2P089

Prediction Solvation Free Energy of Organic Compounds: Molecular Dynamics Simulation and QSPR Model Approach

(Kanazawa University*, Bandung Institute of Technology**, Japan Advanced Institute of Science and Technology***)
○Sri R. Natasia*,**, Hiroaki Saito*, Taku Mizukami***, Kazutomo Kawaguchi*, and Hidemi Nagao*

[Introduction] In recent years, data mining approaches have been adopted widely in many fields outside computer science. For example in chemical physics area, this approach is used to predict some physical properties of a compound such as solvation free energy. One of predictive method and provide a promising result is QSPR model.

Our purpose is to find correlation between experimental and calculated data on chemical material by using clustering technique combined with QSPR model. The clustering technique will be applied in selecting molecular descriptors, which are features representation of chemical material.

In this work, solvation free energy values of 14 organic compounds containing diverse organic functions were calculated using energy representation (ER) method [1]. As calculation of solvation free energy using this method was obtained in combination with MD simulation [2], previously we conducted three MD simulations for each organic compound. These three MD simulations were solution, pure solvent, and isolated solute systems. The generated trajectories from the simulations were then sampled to construct a histogram of solute-solvent pair interaction energy.

The values of solvation free energy of these organic compounds will also be predicted using QSPR model. The further aim of this study is to provide a better QSPR model based on MD simulation results in predicting solvation free energy of organic compounds.

[Material and Method] We computed values of solvation free energy of 14 neutral and structurally different organic compounds in water solvent. The compounds containing diverse organic functions; alkanes, alkenes, alkynes, aromatics, alcohols, aldehydes, ketones, amines, ethers, and esters.

All MD simulations were performed in GROMACS 4.6.5 using GAFF parameter for organic molecules which assigned by Antechamber program in AMBER tools. AM1-BCC charge model was used to calculate atomic point. The coordinate file, in SYBYL mol2 format [3], formerly loaded to Antechamber program in order to set these parameters. After generating AMBER topology and coordinate files, these files were converted to GROMACS topology and coordinate files using acpype conversion script [4]. These organic molecules were then solvated using GROMACS utilities in a simulation box of 32 Å x 32 Å x 32 Å consisting of a fixed 1000 water molecules of TIP3P model. For solution and pure solvent systems, periodic boundary condition were applied. The configuration were generated through NPT ensemble at 300 K and 1 bar using Nose-Hoover thermostat and Parinello-Rahman barostat at time constant of 1 ps. The Lennard-Jones potential was also used for intermolecular interaction with cutoff of 13.5 Å. The Coulomb interaction was handled by PME method with the interpolation order of 6.

The simulation was performed for 100 ps and was sampled every 5 fs for solution system. Meanwhile for pure solvent and isolated solute systems, we conducted 10 ps and 10 ns simulations which were sampled every 100 fs, respectively. All trajectories of the systems were then loaded to ER program. Firstly, the pair interaction between the solute and the solvent in solution system was calculated. Next as the reference system, the configuration for the solute and the solvent systems were constructed independently by inserting the solute into a test particle.

The histogram of solute-solvent interaction energy was then constructed, for both solution and reference systems. The latter, these histograms and correlation matrix of reference system were used as the inputs of a functional for calculation of solvation free energy.

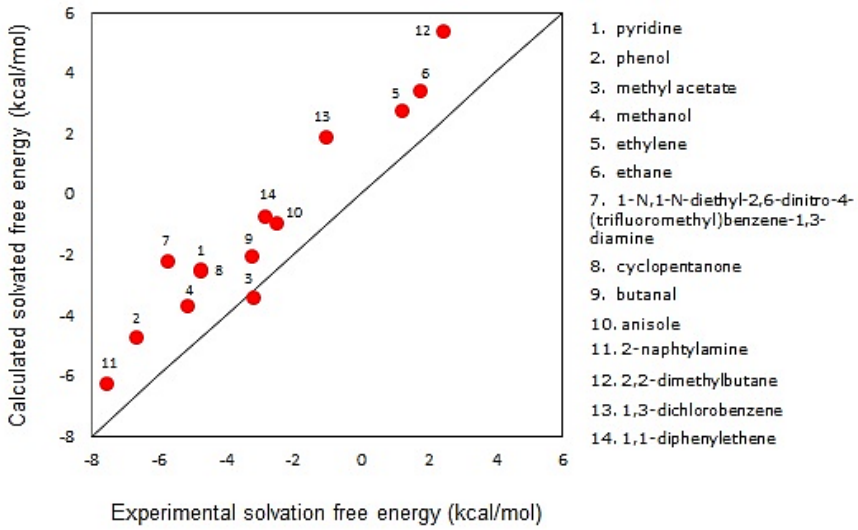

Figure1. Calculated solvation free energy versus experimental data

[**Results**] The calculated values of solvation free energy of 14 organic compounds and corresponding experimental data [3] are shown in Figure 1. Although the trend is a bit over estimated and shows the similar tendency, the values were successfully calculated by the difference is only a few kcal/mol.

We calculated SASA from hydrophobic and hydrophilic parts of molecule. We discovered that almost of all molecules consist of more hydrophobic than hydrophilic atoms, yet it is still soluble since both of parts are accessible by water molecules. It is also found that if one of the parts not accessible, the molecules can not dissolve in water solvent thus the calculated value of solvation free energy is positive. The average of surface area of hydrophobic and hydrophilic parts of 14 organic compounds are listed in Table 1.

Table 1 The Average of SASA from hydrophobic and hydrophilic parts of molecules

Organic compound	Number of atom	Hydrophilic Area (Å^2)	Hyrophobic Area (Å^2)
pyridine	11	15.40	3.26
phenol	13	6.44	13.80
methyl acetate	11	5.06	13.24
methanol	6	3.77	8.01
ethylene	6	12.97	0.00
ethane	8	0.00	13.37
1-N,1-N-diethyl-2,6-dinitro-4-(trifluoromethyl)benzene-1,3-diamine	35	23.60	16.85
cyclopentanone	14	4.42	14.69
butanal	13	6.80	12.19
anisole	16	3.50	19.79
2-naphtylamine	20	3.77	22.60
2,2-dimethylbutane	20	0.00	22.08
1,3-dichlorobenzene	12	0.00	24.77
1,1-diphenylethene	26	4.52	28.17

We will also construct a QSPR model to predict the values of solvation free energy. This model is developed using molecular descriptors from the compounds that generated by a chemical program package. A clustering technique will be applied in order to get a useful and informative features from the molecular descriptor. Furthermore, the predicted values from this model will then be validated by the experimental and simulation results.

Keyword: solvation free energy, organic compounds, molecular dynamic simulation, QSPR

References:

[1] Matubayasi N, Nakahara M. J. Chem. Phys. 113: 6070, 2000.
[2] Y Karino, MV Fedorov, N Matubayasi. Chem. Phys. Letters 496: 351-355, 2010.
[3] DL Mobley, et. al. J. Chem. Theory Comput. 5(2): 350-358, 2009.
[4] AW Sousa da Silva, WF Vranken. BMC Research Notes 5: 367, 2012.