Fluid Phase Equilibria 276 (2009) 37–45

Contents lists available at ScienceDirect

Fluid Phase Equilibria

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

Use of GAMESS/COSMO program in support of COSMO-SAC model applications in phase equilibrium prediction calculations

Shu Wang[a], Shiang-Tai Lin[b], Suphat Watanasiri[a], Chau-Chyun Chen[a,*]

[a] Aspen Technology Inc., 200 Wheeler Road, Burlington, MA 01803, United States
[b] Department of Chemical Engineering, National Taiwan University, Taipei, Taiwan, Republic of China

ARTICLE INFO

Article history:
Received 17 March 2008
Received in revised form 30 July 2008
Accepted 30 September 2008
Available online 14 October 2008

Keywords:
GAMESS
COSMO
COSMO-SAC
Group-contribution methods
Phase equilibrium

ABSTRACT

Thermodynamic models based on conductor-like screening models (COSMO) offer viable alternatives to existing group-contribution methods for the prediction of phase equilibria. Normally a COSMO-based model requires input of the distribution of screening charges on the molecular surface, aka. the sigma profile, determined from a specific quantum chemistry program and settings. For example, the COSMO-SAC model requires input of DMol^3 generated sigma profiles. In this paper, we investigate the proper settings for an open-source quantum chemistry package GAMESS in order to generate sigma profiles to be used directly in the COSMO-SAC model. The phase behaviors (VLE and VLLE) of 45 binary mixtures from 10 commonly used solvents and the solubilities of 4 complex drug compounds in these solvents calculated from DMol^3 and GAMESS generated sigma profiles are compared. While noticeable fine structure differences are observed in the individual sigma profiles for the same chemical compound generated from the two packages, it is found that the accuracy in the VLE/VLLE and solubility predictions from the two packages are comparable. Based on the systems we studied here, the open-source GAMESS/COSMO program with proper program settings could be used as an alternative sigma profile generation source in support of COSMO-SAC model applications in phase equilibrium prediction calculations.

1. Introduction

Knowledge of phase equilibrium is of paramount importance in the design, development, and optimization of chemical processes. With recent advances in computational chemistry, *a priori* predictions of mixture phase equilibria, i.e., without input of any experimental data, are becoming possible [1]. This is evidenced by the biennial competition on “Industrial Fluid Properties Simulation Challenge” hosted by National Institute of Standards and Technology (NIST) in the USA [2]. Computational chemistry approach provides a “bottom-up” route, i.e., solving a problem based on electronic and molecular interactions, and can serve as a complementary for most “top-down” experimental measurements, from which fundamental principles are deduced. While classical group contribution estimation methods may offer robust predictions, the quantum mechanical methods present a viable, though less reliable, alternative to estimate the phase behavior without requirements on extensive experimental data as training systems to identify group contribution parameters.

The conductor-like screening model for real solvent (COSMO-RS) [3–5] and its variants such as COSMO-SAC [6] and COSMO-RS(ol) [7] utilize the results of quantum mechanical solvation calculations and provide satisfactory predictive power for describing phase equilibria over a wide range of chemical species. The accuracy and efficiency of such methods make them promising candidates in real industrial applications. In COSMO-based models, the activity coefficient of a species in liquid mixture is determined by consideration of molecular surface interactions. Such interactions are assumed to be dominated by local (contacting surfaces) electrostatic interactions. Klamt suggested that the electrostatic energy of contacting surfaces can be well approximated using the screening charges on the molecule surface when the molecule is immersed in a perfect conductor (aka. COSMO calculations to generate σ-profiles).

It was suggested [5] that the description of surface screening charges can have a significant impact on the accuracy of phase equilibria predictions. Therefore, while many quantum mechanical packages are capable of performing COSMO calculations (e.g., DMol^3 [8], GAMESS [9], Gaussian [10], MOPAC [11], TURBOMOLE [12], etc.), COSMO-based models usually require the quantum calculations consistently from one specific program and parameter setting. For example, TURBOMOLE originally developed by Prof. Reinhart Ahlrichs at the University of Karlsruhe, and now is main-

---

* Corresponding author. Tel.: +1 781 221 6420; fax: +1 781 221 6410.
E-mail address: chauchyun.chen@aspentech.com (C.-C. Chen).

0378-3812/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.fluid.2008.09.021

tained by the Klamt group [13], is recommended for the generation of sigma profiles for COSMO-RS and COSMOtherm program [14] (COSMOtherm can use Gaussian as well, with necessary modification), and DMol^3 by Accelrys is used with COSMO-SAC model. Recently Mu et al. [15] published a comprehensive article comparing sigma profiles generated from TURBOMOLE, DMol^3 and Gaussian 03. They showed how different sigma profiles derived from different quantum packages, and how predictions of thermodynamic properties would be affected. Interestingly, based on their study, the differences between the various COSMO-RS/SAC variations and quantum chemistry packages are small for most types of mixtures.

The purpose of this investigation, parallel to Mu's work of COSMO-based methods, is to evaluate the usability of open-source quantum chemistry package GAMESS in performing COSMO calculations, thus providing an alternative channel of generating sigma profiles using resources other than DMol^3 or TURBOMOLE. In this article, 10 commonly used solvents are considered: water, acetone, dimethyl sulfoxide (DMSO), *n*-hexane, 1-octanol, methanol, ethyl acetate, *N*,*N*-dimethylformamide (DMF), acetic acid and acetonitrile. These 10 solvents are selected to cover solvents of various hydrophobicity, polarity, solvation strength, and hydrophilicity. Their sigma profiles are first generated from GAMESS/COSMO calculation, and then compared with those generated from DMol^3/COSMO in VT-COSMO database [16]. We also develop guidelines for performing COSMO calculation in GAMESS. Phase diagrams of all binary systems of these solvents are computed using the COSMO-SAC model in Aspen Plus Version 2006.5, and compared with experimental data. Finally, we carry out solubility predictions for large drug molecules using sigma profiles generated from GAMESS/COSMO.

2. Theory: GAMESS and COSMO

In quantum chemistry, molecular properties are obtained by solving many (N)-electron time independent Schrödinger equation

$$\mathbf{H}\Psi = [\mathbf{T}+\mathbf{U}+\mathbf{V}]\boldsymbol{\Psi} = \left[\sum_{i}^{N} -\frac{\hbar^2}{2m}\nabla_i^2 + \sum_{i<j} U(\vec{r}_i,\vec{r}_j) + \sum_{i}^{N} V(\vec{r}_i)\right]\boldsymbol{\Psi} = \mathbf{E}\Psi \qquad (1)$$

where $\mathbf{H}$ is the electronic molecular Hamiltonian, N is the total number of electrons of the molecule, $\mathbf{T}$ is the kinetic operator for electron, $\mathbf{U}$ is the operator for electron–electron interaction, and $\mathbf{V}$ is the operator for electron–external field interactions. The operators $\mathbf{T}$ and $\mathbf{U}$ are universal operators as they are the same for any system, once the level of theory is defined. The $\mathbf{V}$ operator is system dependent; the formulation of $\mathbf{V}$ changes with systems that the molecule is in, such as vacuum, solvent, magnet, etc. For an isolated system, the external operator contains only electron–nuclei interactions, i.e.,

$$V(\vec{r}_i) = \sum_{A}^{M} \frac{Z_A}{r_{iA}} \qquad (2)$$

where r_{iA} is the separation distance between electron i and nucleus A of charge Z_A. For a solvated system, solvent effects can be replaced by the apparent surface charges (or screening charges if solvated in a perfect conductor) located at the solute-solvent boundary. In such a case, the external operator contains additional terms from the apparent charges

$$V(\vec{r}_i) = \sum_{A}^{M} \frac{Z_A}{r_{iA}} + \sum_{s}^{n} \frac{q_s}{r_{is}} \qquad (3)$$

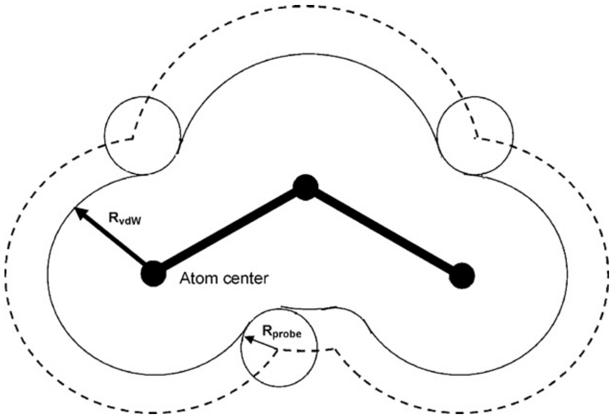

Fig. 1. Schematic drawing of SAS (dashed line) and SES (solid line): SAS (solvent-accessible surface) is described by the center of the probe (with a radius of R_{probe}) as it rolls along the atoms vdW spheres. The SES (solvent-excluded surface) is described by the closest point of the solvent probe as it rolls along the atoms vdW spheres, which compose of contact (convex) surfaces and reentrant (concave) surfaces.

where r_{is} is the separation distance between electron i and apparent surface charge s of charge q_s. Therefore, to proceed with the quantum mechanics (QM) calculation, one needs to provide the value of q_s and its location in space.

The locations of the surface charges are determined via some cavity and surface construction algorithm. When a solute molecule is placed into solvents, there are often spaces near the solute surface that the solvent molecule cannot penetrate into. In general, the solvent-accessible surface (SAS) and solvent-excluded surface (SES) are two of the most commonly used surface definitions. Consider a spherical probe rolling on the van der Waals surface of the solute. The trace of the probe center defines the SAS, and the contact between the probe and the solute defines the SES, as shown in Fig. 1. In GAMESS COSMO module [17–19], a procedure similar to GEPOL algorithm [20] has been used by Klamt [5] to create the SAS with the vdW radii of the solute atom, and the probe with 1.2 Å radius (R^{solv} = 1.2 Å). This SAS is then projected toward nuclei with a distance δ^{SC} (∼1.0 Å), since the effective screening charges will not be at the centers of the solvent molecules [21]. In this work, we set this distance equal to the probe radius. In this way, the SAS reduces to the SES. The advantage of this setting is that the radius of the probe would have little or no effect on the construction of the surface and the sigma profile. Detailed discussion can be found later in this article (see Section 4.2).

In COSMO calculations assuming infinite permittivity, the screening charges $\mathbf{q}$ distributed on the molecular surface are determined by the boundary condition of vanishing electrostatic potential on the cavity surface due to the existence of surrounding perfect conductor:

$$0 = \boldsymbol{\Phi}_{\text{total}} = \boldsymbol{\Phi}_{\text{s,in}} + \boldsymbol{\Phi}_{\sigma,\text{in}} = \mathbf{BQ} + \mathbf{Aq} \qquad (4)$$

where $\boldsymbol{\Phi}_{\text{s,in}}$ is the electrostatic surface potential arising from the solute charge distribution inside the cavity, which is the product of the Coulomb interaction matrix $\mathbf{B}$ and the source charge (solute nuclei and electrons) vector $\mathbf{Q}$. Here $\boldsymbol{\Phi}_{\sigma,\text{in}}$ is the reaction field potential resulting from the screening charges, which is modeled in COSMO as the product of a Coulomb interaction matrix $\mathbf{A}$ and the surface charge vector $\mathbf{q}$ as shown in the third equality of Eq. (4). $\mathbf{A}$ and $\mathbf{B}$ are predetermined interaction matrices and only depend on the coordinates of the segments on the surface and the coordinates of the nuclei of the solute molecule. Detailed generation of matrices $\mathbf{A}$ and $\mathbf{B}$ can be found in the Refs. [21,22]. In principle, the potential $\boldsymbol{\Phi}_{\text{s,in}}$ on any space position can be evaluated from wave function or charge distribution on quantum mechanics. However, to