Available online at www.sciencedirect.com
ScienceDirect
Tetrahedron 64 (2008) 2216–2224
Tetrahedron
www.elsevier.com/locate/tet

Exploration of quantitative structure–property relationships (QSPR) for the design of new guanidinium ionic liquids

Gonçalo V.S.M. Carrera [a,b], Luís C. Branco [a,b], Joao Aires-de-Sousa [a,*], Carlos A.M. Afonso [b,*]

[a] *REQUIMTE, CQFB, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal*
[b] *CQFM, Instituto Superior Técnico, Universidade Técnica de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal*

Received 12 October 2007; received in revised form 5 December 2007; accepted 10 December 2007
Available online 14 December 2007

Abstract

Computer-aided design of new guanidinium salts was explored and experimentally tested, en route to the discovery of new ionic liquids. Quantitative structure–property relationships were established to predict the mp of guanidinium salts of four different anionic families (Cl^-, BPh_4^-, Br^-, and I^-). Models were built with a data set of 101 salts and counterpropagation neural networks. Predictions for an independent test set were obtained with R^2=0.815, and a fivefold cross-validation procedure yielded R^2=0.742. Assisted by the models, six new guanidinium salts were prepared, and the measured melting properties were reasonably in accordance with the predictions. One of the new chloride salts is liquid at room temperature, and three tetraphenylborate salts have mp values lower than those previously available in the data set for that anion.

1. Introduction

The importance of ionic liquids (IL) as potential green alternatives to classic volatile organic solvents[1–3] primarily relates to their almost negligible vapor pressure,[4] low flammability, and easiness of recycling. Beyond green properties, ILs have found a diversity of applications due to their high ionic conductivity, large electrochemical window,[5] and high range of densities,[6] and viscosities.[7] However, other properties also have to be considered in the context of sustainable chemistry—and have just started to be addressed—such as toxicity and life cycle features.[8,9] The design of an IL requires the optimization of the chemical structure, to render the profile of properties compatible with the desired application. This demands a pool of available structural families of ILs, as well as the exploration of almost unending combinations of cations and anions.

The strongly basic guanidine moiety, present in natural products incorporating the arginine amino acid,[10] was recently described as a new generation of ionic liquids.[11,12] Hexa-alkyl substituted guanidinium salts show physico-chemical properties comparable and complementary to the generally used 1,3-dialkylimidazolium-based ionic liquids. Guanidinium-based ionic liquids were used as reaction media, for example, in the oxidation of benzyl alcohols,[13] or in Heck reactions facilitating the β-hydride elimination and acting as a ligand to stabilize activated Pd species.[14] Chiral anions were coupled to guanidinium cations yielding ionic liquids, which (as solvent) could induce asymmetry in catalytic Rh(II) carbenoid C–H insertion and Sharpless dihydroxylation.[12] In a different field, an iodide guanidinium ionic liquid was applied as an iodide resource for dye-sensitized nanocrystalline solar cells.[15]

There is a considerable universe of possible salts to obtain by combining different alkyl groups in hexa-alkyl guanidinium cations with different anions. The obvious first requirement for a salt to be an ionic liquid is a low mp. The ability to predict the mp from the molecular structure is of potential high value in the discovery of new ionic liquids, allowing laboratory efforts to focus on structures exhibiting optimum predicted properties. Prediction and explanation of IL mps have been explored with theoretical quantum calculations.[16] Differently, several attempts have been made to apply QSPR (Quantitative Structure–Property Relationship) techniques to model the mp

* Corresponding authors.
E-mail addresses: goncalo.carrera@dq.fct.unl.pt (G.V.S.M. Carrera), lbranco@dq.fct.unl.pt (L.C. Branco), jas@fct.unl.pt (J. Aires-de-Sousa), carlosafonso@ist.utl.pt (C.A.M. Afonso).

0040-4020/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tet.2007.12.021

of pyridinium,[17] ammonium,[18] imidazolium,[19–21] and triazolium[22] salts. Varnek et al.[23] assessed the performance of QSPR techniques in this endeavor by applying the most common linear and non-linear machine learning methods, and different types of commonly used molecular descriptors, to a large data set of structurally diverse IL. They concluded that only a moderate accuracy of predictions could be achieved (RMSE of prediction for independent test sets in the range 37.5–46.4 °C), which was suggested to result from the quality of the experimental data as well as to difficulties to take into account the structural features of IL in the solid state (polymorphic effects, eutectics, glass formation). Here we explore QSPR models for the mp of guanidinium salts using a data set of 101 compounds that cover mp values from −76 °C to 322 °C, and include four different anionic families (Cl^-, Br^-, I^-, and BPh_4^-). The counterpropagation neural network (CPG NN) was employed as the machine learning technique. For a model to assist in the discrimination between potential ionic liquids, and high-mp salts, it should be trained with a data set including both salts with mp values above and below room temperature. The large proportion of compounds in the data set exhibiting mp above room temperature reflects the current inexistence of many guanidinium ionic liquids. The aim of this study was thus to make use of the available information to design new salts with mp values as low as possible.

A CPG NN performs a non-linear projection of the molecular structures on a 2D surface allowing for an easy visualization of relationships between structural features of the cation and the mp, as well as the influence of the anion on the mp. CPG NNs allow for multidimensional output, and can be trained with data sets including non-available data for some output values of some objects. These are the main reasons CPG NNs were used in our study. We need multidimensional output (the mp values of salts with four possible anions), and the mp value is almost never available for all the four anions of a cation.

Assisted by the obtained model, six new salts were proposed that were eventually synthesized and tested in the laboratory.

2. Computational methods

2.1. Data set

A data set of 101 guanidinium salts were collected from the literature[24] including four different anionic families (Cl^-, Br^-, I^-, and BPh_4^-), and covering a range of mp values from −76 to 322 °C. The range of mp values for each anion class is as follows: chlorides between −76 and 293 °C, tetraphenylborates between 127 and 322 °C, bromides between 93 and 214 °C, and iodides between 46 and 300 °C. The data set were divided into a training set with 81 objects (31 chlorides, 15 tetraphenylborates, 9 bromides, and 26 iodides) and a test set with 20 objects (8 chlorides, 6 tetraphenylborates, 2 bromides, and 4 iodides). In order for a model to recognize room temperature salts, three RT liquid chloride salts were included in the training set although they have no defined mp, but a glass transition temperature (T_g). The full data set is available as Supplementary data.

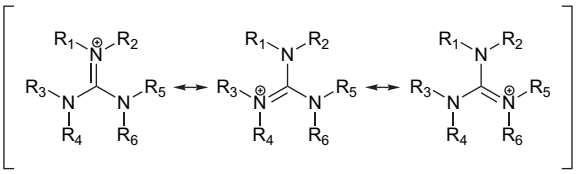

Figure 1. Three resonance structures of a guanidinium cation.

For the computation of molecular descriptors, each guanidinium cation was represented by three resonance structures to enable the comparison of guanidinium salts with asymmetric distribution of substituents (Fig. 1). The final prediction for the mp of a guanidinium salt was thus obtained as the average value of the predictions for the three resonance structures.

2.2. Molecular descriptors

The mp of salts, such as those studied here, mostly reflects the size, symmetry, and branching of the cation, and could be influenced by hydrogen bonds between cation and anion.[25] Molecular descriptors were therefore selected, or designed, to encode such features.

A series of 92 molecular descriptors were initially calculated to represent the structure of the cation. DRAGON 3.0[26] was used to compute 52 descriptors, including 14 constitutional descriptors related to the number of bonds (double, triple, aromatic, rotatable) and number of specific atoms (C, N, O, H), four functional-group descriptors accounting for H-bond donor and acceptor sites, and also related to the degree of substitution of aromatic rings, 33 topological descriptors related to structural flexibility and symmetry, and an atom-centered fragment descriptor (represents the number of H atoms connected to an heteroatom, here corresponding to the number of unsubstituted positions of the guanidinium). With PETRA 3.1[27] software, 28 additional descriptors were calculated to encode molecular size, aromatic and aliphatic indices, and the distribution of linear flexible chains. Finally, 12 new descriptors were designed to encode the size of the fragments bonded to the guanidinium N atoms. For example, six descriptors are the number of non-hydrogen atoms in the six substituents of the guanidinium core. For the calculation of these descriptors, the N atoms of the core were ranked according to the number of non-hydrogen atoms in their two substituents, and then the two substituents of each N atom were ranked according to the same criterion. In case the number of non-hydrogen atoms cannot distinguish substituents, molecular weight and then size of linear chain were used as criteria. The full list of descriptors is available as Supplementary data.

2.3. Counterpropagation neural networks

A counterpropagation neural network (CPG NN) consists of a 2D grid of neurons organized in an input layer and an output layer (Fig. 2).[28] Each neuron of the input layer has

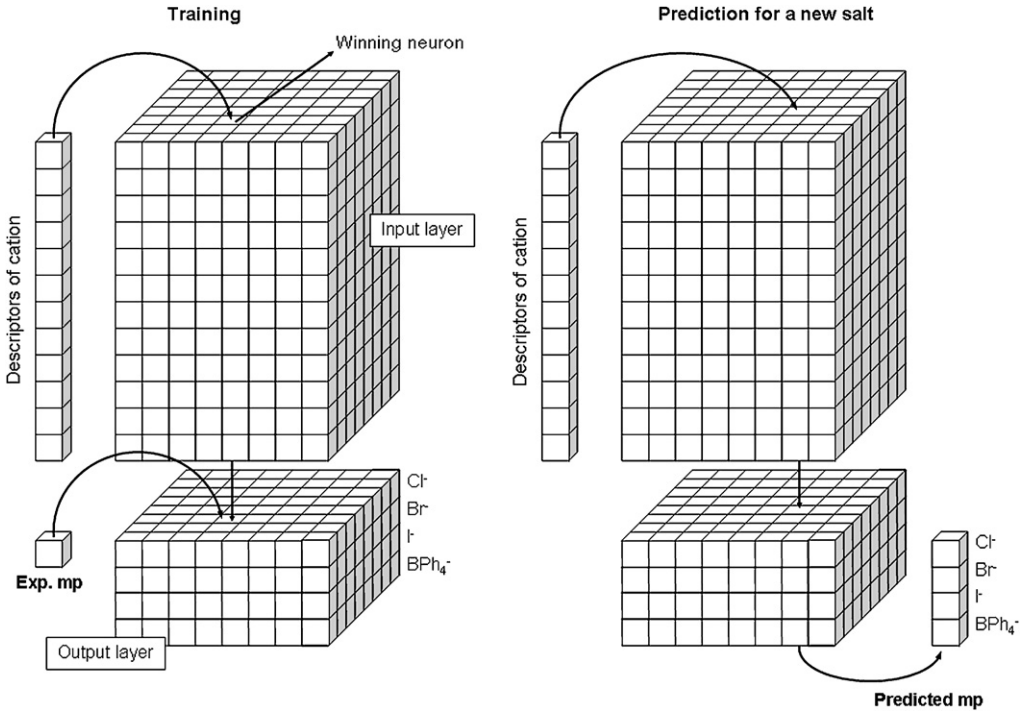

Figure 2. Learning and prediction in a counterpropagation neural network (CPG NN).

a number of weights (represented by small boxes in Fig. 2) equal to the number of molecular descriptors. Neurons of the output layer have as many weights as the number of output values to learn. In our study, the structure of the cation is encoded with molecular descriptors and submitted to the input layer, while the information concerning the mp for their salts is stored in the output layer. Because we have four possible anions, each output neuron has four weights.

Before the training, all weights take random numbers; during the training the objects from the training set are iteratively submitted to the network and the weights are adapted. When an object (a cation) is presented to the net, it activates the neuron (winning neuron) with the input weights most similar to the input vector (molecular descriptors); then the input weights are adjusted to become even closer to the molecular descriptors of the cation. The neurons in the neighborhood of the winning neuron are also corrected, the extent of adjustment depending on the topological distance to the central neuron. In the output layer, the neuron corresponding to the winning neuron has one of its weights adapted—the weight corresponding to the anion of the presented salt—to become closer to the experimental mp. All the objects of the training set are submitted to the network a predefined number of times. After the training process, the network is able to make predictions for new salts. The descriptors of the cation are submitted, a winning neuron is found, and the output weights are retrieved as predictions (Fig. 2). Identification of the neurons activated by objects of a data set yields a map with those objects distributed on the surface of the CPG NN. Visualization of the weights at the output layers shows which regions yield which output values.

In this work, the molecular descriptors were scaled linearly between zero and one and CPG NN of size 18×18 were trained over 100 cycles with an initial learning span of 8, using in-house developed software based on JATOON[29,30] Java applets. The training procedure typically took less than 30 s using a PC with a Pentium™ IV 2.66 GHz CPU.

2.4. Validation methods

Besides validation with an independent test set, the models were also validated by y-randomization, and by fivefold cross-validation.

In the y-randomization test, the values of the mp for the objects of the training set were scrambled, and CPG NNs were trained using those values. Predictions were then obtained for the objects of the test set. The test was run five times, with different random assignment of mp values to guanidinium salts of the training set.

3. Results and discussion

3.1. QSPR studies

A CPG NN was trained with the whole profile of molecular descriptors as the input and the mp values as the output. The output layer included four levels (one for each anionic family). During the training, submission of a salt always influenced the weights at the input layer, but only caused correction of output weights at the level corresponding to the anion of the salt. The predictive performance of the model, both for the training set and the test set, is presented in Table 1. To overcome

Table 1
Prediction of mp for guanidinium salts of four anionic families with CPG NN models

		Training set		Test set	
		CPG NN	Ensemble of five CPG NN	CPG NN	Ensemble of five CPG NN
All	R^2	0.866	0.865	0.791	0.815
	rms	29.49	30.10	25.48	23.95
	SP	0.920	0.9279	0.865	0.8145
	$\bar{e}$	19.96	20.43	22.01	20.33
	n Objects		81		20
Chloride	R^2	0.864	0.8655	0.761	0.670
	rms	34.62	36.78	22.15	25.21
	SP	0.880	0.883	0.905	0.786
	$\bar{e}$	23.13	25.90	18.65	20.37
	n Objects		31		8
BPh_4^-	R^2	0.798	0.8489	0.880	0.909
	rms	24.03	21.36	30.87	25.78
	SP	0.877	0.786	0.857	0.886
	$\bar{e}$	18.87	16.33	27.58	23.16
	n Objects		15		6
Bromide	R^2	0.926	0.895	—	—
	rms	11.85	13.92	—	—
	SP	0.942	0.892	—	—
	$\bar{e}$	8.66	9.24	—	—
	n objects		9		2
Iodide	R^2	0.776	0.789	0.721	0.903
	rms	30.51	29.64	24.23	20.30
	SP	0.808	0.828	0.4	0.8
	$\bar{e}$	21.38	20.16	22.1	18.06
	n Objects		26		4

R^2: square of the Pearson correlation coefficient; rms: root mean square error; SP: Spearman correlation coefficient; $\bar{e}$: mean absolute error.

fluctuations induced by the random factors influencing the training, five independent CPG NNs were trained with the same objects, generating an ensemble of CPG NNs, and the final predictions were obtained by averaging the output values from individual nets. Results are also included in Table 1.

Good correlations were obtained for both training and test sets, with the ensemble of networks yielding improved results for the whole test set, for the BPh_4^- series, and for the iodide series. A graphical representation of the predicted against experimental values is provided in Figure 3. It is observed that errors did not increase if anionic families are considered individually. Absolute errors of 20 °C were typically obtained.

The model was further validated with fivefold cross-validation. The whole data set was randomly divided into five groups, and in five independent experiments a model was trained with four groups obtaining predictions for the group left out. Predictions for the objects with the highest and lowest mp within each anionic family were discarded (when they are left out of the training, their mp are outside the training range). An R^2 value of 0.742 was obtained for the fivefold cross-validation.

In y-randomization tests, the five models trained with random values for the mp, predicted the test set with R^2 values of 0.0625, 0.2731, 0.0102, 0.0426, and 0.0451.

Selection of descriptors was attempted with a Kohonen neural network.[31] The matrix of salts and their descriptors was transposed, transforming descriptors into objects, and salts into descriptors. After a 10×10 Kohonen neural network was trained, one object was extracted from each occupied neuron. In this way, 16 descriptors were selected. Ensembles of five CPG NNs trained on the basis of the selected descriptors to predict mp achieved moderate accuracy (R^2 of 0.600 in fivefold cross-validation), but could never attain the results obtained with all the descriptors.

These results illustrate the difficulty in finding individual molecular descriptors, or a small group of descriptors, with the ability to build robust predictive models of mp that can sustain external validation. In our experiments, one compound (*N*,*N*-dimethyl-*N*′,*N*′-cyclopentamethylene-*N*″,*N*″-cyclopentamethylene guanidinium chloride) was initially included in the data set and later removed from the test set as it has a skeleton not available in the chloride series, activated an empty region of the chlorides map, and exhibited a mp much higher than predicted. The models presented here shall be seen more as a way to estimate the mp on the basis of comparisons with the molecular descriptor profile of known salts. In further discussions, we are referring to the models trained with the entire profile of descriptors, as they allowed for the better predictions. The structure and mp values for compounds discussed in the text are represented in Figure 4.

CPG NNs provide an easy way to visualize relationships between structures and their properties. Figure 5 represents the four output levels of a CPG NN, with ranges of weight values assigned to colors. Output weights correspond to the

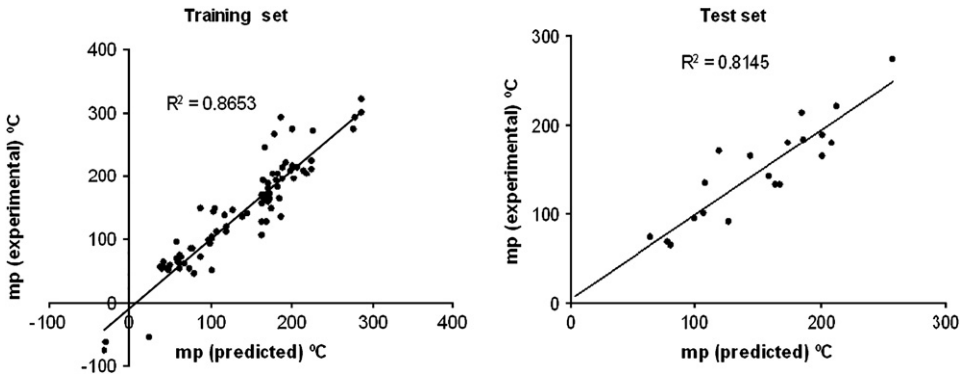

Figure 3. Comparison between experimental mp values and predictions obtained by an ensemble of five CPG NNs.

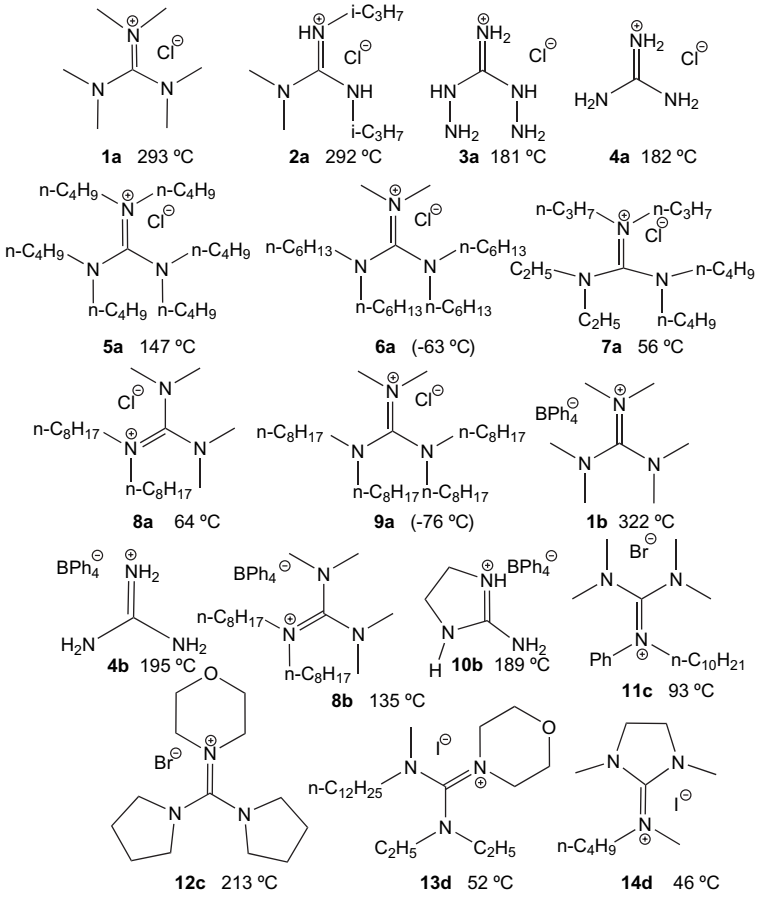

Figure 4. Molecular structures and mp values for compounds discussed in the text. Values in parentheses represent T_g values.

mp values that are produced as predictions of the network when the corresponding neurons are activated. Neurons displayed with blue colors are associated with high mp, while red colors are associated with low mp values. At the same time, inspection of the cationic structures activating neurons in different regions of the map allowed for the identification of regions broadly corresponding to types of structures.

The colors associated with the different regions of the map are highly correlated in the four maps (corresponding to the four anionic families), suggesting that the relationships between the mp and the structural features of the cation are similar for all anions. It is noted that the surface of a CPG NN map has a toroidal topology, i.e., the right and left neurons are considered to be adjacent, as well as the top and bottom neurons. In this way, all neurons have the same number of neighbors.

It was observed that cations activating regions around dark blue neurons (high mp) typically correspond to guanidinium structures bearing unsubstituted positions and small alkyl groups. Examples are salts **1a** and **1b** (neuron R3, Fig. 5a and b, respectively), **2a** (neuron Q5, Fig. 5a), **3a** (neuron Q17, Fig. 5a), **4a** and **4b** (neuron R18, Fig. 5a and b, respectively), **10b** (neuron A16, Fig. 5b), and **12c** (neuron C18, Fig. 5c). Molecular structures of the mentioned examples are displayed in Figure 4.

Pale blue regions, corresponding to intermediate mp values, were typically associated with guanidinium structures with a symmetric distribution of small alkyl chains. An example is hexabutyl guanidinium chloride **5a** (neuron F5, Fig. 5a).

The regions corresponding to red colors (low mp) were activated by guanidinium structures typically incorporating large and small alkyl substituents, such as **6a** (neuron F8, Fig. 5a), **7a** (neuron B10, Fig. 5a), **8a** and **8b** (neuron G11, Fig. 5a and b, respectively), **11c** (neuron I14, Fig. 5c), **13d** (neuron L12, Fig. 5d), and **14d** (neuron P13, Fig. 5d).

The influence of some major structural features of the cation on the mp can be interpreted in terms of known effects for other cations.[21] Large alkyl chains decrease charge density (and electrostatic forces) and promote conformational flexibility, both factors decreasing the mp. In the opposite direction, symmetry allows for better crystal packing, and increase mp.

A trend is observed for the chloride salts to exhibit lower mp values than other anionic classes. However, the distribution of color in the four maps is very similar, which means that the cationic structures giving rise to the lowest (or highest) mp values in one anionic class, generally also give rise to the lowest (or highest) mp values with another anion. The main differences, which appear at the top and bottom center of the maps, are mainly due to different compositions of the data sets for the four anions, i.e., a cation may be available with one anion

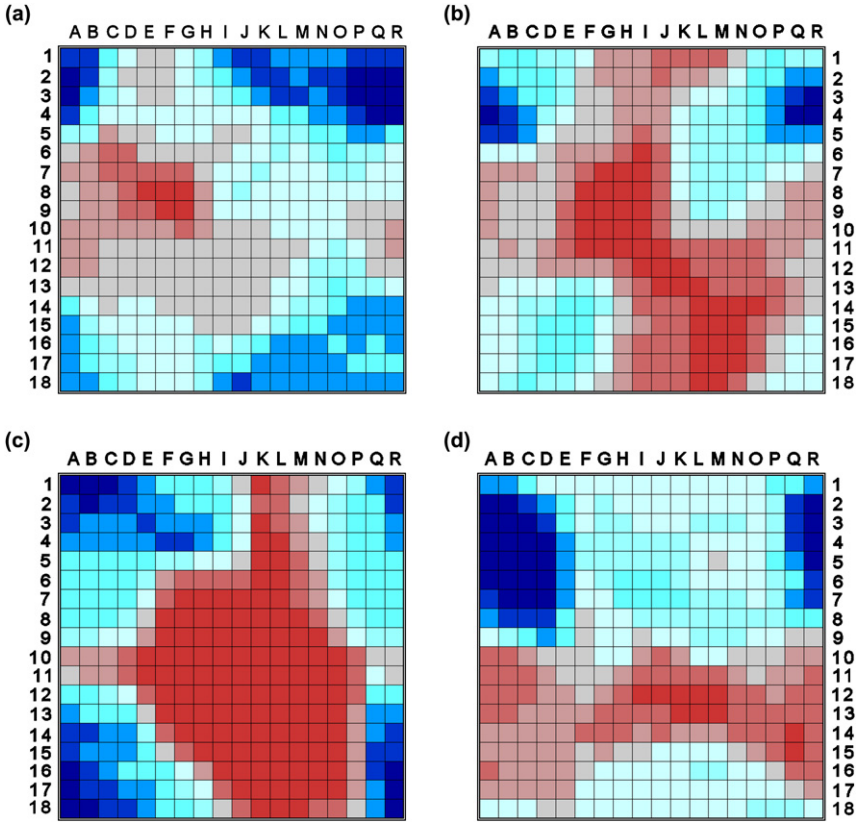

Figure 5. Representation and comparison of the four output levels of the CPG neural network, corresponding to the four anionic families, a: chloride, b: tetraphenylborate, c: bromide and d: iodide. Colors encode weight values—blue neurons represent high mp while red neurons represent low mp. More details on the mapping are provided as Supplementary data.

but not with another—for example, the top center region is particularly well represented by iodide salts.

3.2. Synthesis and testing of new guanidinium salts

Based on the ensemble of CPG NNs, six new salts were proposed to prepare and test in the laboratory: three chloride salts (**15a**, **16a**, and **17a**) and three BPh_4^- salts (**15b**, **16b**, and **6b**). The cations of **16a**, **16b**, and **6b** activate neurons with strongly low predicted mp. For example, in the maps of Figure 5, they activate the neuron F8, a red neuron also activated by **9a** (T_g=−76 °C) and **6a** (T_g=−63 °C). These three cationic structures have four large alkyl substituents and two methyl substituents, cations of **16a**, **16b** exhibiting two ramified chains. Cations of **15a**, **15b** (both J8, Fig. 5a and b, respectively), and **17a** (Q7, Fig. 5a) activate neurons in regions of intermediate mp values. While the structures of **15** have four methyl substituents and two octadecyl alkyl chains, the structure of **17a** is quite symmetrical with two methyl substituents and four C3 chains with a double bond.

The six compounds were prepared according to a published procedure,[11] and their mp properties were measured (Fig. 6). The values for the chloride salts were well in accordance with predictions (average error 25 °C), **16a** showing an

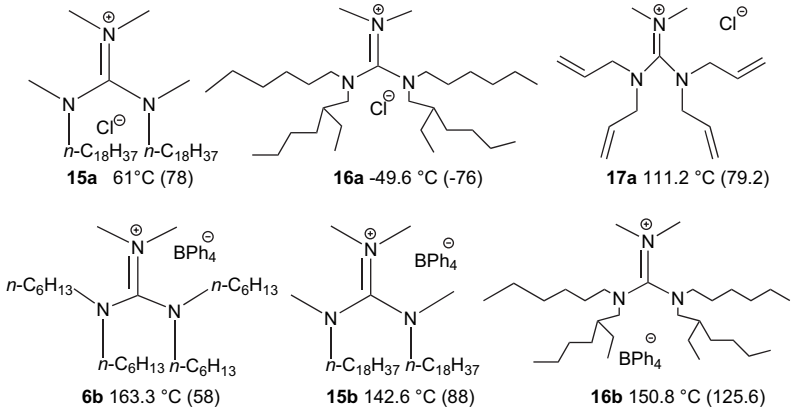

Figure 6. New synthesized guanidinium salts and their predicted and experimental (in parentheses) mp values.

extremely low T_g value, and **15a** and **17a** with intermediate values. (The structure of **15** activates a neuron in an unstable region, with very few occupied neurons, and their predictions are thus less reliable.) The lowest mp value in the data set for BPh_4^- salts is 127 °C, so predictions lower than this value are impossible. For the three synthesized salts **15b**, **16b**, and **6b**, the predicted mp values are within 36 °C of this limit. We were rewarded to observe that the experimental values for the new compounds are lower than those available in the training set.

4. Experimental

4.1. General remarks

All commercial reagents were used as supplied. Anhydrous dichloromethane and triethylamine were previously distilled from CaH_2 under argon atmosphere before their use in the preparation of guanidinium chlorides. ^1H and ^{13}C NMR spectra were recorded on a Bruker AMX400 spectrometer. Chemical shifts are reported downfield in parts per million from a tetramethylsilane reference. IR spectra were recorded on a Jasco FT/IR-430 spectrometer as thinly dispersed films. The samples for elemental analysis were performed by Laboratório de Análises at REQUIMTE, Departamento de Química Faculdade de Ciências e Tecnologia (Monte de Caparica, Portugal). The removal of the volatile compounds in each guanidinium salt was accomplished under high vacuum (6×10^{-6} mbar) at room temperature for 5 h. The calorimetric measurements were performed with a 2920MDSC system from TA Instruments Inc. Dry high-purity He gas with a flow rate of 30 $cm^3\,min^{-1}$ was purged through the sample. Cooling was accomplished with the liquid nitrogen-cooling accessory (LNCA), which provides automatic and continuous programmed sample cooling down to −150 °C. The baseline was calibrated by scanning the temperature domain of the experiments with an empty pan. The temperature calibration was performed taking the onset of the endothermic melting peak of several calibration standards: *n*-decane (T_m=243.75 K), *n*-octadecane (T_m=301.77 K), hexatriacontane (T_m=347.30 K), indium (T_m=430.61 K) and tin (T_m=506.03 K). The organic standards were high-purity Fluka products, while the metal standards were supplied by TA Instruments Inc. The enthalpy was calibrated with indium (melting enthalpy ΔH_m= 28.71 $J\,g^{-1}$).

4.2. General procedure for the preparation of guanidinium chlorides

All the guanidinium chlorides were prepared following a described procedure.[11] To a stirred mixture of *N*,*N*-dimethyl-phosgeniminium chloride (3.4 mmol) in anhydrous dichloromethane (10 mL) at 0 °C (ice bath) and under argon atmosphere was added dropwise a mixture of secondary amine (2.1 equiv) and anhydrous triethylamine (2.01 mL, 2.2 equiv) dissolved in anhydrous dichloromethane (14 mL). After 30 min at 0 °C, the reaction mixture was stirred for 4 h at room temperature. For each case, are specified the quantities and the workup procedure used.

4.2.1. N,N,N′,N″-Tetramethyl-N′,N″-dioctadecyl guanidinium chloride (*15a*)

Prepared following the general procedure starting with 0.546 g (3.36 mmol) of phosgeniminium chloride and 2 g (7.05 mmol) of methyloctadecyl-amine. As workup, the mixture was filtered, and the obtained solid washed with dichloromethane. The solvent of the combined organic mixture was removed under vacuum and water (12 mL) added to the mixture. The aqueous phase was extracted with diethyl ether (2×10 mL). The combined organic layers were washed with hydrochloric acid (1 M) aqueous solution. The solvent was removed under vacuum with the guanidinium salt obtained as white crystals (0.93 g, 40%). Mp 78 °C determined by DSC; IR ν_{max} (film): 3402, 3248, 2920, 2850, 1593, 1566, 1469, 1412, 1377, 1304, 1254, 1207, 1157, 1138, 1095, 1068, 903, 721 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ: 0.85 (6H, t, J=6.8 Hz), 1.22 (64H, m), 2.92–3.10 (16H, m); ^{13}C NMR (100 MHz, $CDCl_3$) δ: 14.03, 22.61, 26.76, 27.52, 29.28, 31.85, 38.62 and 38.86 (rotamers), 40.38 and 41.36 (rotamers), 52.80, 163.73. Analysis: calcd for $C_{41}H_{86}ClN_3\cdot1.5H_2O$: N 6.15, C 72.03, H 13.12. Found: N 6.24, C 72.06, H 13.45.

4.2.2. (2-Ethyl-hexyl)-hexyl-amine

Prepared in a two-step procedure. (2-Ethyl-hexyl)-amine (1 g, 7.8 mmol) was dissolved in anhydrous dichloromethane (20 mL) in the presence of 4 Å molecular sieves (4 g) and hexanal (1.1 g, 10.78 mmol) was added at room temperature. The mixture was stirred for 24 h over argon atmosphere. The mixture was filtered and the solvent of the collected organic phase was removed under vacuum. The mixture was dissolved in anhydrous ethyl ether (30 mL) and $LiAlH_4$ (0.61 g, 16 mmol) was added at 0 °C (ice bath). The mixture was stirred for 24 h at room temperature under argon atmosphere. As workup, the $LiAlH_4$ was neutralized with dropwise addition of water to the reaction mixture at 0 °C (ice bath) until a white precipitate appears consolidated in the bottom of the balloon. The mixture was filtered and the solvent of the collected organic phase was removed under vacuum. The resultant mixture was distilled under vacuum, with the secondary amine that distilled at 66 °C (0.3 mmHg) obtained as a transparent liquid (1 g, 60%). ^1H NMR (400 MHz, $CDCl_3$) δ: 0.83–0.89 (9H, m), 1.26–1.34 (14H, m), 1.40–1.49 (3H, m), 2.47–2.48 (2H, d, 6.2 Hz), 2.55–2.58 (2H, t, 7.28 Hz); ^{13}C NMR (100 MHz, $CDCl_3$) δ: 10.83, 14.02, 14.09, 22.63, 23.11, 24.52, 27.09, 29.01, 30.00, 31.80, 32.86, 39.34, 50.36, 53.29.

4.2.3. N,N-Dimethyl-N′,N″-dihexyl-N′,N″-bis(2-ethylhexyl) guanidinium chloride (*16a*)

Prepared following the general procedure using (2-ethylhexyl)-hexyl-amine (1.95 g, 9.1 mmol) and phosgeniminium chloride (0.705 g, 4.34 mmol). As workup, the mixture was filtered, and the obtained solid washed with dichloromethane. The solvent of the combined organic mixture was removed under vacuum with water (15 mL) added to the mixture. The

aqueous phase was extracted three times with diethyl ether (3×25 mL). The combined organic layers were washed with hydrochloric acid (1 M) aqueous solution. The solvent was removed under vacuum with the guanidinium salt obtained as yellow oil (1.419 g, 64%). IR ν_{max} (film): 3417, 2927, 2857, 1643, 1577, 1535, 1462, 1377, 1246, 1188, 1138, 1061, 879, 783, 729 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ: 0.80–0.90 (18H, m), 1.16–1.27 (28H, m), 1.41–1.53 (4H, m), 1.93–1.96 (2H, m), 2.76 (6H, s), 3.05 (4H, d, J=7.2 Hz), 3.17 (4H, m); ^{13}C NMR (100 MHz, $CDCl_3$) δ: 10.04 and 10.45 (rotamers), 13.83, 13.91 and 13.96 (rotamers), 22.41 and 22.78 (rotamers), 23.71 and 23.58 (rotamers), 26.41 and 26.55 (rotamers), 27.61, 28.14 and 28.53 (rotamers), 30.31 and 30.46 (rotamers), 31.14 and 31.45 (rotamers), 35.88, 38.70, 48.21, 48.68, 50.69, 51.05, 165.17. Analysis: calcd for $C_{31}H_{66}ClN_3 \cdot 1.25H_2O$: N 7.80, C 69.10, H 12.81. Found: N 7.52, C 68.96, H 13.35.

4.2.4. *N,N-Dimethyl-N′,N′,N″,N″-tetraallyl guanidinium chloride (**17a**)*

Prepared following the general procedure from diallylamine (2 g, 20.6 mmol) and phosgeniminium chloride (1.59 g, 9.8 mmol). As workup, the mixture was filtered, and the obtained solid washed with dichloromethane. The solvent of the combined organic mixture was removed under vacuum with an aqueous NaOH solution (2 M; 16 mL) added to the mixture. The aqueous phase was washed twice with diethyl ether (2×22 mL), and the pH adjusted to 7 by dropwise addition of hydrochloric acid (37%). The water was removed under vacuum. The guanidinium salt was obtained as a pale brown solid (1.8 g, 65%). Mp: 79.2 °C determined by DSC; IR ν_{max} (film): 3398, 3082, 2978, 2920, 1643, 1589, 1539, 1419, 1338, 1277, 1234, 1146, 1061, 999, 933, 876, 671 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ: 3.06 (6H, s), 3.78–3.96 (8H, m), 5.28–5.32 (8H, m), 5.76 (4H, m); ^{13}C NMR δ: 40.97, 52.64 and 52.89 (rotamers), 121.57, 131.11, 163.03. Analysis: calcd for $C_{15}H_{26}ClN_3 \cdot 1.7H_2O$: N 13.36, C 57.29, H 9.42. Found: N 13.48, C 56.91, H 9.63.

4.3. *General procedure for the substitution of chloride with tetraphenylborate*

The guanidinium chloride was dissolved in dichloromethane and $NaBPh_4$ (specific number of equivalents for each case) was added and stirred for 24 h. The resulting solid was collected by filtration and washed with dichloromethane; the combined organic layers were dried ($MgSO_4$), filtered, partially evaporated, and eluted with dichloromethane through a column of silica gel and activated carbon. The solvent was removed under vacuum to give the corresponding guanidinium tetraphenylborate salt. For each case, are specified the quantities and the workup procedure used.

4.3.1. *N,N,N′,N″-Tetramethyl-N′,N″-dioctadecyl guanidinium tetraphenylborate (**15b**)*

Prepared from *N,N,N′,N″*-tetramethyl-*N′,N″*-dioctadecyl guanidinium chloride (0.2 g, 0.29 mmol) in 15 mL of dichloromethane according to the general procedure for the substitution of chloride with tetraphenylborate (0.184 g, 0.585 mmol). As workup, the mixture was filtered and the obtained solid washed with dichloromethane. The combined organic phases were concentrated and eluted with dichloromethane through a silica gel and activated carbon column. The solvent was removed under vacuum. The guanidinium salt was obtained as a white solid (0.164 g, 59.7%). Mp 88 °C determined by DSC; IR ν_{max} (film): 3051, 2920, 2850, 1566, 1466, 1408, 1265, 1149, 1061, 1030, 895, 849, 741, 706, 606 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ: 0.89 (6H, t, J=6.92 Hz), 1.27 (64H, m), 2.48 (12H, m), 2.81 (4H, m), 6.89 (4H, t, J=6.92 Hz), 7.03 (8H, t, J=7.2 Hz), 7.42 (8H, m); ^{13}C NMR δ: 14.07, 22.65, 26.68, 27.47, 29.17, 29.32, 29.55, 29.67, 31.88, 37.68, 39.91 and 40.20 (rotamers), 52.62, 121.69, 125.51, 136.17, 163.34, 163.83, 164.32, 164.81. Analysis: calcd for $C_{65}H_{106}BN_3$: N 4.47, C 83.02, H 11.36. Found: N 4.53, C 82.73, H 11.63.

4.3.2. *N,N-Dimethyl-N′,N′,N″,N″-tetrahexyl guanidinium tetraphenylborate (**6b**)*

Prepared from *N,N*-dimethyl-*N′,N′,N″,N″*-tetrahexyl guanidinium chloride (1 g, 2.17 mmol) and sodium tetraphenylborate (1.39 g, 4.37 mmol) in 20 mL of dichloromethane, according to the general procedure for the substitution of chloride with tetraphenylborate. As workup, the mixture was filtered and the obtained solid washed with dichloromethane. The combined organic phases were concentrated and eluted with dichloromethane through a silica gel and activated carbon column. The solvent was removed under vacuum. The guanidinium salt was obtained as a brown solid (1 g, 62%). Mp 58 °C determined by DSC; IR ν_{max} (film): 3055, 2927, 2858, 1574, 1543, 1462, 1377, 1265, 1142, 1034, 845, 733, 706, 613 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ: 0.83–0.92 (12H, t, J=6.8 Hz), 1.19–1.32 (32H, m), 2.21 (6H, s), 2.63–2.97 (8H, m), 6.91 (4H, t, J=7.16 Hz), 7.06 (8H, t, J=7.2), 7.48 (8H, m); ^{13}C NMR δ: 13.88, 22.32 and 22.44 (rotamers), 26.62, 27.32, 27.83, 3.25, 31.33, 39.78, 49.29, 50.10, 121.71, 125.64, 136.12, 162.79, 163.40, 163.89, 164.38, 164.87. Analysis: calcd for $C_{51}H_{78}BN_3 \cdot 0.1H_2O$: N 5.63, C 82.13, H 10.57. Found: N 5.59, C 82.00, H 10.68.

4.3.3. *N,N-Dimethyl-N′,N″-dihexyl-N′,N″-bis(2-ethylhexyl) guanidinium tetraphenylborate (**16b**)*

Prepared from *N,N*-dimethyl-*N′,N″*-dihexyl-*N′,N″*-bis(2-ethylhexyl) guanidinium chloride (0.5 g, 0.97 mmol) in 11 mL of dichloromethane according to the general procedure for the substitution of chloride with tetraphenylborate (1.495 g, 2.9 mmol). As workup, the mixture was filtered and the obtained solid washed with dichloromethane. The combined organic phases were concentrated and eluted with dichloromethane through a silica gel and activated carbon column. The solvent was removed under vacuum. The guanidinium salt was obtained as a brown solid (0.63 g, 82%). Mp 125.6 °C determined by DSC; IR ν_{max} (film): 3516, 3157, 3055, 2929, 2858, 1579, 1537, 1502, 1464, 1404, 1379, 1269, 1244, 1184, 1146, 1066, 1032, 868, 847, 735, 706, 665, 611 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ: 0.75–0.84

(18H, m), 1.15–1.21 (28H, m), 1.33–1.42 (4H, m), 1.81–1.86 (2H, m), 2.71 (6H, s), 2.96 (4H, d, J=7.24 Hz), 3.03–3.25 (4H, m), 7.27–7.29 (4H, m), 7.35–7.39 (8H, m), 7.73–7.79 (8H, m). ^{13}C NMR δ: 10.04 and 10.46 (rotamers), 13.97, 22.43, 22.53, 23.56 and 23.71 (rotamers), 26.46, 26.52, 27.66, 28.15 and 28.56 (rotamers), 30.29, 30.48, 31.15 and 31.50 (rotamers), 35.92, 37.14, 38.58, 48.19, 48.51, 50.69 and 50.88 (rotamers), 115.46, 119.81, 121.60, 125.47, 127.09, 127.18, 127.40, 127.81, 128.26, 128.68, 129.33, 130.26, 130.90, 134.10, 134.70, 135.58, 136.23, 165.87 Analysis: calcd for $C_{55}H_{86}BN_3 \cdot 0.9H_2O$: N 5.20, C 80.87, H 10.80. Found: N 4.58, C 80.51, H 10.52.

5. Conclusion

Six new guanidinium salts were synthesized, and exhibited mp values in reasonable accordance with developed QSPR predictions. One of the salts is a room temperature ionic liquid.

QSPR models were based on CPG NNs that learned relationships between the structural profile of guanidinium cations (represented by 92 descriptors) and the mp of the corresponding salts incorporating one of four possible anions. The models were validated with an independent test set, fivefold cross-validation and y-randomization yielding acceptable predictions. The model is limited to salts of bromide, chloride, iodide, and tetraphenylborate, and to cations within the domain of the training set (which included guanidinium cations with the six binding sites of N atoms linked to hydrogen or to alkyl chains, unfunctionalized or including cyclic ethers, amines, or aromatic rings). Average errors of 20–30 °C are expected. It was not possible to find a small group of descriptors with a strong ability to predict the mp values.

The CPG NN approach allowed to process four series of anions simultaneously, producing four aligned maps that can be compared to assess similarities between anionic series in terms of their structure–mp relationships.

Acknowledgements

The authors thank Fundação para a Ciência e Tecnologia (POCI 2010) and FEDER for financial support (POCI/QUI/57735/2004, SFRH/BD/18354/2004 and SFRH/BPD/24969/2005). Molecular Networks GmbH (Erlangen, Germany) is acknowledged for access to the PETRA software package.

Supplementary data

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.tet.2007.12.021.

References and notes

1. (a) Welton, T. *Chem. Rev.* **1999**, *99*, 2071; (b) Dupont, J.; de Souza, R. F.; Suarez, P. A. Z. *Chem. Rev.* **2002**, *102*, 3667; (c) Sheldon, R. *Chem. Commun.* **2001**, 2399.
2. Earle, M. J.; Seddon, K. R. *Pure Appl. Chem.* **2000**, *72*, 1391.
3. Seddon, K. R. *J. Chem. Technol. Biotechnol.* **1997**, *68*, 351.
4. Earle, M. J.; Esperança, J. M. S. S.; Gilea, M. A.; Lopes, J. N. C.; Rebelo, L. P. N.; Magee, J. W.; Seddon, K. R.; Widegren, J. A. *Nature* **2006**, *439*, 831.
5. Buzzeo, M. C.; Evans, R. G.; Compton, R. G. *ChemPhysChem* **2004**, *5*, 1106.
6. Ye, C.; Shreeve, J. M. *J. Org Chem.* **2004**, *69*, 6511.
7. Bonhôte, P.; Dias, A. P.; Papageorgiou, N.; Kalyanasundaram, K.; Gratzel, M. *Inorg. Chem.* **1996**, *35*, 1168.
8. Wells, A. S.; Coombe, V. T. *Org. Process Res. Dev.* **2006**, *10*, 794.
9. Jastorff, B.; Mölter, K.; Behrend, P.; Bottin-Webber, U.; Filser, J.; Heimers, A.; Ondruschka, B.; Ranke, J.; Schaefer, M.; Schröder, H.; Stark, A.; Stepnowski, P.; Stock, F.; Störmann, R.; Stolte, S.; Welz-Biermann, U.; Ziegert, S.; Thöming, J. *Green Chem.* **2005**, *7*, 362.
10. Overman, L. E.; Rhee, Y. H. *J. Am. Chem. Soc.* **2005**, *127*, 15652.
11. Mateus, N. M. M.; Branco, L. C.; Lourenço, N. M. T.; Afonso, C. A. M. *Green Chem.* **2003**, *5*, 347.
12. Branco, L. C.; Gois, P. M. P.; Lourenço, N. M. T.; Kurteva, V. B.; Afonso, C. A. M. *Chem. Commun.* **2006**, 2371.
13. Xie, H.; Zhang, S.; Duan, H. *Tetrahedron Lett.* **2004**, *45*, 2013.
14. Li, S.; Lin, Y.; Xie, H.; Zhang, S.; Xu, J. *Org. Lett.* **2006**, *8*, 391.
15. Wang, P.; Zakeeruddin, S. M.; Grätzel, M.; Kantlehner, W.; Mezger, J.; Stoyanov, E. V.; Scherr, O. *Appl. Phys. A* **2004**, *79*, 73.
16. Recent examples are: (a) Krossing, I.; Slattery, J. M.; Daguenet, C.; Dyson, P. J.; Oleinikova, A.; Weingärtner, H. *J. Am. Chem. Soc.* **2006**, *128*, 13427; (b) Alavi, S.; Thompson, D. L. *J. Chem. Phys.* **2005**, *122*, 154704; (c) Katsyuba, S. A.; Zvereva, E. E.; Vidiš, A.; Dyson, P. J. *J. Phys. Chem. A* **2007**, *111*, 352.
17. (a) Carrera, G.; Aires-de-Sousa, J. *Green Chem.* **2005**, *7*, 20; (b) Katritzky, A. R.; Lomaka, A.; Petrukhin, R.; Jain, R.; Karelson, M.; Visser, A. E.; Rogers, R. D. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 71.
18. Eike, D. M.; Maginn, E. J. *Green Chem.* **2003**, *5*, 323.
19. Katritzky, A. R.; Jain, R.; Lomaka, A.; Petrukhin, R.; Karelson, M.; Visser, A. E.; Rogers, R. D. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 225.
20. Sun, N.; He, X.; Dong, K.; Zhang, X.; Lu, X.; He, H.; Zhang, S. *Fluid Phase Equilib.* **2006**, *246*, 137.
21. López-Martin, I.; Burello, E.; Davey, P. N.; Seddon, K. R.; Rothenberg, G. *ChemPhysChem* **2007**, *8*, 690.
22. (a) Trohalaki, S.; Pachter, R. *QSAR Comb. Sci.* **2005**, *24*, 485; (b) Trohalaki, S.; Pachter, R.; Drake, G. W.; Hawkins, T. *Energy Fuels* **2005**, *19*, 279.
23. Varnek, A.; Kireeva, N.; Tetko, I. V.; Baskin, I. I.; Solov'ev, V. P. *J. Chem. Inf. Model.* **2007**, *47*, 1111.
24. (a) Kantlehner, W.; Haug, E.; Mergen, W. W.; Speh, P.; Maier, T.; Kapassakalidis, J. J.; Bräuner, H. J.; Hagen, H. *Liebigs Ann. Chem.* **1984**, 108; (b) Linton, B.; Hamilton, A. D. *Tetrahedron* **1999**, *55*, 6027; (c) Brunelle, D. J.; Haitko, D. A. U.S. Patent US5,132,423, 1992; (d) Aldrich web site: http://www.sigmaaldrich.com; (e) Kantlehner, W.; Hagen, H. Deutsch Patent DE2,716,477, 1978; (f) Gleish, A.; Schidtchen, F. P. *Chem. Ber.* **1971**, *104*, 2158; (g) Isobe, T.; Fukuda, K.; Ishikawa, T. *J. Org Chem.* **2000**, *65*, 7770; (h) Beilstein database.
25. Chiappe, C.; Pieraccini, D. *J. Phys. Org. Chem.* **2005**, *18*, 275.
26. Details about DRAGON descriptors can be consulted on the web site: http://www.talete.mi.it/dragon_exp.htm.
27. PETRA can be tested on the website http://www2.chemie.uni-erlangen.de and is developed by Molecular Networks GmbH, http://www.mol-net.de
28. Zupan, J.; Gasteiger, J. *Neural Networks in Chemistry and Drug Design*, 2nd ed.; Wiley-VCH: Weinheim, 1999.
29. Aires-de-Sousa, J. *Chemom. Intell. Lab. Syst.* **2002**, *61*, 167.
30. The JATOON applets are available at http://www.dq.fct.unl.pt/staff/jas/jatoon.
31. Vracko, M. *Curr. Comput.-Aided Drug Des.* **2005**, *1*, 73.