Fluid Phase Equilibria 452 (2017) 58–68

Contents lists available at ScienceDirect

Fluid Phase Equilibria

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

A high-throughput computing procedure for predicting vapor-liquid equilibria of binary mixtures – Using carbon dioxide and n-alkanes as examples

Fenglei Cao [a], Zheng Gong [a], Yanze Wu [a], Huai Sun [a, b, *]

[a] School of Chemistry and Chemical Engineering and Key Laboratory of Scientific and Engineering Computing of Ministry of Education, Shanghai Jiao Tong University, Shanghai 200240, China
[b] State Key Laboratory of Inorganic Synthesis & Preparative Chemistry, College of Chemistry, Jilin University, Changchun 130012, China

ARTICLE INFO

Article history:
Received 3 May 2017
Received in revised form
15 August 2017
Accepted 23 August 2017
Available online 26 August 2017

Keywords:
High-throughput computing
Vapor-liquid equilibria
Binary mixture

ABSTRACT

An automatic high-throughput computing (HTC) procedure is implemented for calculating vapor-liquid equilibrium (VLE) curves of binary systems using molecular dynamics simulation and coarse-grained force field. The HTC procedure builds simulation models, carries out the simulations, and validates the simulations automatically. The procedure is demonstrated by calculating the VLE curves for 16 binary mixtures of carbon dioxide and n-alkanes on 856 state points in the temperature range of 277.2–420.0 K and pressure range of 0.3–25.0 MPa. The averaged uncertainties in predictions are 0.119 MPa for pressure, 0.007 and 0.029 for liquid and vapor CO_2 molar fractions, and 0.005 and 0.010 for liquid and vapor densities respectively. The Validation against experiment data on 10 binary systems and 316 state points shows that the predictions are accurate with average deviation of about 5% for CO_2 mole fraction and 3.5% for saturated density at pressure range to $P < 0.9P_c$. This automatic procedure can only be used for the prediction of the vapor-liquid equilibrium. The computed data can be obtained from http://sun.sjtu.edu.cn/msd.

1. Introduction

The use of molecular simulations to predict thermodynamic properties of fluids has drawn considerable interests in the past decades. For example, a series of biannual competitions of the Industrial Fluid Properties Simulation Collective (IFPSC) have been organized for the past 15 years [1]. Along with the steady advance of computer technology, simulation methods, software and force fields, there have been more and more successful cases of simulations that predict thermodynamic properties with precisions comparable with experimental measurements. However, in order to establish the simulation as a reliable tool, calculation protocol which includes both simulation method and force field parameter must be validated by a sufficiently large number of chemical classes. Taking advantage of computer technology, this task can be done by high-throughput computing (HTC) so that large number of systems can be screened and any flaws in the simulation methods or force field parameters can be detected and corrected systematically. Once the protocol is validated, it can be used to generate data efficiently to complement the experimental measurements which are often difficult, expensive and environment unfriendly.

Several HTC procedures have been reported for screening materials of CO2 capture [2,3], hydrogen storage [4] and gas separations [5], or for computing protein-protein, protein-ligand interactions [6–9]. In this work, we developed a HTC procedure to calculate the vapor-liquid equilibria (VLE) of binary mixtures. Due to the ample experimental data available, binary mixtures of carbon dioxide (CO_2) and hydrocarbons of different chain lengths were used to validate the procedure. The solubility of CO_2 in hydrocarbons [10,11] is important for many industrial processes including enhanced oil recovery (EOR), CO_2 capture and storage (CCS), natural gas transportation, and crude-oil processing. From a theoretical point of view, the homologous series of CO_2/*n*-alkane binary mixtures provides an excellent opportunity to investigate fluid phase behavior. Although the VLE behavior is commonly described by using thermodynamic models [12–20], molecular simulations have

* Corresponding author. School of Chemistry and Chemical Engineering and Key Laboratory of Scientific and Engineering Computing of Ministry of Education, Shanghai Jiao Tong University, Shanghai 200240, China.
E-mail address: huaisun@sjtu.edu.cn (H. Sun).

http://dx.doi.org/10.1016/j.fluid.2017.08.021
0378-3812/

become an alternative to predict the VLEs based on microscopic information. The VLE curves of CO_2/*n*-alkane mixtures have been predicted using Gibbs ensemble Monte Carlo (GEMC) simulations using a united-atom (UA) force field [21] and molecular dynamics (MD) simulations using coarse-grained (CG) force fields [20,22,23].

This work was focused on the automation of the HTC procedure, which should build the simulation models, carry out the simulations, detect and correct problems, and analyze the outputs automatically so that the simulations could be carried out repeatedly with minimum human interference. The details are explained in section 2 of this paper. The predictions were based on the CG force field which was developed using a free energy Lennard-Jones 12-6 (FE-12-6) function in describing the non-bonded interactions [24] and extended to the mixtures of alkanes with nitrogen, oxygen, and CO_2 and tested on predictions of VLE for binary mixtures of methane/*n*-butane, N_2/*n*-decane and CO_2/*n*-decane [25]. Based on our previous work, MD method which is more efficient than GEMC in simulating VLE systems using the CG model, was used in the simulations. The validations and predictions are summarized in section 3. About 37% of the computed state points have experimental counterparts to be compared with, which servers the purpose of validation. Other 63% part are predictions, which complement the missing experimental data. This paper is finished by presenting a short remarks and outlook in Section 4.

2. The automated HTC procedure

A workflow diagram of the automated HTC procedure is illustrated in Fig. 1. The procedure has five major functions: reading the input of the user, building simulation boxes and system representations, simulation of these systems, validation of the accuracy of the simulations and printing calculated property data in the output. Among the five functions, the validation is the central concern of automation, as detailed in the figure. The validation procedure includes tests on primary equilibrium, configuration, secondary equilibrium and precision. Any failure of these validation tests prompts the workflow to build a new system or to repeat the simulation with corrected parameters that are determined based on an analysis of the failure.

2.1. Input

To minimize human operations, common simulation parameters such as simulation box sizes, parameters for molecular dynamics simulations, and force field parameters are fixed to pre-optimized values as default. Only molecule-specific parameters and state-point range are required for each job. For calculations of binary mixtures to determine the VLE curves, the specific parameters are the types and numbers of molecule, and the temperature range. The pressures are determined by the simulations as explained below.

2.2. Builder

The coarse-grained model developed and validated in our previous study [24,25] was used in this work. Each gas molecule, CO_2 and light hydrocarbons (C_nH_{2n+2}, $n \leq 4$), is represented by a single CG bead. Larger *n*-alkane molecules (C_nH_{2n+2}, $n \geq 5$) are represented by the 3 + 2 mapping rule using three bead types: two terminal beads, C2E (CH_3CH_2-) and C3E ($CH_3CH_2CH_2$-), and a repeating chain bead C3M ($-CH_2CH_2CH_2$-). Based on the number of carbon atoms (*n*), the linear n-alkanes can be built as one of three cases:

$$n = \begin{cases} 3k & \text{C3E(C3M)}_{k-2}\text{C3E} \\ 3k+1 & \text{C2E(C3M)}_{k-1}\text{C2E} \\ 3k+2 & \text{C2E(C3M)}_{k-1}\text{C3E} \end{cases} \quad \begin{matrix} n = 5,\ 6,\ 7,\ \ldots\ldots \\ k = 1,\ 2,\ 3,\ \ldots\ldots \end{matrix} \tag{1}$$

As examples shown in Fig. 2, *n*-hexane is represented by a linear chain of two C3E beads; *n*-nonane comprises of two C3E beads and one C3M bead; *n*-decane includes two C3E beads and two C3M beads, and *n*-undecane is modeled as a linear chain of one C2E bead, two C3M beads and one C3E bead.

The dimensions of the simulation box are common parameters for all calculations reported in this work. A slab model which contains one liquid and one vapor phase is used, and each phase is large enough in the direction normal to the interface so that there is

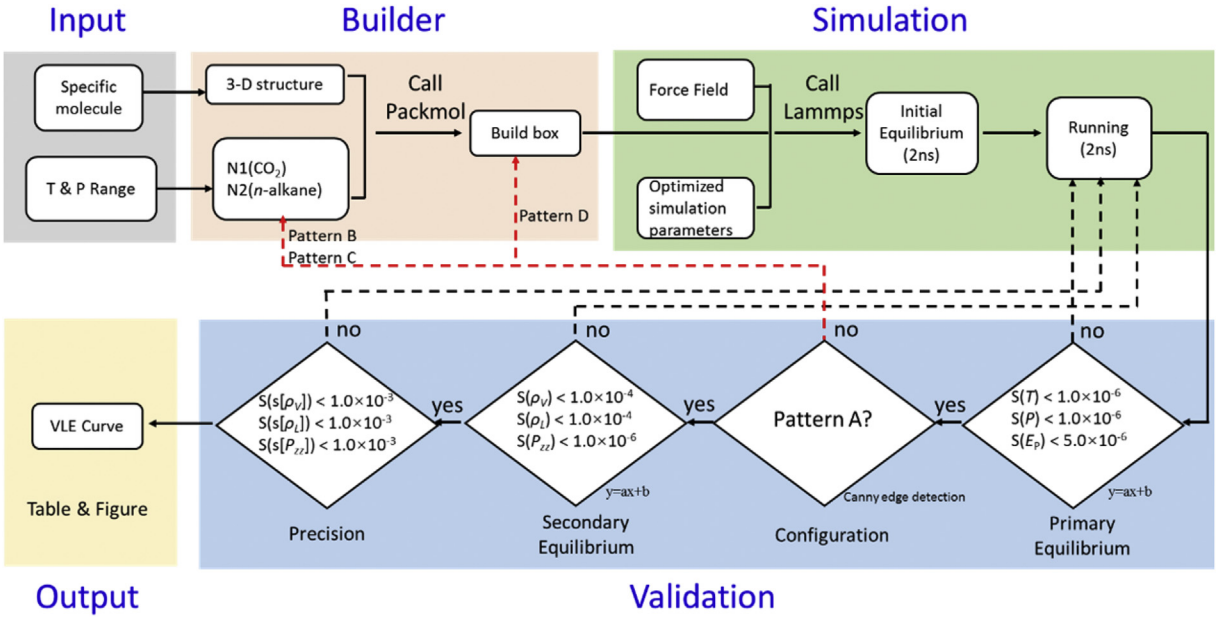

Fig. 1. The workflow of high-throughput computing procedure. The five functions – Input, Builder, Simulation, Validation and Output are represented by five blocks. Among them the Validation is the key component of automation.