



SciForum
MOL2NET

New insights to understand the CoMFA and CoMSIA analysis within the framework of Density Functional Theory. Toward a generalized methodology.

Alejandro Morales-Bayuelo^{1,*}.

¹Fondo Nacional de Desarrollo Científico y tecnológico (FONDECYT), Proyecto postdoctoral N^o 3150035, Talca, Chile.

*Author to whom correspondence should be addressed; E-Mail: almorales@utalca.cl.

Abstract: Currently, the three-dimensional quantitative structure-activity relationship (3D QSAR) models have many applications; however due to the complexity to understand its results is necessary postulate new methodologies. In this sense, this work postulates a generalized version joining the quantum similarity field and chemical reactivity descriptors within the framework of density functional theory.

This generalized methodology can be applied to understand the biological activity on a molecular set taking a reference compound. In this sense, this methodology allows study the CoMFA and CoMSIA results in term of quantum similarity and chemical reactivity. In this form, is possible study steric and electrostatic effect on local substitutions. Considering that these methodologies can be used when the receptor is known or even when it is not known.

Keywords: Comparative Molecular Similarity Field Analysis (CoMFA), Comparative Molecular Similarity Indexes Analysis (CoMSIA), 3D-QSAR, Molecular Quantum Similarity (MQS), Chemical Reactivity Descriptors, Density Functional Theory (DFT).

Introduction: Taking into account that the CoMFA and CoMSIA analysis have many applications in the three-Dimensional Quantitative Structure-Activity Relationships (3D QSAR) studies [1], in this work are presented a new considerations about these methodologies within the Density Functional Theory (DFT) context.

In the DFT framework, the MQS is a field very applied, was introduced by Carbó and co-workers [2-6]. In the MQS field the key variable is the density function [9-11]; therefore it reasonable think that

can be related with the chemical reactivity descriptors such as chemical hardness (η), softness (S), electrophilicity (ω) and Fukui Functions. Therefore, employing this hybrid methodology (joining the MQS and chemical reactivity) is hope show new insight on the understanding of the CoMFA and CoMSIA results within the DFT context. In this form, the main aim of this work is presents new relationship between the MQS and Chemical reactivity that can be applied on the CoMFA and CoMSIA analysis.

Results and Discussion:

2. New insights proposed.

2.1. Quantum object sets (QOS)

Considering the follow set: $Z = \{z_I | I = 1, N\}$, being N the set cardinality, is the Cartesian product of two sets: $Z = \{P \times M\}$, where $P = \{p_I | I = 1, N\}$ is the set of the objects and $M = \{m_I | I = 1, N\}$ the set of tags. Therefore we can write: $\forall I = 1, N: z_I = (p_I; m_I) \in Z$ [12,13]. In this sense, a (QOS) is a tagged set: $Q = P \times S$ made by a set of submicroscopic objects and a set of quantum mechanical Density Function (DF): $S = \{\rho_I | I = 1, N\}$, as elements of the tag set.

Of this form, we can use the Hilbert semispace tag set S and define a central averaged DF using the expression:

$$\rho_c = N^{-1} \sum_I \rho_I \rightarrow \langle \rho_c \rangle = N^{-1} \sum_I \langle \rho_I \rangle = N^{-1} \sum_I v_I = v_c \quad (1)$$

with the DF Minkowski norms being define as:

$$\forall I = 1, N: \langle \rho_I \rangle = \int_D \rho_I(r) dr = v_I \quad (2)$$

Therefore, the centroid DF can be seen as a function describing the arithmetic average of the number of particles v_c of all quantum objects involved. To relate the shape functions associated to the quantum set, we define tag set H associated to the DF set S, therefore we have:

$$S = \{\rho_I | I = 1, N\} \rightarrow \forall \langle \rho_I \rangle = v_I \wedge \sigma_I = v_I^{-1} \rho_I \wedge \langle \sigma_I \rangle = 1 \rightarrow H = \{\sigma_I | I = 1, N\} \quad (3)$$

From this equation 3, we can write the shape centroid function as:

$$\sigma_c = N^{-1} \sum_I \sigma_I \rightarrow \langle \sigma_c \rangle = N^{-1} \sum_I \langle \sigma_I \rangle = N^{-1} N = 1 \quad (4)$$

2.2 Local Molecular Quantum Similarity Measure. A generalized version.

The quantum similarity measure Z_{AB} between compounds A and B, with electron density $\rho_A(r_1)$ and $\rho_B(r_2)$ respectively, can be understood using the minimizing of the expression for the Euclidean distance as [12-15]:

$$D_{AB} = \left(\int |\rho_A(r) - \rho_B(r)|^2 dr \right)^{1/2} = \left(\int (\rho_A(r_1))^2 dr_1 + \int (\rho_B(r_2))^2 dr_2 - 2 \int \rho_A(r_1) \rho_B(r_2) dr_1 dr_2 \right)^{1/2} \quad (5)$$

$$= \sqrt{Z_{AA} + Z_{BB} - 2Z_{AB}}$$

Where Z_{AB} is the overlap integral between the electron density of the compound A and B into the (QOS), Z_{AA} and Z_{BB} are the self-similarity of compounds A and B [16].

In this researcher we have used the Carbó index due to that is very used in the quantum similarity context [12-15]:

$$I_{AB} = \frac{\int \int \rho_A(r_1) \rho_B(r_2) dr_1 dr_2}{\sqrt{\left(\int \rho_A(r_1) dr_1 \right)^2 \left(\int \rho_B(r_2) dr_2 \right)^2}} \quad (6)$$

As the main structural difference on the molecules used by our group in the previous work are local differences [1]; the similarity features can be associated from the local point of view, in this order of ideas is used the Hirshfeld approach to study the local quantum similarity.

With the aim to obtain a generalized Hirshfeld approach to our systems, considering the electron density $\rho(r)$ in contribution $\rho_{x^1}(r)$, where x is an atom. These contributions allow define a concept of atom in a reference system and study its (dis)similarity on a molecular set (i.e.; substituent effect analysis). On the other hand, these contributions are proportional to the weight $w_C(r)$ of the electron density of the isolated compound in the so-called *promolecular density* [17-20]. The promolecular density is defined as:

$$\rho_{x^1}^{Prom}(r) = \sum_y \rho_y^0(r) \quad (7)$$

To calculate the contribution of an atom (x) in the electron density in a molecule A $\rho_A(r)$ we have:

$$\rho_{x^1}(r) = w_{x^1}(r) \rho_A(r) \quad (8)$$

In this form, the weight ($w_x(r)$) is obtained as:

$$w_{x^1}(r) = \frac{\rho_{x^1}^0(r)}{\sum_y \rho_y^0(r)} \quad (9)$$

Here $\rho_{x^1}^0(r)$ is the electron density of the isolated carbon atom x^1 , (i.e.; the reference electron density) [21]. In this sense, the contribution atomic of other carbon atom (x^2) in a molecule B is obtained as:

$$\rho_{x^2,B}(r) = w_{x^2}(r)\rho_B(r) \quad (10)$$

with

$$w_{x^2,B} = \frac{\rho_{x^2,B}^0(r)}{\sum_y \rho_y^0(r)} \quad (11)$$

So we can write the contribution of the asymmetric carbon atom products $\rho_A(r)\rho_B(r)$ as:

$$\rho_{x,AB}(r) = w_{x,AB}(r)\rho_A(r)\rho_B(r) \quad (12)$$

Using the equations (7-12) we can write the numerator Z_{AB} in the Carbó index (equation 6) as:

$$Z_{A,B}^{Local,x} = \frac{Z_{AB}}{\sqrt{Z_{AA}Z_{BB}}} = \frac{\iint w_{x,AB}\rho_A(r)\rho_B(r)dr_A dr_B}{\sqrt{\left(\int w_{x,A}(r)\rho_A(r)dr_A\right)^2 \left(\int w_{x,B}(r)\rho_B(r)dr_B\right)^2}} \quad (13)$$

The equation 10 is a generalized Hirshfeld approach to our systems [1], where $x=$ is an atom, therefore we can write the global index (equation 6) as local contributions. In this context, is possible study the local similarity and the substituent effects on some reference compound into the (QOS).

2.3 Reactivity descriptors.

The CoMFA and CoMSIA analysis are understand in terms of physical-chemistry properties such as electrostatic, hydrophobic and hydrogen-bond donor or acceptor properties, these properties can be related with global chemical descriptors as chemical potential, hardness, electrophilicity index and local reactivity descriptors as the Fukui Functions [1]. In the DFT context, the global reactivity indexes give information about the reactivity or stability of a chemical system front to external perturbations.

The chemical potential (μ) can be understood as the tendency that have the electrons to exit of the electron cloud and is calculate according to the equation:

$$\mu \approx \frac{\varepsilon_H + \varepsilon_L}{2} \quad (14)$$

Where (ε_H) is the energy of the (HOMO) and (ε_L) is the energy of the (LUMO) [22, 23]. Using the equation 14 the chemical hardness is defined according to Pearson et. al. [24].

$$\eta \approx \varepsilon_L - \varepsilon_H \quad (15)$$

From the equation (12), we obtain the softness [25] as:

$$S = \frac{1}{\eta} \quad (16)$$

Finally, using the equations 14 and 15 is defined the electrophilicity index (ω) [25, 26]. This index is understood as the measure of the stabilization energy of the system when it is saturated by electrons from the external environment and is calculated as follows:

$$\omega = \frac{\mu^2}{2\eta} \quad (17)$$

Finally, the Fukui Functions (equation 18 and 19, $f(r)$) are defined as the derivative of the electronic density with respect to the number of electrons at constant external potential:

$$f_k^+ \approx \int_k [\rho_{N+1}(\vec{r}) - \rho_N(\vec{r})] = [q_k(N+1) - q_k(N)] \quad (18)$$

$$f_k^- \approx \int_k [\rho_N(\vec{r}) - \rho_{N-1}(\vec{r})] = [q_k(N) - q_k(N-1)] \quad (19)$$

Where q_k refers to the electron population at k^{th} atomic site in a molecule. (f_k^+) governing the susceptibility for nucleophilic attack and (f_k^-) governing the susceptibility for electrophilic attack [27-30].

2.4 Quantum Operators to calculate Local Similarity.

One the most operator used in quantum similarity is the Dirac delta distribution $\Omega(r_1, r_2) = \delta(r_1, r_2)$ [31] so called overlap molecular quantum similarity measure and relates the volume associated with the overlap of the two densities $\rho_A(r)$ and $\rho_B(r)$:

$$\begin{aligned}
Z_{A,B}^{Local,x}(\Omega) &= \frac{Z_{AB}(\Omega)}{\sqrt{Z_{AA}(\Omega)Z_{BB}(\Omega)}} \\
&= \frac{\iint w_{x,AB} \rho_A(r) \delta(r_1 - r_2) \rho_B(r) dr_A dr_B}{\sqrt{\left(\int w_{x,A}(r) \rho_A(r) dr_A\right)^2 \left(\int w_{x,B}(r) \rho_B(r) dr_B\right)^2}} \\
&= \frac{\iint w_{x,AB} \rho_A(r) \rho_B(r) dr_A dr_B}{\sqrt{\left(\int w_{x,A}(r) \rho_A(r) dr_A\right)^2 \left(\int w_{x,B}(r) \rho_B(r) dr_B\right)^2}}
\end{aligned} \tag{20}$$

Using this equation 20, is possible obtain information about the electron concentration in the molecule and indicates the degree of overlap between the compared compounds.

Another operator very used in quantum chemistry is the Coulomb operator $\Phi(r_1, r_2)$, defined as $\Phi(r_1, r_2) = |r_1 - r_2|^{-1}$, this operator represents the electronic coulomb repulsion energy between molecular densities $\rho_A(r)$ and $\rho_B(r)$ as:

$$\begin{aligned}
Z_{A,B}^{Local,x}(\Phi) &= \frac{Z_{AB}(\Phi)}{\sqrt{Z_{AA}(\Phi)Z_{BB}(\Phi)}} \\
&= \frac{\iint w_{x,AB} \rho_A(r) \left(1/|r_1 - r_2|\right) \rho_B(r) dr_A dr_B}{\sqrt{\left(\int w_{x,A}(r) \rho_A(r) dr_A\right)^2 \left(\int w_{x,B}(r) \rho_B(r) dr_B\right)^2}} \\
&= \frac{\iint w_{x,AB} \rho_A(r) \rho_B(r) dr_A dr_B}{\sqrt{\left(\int w_{x,A}(r) \rho_A(r) dr_A\right)^2 \left(\int w_{x,B}(r) \rho_B(r) dr_B\right)^2}}
\end{aligned} \tag{18}$$

The Carbó index is restricted to the range (0,1) where $C_{AB}=0$ means dis(similarity) and $C_{AB}=1$ self-similarity, according to the Schwartz integral.

$$\left[\int \rho_A(r) \rho_B(r) dr \right]^2 \leq \int \rho_A^2(r) dr \int \rho_B^2(r) dr \tag{19}$$

2.5 Quantum Similarity Matrix.

The quantum similarity Matrix can be associated to a $[N \times N]$ metric associated to a (QOS) tag set made of quantum mechanical density function $S = \{\rho_I | I = 1, N\}$ as:

$$\mathbf{Z} = \begin{bmatrix} \langle \mathbf{z}_1 | \\ \langle \mathbf{z}_2 | \\ \vdots \\ \langle \mathbf{z}_N | \end{bmatrix} = [\langle \mathbf{z}_1 | \langle \mathbf{z}_2 | \cdots \langle \mathbf{z}_N |] \tag{20}$$

In the equation 20 there are equivalence between rows and columns. In this sense, we have:

$$\begin{aligned}
 \mathbf{Z} &= \{Z_{IJ} = \langle \rho_I | \rho_J \rangle | I, J = 1, N \} \\
 \wedge |z_I\rangle &= \{Z_{JI} | J = 1, N \} \\
 \wedge \langle z_I | &= \{Z_{JI} | J = 1, N \}
 \end{aligned}
 \tag{21}$$

Another property important of the matrix \mathbf{Z} is its symmetry, according to:

$$\mathbf{Z} = \mathbf{Z}^T \rightarrow \forall I, J : Z_{IJ} = Z_{JI}
 \tag{22}$$

Taking in account these properties associated to the similarity matrix, we can express the local molecular similarity measures using the overlap and coulomb operators (equations 18 and 19).

2.6 Joining QS and chemical reactivity

According to Carbó et al. [32] it is possible to consider a set of specific vectors and to associate a center for this QOS. Therefore a QOS represented by Fukui Functions can be defined as:

$$M = \{|I\rangle | I = 1, N \}
 \tag{23}$$

In equation 23, the first order densities can be constructed by a set of molecular orbital (MO) of shape functions contributions as:

$$P = \{\sigma_I = |I\rangle \langle I| | I = 1, N \}
 \tag{24}$$

The P elements correspond to the squared MO modulus. Using these consideration we can relate the frontier orbital (HOMO and LUMO) on the QOS. Defining $\{w_I\}$ as the number of occupation in the MOs, we can construct a linear combination of P to the first order density functional as [33]:

$$\rho = \sum_I w_I \sigma_I
 \tag{25}$$

with

(i) ν is the number of electrons: $\sum_I w_I = \nu$

(ii) where the Minkowski norms of the elements of the shape function set P are normalized to unity, belonging to the MO set normalization ($\forall I : \langle \sigma_I | \sigma_I \rangle = 1$).

Therefore we can define a centroid shape function using an average function.

$$\sigma_c = N^{-1} \sum_I \sigma_I \rightarrow \langle \sigma_c \rangle = 1 \quad (26)$$

In this sense, each elements of set P can be compared with the centroid function and can be built as:

$$\forall I : \theta_I = \sigma_I - \sigma_c \rightarrow Z = \{\theta_I | I = 1, N\} \quad (27)$$

Finally the Minkowski pseudonorm of the centroid shape function set **Z** can be written as:

$$\forall I : \theta_I = \langle \sigma_I - \sigma_c \rangle = \langle \sigma_I \rangle - \langle \sigma_c \rangle = 0 \quad (28)$$

Therefore the shifted elements have a null Minkowski pseudonorm. Where the shape function is comprised of N linearly independent elements. Using these relations we can make quantum similarity using the Fukui Functions on the QOS taking in account a reference compound.

$$Z_{AB}(f^{+/-}(r)) = \frac{[f^{+/-}(r)]_A [f^{+/-}(r)]_B}{\sqrt{[f^{+/-}(r)]_A^2 [f^{+/-}(r)]_B^2}} \quad (29)$$

Using the equation 29, we can build scales of convergence quantitative [1], among other. This equation shows a possible join between quantum similarity and chemical reactivity and can be used to calculate the quantum similarity on the local chemical reactivity (Fukui functions). These measures can be related with the contour maps generated by the CoMFA and CoMSIA results. The advantage of the proposed methodology is that shows a possible way to quantify the biological activity of the compounds additional to the reported by the 3D-QSAR studies.

Conclusions: In this work are reported new insights about the relationship between quantum similarity and chemical reactivity in a generalized form. This hybrid methodology, allow us study the steric and electrostatic effects in form of the scales of convergence quantitative; also substituent effects among others (see *J 2015 J. Mol. Model.* 21, 156 and *Mol2Net, 2015, 1 (Section B), pages 1-13, Proceedings*).

In this sense, the CoMFA and CoMSIA results can be modeled joining MQS and chemical reactivity; in this context these outcomes can be applied in QSAR correlations and docking studies to understand the biological activity of some molecular set. Taking into account that this methodologies can be used when the receptor is known or even when it is not known.

Conflicts of Interest: *The author declare no conflict of interest*

Acknowledgements: Thanks to the Universidad de Talca (CBSM)) for the continuous support to this investigation and finally to the **postdoctoral project N° 3150035 (FONDECYT, CHILE)**.

References:

1. Morales-Bayuelo, A.; Ayazo, H.; Vivas-Reyes, R. *Eur. J. Med. Chem.* 2010, **45**, 4509-4522.
2. Carbó-Dorca R, Arnau M, Leyda L 1980 *Int. J. Quant. Chem.* **17**, 1185.
3. Amat L, Carbó-Dorca R 2002 *Int. J. Quant. Chem.* **87**, 59.
4. Gironés X, Carbó-Dorca R 2006 *QSAR Comb. Sci.* **25**, 579.
5. Carbó-Dorca R, Gironés X 2005 *Int. J. Quant. Chem.* **101**,8.
6. Bultinck P, Rafat M, Ponc R, Gheluwe B V, Carbó-Dorca R, Popelier P 2006 *J. Phys. Chem. A.* **110**, 7642.
7. Robert D, Amat L, Carbó-Dorca R 1999 *J. Chem. Inf. Comp. Sci.* **39**, 333.
8. (a) Morales-Bayuelo A, Vivas-Reyes R. 2013 *J. Math. Chem.* **51**, 125. (b) Morales-Bayuelo A, Vivas-Reyes R. 2013 *J. Math. Chem.* **51**, 1835. (c) Morales-Bayuelo A, Torres J, Vivas-Reyes R **2012** *J. Theor. Comput. Chem.* **11**, 1. (d) Morales-Bayuelo A, Vivas-Reyes R. **2014** *J. Quant. Chem.* Article ID 239845, 19 pages. (e) Morales-Bayuelo A, Vivas-Reyes R. *J. Quant. Chem.* **2014**, Article ID 850163, 12 pages.
9. Parr R G, Yang W 1989 *Density Functional Theory of Atoms and Compounds*; Oxford University Press: New York.
10. Geerlings P, De Proft F, Langenaeker W, 2003 *Chem. Rev.* **103**, 1793.
11. R. Carbó-Dorca, *J. Math. Chem.* **23**, 353 (1998)
12. R. Carbó-Dorca, *Adv. Molec. Simil.* **2**, 43–72 (1998)
13. Carbó-Dorca R, Arnau M, Leyda L 1980 *Int. J. Quant. Chem.* **17**, 1185.
14. Amat L, Carbó-Dorca R 2002 *Int. J. Quant. Chem.* **87**, 59.
15. Gironés X, Carbó-Dorca R 2006 *QSAR Comb. Sci.* **25**, 579.
16. Carbó-Dorca R, Gironés X 2005 *Int. J. Quant. Chem.* **101**,8.
17. Bultinck P, Gironés X, Carbó-Dorca R (2005) *Rev. Comput. Chem.* **21**,127.
18. Hirshfeld F L 1977 *Theor. Chim. Acta.* **44**, 129.
19. De Proft F, Van Alsenoy C, Peeters A, Langenaeker W, Geerlings P 2002 *J. Comput. Chem.* **23**, 1198.
20. Randić M, Johnson M A, Maggiora G M 1990. In *Concepts and Applications of Molecular Similarity, Design of Compounds with Desired Properties*. Eds., Wiley-Interscience. New York. **77**.
21. Boon G, Van Alsenoy C, De Proft F, Bultinck P, Geerlings P **2005** *J. Mol. Struct.* **727**, 49.
22. Ayers P W, Anderson J S M, Bartolotti L J 2005 *Int. J Quant. Chem.* **101**, 520.
23. Harbola MK, Chattaraj PK, Parr RG 1991 *Isr. J. Chem.* **31**, 395.
24. Pearson R G **1997** *Chemical Hardness; Applications from Compounds to Solids*; Wiley-VHC, Verlag GMBH: Weinheim, Germany.
25. Yang W T, Parr R G 1985 *Proc. Natl. Acad. Sci.* **82**, 6723.
26. Chattaraj PK, Sarkar U, Roy DR 2006 *Chem. rev.* **106**, 2065.
27. Ayers P, Parr R G 2000 *J. Am. Chem. Soc.* **122**, 2010.
28. Galván M, Pérez P, Contreras R, Fuentealba P 1999 *Chem. Phys. Lett.* **30**, 405.
29. Mortier W J, Yang W 1986 *J. Am. Chem. Soc.* **108**, 5708.
30. Fuentealba P, Pérez P, Contreras R 2000 *J. Chem. Phys.* **113**, 2544.
31. Oliveira M, Cenzi G, Nunes RR, Andrighetti, Valadão D MS, Reis C, Oliveira Simões CM, Nunes RJ, Júnior MC, Taranto AG, Sanchez BAM, Viana GHR Varotti FP 2013 *Molecules.* **18**, 15276.
32. Carbó-Dorca R 2013 *J Math Chem.* **51**, 289.

33. Carbó-Dorca R 2008 J Math Chem. **44**, 621.

Notes: MDPI do not released one specific template for Sciforum conference. Consequently, this is not official template released by MDPI. In principle, the papers may be presented without specific format. However, the chairperson and the secretariat of the conference decided to create this template to give homogeneity to Mol2Net works. As is, this template should be used only for Mol2Net conference. Please, do not use this template for MDPI journals. Please, delete these notes before saving your final version. Type of the Paper (Proceeding, Letter, Expert Opinion, Communication, etc.). MDPI generates doi upon author request.

YouTube link: this option is only for those authors with welcome videos and/or oral presentations, plenary conferences uploaded to Mol2Net YouTube site.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions defined by MDPI AG, the publisher of the Sciforum.net platform. Sciforum papers authors the copyright to their scholarly works. Hence, by submitting a paper to this conference, you retain the copyright, but you grant MDPI AG the non-exclusive and unrevocable license right to publish this paper online on the Sciforum.net platform. This means you can easily submit your paper to any scientific journal at a later stage and transfer the copyright to its publisher (if required by that publisher). (<http://sciforum.net/about>).