Development of an a Priori Ionic Liquid Design Tool. 1. Integration of a Novel COSMO−RS Molecular Descriptor on Neural Networks

ARTICLE *in* INDUSTRIAL & ENGINEERING CHEMISTRY RESEARCH · JUNE 2008

Impact Factor: 2.59 · DOI: 10.1021/ie800056q

4 AUTHORS, INCLUDING:

Jose S. Torrecilla
Complutense University of Madrid
127 PUBLICATIONS **1,818** CITATIONS

SEE PROFILE

Development of an a Priori Ionic Liquid Design Tool. 1. Integration of a Novel COSMO-RS Molecular Descriptor on Neural Networks

José Palomar,*,† José S. Torrecilla,‡ Víctor R. Ferro,† and Francisco Rodríguez‡

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

An innovative computational approach is proposed to design ionic liquids (ILs) based on a new a priori molecular descriptor of ILs, derived from quantum-chemical COSMO-RS methodology. In this work, the charge distribution on the polarity scale given by COSMO-RS is used to characterize the chemical nature of both the cations and anions of the IL structures, using simple molecular models in the calculations. As a result, a novel a priori quantum-chemical parameter, $S_{\sigma\text{-profile}}$, is defined for 45 imidazolium-based ILs, as a quantitative numerical indicator of their electronic structures and molecular sizes. Subsequently, neural networks (NNs) are successfully applied to establish a relationship between the electronic information given by the $S_{\sigma\text{-profile}}$ molecular descriptor and the density properties of IL solvents. As a consequence, we develop here an a priori computational tool for screening ILs with required properties, using COSMO-RS predictions to NN design and optimization. Current methodology is validated following a classical quantitative structure−property relationship scheme, which is the main aim of this work. However, a second part of the current investigation will be devoted to a more useful design strategy, which introduces the desired IL properties as input into inverse NN, resulting in selections of counterions as output, i.e., directly designing ILs on the computer.

1. Introduction

The room temperature ionic liquids (ILs) have gained popularity in the last years as suitable green solvents because of their unique properties: negligible vapor pressure, high thermal and chemical stabilities, a wide liquid-state range, and high solvent capacity.[1,2] Thus, the application of these novel solvents to organic synthesis, catalysis, electrochemistry, biocatalysis, material science, and separation processes is being extensively investigated.[3–5] 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, the practical application to industrial processes of the large number of possible ILs is still limited by the scarce available experimental data. A tremendous amount of work is being carried out in the experimental determination of the thermophysical properties of ILs and their mixtures. In this sense, the recently presented IUPAC IL database (ILThermo) provides up-to-date information on IL publications with more than 40 000 total data points at the moment.[6] At this stage of development of ILs, the application of predictive models to estimate their thermophysical properties is of great interest.[7] Indeed, theoretical applications to predict IL properties have also highlighted growth in the last years, ranging from physically founded methods such as (i) ab initio quantum mechanics (QM), which have been successfully applied to predict vibrational and NMR spectra of ILs, the thermodynamics of ion-pair formation, molecular structures, and hydrogen-bonding interactions between counterions,[8–14] (ii) molecular dynamics (MD), which, using atomistic force fields, have been shown as an excellent tool for predicting reliable thermodynamic and transport properties of IL systems, including crystal structure, melting point, heat capacity, solubility, viscosity, diffusivity, and electric conductivity,[15–19] and (iii) conductor-like screening model for real solvents (COSMO-RS), which has demonstrated a general suitability for calculating density, vapor pressure, activity coefficients and phase equilibria data of IL systems,[20–33] to quantitative structure−property relationship (QSPR) models, using statistical analysis to model IL properties (such as melting point) based on different molecular descriptors (constitutional, topological, geometrical, electrostatic, quantum mechanical, and thermodynamic);[34–40] to classical thermodynamic models, such as NRTL, UNIQUAC, or UNIFAC.[20,27,29,41–43] A basic volumetric property such as the density of solvents may be used as an example of the state-of-art of the IL simulation field. To our knowledge, calculations of IL volumetric properties have recently been performed by a variety of methodologies. (i) QM: Banerjee et al.[43] applied a polarizable continuum model to derive accurately the volume and surface of the isolated molecules of imidazolium-based ILs. In addition, QM methods have been widely used to calculate molecular parameters of IL structures, which were applied to MD, QSPR, and other computational approaches. (ii) MD: Densities of ILs have been simulated using MD,[19,44–46] in general to validate new force fields developed specifically for IL systems, given an accuracy of about 1−5%. (iii) COSMO-RS: Recently, we presented accurate specific density and molar liquid volume calculations of 45 imidazolium-based ILs using the COSMO-RS method (error among 0.1−8%).[33] (iv) QSPR methods: Multiple regression analysis has been used to perform QSPR modeling of the density of a short number of ILs, using quantum-chemical-derived descriptors, with high reliability (error <4%).[36] On the other hand, densities of ILs have been estimated using the critical properties of the compound, previously determined by group contribution methods.[48] Other approximations reported estimated accurate and rapid densities of ILs by group additivity using volume parameters developed for salts and ILs.[49] In summary, it is now clear that simulations are an important complement to experiments to investigate the behavior of these commonly called "green solvents". It is

* To whom correspondence should be addressed. Tel: 34 914 976 938. Fax: 34 914 973 516. E-mail: pepe.palomar@uam.es.
† Universidad Autónoma de Madrid.
‡ Universidad Complutense de Madrid.

10.1021/ie800056q CCC: $40.75
Published on Web 06/04/2008

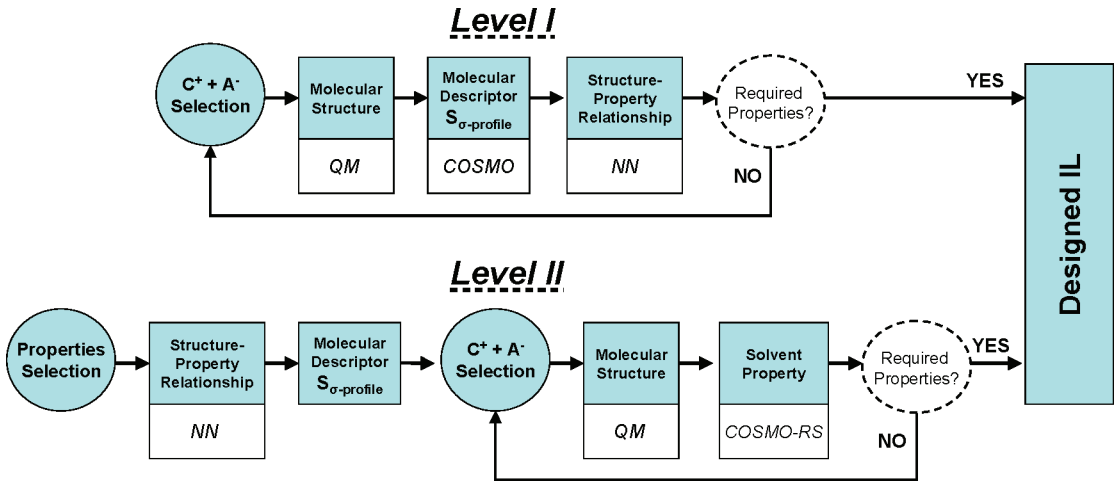

Figure 1. IL design: General scheme.

noteworthy, however, that most computational studies have been performed on a few ILs consisting of the most frequent cation (C^+) and anion (A^-) combinations in order to (i) validate the theoretical method by comparison between calculated and experimental properties (i.e., what C^+A^- property?) and (ii) understand the properties of IL with respect to their underlying molecular structure (i.e., why C^+A^- property?). However, at the moment, it has not yet been proven possible to predict which ion combinations will lead to a given set of properties (i.e., which C^+ and A^- selection?), even including new systems that have not yet been synthesized or measured. Such a predictive method would avoid reliance on the search of new useful ILs on chemical intuition or extrapolation from known compounds.

In this work, we present the first results of the development of a new a priori mathematical tool to determine ILs with appropriate characteristics. In the general scheme labeled as Level II in Figure 1, we adopt an innovative strategy to directly design ILs on the computer, starting from the desirable properties. In this proposal, neural networks (NNs) are applied to establish a relationship between quantum-chemical electronic information of IL structures and their predicted properties but with the aim of obtaining possible C^+A^- combinations with required properties. In others words, the computational approach schemed in Level II introduces the desired IL properties as input, resulting in the selection of counterions as output. Clearly, adequate information must be initially defined for performing such an analysis reliably: (a) IL property values to NN design and optimization and (b) molecular descriptors characteristic of IL systems to be modeled. For this purpose, Level I of Figure 1 here is established as a classical QSPR scheme to validate each step of the proposed computational approach. Concretely, we report in this work the results obtained in Level I for screening ILs with a specific density. In Level I, the required property value is predicted by a NN mathematical algorithm. A NN is a statistical tool with the capability of relating the input and output parameters. It learns from examples through iterations, without requiring a prior knowledge of the relationships between the process parameters.[50] Its structure is relatively simple, with connections in parallel and sequence between neurons. This means a short computing time and a high potential of robustness and adaptive performance. In the current computational approach, a NN will be tested as a structure−property relationship statistical tool. To our knowledge, the NN capability to model physicochemical properties of ILs based on nonlinear dynamics has been at the moment scarcely analyzed.[40,51] Because our objective is to develop an a priori IL design tool to overcome the insufficient available experimental data, we propose to use the COSMO-RS method[52] for predicting the IL properties needed for NN design. We recently showed the capability of the COSMO-RS method to predict accurate specific densities of a representative series of imidazolium-based ILs.[33] In addition, we suggested new interesting applications of the COSMO-RS methodology. Thus, this method provides an easy visualization of the charge distribution (σ) on the molecular surface of the compounds through the histogram function σ-profile, whose information is lately converted into standard thermodynamic data of the mixture. In other words, the σ-profile includes the main chemical information necessary to predict the possible electrostatic, hydrogen-bonding, and dispersion interactions of the compound in a fluid. Using this point of view, we evidenced that the charge distribution area ($S_{\sigma\text{-profile}}$) below σ-profiles of IL structures could be used to quantitatively characterize the chemical nature of both the cation and anion.[33] Consequently, one main contribution of this work is the proposal of using this COSMO-RS parameter [$S_{\sigma\text{-profile}}$] as a novel simple quantitative molecular descriptor for IL solvents, to be applied on a structure−property analysis with a NN algorithm for solvent properties of ILs, such density in this case.

The paper is organized following the successive steps necessary in Level I to find an IL with an appropriate density value. In the first section, two different molecular models are applied to simulate 45 imidazolium-based ILs, in order to analyze their characteristic σ-profiles: (i) the C+A model considers independent counterions of IL and (ii) the CA model consists of considering the ion-paired structure to simulate the IL compound. The second section is devoted to the definition and quantification of the a priori $S_{\sigma\text{-profile}}$ molecular descriptor to characterize the chemical nature of the counterions forming the IL species. Some preliminary validations of $S_{\sigma\text{-profile}}$ as a molecular descriptor of ILs will be performed by comparison to the experimental molar volume of the compounds. Subsequently, we adopt a computational procedure based on COSMO-RS calculations to predict the solvent properties in the current IL design tool. In the next section, NNs will be designed and optimized to estimate the density value as a function of $S_{\sigma\text{-profile}}$ molecular descriptors, using for this purpose the information corresponding to the database of those 45 ILs. The NN topology consists of eight input nodes ($S_{\sigma\text{-profile}}$) and one output neuron for the estimated density property at 298 K. Depending on the learning sample used (CA and C+A models), two NNs are evaluated with both criteria, the predictive capability of the computational approaches and their computational time cost.

Table 1. Density Values for Imidazolium-Based IL Compounds at 298 K Predicted by Linear Regression (Eq 5) between the Available Experimental Data and the Calculated COSMO-RS Values[a]

	ρ (kg/m^3)				
	Mmim	Emim	Bmim	Hxmim	Omim
Cl^-	1258 (solid)	1196 (1186)	1120 (1080)	1053 (1040)	1011 (1007)
Br^-	1681 (solid)	1561 (solid)	1389	1298	1217
BF_4^-	1277 (solid)	1224 (1270)	1149 (1204)	1096 (1148)	1050 (1103)
$EtSO_4^-$	1307	1251 (1234)	1181	1126	1087
$MeSO_4^-$	1358 (1328)	1298	1206 (1211)	1148	1113
$CF_3SO_3^-$	1459	1383 (1384)	1291 (1301)	1204	1154
PF_6^-	1472 (solid)	1399 (solid)	1294 (1366)	1231 (1293)	1174 (1236)
$FeCl_4^-$	1580	1505	1410 (1380)	1329 (1333)	1270 (1280)
Tf_2N^-	1587 (1570)	1533 (1521)	1444 (1437)	1378 (1370)	1321 (1325)

[a] For comparison, experimental data in parentheses were collected from the ILthermo database.[6]

Finally, the suitability of the developed IL design tool will be validated by the prediction of density values corresponding to new ILs, not included in the initial NN design, but whose density has been reported in the bibliography. Those ILs imply the selection of new cations or new anions in a supposed screening performed in Level I of Figure 1 to obtain a desired density.

2. Computational Details

Molecular Structures. All molecular geometries (cationic, anionic, and ion-paired species of ILs) were optimized at the B3LYP/6-31++G** computational level in the ideal gas phase using the quantum-chemical *Gaussian03* package.[53] Vibrational frequency calculations were performed for each case to confirm the presence of an energy minimum.

σ-Profiles. *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 the BVP86/TZVP/DGA1 level of theory.[54–56] Subsequently, COSMO files were used as input in the *COSMOtherm*[57] statistical thermodynamic code to obtain the histogram function σ-profile of each species.

COSMO-RS Properties. *COSMOtherm* software is applied to calculate the specific density of the neutral ion-paired species studied at 298 K using as input the previously obtained COSMO files at the BVP86/TZVP/DGA1 level of theory. According to our chosen method, the functional, and the basis set, 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. This information is supplied in the Supporting Information.

The liquid density of pure IL compounds is computed by the current version of the *COSMOtherm* software, as it was described elsewhere.[33,58] The experimental data of pure ILs are available, in most of the cases, in the recently presented IUPAC IL database (ILThermo)[6] and were chosen on the basis of the lowest assigned uncertainty among the reported data.

NN Design. The NN consists of several neurons arranged in two or more layers. Each neuron receives information on all of the neurons of the previous layer. Each connection is controlled by a weight that modules the output of the neuron before inputting its numerical content into a neuron in the following layer. The neurons are information-processing units that are fundamental for the NN operation. The inputs of each neuron are added by activation function, and the result is transformed by a transfer function, which serves to limit the amplitude of the neuron output. When the NN parameters are adjusted, the NN is able to learn from its environment. The process where the weights are optimized is called the learning process.[50,59]

Depending on the way the neurons are connected, the activation function inside the neurons and the learning method used, several NN types can be defined. A multilayer perceptron (MLP) is used here. The MLP model is a feed-forward network and a supervised network. The training algorithm used to optimize the weights is based on a back-propagation (BP) algorithm.[50]

Because the NN estimates the density at 298 K of ILs using eight values of σ-profile values, the learning, verification, and validation samples consist of nine rows (one for each σ-profile and density of an IL) and as many columns as ILs tested. The only difference between the verification and learning samples is that the latter is composed of 90% of the data and the former of the remaining 10% of 45 ILs (Table 1). Taking into account that every datum of the verification sample should be interpolated within the learning range, the data were randomly distributed in both samples. Finally, the NN was validated using the validation samples; it consists of σ-profile and density at 298 K of four new ILs.

Topology, transfer, and training functions used and parameters of the NN model are optimized in the MLP design process. The MLP model used consists of three layers (input, hidden, and output), a topology widely used to treat several problems. It has as many input nodes as σ-profile used and an output neuron to estimate the density at 298 K. The hidden neuron number (HNN) should be fixed by optimization techniques. Given that the range of most σ-profiles and a sigmoid function are bounded between 0 and 1, the sigmoid function has been used as the MLP transfer function.[50,59]

The NN training is made by the application of the BP algorithm. The Bayesian Regularization (trainBR) training function was selected. The generalization power of trainBR is higher than those of other training functions and, in addition, it avoids the overfitting and overtraining when the small learning sample is used.[59] TrainBR minimizes a linear combination of squared errors (eq 1) and weights by the computation of the Hessian matrix of this combination.[59,60] Their parameters are the learning coefficient (Lc), learning coefficient decrease (Lcd), and learning coefficient increase (Lci). The Lc parameter is similar to "h" in Newton's method (often called the Newton−Raphson method). Lcd and Lci control the value of Lc depending on the MLP model performance. To avoid an overfitting of the NN model, the learning process was repeated while the verification mean square error (MSE), defined by eq 1, was decreased. A deep description of the calculation process is described in the literature.[61–64]

$$\mathrm{MSE} = \frac{1}{N}\sum_{k=1}^{N}(r_k - y_k)^2 \qquad (1)$$

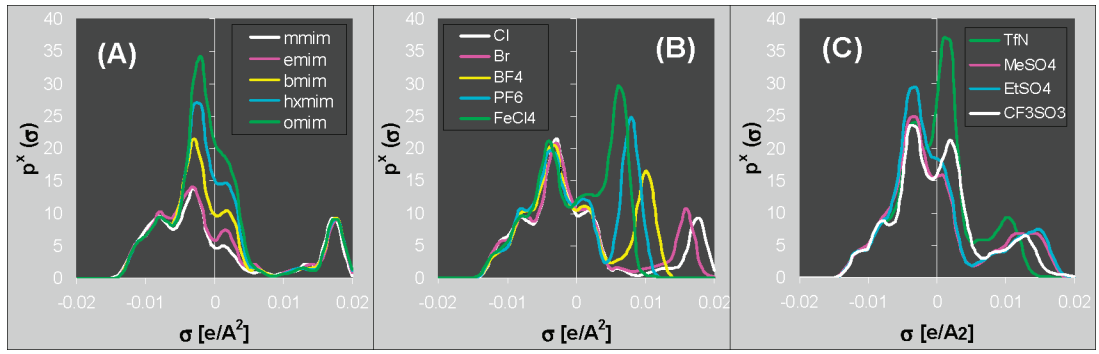

Figure 2. σ-Profile employed for benchmarking (A) 1-alkyl-3-methylimidazolium chloride compounds and (B and C) 1-butyl-3-methylimidazolium compounds with different anions, obtained using the CA model.

In eq 1, N, y_k, and r_k are respectively the number of data sets of the database ($N = 45$), the response of each output neuron, and the respective real output response.

The HNN and NN parameters are optimized by an experimental design based on the Box−Wilson Central Composite Design 2^4 + Star Points, commonly called a central composite design. This design contains an imbedded factorial design with center points that is augmented with a group of star points that allows the estimation of curvature. The number of experiments is quantified by $2n+ 2(n + 1)$, where n represents the number of factors.[65,66] The experimental factors analyzed were Lc, Lcd (between 1 and 0.001), and Lci (between 2 and 100).[67] Taking the learning sample size into account, the HNN range was selected (between 5 and 20).[68] The responses of the experimental design were the mean prediction error (MPE; eq 2) and the correlation coefficient (predicted vs real values, R^2). Both indexes are easily computed and provide a good description of the predictive performance of the network.[69]

$$\mathrm{MPE} = \frac{1}{N}\sum_{k=1}^{N}\frac{|r_k - y_k|}{r_k} \times 100 \qquad (2)$$

Because the main goal is to have a NN that predicts results with the highest accuracy as possible, the considerations taking into account to analyze the experimental design were to obtain the least MPE with the highest values of the correlation coefficient and P-value (PV). PV is a measure of the evidence of the null hypothesis (both predicted and real series are equal, H0). The PV threshold was set at 0.05; if the value of PV obtained is greater than 0.05, the null hypothesis is fulfilled. The higher the PV, the more evidence of H0.[70]

3. Results

3.1. Molecular Structure and σ-Profile of ILs. The first step in the current development is constructing a proper molecular model to simulate ILs. The IL compounds selected for this analysis were generated by combining the cations 1-methyl- ($Mmim^+$), 1-ethyl- ($Emim^+$), 1-butyl- ($Bmim^+$), 1-hexyl- ($Hxmim^+$), and 1-octyl-3-methylimidazolium ($Omim^+$) with the anions chloride (Cl^-), bromide (Br^-), tetrafluoroborate (BF_4^-), tetrachloroferrate ($FeCl_4^-$), hexafluorophosphate (PF_6^-), methylsulfate ($MeSO_4^-$), trifluoromethanesulfonate ($CF_3SO_3^-$), ethylsulfate ($EtSO_4^-$), and bis(trifluoromethanesulfonyl)imide (Tf_2N^-). We have considered two different structural approximations for obtaining the σ-profile of ILs in this work:

I. CA Model. This molecular model consists of considering the ion-paired structure to simulate the IL compound. In this case, the σ-profile of the IL compound is obtained for the optimized molecule as a whole. The CA model has been extensively used in QM studies of ILs because it includes a simple representation of the interaction between the cation and anion of the compound. Thus, the CA model has been successfully applied in quantum-chemical calculations to predict IR/Raman and NMR spectra of ILs, the thermodynamics of IL formation, the location of the anion with respect to the cation in IL, and the existence of hydrogen-bonding interactions between counterions.[8−14] In addition, the CA model has been applied to reliably predict the volumetric properties, vapor pressure, and vapor−liquid equilibrium (VLE) data of IL systems using the COSMO-RS method.[30−33]

II. C+A Model. An alternative molecular model consists of considering the independent counterions. In this simpler model, the screening charge densities of the cation (C^+) and anion (A^-) are calculated separately and then combined into a single σ-profile to define the IL compound. The C+A model presents the advantage of being much less computationally time-consuming than that based on ion-paired structures (CA model). However, it is not able to describe the effect of counterion interactions on the σ-profile of IL compounds. This molecular model has been the most frequently used in the available applications of COSMO-RS to the study of ILs, given a general suitability to calculate activity coefficients and phase equilibria data (VLE and LLE) of these systems.[20−29]

As a starting point of our analysis, Figure 2A presents the σ-profile for the 1-alkyl-3-methylmidazolium chloride cation series obtained using the CA model. As can be observed, the σ-profile of the 1-alkyl-3-methylimidazolium chloride compounds presents a peak at ~0.018 e/Å^2, which corresponds to the negatively charged Cl^- anion. The polarity of this peak is almost not modified by the cation size. Because this peak is at the high-polarity region, the Cl^- anion can be considered as a hydrogen-bond segment. On the negative side, we observe some unresolved and low peaks at lower values than −0.01 e/Å^2. These peaks are not affected by the alkyl chain. They are related to the hydrogen atoms of the imidazolium ring; thereby they may contribute to hydrogen bonds as acceptors. Finally, the distribution of the charge densities around zero (-0.01 e/Å^2 < σ < $+0.01$ e/Å^2) corresponds to the nonpolar alkyl groups of the cation, being those for positive and negative signals assigned to carbon and hydrogen atoms, respectively. As Figure 2A manifests, the higher number of carbon atoms in the alkyl chain implies an increase of the distribution of the charge densities around the nonpolar area. Then, the σ-profiles for the XmimCl series obtained by the CA model indicate that the lengthening of the alkyl chain almost does not change the chemical nature of the compound, but it increases the amount of charge with this polarization on the molecular surface. In addition, we can analyze the σ-profile description of the anion series in Figure

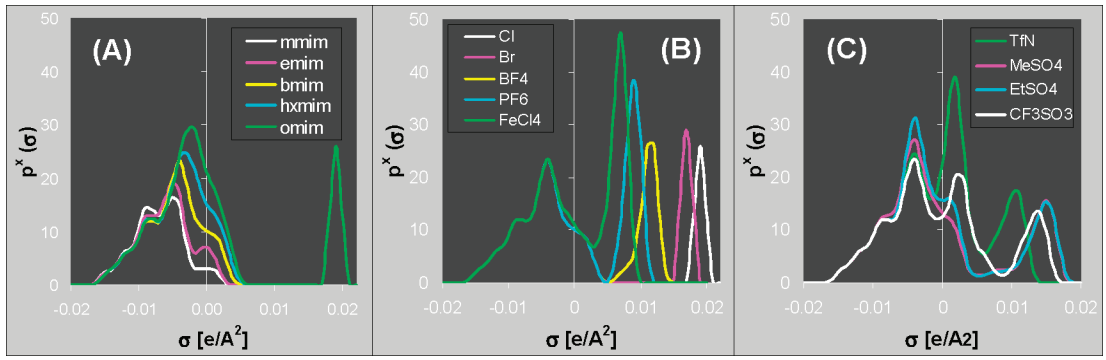

Figure 3. σ-Profile employed for benchmarking (A) 1-alkyl-3-methylimidazolium chloride compounds and (B and C) 1-butyl-3-methylimidazolium compounds with different anions, obtained using the C+A model.

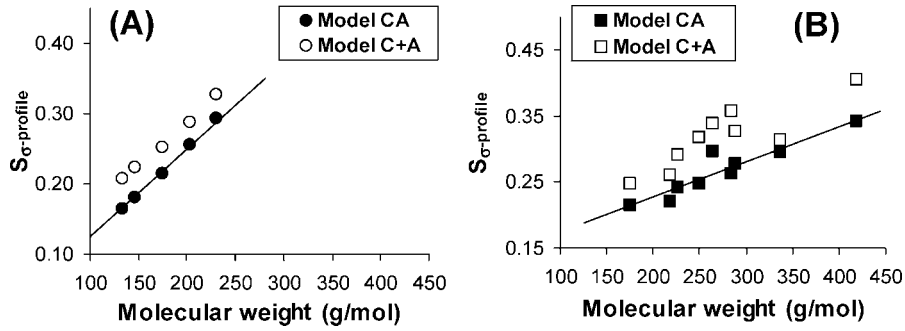

Figure 4. Comparison of the molecular weights of ILs with the values of the charge distribution area ($S_{\sigma\text{-profile}}$) below the σ-profile in the range ($-0.02 \leftrightarrow +0.02$ e/Å^2) for (A) the 1-alkyl-3-methylimidazolium chloride series and (B) 1-butyl-3-methylimidazolium compounds with Cl^-, BF_4^-, PF_6^-, $FeCl_4^-$, $MeSO_4^-$, $CF_3SO_3^-$, $EtSO_4^-$, and Tf_2N^- anions, obtained by the CA and C+A models.

2B, obtained by the CA model for the 1-butyl-3-methylimidazolium anion series with Cl^-, Br^-, BF_4^-, PF_6^-, and $FeCl_4^-$ groups. In this case, the distribution of the charge densities in the negative side—assignable to the 1-butyl-3-methylimidazolium cation—is very similar for all compounds, suggesting that the polarity of the cation fragment is scarcely modified by the anion effect. In contrast, the different chemical nature of the anion leads to different locations of its corresponding characteristic peak at the σ-profile. Thus, the peak location at the polarity range in the σ-profile manifests a hydrogen-bond acceptor capacity of each anion fragment in the order Cl^- (0.018 e/Å^2) > Br^- (0.016 e/Å^2) > BF_4^- (0.010 e/Å^2) > PF_6^- (0.008 e/Å^2) > $FeCl_4^-$ (0.006 e/Å^2). In the case of the 1-butyl-3-methylimidazolium anion series with Tf_2N^-, $MeSO_4^-$, $EtSO_4^-$, and $CF_3SO_3^-$ groups (Figure 2C), the σ-profiles evidence the presence of clearly electron-donor fragments with peaks located at the 0.01−0.02 e/Å^2 range but also the charge distribution in the nonpolar region (-0.01 e/Å^2 < σ < $+0.01$ e/Å^2) associated with carbon and hydrogen atoms. It is noteworthy that the peaks corresponding to different anionic groups also present very different intensities in the σ-profile.

On the other hand, Figure 3A presents the σ-profile for the 1-alkyl-3-methylmidazolium chloride cation series obtained using the C+A model. As can be observed, the locations and intensities of the peaks corresponding to the cation species are very similar to those obtained using the CA model. The main difference found is the higher amount of total polarizable charge in the independent cation. In contrast, the σ-profiles of the 1-butyl-3-methylimidazolium anion series obtained by the C+A model (Figures 3B,C) show some significant variations. For example, Figure 3B indicates a similar polarity anion order for the series of Cl^-, Br^-, BF_4^-, PF_6^-, and $FeCl_4^-$ groups but with the location of the characteristic peaks corresponding to the anion located at higher polarity values in the C+A model case and with remarkable higher intensities. These results indicate the stronger hydrogen-acceptor capacity of the anion in the absence of the cation, which is a logical finding because the optimized ion-paired structures generally present a hydrogen bond between counterions.

3.2. Molecular Descriptor for ILs. The histogram function σ-profile given by COSMO-RS methodology[52] includes the main chemical information necessary to predict the possible interactions of one compound in a fluid. Therefore, it seems reasonable to establish the hypothesis of using the σ-profile of IL structures as a numerical molecular descriptor in structure−property analysis.

In the following, we investigate the possibility of applying the σ-profile data to quantitatively characterize the solvent behavior of a wide series of 45 imidazolium-based ILs. As Figures 2A and 3A manifest, the higher number of carbon atoms in the alkyl chain of the cation implies an increase of the distribution of the charge densities around the nonpolar area. Because the total electronic density of a molecule is directly related to its molecular weight, we will analyze whether the probabilistic surface charge distribution ($S_{\sigma\text{-profile}}$) area below the σ-profile can be used as a quantitative parameter of the imidazolium alkyl size. Very interestingly, Figure 4A shows a linear relationship between the molecular weight of the compound and $S_{\sigma\text{-profile}}$ values calculated in the total -0.02 e/Å^2 < σ < $+0.02$ e/Å^2 range [$S_{\sigma\text{-profile}}$(total)] of the σ-profile for the 1-alkyl-3-methylmidazolium chloride (XmimCl) cation series given by the CA model. It seems that the total amount of polar charge in the σ-profile can as well be used as a quantitative parameter to describe the anion contribution to the molecular weight. In fact, Figure 4B also shows a linear relationship between the molecular weight and the $S_{\sigma\text{-profile}}$(total) values of the σ-profile for the 1-butyl-3-methylmidazolium anion series

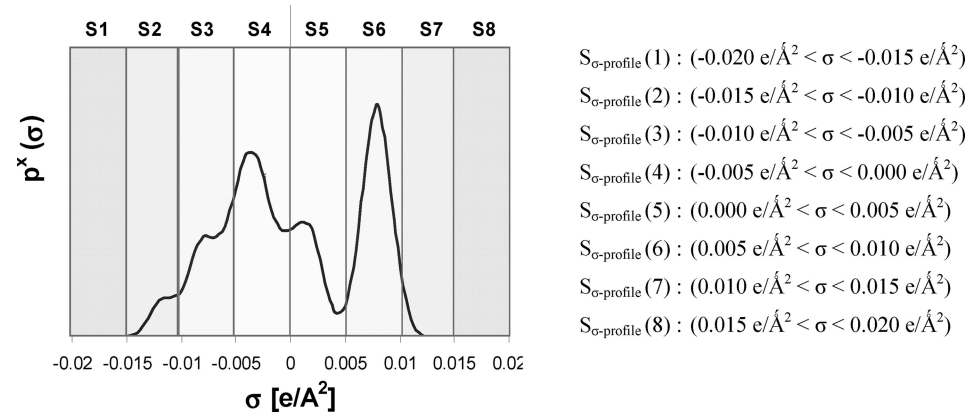

Figure 5. Partition perfomed on the polarity scale of the σ-profile to define the eigth $S_{\sigma\text{-profile}}(1-8)$ parameters evaluated as possible molecular descriptor values for IL compounds.

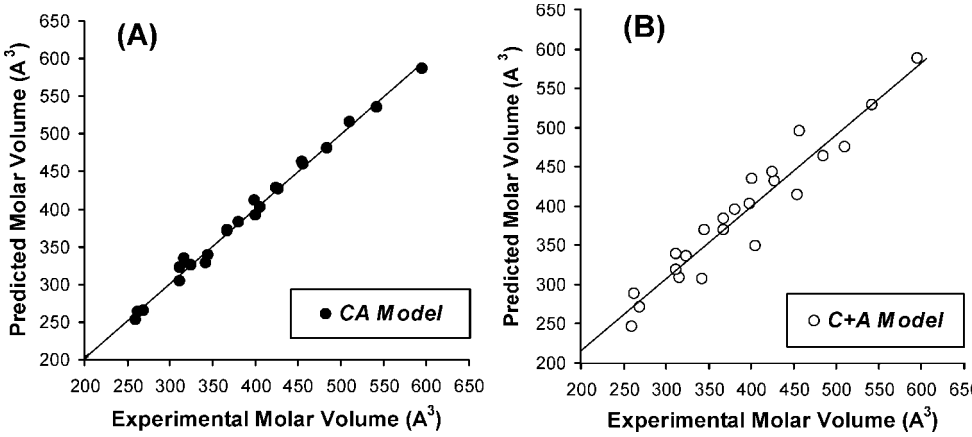

Figure 6. Comparison of experimental molar volumes to those estimated by a multilinear regression using the eight values of $S_{\sigma\text{-profile}}(1-8)$ obtained by (A) the CA model (eq 3) and (B) the C+A model (eq 4).

given by the CA model, including the nine different anions of the study. Clearly, linear trends of Figure 4A,B present different slopes, which indicates that the $S_{\sigma\text{-profile}}$(total) value is differently related to the molecular weight as a function of the nature of the functional groups included, i.e., the location of the electronic density on the polarity σ scale. On the other hand, Figure 4A also shows a good linear relationship between the molecular weight and the $S_{\sigma\text{-profile}}$(total) values calculated using the C+A model for the cation series. The main difference found is the slightly higher value of the intercept in the C+A model, which is related to the higher amount of total polarizable charge in the independent cations. Furthermore, Figure 4B clearly shows a clear trend between the molecular weights and the $S_{\sigma\text{-profile}}$(total) values calculated for the 1-butyl-3-methylmidazolium anion series given by the C+A model, even when this molecular approach provides a stronger dispersion of the data than the CA model. Therefore, we observe that the $S_{\sigma\text{-profile}}$ parameter, which quantifies the charge distribution area below the σ-profile in a range of polarity, can be used as a quantitative indicator of the total electronic density of the compound, if the complete σ range is included in the calculation.

On the basis of these results, we propose to use $S_{\sigma\text{-profile}}$ as a unique a priori parameter to characterize the different chemical contributions of the cation and anion to the IL compound. For this purpose, we will consider eight different regions in the σ-profile to evaluate the charge distribution area, $S_{\sigma\text{-profile}}(1-8)$, as is shown in Figure 5. The $S_{\sigma\text{-profile}}(1-8)$ values of the 45 ILs are provided in the Supporting Information. In order to perform a preliminary validation of our hypothesis, the $S_{\sigma\text{-profile}}(1-8)$ values given by the CA model will be used to describe a fluid property of pure ILs as the experimental molar volume ($\tilde{V}$). The volumetric properties of pure IL are an appropriate selection for this analysis because they must depend on the size of the constituent counterions and also the physicochemical intermolecular interactions in IL media must play a determinant role on them. Therefore, the molar liquid volume ($\tilde{V}$) of 24 imidazolium-based ILs (derived from the experimentally measured density)[6] has been correlated by multilinear regression to $S_{\sigma\text{-profile}}(1-8)$ parameters estimated in the eight σ-profile intervals given by the CA model, obtaining the following expression:

$$\tilde{V}\,(\text{Å}^3) = -52.8 + 892S_{\sigma-\text{profile}}(4) + 2198S_{\sigma-\text{profile}}(5) + 2824S_{\sigma-\text{profile}}(6) + 3630S_{\sigma-\text{profile}}(7) + 4605S_{\sigma-\text{profile}}(8) \quad (3)$$

$$R^2 = 0.99$$

Then, we find that the experimentally derived molar volumes of 24 ILs are remarkably well correlated by multilinear regression to five values of the $S_{\sigma\text{-profile}}$ parameter defined in the range $-0.010 \leftrightarrow +0.02$ e/Å^2 of the σ-profile [$S_{\sigma\text{-profile}}(4-8)$]. In this case, the three $S_{\sigma\text{-profile}}(1)$, $S_{\sigma\text{-profile}}(2)$, and $S_{\sigma\text{-profile}}(3)$ parameters do not have statistical relevance for the series of 24 imidazolium-based ILs studied because their values are almost equal for all of them. In order to easily visualize the descriptive capacity of the $S_{\sigma\text{-profile}}$ parameters given by the CA model, Figure 6A shows the good agreement between experimental molar volumes and those estimated using eq 3, for which the root-mean-square deviation (rmsd) of the 24 ILs is lower than 1.6%. In addition, we have performed a multilinear analysis to

relate the experimentally derived molar liquid volume ($\tilde{V}$) of 24 imidazolium-based ILs with the $S_{\sigma\text{-profile}}(1-8)$ parameters estimated using the C+A model for these compounds. The results of this fit are summarized by the following expression:

$$\tilde{V}(\text{Å}^3) = 314 + 2363S_{\sigma-\text{profile}}(4) + 590S_{\sigma-\text{profile}}(5) + 1338S_{\sigma-\text{profile}}(6) + 2512S_{\sigma-\text{profile}}(7) + 5486S_{\sigma-\text{profile}}(8) \quad (4)$$

$$R^2 = 0.90$$

Clearly, the poorer correlation of eq 4 results in a stronger dispersion of the data in Figure 6B, in which are presented the experimental molar volumes versus those predicted by eq 4 using the $S_{\sigma\text{-profile}}$ values given by the C+A model for those imidazolium-based ILs whose experimental density is available in the bibliography.[6] In any case, we also obtain a reasonable prediction of the molar volume using the simple $S_{\sigma\text{-profile}}(1-8)$ parameters calculated with the C+A model, with a value of the rmsd of 5.1% for a datasheet of 24 ILs.

In summary, our results suggest that the charge distribution area, $S_{\sigma\text{-profile}}$, estimated in eight σ-polarity intervals of the σ-profile (given by the COSMO-RS method) is an adequate molecular descriptor for ILs, defined with the requirements of being (i) a unique a priori quantum-chemical parameter, (ii) able to qualitatively characterize the chemical nature of a huge variety of cations and anions, and (iii) a quantitative numerical indicator of the electronic structure and molecular size of the compound. In addition, the $S_{\sigma\text{-profile}}$ descriptor presents additional advantages because the σ-profile histogram of the IL compound is covered in a restricted scale of polarity (σ), which can always be improved by more realistic molecular models of ILs and whose calculation is very fast and simple, once the molecular model is optimized.

3.3. Solvent Property Prediction for ILs. The construction of Level I in the general scheme shown in Figure 1 requires the accessibility of sufficient property data of ILs to NN design. Because it should lead to an a priori computational procedure, we propose a thermodynamic COSMO-RS model based on QM calculations to predict the required solvent properties. The COSMO-RS method has been shown to be able to generally describe the structural effects associated with the cation and anion selection, providing a very good prediction of the experimental density of imidazolium-based ILs by COSMO-RS, with a value of the rmsd of 3.0% and a next linear relationship with a high correlation coefficient ($R^2 = 0.99$):[33]

$$\rho = -222 + 1.14\rho_{\text{COSMO-RS}} \quad (5)$$

Table 1 collects the density values predicted using eq 5 for 45 IL compounds, together with the available experimental data for comparison. The rmsd of the 24 ILs with experimentally measured density is again lower than 1.7%. These density values of ILs derived from COSMO-RS calculations (Table 1) will be used in the structure−property analysis performed by a NN in the next section.

3.4. Structure−Property Relationships. The following step consists of implementing the $S_{\sigma\text{-profile}}$ as a numerical molecular descriptor of ILs in the NN design and optimization, as the final phase in the development of the design IL tool schemed in Level I of Figure 1.

In this work, depending on the learning sample used (CA and C+A models), two NNs ($NN_{CA\ model}$ and $NN_{C+A\ model}$, respectively) are developed. Sigmoid and TrainBR transfer and training functions were respectively used in both NNs. The input layer of the MLP model consists of eight input nodes [$S_{\sigma\text{-profile}}(1-8)$] and an output neuron for estimating the density at 298 K. In both NNs, HNNs, Lc, Lcd, and Lci were optimized following the calculation process and taking the considerations described in the NN Design section into account. Using the verification sample, the optimized HNNs, Lc, Lcd, and Lci values were tested by calculating the statistical parameters (Table 2). In view of these statistical results, both $NN_{CA\ model}$ and $NN_{C+A\ model}$ provide good estimations of the IL densities based on the corresponding $S_{\sigma\text{-profile}}$ descriptors, which are derived solely from QM calculations. However, $NN_{CA\ model}$ estimates slightly more accurately the density values calculated by COSMO-RS with a rmsd of 0.5%, compared to that the error value of $NN_{C+A\ model}$ (1.0%). The estimations of $NN_{CA\ model}$ and $NN_{C+A\ model}$ versus those predicted by COSMO-RS for the verification sample are shown in Figure 7. Comparing their slopes, the closest to unity is also taken by $NN_{CA\ model}$. Therefore, taking the statistical results into account, the $NN_{CA\ model}$ is the most adequate NN model to estimate the density values of this series of ILs because the CA model introduces the interaction between counterions of IL in the QM calculation. 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 that to make the CA model database.

Table 2. Parameters of the NNs and Final Statistical Results Using the Verification Samples

optimized parameters of the NNs	$NN_{CA\ model}$	$NN_{C+A\ model}$
HNN	5	8
Lc	0.834	0.400
Lcd	0.004	0.200
Lci	2	75
StatisticalResults		
MPE (%)	0.8	2
R^2	0.995	0.936
PV	0.752	0.526

It should be indicated that current computational approaches of Level I are not dependent on the experimental data because measured density values were only introduced to improve the COSMO-RS predictions. As a consequence, there is no limitation with respect to the scale and diversity of the data that can be used for the NN learning.

3.5. Application of NN to a Real Case. Finally, the suitability of the developed IL design tool of Level I will be validated by its application to new ILs, not included in the initial IL series used for NN design. Those ILs imply the selection of two new cations [the 1-nonyl-3-methylimidazolium ($Nmim^+$)

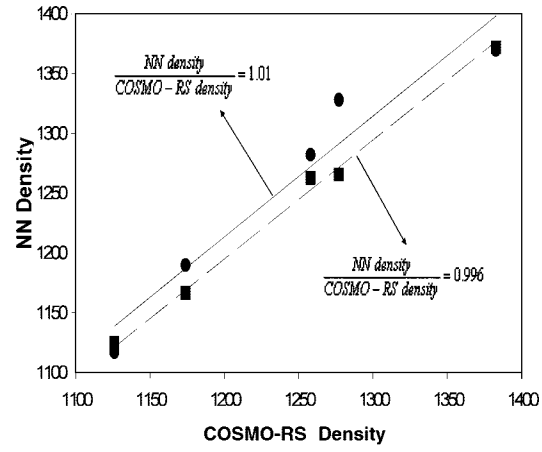

Figure 7. NN model performance (■ and ● are density values from CA and C+A models, respectively; - - - and ─ lines represent the corresponding linear correlations of CA and C+A models, respectively).

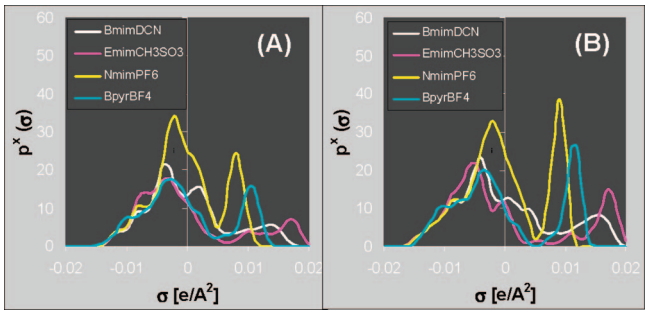

Figure 8. σ-Profiles obtained for the new ILs used in the validation test, obtained by (A) the CA model and (B) the C+A model.

Table 3. Experimental and Estimated Molar Volumes from the Eight Values of $S_{\sigma\text{-profile}}$(1−8) Obtained by CA (Eq 3] and C+A [Eq 4] Models for the New ILs Used in Validation Tests

	V (Å^3)		
	experimental	CA model	C+A model
BmimDCN	322	345	309
$EmimMeSO_3$	275	277	176
$NmimPF_6$	503	490	528
$BpyrBF_4$	305	303	320

cation, which is an extension of the 1-alkyl-3-methylimidazolium series, and 1-butylpyridinium ($Bpyr^+$), which is a completely different cationic species] and two new anions [dicyanamide (DCN^-), a frequently used anion with a clearly different nature, and methanesulfonate ($CH_3SO_3^-$), which implies a variation with respect to $CF_3SO_3^-$] combined to form four new compounds: BmimDCN, $EmimCH_3SO_3$, $NmimPF_6$, and $BpyrBF_4$. Following the scheme of Level I, the validation first involves performing the QM optimizations of the new IL structures by CA and C+A models, which determine the σ-profiles characteristic of the compounds. Figure 8 shows some significant variations in the distribution of the polar charge of these new ILs with respect to those shown in Figures 2 and 3, given respectively by CA and C+A models. For example, the peak related to oxygen atoms of $CH_3SO_3^-$ is located at a clearly more basic position than that of $CF_3SO_3^-$, whereas the peak of nitrogen atoms of DCN^- appears among those of the halide atoms of Br^- and BF_4^-. In addition, the σ-profile of $Bpyr^+$ is similar than those of imidazolium cations, but with the presence of more electronic charge in the acidic region, whereas in $Nmim^+$, the addition of one methyl group to the alkyl substituent is described by an increase of the peak intensity in the nonpolar area. As a consequence of these differences, the estimation of $S_{\sigma\text{-profile}}$ parameters for these new IL compounds provides new volume values from eqs 3 and 4, which are compared to those obtained from experimental data in Table 3. The rmsds obtained for the four new ILs are 4% and 10% using the CA and C+A models, respectively. Finally, in order to validate the prediction capability of the optimized NNs, a new validation datum was made using the new series of ILs, formed by the corresponding eight values of $S_{\sigma\text{-profile}}$, for each molecular model. The experimental density and their estimations by COSMO-RS (eq 5), $NN_{CA\ model}$, and $NN_{C+A\ model}$ are shown in Table 4. Taking the prediction carried out by COSMO-RS into account, using $NN_{CA\ model}$ in the verification and validation samples, the MPE values are similar (<0.8% and <0.85%, respectively), and using $NN_{C+A\ model}$, the MPE value slightly increases from 2 to 3%. In any case, the proposed computational approaches, based on the $S_{\sigma\text{-profile}}$ molecular descriptor integrated on NN statistical tools, are able to predict reasonable density values of new ILs, as is shown in Table 4.

Table 4. Experimental Density and Their Estimations by COSMO-RS (Eq 5) and NN Models Using the New ILs as Validation Samples

	ρ (kg/m^3)			
	experimental	COSMO-RS	$NN_{CA\ model}$	$NN_{C+A\ model}$
Bmim DCN	1058	1067	1094	1125
Emim $MeSO_3$	1244	1299	1291	1275
Nmim PF_6	1202	1194	1191	1220
Bpyr BF_4	1214	1212	1212	1240

4. Conclusions

An innovative computational approach has been proposed to design ILs directly on the computer, using COSMO-RS methodology and NN statistical tools. Concretely, in this work we propose and validate the molecular descriptors, the solvent property prediction method, and the NN design procedure needed to develop a current IL design tool. Thus, the distribution of the polar electronic charge on the polarity σ scale of a histogram function σ-profile given by COSMO-RS was proposed to characterize the chemical nature of both the cation and anion of the IL structures. As a result, a new molecular descriptor, $S_{\sigma\text{-profile}}$, is defined and estimated in eight σ-polarity intervals of the σ-profile for 45 imidazolium-based ILs. The $S_{\sigma\text{-profile}}$ descriptor, an a priori quantum-chemical parameter, was shown to be capable of describing qualitatively and quantitatively the electronic structure and molecular size of a huge variety of counterion combinations, by comparisons to experimental molecular weights and molar volumes of compounds. Once the $S_{\sigma\text{-profile}}$ molecular descriptor was validated, their values for 45 ILs were successfully related by NNs to the density values of these solvents, predicted from COSMO-RS calculations. The performance of this computational approach was demonstrated by density estimations of new ILs, using their $S_{\sigma\text{-profile}}$ descriptor in NN analysis. In the current study, two very simple molecular models for obtaining $S_{\sigma\text{-profile}}$ were validated. When ion-paired structures (CA model) are used, a more reliable description of the intermolecular interactions in a pure IL fluid is obtained. However, a model of independent ions (C+A model) presents the clear advantage of a reasonable reliability at much less computational cost.

This work has been performed following a classical QSPR procedure. The second part of the current investigation will report the application of inverse NNs to introduce the required properties as input, resulting in the $S_{\sigma\text{-profile}}$ molecular descriptor as output, i.e., giving the way to design directly ILs on the computer.

Acknowledgment

The authors are grateful to the "Ministerio de Educación y Ciencia" and "Universidad Autónoma de Madrid-Comunidad de Madrid" for financial support (Projects CTQ2006-04644 and CCG06-UAM/AMB-0494). J.S.T. was supported by a Ramón y Cajal research contract from the "Ministerio de Educación y Ciencia" in Spain. We are very grateful to Centro de Computación Científica de la Universidad Autónoma de Madrid for computational facilities.

Supporting Information Available: Tables of COSMO-RS parameters and values of COSMO-RS molecular descriptors. This material is available free of charge via the Internet at http://pubs.acs.org.

Literature Cited

(1) Welton, T. Room-temperature ionic liquids. Solvents for synthesis and catalysis. *Chem. Rev.* **1999**, *99*, 2071.

(2) Sheldon, K. R. Catalytic reactions in ionic liquids. *Chem. Commun.* **2001**, *23*, 2399.

(3) Wasserscheid, P.; Welton, T. *Ionic Liquids in Synthesis*; Wiley-VCH: Weinheim, Germany, 2003.

(4) Rogers, R. D.; Seddon, K. R. *Ionic Liquids as Green Solvents*; ACS Symposium Series 856; American Chemical Society: Washington, DC, 2003.

(5) Rogers, R. D.; Seddon, K. R. *Ionic Liquids IIIB: Fundamentals, Properties, Challenges and Opportunities*; ACS Symposium Series 902; American Chemical Society: Washington, DC, 2005.

(6) http://ilthermo.boulder.nist.gov.

(7) Rogers, R. D.; Voth, G. A. Ionic liquids. *Acc. Chem. Res.* **2007**, *40*, 1077.

(8) Katsyuba, S. A.; Dyson, P. J.; Vandyukova, E. E.; Chernova, A. V.; Vidis, A. Application of density functional theory and Vibrational spectroscopy toward the rational design of ionic liquids. *J. Phys. Chem. A* **2007**, *111*, 352.

(9) Berg, R. W.; Deetlefs, M.; Seddon, K. R.; Shim, I.; Thompson, J. M. Raman and ab initio studies of simple and binary 1-alkyl-3-methylimidazolium ionic liquids. *J. Phys. Chem. B* **2005**, *109*, 19018.

(10) Tsuzuki, S.; Tokuda, H.; Hayamizu, K.; Watanabe, M. Magnitude and directionality of interaction in ion pairs of ionic liquids: Relationship with ionic conductivity. *J. Phys. Chem. B* **2005**, *109*, 16474.

(11) Gutowski, K. E.; Holbrey, J. D.; Rogers, R. D.; Dixon, D. A. Prediction of the formation and stabilities of energetic salts and ionic liquids based on ab initio electronic structure calculations. *J. Phys. Chem. A* **2005**, *109*, 23196.

(12) Meng, Z.; Dölle, A.; Carper, W. R. Gas phase model of an ionic liquid. *J. Mol. Struct.* **2002**, *585*, 119.

(13) Dong, K.; Zhang, S.; Wang, D.; Yao, X. Hydrogen Bonds in Imidazolium Ionic Liquids. *J. Phys. Chem. A* **2006**, *110*, 9775.

(14) Palomar, J.; Ferro, V. R.; Gilarranz, M. A.; Rodriguez, J. J. A Computational Approach to Nuclear Magnetic Resonance in 1-Alkyl-3-Methylimidazolium Ionic Liquids. *J. Phys. Chem. A* **2007**, *111*, 168.

(15) Padua, A. A. H.; Costa Gomes, M. F.; Canongia Lopes, J. N. A. Molecular Solutes in Ionic Liquids: A Structural Perspective. *Acc. Chem. Res.* **2007**, *40*, 1087.

(16) Shim, Y.; Jeong, D.; Maniari, S.; Choi, M. Y.; Kim, H. J. Solvation, Solute Rotation and Vibration Relaxation, and Electron-Transfer Reactions in Room-Temperature Ionic Liquids. *Acc. Chem. Res.* **2007**, *40*, 1130.

(17) Lynden-Bell, R. M.; Del Popolo, M. G.; Youngs, T. G. A.; Kohanoff, J.; Hanke, C. G.; Harper, J. B.; Pinilla, C. C. Simulations of Ionic Liquids, Solutions, and Surfaces. *Acc. Chem. Res.* **2007**, *40*, 1138.

(18) Hardacre, C.; Holbrey, J. D.; Nieuwenhuyzen, M.; Youngs, T. G. A. Structure and Solvation in Ionic Liquids. *Acc. Chem. Res.* **2007**, *40*, 1146.

(19) Maggin, E. J. Atomistic Simulation of the Thermodynamic and Transport Properties of Ionic Liquids. *Acc. Chem. Res.* **2007**, *40*, 1200.

(20) Kato, R.; Gmehling, J. Systems with ionic liquids: Measurement of VLE and γ^{∞} data and prediction of their thermodynamic behavior using original UNIFAC, mod. UNIFAC(Do) and COSMO-RS(Ol). *J. Chem. Thermodyn.* **2005**, *37*, 603.

(21) 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.

(22) Marsh, K. N.; Boxall, J. A.; Lichtenthaler, R. Room temperature ionic liquids and their mixtures—a review. *Fluid Phase Equilib.* **2004**, *1*, 93.

(23) 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.

(24) 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.

(25) 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.

(26) Lei, Z.; Arlt, W.; Wasserscheid, P. Separation of 1-hexene and *n*-hexane with ionic liquids. *Fluid Phase Equilib.* **2006**, *241*, 290.

(27) Sahandzhieva, 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.

(28) 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.

(29) 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.

(30) 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.

(31) 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.

(32) Diedenhofen, M.; Klamt, A.; Marsh, K.; Schäfere, 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.

(33) 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.

(34) Katritzky, A. R.; Lomaka, A.; Petrukhin, R.; Jain, R.; Karelson, M.; Visser, A. E.; Rogers, R. D. QSPR Correlation of the Melting Point for Pyridinium Bromides, Potential Ionic Liquids. *J. Chem. Inf. Comput. Sci.* **2002**, *42* (1), 71.

(35) Eike, D.; Brennecke, J.; Maginn, E. Predicting Melting Points of Quaternary Ammonium Ionic Liquids. *Green Chem.* **2003**, *5*, 323.

(36) Trohalaki, S.; Pachter, R.; Drake, G.; Hawkins, T. Quantitative Structure−Property Relationships for Melting Points and Densities of Ionic Liquids. *Energy Fuels* **2005**, *19*, 279.

(37) Carrera, G.; Aires-de-Sousa, J. Estimation of Melting Points of Pyridinium Bromide Ionic Liquids with Decision Trees and Neural Networks. *Green Chem.* **2005**, *7*, 20.

(38) Sun, N.; He, X.; Dong, K.; Zhang, X.; Lu, X.; He, H.; Zhang, S. Prediction of the melting points for two kinds of room temperature ionic liquids. *Fluid Phase Equilib.* **2006**, *246*, 137.

(39) López-Martin, I.; Burello, E.; Davey, P. N.; Seddon, K. R.; Rothenberg, G. Anion and Cation Effects on Imidazolium Salt Melting Points: A Descriptor Modelling Study. *Chem. Phys. Chem.* **2007**, *8*, 690.

(40) Varnek, A.; Kireeva, N.; Tetko, I. V.; Baskin, I. I.; Solovev, V. P. Exhaustive QSPR Studies of a Large Diverse Set of Ionic Liquids: How Accurately Can We Predict Melting Points? *J. Chem. Inf. Comput. Sci.* **2007**, *47*, 1111.

(41) Belveze, L. S.; Brennecke, J. F.; Stadtherr, M. A. Modeling of Activity Coefficients of Aqueous Solutions of Quaternary Ammonium Salts with the Electrolyte-NRTL Equation. *Ind. Eng. Chem. Res.* **2004**, *43* (3), 815.

(42) Nebig, S.; Bolts, R.; Gmehling, J. Measurement of vapor−liquid equilibria (VLE) and excess enthalpies (H−F) of binary systems with 1-alkyl-3-methylimidazolium bis(trifluoromethylsulfonyl)imide and prediction of these properties and γ^{∞} using modified UNIFAC (Dortmund). *Fluid Phase Equilib.* **2006**, *258*, 168.

(43) Banerjee, T.; Singh, M. K.; Sahoo, R. K.; Khanna, A. Volume, surface and UNIQUAC interaction parameters for imidazolium based ionic liquids via Polarizable Continuum Model. *Fluid Phase Equilib.* **2005**, *234*, 64.

(44) Urahata, S. M.; Ribeiro, M. C. C. Structure of ionic liquids of 1-alkyl-3-methylimidazolium cations: A systematic computer simulation study. *J. Chem. Phys.* **2004**, *120*, 1855.

(45) Canongia Lopes, J. N.; Deschamps, J.; Pádua, A. A. H. Modeling Ionic Liquids of the 1-Alkyl-3-methylimidazolium Family Using an All-Atom Force Field. Ionic Liquids IIIB: Fundamentals, Progress, Challenges, and Opportunities; Transformations and Processes. *ACS Symp. Ser.* **2005**, *901*, 134.

(46) Liu, Z.; Wu, X.; Wang, W. A novel united-atom force field for imidazolium-based ionic liquids. *Phys. Chem. Chem. Phys.* **2006**, *9*, 1096.

(47) Deetlefs, M.; Seddon, K. R.; Shara, M. Predicting physical properties of ionic liquids. *Phys. Chem. Chem. Phys.* **2006**, *8*, 642.

(48) Valderrama, J. O.; Robles, P. A. Critical properties, normal boiling temperaturas and acentric factors of fifty ionic liquids. *Ind. Eng. Chem. Res.* **2007**, *46*, 1338.

(49) Ye, C.; Shreeve, J. M. Rapid and accurate estimation of densities of room temperature ionic liquids and salts. *J. Phys. Chem. A* **2007**, *111*, 1456.

(50) Maren, A. J.; Harston, C. T.; Pap, R. M. *Handbook of neural computing applications*; Academic Press, Inc.: San Diego, CA, 1990.

(51) Torrecilla, J. S.; Fernandez, A.; Garcia, 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.

(52) Klamt, A. *COSMO-RS: From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*, 1st ed.; Elsevier: Amsterdam, The Netherlands, 2005.

(53) 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.

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

(55) 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.

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

(57) *COSMOtherm C2.1*, release 01.06; GmbH&CoKG: Leverkusen, Germany, 2003; http://www.cosmologic.de.

(58) *COSMOtherm Users Manual*, version C2.1 Release 01.06.

(59) Demuth, H.; Mark, B.; Hagan, M. T. *Neural Network Toolbox for Use with MATLAB; Neural Network Toolbox User's Guide*; The MathWorks Inc.: Natick, MA, 2005.

(60) Hirschen, K.; Schäfer, M. Bayesian regularization neural networks for optimizing fluid flow processes. *Comput. Methods Appl. Mech. Eng.* **2006**, *195*, 481.

(61) Torrecilla, J. S.; Aragón, J. M.; Palancar, M. C. Modeling the Drying of a High-Moisture Solid with an Artificial Neural Network. *Ind. Eng. Chem. Res.* **2005**, *44*, 8057.

(62) Palancar, M. C.; Aragon, J. M.; Torrecilla, J. S. pH-Control System Based on Artificial Neural Networks. *Ind. Eng. Chem. Res.* **1998**, *37*, 2729.

(63) Torrecilla, J. S.; Otero, L.; Sanz, P. D. Artificial neural networks: a promising tool to design and optimize high-pressure food processes. *J. Food Eng.* **2005**, *69*, 299.

(64) Torrecilla, J. S.; Mena, M. L.; Yáñez-Sedeño, P.; García, J. Quantification of Phenolic Compounds in Olive Oil Mill Wastewater by Artificial Neural Network/Laccase Biosensor. *J. Agric. Food Chem.* **2007**, *55*, 7418.

(65) Wsól, V.; Fell, A. F. Central composite design as a powerful optimisation technique for enantioresolution of the *rac*-11-dihydrooracin—the principal metabolite of the potential cytostatic drug oracin. *J. Biochem. Biophys. Methods* **2002**, *54*, 377.

(66) Lucchesi, M. E.; Smadja, J.; Bradshaw, S.; Louw, W.; Chemat, F. Solvent free microwave extraction of Elletaria cardamomum L.: A multivariate study of a new technique for the extraction of essential oil. *J. Food Eng.* **2007**, *79*, 1079.

(67) Vacic, V. Summary of the training functions in Matlab's NN toolbox. http://www.cs.ucr.edu/~vladimir/cs171/nn_summary.pdf, 2005.

(68) Sun, Y.; Peng, Y.; Chen, Y.; Shukla, A. J. Application of artificial neural networks in the design of controlled release drug delivery systems. *Adv. Drug Delivery Rev.* **2003**, *55*, 1201.

(69) Sheiner, L. B.; Beal, S. Some suggestions for measuring predictive performance. *J. Pharm. Biophys.* **1981**, *9*, 503.

(70) Collet, D. *Modelling binary data*, 2nd ed.; Chapman & Hall/CRC: London, U.K., 2003.

Received for review January 14, 2008
Revised manuscript received April 6, 2008
Accepted April 14, 2008

IE800056Q