DOI: 10.1002/chem.200601101

Neural Networks as a Tool To Classify Compounds According to Aromaticity Criteria

Mercedes Alonso* and Bernardo Herradón*[a]

Abstract: Aromaticity is a fundamental concept in chemistry, with many theoretical and practical implications. Although most organic compounds can be categorized as aromatic, non-aromatic, or antiaromatic, it is often difficult to classify borderline compounds as well as to quantify this property. Many aromaticity criteria have been proposed, although none of them gives an entirely satisfactory solution. The inability to fully arrange organic compounds according to a single criterion arises from the fact that aromaticity is a multidimensional phenomenon. Neural networks are computational techniques that allow one to treat a large amount of data, thereby reducing the dimensionality of the input set to a bidimensional output. We present the successful applications of Kohonen's self-organizing maps to classify organic compounds according to aromaticity criteria, showing a good correlation between the aromaticity of a compound and its placement in a particular neuron. Although the input data for the training of the network were different aromaticity criteria (stabilization energy, diamagnetic susceptibility, NICS, NICS(1), and HOMA) for five-membered heterocycles, the method can be extended to other organic compounds. Some useful features of this method are: 1) it is very fast, requiring less than one minute of computational time to place a new compound in the map; 2) the placement of the different compounds in the map is conveniently visualized; 3) the position of a compound in the map depends on its aromatic character, thus allowing us to establish a quantitative scale of aromaticity, based on Euclidean distances between neurons, 4) it has predictive power. Overall, the results reported herein constitute a significant contribution to the longstanding debate on the quantitative treatment of aromaticity.

Keywords: ab initio calculations · aromaticity · density functional calculations · neural networks · Sammon map

Introduction

The term aromaticity is rooted in structural organic chemistry, and it refers to the existence of some properties similar to those of benzene,[1] namely electron cyclic delocalization with energetic stabilization.[2] Aromaticity is a concept that allows organic compounds to be classified in three wide groups: aromatic, non-aromatic, and antiaromatic. Originally, the aromatic character was used mainly to explain reactivity and (thermodynamic) stability. Afterwards, aromaticity was found to influence a variety of chemical and chemico-physical properties;[3,4] which, in turn, were used to classify compounds according to their aromatic character. Although from a qualitative point of view, most organic compounds[5] can be sorted according to their aromaticity, the quantification of this property as well as the classification of borderline compounds is more challenging.

Whereas aromaticity is one of the most frequently used terms, it is vaguely defined and there is no unequivocal quantitative scale. Several chemical,[6–8] energetic,[9–14] magnetic,[15–18] electronic (delocalization),[19–23] and structural criteria[24–26] have been used with the objective to quantify aromaticity, but none of these criteria is universal. This seems to indicate that aromaticity is a multidimensional phenomenon that can not be gauged by a single criterion.[27]

The recent development of supramolecular chemistry,[28–29] as well as its consequences in the design of biologically active[30] or functional materials,[31] has stimulated a great deal of research on noncovalent bonding in organic chemistry, including the types of interacting species and their relative orientations. This research has shown that aromatic compounds are versatile building blocks for the generation

[a] M. Alonso, Dr. B. Herradón
Instituto de Química Orgánica General, C.S.I.C.
Juan de la Cierva 3
28006 Madrid (Spain)
Fax: (+34) 916-544-853
E-mail: herradon@iqog.csic.es
mercuea@iqog.csic.es

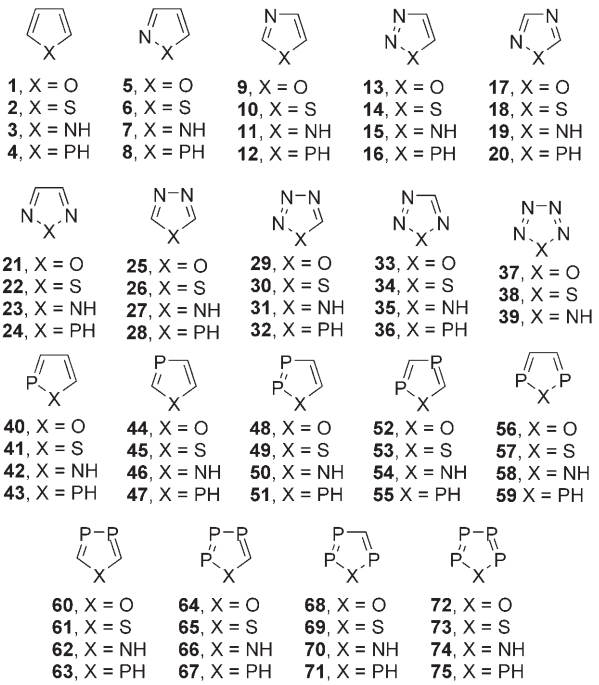
Supporting information for this article is available on the WWW under http://www.chemeurj.org/ or from the author.

of supramolecular structures[32] and, as a consequence, the classification and quantification of aromaticity is becoming essential.[33]

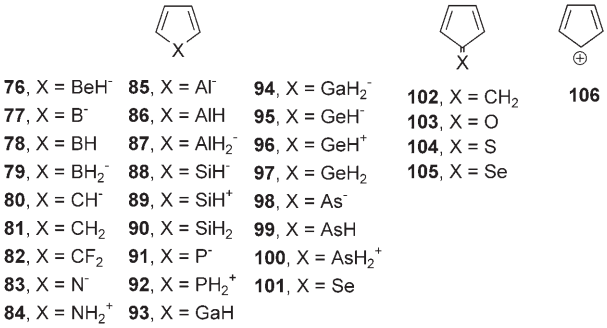
In connection with our experimental and computational studies on different aspects (synthesis,[34] structure,[35] and biological[36] properties) of aromatic compounds, we were interested in developing a simple method for classifying aromatic compounds; which, in turn, might be the starting point for a quantitative treatment of aromaticity. We proposed that the multi-dimensional character of aromaticity might be understood by using neural networks.

The use of artificial neural networks is a technique for processing data that has been applied recently in chemistry[37–39] and in other fields.[40] We reasoned that pattern classifications with neural networks are convenient alternative methods to achieve the objective to classify organic compounds based on their aromatic character, and that they would be a useful starting point to quantify this property. To this end, we have used the self-organizing maps (SOMs) developed by Kohonen.[41] These maps permit "unsupervised learning"; thus, we do not need to use any output data in this approach to make a classification of compounds based on the input data (the values of the different aromaticity criteria).

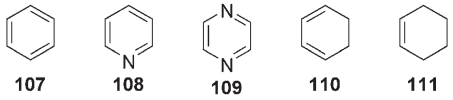
Herein, we report the successful application of SOMs to the classification of organic compounds according to their aromatic/antiaromatic character; which, in turn, is the starting point for a quantitative treatment of aromaticity.



Scheme 1. Five-membered heterocyclic compounds (**1**–**75**) used in the training of the neural network.



Scheme 2. Five-membered rings of the type C_4H_4X (**76**–**105**) and the cyclopentadienyl cation (**106**) used in the training and validation of the neural network.



Scheme 3. Six-membered ring systems used to validate the neural network.

Computational Methods

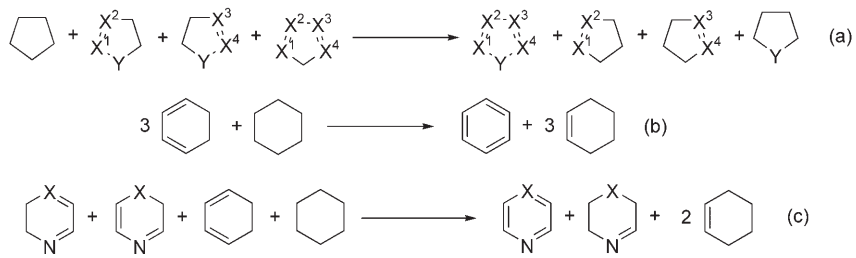
A comprehensive set of 106 five-membered cyclic compounds (**1**–**106**, Scheme 1 and Scheme 2) and five six-membered molecules (compounds **106**–**111**, Scheme 3) was used to train and validate the network. Each compound is represented by four independent descriptors that are widely used to quantify aromaticity: aromatic stabilization energy (ASE), magnetic susceptibility exaltation (Λ) and NICS, computed at the ring center and at 1 Å above the molecular plane [NICS(1)]. The aromaticity indexes for **1**–**105** were taken from a comprehensive study recently published by Cyrañski et al.[42] The calculations for the rest of the molecules were performed with the Gaussian 98 program.[43]

The geometries of compounds **106**–**111** were optimized and characterized by harmonic vibrational frequency computation at the MP2/6-31G(d) level of theory, which showed that all the structure were minima on the potential energy surface. In concordance with the results reported by Cyrañski et al.,[42] the homodesmotic reaction indicated in Scheme 4a was used to evaluate the ASE of the singlet cyclopentadienyl cation. This equation is a strain-balanced homodesmotic approach since all reference compounds are five-membered rings computed in their most stable conformation.[10b] To estimate the ASE of the six-membered systems, we employed the homodesmotic and strain-balanced equation indicated in Scheme 4b, based on 1,3-cyclohexadiene, as proposed by Schleyer and Pülhofer.[13] This reaction can be adapted to obtain reliable ASEs for six-membered heterocycles (Scheme 4c). The energies were corrected by the MP2/6-31G(d) zero-point energies. Systems with strongly positive ASEs are aromatic, while those with strongly negative values are antiaromatic.

The magnetic susceptibility exaltation (Λ) is defined as the difference between the magnetic susceptibility of a compound (χ_M) and a reference one without cyclic electron delocalization ($\chi_{M'}$) [Eq. (1)].[17] The exaltations were obtained from the reactions indicated in Scheme 4. The magnetic susceptibilities were computed using the CSGT method[44] at the HF/6-311+G** level of theory. The exaltations are negative (diamagnetic) for aromatic compounds and positive (paramagnetic) for antiaromatic compounds.

$$\Lambda = \chi_M - \chi_{M'} \quad (1)$$

The nucleus independent chemical shift (NICS) is defined as the negative value of the absolute magnetic shielding computed at the ring center or another interesting point of the system.[16] NICS were calculated at the ring critical point, the position of minimal charge density in the ring



Scheme 4. Homodesmotic and strain-balanced reactions used to evaluate the aromatic exaltation energies (ASE) and magnetic susceptibility exaltations (Λ) of singlet cyclopentadienyl cation and five-membered heterocycles (a), six-membered carbocycles (b), and six-membered heterocyclic compounds (c).

plane, as suggested by Cossio and co-workers,[45] and 1 Å above it. The NICS(1) values computed 1 Å above the molecular plane are considered to better reflect the π-electron effects.[16b] The GIAO/HF/6-311+G** method was used for the NICS calculations.[46] Rings with highly negative values of NICS and NICS(1) are quantified as aromatic, whereas those with positive values are antiaromatic.

Other useful aromaticity descriptor is the harmonic oscillator model of aromaticity (HOMA) index, defined by Kruszewski and Krygowski [Eq. (2)].[26]

$$\mathrm{HOMA} = 1-\frac{\alpha}{n}\sum_{i=1}^{n}(R_{opt}-R_i)^2 \qquad (2)$$

Where n is the number of bonds taken into the summation, and α is an empirical constant fixed to give HOMA = 0 for a model non-aromatic system and HOMA = 1 for a system with all bonds equal to an optimal value R_{opt}, assumed to be realized for a fully aromatic system. R_i is the running bond length. Since this structure-based index can not be applied to all systems due to a lack of parameters, it could not be used as descriptor in the general classification. However, we have also employed HOMA values to generate neural networks from a limited dataset (see below).

Kohonen neural network (self-organizing maps, SOM): Once the aromaticity descriptors have been obtained, we have generated a family of input vectors that represent each compound of the training dataset.

A Kohonen self-organizing map (SOM) is an unsupervised neural network that can be used to create a projection of objects from a higher dimensional space onto a lower dimensional space, usually two-dimensional, while preserving topological relations as faithfully as possible. The basic purpose of a self-organizing map is to classify, to cluster, and to visualize multivariate data.

A SOM consists of two layers of neurons: input and output layers. The input layer contains m neurons corresponding to m molecular descriptors. The output layer is usually a two-dimensional geometrical arrangement of n neurons. The m neurons of the input layer are all connected to each of the n neurons of the output layer by weights vectors w_{ji} (Figure 1).

Each compound p in the training set is represented by a vector $\boldsymbol{X_p}$ (x_{p1}, x_{p2},⋯, x_{pm}) and each neuron j in the output layer is characterized by a weight vector $\boldsymbol{W_j}$ (w_{j1}, w_{j2},⋯, w_{jm}), where m is the number of molecular descriptors employed.

The training is an iterative process during which the weight vectors are adjusted to become more similar to the training data. Initially, the weight vectors are set to random values. Then, an input vector is presented to all neurons of the network and is mapped into *one* neuron, the "winner". The selection of the matching neuron is usually made by comparison of the Euclidean distance d_j between the input vector $\boldsymbol{X_p}$ and all the weight vectors $\boldsymbol{W_j}$ [Eq. (3)].

$$d_j = \sum_{i=1}^{m}(x_{pi}-w_{ji})^2 \qquad (3)$$

The neuron j having the shortest distance to the input vector $\boldsymbol{X_p}$ is the winner. Then, the weights of the winning node and the neighboring neurons nodes are modified, according to Equation 4.

$$w_{ji}^{new} = w_{ji}^{old} + h_{ci}(t)\cdot(x_{pi}-w_{ji}(t)) \qquad (4)$$

Where t is the iteration number. $h_{ci}(t)$ is the neighborhood kernel and it determines which neurons are neighbors and how such neighboring neurons will be modified. The same procedure is repeated for all objects. Thus, a mapping of the objects onto a two-dimensional space is obtained that reflects the topology, the arrangements of the objects in the m-dimensional space.

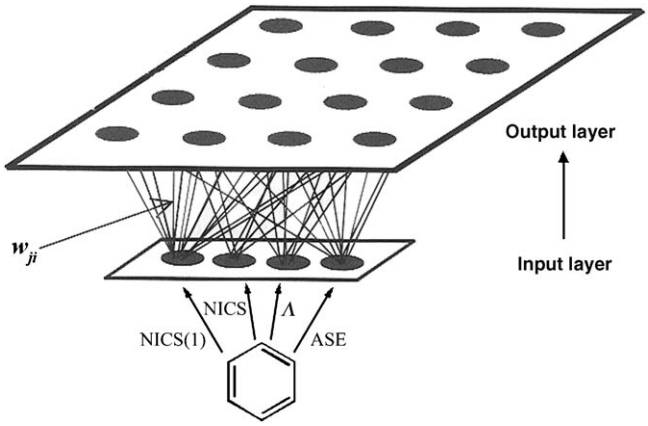


Figure 1. Architecture of a Kohonen neural network.

The Kohonen networks were obtained with the SOM_PAK program.[47] About 50 SOMs were trained by varying both the map size (number of neurons) and training parameters. Two different topologies, namely rectangular and hexagonal, were also tested and it was found that the hexagonal lattice was better for visual inspection. Although all trained networks gave similar results, a hexagonal lattice with 13×10 neurons was selected based on the appearance of the clusters in the map. *Bubble* function was used as neighborhood kernel. Training was done in two phases: an ordering phase with 2000 steps and a self-organizing phase with 20000 steps. During the first cycle, 2000 training steps were carried out and the learning rate and initial neighborhood were set to 1 and 10, respectively. The parameter values for the second cycle were 0.05 and 1, respectively.[48]

Results and Discussion

Classification of five-membered systems using SOMs: We have applied SOMs for the classification of organic compounds according to their aromatic character. The aim of this research was to elaborate a pattern that contains the most frequently used and readily obtained aromaticity indicators and to use this as a basis for quantifying aromaticity.

Since aromaticity is a multidimensional property, a classification taking into account the main aromaticity criteria is necessary. We have employed the most widely used indices of aromaticity to represent each molecule: an energetic-

based index (ASE) and three magnetic-based indices, Λ, NICS, computed at the ring center, and NICS(1), calculated 1 Å above the molecular plane. The most accepted and effective structure-based criteria of aromaticity, HOMA, can not be used as descriptor in the general classification because it can not be estimated in systems containing N–S, P–N, P–P, P–O, P–S, and C–metal bonds due to the lack of parameters. However, to check the influence of this structural index in the resulting classification, two different SOMs were performed with a dataset consisting of 56 compounds, whose HOMAs are available. These results will be presented later.

Scheme 1 and Scheme 2 show the molecular structures of the 106 five-membered ring systems used in the training and validation of the network. This comprehensive dataset, used in previous aromaticity studies,[42] contains a wide range of aromatic, non-aromatic, and antiaromatic compounds. The values of the ASE, Λ, NICS, NICS(1), and HOMA for each system are collected in Table 1. These descriptors are used as a multidimensional input for the SOM which transforms this data into a 2D map, preserving the essential topological features of the data. The SOM automatically clusters the similar compounds based on these descriptors.

Figure 2 shows the map obtained by training a self-organizing network with the four aromaticity descriptors of the first 100 compounds of the Table 1. The 2,3,4-tetraza derivatives (**37**, **38**, and **39**) were excluded from the training set since two parameters could not be computed. Systems **80**, **81**, and **106** were used to test the SOM obtained.

We have chosen a hexagonal network with 130 neurons arranged in a 13×10 grid. According to Chen and Gasteiger,[49] the networks having between one and three times as many neurons as compounds in the dataset perform very well. Each small hexagon represents a neuron: the number within a neuron is the number of the compound, as indicated in Scheme 1 and Scheme 2, mapped into it. Although there are more neurons in the output layer than compounds in the dataset, there are some compounds mapped into the same node, emphasizing that the neural network has recognized the close similarity of these compounds.

Clusters on the map are better detected by analyzing the U-matrix, that is the matrix of distances between adjacent units of the SOM obtained (Figure 3). The distances between the neighboring neurons are visualized by grey levels: white areas represent nodes that are close to each other, while black areas represent nodes that are far apart from those around it. Therefore, clusters are white zones surrounded by black boundaries. In this case, three large clusters can be identified on the left-hand side, on the upper right-hand side, and on the lower right-hand side. All the compounds placed on the left cluster follow the four aromaticity criteria: positive ASE values, negative Λ, and highly negative values of NICS and NICS(1); that is aromatic compounds. On the contrary, the systems placed on the lower right side exhibit an energetic destabilization, an exalted paramagnetic susceptibility and positive values of NICS and NICS(1); that is antiaromatic compounds. The rings located on the upper right side show intermediate values for these descriptors. Therefore, from these four descriptors, the network clusters the extensive set of 100 five-membered systems into three large families: aromatic, non-aromatic, and antiaromatic. Moreover, the SOM places compounds with similar, but non-identical, aromatic character in neighboring neurons, *creating a smooth transition of different aromaticity degrees over the whole map*.

The visualization of the components of the SOM (Figure 4) show clearly how the neurons (here representing the compounds) are ordered progressively depending on their weights. The correlations and relationships between the different aromaticity indices and the spatial localization of a compound are easily visualized by using the component planes.

The quantitative character of the SOM obtained is better illustrated in a Sammon map (Figure 5).[50] Sammon mapping is an iterative method that generates a nonlinear projection from the *n*-dimensional input vectors to two-dimensional points on a plane; the distances between the image vectors tend to approximate to Euclidean distances of the input vectors. It has been used to represent the SOM in a proportional scale that allows visualizing the shape of the clusters and the relative distance between them.

Figure 5 illustrates fairly well the quantitative character of the classification obtained. The scale shows the degree of aromaticity associated with each system depending on its position on the map. The neuron activated by 1*H*-1,2,3-triazole (**15**) represent the highest degree of aromaticity, whereas the one activated by borole (**78**) represent the highest degree of antiaromaticity. Aromaticity decreases gradually from 1*H*-1,2,3-triazole (**15**) to borole (**78**). Consequently, depending on the Euclidean distance to the neuron activated by **15**, the five-membered compounds can be divided into different groups according to their aromatic character. The values of the different aromaticity indices for each category are collected in Table 2.

Besides the general features of the map indicated in Figure 5, other interesting conclusions can be extracted from the classification pattern of aromaticity, which are discussed below.

All the compounds classified as aromatic have been totally discriminated from the other two classes by the Kohonen network. All these compounds are energetically stabilized, exhibit diamagnetic susceptibility exaltations Λ and negative values of NICS and NICS(1) (Table 2). Therefore, the classification obtained satisfies the definition proposed by Krygowski and co-workers,[3a,c] that is, the fully aromatic systems are those cyclic π-electronic systems that follow all the main aromatic criteria.

Both the U-matrix map (Figure 3) and the Sammon map (Figure 5) show a subcluster within the aromatic compounds, formed by phospha derivatives of phosphole (compounds **75**, **67**, **71**, **59**, **55**, **51**, **43**, and **63**), the silacyclopentadienyl anion (**88**), and the boracyclopentadienyl anion (**77**). Table 2 shows that these pyramidal five-membered heterocycles exhibit highly negative values of the magnetic indices, in spite

Table 1. Calculated ASE [kcal mol^{-1}], Λ [ppmcgs], NICS and NICS(1) [ppm], and HOMA for the five-membered rings.[a]

Compd	ASE	Λ	NICS	NICS(1)	HOMA[b]	Compd	ASE	Λ	NICS	NICS(1)	HOMA[b]
1	14.77	−2.90	−12.31	−9.36	0.298 (0.778)	**57**	16.02	−7.91	−13.07	−11.99	
2	18.57	−7.00	−13.80	−10.79	0.891 (0.900)	**58**	19.24	−6.41	−11.84	−10.85	
3	20.57	−6.48	−14.86	−10.60	0.876 (0.895)	**59**	7.97	−9.75	−10.00	−10.28	
4	3.20	−1.68	−5.43	−5.97	0.236 (0.557)	**60**	12.18	−3.33	−12.41	−11.12	
5	17.29	−2.71	−12.36	−10.58	0.527	**61**	16.75	−9.64	−13.11	−12.37	
6	20.18	−7.13	−13.96	−11.66		**62**	19.47	−6.96	−14.45	−11.97	
7	23.70	−7.09	−14.75	−11.93	0.926	**63**	4.11	−4.90	−6.88	−8.48	
8	3.34	−1.54	−5.65	−6.84		**64**	11.84	−2.93	−12.73	−11.37	
9	12.37	−1.83	−11.31	−9.45	0.332	**65**	15.23	−10.25	−14.12	−13.21	
10	17.43	−6.21	−13.10	−11.37	0.905	**66**	18.38	−6.82	−14.42	−12.35	
11	18.78	−5.18	−13.85	−10.83	0.908	**67**	7.22	−11.47	−12.42	−11.75	
12	3.01	−1.20	−3.78	−6.25	0.276	**68**	12.72	−1.19	−11.02	−10.34	
13	17.20	−1.57	−12.97	−11.99	0.443	**69**	14.53	−8.77	−13.60	−12.91	
14	20.48	−7.75	−14.38	−13.72		**70**	17.96	−6.11	−12.42	−11.63	
15	24.37	−6.67	−14.90	−13.51	0.931	**71**	8.93	−12.38	−11.02	−11.88	
16	2.56	−0.98	−4.13	−8.56		**72**	12.30	−2.45	−13.37	−11.97	
17	14.23	−1.33	−11.51	−10.40	0.553	**73**	12.79	−10.57	−15.00	−14.38	
18	18.28	−6.31	−13.47	−11.96		**74**	17.12	−5.98	−14.62	−12.98	
19	21.33	−5.29	−13.66	−11.84	0.940	**75**	11.24	−20.82	−17.22	−14.93	
20	3.04	−1.08	−4.62	−7.16		**76**	−7.78	10.19	9.13	4.04	(−0.166)
21	20.19	−1.58	−12.72	−12.52	0.677	**77**	9.05	−13.48	−12.65	−6.92	(0.420)
22	22.67	−7.60	−14.52	−12.96		**78**	−22.49	16.09	17.22	9.24	(−0.595)
23	22.66	−7.91	−14.83	−13.61	0.960	**79**	−0.24	−0.20	0.12	−2.79	(0.281)
24	3.14	−1.34	−5.48	−7.64		**82**	−11.88	6.65	3.36	0.48	
25	7.78	−0.59	−10.74	−10.00	0.243	**83**	19.56	−9.43	−13.26	−11.03	0.844 (0.818)
26	13.69	−5.34	−13.00	−12.34	0.849	**84**	−2.05	1.58	−5.18	−5.27	−0.308 (0.135)
27	14.96	−3.50	−13.13	−11.52	0.823	**85**	−6.87	8.93	5.56	1.18	(0.058)
28	1.80	−0.88	−2.94	−6.97	0.025	**86**	−9.98	13.05	6.35	3.06	(−0.261)
29	9.65	0.42	−12.94	−12.29	0.413	**87**	−2.07	3.78	2.84	−0.04	(0.007)
30	14.72	−6.57	−15.18	−14.65		**88**	9.30	−8.92	−9.09	−7.90	(0.792)
31	18.26	−3.48	−14.79	−14.12	0.897	**89**	−26.58	18.60	12.42	7.66	(−0.664)
32	1.51	−1.27	−4.20	−8.69		**90**	−4.61	4.06	1.07	−1.41	(−0.035)
33	18.71	−0.12	−13.84	−13.84	0.586	**91**	23.12	−9.78	−13.41	−11.03	0.730 (0.859)
34	21.62	−7.85	−15.49	−14.96		**92**	−8.31	4.17	−0.70	−2.56	0.047 (0.016)
35	26.49	−6.99	−14.96	−14.64	0.960	**93**	−9.97	13.35	6.69	3.18	(−0.300)
36	2.24	−0.48	−4.92	−9.21		**94**	−0.96	3.45	1.83	−0.52	(−0.059)
40	13.19	−1.60	−11.38	−9.34		**95**	4.88	−2.66	−4.29	−4.92	(0.626)
41	17.45	−7.21	−13.51	−11.40		**96**	−23.92	18.48	11.33	6.90	(−0.628)
42	20.31	−6.12	−13.55	−10.77		**97**	−2.97	3.74	0.35	−1.51	(0.037)
43	4.97	−4.91	−7.38	−7.73		**98**	22.21	−10.75	−12.88	−10.60	(0.877)
44	13.50	−2.99	−11.93	−10.26	0.326	**99**	1.71	−0.08	−3.93	−4.62	(0.447)
45	17.01	−8.40	−13.04	−11.59	0.854	**100**	−6.55	4.12	−1.12	−2.30	(0.010)
46	19.91	−6.85	−14.26	−11.33	0.829	**101**	16.74	−7.43	−12.81	−10.01	(0.878)
47	3.03	−2.74	−5.34	−6.90	0.378	**102**	−3.06	1.01	−0.72	−3.42	−0.142 (0.280)
48	12.23	−2.01	−11.94	−10.36		**103**	−14.65	9.05	9.63	2.81	−1.255 (−0.326)
49	15.14	−8.79	−13.89	−12.24		**104**	−11.96	10.48	12.60	3.46	−0.454 (0.031)
50	19.17	−6.70	−14.03	−11.61		**105**	−11.44	12.44	13.49	3.79	−0.307 (0.092)
51	4.25	−5.52	−8.92	−9.17		**37**			−16.16	−15.34	0.500
52	12.14	−1.65	−11.08	−10.14		**38**			−18.40	−17.48	
53	16.14	−8.37	−12.90	−12.16		**39**			−16.76	−16.59	0.950
54	18.85	−6.16	−12.89	−11.39		**80**	22.05	−10.15	−13.99	−10.25	0.736 (0.736)
55	6.18	−7.08	−8.50	−9.34		**81**	0.00	0.00	−3.18	−4.82	−0.780 (0.306)
56	12.69	−1.11	−10.25	−9.28		**106**	−60.25	36.41	52.17	36.29	−1.050

[a] The structures of compounds **1**–**106** are indicated in Scheme 1 and Scheme 2. The data for **1**–**105** were taken from reference [42]. [b] Data for HOMA based on three C−C bonds are given in parentheses.

of having a relatively small energetic stabilization. The size of these molecules is greater than the other aromatic systems and it is well-known that the magnetic indices are highly dependent on the ring size, especially the diamagnetic susceptibility exaltation. Within this subcluster, the planar 1*H*-pentaphosphole (**75**) is classified as the most aromatic phosphole derivative since it is furthest from both the non-aromatic and antiaromatic classes (see Figure 5). It is the unique planar phosphole and it has been demonstrated that the aromaticity in polyphospha systems increases on decreasing the pyramidalization of the phosphorus atom.[51]

The relative aromaticity of furan (**1**), thiophene (**2**), pyrrole (**3**), and phosphole (**4**) has been discussed extensively.[10b,52] The SOMs classify pyrrole as the most aromatic

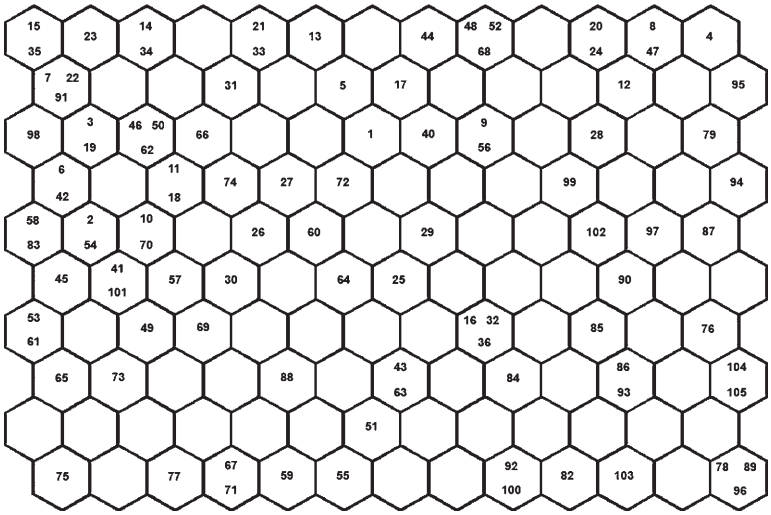


Figure 2. Kohonen map trained for the classification of five-membered systems. The numbers into the neurons corresponds to the compounds shown in Scheme 1 and Scheme 2.

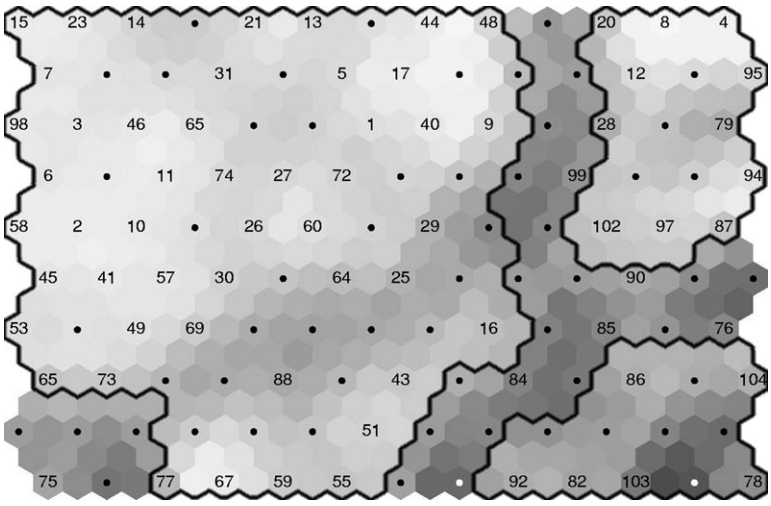


Figure 3. U-matrix map of the SOM trained with the aromaticity descriptors. The distances between the neighbouring neurons are visualized by grey levels. Darker hexagons indicate a larger distance.

compound in the series, followed by thiophene, furan and, finally, phosphole. Interestingly, the diamagnetic susceptibility exaltation Λ and NICS(1) give an inconsistent order. The phosphole is placed in the group of the non-aromatic compounds, clearly separated from the fully aromatic systems. This result does not agree with the previous conclusion that "phosphole along with cyclopentadiene are borderline aromatics".[52]

The aromaticity order pyrrole > thiophene > furan > phosphole is maintained in the rest of their polyaza and polyphospha derivatives. Moreover, it is well known that aromaticity increases when the difference in electronegativity between a heteroatom and its neighboring atoms diminishes.[53] Accordingly, the most aromatic five-membered systems in the training dataset are 1*H*-1,2,3-triazole (**15**) and 2*H*-1,2,3,4-tetrazole (**35**), since the replacement of a CH by a nitrogen atom at position 2 causes a substantial increase of aromaticity relative to pyrrole.

Validation and predictive power of the neural network: The trained Kohonen network can be used to classify new compounds according to their aromatic character as well as to predict their degree of aromaticity. To test the applicability of the method, eight compounds with a variety of structural features were selected: three five-membered rings (the highly aromatic cyclopentadienyl anion (**80**), the highly antiaromatic singlet cyclopentadienyl cation (**106**), and the cyclopentadiene (**81**)) and five six-membered rings (Scheme 3). These compounds were not used for the training of the neural network but only to validate it. Figure 6 shows the trained SOM with the neurons color coded. A quantitative assessment of the degree of aromaticity is given by the color scale in Figure 6. White neurons represent the cluster boundaries determined by the U-matrix map as discussed above.

The cyclopentadienyl anion (**80**) is correctly classified together with compound **98** in the region of the highly aromatic compounds. On the other hand, the singlet cyclopentadienyl cation (**106**), the most antiaromatic five-membered ring system, according the four descriptors, is mapped into the neuron which represents the highest degree of antiaromaticity. Finally, we have chosen the controversial cyclopentadiene (**81**), because its possible aromaticity, through the 2π hyperconjugative contribution of the CH_2, has been long discussed.[16d,52,54] According to the classification pattern of the SOM, cyclopentadiene is a non-aromatic conjugated diene in conformity with a very recent analysis.[55,56] Therefore, the neural network is able to discriminate **81** from the fully aromatic systems; although cyclopentadiene exhibits a small diamagnetic ring current but not a significant energetic stabilization,[55] it is not aromatic. This result agrees with the conclusion that the "ring current is necessary but certainly not a sufficient condition for aromaticity".[56]

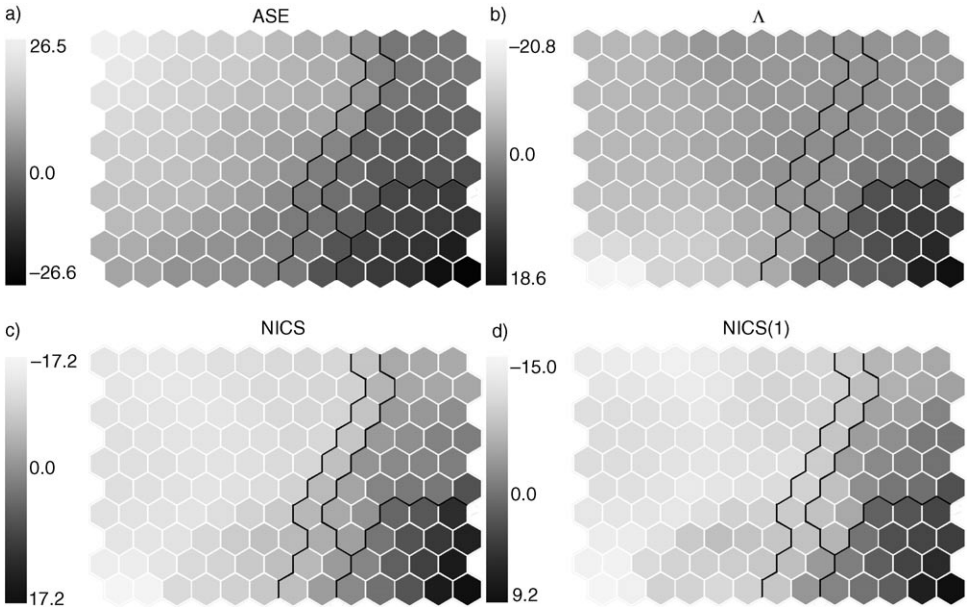


Figure 4. Component visualization of the SOM: a) ASE, b) Λ, c) NICS, and d) NICS(1). The values of each descriptor are represented by grey levels.

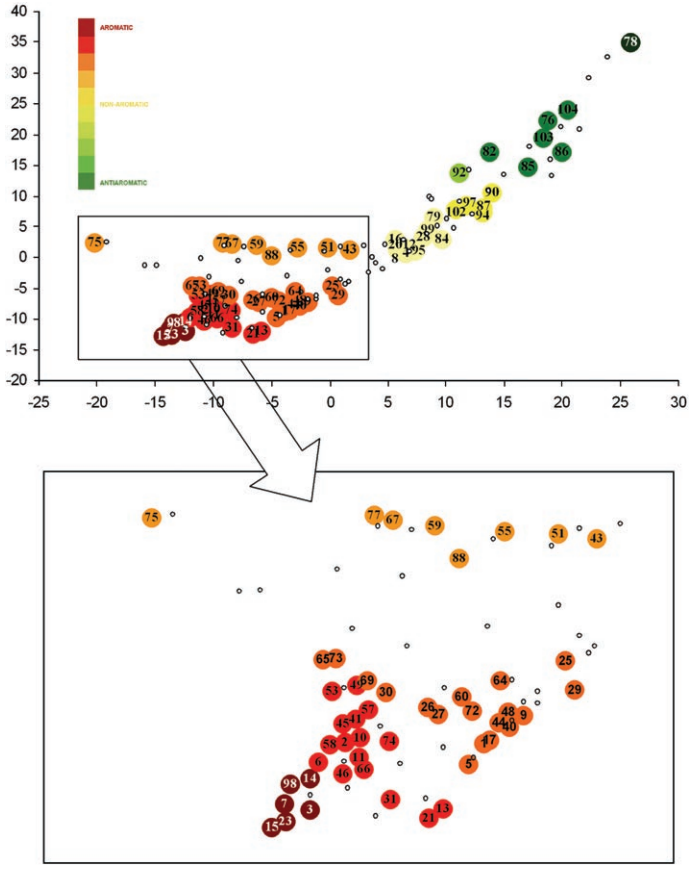


Figure 5. Sammon map obtained for the training dataset. The numbers refer to the compounds indicated in Scheme 1 and Scheme 2. The distance between the neurons corresponds approximately to Euclidean distances of the input pattern (aromaticity criteria). The colored scale corresponds to aromatic/antiaromatic features (Table 2), and the axis scales are in arbitrary units.

An advantage of SOMs is that even compounds lacking some of the descriptors (that is, having incomplete input vectors) can be classified. This is the case of the 1,2,3,4,5-oxatetrazole (**37**), 1,2,3,4,5-thiatetrazole (**38**), and 1*H*-1,2,3,4,5-pentazole (**39**), whose ASE and Λ could not be evaluated by Equation (1) since that a reference compound (2,3,4,5-tetrazacyclopentadiene) could not be optimized.[10b] With only NICS and NICS(1) as descriptors, these systems are placed in the same neuron as 1*H*-pentaphosphole (**75**), that is weakly aromatic. In this case, the Kohonen network looks for the highest similarity with regard to the available descriptors.

We have also tested the ability of the SOM to classify and quantify the aromaticity of six-membered cyclic compounds. The aromaticity indices have a series of disadvantages when comparing rings of different size. Thus, magnetic susceptibility exaltation depends heavily on the ring size and, on the other hand, ASE is strongly dependent on the reaction scheme employed for its evaluation, which is different for five-membered and six-membered rings. Despite these inconveniences, we have evaluated five six-membered ring compounds (Scheme 3), namely benzene (**107**), pyridine (**108**), pyrazine (**109**), cyclohexadiene (**110**) and cyclohexene (**111**). Table 3 collects the values of the descriptors of these molecules, and Figure 6 depicts their placements on the map. Benzene, the model aromatic compound, together with pyridine and pyrazine, are properly mapped into the neuron that represents the highest aromaticity. Cyclohexene and cyclohexadiene are well classified as non-aromatic compounds. A close inspection to the descriptors of **107**, **108**, and **109** shows that benzene is clearly more aromatic than pyrazine. However, they are mapped into the same neuron. The problem is that the values of the ASE of **107** and **108** are out of the range of the ASE values employed during the training of the network. Accordingly, any compound whose descriptors indicate a higher aromaticity than **15** and **35** will be mapped onto the same neuron since this node represents the highest aromaticity. Although the SOM is able to classify six-membered rings according to their aromatic character, it would be advisable to generate another classification for six-membered rings. A potential solution to this issue would be to make a larger map and employ six-membered cyclic compounds during the network training. Nevertheless, as the number of neurons increases, the similarity perception capability of the SOM decreases. The number of neurons chosen (130) is a good compromise for making use of both

Table 2. Range of values for d_j, ASE [kcal mol^{-1}], Λ [ppmcgs], NICS [ppm], and NICS(1) [ppm] within each category.[a]

Category	d_j[b]	ASE	Λ	NICS	NICS(1)
aromatic	0.0↔3.2	26.5↔20.5	−10.8↔−5.3	−15.5↔−12.9	−15.0↔−10.6
	4.1↔8.3	20.3↔15.1	−9.6↔−0.1	−14.8↔−11.8	−14.1↔−10.0
	8.8↔14.5	17.3↔11.8	−10.6↔0.4	−15.2↔−10.3	−14.7↔−9.3
	16.0↔21.4	11.2↔4.1	−20.8↔−4.9	−17.2↔−6.9	−14.9↔−6.9
non-aromatic	23.9↔30.0	4.9↔−2.1	−2.7↔1.6	−5.7↔0.1	−9.2↔−2.8
	31.6↔35.4	−1.0↔−4.6	1.0↔4.1	−0.7↔2.8	−3.4↔0.0
antiaromatic	36.8↔38.2	−6.6↔−8.3	4.1↔4.2	−0.7↔−1.1	−2.3↔−2.6
	40.1↔49.1	−6.9↔−14.7	6.7↔13.1	3.4↔13.5	0.5↔4.0
	54.9↔61.7	−22.5↔−26.6	16.1↔18.6	11.3↔17.2	6.9↔9.2

[a] The compounds in the Sammon map are coloured according to this scale. [b] d_j is the Euclidean distance between the weight vectors of each neuron and the neuron activated by 1*H*-1,2,3-triazole (**15**).

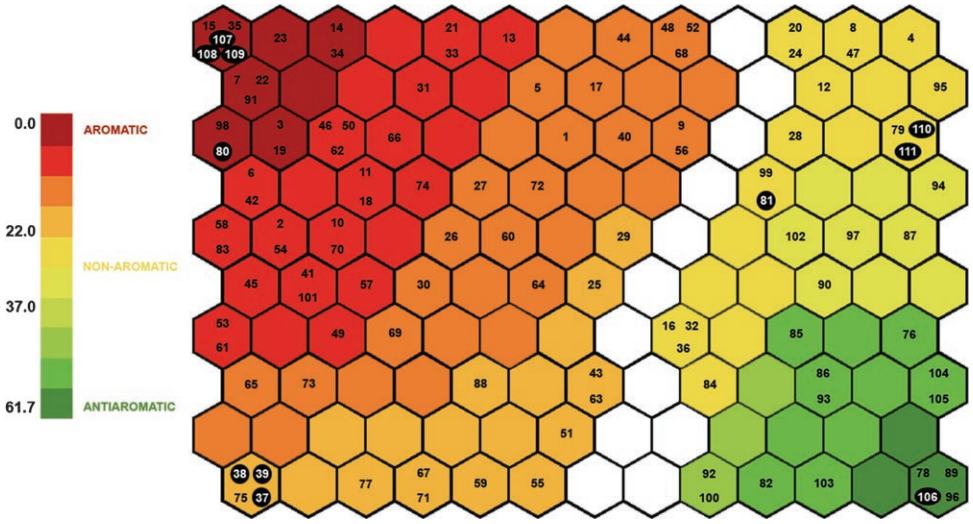


Figure 6. Self-organizing map obtained for the classification of 106 five-membered systems (**1**–**106**) and five six-membered rings (**107**–**111**). Black circles represent the compounds used in the validation of the network. The scale indicates the Euclidean distance between the weight vectors of each neuron and the neuron activated by compounds **15** and **35** (the most aromatic compounds in the training dataset). White neurons represent the cluster boundaries.

Table 3. Calculated ASE [kcal mol^{-1}], Λ [ppmcgs], NICS and NICS(1) [ppm], and HOMA for the six-membered cyclic compounds **107**–**111**.

Compound	ASE	Λ	NICS	NICS(1)
107	35.84	−15.36	−9.55	−11.26
108	33.23	−12.97	−8.10	−11.09
109	25.80	−11.06	−6.23	−11.15
110	2.15	0.76	1.45	−1.26
111	0.00	0.00	−0.77	−1.85

similarity perception and interpolation capabilities of the Kohonen neural network. Furthermore, we have not attempted to train the network with six-membered rings due to the absence of an extensive dataset with enough quality and homogeneity.

Finally, we have analyzed the influence of the structure-based index HOMA in the resultant classification. Two different SOM analyses were performed with a dataset of 56 compounds from which this index is available (see Table 1). In the first SOM, each system is represented by five descriptors including the structural index, whereas in the second one we have used only energetic and magnetic indices to represent each compound. The corresponding Sammon maps are shown in Figure 7. The distribution of the compounds hardly changes on adding the structural descriptor. In fact, the component planes for ASE and HOMA show that these variables behave in a similar way (Figure 8).

Conclusions

We have reported herein a methodology that classifies organic compounds according to their aromatic character. The approach developed is based on a Kohonen network and a set of indices of aromaticity that represent the molecules. Since the different physical properties described by the corresponding aromaticity criteria will, in general, not lead to the same classification of compounds, a pattern taking account of the most widely accepted aromaticity parameters is necessary. We have demonstrated that the SOM is able to cluster an extensive dataset of five-membered systems into three classes (aromatic, non-aromatic, and antiaromatic), based on the values of the ASE, diamagnetic exaltation Λ, NICS, and NICS(1) of each compound. Furthermore, the weight distances of the SOM (as shown in Sammon maps) allow us to quantify the degree of aromaticity associated to each molecule. The relationships between the different compounds are straightforwardly visualized on a two-dimensional map. As pointed out by Gasteiger and Chen in a different application of SOMs,[49] a two-dimensional classification can reflect much better the results of the different influences on the property (aromaticity in this case) of a compound. Different directions in such a map represent different types of similarity between the com-

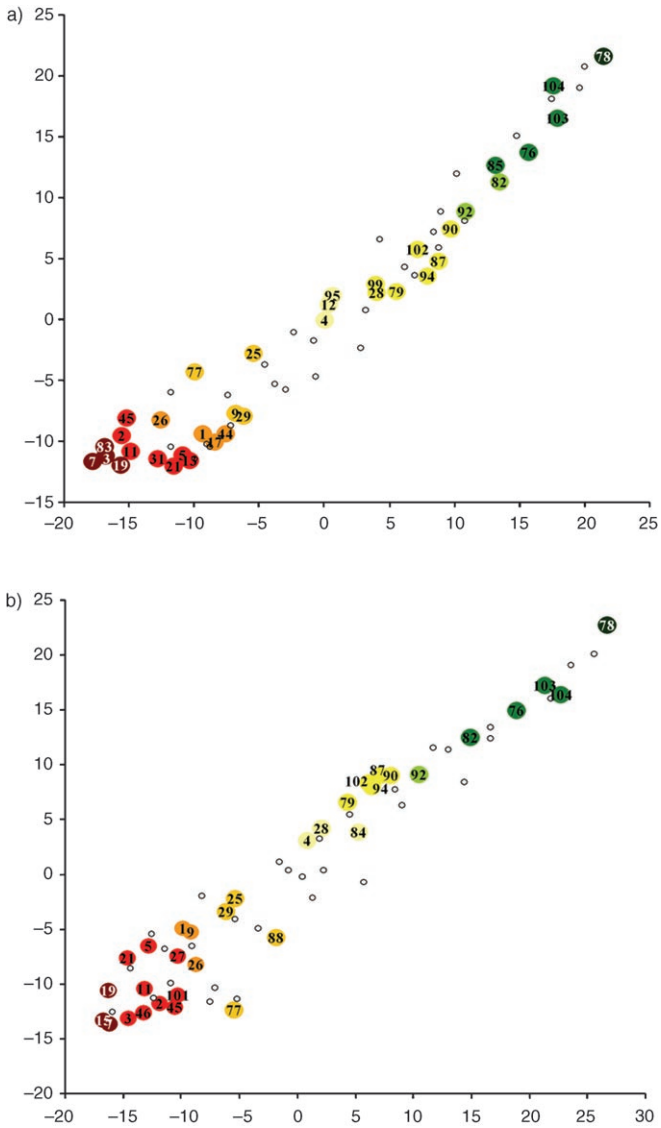


Figure 7. Sammon map obtained for the training dataset consisted of 56 compounds using: a) five descriptors including HOMA and b) four descriptors.

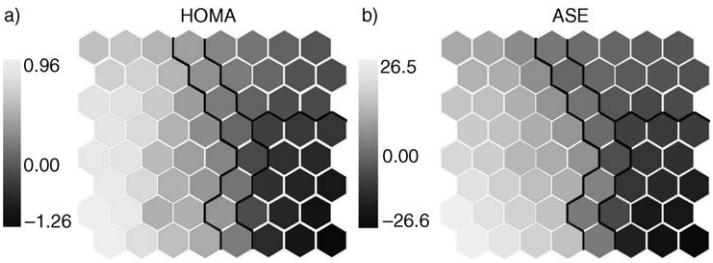


Figure 8. Component visualization of the SOM: a) ASE and b) HOMA. The values of each descriptor are represented by grey levels.

pounds and different distances indicate different degrees of similarity.

The most powerful advantages of the present method are that it allows one visually to classify organic compounds according their aromatic character and, even, to estimate quantitatively this property; additionally, it has predictive power. Since the computer program is easy to use, very fast,[57] and readily extrapolated to other aromatic topologies (different sized rings, fused rings, etc.), the methodology reported herein can be an important contribution to the long debate on the quantitative treatment of aromaticity.

Acknowledgements

We would like to thank Prof. M. K. Cyrañski, Prof. T. M. Krygowski, Prof. A. R. Katritzky, and Prof. P. von R. Schleyer, authors of the paper in reference [42], for providing the bibliographic data for five-membered rings used in the training of the network. Dr. A. Chana is thanked for helpful discussions on neural networks. We thank J. D. Guillén for helpful advice with statistics (see Supporting Information). This work was financially supported by MEyC (CTQ2004-01978). M.A. thanks MEyC for a fellowship. We also thank the Centro de Supercomputación de Galicia (CESGA) for computational time at the COMPAQ HPC 320 supercomputer.

[1] For a historical overview, see: A. T. Balaban, P. von R. Schleyer, H. S. Rzepa, *Chem. Rev.* **2005**, *105*, 3436–3447.
[2] For a more detailed definition of aromaticity, see: P. von R. Schleyer, H. J. Jiao, *Pure Appl. Chem.* **1996**, *68*, 209–218.
[3] a) T. M. Krygowski, M. K. Cyrański, Z. Czarnocki, G. Hafelinger, A. R. Katritzky, *Tetrahedron* **2000**, *56*, 1783–1796; b) A. T. Balaban, D. C. Oniciu, A. R. Katritzky, *Chem. Rev.* **2004**, *104*, 2777–2812; c) T. M. Krygowski, B. T. Stepien, *Chem. Rev.* **2005**, *105*, 3482–3512.
[4] At the beginning of organic chemistry, aromatic compounds were classified according to their odor and smell (biological properties). Since many aromatic compounds possess biological activity, it is likely that aromatic character influences the biological properties, for some discussions, see: a) V. G. S. Box, F. Jean-Mary, *J. Mol. Model.* **2001**, *7*, 334–342; b) N. Zacharias, D. A. Dougherty, *Trends Pharmacol. Sci.* **2002**, *23*, 281–287; c) P. Cysewski, *J. Mol. Struct.* **2005**, *714*, 29–34.
[5] For a review dealing with the aromaticity of all-metal clusters, see: A. I. Boldyrev, L. S. Wang, *Chem. Rev.* **2005**, *105*, 3716.
[6] P. J. Garratt, *Aromaticity*, 3rd ed., Wiley, New York, **1986**.
[7] While aromaticity is a property of a molecule in its ground state; reactivity depends on the energy difference between ground and transition state; therefore, no reliable criteria based on reactivity can be established; although some attempts have been done, see: I. Morao, I. H. Hillier, *Tetrahedron Lett.* **2001**, *42*, 4429–4431.
[8] The aromatic character of a transition state can be used to define a reaction mechanism, for selected examples, see: E. Matito, J. Poater, M. Durán, M. Solá, *J. Mol. Struct. Theochem* **2005**, *727*, 165–171, and references therein.
[9] M. K. Cyrański, *Chem. Rev.* **2005**, *105*, 3773–3811.
[10] Resonance energy (RE): a) L. Pauling, G. W. Wheland, *J. Chem. Phys.* **1933**, *1*, 362–374; b) for a critical comparison of RE and ASE, see: M. K. Cyrański, P. von R. Schleyer, T. M. Krygowski, H. J. Jiao, G. Hohlneicher, *Tetrahedron* **2003**, *59*, 1657–1665; c) for recent developments, see: I. Fishtik, R. Datta, *J. Phys. Chem. A* **2003**, *107*, 10471; d) Y. R. Mo, P. von R. Schleyer, *Chem. Eur. J.* **2006**, *12*, 2009–2020.
[11] Aromatic stabilization energy (ASE) using homodesmotic reactions: P. George, M. Trachtman, C. W. Bock, A. M. Brett, *J. Chem. Soc. Perkin Trans. 2* **1976**, 1222–1227.
[12] Aromatic stabilization energy (ASE) using isodesmotic reactions: C. H. Suresh, N Koga, *Chem. Phys. Lett.* **2006**, *419*, 550–556; and references therein.
[13] Isomerization stabilization energy (ISE): P. von R. Schleyer, F. Pülhofer, *Org. Lett.* **2002**, *4*, 2873–2876.

[14] Energetic criteria based on graph theory: a) I. Gutman, M. Milun, N. Trinajstič, *J. Am. Chem. Soc.* **1977**, *99*, 1692; b) M. Randič, *Chem. Rev.* **2003**, *103*, 3449; A. T. Balaban, M. Randič, *J. Mathem. Chem.* **2005**, *17*, 443–453; c) for a related approach, named circuit resonance energy, see: J. Aihara, *J. Am. Chem. Soc.* **2006**, *128*, 2873–2879.

[15] T. Heine, C. Corminboeuf, G. Seifert, *Chem. Rev.* **2005**, *105*, 3889–3910.

[16] Nucleus-independent chemical shifts (NICS): a) P. von R. Schleyer, C. Maerker, A. Dransfeld, H. Jiao, N. J. R. van Eikema Hommes, *J. Am. Chem. Soc.* **1996**, *118*, 6317–6318; b) NICS (1) is the nucleus-independent chemical shift at points 1 Å above the ring center: P. von R. Schleyer, H. J. Jiao, N. Hommes, V. G. Malkin, O. L. Malkina, *J. Am. Chem. Soc.* **1997**, *119*, 12669–12670; c) for a comprehensive review on the use of the NICS as aromaticity criteria, see: Z. Chen, C. S. Wannere, C. Corminboeuf, R. Puchta, P. von R. Schleyer, *Chem. Rev.* **2005**, *105*, 3842–3888. d) Recently, the plot of the NICS versus the distance (NICS-scan) has been reported as aromaticity criteria: A. Stanger, *J. Org. Chem.* **2006**, *71*, 883–893.

[17] Magnetic susceptibility exaltation: H. J. Dauben, J. D. Wilson, J. L. Layti, *J. Am. Chem. Soc.* **1968**, *90*, 811–813.

[18] Current density: P. Lazzeretti, *Prog. Nucl. Magn. Reson. Spectrosc.* **2000**, *36*, 1–88.

[19] a) J. Poater, M. Duran, M. Sola, B. Silvi, *Chem. Rev.* **2005**, *105*, 3911–3947; b) G. Merino, A. Vela, T. Heine, *Chem. Rev.* **2005**, *105*, 3812–3841.

[20] Electron localization function (ELF): a) A. D. Becke, K. E. Edgecombe, *J. Chem. Phys.* **1990**, *92*, 5397–5403; b) D. B. Chesnut, L. Bartolotti, *Chem. Phys.* **2000**, *253*, 1–11; c) B. Silvi, *Phys. Chem. Chem. Phys.* **2004**, *6*, 256–260; d) J. C. Santos, J. Andres, A. Aizman, P. Fuentealba, *J. Chem. Theory Comp.* **2005**, *1*, 83–86.

[21] Para-delocalization index (PDI): a) J. Poater, X. Fradera, M. Duran, M. Solà, *Chem. Eur. J.* **2003**, *9*, 400–406; b) J. Poater, X. Fradera, M. Duran, M. Solà, *Chem. Eur. J.* **2003**, *9*, 1113–1122.

[22] Aromatic fluctuation index (FLU): E. Matito, M. Duran, M. Solà, *J. Chem. Phys.* **2005**, *122*, 014109.

[23] Topography of the molecular electrostatic potential at critical points: C. H. Suresh, S. R. Gadre, *J. Org. Chem.* **1999**, *64*, 2505–2512.

[24] T. M. Krygowski, M. K. Cyrański, *Chem. Rev.* **2001**, *101*, 1385–1419.

[25] Structural aromaticity index: C. W. Bird, *Tetrahedron* **1985**, *41*, 1409–1414.

[26] Harmonic oscillator model of aromaticity (HOMA): a) J. Kruszewski, T. M. Krygowski, Tetrahedron Lett. **1972**, 3839–3842; b) T. M. Krygowski, *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 70–78.

[27] The multidimensional character of aromaticity has been noted in several papers, see: a) A. R. Katritzky, M. Karelson, S. Sild, T. M. Krygowski, K. Jug, *J. Org. Chem.* **1998**, *63*, 5228–5231; b) I. Alkorta, J. Elguero, *New J. Chem.* **1999**, *23*, 951–954; c) J. Van Droogenbroeck, C. Van Alsenoy, F. Blockhuys, *J. Phys. Chem. A* **2005**, *109*, 4847; and ref. [16c].

[28] J. W. Steed, J. L. Atwood, *Supramolecular Chemistry*, Wiley, Chichester, **2000**.

[29] In a broad sense, supramolecular chemistry deals with many different aspects, both in solid and solution states, including host–guest interactions, activation/deactivation of bio-macromolecules, and crystal packing. While the subject of chemistry is the molecule, the subject of supramolecular chemistry is the assembly of molecules, which has been termed "supramolecule" or "supermolecule" (G. R. Desiraju, *Angew. Chem.* **1995**, *107*, 2541–2558; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 2311–2327).

[30] a) O. Almarsson, M. J. Zaworotko, *Chem. Commun.* **2004**, 1889–1896; b) R. Paulini, K. Müller, F. Diederich, *Angew. Chem.* **2005**, *117*, 1820–1839; *Angew. Chem. Int. Ed.* **2005**, *44*, 1788–1805.

[31] a) V. Balzani, A. Credi, F. M. Raymo, J. F. Stoddart, *Angew. Chem.* **2000**, *112*, 3484–3491; *Angew. Chem. Int. Ed.* **2000**, *39*, 3348–3391; b) D. N. Reinhoudt, M. Crego-Calama, *Science* **2002**, *295*, 2403–2407.

[32] a) C. A. Hunter, K. R. Lawson, J. Perkins, C. J. Urch, *J. Chem. Soc. Perkin Trans. 2* **2001**, 651–669; b) E. A. Meyer, R. K. Castellano, F. Diederich, *Angew. Chem.* **2003**, *115*, 1244–1287; *Angew. Chem. Int. Ed.* **2003**, *42*, 1210–1250; c) M. Nishio, *CrystEngComm* **2004**, *6*, 130–158.

[33] M. Guell, J. Poater, J. M. Luis, O. Mo, M. Yañez and M. Solà, *ChemPhysChem.* **2005**, *6*, 2552–2561.

[34] a) A. Morcuende, M. Ors, S. Valverde, B. Herradon, *J. Org. Chem.* **1996**, *61*, 5264–5270; b) F. Sanchez-Sancho, E. Mann, B. Herradon, *Adv. Synth. Catal.* **2001**, *343*, 360–368; c) A. Montero, F. Albericio, M. Royo, B. Herradon, *Org. Lett.* **2004**, *6*, 4089–4092; and references therein.

[35] a) N. J. Heaton, P. Bello, B. Herradon, A. del Campo, J. Jimenez-Barbero, *J. Am. Chem. Soc.* **1998**, *120*, 12371–12384; b) E. Mann, A. Montero, M. A. Maestro, B. Herradon, *Helv. Chim. Acta* **2002**, *85*, 3624–3638; c) E. Mann, J. Mahia, M. A. Maestro, B. Herradon, *J. Mol. Struct.* **2002**, *641*, 101–107; d) A. Chana, M. A. Concejero, M. de Frutos, M. J. Gonzalez, B. Herradon, *Chem. Res. Toxicol.* **2002**, *15*, 1514–1526; e) B. Herradon, A. Montero, E. Mann, M. A. Maestro, *CrystEngComm* **2004**, *6*, 512–521; and references therein.

[36] a) E. Mann, A. Chana, F. Sanchez-Sancho, C. Puerta, A. Garcia-Merino, B. Herradon, *Adv. Synth. Catal.* **2002**, *344*, 855–867; b) J. M. Navas, A. Chana, B. Herradon, H. Segner, *Environ. Toxicol. Chem.* **2003**, *22*, 830–836; c) A. Montero, E. Mann, A. Chana, B. Herradon, *Chem. Biodiversity* **2004**, *1*, 442–457; d) A. Montero, M. Alonso, E. Benito, A. Chana, E. Mann, J. M. Navas, B. Herradon, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2753–2757.

[37] For a comprehensive, didactic book, see: J. Zupan, J. Gasteiger, *Neural Networks for Chemists. An Introduction,* Wiley-VCH, Weinheim, **2003**.

[38] For recent applications, see: a) J. Gasteiger, A. Teckentrup, L. Terfloth, S. Spycher, *J. Phys. Org. Chem.* **2003**, *16*, 232–245; b) R. Guha, J. R. Serra, P. C. Jurs, *J. Mol. Graphics Modell.* **2004**, *23*, 1–14; c) S. Spycher, E. Pellegrini, J. Gasteiger, *J. Chem. Inf. Comput. Sci. J. Chem. Inf. Mod.* **2005**, *45*, 200–208; d) J. Aires-de-Sousa, J. Gasteiger, *J. Comb. Chem.* **2005**, *7*, 298–301; e) R. Guha; P. C. Jurs, *J. Chem. Inf. Comput. Sci. J. Chem. Inf. Model.* **2005**, *45*, 800–806; f) A. Ceroni, P. Frasconi, G. Pollastri, *Neural Networks* **2005**, *18*, 1029–1039; g) J. Fayos, L. Infantes, F. H. Cano, *Cryst. Growth Des.* **2005**, *5*, 191–200; h) M. Randić, D. Butina, J. Zupan, *Chem. Phys. Lett.* **2006**, *419*, 528–532; i) A. Asikainen, M. Kolehmainen, J. Ruuskanen, K. Tuppurainen, *Chemosphere* **2006**, *62*, 658–673.

[39] Recently, the application of Kohonen neural networks has been used to classify and cluster a set of 88 derivatives of benzene, naphthalene, and phenanthrene according to their electronic structures, see: J. J. Panek, A. Jezierska, M. Vračko, *J. Chem. Inf. Mod.* **2005**, *45*, 264–272.

[40] S. Haykin, *Neural Networks. A Comprehensive Foundation*, 2nd ed., Prentice-Hall, New Jersey, **1999**.

[41] T. Kohonen, *Self-Organizing Maps*, 3rd ed., Springer, Berlin, **2001**.

[42] M. K. Cyrański, T. M. Krygowski, A. R. Katritzky, P. von R. Schleyer, *J. Org. Chem.* **2002**, *67*, 1333–1338.

[43] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle and J. A. Pople, Gaussian 98, revison A3, Pittsburgh PA, **1998**.

[44] T. A. Keith, R. F. W. Bader, *Chem. Phys. Lett.* **1993**, *210*, 223–231.

[45] I. Morao, B. Lecea, F. P. Cossio, *J. Org. Chem.* **1997**, *62*, 7033–7036.

[46] K. Wolinski, J. F. Hinton, P. Pulay, *J. Am. Chem. Soc.* **1990**, *112*, 8251–8260.

[47] T. Kohonen, J. Hynninen, J. Kangas, J. Laaksonen, *SOM PAK: The Self-Organizing Map program package*, version 3.1, Helsinki University of Technology, Laboratory of Computer and Information Science **1995**; web address: http://citeseer.ist.psu.edu/kohonen96 som.html (last accessed on July 27th, 2006).
[48] This self-organizing map can be supplied on request.
[49] L. Chen, J. Gasteiger, *J. Am. Chem. Soc.* **1997**, *119*, 4033–4042.
[50] J. S. Sammon, *IEEE Trans. Computers* **1969**, *C-18*, 401–409.
[51] a) A. Dransfeld, L. Nyulaszi, P. von R. Schleyer, *Inorg. Chem.* **1998**, *37*, 4413–4420; b) L. Nyulaszi, *Chem. Rev.* **2001**, *101*, 1229–1246.
[52] P. von R. Schleyer, P. K. Freeman, H. Jiao, B. Goldfuss, *Angew. Chem.* **1995**, *107*, 332–335; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 337–340.
[53] V. I. Minkin, M. N. Glukhovtsev, B. Y. Simkin, *Aromaticity and Antiaromaticity: Electronic and Structural Aspects*, Wiley, New York, **1994**.
[54] L. Nyulászi, P. von R. Schleyer, *J. Am. Chem. Soc.* **1999**, *121*, 6872–6875.
[55] Recently, Mo and Schleyer have reported[10d] that cyclopentadiene has an extra cyclic resonance energy (ECRE) of 7.7 kcal mol^{-1}, which indicates a certain aromaticity character, but which is quite small as compared with aromatic compounds, such a cyclopentadienyl anion (42.8 kcal mol^{-1}) or pyrrole (34.2 kcal mol^{-1}). It is worth noting that although cyclopentadiene is placed in the region of non-aromatic compounds, its neuron is the nearest to the aromatic region.
[56] A. Stanger, *Chem. Eur. J.* **2006**, *12*, 2745–2751.
[57] The generation of each network requires a very short computational time (less than 1 min).

Received: August 30, 2006
Published online: February 26, 2007

