THERMODYNAMICS AND CHEMICAL ENGINEERING DATA
Chinese Journal of Chemical Engineering, **17**(3) 468—472 (2009)

Position Group Contribution Method for Estimation of Melting Point of Organic Compounds

WANG Qiang (王强)[1,*], MA Peisheng (马沛生)[2] and NENG Shifeng (能士峰)[1,2]
[1] School of Material Science and Chemical Engineering, Tianjin University of Science and Technology, Tianjin 300457, China
[2] School of Chemical Engineering and Technology, Tianjin University, Tianjin 300072, China

Abstract A new method is proposed based on the position group contribution additivity for the prediction of melting points of covalent compounds. The characteristics of this method are the use of position distribution function, which could distinguish between most isomers including *cis* or *trans* structure of organic compounds. Contributions for hydrocarbons and hydrocarbon derivatives containing oxygen, nitrogen, chlorine, bromine and sulfur, are given. Results are compared with those by the most commonly used estimating methods. The average derivation for prediction of normal melting temperature of 730 compounds is 14.46 K, compared to 29.33 K with the method of Joback, and 27.81 K with the method of Constantinou-Gani. The present method is not only more accurate, but also much simpler and more stable.

Keywords melting point, prediction, position group contribution

1 INTRODUCTION

Knowledge of physicochemical properties of organic compounds is essential for predicting their behavior. Melting point is one of the most widely used physical properties in chemical engineering design and experiments. Although vast amount of experimental data are available in literature, but the presence of impurity or thermal instability of some compounds may thwart the experimental measurement. Under such conditions, mathematical models could be used to provide a reasonable estimation of the properties [1–8]. Some of the group contribution methods for estimation of melting points were developed [5–7]. Joback-Reid developed a first-order group contribution method with 40 groups for organic compounds containing halogens, oxygen, nitrogen and sulfur [5]. Constantinou-Gani introduced a two-level group contribution scheme consisting of 63 first-order groups and 40 second-order groups [6]. Marrero-Gani modified this model using the third-order groups that could account for more complex heterocyclic and large polyfunctional alicyclic compounds [7].

Methods of quantitative structure-property relationship (QSPR) also could be used to estimate melting points [8–10]. This method is based on specific families of compounds, such as alkanes [11], aldehydes, amines, and ketones [12], benzenes [10] and pyridinium bromides [13]. Katritzky [14] reviewed the accuracy of QSPR methods from very well (standard deviation = 15.1 K for *n*-alkanes) to similar accuracy as group contributions (standard deviation = 36.1 K for pyridines). Chickos *et al.* [15] developed a parabolic model for estimating the melting points of homologous series. Karthikeyyan *et al.* [16] reported prediction of melting point based on diverse compound data and artificial neutral network. Jain and Yalkowsky [1, 3] compared the melting point predictions from UPPER to MPBPWIN [3] method for over 2200 organic compounds. The mean absolute error (MAE) in prediction using UPPER is 30.1 K while MPBPWIN gives an AAE of 44.5 K.

The mean absolute errors, number of compounds and compound categories by some of the methods discussed above are listed in Table 1.

Table 1 Mean absolute errors ($\bar{\varepsilon}$) in general estimation methods for melting point

Method	Number of compounds	$\bar{\varepsilon}$ /K	Family
Tsakanikas *et al.* [17]	72	15.7	alkanes
Simamora *et al.* [18]	123	47.6	rigid aromatic
Krzyzaniak *et al.* [19]	497	30.3	non H-bonding aliphatic
Zhao *et al.* [20]	794	39.9	aliphatic
Yalkowsky *et al.* [4]	918	31.7	from boiling point
Joback *et al.* [5]	1117	43.0	organic
Constantinou *et al.* [6]	1007	34.6	organic
Marrero *et al.* [7]	1170	32.4	organic
UPPER [21]	2200	30.1	organic
MPBPWIN [1,3]	2200	44.5	organic

As our preliminary work on the critical temperature [22], critical pressure [23], critical volume [24] and the normal boiling point [25] showed significant improvements with respect to the published methods, it was decided to invest further effort into the development of a new group contribution method for melting point of organics. We proposed a position group contribution method for the prediction of critical parameters of organic compounds, in which only the knowledge of their chemical structure is required [22–25]. A

Received 2008-11-11, accepted 2009-02-25.
* To whom correspondence should be addressed. E-mail: wang_q@tust.edu.cn

most important point must be claimed is that all these properties T_c, P_c, V_c and t_b mentioned above and t_m in this work, are estimated by the same universal position distribution function proposed.

2 EXPERIMENTAL DATA

Total 730 compounds containing carbon, hydrogen, oxygen, nitrogen, chlorine, bromine and sulphur were used for the determination of group contributions. The detail list is given in Fig. 1, and includes linear and branched alkanes (158), cycloalkanes (41), alkenes (63), aromatics (40), alcohols (66), aldehydes and ketones (67), acids (66), phenols and ether oxides(19), esters (126), chloro and bromoalkanes (25), amines (27), pyridines (10), alkane thiols and thioethers (22).

The experimental data, from the TRC Thermodynamic Tables [26], give the critical properties, normal boiling points and melting points for a large number of hydrocarbons and derivatives. The handbooks by Ma [27] and Poling *et al*. [28] also provide an abundance of property data about organic compounds.

3 COMPUTATIONS

3.1 Calculation method

The first step was testing correlations to represent the properties. Only one-parameter contribution was considered for each group at first. The melting point function was constructed by the contributions from all the groups as well as position distribution function. Benson's second order groups and a few of the third order groups were applied into this works. The details of the group classification are described in Table 2. Here, the melting point is expressed as follows [22].

$$T_m = T_{m0} + \sum_i A_i N_i + \sum_j A_j \tanh\left(N_j / N\right) + \sum_k A_k P_k + a_1 \exp\left(1/M_w\right) + a_2 \exp\left(1/N\right) \tag{1}$$

$$N = \sum_i N_i + \sum_j N_j \tag{2}$$

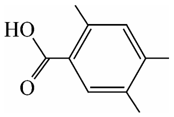
Where parameter A_i or A_j reveals i or j group contributions, which allowed the minimization of the residual estimation error by regression. N_i represents the number of groups in which carbon element forms the centre of the group, N_j represents the number of groups in which non-carbon element forms the center, N is the total number of groups, P_k characterizes position factor, T_{mo} is 5963.486 K, and M_W is the molecular weight.

Moreover, the position corrections were taken into account due to the longer distance interactions. Some corrections for interactions through benzene or pyridine rings were obtained, which was the result in a better distinction between series of chain-branched aromatic isomers. Parameters for five- and six-membered saturated hydrocarbon rings were also determined. We present the correlations that give the best estimations in Table 3.

3.2 Application examples

Some examples for computation of T_m for the organics are given follows.

Example 1 Estimation of the melting point of 2,4,5-trimethylbenzoic acid



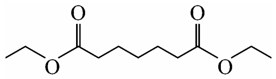
This compound is decomposed into the position groups as follows:

1 CO—(CH)(O); 2 Cb—(H);
3 Cb—(C); 3 C—(Cb)(H)$_3$;
1 O—(CO)(H); 1 Cb—(COOH)
Total number of group: $N = 11$
2 *ortho* correction; 2 *meta* correction
Molecular weight: $M_W = 164.21$

$$\begin{aligned} Z_c &= 5963.486 + 49.96\times 3 + 34.246\times\tanh(1/11) + \\ &6.224\times\tanh(2/11) - 331.51\times\tanh(3/11) + \\ &369.423\times\tanh(1/11) + 369.423\times\tanh(1/11) + \\ &1181.043\times\tanh(1/11) + 0.777\times 2 - \\ &7.374\times 2 + 51.127\times\exp(1/11) - \\ &5758.997\times\exp(1/164.21) = 418.54\ \text{K} \end{aligned}$$

The calculated result is 418.54 K while the experimental melting point is 425.15 K.

Example 2 Estimation of the melting point of diethyl pimelate



This compound is decomposed into the position groups as follows:

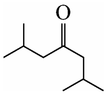
2 C—(CH$_2$)(H)$_3$; 3 C—(C)$_2$(H)$_2$;
2 C—(C)(CO)(H)$_2$; 2 C—(C)(O)(H)$_2$;
2 CO—(CH$_2$)(O); 2 O—(CH$_2$)(CO)
Total number of group: $N = 13$
Molecular weight: $M_W = 216.29$

From the contributions in Table 2, the melting point can be calculated from Eq. (1).

$$\begin{aligned} Z_c &= 5963.486 - 31.996\times 2 + 7.229\times 3 + \\ &35.135\times 2 - 21.4\times 2 - 67.931\times\tanh(2/13) + \\ &187.022\times\tanh(2/13) + 51.127\times\exp(1/13) - \\ &5758.997\times\exp(1/216.29) = 236.35\ \text{K} \end{aligned}$$

The calculated result is 236.36 K while the experimental melting point is 249.15 K.

Example 3 Estimation of the melting point of 2,6-dimethyl-4-heptanone



This compound is decomposed into the position groups as follows:

4 C—(CH)(H)$_3$; 2 C—(C)$_3$(H);
2 C—(C)(CO)(H)$_2$; 1 CO—(C)$_2$
Total number of group: $N = 9$

The position of —(CH)< group: the position factor $P = 2+6 = 8$

The position of (CO)— group: the position factor $P = 4$

Molecular weight: $M_W = 142.24$

From the contributions in Table 2, The melting point is estimated by Eq. (1):

$$\begin{aligned} Z_c &= 5963.486 - 20.213\times 4 + 19.312\times 2 + \\ &35.135\times 2 + 14.023\times \tanh(1/9) - \\ &3.791\times 8 - 0.658\times 4 + 51.127\times \exp(1/9) - \\ &5758.997\times \exp(1/142.24) = 217.63\ \mathrm{K} \end{aligned}$$

The calculated result is 217.63 K, while the experimental melting point is 227.15 K.

Table 2 Position group contributions for the prediction of T_m

Group	*A*	Group	*A*
C—(CH$_3$)(H)$_3$	–1.846	C—(C)(Br)(H)$_2$	–42.806
C—(CH$_2$)(H)$_3$	–31.996	C—(C)$_2$(Br)(H)	30.211
C—(CH)(H)$_3$	–20.213	C—(C)$_3$(Br)	107.607
C—(C)(H)$_3$	–19.810	C—(C)(Cl)(H)$_2$	–25.042
C—(C)$_2$(H)$_2$	7.229	C—(C)$_2$(Cl)(H)	–22.104
C—(C)$_3$(H)	19.312	C—(C)$_3$(Cl)	98.846
C—(C)$_4$	80.693	C—(C)(Cl)$_2$(H)	–8.873
Cd—(H)(O)	118.537	C—(S)(H)$_3$	–57.954
Cd—(H)$_2$	–211.880	C—(C)(S)(H)$_2$	–29.606
Cd—(C)(H)	12.761	C—(C)$_2$(S)(H)	–29.485
C—(Cd)(C)(H)$_2$	–9.637	C—(C)$_3$(S)	65.204
C—(Cd)(H)$_3$	–22.920	Cb—(N)	–270.079
Cd—(C)$_2$	258.506	C—(N)(H)$_3$	–57.293
C—(Cd)(C)$_2$(H)	8.663	C—(C)(N)(H)$_2$	–12.077
Cd—(Cd)(H)	43.026	C—(C)$_2$(N)(H)	–11.457
C—(Cd)(O)(H)$_2$	–30.612	C—(C)$_3$(N)	23.152
C—(O)(H)$_3$	–31.772	C—(C)(CN)(H)$_2$	4.921
C—(CO)(H)$_3$	–7.451	C—(C)$_2$(CN)(H)	50.171
C—(C)(CO)(H)$_2$	35.135	C—(C)$_3$(CN)	147.640
C—(C)$_2$(CO)(H)	51.707	N—(C)(H)$_2$	277.730
C—(C)$_3$(CO)	140.414	N—(C)$_2$(H)	488.698
C—(C)(O)(H)$_2$	–21.400	N—(C)$_3$	625.880
C—(C)$_2$(O)(H)	–5.093	N—(Cb)(H)$_2$	2341.412
C—(C)$_3$(O)	122.479	N—(Cb)(C)(H)	2665.735
CO—(CH$_3$)(O)	7.884	N—(Cb)(C)$_2$	3997.599
CO—(CH$_2$)(O)	–67.931	NI—(Cb)$_2$	263.807
CO—(CH)(O)	34.246	S—(C)(H)	82.567
CO—(C)(O)	–126.384	S—(C)$_2$	363.527
CO—(O)(H)	–31.753	*ortho* correction①	0.777
CO—(C)(H)	–54.751	*meta* correction①	–7.374
CO—(C)$_2$	14.023	cyclopropane correction	–13.306
CO—(Cd)(O)	195.607	cyclobutane correction	–13.011
Cb—(H)	6.224	cyclopentane correction	–64.530
Cb—(C)	–331.510	cyclohexane correction	–39.622
C—(Cb)(H)$_3$	49.960	*Co*②	–12.837
C—(Cb)(C)(H)$_2$	43.526	*Cm*②	3.941
C—(Cb)(C)$_2$(H)	55.777	*Cp*②	–4.055
C—(Cb)(C)$_3$	86.941	—(CH)< position factor	–3.791
Cb—(O)	138.092	>(C)< position factor	–7.988
Cb—(COOH)	1181.043	double bond position factor	3.355
O—(Cb)(H)	550.342	Hydroxyl position factor	12.473
O—(C)(H)	108.175	*trans* or *cis* structure factor③	–4.361
O—(C)$_2$	360.491	carbonyl position factor	–0.658
O—(CO)(CH$_3$)	119.462	phenol position factor	2.058
O—(CO)(CH$_2$)	187.022	O—(C)$_3$ position factor	–36.162
O—(CO)(CH)	229.497	exp(1/*N*)	51.127
O—(CO)(H)	369.423	exp(1/M_w)	–5758.997
T_{m0}	5963.486		

① *Ortho* and *meta* corrections consider interactions between alkyl chains through a benzene ring.

② Corrections for pyridines: *Co*, *Cm* and *Cp* pyridine corrections take into account alkyl ligands in position *ortho*, *meta* and *para* with respect to the N element, respectively.

③ *Trans* or *cis* correction: *cis*-structure correction is 1 and *trans* structure is −1 .

Note: The formalism of Benson groups is described in Ref. [1]. The first symbol represents the element that forms the centre of the group. The symbols between parentheses represent the elements to which it is linked. Usual symbols are used to represent the elements in their normal valence state. Elements in other valence states are distinguished by using additional characters; furthermore, different symbols represent multiple bonded carbons, depending on the element at the other end of the multiple bond: Cd, carbon forming a double bond with another carbon; Cb, carbon involved in a benzene or a pyridine ring; CO, C O group; CN, C≡N group; NI, nitrogen of the imide (CN) function. Also used for the nitrogen of pyridine derivatives. The pyridine ring is considered as five Cb and one NI. For instance, pyridine will be decomposed into: one NI—(Cb)$_2$, five Cb—(H). Unlike Benson, who included contributions for gauche interactions to ideal gas phase properties, no contribution for the gauche interaction to melting points and vaporization enthalpies were considered in this work.

4 RESULTS AND DISCUSSION

The results obtained by the new position group contribution method were compared to the experimental melting points, the values predicted by the method of Joback, which is the most extended commonly used method, and those by the method of Constantinou and Gani. The results for the whole set of reference compounds are presented in Table 3, which we also summarize the mean absolute errors ($\overline{\varepsilon}$) and mean absolute percentage error ($\overline{\delta}$) between experimental and predicted values:

Table 3 Comparison of T_m predicted by our model with those methods of Joback and Constantinou

Chemical family	No. of compounds	Joback		Constantinou		This study	
		$\bar{\varepsilon}$/K	$\bar{\delta}$/%	$\bar{\varepsilon}$/K	$\bar{\delta}$/%	$\bar{\varepsilon}$/K	$\bar{\delta}$/%
alkanes	158	19.87	9.68	22.31	13.81	13.89	7.15
cycloalkanes	41	21.46	11.12	24.13	13.52	16.86	8.56
alkenes	63	15.75	10.93	12.96	8.80	10.12	6.61
aromatics	40	30.42	14.08	32.20	14.52	22.44	9.91
ketones and aldehydes	67	16.18	6.57	15.98	7.19	9.06	3.89
alcohols	66	42.63	21.85	36.27	18.76	23.75	11.02
acids	66	60.31	17.82	57.02	16.20	20.49	5.91
phenols and Ether oxides	19	39.64	14.53	18.52	7.30	12.25	4.82
esters	126	35.19	14.22	28.95	12.11	11.19	4.79
bromo, chloroalkanes and nitriles	25	31.30	16.63	19.96	11.53	8.27	4.89
amines and anilines	27	20.64	9.91	55.10	24.72	8.44	4.10
Pyridines	10	26.14	11.31	28.20	12.26	15.39	6.44
alkane thiols and thio ethers	22	21.47	11.70	31.92	19.29	14.76	8.55
overall	730	29.33	12.88	27.81	12.30	14.46	6.67

$$\varepsilon(\mathrm{K}) = \left|T_{\mathrm{m,exp}} - T_{\mathrm{m,pred}}\right| \tag{3}$$

$$\bar{\varepsilon}(\mathrm{K}) = \frac{1}{N}\sum_{n}\left|T_{\mathrm{m,exp}} - T_{\mathrm{m,pred}}\right| \tag{4}$$

$$\delta(\%) = \left|\frac{T_{\mathrm{m,exp}} - T_{\mathrm{m,pred}}}{T_{\mathrm{m,exp}}}\right| \times 100 \tag{5}$$

$$\bar{\delta}(\%) = \frac{1}{N}\sum_{n}\left|\frac{T_{\mathrm{m,exp}} - T_{\mathrm{m,pred}}}{T_{\mathrm{m,exp}}}\right| \times 100 \tag{6}$$

Figure 1 reports results for the 730 compounds considered in this study and compares our prediction for melting points with experimental data.

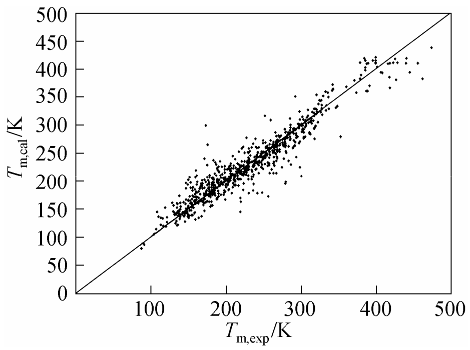


Figure 1 Comparison of melting point temperature calculated from this study with experimental values

One complementary set of position group contributions was developed for the predictive estimation of melting points of organic compounds. The second order groups defined by Benson, which have long been recognized as more accurate than the first order groups for thermochemical predictions, were used for this purpose. Position factor could distinguish between most isomers including *cis*- or *trans*-structure of organic compounds for their properties. Contributions for compounds containing carbon, hydrogen, oxygen, nitrogen, sulphur, chlorine and bromine were reported, and the results were evaluated with respect to other predictive methods. The presents melting points show an average deviation about 14.46 K, much less than those from the well known method of Joback-Reid and Constantinou-Gani well as all other estimation methods.

5 CONCLUSIONS

A new position group contribution method recently proposed by our laboratory for the estimation of the critical temperature, critical pressure, critical volume and normal boiling point is extended to the prediction of melting point of organics. Contributions for compounds containing carbon, hydrogen, oxygen, nitrogen, chlorine and sulphur are reported, and the position distribution function is developed, which could distinguish between the thermodynamic properties of most isomers of organic compounds including *cis*- and *trans*- or *Z*- and *E*- structures. The results indicate that our model provides very satisfactory results. The average difference for melting point predictions of 730 organic compounds is 14.46 K and 6.67 %. The higher prediction accuracy of the proposed method shown in our previous works and this work suggests that it is possible to use a same framework to predict the critical properties of organic compounds containing various functional groups.

REFERENCES

1 Jain, A., Yalkowsky, S.H., "Estimation of melting points of organic compounds-II", *J. Pharmaceutical Sci.*, **95**, 2562−2618 (2006).

2 Li, H.D., Higashi, H., Tamura, K., "Estimation of boiling and melting points of light, heavy and complex hydrocarbons by means of a modified group vector space method", *Fluid Phase Equilibria*, **239**, 213−222 (2006).

3 Jain, A., Yalkowsky, S.H., "Comparison of two methods for estimation of melting points of organic compounds", *Ind. Eng. Chem. Res.*, **46** (8), 2589−2592 (2007).

4 Sanghvi, R., Yalkowsky, S.H., "Estimation of normal boiling point of organic compounds", *Ind. Eng. Chem. Res.*, **45**, 2856−2861 (2006).

5 Joback, K.G., Reid, R., "Estimation of pure-component properties from group-contributions", *Chem. Eng. Commun.*, **57**, 233−243 (1987).

6 Constantinou, L., Gani, R., "A new group contribution method for estimation of properties of pure compounds", *AIChE J.*, **40**, 1697−1710 (1994).

7 Marrero, J., Gani, R., "Group-contribution based estimation of pure component properties", *Fluid Phase Equil.*, **183/184**, 183−208 (2001).

8 Guha, R., Jurs, P.C., "Determining the validity of a QSAR model", *J. Chem. Inf. Model*, **45**, 65−73 (2005).

9 Trohalaki, S., Patchter, R., Drake, G.W., Hawkins, T., "Quantitative structure property relationships for melting point and densities of ionic liquids", *Energy & Fuels*, **19**, 279−284 (2005).

10 Katritzky, A.R., Maran, U., Karelson, M., Lobanov, V.S., "Prediction of melting points for the substituted benzenes: A QSPR approach", *J. Phys. Chem. Ref. Data*, **37**, 913−919 (1997).

11 Charton, M., Charton, B., "Quantitative description of structural effects on melting points of substituted alkanes", *J. Phys. Org. Chem.*, **7**, 196−206 (1994).

12 Katritzky, A.R., Gordeeva, E.V., "Traditional topological indices *vs* electronic, geometrical, and combined molecular descriptors in QSAR/QSPR research", *J. Chem. Inf. Comput. Sci.*, **33**, 835−857 (1993).

13 Katritzky, A.R., Lomaka, A., Petrukhin, R., Jain, R., "QSPR correlation of the melting point for pyridinium bromides, potential ionic liquids", *J. Chem. Inf. Comput. Sci.*, **42**, 71−75 (2002).

14 Katritzky, A.R., Jain, R., Lomaka, A., Petrukhin, R., Maran, U., Karelson, M., "Perspective on the relationship between melting points and chemical structure", *Crystal Growth and Design*, **1**, 261−265 (2001).

15 Chickos, J.S., Nichols, G., "Simple relationships for the estimation of melting temperatures of homologous series", *J. Chem. Eng. Data.*, **46**, 562−573 (2001).

16 Karthikeyyan, M., Glen, R.C., Bender, A., Bender, A., "General melting point prediction based on a diverse compound data set and artificial neutral networks", *J. Chem. Inf. Model*, **45**, 581−590 (2005).

17 Tsakanikas, P.D., Yalkowsky, S.H., "Estimation of melting point of flexible molecules: Aliphatic hydrocarbons", *Toxic. Environ. Chem.*, **17**, 19−33 (1988).

18 Simamora, P., Yalkowsky, S.H., "Group contribution methods for predicting the melting point and boiling point of aromatic compounds", *Ind. Eng. Chem. Res.*, **33**, 1405−1409 (1994).

19 Krzyzaniak, J.F., Myrdal, P., Simamora, P., Yalkowsky, S.H., "Boiling point and melting point prediction for alphatic nonhydrogen-bonding compounds", *Ind. Eng. Chem. Res.*, **34**, 2530−2535 (1995).

20 Zhao, L., Yalkowsky, S.H., "A combined group contribution and molecular geometry approach for predicting melting point of alphatic compounds", *Ind. Eng. Chem. Res.*, **38**, 3581−3584 (1999).

21 Yalkowsky, S.H., Dannenfelser, R.M., Myrdal, P.B., Simamora, P., Mishra, D.S., "Unified physical property estimation relationships (UPPER)", *Chemosphere*, **28**, 1657−1673 (1994).

22 Wang, Q., Ma, P.S., Jia, Q., "Position group contribution method for the prediction of critical temperature of organic compounds", *J. Chem. Eng. Data*, **53**, 1103−1109 (2008).

23 Wang, Q., Jia, Q., Ma, P.S., "Position group contribution method for the prediction of critical pressure of organic compounds", *J. Chem. Eng. Data*, **53**, 1877−1885 (2008).

24 Jia, Q., Wang, Q., Ma, P.S., "Position group contribution method for the prediction of critical volume of organic compounds", *J. Chem. Eng. Data*, **53**, 2606−2612 (2008).

25 Wang, Q., Ma, P.S., "Position group contribution method for the prediction of normal boiling point of organic compounds", *Chin. J. Chem. Eng.*, **17** (2), 254−258 (2009).

26 Frenkel, M., Gadalla, N.M., Hall, K.R., Hong, X., Marsh, K.N., Wilhoit, R.C., TRC Thermodynamic Tables: Hydrocarbon; Non-Hydrocarbon, Thermodynamic Research Center, The Texas A&M University System (1997).

27 Ma, P.S., Handbook of Property Data of Organic Compound, Chemical Industry Press, Beijing (2006).

28 Poling, E.B., Prausnitz, J.M., O'Connell, P.J., The Properties of Gases and Liquids, 5th edition, McGraw-Hill, Inc., New York (2001).