



## Prediction of Thermodynamic Functions of Solvation by Dispersion-Corrected Density Functional Theory Calculations

Masao Fujisawa<sup>1\*</sup>, Hirohito Ikeda<sup>2</sup>, Tomonori Ohata<sup>2</sup>, Miho Yukawa<sup>2</sup>, Hatsumi Aki<sup>2</sup> and Takayoshi Kimura<sup>3</sup>

<sup>1</sup>Department of Biotechnological Science, Kinki University, Japan

<sup>2</sup>Department of Pharmaceutical Science, Fukuoka University, Japan

<sup>3</sup>Department of Chemistry, Kinki University, Japan

### ABSTRACT

The predicted hydration Gibbs energies were determined using vibrational frequency calculations at the B97D3/cc-pvtz level and APFD/cc-pvtz levels. These predicted values well reproduced the experimental hydration Gibbs energies. The APFD-predicted hydration Gibbs energies reproduced the experimental values better than the APF and B97D3 functions. The predicted solvation Gibbs energies of naproxen in several solvents were calculated at the B97D3/cc-pvtz and APFD/cc-pvtz levels and well reproduced the experimental solvation Gibbs energies.

**Keywords:** Solvation Gibbs energy; DFT-D; Vibrational frequencies; APFD; B97D3

### INTRODUCTION

Solvation phenomena play a significant role in chemical reactions and biomolecular recognition; however, it can be very difficult to determine the related thermodynamic quantities. For example, one must know the temperature dependence of the vaporization enthalpy or the vapor pressure to calculate the solvation Gibbs energy. Moreover, the dispersion energy is an important contributor to the solvation energy of a solute. Theoretical methods, including explicit or continuum solvent models, have been developed and applied in an attempt to predict the solvation Gibbs energy [1-5]. In explicit solvent models, there was an attempt to determine the hydration Gibbs energies computed using the molecular dynamics simulation and the energy-representation theory of solvation [6]. In this study, the solvation Gibbs energies for several solutes were calculated using the dispersion-corrected density functional theory (DFT-D) method; the results were then compared with experimental thermodynamic data. A conformation search was performed to decide the lowest energy structure. The obtained lowest energy conformer was the initial structure used for the DFT-D calculations in the gas phase. Vibrational frequencies were calculated for these optimized geometries at the same level of theory in the gas phase. The solute geometries were optimized in water using DFT-D functions. Similarly, vibrational frequencies were calculated for these optimized geometries at the same level of theory in water. The Gibbs energies of the solutes were determined from the vibrational frequencies in the gas phase and in water.

### EXPERIMENTAL SECTION

#### Geometry Optimization and Frequency Computation

We assumed that all relevant interactions of the perfectly screened molecules could be expressed as. A conformation search was performed with the MMFF94s force field using the COMPLEX6.2 program [7] to determine the lowest energy structure of the solute. This program has the following feature: the lowest energy conformer does not depend on the initial user-input structure. The obtained lowest energy conformer was the initial structure used for the DFT calculations. All DFT and DFT-D calculations were performed using the Gaussian 09 program [8]. The solute geometries were optimized in the gas phase using a polarized valence triple-zeta basis set (cc-pvtz) and B97D3 [9], and Grimme's modified B97 functional, which includes the empirical D3 dispersion correction with the Becke-Johnson damping method. In addition, optimization in the gas phase was performed using the Austin-Frisch-Petersson functional with dispersion correction (APFD) [10]. The vibrational frequencies were calculated for these optimized geometries at the same level of theory (i.e., B97D3/cc-pvtz and APFD/cc-pvtz). Similarly, the geometries of the solutes were optimized at the B97D3/cc-pvtz and APFD/cc-pvtz levels using the SMD solvation model [11], and the frequencies for

these geometries were calculated at the same levels of theory. For the SMD method, the solvation Gibbs energy is defined as the electrostatic and non-electrostatic interactions. This model attempts to use the electron density to estimate the solvent-accessible surface area and the atomic surface tensions to determine the electrostatic interaction. The electrostatic interaction of the solvation Gibbs energy was estimated by the IEF-PCM model. All of the thermodynamic quantities were determined using frequency calculations that included the zero-point energy.

### Solvation Gibbs energy

The internal thermal energy can also be obtained from Eq (1),

$$E = Nk_B T^2 \left( \frac{\partial \ln q}{\partial T} \right)_V \quad (1)$$

The total internal energy was determined from Eq (2),

$$E_{\text{tot}} = E_t + E_f + E_e + E_v, \quad (2)$$

where  $E_t$  is the translational energy,  $E_f$  is the rotational energy,  $E_e$  is the electronic energy, and  $E_v$  is the vibrational energy. The correlation to the enthalpy was determined according to Eq (3),

$$H_{\text{corr}} = E_{\text{tot}} + K_B T. \quad (3)$$

The entropy can be obtained from Eq (4),

$$S = Nk_B + Nk_B \ln \left( \frac{q(V, T)}{N} \right) + Nk_B T \left( \frac{\partial \ln q}{\partial T} \right)_V \quad (4)$$

The total entropy was determined from Eq (5),

$$S_{\text{tot}} = S_t + S_f + S_e + S_v \quad (5)$$

and the thermal correction to the Gibbs free energy was determined by Eq (6),

$$G_{\text{corr}} = H_{\text{corr}} - TS_{\text{tot}} \quad (6)$$

The total Gibbs energy was determined according to Eq (7),

$$G = \square_0 + G_{\text{corr}}, \quad (7)$$

where  $\square_0$  is the total electronic energy. Finally, the solvation Gibbs energy was determined from Eq (8),

$$\square G_{\text{solv}} = G(\text{in solvent}) - G(\text{in gas}). \quad (8)$$

## RESULTS AND DISCUSSION

The predicted hydration Gibbs energies were determined at the B97D3/cc-pvtz and APFD/cc-pvtz levels, and the experimental hydration Gibbs energies [12] are given in Table 1 along with the experimental hydration Gibbs energies for comparison.

To determine the accuracy of the predicted Gibbs energies, the root mean square error (RMSE) values, as defined in Eq (9) [13], were calculated and are listed in Graph 1, Graph 2, and Table 2.

$$RMSE = \sqrt{\frac{\sum (\text{observed} - \text{predicted})^2}{n}} \quad (9)$$

Table 1: Predicted hydration Gibbs energies of organic compounds at 298.15 K; All data are given in kJ mol<sup>-1</sup> \*ref 12

Molecules	Prediction		Experiment*
	APFD	B97D3	
Methane	7.8	8.5	9.37
Ethane	5.7	6.3	7.66
Propane	7.6	6.7	8.18
Butane	5.7	6.7	8.7
Pentane	7.6	9.6	9.76
Hexane	7.3	8.8	10.4
Heptane	9.1	9	10.96
Octane	8	9.7	12.1
Cyclopropane	-1.1	-0.3	3.13
Cyclopentane	5	4	5.02
Cyclohexane	3.9	5	5.14
Cycloheptane	3.2	4.3	3.33
Cyclooctane	3.5	4.3	3.58
Ethene	5	6	5.32
Propene	3.4	4.7	5.31
Butene	3.5	4.9	5.77
Pentene	4.2	5.7	6.96
Hexene	3.8	4.9	7.02
Octene	6.2	10	9.08
Butyne	-3.6	-2.2	-0.68
Pentyne	-4	-2.3	0.06
Hexyne	-2.9	-1.2	1.2
Heptyne	0	2.3	2.51
Octyne	-2.4	-4.5	2.97

Table 2: Predicted hydration Gibbs energies of organic compounds at 298.15 K (continued); All data are given in kJ mol<sup>-1</sup> \*ref 12

Molecules	Prediction		Experiment*
	APFD	B97D3	
Nonyne	-1.4	4.7	4.4
2-butanone	-16.2	-14.3	-15.22
2-pentanone	-19.1	-15.4	-14.76
3-pentanone	-14.5	-13.1	-14.28
2-hexanone	-10.7	-16.1	-13.76
2-heptanone	-18.7	-12.4	-12.72
4-heptanone	-15.1	-10	-12.24
2-octanone	-13.6	-14.8	-12.06
5-nonanone	-16.9	-13.2	-11.18
1,4-pentadiene	2.3	4.3	2.57
1,5-hexadiene	1	2.7	3.94
o-xylene	-3.2	-0.2	-3.77
m-xylene	-4.1	0.3	-3.51
p-xylene	-2.6	0.1	-3.27
Benzene	-4.5	-1.9	-0.87
1,4-dichlorobenzene	-5.1	-3.7	-4.22
1,3-dichlorobenzene	-2.4	-2.9	-4.11
1,2-dichlorobenzene	-4.8	-3.3	-5.71
1,2,4-trimethylbenzene	-1.9	1.7	-3.6
Ethylbenzene	-5.9	-2.5	-3.33
Butylbenzene	-2.4	-7.2	-1.66
Morpholine	-31.4	-30	-30.02
Pyridine	-17.5	-15.4	-18.58
3-ethylpyridine	-19.7	-17.9	-19.97
4-ethylpyridine	-20.8	-13.8	-20.65
2,3-dimethylpyridine	-21.1	-13.3	-20.19
3,4-dimethylpyridine	-28.1	-16.1	-21.64
4,5-dimethylpyridine	-27.1	-13.9	-20.26

Table 3: Predicted hydration Gibbs energies of organic compounds at 298.15 K (continued); all data given in kJ mol<sup>-1</sup>

\*ref 12

Molecules	Prediction		Experiment*
	APFD	B97D3	
Anthracene	-11.9	-7.2	-17.7
Naphthalene	-8.1	-4.5	-10.01
3-hydroxybenzaldehyde	-35	-32.2	-34.99
4-hydroxybenzaldehyde	-38.5	-34.8	-38.48
3-pentanol	-16	-13.6	-18.22
3-hexanol	-13.9	-13.5	-17.05
4-heptanol	-13.8	-10.8	-16.76
2methyl2pentanol	-13.8	-11.8	-16.44
4methyl2pentanol	-17.5	-15.1	-16.64
2methyl3pentanol	-16.3	-13.8	-16.26
1,2-ethanediol	-40.1	-33.3	-40.2
Phenol	-22.7	-22	-27.68
Cyclohexanol	-21	-18.9	-22.9
Cycloheptanol	-19.7	-18	-22.95
2-methoxyethanol	-22	-21.9	-28.31
2-ethoxyethanol	-23.8	-23.8	-27.64
2-butoxyethanol	-25	-22.4	-26.22
Propylamine	-12.7	-15.6	-18.37
Butaneamine	-12.3	-11	-17.97
Pentylamine	-13.6	-12	-17.14
Hexylamine	-13.6	-12.4	-16.87
Propanal	-13.1	-11.4	-14.66
Butanal	-13.4	-11	-14.4
Pentanal	-13.2	-11.5	-13.29
Hexanal	-12.5	-10.9	-12.68
Heptanal	-12.7	-10.2	-11.76
Octanal	-11.6	-21.18	-11.18
Nonanal	-8.7	-18.46	-9.56
Acetonitrile	-17.1	-32.38	-16.28

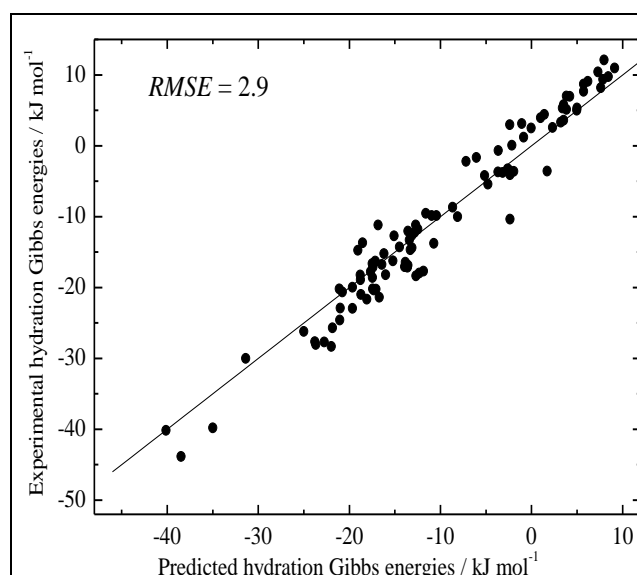


Figure 1: The correlation between the predicted hydration Gibbs energies at the APFD/cc-pvtz level and the experimental hydration Gibbs energies

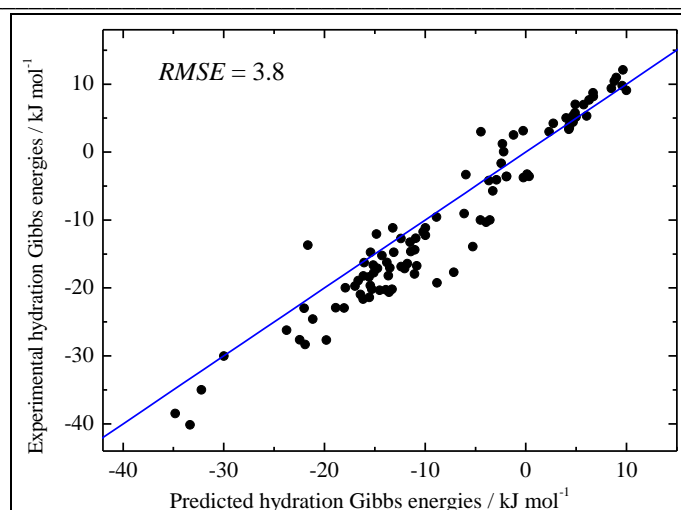


Figure 2: The correlation between the predicted hydration Gibbs energies at the B97D3/cc-pvtz level and the experimental hydration Gibbs energies

Table 4: The root mean square error (RMSE) for the predicted hydration Gibbs energies

Calculation level	RMSE
APFD/cc-pvtz	2.7
B97d3/cc-pvtz	3.8

The predicted hydration Gibbs energies of the normal alkanes comparatively reproduced the experimental values. The experimental hydration Gibbs energies of the normal alkanes became larger with an increasing number of carbon atoms in the normal alkanes. However, this increasing trend was not altogether reproduced. The hydration Gibbs energies predicted using the APFD functions for the cycloalkanes, except cyclopropane, correctly reproduced the experimental values. Similarly, the B97D3 function moderately reproduced the experimental cycloalkane values, except for cyclopropane. Both functions moderately reproduced the experimental values for the ketones and aromatic hydrocarbons, and the predicted values for the xylenes correctly reproduced the experimental values. The predicted hydration Gibbs energies using the B97D3 function of the normal alkenes correctly reproduced the experimental values, while both functions only slightly reproduced the experimental values of the normal alkynes. The predicted hydration Gibbs energies using the APFD function of the normal alkane-1-ols from  $N_c=5$  to 8 and the normal aldehydes considerably reproduced the experimental values. However, both functions only slightly reproduced the experimental values for the hydration Gibbs energies for the amines.

To clarify the effect of the dispersion-corrected function, the predicted solvation Gibbs energies of the *n*-alkane-1-ols were calculated at the B97D3/cc-pvtz, APFD/cc-pvtz, and APF/cc-pvtz levels, and are listed in Table 3.

Table 5: Predicted hydration Gibbs energies of *n*-alkane-1-ols at 298.15 K. All data are given in kJ mol<sup>-1</sup>

\*ref 12

Molecules	Prediction			Experiment*
	APFD	APF	B97D3	
Methanol	-16.7	-16.7	-15.5	-21.4
Ethanol	-18.7	-18.2	-16.4	-21
1-Propanol	-17.4	-16.8	-15.3	-21
1-Butanol	-19.1	-18.5	-17	-19.75
1-Pentanol	-18.8	-18.1	-16.6	-18.93
1-Hexanol	-18.8	-18	-16.1	-18.26
1-Heptanol	-17.7	-17.7	-15.1	-17.76
1-Octanol	-17.4	-16.1	-14.7	-17.13

To confirm the applicability in solvents other than water, the solvation Gibbs energy of naproxen was calculated at the B97D3/cc-pvtz and APFD/cc-pvtz level, and the results are listed with the experimental values [14] in Table 4.

The solvation Gibbs energy of naproxen in organic solvents and water was predicted; the B97D3 method better reproduces the experimental values than the APFD method (the RMSE values are shown for each method in Table 5). Consequently, the solvation Gibbs energies that compare well with the experimental values can be predicted using the DFT-D method.

Table 6: Predicted solvation Gibbs energies of naproxen at 298.15 K. All data are given in kJ mol<sup>-1</sup>

\*ref 14

Solvent	Prediction		Experiment
	APFD	B97D3	
Benzene	-40.2	-43.5	-44.6
<i>n</i> -Hexane	-35.6	-39.1	-41.3
Ethanol	-50.5	-48.6	-47.3
1-Propanol	-51.4	-50.9	-47.6
1-Butanol	-49.9	-48.1	-47.9
1-Pentanol	-49.1	-47.4	-48
1-Hexanol	-47.7	-45.9	-48.3
1-Heptanol	-46.2	-44.5	-48.8
1-Octanol	-40.2	-43.2	-48

Table 7: The root mean square error (RMSE) for the predicted solvation Gibbs energies

Calculation level	RMSE
APFD/cc-pvtz	4.1
B97d3/cc-pvtz	2.7

In conclusion, the predicted hydration Gibbs energies, which were determined by performing DFT-D vibrational frequency calculations, reproduced the experimental hydration Gibbs energies. In particular, the APFD functions correctly reproduced the experimental values. The predicted solvation Gibbs energies using this method well reproduced the experimental solvation Gibbs energies. This method is also applicable in solvents other than water.

## REFERENCES

- [1] E Gallicchio; LY Zhang; RM Levy. *J. Comput. Chem.*, **2002**, 23(5), 517.
- [2] D Riccardi; H-B Guo; JM Parks; B Gu; L Liang; JC Smith. *J. Chem. Theory Comput.*, **2013**, 9(1), 555.
- [3] D Shivakumar; J Williams; Y Wu; W Damm; J Shelley; W Sherman. *J. Chem. Theory Comput.*, **2010**, 6(5), 1509.
- [4] Y Takano; KN Houk. *J. Chem. Theory Comput.*, **2005**, 1(1), 70.
- [5] CJ Cramer; DG Truhlar. *Science*, **1992**, 256(5054), 213.
- [6] N Matubayasi; M Nakahara. *J. Chem. Phys.*, **2003**, 119(18), 9686.
- [7] Conflex 6.2, Conflex Corp., Japan.
- [8] Gaussian 09, Revision A.1. MJ Frisch; GW Trucks; HB Schlegel; GE Scuseria; MA Robb; JR Cheeseman; G Scalmani; V Barone; B Mennucci; GA Petersson; H Nakatsuji; M Caricato; X. Li; HP Hratchian; AF Izmaylov; J Bloino; G Zheng; JL Sonnenberg; M Hada; M Ehara; K Toyota; R Fukuda; J Hasegawa; M Ishida; T Nakajima; Y Honda; O Kitao; H Nakai; T Vreven; JA Montgomery, Jr.; JE Peralta; F Ogliaro; M Bearpark; JJ Heyd; E Brothers; KN Kudin; VN Staroverov; R Kobayashi; J Normand; K Raghavachari; A Rendell; JC Burant; SS Iyengar; J Tomasi; M Cossi; N Rega; JM Millam; M Klene; JE Knox; JB Cross; V Bakken; C Adamo; J Jaramillo; R Gomperts; RE Stratmann; O Yazyev; AJ Austin; R Cammi; C Pomelli; JW Ochterski; RL Martin; K Morokuma; VG Zakrzewski; GA Voth; P Salvador; JJ Dannenberg; S Dapprich; AD Daniels; Ö Farkas; JB Foresman; JV Ortiz; J Cioslowski; DJ Fox. Gaussian, Inc., Wallingford CT, 2009.
- [9] A Austin; GA Petersson; MJ Frisch; FJ Dobek; G Scalmani; K Throssell. *J. Chem. Theory Comput.*, **2012**, 8(12), 4989.
- [10] S Grimme; S Ehrlich; L Goerigk. *J. Comput. Chem.*, **2011**, 32(7), 1456.
- [11] AV Marenich; CJ Cramer; DG Truhlar. *J. Phys. Chem. B*, **2009**, 113(18), 6378.
- [12] S Cabani; P Gianni; V Mollica; L Lepori. *J. Soln. Chem.*, **1981**, 10(8), 563.
- [13] SG, Machatha; P Bustamante; SH Yalkowsky. *Int. J. Pharm.*, **2004**, 283(1-2), 83.
- [14] GL Perlovich; SV Kurkov; AN Kinchin; A Bauer-Brandl. *Eur. J. Pharm. Biopharm.*, **2004**, 57(2), 411.