Quantum-Chemical Foundations of the Topological Substructural Molecular Design

Ernesto Estrada*

Complex Systems Research Group, RIAIDT & Department of Organic Chemistry, Faculty of Pharmacy, Edificio CACTUS, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain

Received: February 5, 2008; Revised Manuscript Received: February 25, 2008

The topological substructural molecular design (TOPS-MODE) approach is formulated as a tight-binding quantum-chemical method. The approach is based on certain postulates that permit to express any molecular property as a function of the spectral moments of certain types of molecular and environment-dependent energies. We use several empirical potentials to account for these intrinsic and external molecular energies. We prove that any molecular property expressed in terms of a quantitative structure−property and structure−activity relationships (QSPR/QSAR) model developed by using the TOPS-MODE method can be expressed as a bond additivity function. In addition, such a property can also be expressed as a substructural cluster expansion function. The conditions for such bond contributions being transferable are also analyzed here. Several new statistical-mechanical electronic functions are introduced as well as a bond−bond thermal Green's function or a propagator accounting for the electronic hopping between pairs of bonds. All these new concepts are applied to the development and application of a new QSAR model for describing the toxicity of polyhalogenated-dibenzo-1,4-dioxins. The QSAR model obtained displays a significant robustness and predictability. It permits an easy structural interpretation of the structure−activity relationship in terms of bond additivity functions, which display some resemblances with other theoretical parameters obtained from first principle quantum-chemical methods.

Introduction

The study of quantitative structure−property and structure−activity relationships (QSPR/QSAR) has a well-deserved reputation in modern chemistry and biochemistry.[1,2] The strategy of QSPR/QSAR modeling is to condense in a mathematical expression the relationship between the structure of molecules and their properties. Physical chemistry plays a central role in this strategy by providing appropriate ways of characterizing the molecular structure and giving a physical support to the QSPR/QSAR models. An exciting new area of research in this field is to find ways of adjusting QSPR/QSAR to the mathematical structure of the physical chemistry methods used, such as the quantum similarity principles of QSPR and QSAR.[3–5] This paper will focus on this kind of research by providing the quantum-chemical formulation of a topological substructural approach to QSPR/QSAR.

The topological substructural molecular descriptors/design (TOPS-MODE) approach is a theoretical method that relates quantitatively the properties of molecules to their hydrophobicity, polar surface area, polarizability, molar refractivity, van der Waals radii, and electronic charges.[6] The QSPR/QSAR models obtained are expressible in terms of bond additivity permitting a clear interpretation of the property/activity in structural terms. This approach has been applied to several QSPR studies for different physicochemical properties of organic molecules[7,8] as well as for ADMET (administration, distribution, metabolism, excretion and toxicity) parameters of drugs and drug-like compounds.[9–11] The QSAR applications of TOPS-MODE include a large variety of biological activities, such as sedative/hypnotic, anticonvulsant, anticancer, antiinflamatory, herbicidal, antibacterial, central nervous system activity, antifungal, and anti-HIV, among others.[12–14]

Here, we reformulate the TOPS-MODE approach as a tight-binding quantum chemical approach. In such reformulation, we use several intrinsic molecular potentials accounting for the electrostatic, polarizability, and steric molecular properties as well as external potentials describing the interactions of a molecule with its environment. Several important features of this approach, such as bond additivity and transferability as well as the cluster expansion of properties, are formalized in the light of the tight-binding quantum chemical approach. The new formulation of TOPS-MODE permits to define several statistical-mechanical parameters which are directly connected to the topology of the molecules. Finally, we illustrate the applications of this approach in QSPR/QSAR by studying the toxicity of a series of polyhalogenated-dibenzo-1,4-dioxins.

Theoretical Formulation

The basic principle for developing quantitative structure−property or structure−activity relationships is that we can express this property/activity in terms of *linear free energy relationships* (LFERs).[15] In general, the property/activity is expressed in terms of functions describing the hydrophobic, steric, and electronic characteristics of the molecules.[15] Here, we reformulate this principle by stating it in terms of the molecular and environment-dependent energies.

Postulate 1. Any molecular property A can be expressed as a function of interaction energies $E(t)$,

$$A = f(E_{\text{intrinsic}}) + f(E_{\text{external}}) = \sum_t a_t E(t) + \alpha \tag{1}$$

where $f(E_{\text{intrinsic}})$ and $f(E_{\text{external}})$ are functions of the intrinsic molecular energy and an external energy depending on the molecular environment, respectively, α is the fitting error, and $E(t)$ are the different types of energies involved in the $f(E_{\text{intrinsic}})$ and $f(E_{\text{external}})$ functions.

---

* E-mail: estrada66@yahoo.com. Fax: +34-981-547077.

10.1021/jp8010712 CCC: $40.75

The following step in the current approach consists in connecting the different types of energies and the density of states. It is known that any type of energy $E(t)$ can be expressed as[16]

$$E(t) = \int_{-\infty}^{\infty} \varepsilon(t)\Gamma[\varepsilon(t)]\, \mathrm{d}\varepsilon(t) \tag{2}$$

where $\Gamma[\varepsilon(t)]$ is a density function and $\varepsilon(t)$ are the discrete values of the energy levels for energy of type t. Then, we assume the existence of the following approximation for the density of states.

Postulate 2. There is always an approximate solution $\hat{\Gamma}[\varepsilon(t)]$ to the true solution $\Gamma[\varepsilon(t)]$ through the relationship[17]

$$\hat{\Gamma}[\varepsilon(t)] = \int \hat{\delta}[\varepsilon'(t), \varepsilon(t)]\Gamma[\varepsilon'(t)]\, \mathrm{d}\varepsilon'(t) \tag{3}$$

where $\hat{\delta}[\varepsilon'(t), \varepsilon(t)]$ is the resolution function, the averaging kernel.

Now, we define the spectral moments of the density function, which are given by[18,19]

$$\mu_k[E(t)] = \int_{-\infty}^{\infty} \varepsilon(t)^k \Gamma[\varepsilon(t)]\, \mathrm{d}\varepsilon(t) \tag{4}$$

where the spectral density function is defined as usual as

$$\Gamma[\varepsilon(t)] = \sum_{j=1}^{n} \delta[\varepsilon(t) - \varepsilon_j(t)] \tag{5}$$

with $\delta(x)$ being the Dirac delta function. The values of $\varepsilon_j(t)$ are the eigenvalues of the Hamiltonian operator $\mathbf{H}(t)$ for a molecule, which will be defined later on. The kth spectral moment of $\mathbf{H}(t)$ can be defined as[18,19]

$$\mu_k[E(t)] = Tr\{\mathbf{H}^k(t)\} = \sum_i \langle i|\mathbf{H}^k(t)|i\rangle = \sum_j [\varepsilon_j(t)]^k \tag{6}$$

Then we can express the function $\hat{\Gamma}[\varepsilon(t)]$ as a linear combination of the moments $\mu_k[E(t)]$,[17]

$$\hat{\Gamma}[\varepsilon(t)] = \sum_i c_i[\varepsilon(t)]\mu_i[E(t)] \tag{7}$$

Now using the expressions 2 and 7, we can consider that the energy can be expressed as a linear combination of its spectral moments,

$$E(t) \approx \int_{-\infty}^{\infty} \varepsilon(t)\hat{\Gamma}[\varepsilon(t)]\, \mathrm{d}\varepsilon(t) = \int_{-\infty}^{\infty} \varepsilon(t)\sum_i c_i[\varepsilon(t)]\mu_i[E(t)]\, \mathrm{d}\varepsilon(t) = \sum_i c_i[\varepsilon(t)]\mu_i[E(t)] \tag{8}$$

We can now use the first postulate in order to express any property as a function of the spectral moments of the different energies,

$$A = \sum_t a_t E(t) = \sum_t a_t \sum_i c_i[\varepsilon(t)]\mu_i[E(t)] + \alpha = \sum_t \sum_i b_i(t)\mu_i[E(t)] + \alpha \tag{9}$$

where $b_i(t)$ is an empirical coefficient multiplying the ith spectral moment of the energy of type t. The assumption that any property/activity can be, in principle, expressed in terms of the spectral moments of certain density functions is the basis of the approach known as TOPS-MODE.[6] The expression (9) states that *any molecular property can be expressed as a function of the spectral moments of certain types of molecular and environment-dependent energies.*

Quantum-Chemical Formulation. Here, we start by considering a united-atom approach to molecular representation. We follow an approach introduced by Franklin,[20] in which each carbon atom and its associated hydrogens are treated as a united atom. If we apply this united-atom approach to methane we can see that it is isoelectronic to neon, with electronic configuration $1s^2$, $2s^2$, $2p_x{}^2 2p_y{}^2 2p_z{}^2$. Similarly, ethane is isoelectronic to F_2, propane to FOF, isobutane to NF_3, neopentane to CF_4, etc.[20] Then, the united-atom approach is equivalent to the use of the so-called "hydrogen-depleted graphs".

At this point, we assume that the wave functions are built up from "pseudobond" orbitals of the system. A typical bond orbital is an orbital centered at the bonds of the system.[21] The use of such orbitals was the basis for the creation of the linear combination of bond orbitals to give molecular orbitals (LCBO-MO) method.[22] However, here we consider orbitals which are centered at two united-atoms,[20] e.g., centered at the line marking the following two united atoms $(CH_3)-(CH_3)$, which is equivalent to be centered at the isoelectronic F−F bond. In this method, the molecular orbitals are created as linear combinations of the pseudobond orbitals.[20]

In order to obtain the discrete values of the energies $\varepsilon(t)$ we solve the secular equations,

$$\sum_{b=1}^{m} c_j(b,t)[H_{jb}(t) - pS_{jb}(t)] = 0,\ j = 1 \ldots m \tag{10}$$

where the coefficients $c_j(b,t)$ arise from the linear combination of pseudobond orbitals, φ_r, centered at the pseudobond b to give molecular orbitals,[20]

$$\psi_j(t) = \sum_{b=1}^{m} c_j(b,t)\varphi_b(t) \tag{11}$$

The summations in (10) and (11) are taken over all m pseudobonds of the molecule. The terms $H_{jb}(t)$ in (10) are the entries of the matrix $\mathbf{H}(t)$, and $\varepsilon_j(t)$ values correspond to the values of the energies of type t for the molecular orbitals. The nontrivial solutions of eq 10 are obtained by solving the determinant equation

$$|\mathbf{H}(t) - \varepsilon(t)\mathbf{S}(t)| = 0 \tag{12}$$

We assume that the Coulomb integral $H_{ii}(t)$ of a pseudobond orbital $\varphi_i(t)$ depends only on the type of the pseudobond in which it resides. The resonance integral $H_{ij}(t)$ between pseudobond orbitals $\varphi_i(t)$ and $\varphi_j(t)$ is assumed to be zero, unless i and j are adjacent pseudobonds, in which case it is taken to be $H_{ij}(t) = q(t)$. We consider that the pseudobond orbitals are orthonormal, thus $S_{ij} = \delta_{ij}$. The Coulomb integrals are set equal to $H_{ii}(t) = p(t) - V_i(t)q(t)$, where $V_i(t)$ is a potential function to be defined later for the specific types of energies. Then if we divide every entry of the secular determinant by $q(t)$, the secular determinant for a molecule having m bonds is written as

$$\begin{vmatrix} \varepsilon(t)-V_1(t) & \delta_{1,2} & \delta_{1,3} & \cdots & \delta_{1,m} \\ \delta_{2,1} & \varepsilon(t)-V_2(t) & \delta_{2,3} & \cdots & \delta_{2,m} \\ \delta_{3,1} & \delta_{3,2} & \varepsilon(t)-V_3(t) & \cdots & \delta_{3,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \delta_{m,1} & \delta_{m,2} & \delta_{m,3} & \cdots & \varepsilon(t)-V_m(t) \end{vmatrix} = 0 \tag{13}$$

where δ_{ij} is equal to one if, and only if, the corresponding bonds are adjacent and $\varepsilon_j(t) = \{p(t) - E(t)\}/q(t)$ are the eigenvalues of the matrix $\mathbf{H}(t) = \mathbf{B} - \mathbf{V}(t)$ and $c_j(b, t)$ are the entries corresponding to the bond m of the eigenvector associated with this eigenvalue. The matrix B is the bond−bond adjacency matrix of the hydrogen-depleted graph representing the molecular skeleton. In this context, the total energy is defined as twice the sum of the energies of the molecular orbitals,

$$E(t) = 2\sum_{j=1}^{m} \varepsilon_j(t) \tag{14}$$

Definition of the Potentials. The next step is to extend the approach developed in the previous section to consider any organic molecule. Then, we consider a similar united-atom approach for the different groups having hydrogen atoms, such as NH_2 and OH, which are isoelectronic with CH_3 and F; NH which is isoelectronic with CH_2 and O; SH which is isoelectronic with Cl; and so forth. A similar approach was already introduced by Franklin.[20] As a consequence, different molecules like isobutane and 2-propanol are isoelectronic with NF_3. This situation makes it strictly necessary to use different pseudobond potentials for distinguishing the different pseudobonds in isoelectronic molecules.

In order to account for intrinsic and intermolecular interactions of a molecule, we consider the following functions,

$$f(E_{\text{intrinsic}}) = f[E(P)] + f[E(L)] + f[E(S)] + f[E(\text{PS})] \tag{15a}$$

$$f(E_{\text{external}}) = f[E(H)] + f[E(\text{MR})] \tag{15b}$$

where P stands for the polarizability, L for the electrostatic, S for the steric, PS for the polar surface, H for the hydrophobic, and MR for the molar refractivity. There are several possible ways of selecting the potentials for the different energy functions, which range from first principle to empirical methods. For the molecular energy functions, we select pseudobond potentials based on atomic polarizability (P), on Gasteiger−Marsili atomic charges (L), van der Waals radii of atoms (S), and atomic polar surface areas (PS).[6] Atomic polarizabilities were obtained from the work of Miller,[23] who calculated them from an additive function based on atomic hybridization components for each atom using experimental values of molecular polarizability. The atomic polar surface areas were obtained using the method of Ertl,[24] who calculated topological polar surface areas as the sum of atomic contributions. The values of the Gasteiger−Marsili atomic charges[25] are calculated for every atom in the molecule using the standard procedure, and the values of van der Waals radii of atoms were taken from Bondi.[26]

In the case of the external molecular functions we use two empirical functions based on the partition coefficient *n*-octanol/water (log P) and on molar refractivity (MR). It is known that log P can be decomposed into a volume term expressing mainly hydrophobic and dispersive solute−solvent interactions, and polar terms expressing electrostatic solute−solvent interactions.[27] Here, we use the atomic contributions to the experimental values of log P determined by Wang et al.[28] For the MR-based potential, we use the atomic contributions to the molar refractivity obtained by Ghose and Crippen.[29] The molar refractivity is calculated from the experimental values of the density and the refractivity index of a compound. Then, MR accounts for the volume occupied by a molecule as influenced by the rest of the molecules surrounding it. Consequently, both log P and MR characterize the interactions of a solute molecule with its environment and can be used for $f(E_{\text{external}})$.

In all cases, the atomic contributions p_i and p_j for the bonded atoms i,j are transformed into a pseudobond contribution p_{ij} as follows

$$p_{ij} = \frac{p_i}{k_i} + \frac{p_j}{k_j} \tag{16}$$

where k_i is the number of pseudobonds incident with the atom i.

Using the different types of potentials mentioned in the previous paragraphs expression 9 can be rewritten in order to express the molecular property A in terms of the spectral moments of the different types of intrinsic and external molecular energies, which is the basic principle of the TOPS-MODE approach,[6]

$$A = \sum_i b_i(L)\mu_i[E(L)] + \sum_i b_i(P)\mu_i[E(P)] + \sum_i b_i(S)\mu_i[E(S)] + \sum_i b_i(PS)\mu_i[E(\text{PS})] + \sum_i b_i(H)\mu_i[E(\text{PS})] + \sum_i b_i(\text{MR})\mu_i[E(\text{MR})] + \alpha \tag{17}$$

Bond Additivity of Molecular Properties. An important characteristic of the current approach is that any property which is expressed in terms of spectral moments of the density function can be expressed as a linear combination of bond contributions. This property is known as the bond (or group) additivity[30,31] of properties and it is proved in the following:

Theorem. Let A be any molecular property, which according to (1) can be expressed as a function of interaction energies. Then A can be expressed as an additivity function of bond contributions $A(b)$.

Proof. We start by considering any property which is expressed in terms of the spectral moments of different types of molecular energies, here designated as t_i

$$A = \sum_k b_k(t_1)\mu_k[E(t_1)] + \sum_k b_k(t_2)\mu_k[E(t_2)] + \cdots + \sum_k b_k(t_n)\mu_k[E(t_n)] + \alpha \tag{18}$$

The spectral moments of any energy type can be expressed in terms of local moments for the bonds of the molecule

$$\mu_k[E(t)] = \sum_{b=1}^{m} \mu_k[E(t), b] \tag{19}$$

where the bond moments are expressed in terms of the molecular orbitals and energies as follows,

$$\mu_k[E(t), b] = \sum_{j=1}^{m} [c_j(b, t)]^2 \varepsilon_j^k(t) \tag{20}$$

Then, substituting (19) in (18), we obtain the property under study in terms of the bond moments of different energy types,

$$A = \sum_k \sum_{b=1}^{m} b_k(t_1)\mu_k[E(t_1), b] + \sum_k \sum_{b=1}^{m} b_k(t_1)\mu_k[E(t_2), b] + \cdots + \sum_k \sum_{b=1}^{m} b_k(t_n)\mu_k[E(t_n), b] + \alpha \tag{21}$$

Now we can sum all the local moments of different types and order corresponding to a given bond b multiplied by their respective coefficients,

$$A(b) = \sum_k b_k(t_1)\mu_k[E(t_1), b] + \sum_k b_k(t_2)\mu_k[E(t_2), b] + \cdots + \sum_k b_k(t_n)\mu_k[E(t_n), b] \quad (22)$$

We call this term $A(b)$ the contribution of the bond b to the property A. Consequently, we can express the global property as a sum of bond contributions, which proves the theorem,

$$A = \sum_{b=1}^{m} A(b) + \alpha \quad (23)$$

Bond Transferability. The bond transferability[30] is based on the assumption that a molecular property associated with a bond b in a molecule has a similar, but not identical, value in another molecule. Then, if we call $A_1(b)$ the contribution of the bond b to the property A in the molecule 1, the contribution of the bond b to the property A in the molecule 2 should be $A_2(b) \approx A_1(b)$. The question in which we are interested here is in what extension a bond contribution $A(b)$ obtained by using the TOPS-MODE approach is transferable.

For the sake of simplicity, we consider here a bond contribution obtained as a linear combination of the spectral moments of a bond matrix representing an alkane. The results obtained for this bond contribution can be easily generalized to a bond property obtained by a more general expression like (21).

Let $V(t) = 0$ in $\mathbf{H}(t) = \mathbf{B} - \mathbf{V}(t)$, so that the bond contribution to the property is expressed as follows,

$$A(b) = \sum_k b_k \mu_k(b) \quad (24)$$

The kth spectral moment $\mu_k(b)$ of the bond b represents the number of closed walks of length k that start and ends at the bond b in the molecule. Closed walks, also known as self-returning walks, have been used in different contexts of theoretical chemistry.[32] The length of a closed walk refers to the number of steps that a particle needs to give for moving from one site to another. For instance, a particle starting at bond i visiting an adjacent bond j and then returning to i completes a closed walk of length $k = 2$. This closed walk of length two can be represented by a fragment formed by two adjacent bonds, here designated as $F_2(b)$. Then, the spectral moment $\mu_2(b)$ is simply the number of fragments $F_2(b)$ in the alkane, or in other words the number of bonds adjacent to the bond b.

In general, any spectral moment can be expressed as a linear combination of structural fragments of the molecule. For instance, the first five spectral moments for a bond b in an alkane are given below, where the symbol $|\cdots|$ represents the number of fragments, and the different fragments $F_i(b)$ are illustrated in Figure 1, where i represents an arbitrary numbering given to fragments in the Figure 1,

$$\mu_2(b) = |F_2(b)| \quad (25)$$

$$\mu_3(b) = 2|F_3(b)| \quad (26)$$

$$\mu_4(b) = |F_2(b)| + 4|F_3(b)| + 6|F_4(b)| + 5|F_5^1(b)| + 2|F_5^2(b)| \quad (27)$$

$$\mu_5(b) = 10|F_3(b)| + 30|F_4(b)| + 2|F_6^1(b)| + 4|F_6^2(b)| + 2|F_6^3(b)| \quad (28)$$

If we substitute the expressions for the spectral moments in terms of fragments in the eq 22 we obtain,

$$A(b) = \sum_i g_i F_i(b) \quad (29)$$

This means that the contribution of the bond b to the property A depends on the environment in which this bond is located. The dependence of $A(b)$ on the molecular environment increases with the value of k in expression 24. Then, if the bond contribution $A(b)$ is expressed in terms of large spectral moments, i.e., large value of k, then $A(b)$ will be included in very large fragments. As a consequence, $A(b)$ is not transferable from one molecule to another.

In summary, bond transferability is very much affected by the order of the spectral moments included in the model describing the property under study. Low order spectral moments favor transferability while higher order moments make the property less transferable.

Cluster Expansion of Properties. The substructural cluster expansion[33] of a property A can be seen as a generalization of the bond additivity of the property. Instead of expressing the property A as a function of bond contributions in the cluster expansion the property is expanded as an additivity function of fragment contributions. Formally, the cluster expansion is expressed as,

$$A = \sum_i f_i(A) F_i \quad (30)$$

where F_i is a molecular fragment and $f_i(A)$ is the corresponding contribution of such fragment to the property A.

The kth spectral moment of the Hamiltonian matrix represents a weighted closed walk of length k between the bonds of the molecule, where the weight associated with the walk is the product of the $\mathbf{H}(t)_{1a,ib}$ interaction elements,[34]

$$\mu_k(t) = \sum_{i1,i2\ldots ik} \mathbf{H}(t)_{i1,i2} \mathbf{H}(t)_{i2,i3} \cdots \mathbf{H}(t)_{ik,i1} \quad (31)$$

These weighted closed walks can be represented pictorially. Then, every weighted closed walk starting (and ending) at a given bond can be associated with a fragment. Consequently, every spectral moment can be expressed as a linear combination of molecular fragments. Thus, by substituting these expressions of spectral moments in terms of molecular fragments into (17) we can obtain a cluster expansion for the property A.

Statistical-Mechanical Molecular Parameters. Another important characteristic of the new formulation of the TOPS-MODE approach is that we can define several statistical-mechanical parameters for a molecule based on the different

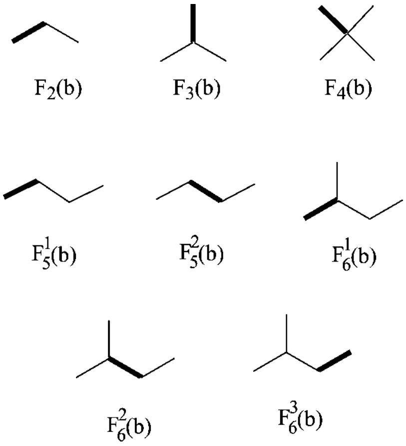


Figure 1. Molecular fragments contributing to the first spectral moments of a Hamiltonian matrix according to the eqs 25–28 where the bold bond corresponds to the bond designated as b.

types of energies previously defined. First, we define an electronic partition function for the different types of energies previously defined, which for the sake of simplicity in the notation will be designated simply by *t*,

$$Z(t) = \sum_{j=1}^{m} \mathrm{e}^{-\beta\varepsilon_j(t)} \tag{32}$$

where $\beta = 1/kT$ is the inverse temperature (T) and k is the Boltzmann constant.[35] Using this partition function, we can define the entropy of the electronic distribution as

$$S(t) = -k\sum_{j=1}^{m} p_j(t)\ \ln p_j(t) = k\sum_{j=1}^{m} p_j(t)[\beta\varepsilon_j(t) + \ln Z(t)] \tag{33}$$

where

$$p_j(t) = \frac{\mathrm{e}^{-\beta\varepsilon_j(t)}}{Z(t)} \tag{34}$$

is the probability of finding the system in the state having energy $\varepsilon(t)$. We can rewrite (33) in the following way

$$S(t) = k\beta\sum_{j=1}^{m} \varepsilon_j(t)p_j(t) + k\ \ln Z(t)\sum_{j=1}^{m} p_j(t) \tag{35}$$

Then, by multiplying by T and reordering the terms, we obtain

$$-\frac{1}{\beta}\ln Z(t) = \sum_{j=1}^{m} \varepsilon_j(t)p_j(t) - TS(t) \tag{36}$$

which by using the known expression $F = H - TS$ permits to identify the enthalpy H and the free energy F of the electronic distribution,

$$H(t) = \sum_{j=1}^{m} \varepsilon_j(t)p_j(t) \tag{37}$$

$$F(t) = -\frac{1}{\beta}\ln Z(t) \tag{38}$$

Electron Propagation. Another interesting question that can be answered under the current theoretical scheme is to know how an impact propagates from one place to another in a molecule considering the different potentials studied here. In this sense, we can consider particle hopping on the molecule in such a way that we can describe how many particles end up at the *q*th bond if we put particles at the *p*th bond of the molecule. The way of measuring such propagation from one bond to another is by considering the thermal Green′s function or the propagator, $G(p,\ q;\ \beta,\ t)$, which is defined as,[36]

$$G(p, q; \beta, t) = \langle p|\mathrm{e}^{-\beta\mathbf{H}(t)}|q\rangle \tag{39}$$

The thermal Green′s function can be expressed in terms of the wave functions for the molecular orbitals with given energy, such that

$$G(p, q; \beta, t) = \sum_{j=1}^{m} c_j(p, t)c_j(q, t)\mathrm{e}^{-\beta\varepsilon_j(t)} \tag{40}$$

where $c_j(p,\ t)$ is the *p*th entry of the *j*th eigenvector of the Hamiltonian matrix $\mathbf{H}(t)$, which has energy $\varepsilon_j(t)$.

Topological Connection. In order to establish a connection between the statistical-mechanical parameters previously defined and the topological roots of TOPS-MODE, we will make use of the spectral moments of the Hamiltonian matrices used in this method. First, we rewrite the partition function (32) as the trace of the exponential of the Hamiltonian matrix, which can be immediately interpreted in term of the spectral moments of the corresponding Hamiltonian,

$$Z(t) = tr[\mathrm{e}^{-\beta\mathbf{H}(t)}] = \sum_{k=0}^{\infty} (-\beta)^k \frac{tr\mathbf{H}(t)^k}{k!} = \sum_{k=0}^{\infty} (-\beta)^k \frac{\mu_k(t)}{k!} \tag{41}$$

Then, we can group together the even- and odd-order spectral moments of the Hamiltonian to give the following expression,

$$Z(t) = \sum_{k=0}^{\infty} (\beta)^{2k}\frac{\mu_{2k}(t)}{(2k)!} - \sum_{k=0}^{\infty} (\beta)^{2k+1}\frac{\mu_{2k+1}(t)}{(2k+1)!} \tag{42}$$

We recall that the spectral moment $\mu_k(t)$ represents the number of weighted closed walks of length k that exist in the molecule. Hence, the first term of the right-hand part of eq 42 represents the number of weighted closed walks of even length in the molecule. The second term represents the number of weighted odd-closed walks. Here, the weighting scheme is carried out in two different ways. First, every bond in the molecule is weighted by the corresponding potential of type *t* according to (31). Second, the spectral moments are weighted in decreasing order of their lengths by dividing them by a factorial term. In this way, the shortest closed walks receive more weight than the longer ones.

In summary, the expression (42) indicates that the partition function is simply the difference between the number of weighted closed walks of even length minus the number of weighted closed walks of odd length. This expression can be simply written in terms of hyperbolic functions as follows,

$$Z(t) = \sum_{j=1}^{m} \cosh[\varepsilon_j(t)] - \sum_{j=1}^{m} \sinh[\varepsilon_j(t)] \tag{43}$$

The first term of the right-hand part of (43) is known as the even subgraph centrality and the left-hand part is the odd subgraph centrality of a graph or network.[37] Their sum corresponds to the subgraph centrality or Estrada index of the graph or network.[38]

In addition, the partition function $Z(t)$ can be expressed as a sum of bond contributions

$$Z(t) = \sum_{i=1}^{m} Z_i(t) \tag{44}$$

where

$$Z(i, t) = \sum_{k=0}^{\infty} (-\beta)^k \frac{\mu_k(i, t)}{k!}$$

Then, we can express $Z(i,\ t)$ in terms of the eigenfunctions and energies of the Hamiltonian matrix $\mathbf{H}(t)$,

$$Z(i, t) = \sum_{j=1}^{m} [c_j(i, t)]^2\mathrm{e}^{-\beta\varepsilon_j(t)} \tag{45}$$

where $c_j(i,\ t)$ is the *i*th entry of the *j*th eigenvector of the Hamiltonian $\mathbf{H}(t)$, which has energy $\varepsilon_j(t)$.

Computational Results

Here, we will illustrate the use of the previous theoretical developments by analyzing the toxicity of polyhalogenated dibenzo-1,4-dioxins (PHDD). These compounds have attracted an intense public and scientific scrutiny in the past few years because of their widespread presence in the environment and a large number of toxic effects produced by them.[39] Among such toxic effects it can be mentioned the hepatotoxicity, porphyria, endocrine effects, immunotoxicity, teratogenicity, and carcinogenecity, among others.[39] PHDDs have been the target of several QSAR studies using a large variety of theoretical approaches.[40]

We study here a data set of 25 polychlorinated, polybrominated, and polychlorinated-brominated dibenzo-1,4-dioxins which were originally reported by So and Karplus,[41] where a QSAR model was reported for the Ah receptor (AhR) binding affinities (pEC_{50}). This data set was recently used by Zheng et al.[42] in a QSAR study using a radial basis neural network. In this study, the authors used a large series of quantum chemical descriptors, such as the final heat of formation, total energy, electronic energy, core−core repulsion energy, dipole moment, orbital energies, linear polarizabilities, hyper-polarizability, electrostatic potential on each atom, and net atomic charges, among others. The best model obtained was using a radial basis neural network (RBNN) with a squared cross-validation correlation coefficient $q^2 = 0.8818$, a predictive residual error sum of squares PRESS = 5.9215 and a root of the mean square error of the cross-validation rmse = 0.4867. With these values of the cross-validation parameters, this QSAR has been claimed as "superior to any other reported QSAR model based on the same activity data set".[42] The final model, however, uses nine independent variables in a highly nonlinear model obtained with an RBNN.

In a previous work, Fraschini et al.[43] determined that the molecular polarizability plays an important role in the origins of PHDDs specific binding to their receptor proteins. Then, for the sake of simplicity, we calculate only the spectral moments of the polarizability Hamiltonian for the 25 PHDDs studied. By using linear regression analysis, we obtained the QSAR model illustrated below:

$$pEC_{50} = 39.03503\mu_0 - 6.2778\mu_2(P) + 3.5028.10^{-2}\mu_4(P) + 3.9752.10^{-6}\mu_{11}(P) - 8.8386.10^{-8}\mu_{13}(P) + 59.4319$$

$$N = 25 \quad r^2 = 0.931 \quad s = 0.427 \quad F(5, 19) = 51.1$$

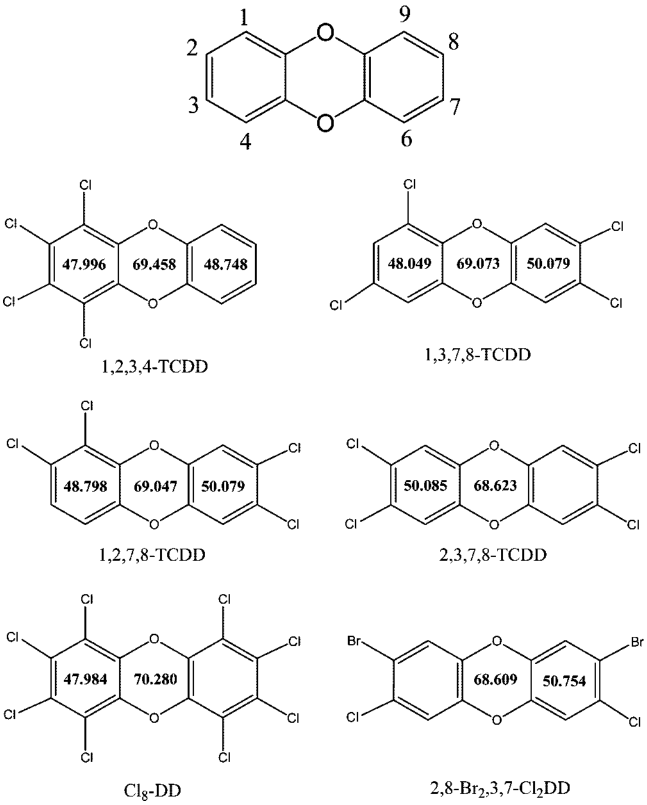
$$q^2 = 0.921 \quad PRESS = 5.8226 \quad RMSE = 0.4826 \quad (46)$$

This model explains 4% more of the cross-validation variance in the experimental toxicity that the model previously reported by Zheng et al.[42] The current model also improves the values of the cross-validation errors, PRESS and RMSE. The cross-validation experiment was carried out by the leave-one-out approach in which one compound is leave out the training set, a model then created and the toxicity of such compounds is predicted with this model. The process is repeated for all the compounds in the data set, which simulates the prediction of the study property for compounds not in the data set. However, from a QSAR perspective, the most important advantages of the current model are that it contains only 5 independent variables (instead of nine in the Zheng et al.[42] model) and that the model permits a clear physical interpretation of the property studied. The toxicity values are given in the Table 1.

To start with the physical interpretation of the QSAR model obtained, we express the model (46) as a linear combination of bond contributions. As a consequence, we obtain a bond additivity expression for the toxicity of PHDDs. In Figure 2, we illustrate six of the compounds studied in the data set. Here, we have grouped the bond contributions into three molecular regions for the clarity of the analysis. The region I corresponds to the phenyl ring at the left of the structure plus the carbon−halogen bonds attached to it. Region II corresponds to the four central carbon−oxygen bonds, and region III is formed by the phenyl ring to the right together with the halogen attached to it. The values reported in Figure 2 are the sum of the contributions of the bonds forming each region.

TABLE 1: Observed and Predicted Values of the Toxicity of Polyhalogenated Dibenzo-1,4-dioxins Using the Best QSAR Model Obtained, Equation 46

name	pEC_{50} observed	pEC_{50} predicted
1-Cl	4.000	4.688
2,8-Cl_2	5.495	5.849
1,2,4-Cl_3	4.886	4.715
1,7,8-Cl_3	6.658	6.050
2,3,7-Cl_3	7.149	6.790
1,2,3,4-Cl_4	5.886	5.150
1,2,7,8-Cl_4	6.796	6.863
1,3,7,8-Cl_4	6.102	6.134
2,3,7,8-Cl_4	8.000	7.731
1,2,3,4,7-Cl_5	5.194	5.570
1,2,3,7,8-Cl_5	7.102	7.186
1,2,4,7,8-Cl_5	5.959	6.077
1,2,3,4,7,8-Cl_6	6.553	6.511
Cl_8	5.000	5.295
2,3,7,8-Br_4	8.824	9.569
2,3-Br_2,7,8-Cl_2	8.830	8.650
2,8-Br_2,3,7-Cl_2	9.350	9.122
2-Br,3,7,8-Cl_3	7.939	8.426
1,3,7,9-Br_4	7.032	7.126
1,3,7,8-Br_4	8.699	8.347
1,2,4,7,8-Br_5	7.770	7.753
1,2,3,7,8-Br_5	8.180	7.961
2,3,7-Br_3	8.932	8.741
2,7-Br_2	7.810	7.912
2-Br	6.530	6.46



1,2,3,4-TCDD

1,3,7,8-TCDD

1,2,7,8-TCDD

2,3,7,8-TCDD

Cl_8-DD

2,8-Br_2,3,7-Cl_2DD

Figure 2. Contributions of three molecular regions (see main text) to the toxicity of some polyhalogenated dibenzo-1,4-dioxins. The contributions are calculated as the sum of bond contributions for the bond corresponding to the specific regions.

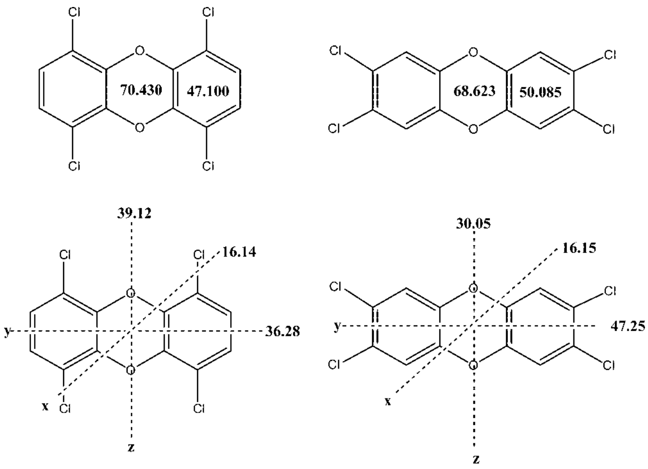


Figure 3. Comparison of the contributions of molecular regions to the toxicity and the polarizability tensor (in SI units, 10^{40} C m^2 V^{-1}) for 1,4,6,9-and 2,3,7,8-tetrachlorodibenzo-*p*-dioxins along the principal axes of polarizability, which originate in the centre of the molecules.

As can be seen for the tetra-chloro-DDs (first four structures) the substitution at positions 2,3 or 7,8 produces the largest contribution of the regions I and III, respectively. Then, the largest contribution for these regions is observed for the most toxic of the four isomers, i.e., 2,3,7,8-tetrachloro-1,4-dibenzodioxin. The second largest contribution for the region of type I is obtained for the substitution at positions 1,2 and the lowest contribution is obtained for the 1,3 substitution. This trend coincides with the toxicity order of these compounds. The substitution 1,2,3,4 produces the lowest contribution to region I for the four isomers, coinciding with the fact that this compound is the least toxic of the four. The contribution of the region II follows a reverse order. That is, the lowest contribution for this region is obtained for the 2,3,7,8 isomer and the largest one for the 1,2,3,4 one. In order to understand these trends we have made the calculation of bond contribution to the toxicity of 1,4,6,9-tetrachloro-1,4-dibenzodioxin. The polarizability tensor for this compound was studied by Fraschini et al.[43] together with that of 2,3,7,8-tetrachloro-1,4-dibenzodioxin. In Figure 3, we illustrate the contributions of the three molecular regions studied in the current work and the values reported by Fraschini et al.[43] of the polarizability at the three principal axes, which originate in the center of the molecule. As can be seen, the polarizability along the axis *y* is significantly larger for the 2,3,7,8 isomer. However, the polarizability is larger along the axis *z* for the 1,4,6,9 isomer. Then, the contribution of the region II to the toxicity might be interpreted as produced mainly by the polarizability along the axis *z*, while the contributions of regions I and III might be produced by the polarizability along the axis *y*.

If we return to the Figure 2, we can observe that the contribution of the region II in the octachloro-DD is almost the same as for the 1,4,6,9-DD, indicating a clear predominance of the effects produced by the chlorines at positions 1,4 and 6,9. In addition, we have included the most toxic compound in the data set studied, i.e., 2,8-dibromo-3,7-dichloro-DD, which indeed has the largest contribution of regions I and III of all the compounds shown in this figure.

In closing, we can say that halogen substitution in DDs has two different effects, which depend on the positions where the substitutions take place. The substitutions along the axis *y* increase the toxic contribution of regions I and III, while substitutions along the axis *z* increases the toxicity of region II. In general, it appears that the toxicity of regions I and III, i.e., toxicity along the axis *y*, is determinant for the global toxicity of these compounds.

Another route to the expression of the toxicity of PHDD in terms of molecular parameters is by considering the statistical mechanic parameters defined previously in this work. We have obtained a good model that expresses pEC_{50} as a function of the average polarizability partition function,

$$\begin{aligned}\mathrm{pEC}_{50} &= 86.423 - 127.279\bar{Z}(P)\\ N = 25 \quad r^2 &= 802 \quad s = 0.656 \quad F(1,23) = 93.3\end{aligned} \tag{47}$$

Then using the relationship between the partition function and the expected value of the energy $\langle E(P)\rangle$,

$$Z(P) = \frac{\sum_{m=1}^{m} \varepsilon_j(P) e^{-\varepsilon_j(P)}}{\langle E(P)\rangle} \tag{48}$$

the QSAR model (47) relates directly the toxicity of PHDDs with the orbital energies and the expected energy based on a polarizability potential. Now, by making use of (43) we can express the toxicity in terms of the hyperbolic functions of the orbital energies,

$$\mathrm{pEC}_{50} = 86.423 - \frac{127.279}{m}\sum_{j=1}^{m} \cosh[\varepsilon_j(P)] + \frac{127.279}{m}\sum_{j=1}^{m} \sinh[\varepsilon_j(P)] \tag{49}$$

which immediately suggests negative contributions of the even closed walks and positive contributions of the odd closed walks of an electron in the PHDD.

Now, by using (44), we can express $Z(P)$ in terms of bond contributions, which can be calculated by means of (45). These numbers, $Z_{pp}(P) = Z_p(P)$, represent the weighted number of closed orbits for an electron starting (and ending) at bond p. Then, we can plot these values together with the values of $G(p, q; \beta = 1, P)$ (Green's function) representing the electronic jumps from bond p to bond q under the polarizability potential. The plots are built as contour plots by using the bond numbers as the x and y coordinates and the values of $G(p, q; \beta = 1, P)$ and $Z_p(P)$ as the z coordinate. These plots are given in Figure 4 for the four tetrachloro-DDs previously studied, where the main diagonal entries represent the values of $Z_p(P)$ and the nondiagonal ones represent $G(p, q; P)$. The structures of these compounds as well as their bond numbering are given in the Figure 5.

As can be seen, the largest electron hopping takes place among the bonds located in any of the three regions previously demarked. If we take the average values of $G(p, q; P)$ for the bonds in regions I, II, and III, we observe that the central region has a value of 6.36 for all compounds in Figure 3. In the case of bisubstituted phenyl rings, the largest average value (in parenthesis) of Green′s function is obtained for the 1,2-dichloro (11.03), followed by the 2,3-dichloro (10.994) and then the 1,3-dichloro (10.392). This pattern reflects the fact that the electron hopping is limited when chlorine atoms are at 1,3 position, which might be due to the lack of the resonance or mesomeric effect between these two groups at meta position to each other. The average values of $Z_p(P)$ for the three regions previously defined follow the same trend of the toxicity of these compounds. For instance, the largest average value for bisubstituted

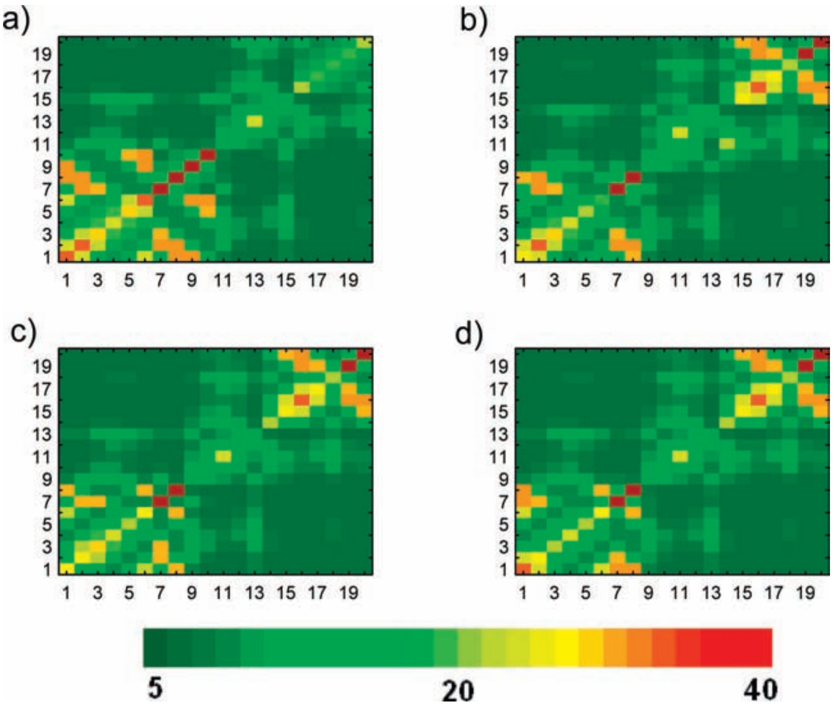


Figure 4. Plot of the values of the bond contributions to the polarizability partition function (diagonal entries) and the Green's functions for the electron hopping from one bond to another (nondiagonal entries) for the different bonds in four tetrachlorodibenzo-1,4-dioxins. The structures and bond numbering are given in Figure 5. Parts a, b, c, and d correspond to structures a, b, c and d in Figure 5, respectively.

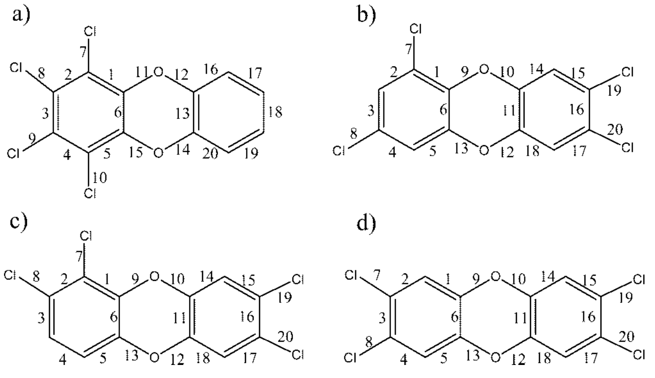


Figure 5. Structures and bond numbering of the four tetrachlorodibenzo-1,4-dioxins represented in Figure 4.

phenyl rings is for the 2,3 (25.63), followed by the 1,3 (25.60) and finally the 1,2 (25.52).

In summary, the model (46) obtained by using the tight-binding quantum-chemical TOPS-MODE approach is not only a high-quality QSAR model but also an easily interpretable one in terms of the structural influence on the property/activity studied. The quantum-chemical formulation of the TOPS-MODE permits this interpretation in a wider context. Such a context combines the more intuitive topological ideas with the more rigorous quantum principles. In addition, such formulation increases the armamentarium of theoretical tools provided by the TOPS-MODE to account for several electronic statistical-mechanical parameters as well as for understanding the electron flowing in the molecules.

Conclusions and Future Outlook

The topological substructural molecular design (TOPS-MODE) approach has been widely proved in QSPR/QSAR studies. The new formulation of this approach in terms of the quantum-chemical formalism gives TOPS-MODE a solid theoretical basis. Accordingly, the TOPS-MODE postulates affirm that any molecular property can be expressed as a function of the spectral moments of certain types of molecular and environment-dependent energies. The functions accounting for the intrinsic and external energies used in the current version of TOPS-MODE cannot be considered as a complete basis for fulfilling this hypothesis. However, the introduction and testing of new energy functions is possible at any stage and does not alter the fundamental theoretical principles introduced in this work.

As we have proved in this work, TOPS-MODE is a quantum-chemical approach based on a tight-binding formalism in which a series of bond potentials account for intra- and intermolecular interactions. The empirical potentials used in the current version of this method can be easily substituted for others derived from first-principle approaches. We have proved that any QSPR/QSAR model obtained by using the current formalism can be expressed as a bond additivity function. In addition, we have established the basis for analyzing the transferability of these bond contributions. In general, most of such contributions are transferable if the moments of the Hamiltonian matrices involved in the QSPR/QSAR models are of relatively low order. The current approach also permits one to express any QSPR/QSAR model as a substructural cluster expansion, which can be useful for identifying molecular regions important for a property/activity. Finally, the new formalism of tight-binding quantum-chemical TOPS-MODE has permitted to define a series of new statistical-mechanical electronic parameters for the molecules. These new parameters keep their topological roots permitting a clear structural interpretation and have been proved as valuable tools in a new QSAR model developed here. The study of the toxicity of polyhalogenated dibenzo-1,4-dioxins has illustrated some of the advantages of using TOPS-MODE in QSPR/QSAR as well as of its new quantum-chemical formulation.

There are some future directions that immediately emerge from the current work. The first is to extend it to account for the three-dimensional geometry of molecules. This can be done by considering potentials that explicitly depend on interatomic distances. Second, we have observed in previous studies that both TOPS-MODE and density functional theory (DFT) permit the definition of Hammett-type reactivity constants.[12a,44] Thus, we believe that another future direction is to study the relationships between TOPS-MODE and the reactivity parameters defined in the context of DFT. This can be done through the use of the known relationship between the density of states and the molecular softness.[45] Finally, we believe that the study of other types of empirical potentials as well as the use of potentials obtained from *first principle* quantum-chemical methods might enhance the current performance of TOPS-MODE.

Acknowledgment. The current work has been possible only because I have been "standing on the shoulders of Giants", to whom this work is dedicated: A. T. Balaban, D. Bonchev, R. Carbó-Dorca, I. Gutman, G. G. Hall, R. Hoffmann, D. J. Klein, R. A. Marcus, M. Randić and N. Trinajstić, who previously developed the basis on which this work stands. Dr. Y. Simón-Manso (NIST, USA) and Prof. N. Hatano (Univ. of Tokyo, Japan) are thanked for revision and critical comments of an early version of this work.

References and Notes

(1) (a) Katritzky, A. R.; Lobanov, V. S.; Karelson, M. *Chem. Soc. Rev.* **1995**, *24*, 279. (b) Lombardo, F.; Gifford, E.; Shalaeva, M. Y. *Mini Rev. Med. Chem.* **2003**, *3*, 861. (c) Kaliszan, R. *Chem. Rev.* **2007**, *107*, 3212.

(2) (a) Winkler, A. A. *Brief. Bioinform.* **2002**, *3*, 73. (b) Brown, N.; Lewis, R. *Curr. Op. Drug Discov. Dev.* **2006**, *9*, 419. (c) Lill, M. A. *Drug Disc. Today* **2007**, *12*, 1013.

(3) Besalú, E.; Gironés, X.; Amat, L.; Carbó-Dorca, R. *Acc. Chem. Res.* **2002**, *35*, 289.

(4) Carbó-Dorca, R.; Amat, L.; Besalú, E.; Gironés, X.; Robert, D. *J. Mol. Struct. (Theochem)* **2000**, *504*, 181.

(5) (a) Ponec, R.; Amat, L; Carbó-Dorca, R *J. Comput.-Aided Mol. Des* **1999**, *13*, 259. (b) Gironés, X.; Carbó-Dorca, R.; Ponec, R. J. *Chem. Inf. Comput. Sci.* **2003**, *43*, 2033. (c) Carbó-Dorca, R. *SAR QSAR Environ. Res.* **2007**, *18*, 265.

(6) Estrada, E.; Uriarte, E. *Curr. Med. Chem.* **2001**, *8*, 1699.

(7) (a) Estrada, E.; Gutierrez, Y. *J. Chromatogr. A* **1999**, *858*, 187. (b) Estrada, E.; Gutierrez, Y.; González, H. J. *Chem. Inf. Comput. Sci.* **2000**, *40*, 1386.

(8) (a) Pérez-González, M.; Morales-Helguera, A. *J. Comp.-Aided Mol. Des.* **2003**, *17*, 665. (b) Pérez-González, M.; Morales-Helguera, A.; Collado, I. G. *Mol. Div.* **2006**, *10*, 109.

(9) (a) Cabrera-Pérez, M. A.; Ruiz-García, A.; Fernández-Teruel, C.; González-Álvarez, I.; Bermejo-Sanz, M. *Eur. J. Pharm. Biopharm.* **2003**, *56*, 197. (b) Cabrera-Pérez, M. A.; Bermejo, M.; Pérez, M.; Ramos, R. J. *Pharm. Sci.* **2004**, *93*, 1701. (c) Cabrera-Pérez, M. A.; Bermejo-Sanz, M.; Ramos-Torres, L.; Grau-Ávalos, R.; Pérez-González, M.; González-Díaz, H. *Eur. J. Med. Chem.* **2004**, *39*, 905. (d) Cabrera-Pérez, M. A.; González, I.; Fernández, C.; Navarro, C.; Bermejo, M. J. *Pharm. Sci.* **2006**, *95*, 589.

(10) (a) Estrada, E.; Uriarte, E.; Molina, E.; Simón-Manso, Y.; Milne, G. W. *J. Chem. Inf. Model.* **2006**, *46*, 2709. (b) Estrada, E.; Patlewicz, G.; Chamberlain, M.; Basketter, D.; Larbey, S. *Chem. Res. Toxicol.* **2003**, *16*, 1226. (c) Estrada, E.; Patlewicz, G.; Gutierrez, Y. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 688.

(11) (a) Pérez-González, M.; Teran-Moldes, M. C.; Fall, Y.; Carlos Dias, L.; Morales-Helguera, *Polymer* **2005**, *46*, 2783. (b) Morales-Helguera, A.; Cabrera-Pérez, M. A.; Pérez-González, M.; Molina-Ruiz, R.; González-Díaz, H. *Bioorg. Med. Chem.* **2005**, *13*, 2477. (c) Morales-Helguera, A.; Cabrera-Pérez, M. A.; Combes, R. D.; Pérez-González, M. *Toxicology* **2006**, *220*, 51.

(12) (a) Estrada, E.; Molina, E. *J. Mol. Graph. Modell.* **2006**, *25*, 275. (b) Estrada, E.; Molina, E.; Uriarte, E. *SAR QSAR Environ. Res.* **2001**, *12*, 445. (c) Estrada, E.; Uriarte, E.; Gutierrez, Y.; Gonzalez, H. *SAR QSAR Environ. Res.* **2003**, *14*, 145.

(13) (a) Estrada, E.; Peña, A.; García-Domenech, R. *J. Comp.-Aided Mol. Des.* **1998**, *12*, 583. (b) Estrada, E.; Peña, A. *Bioorg. Med. Chem.* **2000**, *8*, 2755. (c) Estrada, E.; Uriarte, E.; Montero, A.; Teijeira, M.; Santana, L.; De Clercq, E. *J. Med. Chem.* **2000**, *43*, 1975. (d) Estrada, E.; Vilar, S.; Uriarte, E.; Gutierrez, Y. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1194. (e) Vilar, S.; Estrada, E.; Uriarte, E; Santana, L.; Gutierrez, Y. *J. Chem. Inf. Comput. Sci.* **2005**, *45*, 502. (f) Estrada, E.; Quincoces, J.; Patlewicz, G. *Mol. Div.* **2004**, *8*, 21.

(14) (a) Pérez-González, M.; Carlos Dias, L.; Morales-Helguera, A.; Morales-Rodríguez, Y.; Gonzaga de Oliveira, L.; Torres-Gómez, L.; González-Díaz, H. *Bioorg. Med. Chem.* **2004**, *12*, 4467. (b) Pérez-González, M.; González-Díaz, H.; Molina-Ruiz, R.; Cabrera-Pérez, M. A.; Ramos-de-Armas, R. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1192. (c) Molina, E.; González-Díaz, H.; Pérez-González, M.; Rodríguez, E.; Uriarte, E. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 515. (d) Cabrera-Pérez, M. A.; Bermejo-Sanz, M. *Bioorg. Med. Chem.* **2004**, *12*, 5833. (e) Pérez-González, M.; Teran-Moldes, M. C. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3077. (f) Saíz-Urra, L.; Pérez-González, M.; Collado, I. G.; Hernández-Galán, R. *J. Mol. Graph. Modell.* **2007**, *25*, 680.

(15) Franke, R. *Theoretical Drug Design Methods*; Elsevier: Amsterdam, 1984.

(16) Halpern, F. R. *Phys. Rev.* **1957**, *107*, 1145.

(17) (a) Gutman, I.; Soldatović, T.; Graovac, A.; Vuković, S. *Chem. Phys. Lett.* **2001**, *334*, 168. (b) Morales, D. A. *J. Math. Chem.* **2005**, *38*, 389.

(18) (a) Cyrot-Lackmann, F. *J. Phys. Chem. Solids* **1968**, *29*, 1235. (b) Gaspard, J. P.; Cyrot-Lackmann, F. *J. Phys. C* **1973**, *6*, 3077. (c) Trias, A.; Kiwi, M.; Weissmann, M. *Phys. Rev. B* **1983**, *28*, 1859.

(19) (a) Gutman, I.; Trinajstić, N. *Chem. Phys. Lett.* **1972**, *17*, 535. (b) Gutman, I.; Trinajstić, N. *Chem. Phys. Lett.* **1973**, *20*, 257. (c) Jiang, Y.; Tang, A.; Hoffmann, R. *Theor. Chim. Acta* **1984**, *66*, 183.

(20) Franklin, J. L. *J. Chem. Phys.* **1954**, *22*, 1304.

(21) (a) Lennard-Jones, J. E. *Proc. R. Soc.* **1949**, *A198*, 1. (b) Hall, G. G.; Lennard-Jones, J. E. *Proc. R. Soc.* **1950**, *A202*, 155. (c) Hall, G. G. *Proc. R. Soc.* **1950**, *A202*, 166. (d) Hall, G. G.; Lennard-Jones, J. E. *Proc. R. Soc.* **1951**, *A205*, 357.

(22) Brown, R. D. *J. Chem. Soc.* **1953**, 2615.

(23) Miller, K. J. *J. Am. Chem. Soc.* **1990**, *112*, 8533.

(24) Ertl, P. *J. Med. Chem.* **2000**, *43*, 3714.

(25) (a) Gasteiger, J.; Marsilli, M. *Tetrahedron Lett.* **1978**, *34*, 3181. (b) Gasteiger, J.; Marsilli, M. *Tetrahedron* **1980**, *36*, 3219.

(26) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441.

(27) El Tayar, N.; Testa, B.; Carrup, P.-A. *J. Phys. Chem.* **1992**, *96*, 1455.

(28) (a) Wang, R.; Fu, Y.; Lai, L. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 615. (b) Wang, R.; Gao, Y.; Lai, L. *Persp. Drug Discov. Des.* **2000**, *19*, 47.

(29) Ghose, A. K.; Crippen, G. M. *J. Chem. Inf. Comput. Sci.* **1987**, *27*, 21.

(30) Allen, T. L.; Shull, H. *J. Chem. Phys.* **1961**, *35*, 1644.

(31) Benson, S. W. *Thermochemical Kinetics*, second ed.; Wiley: New York, 1976.

(32) (a) Marcus, R. A. J. *Chem. Phys.* **1965**, *43*, 2643. (b) Randić, M. *J. Comput. Chem.* **1980**, *1*, 386. (c) Bonchev, D.; Kier, L. B. *J. Math. Chem.* **1992**, *9*, 75. (d) Ivanciuc, O.; Balaban, A. T. *J. Math. Chem.* **1992**, *11*, 155. (e) Bonchev, D.; Gordeeva, E. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 383.

(33) (a) Klein, D. J. *Int. J. Quantum Chem.* **1986**, *S20*, 153. (b) Klein, D. J.; Schmalz, T. G.; Bytautas, L. *SAR QSAR Environ. Res.* **1999**, *10*, 131.

(34) (a) Burdett, J. K.; Lee, S. *J. Am. Chem. Soc.* **1985**, *107*, 3050. (b) Burdett, J. K.; Lee, S.; McLarman, T. J. *J. Am. Chem. Soc.* **1985**, *107*, 3083. (c) Burdett, J. K. *Struct. Bonding (Berlin)* **1987**, *65*, 29. (d) Lee, S.; Hoistad, L. M.; Carter, S. T. *New. J. Chem.* **1992**, *16*, 651. (e) Burdett, J. K. *J. Mol. Struct. (Theochem)* **1995**, *336*, 115.

(35) Atkins, P. W. *Physical Chemistry*; Oxford Univ. Press: Oxford, 1998.

(36) Morse, P. M.; Feshbach, H. *Methods in Theoretical Physics*; McGraw-Hill: New York, 1953; Chapter 7.

(37) (a) Estrada, E.; Rodríguez-Velázquez, J. A. *Phys. Rev. E* **2005**, *71*, 056103. (b) Estrada, E.; Rodríguez-Velázquez, J. A. *Phys. Rev. E* **2005**, *72*, 046105. (c) Estrada, E.; Hatano, N. *Chem. Phys. Lett.* **2007**, *439*, 247.

(38) (a) de la Peña, J. A.; Gutman, I.; Rada, J. *Linear Algebra Appl.* **2007**, *427*, 70. (b) Gutman, I.; Radenković, S. Z. *Naturforsch.* **2007**, *62a*, 254. (c) Gutman, I.; Graovac, A. *Chem. Phys. Lett.* **2007**, *436*, 294. (d) Gutman, I.; Furtula, B.; Marković, V.; Glišić, B. *Z. Naturforsch.* **2007**, *62a*, 495. (e) Gutman, I.; Radenković, S.; Graovac, A.; Plavšić, D. *Chem. Phys. Lett.* **2007**, *446*, 233. Carbó-Dorca, R. *J. Math. Chem.*, available online Nov 1, 2007, http://dx.doi.org/10.1007/s10910-007-9314-y.

(39) (a) Safe, S. H. *Rev. Pharmacol. Toxicol.* **1986**, *26*, 371. (b) Safe, S. H. *Crit. Rev. Toxicol.* **1990**, *21*, 51.

(40) (a) Kobayashi, S.; Kitadai, M.; Sameshima, K.; Ishii, Y.; Tanaka, A. *J. Mol. Struct.* **1999**, *475*, 203. (b) Asatryan, R. S.; Mailyan, N. Sh.; Khachatryan, L.; Dellinger, B. *Chemosphere* **2002**, *48*, 227. (c) Arulmozhiraja, S.; Morita, M. *Chem. Res. Toxicol.* **2004**, *17*, 348. (d) Ashek,

A.; Lee, C.; Park, H.; Cho, S. J. *Chemosphere* **2006**, *65*, 521. (e) Gu, C. G.; Jiang, X.; Ju, X. H.; Yang, X. L.; Yu, G. F. *SAR QSAR Environ. Res.* **2007**, *18*, 603.

(41) So, S. S.; Karplus, M. *J. Med. Chem.* **1997**, *40*, 4360.

(42) Zheng, G.; Xiao, M.; Lu, X. H. *Anal. Bioanal. Chem.* **2005**, *383*, 810.

(43) Fraschini, E.; Bonati, L.; Pitea, D. *J. Phys. Chem.* **1996**, *100*, 10564.

(44) Simón-Manso, Y. *J. Phys. Chem. A* **2005**, *109*, 2006.

(45) (a) Yang, W.; Parr, R. G. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 6723. (b) Cohen, M. H.; Ganduglia-Pirovano, M. V.; Kudrnovský, J. *J. Chem. Phys.* **1994**, *101*, 8988. (c) Kohn, W.; Becke, A. D.; Parr, R. G. *J. Phys. Chem.* **1996**, *100*, 12974.

JP8010712