Three-Dimensional Generalized Graph Matrix, Harary Descriptors, and a Generalized Interatomic Lennard-Jones Potential

Ernesto Estrada*

Molecular Informatics, X-rays Unit, RIAIDT, Edificio CACTUS, University of Santiago de Compostela, Santiago de Compostela 15782, Spain

Received: February 26, 2004; In Final Form: April 26, 2004

The generalized graph matrix is extended to consider three-dimensional (3D) interatomic distances between pairs of atoms in molecules. It is used to (re)define some topographic descriptors, such as 3D-Wiener and 3D-Harary numbers. Harary numbers are generalized to a wide set of descriptors, and they are adapted for considering only nonbonded pairs of atoms. These Harary numbers for nonboded pairs of atoms are identified as repulsion potentials of the Mie type providing a physical interpretation to these descriptors. This formalism is adopted for generalizing Lennard-Jones (LJ) potentials using topological parameters. LJ parameters of three united-atom (UA) force fields, TIPS−UA, PRF−UA, and TraPPE-UA, are redefined using vertex degrees of pseudoatoms. High correlation coefficients (>0.99) are obtained between original LJ parameters and those derived from vertex degrees for linear and branched alkanes. These results make some links between some graph theoretical parameters used in structure−property relations and well-known interatomic potentials used in computational chemistry force fields.

Introduction

The use of graph theoretical concepts and methods is a common practice in several branches of physics, chemistry, and biology. Applications in statistical and theoretical physics[1] and the recent study of Anderson localization on graphs, i.e., quantum graphs,[2] are a couple of examples of the fruitful use of graph theory in physics. Many of the cheminformatics[3] and bioinformatics[4] tools used for coding, storing, manipulating, and analyzing molecular and macromolecular structures are also based on graph theoretical algorithms. In physical chemistry, the use of graph theory has played an important role in areas ranging from the study of electronic properties of conjugated systems,[5] statistical thermodynamics,[6] molecular spectroscopy, and quantum chemistry.[7]

The generation of molecular descriptors using graph theoretical concepts has grown as a parallel area of research in physical chemistry.[8] Although these descriptors, known as topological indices, characterize only two-dimensional (2D) molecular attributes, they have been very useful in several real world applications. They include the design of new biologically active molecules[9] and recent applications in chemical genomic profiling of biological networks.[10] In some cases, these indices have been extended to three-dimensional (3D) descriptors, which are known as topographic indices.[11−15] However, most of these descriptors, 2D and 3D, have been developed in an ad hoc way, which have kept them elusive to physical interpretations. On the other hand, most of these descriptors appear to be disconnected from each other, even from a mathematical point of view. This lack of unity gives the false impression that these indices are totally unrelated, which in many cases is not true, and stimulate the proliferation of new descriptors, which in some cases are unnecessary.

Recently, we have introduced a generalized graph matrix which encompasses several of the mathematical formalisms used in chemical graph theory into one unified scheme.[16−18] We have shown that the use of this generalized graph matrix has implications to the study of conjugated systems,[18] giving a solution to the problem of isospectral (molecular) graphs, it is useful for studying geometrical invariants associated with (chemical) graphs[18] and permits to define several of the well-known topological indices using the same graph invariant.[16−18] Among these descriptors, we can mention Wiener[19] and Harary numbers,[20,21] connectivity index,[22] Balaban index,[23] and Zagreb indices,[24] among others. One of the obvious advances of this unified approach to the use of graph theory in physical chemistry is related to the optimization and interpretation of these molecular descriptors.

In this work, we will extend the generalized graph matrix to consider 3D molecular features by using geometrical instead of topological distances. However, we will keep topological information at the same time as the geometrical one, which will be useful for studying some “through bond” molecular interactions. An obvious application of this new generalized 3D matrix is in the (re)definition of some topographic descriptors, such as 3D Wiener and Harary numbers. We will be concentrating on the Harary numbers[25] as we will find some important links between them and interatomic potentials of the Mie type,[26] such as the Lennard-Jones (LJ) potential.[27] Taking advantage of this interrelation, we will give a physical interpretation of the Harary numbers and we will introduce a generalized LJ potential. In the last case, we will be able to present a LJ potential for alkanes (linear and branched) based on the combination of topological and geometrical information. By using some topological parameters for defining LJ parameters, such as size and well depth, we are able to reproduce some of the LJ potentials used in well know united-atoms (UA) force fields.

Theoretical Developments

Let $G = (V,E)$ be a molecular graph, where V and E are the vertex and edge sets representing atoms and covalent bonds,

---

* To whom correspondence should be addressed. E-mail: estrada66@yahoo.com. Fax: 34-981-547 077.

10.1021/jp049139c CCC: $27.50

respectively. Let N be the number of atoms and M the number of bonds in the molecule. We will start by giving the definition of the generalized graph matrix, which was previously defined by us.[16]

Definition 1. Let $\Gamma(x,v) = [g_{ij}(x,v)]_{N\times N}$ be the generalized graph matrix, which is defined as a square symmetric matrix with elements g_{ij}:

$$g_{ij} = \begin{cases} 1 & \text{if } d_{ij} = 1 \\ (d_{ij}x^{d_{ij}-1})^{v} & \text{if } i \neq j;\ d_{ij} > 1 \\ 0 & \text{otherwise} \end{cases}$$

where x and v are not simultaneously equal to zero. According to this formalism, several topological indices can be defined using the same graph invariant.[16-18] Here we will concentrate mainly on those descriptors based on distances, such as Wiener[19] and Harary indices,[20,21] which are defined as

$$W = \frac{1}{2}[\mathbf{u}\Gamma(x=1,v=1)\mathbf{u}^{\mathrm{T}}] = \sum_{i<j} d_{ij} \tag{1}$$

$$H_1 = \frac{1}{2}[\mathbf{u}\Gamma(x=1,v=-1)\mathbf{u}^{\mathrm{T}}] = \sum_{i<j}\frac{1}{d_{ij}} \tag{2}$$

$$H_2 = \frac{1}{2}[\mathbf{u}\Gamma(x=1,v=-2)\mathbf{u}^{\mathrm{T}}] = \sum_{i<j}\frac{1}{(d_{ij})^2} \tag{3}$$

where $\mathbf{u}$ is a unit vector and $\mathbf{u}^{\mathrm{T}}$ its transpose.

It is straightforward to generalize Harary indices to a series of descriptors of the form

$$H_m = \frac{1}{2}[\mathbf{u}\Gamma(x=1,v=-m)\mathbf{u}^{\mathrm{T}}] = \sum_{i<j}\frac{1}{(d_{ij})^m} \tag{4}$$

We are going to extend these indices to consider three-dimensional interatomic distances instead of topological ones. First, we will define the generalized graph matrix for the 3D case.

Definition 2. Let $\Theta(x,v) = [f_{ij}(x,v)]_{N\times N}$ be the generalized 3D graph matrix, which is defined as a square symmetric matrix with elements f_{ij}

$$f_{ij} = \begin{cases} (r_{ij})^{v} & \text{if } d_{ij} = 1 \\ (r_{ij})^{v}x^{f(d_{ij})} & \text{if } d_{ij} > 1 \\ 0 & \text{otherwise} \end{cases}$$

where x and $f(d_{ij})$ are not simultaneously equal to zero and $f(d_{ij})$ is a function of the topological distance, defined as the shortest path between vertexes i and j. We have previously used the simple function $f(d_{ij}) = d_{ij} - 1$ for the 2D case.

The 3D Wiener[12] and Harary numbers are then written as follows:

$${}^{3\mathrm{D}}W = \frac{1}{2}[\mathbf{u}\Theta(x=1,v=1)\mathbf{u}^{\mathrm{T}}] = \sum_{i<j} r_{ij} \tag{5}$$

$${}^{3\mathrm{D}}H_m = \frac{1}{2}[\mathbf{u}\Theta(x=1,v=-m)\mathbf{u}^{\mathrm{T}}] = \sum_{i<j}\frac{1}{(r_{ij})^m} \tag{6}$$

Harary Indices and Mie Potential

If we take the interatomic distances between bonded atoms as constants or having small variations, we can consider the generalized topographic matrix only for nonbonded atoms

$$\Theta_{\mathrm{NB}}(1,-m) = \Theta(x=1,v=-m) - \Theta(x=0,v=-m) \tag{7}$$

The matrix Θ_{NB} considers all nonbonded pairs of atoms in a molecule. However, it is common practice in the study of van der Waals interatomic potentials to exclude interactions between a pair of atoms separated not only by one but also by two (and even by three) bonds. In these cases, Θ_{NB} can be built by considering that $x = 0$ for $d_{ij} > 2$ (or $d_{ij} > 3$) and $x = 1$ otherwise in the second matrix of the right part of eq 7.

A topographic (3D) Harary number for non bonded atoms is then defined as follows:

$${}^{3\mathrm{D}}H_{\mathrm{NB}}(-m) = \frac{1}{2}[\mathbf{u}\Theta_{\mathrm{NB}}(x=1,v=-m)\mathbf{u}^{\mathrm{T}}] = \sum_{i<j}^{\mathrm{NB}}\frac{1}{(r_{ij})^m} \tag{8}$$

If we plot $H_{\mathrm{NB}}(-m)$ as a function of r_{ij}, we obtain a typical graphic for an interatomic repulsion potential. In a similar way, we can define another potential simulating the attractive interaction between nonbonded pairs of atoms in the following form:

$$H_{\mathrm{NB}}(-n) = \frac{1}{2}[\mathbf{u}\Theta_{\mathrm{NB}}(x=1,v=-n)\mathbf{u}^{\mathrm{T}}] = -\sum_{i<j}^{\mathrm{NB}}\frac{1}{(r_{ij})^n} \tag{9}$$

where

$$\Theta_{\mathrm{NB}}(1,-n) = \Theta(x=0,v=-m) - \Theta(x=1,v=-m) \tag{10}$$

Here again, in case that interactions between atoms separated by two or three bonds are not considered, we have to take $x = 0$ for $d_{ij} > 2$ or $d_{ij} > 3$, respectively, in the first matrix of the right part of eq 10. Note that here the matrices in the right part of the equation are in a different order compared to those in eq 7, which introduces the minus sign in eq 9.

It is now straightforward to realize that the Mie potential[26] for interatomic interactions is a particular case of the sum of $H_{\mathrm{NB}}(-m)$ and $H_{\mathrm{NB}}(-n)$ when $m > n > 3$

$$u(r)_{\mathrm{Mie}} = \gamma\cdot H_{\mathrm{NB}}(-m) + \mu\cdot H_{\mathrm{NB}}(-n) = \sum_{i<j}^{\mathrm{NB}}\frac{\gamma}{(r_{ij})^m} - \frac{\mu}{(r_{ij})^n} \qquad m > n > 3 \tag{11}$$

where γ and μ are arbitrary constants. A particularly very well-known case of this potential is the (12-6) LJ potential,[27] which will be considered later on in this work, where the arbitrary constants are substituted by parameters with physical meaning.

Interpretation of Harary Numbers

According to the original definition of the Harary numbers, these descriptors can be considered as the sum of two terms, one consisting of the sum of powers of inverse distances between bonded atoms and the other summing powers of inverse distances between pairs of nonbonded atoms

$$H_m = \sum_{i<j}^{\mathrm{ALL}}\frac{1}{(d_{ij})^m} = \sum_{i<j}^{\mathrm{Bonded}}\frac{1}{(d_{ij})^m} + \sum_{i<j}^{\mathrm{NB}}\frac{1}{(d_{ij})^m} \tag{12}$$

It is evident that in the case of the 2D Harary indices the first term is simply the number of bonds, M, in the molecule

$$H_m = M + \sum_{i<j}^{\mathrm{NB}} \frac{1}{(d_{ij})^m} \qquad (13)$$

If we consider that in the 3D case (pseudo)atoms are connected by bonds with a fixed length L, we will have

$$^{3\mathrm{D}}H_m = LM + \sum_{i<j}^{\mathrm{NB}} \frac{1}{(r_{ij})^m} \qquad (14)$$

The second term in eqs 13 and 14 represents the energetic of nonbonding interatomic repulsions in a molecule. In consequence, we can consider that the first term represents the energy of bonding interactions, which here are taken to be proportional to the number of bonds in the molecule. Hence, Harary numbers are a sort of repulsion potentials considering bonding and nonbonding interatomic interactions in a molecule, the first being proportional to the number of bonds and the second being proportional to the sum of powers of inverse interatomic distances without any differentiation between (pseudo)atom types, e.g., $\gamma = \mu = 1$ for any pair of (pseudo)atoms in the Mie potential.

Generalized Lennard-Jones Potential

In the most general case in which atoms of different types are considered, we need to substitute the terms γ and μ in the Mie potential by parameters with physical meaning for the different pairs of atoms interacting in the molecule. This kind of potentials are known as LJ potentials and they can be written for a pair of atoms as follows:

$$u(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^m - \left(\frac{\sigma}{r}\right)^n\right] \qquad (15)$$

where ϵ is the well depth and σ is the distance at which $u(r) = 0$. The LJ potential is used in simulating interatomic interactions between nonbonded pairs of atoms. The most common case is the (12-6) potential, i.e., $m = 12$ and $n = 6$, which is used in computational chemistry softwares such as CVFF,[28] CHARMM,[29] TRIPOS,[30] SHAPES,[31] UFF,[32] ECEPP,[33] AMBER,[34] and OPLS.[35] However, other variants of this potential are also used in several force fields. For instance, CFF,[36] QMFF,[37] and ESFF[38] use a (9-6) function. Also a (12-10) LJ potential is used to simulate hydrogen bonding in ECEPP[33] and AMBER.[34]

To study LJ potentials, we introduce a change of variable in $\Gamma'(x,\upsilon)$ so that x is substituted by y or z and $\upsilon = 1$. Thus, we obtain the following matrices: $\Gamma'(y,1)$ and $\Gamma'(z,1)$. Using these matrices, two new graph-theoretical vectors are defined as follows.[16]

Definition 3: Let $\mathbf{y}(w, y, q)$ and $\mathbf{z}(s, z, r)$ be two vectors of order N whose elements y_i and z_i are defined as follows:

$$y_i = (w_i + \sum_j g_{ij}(y,1))^q \qquad z_i = (s_i + \sum_j g_{ij}(z,1))^r \qquad (16)$$

where w_i and s_i are weights to be assigned to the corresponding vertex.

Let us introduce the following change into the original definition of the $\mathbf{y}(w, y, q)$ and $\mathbf{z}(s, z, r)$ vectors by including a new parameter b, which transforms the vectors given by eq 16 into $\mathbf{y}(w, y, b, q)$ and $\mathbf{z}(s, z, b, r)$

$$y_i = \{w_i + A[\sum_j g_{ij}(y,1)]^b\}^q \text{ and } z_i = \{s_i + C[\sum_j g_{ij}(z,1)]^b\}^r \qquad (17)$$

where A and C are constants that can be obtained by multiplying $\Gamma'(y,1)$ by $A^{1/b}$ and $C^{1/b}$, respectively. The reason for this new definition of the $\mathbf{y}(w, y, b, q)$ and $\mathbf{z}(s, z, b, r)$ vectors will be evident later on in this work. Using these vectors, we can introduce a generalized interatomic repulsion potential applying a vector−matrix−vector (VMV) multiplication procedure.[39,40] For the sake of simplicity in the rest of this work, we will consider that $\mathbf{y}(w, y, b, q) = \mathbf{z}(s, z, b, r)$. Thus, a generalized interatomic repulsion potential can be obtained as follows:

$$u_G(\mathrm{rep}) = \frac{1}{2}[\mathbf{y}(w,y,b,q)\cdot\Theta_{\mathrm{NB}}(x, -m)\cdot\mathbf{y}^{\mathrm{T}}(w,y,b,q)] = \sum_{i<j}^{\mathrm{NB}} \frac{(y_i y_j)^q \cdot x^{f(d_{ij})}}{(r_{ij})^m} \qquad (18)$$

Here, y_i and y_j are the components of the $\mathbf{y}$ vector for atoms i and j, respectively. The term $x^{f(d_{ij})}$ accounts for sort of "through-bond" interactions,[41] which can be useful in simulating some interatomic effects transmitted through the molecular skeleton, such as electronic effects. In those cases in which these effects are not taken into consideration we will simply take $x^{f(d_{ij})} = 1$.

In a similar way as in eq 18, we can introduce a generalized attraction potential of the following form:

$$u_G(\mathrm{attr}) = \frac{1}{2}[\mathbf{y}(w,y,b,q')\cdot\Theta_{\mathrm{NB}}(x', -n)\cdot\mathbf{y}^{\mathrm{T}}(w,y,b,q')] = -\sum_{i<j}^{\mathrm{NB}} \frac{(y'_i y'_j)^{q'} \cdot x'^{f'(d_{ij})}}{(r_{ij})^n} \qquad (19)$$

These expressions will be applied here for studying some LJ potentials implemented in force fields currently in use in computational chemistry.

Electrostatic Interactions

Let $\mathbf{q} = (q_1 q_2 \cdots q_N)$ be a vector whose elements are the atomic charges for the different atoms in the molecule. Then the electrostatic interaction potential for nonbonded pairs of atoms is given by

$$u_{\mathrm{elec}} = \frac{1}{2}[\mathbf{q}\Theta(x = 1, \upsilon = -1)\mathbf{q}^{\mathrm{T}}] = \sum_{i<j} \frac{q_i q_j}{r_{ij}} \qquad (20)$$

In several force fields, the electrostatic interactions of atoms separated by three bonds, "1-4 interactions" are reduced by a multiplicative scale factor.[34,42] These scaling factors are also necessary for LJ interactions. For instance, OPLS and AMBER force fields use scaling factors of $^1/_2$ and $^1/_8$, respectively.[43] This kind of 1-4 interaction can be expressed through the use of our topological distance function as follows:

$$u_{\mathrm{elec}} = \frac{1}{2}[\mathbf{q}\Theta(x < 1, \upsilon = -1)\mathbf{q}^{\mathrm{T}}] = \sum_{i<j} \frac{q_i q_j}{r_{ij}} \cdot x^{f(d_{ij})} \qquad (21)$$

where we have to take $x < 1$ in order to simulate the reduction of electrostatic interactions due to the "1-4 effect".

This multiplicative scale factor only appears for pairs of atoms separated by three bonds. Consequently, we can use a Lorentzian-type function for $f(d_{ij})$ in such a way that it takes the value of 1 if the pair of atoms is separated by three bonds or zero otherwise

$$f(d_{ij}) = \frac{1}{1 + \beta(d_0 - d_{ij})^2} \tag{22}$$

where $d_0 = 3$ is the number of bonds (topological distance) at which the effect is transmitted, d_{ij} is the number of bonds separating the pair of atoms, i.e., its topological distance, and $\beta \gg 1$, e.g., $\beta = 10^6$. This function takes a value of $f(d_{ij}) = 1$ for atoms separated at topological distance of three or almost zero ($\approx 10^{-6}$ in this case) otherwise. This function can also be used in eqs 18 and 19 for introducing the scaling multiplicative factor in the LJ potential.

In united-atom approaches, such as those we will analyze here, the electrostatic interaction potential is not taken into account because the CH_n groups are considered as neutral ($q_i = 0$).[44] This is supported by ab initio calculations[45] on *n*-alkanes and by the negligible dipole moments in branched alkanes.[46]

United-Atoms LJ Potentials

Several united-atom (UA) force field models have been introduced for the simulation of organic compound in liquid phase. Most of them use optimized sets of LJ parameters for describing interatomic interactions between nonbonded pairs of atoms. In the particular case of UA description of normal and branched alkanes, several sets of such parameters have been introduced in the literature for force fields such as PRF-UA,[47] OPLS,[35] GK,[48] TIPS−UA,[49] UNICEPP,[50] and TraPPE-UA,[51,52] among others. The LJ potential in these cases has the following appearance:

$$u(r_{ij}) = \sum_{i<j} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{23}$$

where interactions are computed using standard Lorentz−Berthelot combining rules[53,54]

$$\sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj}) \text{ and } \epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}} \tag{24}$$

Our approach consists of substituting the LJ parameters used for alkanes by expressions based on topological parameters such as vertex degrees, giving formulas of the following form:

$$u_G(r,\delta) = \sum_{i<j}^{\mathrm{NB}} (\delta_i^{CD}\delta_j^{CD})\left[\left(\frac{\delta_i^{AB}\delta_j^{AB}}{r_{ij}}\right)^{12} - \left(\frac{\delta_i^{AB}\delta_j^{AB}}{r_{ij}}\right)^{6}\right] \tag{25}$$

This is possible by using our generalized matrix approach with the following set of parameters:

$$u_G(\mathrm{rep}) = \frac{1}{2}[\mathbf{y}(B,0,b,m)\cdot\Theta_{\mathrm{NB}}(1,-m)\cdot\mathbf{y}^{\mathrm{T}}(B,0,b,m)] = \sum_{i<j}^{\mathrm{NB}}\left(\frac{\delta_i^{AB}\delta_j^{AB}}{r_{ij}}\right)^{m} \tag{26}$$

$$u_G(\mathrm{attr}) = \frac{1}{2}[\mathbf{y}(B,0,b,n)\cdot\Theta_{\mathrm{NB}}(1,-n)\cdot\mathbf{y}^{\mathrm{T}}(B,0,b,n)] = -\sum_{i<j}^{\mathrm{NB}}\left(\frac{\delta_i^{AB}\delta_j^{AB}}{r_{ij}}\right)^{n} \tag{27}$$

where $\delta_i^{AB} = B + A(\delta_i)^b$ with A, B, and b being obtained by regression analysis. We first make the following approximation to transform the Lorentz−Berthelot combining rules[53,54] from a sum to a multiplication

$$\sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj}) \approx \sqrt{\sigma_{ii}\sigma_{jj}} \tag{28}$$

This approximation is valid when the difference between σ_{ii} and σ_{jj} is small, which is the case in most of the force fields analyzed here. Then, we obtain the parameters A, B, and b by regression analysis of $\sqrt{\sigma_{ii}}$ versus δ_i. In general, correlation coefficients higher than 0.99 were obtained as illustrated below. In the case of the potential well depth, we can use the following expression:

$$2(\mathbf{z}\Theta_{\mathrm{NB}}(1,0)\mathbf{z}^{\mathrm{T}}) = 4\sum_{i<j}\delta_i^{CD}\delta_j^{CD} \tag{29}$$

where $\delta_i^{CD} = D + C(\delta_i)^b$ and the C, D, and b parameters are again obtained from regression analysis of $\sqrt{\epsilon_{ii}}$ versus δ_i.

The first force field potential we are going to analyze is the TIPS-UA developed by Jorgensen.[49] In this case, both LJ parameters can be substituted by linear functions of the vertex degrees. The corresponding equations are given below (σ in Å and ϵ in kcal/mol):

$$\delta_i^{AB} = \sqrt{\sigma_{ii}} = 1.9086 + 0.0494(\delta_i) \quad \text{and} \quad \delta_i^{CD} = \sqrt{\epsilon_{ii}} = 0.5036 - 0.086(\delta_i)$$

In Table 1, we give the values reported by Jorgensen for these parameters[49] and those calculated from the topological functions. In both cases, the linear correlation coefficients are higher than 0.99 showing that TIPS-UA LJ parameters for alkanes can be substituted by topological terms.

The second force field analyzed is that developed by Poncela, Rubio, and Freire, (PRF-UA).[47] In this case, the LJ size for pairs of atoms increases from CH_3 to C, which is contrary to the trend observed in the TIPS-UA force field. In this case, the LJ parameters are given by the following expressions (σ in Å and ϵ/k_B in K, where k_B is the Boltzmann's constant):

$$\delta_i^{AB} = \sqrt{\sigma_{ii}} = 2.0029 - 0.0068(\delta_i)^3 \quad \text{and} \quad \delta_i^{CD} = \sqrt{\epsilon_{ii}} = 12.073 - 2.1943(\delta_i)$$

In Table 2, we give the values of LJ parameters calculated by our approach and those originally reported by PRF.[47]

The last force field to be analyzed is the TraPPE-UA potential developed by Martin and Siepmann.[51,52] The LJ size and well depth are approximated here by the following expressions in terms of vertex degrees (σ in Å and ϵ/k_B in K, where k_B is the Boltzmann's constant):

$$\delta_i^{AB} = \sqrt{\sigma_{ii}} = 1.9165 + 0.00951(\delta_i)^3 \quad \text{and} \quad \delta_i^{CD} = \sqrt{\epsilon_{ii}} = 12.937 - 3.1197(\delta_i)$$

The correlation coefficient between TraPPE-UA parameters[51,52]

TABLE 1: Values of Lennard-Jones (LJ) Size and Well Depth Parameters for Different Pairs of Pseudoatoms as Defined by TIPS-UA Force Field and Derived from the Generalized Topological LJ Potential

UA-pair	σ_{ij}	$\delta_i^{AB}\,\delta_j^{AB}$	ϵ_{ij}	$\delta_i^{CD}\,\delta_j^{CD}$
$CH_3...CH_3$	3.86	3.83	0.18	0.174
$CH_3...CH_2$	3.92	3.93	0.141	0.138
$CH_3...CH$	4.055	4.03	0.095	0.102
$CH_3...C$	4.15	4.12	0.073	0.067
$CH_2...CH_2$	3.98	4.03	0.11	0.110
$CH_2...CH$	4.115	4.13	0.074	0.081
$CH_2...C$	4.21	4.228	0.057	0.053
CH...CH	4.25	4.23	0.05	0.060
CH...C	4.345	4.33	0.039	0.039
C...C	4.44	4.44	0.03	0.025
R[a]		0.991		0.992

[a] Correlation coefficient for the linear regression between σ_{ij} vs $\delta_i^{AB}\,\delta_j^{AB}$ and ϵ_{ij} vs $\delta_i^{CD}\,\delta_j^{CD}$. See text for *A*, *B*, *C*, and *D* parameters.

TABLE 2: Values of Lennard-Jones (LJ) Size and Well Depth Parameters for Different Pairs of Pseudoatoms as Defined by PRF-UA Force Field and Derived from the Generalized Topological LJ Potential

UA-pair	σ_{ij}	$\delta_i^{AB}\,\delta_j^{AB}$	ϵ_{ij}	$\delta_i^{CD}\,\delta_j^{CD}$
$CH_3...CH_3$	4.02	3.98	96	97.6
$CH_3...CH_2$	3.87	3.89	73.96	75.9
$CH_3...CH$	3.69	3.63	58.79	54.2
$CH_3...C$	3.23	3.13	29.39	32.6
$CH_2...CH_2$	3.72	3.80	57	59.0
$CH_2...CH$	3.54	3.54	45.3	42.2
$CH_2...C$	3.08	3.05	22.65	25.3
CH...CH	3.36	3.31	36	30.1
CH...C	2.90	2.85	18	18.1
C...C	2.44	2.45	9	10.9
R[a]		0.995		0.993

[a] Correlation coefficient for the linear regression between σ_{ij} vs $\delta_i^{AB}\,\delta_j^{AB}$ and ϵ_{ij} vs $\delta_i^{CD}\,\delta_j^{CD}$. See text for *A*, *B*, *C*, and *D* parameters.

TABLE 3: Values of Lennard-Jones (LJ) Size and Well Depth Parameters for Different Pairs of Pseudoatoms as Defined by TraPPE-UA Force Field and Derived from the Generalized Topological LJ Potential

UA-pair	σ_{ij}	$\delta_i^{AB}\,\delta_j^{AB}$	ϵ_{ij}	$\delta_i^{CD}\,\delta_j^{CD}$
$CH_3...CH_3$	3.75	3.710	98.00	96.4
$CH_3...CH_2$	3.85	3.838	67.14	65.8
$CH_3...CH$	4.215	4.186	31.30	35.1
$CH_3...C$	5.075	4.863	7.00	4.5
$CH_2...CH_2$	3.95	3.970	46.00	44.9
$CH_2...CH$	4.315	4.330	21.45	23.9
$CH_2...C$	5.175	5.032	4.80	3.1
CH...CH	4.68	4.723	10.00	12.8
CH...C	5.54	5.488	2.24	1.65
C...C	6.1	6.376	0.50	0.21
R[a]		0.996		0.998

[a] Correlation coefficient for the linear regression between σ_{ij} vs $\delta_i^{AB}\,\delta_j^{AB}$ and ϵ_{ij} vs $\delta_i^{CD}\,\delta_j^{CD}$. See text for *A*, *B*, *C*, and *D* parameters.

and those obtained by our topological approach are 0.996 and 0.998 for the size and well depth, respectively (see Table 3).

Parameters *B* and *D* in the expressions $\delta_i^{AB} = B + A(\delta_i)^b$ and $\delta_i^{CD} = D + C(\delta_i)^b$ can be interpreted as the size and well depth of the pseudoatom for which vertex degree is zero, i.e., CH_4. In the case of TIPS-UA, there is no value reported for these parameters for methane.[49] In PRF-UA, the values of these parameters are $\sigma(CH_4) = 4.10$ and $\epsilon/k_B(CH_4) = 140$.[46] According to expressions based on topological vertex degree, these parameters are $\sigma(CH_4) = 4.01$ and $\epsilon/k_B(CH_4) = 146$, which are in very good agreement with the reported values. Note that methane parameters were not used in the regression analysis for deriving δ_i^{AB} and δ_i^{CD}. In the case of TraPPE-UA, these parameters are[51] $\sigma(CH_4) = 3.73$ and $\epsilon/k_B(CH_4) = 148$. Here there is more disagreement with the parameters obtained from the topological approach, which exaggerate the value of $\epsilon/k_B(CH_4)$: $\sigma(CH_4) = 3.67$ and $\epsilon/k_B(CH_4) = 167$. Using these relations, we can rewrite the expressions for the size and well depth in our topological approach as follows:

$$\delta_i^{AB} = \sqrt{\sigma_i} = \sqrt{\sigma(CH_4)} + A(\delta_i)^b \quad \text{and}$$

$$\delta_i^{CD} = \sqrt{\epsilon_i} = \sqrt{\sigma(CH_4)} + C(\delta_i)^b$$

Other empirical descriptions of the LJ parameters have also been reported in the literature. For instance, in the PRF-UA force field, the authors have proposed the following empirical rules for such parameters:[47]

$$\epsilon_{\text{site1-site2}} = n_{HH}\epsilon_{HH} + n_{HC}\epsilon_{HC} + n_{CC}\epsilon_{CC} \qquad (30)$$

where n_{HH}, n_{HC}, and n_{CC} are the number of interactions between the given atoms belonging to the two different sites. For instance, $n_{HH} = 0$, 1, 4, and 9 for the CC, CHCH, CH_2CH_2, and CH_3CH_3 interactions. ϵ_{HH}, ϵ_{HC}, and ϵ_{CC} are the empirical atomic contributions. In the case of the σ parameter, it is estimated as[47]

$$\sigma_{\text{site}}^{\;3} = \sum_i n_i \sigma_i^{\;3} \qquad (31)$$

where n_i is the number of *i* atoms in the site and σ_i is an empirical diameter contribution and the sum is carried out over all H and C atoms in the site.

Conclusions

The generalized graph matrix has allowed us to bring some links between apparently disconnected formalisms used in physical chemistry. Here, an extension of this matrix to consider 3D distances between atoms in molecules has allowed us to relate some topological/topographic descriptors, i.e., Harary indices, with well-known interatomic potentials, such as Mie and LJ potentials. The finding of these interconnections represents advances for both fields. In the first place, it allows us to provide some physical interpretation to known molecular descriptors, which are used in modeling and predicting physicochemical and biological properties. These descriptors can also be optimized to describe some physicochemical or biological properties in an efficient way using these interconnections.

On the other hand, the use of vertex degrees in the expression of the LJ potential constitutes a simplification of the site parameters and represents a useful way of deriving them in a general form. We have previously shown that the use of the 3D generalized graph matrix has permitted to obtain these parameters in a simple way for different force fields. It is straightforward to realize that this approach permits us to express the LJ potentials used in the different force fields studied here as particular cases of the generalized topological one. More developments based on this generalized 3D graph matrix can also be found by approaching the methods previously developed to show the interplays between graph theory and three-dimensional geometry.[54]

On the other hand, our formulas for LJ sizes and well depths using vertex degrees are in agreement with physical intuition as this topological parameter is related to the size of the corresponding pseudoatom.[55] In a previous work, we have shown that vertex degrees are proportional to the accessibility perimeter of atoms in a molecule, which can be computed from the van der Waals and covalent radii of the atom and the overlapping angle between the van der Waals circumferences of bonded atoms.[55]

Acknowledgment. Thanks are given to Ministerio de Ciencia y Tecnología, Spain for partial financial support under the “Ramón y Cajal” Program.

References and Notes

(1) *Graph Theory and Theoretical Physics*; Harary, F., Ed.; Academic Press: New York, 1967.

(2) Schanz, H.; Smilansky, U. *Phys. Rev.* **2000**, *84*, 1427.

(3) *Handbook of Cheminformatics: From Data to Knowledge*; Gasteiger, J., Ed.; Wiley-VCH: Berlin, 2003.

(4) Setubal, J.; Meidanis, J. *Introduction to Computational Molecular Biology*; PWS Publishing Co.: Boston, MA, 1997.

(5) Randić, M. *Chem. Rev.* **2003**, *103*, 3449.

(6) *Chemical Applications of Graph Theory*; Balaban, A. T., Ed.; Academic Press: London, 1976.

(7) Balasubramanian, K. *Chem. Rev.* **1985**, *85*, 599.

(8) *Topological Indices and Related Descriptors in QSPR and QSAR*; Devillers, J.; Balaban, A. T., Eds.; Gordon and Breach: Reading, 1999.

(9) Grassy, G.; Calas, B.; Yasri, A.; Lana, R.; Woo, J.; Iyer, S.; Kaczorek, M.; Foc'h, R.; Buelow, R. *Nat. Biotech*. **1998**, *16*, 748.

(10) Haggarty, S. J.; Clemons, P. A.; Schreiber, S. L. *J. Am. Chem. Soc*. **2003**, *125*, 10543.

(11) Randić, M.; Jerman-Blazić, B.; Trinajstić, N. *Comput. Chem*. **1990**, *14*, 237.

(12) Bogdanov, B.; Nikolić, S.; Trinajstić, N. *J. Math. Chem*. **1989**, *3*, 299.

(13) Diudea, M. V.; Horvath, D.; Graovac, A. *J. Chem. Inf. Comput. Sci*. **1995**, *35*, 129.

(14) Estrada, E. *J. Chem. Inf. Comput. Sci*. **1995**, *35*, 708.

(15) Randić, M.; Razinger, M. *J. Chem. Inf. Comput. Sci*. **1995**, *35*, 140.

(16) Estrada, E. *Chem. Phys. Lett*. **2001**, *336*, 248.

(17) Estrada, E. In *Topology in Chemistry*; Rouvray, D., King, B. R., Eds.; Horwood Pub. Ltd.: Chichester, U.K., 2002.

(18) Estrada, E. *J. Phys. Chem. A* **2003**, *107*, 7482.

(19) Wiener, H. *J. Am. Chem. Soc*. **1947**, *69*, 17. For review, see: Nikolić, S.; Trinajstić, N.; Mihalić, Z. *Croat. Chem. Acta* **1995**, *68*, 105.

(20) Plavšić, D.; Nikolić, S.; Trinajstić, N.; Mihalić, Z. *J. Math. Chem*. **1993**, *12*, 235.

(21) Ivanciuc, O.; Balaban, T. S.; Balaban, A. T. *J. Math. Chem*. **1993**, *12*, 309.

(22) Randić, M. *J. Am. Chem. Soc*. **1975**, *97*, 6609.

(23) Balaban, A. T. *Chem. Phys. Lett*. **1982**, *89*, 399.

(24) Gutman, I.; Ruscić, B.; Trinajstić, N. Wilcox, C. F. *J. Chem. Phys*. **1975**, *62*, 3399.

(25) Lučić, B.; Miličević, A.; Nikolić, S.; Trinajstić, N. *Croat. Chem. Acta* **2002**, *75*, 847.

(26) Mie, G. *Ann. Phys. Leipzig* **1903**, *11*, 657.

(27) Lennard-Jones, J. E. *Physica* **1937**, *4*, 941.

(28) Lifson, S.; Haggler, A. T.; Dauber, P. *J. Am. Chem. Soc*. **1979**, *101*, 5111.

(29) Brooks, S.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem*. **1983**, *4*, 187.

(30) Clark, M.; Cramer, R. D., III.; van Opdenbosch, N. *J. Comput. Chem*. **1989**, *10*, 982.

(31) Allured, V. S.; Kelly, C. M.; Landis, C. R.; *J. Am. Chem. Soc*. **1991**, *113*, 1.

(32) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A., III.; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024.

(33) Nemethy, G.; Gibsen, K. D.; Palmer, K. A.; Yoon, C. N.; Paterlini, G.; Zgari, A.; Rumsey, S.; Scheraga, H. A. *J. Phys. Chem*. **1992**, *96*, 6472.

(34) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. *J. Am. Chem. Soc*. **1995**, *117*, 5179.

(35) Damm, W.; Frontera, A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Comput. Chem*. **1997**, *18*, 1955.

(36) Hwang, M. J.; Stockfisch, T. P.; Hagler, A. T. *J. Am. Chem. Soc.* **1994**, *116*, 2515.

(37) Maple, J. R.; Hwang, M. J.; Stockfisch, T. P.; Dinur, U.; Waldman, M.; Ewig, C. S.; Haggler, A. T. *J. Comput. Chem*. **1994**, *15*, 162.

(38) Barlow, S.; Rohl, A. L.; Shi, S.; Freeman, C. M.; O'Hare, D. *J. Am. Chem. Soc*. **1996**, *118*, 7579.

(39) Estrada, E.; Rodríguez, L.; Gutierrez, A. *Comm. Math. Comput. Chem*. (*MATCH*) **1997**, *35*, 145.

(40) Estrada, E.; Rodríguez, L. *Comm. Math. Comput. Chem*. (*MATCH*) **1997**, *35*, 157.

(41) Hoffmann, R.; Imamura, A.; Hehre, W. J. *J. Am. Chem. Soc.* **1984**, *106*, 6638.

(42) Smith, J. C.; Karplus, M. *J. Am. Chem. Soc*. **1992**, *114*, 801.

(43) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. *J. Am. Chem. Soc.* **1984**, *106*, 6638.

(44) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.

(45) Jorgensen, W. L. *J. Phys. Chem*. **1983**, *87*, 5304.

(46) Durig, J. R.; Compton, D. A. C. *J. Phys. Chem*. **1979**, *83*, 265.

(47) Poncela, A.; Rubio, A. M.; Freire, J. J. *Mol. Phys*. **1997**, *91*, 189.

(48) Gelin, B. R.; Karplus, M. *Biochemistry* **1979**, *18*, 1256.

(49) Jorgensen, W. L. *J. Am. Chem. Soc*. **1981**, *103*, 335.

(50) Dunfield, L. G.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem*. **1978**, *82*, 2609.

(51) Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1998**, *102*, 2569.

(52) Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1999**, *103*, 4508.

(53) Lorentz, H. A. *Ann. Phys*. **1881**, *12*, 127.

(54) Berthelot, D. C. R. *Hebd. Séans. Acad. Sci., Paris* **1898**, *126*, 1703.

(55) *From Chemical Topology to Three-Dimensional Geometry*; Balaban, A. T., Ed.; Plenum Press: New York, 1997.

(56) Estrada, E. *J. Phys. Chem. A* **2002**, *106*, 9085.