
63. M. Solimannejad, I. Alkorta, *Chem. Phys.* **2006**, *324*, 459–464.
64. B. G. Oliveira, R. C. M. U. Araújo, M. N. Ramos, *Struct. Chem.* **2008**, *19*, 185–189.
65. B. G. Oliveira, R. C. M. U. Araújo, M. N. Ramos, *Struct. Chem.* **2008**, *19*, 665–670.
66. D. Hugas, S. Simon, M. Duran, *Struct. Chem.* **2005**, *16*, 257–263.
67. M. Meot-Ner (Mautner), *Chem. Rev.* **2005**, *105*, 213–284.
68. C. Desfrançois, S. Carles, J. P. Schermann, *Chem. Rev.* **2000**, *100*, 3943–3962.
69. S. K. Panigrahi, G. R. Desiraju, *J. Biosci.* **2007**, *34*, 677–691.
70. S. Wojtulewski, S. J. Grabowski, *Chem. Phys. Lett.* **2003**, *378*, 388–394.
71. S. J. Grabowski, *J. Mol. Struct.* **2001**, *562*, 137–143.
72. S. J. Grabowski, *J. Mol. Struct.* **2000**, *553*, 151–156.
73. B. G. Oliveira, R. C. M. U. Araújo, A. B. Carvalho, E. F. Lima, W. L. V. Silva, M. N. Ramos, A. M. Tavares, *J. Mol. Struct. (THEOCHEM)* **2006**, *775*, 39–45.
74. B. G. Oliveira, M. L. A. A. Vasconcellos, *J. Mol. Struct. (THEOCHEM)* **2006**, *774*, 83–88.
75. B. G. Oliveira, F. S. Pereira, R. C. M. U. Araújo, M. N. Ramos, *Chem. Phys. Lett.* **2006**, *427*, 181–184.
76. B. G. Oliveira, R. C. M. U. Araújo, A. B. Carvalho, M. N. Ramos, *Chem. Phys. Lett.* **2007**, *433*, 390–394.
77. B. G. Oliveira, R. C. M. U. Araújo, F. S. Pereira, E. F. Lima, W. L. V. Silva, A. B. Carvalho, M. N. Ramos, *Quim. Nova* **2008**, *31*, 1673–1679.
78. R. F. W. Bader, *Theor. Chem Acc.* **2002**, *107*, 381–382.
79. J. R. Mohallem, *Theor. Chem. Acc.* **2002**, *107*, 372–374.
80. L. D. Site, *Theor. Chem. Acc.* **2002**, *107*, 378–380.
81. E. S. Kryachko, *Theor. Chem. Acc.* **2002**, *107*, 375–377.
82. C. L. Perrin, *J. Am. Chem. Soc.* **1991**, *113*, 2865–2868.
83. P. T. A. Galek, N. C. Handy, A. J. Cohen, G. K.-L. Chan, *Chem. Phys. Lett.* **2005**, *404*, 156–163.
84. P. Cassam-Chenaï, D. Jayatilaka, *Theor. Chem. Acc.* **2001**, *105*, 213–218.

Povzetek

Intermolekularne elektronske gostote v kompleksih $C_2H_4O\cdots HX$ (X = F, CN, NC, and CCH), vsebujočih vodikovo vez, so bile določene s pomočjo izračunov izvedenih s teorijo gostotnih funkcionalov (DFT). B3LYP funkcional in Popleov 6-31G bazni set z razcepljenimi valencami sta bila združena v (11) valenc, (++) difuzne in (d,p) polarizacijske funkcije z namenom določitve optimiziranih geometrij $C_2H_4O\cdots HX$ kompleksov. S pomočjo teh geometrij smo s statističnima metodama hierarhične analize klastrov (HCA) in dvostopenjskim faktorialnim načrtovanjem (TLFD) raziskali intermolekularne elektronske gostote $\rho(r,x)^{QTAIM}$ izračunane iz kvantne teorije atomov v molekulah (QTAIM). Čeprav je dobro znano, da je QTAIM pristop vsaj delno odvisen od *ab initio* baznih setov, je naša raziskava razkrila dobro korelacijo med $\bar{\rho}(r,x)$ in vrednostmi, ocenjenimi s kemometrično analizo ($\rho(r,x)^{EST}$), izračunanimi v bistvu z valenčnimi (11) in difuznimi (++) funkcijami.