Electronic Supplementary Material

A computational toolbox for molecular property prediction based on quantum mechanics and quantitative structure-property relationship

Qilei Liu, Yinke Jiang, Lei Zhang (✉), Jian Du

Institute of Chemical Process Systems Engineering, School of Chemical Engineering, Dalian University of Technology, Dalian 116024, China

E-mail: keleiz@dlut.edu.cn

Appendixes

Appendix A

Some commonly used QM methods for predictions of thermodynamic properties are listed in Table A1. The experimental ΔH_{vap}^{θ}, ΔH_{sub}^{θ}, ΔG_{vap}^{θ}, ΔG_{sub}^{θ}, ΔS_{vap}^{θ}, ΔS_{sub}^{θ} and recommended spin multiplicity for each element being in its most stable state at 1 bar and 298.15 K are listed in Table A2.

Table A1. Some commonly used QM methods and basis sets for predictions of thermodynamic properties.

QM method	Computational cost	Prediction accuracy	Molecular size limitation
G4	Highest	Highest	Less than 30 atoms
G4(MP2)	Very high	Very high	Less than 30 atoms
CBS-QB3	High	High	Less than 30 atoms
G3(MP2)	High	High	Less than 30 atoms
CBS-4M	Not high	Not high	More than 30 atoms
Unpackaged hybrid method	Low	High	More than 30 atoms

Note: the methods of G4, G4(MP2), CBS-QB3, G3(MP2) and CBS-4M are all packaged hybrid methods, which are directly called in Gaussian software, while this paper employs the unpackaged

hybrid method ("B3LYP/6-31G(d)" for geometry optimization and frequency analysis, and "M062X/def2TZVP em=GD3" for single point energy calculation) to predict thermodynamic properties ($\boldsymbol{H_g^\theta}$, $\boldsymbol{G_g^\theta}$, $\boldsymbol{S_g^\theta}$) based on our knowledge.

Table A2. The experimental $\boldsymbol{\Delta H_{vap}^\theta}$, $\boldsymbol{\Delta H_{sub}^\theta}$, $\boldsymbol{\Delta G_{vap}^\theta}$, $\boldsymbol{\Delta G_{sub}^\theta}$, $\boldsymbol{\Delta S_{vap}^\theta}$, $\boldsymbol{\Delta S_{sub}^\theta}$ and recommended spin multiplicity for each element being in its most stable state at 1 bar and 298.15 K.

Element	$\boldsymbol{\Delta H_{vap}^\theta}$ **or** $\boldsymbol{\Delta H_{sub}^\theta}$**/kJ/mol**	$\boldsymbol{\Delta G_{vap}^\theta}$ **or** $\boldsymbol{\Delta G_{sub}^\theta}$**/kJ/mol**	$\boldsymbol{\Delta S_{vap}^\theta}$ **or** $\boldsymbol{\Delta S_{sub}^\theta}$**/J/mol/K**	**Spin multiplicity**
$H_2(g)$	0.0	0.0	0.0	1
$C(s)$	716.7	671.3	152.4	3
$N_2(g)$	0.0	0.0	0.0	1
$O_2(g)$	0.0	0.0	0.0	3
$F_2(g)$	0.0	0.0	0.0	1
$P(s)$	316.5	280.1	122.0	4
$S(s)$	277.2	236.7	135.8	3
$Cl_2(g)$	0.0	0.0	0.0	1
$Br_2(l)$	30.9	3.1	93.3	1
$I_2(s)$	62.4	19.3	144.5	1

Appendix B

Table B1. The raw data of comparison results between QM calculated values and experimental data for $\Delta_f H_g^\theta$, $\Delta_f G_g^\theta$ and S_g^θ at 298.15 K of 22 representative compounds (blank means experimental data are unavailable in database).

No.	**Compound**	**CAS number**	**Calculated** $\boldsymbol{\Delta_f H_g^\theta}$**/kJ/mol**	**Experimental** $\boldsymbol{\Delta_f H_g^\theta}$**/kJ/mol**	**Calculated** $\boldsymbol{\Delta_f G_g^\theta}$**/kJ/mol**	**Experimental** $\boldsymbol{\Delta_f G_g^\theta}$**/kJ/mol**	**Calculated** $\boldsymbol{S_g^\theta}$**/J/mol/K**	**Experimental** $\boldsymbol{S_g^\theta}$**/J/mol/K**
1	Methane	74-82-8	-73.7	-74.6	-58.7	-50.5	206.7	186.3
2	1-Octene	111-66-0	-99.3	-81.4	77.0	104.2	423.0	462.5
3	Ethylene	74-85-1	51.5	52.5	61.9	68.4	218.9	219.3
4	1-Butyne	107-00-6	157.2	165.2	183.3	202.1	289.2	290.8
5	1-Bromohexane	111-25-1	-164.7	-150.1	-19.2		413.9	
6	Bromoethane	74-96-4	-70.0	-63.6	-37.7	-23.9	286.4	286.7
7	Propanol	71-23-8	-261.0	-256.0	-167.9	-161.8	300.8	322.7
8	Ether	60-29-7	-268.6	-252.0	-147.4	-122.3	333.4	342.7
9	Benzaldehyde	100-52-7	-52.4	-37.0	-11.8		332.5	
10	Acetaldehyde	75-07-0	-164.4	-166.1	-133.2	-133.0	251.5	263.8
11	Acetone	67-64-1	-220.5	-217.0	-161.2	-152.7	283.8	295.3
12	Methyl ethyl ketone	78-93-3	-247.2	-238.6	-159.0		313.7	339.9
13	Hexanoic acid	142-62-1	-509.2	-512.0	-343.9		411.4	
14	Acetic acid	64-19-7	-411.7	-432.2	-360.2	-374.2	285.6	283.5
15	Ethyl acetate	141-78-6	-457.5	-445.4	-350.7		353.7	362.8
16	Methyl formate	107-31-3	-368.1	-357.4	-316.1	-297.2	284.4	285.3
17	Propionitrile	107-12-0	46.4	51.5	84.1		284.4	
18	Acetonitrile	75-05-8	67.3	74.0	77.0	91.9	251.4	243.4
19	Propylamine	107-10-8	-81.2	-70.0	28.6	39.8	303.1	325.1

20	Methanethiol	74-93-1	-27.6	-22.9	-18.9	-9.9	253.2	255.1
21	Water	7732-18-5	-226.9	-241.8	-215.6	-228.6	194.6	188.8
22	Carbon dioxide	124-38-9	-390.5	-393.5	-394.3	-394.4	214.1	213.8

Appendix C

Table C1. The COSMO-SAC model parameters[39].

Parameter	**Value**	**Unit**
a_{eff}	6.4547	Å^2
A_{ES}	5501.09	$(\text{kcal/mol}) \cdot (\text{Å}^4/\text{e}^2)$
B_{ES}	1.8822×10^8	$(\text{kcal/mol}) \cdot (\text{Å}^4/\text{e}^2) \cdot \text{K}^2$
c_{OH-OH}	4900.06	$(\text{kcal/mol}) \cdot (\text{Å}^4/\text{e}^2)$
c_{OT-OT}	1378.62	$(\text{kcal/mol}) \cdot (\text{Å}^4/\text{e}^2)$
c_{OH-OT}	4211.67	$(\text{kcal/mol}) \cdot (\text{Å}^4/\text{e}^2)$