Computers and Chemical Engineering 81 (2015) 310–322

Contents lists available at ScienceDirect

Computers and Chemical Engineering

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

Multivariate characterization, modeling, and design of ionic liquid molecules

Subin Hada, Robert H. Herring III, Sarah E. Davis, Mario R. Eden*

Department of Chemical Engineering, Auburn University, Auburn, AL 36849, USA

ARTICLE INFO

Article history:
Received 17 January 2015
Received in revised form 7 April 2015
Accepted 8 April 2015
Available online 18 April 2015

Keywords:
Ionic liquid design
Latent variable models
Principal component analysis
Density functional theory
QSPR
CAMD

ABSTRACT

Ionic liquids that have tailored structures with an array of unique functional properties can have important applications in areas such as CO_2 capture and sequestration, sulfur removal from fuels, energy storage, biomass pretreatment, and chemical separations. Within a computer-aided molecular design (CAMD) framework, a characterization based method was combined with chemometric techniques in a reverse problem formulation to design ionic liquid (IL) structures corresponding to particular physical properties. Infrared spectra generated from density functional theory (DFT) simulations were used for capturing information on molecular architecture and calibration of latent variable property models to synthesize ILs in a logical and systematic methodology.

1. Introduction

Developing techniques and systematic methodologies to discover and optimize the molecular architecture of chemical products that possess and deliver both the consumer specified attributes and environmentally acceptable characteristics has been a primary focus of research in the process systems engineering (PSE) community. With rising costs and stringent regulations from environmental, health, and safety (EHS) hazards associated with toxicity or waste disposal, the design of environmentally benign solvents and alternative media for extraction and purification are new challenges within chemical product design.

One class of novel compounds being studied for such applications is ionic liquids (ILs) (Ayala et al., 2005; Holbrey and Seddon, 1999; Karunanithi and Mehrkesh, 2013; McLeese et al., 2010; Turner et al., 2003). They have become the subject of an increasing number of investigations due to their unique properties such as high polarity, stability at high temperature, flame resistance, and negligible vapor pressure (Ranke et al., 2007). Ionic liquids that have tailored structures with an array of unique functional properties can have important applications in areas such as CO_2 capture and sequestration, sulfur removal from fuels, energy storage, biomass pre-treatment, and chemical separations. Through variation of both cation and anion, particular ionic liquids with tunable physical properties can be tailored. For example, the miscibility of ionic liquids with water or organic solvents can be varied with alkyl chain lengths on the cation and the type of anion present (Wasserscheid and Keim, 2000).

It is estimated that over 100 trillion unique cation/anion combinations are possible for use as room temperature ionic liquids, the majority of which have never been synthesized (Turner et al., 2003). The traditional experimental trial-and-error approach of searching through this large molecular space is unrealistic as it is both time and labor intensive. Thus, it is essential to develop a logical and systematic approach of selectively choosing a given ionic pair that matches a set of desired physico-chemical property targets. An example of possible combinations of cation, anion, and the alkyl chain of the side chain attached to the cation is represented graphically in Fig. 1.

Unlike database searching of chemical products which is limited by the number of existing tabulated alternatives, computer-aided molecular design (CAMD) facilitates the application of computer algorithms to solve the mathematical formulations of new chemical product designs with desired properties (Gani, 2004; Maranas, 1996; Roughton et al., 2012; Weis and Visco, 2010). To accomplish this, CAMD generally utilizes group contribution-based property prediction methods to evaluate the generated compound with respect to a specified set of desired properties

* Corresponding author. Tel.: +1 334 844 2064; fax: +1 334 844 2063.
E-mail address: edenmar@auburn.edu (M.R. Eden).

http://dx.doi.org/10.1016/j.compchemeng.2015.04.009
0098-1354/

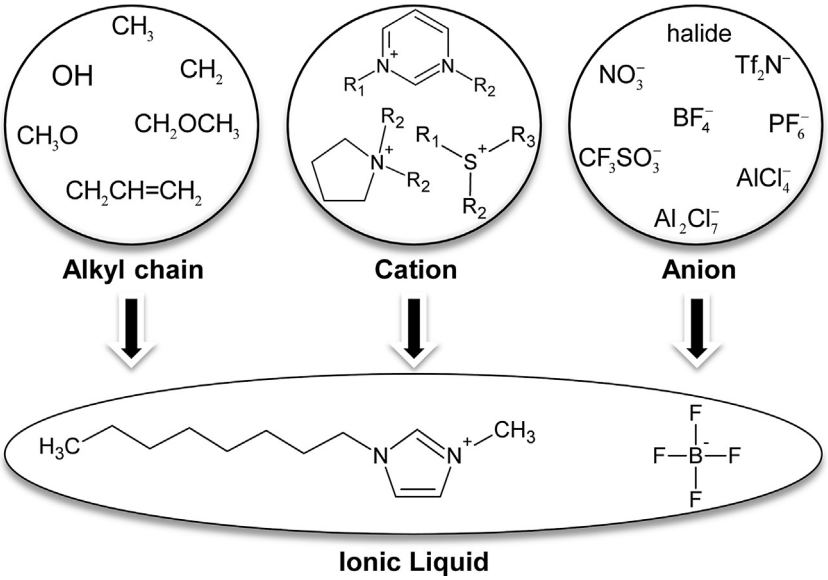

Fig. 1. Selection of anions, cations, and side chains attached in cations for a task specific ionic liquid application.

(Chemmangattuvalappil et al., 2010; Constantinou and Gani, 1994; Eljack and Eden, 2007; Harper et al., 1999; Joback, 1989; Marrero and Gani, 2001).

Group contribution methods (GCM) use predefined fragments (group, bond, atom, etc.) to represent a molecule and the properties of the molecule are estimated by summing all the contributions from each fragment that make up the molecule (Constantinou and Gani, 1994; Joback and Reid, 1987; Marrero and Gani, 2001). Group contribution methods (GCM) are simple and powerful property prediction methods. However, any application of group contribution relies on the availability of atom type, molecular group type, or type of chemical bonding present to describe the structure as well as tables giving the property contributions of each group. There are many property parameters or consumer attributes for which group contribution (GC) data is not available. For example, GCM does not exist for properties such as cetane number (CN), and property parameters to estimate properties of ionic liquid groups are limited.

Recently, a characterization-based group contribution method (cGCM) has been demonstrated to predict physicochemical properties and design of biodiesel additive molecules (Hada et al., 2014). The characterization-based group contribution method (cGCM) utilizes spectroscopy-based characterization techniques such as infrared (IR) and near infrared (NIR) to generate numerical descriptors of molecular architecture in terms of IR/NIR frequency (Solvason, 2011). Since the cGCM utilizes the latent property parameters based on characterization data instead of regression-based property parameters, the developed method and techniques alleviated the shortcoming of the conventional group contribution methods (GCM) mentioned above (Hada et al., 2014).

IR/NIR spectra of molecules can contain large quantities of descriptor data involving the information on molecular architecture. Managing such complexity of information to design chemical products and to build appropriate models for a specific application requires a systematic method for capturing important features of the molecular architecture. To decompose large quantities of information, and identify systematic patterns and important features of the molecular architecture in such multivariate data, multivariate statistical techniques such as principal component analysis (PCA) are used (Eriksson et al., 2006; Wold, 1995).

This paper extends the group contribution methods based on characterization (cGCM) data by investigating the use of density functional theory (DFT) based simulation techniques to generate the required IR spectra as molecular descriptors with which to develop predictive property models that can be used for the reverse design of ionic liquids. Such techniques can be particularly beneficial when experimental spectroscopic data is not available to estimate properties of chemical products such as ionic liquids, thereby extending the capabilities of a design method based on characterization techniques.

2. Methodology

Within the computer-aided molecular design (CAMD) framework (Achenie et al., 2002; Eden et al., 2004; Eljack et al., 2008; Harper et al., 1999; McLeese et al., 2010), a characterization-based method was combined with chemometric (Hada et al., 2014; Solvason, 2011) and computational chemistry techniques (Scott and Radom, 1996) in a reverse problem formulation (Eden et al., 2004) to develop a logical and systematic approach of selectively choosing a given ionic pair that matches a set of desired physical property targets. The following sections describe the integration of these techniques within a detailed framework exemplified by the design of ionic liquids with tailored properties.

2.1. Characterization

Multivariate characterization techniques like infrared (IR) spectroscopy are utilized to generate numerical descriptors of molecular architecture in terms of IR frequency of a set of ionic liquids. 23 training set ionic liquid molecules and their three properties (dynamic viscosity, density, and melting temperature) were compiled from the IUPAC Ionic Liquid database (ILThermo) (Dong et al., 2007) and listed in Table A1 in the Appendix. In general, larger training sets are desired, when available, to have an adequate representative sample of molecules that can produce reliable and stable prediction models.

The choice of 23 samples in this proof of concept paper is limited by the availability of three properties for each of the ionic pairs in the ILThermo database. In addition, the choice of the three selected properties of ILs compared to the other interesting properties such as polarity, conductance, etc., is because the former are known to have expressions that follow linear mixing rules when applied in group contribution based design frameworks. The choice of using