Green Chemistry

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/greenchem

Towards designing environmentally safe ionic liquids: Influence of cation structure

Anita Sosnowska[1], Maciej Barycki[1], Monika Zaborowska[1], Anna Rybinska[1], and Tomasz Puzyn[1*]

[1] *Laboratory of Environmental Chemometrics, Institute for Environmental and Human Health Protection, Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk Poland*

***Corresponding author**: **Prof. Tomasz Puzyn**,
Laboratory of Environmental Chemometrics, Faculty of Chemistry, University of Gdańsk,
ul. Wita Stwosza 63, 80-308 Gdańsk, Poland.
Tel.: +48 58 523 52 48; E-mail address: t.puzyn@qsar.eu.org

Abstract

Ionic liquids (ILs) are considered to be an excellent substituent of organic solvents, commonly used nowadays. However, introducing new classes of compounds always attends with possibility of environmental threat and health issues they may induce.

This work was aimed at examining, which structural features of ILs are responsible for their toxicity. There are examples of structure-activity relationship models for ILs in the literature. However, in our approach, we have analysed this issue globally, for a wide range of ionic liquids and their toxicity measured by multiple toxicological tests (multiple endpoints).

We have collected the experimentally measured available literature data on toxicity of ILs for various organisms. Then, by employing Principle Component Analysis (PCA), we examined structural similarity of 375 different ionic liquids having six different types cations (namely: imidazolium, ammonium, phosphinium, pirydinium, pyrolidinium and sulfonium). For expressing the structural features of studied ILs we used Weighted Holistic Invariant Molecular (WHIM) descriptors, calculated for cations and anions separately. Geometry of each structure was optimized at the level of semi-empirical PM7 method. Toxicological response was thereafter analyzed in the space of the first and second principle components.

We pointed out that for most of the tested cases, there is a strong relationship between the variance in observed toxicity and the cations' descriptors. We also proved the anions' influence on ILs toxicity to be less meaningful. After repeating PCA using only the cations' descriptors we proved that the toxicity of ILs against selected targets, depends mostly on the size and branching of the cation. On this basis, we have proposed a Toxicity Ranking Index based on structural similarity of Cations (TRIC) for initial toxicity screening studies of ILs. It should be mentioned however, that the use of TRIC is limited to the prediction of toxicity endpoints used to its development. The use of TRIC as a preliminary toxicity indicator would provide a general view on ILs' toxicological potential and its predictions may provide

valuable conclusions, which, taken under consideration, may help to simplify the procedure of designing new greener and safer ionic liquids in future.

Keywords: toxicity, ionic liquids, Principle Component Analysis, TRIC

1. Introduction

Ionic liquids (ILs) form an exciting group of chemical substances that have a great potential to improve the development of organic chemistry and chemical technology.[1, 2] These compounds may be applied as solvents in various chemical processes including synthesis[1], catalysis[3] and biocatalysis[4], separation techniques, etc.[5] Moreover, they can be used as electrolytes in many products of novel technology.[6] Currently used ionic liquids are usually composed of large and asymmetrical cations and smaller inorganic or organic anions. ILs have a unique array of physiochemical properties that make them suitable in numerous applications in which conventional organic solvents are not sufficiently effective or not applicable.

Ionic liquids are often referred to be "green solvents", but the opinion has been nowadays questioned. According to the current studies, ionic liquids exhibit toxicity to bacteria,[7] green algae,[8] mammalian cells,[9] duckweed[10] and freshwater crustacean (*Daphnia magna).*[11] Moreover, ILs were considered as environmentally friendly because of their negligible vapor pressure. Although they cannot be spread in the air, many of them offer some level of water solubility, even in case of hydrophobic compounds.[12, 13] Therefore, ILs can be transported in the environment through water and soil. They may enter the aquatic environment by accidental spills or effluents. Physiochemical properties that make ILs of great industrial interest (i.e. thermal stability, high chemical persistence), may be also responsible for potential problems with degradation and/or persistence in environment.[10] Thus, it becomes necessary to examine the potential negative impact of ILs on human and other organisms health and the possibility of transport of these compounds in the environment, their interaction with the components of soils and sediments and their stability or susceptibility to degradation processes.

The unlimited possibility of modifying the structure of cation and anion of a given

ionic liquid delivers a countless number of potential derivatives having different toxicological activities and physicochemical properties. Experimental studies for such a large set of chemicals are very time-consuming and expensive, thus it is impossible to conduct a comprehensive evaluation of risk for those chemicals. In this case, computational methods such as multivariate explorative chemometric techniques could be used, as an alternative to expensive experiments. The methods may be employed to perform chemical screening and obtain a ranking, according to the studied phys/chem properties, reactivity, and/or activity.

There are already many contributions describing the relationship between the structure of ionic liquids and its toxicity using QSAR (Quantitative Structure-Activity Relationship) approach.[14-23] Each of them brings an evaluated tool for predicting ILs toxicity against one, specific target, by estimating the concentration of IL (so-called endpoint), that causes specific response of tested target (e.g., concentration of ILs causing death of 50% of tested bacteria *E. coli*). This proves that the ILs toxic potential depends strongly on its structure. Based on this, we decided to perform more extensive analysis, aimed at examining, whether the structure of ionic liquids determines their toxic behaviour globally and, if positive, which particular structural features are responsible for IL's toxicity in general, for all the studied endpoints. Such conclusions would be particularly useful for designing new, greener and safer ILs. Powered by the obtained results, we have proposed a simple Toxicity Ranking Index based on structural similarity of Cations (TRIC) to be used for further toxicity screening studies of ILs.

2. Methodology

2.1. Ionic liquids

This study was performed based on a set of 375 ionic liquids. The selection was based on availability of experimental data, obtained from a single experiment, or few measurements conducted under the same conditions (preferably, by the same experimentalists). Particular

ionic liquids differed by both cation and anion (Figure 1). The studied ILs had six different types cations (namely: imidazolium, ammonium, phosphinium, pirydinium, pyrolidinium and sulfonium), and 64 different anions. All studied 375 liquids are listed in Supplementary Material.

[Insert Figure 1 about here]

2.2 Experimental data

Toxicity data for ILs, experimentally measured for nine different endpoints (Table 1), were collected from the available sources (databases and publications). We ensured that all results within series of given ILs' toxic concentrations have been obtained with use of the same protocol, with the same experiment condition. This was important, because of the need of eliminating unnecessary additional variance in the dataset. Detailed data collected for ionic liquids used in this project are available in Supplementary Material. It's worth noting that the experimental data have not been available for about 79% of the studied ILs.

Table 1. Toxicological tests investigated in this study along with the endpoints and numbers of ionic liquids, for which data were collected

Toxicological test	Endpoint	Number of ILs	References
Acetylcholinesterase inhibition test	EC_{50}[A]	236	19, 24-26
Vibrio fischeri illumination inhibition test	EC_{50}[A]	57	26-29
Escherichia coli growth inhibition test	MIC[B]	89	30-33
Pseudokirchneriella subcapitata growth inhibition test	EC_{50}[A]	10	34
Scenedesmus vacuolatus reproduction inhibition test	EC_{50}[A]	38	26, 28, 29
Rat cell line IPC-81viability test	EC_{50}[A]	242	19, 24-27, 35, 36
Human cell line HeLa viability test	EC_{50}[A]	21	37
Human cell line MCF7 viability test	IC_{50}[C]	13	38
Daphnia Magna immobilization test	EC_{50}[A]	15	34, 39

[A] – concentration of a chemical inducing a specified response to 50% of the tested population, after a specified exposure duration, [B] – minimal concentration inhibiting the growth of a tested micro-organism population, after a specified exposure duration, [C] – concentration of a chemical causing an inhibition of a specified process in 50% of the tested population, after a specified exposure duration

2.3 Molecular descriptors

Towards the needs of further analysis, we had to translate the information on the structure of chemicals into numerical variables we could operate on. The most common approach to achieve that goal is to calculate a set of mathematical indices, called molecular descriptors, which express particular structural features. In order to describe molecular structure of the studied ILs in the numeric manner, we applied the Weighted Holistic Invariant Molecular (WHIM) descriptors.[40] They are designed to express the information about three-dimensional features of molecules, namely: molecular shape, size, symmetry and atom distribution within a molecule, with respect to invariant reference frames.

WHIM descriptors are calculated from a matrix of scores, that is created by performing Principal Component Analysis, PCA[41] (for general information on PCA please also refer to section 2.4) on the centered molecular coordinates. The covariance matrix used in PCA is weighted according to one of six different weighting schemes [40], following the equation (1):

$$s_{jk} = \frac{\sum_{i=1}^{n} w_i (q_{ij} - \bar{q}_j)(q_{ik} - \bar{q}_k)}{\sum_{i=1}^{n} w_i} \qquad (1)$$

where: s_{jk} is the value of covariance between coordinates *j* and *k*, n is the number of atoms, w_i is the chosen weight, q_{ij} is the *j*-th coordinate of *i*-th atom, q_{kj} is the *k*-th coordinate of *i*-th atom, $\bar{q}_j$ is the average of *j*-th coordinates and $\bar{q}_k$ is the average of *k*-th coordinates. The six weighting schemes used in WHIM approach include weighting by (i) molecular mass, (ii) van der Waals volume, (iii) Mulliken electronegativity, (iv) polarizability, (v) electrotopological indices by Kier and Hall and (v) the use of unweighted values ($w_i = 1$).

WHIM descriptors are consisted of two types of indices: (a) directional – calculated from the scores of each individual principle component, (b) and non-directional – derived

straight from the directional group, related to a holistic view of the molecule. There are 66 directional and 33 non-directional WHIM descriptors in total. For more details on WHIM calculation procedure one can refer to Gramatica.[40, 42]

In our work, molecular structures of ILs were built in MOLDEN[43] software. The structure of cation and anion was built and then optimized separately for each IL. Structure optimization (according to the minimal energy gradient) was performed in MOPAC[44] software at the level of semi-empirical PM7 method[45]. It has been already proved that semi-empirical calculations at the level of PM6 method are sufficient for similar types of studies.[46] PM7 leads for even more correct results, since it has been parameterized for a larger array of chemical species.[44] [25] The optimized molecular structures were thereafter imported to DRAGON software[47] for calculating WHIM descriptors. After excluding descriptors having constant and near-to-constant values (17 excluded descriptors in total), we received a matrix, where each of 375 ionic liquids is described by 164 WHIM descriptors (82 descriptors specific for cation, and 82 for anion).

2.4 Principle Component Analysis (PCA)

Principal Component Analysis (PCA) is a statistical tool, commonly used for reducing data complexity (data compression), creating new set of uncorrelated vectors from original data set, and analyzing similarities between the studied objects.[48] In this method, new variables, called principle components (PCs), are developed as linear combinations of the original ones. [41, 49, 50] The first PC explains the largest possible amount of the variance in the original data matrix, the second PC explains the largest possible variance unexplained by the first PC and so on. In effect, every object from the original matrix is described by a set of principal components instead of the original variables. The percentage of the total variance explained by PCs is decreasing with the increasing number of PC. Thus, the number of PCs is

usually much lower than the number of the original variables, since the total variance in the data is "compressed" in few first principal components. Moreover, all PCs are orthogonal (uncorrelated each other) by definition, which is very useful whenever possible similarities between the objects are of interest.[50]

In this work, PCA approach was adopted in order to group the studied ionic liquids based on their structural similarity and then to search for suggestions on possible relationships between ionic liquids structure and their toxicological response. We have presented the structures of ILs in the space of first and second principle components (score plot), in accordance with the demonstrative criterion.[51] We assumed that the objects (ionic liquids) located closely each other on the plot were structurally similar. The physical interpretation was assign to each PC based on the Malinowski's rule (only the contributions of descriptors having the normalized loadings higher than 0.7 were significant).[51]

In the next step, we transferred the collected toxicity data into a range scale, in which the ranges correspond to the standardized values of the endpoint (Figure 2). Finally, we assigned colors to the ranges and then colored markers representing particular ILs on the score plots derived from PCA. Structurally driven grouping of ILs having the same range of toxicity (the same color markers) would prove the dependence between structure and toxicity.

[Insert Figure 2 about here]

3. Results and discussion

3.1. Similarity analysis of ionic liquids based on anion and cation WHIM descriptors

In the first step, we explored the distribution of the selected 375 ionic liquids in the space of their structural descriptors (164 WHIM descriptors) with PCA. First two principal components (PC1 and PC2) explained together 50% (33% + 17%) of the total variance in the

data. As mentioned (in section 2.4), physical interpretation of a given PC can be assigned based on the contributions of the original descriptors to that PC (loadings values).

[Insert Figure 3 about here]

Along with the loading values (data presented in Supplementary Material 1) we concluded that PC1 represents the size, shape, accessibility and symmetry of the anion. Interestingly, one can recognize seven major groups of ionic liquids along PC1 (Figure 3A). Group **A** contains ILs with small halogen anions (Cl^-, Br^-, I^-); group **B** consists on ILs with hexafluorophophate (PF_6^-) and tetrafluoroborate (BF_4^-) anions; group **C** – ILs with acetate and 2-hydroxypropanoate anions; group **D** – sulfate ionic liquids; group **E** – contains ILs with bis(trifluoromethylosulfonyl)imide; group **F** – phosphate ionic liquids; and group **G** – ILs with bis[1,2-benzenediolato(2-)]borate anion. When looking at those groups, one can conclude that the size, shape and accessibility of the anion increases with the increasing values of PC1. For instance (see arrow **I**), in case of IL having the same cation, the size, shape and accessibility of the anion increases starting from a small halogen (**262**: trihexyltetradecylphosphonium chloride), through larger imide (**252**: trihexyltetradecylphosphonium bis(trifluoromethylsulfonyl)imide), up to the highly branched phosphinate anion (**257**: trihexyltetradecylphosphanium bis(2,4,4-trimethylpentyl)phosphinate).

On the contrary, PC2 is related to the size of the cation in the ionic liquid. For example, the arrows **II** and **III** describe the increase in size and linearity of cations, when anions are the same or very similar. In the first case (arrow **II**), there is a series of the ammonium ILs with halogens anions (**187**: tetraethylammonium bromide; **194**: benzyldecyldimethylammonium chloride; **220**: benzylhexadecyldimethylammonium

chloride). In the second case (arrow **III**), there are imidazolium ILs with 2-hydroxypropanoate anion (**122**: 1-methylimidazolium 2-hydroxypropanoate; **148**: 1-(pentoxymethyl)imidazolium 2-hydroxypropanoate; and **145**: 1-(dodecyloxymethyl)imidazolium 2-hydroxypropanoate).

In summary, in effect of performing PCA, we have extracted two principal components (PC1 and PC2) that described the most significant part of structural variance in the studied group of ILs. PC1 was related to the selected features of anions, whereas PC2 – of cations.

3.2. Toxicity of ionic liquids: cation or anion effect?

The next step of our research was to verify, whether the structural similarity determines the similarity in toxicity of investigated ILs. For this purpose, we colored data points from Figure 3A, for which the experimental data had been available, with using the color scheme that represents the standardized toxicity (see Figure 2). This was performed for each of the studied endpoint (see Table 1) separately. In this way, we tried identifying systematic patterns in the data that might suggest, which structural features are mainly responsible for the observed toxicity.

For example, in case of inhibition growth of bacteria *Eschericha coli* (Figure 3B) the lowest values of toxicity (lowest inhibition) are is observed for the low PC2 scores values (dark blue and blue circles). With the increasing score values of PC2 (size of the cation) one can observe the increase of toxicity (the color of markers is changing up to yellow, orange and red). Additionally, there are no significant trends along with the PC1 axis. Remembering that PC1 represents the size, shape, accessibility and symmetry of anions, whereas the PC2 is related to the size of cations, one can conclude that the structure of cation is the factor that

mainly determines the toxicity of studied ILs to *E. coli*. Interestingly, similar results have been obtained from analysis of the remaining endpoints (see Supplementary Materials).

We have compared our observations with other studies presented in the literature. There are some interesting contributions available that are focused on the effect of the alkyl side chain length in cations (e.g. methylimidazolium, pyridinium etc.) on toxicity of ILs to various biological systems. The side-chain effect has been noticed in studies of enzyme (acetylcholinesterase) inhibition,[24, 52, 53] antibacterial activity,[7, 27, 30, 54] toxicity to algae,[55] cytotoxicity to leukemia rat cell line IPC-81[25, 36, 38] and to human cell line HeLa,[56] as well as in a study of ILs toxicity to invertebrates[57]. All these contributions conclude that the increase of alkyl side chain length in a cation results in the increase in toxicity.

There are also a number of works describing the anion contribution into ILs' toxic behavior.[29, 58, 59] For example, Stolte et al.[35] tested toxicity of a series of ILs containing the same cation and various anions on leukemia rat cell line IPC-81. They concluded that, among 27 different anions, there were some anions that did not affect the toxicity, whereas some of them noticeably influenced the toxicity. Moreover, there are first, successful attempts to indicate anions' influence on the toxicity, using QSAR approach.[14, 23] To determine whether the influence of the cation or the anion is more important in the context of designing new, greener and safer ILs we conducted the series of the following analysis and comparisons.

First, we choose four different endpoints, for which experimentally measured toxicity data were available for sufficiently large groups of ILs with different combinations of cations and anions. These were: (A) EC_{50} for rat leukemia cell line viability, (B) EC_{50} for acetylcholinesterase inhibition, (C) EC_{50} for bacteria *Vibrio fischeri* illumination inhibition, and (D) MIC for bacteria *Escherichia coli* growth inhibition. Then, in case of each endpoint, we analyzed a series of observations for (i) ILs containing the same anion and different cations, and (ii) ILs containing the same cation and different anions.

246

247 *[Insert Figure 4 about here]*

248

249 ILs containing the same anion but different cations are covering wider range of
250 toxicity variability in case of each endpoint (Figure 4). This suggests that the structure of
251 cation always significantly alters the toxicity of ILs. We also observed that toxicity of ILs
252 containing the same cation and different anions does not vary as much as in the previous case.
253 It is best noticeable for EC_{50} for *Escherichia coli* growth inhibition (Figure 4B), EC_{50} for
254 acetylcholinesterase inhibition (Figure 4C) and EC_{50} for *Vibrio fischeri* illumination inhibition
255 (Figure 4D). Only in case of toxicity against rat cell line (EC_{50} for viability) (Figure 4A) one
256 can observe a significant influence of the anion. However, it never exceeds the influence of
257 cation.

258 Summarizing, the influence of anion on toxicity of ILs is not always negligible, but
259 usually smaller (or even much smaller) than the influence of cation. This observation is useful
260 from the practical viewpoint, for designing new, greener and safer ILs. Thus, our further
261 systematic chemometrical analysis, presented in this work, was conducted in the light of these
262 conclusions.

263

264 **3.3. Analysis of the influence of the cation's structure on toxicity of ILs**

265 Similarly to the investigation presented in Section 3.1, we performed PCA analysis of
266 375 ILs within the space defined by WHIM descriptors. But, this time, only a set of 82 cation
267 descriptors was taken into account. The first two principal components explain together
268 58.2% (35.2% + 23.0%) of the total variance in the structural descriptors of the cation. PC1
269 represents molecular size of the cation, whereas PC2 – molecular shape of the cation (for the

values of loadings that have original WHIM descriptors to each of the PCs, please refer to Supplementary Material 1).

When analyzing a score plot (Figure 5A), one can notice the existence of a simple relationship between the score values of PCs and the structure of ILs. Size of the cation increases with increasing values of PC1. For instance, arrow **A** represents the increase in size of selected ammonium ILs, namely: **211** (2-hydroxyethylammonium formate) < **194** (benzyldecyldimethylammonium chloride) < **220** (benzylhexadodecyldimethyloammonium chloride). Similarly, arrow **B** highlights the large change in the size of imidazolium ILs: **026** (1-methylimidazolium chloride) < **022** (1-decyl-3-methylimidazolium chloride) < **080** (3-methyl-1-octadecylimidazolium chloride), whereas arrow **C** highlights the large change in the size of the phosphonium ILs: **266** (buthyltrihexylphosphonium bromide) < **273** (trihexyloctylphosphonium chloride) < **270** (hexadecyltrihexylphosphonium chloride).

[Insert Figure 5 about here]

PC2 expresses molecular shape of the cation. With the increasing values of PC2 the increase of linearity of the cation is observed. For example, along with arrow **I** (Figure 5A)**,** the shape of cations is changing from **201** (tetrabutyloammonium bromide) that has a spheric shape of the cation, though **222** (2-hydroxyethylodimethylammonium acetate) to **209** (bis(2-methoxyethyl)ammonium sulfamate) that has a cation with nearly linear shape.

Also in this exercise, we colored data points, for which the experimental data had been available, with using the color scheme that represents the standardized toxicity (Figure 2). It should be noted that the first principal component tends to discriminate the toxicity of ionic liquids. ILs with higher (positive) score values of PC1 (see Figure 5B) have higher toxicity to *Escherichia coli* (yellow and orange markers), while the compounds with lower (negative)

PC1 scores are characterized by lower toxic activity (green and dark blue circles). Taking these results into account, we can define that the toxicity of ILs is determined mainly by the molecular size of the cation from which the liquid is composed. The influence of shape (linearity) of the cation is minor.

We have observed the same trends for almost all of the studied endpoints (EC_{50} for *Vibrio fischeri* illumination inhibition, EC_{50} for *Pseudokirchneriella subcapitata* growth inhibition, EC_{50} for *Scenedesmus vacuolatus* reproduction inhibition, EC_{50} for human cell line HeLa viability, IC_{50} for human cell line MCF7 viability and EC_{50} for *Daphnia magna* immobilization). The only exceptions were EC_{50} for acetylcholinesterase inhibition and EC_{50} for rat cell line IPC-81 viability. In the first case, no trends in toxicity within the space of cation descriptors have been observed. We believe, that the main reason of the observed lack of the relationship between the cation's structure and activity might be a different mode of action. Enzyme inhibition is based on a chemical's interaction with the enzyme's active center, whereas, in the other examined cases, the toxic responses were strongly dependent on the cellular uptake.[60] In the second test (toxicity to rat cell line IPC-81), the trend was mostly the same, but with few exemptions (ionic liquids having very small values of PC2). The exceptions observed in this test might be particular-case specific. This, however, should be confirmed by performing additional experiments.

3.4. Analysis of the influence of the anion's structure on toxicity of ILs

As noticed before, in some cases, the structure of anion may also have an important influence on toxicity of ILs. Thus, we conducted similar series of analysis, but restricted to anion's descriptors only. The analysis was performed to determine, whether there is a noticeable, more general trend in toxicity, related to specific structural features of ILs.

Our strategy was similar to the one we used for cations analysis. In the first step, we performed principle component analysis for all 375 ILs using 82 WHIM descriptors of anions structure. Coloring data points, for which the experimental data had been available, with using the color scheme that represents the standardized toxicity, indicated intensity of the toxicological response. This time, first two principle components explained 76.2% (64.1% + 12.7%) of the total variance.

However, after performing PCA and analyzing results for each of the nine endpoints, we did not recognize any pattern that might suggest the existence of a systematic relationship between the anion's structure and toxicity of ILs (data presented in Supplementary Material 1). This additionally proves that the structural features of cation in fact, mainly drive the toxicity of ILs. Thus, the application of anion descriptors only to analyze relationships between the structure and toxicity might be insufficient to observe any general trends.

Because of that, in the next step, we performed a series of analysis using only ILs consisting of the same cation, but differing by anions. Due to substantial lack of the available data in the literature, the analysis could be performed on the four following endpoints: (A) EC_{50} for rat leukemia cell line viability, (B) EC_{50} for acetylcholinesterase inhibition, (C) EC_{50} for bacteria *Vibrio fischeri* illumination inhibition, and (D) MIC for bacteria *Escherichia coli* growth inhibition (for detailed results please refer to Supplementary Material 1).

We discovered very clear relationship between structural features and EC_{50} to rat leukemia cell line viability in a group of ILs consisting of 1-hexyl-3-methylimidazolium cation and different anions (Figure 6). The toxicity is increasing with the increasing score values of the first principal component (PC1). Since PC1 represents such descriptors as size, shape, accessibility and symmetry of the anion (for the values of descriptors loadings to PC1 please refer to Supplementary Material 1), larger anions, having more symmetrical, better-

343 developed structure are responsible for the increase of the toxicity of studied ionic liquids.
344 This observations are in good accordance with Stolte at al.[35]
345
346 *[Insert Figure 6 about here]*
347
348 Similar trends might be observed in two other cases: (i) in a group of ILs with 1-butyl-3-
349 methylimidazolium and 1-ethyl-3-methylimidazolium cation and different anions (the same
350 endpoint: EC_{50} for rat leukemia cell line viability) and (ii) in a group of ILs with 1-butyl-3-
351 methylimidazolium cation and different anions (different endpoint: EC_{50} for bacteria *Vibrio*
352 *fischeri* illumination inhibition). Interestingly, the analysis performed on a group of ILs with
353 trihexyltetradecylphosphinium cation, exhibited the opposite trend in case of MIC for
354 *Escherichia coli* growth inhibition: this time the presence of smaller anions with simpler
355 structure (i.e., Cl^- and $BF4^-$) increased the toxicity of the studied ILs.
356 Analyses performed on other groups of ILs in combination with other endpoints (e.g.,
357 acetylcholinesterase inhibition of ILs containing 1-ethyl-3-methylimidazolium cation and
358 different anions) show no influence of the anion's structure on the studied toxicity. This leads
359 to the conclusion that the influence of anion's structure on toxicity of ILs is both group-
360 specific and/or endpoint-specific. It stands in the good accordance with the experimental
361 results obtained for ILs so far. For instance, Kumar et al.[38] indicated, that influence of the
362 anions is less significant with increasing alkyl chain length in the cation, and also noticed, that
363 fluoride containing anions alters the ILs' toxicity less when combined with pirydinium cation
364 than with pyrolidinium or piperidinium. Therefore, whether one intends to design a new
365 greener and safer ionic liquid, one should consider the structure of cation first. Then, the
366 designer should carefully study the influence of the anion's structure on the toxicity with use

of several diversified endpoints to finally select the best candidate to be synthesized. To this final verification, one may also use existing QSAR models, based on anions descriptors.

3.5 Toxicity Ranking Index based on structural similarity of Cations (TRIC)

Since the structure of cation more significantly affects the toxicity of ILs, we have developed a novel measure for ranking ILs according to their expected toxicity. The new measure (TRIC - Toxicity Ranking Index of Cations) is based on the score values of PC1 resulted from principal component analysis (PCA) performed on the pool of cation's WHIM descriptors. TRIC is determined mainly by the molecular size of the cation. The TRIC metric could be used as a preliminary, fast method for screening ILs and ranking them according to their toxicity. There are two important advantages of possible using TRIC. First, it starts only from the molecular structure of the ILs cation (any experiments are not required). Second, toxicity of new ionic liquids (if only they are similar to the studied data set) can be initially estimated based on the TRIC value. The use of TRIC as a preliminary toxicity indicator would greatly simplify the procedure of designing safe ionic liquids and would allow avoiding time-consuming process of extensive experimental research (potentially toxic ILs could be eliminated at the stage of *in silico* design, even before the synthesis). However, we expect its best performance for ionic liquids, having the structure of cation similar to those cations analyzed in this study. The applicability of TRIC for estimating toxicity of less common ILs, with unusual cations (e.g., magnetic anions: $[FeCl_4]$, $[GdCl_6]$, $[CoCl_4]$ or $[MnCl_4]$)[61], might be less reliable.

It should be mentioned that PCA has been already successfully used for screening and ranking of other chemicals, in the context of risk assessment.[62,63,64,65,66]. In this study we have demonstrated also the usefulness of PCA for developing toxicity ranking of ionic liquids.

4. Conclusions

Although the toxicity mechanism is a very complicated and individual matter in case of every organism/cell, we came out with some general observations related to toxicity of ionic liquids. We confirmed that the toxicity of ionic liquids is, in most cases, strongly depends on the structure of its cation. Moreover, the size of the cation in the ionic liquid structure is much more important that its shape (linearity).

The influence of anion's structure on toxicity of ILs is both group-specific and endpoint-specific. This means that the use of simpler and smaller structure of the anion may result in either increasing or decreasing toxicity of IL, dependently on particular endpoint, or even might not have any influence on that.

The conclusions presented above have a very practical meaning for designing new, greener and safer ionic liquids. The designers should select the least toxic cation first, and then they should consider the influence of the anion's structure on the toxicity in the group based on the selected (the same) cation. The final choice should be based on the results of several diversified toxicity tests performed on the family of ILs based on the selected cation, or, whenever possible, on the predictions obtained from appropriate QSAR models, developed for the group of ILs based on the same cation, but differing by the anion.

We propose Toxicity Ranking Index of Cations (TRIC) – a simple metric of cation-related toxicity of ILs. The use of TRIC provides a time-saving and convenient way of performing the first step (cation selection) in synthesis of ILs that are safe-by-design, without necessity of conducting expensive experiments. However, one should be aware, that our results are based on a restricted group of tested targets and predictions obtained by using TRIC have mostly informative, overall character. Trends we captured throughout our analysis, are valuable as general marks, for recognizing probably least toxic ILs, but for the

full assessment of ILs' toxicity, additional, more complex toxicological tests would be needed.

We believe that our results and the proposed screening method with use of TRIC would contribute to future, accurate quantitative prediction of ILs' toxicity, and therefore designing and/or selecting ionic liquids that are safe for human and the environment.

Acknowledgements

This material is based on research sponsored by the Polish National Science Center (project CRAB, grant no. UMO-2012/05/E/NZ7/01148).

References

1. P. Wasserscheid and T. Welton, *Ionic Liquids in Synthesis*, Wiley-VCH, Weinheim, Germany, 2002.
2. Seddon K., *Green Chemm*, 2002, G35-G36.
3. T. Welton, *Chem Rev*, 1999, 99, 2071-2084.
4. R. A. L. Sheldon, R.M.; Sorgedrager, M.J.; van Rantwijk, F.; Seddon,K.R.;, *Green Chem*, 2002, 147-151.
5. A. Berthod, M. J. Ruiz-Angel and S. Carda-Broch, *J Chromatogr A*, 2008, 1184, 6-18.
6. F. Endres, *Chemphyschem*, 2002, 3, 145-154.
7. J. Pernak, K. Sobaszkiewicz and I. Mirska, *Green Chemistry*, 2003, 5, 52-56.
8. A. Latala, P. Stepnowski, M. Nedzi and W. Mrozik, *Aquat Toxicol*, 2005, 73, 91-98.
9. R. F. M. Frade, A. Matias, L. C. Branco, C. A. M. Afonso and C. M. M. Duarte, *Green Chemistry*, 2007, 9, 873.
10. B. Jastorff, K. Mölter, P. Behrend, U. Bottin-Weber, J. Filser, A. Heimers, B. Ondruschka, J. Ranke, M. Schaefer, H. Schröder, A. Stark, P. Stepnowski, F. Stock, R. Störmann, S. Stolte, U. Welz-Biermann, S. Ziegert and J. Thöming, *Green Chemistry*, 2005, 7, 362.
11. T. P. Pham, C. W. Cho and Y. S. Yun, *Water Res*, 2010, 44, 352-372.
12. J. L. Anthony, E. J. Maginn and J. F. Brennecke, *Journal Phys Chem B*, 2001, 10942-10949.
13. J. R. McFarlane, W.B.; Luo, H.; Hunt, R.D.; Depaoli, D.W.; Ren, R.X., *Sep. Sci. Technol* 2005 1245-1265.
14. M. H. Fatemi and P. Izadiyan, *Chemosphere*, 2011, 84, 553-563.
15. M. Ismail Hossain, B. B. Samir, M. El-Harbawi, A. N. Masri, M. I. Abdul Mutalib, G. Hefter and C. Y. Yin, *Chemosphere*, 2011, 85, 990-994.
16. R. N. Das and K. Roy, *Chemosphere*, 2014, 104, 170-176.

17. K. Roy, R. N. Das and P. L. A. Popelier, *Chemosphere*, 2014, 112, 120-127.
18. S. Bruzzone, C. Chiappe, S. E. Focardi, C. Pretti and M. Renzi, *Chemical Engineering Journal*, 2011, 175, 17-23.
19. J. S. Torrecilla, J. Garcia, E. Rojo and F. Rodriguez, *Journal of hazardous materials*, 2009, 164, 182-194.
20. K. Roy and R. N. Das, *Journal of hazardous materials*, 2013, 254-255, 166-178.
21. F. Yan, S. Xia, Q. Wang and P. Ma, *Industrial & Engineering Chemistry Research*, 2012, 51, 13897-13901.
22. F. Yan, S. Xia, Q. Wang and P. Ma, *Journal of Chemical & Engineering Data*, 2012, 57, 2252-2257.
23. Á. Irabiena, A. Gareaa and P. Luis, *19th European Symposium on Computer Aided Process Engineering - Elsevier*, 2009, 63-67.
24. J. Ranke, ;, S. Stolte, R. Stormann, J. Arning and B. Jastorff, *Chem. Rev.*, 2007, 2183-2006.
25. M. Stasiewicz, E. Mulkiewicz, R. Tomczak-Wandzel, J. Kumirska, E. M. Siedlecka, M. Golebiowski, J. Gajdus, M. Czerwicka and P. Stepnowski, *Ecotoxicol Environ Saf*, 2008, 71, 157-165.
26. Centre for Environmental Research and Sustainable Technology (UFT) (www.Il-eco.uft) (Access December 2013).
27. J. Ranke, K. Mölter, F. Stock, U. Bottin-Weber, J. Poczobutt, J. Hoffmann, B. Ondruschka, J. Filser and B. Jastorff, *Ecotoxicology and Environmental Safety*, 2004, 58, 396-404.
28. S. Stolte, M. Matzke, J. Arning, A. Böschen, W.-R. Pitner, U. Welz-Biermann, B. Jastorff and J. Ranke, *Green Chemistry*, 2007, 9, 1170.
29. M. Matzke, S. Stolte, K. Thiele, T. Juffernholz, J. Arning, J. Ranke, U. Welz-Biermann and B. Jastorff, *Green Chemistry*, 2007, 9, 1198.
30. A. Cieniecka-Rosłonkiewicz, J. Pernak, J. Kubis-Feder, A. Ramani, A. J. Robertson and K. R. Seddon, *Green Chem*, 2005, 855-862.
31. X.-D. Hou, Q.-P. Liu, T. J. Smith, N. Li and M.-H. Zong, *PLoS ONE*, 2013, 8, doi:10.1371/journal.pone.0059145.
32. J. Pernak, J. Kalewska, H. Ksycińska and J. Cybulski, *Eur. J. Med. Chem.*, 2001, 36, 899-907.
33. J. Pernak, I. Goc and I. Mirska, *Green Chem*, 2004, 6, 323.
34. C. Pretti, C. Chiappe, I. Baldetti, S. Brunini, G. Monni and L. Intorre, *Ecotoxicol Environ Saf*, 2009, 72, 1170-1176.
35. S. Stolte, J. r. Arning, U. Bottin-Weber, M. Matzke, F. Stock, K. Thiele, M. Uerdingen, U. Welz-Biermann, B. Jastorff and J. Ranke, *Green Chemistry*, 2006, 8, 621.
36. S. Stolte, J. r. Arning, U. Bottin-Weber, A. Muller, W.-R. Pitner, U. Welz-Biermann, B. Jastorff and J. Ranke, *Green Chemistry*, 2007, 9, 760.
37. X. Wang, C. A. Ohlin, Q. Lu, Z. Fei, J. Hu and P. J. Dyson, *Green Chemistry*, 2007, 9, 1191.
38. R. A. Kumar, N. Papaiconomou, J. M. Lee, J. Salminen, D. S. Clark and J. M. Prausnitz, *Environmental toxicology*, 2009, 24, 388-395.
39. A. S. Wells and V. T. Coombe, *Organic Process Research & Development*, 2006, 10, 794-798.
40. R. Todeschini and P. Gramatica, *Quantitative Structure-activity Relationships*, 1997, 16, 113-119.
41. I. T. Jolliffe, *Principal Component Analysis, Second Edition*, Springer, New York, USA, 2002.

42. R. Todeschini and P. Gramatica, *SAR and QSAR in Environmental Research*, 1997, 7, 89-115.
43. G. Schaftenaar and J. H. Noordik, *J. Comput.-Aided Mol. Design*, 2000, 14, 123-134.
44. J. J. P. Stewart, MOPAC, Stewart Computional Chemistry, Colorado Springs, CO, USA, 2012.
45. J. J. P. Stewart, *Journal of molecular modeling*, 2013, 19, 1-32.
46. T. Puzyn, N. Suzuki, M. Haranczyk and J. Rak, *Journal of chemical information and modeling*, 2008, 48, 1174-1180.
47. DRAGON, (Software for Molecular Descriptor Calculation), Talete srl, 6.0.30 edn., 2013.
48. D. Zuba and A. Parczewski, *Chemometria w analityce*, Instytut Ekspertyz Sadowych, Krakow, Poland, 2008.
49. P. Praus, *Central European Journal of Chemistry*, 2005, 3, 731-741.
50. H. Abdi and L. J. Williams, *Wiley Interdisciplinary Reviews: Computational Statistics*, 2010, **2**, 433-459.
51. J. Mazerski, *Podstawy Chemometrii*, Wydawnictwo Politechniki Gdańskiej, Gdańsk, 2000.
52. F. Stock, J. Hoffmann, J. Ranke, R. Stormann, B. Ondruschka and B. Jastorff, *Green Chemistry*, 2004, 6, 286.
53. J. Arning, S. Stolte, A. Boschen, F. Stock, W. R. Pitner, U. Welz- Biermann, B. Jastorff and J. Ranke, *Green Chem*, 2008, 47-58.
54. K. M. Docherty and J. C. F. Kulpa, *Green Chemistry*, 2005, 7, 185.
55. C. W. Cho, Y. C. Jeon, T. P. Pham, K. Vijayaraghavan and Y. S. Yun, *Ecotoxicol Environ Saf*, 2008, 71, 166-171.
56. P. Stepnowski, A. C. Składanowski, A. Ludwiczak and E. Łaczyńska, *Human & Experimental Toxicology*, 2004, 23, 513-517.
57. M. T. Garcia, N. Gathergood and P. J. Scammells, *Green Chem*, 2005, 9-14.
58. K. S. Egorova and V. P. Ananikov, *ChemSusChem*, 2014, 7, 336-360.
59. S. Steudte, P. Stepnowski, C. W. Cho, J. Thoming and S. Stolte, *Chem Commun (Camb)*, 2012, 48, 9382-9384.
60. D. Zhao, Y. Liao and Z. Zhang, *CLEAN – Soil, Air, Water*, 2007, 35, 42-48.
61. R. F. Frade, S. Simeonov, A. A. Rosatella, F. Siopa and C. A. Afonso, *Chemosphere*, 2013, 92, 100-105.
62. M. Vighi, P. Gramatica, F. Consolaro and R. Todeschini, *Ecotoxicol Environ Saf*, 2001, 49, 206-220.
63. E. Papa, F. Battaini and P. Gramatica, *Chemosphere*, 2005, 58, 559-570.
64. P. Gramatica and E. Papa, *QSAR & Combinatorial Science* 2005, 953-960.
65. P. Gramatica, P. Pilutti and E. Papa, *SAR QSAR Environ Res*, 2002, 13, 743-753.
66. P. Gramatica, E. Papa and S. Pozzi, *Fresenius Environ Bull* 2004, 1204-1209.

Figure captions

Figure 1. Schematic representation of cations' structures. **A** – imidazolium, **B** – ammonium, **C** – phosphinium, **D** – pirydinium, **E** – pyrolidinium and **F** – sulfonium. **R1-4** – substituents of different types.

Figure 2. Color scale used for representing the standardized toxicity of ionic liquids.

Figure 3. Score plots from the two principal component analysis performed for 375 ionic liquids. Panel A: Structural similarities of ILs in the space of cation and anion WHIM descriptors. Panel B: Similarity analysis in relation to the toxicity to *Escherichia coli*. Color codes as in Figure 2: Red circles correspond to the highest values of the toxicity, the dark blue circles represent ionic liquids with the lowest toxicity. Green and yellow circles represent moderate toxicity.

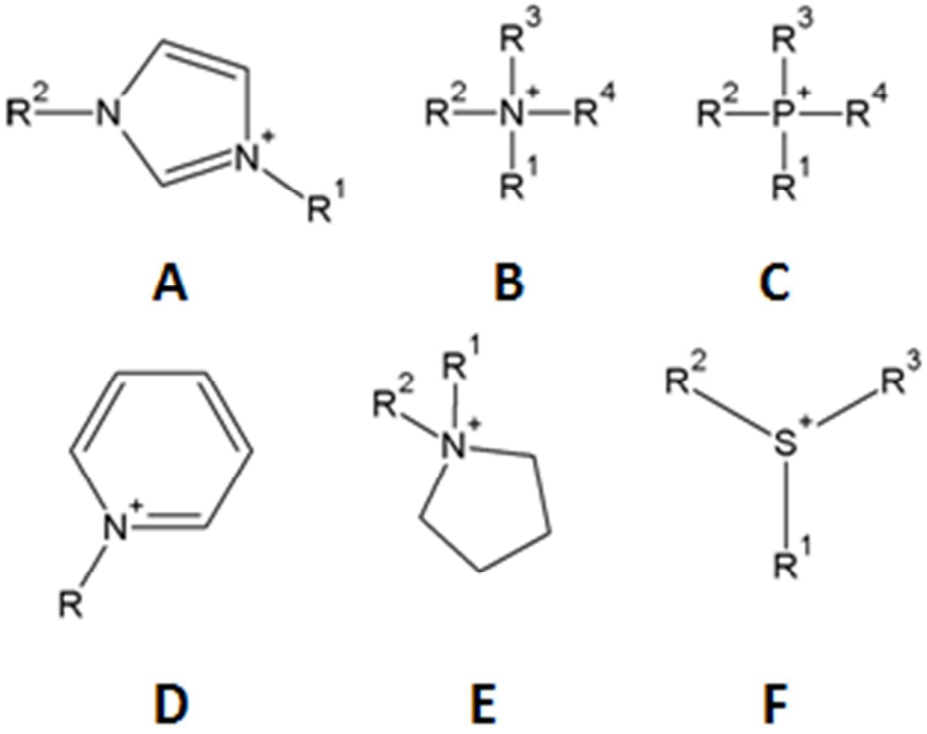
Figure 4. Toxicity ranges of ionic liquids families based on the same anions and the same cations. Panel A: toxicity to rat leukemia cell line. Panel B: acetylcholinesterase inhibition. Panel C: toxicity to bacteria *Vibrio fischeri*. Panel D: toxicity to bacteria *Escherichia coli.*

Figure 5. Score plots from the two principal component analysis performed for 375 ionic liquids. Panel A: Structural similarities of ILs in the space of cation WHIM descriptors. Panel B: Similarity analysis in relation to the toxicity to *Escherichia coli*. Color codes as in Figure 2: Red circles correspond to the highest values of the toxicity, the dark blue circles represent ionic liquids with the lowest toxicity. Green and yellow circles represent moderate toxicity.

Figure 6. Score plots from the two principal component analysis performed for 375 ionic liquids. Panel A: Structural similarities of ILs in the space of anion WHIM descriptors for the family of ILs based on 1-hexyl-3-methylimidazolium cation. Panel B: Similarity analysis in relation to toxicity to rat leukemia cell line. Color codes as in Figure 2: Red circles correspond to the highest values of the toxicity, the dark blue circles represent ionic liquids with the lowest toxicity. Green and yellow circles represent moderate toxicity.

544

Figure 1.



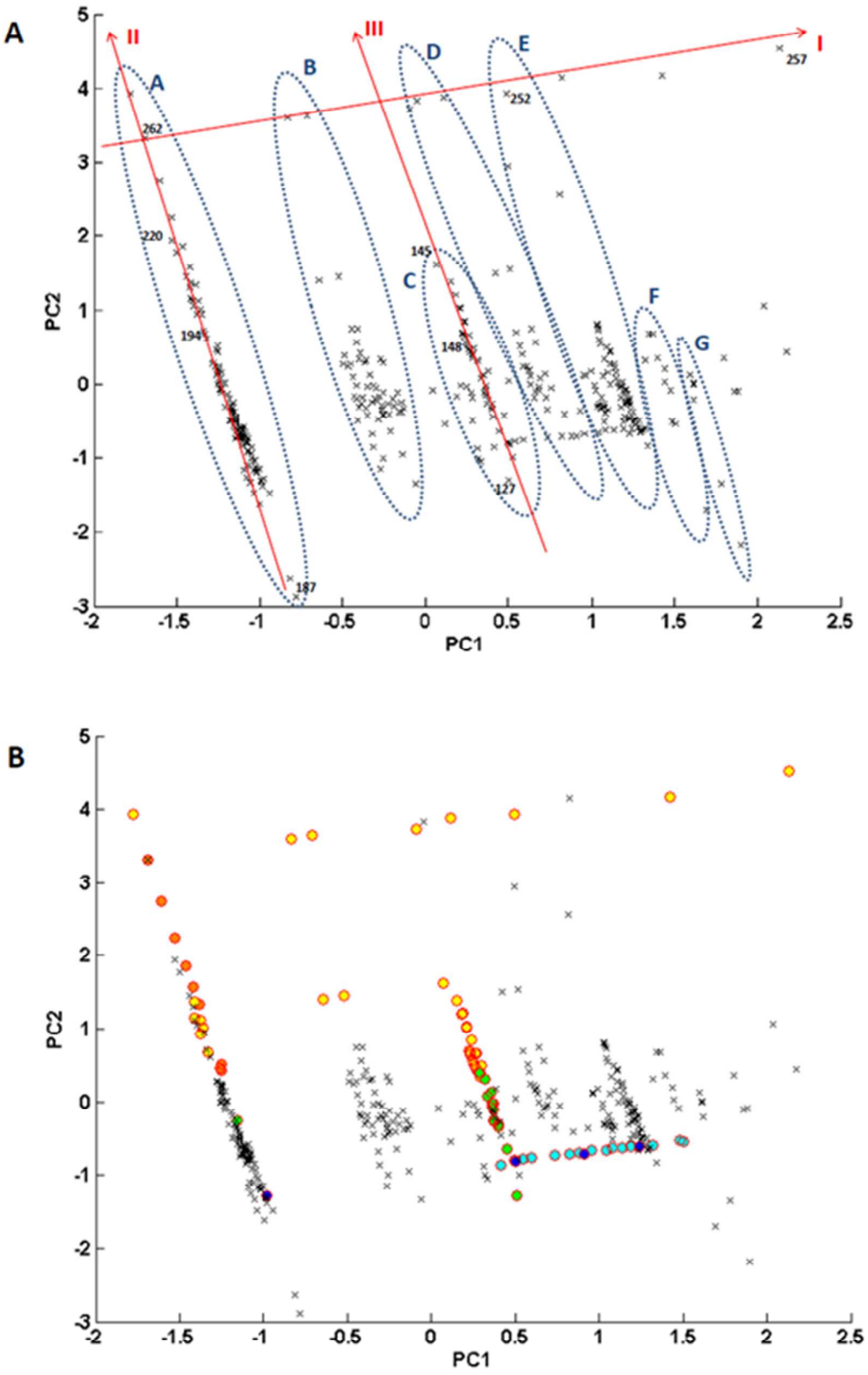
545

546

Figure 2.

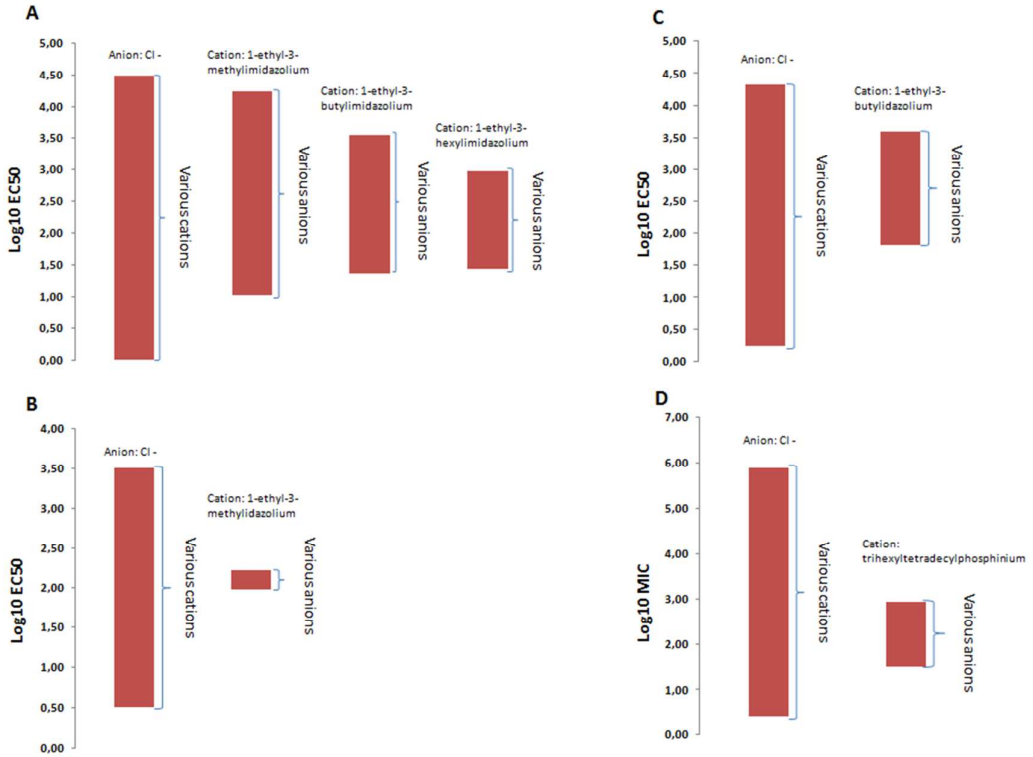
547

Figure 3.



548

Figure 4.



549

Figure 5.

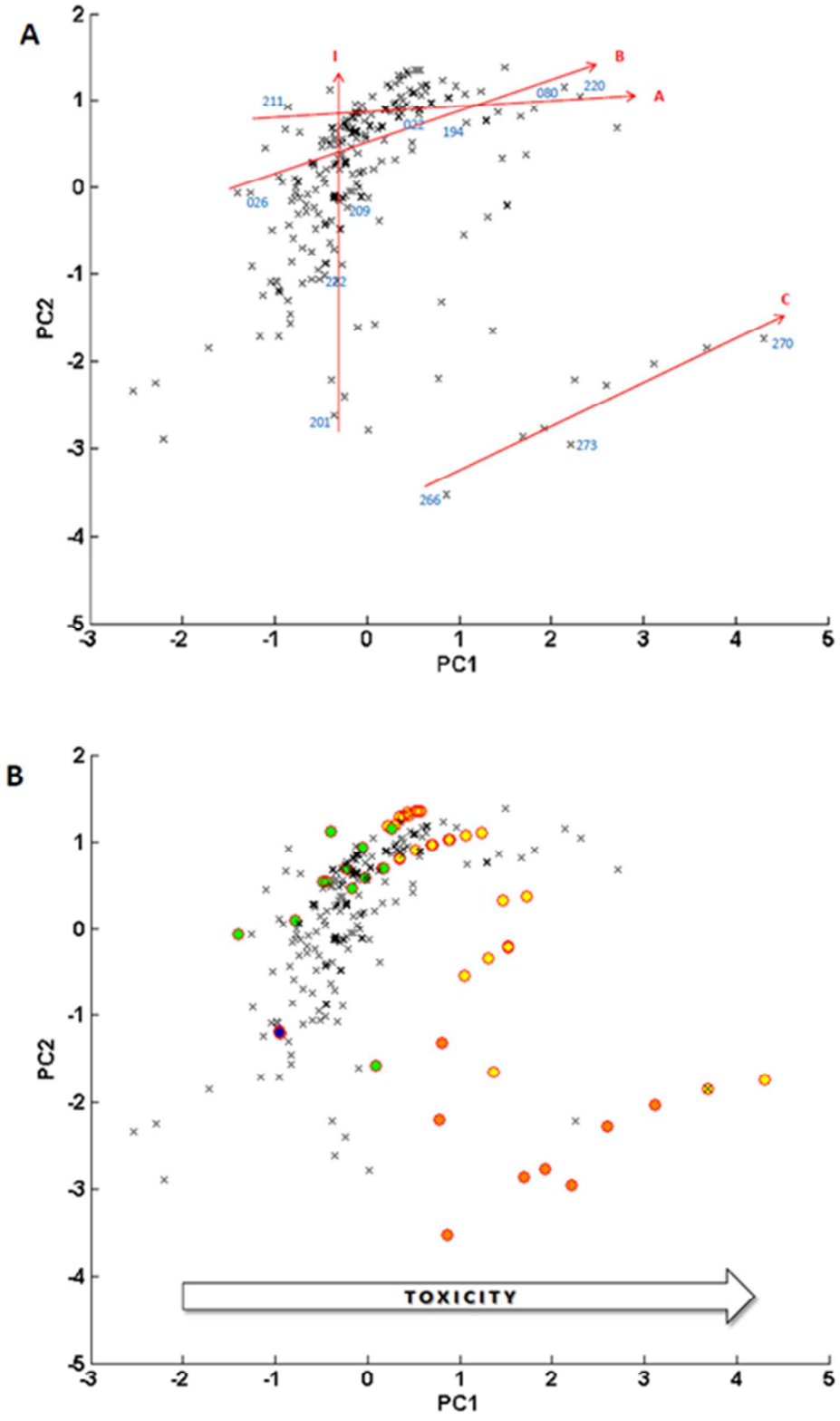
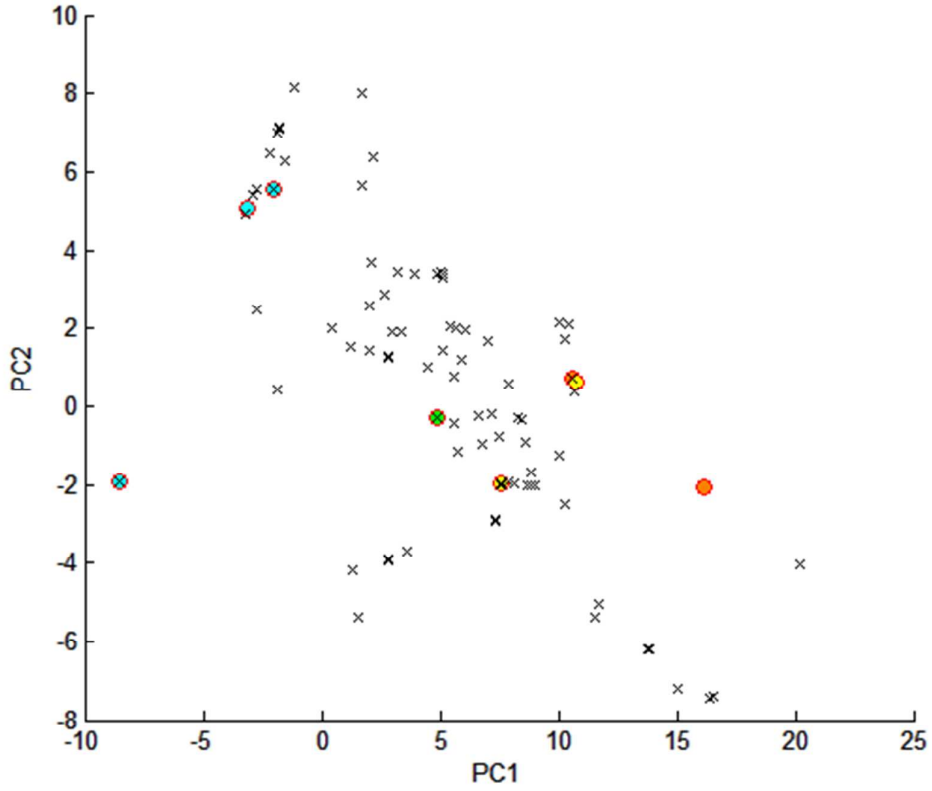
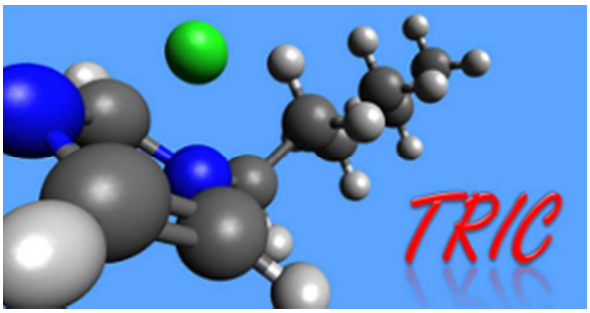


Figure 6.



Novelty:

Toxicity Ranking Index for ionic liquids based on structural similarity of Cations (TRIC), based on data from different toxicological tests.



Colour graphic to be used in a table of contents.