

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/231389741>

# Development of an a Priori Ionic Liquid Design Tool. 2. Ionic Liquid Selection through the Prediction of COSMO-RS Molecular Descriptor by Inverse Neural Network

ARTICLE *in* INDUSTRIAL & ENGINEERING CHEMISTRY RESEARCH · FEBRUARY 2009

Impact Factor: 2.59 · DOI: 10.1021/ie8009507

---

CITATIONS

42

---

READS

29

4 AUTHORS, INCLUDING:



[Jose S. Torrecilla](#)

Complutense University of Madrid

127 PUBLICATIONS 1,823 CITATIONS

SEE PROFILE

Article

**Development of an a Priori Ionic Liquid Design Tool.  
2. Ionic Liquid Selection through the Prediction of  
COSMO-RS Molecular Descriptor by Inverse Neural Network**

Jose# Palomar, Jose# S. Torrecilla, Vi#ctor R. Ferro, and Francisco Rodri#guez

*Ind. Eng. Chem. Res.*, **2009**, 48 (4), 2257-2265 • DOI: 10.1021/ie8009507 • Publication Date (Web): 22 January 2009

Downloaded from <http://pubs.acs.org> on February 13, 2009

**More About This Article**

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

[View the Full Text HTML](#)



**ACS Publications**  
High quality. High impact.

## Development of an a Priori Ionic Liquid Design Tool. 2. Ionic Liquid Selection through the Prediction of COSMO-RS Molecular Descriptor by Inverse Neural Network

José Palomar,<sup>\*,†</sup> José S. Torrecilla,<sup>‡</sup> Víctor R. Ferro,<sup>†</sup> and Francisco Rodríguez<sup>‡</sup>

Sección de Ingeniería Química (Dpt. Química Física Aplicada), Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain, and Departamento de Ingeniería Química, Universidad Complutense de Madrid, 28040 Madrid, Spain

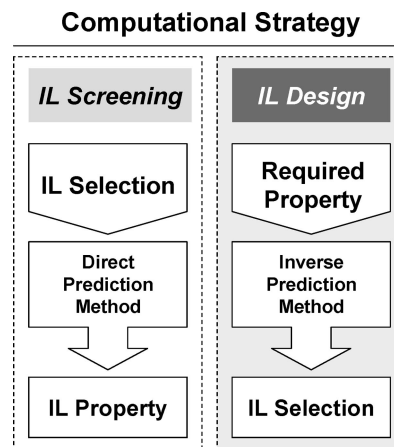
In this work, the a priori computational tool for screening ILs, developed in previous part 1, is extended to the simultaneous prediction of a set of IL properties for 45 imidazolium-based ILs. In addition, current part 2 reports the development of a more useful design strategy, which introduces the target IL properties as input, resulting in the selections of counterions as output, that is directly designing ILs on the computer. For this purpose, inverse neural networks are used to estimate the  $S_{\sigma\text{-profile}}$  molecular descriptor of a potential IL solvent by the specification of its required properties, following a reverse quantitative structure–property relationship scheme. Subsequently, a statistical tool based on Euclidean distances is developed to select an adequate set of anion+cation combinations that fulfill the estimated  $S_{\sigma\text{-profile}}$  values, to obtain, in this case, the tailor-made ILs. Finally, the proposed computational tool for designing ILs is applied in liquid–liquid extraction of a system model (toluene/*n*-heptane). In view of the inherent modular nature of ILs, the proposed methodology is here used in the formulation of IL mixtures to enhance the performance of extractive solvents in the aromatic/aliphatic separation.

### 1. Introduction

The room temperature ionic liquids (ILs) have gained popularity in recent years due to their unique properties: negligible vapor pressure, high thermal and chemical stabilities, wide liquid state range, and high solvent capacity.<sup>1,2</sup> Thus, the application of these novel solvents to organic synthesis, catalysis, electrochemistry, biocatalysis, material science, and separation processes is being extensively investigated.<sup>3–5</sup> A main advantage of these new solvents is that the cation and anion can be selected among a huge diversity to obtain an appropriate IL for a specific purpose. Therefore, ILs are often referred to as *designer solvents*. However, there are more than 1000 different ILs commercially available, and it is considered that there are  $>10^4$  counterion combinations possible. For the development of ionic liquids, the application of predictive models to estimate their thermo-physical properties is of great interest, principally for the proper selection of the component ions.<sup>6</sup> Indeed, theoretical applications to predict IL properties have highlighted a growing trend in the past few years.<sup>7</sup> It is noteworthy, however, that in general those computational studies have been performed on a few common ILs to i) validate the theoretical method by comparison between calculated and experimental properties (*i.e.* *what IL property?*), or ii) understand the properties of ILs with respect to their underlying molecular structures (*i.e.* *why IL property?*). Concerning the design of tailor-made ILs, most of available theoretical studies were centered on screening among a wide variety of cations ( $C^+$ ) and anions ( $A^-$ ) to obtain a counterion combination with the required properties,<sup>8–12</sup> following the computational approach named as “*IL Screening*” in Scheme 1. However, at the moment, it has not yet proven possible to predict directly on the computer which cation ( $C^+$ ) and anion ( $A^-$ ) combinations will lead to a given set of properties (*i.e.*

*which IL selection?*). In this work, we present an a priori mathematical tool to directly design ILs on the computer, starting from the required properties. This method is identified as “*IL Design*” in Scheme 1. To our knowledge, only one attempt has been reported to design ILs with required properties (conductivity and viscosity) using an iterative quantitative structure–property relationship (QSPR) model coupled with descriptors of group contribution type.<sup>13</sup> In our proposal, neural networks (NNs) are used to establish relationship between quantum-chemical COSMO-RS<sup>14</sup> electronic information of IL structures and their solvent properties, but with the aim of obtaining possible  $C^+A^-$  combinations with target characteristics, following the innovative strategy schemed as *Level II* in Figure 1. Clearly, adequate information must be initially given for performing such analysis reliably: a) IL property values to neural network (NN) design and optimization; and b) molecular descriptors characteristic of IL systems to be modeled. Therefore, in previous part 1 of this work we proposed a new  $S_{\sigma\text{-profile}}$  molecular descriptor<sup>15</sup> based on COSMO-RS methodology,

Scheme 1



\* To whom correspondence should be addressed. Phone: 34 91 4976938. Fax number: 34 91 4973516. E-mail: pepe.palomar@uam.es.

<sup>†</sup> Universidad Autónoma de Madrid.

<sup>‡</sup> Universidad Complutense de Madrid.

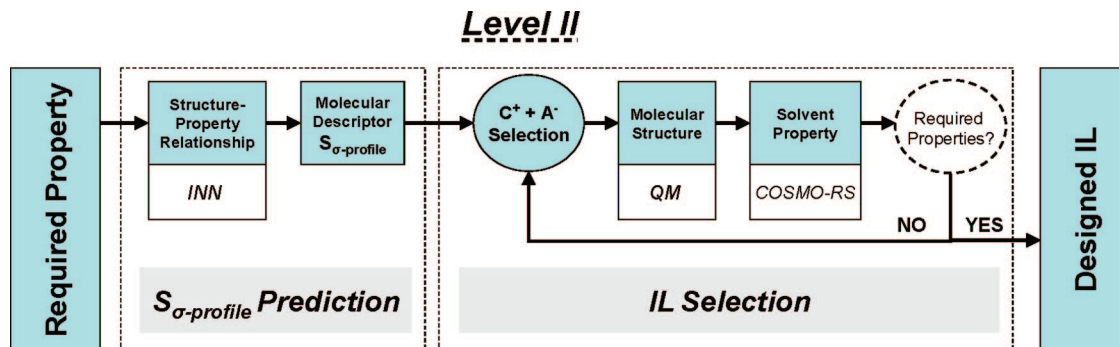


Figure 1. Ionic liquid design tool: *IL Design* strategy schemed as *Level II*.

whose  $S_{\sigma\text{-profile}}$ (1–8) values were successfully related to density of 45 imidazolium based ILs by NNs, following a QSPR procedure (*Level I* of Figure 1 in ref 7). The property values needed to NN design and optimization were calculated by the quantum-chemical COSMO-RS method, which has been shown to calculate density, activity coefficients, and phase-equilibrium data of IL systems suitability.<sup>15–27</sup>

In this work, we report the results of the development of the computational approach schemed as *Level II* in Figure 1, which introduces the required IL properties as input, resulting in the selection of counterions as output, that is directly designing ILs on the computer. The proposed *IL Design* tool consists of two steps: First, inverse neural network (INN) is used in reverse QSPR analysis to obtain the  $S_{\sigma\text{-profile}}$ (1–8) values of the molecular descriptor according to the input set of solvent properties ( $S_{\sigma\text{-profile}}$  Prediction in Figure 1). Subsequently, using the  $S_{\sigma\text{-profile}}$  (1–8) parameters estimated by INN, an IL, or a reduced ILs group can be selected from a database that fulfils previous requirements (physical premises) applying a simple statistical tool (*IL Selection* in Figure 1). If the  $S_{\sigma\text{-profile}}$  values estimated by INN do not resemble those included in the database of known ILs, new cations or anions or additional functional groups could be implemented in such a database to obtain an IL with the needed  $S_{\sigma\text{-profile}}$ . It is important to remark that due to the additive character of the  $S_{\sigma\text{-profile}}$  parameter even more than one cation or anion – IL mixtures – could be considered to fulfill the predicted  $S_{\sigma\text{-profile}}$ . Finally, the selected IL structures will be optimized by quantum-chemical calculations and used in standard COSMO-RS procedure to confirm their required solvent properties, as final validation step (*Level II* in Figure 1). In summary, *Level II* considers first the target criteria, and then the IL is designed by appropriate choice of the ions to deliver the predicted  $S_{\sigma\text{-profile}}$  molecular descriptor, giving a way to systematically characterize the influence of the cation and anion on IL properties. As a result, we develop an a priori computational tool to obtain tailor-made ILs, even including new ILs or their mixtures that have not been synthesized or measured yet. This would avoid the search of new useful ILs based on chemical intuition or by extrapolation from known compounds.

The following successive steps are necessary to develop *Level II* in Figure 1. Toluene/heptane separation by liquid–liquid extraction is considered here as a case of reference for these IL applications. Then a set of thermodynamic properties were chosen, including those of i) pure solvents, density ( $\rho_{\text{IL}}$ ); ii) binary mixtures, solubility of IL in *n*-heptane, measured as molar fraction of IL in *n*-heptane ( $\Pi_{\text{IL-C7}}$ ) and solubility of heptane in IL, measured in terms of molar fraction of IL ( $\Pi_{\text{C7-IL}}$ ); and iii) ternary systems, partition coefficients of toluene between IL and heptane phases ( $\phi_{\text{TOLUENE}}$ ). In a preliminary section, the *IL*

*Screening* tool previously developed in part 1 is extended to the simultaneous prediction of various solvent properties ( $\rho_{\text{IL}}$ ,  $\Pi_{\text{C7-IL}}$ ,  $\Pi_{\text{IL-C7}}$ , and  $\phi_{\text{TOLUENE}}$ ), using solely an  $S_{\sigma\text{-profile}}$  descriptor for 45 imidazolium-based ILs. Following the procedure of part 1, two different molecular models are applied to simulate the IL compounds, to quantify their  $S_{\sigma\text{-profile}}$  values: i) C+A model considers independent counterions of IL and ii) CA model considers the ion-paired structure to simulate the ionic liquid compound. The second section is devoted to design and optimize the inverse neural networks (INNs) for estimating  $S_{\sigma\text{-profile}}$  molecular descriptors as a function of input property values, using for this purpose the previously calculated COSMO-RS property data. Six inverse neural network models are designed to estimate the  $S_{\sigma\text{-profile}}$  parameters, which differ in the number of IL properties (from one to four) and in the molecular model (CA and C+A) used. Subsequently, the suitability of the developed INNs is demonstrated by the reliable prediction of  $S_{\sigma\text{-profile}}$  descriptors corresponding to four new ILs (validation sample), whose type of cation or anion was not included in the initial INN design. It implies that the developed computational tool designed is able to generate the electronic characteristics – in terms of the  $S_{\sigma\text{-profile}}$  molecular descriptor – of a potential IL solvent with the requirements of possessing some specific properties; that is the possibility to invent new IL compounds. Subsequently, an adequate statistical tool using Euclidian distances is proposed to select the most adequate ILs of a known database, using the estimated  $S_{\sigma\text{-profile}}$ (1–8) values. In the third section of this work, the proposed a priori IL design tool in *Level II* is validated with a liquid–liquid extraction example, using multicomponent IL mixtures or reformulated ILs for tuning solvent performance. The final solution offers a set of desirable combinations of ionic components, which satisfy the separation process requirements.

## 2. Computational Details

**Molecular Structures.** All molecular geometries (cationic, anionic and ion-paired species of ILs and organic molecules) were optimized at B3LYP/6-31++G\*\* computational level in the ideal gas phase using quantum chemical *Gaussian03* package.<sup>28</sup> Vibrational frequency calculations were performed for each case to confirm the presence of an energy minimum.

**$S_{\sigma\text{-profile}}$  Descriptor.** *Gaussian03* was used to compute the COSMO files. The ideal screening charges on the molecular surface for each species were calculated by the continuum solvation COSMO model using BVP86/TZVP/DGA1 level of theory.<sup>29–31</sup> Subsequently, COSMO files were used as input in *COSMOtherm*<sup>32</sup> statistical thermodynamic code to obtain the histogram function  $\sigma\text{-profile}$  of each species for both CA and C+A models. Then the  $S_{\sigma\text{-profile}}$  molecular descriptor of ILs was

**Table 1. Main Format Characteristic of Learning, Verification, and Validation Samples Depending on the NN Used in This Work**

	sample type	input	output	number of data set	number of Rows
<b>Level I:</b> DNN <sub>CA</sub> , DNN <sub>C+A</sub>	leaning	8 [ $S_{\sigma\text{-profile}}(1-8)$ ]	4 <sup>a</sup>	40	12
	verification	8 [ $S_{\sigma\text{-profile}}(1-8)$ ]	4 <sup>a</sup>	5	12
	validation	8 [ $S_{\sigma\text{-profile}}(1-8)$ ]	4 <sup>a</sup>	4	12
<b>Level II:</b> INN <sub>CA</sub> <sup>1</sup> , INN <sub>C+A</sub> <sup>1</sup>	learning	1 ( $\rho_{\text{IL}}$ )	8 [ $S_{\sigma\text{-profile}}(1-8)$ ]	40	9
	verification	1 ( $\rho_{\text{IL}}$ )	8 [ $S_{\sigma\text{-profile}}(1-8)$ ]	5	9
	validation	1 ( $\rho_{\text{IL}}$ )	8 [ $S_{\sigma\text{-profile}}(1-8)$ ]	4	9
INN <sub>CA</sub> <sup>2</sup> , INN <sub>C+A</sub> <sup>2</sup>	learning	2 ( $\rho_{\text{IL}}$ , $\Pi_{\text{IL/C7}}$ )	8 [ $S_{\sigma\text{-profile}}(1-8)$ ]	40	10
	verification	2 ( $\rho_{\text{IL}}$ , $\Pi_{\text{IL/C7}}$ )	8 [ $S_{\sigma\text{-profile}}(1-8)$ ]	5	10
	validation	2 ( $\rho_{\text{IL}}$ , $\Pi_{\text{IL/C7}}$ )	8 [ $S_{\sigma\text{-profile}}(1-8)$ ]	4	10
INN <sub>CA</sub> <sup>4</sup> , INN <sub>C+A</sub> <sup>4</sup>	learning	4 <sup>a</sup>	8 [ $S_{\sigma\text{-profile}}(1-8)$ ]	40	12
	verification	4 <sup>a</sup>	8 [ $S_{\sigma\text{-profile}}(1-8)$ ]	5	12
	validation	4 <sup>a</sup>	8 [ $S_{\sigma\text{-profile}}(1-8)$ ]	4	12

<sup>a</sup>  $\rho_{\text{IL}}$ ,  $\Pi_{\text{C7-IL}}$ ,  $\Pi_{\text{IL-C7}}$ ,  $\phi_{\text{TOLUENE}}$ .

evaluated by the estimation of the probabilistic surface charge distribution area below  $\sigma$ -profile on eight different polarity regions [ $S_{\sigma\text{-profile}}(1-8)$ ], following the procedure described in part 1 of this work.<sup>7</sup> The  $S_{\sigma\text{-profile}}(1-8)$  values of the 45 ILs for both CA and C + A models are provided as Supporting Information.

**Thermodynamic Properties.** COSMOtherm software is applied to calculate the thermodynamic properties ( $\rho_{\text{IL}}$ ,  $\Pi_{\text{C7-IL}}$ ,  $\Pi_{\text{IL-C7}}$ , and  $\phi_{\text{TOLUENE}}$ ) of ILs and their mixtures at 298 K using as input the previously obtained COSMO files for the neutral ion-paired species. According to our chosen method, we used the corresponding parametrization (BP\_TZVP\_C21\_0105) that is required for the calculation of physicochemical data and contains intrinsic parameters of COSMOtherm as well as element specific parameters (Table S1 in the Supporting Information).

**Direct Neural Network (DNN) Design.** A detailed information related with direct neural networks used in this work and their optimization procedures were provided in previous part 1 (called neural networks) devoted to the development of *Level I*.<sup>7,33</sup> In this case, the density ( $\rho_{\text{IL}}$ ), mutual solubility of ILs/*n*-heptane ( $\Pi_{\text{C7-IL}}$ ,  $\Pi_{\text{IL-C7}}$ ), partition coefficients of toluene ( $\phi_{\text{TOLUENE}}$ ) in ternary mixtures composed of heptane, toluene, and IL, and the  $\sigma$ -profiles of ILs for CA and C+A models (calculated by COSMO-RS) were correlated by a multilayer perceptron (MLP) neural network model.<sup>7,34</sup>

**Inverse Neural Network (INN) Design.** The design (type of neural network, number of layers, activation, transfer, and training functions) and optimization process of INN are the same as those of the DNN.<sup>7,35</sup> The main difference between both neural network models consists of the input and output data used in the INN models and the NNs topology (input, hidden, and output neurons). Here, eight output values have been estimated (one for each of the eight different polarity regions [ $S_{\sigma\text{-profile}}(1-8)$ ]) from three different groups of input values. Given the corresponding three groups of properties [i] pure solvents, density ( $\rho_{\text{IL}}$ ); ii) binary systems, mutual solubilities IL/heptane ( $\Pi_{\text{C7-IL}}$ ,  $\Pi_{\text{IL-C7}}$ ); and iii) ternary systems, partition coefficient of toluene between IL/heptane phases ( $\phi_{\text{TOLUENE}}$ ), three INNs have been designed, optimized, and applied for CA and C+A models (Table 1): The first group of models (INN<sub>CA</sub><sup>1</sup> and INN<sub>C+A</sub><sup>1</sup>) uses the density of IL as input variable (one input and eight output neurons). The second group is formed by two NNs (INN<sub>CA</sub><sup>2</sup> and INN<sub>C+A</sub><sup>2</sup>) that estimate the  $S_{\sigma\text{-profile}}(1-8)$  parameters using the density and solubility of IL in *n*-heptane (two input and eight output neurons). Finally, two more NNs (INN<sub>CA</sub><sup>4</sup> and INN<sub>C+A</sub><sup>4</sup>) are used to estimate the descriptor values using the density, mutual solubility of ILs/*n*-heptane, and partition coefficients of toluene in the aforementioned ternary mixtures (four input and eight output neurons).

As can be seen in Table 1, although all learning and verification samples utilize different formats, in every case, the data have been randomly divided in the same way as in DNN (40 and 5 ILs, respectively).

Every NN model used in this work has been designed using *Matlab* version 7.01.24704 (R14)<sup>36</sup> and the statistical analysis were carried out by *Statgraphics 5.1*.<sup>37</sup>

**Data Set.** The IL compounds in the data sets were divided into learning, verification, and validation samples which are made by 40, 5, and 4 ILs, respectively. These samples can be classified into two groups: i) Learning and verification samples used for the DNN and INN models generation, which include 45 imidazolium-based IL generated by combining the cations 1-methyl- (Mmim<sup>+</sup>), 1-ethyl- (Emim<sup>+</sup>), 1-butyl- (Bmim<sup>+</sup>), 1-hexyl- (Hxmim<sup>+</sup>), and 1-octyl-3-methylimidazolium (Omim<sup>+</sup>) with the anions chloride (Cl<sup>-</sup>), bromide (Br<sup>-</sup>), tetrafluoroborate (BF<sub>4</sub><sup>-</sup>), tetrachloroferrate (FeCl<sub>4</sub><sup>-</sup>), hexafluorophosphate (PF<sub>6</sub><sup>-</sup>), methylsulfate (MeSO<sub>4</sub><sup>-</sup>), trifluoromethanesulfonate (CF<sub>3</sub>SO<sub>3</sub><sup>-</sup>), ethylsulfate (EtSO<sub>4</sub><sup>-</sup>), and bis(trifluoromethanesulfonyl)imide (Tf<sub>2</sub>N<sup>-</sup>); and ii) A validation sample set used for the evaluation of the DNN and INN generated models, which includes 4 new ILs generated by combination of the cations 1-nonyl-3-methylimidazolium (Nmim<sup>+</sup>) and 1-butyl-pyridinium (Bpyr<sup>+</sup>) with the anions dicyanamide (DCN<sup>-</sup>) and methanesulfonate (CH<sub>3</sub>SO<sub>3</sub><sup>-</sup>), all having chemical differences with respect to the learning and verification samples.

### 3. Results

**3.1. Extending Level I.** The first step in current development is to extend the previously developed *Level I* in part 1 to the simultaneous prediction of a set of IL solvent properties of interest. For this purpose, we design a mathematical function ( $\Psi$ , eq 1) based on DNN for estimating simultaneously four different properties  $\rho_{\text{IL}}$ ,  $\Pi_{\text{C7-IL}}$ ,  $\Pi_{\text{IL-C7}}$ , and  $\phi_{\text{TOLUENE}}$  values from the a priori  $S_{\sigma\text{-profile}}$  molecular descriptor described before.<sup>7,15</sup> In addition, the direct mathematical procedure of the DNN is validated using C+A and CA models. That is, the estimations of  $\rho_{\text{IL}}$ ,  $\Pi_{\text{C7-IL}}$ ,  $\Pi_{\text{IL-C7}}$ , and  $\phi_{\text{TOLUENE}}$  values (output data, Table 1) from  $S_{\sigma\text{-profile}}(1-8)$  parameters (input data, Table 1) is carried out (vide infra) in C+A and CA databases. Then in both neural networks (DNN<sub>CA</sub> and DNN<sub>C+A</sub>) the parameters Learning coefficient (Lc), learning coefficient decrease (Lcd), and learning coefficient increase (Lci) were optimized following the calculation process described above, giving the results in Table 2. Lc is similar to the *h* parameter in Newton's method, often called Newton–Raphson method or the Newton–Fourier method, Lcd and Lci parameters control the Lc value as a function of the difference between real and estimated values. Using the verification sample of five ILs, the optimized hidden neurons



**Table 2. Main Characteristics of the DNNs Used in Level I and Statistical Results of Estimations Corresponding to Verification+Validation Sample**

	Properties: $\rho_{IL}$ , $\Pi_{C7-IL}$ , $\Pi_{IL-C7}$ , $\phi_{TOLUENE}$	
	DNN <sub>CA</sub>	DNN <sub>C+A</sub>
HNN	10	10
Lc	0.0010	0.5000
Lcd	1.0000	0.5000
Lci	10	85
Statistical Parameters		
$R^2$	0.997	0.988
rmsd (%)	2.3	3.7
$\sigma$ (%)	0.23	0.32

number (HNNs), Lc, Lcd, and Lci values were tested calculating the statistical parameters, as seen in Table 2. In view of these results, both DNN<sub>CA</sub> and DNN<sub>C+A</sub> provide good estimations of  $\rho_{IL}$ ,  $\Pi_{C7-IL}$ ,  $\Pi_{IL-C7}$ , and  $\phi_{TOLUENE}$  values based on the corresponding  $S_{\sigma-profile}$  descriptors, which are derived solely from quantum-chemical calculations. The normalized estimations of DNN<sub>CA</sub> and DNN<sub>C+A</sub> versus those normalized COSMO-RS values for verification sample are shown in Figure 2. The values of the four  $\rho_{IL}$ ,  $\Pi_{C7-IL}$ ,  $\Pi_{IL-C7}$ , and  $\phi_{TOLUENE}$  properties have been linearly normalized between 0 and 1 to facilitate their mutual comparison. In addition, Figure 2 also presents the predictions for validation sample, which corresponds to four new ILs, not included in the initial IL series used for DNNs design. The main statistical parameters obtained for the verification+validation sample with DNN<sub>CA</sub> are root mean standard deviation (rmsd), less than 2.3%, and correlation coefficient ( $R^2$ ), higher than 0.996. DNN<sub>CA</sub> estimates slightly more accurately the property values calculated by COSMO-RS than DNN<sub>C+A</sub> model (rmsd < 3.7,  $R^2$  > 0.988) (also Figure 2). This difference is based on that the C+A model does not describe the effect of counterions interaction of IL compounds, and therefore the correlation between the descriptor and the IL properties is more hidden. However, it is also important to bear the computational efforts in mind: In the systems studied here, the average computational time cost to calculate the C+A model database is 7 times less than to make the CA model database.

Current results show that the  $S_{\sigma-profile}$  descriptor includes the compound information necessary to characterize the IL solvent behavior in pure fluid and binary/ternary mixtures, suggesting the high potential application of the proposed  $S_{\sigma-profile}$  descriptor to other IL properties of interest.

**3.2. Development Level II.** The principal aim of this work is to define a useful IL design tool able to select one or a reduced group of ILs with specific physical properties, following the strategy schemed as *Level II* in Figure 1. The proposed computational approach consists of two steps:

**$S_{\sigma-profile}$  Prediction.** Inverse neural networks (INNs) have been used, for the first time, to estimate the molecular descriptor  $S_{\sigma-profile(1-8)}$  values from a set of specified IL properties, in this case density of ILs ( $\rho_{IL}$ ), *n*-heptane solubility in a IL ( $\Pi_{C7-IL}$ ), IL solubility in *n*-heptane ( $\Pi_{IL-C7}$ ) and partition coefficient of toluene in ternary mixtures composed of *n*-heptane, IL, and toluene ( $\phi_{TOLUENE}$ ).

In the previous section, DNN models were successfully tested, obtaining an adequate nonlinear relation ( $\Psi$ ) between independent ( $\rho_{IL}$ ,  $\Pi_{C7-IL}$ ,  $\Pi_{IL-C7}$ , and  $\phi_{TOLUENE}$ ) and dependent ( $\sigma-profile$ ) variables, as seen in eq 1. Taking into account the main characteristics of the DNN,  $\Psi$  is based on sigmoid function which relates the output with the input in every neuron.

$$\text{properties} = \psi(S_{\sigma-profile}(1-8)) \quad (1)$$

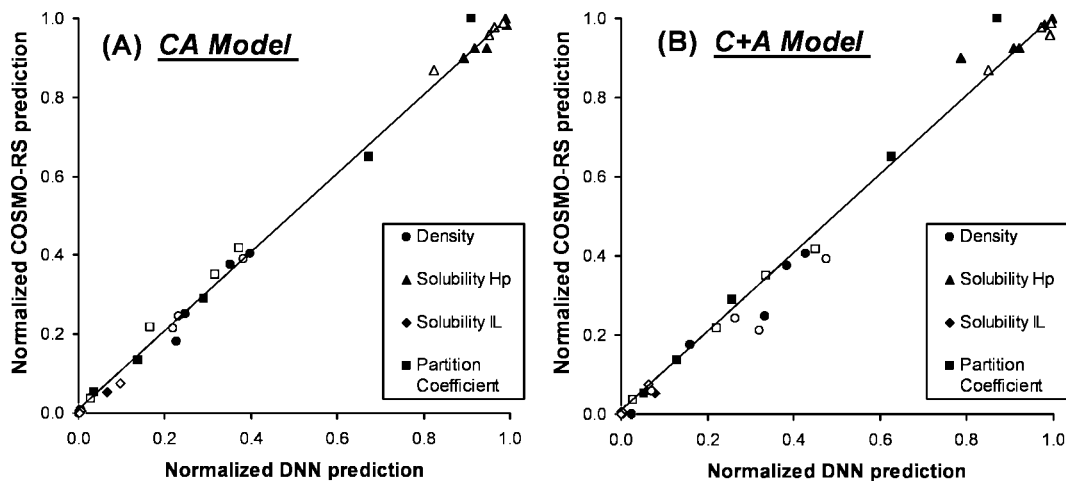
Now, as the first stage on the seeking of the tailor-made IL, it is necessary found the inverse function ( $\Omega$ ), that is the relation between independent [ $S_{\sigma-profile(1-8)}$ ] and dependent ( $\rho_{IL}$ ,  $\Pi_{C7-IL}$ ,  $\Pi_{IL-C7}$ , and  $\phi_{TOLUENE}$ ) variables, as seen in eq 2.  $\Psi$  and  $\Omega$  functions are based on the same nonlinear function (sigmoid). Second, when the  $\Omega$  function is known, for a given properties ( $\rho_{IL}$ ,  $\Pi_{C7-IL}$ ,  $\Pi_{IL-C7}$ , and  $\phi_{TOLUENE}$ ), the  $S_{\sigma-profile}$  parameters could be estimated, and then, fixed the chemical characteristics of the IL compound to be selected.

$$S_{\sigma-profile}(1-8) = \Omega(\text{properties}) \quad (2)$$

To find the  $\Omega$  function, it is necessary to calculate the inverse of the direct neural network model ( $\Psi$ ).<sup>34</sup> Therefore, depending on the properties and the database (CA or C+A models) used to estimate the descriptor values, six inverse neural network models have been designed (Table 1). First, only using one property (density,  $\rho_{IL}$ ) of the ILs, two of them (INN<sub>CA</sub><sup>1</sup> and INN<sub>C+A</sub><sup>1</sup>) were designed; second, using two properties ( $\rho_{IL}$ ,  $\Pi_{C7-IL}$ ), two more INNs (INN<sub>CA</sub><sup>2</sup> and INN<sub>C+A</sub><sup>2</sup>) were built up; and finally the two last ones INNs (INN<sub>CA</sub><sup>4</sup> and INN<sub>C+A</sub><sup>4</sup>) were designed using all the four properties tested ( $\rho_{IL}$ ,  $\Pi_{C7-IL}$ ,  $\Pi_{IL-C7}$ , and  $\phi_{TOLUENE}$ ). The design and optimization of these six models have been carried out following the mathematical procedures exposed in Computational Details. As a result, Table 3 shows the values of the main parameters of INNs and their statistical results. These have been calculated comparing the  $S_{\sigma-profile(1-8)}$  values given by COSMO-RS from the verification+validation sample and their respective estimations by INN models. Given the different nature of  $\Psi$  and  $\Omega$  functions (eqs 1 and 2, respectively), the optimum parameters values of DNN and INN models are not comparable. As can be observed, all INNs provide good estimations of  $S_{\sigma-profile}$  descriptors in Figure 3. Table 3 reports the main characteristics of the INNs used in *Level II* and statistical results of estimations corresponding to the verification+validation sample. As it was expected, the quality of correlations improves with the number of properties used, that is with the number of depending variables. It is noteworthy, anyhow, that INN models were able to predict accurately eight  $S_{\sigma-profile(1-8)}$  parameters using only one property – in this case, density – as target input (Table 3). Again, INN<sub>CA</sub> estimates slightly more accurately the descriptor values given by COSMO-RS compared to INN<sub>C+A</sub>, but such deviations are scarcely significant. It should be noted the meaning of the correct prediction for validation sample by INN<sub>CA</sub> and INN<sub>C+A</sub> because the ILs included in this sample contain new cations or anions, which were previously unknown for INN models. In other words, INNs successfully generate the polar electronic information of a potential IL, which provides it the chemical characteristic to obtain a solvent with the properties specified in the design.

**IL Selection.** Once the  $\Omega$  function (eq 2) has been found, for a given group of properties ( $\rho_{IL}$  or  $\rho_{IL}$  and  $\Pi_{C7-IL}$ ,  $\Pi_{IL-C7}$  or  $\rho_{IL}$ ,  $\Pi_{C7-IL}$ ,  $\Pi_{IL-C7}$ , and  $\phi_{TOLUENE}$ ) inside the learning range, the  $S_{\sigma-profile(1-8)}$  have been adequately estimated. Then, using these descriptor values, the most adequate IL of a known database used can be selected using a simple statistical tool. In this work, the selection of the appropriate IL for the needed  $S_{\sigma-profile(1-8)}$ , estimated from the target properties, was carried out using the Euclidean distance.

The Euclidean distance is defined as the straight line distance between two points, that is it is the shortest distance between two points. In a plane, with  $p_1$  at ( $x_1$ ,  $y_1$ ) and  $p_2$  at ( $x_2$ ,  $y_2$ ), the distance is  $[(x_2 - x_1)^2 + (y_2 - y_1)^2]^{1/2}$ .<sup>36</sup> This expression is widely used to calculate the modulus of vectors or the hypotenuse of a right triangle by Pythagoras' theorem. In



**Figure 2.** Performance of DNN models for the simultaneous prediction of four IL properties ( $\rho_{IL}$ ,  $\Pi_{C7-IL}$ ,  $\Pi_{IL-C7}$ , and  $\phi_{TOLUENE}$ ) of verification (filled icons) and validation (open icons) samples using the (A) CA model and the (B) C+A model. Scales are normalized.

**Table 3.** Main Characteristics of the INNs Used in Level II and Statistical Results of Estimations Corresponding to Verification+Validation Sample

	$\rho_{IL}$		$\rho_{IL}, \Pi_{C7-IL}$		$\rho_{IL}, \Pi_{C7-IL}, \Pi_{IL-C7}, \phi_{TOLUENE}$	
	INN <sub>CA</sub> <sup>1</sup>	INN <sub>C+A</sub> <sup>1</sup>	INN <sub>CA</sub> <sup>2</sup>	INN <sub>C+A</sub> <sup>2</sup>	INN <sub>CA</sub> <sup>4</sup>	INN <sub>C+A</sub> <sup>4</sup>
HNN	10	10	10	10	10	10
Lc	0.5005	0.0010	1.0000	1.0000	1.0000	1.0000
Lcd	0.5005	0.0010	0.0010	0.0010	0.0010	0.0010
Lci	40	60	50	50	50	50
Statistical Parameters						
R <sup>2</sup>	0.993	0.987	0.997	0.973	>0.999	>0.999
rmsd (%)	3.9	5.9	0.08	0.1	0.008	0.01
$\sigma$ (%)	0.15	0.16	0.06	0.18	~0	~0

addition, this distance can be used to compare quantitatively which point from a group of points is more similar to a given point (the less the distance is, the more similar the two points are). With such an objective, this tool has been applied here to quantify the similarity of the estimated  $S_{\sigma\text{-profile}}$  by INNs and the corresponding values of the descriptor from database (provided by COSMO-RS) in an 8D space (one for each  $S_{\sigma\text{-profile}}(1-8)$ ). For instance, for a given set of properties, the INN estimates the molecular descriptor, and then, the Euclidean distances,  $d(S_{\sigma\text{-profile}}^{\text{estimated}}, S_{\sigma\text{-profile}}^{\text{database}})_{IL}$ , between the  $S_{\sigma\text{-profile}}(1-8)$  estimations by INN and all of the corresponding COSMO-RS values for the 49 ILs from database are calculated by eq 3.<sup>36</sup> Taking into account that the IL from the database that presents the least distance is the most similar IL, the most suitable IL can be selected.

$$d(S_{\sigma\text{-profile}, i}^{\text{estimated}}, S_{\sigma\text{-profile}, i}^{\text{database}})_{IL} = \sqrt{\sum_{j=1}^8 (S_{\sigma\text{-profile}, i}^{\text{estimated}} - S_{\sigma\text{-profile}, i}^{\text{database}})^2} \quad (3)$$

In eq 3,  $S_{\sigma\text{-profile}, i}^{\text{estimated}}$  and  $S_{\sigma\text{-profile}, i}^{\text{database}}$  are the descriptors calculated by INN models and those of the database used (COSMO-RS), respectively. To verify if the proposed statistical tool is able to select the appropriate IL for a given property, the validation sample was used, which includes new ILs lately included in the CA or C+A databases but neither in the learning nor verification samples of the INN designs. For these ILs, the  $S_{\sigma\text{-profile}, i}^{\text{estimated}}$  values were estimated by  $\Omega$  function (eq 2) and then these values were compared with  $S_{\sigma\text{-profile}, i}^{\text{database}}$  given by COSMO-RS for the six INN models designed. As example of the results, the estimated  $S_{\sigma\text{-profile}}(1-8)$  obtained using the four properties of bpyrBF<sub>4</sub> in INN<sub>C+A</sub><sup>4</sup> and INN<sub>CA</sub><sup>4</sup> have been com-

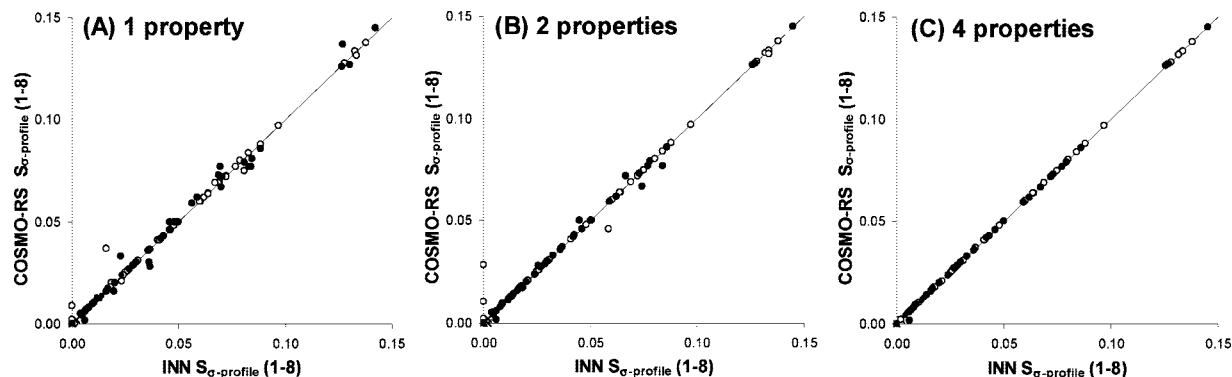
pared to the COSMO-RS descriptors of each IL from C+A and CA database, respectively. The obtained Euclidean distances,  $d(S_{\sigma\text{-profile}}^{\text{estimated}}, S_{\sigma\text{-profile}}^{\text{database}})_{IL}$ , for the 49 ILs in the database are shown in Figure 4. As can be seen, the least one is the distance between estimated and real molecular descriptor of bpyrBF<sub>4</sub> with both C+A and CA databases. In addition, other similar ILs could be also selected, namely bmimBF<sub>4</sub> and emimBF<sub>4</sub>. The higher Euclidean distances between bpyrBF<sub>4</sub> descriptors and those of other ILs observed in Figure 4 for the CA database are assignable to the higher suitability of the CA molecular model to differentiate IL electronic characteristics. The developed statistical tool obtained exhibits identical satisfactory behavior for the four ILs of validation sample in both CA and C+A databases, being the obtained Euclidean distances shown in Table 4. As can be seen, in both databases the minimum distance decreases with the number of properties used.

Of course, a more extended database – with additional cations, anions, or functional groups – could be necessary to obtain a new IL structure with the  $S_{\sigma\text{-profile}}(1-8)$  required. However, these results confirm that current approach based on Euclidean distances is an adequate preliminary tool to select an IL with specific physical properties from previously known databases, such as CA and C+A with 49 ILs used here.

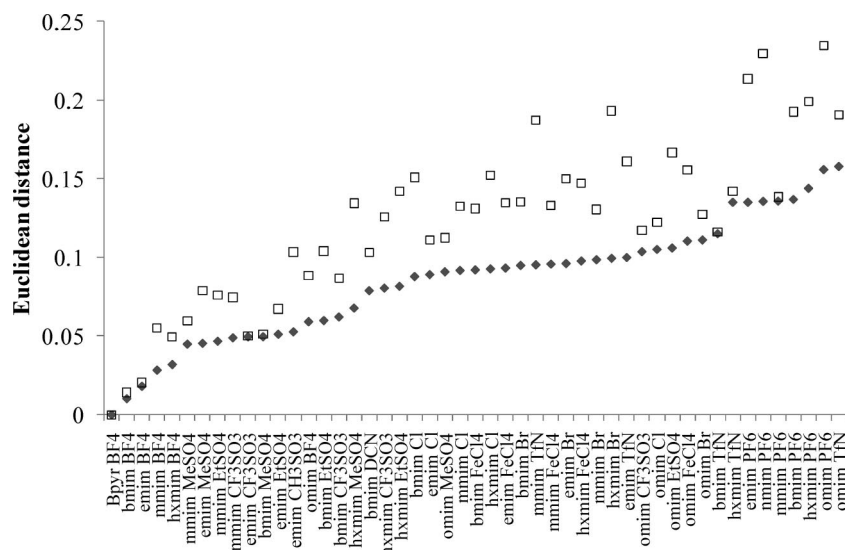
**3.3. Application Example.** The suitability of IL solvents for the liquid–liquid (LL) extraction of toluene from *n*-heptane mixtures has been experimentally investigated.<sup>38–41</sup> A recent process evaluation showed that the investments of this separation can be reduced using ILs with a high aromatic distribution coefficient and a reasonable aromatic/aliphatic selectivity.<sup>42</sup> Obviously, only a limited group of ILs (~20) have been investigated among the large number of possible combinations (>10<sup>4</sup>). For this reason, the proposed methodology is illustrated here with the design of task-specific ILs for aromatic/aliphatic separation. Following the criteria for solvent design in LL extractions, properties have been analyzed as numerical indicators of solvent performance: i) density ( $\rho$ ), ii) partition coefficient at infinite dilution ( $\phi_{TOLUENE}$ ), and iii) selectivity at infinite dilution ( $\beta_{Toluene/Heptane}$ ), which is approximated by the expression:

$$\beta_{Toluene/Heptane} = \frac{X_{Heptane}^{\text{RaffinatePhase}} X_{Toluene}^{\text{ExtractPhase}}}{X_{Heptane}^{\text{ExtractPhase}} X_{Toluene}^{\text{RaffinatePhase}}} \approx \frac{(1 - \Pi_{IL-C7})}{(1 - \Pi_{C7-IL})} \varphi_{Toluene} \quad (4)$$

Figure 5 shows the values of these indicators computed by the COSMO-RS method at 298 K, which qualitative resembles



**Figure 3.** Performance of INN model for the estimation of  $S_{\sigma-profile}(1-8)$  descriptor parameters of CA model (filled icons) and C+A model (open icons) fixing as input data: (A) One property ( $\rho_{IL}$ ); (B) two properties ( $\rho_{IL}$ ,  $\Pi_{IL-C7}$ ); and (C) four properties ( $\rho_{IL}$ ,  $\Pi_{C7-IL}$ ,  $\Pi_{IL-C7}$ , and  $\phi_{TOLUENE}$ ).



**Figure 4.** Euclidean distance between  $S_{\sigma-profile}(1-8)$  estimated by  $INN_{C+A}^4$  (filled diamonds) and  $INN_{CA}^4$  (open squares) for bpyrBF<sub>4</sub> properties and the COSMO-RS values for the 49 ILs in the C+A and CA databases, respectively.

**Table 4.** Euclidean Distance between the  $S_{\sigma-profile}(1-8)$  Values Estimated by INN Models and Those Given by COSMO-RS for the Validation Sample

	BmimDCN	EmimMeSO <sub>3</sub>	BpyrBF <sub>4</sub>	NmimPF <sub>6</sub>
$INN_{C+A}^1$	0.00901	0.02204	0.00270	0.00216
$INN_{C+A}^2$	0.00000	0.00001	0.00030	0.00283
$INN_{C+A}^2$	~0	~0	~0	~0
$INN_{CA}^1$	0.00430	0.01174	0.01038	0.01219
$INN_{CA}^2$	0.00042	0.00268	0.00123	0.00700
$INN_{CA}^4$	0.00002	0.00003	0.00001	0.00000

the experimental data reported by Meindersma et al.<sup>42</sup> for mixtures of 10% (*v/v*) toluene in heptane at 313 K. The solvent sulfolane is used as benchmark because it is widely used as an aromatic/aliphatic extractive solvent in the industry. Therefore, attending to its COSMO-RS estimated indicators, suitable ILs for this separation must show  $\phi_{TOLUENE} \geq 0.94$  and/or  $\beta_{Toluene/Heptane} \geq 14$  and  $\rho \geq 1285 \text{ kg/m}^3$ . As can be seen in Figure 5, there are only 10 ILs, among the 45 imidazolium based ILs studied, that have partition coefficient of toluene greater than or equal to 0.94 and selectivity greater than or equal to 14. Thus, it is observed that the commonly noted hydrophobic anions (as  $Tf_2N^-$  or  $PF_6^-$ ; however,  $FeCl_4^-$  must be excluded due to its high reactivity) confer to ILs the highest toluene partition coefficients. This  $\phi_{TOLUENE}$  value decreasing in the order  $Tf_2N^- > PF_6^- > CF_3SO_3^- > CH_3SO_4^- > C_2H_5SO_4^- > BF_4^- > Br^- > Cl^-$  (Table S3 in the Supporting Information). In addition, the type of anion has

great influence on density, increasing this value in the order  $Br^- > Tf_2N^- > PF_6^- > CF_3SO_3^- > CH_3SO_4^- > C_2H_5SO_4^- > BF_4^- > Cl^-$ , but this scarcely affects the selectivity. Regarding the effect of the cation on these IL properties, it is found that the increase of the alkyl chain of cation implies significantly higher  $\phi_{TOLUENE}$  values but also a great loss of selectivity and lower density (Table S3 in the Supporting Information). As a result, Figure 5 shows that most of studied ILs with  $\phi_{TOLUENE} > 0.94$  present values of  $\beta_{Toluene/Heptane} < 14$  and disperse density. In other words, it is difficult to implement all targeted properties in one single salt. Two alternative approaches have been proposed to be implement in these cases: i) Using a multicomponent approach to utilize the solvent behavior of known ILs to obtain IL mixtures with the required properties; and ii) Modifying the structure of the designed target with additional anion, cation or functional groups, that is designing new ILs.<sup>43</sup> It is remarkable that both approaches would be affordable by the proposed IL design tool using the Euclidean distance based tool with adequate databases.

As an application example, the proposed methodology is following demonstrated with this aromatic/aliphatic separation by the approach of designing IL mixtures. First, the required target properties are selected as key factors in solvent performance for two cases: *Solvent 1* [ $\phi_{TOLUENE} = 2.4$ ;  $\beta_{Toluene/Heptane} = 14$  and  $\rho = 1450 \text{ kg/m}^3$ ], which corresponds to an



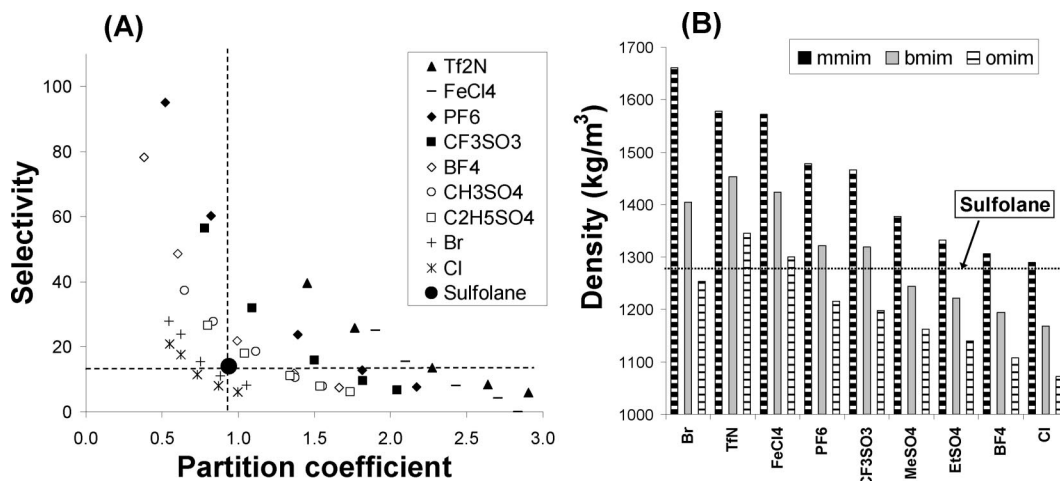


Figure 5. COSMO-RS simulation of solvent selection indicators for toluene/*n*-heptane separations with ILs.

Table 5. Design Results of Liquid–Liquid Extraction for the Toluene/*n*-Heptane System with Ternary IL Mixtures for the CA model

	prediction	$\phi^a$	$\beta^b$	S(1)	S(2)	S(3)	S(4)	S(5)	S(6)	S(7)	S(8)
solvent 1	INN <sub>CA</sub> <sup>4</sup>	2.4	14	0.004	0.025	0.111	0.182	0.154	0.139	0.039	0.000
EmimTf <sub>2</sub> N+OmimTf <sub>2</sub> N	COSMO-RS	2.5	13	0.000	0.026	0.096	0.188	0.156	0.110	0.029	0.000
Solvent 2	INN <sub>CA</sub> <sup>4</sup>	1.8	20	0.000	0.024	0.099	0.261	0.259	0.071	0.044	0.000
BmimTf <sub>2</sub> N+BmimPF <sub>6</sub>	COSMO-RS	1.9	17	0.000	0.026	0.093	0.241	0.250	0.066	0.046	0.000

<sup>a</sup>  $\phi_{\text{TOLUENE}}$ ; <sup>b</sup>  $\beta_{\text{Toluene/Heptane}}$ .

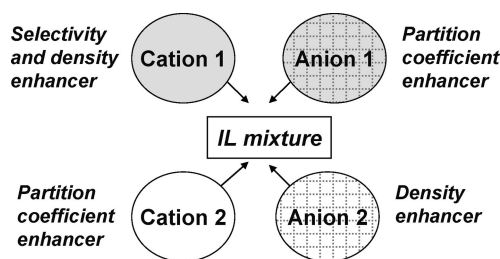


Figure 6. Concept of multicomponent IL mixtures for enhancement of solvent properties.

enhanced solvent with high partition coefficient, high density, and acceptable selectivity; and *Solvent 2* [ $\phi_{\text{TOLUENE}} = 1.8$ ;  $\beta_{\text{Toluene/Heptane}} = 20$ ] and  $\rho = 1400 \text{ kg/m}^3$ ], which presents an intermediate performance with a partition coefficient twice that of sulfolane ( $\phi_{\text{TOLUENE}} > 0.94$ ) but with enhanced selectivity ( $\beta_{\text{Toluene/Heptane}} > 14$ ) and density ( $\rho > 1285 \text{ kg/m}^3$ ). Following *Level II* scheme, the required property values of *Solvents 1* and *2* are used in INN<sub>CA</sub><sup>4</sup> and INN<sub>C+A</sub><sup>4</sup> models to obtain the  $S_{\sigma\text{-profile}}(1-8)$  values corresponding to these potential solvents (Table 5). Once  $S_{\sigma\text{-profile}}$  estimation is done, the IL selection will be performed using the concept of multicomponent approach in Figure 6.<sup>42</sup> Thus, the statistical tool based on Euclidean distances will be adapted to obtain the most appropriate mixtures of ILs (defined as combinations of two ILs with one common ion, that is ternary IL mixture) for both CA and C+A databases. Table 6 shows the final solution of the IL selection procedure as a set of ternary IL mixtures obtained with the least Euclidean distances among all the possible ternary IL mixtures with the database including all IL compounds considered. Comparing the mean Euclidean distance values of both databases, the combinations of ILs EmimTf<sub>2</sub>N+OmimTf<sub>2</sub>N and BmimTf<sub>2</sub>N+BmimPF<sub>6</sub> are selected as potential solvents with the target properties of *Solvent 1* and *Solvent 2*, respectively. To verify the IL design results in *Level II*, the  $S_{\sigma\text{-profile}}(1-8)$  values and thermodynamic properties of EmimTf<sub>2</sub>N+OmimTf<sub>2</sub>N and

Table 6. Ternary IL Mixture Selection Using Euclidean Distances from the Estimated  $S_{\sigma\text{-profile}}(1-8)$  Values for *Solvents 1* and *2* with Target Properties

Solvent 1							
CA model				C+A model			
ion 1	ion 2	ion 3	ion 1	ion 1	ion 2	ion 3	Euc. dist.
Emim	Omim	Tf <sub>2</sub> N	0.058	emim	omim	Tf <sub>2</sub> N	0.006
Mmim	Omim	Tf <sub>2</sub> N	0.062	emim	nmim	Tf <sub>2</sub> N	0.041
CF <sub>3</sub> SO <sub>3</sub>	Hxmim	Tf <sub>2</sub> N	0.066	mmim	omim	Tf <sub>2</sub> N	0.064
DCN	Omim	Tf <sub>2</sub> N	0.067	emim	nmim	Tf <sub>2</sub> N	0.118

Solvent 2							
CA model				C+A model			
ion 1	ion 2	ion 3	Euc. dist.	ion 1	ion 2	ion 3	Euc. dist.
Tf <sub>2</sub> N	PF <sub>6</sub>	bmim	0.060	Tf <sub>2</sub> N	PF <sub>6</sub>	bmim	0.0040
Tf <sub>2</sub> N	FeCl <sub>4</sub>	Bmim	0.062	Tf <sub>2</sub> N	PF <sub>6</sub>	nmim	0.0204
Tf <sub>2</sub> N	FeCl <sub>4</sub>	pyrbut	0.065	Tf <sub>2</sub> N	PF <sub>6</sub>	emin	0.0361
CF <sub>3</sub> SO <sub>3</sub>	FeCl <sub>4</sub>	hxmim	0.067	CF <sub>3</sub> SO <sub>3</sub>	PF <sub>6</sub>	hxmim	0.0369

BmimTf<sub>2</sub>N+BmimPF<sub>6</sub> mixtures are calculated by COSMO-RS method and compared in Table 5 for CA model. Then, the performance of the developed IL design tool schemed in *Level II* is demonstrated by the good agreement between the introduced target properties and  $S_{\sigma\text{-profile}}(1-8)$  estimations with those values calculated by COSMO-RS for the IL mixtures obtained (Table 5).

Clearly, the alternative of modifying the IL structures to design *Solvents 1* and *Solvent 2* with target criteria would require to amplify the current database by the addition of new functional groups or ionic components. As a simple example, the addition of phenyl substituent in the CA database results in the proposal of Phenyl+EmimPF<sub>6</sub> mixture as a candidate for *Solvent 2*, which suggests the 1-benzyl-3-ethylimidazolium hexafluorophosphate compound as a new IL with target properties. The subsequent validation by COSMO-RS calculations shows aromatic/aliphatic extractive indicators [ $\phi_{\text{TOLUENE}} = 1.9$ ;  $\beta_{\text{Toluene/Heptane}} = 19\%$  and  $\rho > 1348 \text{ kg/m}^3$ ] close to *Solvent 2* values. Such approach of designing new

ILs by extending the CA and C+A databases, with new groups and ions, will be object of a detailed study in future investigations.

#### 4. Conclusions

An innovative computational approach has been developed to design ionic liquids (ILs) directly on the computer, by the integration of COSMO-RS methodology, and neural network statistical tools. In the current approach, a set of target properties are introduced as inputs in an inverse QSPR model to estimate the  $S_{\sigma\text{-profile}}(1-8)$  parameters, which are used as molecular descriptors of the solvent IL behavior. Once that the  $S_{\sigma\text{-profile}}$  descriptor is calculated for a potential solvent with required properties, statistical tools based on Euclidean distances are applied to select, from a database, the most adequate combination of ionic components with the needed  $S_{\sigma\text{-profile}}(1-8)$  values. At the end, the solvent properties of the tailor-made IL are calculated by the quantum-chemical COSMO-RS method to verify the reliability of the design. In this work, the suitability of the developed IL design tool has been demonstrated by the proposal of ternary IL mixtures with enhanced solvent properties for the liquid-liquid (LL) extraction of toluene from *n*-heptane, as a reference of aliphatic/aromatic separation.

The current approach may contribute to the evolution of ILs because it considers first the target criteria of the required material and then the ILs are computationally designed by the proper selection of ions to deliver these properties. The inherent modular nature of ILs can be fruitfully handled by the additive  $S_{\sigma\text{-profile}}$  molecular descriptor to progress in the optimization of IL behavior for specific applications. In the near future, the proposed methodology will be extended to design ILs with additional physical, chemical, or biological properties.

#### Acknowledgment

The authors are grateful to the "Ministerio Ciencia e Innovación" and "Universidad Autónoma de Madrid-Comunidad de Madrid" for financial support (projects CTQ2006-04644 and CCG07-UAM/AMB-1791). José S. Torrecilla was supported by Ramón y Cajal research contract from the "Ministerio Ciencia e Innovación" in Spain. We are very grateful to "Centro de Computación Científica de la Universidad Autónoma de Madrid" for computational facilities.

**Supporting Information Available:** COSMO-RS parameters used in this study. This material is available free of charge via the Internet at <http://pubs.acs.org>.

#### Literature Cited

- Welton, T. Room-temperature ionic liquids. Solvents for synthesis and catalysis. *Chem. Rev.* **1999**, 99, 2071.
- Wasserscheid P.; Welton, T. *Ionic Liquids in Synthesis*; Wiley-VCH: Weinheim, 2003.
- Rogers, R. D.; Seddon, K. R. *Ionic Liquids as Green Solvents*; ACS Symposium Series 856, American Chemical Society, 2003.
- Rogers, R. D.; Seddon, K. R. *Ionic Liquids IIIB: Fundamentals, Properties; Challenges and Opportunities*, ACS Symposium Series 902, American Chemical Society, 2005.
- Han, X.; Armstrong, D. W. Ionic liquids in separations. *Acc. Chem. Res.* **2007**, 40, 1079.
- Rogers, R. D.; Voth, G. A. Ionic liquids. *Acc. Chem. Res.* **2007**, 40, 1077.
- Palomar, J.; Torrecilla, J. S.; Ferro, V. R.; Rodríguez, F. Development of an a priori ionic liquid design tool. I: Integration of a novel COSMO-RS molecular descriptor on neural networks. *Ind. Eng. Chem. Res.* **2008**, 47, 4523.
- Jork, C.; Kristen, C.; Pieraccini, D.; Stark, A.; Chiappe, C.; Beste, Y. A.; Arlt, W. Tailor-made ionic liquids. *J. Chem. Thermodyn.* **2005**, 37, 537.
- Lei, Z.; Arlt, W.; Wasserscheid, P. Separation of 1-hexene and *n*-hexane with ionic liquids. *Fluid Phase Equilib.* **2006**, 241, 290.
- Meindersma, G. W.; Galán Sánchez, L. M.; Hansmeier, A. R.; Haan, A. B. Application of task-specific ionic liquids for intensified separations. *Monats. Chem.* **2007**, 138, 1125.
- Lei, Z.; Arlt, W.; Wasserscheid, P. Selection of entrainers in the 1-hexene/*n*-hexane system with a limited solubility. *Fluid Phase Equilib.* **2007**, 260, 29.
- Lei, Z.; Chen, B.; Li, C. COSMO-RS modeling on the extraction of stimulant drugs from urine sample by the double actions of supercritical carbon dioxide and ionic liquid. *Chem. Eng. Sci.* **2007**, 62, 3940.
- Matsuda, H.; Yamamoto, H.; Kurihara, K.; Tochigi, K. Computer-aided reverse design for ionic liquids by QSPR using descriptors of group contribution type for ionic conductivities and viscosities. *Fluid Phase Equilib.* **2007**, 261, 434.
- Klamt, A. *COSMO-RS: From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*, 1st Edition; Elsevier: Amsterdam, 2005.
- Palomar, J.; Ferro, V. R.; Torrecilla, J. S.; Rodríguez, F. Density and molar volume predictions using COSMO-RS for ionic liquids. An approach to solvent design. *Ind. Eng. Chem. Res.* **2007**, 46, 6041.
- Kato, R.; Gmehling, J. Systems with ionic liquids: Measurement of VLE and  $\gamma^\infty$  data and prediction of their thermodynamic behavior using original UNIFAC, mod. UNIFAC(Do) and COSMO-RS(OL). *J. Chem. Thermodyn.* **2005**, 37, 603.
- Diedenhofen, M.; Eckert, F.; Klamt, A. Prediction of infinite dilution activity coefficients of organic compounds in ionic liquids using COSMO-RS. *J. Chem. Eng. Data* **2003**, 3, 475.
- Marsh, K. N.; Boxall, J. A.; Lichtenthaler, R. Room temperature ionic liquids and their mixtures - a review. *Fluid Phase Equilib.* **2004**, 1, 93.
- Domanska, U.; Pobudkowska, A.; Eckert, F. Liquid-liquid equilibria in the binary systems (1,3-dimethylimidazolium, or 1-butyl-3-methylimidazolium methyl sulfate + hydrocarbons). *Green Chem.* **2006**, 3, 268.
- Domanska, U.; Pobudkowska, A.; Eckert, F. (Liquid+liquid) phase equilibria of 1-alkyl-3-methylimidazolium methyl sulfate with alcohols, or ethers, or ketones. *J. Chem. Thermodyn.* **2006**, 38, 685.
- Sahandzheva, K.; Tuma, D.; Breyer, S.; Pérez-Salado Kamps, A.; Maurer, J. Liquid-liquid equilibrium in mixtures of the ionic liquid 1-*n*-butyl-3-methylimidazolium hexafluorophosphate and an alkanol. *J. Chem. Eng. Data* **2006**, 51, 1516.
- Banerjee, T.; Singh, M. K.; Khanna, A. Prediction of binary VLE for imidazolium based ionic liquid systems using COSMO-RS. *Ind. Eng. Chem. Res.* **2006**, 45, 3207.
- Banerjee, T.; Khanna, A. Infinite dilution activity coefficients for trihexyltetradecyl phosphonium ionic liquids: Measurements and COSMO-RS prediction. *J. Chem. Eng. Data* **2006**, 51, 2170.
- Freire, M. G.; Santos, L. M. N. B. F.; Marrucho, I. M.; Coutinho, J. A. P. Evaluation of COSMO-RS for the prediction of LLE and VLE of alcohols + ionic liquids. *Fluid Phase Equilib.* **2007**, 255, 167.
- Lei, Z.; Arlt, W.; Wasserscheid, P. Selection of entrainers in the 1-hexene/*n*-hexane system with a limited solubility. *Fluid Phase Equilibria* **2007**, 260, 29.
- Diedenhofen, M.; Klamt, A.; Marsh, K.; Schäfer, A. Prediction of the vapor pressure and vaporization enthalpy of 1-*n*-alkyl-3-methylimidazolium-bis-(trifluoromethanesulfonyl) amide ionic liquids. *Phys. Chem. Chem. Phys.* **2007**, 9, 4653.
- Freire, M. G.; Carvalho, P. J.; Gardas, R. L.; Marrucho, I. M.; Santos, L. M. N. B. F.; Coutinho, J. A. P. Mutual solubilities of water and the [C(n)mim][Tf(2)N] hydrophobic ionic liquids. *J. Phys. Chem. B* **2008**, 112, 1604.
- Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.;

Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian03*, revision B.05; Gaussian, Inc.: Wallingford, CT, 2004.

(29) Perdew, J. P. Density-functional approximation for the correlation physics of the inhomogeneous electron gas. *Phys. Rev. B* **1986**, *33*, 8822.

(30) Sosa, C. J.; Andzelm, B. C.; Elkin, E.; Wimmer, K.; Dobbs, D.; Dixon, D. A. A local density functional study of the structure and vibrational frequencies of molecular transition-metal compounds. *J. Phys. Chem.* **1992**, *96*, 6630.

(31) Schaefer, A.; Huber, C.; Ahlrichs, R. Fully optimized contracted Gaussian basis sets of triple- $\zeta$  valence quality for atoms Li to Kr. *J. Chem. Phys.* **1994**, *100*, 5829.

(32) *COSMOtherm C2.1* Release 01.06; GmbH&CoKG: Leverkusen, Germany, 2003; <http://www.cosmologic.de>.

(33) Torrecilla, J. S.; Fernández, A.; García, J.; Rodríguez, F. Determination of 1-ethyl-3-methylimidazolium ethylsulfate ionic liquid and toluene concentration in aqueous solutions by artificial neural network/UV spectroscopy. *Ind. Eng. Chem. Res.* **2007**, *46*, 3787.

(34) Demuth, H.; Beale, M.; Hagan, M. *Neural Network Toolbox for Use with MATLAB® User's Guide*, version 4.0.6, ninth printing revised for version 4.0.6 (release 14SP3); 2005.

(35) Palancar, M. C.; Aragón, J. M.; Torrecilla, J. S. pH-Control system based on artificial neural networks. *Ind. Eng. Chem. Res.* **1998**, *37*, 7–2729.

(36) Demuth, H.; Beale, M.; Hagan, M. *Neural Network Toolbox for Use with MATLAB® User's Guide*, version 5, ninth printing revised for version 5.1 (release 2007b); 2007 (online only).

(37) *NIST/S EMATECH e-Handbook of Statistical Methods*, <http://www.itl.nist.gov/div898/handbook/>, 2008.

(38) Meindersma, G. W.; Podt, A. J. G.; de Haan, A. B. Selection of ionic liquids for the extraction of aromatic hydrocarbons from aromatic/aliphatic mixtures. *Fuel Process. Technol.* **2005**, *5887*, 59.

(39) N. Deenadayalu, N.; Ngongo, K. C.; Letcher, T. M.; Ramjugernath, D. Liquid–liquid equilibria for ternary mixtures (an ionic liquid plus benzene plus heptane or hexadecane) at  $T = 298.2$  K and atmospheric pressure. *J. Chem. Eng. Data* **2006**, *51*, 988.

(40) Meindersma, G. W.; Podt, A. J. G.; de Haan, A. B. Ternary liquid–liquid equilibria for mixtures of toluene + *n*-heptane + an ionic liquid. *Fluid Phase Equilib.* **2006**, *247*, 58.

(41) Arce, A.; Earle, M. J.; Rodríguez, H.; Seddon, K. R. Separation of aromatic hydrocarbons from alkanes using the ionic liquid 1-ethyl-3-methylimidazolium bis{(trifluoromethyl) sulfonyl} amide. *Green Chem.* **2007**, *9*, 70.

(42) Meindersma, G. W.; Galán Sánchez, L. M.; Hansmeier, A. R.; Haan, A. B. Application of task-specific ionic liquids for intensified separations. *Monatsh. Chem.* **2007**, *138*, 1125.

(43) Smiglak, M.; Metlen, A.; Rogers, R. D. The second evolution of ionic liquids: From solvents and separations to advanced materials-energetic examples from the ionic liquid cookbook. *Acc. Chem. Res.* **2007**, *40*, 1182.

Received for review June 17, 2008

Revised manuscript received November 18, 2008

Accepted December 5, 2008

IE8009507