9th International Electronic Conference on Synthetic Organic Chemistry. ECSOC-9. 1-30 November 2005.

http://www.usc.es/congresos/ecsoc/9/ECSOC9.HTM & http://www.mdpi.net/ecsoc/

[G011]

Periodic Table of Local Anaesthetics (Procaine Analogues)

Francisco Torrens[1*] and Gloria Castellano[2]

[1] Institut Universitari de Ciència Molecular, Universitat de València, Dr. Moliner 50, E 46100 Burjassot (València), Spain. . Tel. +34 963 543 182, Fax +34 963 543 156, E mail Francisco.Torrens@uv.es

[2] Facultad de Ciencias Experimentales, Universidad Católica de Valencia San Vicente Mártir, Guillem de Castro 106, E 46003 València, Spain

Abstract: *Algorithms for classification and taxonomy are proposed based on criteria, e.g. information entropy and its production. The feasibility of replacing a given anaesthetic by similar ones in the composition of a complex drug is studied. Some local anaesthetics currently in use are classified using characteristic chemical properties of different portions of their molecules. Many classification algorithms are based on information entropy. When applying these procedures to sets of moderate size, an excessive number of results appear compatible with data and this number suffers a combinatorial explosion. However, after the equipartition conjecture, one has a selection criterion between different variants resulting from classification between hierarchical trees. According to this conjecture, for a given charge or duty, the best configuration of a flowsheet is the one in which the entropy production is most uniformly distributed. Information entropy analysis of the structural parameters and principal component analysis (PCA) of the anaesthetics permit classifying them and agree. A periodic table of anaesthetics is built. The periodic law has not the rank of the laws of physics: (1) the properties of anaesthetics are not repeated; (2) the order relationships are repeated with exceptions. The proposed statement is: The relationships that any anaesthetic p has with its neighbour p + 1 are approximately repeated for each period.*

Keywords: Periodic property; Periodic table; Periodic law; Classification; Information entropy; Equipartition conjecture; Principal component analysis; Cluster analysis; Structural parameter; Local anaesthetic; Procaine analogue

Introduction

For hundreds of years, surgeons had to work fast to minimize shock and pain to their patients [1]. They gave alcohol, *e.g*. rum, to their patients to drink to deaden the pain, and internal operations were not possible. The original inhalation anaesthetic, diethyl ether, was first used by a Massachusetts dentist, W. T. G. Morton (1846); James Simpson, a Scottish doctor, used chloroform on a patient during an operation (1847). Nitrous oxide (N_2O) *laughing gas* was also used as an anaesthetic (XIX century). Anaesthetics really became widely accepted after John Snow gave chloroform to Queen Victoria during the birth of Prince Leopold (1853). There are two types of anaesthetics, *viz*. local (*i.e*. surface acting) nerve blocking anaesthetics for use in *e.g*. dentistry as well as epidurals for use in *e.g*. childbirth,

and general (which act on the whole body). For operations today, a combination of anaesthetics is used to afford muscle relaxation, *e.g*. curare, a general anaesthesia, *e.g*. N_2O/O_2 mixture, followed by halothane and analgesics for pain relief. Of the traditional anaesthetics, chloroform, which can cause liver damage, has only a small margin between its anaesthetic dose and a fatal one, while ether is inflammable and N_2O is safe but does not produce deep anaesthesia. Chemists have developed better anaesthetics (World War II). What chemists look for in a good anaesthetic are: (1) a wide safety margin between effective and toxic dose, (2) chemically unreactive, (3) no unpleasant side effects like a bad taste or nausea when waking, (4) no side effects on the body, especially vital organs like the heart, (5) non flammable, (6) fast acting and (7) low boiling point. Having halogen atoms in the molecule helps promote anaesthesia and also reduces inflammability, but too much Cl atoms cause toxicity. The presence of F atoms helps to reduce the boiling point and also increase stability, because the C–F bond is so strong. Halothane (fluothane) has been used regularly as an anaesthetic in UK hospitals (1957). Anaesthetics block the passage of nerve impulses to the brain. The more soluble a particular inhalation anaesthetic is in hydrocarbons and fats (lipids), the better it is as an anaesthetic. The inhaled anaesthetic probably dissolves in the lipid-rich neuronal cell membranes near the nerve endings, and changes their volume or fluidity. Normally, when the resting cell receives an electrochemical signal, a membrane opens and sodium ions pass through to the interior of the cell, allowing the transmission of a nerve impulse. The dissolved anaesthetic prevents the channel opening when the message is received and, therefore, blocks the signal to the brain.

Local anaesthetics are amphiphile molecules of tertiary amines, and some of them have colloidal properties in aqueous solution. They are classified into the ester type and amide type by the difference in the chain that binds the hydrophobic group and the hydrophilic group in a molecule. The anaesthetic potency of these drugs is significantly dependent on the hydrophobicity of the molecules. The great majority of the local anaesthetics currently used in medical practice have in common a lipophilic portion, generally an aromatic system, an intermediate aliphatic C chain and a hydrophilic portion, frequently the substituted amine group. Although the mechanism of action at the molecular level is not fully cleared [2,3], it was considered that the balance between the lipophilic and hydrophilic portions influence significantly the biological activity, modulating its local anaesthetic potency [4,5]. On the other hand, some authors have considered that the electronic distribution of the carbonyl group C=O, present in the majority of local anaesthetics, has an important role for the establishment of this activity. Thus, it is proposed that substituent groups present in the aromatic ring affect the local anaesthetic activity, by its effects hydrophobic and of polar nature. Once it is known that the inductive and resonance effects affect directly the electronic density on the carbonylic O atom, as consequence it is proposed that the carbonyl group polarity can be, in principle, modulating the local anaesthetic

activity. The biological activity of drugs, in particular local anaesthetics, can be considered as the result of the interactions of these with the biophase. The drug–receptor interactions depend, by its turn, of the physicochemical properties of the compound, determining and modulating the forces of chemical nature present in these interactions.

A simple computerized algorithm useful for establishing a relationship between chemical structures and their biological activities or significance is proposed and exemplified here (*cf*. Reference 6 or 7 for a review). The starting point is to use an informational or configurational entropy for pattern recognition purposes. This entropy is formulated on the basis of a *matrix of similarity* between two chemical or biochemical species. The presented example shows a classification of local anaesthetics on the basis of their similarity with procaine [8]. As entropy is weakly discriminating for classification purposes, the more powerful concept of *entropy production* and the *equipartition conjecture of entropy production* are introduced [9]. Learning potentialities of the code have also been developed.

Computational Method

The key problem in classification studies is to define *similarity indices* when several criteria of comparison are involved. The first step in quantifying the concept of similarity for molecules of local anaesthetics is to list the most important portions of such molecules. Furthermore, the *vector of properties* $\vec{i} = \langle i_1, i_2, \ldots i_k, \ldots \rangle$ should be associated to every local anaesthetic *i*, whose components correspond to different characteristic groups of the molecule of anaesthetic, in a hierarchical order according to the expected importance of their pharmacological potency. If the *m th* portion of the molecule is pharmacologically more significant for the anaesthetic effect than the *k th* portion, then $m < k$. The components i_k are "1" or "0" according to whether a similar (or identical) portion of rank *k* is present or absent in anaesthetic *i* compared with the reference anaesthetic. Our analysis includes such chemical compounds that fit the following general scheme: (lipophilic portion)–(intermediate chain)–(hydrophilic portion), since these are the most numerous and have the widest range of uses among the species used in practice of local anaesthesia [10]. The lipophilic portion normally consists of at least one phenyl radical, the hydrophilic portion is most often a secondary or tertiary amine, and the intermediate chain commonly has an ester or amide linkage [11]. It is assumed that the *structural elements* of a local anaesthetic molecule can be *ranked*, according to their contribution to anaesthetic potency, in the following order of decreasing importance: lipophilic portion > hydrophilic portion > intermediate chain > number of nitrogen atoms > number of oxygen atoms. Procaine was selected as a *reference* anaesthetic in this work. In the case of procaine, the lipophilic portion is a phenyl radical, the hydrophilic portion is an amine, the intermediate chain is an ester, there are two N atoms and two O atoms. Obviously the vector associated to procaine is <11111>. The vector <11110> is associated to benoxinate (*cf*. Table 1) since there are three oxygens in this case. The vector <10101>

is associated to benzocaine since the hydrophilic partition is not an amine and there is one nitrogen in this case. Table 1 contains the vectors associated in the same manner to 27 anaesthetic agents.

Table 1. Vector properties of local anaesthetics analogues of procaine.

1. benoxinate	<11110>	15. lidocaine	<11010>
2. benzocaine	<10101>	16. mepivacaine	<11010>
3. bupivicaine	<11010>	17. piperocaine	<11101>
4. butacaine	<11111>	18. pramoxine	<11000>
5. butamben	<10101>	19. prilocaine	<11010>
6. 2 chloroprocaine	<11111>	20. procaine	<11111>
7. cocaine	<11100>	21. proparacaine	<11110>
8. cyclomethycaine	<11100>	22. propoxycaine	<11110>
9. dibucaine	<01001>	23. tetracaine	<11111>
10. dimethisoquin	<01010>	24. tocainide	<11010>
11. diperodon	<11000>	25. mexiletine	<11000>
12. dyclonine	<11001>	26. propanolol	<01001>
13. etidocaine	<11010>	27. phenytoin	<10011>
14. hexylcaine	<11101>		

Let us denote by r_{ij} ($0 \le r_{ij} \le 1$) the similarity index of two anaesthetics associated to the vectors $\vec{i}$ and $\vec{j}$, respectively. The relation of similitude is characterized by a *similarity matrix* $\bar{\bar{R}} = [r_{ij}]$. The similarity index between two anaesthetics $\vec{i} = <i_1, i_2, \ldots i_k, \ldots>$ and $\vec{j} = <j_1, j_2, \ldots j_k, \ldots>$ is defined as:

$$r_{ij} = \sum_k t_k (a_k)^k \quad (k = 1, 2, \ldots) \qquad (1)$$

where $0 \le a_k \le 1$ and $t_k = 1$ if $i_k = j_k$, but $t_k = 0$ if $i_k \neq j_k$ $t_k = 0$ if $i_k \neq j_k$. This definition assigns a weight $(a_k)^k$ to any property involved in the description of molecules i or j.

Classification Algorithm

The *grouping algorithm* uses the *stabilized* matrix of similarity, obtained by applying the *max-min composition rule* o defined by:

$$\left(\bar{\bar{R}} o \bar{\bar{S}}\right)_{ij} = \max_k \left[\min_k \left(r_{ik}, s_{kj}\right)\right] \qquad (2)$$

where $\bar{\bar{R}} = [r_{ij}]$ and $\bar{\bar{S}} = [s_{ij}]$ are matrices of the same type, and $\left(\bar{\bar{R}} o \bar{\bar{S}}\right)_{ij}$ is the (i,j) th element of the matrix $\bar{\bar{R}} o \bar{\bar{S}}$ [12]. It can be shown that when applying this rule iteratively so that $\bar{\bar{R}}(n+1) = \bar{\bar{R}}(n) o \bar{\bar{R}}$, there exists an integer n such that: $\bar{\bar{R}}(n) = \bar{\bar{R}}(n+1) = \cdots$ The resulting matrix $\bar{\bar{R}}(n)$ is called the *stabilized similarity matrix*. The importance of stabilization lies in the fact that in the classification process, it will generate a partition in disjoint classes. From now on, it is understood that the stabilized matrix is used and designated by $\bar{\bar{R}}(n) = [r_{ij}(n)]$. The *grouping rule* is the following: i and j are assigned to the same class if $r_{ij}(n) \geq b$. The class of i noted $\hat{i}$ is the set of species j that satisfies the rule $r_{ij}(n) \geq b$. The matrix of classes is:

$$\bar{\bar{R}}(n) = [r_{\hat{i}\hat{j}}] = \max_{s,t}(r_{st}) \quad (s \in \hat{i},\ t \in \hat{j}) \qquad (3)$$

where s stands for any index of a species belonging to the class $\hat{i}$ (similarly for t and $\hat{j}$). Rule (3) means finding the largest similarity index between species of two different classes. The *information entropy* associated with the matrix of similarity $\bar{\bar{R}}$ is:

$$h\left(\bar{\bar{R}}\right) = -\sum_{i,j} r_{ij} \ln r_{ij} - \sum_{i,j}\left(1 - r_{ij}\right)\ln\left(1 - r_{ij}\right) \qquad (4)$$

Denote also by C_b the set of classes and by $\bar{\bar{R}}_b$ the matrix of similarity at the grouping level b. The information entropy satisfies the following properties.

1. $H\left(\bar{\bar{R}}\right) = 0$ if $r_{ij} = 0$ or $r_{ij} = 1$.
2. $H\left(\bar{\bar{R}}\right)$ is maximum if $r_{ij} = 0.5$, *i.e*., when the imprecision is maximum.
3. $H\left(\bar{\bar{R}}_b\right) \leq H\left(\bar{\bar{R}}\right)$ for any b, *i.e*., classification leads to a loss of entropy.
4. $H\left(\bar{\bar{R}}_{b_1}\right) \leq H\left(\bar{\bar{R}}_{b_2}\right)$ if $b_1 < b_2$, *i.e*., the entropy is a monotone function of the grouping level b.

The Equipartition Conjecture of Entropy Production

In the classification algorithm, every *hierarchical tree* corresponds to a dependence of entropy on the grouping level, and thus an h–b diagram can be obtained. The Tondeur and Kvaalen [9] *equipartition conjecture of entropy production* is proposed as a selection criterion among different variants resulting from classification among hierarchical trees. According to this conjecture, for a given charge or duty, the best configuration of a flowsheet is the one in which entropy production is most uniformly distributed, *i.e*. closest to a kind of equipartition. One proceeds

here by analogy using *information entropy* instead of thermodynamic entropy. Equipartition implies a linear dependence, that is a constant production of entropy along the b scale, so that the *equipartition line* is described by:

$$h_{eqp} = h_{max}\, b \tag{5}$$

Indeed, since the classification is discrete, a realistic way of expressing equipartition would be a regular staircase function. The best variant is chosen to be that minimizing the sum of squares of the deviations:

$$SS = \sum_{b_i} \left(h - h_{eqp}\right)^2 \tag{6}$$

Learning Procedure

Learning procedures similar to those encountered in *stochastic methods* are implemented as follows [13]. Consider a given partition in classes as *good* or ideal from practical or empirical observations. This corresponds to a *reference* similarity matrix $\bar{\bar{S}} = [s_{ij}]$ obtained for equal weights $a_1 = a_2 = \ldots = a$ and for an arbitrary number of fictious properties. Next consider the same set of species as in the good classification and the actual properties. The similarity degree r_{ij} is then computed with Equation (1) giving the matrix $\bar{\bar{R}}$. The number of properties for $\bar{\bar{R}}$ and $\bar{\bar{S}}$ may differ. The learning procedure consists in trying to find classification results for $\bar{\bar{R}}$ as close as possible to the *good* classification. The first weight a_1 is taken constant and only the following weights a_2, a_3,… are subjected to random variations. A new similarity matrix is obtained using Equation (1) and the new weights. The distance between the partitions in classes characterized by $\bar{\bar{R}}$ and $\bar{\bar{S}}$ is given by:

$$D = -\sum_{ij}\left(1 - r_{ij}\right)\ln\frac{1 - r_{ij}}{1 - s_{ij}} - \sum_{ij} r_{ij}\ln\frac{r_{ij}}{s_{ij}} \qquad \forall\, 0 \le r_{ij}, s_{ij} \le 1 \tag{7}$$

The result of the algorithm is a set of weights allowing adequate classification. Such a procedure has been applied in the synthesis of complex flowsheets using of information entropy [14].

Calculation Results and Discussion

In the present report 27 local anaesthetics analogues of procaine (*cf*. Table 1) have been studied. The analysis includes such chemical compounds that fit the following general scheme: lipophilic portion–intermediate chain–hydrophilic portion, since among the species used in practice of local anaesthesia, these are the most numerous and have the widest range of uses. The lipophilic portion normally consists of at least one phenyl radical; the hydrophilic portion is most often a secondary or tertiary amine; the intermediate chain commonly has an ester or amide linkage. Using the grouping rule in the drug-design case with equal weights $a_k = 0.5$, for $0.94 \le b_1 \le 0.96$ the following set of

classes are obtained:

$Cb1$ = (1,21,22)(2,5)(3,13,15,16,19,24)(4,6,20,23)(7,8)(9,26)(10)(11,18,25)(12)(14,17)(27)

The 11 classes are obtained with the associated entropy: $h\left(\overline{\overline{R}}_{b_1}\right)$ = 58.86. The dendrogram (binary tree) [15,16] matching to $<i_1,i_2,i_3,i_4,i_5>$ and C_{b1}, illustrated [17] in Figure 1, provides a binary taxonomy of Table 1, which separates the same 11 classes.

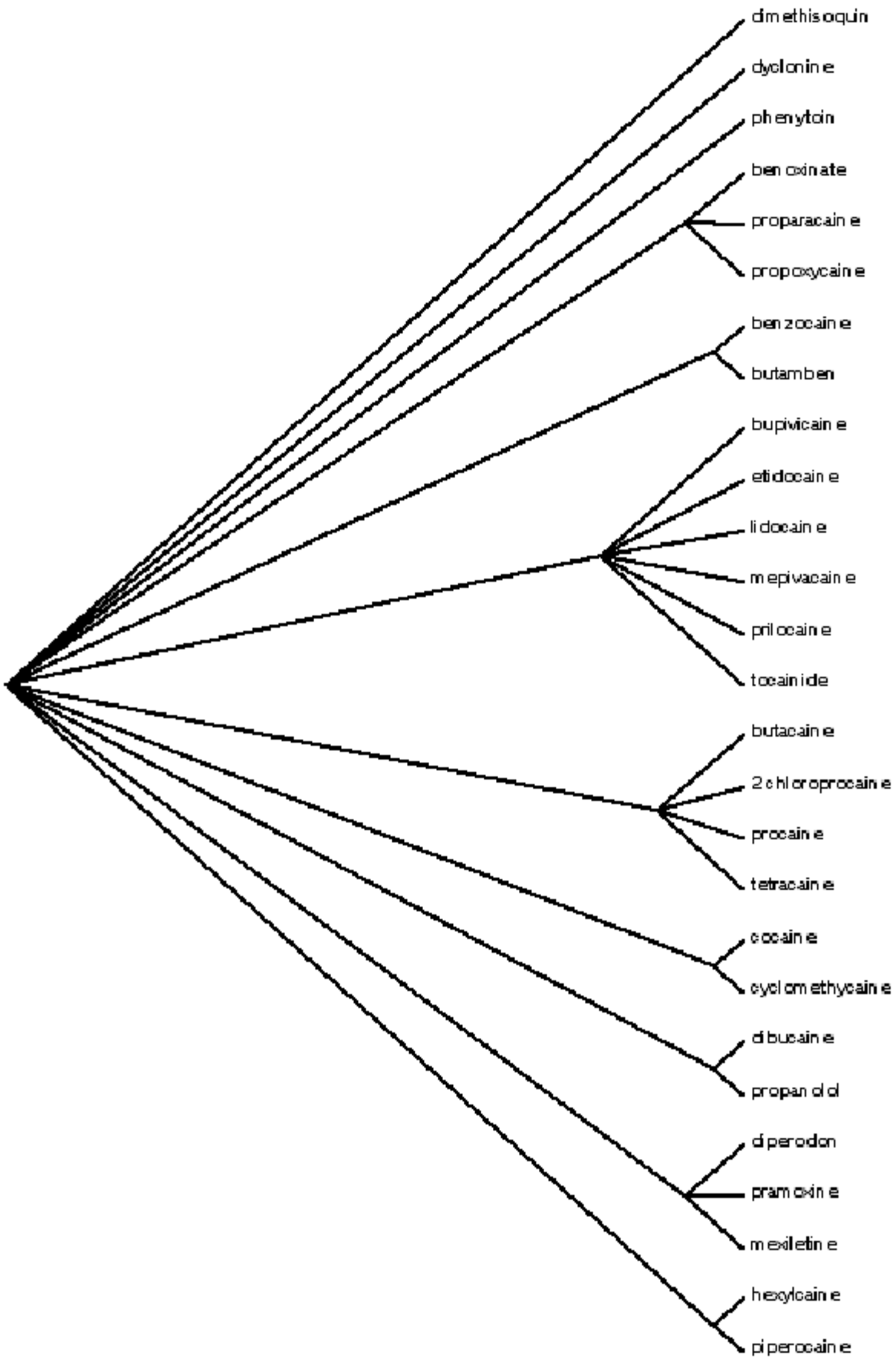

Figure 1. Dendrogram for the local anaesthetics analogues of procaine at level b_1.

The radial tree for the local anaesthetics relating to $<i_1,i_2,i_3,i_4,i_5>$ and C_{b1} (*cf*. Figure 2) separates the same 11 classes.

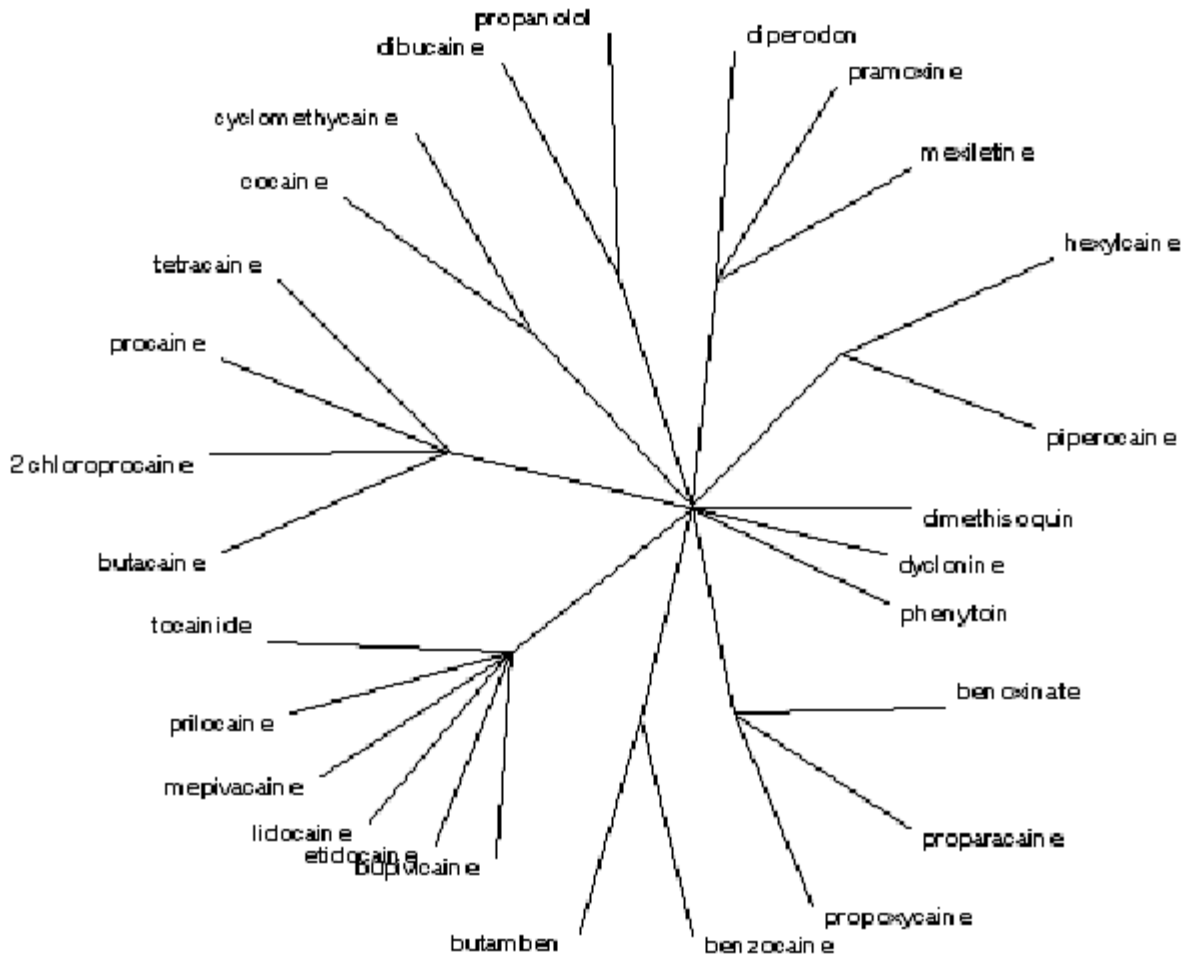

Figure 2. Radial tree for the local anaesthetics analogues of proacine at level b_1.

At level b_2 with $0.85 \leq b_2 \leq 0.87$ the set of classes is

$$Cb2 = (1,4,6,7,8,14,17,20,21,22,23)(2,5)(3,11,12,13,15,16,18,19,24,25)(9,10,26)(27)$$

The five classes result in this case and the entropy is $h\left(\bar{\bar{R}}_{b_2}\right)$ = 12.20. The dendrogram matching to $<i_1,i_2,i_3,i_4,i_5>$ and C_{b2} (*cf*. Figure 3) separates the same five classes.

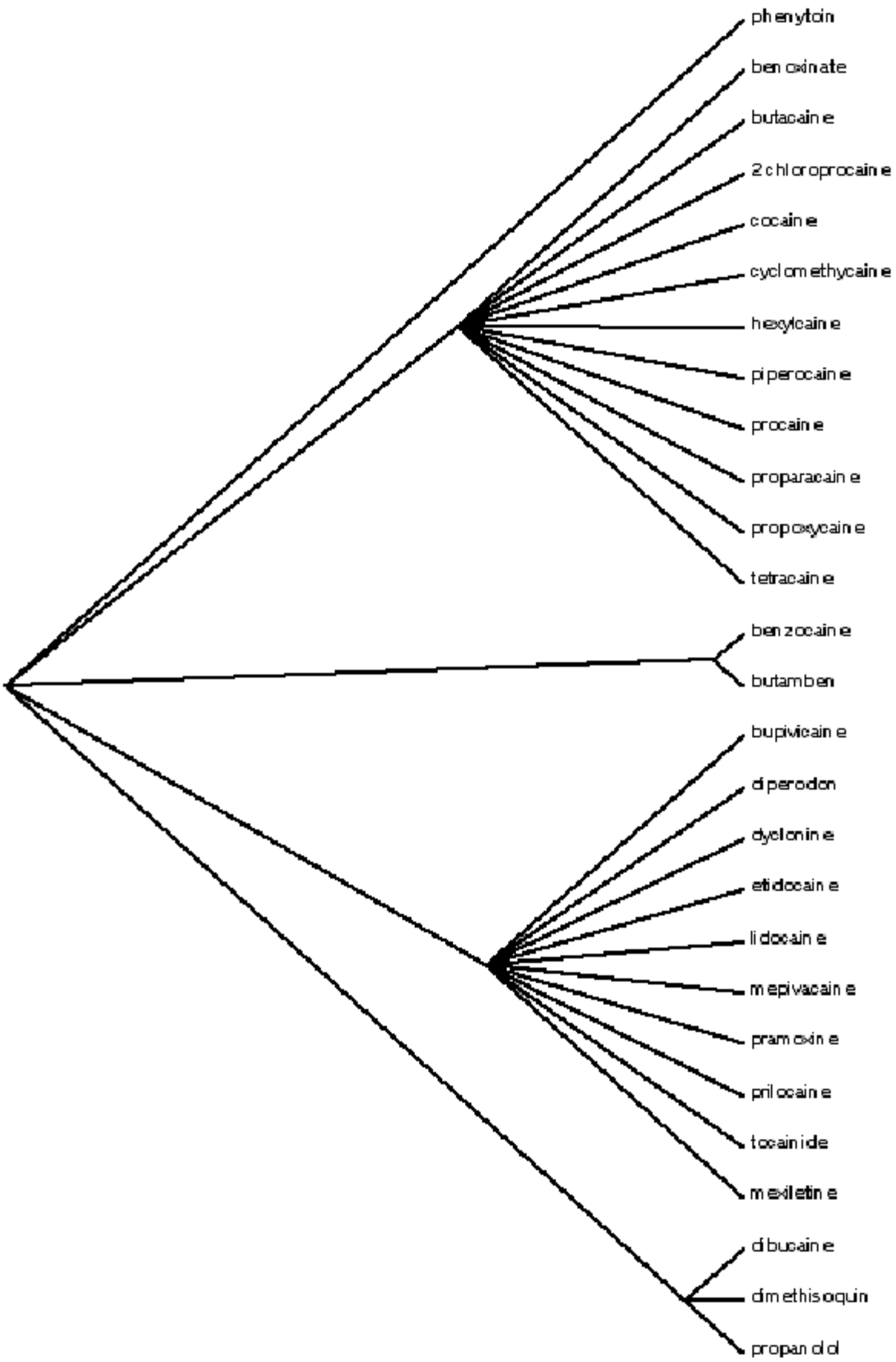

Figure 3. Dendrogram for the local anaesthetics analogues of procaine at level b_2.

The radial tree for the local anaesthetics relating to $<i_1,i_2,i_3,i_4,i_5>$ and C_{b2} (*cf*. Figure 4) separates the same five classes. Naturally anaesthetics 4, 6, 20 and 23, *i.e*., butacaine, 2 chloroprocaine, procaine and tetracaine, belong to the same class at any grouping level *b*, except at the highest level above which each class contains only one species. A high degree of similarity is found for 9 and 26, *i.e*., dibucaine and propanolol, 2 and 5, *i.e*., benzocaine and butamben. A natural trend is to interchange the similar anaesthetic agents in the composition of complex drugs. A detailed classification at level b_1 into 11 classes and a less detailed classification at a lower level b_2 into five classes can be selected, taking into account the amount of entropy variation. The resulting partition into classes could be compared with others considered as *good*. When differences appear, *e.g*. if there are five classes as in C_{b2} and if species 2 and 5 belong to different classes, an adaptive modification of the weights of properties should be considered. Therefore, assuming that the number of O atoms is unimportant with regard to the anaesthetic classification, this property is omitted and an adaptive search is executed resulting in new weights, *viz*. $a_1 = 0.50$, $a_2 = 0.53$, $a_3 = 0.46$, and $a_4 = 0.46$. Once the weights have been established, the algorithm could be used for classification of similar but as yet unclassified molecules.

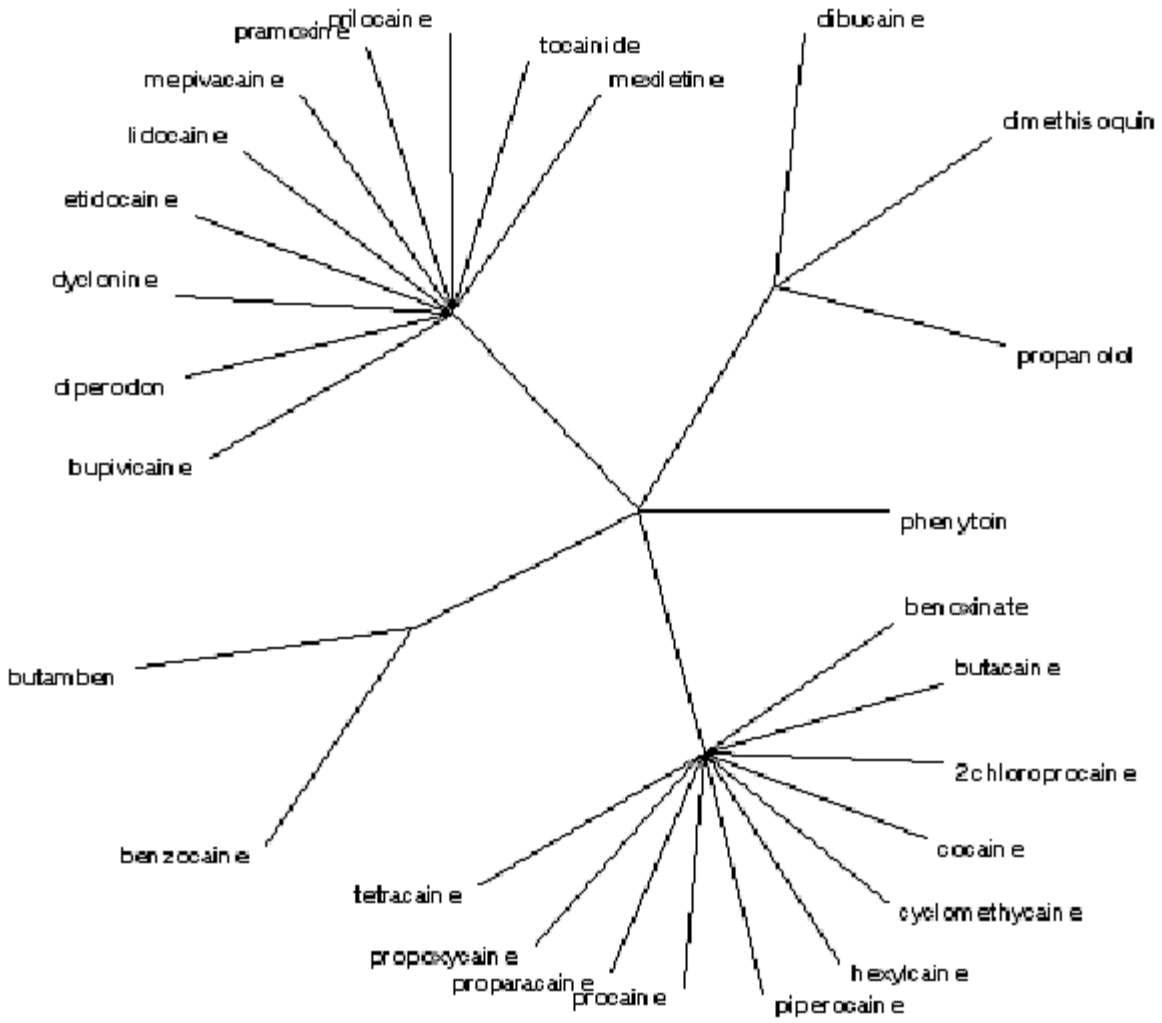

Figure 4. Radial tree for the local anaesthetics analogues of procaine at level b_2.

SplitsTree is a program for analyzing cluster analysis (CA) data [18]. Based on the method of *split decomposition*, it takes as input a *distance matrix* or a set of CA data, and produces as output a graph that represents the relationships between the taxa. For ideal data this graph is a tree, whereas less ideal data will give rise to a tree-like network, which can be interpreted as possible evidence for different and conflicting data. Further, as split decomposition does not attempt to force data onto a tree, it can provide a good indication of how *tree*-like given data are. The splits graph for the local anaesthetics (*cf*. Figure 5) reveals no conflicting relationship between the anaesthetics. Moreover, it is in agreement with both dendrograms and binary trees (Figures 1–4).

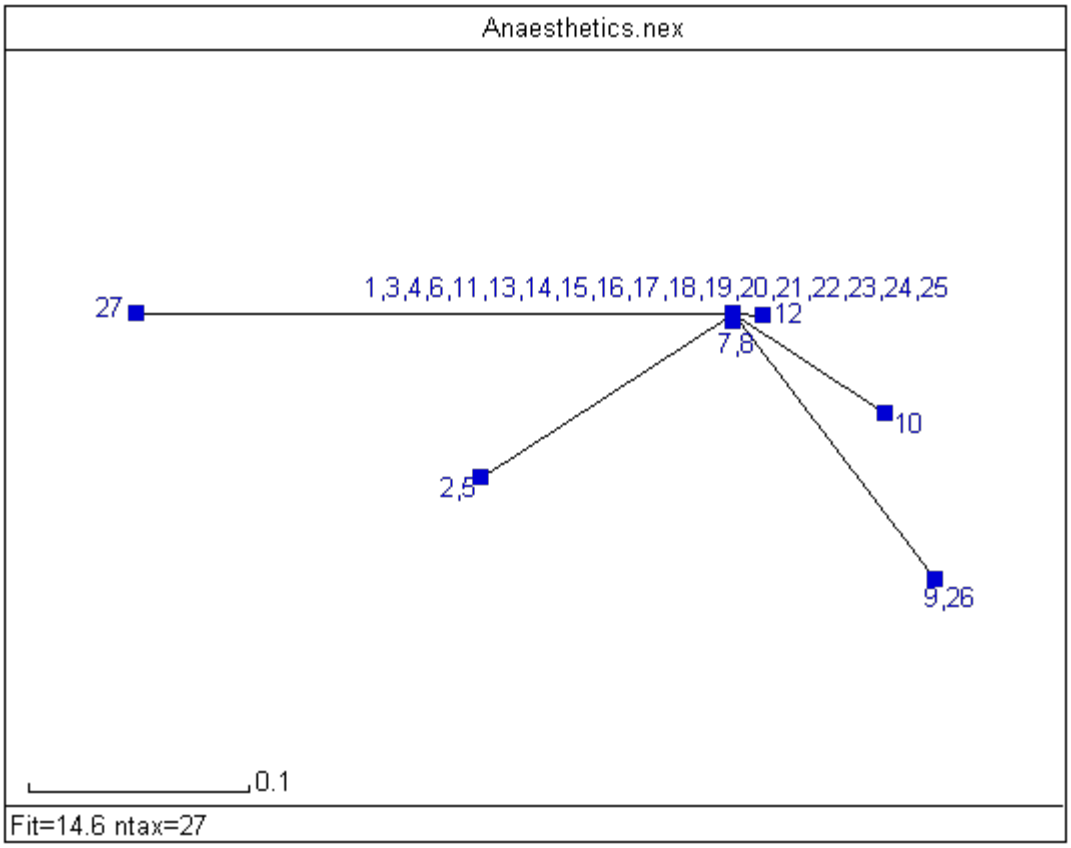

Figure 5. Splits graph for the local anaesthetics analogues of procaine.

The correlation matrix (27´27) is no meaningful because four local anaesthetics analogues of procaine, *viz*. butacaine, 2 chloroprocaine, procaine and tetracaine (Entries 4, 6, 20 and 23 in Table 1) show a vector with null standard deviation (<11111>). After elimination of these four compounds, the new correlation matrix (23´23) is properly calculated. The intercorrelations between the 23 anaesthetics are illustrated in the partial correlation diagram, which contains 25 high partial correlations $r \geq 0.75$ (*cf*. Figure 6, *red lines*), 69 medium partial correlations $0.50 \leq r < 0.75$ (*orange lines*), and 22 low partial correlations $0.25 \leq r < 0.50$ (*yellow lines*). Pairs of anaesthetics with high partial correlation appear close in dendrograms, binary trees and splits graph (Figures 1–5).

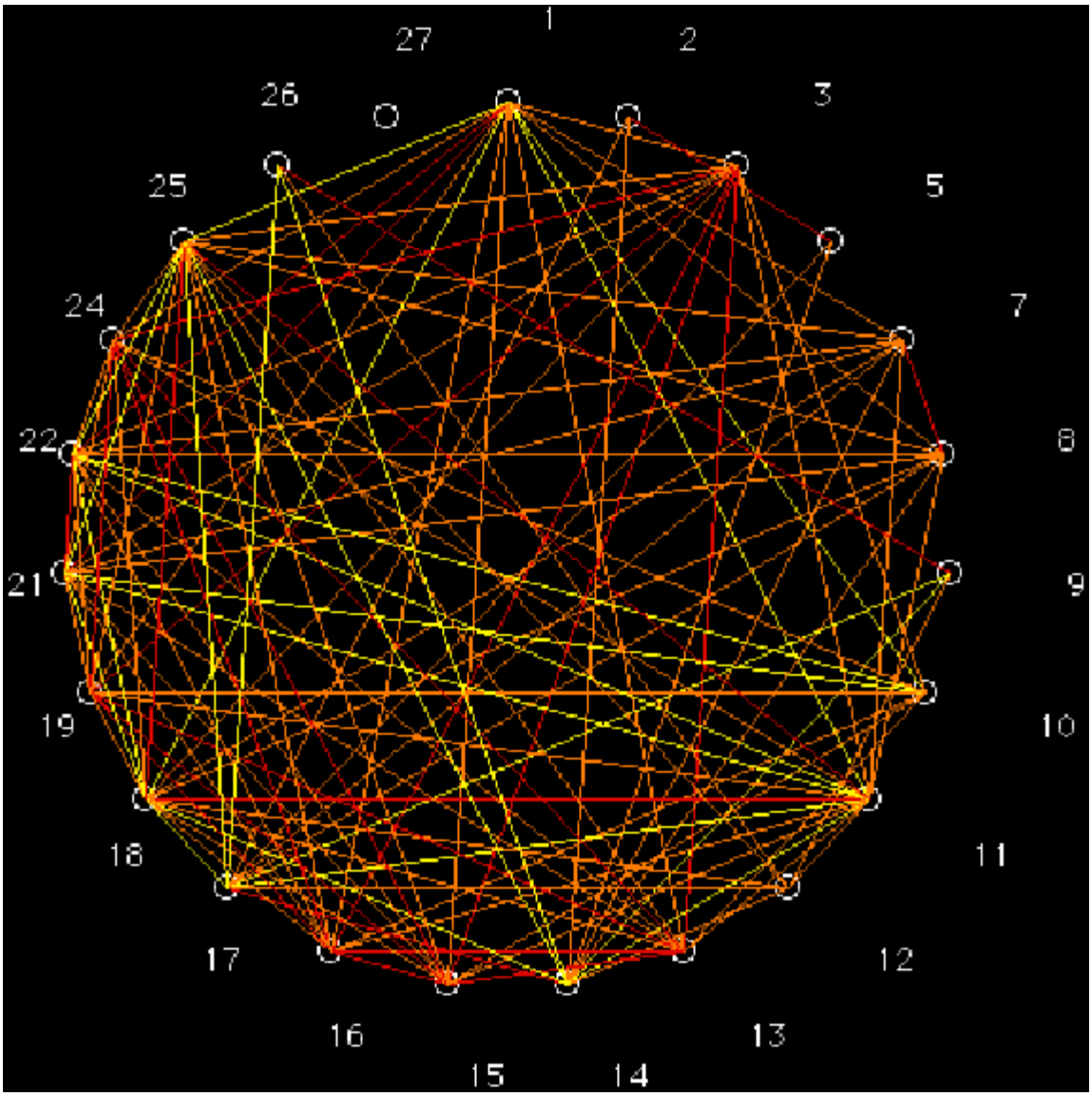

Figure 6. Partial correlation diagram: High (*red*), medium (*orange*) and low (*yellow*) correlations.

A principal component analysis (PCA) [19] has been carried out for the local anaesthetics. The importance of PCA factors $F_{1–5}$ for $\{i_1,i_2,i_3,i_4,i_5\}$ is collected in Table 2. In particular, the use of only the first factor F_1 explains 34% of the variance (66% error); the combined use of the first two factors $F_{1–2}$ explains 61% of the variance (39% error); the use of the first three factors $F_{1–3}$ explains 77% of the variance (23% error).

Table 2. Importance of the Principal Component Analysis Factors for the vector property.

Factor	**Eigenvalue**	**Percentage of variance**	**Cumulative percentage of variance**
F_1	1.69801410	33.96	33.96
F_2	1.35390023	27.08	61.04
F_3	0.81276940	16.25	77.29

F_4	0.77483947	15.50	92.79	
F_5	0.36047680	7.21	100.00	

The PCA factor loadings are shown in Table 3.

Table 3. Principal component analysis loadings for the vector property of local anaesthetics.[a]

Property	PCA factor loadings				
	F_1	**F_2**	**F_3**	**F_4**	**F_5**
i_1	-0.102	**0.737**	-0.101	0.442	0.490
i_2	0.531	0.007	**0.751**	-0.128	0.370
i_3	-0.459	0.463	0.533	-0.200	-0.501
i_4	0.340	0.442	-0.371	**-0.743**	0.000
i_5	-0.617	-0.216	0.063	-0.443	0.610

[a] Loadings greater than 0.7 are boldfaced.

The PCA $F_{1–5}$ profile for the vector property is listed in Table 4. In particular, for F_1 and F_5 variable i_5 has the greatest weight in the profile; however, F_1 cannot be reduced to three variables $\{i_2,i_3,i_5\}$ without a 13% error. For F_2 variable i_1 has the greatest weight; notwithstanding, F_2 cannot be reduced to three variables $\{i_1,i_3,i_4\}$ without a 5% error. For F_3 variable i_2 has the greatest weight; furthermore, F_3 can be reduced to three variables $\{i_2,i_3,i_4\}$ with an error of only 1%. For F_4 variable i_4 has the greatest weight; nevertheless, F_4 cannot be reduced to three variables $\{i_1,i_4,i_5\}$ without a 6% error. $F_{1–2–3–4–5}$ can be considered as linear combinations of $\{i_2,i_3,i_5\}$, $\{i_1,i_3,i_4\}$, $\{i_2,i_3,i_4\}$, $\{i_1,i_4,i_5\}$ and $\{i_1,i_3,i_5\}$ with 13%, 5%, 1%, 6% and 14% errors, respectively.

Table 4. Profile of the principal component analysis factors for the vector property.[a]

Factor	Percentage of i_1	Percentage of i_2	Percentage of i_3	Percentage of i_4	Percentg. of i_5
F_1	1.05	28.21	21.06	11.56	38.12
F_2	**54.34**	0.01	21.43	19.55	4.68
F_3	1.02	**56.46**	28.36	13.76	0.40

F_4	19.57	1.63	4.01	**55.13**	19.65
F_5	24.03	13.69	25.13	0.00	37.16

[a] Percentages greater than 50% are boldfaced.

In the F_2– F_1 plot (*cf*. Figure 7), those local anaesthetics analogues of procaine with the same vector property come out superimposed. Five classes of anaesthetics are clearly distinguished: class 1 with 11 units ($0 \approx F_1 < F_2$, *top*), class 2 (10 units, $F_1 > F_2$, *right*), class 3 (2 units, $F_1 << F_2 \approx 0$, *left*), class 4 (1 unit, $-1 \approx F_1 < F_2 \approx 0$, *middle*) and class 5 (3 units, $F_1 >> F_2$, *bottom*). These classification is in agreement with dendrograms, binary trees, splits graph and partial correlation diagram (Figures 1–6).

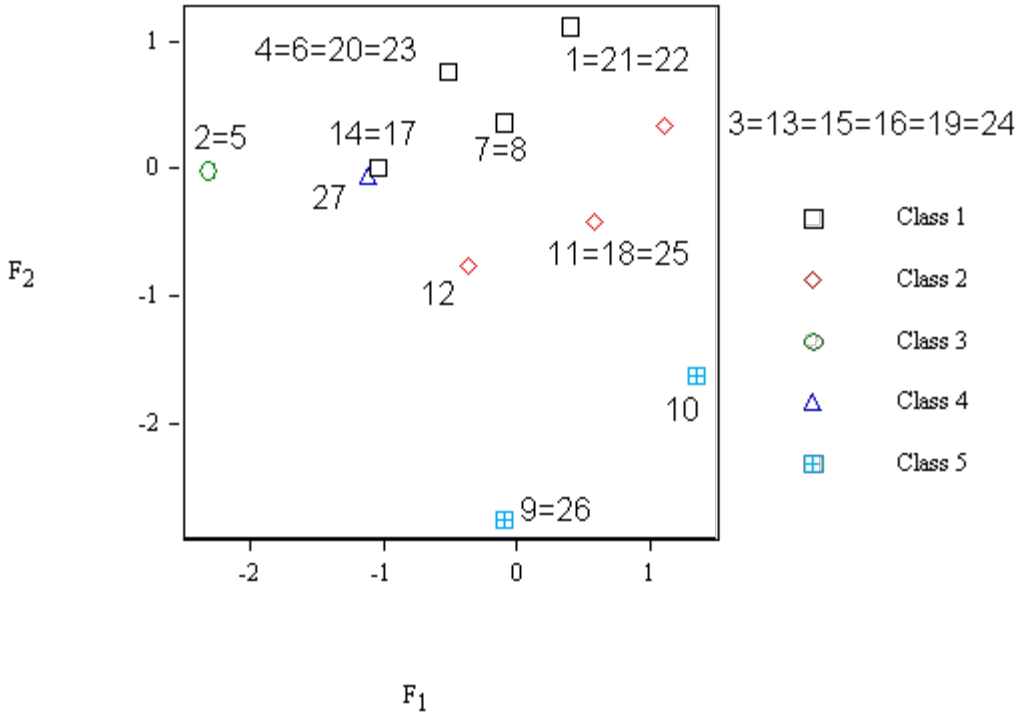

Figure 7. Principal component analysis F_2 *vs*. F_1 plot for the local anaesthetics.

The recommended format for the periodic table (PT) of the local anaesthetics analogues of procaine is listed in Table 5. Local anaesthetics are classified first by i_5, then by i_4, i_3, i_2 and, finally, by i_1. Periods of five units are assumed. Group g010 stands for $<i_1,i_2,i_3>$ = <010>, *viz*. <01001> (dibucaine, propanolol), and <01010> (dimethisoquin), group g100, for $<i_1,i_2,i_3>$ = <100>, *viz*. <10011> (phenytoin), *etc*. The local anaesthetics in the same column of Table 5 appear close in dendrograms, radial trees, splits graph, partial correlation diagram and PCA (Figures

1–7).

Table 5. Table of periodic properties for local anaesthetics analogues of procaine.

g010	**g100**	**g101**	**g110**	**g111**
			diperodon, pramoxine, mexiletine	cocaine, cyclomethycaine
dibucaine, propanolol		benzocaine, butamben	dyclonine	hexylcaine, piperocaine
dimethisoquin			bupivicaine, etidocaine, lidocaine, mepivacaine, prilocaine, tocainide	benoxinate, proparacaine, propoxycaine
	phenytoin			butacaine, 2 chloroprocaine, procaine, tetracaine

Figure 8 exhibits the variation of the vector property as a function of the structural parameters $\{i_1,i_2,i_3,i_4,i_5\}$ for local anaesthetics. The lines for the structural parameters i_4 and i_5 appear superimposed, what agrees with a PT of properties with vertical groups defined by $\{i_1,i_2,i_3\}$ and horizontal periods described by $\{i_4,i_5\}$.

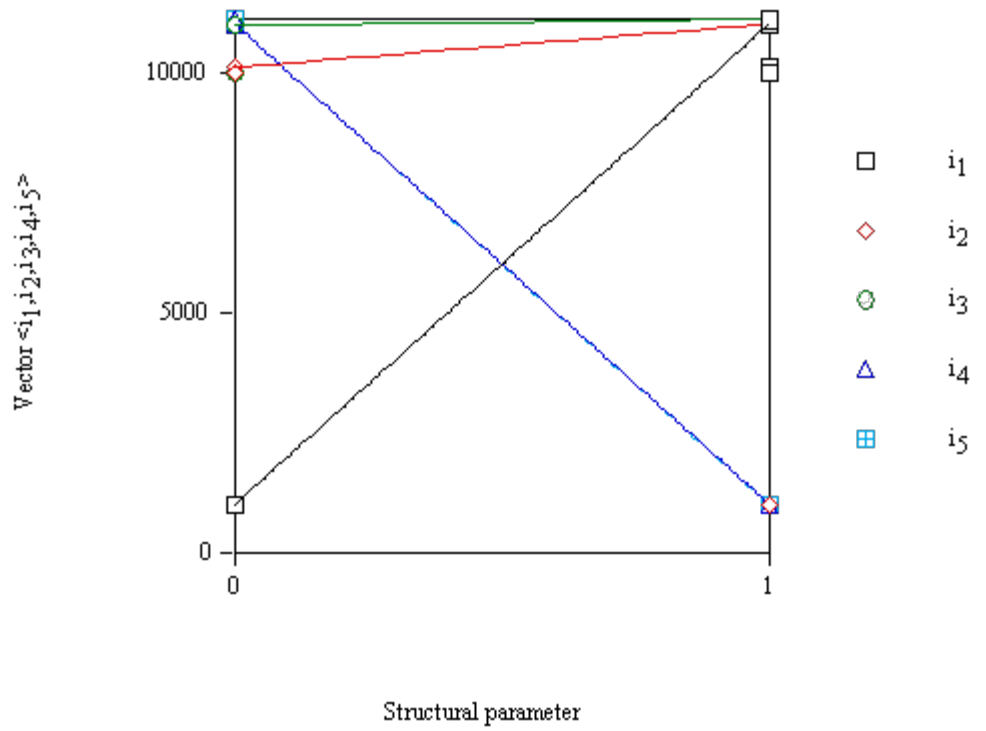

Figure 8. Variation of the vector property of local anaesthetics *vs*. counts $\{i_1,i_2,i_3,i_4,i_5\}$.

The variation of the vector property $P = <i_1,i_2,i_3,i_4,i_5>$ as a function of the number of the group in PT (*cf*. Figure 9) for local anaesthetics reveals that the minima correspond to anaesthetics with $\{i_1,i_2,i_3\}$ values of <010> (group g010). The corresponding function $P(i_1,i_2,i_3,i_4,i_5)$ reveals a series of *waves*, clearly limited by maxima or minima, which suggest a periodic behaviour that recalls the form of a trigonometric function. For $<i_1,i_2,i_3,i_4,i_5>$ two minima are clearly shown. The distance in $\{i_1,i_2,i_3,i_4,i_5\}$ units between each pair of consecutive minima is five, which coincides with the local anaesthetic sets belonging to the same group in PT and in the successive periods. The minima occupy analogous positions in the curve and are in phase. The representative points in phase should correspond to the elements of the same group in PT. For the $<i_1,i_2,i_3,i_4,i_5>$ minima there is coherence between both representations; however, the consistency is not general. The comparison of the *waves* shows two differences: (1) periods 1–2 show some sawtooth-like structures with marked discontinuities in $P(i_1,i_2,i_3,i_4,i_5)$; (2) periods 3–4 are also sawtooth-like although much less marked. The most characteristic points of the plot are the minima, and correspond to the anaesthetics of group g010. Their $<i_1,i_2,i_3,i_4,i_5>$ values are not repeated as the periodic law (PL) states but decrease regularly.

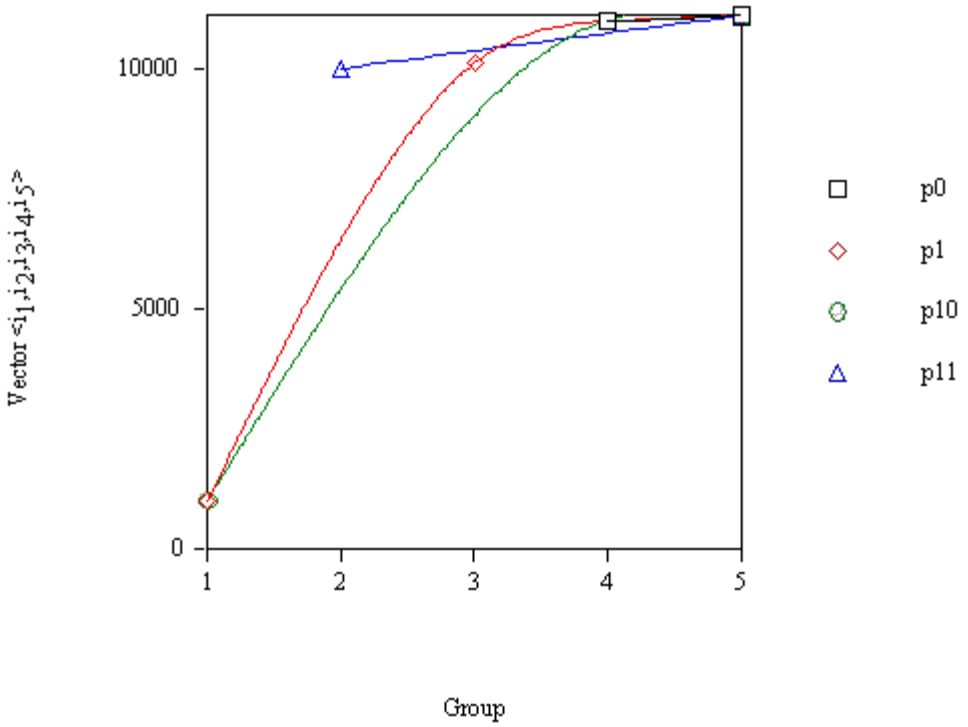

Figure 9. Variation of the vector property of local anaesthetics *vs*. group number.

An empirical function $P(i_1,i_2,i_3,i_4,i_5)$ reproduces, with enough precision, the different $<i_1,i_2,i_3,i_4,i_5>$ values for the anaesthetics. A minimum value of a function $P(p)$ has meaning only if it is compared with those for the former $P(p-1)$ and later $P(p+1)$ points, needing to fulfil:

$$P_{\min}(p) < P(p-1)$$
$$P_{\min}(p) < P(p+1) \qquad (8)$$

Order relations (8) should repeat at determined intervals equal to the values of the period size and are equivalent to:

$$P_{\min}(p) - P(p-1) < 0$$
$$P(p+1) - P_{\min}(p) > 0 \qquad (9)$$

As relations (9) are valid only for minima more general others are desired for all the values of p. Therefore, the differences $P(p+1) - P(p)$ are calculated assigning each of their values to anaesthetic p. Naming this value $D(p)$:

$$D(p) = P(p+1) - P(p) \qquad (10)$$

Instead of $D(p)$ the $R(p) = P(p+1)/P(p)$ values can be taken assigning them to anaesthetic p. If PL were general, the elements belonging to the same group occupying analogous positions in the different *waves* would satisfy:

$$D(p) > 0 \text{ or } D(p) < 0 \qquad (11)$$
$$R(p) > 1 \text{ or } R(p) < 1 \qquad (12)$$

However, the results show that this is not the case so that PL is not general, existing some anomalies; *e.g*., the variation

of $D(p)$ *vs*. group number in Figure 10 presents lack of coherence between the $<i_1,i_2,i_3,i_4,i_5>$ Cartesian and PT representations. If consistency were rigorous all the points in each period would have the same sign. In general there is a tendency in the points to give $D(p) < 0$ especially for the greater groups. In detail, however, there are irregularities in which the anaesthetics for successive periods are not always in phase.

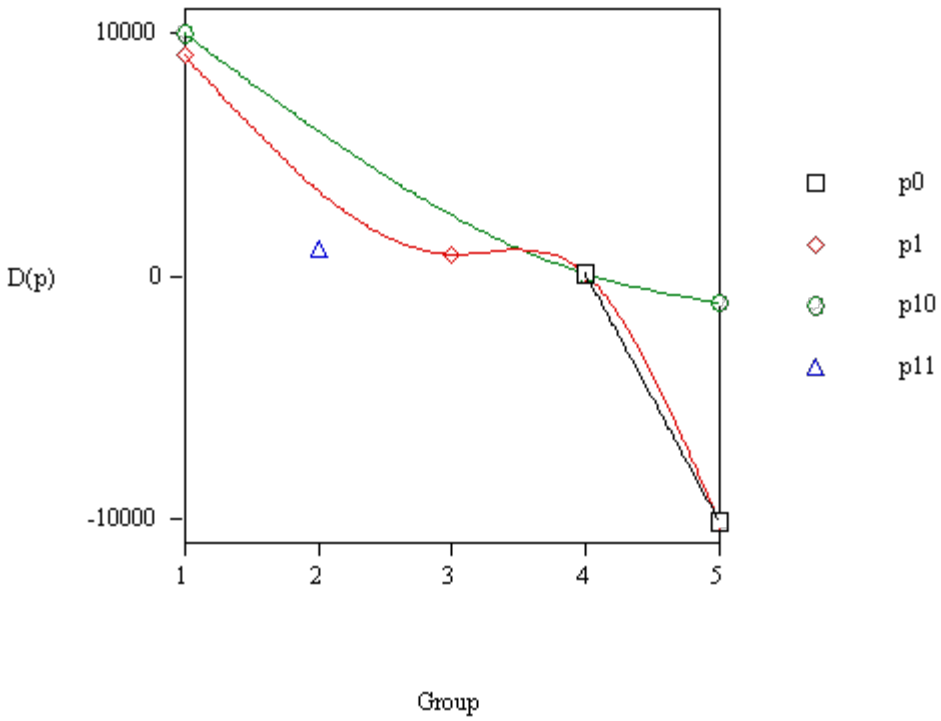

Figure 10. Variation of $D(p) = P(p+1) – P(p)$ *vs*. group number. P is the vector property.

The change of $R(p)$ *vs*. group number in Figure 11 shows lack of constancy between the Cartesian and PT charts. If steadiness were exact, all the points in each period would be either lower or greater than one. There is a propensity in the points to give $R(p) > 1$ particularly for the smaller groups. Notwithstanding, there are incongruities in which the anaesthetics for consecutive waves are not always in phase.

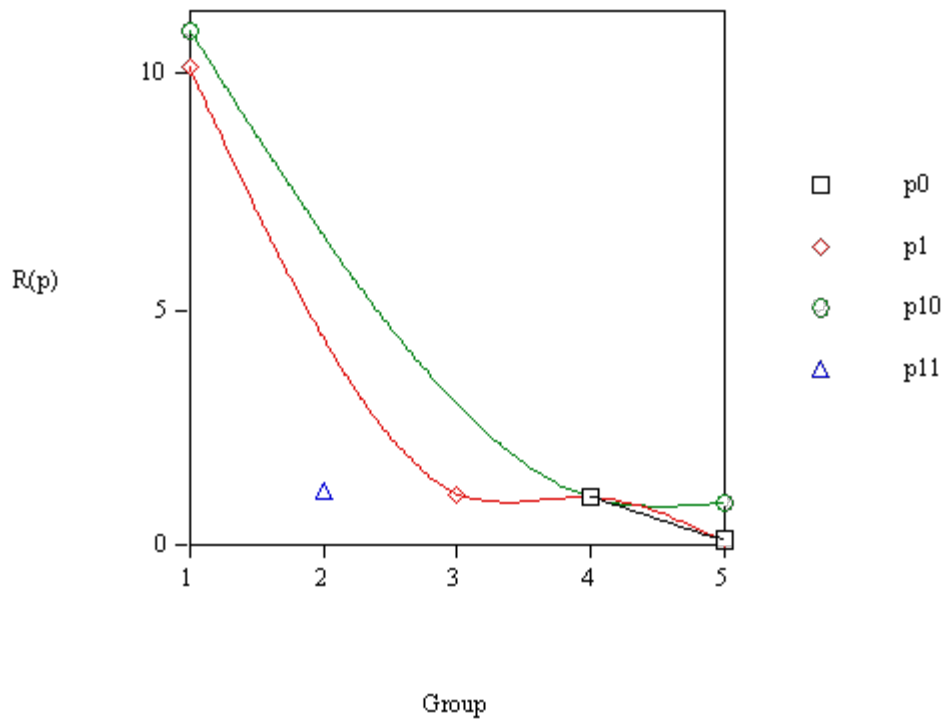

Figure 11. Variation of $R(p) = P(p+1)/P(p)$ *vs*. group number. P is the vector property.

Conclusion

From the present results and discussion the folllowing conclusions can be drawn.

1. Many algorithms for classification and taxonomy are based on *information entropy*. When applying these procedures to sets of moderate size, an excessive number of results appear compatible with data, and this number suffers a combinatorial explosion. However, after the *equipartition conjecture* one has a selection criterion between different variants resulting from classification between hierarchical trees. According to this conjecture, for a given charge or duty, the best configuration of a flowsheet is the one in which the entropy production is most uniformly distributed.

2. Several criteria have been selected to reduce the analysis to a manageable quantity of structures from the large set of local anaesthetics. They refer to the structural parameters related with the lipophilic portion, hydrophilic portion, *etc*.

3. Information entropy analysis of the structural parameters and principal component analysis of the local anaesthetics permit classifying them and agree.

4. The area of clustering is notoriously difficult; *e.g*., although oranges and apples seem to have significant differences, they are both fruit. Is a pomegranate more like an apple or is it more like an orange? When the clustering

problem is poorly specified, or the variation within each cluster is greater than that between different clusters, meaningful clustering often becomes almost impossible. Progression in the development of new methods is hampered by the lack of *gold standards*, against which to judge the quality of any clustering exercise. An understanding of both the chemistry and the computational methods is essential for tackling the associated *data mining* tasks, without being distracted by the abundant fool's gold. If a small number of clusters of data are easy to fit, the predictive ability of the model could be guaranteed only if the deviations inside the clusters do not diverge [20].

5. The periodic law has not the rank of the laws of physics: (1) the properties of the local anaesthetics are not repeated; perhaps, their chemical character; (2) the order relationships are repeated, with exceptions. The analysis forces the statement: The relationships that any anaesthetic p has with its neighbour $p + 1$ are approximately repeated for each period. Periodicity is not general; however, if a natural order of the anaesthetics is accepted the law must be phenomenological.

Acknowledgements

The authors acknowledge financial support from the Spanish MEC DGI (Project No. CTQ2004 07768 C02 01/BQU) and Generalitat Valenciana (DGEUI INF01 051 and INFRA03 047, and OCYT GRUPOS03 173).

References

1. Cotton, S. Soundbite Molecules. *Educ. Chem*. **1999**, *36*, 62-62.

2. Caliendo, G.; di Carlo, R.; Greco, G.; Grieco, P.; Meli, R.; Novellino, E.; Perissutti, E.; Santagada, V. Synthesis, Local Anesthetic Activity and QSAR Studies for a Set of *N* [2 (alkylamino)ethyl]benzotriazol *x* yl Acetamides. *Eur. J. Med. Chem*. **1995**, *30*, 603-608.

3. Caliendo, G.; Greco, G.; Grieco, P.; Raso, G. M.; Meli, R.; Novellino, E.; Perissutti, E.; Santagada, V. Synthesis and Pharmacological Evaluation of a Set of *N* [2 (alkylamino)ethyl]benzotriazol *X* yl Isobutyramides Acting as Local Anesthetics. *Eur. J. Med. Chem*. **1996**, *31*, 99-104.

4. Do Amaral, A. T.; Oliverira, A. C.; Neidlein, R.; Gallacci, M.; Caprara, L.; Miyazaki, Y. Physicochemical Parameters Involved in the Lethal Toxicity of *N,N* [(dimethylamino)ethyl]-4 substituted Benzoate Hydrochlorides: A QSAR Study. *Eur. J. Med. Chem*. **1997**, *32*, 433-443.

5. Tavares, L. C.; do Amaral, A. T. QSPR/QSAR in *N* [(dimethylamine)methyl] Benzamides Substituents Groups Influence upon Electronic Distribution and Local Anaesthetics Activity. *Bioorg. Med. Chem*. **2004**, *12*, 1377-1381.

6. Varmuza, K. *Pattern Recognition in Chemistry* ; Springer: New York, 1980.

7. Benzecri, J. P. *L'Analyse des Données I. La Taxonomie* ; Dunod: Paris, 1984.

8. Adams, H. J.; Ronfeld, R. A.; Takman, B. H. In *Handbook of CNS Agents of Local Anaesthetics* ; Verderame, M., Ed., CRC: Boca Raton, 1986.

9. Tondeur, D.; Kvaalen, E. Equipartition of Entropy Production. An Optimality Criterion for Transfer and Separation Processes. *Ind. Eng. Chem., Fundam*. **1987**, *26*, 50-56.

10. Kimura, M. *New Scientist* **1986**, *11*, 41-41.

11. Corriou, J. P.; Iordache, O.; Tondeur, D. Classification of Biomolecules by Information Entropy. *J. Chim. Phys. Phys.-Chim. Biol*. **1991**, *88*, 2645-2652.

12. Kaufmann, A. *Introduction à la Théorie des Sous-ensembles Flous* ; Masson: Paris, 1975; Vol. 3.

13. White, H. *AI Expert* **1989**, *12*, 48-48.

14. Iordache, O.; Corriou, J. P.; Garrido-Sánchez, L.; Fonteix, C.; Tondeur, D. Neural Network Frames. Application to Biochemical Kinetic Diagnosis. *Comput. Chem. Eng*. **1993**, *17*, 1101-1113.

15. Tryon, R. C. *J. Chronic Dis*., **1939**, *20*, 511-524.

16. Jarvis, R. A.; Patrick, E. A. Clustering Using a Similarity Measure Based on Shared Nearest Neighbors. *IEEE Trans. Comput*., **1973**, *C22*, 1025-1034.

17. Page, R. D. M. Program TreeView. Universiy of Glasgow, 2000.

18. Huson, D. H. SplitsTree: Analizing and Visualizing Evolutionary Data. *Bioinformatics* **1998**, *14*, 68-73.

19. Hotelling, H. Analysis of a Complex of Statistical Variables into Principal Components. *J. Educ. Psychol*., **1933**, *24*, 417-441.

20. He, L.; Jurs, P. C. Assessing the Reliability of a QSAR Model's Predictions, *J. Mol. Graphics Model*. **2005**, *23*, 503-523.