Article

Predicting Thermal Decomposition Temperature of Binary Imidazolium Ionic Liquid Mixtures from Molecular Structures

Hongpeng He, Yong Pan,* Jianwen Meng, Yongheng Li, Junhong Zhong, Weijia Duan, and Juncheng Jiang

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Ionic liquids (ILs) have been regarded as "designer solvents" because of their satisfactory physicochemical properties. The 5% onset decomposition temperature ($T_{d,5\%onset}$) is one of the most conservative but reliable indicators for characterizing the possible fire hazard of engineered ILs. This study is devoted to develop a quantitative structure−property relationship model for predicting the $T_{d,5\%onset}$ of binary imidazolium IL mixtures. Both in silico design and data analysis descriptors and norm index were employed to encode the structural characteristics of binary IL mixtures. The subset of optimal descriptors was screened by combining the genetic algorithm with the multiple linear regression method. The resulting optimal prediction model was a four-variable multiple linear equation, with the average absolute error

(AAE) for the external test set being 12.673 K. The results of rigorous model validations also demonstrated satisfactory model robustness and predictivity. The present study would provide a new reliable approach for predicting the thermal stability of binary IL mixtures.
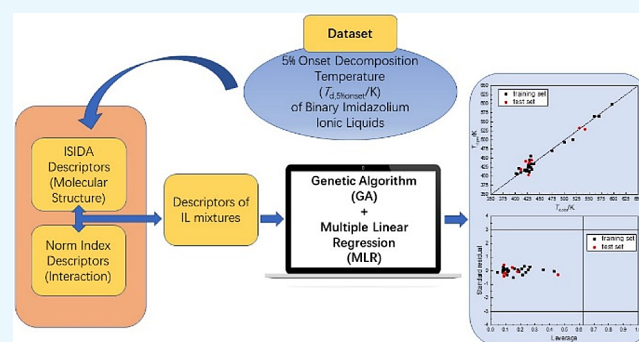
1. INTRODUCTION

By definition, ionic liquids (ILs) are novel green solvents composed of organic cations and organic or inorganic anions, which are liquid at room temperature. They are labeled as "designer solvents" because of their excellent physical and chemical properties such as high conductivity, low melting point, remarkable thermal stability, and good solubility.[1−3] ILs have achieved a rapid development and are widely applied in various fields such as catalysis,[4] carbon capture,[5] electrolytes in batteries,[6] and pharmaceutical processing.[7]

However, it is still difficult to find the ideal IL that satisfies the deeper demand of desired properties. Since each new IL requires a complete property analysis and registration, it can be costly and time-consuming to study and develop a new IL. A feasible and alternative solution is the IL mixtures. IL mixtures, which are also known as double salt ILs (DSILs),[8] make it possible to further fine-tune their properties. DSILs often exhibit different physicochemical properties from pure ILs since new ionic associations generate. By changing the combinations of ILs, the desired IL mixtures can be designed and synthesized for certain properties. A great deal of research has been published on the applications of IL mixtures in multiple fields, e.g., electrochemistry,[9,10] carbon dioxide capture,[11,12] catalysis,[13] and extraction.[14]

Since many industrial processes generate substantial heat or need to be carried out under high-temperature conditions, high

requirements are put forward for the thermal stability of ILs. Although the ILs are almost nonflammable, there still exists thermal decomposition. During high-temperature industrial processes, the ILs may change their structures and thermally decompose, which leads to unwanted byproducts and unexpected accidents.[15] Therefore, the thermal decomposition temperature (T_d) is commonly referred to characterize the thermal hazards of ILs.[16−19] Furthermore, compared to the traditional onset decomposition temperature ($T_{d,onset}$), the temperature at 5% decomposition ($T_{d,5\%onset}$) is considered as a more realistic and conservative indicator to determine the maximum permissible working temperature during the operational process.[20,21]

Extensive research on $T_{d,5\%onset}$ of IL mixtures has been performed.[12,22−24] All of them tested the thermal stability by experimental studies, which requires a lot of manpower and material resources. Furthermore, for those samples with poisonous and harmful substances, there remains difficulties and hazards in measurement. Currently, there is a trend that

Table 1. Descriptors Selected for the Present Model for the Prediction of IL Mixtures[a]

structure descriptor	representation	ionic type	mixing rule	interaction descriptor	representation
X_1	C($-$N$'$)	cation	$\lvert x_1 d_1 - x_2 d_2 \rvert$	X_4	$\mathrm{norm_{in1}}(\mathbf{CM}_{2,1},2)$
X_2	F($-$C$'$)	anion	$\sqrt{x_1}\,d_1 + \sqrt{x_2}\,d_2$	/	/
X_3	F($-$C$'$)	anion	$(1 - \Delta x)\Delta d$	/	/

[a]Here, N, C, and F represent the kinds of atoms and "$-$" represent the single bond, with those outside of brackets denoting augmented atoms and those in brackets denoting their neighbor atoms and bonds

the quantitative structure−property relationship (QSPR) method is widely employed to predict the physicochemical properties of chemicals. Compared with the experimental method, the QSPR is apparently a more convenient and efficient approach.

In the past few years, considerable efforts have been made to predict the T_d of pure ILs and IL composites. Lazzús established a theoretical prediction model for the T_d of 198 ILs based on a group contribution method.[25] The data set contains 27 cationic groups and 31 anionic groups, and the resulting model provided an alternative approach to predict the T_d of ILs. Venkatraman and Alsberg employed partial least squares and random forests (RF) to develop models for predicting T_d of 995 diverse ILs that consisted of 461 cations and 119 anions.[26] Electronic, thermodynamic, and geometrical descriptors calculated from semiempirical PM6 were selected, and the RF model showed better predictive performances. Zhao et al. established a QSPR model to describe T_d values for a total of 168 ILs.[27] They used the DRAGON program (version 6) to calculate descriptors and employed genetic algorithm with the multiple linear regression (GA-MLR) to filter the most relevant variables. Both the structure for single ions and the cation−anion interactions were considered, and the results showed satisfactory predictability. Zeeshan et al. combined 29 diverse imidazolium ILs with two diverse metal−organic frameworks (MOFs) and proposed QSPR models based on the MLR method to predict T_d values.[28] The results demonstrated satisfactory prediction of T_d values of IL/MOF composites, and they found that ILs with the $[\mathrm{NTf_2}]^-$ anion possess higher thermal stability in the composites. Recently, Duan et al. developed the QSPR model for 158 different imidazolium ILs via norm index descriptors.[29] The descriptors they employed not only characterize the structures of anions and cations but also describe the interaction between anions and cations.

However, with more research on IL mixtures, there is still no prediction model for the $T_{d,5\%\text{onset}}$ of IL mixtures because of the lack of experimental data. In this study, the data set was selected from previous publications, aiming to develop a simple and reliable model to predict the decomposition temperatures of imidazolium IL mixtures via the QSPR method. In this study, descriptors for both structural characteristics of ions and cation−anion interactions were calculated and combined together to represent the characteristics of the selected imidazolium IL mixtures. In particular, in silico design and data analysis (ISIDA), which represents the molecular structure based on a finite number of topological fragments,[30,31] was employed to explore the relationship between target properties and ionic structures. Besides, the norm index descriptor was utilized to describe the interaction between cations and anions.[29,32,33] The QSPR modeling was carried out in this study for a total of 31 different imidazolium IL mixtures, and various rigorous model validation strategies were then performed to ensure the model performance. Through our

work, it was expected to provide an efficient approach to evaluate the thermal stability of IL mixtures and provide some guidance for experimental design and mixing of ILs.

2. RESULTS AND DISCUSSION

2.1. Results of Prediction. The most optimal subset was selected after the process of GA-MLR. In this study, three structural descriptors with "augmented atoms" type "IIAB" (including a cation descriptor and two anion descriptors) and an interaction descriptor were obtained to encode the characteristics of imidazolium IL mixtures. To analyze the contribution of substructures, the representations of these selected descriptors are listed in Table 1. The optimal MLR model is shown as follows:

$$\text{Model:} \quad Y = -32.615X_1 + 23.264X_2 - 20.285X_3 + 11.392X_4 + 384.437$$

$$R^2 = 0.969,\ Q^2_{\mathrm{LOO}} = 0.949,\ \mathrm{SE} = 9.666,\ n = 24,\ F = 150.113$$

$$(1)$$

where R^2 represents the squared correlation coefficient, Q^2_{LOO} represents the value of leave-one-out (LOO) cross-validation (CV), SE represents the standard error of the model, n represents the number of mixtures in the training set, and F represents the Fischer F-ratio.

Then, the predictive ability of the model was tested by comparing the predicted and observed values, which is shown in Figure 1. The predicted values are presented in the Supporting Information (Table S1). The plot clearly depicts that the majority of data points are located in the diagonal line, which indicates a good correlation between the predicted and observed values of IL mixtures. Table 2 summarizes the main
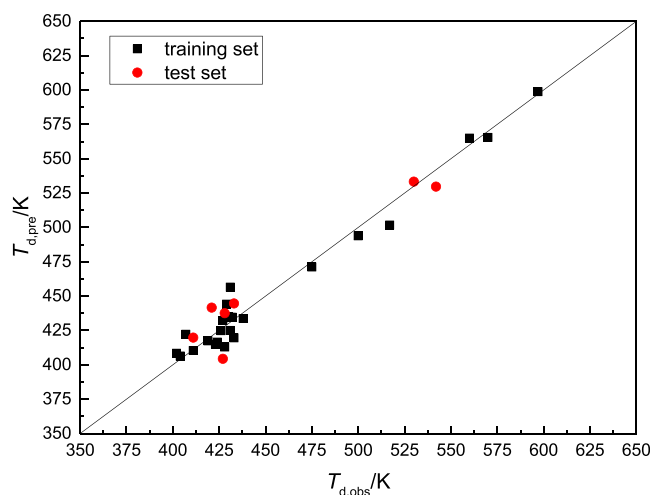

Figure 1. Plot of the predicted vs observed values of the $T_{d,5\%\text{onset}}$ of binary IL mixtures.

Table 2. Main Statistical Parameters of the Obtained MLR Model

statistical parameters	training set	test set
R^2	0.969	0.923
Q^2_{LOO}	0.949	
Q^2_{EXT}		0.924
RMSE	9.463	14.152
AAE	7.294 K	12.673 K
n	24	7

statistical parameters of the MLR model. The result shows that the $T_{d,5\%onset}$ values of IL mixtures can be effectively predicted by the model equation with an AAE of 7.294 K. In addition, the predicted percentage error of 31 samples is shown in Figure 2. The distribution showed the specific number of
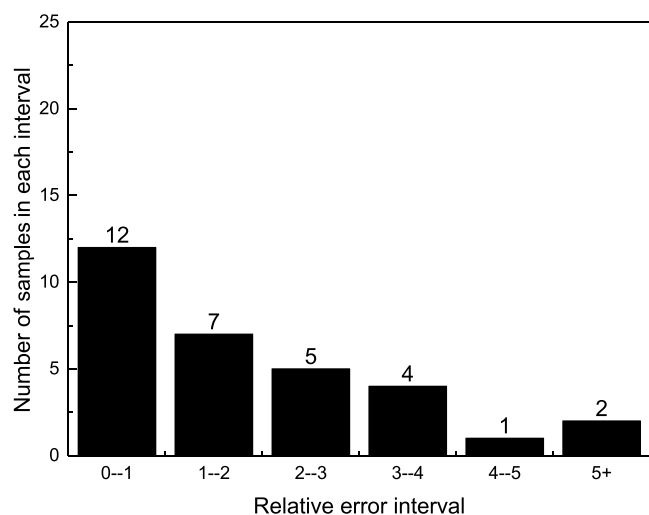

Figure 2. Percent errors obtained by the model and the number of mixtures in each range.

samples in each error interval. The obtained minimum relative error is 0.14%, the maximum relative error is 5.91%, and the average relative error is 1.93%.

2.2. Model Stability Validation and Result Analysis. For the present model, the Y-randomization test was employed as a validation approach to determine model stability. It was carried out 100 times for the training set samples in this paper. The visual summary of the Y-randomization test is performed by demonstrating the relationship between the R^2 values of randomized models and the model frequency (Figure 3). The obtained maximum, minimum, and average R^2 values of the randomized model are 0.465, 0.0153, and 0.168, respectively, while the value of standard deviation (SD) is 0.110. Since the difference of R^2 between the original model and the mean highest random is much greater than 3SD, the result indicates a negligible chance correlation between the randomly shuffled models and the original model, which emphasizes the unique connection between the developed model and the selected variable set.

Furthermore, the predicted residuals and observed values are analyzed in Figure 4. Evidently from the diagram, all predicted residuals are distributed on both sides of the zero baseline uniformly and randomly, which indicates that no systematic errors occurred during the establishment of the MLR model.
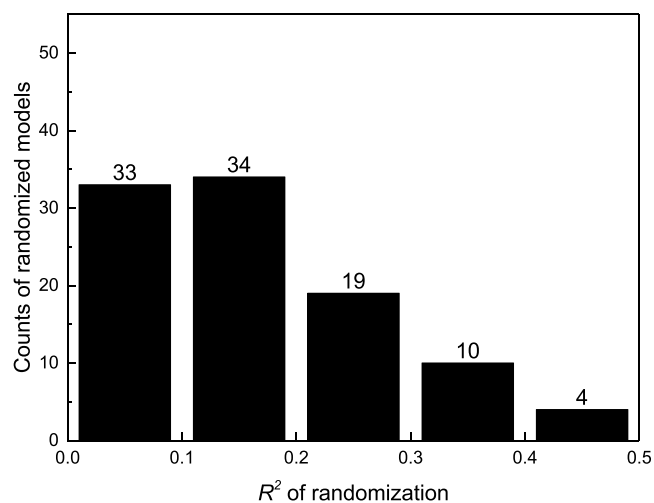
Figure 3. Histogram of R^2 of randomization vs frequency of occurrence of the randomized models.
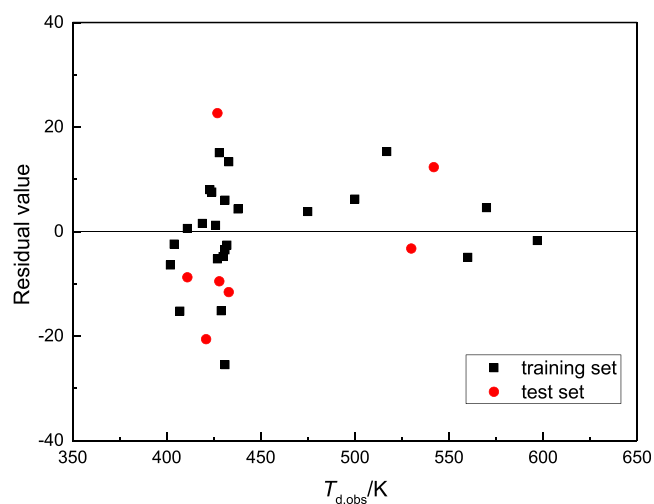
Figure 4. Plots of the residual vs the observed $T_{d,5\%onset}$ values of binary IL mixtures.

With the rigorous validation of multiple strategies, it can be concluded that the obtained model shows excellent robustness and predictability. The QSPR model derived from molecular structural descriptors can efficiently and conveniently predict the $T_{d,5\%onset}$ values of IL mixtures.

2.3. Applicability Domain of the Present Model. The domain application of the QSPR model is defined as a chemical structure space, which is characterized by the properties of the compounds in the training set. The Williams plot is employed to analyze the domain application and is shown in Figure 5. The applicability domain (AD) is depicted as a squared area within ±3 standard deviations and a leverage threshold h^* of 0.625.

As can be easily observed in Figure 5, all predicted values of IL mixtures are inside the area. Since all IL mixtures are within the residual range, the developed model is considered reliable and can satisfactorily predict $T_{d,5\%onset}$ of imidazolium IL mixtures within the application ranges.

3. CONCLUSIONS

In this work, the QSPR approach was employed to determine the relationship between the thermal stability and molecular
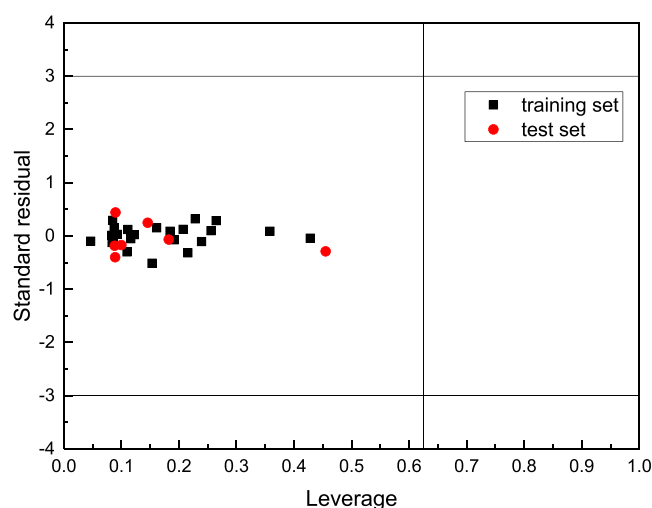
Figure 5. Williams plot describing the AD of the QSPR model for the $T_{d,5\%onset}$ of the binary IL mixtures ($h^* = 0.625$).

structures of binary imidazolium IL mixtures. After carrying out internal and external validations, the results showed that the predictive model has been successfully developed with satisfactory model performance. To sum up, this work has the following main findings:

(1) The relationship between the thermal stability of binary IL imidazolium mixtures and their molecular structures was investigated in this study, for the first time, and the corresponding prediction model was proposed to predict the $T_{d,5\%onset}$ of the IL mixtures.

(2) The norm index descriptors combined with ISIDA descriptors were employed to extensively and effectively characterize both molecular structure features and interactions of binary IL mixtures, which lead to the resulting QSPR model with satisfactory prediction performance.

(3) The developed model was conceptually simple, convenient for application, and would also be reasonably expected to reliably predict the decomposition temperature of binary imidazolium IL mixtures with satisfactory robustness and predictivity, from only their molecular structures. Additionally, this study could also provide some guidance for synthesizing or designing IL mixtures for safety purposes.

4. MATERIALS AND METHODS

4.1. Data Set. The observed 5% onset decomposition temperatures of binary imidazolium IL mixtures were selected from previous publications.[12,23] The data set consists of 31 data points containing 4 pure imidazolium ILs and 27 mixtures of binary imidazolium ILs. The $T_{d,5\%onset}$ values of these mixtures ranged from 402 to 597 K. The details of 31 binary IL mixtures are listed in Table S2 of the Supporting Information.

4.2. Calculation of Molecular Descriptors. *4.2.1. Calculation of Descriptors for Ions.* The ISIDA fragment descriptor, which is also realized as the substructure molecular fragment (SMF) descriptor, is defined as the number of topological fragments of the molecular structure. There are two main classes of molecular subgraphs: (I) "sequences" and (II) "augmented atoms." The "sequences" means the shortest path of two atoms connected by chemical bonds. There are three

types of "sequences," which are performed as atom types (A), bond types (B), and atom and bond types (AB). For each type of sequence, the minimal and maximal lengths of the path are defined as n_{min} and n_{max}, respectively. In this study, the length of the sequence corresponds to $n_{min} = 2$ and $n_{max} = 8$. The selected atoms with their environment are defined as "augmented atoms," which also includes neighbor atom and bond types (AB), atom types (A), or bond types (B).

Since ILs are composed of anions and cations, it is necessary to calculate descriptors for cations and anions separately. The detailed descriptor calculation process is performed as follows: first, the 2D chemical structures of each cation and anion were drawn using MarvinSketch (version 15.6.29.0) and optimized based on the clean in 2D method. Structure data files (SDF) including two cations and three anions were then, respectively, imported into EdiSDF, a program of the ISIDA system, to generate SDF records of cations and anions. The described SMF descriptors were generated by the SMF module of the ISIDA system.

The mixture descriptor for the IL mixture was developed by combining the descriptors of diverse cations and anions separately. The 12 proposed formulas that take into account their respective mole fractions were used to calculate mixed-cation descriptors and mixed-anion descriptors[34] (Table 3). Then, descriptors for constituted species were concatenated to describe the structure of the multiple-ion system.[35]

Table 3. List of Mixture Formulas[a]

no.	formula
(1)	$D = x_1 d_1 + x_2 d_2$
(2)	$D = |x_1 d_1 - x_2 d_2|$
(3)	$D = \sqrt{(x_1 d_1)^2 + (x_2 d_2)^2}$
(4)	$D = x_1^2 d_1 + x_2^2 d_2$
(5)	$D = \sqrt{x_1}\, d_1 + \sqrt{x_2}\, d_2$
(6)	$D = (x_1 d_1 + x_2 d_2)^2$
(7)	$D = (1 - \Delta x)\Delta d$
(8)	$D = (1 - \Delta x^2)\Delta d$
(9)	$D = (1 - \Delta x)^2 \Delta d$
(10)	$D = (d_1 + d_2)/2$
(11)	$D = (d_1 - d_2)^2$
(12)	$D = |d_1 - d_2|$

[a]D is the descriptor of mixture; x_1 and x_2 are the molar fractions of components 1 and 2; respectively; d_1 and d_2 are ISIDA descriptor values for components 1 and 2, respectively; Δx is the absolute value of the difference between x_1 and x_2; and Δd is the absolute value of the difference between d_1 and d_2.

4.2.2. Calculation of Interaction Descriptors. Since ILs consist of cations and anions, their thermal stability and the character of the cations and anions are intimately connected. Thus, not only the structural characteristics of ions should be taken into account, but also the contribution of interaction is indispensable. The molecular structure can be intuitively represented by the SMF module, while the interactions between the cations and anions cannot be represented because of the limitation of the software. Considering the diverse associations that possibly exist in a multiple-ion system of IL mixtures, the descriptors based on the norm index were proposed to describe each cation–anion interaction of imidazolium IL mixtures in this work.

First, the 3D structure of each cation and anion was achieved by Chemdraw (version 14), with the optimization based on the MM2 module (the class 1 Allinger molecular mechanics program). For further optimization, the Gaussian (version GaussView 6.0.16) was employed to carry out density functional theory (DFT) M06-2X functional calculation on the basis of 6-311+G (d,p). Then, the optimized molecular structures were characterized by the matrix norm index.[33]

The calculation process of distance matrices and property matrices is shown as follows:

$$\mathbf{MS} = [a_{ij}] \quad a_{ij} = \begin{cases} n \text{ if } i \neq j \\ 0 \text{ if } i = j \end{cases} \tag{2}$$

n is the path between atoms i and j.

$$\mathbf{MD} = [a_{ij}] \quad a_{ij} = \begin{cases} 1/r_{ij} \text{ if } i \neq j \\ 0 \text{ if } i = j \end{cases} \tag{3}$$

r_{ij} is the Euclidean spatial distance between atoms i and j.

$$\mathbf{MP}_1 = [\text{AW/Ra}] \tag{4}$$

$$\mathbf{MP}_2 = [\text{EN} \times \text{Ra}] \tag{5}$$

$$\mathbf{MP}_3 = [\exp(\text{AW/MW})] \tag{6}$$

$$\mathbf{MP}_4 = [\text{AQ/AW}] \tag{7}$$

$$\mathbf{MP}_5 = [1/(1 + \exp(\text{BD}))] \tag{8}$$

where AW, Ra, EN, MW, AQ, and BD are atom weight, atom radius, electronegativity, molecular weight, atom charge, and branching degree, respectively.

Then, the distance matrices and the property matrices are combined as **CM** matrices, which are calculated as follows:

$$\mathbf{CM}_{1,m} = \begin{bmatrix} \mathbf{MP}_m \\ \mathbf{MS} \end{bmatrix} \tag{9}$$

$$\mathbf{CM}_{2,m} = \begin{bmatrix} \mathbf{MP}_m \\ \mathbf{MD} \end{bmatrix} \tag{10}$$

$$\mathbf{CM}_{3,m} = [\mathbf{MP}_m^{\mathrm{T}} \times \mathbf{MP}_m + \mathbf{MS}] \tag{11}$$

$$\mathbf{CM}_{4,m} = [\mathbf{MP}_m^{\mathrm{T}} \times \mathbf{MP}_m + \mathbf{MD}] \tag{12}$$

where m = 1, 2, 3, 4, and 5, which correspond to matrices $\mathbf{MP}_1$ to $\mathbf{MP}_5$. Furthermore, the three norm indexes are defined in eqs 12−14:

$$\text{norm}(\mathbf{CM}, 1) = \max_j \left\{ \sum_{i=1}^{p} |\mathbf{CM}_{i,j}| \right\}, \quad j = 1, 2, ..., q \tag{13}$$

$$\text{norm}(\mathbf{CM}, 2) = \sqrt{(\max(\lambda_i(\mathbf{CM}^{\mathrm{H}} \times \mathbf{CM})))} \tag{14}$$

$$\text{norm}(\mathbf{CM}, 3) = \sqrt{\left(\sum_{j}^{q} \sum_{i}^{p} \mathbf{CM}_{ij}^2 \right)} \tag{15}$$

where p and q represent the number of rows and columns of matrix **CM**, respectively. λ_i is the eigenvalue of the matrix, and the $\mathbf{CM}^{\mathrm{H}}$ is defined as the Hermite matrix of the matrix.

Considering that the thermal stability is related to the various interactions of the binary IL system, the descriptors for multiple cation−anion interactions are further proposed based on the original calculation for single ILs,[29,33] which are defined as follows:

$$\text{norm}_{\text{in1}}(\mathbf{CM}, w)$$
$$= \sqrt{\text{CF} \times \text{norm}_{\text{Ca1}}(\mathbf{CM}, w) + \text{AF} \times \text{norm}_{\text{An1}}(\mathbf{CM}, w)} \tag{16}$$

$$\text{norm}_{\text{in2}}(\mathbf{CM}, w)$$
$$= \sqrt{\text{CF} \times \text{norm}_{\text{Ca1}}(\mathbf{CM}, w) + \text{AF} \times \text{norm}_{\text{An2}}(\mathbf{CM}, w)} \tag{17}$$

$$\text{norm}_{\text{in3}}(\mathbf{CM}, w)$$
$$= \sqrt{\text{CF} \times \text{norm}_{\text{Ca2}}(\mathbf{CM}, w) + \text{AF} \times \text{norm}_{\text{An2}}(\mathrm{CM}, w)} \tag{18}$$

$$\text{norm}_{\text{in4}}(\mathbf{CM}, w)$$
$$= \sqrt{\text{CF} \times \text{norm}_{\text{Ca2}}(\mathbf{CM}, w) + \text{AF} \times \text{norm}_{\text{An1}}(\mathbf{CM}, w)} \tag{19}$$

In this study, all the four interaction descriptors $\text{norm}_{\text{in1}}(\mathbf{CM}, w)$, $\text{norm}_{\text{in2}}(\mathbf{CM}, w)$, $\text{norm}_{\text{in3}}(\mathbf{CM}, w)$, and $\text{norm}_{\text{in4}}(\mathbf{CM}, w)$ were proposed to represent the cation−anion interactions individually, and then all those descriptors were concatenated to describe the characteristics of multiple-interaction system, where w = 1, 2, and 3, corresponding to $\text{norm}(\mathbf{CM},1)$, $\text{norm}(\mathbf{CM},2)$, and $\text{norm}(\mathbf{CM},3)$, and CF and AF refer to the mass fractions of the cations and anions, respectively. The $\text{norm}_{\text{Ca1}}(\mathbf{CM}, w)$, $\text{norm}_{\text{Ca2}}(\mathbf{CM}, w)$ and $\text{norm}_{\text{An1}}(\mathbf{CM}, w)$, $\text{norm}_{\text{An2}}(\mathbf{CM}, w)$ refer to the norm indexes of matrices calculated from the cations and anions, individually.

4.3. Descriptor Selection and Model Development. Based on the above programs, descriptors for IL mixtures were obtained. Determining the optimal descriptors that are most relevant to the target attribute is a critical step in QSPR modeling. The GA has been proved to be an efficient method applied for variable selection. In this paper, the optimal subset of descriptors that accurately characterize structural features corresponding to their target properties was selected by combining the GA-MLR method. The selection process of GA-MLR was implemented by Materials Studio (version 8.0).

4.4. Model Validation. To ensure the model reliability of resulted QSPR models, model validation can be absolutely necessary. Multiple verification strategies were employed for both internal and external validation in the present work.

The squared correlation coefficient (R^2) is commonly used to measure the fit between the observed and predicted values. The model has good fit if the R^2 value is greater than 0.6. The root-mean-square error (RMSE), the SE, and the AAE were also presented to evaluate the predictive capability.[36]

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (y_i - y_0)^2}{n}} \tag{20}$$

$$\text{SE} = \sqrt{\frac{\sum_{i=1}^{n} (y_i - y_0)^2}{n - 1}} \tag{21}$$

$$\text{AAE} = \frac{\sum_{i=1}^{n} |y_i - y_0|}{n} \qquad (22)$$

where y_i is defined as the observed $T_{d,5\%onset}$ value, y_0 represents the predicted $T_{d,5\%onset}$ value, and n is the number of the IL mixtures.

The CV is considered as an approach of great use to verify the robustness of the resulting model. The internal predictive ability was tested by the CV result (Q^2) in internal validation. The most economical LOO CV (Q^2_{LOO}) was utilized in this work,[37] which is calculated as follows:

$$Q^2_{LOO} = 1 - \frac{\sum_{i=1}^{training} (y_i - y_0)^2}{\sum_{i=1}^{training} (y_i - \bar{y})^2} \qquad (23)$$

where y_i, y_0, and $\bar{Y}$ represent, respectively, the observed, predicted, and mean observed $T_{d,5\%onset}$ values of the IL mixtures in the training set.

To determine the generalizability and the prediction ability of the resulting QSPR model, external validation is indispensable. The data set was divided into a training set and an external test set. The training set is used for descriptor selection and model development, while the test set is for model external validation. There are three main strategies commonly applied for data set partition of mixtures, including "Points out," "Compounds out," and "Mixtures out." The "Points out" strategy was employed in this work.

In this study, the data set was randomly divided into a training set with 24 samples (77.4% of the data set) and a test set (22.6% of the data set) with seven samples based on the "Points out" strategy. Then, the squared correlation coefficient for external validation (Q^2_{ext}) was calculated to indicate the predictive capability of the established QSPR model. The calculation process is as follows:

$$Q^2_{ext} = 1 - \frac{\sum_{i=1}^{test} (y_i - y_0)^2}{\sum_{i=1}^{test} (y_i - \bar{y}_{tr})^2} \qquad (24)$$

where y_i and y_0 represent the observed and predicted $T_{d,5\%onset}$ values of the IL mixtures in the test set, and $\bar{y}_{tr}$ represents the mean observed $T_{d,5\%onset}$ values of the IL mixtures in the training set.

Furthermore, the Y-randomization test was applied to verify the stability of the prediction model. It randomly scrambles the dependent variables and keeps independent variables unchanged. Then, the performance of the presented model is compared with the original model. The operation is repeated 50−100 times through the program. If the randomly shuffled models present much lower R^2 values than the original model, it can be considered that no chance correlation exists in the modeling process.

4.5. AD. An explicitly defined AD is required for any resulting QSPR model according to OECD principle 3.[38] The AD is a theoretical area defined by the molecular similarity of the training set and determines if the developed model can reliably predict the valid range of new chemicals. The popular Williams plot is applied as a practical tool for the definition of the AD. It is presented as a two-dimensional scatter plot determining whether the chemicals of the developed model are located in or out of the AD.

The structural similarity between the sample and the training set is represented by the leverage value h_i, which is defined as

$$h_i = X_i(X^T X)^{-1} X_i^T (i = 1, 2, 3, ..., n) \qquad (25)$$

where X_i represents the descriptor row-vector of descriptors for the ith sample, and X represents the matrix of descriptors for all samples in the training set. The warning leverage value h^* is defined as follows:

$$h^* = \frac{3(k + 1)}{m} \qquad (26)$$

where k represents the number of selected descriptors in the developed model, and m represents the number of samples in the training set. If the leverage value (h_i) of the sample is higher than the warning leverage value (h^*), the predicted result of the compound is considered out of the valid range of application and the prediction is considered unreliable.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.1c00846.

> The predicted and observed $T_{d,5\%onset}$ values, as well as the values of four employed descriptors in the developed model (Table S1); A complete list of the detailed compositions of mixtures as well as their corresponding observed $T_{d,5\%onset}$ values (Table S2) (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

Yong Pan − *Jiangsu Key Laboratory of Hazardous Chemicals Safety and Control, College of Safety Science and Engineering, Nanjing Tech University, Nanjing 211816, China;* orcid.org/0000-0002-6544-1443; Phone: +86-25-58139873; Email: yongpan@njtech.edu.cn

Authors

Hongpeng He − *Jiangsu Key Laboratory of Hazardous Chemicals Safety and Control, College of Safety Science and Engineering, Nanjing Tech University, Nanjing 211816, China*

Jianwen Meng − *Jiangsu Key Laboratory of Hazardous Chemicals Safety and Control, College of Safety Science and Engineering, Nanjing Tech University, Nanjing 211816, China*

Yongheng Li − *Jiangsu Key Laboratory of Hazardous Chemicals Safety and Control, College of Safety Science and Engineering, Nanjing Tech University, Nanjing 211816, China*

Junhong Zhong − *Jiangsu Key Laboratory of Hazardous Chemicals Safety and Control, College of Safety Science and Engineering, Nanjing Tech University, Nanjing 211816, China*

Weijia Duan − *Jiangsu Key Laboratory of Hazardous Chemicals Safety and Control, College of Safety Science and Engineering, Nanjing Tech University, Nanjing 211816, China*

Juncheng Jiang − *Jiangsu Key Laboratory of Hazardous Chemicals Safety and Control, College of Safety Science and Engineering, Nanjing Tech University, Nanjing 211816, China; School of Environment & Safety Engineering, Changzhou University, Changzhou 213164, China*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acsomega.1c00846

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Dong, K.; Liu, X.; Dong, H.; Zhang, X.; Zhang, S. Multiscale Studies on Ionic Liquids. *Chem. Rev.* **2017**, *117*, 6636−6695.

(2) Javed, F.; Ullah, F.; Zakaria, M. R.; Akil, H. M. An Approach to Classification and Hi-Tech Applications of Room-Temperature Ionic Liquids (RTILs): A Review. *J. Mol. Liq.* **2018**, *271*, 403−420.

(3) Welton, T. Ionic Liquids: A Brief History. *Biophys. Rev.* **2018**, *10*, 691−706.

(4) Keglevich, G.; Zsuzsa Kiss, N.; Radai, Z. An Overview of the Applications of Ionic Liquids as Catalysts and Additives in Organic Chemical Reactions. *Curr. Org. Chem.* **2017**, *22*, 533−556.

(5) Zhang, X.; Zhang, X.; Dong, H.; Zhao, Z.; Zhang, S.; Huang, Y. Carbon Capture with Ionic Liquids: Overview and Progress. *Energy Environ. Sci.* **2012**, *5*, 6668−6681.

(6) Vogl, T.; Menne, S.; Kühnel, R. S.; Balducci, A. The Beneficial Effect of Protic Ionic Liquids on the Lithium Environment in Electrolytes for Battery Applications. *J. Mater. Chem. A* **2014**, *2*, 8258−8265.

(7) Egorova, K. S.; Ananikov, V. P. Fundamental Importance of Ionic Interactions in the Liquid Phase: A Review of Recent Studies of Ionic Liquids in Biomedical and Pharmaceutical Applications. *J. Mol. Liq.* **2018**, *272*, 271−300.

(8) Chatel, G.; Pereira, J. F. B.; Debbeti, V.; Wang, H.; Rogers, R. D. Mixing Ionic Liquids-"simple Mixtures" or "Double Salts"? *Green Chem.* **2014**, *16*, 2051−2083.

(9) Taige, M. A.; Hilbert, D.; Schubert, T. J. S. Mixtures of Ionic Liquids as Possible Electrolytes for Lithium Ion Batteries. *Z. Phys. Chem.* **2012**, *226*, 129−139.

(10) Zistler, M.; Wachter, P.; Wasserscheid, P.; Gerhard, D.; Hinsch, A.; Sastrawan, R.; Gores, H. J. Comparison of Electrochemical Methods for Triiodide Diffusion Coefficient Measurements and Observation of Non-Stokesian Diffusion Behaviour in Binary Mixtures of Two Ionic Liquids. *Electrochim. Acta* **2006**, *52*, 161−169.

(11) Wang, M.; Zhang, L.; Gao, L.; Pi, K.; Zhang, J.; Zheng, C. Improvement of the CO2 Absorption Performance Using Ionic Liquid $[NH_2emim][BF_4]$ and $[emim][BF_4]/[bmim][BF_4]$ Mixtures. *Energy Fuels* **2013**, *27*, 461−466.

(12) Pinto, A. M.; Rodríguez, H.; Colón, Y. J.; Arce, A.; Soto, A. Absorption of Carbon Dioxide in Two Binary Mixtures of Ionic Liquids. *Ind. Eng. Chem. Res.* **2013**, *52*, 5975−5984.

(13) Lee, S. H.; Ha, S. H.; Hiep, N. M.; Chang, W. J.; Koo, Y. M. Lipase-Catalyzed Synthesis of Glucose Fatty Acid Ester Using Ionic Liquids Mixtures. *J. Biotechnol.* **2008**, *133*, 486−489.

(14) Zhao, R. S.; Wang, X.; Li, F. W.; Wang, S. S.; Zhang, L. L.; Cheng, C. G. Ionic Liquid/Ionic Liquid Dispersive Liquid-Liquid Microextraction. *J. Sep. Sci.* **2011**, *34*, 830−836.

(15) Akçay, A.; Balci, V.; Uzun, A. Structural Factors Controlling Thermal Stability of Imidazolium Ionic Liquids with 1-n-Butyl-3-Methylimidazolium Cation on γ-Al_2O_3. *Thermochim. Acta* **2014**, *589*, 131−136.

(16) Navarro, P.; Larriba, M.; García, J.; Rodríguez, F. Thermal Stability and Specific Heats of $\{[emim][DCA] + [emim][TCM]\}$ Mixed Ionic Liquids. *Thermochim. Acta* **2014**, *588*, 22−27.

(17) Clough, M. T.; Crick, C. R.; Gräsvik, J.; Hunt, P. A.; Niedermeyer, H.; Welton, T.; Whitaker, O. P. A Physicochemical Investigation of Ionic Liquid Mixtures. *Chem. Sci.* **2015**, *6*, 1101−1114.

(18) Larriba, M.; Navarro, P.; Beigbeder, J. B.; García, J.; Rodríguez, F. Mixing and Decomposition Behavior of $\{[4bmpy][Tf_2N] + [emim][EtSO_4]\}$ and $\{[4bmpy][Tf_2N] + [emim][TFES]\}$ Ionic Liquid Mixtures. *J. Chem. Thermodyn.* **2015**, *82*, 58−75.

(19) Huang, G.; Lin, W. C.; He, P.; Pan, Y.; Shu, C. M. Thermal Decomposition of Imidazolium-Based Ionic Liquid Binary Mixture: Processes and Mechanisms. *J. Mol. Liq.* **2018**, *272*, 37−42.

(20) Smiglak, M.; Reichert, W. M.; Holbrey, J. D.; Wilkes, J. S.; Sun, L.; Thrasher, J. S.; Kirichenko, K.; Singh, S.; Katritzky, A. R.; Rogers, R. D. Combustible Ionic Liquids by Design: Is Laboratory Safety Another Ionic Liquid Myth? *Chem. Commun.* **2006**, *24*, 2554−2556.

(21) Pena, C. A.; Soto, A.; King, A. W. T.; Rodríguez, H. Improved Reactivity of Cellulose via Its Crystallinity Reduction by Non-dissolving Pretreatment with an Ionic Liquid. *ACS Sustainable Chem. Eng.* **2019**, *7*, 9164−9171.

(22) Pinto, A. M.; Rodríguez, H.; Arce, A.; Soto, A. Carbon Dioxide Absorption in the Ionic Liquid 1-Ethylpyridinium Ethylsulfate and in Its Mixtures with Another Ionic Liquid. *Int. J. Greenhouse Gas Control* **2013**, *18*, 296−304.

(23) Pinto, A. M.; Rodríguez, H.; Arce, A.; Soto, A. Combined Physical and Chemical Absorption of Carbon Dioxide in a Mixture of Ionic Liquids. *J. Chem. Thermodyn.* **2014**, *77*, 197−205.

(24) Stolarska, O.; Soto, A.; Rodríguez, H.; Smiglak, M. Thermal Behaviour of Mixtures of 1-Alkylpyridinium Halides with and without a Common Ion. *J. Mol. Liq.* **2018**, *268*, 781−790.

(25) Lazzús, J. A. A Group Contribution Method to Predict the Thermal Decomposition Temperature of Ionic Liquids. *J. Mol. Liq.* **2012**, *168*, 87−93.

(26) Venkatraman, V.; Alsberg, B. K. Quantitative Structure-Property Relationship Modelling of Thermal Decomposition Temperatures of Ionic Liquids. *J. Mol. Liq.* **2016**, *223*, 60−67.

(27) Zhao, X.; Pan, Y.; Jiang, J.; Xu, S.; Jiang, J.; Ding, L. Thermal Hazard of Ionic Liquids: Modeling Thermal Decomposition Temperatures of Imidazolium Ionic Liquids via QSPR Method. *Ind. Eng. Chem. Res.* **2017**, *56*, 4185−4195.

(28) Zeeshan, M.; Nozari, V.; Keskin, S.; Uzun, A. Structural Factors Determining Thermal Stability Limits of Ionic Liquid/MOF Composites: Imidazolium Ionic Liquids Combined with CuBTC and ZIF-8. *Ind. Eng. Chem. Res.* **2019**, *58*, 14124−14138.

(29) Duan, W.; Pan, Y.; He, H.; Zhao, S.; Zhao, X.; Jiang, J.; Shu, C. M. Prediction of the Thermal Decomposition Temperatures of Imidazolium ILs Based on Norm Indexes. *J. Mol. Liq.* **2020**, *315*, No. 113780.

(30) Katritzky, A. R.; Kuanar, M.; Fara, D. C.; Karelson, M.; Acree, W. E.; Solov'ev, V. P.; Varnek, A. QSAR Modeling of Blood: Airand Tissue: AirPartition Coefficients Using Theoretical Descriptors. *Bioorg. Med. Chem.* **2005**, *13*, 6450−6463.

(31) Varnek, A.; Kireeva, N.; Tetko, I. V.; Baskin, I. I.; Solov'ev, V. P. Exhaustive QSPR Studies of a Large Diverse Set of Ionic Liquids: How Accurately Can We Predict Melting Points? *J. Chem. Inf. Model.* **2007**, *47*, 1111−1122.

(32) He, W.; Yan, F.; Jia, Q.; Xia, S.; Wang, Q. QSAR Models for Describing the Toxicological Effects of ILs against Staphylococcus Aureus Based on Norm Indexes. *Chemosphere* **2018**, *195*, 831−838.

(33) He, W.; Yan, F.; Jia, Q.; Xia, S.; Wang, Q. Prediction of Ionic Liquids Heat Capacity at Variable Temperatures Based on the Norm Indexes. *Fluid Phase Equilib.* **2019**, *500*, No. 112260.

(34) Gaudin, T.; Rotureau, P.; Fayet, G. Mixture Descriptors toward the Development of Quantitative Structure-Property Relationship Models for the Flash Points of Organic Mixtures. *Ind. Eng. Chem. Res.* **2015**, *54*, 6596−6604.

(35) Billard, I.; Marcou, G.; Ouadi, A.; Varnek, A. In Silico Design of New Ionic Liquids Based on Quantitative Structure-Property Relationship Models of Ionic Liquid Viscosity. *J. Phys. Chem. B* **2011**, *115*, 93−98.

(36) Tropsha, A.; Gramatica, P.; Gombar, V. K. The Importance of Being Earnest: Validation Is the Absolute Essential for Successful Application and Interpretation of QSPR Models. *QSAR Comb. Sci.* **2003**, *22*, 69−77.

(37) Konovalov, D. A.; Llewellyn, L. E.; Heyden, Y. V.; Coomans, D. Robust Cross-Validation of Linear Regression QSAR Models. *J. Chem. Inf. Model.* **2008**, *48*, 2081−2094.

(38) Gramatica, P. Principles of QSAR Models Validation: Internal and External. *QSAR Comb. Sci.* **2007**, *26*, 694−701.