Journal of Molecular Liquids 254 (2018) 231–240

Contents lists available at ScienceDirect

Journal of Molecular Liquids

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

Exploration of automatic learning to establish relationships between the molecular structure of chiral ionic liquids and the specific optical rotation

Mengyao Chen [a], Kaixia Xiao [a], Tanfeng Zhao [a], Yanmei Zhou [a], Qingyou Zhang [a,*], Joao Aires-de-Sousa [b,*]

[a] Henan Engineering Research Center of Industrial Circulating Water Treatment, College of Chemistry and Chemical Engineering, Henan University, Kaifeng 475004, PR China
[b] LAQV and REQUIMTE, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

ARTICLE INFO

Article history:
Received 12 October 2017
Received in revised form 11 January 2018
Accepted 16 January 2018

Keywords:
Ionic liquids
Machine learning
Specific optical rotation
Chirality
Molecular descriptors

ABSTRACT

The experimental assignment of absolute configurations of chiral compounds is generally expensive and time consuming. Theoretical prediction of optical rotations can greatly assist in this endeavor. Herein a chemoinformatics exploration is reported to automatically distribute ionic liquids (IL) in a counterpropagation neural network incorporating a Kohonen layer for processing descriptors of cations, and five output layers to store optical rotations of ILs with five different anions. Cations were represented by chiral descriptors based on parities of the chiral centers assigned according to several steric and physicochemical atomic properties of the ligands. A data set with 106 chiral ionic liquids and their enantiomers was used covering a range of specific optical rotation between −147.3° and +147.3°. The obtained maps reveal relationships between the molecular structure of cations, the anions, and the optical rotation. The correct assignment of the sign of the specific optical rotations was achieved for 20 of the 22 pairs of enantiomers in an independent test set. CPG NNs estimated the value of the specific rotation with a RMS error of 22°, but more accurate predictions were obtained with a random forest that yielded a RMS error of 11°.

1. Introduction

Ionic liquids are salts entirely composed of anions and cations in the liquid state below a low temperature threshold (e.g., 100 °C). The reduced volatility makes them potential environment-friendly solvents and many of their distinctive properties have been explored, such as liquid temperature range, solubility of organic and inorganic compounds, electrical conductivity, thermal stability and oxidation resistance. Due to excellent physical and chemical properties, ionic liquids have been widely applied in organic synthesis, electrochemistry, separation and extraction, clean oxidation, catalysis and other fields [1]. Chiral ionic liquids have been used for chiral recognition and enantiomeric separation [2], as co-catalysts for asymmetric reactions [3] or as a tailoring agent in the synthesis of materials [4]. Different properties of individual enantiomers can only be revealed in the presence of an external chiral element.

Specific rotation ([α]) is a classic property of chiral compounds, enantiomers exhibiting opposite specific rotations. It can be experimentally measured with a polarimeter, and is commonly used to distinguish enantiomers, to assess the enantiomeric composition of mixtures or to determine the concentration of chiral compounds. Optical rotations can be used to assign the absolute configuration of unknown compounds based on theoretical methods to predict the sign of rotation from the geometrical configuration [5] – if the experimental specific rotation match the prediction, the absolute configuration is inferred to be the same.

Machine learning methodologies can automatically process molecular data, for the recognition of patterns, identification of common features in series, detection of trends, identification of outliers, and estimation of unknowns. Molecular data concerning chiral compounds and chiral properties must be encoded with chiral molecular descriptors, i.e. descriptors yielding different values for opposite enantiomers. Several approaches have been reported for chiral descriptors [6]. Some of us have been involved in the development of conformation-independent and conformation-dependent chirality codes derived from radial distribution functions [7,8], physicochemical atomic stereo descriptors (PAS) [9] and simplified derivatives [10–12].

Current typical collections of chiral ionic liquids consist of series with few cation motifs, small variations in the cation structure and variations in anions (typically few distinct anions, often monoatomic species). Here we explore the possibility of processing such data sets with counterpropagation neural networks (CPG NN), which distribute objects on a self-organizing map, based on object attributes, and link the map to several output layers where information is stored related to the corresponding properties.

We combine PAS chiral descriptors (to encode the molecular structures of IL cations) with a CPG NN designed with several output layers,

* Corresponding authors.
E-mail addresses: qingyou@vip.henu.edu.cn (Q. Zhang), joao@airesdesousa.com (J. Aires-de-Sousa).

https://doi.org/10.1016/j.molliq.2018.01.086
0167-7322/

each corresponding to a different IL anion, to visualize relationships between the molecular structures and their specific optical rotations, as well as to estimate specific rotations. PAS descriptors were proposed [9] for the investigation of relationships between the chirality of molecular structures and observable properties, as they are derived from physicochemical and steric atomic properties and are less prone to the artifacts of stereochemical specification schemes developed for nomenclature purposes. Several machine learning alorithms are explored to estimate specific rotations from the PAS descriptors.

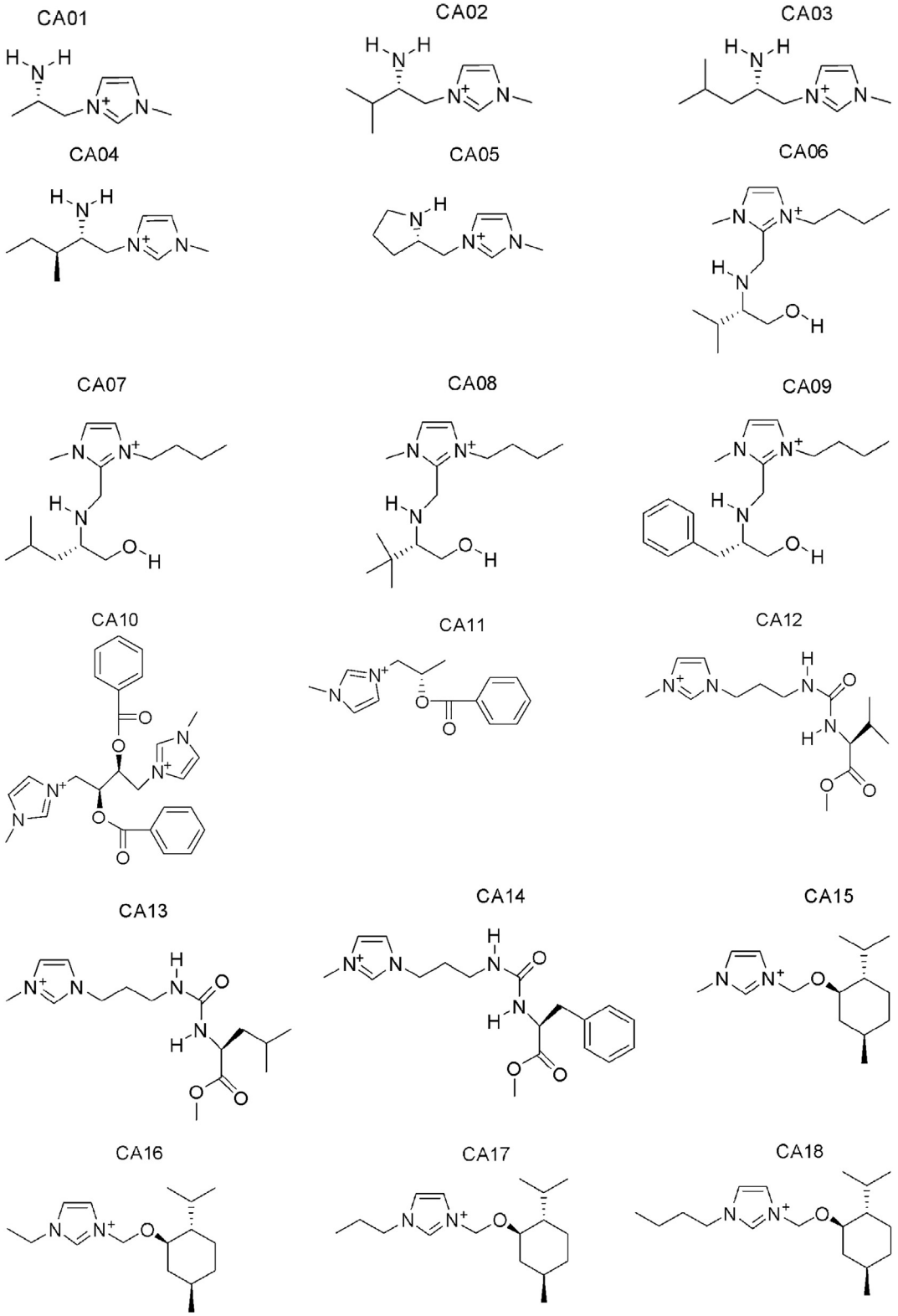

Fig. 1. The structures of the 46 cations.