Fluid Phase Equilibria 475 (2018) 25–36

Contents lists available at ScienceDirect

Fluid Phase Equilibria

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

Computational prediction of cellulose solubilities in ionic liquids based on COSMO-RS

Yunhan Chu [a], Xiangping Zhang [b], Magne Hillestad [a], Xuezhong He [a, *]

[a] Department of Chemical Engineering, Norwegian University of Science and Technology, NO-7491, Trondheim, Norway
[b] Beijing Key Laboratory of Ionic Liquids Clean Process, Institute of Process Engineering, Chinese Academy of Sciences, P.O. Box 353, Beijing, 100190, China

ARTICLE INFO

Article history:
Received 15 June 2018
Received in revised form
18 July 2018
Accepted 26 July 2018
Available online 27 July 2018

Keywords:
Ionic liquids
Cellulose solubility
COSMO-RS
Activity coefficient
Excess enthalpy

ABSTRACT

A computational approach is presented for prediction of cellulose solubilities in ionic liquids (ILs) based on COSMO-RS (Conductor-like Screening Model for Real Solvents). Thermodynamically stable molecular structures were optimized from 2D structures of cellulose and ILs following specific force-field based search of conformation lowest in energy and quantum chemical optimizations of molecular geometry. The thermodynamic property of logarithmic activity coefficient ($\ln\gamma$) and excess enthalpy (H^E) were calculated by COSMO-RS based on the COSMO molecular surfaces of cellulose and ILs to qualitatively predict the ability of ILs for cellulose dissolution. To evaluate the method, four sets of ILs were used to calculate $\ln\gamma$ and H^E based on four different cellulose models. The goodness-of-fit of linear regressions between the experimental cellulose solubilities and the calculated $\ln\gamma$ and H^E shows that $\ln\gamma$ is more reliable than H^E for prediction of the dissolving power of ILs to dissolve cellulose. However, H^E is more suitable for prediction of the dissolution ability of halogen-based ILs. Moreover, all the cellulose models gave comparably good prediction results regarding of the dissolving power of ILs based on the calculated $\ln\gamma$, but the cellobiose model was identified as the optimal model due to the relatively higher prediction ability (R^2) across different IL datasets. The approach is time efficient and robust, which provides a novel method for large-scale screening of ILs for cellulose dissolution.

1. Introduction

Biomass is an abundant renewable resource on the earth [1]. Growing concerns about sustainability and environmental protection brings large amount of attention to efficient conversion of biomass into valuable products such as biofuels, chemicals and biomaterials, as the concept of biorefinery suggests [2–4]. Plants and plant-based biomass (lignocellulose) contains three major components: cellulose, hemicellulose and lignin [5]. Cellulose, as the largest component of lignocellulose and a sustainable raw material, is widely used in producing paper, fiber, membrane, and other commodity materials and chemicals, meanwhile, it has a high global quantity (~700,000 billion tons) and only a small amount (~0.1 billion tons) is currently used for production, which leaves a large space for wide employment [6,7]. For the full utilization of cellulose, a prime step is to dissolve cellulose. Cellulose is a type of linear-chain biopolymer consisting of several hundred to over ten thousand β-(1−>4)-linked glucose repeating units [6,8]. The linear structure enables numerous cellulose strands to be packed into crystalline fibrils through enormous number of intramolecular and intermolecular hydrogen bonds [9], which makes cellulose highly resistant to chemical modification and difficult to dissolve. Ionic liquids (ILs) as a new type of green solvents, provides the possibility to dissolve cellulose [8,10] and lignocellulosic biomass [5,11] in a clean process, which has attracted a great deal of academic and industrial interest. These solvents are salts generally composed of a large, low-symmetry and non-reactive organic cation and an organic or inorganic anion that can largely tune their physical and chemical properties [12,13]. Compared to conventional solvents [14–19], ILs show remarkable properties such as low vapor pressure, high thermal stability, non-flammability, non-volatility and low toxicity. Moreover, by proper selection of cation − anion combinations, ILs can be tailored with desired properties (e.g. viscosity, density and solubility) [20,21].

Appropriate ILs can be selected as cellulose solvents. In the recent years, more than 60 ILs have been experimentally examined for their solubilities of cellulose [8]. However, there is a vast

* Corresponding author.
E-mail address: xuezhong.he@ntnu.no (X. He).

https://doi.org/10.1016/j.fluid.2018.07.026
0378-3812/

number of ILs available, and selection of the most appropriate ILs with desired solubilities from a large pool of possible ILs for a specific task requires long time and many experiments, which highlights the advantage of using rapid and accurate *a priori* screening tools to predict the dissolution capacity of ILs. COSMO-RS (Conductor-like Screening Model for Real Solvents) as a physically founded approach combining quantum chemistry and statistical thermodynamics, has recently attracted great attention for calculation of thermodynamic properties [22–27]. COSMO-RS integrates dominant interactions of H-bonds, misfits, and van der Waals forces in the ionic liquid systems to characterize multiple solvation interactions, and can be used for performing mixture calculations at various temperatures [28]. With COSMO molecular surface generated by quantum chemical methods (QM), COSMO-RS calculates the charge distribution (σ-profile) on the molecular surface and chemical potential (μ) of molecule in a liquid solvent or mixture based solely on information concerning the 3D molecular structure. The computed chemical potentials are the basis for prediction of other thermodynamic equilibrium properties such as activity coefficient (γ), excess enthalpy (H^E) and solubility. By reducing the thermodynamics of a mixture to the interaction of a number of surface segments, COSMO-RS gains full predictivity from underlying QM calculations, which enables high-speed screening of compounds for desired properties.

COSMO-RS has a general suitability for prediction of properties of IL systems [27,29]. Previous publications have also demonstrated the applicability of COSMO-RS in prediction of solubilities of biopolymers [30–32] and drug molecules [28] in ILs based on COSMO surfaces of different molecular models. Given that an appropriate database of QM-COSMO calculations is available, COSMO-RS are well suited for the task of screening large numbers of solvents or solutes. On the other hand, the impact of conformations of cations/anions/solutes on the prediction quality of COSMO-RS is well recognized that even two different conformations with nearly identical energies for same molecule can result in complete different predictions of thermodynamic properties [33,34]. Thus, it is important to carefully select molecular model and conformation in order to obtain quantitatively and qualitatively correct predictions. Sampling of multiple relevant conformers according to their Boltzmann distribution of free energy can be a good solution [30]. In this work, we were interested in examining the solubilities of cellulose in ILs by the use of a 'medium' thermodynamically stable conformational state [35] both for cellulose and IL in COSMO-RS calculation, which may help to rapidly reduce computation time. The thermodynamically stable molecular conformation was obtained through specific force-field based search of lowest-energy conformer and QM-based geometry optimization. The ADF COSMO-RS [36–39] was facilitated to conduct the DFT (density functional theory) optimization of molecular geometry and calculations of thermodynamic properties for prediction of cellulose solubilities in ILs on the availability of convenient scripting tools. In order to evaluate our approach, four different cellulose models (i.e. glucose, cellobiose, cellotriose and cellotetraose) were selected to represent cellulose, and four diverse sets of ILs were used to calculate the lnγ of cellulose in ILs and the H^E of mixture at different temperatures. The calculated lnγ and H^E were compared with available experimental data of cellulose solubilities by linear regression, and the best cellulose model for prediction of cellulose solubilities in ILs was selected based on the goodness-of-fit of the regressions.

2. Methods

2.1. Theoretical basis for COSMO-RS calculation

Dissolution of cellulose is a result of disruption of the inter- and intra-molecular hydrogen bonding, and the formation of new hydrogen bonds mainly between the carbohydrate hydroxyl protons and the anions of the ILs [35,40–42]. The logarithmic activity coefficient (lnγ) of cellulose in ILs at infinite dilution and H^E of mixture (0.8:0.1:0.1 in mole fraction for biopolymer, IL cation and anion respectively) which are supposed to have correlations with cellulose solubility were calculated by COSMO-RS to qualitatively measure the dissolution ability of ILs.

The activity coefficients is associated with the affinity of solutes to solvents resulting from differences in the strength and nature between molecules caused by local pairwise interactions of surface segments due to the electrostatic misfit, hydrogen bonding and van der Waals dispersive forces [34]. With the chemical potential (μ_i^S) of a compound in a mixture being calculated, the activity coefficient (γ_i^S) at a given temperature (T) can be derived by the following equation [23]:

$$\gamma_i^S = \exp\left\{\frac{\mu_i^S - \mu_i^X}{RT}\right\} \quad (1)$$

where μ_i^S and μ_i^X are the potential of compound *i* in the reference state of the mixture and the pure compound, respectively.

H^E is a fundamental thermodynamic property (defined as the temperature derivative of the Gibbs free energy), which gives an insight into the behavior of the species in solution and can act as a sensitive measure of the intermolecular interactions within the mixture [33]. In the current version of ADF COSMO-RS, the H^E (kJ/mol) is calculated based on Gibbs-Helmholtz equation:

$$H^E = -T^2\frac{\partial(G^E/T)}{\partial T} \quad (2)$$

where G^E is excess Gibbs free energy (kJ/mol), and T is temperature (K).

2.2. Geometry optimization and COSMO-RS calculations

The COSMO-RS calculations were based on the ADF COSMO result files [38], which are the results of COSMO quantum mechanical calculation on the basis of information solely concerning the molecular structure. The molecular structure used for COSMO calculation were generated through a computational route including following steps: 1) the molecules of cellulose models and ILs were sketched as 2D structures using Marvinsketch [43]; 2) the structures of cellulose models were subjected to a 3D conversion from 2D by Molconvert [43] and a conformational search at energy minimization based Dreiding force field [44] by Cxcalc [43]; and 3) the obtained lowest-energy conformer of celluloses were optimized with geometry based on PM6 [45] method of MOPAC [46,47] – and this saves time of DFT optimization at the late stage that the SCF (self-consistent field) gets converged at a smaller amount of geometry cycles. In contrast, the ILs were conducted with searching a lowest-energy conformer based on the universal force field (UFF) [48] implemented in OpenBabel [49]. The PM6 method will deform the structures of some cations thus was not used in optimization of ILs. The obtained geometries of cellulose and ILs were further optimized to obtain the COSMO-RS result files by DFT method using the main parametrization GGA:BP/TZP in ADF [36]. The σ-profiles, σ-potentials, and thermodynamic properties such as lnγ and H^E were then calculated at a given temperature by COSMO-RS based on the obtained COSMO result files. A self-developed scripting tool called MoDoop was utilized to automatize the overall computational workflow, which imports the 2D structures of cellulose and ILs and exports the COSMO-RS calculated σ-profiles, σ-potentials and thermodynamic properties. Here, generation of COSMO-RS