Journal of Molecular Liquids 223 (2016) 60–67

Contents lists available at ScienceDirect

Journal of Molecular Liquids

journal homepage: www.elsevier.com/locate/molliq

Quantitative structure-property relationship modelling of thermal decomposition temperatures of ionic liquids

Vishwesh Venkatraman, Bjørn Kåre Alsberg*

Department of Chemistry, Norwegian University of Science and Technology, 7491Trondheim, Norway

ARTICLE INFO

Article history:
Received 2 May 2016
Received in revised form 3 August 2016
Accepted 6 August 2016
Available online 9 August 2016

Keywords:
QSPR
Ionic liquids
Thermal decomposition
Machine learning

ABSTRACT

The thermal decomposition point for ionic liquids (ILs) is an essential property that imposes an upper operating limit for many applications. Since the decomposition of ILs can lead to unwanted byproducts, it is desirable to improve their thermophysical properties and create more application specific compounds. With a view to rapidly estimate these properties of interest, approaches based on quantitative structure-property relationship (QSPR) models have been relatively successful but somewhat restricted to small datasets with limited diversity. Here, we investigate the effectiveness of a wide range of electronic, thermodynamic and geometrical descriptors derived from semi-empirical PM6 calculations to estimate the thermal decomposition temperatures of 995 diverse ILs comprising 461 cations and 119 anions. Of the two regression schemes used: partial least squares and random forests, the latter yielded slightly improved performances ($R^2_{cv} = 0.81, R^2_{test} = 0.77$). Analysis based on variable importance indicates that the anion-specific descriptors such as nucleophilicity and size significantly influence the thermal stabilities which are in agreement with experimental observations.

1. Introduction

Owing to their attractive properties such as negligible vapour pressure, large liquid range, thermal stability etc., ionic liquids (ILs) have been used in a number of application areas such as carbon capture [1], electrolytes in batteries [2], biomass processing [3], fuel cells [4] and medicine [5] to name a few. Furthermore, by selecting suitable combinations of the cations and anions, a vast number of potential candidate ILs (10^6–10^{18}) can be produced [6]. However, the laborious and often expensive synthesis and characterization of even a small subset of the possible combinations poses significant challenges. In order to make the search for ILs with specific properties sufficiently tractable, computational predictive models provide a viable alternative [7].

The thermophysical properties of ILs are of particular interest for heat transfer applications that range from low temperature refrigeration systems to solar energy collection at high temperatures [8]. Many of these applications require the ILs to be recycled and any degradation will negatively impact their efficacy. The thermal stability of the ILs is typically studied using thermogravimetric analysis (TGA) which yields the onset temperature of decomposition (T_d) and imposes an upper limit on the liquid range [9,10]. Theoretical and/or computational methods that can help predict the thermal behaviour and facilitate safe use would be of great practical interest. Although ab initio molecular dynamics (MD) [11] and density functional theory (DFT) [12,13] based methods have been used to investigate the kinetics and mechanisms of thermal degradation, the high computational cost limits their application. In comparison, the application of quantitative structure-property relationship (QSPR) models provide a relatively low-cost and computationally efficient protocol to screen new candidate ILs. Since these approaches yield a mathematical/statistical model correlating descriptors derived from the molecular structure with known molecular/material properties [7,14,15], they may also provide an improved understanding of the phenomena under investigation.

Thermal stability is seen to be strongly dependent on the structure of the IL, in particular the coordinating ability of the anion [10,16]. This knowledge has been used extensively in the development of important IL properties such as melting points [17], heats of fusion [18] and heat capacity [19]. In this work, we investigate the use of structure-property models to estimate thermal decomposition temperatures (corresponding to 10% weight loss) of a structurally diverse dataset of 995 ILs, compiled from various literature sources. Previous studies have reported regression models based on group contribution [20,21] schemes, topological indices [22] and three-dimensional (3D) descriptors [23] for datasets ranging between 150 and 600 datapoints (see Table S1 in the Supplementary material).

* Corresponding author.
E-mail address: bjorn.k.alsberg@ntnu.no (B. Alsberg).

http://dx.doi.org/10.1016/j.molliq.2016.08.023
0167-7322/

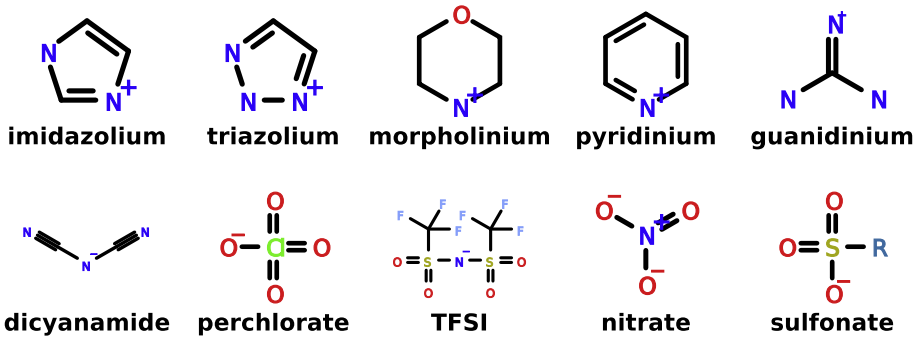

Fig. 1. Prominent cations (top row) and anions (bottom row) found in the dataset.

In addition to a much larger dataset covering more recent literature and a more diverse set of chemical moieties, here, we make use of descriptors [24] derived from semi-empirical MOPAC [25] calculations. In particular, we examine the influence of reactivity indices that provide information about the molecular interactions that influence thermal stability [26,27]. Our results show that the models based on non-linear random forests outperform those based on linear partial least squares regression.

2. Methods and materials

2.1. Dataset

Experimental thermal decomposition temperatures (T_d) for a diverse set of 995 ILs comprising 461 cations and 119 anions were compiled from the literature (a primary source was [28]). The prominent cations and anion groups are shown in Fig. 1. The names of the ILs, the experimental T_d and associated references are listed in Table S2 in the Supplementary material. The temperatures span the range 350 to 760 K (rounded to the nearest integer) with a majority of the values concentrated between 450 and 700 K (see Fig. F1 in the Supplementary material). The distribution of the temperatures summarized by the minimum, median, maximum and the first and third quartiles are shown as boxplots in Figs. 2 and 3 for the prominent cations (imidazolium, triazolium, amino acids, phosphonium) and anions (perchlorate, halide, bis(trifluoromethanesulfonyl)amide (TFSI)) present in the data.

2.2. Descriptors

We have focused mainly on 3D and topographical descriptors that can be calculated independent of size, as the predictor variables. The 3D structures of the individual cations and anions were subjected to geometry optimization and force calculations using the *PM*6 Hamiltonian [25]. In a recent study [29], the *PM*6 geometry optimization method has been shown to have predictive capabilities comparable with that of more computationally expensive DFT methods. Various quantum chemical and molecular orbital based descriptors such as the HOMO/LUMO energies, charges, polarizabilities and superdelocalizabilities,vibrational frequency based eigenvalue (EVA) indices thermodynamic descriptors (entropies, enthalpies), charge partial surface area (CPSA), property weighted radial distribution function (RDF), and 3D molecular representations of structure based on electron diffraction descriptors (3D-MoRSE) descriptors and other geometrical indices (such as the globularity and weighted holistic invariant molecular descriptors (WHIM)) were calculated using the open source software KRAKENX [24] (downloadable from www.krakenminer.com) developed in our group. Table S5 in the Supplementary material provides a description of the variables. The generated descriptors have been found to be successful in various applications such as predicting solar power conversion efficiencies, viscosity, density and pK_a [24,30-32]. The descriptors were calculated independently for the cations and anions and interaction effects were not considered [33]. The descriptor names were prefixed with “Cation” and “Anion” in order to differentiate between those

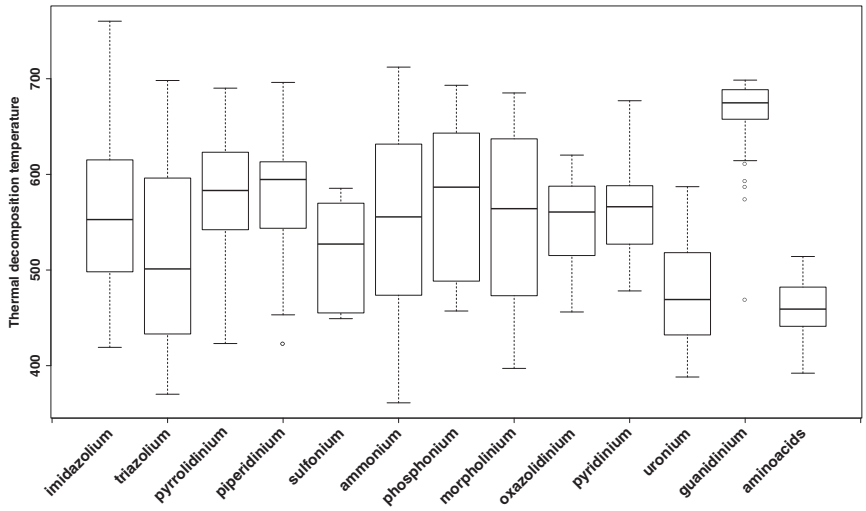

Fig. 2. Boxplot showing the distribution of the T_d for prominent cations. The “whiskers” above and below the box indicate the positions of the minimum and maximum. See the supporting information for details.