Modeling Antimalarial Activity: Application of Kinetic Energy Density Quantum Similarity Measures as Descriptors in QSAR

Xavier Gironés, Ana Gallegos, and Ramon Carbó-Dorca*

Institute of Computational Chemistry, University of Girona, Campus Montilivi, 17071 Girona, Catalonia, Spain

Received May 30, 2000

In this work, is studied the application, within a quantum similarity framework, of the recently described Kinetic Energy Density Function in the evaluation of the antimalarial activity. First, this new type of Density Function is briefly presented from its theoretical foundations, and its inclusion in the molecular quantum similarity is discussed afterward. The application of Kinetic Energy-based Quantum Similarity Measures to QSAR is tested with 2 molecular sets composed of artemisinin derivatives, in which the 50% inhibition of synthesis and reduction of hidrofolate (IC_{50}) in different *Plasmodium falciparum* clones are analyzed. Satisfactory correlations are obtained for all antimalarial activities in all studied molecular sets. Molecular Quantum Similarity analysis provides a consistent, unbiased, and homogeneous set of molecular descriptors and is a feasible alternative to the use of classical physicochemical descriptors.

INTRODUCTION

Quantitative Structure−Activity Relationships (QSAR) find their origin in the XIX century when, in 1863, the toxicity of different alcohols in mammals was observed to increase with their decreasing solubility in water.[1] Since then, several attempts to relate the molecular structure to biological responses[2−4] or to physicochemical features[5−7] were postulated. However, present day formulation of QSAR is due to the work of Hansch and Fujita[8] in 1964. The basis of this method, and consequently of all later inspired ones, consists of assuming that the biological activity is the result of the contribution of many factors behaving independently. A structural descriptor characterizes each contribution, and the biological activity of a set of compounds is adjusted to a multilinear model. The most used descriptors in this approach are those related to hydrophobicity (log P), Hammett's constant (σ), and lipophilicity (π). Other molecular features, also taken into account in this scope when deriving QSARs, are partial charges, orbital energies,[9] polar interactions,[10] electrostatic potentials,[11] and steric parameters.[12] Parallel to these techniques and avoiding the use of physicochemical parameters, molecular topology was developed. This theory considers the biological activity related to the molecular topological features, which are numerically represented through connectivity and distance indices. Even though the seminal works in this field are also dated in the XIX century,[13] modern formulation is due to Wiener[14] and later continued by Kier and Hall[15] and Randic.[16] Other contributions also involve Hosoya, Balaban and Platt.[17] Several topological indices have also been successfully applied to QSAR.[18,19]

A new trend in QSAR studies appeared in 1988 due to the introduction of three-dimensional parameters, which could account for the influence of conformation and stereochemistry. These new techniques, commonly known as 3D-QSAR, use to involve extensive calculations in the space surrounding the molecules and thus, alignment of the molecules according to a common pharmacophore is required. The first model of this kind was the COMFA analysis[20] and it is one of the most used nowadays. Since then, several 3D-QSAR techniques have been presented[21−26] including those based on similarity concepts.[27−41]

Among these previously quoted methodologies dealing with similarity in QSAR studies, Molecular Quantum Similarity Measures (MQSM)[42−48] stand as an efficient tool to solve actual chemical problems. MQSM methodology has been successfully applied within pharmacological[49−54] and toxicological[55−57] problems. MQSM are based on the concept of molecular similarity[58,59] and are related to a self-evident molecular similarity principle: "the most similar two molecules are, the more similar properties will possess". This last statement requires a procedure to compare the molecules, and MQSM easily establish the degree of similarity based on the electronic distribution, namely density functions (DF), of the molecular structures.

Traditionally, MQSM have been computed using electronic DF. However, nothing opposes the use of the recently described extended DF[60−62] into the general formulation of MQSM. Among the assorted set of different extended DF, Kinetic Energy (KE) DF may constitute a suitable basis for the generation of molecular descriptors. In this work, it is intended to use Kinetic Energy-based MQSM to correlate the antimalarial activity of two series of compounds.

Malaria is an infectious disease endemic in many parts of the world.[63] Mostly located in tropical and subtropical areas, it is estimated that now between 300 and 500 million people have malaria. Therapies in these regions are being made difficult by the continuing spread of drug-resistant *Plasmodium falciparum* clones, due to the increasing immunity against traditional medications such as chloroquine,[64] and thus some reports indicate that this problem could perhaps be controlled by periodic introduction of new antimalarial drugs.[65,66] In present times, a new class of nonalkaloidal

* Corresponding author phone: +34 972418367; fax: +34 972 418356; e-mail: director@iqc.udg.es.

10.1021/ci0004558 CCC: $19.00
Published on Web 09/16/2000

trioxane compounds is becoming widespread: artemisinin and the semisynthetic ether and ester derivatives are reaching the general affected population. Artemisinin, a natural occurring peroxidyc cadinane sesquiterpene possessing good antimalarial activity[67,68] has been the subject of many total synthesis[69] and structure−activity relationships studies;[70−73] however, up to date few QSAR studies involving antimalarial compounds have appeared in the literature (see for example ref 74).

As stated before, much of this intense interest stems from the need for new antimalarial drugs with unconventional structures and novel modes of action to be used for the treatment of pervasive strains of drug-resistant *P. falciparum*. QSAR studies represent one of the best methodologies in computer-based drug design, offering valuable information about biological activity and providing a computationally inexpensive methodology to the design of potential bioactive drugs. Thus, this paper is devoted to this task, being consequently structured as follows: first, the KE DF is defined. Next, its inclusion in the MQSM framework is discussed, and afterward the statistical treatment of the resulting similarity matrices is exposed. As computational examples, the molecular sets composed by 18 and 15 artemisinin derivatives are presented and the corresponding QSAR analysis performed. Finally, the conclusions of this work are given.

MATERIAL AND METHODS

Kinetic Energy MQSM and Carbó Indices. KE DF arise from the mathematical problem of the inclusion of the second-order differential operator ∇^2, accounting for the expectation value of KE, within quantum mechanical theory. As quantum mechanical KE expectation values do not fulfill the usual expectation value statistical formalism, which other nondifferential operators do, the following expression (1) might be used in order to obtain a statistically consistent probability distribution

$$2\langle K\rangle = -\int \Psi^* \nabla^2 \Psi d\mathbf{r} = \int (\nabla\Psi)^* (\nabla\Psi) d\mathbf{r} = \int \kappa(\mathbf{r}) d\mathbf{r} \quad (1)$$

where $\langle K\rangle$ is the KE expectation value and Ψ any state wave function. The corresponding expression for the KE DF, distribution within Molecular Orbital (MO), theory can be written as

$$\kappa(\mathbf{r}) = \sum_i w_i |\nabla \varphi_i(\mathbf{r})|^2 \quad (2)$$

where $\{\varphi_i\}$ are the system's MOs and $\{w_i\}$ are the corresponding occupation numbers. This last equation provides the feasible inclusion of KE DF in the MQSM framework, resulting in the KE MQSM, between two Quantum Objects *A* and *B*

$$Z_{AB} = \int \kappa_A(\mathbf{r}) \kappa_B(\mathbf{r}) d\mathbf{r} \quad (3)$$

where Z_{AB} is the resulting KE MQSM and the *Overlap*[43,44,46,47] operator has been assumed. Z_{AB} provides the quantitative measure of resemblance between two molecules based on their KE DF, however, as Z_{AB} is an absolute magnitude, a transformation by means of the Carbó Index[42] is applied in order to obtain a relative value of similarity between the interval [0,1]. That is the more similar both molecules are, the closer to 1 the value of the Carbó Index will be and vice versa. In this way, all measures computed as in eq 3 may be transformed into Carbó Index[42] using the following expression:

$$C_{AB} = Z_{AB}(Z_{AA}Z_{BB})^{-1/2} \quad (4)$$

and, in the general case of QSAR studies, a set composed of *n* molecules can be submitted to this procedure.[48,49] The resulting values of the Carbó Index for every molecular pair can be collected in matrix form, whose columns could be thereafter used as molecular descriptors.

Promolecular Atomic Shell Approximation (ASA). To avoid expensive theoretical computations, a promolecular ASA[75−78] is used here in order to construct the electronic DF. This approach views molecular density as a sum of discrete atomic density contributions, which are taken as linear combinations of 1*S* Gaussian functions, previously fitted to atomic ab initio DF. ASA DF are mathematically expressed as

$$\rho_A^{ASA}(\mathbf{r}) = \frac{1}{N_e}\sum_{a\in A} Z_a \rho_a^{ASA}(\mathbf{r}) \wedge \rho_a^{ASA}(\mathbf{r}) = \sum_{i\in a} w_i |S_i(\mathbf{r}-\mathbf{R_I})|^2 \quad (5)$$

where N_e is the number of electrons of molecule *A*, Z_a is the atomic number of atom *a*, w_i are the convex fitted coefficients of the linear expansion,[76−78] and S_i are the 1*S* Gaussian functions, which are expressed in turn as

$$S_i(\mathbf{r}-\mathbf{R_I}) = N(\alpha_i)\exp(-\alpha_i|\mathbf{r}-\mathbf{R_I}|^2) \quad (6)$$

Once the overall atomic densities are built, each molecular density function can be constructed by adding as appropriate these elementary pieces together within a given geometry, obtained from experimental determinations or previous theoretical calculations. Since it has been proved that regular MQSM from molecular densities built in this way differ by less than 2% from ab initio densities,[76,78] their use is clearly justified.

Starting from the promolecular ASA DF framework, the gradient vector appearing in eq 2 can be written as

$$\nabla\varphi_i(\mathbf{r}) = \nabla(N(\alpha_i)\exp(-\alpha_i|\mathbf{r}-\mathbf{R_I}|^2)) = -2\alpha_i|\mathbf{r}-\mathbf{R_I}|S_i(\mathbf{r}-\mathbf{R_I}) \quad (7)$$

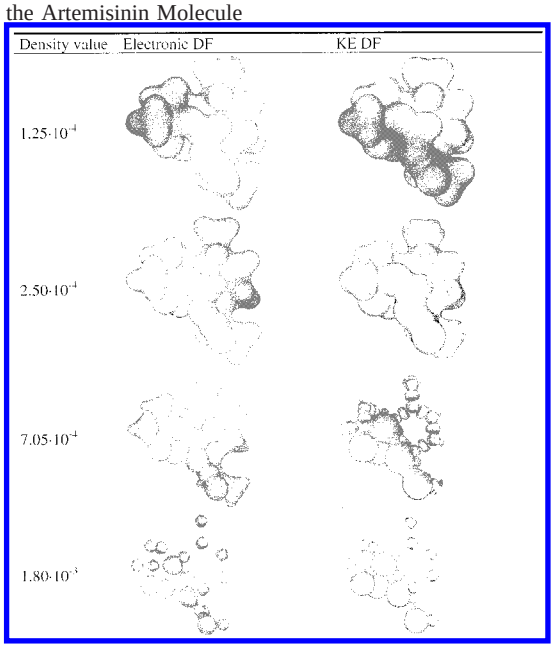
so a Promolecular ASA KE DF can be constructed as

$$\kappa(\mathbf{r}) = 4\sum_i w_i \alpha_i^2 |\mathbf{r}-\mathbf{R_I}|^2 |S_i(\mathbf{r}-\mathbf{R_I})|^2 \quad (8)$$

To make a clearer picture of the behavior of KE DF, a comparative analysis of both electronic and KE DF is presented in Table 1.

As seen in Table 1, at low and high DF values both representations behave almost in the same way, reflecting the molecular shape and the atomic locations, respectively. However, at intermediate values some relevant differences become visible, like an oversize of the heavier atoms and the appearance of interatomic maxima. These features

Table 1. Graphical Comparative Analysis of Electronic and KE DF over the Artemisinin Molecule



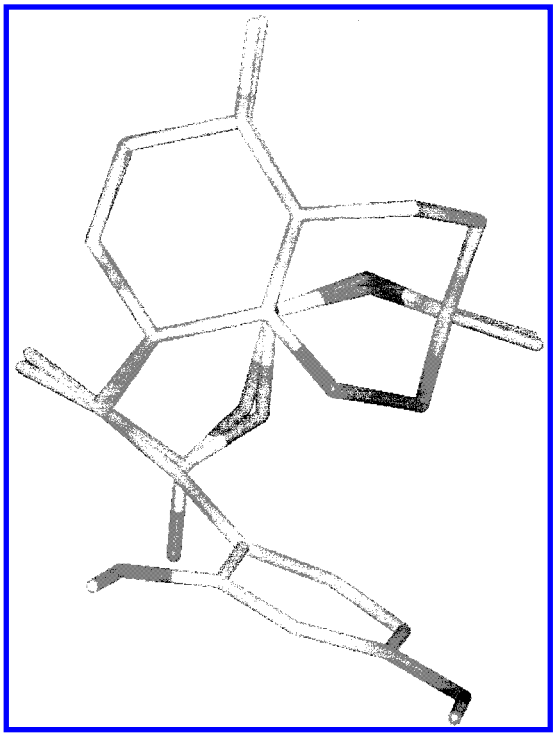


Figure 1. Pairwise molecular alignment between artemisinin and a derivative (compound number **2**).

preclude a different description of the electronic distribution based on KE concepts.

Molecular Alignment. MQSM are dependent on the relative position of the molecules under comparison. Although there have been described superposition procedures based on MQSM, in the present work, the simpler Topo-Geometrical Approach (TGSA) superposition[79] has been used to perform the pairwise alignments. This molecular superposition method overlays the molecules according to the maximal common substructure shared by the analyzed compounds. As illustrative examples of TGSA, a pairwise superposition between artemisinin and a derivative (second compound of the first set) as well as the global superposition based on artemisinin acting as a template are presented in Figures 1 and 2, respectively.

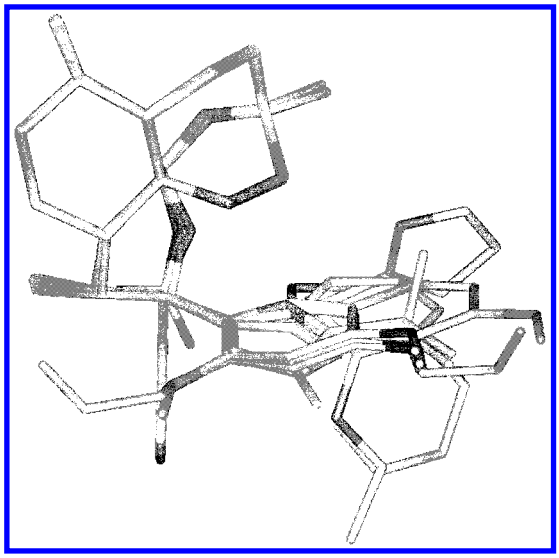


Figure 2. Molecular alignment solution for a set of 17 synthetic 1,2,4-trioxanes over artemisinin.

Table 2. Upper Triangle of the Carbó Index Matrix for Artemisinin Used To Compare the Different Computational Optimization Methodologies

Sybyl	AM1	HF/3-21 g*
1.000	0.978	0.985
	1.000	0.980
		1.000

As observed in Figure 1, both structures are aligned around the peroxy bridge present in artemisinin and the rest of the molecular structures belonging to the set. This fact is again evidenced observing Figure 2, where the entire molecular set is overlaid according to this substructure. Hence, the molecular skeleton has been detected and superposed.

Molecular Modeling. All molecular structures have been constructed using the PC Spartan software package[80] and optimized with the built-in *Sybyl* Molecular Mechanics force field. This choice was made after performing a comparative analysis between different optimization methodologies over the artemisinin molecule. The comparison was carried out using the following methodologies: *Sybyl* Molecular Mechanics force field and AM1, both included in PC Spartan,[80] and using a direct ab initio method, which in this case corresponds to a restricted Hartree−Fock with the 3-21G* basis set, implemented in Gaussian 94.[81] From the optimized molecular coordinates of artemisinin, the electronic DF has been constructed within the previously discussed promolecular ASA, and MQSM involving the Coulomb operator[44] have been carried out. The choice of Coulomb operator is based on a previous work[82] where a comparative analysis was also performed to choose the best geometry optimization method. The quantum similarity results over the artemisinin molecule are given in terms of Carbó Indices, resulting from the comparison of the different geometries of artemisinin and are presented in Table 2.

From the results shown in Table 2, it is evidenced that in this case all methods lead to very close structures. However, there is a very important difference in computational time

required to complete the calculations: whereas the *Sybyl* process was completed in a few seconds, AM1 and ab initio methods lasted some minutes and several hours, respectively. In this way, the *Sybyl* methodology was chosen to optimize the geometry due to its accuracy and efficiency regarding these molecular sets.

Treatment of Quantum Similarity Matrices and Model Building. Common chemometric tools are applied to deal with the resulting quantum similarity matrices. Principal Component Analysis (PCA)[83] is an appropriate technique to transform (dis)-similarity data. Taking the Carbó Index matrix, C ($n \times n$) being n the number of molecules in the studied set, as a starting point, PCA allows deriving a new $n \times n$ coordinate matrix in a new multidimensional space. These coordinates are called principal components (PCs) of the system and constitute the molecular descriptors used in this work.

Quantitative models are built up by means of multilinear regression using these PCs as independent *X*-variables. If a maximum number of descriptors, k, when constructing a QSAR model has to be chosen, a Nested Do Loop (NDL)[84−87] of dimension k over the n PC set is carried out. Each generated model is tested by means of its predictive capacity, evaluated by the squared leave-one-out cross-validation correlation coefficient (q^2), expressed as in eq 9

$$q^2 = \frac{SD - PRESS}{PRESS} \tag{9}$$

where SD is the quadratic deviation of the experimental values compared to their mean (expressed as $SD = \sum_i^n (y_i - \bar{y})^2$) and $PRESS$ stands for the predictive residual sum of squares (formulated as $PRESS = \sum_i^n (y_i - \hat{y})^2$). The chosen model, for k descriptors, is the one that maximizes the value of the q^2 parameter. The models obtained are evaluated by goodness-of-fit (r^2), standard deviation (σ_N) coefficients, and the previously mentioned q^2, derived from the leave-one-out cross-validation procedure.[88] Cross-validation consists of removing one molecule from the set, recalculating the model with the remaining elements, and predicting the value of the removed compound. This procedure is repeated for each member of the set and allows testing the predictive capacity of the proposed models by means of the squared cross-validated residuals (PRESS). A value of $q^2 > 0.5$ is generally accepted as satisfactory.[89]

Finally, the randomization test[90] is adopted to detect possible chance correlations. In this statistical technique, the dependent *Y*-variables are randomly permuted in their positions, and new predictive models are built with the altered vectors. If a real structure−activity exists, the only satisfactory results should be obtained for the correctly ordered *Y*-variables. Otherwise, even if the obtained model seems to correctly describe the system, it cannot be considered statistically significant because it is built up from an excess of parameters, able to correlate any data set, and thus it should be disavowed.

RESULTS AND DISCUSSION

Activity Data. For the set composed of 18 antimalarial compounds, the analyzed property is the nanomolar concentration of the drug able to inhibit 50% synthesis and reduction of hydrofolate in the NF54 strain (choloroquine sensitive) of *P. falciparum in vitro* (IC_{50}),[91] as reported in a previous work.[92] The molecular structures, as well as the activities,

Table 3. Molecular Structures and Biological Activities of a Set of 18 Artemisinin Derivatives

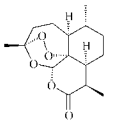
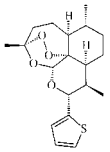
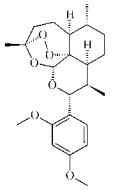
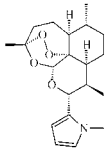
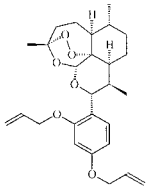
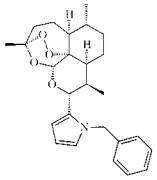
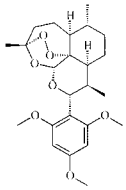
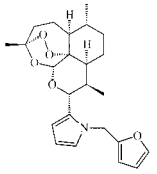
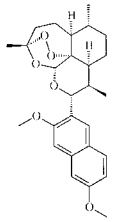
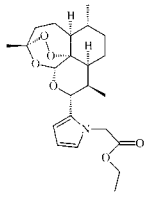
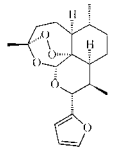
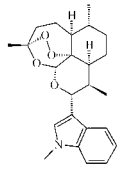
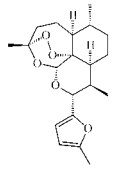
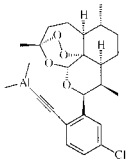
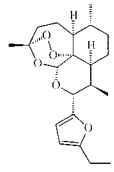
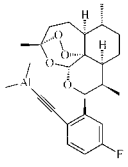
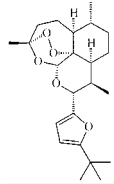
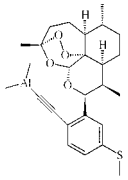
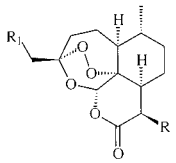
n°	Structure	IC_{50}	n°	Structure	IC_{50}
1		9.908	10		5.105
2		4.198	11		4.609
3		6.607	12		15.996
4		7.798	13		9.397
5		8.995	14		9.099
6		1.400	15		4.000
7		5.200	16		10.990
8		8.590	17		8.299
9		10.000	18		8.395

Table 4. Molecular Structures and Biological Activities of a Set of 15 Artemisinin Derivatives

Molecular Structure



n	R_1	R	log IC_{50} Relative activity D6	W2
1	H	CH_3	2.000	2.000
2	CH_3	H	1.944	2.049
3	CH_3CH_2	H	3.323	2.828
4	$CH_3(CH_2)_2$	H	1.301	1.255
5	$(CH_3)_2CH$	H	1.724	1.653
6	EtO_2CCH_2	H	2.365	2.365
7	$C_6H_5CH_2$	H	0.477	0.000
8	p-$ClC_6H_4(CH_2)_2$	H	2.057	2.104
9	$C_6H_4(CH_2)_3$	H	2.342	2.449
10	CH_3	$CH_3(CH_2)_3$	2.265	2.410
11	$CH_3(CH_2)_2$	$CH_3(CH_2)_3$	1.447	1.519
12	$C_6H_5CH_2$	$CH_3(CH_2)_3$	0.000	0.000
13	p-$ClC_6H_4(CH_2)_2$	$CH_3(CH_2)_3$	1.633	1.724
14	$C_6H_4(CH_2)_3$	$CH_3(CH_2)_3$	1.591	1.681
15	EtO_2CCH_2	$CH_3(CH_2)_3$	3.141	3.359

are presented in Table 3. For the second series, the studied property is also IC_{50}; however, the activity has been measured relative to the artemisinin value when tested in vitro in human blood over *P. falciparum* Indochina (W2) and Sierra Leone (D6) clones,[93] following a procedure proposed by Desjardins.[94,95] These clones present an interesting resistance to drugs: whereas the W2 strain is chloroquine-resistant and mefloquine-sensitive, the D6 breed is sensitive to chloroquine and resists mefloquine. This fact allows evaluating the effect and mechanism of action of a novel drug in this widely mutated parasite. The relative potency of these compounds has been adjusted according to IC_{50} values and then multiplied by the ratio between molecular weight of the analogue and molecular weight of artemisinin. These activities and the corresponding molecular structures are compiled in Table 4. A logarithmic scaling has been applied to the last set of activities due to the wide range of values present in ref 93.

Quantitative Structure−Activity Relationships. MQSM matrices have been constructed using the formula present in eq 3 and scaled using the Carbó Index notation present in eq 4. To quantify the influence on antimalarial activity of the diverse molecules according to their structure, the quantum similarity matrix deals with multivariate analysis techniques as detailed previously. These simple mathematical manipulations yield to reliable models that relate the molecular descriptors to the antimalarial activity.

A quantitative relationship between the structural descriptors and the property is established by means of a multilinear regression, as pointed out in the theoretical section. Valuable models are achieved using few descriptors, as presented in Table 5. Optimal values of $r^2 = 0.754$ and $q^2 = 0.561$ for the NF54 strain are obtained when using four descriptors. As an illustrative example, Figure 3 presents how the first two most predictive PCs spread the molecules in the 2D space.

As seen in Figure 3, both *PC 16* and *2* behave differently. *PC 16* scatters the molecules according to the weight of the substitution, roughly collapsing the heavy substitutions nearby the origin and distributing the light ones along the sides. *PC 2* mostly acts as a substituent discriminator. In the negative part of *PC 2*, most of the compounds presenting oxygen are present (**2**, **3**, **4**, **5**, **6**, **7**, **8**), those compounds having an aluminum atom are approximately located in the middle of the axis (**16**, **17**, **18**), and finally the nitrogenated substitutions are present in the positive part of *PC 2* (**11**, **12**, **13**, **14**). Exceptions to this rule are compound **1**, which only contains a ketonic substitution, compound **9**, which contains oxygen but is located with the nitrogen group, and compound **15**, which contains a nitrogen and it is located too close to the aluminum containing substitutions. Compound **10** contains a sulfur and is located between the oxygen containing rings and the heavier substitutions with aluminum.

Table 5. Statistical Parameters for the Proposed QSAR Models

P. falciparum strain	no. of PCs	PCs used	% variance explained	r^2	q^2	σ_N
NF54	4	2,4,9,16	27.775	0.754	0.561	2.140
D6	4	4,6,14,15	19.503	0.767	0.520	0.580
W2	4	1,4,14,15	37.693	0.821	0.576	0.575

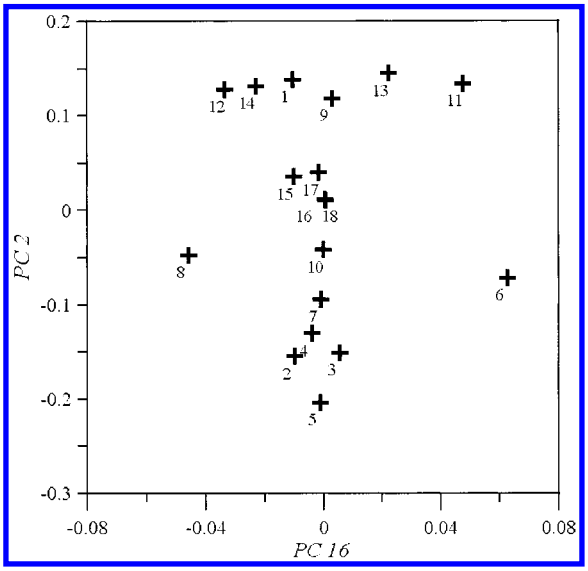


Figure 3. 2D distribution of the antimalarial compounds according to the first two chosen PCs.

Both biological activities of the second series are also correctly predicted using four PCs, achieving values of $r^2 = 0.767$ and $q^2 = 0.520$ for the D6 clone and $r^2 = 0.821$ and $q^2 = 0.576$ for the W2 strain. The equations of the optimal models are presented in eqs 10−12 and the statistical parameters in Table 5.

$$IC_{50}^{NF54} = 11.956\mathbf{x_2} - 12.179\mathbf{x_4} - 16.280\mathbf{x_9} - 74.446\mathbf{x_{16}} + 7.700 \quad (10)$$

$$\log IC_{50}^{D6} = 1.900\mathbf{x_4} + 4.620\mathbf{x_6} - 32.563\mathbf{x_{14}} + 5.965\mathbf{x_{15}} + 1.826 \quad (11)$$

$$\log IC_{50}^{W2} = 4.276\mathbf{x_1} + 2.507\mathbf{x_4} - 30.241\mathbf{x_{14}} + 5.822\mathbf{x_{15}} + 1.841 \quad (12)$$

As seen in Table 5, not always the PCs accounting for the maximal variance are those related to the activity. In the

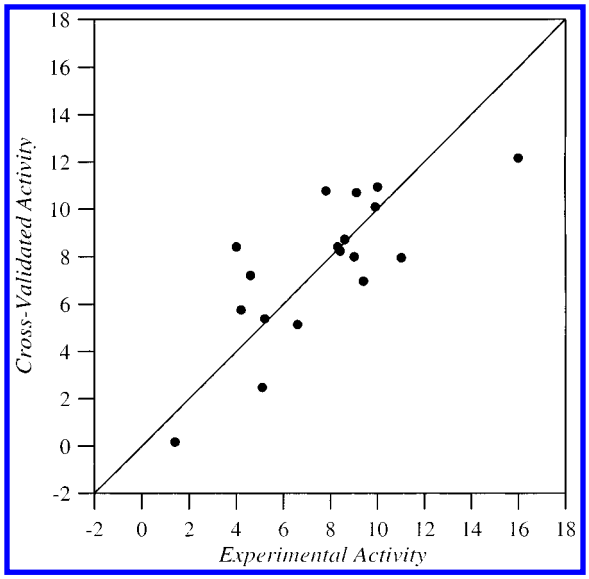


Figure 4. Cross-validated versus experimental antimalarial activity values over *P. falciparum* NF54.

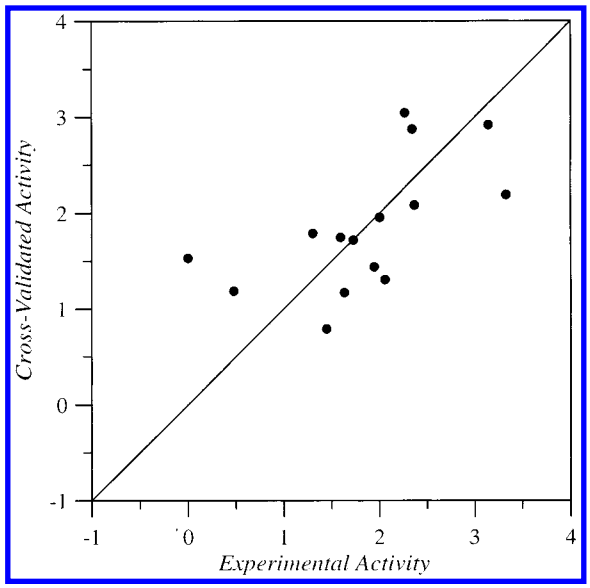


Figure 5. Cross-validated versus experimental antimalarial activity values over *P. falciparum* D6.

studied molecular sets, even the chosen PCs explain less than 40% of the whole variance, and they are able to provide an acceptable description.

Figures 4 to 6 present the plots of the cross-validated activities for the three models consecutively. As observed, most of the predicted values are in good agreement with the experimental ones with acceptable error.

Finally, the proposed models were subjected to a random test procedure to assess that they are not built up from an excess of parameters. Thus, in each molecular set, a hundred activity vectors were generated from randomized permutations of the original ordered one. All proposed models present a clear separation between the original solution and the randomized ones, which clearly achieve values of q^2 below 0.5 or furthermore being negative. As an example, Figure 7 presents the compilation of the results for the *P. falciparum* NF54 set. Similar results are obtained for the remaining studied systems.

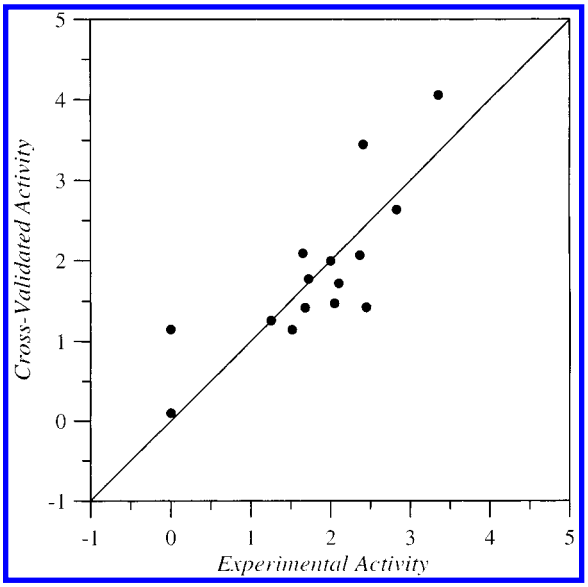


Figure 6. Cross-validated versus experimental antimalarial activity values over *P. falciparum* W2.

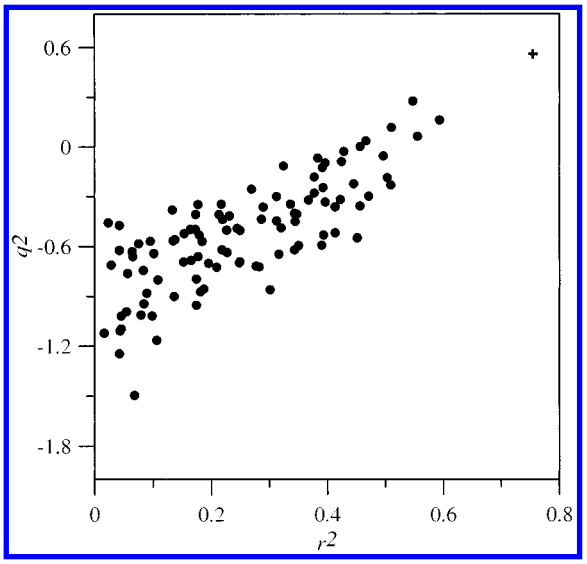


Figure 7. Randomization test for the optimal *P. falciparum* NF54 model. The randomized responses have been marked with dots, and the correctly ordered activity with a cross.

CONCLUSIONS

In the present study, Kinetic Energy based molecular quantum similarity measures have been applied to correlate systematically the antimalarial activity of various artemisinin derivatives. Quantitative structure−activity relationships provide valuable information about the biological behavior of potential drugs and might be of help in the development of new therapies in the fight against malaria. In addition, this approach carries low computational costs, which constitutes another interesting advantage to medical research.

ACKNOWLEDGMENT

This research has been partially supported by the European Commission Project number ENV4-CT97-0508. One of us (X. Gironés) benefits from a predoctoral fellowship from the University of Girona. Financial support from the *Fundació María Francisca de Roviralta* is also acknowledged. Some

comments by the evaluators have been helpful to ameliorate the structure of this paper, hence the authors wish to express their gratitude to them.

REFERENCES AND NOTES

(1) Quoted in: Borman, S. New QSAR Techniques Eyed for Environmental Assessments. *Chem. Eng. News* **1990**, *68*, 20−23.

(2) Crum-Brown. A.; Fraser, T. R. On the connection between chemical constitution and physiological action. Part 1. On the physiological action of the ammonium bases, derived from Strychia, Brucia, Thebaia, Codeia, Morphia and Nicotia. *Trans. Royal Soc. Edinburgh* **1868**−**1869**, *25*, 151−203.

(3) Meyer, H. Theorie der Alkoholnarkose; welche Eigenschaft die Anästhetica bedignt ihre narkotische Wirkung. *Arch. Exp. Pathol. Pharmakol.* **1899**, *42*, 109−118.

(4) Ferguson, J. The use of chemical potentials as indicators of toxicity. *Proc. Royal Soc. London* **1939**, *B127*, 387.

(5) Hammett, L. P. The Effect of Structures upon the Reactions of Organic Compounds. Benzene Derivatives. *J. Am. Chem. Soc.* **1937**, *59*, 96−103.

(6) Swain, C. G.; Lupton, E. C., Jr. Field and Ressonance Effects Components of Substituent Effects. *J. Am. Chem. Soc.* **1968**, *90*, 4328.

(7) Free, S. M.; Wilson, J. W. A mathematical contribution to structure−activity studies. *J. Med. Chem.* **1964**, *7*, 395−399.

(8) Hansch, C.; Fujita, T. $\rho-\sigma-\pi$ Analysis. A Method for the Correlation of Biological Activity and Chemical Structure. *J. Am. Chem. Soc.* **1964**, *86*, 148−152.

(9) Kubinyi, H. *QSAR: Hansch Analysis and Related Approaches*; VCH: Weinheim, 1993.

(10) Martin, Y. C. *Quantitative Drug Desing: A Critical Introduction*; Marcel Dekker: New York, 1978.

(11) *Chemical Application of Atomic and Molecular Electrostatic Potentials*; Politzer, P., Truhlar D. G., Eds.; Plenum: New York, 1981.

(12) Verloop, A.; Hoogenstraaten, W.; Tipker, J. In *Drug Design, Vol. VII*; Ariens, E. J., Ed.; Academic Press: New York, 1976; p 165.

(13) Sylvester, J. J. *Am. J. Math.* **1874**, *1*, 64.

(14) Wiener, H. Structural determination of boiling points. *J. Chem. Phys.* **1947**, *69*, 17−20.

(15) Kier, L. B.; Hall, L. H.; Murray, W. J.; Randic, M. Molecular connectivity. I: Relationship to nonspecific local anaesthesia. *J. Pharm. Sci.* **1975**, *64*, 1971−1974.

(16) Randic, M. On Characterization of Molecular Branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609−6615.

(17) Trinajstic, N. *Chemical graph theory*; CRC Press: Boca Raton, 1992.

(18) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press: New York, 1976.

(19) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure−Activity Analysis*; Research Studies Press: Letchwork, 1986.

(20) Cramer, R. D., III; Paterson, D. E.; Brunce, J. D. Comparative Molecular Field Analysis (COMFA). 1 Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1998**, *110*, 5959−5967.

(21) Klebe, G.; Abraham, U.; Mietzner, T. Molecular Similarity Indices in a Comparative Analysis (CoMSIA) of Drug Molecules to Correlate and Predict their Biological Activity. *J. Med. Chem.* **1994**, *37*, 4130−4146.

(22) Silverman, B. D.; Platt, D. E. Comparative Molecular Moment Analysis (CoMMA): 3D-QSAR without molecular superposition. *J. Med. Chem.* **1996**, *39*, 2129−2140.

(23) Jain, A. N.; Koile, K.; Chapman, D. Compass: Predicting Biological Activities from Molecular Surface Properties. Performance Comparisons on a Steroid Benchmark. *J. Med. Chem.* **1994**, *37*, 2315−2327.

(24) Bravi, G.; Gancia, E.; Mascagni, P.; Pegna, M.; Todeschini, R.; Zaliani, A. MS−WHIM, a new 3D theoretical descriptors derived from molecular surface properties: A comparative 3D-QSAR study in a series of steroids. *J. Comput.-Aided Mol. Design* **1997**, *11*, 79−92.

(25) Kellogg, G. E.; Kier, L. B.; Gaillard, P.; Hall, L. H. E-state fields: Applications to 3D QSAR. *J. Comput.-Aided Mol. Design* **1996**, *10*, 513−520.

(26) Chen, H.; Zhou, J.; Xie, G. PARM: A Genetic Evolved Algorithm to Predict Bioactivity. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 243−250.

(27) Martín, M.; Sanz, F.; Campillo, M.; Pardo, L.; Pérez, J.; Turmo, J.; Aulló, J. M. Quantum Chemical Structure−Activity Relationships on β-Carbolines as Natural Monoamine Oxidase Inhibitors. *Int. J. Quantum Chem.* **1983**, *23*, 1643−1652.

(28) Sanz, F.; Martín, M.; Lapeña, F.; Manaut, F. Quantitative structure−activity relationships on MAO substrates by means of quantum chemical properties. *Quant. Struct.-Act. Relat.* **1986**, *5*, 54−57.

(29) Sanz, F.; Manaut, F.; José, J.; Segura, J.; Carbó, M.; De la Torre, R. Automatic determination of MEP patterns of molecules and its application to cafeine metabolism inhibitors. *J. Mol. Struct. (THEOCHEM)* **1988**, *170*, 171−180.

(30) Luque, F. J.; Sanz, F.; Illas, F.; Pouplana, R.; Smeyers, Y. G. Relationships between the activity of some H2-receptor agonists of histamine and their ab initio molecular electrostatic potential (MEP) and electron density comparison coefficients. *Eur. J. Med. Chem.* **1988**, *23*, 7−10.

(31) Sanz, F.; Manaut, F.; Dot, T.; López de Briñas, E. Complete or partial comparison of molecular electrostatic potential distributions? Some tests with 5-HT ligands. *J. Mol. Struct. (THEOCHEM)* **1992**, *256*, 287−293.

(32) Sanz, F.; Manaut, F.; Rodriguez, J.; Lozoya, E.; Lopez de Briñas, E. MEPSIM: A computational package for analysis and comparison of molecular electrostatic potentials. *J. Comput.-Aided Mol. Design* **1993**, *7*, 337−347.

(33) Burt, C.; Huxley, P.; Richards, W. G. The Application of Molecular Similarity Calculations. *J. Comput. Chem.* **1990**, *11*, 1139−1146.

(34) Richard, A. M. Quantitative Comparison of Molecular Electrostatic Potentials for Structure Activity Studies. *J. Comput. Chem.* **1991**, *12*, 959−969.

(35) Rum, G.; Herndon, W. C. Molecular Similarity Concepts. 5. Analysis of Steroid-Protein Binding Constants. *J. Am. Chem. Soc.* **1991**, *113*, 9055−9060.

(36) Herndon, W. C.; Rum, G. Three- Dimensional Topological Descriptors and Similarity of Molecular Structures: Binding Affinities of Corticosteroids. In *QSAR and Molecular Modeling: Concepts, Computational Tools and Biological Applications: Proceedings of the 10th European Symposium on Structure−Activity Relationships, QSAR and Molecular Modelling*; Sanz, F., Giraldo, J., Manaut, F., Eds.; Prous Science: Madrid, 1995; pp 380−384.

(37) Good, A. C.; So, S.-S.; Richards, W. G. Structure−Activity Relationships from Molecular Similarity Matrices. *J. Med. Chem.* **1993**, *36*, 433−438.

(38) Good, A. C.; Peterson, S. J.; Richards, W. G. QSAR's from Similarity Matrices. Technique Validation and Application in the Comparison of Different Similarity Evaluation Methods. *J. Med. Chem.* **1993**, *36*, 2929−2937.

(39) Good, A. C.; Richards, W. G. The extension and application of Molecular Similarity to Drug Design. *Drug Information J.* **1996**, *30*, 371−388.

(40) Good, A. C.; Hodgkin, E. E.; Richards, W. G. the Utilization of Gaussian Functions for the Rapid Evaluation of Molecular Similarity. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 188−191.

(41) Good, A. C.; Richards, W. G. Rapid Evaluation of Shape Similarity Using Gaussian Functions. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 112−116.

(42) Carbó, R.; Arnau, J.; Leyda, L. How similar is a molecule to another? An electron density measure of similarity between two molecular structures. *Int. J. Quantum Chem.* **1980**, *17*, 1185−1189.

(43) Carbó-Dorca, R.; Besalú, E. A general survey of molecular quantum similarity. *THEOCHEM* **1998**, *451*, 11−23.

(44) Carbó-Dorca, R.; Amat, L.; Besalú, E.; Lobato, M. Quantum Similarity. In *Advances in Molecular Similarity*; Carbó-Dorca, R., Mezey, P. G., Eds.; JAI Press: Greenwich, CT, 1998; Vol. 2.

(45) Carbó-Dorca, R. Fuzzy sets and Boolean tagged sets; vector semispaces and convex sets; quantum similarity measures and ASA density functions; diagonal vector spaces and quantum chemistry. In *Advances in molecular similarity*; Carbó-Dorca, R., Mezey, P. G., Eds.; JAI Press: Greenwich, CT, 1998; Vol. 2, pp 1−42

(46) Besalú, E.; Carbó, R.; Mestres, J.; Solà, M. Foundations and recent developments on molecular quantum similarity. *Top. Curr. Chem.* **1995**, *173*, 31−62.

(47) Carbó, R.; Besalú, E.; Theoretical foundations of quantum molecular similarity. In *Molecular similarity and reactivity: from quantum chemical to phenomenological approaches*; Carbo, R., Ed.; Kluwer: Amsterdam, 1995; pp 3−30.

(48) Carbó-Dorca R.; Besalú, E.; Amat, L.; Fradera, X. Quantum molecular similarity measures: concepts, definitions and applications to quantitative structure-properties relationships. In *Advances in molecular similarity*; Carbó-Dorca, R., Mezey, P. G., Eds.; JAI Press: Greenwich, CT, 1996; Vol. 1, pp 1−42.

(49) Carbó, R.; Besalú, E.; Amat, L.; Fradera, X. Quantum molecular similarity measures (QMSM) as a natural way leading towards a theoretical foundation of quantitative structure-properties relationships. *J. Math. Chem.* **1995**, *18*, 237−246.

(50) Fradera, X.; Amat, L.; Besalú, E.; Carbó-Dorca, R. Application of molecular quantum similarity to QSAR. *Quant. Struct.-Act. Relat.* **1997**, *16*, 25−32.

(51) Lobato, M.; Amat, L.; Carbó-Dorca, R. Structure−activity relationships of a steroid family using quantum similarity measures and topological quantum similarity indices. *Quant. Struct.-Act. Relat.* **1997**, *16*, 465−472.

(52) Amat, L.; Robert, D.; Besalú, E.; Carbó-Dorca, R. Molecular quantum similarity measures tuned 3D QSAR: an antitumoral family validation study. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 624−631.
(53) Robert, D.; Amat, L.; Carbó-Dorca, R. Three-dimensional quantitative structure−activity relationships from tuned molecular quantum similarity measures. Prediction of the corticosteroid-binding globulin binding affinity for a steroid family. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 333−344.
(54) Robert, D.; Gironés, X.; Carbó-Dorca, R. Facet diagrams for quantum similarity data. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 597−610.
(55) Gironés, X.; Amat, L.; Robert, D.; Carbó-Dorca, R. Use of electron−electron repulsion energy as a molecular descriptor in QSAR and QSPR studies. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 477−485.
(56) Robert, D.; Carbó-Dorca, R., Aromatic compounds aquatic toxicity QSAR using quantum similarity measures. *SAR QSAR Environ. Res.* **1999**, *10*, 401−422.
(57) Gironés, X.; Amat, L.; Carbó-Dorca, R. Using molecular quantum similarity measures as descriptors in quantitative structure-toxicity relationships. *SAR QSAR Environ. Res.* **1999**, *10*, 545−556.
(58) Dean, P. M. Molecular Similarity. *In 3D QSAR in Drug Design: Theory, Methods and Applications*; Kubinyi, H., Ed.; ESCOM Science Publishers B. V.: Leiden, The Netherlands, 1993; pp 150−172.
(59) Richards, W. G. Molecular Similarity and Dissimilarity. In *Modeling of Biomolecular Structures and Mechanisms*; Pullman, A., Jortner, J., Pullman, B., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1995; pp 365−369.
(60) Carbó-Dorca, R.; Amat, L.; Besalú, E.; Gironés, X.; Robert, D. Quantum Mechanical Origin of QSAR: Theory and Applications. *J. Mol. Struct.* (*THEOCHEM)* **2000**, *504*, 181−228.
(61) Carbó-Dorca, R.; Besalú, E.; Gironés, X. Extended Density Functions and Quantum Chemistry. *Adv. Quan. Chem.* in press.
(62) Carbó-Dorca, R.; Amat, L.; Besalú, E.; Gironés, X.; Robert, D. Quantum Molecular Similarity: Theory and Applications to the Evaluation of Molecular Properties, Biological Activities and Toxicity. Proceedings of the 4th Girona Seminar on Molecular Similarity; Kluwer: in press.
(63) TDR News (News from the WHO Division of Control of Tropical Diseases) **1994**, *46*, 5.
(64) Peters, W. *Chemotherapy and Drug Resistance in Malaria*, 2nd ed.; Academic Press: London, 1987.
(65) Oaks, S. C.; Mitchell, V. S.; Pearson, G. W.; Carpenter, C. C. *MALARIA Obstacles and Opportunities;* National Academic Press: Washington, DC, 1991.
(66) Oduola, A. M. J.; Omitowoju, G. O.; Gerena, L.: Kyle, D. E.; Milhous, M. K.; Sowunmi, A.; Salako, L. A. Reversal of Mefloquine Resistance with Penfluridol in Isolates of Plasmodium Falciparum from South-West Nigeria. *Trans. R. Soc. Trop. Med. Hyg.* **1993**, *87*, 81−83.
(67) Klayman, D. L. Qinghaosu (Artemisinin): an antimalaric drug from China. *Science* **1985**, *228*, 1049−1055.
(68) Luo, X.-D.; Shen, C.-C. The chemistry, pharmacology, and clinical applications of qinghaosu (artemisinin) and its derivatives. *Med. Res. Rev.* **1987**, *7*, 29−52.
(69) Zhou, W.-S.; Xu, X.-X. Total Synthesis of the Antimalarial Sesquiterpene Peroxide Qinghaosu and Yingzhaosu A. *Acc. Chem. Res.* **1994**, *27*, 211−216.
(70) Zaman, S. S.; Sharma, R. P. Some Aspects of the Chemistry and Biological Activity of Artemisinin and Related Antimalarials. *Heterocycles* **1991**, *32*, 1593−1638.
(71) Jung, M. Current Developments in the Chemistry of Artemisinin and Related Compounds. *Curr. Med. Chem.* **1994**, *1*, 35−49.
(72) Avery, M. A.; Gao, F.; Mehrotra, S.; Chong, W. K.; Jennings-White, C. *The Organic and Medicinal Chemistry of Artemisinin and Analogues*. *Research Trends Trivandrum*; India, 1993.
(73) Posner, G. H.; Ploypradith, P.; Parker, M. H.; O'Dowd, H.; Woo, S.-H.; Northrop, J.; Krasavin, M.; Dolan, P.; Kensler, T. W.; Xie, S.; Shapiro, T. A. Antimalarial, Antiproliferative, and Antitumor Activities of Artemisinin-Derived, Chemically Robust, Trioxane Dimers. *J. Med. Chem.* **1999**, *42*, 4275−4280.
(74) Grigorov, M.; Weber, J.; Tronchet, J. M. J.; Jefford, C. W.; Milhous, W. K.; Maric D. A QSAR Study of the Antimalarial Activity of Some Synthetic 1,2,4-Trioxanes. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 124−130.
(75) Constants, P.; Carbó, R. Atomic shell approximation: electron density fitting algorithm restricting coefficients to positive values. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 1046−1053.
(76) Amat, L.; Carbó-Dorca, R. Quantum similarity measures under atomic shell approximation: first-order density fitting using elementary Jacobi rotations. *J. Comput. Chem.* **1997**, *18*, 2023−2039.
(77) Amat, L.; Carbó-Dorca, R. Fitted electronic density functions from H to Rn for use in quantum similarity measures: cis-diamminedichloroplatinum(II) complex as an application example. *J. Comput. Chem.* **1999**, *20*, 911−920.
(78) Amat, L.; Carbó-Dorca, R. Molecular Electronic Density Fitting Using Elementary Jacobi Rotations under Atomic Shell Approximation. *J. Chem. Inf. Comput. Sci.* in press.
(79) Gironés, X.; Robert, D.; Carbó-Dorca, R. TGSA: a molecular superposition program based on Topo-Geometrical Considerations. IT-IQC-00-11. *J. Comput. Chem.* in press.
(80) PC Spartan Pro, version 1.4; Wavefunction, Inc.: Irvine, 1997.
(81) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez C.; Pople, J. A. Gaussian-94, Revision E.2; Gaussian, Inc.: Pittsburgh, PA, 1995.
(82) Robert, D.; Gironés, X.; Carbó-Dorca, R. Quantification of the Influence of Single-Point Mutations on Haloalkane Dehalogenase Activity: A Molecular Quantum Similarity Study. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 839−46.
(83) Cox, T. F.; Cox, M. A. A. *Multidimensional Scaling*; Chapman & Hall: London, 1994.
(84) Carbó, R.; Besalú E. Nested Summation Symbols and Perturbation Theory. *J. Math. Chem.* **1993**, *13*, 331−342.
(85) Carbó, R.; Besalú, E. Definition, Mathematical Examples and Quantum Chemical Applications of Nested Summation Symbols and Logical Kronecker Deltas. *Comput. Chem*. **1994**, *18(2)*, 117−126.
(86) Carbó, R.; Besalú, E. Definition and Quantum Chemical Applications of Nested Summation Symbols and Logical Functions: Pedagogical Artificial Intelligence Devices for Formulae Writing, Sequential Programming and Parallel Implementation. *J. Math. Chem.* **1995**, *18*, 37−72.
(87) Carbó, R.; Besalú, E. Application of Nested Summation Symbols to Quantum Chemistry: Formalism and Programming Techniques. In *Strategies and Applications in Quantum Chemistry: from Astrophysics to Molecular Engineering*; Defranceschi, M., Ellinger, Y., Eds.; Kluwer Academic Publishers: Amsterdam, 1996; pp 229−248.
(88) Allen, D. M. The relationship between variable selection and data augmentation and a method for prediction. *Technometrics* **1974**, *16*, 125−127.
(89) Clementi, S.; Wold, S. How to choose the proper statistical method. In *Chemometric Methods in Molecular Design*; Van der Waterbeemd, H., Ed.; VCH: New York, 1995; pp 319−338.
(90) Wold, S.; Erikson, L. Statistical validation of QSAR results. In *Chemometric Methods in Molecular Design*; Van der Waterbeemd, H., Ed.; VCH: New York, 1995; pp 309−318.
(91) Posner, G. H.; Parker, M. H.; Northrop, J.; Elias, J. S.; Ploypradith, P.; Xie, S.; Shapiro, T. A. Orally Active, Hydrolytically Stable, Semisynthetic, Antimalarial Trioxanes in the Artemisinin Family. *J. Med. Chem.* **1999**, *42*, 300−304.
(92) Posner, G. H.; González, L.; Cumming, J. N.; Klinedinst, D.; Shapiro, T. A. Synthesis and Antimalarial Activity of Heteroatom-Containing Bicyclic Endoperoxides. *Tetrahedron* **1997**, *53*, 37−50.
(93) Avery, M. A.; Mehrotra, S.; Bonk, J. D.; Vroman J. A.; Goins, D. K.; Miller, R. Structure−Activity Relationships of the Antimalarial Agent Artemisinin. 4. Effect of Substitution at C-3. *J. Med. Chem.* **1996**, *39*, 2900−2906.
(94) Desjardins, R. E.; Canfield, C. J.; Haynes, D. E.; Chulay, J. D. Quantitative Assessment of Antimalarial Activity In Vitro by a Semi-automated Microdilution Technique. *Antimicrob. Agents Chemother.* **1979**, *16*, 710−718.
(95) Milhous, W. K.; Weatherly, N. F.; Bowdre, J. H.; Desjardins, R. E. In Vitro Activities of and Mechanisms of Resistance to Antifolate Antimalarial Drugs. *Antimicrob. Agents Chemother.* **1985**, *27*, 525−530.

CI0004558