

Periodic Classification of Local Anaesthetics (Procaine Analogues)

Francisco Torrens ^{1,*} and Gloria Castellano ²

1 Institut Universitari de Ciència Molecular, Universitat de València, Edifici d'Instituts de Paterna, P. O. Box 22085, E-46071 València, Spain. <http://www.uv.es/~uiqt/index.htm>. Tel. +34 963 544 431, Fax +34 963 543 274

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

* Author to whom correspondence should be addressed. E-mail: francisco.torrens@uv.es

Received: 16 December 2005 / Accepted: 26 January 2006 / Published: 31 January 2006

Abstract: Algorithms for classification are proposed based on criteria (*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 and principal component analyses agree. The periodic law of anaesthetics 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, local anaesthetic, procaine analogue.

1. Introduction

For hundreds of years, surgeons had to work fast to minimize shock and pain to their patients [1]. Local anaesthetics are amphiphile molecules of tertiary amines, and some of them have colloidal properties in aqueous solution. They are classified into the ester and amide types, 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, 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 was proposed that substituent groups present in the aromatic ring affect the local anaesthetic activity, by its effects hydrophobic and of polar nature. Once it was known that the inductive and resonance effects affect directly the electronic density on O(=C) atom, as consequence it was proposed that C=O 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. Many local anaesthetics in clinical use are basically tertiary amine compounds. They are classified as being of the ester (benzocaine, 2-chloroprocaine, procaine, tetracaine) and amide types (bupivacaine, dibucaine, etidocaine, lidocaine, mepivacaine, prilocaine, S-ropivacaine). These drugs exist in both positively charged and uncharged forms under normal *in vivo* conditions [6–8]. The most important clinical properties of local anaesthetics are potency, onset, duration of action as well as relative blockade of sensory and motor fibres. These qualities are related primarily to the physicochemical properties of the various compounds. In general, lipid solubility determines the relative intrinsic potency of the various agents, while protein binding influences the duration of anaesthesia and pK_a is correlated with the onset of action. The local anaesthetics for infiltration, peripheral nerve blockade and extradural anaesthesia can be classified into three groups: agents of low potency and short duration (procaine, 2-chloroprocaine), agents of moderate potency and duration (lidocaine, mepivacaine, prilocaine), as well as agents of high potency and long duration (tetracaine, bupivacaine, etidocaine, S-ropivacaine). These local anaesthetics also vary in terms of onset: 2-chloroprocaine, etidocaine, lidocaine, mepivacaine and prilocaine have a rapid onset, while procaine, tetracaine and bupivacaine are characterized by a longer latency period; bupivacaine presents a moderate onset.

General anaesthetics, including inhaled agents, *e.g.* halothane, isoflurane and desflurane, as well as intravenous agents, *e.g.* propofol, are now generally considered to act by modulating the effects of γ -aminobutyric acid (GABA) and type A ($GABA_A$) receptors and/or by affecting other ion channels [9,10]. General anaesthetics that do not act at $GABA_A$ receptors are NO_2 , Xe and ketamine. The theory that *inhaled anaesthetics probably dissolve in the lipid-rich neuronal cell membranes near the nerve endings and changes their volume or fluidity* is not well supported by experimental evidence. For

example, such theories fail to account for differences in potency for stereoisomers of general anaesthetics [11]. Local anaesthetics act by blocking voltage-gated Na^+ channels although some agents, *e.g.* propranolol, affect other proteins as well.

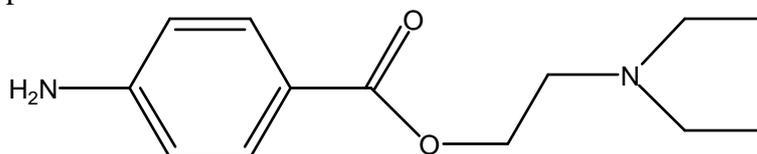
A simple computerized algorithm useful for establishing a relationship between chemical structures and their biological activities or significance is proposed and exemplified here (*cf.* Reference 12 or 13 for a review). The starting point is to use an informational or configurational entropy for pattern 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 [14]. As entropy is weakly discriminating for classification purposes, the more powerful concept of *entropy production* and its *equipartition conjecture* are introduced [15]. Learning potentialities of the code have also been developed. Since molecules are more naturally described *via* a varying size structured representation, the study of general approaches to the processing of structured information is needed. Section 2 presents the computational method. Section 3 describes the classification algorithm. Section 4 exposes the equipartition conjecture of entropy production. Section 5 analyzes the learning procedure. Section 6 presents and discusses the calculation results. Section 7 summarizes the conclusions.

2. 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* $\vec{i} = \langle i_1, i_2, \dots, i_k, \dots \rangle$ should be associated to each 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 [16]. 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 [17]. 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. The lipophilic portion is primarily responsible for the lipid solubility that allows diffusion across the nerve cell membrane, determining the intrinsic potency of local anaesthetics [6,7]. Both the lipophilic and hydrophilic portions determine protein-binding characteristics, which are felt to be the primary determinant of anaesthesia duration [8].

Procaine or novocaine {4-aminobenzoic acid [2-(diethylamino)ethyl] ester, $4\text{-H}_2\text{NC}_6\text{H}_4\text{CO}_2[\text{CH}_2\text{CH}_2\text{N}(\text{C}_2\text{H}_5)_2]$, PR} is a tertiary amine, containing a primary amino group linked

to an aromatic ring (*cf.* Scheme 1). Consequently, it may exist as a neutral molecule (PR), a monocation ($4\text{-H}_2\text{NC}_6\text{H}_4\text{CO}_2[\text{CH}_2\text{CH}_2\text{NH}^+(\text{C}_2\text{H}_5)_2]$, PRH^+), or a dication ($4\text{-H}_2\text{NH}^+\text{C}_6\text{H}_4\text{CO}_2[\text{CH}_2\text{CH}_2\text{NH}^+(\text{C}_2\text{H}_5)_2]$, PRH_2^{2+}). In 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, its associated vector is $\langle 11111 \rangle$. In this work, procaine was selected as a *reference* anaesthetic because it was the first synthetic compound successfully used for regional anaesthesia and, in this and most other local anaesthetics, the lipophilic portion consists of at least one phenyl radical, the hydrophilic portion is a secondary or tertiary amine, and/or the intermediate chain has an ester linkage. This improves the quality of the classification for those anaesthetics similar to procaine. The selection as reference of an anaesthetic dissimilar to procaine, *e.g.* dibucaine, would not improve the quality of the classification for those anaesthetics similar to procaine. Furthermore, Covino results included both ester (similar to procaine) and amide (similar to prilocaine) anaesthetics [8]; the inclusion of the results in the classification described below improves the quality of the taxonomy, for both types of anaesthetics.



Scheme 1. Molecular structure of a local anaesthetic procaine neutral molecule.

Table 1 contains the vectors associated to 28 local anaesthetics. The $\langle 11110 \rangle$ vector is associated to benoxinate (Table 1), since there are three O atoms in this case. The $\langle 10101 \rangle$ vector is associated to benzocaine since the hydrophilic partition is not an amine, and there is one N atom in this case.

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

1. benoxinate	$\langle 11110 \rangle$	15. lidocaine	$\langle 11010 \rangle$
2. benzocaine	$\langle 10101 \rangle$	16. mepivacaine	$\langle 11010 \rangle$
3. bupivacaine	$\langle 11010 \rangle$	17. piperocaine	$\langle 11101 \rangle$
4. butacaine	$\langle 11111 \rangle$	18. pramoxine	$\langle 11000 \rangle$
5. butamben	$\langle 10101 \rangle$	19. prilocaine	$\langle 11010 \rangle$
6. 2-chloroprocaine	$\langle 11111 \rangle$	20. procaine	$\langle 11111 \rangle$
7. cocaine	$\langle 11100 \rangle$	21. proparacaine	$\langle 11110 \rangle$
8. cyclomethycaine	$\langle 11100 \rangle$	22. propoxycaine	$\langle 11110 \rangle$
9. dibucaine	$\langle 01001 \rangle$	23. tetracaine	$\langle 11111 \rangle$
10. dimethisoquin	$\langle 01010 \rangle$	24. tocainide	$\langle 11010 \rangle$
11. dipherodon	$\langle 11000 \rangle$	25. mexiletine	$\langle 11000 \rangle$
12. dyclonine	$\langle 11001 \rangle$	26. propanolol	$\langle 01001 \rangle$
13. etidocaine	$\langle 11010 \rangle$	27. phenytoin	$\langle 10011 \rangle$
14. hexylcaine	$\langle 11101 \rangle$	28. S-ropivacaine	$\langle 11010 \rangle$

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

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

where $0 \leq a_k \leq 1$ and $t_k = 1$ if $i_k = j_k$, but $t_k = 0$ 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 .

3. Classification Algorithm

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

$$(\bar{R} \circ \bar{S})_{ij} = \max_k [\min_k (r_{ik}, s_{kj})] \quad (2)$$

where $\bar{R} = [r_{ij}]$ and $\bar{S} = [s_{ij}]$ are matrices of the same type, and $(\bar{R} \circ \bar{S})_{ij}$ is the (i,j) -th element of the matrix $\bar{R} \circ \bar{S}$ [18]. It can be shown that when applying this rule iteratively so that $\bar{R}(n+1) = \bar{R}(n) \circ \bar{R}$, there exists an integer n such that: $\bar{R}(n) = \bar{R}(n+1) = \dots$. The resulting matrix $\bar{R}(n)$ 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 $\bar{R}(n) = [r_{ij}(n)]$. 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:

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

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.

In information theory, the *information entropy* h measures the surprise that the source emitting the sequences can give [19,20]. For a single event occurring with probability p , the degree of surprise is proportional to $-\ln p$. Generalizing the result to a random variable X (which can take N possible values x_1, \dots, x_N with probabilities p_1, \dots, p_N), the average surprise received on learning the value of X is $-\sum p_i \ln p_i$. The information entropy associated with the matrix of similarity \bar{R} is:

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

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

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

4. The Equipartition Conjecture of Entropy Production

In the classification algorithm, each *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 [15] *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:

$$h_{\text{eqp}} = h_{\text{max}} b \quad (5)$$

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:

$$SS = \sum_{b_i} (h - h_{\text{eqp}})^2 \quad (6)$$

5. Learning Procedure

Learning procedures similar to those encountered in *stochastic methods* are implemented as follows [21]. Consider a given partition in classes as *good* or ideal from practical or empirical observations. This corresponds to a *reference* similarity matrix $\overline{\overline{S}} = [s_{ij}]$ obtained for equal weights $a_1 = a_2 = \dots = a$ and for an arbitrary number of fictitious 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 $\overline{\overline{R}}$. The number of properties for $\overline{\overline{R}}$ and $\overline{\overline{S}}$ may differ. The learning procedure consists in trying to find classification results for $\overline{\overline{R}}$ 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, \dots 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 $\overline{\overline{R}}$ and $\overline{\overline{S}}$ is given by:

$$D = - \sum_{ij} (1 - r_{ij}) \ln \frac{1 - r_{ij}}{1 - s_{ij}} - \sum_{ij} r_{ij} \ln \frac{r_{ij}}{s_{ij}} \quad \forall 0 \leq r_{ij}, s_{ij} \leq 1 \quad (7)$$

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 [22].

6. Calculation Results and Discussion

In the present report 28 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. The matrix of Pearson correlation coefficients between each pair of vector properties $\langle i_1, i_2, i_3, i_4, i_5 \rangle$ of the 28 anaesthetics has been calculated. The intercorrelations are illustrated in the partial correlation diagram, which contains 133 high partial correlations ($r \geq 0.75$, *cf.* Figure 1, *red lines*), 76 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 correlations have a similar vector property (Table 1). However, the results (Figure 1) should be taken with care, because four compounds show the constant $\langle 11111 \rangle$ vector (Entries 4, 6, 20 and 23 in Table 1), for which the null standard deviation causes high partial correlations ($r = 1$) with any local anaesthetic, which is an artifact.

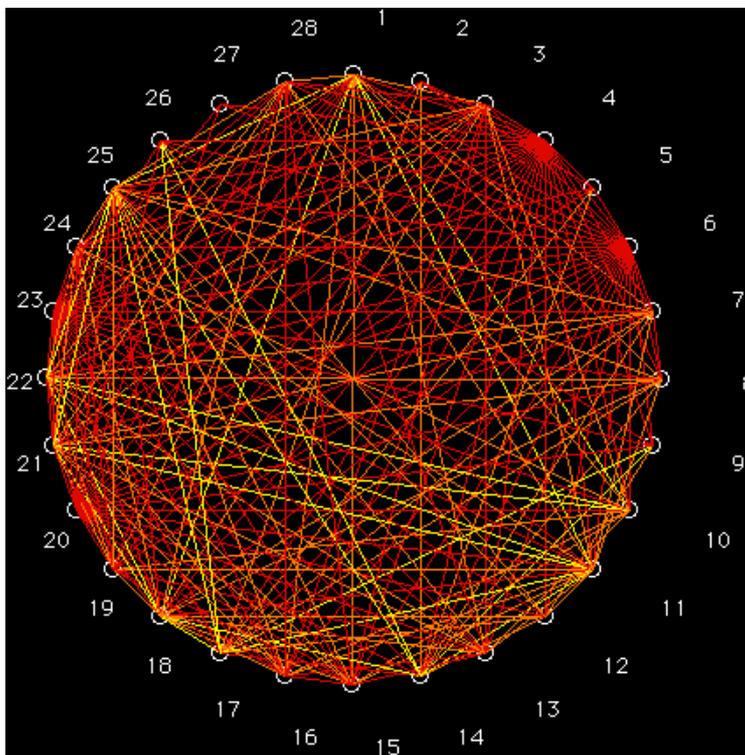


Figure 1. Partial correlation diagram: High (red), medium (orange) and low (yellow) correlations.

Using the grouping rule in the drug-design case with equal weights $a_k = 0.5$, for $0.94 \leq b_1 \leq 0.96$ the following set of classes are obtained [17]:

$$C_{b_1} = (1,21,22)(2,5)(3,13,15,16,19,24,28)(4,6,20,23)(7,8)(9,26)(10)(11,18,25)(12)(14,17)(27)$$

The 11 classes are obtained with the associated entropy $h\left(\overline{R}_{b_1}\right) = 58.86$. The dendrogram (binary tree) [23,24] matching to $\langle i_1, i_2, i_3, i_4, i_5 \rangle$ and C_{b_1} is illustrated [25] in Figure 2. It provides a binary taxonomy of Table 1, which separates the same 11 classes. In particular, the ester (benzocaine, 2-chloroprocaine, procaine, tetracaine) and amide local anaesthetics (bupivacaine, dibucaine, etidocaine, lidocaine, mepivacaine, prilocaine, S-ropivacaine) are always grouped in different classes. The agents of low potency and short duration (procaine, 2-chloroprocaine) are separated from the agents of high potency and long duration (bupivacaine, etidocaine, S-ropivacaine), while the agents of moderate potency and duration (mepivacaine, prilocaine) are classified together with the latter. Those anaesthetics belonging to the same class appear highly correlated in the partial correlation diagram (Figure 1), in agreement with previous results obtained for the first 27 entries in Table 1 [17].

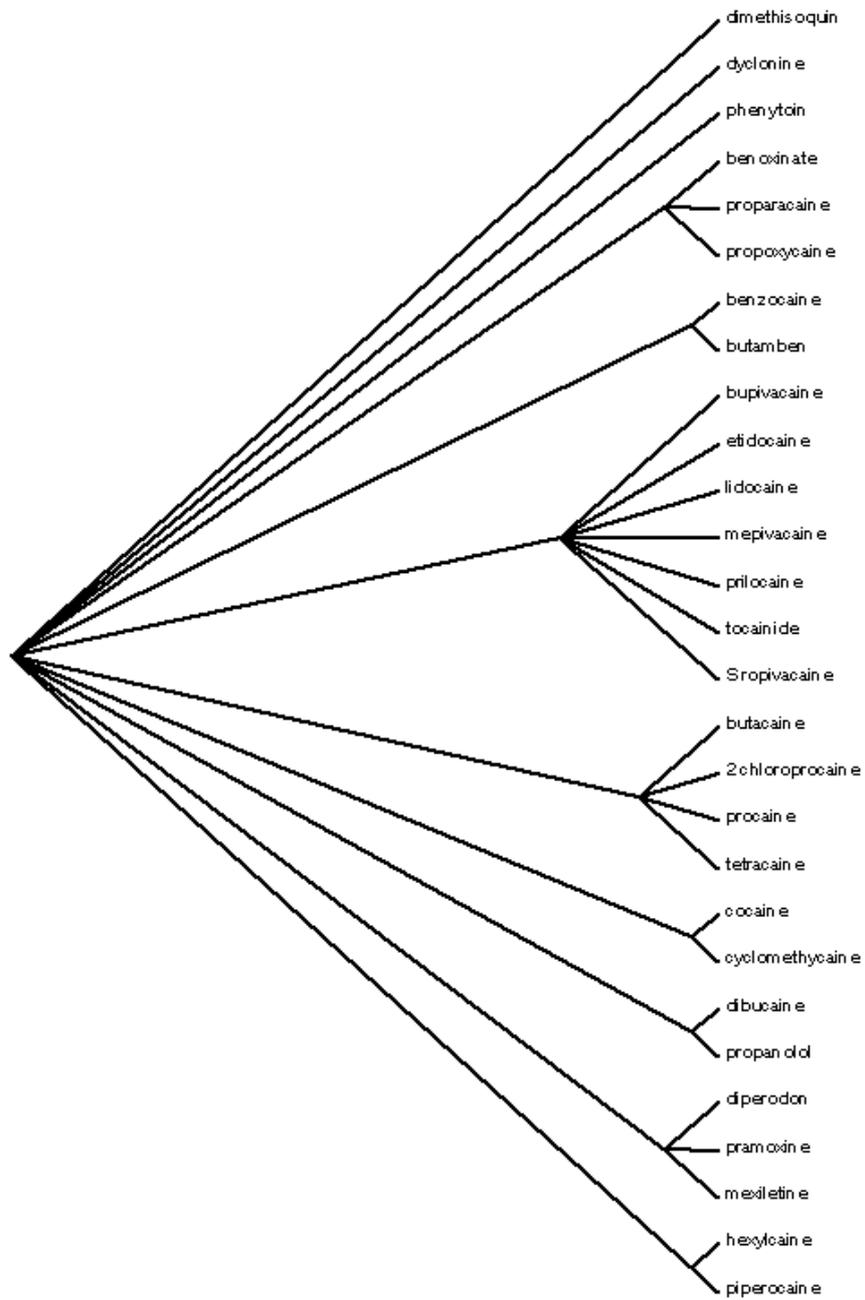


Figure 2. Dendrogram for the local anaesthetics analogues of procaine at level b_1 .

The radial tree for the local anaesthetics relating to $\langle i_1, i_2, i_3, i_4, i_5 \rangle$ and C_{b_1} (cf. Figure 3) separates the same 11 classes, in agreement with the partial correlation diagram, dendrogram (Figures 1–2) and previous results obtained for the first 27 entries in Table 1 [17].

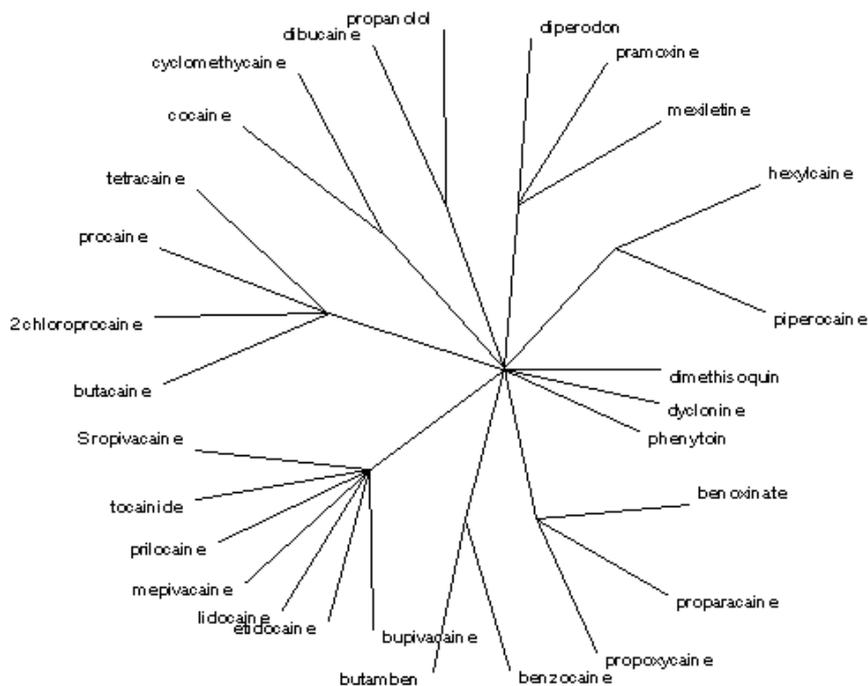


Figure 3. Radial tree for the local anaesthetics analogues of procaine at level b_1 .

At level b_2 with $0.85 \leq b_2 \leq 0.87$ the set of classes is [17]

$$C_{b_2} = (1,4,6,7,8,14,17,20,21,22,23)(2,5)(3,11,12,13,15,16,18,19,24,25,28)(9,10,26)(27)$$

Five classes result in this case; the entropy is $h(\overline{R}_{b_2}) = 12.20$. The radial tree matching to $\langle i_1, i_2, i_3, i_4, i_5 \rangle$ and C_{b_2} (cf. Figure 4) separates the same five classes, in agreement with the partial correlation diagram, dendrogram, binary tree (Figures 1–3) and previous results obtained for the first 27 entries in Table 1 [17]. A high degree of similarity is found for Entries 9 and 26 (*i.e.* dibucaine and propanolol), as well as Entries 2 and 5 (*i.e.* benzocaine and butamben). Again, the ester and amide local anaesthetics are grouped in different classes; the agents of low potency and short duration are separated from the agents of high–medium potency and long–medium duration. The lower level b_2 classification process shows lower entropy and, therefore, may be more parsimonious. The classification model divides the point process into two components, *viz.* signal, and noise; the lower-level b_2 may have greater signal-to-noise ratio than the higher-level b_1 classification process. Naturally, Entries 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 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.

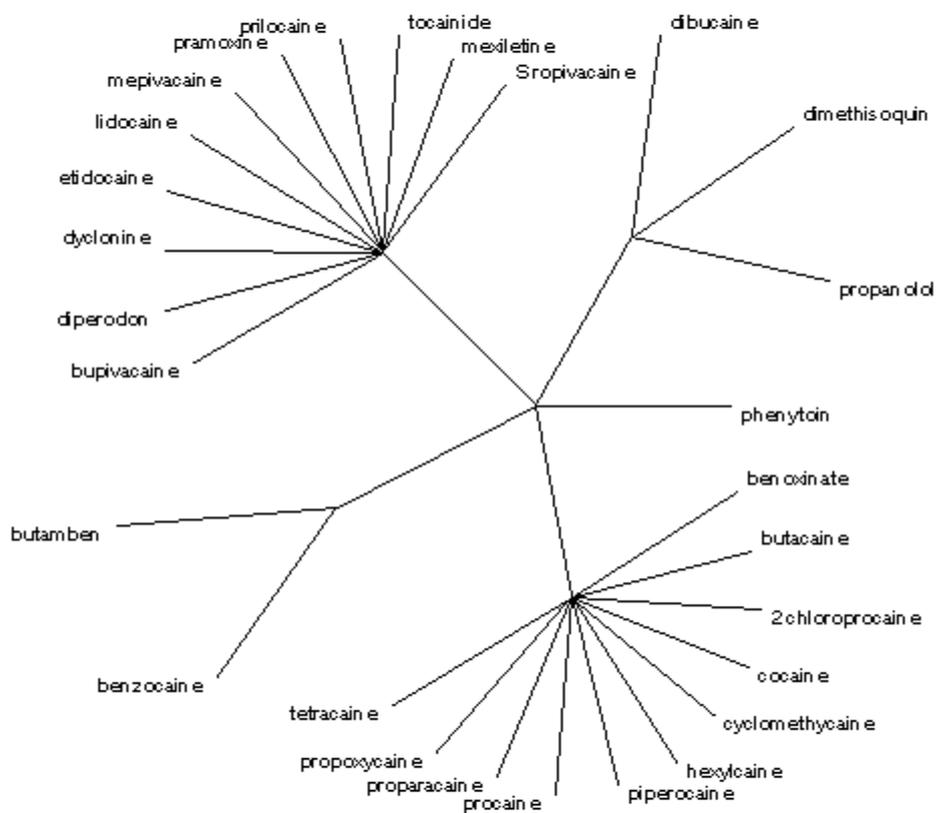


Figure 4. Radial tree for the local anaesthetics analogues of procaine at level b_2 .

A comparative analysis of the set containing from one to 11 classes is summarized in Table 2, in agreement with previous results obtained for the first 27 entries in Table 1 [17].

Table 2. Classification level, number of classes and entropy for the local anaesthetics.

Classification level b	Number of classes	Entropy h
0.96	11	59.65
0.93	8	31.31
0.87	5	12.00
0.78	4	7.23
0.75	3	3.95
0.56	2	1.66
0.25	1	0.14

From the set containing from one to 11 classes (Table 2), the radial tree matching to $\langle i_1, i_2, i_3, i_4, i_5 \rangle$ and C_{b_1-11} (cf. Figure 5) separates the same five and 11 classes, in agreement with the partial correlation diagram, dendrogram, binary trees (Figures 1–4) and previous results obtained for the first 27 entries in Table 1 [17]. Again, the ester and amide local anaesthetics are grouped in different classes; the agents of low potency and short duration are separated from the agents of high–medium potency and long–medium duration.

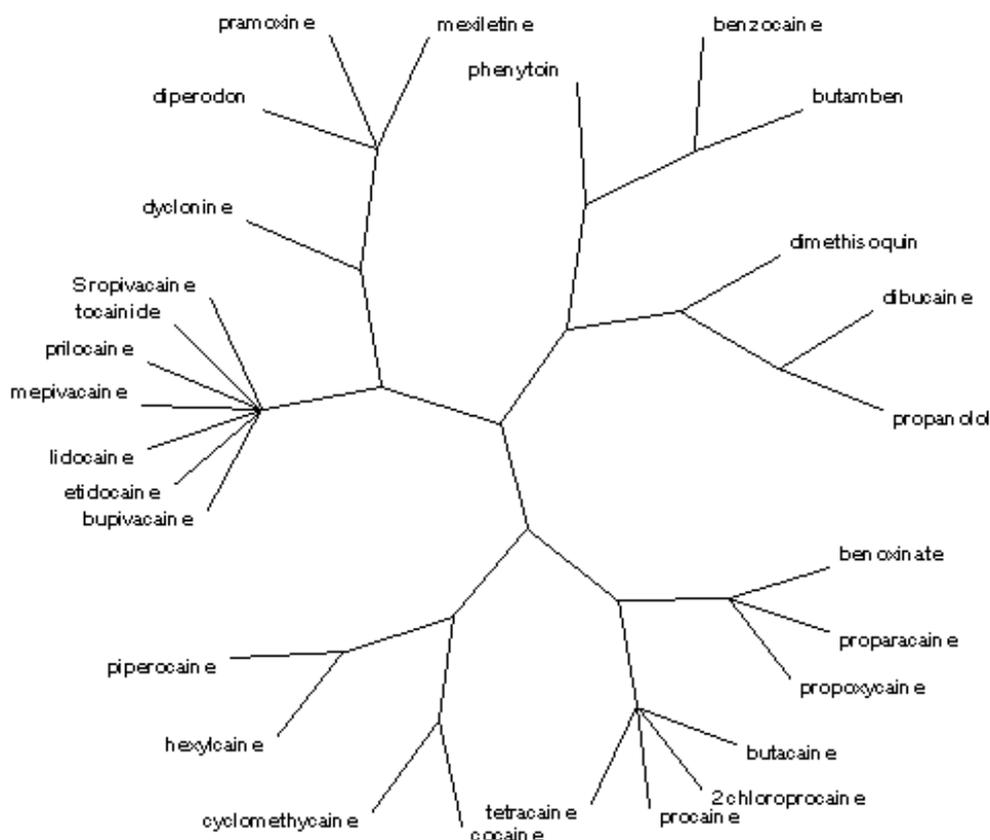


Figure 5. Radial tree for the local anaesthetics analogues of procaine from 1 to 11 classes.

The resulting partition into classes compares well with other from Covino considered as *good* [8]. He compared three ester, *viz.* 2-chloroprocaine, procaine, and tetracaine, as well as five amide local anaesthetics, *viz.* bupivacaine, etidocaine, lidocaine, mepivacaine, and prilocaine, based on chemical configuration (aromatic lipophilic group, intermediate chain and amine hydrophilic group), four physicochemical properties (molecular weight, pK_a , partition coefficient and protein binding), as well as three pharmacological properties (onset, relative potency and duration). The onset is determined primarily by pK_a . The percentage of a local anaesthetic that is present in the neutral form, when injected to tissue of pH 7.4, decreases with pK_a , according to the equation of Henderson–Hasselbalch: $pH = pK_a + \log ([PR]/[PRH^+])$. The potency is determined primarily by lipid solubility, which increases with partition coefficient. Both lipid solubility and partition coefficient are mainly due to the neutral forms. Different conformations have different partition coefficients, lipid solubilities and potencies. It would be of interest to study the effect of different intermediate chain lengths. In particular, the presence of a double bond in a chain would increase rigidity and enhance potency; *e.g.*, the conjugated enol group in 3-phenyl-2-propen-1-ol determines a greater membrane permeability, with respect to 3-phenyl-1-propanol [26]. On the one hand, esters are hydrolyzed easily and are relatively unstable in solution; on the other, amides are much more stable. In the body, the amino esters are hydrolyzed in plasma by the enzyme cholinesterase, whereas the amino amides undergo enzymatic degradation in the liver.

The inclusion of this comparison [8] in the radial tree of the present work (*cf.* Figure 6) is in agreement with the partial correlation diagram, dendrogram, binary trees (Figures 1–5) and previous results obtained for the first 27 entries in Table 1 [17]. The classification scheme from 1 to 11 levels is conserved after the addition of Entry 28 (S-ropivacaine) and local anaesthetic S-bupivacaine. In particular, Fawcett *et al.* compared S-bupivacaine with racemic bupivacaine [27]. S-ropivacaine is structurally close to bupivacaine; the main difference is that the former is a pure S-(–) enantiomer where the latter is a racemate. Again, the ester and amide local anaesthetics are grouped in different classes; the agents of low potency and short duration are separated from the agents of high–medium potency and long–medium duration. Moreover, the classification presents lower bias and greater precision, resulting in lower divergence with respect to the original distribution. Therefore, the approach is quite general. However, the inclusion of other local anaesthetics could change the detail, *i.e.* subsequent classifications with more than 11 levels. A natural trend is to interchange similar anaesthetics in the composition of complex drugs, *e.g.* the eutectic mixture of local anaesthetics (EMLA[®], lidocaine–prilocaine 2.5/2.5% w [28]). However, mixtures of dissimilar anaesthetics are also used, *e.g.*, betacaine-LA (lidocaine–prilocaine–dibucaine) [29] and S-caine (1:1 lidocaine–tetracaine eutectic mixture) [30].

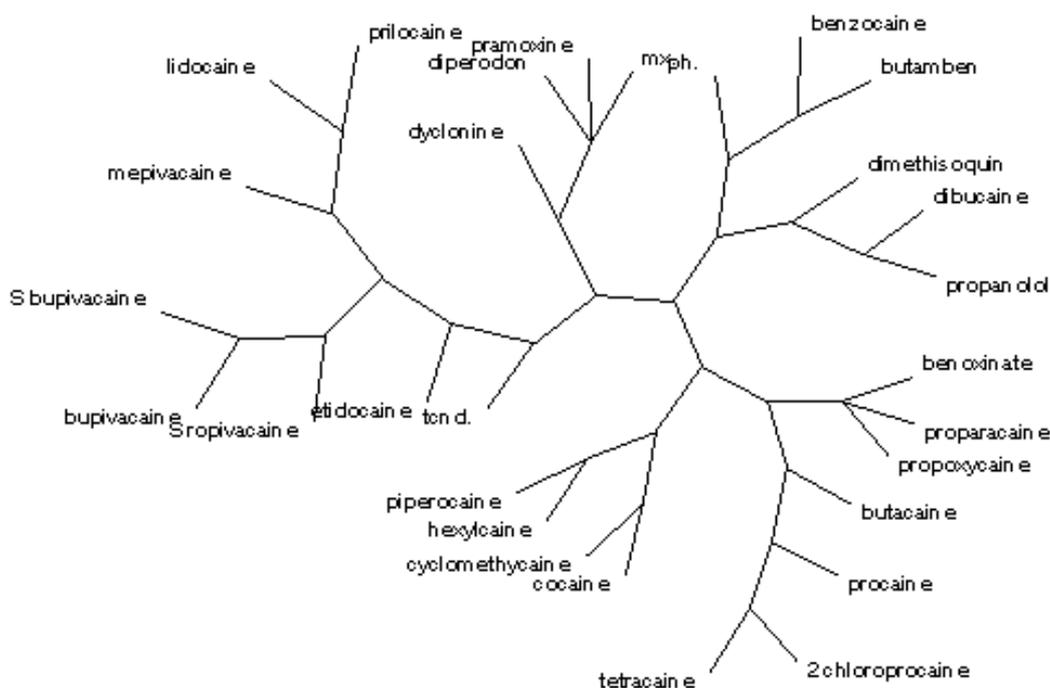


Figure 6. Radial tree for anaesthetics including physicochemical and pharmacological properties.

The predictions for topical anaesthetics and ice, both not included in the models, are included in Table 3. The predictions have been compared with experimental results [29,30]. The relative potency is obtained from the mean pain scores after application of topical anaesthetics for 60 minutes [29,30]. ELA-max is superior to tetracaine and betacaine-LA at 60 minutes, while EMLA[®] is superior to betacaine-LA at 60 minutes, which is in partial agreement with our prediction. The relative potency after removal is obtained from the mean pain scores 30 minutes after removal of the topical

anaesthetics [29,30]. ELA-max and EMLA[®] are superior to tetracaine and betacaine-LA 30 minutes after the 60-minute application period, which is in partial agreement with our prediction. Increased anaesthetic benefit is obtained 30 minutes after removal, which suggests that a reservoir of anaesthetic is located and stored in the upper skin layers during application, providing additional anaesthetic benefit after removal (Table 3). Although EMLA[®] is more potent than ice, ice has advantages in easy of use, fast action, and is less expensive than EMLA[®] [31].

Table 3. Predictions for topical anaesthetics and ice both not included in the models.

Anaesthetic	Ingredients	Prediction			Experiment	
		Onset	Potency	Duration	Relative potency ^a	Rel. pot. after ^b
Betacaine-LA	lidocaine: prilocaine: dibucaine	rapid	moderate	moderate	1.0	1.0
ELA-max	4% lidocaine	rapid	moderate	moderate	1.5	1.5
ELA-max 5	5% lidocaine	rapid	moderate	moderate	–	–
EMLA cream	2.5% lidocaine: 2.5% prilocaine	rapid	moderate	moderate	1.4	1.5
Tetracaine gel	4% tetracaine gel	slow	high	long	1.2	1.1
Amethocaine	4% tetracaine	slow	high	long	–	–
Topocaine	4% lidocaine	rapid	moderate	moderate	–	–
S-caine	2.5% lidocaine: 2.5% tetracaine	moderate	moderate- high	moderate- high	–	–
Ice ^c		moderate	low	low	<1.4	<1.5

^a From mean pain scores after application of topical anaesthetics for 60 minutes [29].

^b From mean pain scores 30 minutes after the 60-minute application period of anaesthetics [29].

^c From mean pain scores after application of topical anaesthetics [31].

SplitsTree is a program for analyzing cluster analysis (CA) data [32]. 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. Furthermore, 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 28 local anaesthetics of Table 1 (*cf.* Figure 7) reveals no conflicting relationship between the anaesthetics. In particular compounds 1, 3, 4, 6, 11, 13–25 and 28 appear superimposed. The splits graph is in general agreement with the partial correlation diagram, dendrogram and binary trees (Figures 1–6). The main difference is the partial fusion of C_{b_2} classes (1,4,6,7,8,14,17,20,21,22,23) and (3,11,12,13,15,16,18,19,24,25,28) corresponding to Figure 4. However, the results (Figure 7) should be taken with care, because the former class includes four compounds with the constant <11111> vector

(anaesthetics 4, 6, 20 and 23), for which the null standard deviation causes a correlation coefficient of $r = 1$ with any local anaesthetic, which is an artifact.

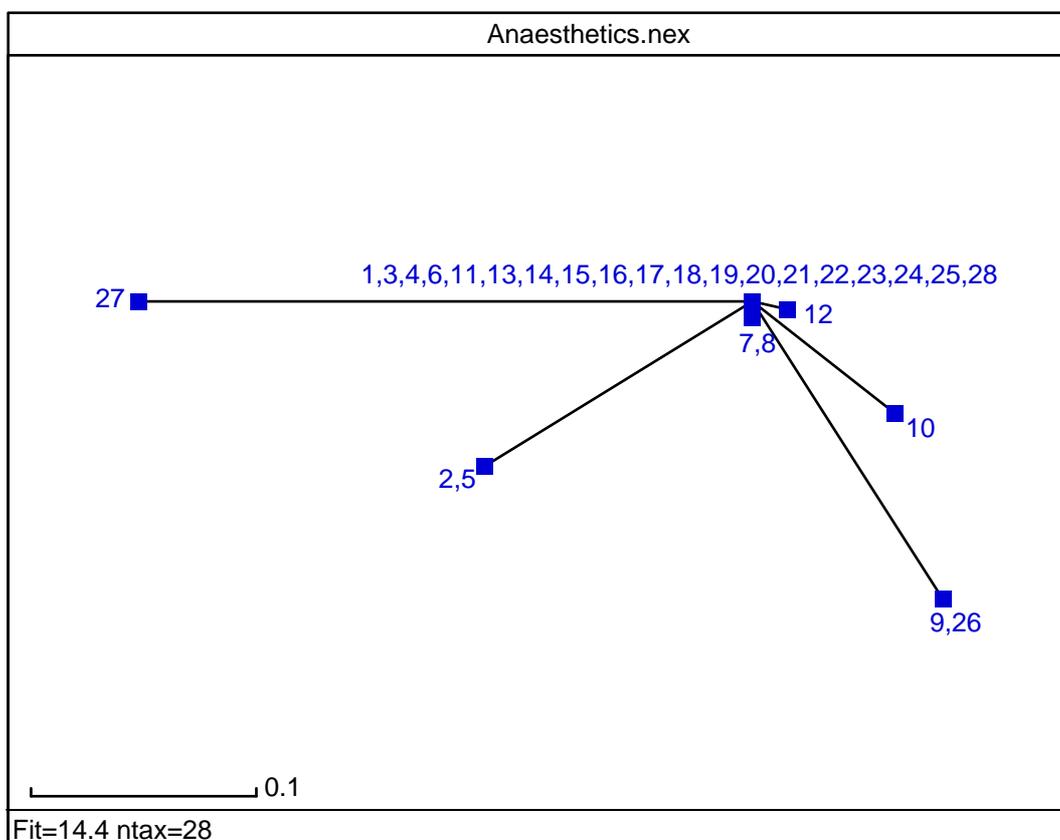


Figure 7. Splits graph for the local anaesthetics analogues of procaine.

A principal component analysis (PCA) [33] 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 4. In particular the use of only the first factor F_1 explains 35% of the variance (65% 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 78% of the variance (22% error).

Table 4. Importance of the principal component analysis factors for the vector property.

Factor	Eigenvalue	Percentage of variance	Cumulative percentage of variance
F_1	1.73558585	34.71	34.71
F_2	1.33757290	26.75	61.46
F_3	0.81501667	16.30	77.76
F_4	0.75678743	15.14	92.90
F_5	0.35503715	7.10	100.00

The PCA factor loadings are shown in Table 5.

Table 5. 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	-0.058	0.756	-0.109	-0.428	0.480
i_2	0.518	-0.034	0.767	0.086	0.367
i_3	-0.453	0.465	0.544	0.184	-0.499
i_4	0.371	0.420	-0.315	0.766	-0.006
i_5	-0.621	-0.186	0.065	0.435	0.621

^aLoadings greater than 0.7 are boldfaced.

The PCA F_{1-5} profile for the vector property is listed in Table 6. 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 14% 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 4% error. For F_3 variable i_2 has the greatest weight; nevertheless, F_3 cannot be reduced to three variables $\{i_2, i_3, i_4\}$ without a 2% error. For F_4 variable i_4 has the greatest weight; however, F_4 cannot be reduced to three variables $\{i_1, i_4, i_5\}$ without a 4% 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 14%, 4%, 2%, 4% and 13% errors, respectively.

Table 6. 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	0.34	26.82	20.48	13.77	38.60
F_2	57.12	0.11	21.61	17.68	3.47
F_3	1.20	58.84	29.64	9.89	0.42
F_4	18.32	0.73	3.37	58.65	18.93
F_5	23.03	13.49	24.89	0.00	38.58

^aPercentages greater than 50% are boldfaced.

In the F_2 – F_1 plot (*cf.* Figure 8), those local anaesthetics analogues of procaine with the same vector property appear superimposed. In particular anaesthetic 27 (class 4) also comes out placed over compounds 14 and 17 (class 1). Five classes of anaesthetics are clearly distinguished: class 1 with 11 units ($0 \approx F_1 < F_2$, *top*), class 2 (11 units, $F_1 > F_2$, *right*), class 3 (2 units, $F_1 \ll F_2 \approx 0$, *left*), class 4 (1 unit, $-1 \approx F_1 < F_2 \approx 0$, *middle*) and class 5 (3 units, $F_1 \gg F_2$, *bottom*). The classification is in agreement with the partial correlation diagram, dendrogram, binary trees and splits graph (Figures 1–7).

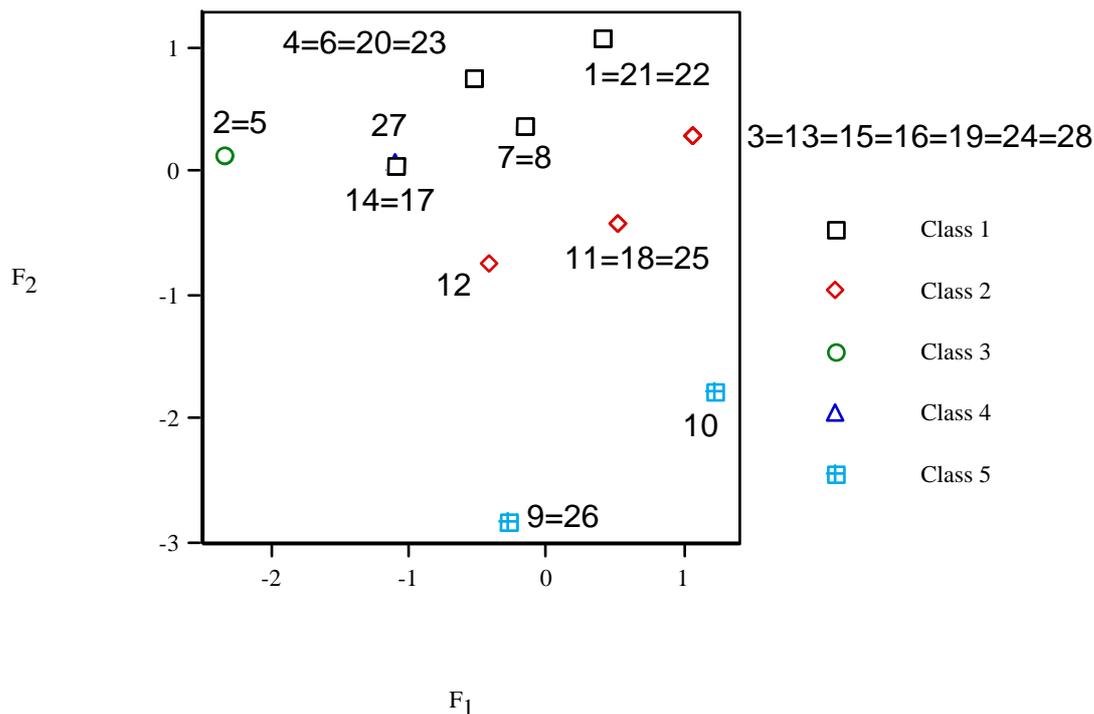


Figure 8. Principal component analysis F_2 vs. F_1 plot for the local anaesthetics.

The recommended format for the periodic table (PT) of the local anaesthetics analogues of procaine is listed in Table 7. 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 $\langle i_1, i_2, i_3 \rangle = \langle 010 \rangle$, viz. $\langle 01001 \rangle$ (dibucaine, propanolol), and $\langle 01010 \rangle$ (dimethisoquin), group g100, for $\langle i_1, i_2, i_3 \rangle = \langle 100 \rangle$, i.e. $\langle 10011 \rangle$ (phenytoin), etc. The local anaesthetics in the same column of Table 7 appear close in the partial correlation diagram, dendrogram, radial trees, splits graph and PCA (Figures 1–8).

Table 7. Table of periodic properties for the local anaesthetics analogues of procaine.

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

Figure 9 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 $\{i_4, i_5\}$.

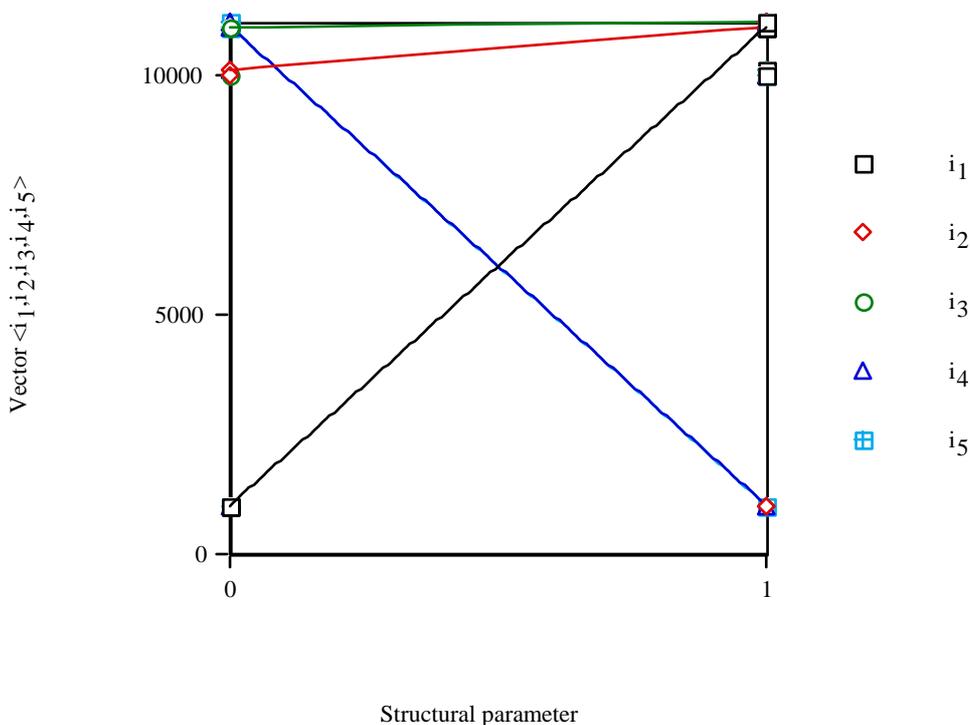


Figure 9. Variation of the vector property of local anaesthetics vs. counts $\{i_1, i_2, i_3, i_4, i_5\}$.

The variation of the vector property $P = \langle i_1, i_2, i_3, i_4, i_5 \rangle$, as a function of the number of the group in PT (cf. Figure 10) for local anaesthetics, reveals that the minima correspond to anaesthetics with $\{i_1, i_2, i_3\}$ values of $\langle 010 \rangle$ (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 $\langle i_1, i_2, i_3, i_4, i_5 \rangle$ 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 $\langle i_1, i_2, i_3, i_4, i_5 \rangle$ 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, which correspond to the anaesthetics of group g010. Their $\langle i_1, i_2, i_3, i_4, i_5 \rangle$ values are not repeated as the periodic law (PL) states but decrease regularly.

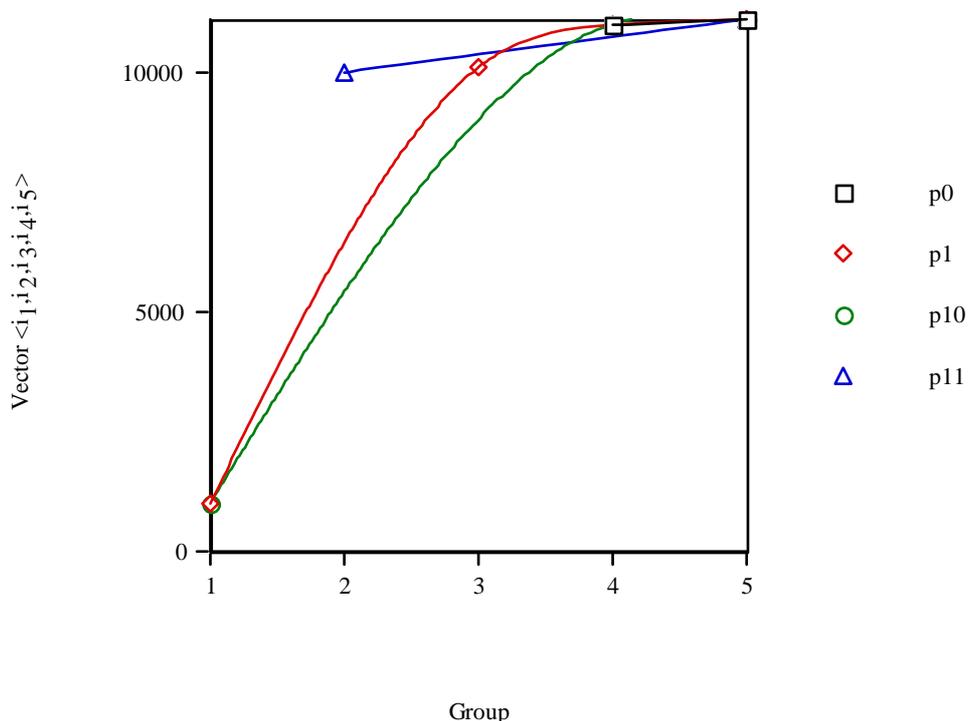


Figure 10. Variation of the vector property of local anaesthetics vs. group number.

An empirical function $P(i_1, i_2, i_3, i_4, i_5)$ reproduces, with enough precision, the different $\langle i_1, i_2, i_3, i_4, i_5 \rangle$ 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:

$$\begin{aligned} P_{\min}(p) < P(p-1) \\ P_{\min}(p) < P(p+1) \end{aligned} \tag{8}$$

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

$$\begin{aligned} P_{\min}(p) - P(p-1) < 0 \\ P(p+1) - P_{\min}(p) > 0 \end{aligned} \tag{9}$$

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 $D(p)$:

$$D(p) = P(p+1) - P(p) \tag{10}$$

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:

$$D(p) > 0 \text{ or } D(p) < 0 \tag{11}$$

$$R(p) > 1 \text{ or } R(p) < 1 \tag{12}$$

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 11 presents lack of coherence between the $\langle i_1, i_2, i_3, i_4, i_5 \rangle$ Cartesian and PT representations. If consistency were rigorous all the points in each

period would have the same sign. In general there is a trend 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.

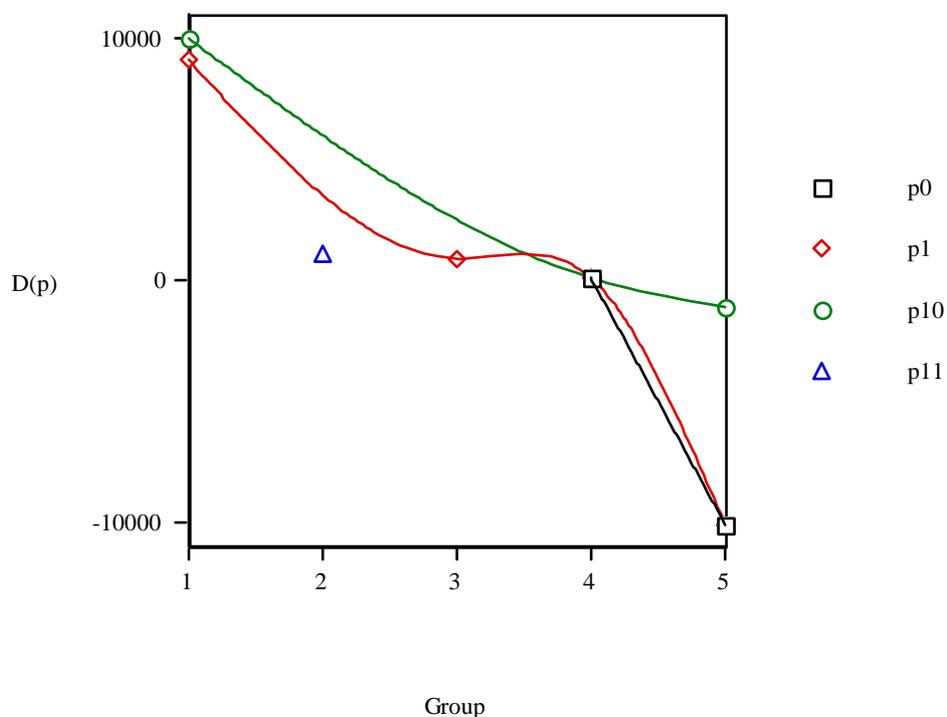


Figure 11. Variation of $D(p) = P(p+1) - P(p)$ vs. group number. P is the vector property.

The change of $R(p)$ vs. group number in Figure 12 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 trend 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.

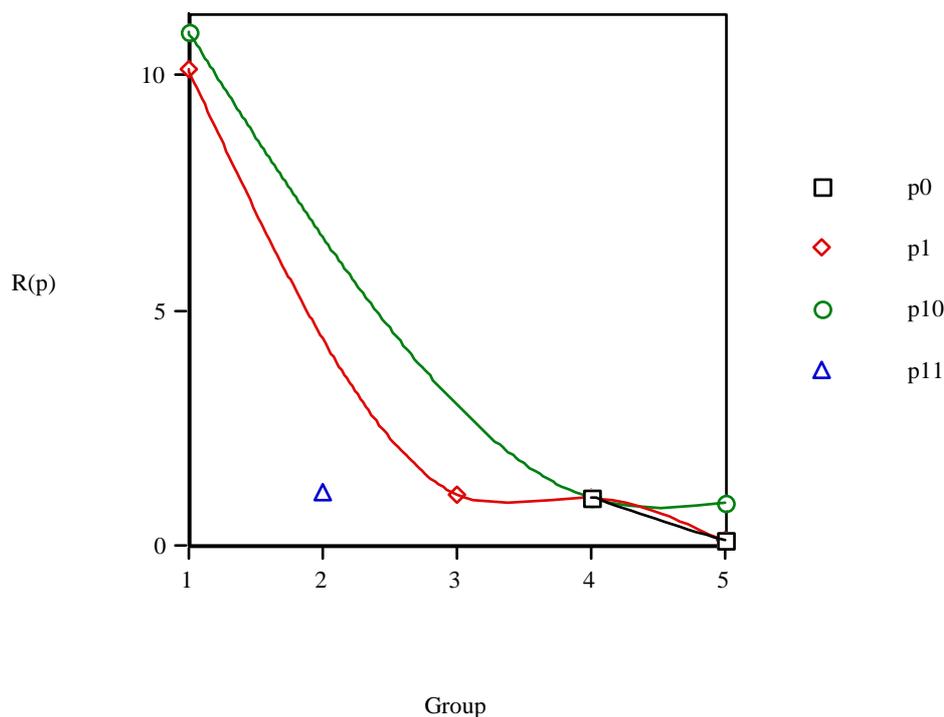


Figure 12. Variation of $R(p) = P(p+1)/P(p)$ vs. group number. P is the vector property.

7. Conclusions

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

1. 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.*

2. Many algorithms for classification 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. The method avoids the problem of other methods that consider continuum variables, because for the four compounds with constant $\langle 11111 \rangle$ vector (anaesthetics 4, 6, 20 and 23), the null standard deviation causes a correlation coefficient of $r = 1$ with any local anaesthetic. The lower level classification processes show lower entropy and, therefore, may be more parsimonious.

3. In this work, an overview of an information entropy approach (based on the equipartition conjecture) to the modelling of complex data in the area of cheminformatics has been presented. Through the proposed method we intended to show that the information entropy-based modelling of complex systems can be effectively equipped with expressive representation of complex data in the form of structured representation. As a result, concerning chemical and biological problems, we have shown that predictions can be done directly from molecular structures, introducing potential benefits in

the current quantitative structure–property/activity relationship (QSPR/QSAR) method. In particular, since universal approximation capabilities of the equipartition conjecture have been proved (specifically for tree-structured domains [15]), the equipartition conjecture can be considered a general tool especially useful to deal with new tasks where the relevance of the *ad hoc* molecular descriptors is unknown. More generally, the present approach can be seen as a paradigmatic example of the studies aiming at extending information entropy techniques to the treatment of various structured classes of data [17]. Our aim here is also to propose the new approach as a general method to tackle various structured problems in the areas of cheminformatics and bionformatics. Main potential developments concern hard tasks in toxicology and bioinformatics whenever is natural to find useful structured representation of chemical/biological data.

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 [34]. As suggested by Senese *et al.* [35] in a different context, the generated clusters can be used to generate different QSAR models in order to obtain better representation of the data. Thus, clustering methods can be used to identify single QSARs representing each of them different information that can be overlooked when trying to represent all the data by only one.

5. Information entropy and principal component analyses permit classifying the local anaesthetics and agree. The ester (benzocaine, 2-chloroprocaine, procaine, tetracaine) and amide type local anaesthetics (bupivacaine, dibucaine, etidocaine, lidocaine, mepivacaine, prilocaine, S-ropivacaine) are always grouped in different classes. The agents of low potency and short duration (procaine, 2-chloroprocaine) are separated from the agents of high potency and long duration (bupivacaine, etidocaine, S-ropivacaine), while the agents of moderate potency and duration (mepivacaine, prilocaine) are classified together with the latter. The final classification is shown more precise and with lower bias. The classification model calculates in each case the contribution of signal and noise.

6. 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.

7. Bupivacaine cardiotoxicity results from prolonged Na^+ channel dwell time of the R-, as compared with the S-, stereoisomer. Bupivacaine, like most aminoamide local anesthetics, has a chiral C atom where the amide linkage joins the hydrophilic tail. Chirality yields two steric forms (S and R) which are spatial mirror images with different receptor kinetics; commercial bupivacaine is the optically inactive racemic (RS) mixture of R- and S-bupivacaine. Ropivacaine is unique in that membrane

separation synthesis exclusively yields the S-(–)-enantiomer, which is a local anesthetic with lower cardiotoxic potential than racemic bupivacaine.

8. As the options for the practitioner continue to grow, the need for studies comparing onset of action, efficacy and safety continues to be of paramount importance. A natural trend is to interchange similar anaesthetics in the composition of complex drugs (EMLA®).

Acknowledgements

The authors acknowledge the referees for their kind comments, as well as 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 and Notes

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. Covino, B. G. Local Anesthesia. *N. Engl. J. Med.* **1972**, *286*, 975-983.
7. Covino, B. G. Local Anesthetic Agents for Peripheral Nerve Blocks. *Anaesthetist* **1980**, *29*(7), 33-37.
8. Covino, B. G. Pharmacology of Local Anaesthetic Agents. *Br. J. Anaesth.* **1986**, *58*, 701-716.
9. Krasowski, M. D.; Harrison, N. L. General Anaesthetic Actions on Ligand-Gated Ion Channels, *Cell. Mol. Life Sci.* **1999**, *55*, 1278-1303.
10. Campagna, J. A.; Miller, K. W.; Forman, S. A. Mechanisms of Actions of Inhaled Anesthetics, *N. Engl. J. Med.* **2003**, *348*, 2110-2124.
11. Franks, N. P.; Lieb, W. R. Molecular and Cellular Mechanisms of General Anaesthesia, *Nature (London)* **1994**, *367*, 607-614.
12. Varmuza, K. *Pattern Recognition in Chemistry* ; Springer: New York, 1980.
13. Benzecri, J.-P. *L'Analyse des Données I. La Taxonomie* ; Dunod: Paris, 1984.
14. Adams, H. J.; Ronfeld, R. A.; Takman, B. H. In *Handbook of CNS Agents of Local Anaesthetics* ; Verderame, M., Ed., CRC: Boca Raton, 1986.

15. Tondeur, D.; Kvaalen, E. Equipartition of Entropy Production. An Optimality Criterion for Transfer and Separation Processes. *Ind. Eng. Chem., Fundam.* **1987**, *26*, 50-56.
16. Kimura, M. *New Scientist* **1986**, *11*, 41-41.
17. Corriou, J. P.; Iordache, O.; Tondeur, D. Classification of Biomolecules by Information Entropy. *J. Chim. Phys. Phys.-Chim. Biol.* **1991**, *88*, 2645-2652.
18. Kaufmann, A. *Introduction à la Théorie des Sous-ensembles Flous*; Masson: Paris, 1975; Vol. 3.
19. Shannon, C. E. A Mathematical Theory of Communication: Part I, Discrete Noiseless Systems. *Bell Syst. Tech. J.* **1948**, *27*, 379-423.
20. Shannon, C. E. A Mathematical Theory of Communication: Part II, The Discrete Channel with Noise. *Bell Syst. Tech. J.* **1948**, *27*, 623-656.
21. White, H. *AI Expert* **1989**, *12*, 48-48.
22. 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.
23. Tryon, R. C. *J. Chronic Dis.*, **1939**, *20*, 511-524.
24. Jarvis, R. A.; Patrick, E. A. Clustering Using a Similarity Measure Based on Shared Nearest Neighbors. *IEEE Trans. Comput.*, **1973**, *C22*, 1025-1034.
25. Page, R. D. M. Program TreeView. University of Glasgow, 2000.
26. Faus, V.; López, A.; Martín, A.; Herráez, M. V Congreso de la Sociedad Española de Farmacia Industrial y Galénica, Valencia, 2001, pp. 161-162.
27. Fawcett, J. P.; Kennedy, J. M.; Kumar, A.; Ledger, R.; Kumara, G. M.; Patel, M. J.; Zacharias, M. Comparative Efficacy and Pharmacokinetics of Racemic Bupivacaine and S-Bupivacaine in Third Molar Surgery. *J. Pharm. Pharmaceut. Sci.*, **2002**, *5*, 199-204.
28. Brodin, A.; Nyquist-Mayer, A.; Wadstein, T. Phase Diagram and Aqueous Solubility of the Lidocaine-Prilocaine Binary System. *J. Pharm. Sci.* **1984**, *73*, 481-484.
29. Friedman, P. M.; Fogelman, J. P.; Nouri, K.; Levine, V. J.; Ashinoff, R. Comparative Study of the Efficacy of Four Topical Anesthetics. *Dermatol. Surg.* **1999**, *25*, 950-954.
30. Friedman, P. M.; Mafong, E. A.; Friedman, E. S.; Geronemus, R. G. Topical Anesthetics Update: EMLA and Beyond. *Dermatol. Surg.* **2001**, *27*, 1019-1026.
31. Kuwahara, R. T.; Skinner, R. B., Jr., EMLA Versus Ice as a Topical Anesthetic. *Dermatol. Surg.* **2001**, *27*, 495-496.
32. Huson, D. H. SplitsTree: Analyzing and Visualizing Evolutionary Data. *Bioinformatics* **1998**, *14*, 68-73.
33. Hotelling, H. Analysis of a Complex of Statistical Variables into Principal Components. *J. Educ. Psychol.*, **1933**, *24*, 417-441.
34. He, L.; Jurs, P. C. Assessing the Reliability of a QSAR Model's Predictions, *J. Mol. Graphics Model.* **2005**, *23*, 503-523.
35. Senese, C. L.; Hopfinger, A. J. A Simple Clustering Technique to Improve QSAR Model Selection and Predictivity: Application to a Receptor Independent 4D-QSAR Analysis of Cyclic Urea Derived Inhibitors of HIV-1 Protease, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 2180-2193.