Estimation of the Normal Boiling Point of Organic Compounds

Ritesh Sanghvi* and Samuel H. Yalkowsky

College of Pharmacy, The University of Arizona, Tucson, Arizona 85721

A semiempirical model has been developed for the estimation of the normal boiling points of organic compounds. The normal boiling point is calculated as the ratio of the enthalpy of boiling to the entropy of boiling. Both these values are estimated independently using a combination of additive group contribution and nonadditive molecular descriptors. A group contribution model is proposed for the estimation of the enthalpy of boiling, based on the experimental data for 1322 structurally diverse organic compounds. The average absolute error associated with this estimation is $0.83 \, \text{kJ/mol}$. A semiempirical model has been developed for the estimation of the entropy of boiling using experimental entropy data for the entire set. This model is a modification of Trouton's rule and takes into account the effect of hydrogen bonding and molecular flexibility. The average absolute error associated with this estimation is $4.7 \, \text{J/(K mol)}$. The average absolute error associated with the estimation of the normal boiling point, calculated as a ratio of the enthalpy of boiling to the entropy of boiling, for the entire data set is $9.3 \, \text{K}$ (2.3%). A good correlation is observed between the estimated and the experimental boiling points with an r^2 of 0.98. This model is easy to use, accurate, and applicable to a wide variety of organic compounds.

Introduction

Knowledge of physicochemical properties of organic compounds is essential for predicting their behavior. The normal boiling point (T_b), the entropy of boiling (ΔS_b), and the enthalpy of boiling (ΔH_b) are important examples of such physical properties. Besides being indicators of the physical state of the compound, T_b and ΔH_b provide information about volatility, i.e., its vapor pressure, and, consequently, its air—water and air—octanol partition coefficients. Since these properties are determinants of the loss due to evaporation and inhalation exposure, they are important in both pharmaceutical and environmental fields. Boiling point has also been used to calculate or estimate other physical properties, such as critical temperature, flash point, molar volume, and chromatographic retention index.

The normal boiling point and the enthalpy of boiling can be either measured experimentally or estimated using mathematical models. The experimental measurements are generally accurate for pure organic compounds. A large amount of experimental data are readily available in the literature. However, lots of inconsistencies have been noticed⁶ in the literature values. The presence of impurity or thermal instability of some compounds may thwart the experimental measurements. In such conditions, mathematical models can be used to provide a reasonable estimate of these properties. Such models provide an important tool to save time and material costs. They also provide a means of predicting the properties of yet-to-be-synthesized compounds, which may be very important in screening processes at a very early stage. These reasons support the need for the estimation of these properties.

Various methods for the estimation of T_b , ΔH_b , and ΔS_b have been proposed. Dearden⁷ reviewed some of these methods. Among the methods based on group contribution approach are those by Joback and Reid,⁸ Constantinou and Gani,⁹ Stein and Brown,¹⁰ and Nannoolal et al.¹¹ Most of these models use a multitier group contribution approach. Wessel and Jurs¹² have reported a model for predicting the T_b of hydrocarbons based

on a neural network. Basak et al.¹³ proposed a model using topological, topochemical, and geometric parameters. Group contribution methods have also been proposed to estimate $\Delta H_{\rm b}$.^{10,14} The entropy of boiling is mostly assumed to be constant and approximated by Trouton's rule. Recently, Kolska et al. have proposed a new group contribution model to estimate $\Delta S_{\rm b}$.¹⁴

In this work, we have developed a semiempirical model for the estimation of T_b . This model is based on a single-tier group contribution approach for the estimation of ΔH_b and a modified Trouton's rule for the estimation of ΔS_b . The normal boiling point is calculated as a ratio of the estimated ΔH_b to ΔS_b .

Theoretical Background

The normal boiling point ($T_{\rm b}$), i.e., the boiling point at 1 atm of a compound, is defined as the temperature at which the vapor pressure of its liquid phase equals 1 atm. It is an outcome of various molecular properties that affect the tendency of a molecule to escape from the liquid surface. The enthalpy of boiling and the entropy of boiling are the change in the respective properties associated with the vaporization of the liquid at its $T_{\rm b}$. At the boiling point, the free energy of transition is equal to zero and the three properties are related as

$$T_{\rm b} = \Delta H_{\rm b} / \Delta S_{\rm b} \tag{1}$$

The entropy of boiling is a measure of the increase in randomness of a system upon boiling. According to Boltzmann's statistical definition, the entropy of boiling is related to the relative probabilities of the existence of the gas and the liquid states as

$$\Delta S_{\rm b} = R \ln(\Omega^{\rm G}/\Omega^{\rm L}) \tag{2}$$

where Ω^L is the number of ways 1 mol of a material can be found that would constitute the liquid phase and Ω^G is the number of ways those molecules can exist as a gas. It is expressed as a sum of translational, rotational, and conformational components, i.e.,

^{*} Corresponding author. E-mail: rsanghvi@pharmacy.arizona.edu. Tel.: (520) 626-4308. Fax: (520) 626-4063.

$$\Delta S_{b} = \Delta S_{b}^{\text{Trans}} + \Delta S_{b}^{\text{Rot}} + \Delta S_{b}^{\text{Conf}}$$
 (3)

The contribution of the translational component predominates in eq 3. Boiling is associated with a large increase in the free volume. As a consequence, the molar volume of the compound undergoes a large increase too, and the molecules gain a large amount of translational freedom upon boiling. For most liquids, the change in the free volume upon boiling is approximately equal to the volume of its gas. Since the volume of a gas is almost constant under normal temperature—pressure conditions, the change in the translational freedom of the molecules is nearly constant. This explains Trouton's rule of constant entropy of boiling of 88 J/(deg mol). The contributions of rotational and conformational freedom to the entropy of boiling are small since compounds have a large degree of rotational and conformational freedom in both the liquid and the gas phase. However, the effect of these contributions cannot be completely neglected. Trouton's rule works exceedingly well for the rigid noninteracting molecules but should be modified to account for the molecular properties such as flexibility and hydrogen-bonding capability that may restrict rotational and conformational freedom in the liquid phase.

The flexibility of a molecule arises from the ability of its atoms to torsionally rotate about single bonds. A flexible molecule has a greater degree of order in the liquid form because it tends to be partially aligned with its neighbors. Upon boiling, the molecules tend to randomize and attain a higher conformational freedom resulting in an increase in entropy. Hydrogen bonding restricts the rotational freedom of the molecules in the liquid phase, whereas it has no effect over the large intermolecular distance in the gas phase. As a consequence, hydrogenbonding compounds typically have a higher entropy of boiling. Thus, the entropy of boiling may be estimated more accurately using the following equation:

$$\Delta S_{\rm b}$$
 = constant + effect of flexibility + effect of hydrogen bonding (4)

The molar enthalpy of vaporization at the boiling point depends on the amount of energy required to vaporize 1 mol of the liquid at its $T_{\rm b}$. The nature of the functional groups constituting the molecule determines the strength of the intermolecular interactions in the liquid state. The enthalpy of boiling is a function of the strength of these interactions and is related to the number and the type of groups comprising the molecule. Thus, it is reasonable to estimate $\Delta H_{\rm b}$ using a group contribution approach,

$$\Delta H_{\rm b} = \sum n_i b_i \tag{5}$$

where n_i is the number of times a group appears in the compound and b_i is its contribution to the enthalpy of boiling.

The experimental value of the enthalpy of vaporization at 298 K ($\Delta H_{\rm v}$) is more commonly available. Our data indicates that, although $\Delta H_{\rm v}$ values may be more accurate as the measurements are performed at lower temperature, these values are less constant than $\Delta H_{\rm b}$ values. Thus, in this study, a group contribution scheme is developed for the estimation of $\Delta H_{\rm b}$ instead of $\Delta H_{\rm v}$. The two values are related by the following equation,

$$\Delta H_{\rm b} = \Delta H_{\rm v} + \int_{298}^{T_{\rm b}} \Delta C_{p,\rm b} \,\mathrm{d}T \tag{6}$$

where the value of $\Delta C_{p,b}$ is the difference between the enthalpy—temperature slopes of the gas phase and the liquid phase. Since

Table 1. Summary of the Average Absolute Errors for the Estimation of ΔS_h

| | | | AAE^a | |
|------------------------|------|------------------------|---------|----------------|
| compounds | n | data range [J/(K mol)] | eq XI | Trouton's rule |
| rigid-non-H-bonding | 287 | 69.7-131.7 | 4.5 | 4.3 |
| flexible-non-H-bonding | 863 | 47.4 - 137.0 | 4.3 | 4.4 |
| rigid-H-bonding | 40 | 78.6 - 127.5 | 7.3 | 12.5 |
| flexible-H-bonding | 132 | 62.4 - 133.4 | 7.2 | 13.8 |
| total data set | 1322 | 47.4-137.0 | 4.7 | 5.6 |

^a AAE = \sum (| experimental value - predicted value |)/n.

liquid-phase enthalpy is more sensitive to temperature, $\Delta C_{p,b}$ values are invariably negative. Because the effect of temperature on $\Delta C_{p,b}$ has been shown to be small, ¹⁵ $\Delta C_{p,b}$ is assumed to be constant and equal to an average value calculated over (298 – T_b). Thus, eq 6 may be simplified as

$$\Delta H_{\rm b} = \Delta H_{\rm v} + [\Delta C_{\rm n,b} (T_{\rm b} - 298)]$$
 (7)

The heat capacity change upon boiling has units of entropy and depends on the same molecular parameters, flexibility, and hydrogen-bonding ability. It can be estimated using the following equation, ¹⁶

$$\Delta C_{n,b} = -56 - 4(\tau) - 40(\text{HBP})$$
 (8)

where τ is the flexibility parameter of the compound and HBP is the hydrogen-bond parameter.

Experimental Section

Estimation of the Entropy of Boiling. The equation for estimation of entropy of boiling is obtained by the multiple linear regression (SPSS version 10) of the experimental entropy data using the flexibility parameter (τ) and the hydrogen-bond parameter (HBP) as independent variables. These parameters are defined below. The equation includes an intercept that represents the ΔS_b value for rigid, non-hydrogen-bonding compounds.

Flexibility Parameter (τ). The flexibility parameter is calculated using the equation proposed by Dannenfelser and Yalkowsky, ¹⁷

$$\tau = (SP3 + 0.5SP2 + 0.5ring - 1)$$
 (9)

where SP3 and SP2 are the total number of nonring, nonterminal sp³ and sp² atoms. Ring indicates the number of independent single, fused, or conjugated ring systems in the molecule. Since terminal atoms and sp hybrid atoms do not contribute to flexibility, they are not counted. The rotation of hydrogen atoms is also ignored. Aliphatic cyclic groups are counted as single-ring systems. Compounds with a negative value for τ are assigned a value of zero.

Hydrogen-Bond Parameter (HBP). The effect of hydrogen bonding was considered as a function of the number of different hydrogen-bonding groups $[-OH, -COOH, and -NH (including -NH and -NH_2)]$ present on the molecule. The HBP is calculated using the equation proposed by Sanghvi and Yalkowsky: 16

$$HBP = \sqrt{[(\#OH) + (\#COOH) + 0.0625(\#NH)]}$$
 (10)

where # is the number of respective groups (aliphatic or aromatic).

Estimation of the Enthalpy of Boiling. The compounds were fragmented using the group breakdown scheme (Table 2)

Table 2. Group Values (b_i) for the Enthalpy of Boiling $(kJ/mol)^a$

| groups | $b_i(X)$ | $b_i(Y)$ | $b_i(YY)$ | groups | $b_i(X)$ | $b_i(Y)$ | $b_i(YY)$ |
|---|------------------|----------|-------------------|---|-----------------------|----------|-----------|
| CIT | 11.005 | 10.701 | Carl | | 2.000 | 1.650 | 1 -1- |
| -CH ₃ | 11.006 | 10.501 | 0.000 | -CH ₂ - | 2.088 | 1.658 | 1.646 |
| -CH< | -7.169 | -8.028 | -8.860 | >C< | -16.722 | -17.780 | -44.34 |
| =CH ₂ | 10.429 | 4.770 | 0.420 | =CH- | 2.41 | 2.255 | |
| =C< | -5.839 | -4.770 | -9.438 | =CH | 9.124 | | |
| =C- | 3.152 | 4.130 | | =C= | 1.852 | | |
| -C-ar | -3.139 | | | >C <bp< td=""><td>-3.844</td><td></td><td></td></bp<> | -3.844 | | |
| CHar | 5.267 | | | >C <br< td=""><td>-0.029</td><td></td><td></td></br<> | -0.029 | | |
| >CH—fus | -1.084 | | | >C <fus< td=""><td>-2.033</td><td></td><td></td></fus<> | -2.033 | | |
| >C=fus | 7.28 | | | >CH ₂ ring | 4.606 | | |
| >CH—ring | -4.064 | | | >CH=ring | 4.396 | | |
| >C <ring< td=""><td>-14.345</td><td></td><td></td><td>=C<ring< td=""><td>-4.072</td><td></td><td></td></ring<></td></ring<> | -14.345 | | | =C <ring< td=""><td>-4.072</td><td></td><td></td></ring<> | -4.072 | | |
| ring-3 | 8.832 | | | ring-4 | 6.845 | | |
| ring-5 | 5.152 | | | ring->5 | 2.742 | | |
| | | | Oxy | | | | |
| 1° OH | 22.487 | 21.176 | | 2° OH | 20.921 | | |
| 3° OH | 20.124 | | | -0- | 2.061 | 2.333 | 0.304 |
| −O−ar | 5.785 | | | —O—ring | 5.855 | | |
| | | | Oxygen ar | | | | |
| -CHO | 16.365 | 15.263 | | -co- | 6.826 | 5.589 | |
| -COOH | 30.100 | 29.191 | | -COO | 6.936 | 5.153 | |
| -HCOO | 16.039 | | | -cooco | 13.415 | | |
| -OCHO | 2.942 | | | HCOOH | 38.117 | | |
| -CO-ring | 10.047 | | | -COO-ring | 20.21 | | |
| -COOCO-ring | 27.451 | | | -COCl ₂ | 24.464 | | |
| | | | Nitro | ogen | | | |
| 1° NH₂ | 15.222 | 17.329 | 1,111 | 2° NH ₂ | 14.051 | | |
| 8° NH₂ | 12.508 | | | -NH- | 4.347 | 6.943 | |
| >N- | -7.454 | -4.190 | -7.911 | =N- | | 9.323 | |
| -N=N- | 3.07 | | | $-NHNH_2$ | 22.332 | | |
| >NNH ₂ | 8.48 | | | -NHNH- | 10.822 | | |
| >NNH— | -6.417 | | | -NH-ar | 15.505 | | |
| -N-ar | 6.925 | | | -NN-ar | 20.780 | | |
| -NH-ring | 9.713 | | | >N—ring | -3.801 | | |
| >PN ₃ | -28.791 | | | -NCBr | 21.093 | | |
| | | | Nitrogen a | | | | |
| -CN | 19.391 | 17.084 | ividogen ai | CNCN | 21.907 | | |
| HCN | 26.013 | 17.004 | | -NCN ar | -4.358 | | |
| | | | N: | | | | |
| $-NO_2$ | 19.835 | 17.529 | Nirogen an | ONO ₂ | 18.262 | | |
| −NO ₂ −NO−ar | 16.343 | 17.329 | | -ONO ₂ | 16.202 | | |
| 110 ai | 10.545 | | N: G 1 | 1.0 | | | |
| -CONH | 29.915 | | Nitrogen, Carbo | on, and Oxygen —CON | 4.618 | | |
| | | | | | | | |
| -NCHO | 15.125 | | | -NCO | 16.992 | | |
| -CONH-ring | 28.038 | | | —CON <ring< td=""><td>11.349</td><td></td><td></td></ring<> | 11.349 | | |
| -OC=N-ring | 8.447 | | | | | | |
| _011 | 15 200 | 15 205 | Sul | | 5 170 | 2 796 | |
| −SH −SS− | 15.308 11.324 | 15.385 | | —S— —S—ar | 5.478 <i>10.41</i> | 3.786 | |
| −SS− −S−ring | 10.729 | | | o al | 10.41 | | |
| ~ 11115 | 10.727 | | O 16 | d Control | | | |
| -scs- | 27.77 | | Sulfur and | и Сагроп | | | |
| | | | C.,1£ | 1 Ovugan | | | |
| -SO ₂ - | 12.539 | 18.856 | Sulfur and | SO— | 19.192 | | |
| $-SO_3$ | 11.925 | 10.050 | | 50 | 17.174 | | |
| 503 | 11.743 | | a .a | 1.0 | | | |
| OCC | 10.204 | | Sulfur, Carbon | , and Oxygen | | | |
| -ocs- | 19.384 | | | | | | |
| | | | Sulfur, Carbon | , and Nitrogen | | | |
| -NCS- | 22.754 | | | | | | |
| | | | Phosphorus a | and Oxygen | | | |
| PO_4 | 6.21 | | F | - 70- | | | |
| | | | Phosphorus, Oxy | gen and Sulfur | | | |
| PO_2S | 0.668 | | i nospilorus, Oxy | PO ₂ S ₂ | -10.016 | | |
| PO ₃ S | -8.783 | | | 1 0202 | 10.010 | | |
| . 0,0 | 5.765 | | | | | | |
| Г | 10.61 | 0.265 | Halo | | 10.714 | 10 100 | |
| -F | 10.61 | 8.265 | | -Cl | 13.714 | 12.139 | |
| –Br | 15.435 | 14.115 | | —I | 18.069 | 16.236 | |
| | | | Proximit | y Effect | | | |
| | | | | | | | |
| HB-5 | 1.128 | | | IHB-6 | 0.396 | | |
| IHB-5 IHB-7 | 1.128 1.966 | | | IHB-6 G-2 | 0.396 -2.242 | | |

^a (X): The fragment is attached to singly bonded groups; (Y): the fragment is attached to one doubly or triply bonded group; and (YY): the fragment is attached to two or more doubly or triply bonded groups. ar: aromatic; bp: atoms that are part of two aromatic rings; br: atoms that are part of more than one aromatic rings; ring: atoms that are part of an aliphatic ring system; and ring- 3,4,5,>5: three, four, five, or more than five-membered aliphatic ring systems, respectively. 1°: Primary aliphatic; 2°: secondary aliphatic; 3°: tertiary aliphatic; IHB-5,6,7: five-, six-, or seven-membered ring systems, respectively, formed due to the formation of intramolecular hydrogen bonding. G-2,3,4: Carbons bound to two, three, or four halogens, respectively. Italicized values are based on <5 occurrences.

proposed by Jain. 19 This scheme takes into account the effect of individual groups and their proximity on the heat of transition. For describing the structure of compounds used in this study, 132 group values were required. This includes nine new groups that were defined to describe the structure of some compounds that are not covered by the scheme of Jain et al. The coefficients of the groups for the estimation of enthalpy of boiling were obtained by multiple linear regression (SPSS version 10) of the experimental $\Delta H_{\rm b}$ values.

Data. The experimental $\Delta H_{\rm v}$ values for 1322 compounds were obtained from Chickos and Acree. 18 These compounds were selected as the enthalpy measurements were performed at 298 K (± 5 K). The experimental ΔH_b^* values for these compounds were calculated using eq 7. The experimental $T_{\rm b}$ values for the compounds were obtained from the literature. The experimental ΔS_b values were calculated as a ratio of ΔH_b * and $T_{\rm b}$. These values were used to develop the model for the estimation of $\Delta S_{\rm h}$. The apparent experimental $\Delta H_{\rm h}$ values were obtained by multiplying the estimated ΔS_b and experimental $T_{\rm b}$ values. These values were used to develop the model for the estimation of $\Delta H_{\rm b}$.

Estimation of the Boiling Point. The boiling points of the compounds were calculated by dividing the estimated $\Delta H_{\rm b}$ values by the respective estimated ΔS_b values.

Results and Discussion

Entropy of Boiling. On the basis of the statistical analysis, the following equation was obtained for the estimation of ΔS_b :

$$\Delta S_{\rm b} = 87 + 0.35(\tau) + 15(\text{HBP})$$
 (11)

The average absolute error of the estimation is 4.7 J/(K mol), corresponding to 5.2%. Table 1 summarizes the results obtained using eq 11. The compounds have been divided into four subsets: rigid-non-hydrogen-bonding; flexible-non-hydrogenbonding; rigid-hydrogen-bonding; and flexible-hydrogenbonding. For comparison, the results obtained using Trouton's rule are included in the table.

The table demonstrates the general applicability of Trouton's rule, particularly of rigid-non-hydrogen-bonding compounds. However, the errors associated with flexible compounds are slightly lower, and those for hydrogen-bonding compounds are substantially lower with eq 11. The intercept of 87 J/(K mol) is very close to Trouton's constant entropy of boiling [88 J/(K mol)], which accounts for the gain in the translational freedom upon boiling.

The effect of increase in conformational and rotational freedom cannot be totally neglected. These motions are somewhat restricted in the liquid phase, particularly for flexible molecules. Intermolecular hydrogen bonding also restricts molecular rotations in the liquid state. These restrictions are relieved upon boiling, since the likelihood of hydrogen bonding is negligible in the gas phase. Intuitively, the effect of different hydrogen-bonding groups on the entropy of boiling depends on their strength. This is reflected in eq 10 which is used to calculate HBP, where the coefficients of -COOH and -OH groups are larger than that of an -NH group, since the former are stronger hydrogen-bonding groups than the latter. Since the -SH group is a very weak hydrogen-bonding group, its effect on the entropy of boiling is insignificant. The effect of hydrogen bonding increases with the number of such groups, but this increase is not linear. The square root term in eq 10 accounts for the fact that the formation of the first hydrogen bond imposes geometric restrictions on the molecule

Table 3. Summary of the Average Absolute Errors for the Estimation of $\Delta H_{\rm h}$

| compounds | n | data range (kJ/mol) | AAE^a |
|------------------------|------|---------------------|---------|
| rigid-non-H-bonding | 287 | 9.7-48.5 | 1.0 |
| flexible-non-H-bonding | 863 | 21.6-60.3 | 0.8 |
| rigid-H-bonding | 40 | 24.2-50.5 | 0.8 |
| flexible-H-bonding | 132 | 26.6-64.0 | 0.7 |
| total data set | 1322 | 9.7 - 64.0 | 0.8 |

^a AAE = \sum (|(experimental value - predicted value)|)/n.

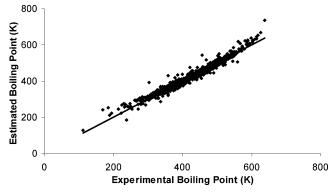


Figure 1. Estimated versus experimental boiling point (K).

to reduce the feasibility of the second hydrogen-bond formation. Molecules containing only hydrogen-bond acceptor groups and no donor groups cannot undergo hydrogen-bond formation. Thus, the presence of groups such as carbonyl, nitro, or halo does not appear to affect the entropy term. A mathematically significant effect of such groups in the estimation was not observed.

Enthalpy of Boiling. The group values for the estimation of the enthalpy of boiling are given in Table 2. The summation of these values is used to calculate the total enthalpy for each compound using eq 5.

The group breakdown scheme of Jain¹⁹ was proposed originally for estimating the enthalpy of melting. An addition of nine new groups was made to describe the structures of some compounds that were not covered by the original scheme. The average absolute error associated with the estimation is 0.83 kJ/mol, which correspond to 2.3%. These results are summarized in Table 3. The considerable accuracy obtained with the estimation of enthalpy of boiling using this scheme indicates that the influence of a group on the process of melting and boiling is qualitatively similar. However, the coefficients for the estimation of enthalpy of boiling are generally larger than those for the estimation of enthalpy of melting.

Normal Boiling Point. The estimated boiling points for the compounds used in this study were calculated by incorporating eqs 5 and 11 into eq 1:

$$T_{\rm b} = [\sum n_i b_i]/[87 + 0.35(\tau) + 15(\text{HBP})]$$
 (12)

The average absolute error of estimation is 9.3 K, corresponding to 2.3%. A correlation coefficient (r^2) of 0.98 was obtained with a standard deviation of 9.8 K. A plot of the predicted versus the experimental boiling points is given in Figure 1.

It is noticed that slightly higher negative errors were associated with high-boiling compounds. The systematic overprediction for such compounds may be due to the underestimation of entropy of boiling. Also for extremely high-boiling compounds, the likelihood for thermal decomposition exists at temperatures near their boiling points.

The results obtained for the estimation of the boiling points are presented in Table 4. The proposed model performs well,

| · · | | | |
|------------------------|------|----------------|---------|
| compounds | n | data range (K) | AAE^a |
| rigid-non-H-bonding | 287 | 111.6-557.5 | 11.4 |
| flexible-non-H-bonding | 863 | 247.0-639.0 | 8.9 |
| rigid-H-bonding | 40 | 266.7-518.0 | 8.1 |
| flexible-H-bonding | 132 | 292.0-581.5 | 7.1 |
| total data set | 1322 | 111.6-639.0 | 9.3 |
| | | | |

^a AAE = \sum (| (experimental value – predicted value) |)/n.

Table 5. Cross-Validation Results of the Proposed Model

| | training set | | test set | |
|------------------|--------------|---------|----------|---------|
| cross validation | n | AAE^a | n | AAE^a |
| round 1 | 1190 | 9.4 K | 129 | 11.1 K |
| round 2 | 1190 | 9.3 K | 130 | 9.9 K |
| round 3 | 1190 | 9.6 K | 128 | 10.1 K |

^a AAE = \sum (| (experimental value – predicted value) |)/n.

especially for the hydrogen-bonding compounds. It should be mentioned that, although the hydrogen-bonding term appears directly in the entropy term, its effect is much higher in the enthalpy-of-boiling calculation. This is in accordance to the fact that hydrogen-bonding compounds have a tendency to boil at higher temperatures than their non-hydrogen-bonding counterparts.

Cross-Validation

The validity of the proposed model was tested by cross-validation. For this purpose, the entire data set was randomized and the first 10% of compounds were used as the test set. The remaining 90% of the compounds (n=1190) were used as the training set to generate the coefficients for the estimation. The results for cross-validation are given in Table 5. The average absolute errors for the training and test sets are comparable for all three rounds. The paired T-test indicates that the experimental and the predicted boiling points for the test compounds were statistically not different ($\alpha=0.05$) for all three rounds.

Conclusion

The proposed semiempirical model enables a simple, accurate, and widely applicable method for estimation of the normal boiling point of organic compounds based on their chemical structure. The enthalpy of boiling can be estimated using the proposed group contribution values. The entropy of boiling can be estimated using a modified Trouton's rule. The normal boiling point is than calculated as the ratio of the two values.

Examples

(a) 1-penten-3-ol.

Experimental T_b : 387.5 K

SP3: 2, SP2: 1, ring: 0, #(OH): 1

$$\Delta S_b = 87 + 0.35(\tau) + 15(HBP)$$

$$= 87 + 0.35(2 + 0.5 + 0 - 1) + 15(1) = 102.53 \text{ J/(K mol)}$$

$$\Delta H_{b} = \sum n_{i}b_{i} = 1(\text{XCH3}) + 1(\text{XCH2}) + 1(\text{YCH}) + 1(\text{XCH=}) + 1(\text{CH2=}) + 1(2^{\circ} \text{ OH})$$

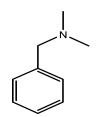
$$= (11.006) + (2.088) + (-8.028) + (2.41) + (10.429) + (20.921) = 38.83 \text{ kJ/mol}$$

Predicted
$$T_{\rm b} = \Delta H_{\rm b}/\Delta S_{\rm b} =$$

 $38.83 \times 1000/102.53 = 378.7 \text{ K}$

Absolute error =
$$8.8 \text{ K} (2.2\%)$$

(b) N,N-Dimethylbenzylamine.



Experimental T_b: 456.0 K

SP3: 2, SP2: 0, ring: 1

$$\Delta S_b = 87 + 0.35(\tau) + 15(HBP)$$

= 87 + 0.35 (2 + 0 + 0.5 - 1) = 87.53 J/(K mol)

$$\Delta H_{\rm b} = \sum n_i b_i = 2({\rm XCH3}) + 1({\rm YCH2}) + 1({\rm XN}) + 5({\rm CHar}) + 1({\rm Car})$$
$$= 2(11.006) + (1.658) + (-7.454) + 5(5.267) + (-3.139) = 39.41 \text{ kJ/mol}$$

Predicted
$$T_{\rm b} = \Delta H_{\rm b}/\Delta S_{\rm b} =$$
 (39.41 × 1000)/87.53 = 450.3 K

Absolute error: 5.7 K (1.3%)

Supporting Information Available: The experimental and the estimated values for the normal boiling point along with the absolute errors are available for all the organic compounds used in this study. This material is available free of charge via the Internet at http://pubs.acs.org.

Literature Cited

- (1) Yalkowsky, S. H.; Dannenfelser, R. M.; Myrdal, P.; Simamora, P. Unified Physical Property Estimation Relationship (UPPER). *Chemosphere* **1994**, 28, 1657–1673.
- (2) Fisher, C. H. Boiling Point Gives Critical Temperature. Chem. Eng. 1989, 96, 157–158.
- (3) Satyanarayana, K.; Kakati, M. C. Note: Correlation of Flash Points. *Fire Mater.* **1991**, *15*, 97–100.
- (4) Sladkov, B. Estimation of Molar Volume of Inorganic Liquids at Boiling Points. J. Appl. Chem. 1991, 64, 2273.
- (5) Herberger, K. Discrimination between Linear and Non-Linear Models Describing Retention Data of Alkylbenzene in Gas-Chromatography. *Chromatographia* **1990**, *29*, 275.
- (6) Rucker, G.; Rucker, C. On topological indices, boiling points, and cycloalkanes. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 788–802.
- (7) Dearden, J. C. Quantitative Structural-Property Relationship for Prediction of Boiling Point, Vapor Pressure, and Melting Point. *Environ. Toxicol. Chem.* **2003**, 22, 1696–1709.
- (8) Joback, K. G.; Reid, R. C. Estimation of pure-component properties from group contributions. *Chem. Eng. Commun.* **1987**, *57*, 233–243.
- (9) Constantinou, L.; Gani, R. New group contribution method for estimating properties of pure compounds. *AIChE J.* **1994**, *40*, 1697–1710.

- (10) Stein, S. E.; Brown, R. L. Estimation of Normal Boiling Point from Group Contribution. J. Chem. Inf. Comput. Sci. 1994, 34, 581-587.
- (11) Nannoolal, Y.; Rarey, J.; Ramjugernath, D.; Cordes, W. Estimation of Pure Component Properties. Part 1. Estimation of Normal Boiling Point of Non-Electrolyte Organic Compounds via Group Contribution and Group Interactions. Fluid Phase Equilib. 2004, 226, 45-63.
- (12) Wessel, M. D.; Jurs, P. C. Prediction of Normal Boiling Points of Hydrocarbons from Molecular Structure. J. Chem. Inf. Comput. Sci. 1995, *35*, 68−76.
- (13) Basak, S. C.; Gute. B. D.; Grunwald, G. D. A Comparative Study of Topological and Geometrical Parameters in Estimating Normal Boiling Point and Octanol/Water Partition Coefficient. J Chem. Inf. Comput. Sci. **1996**, 36, 1054-1060.
- (14) Kolska, Z.; Ruzicka, V.; Gani, R. Estimation of the Enthalpy of Vaporization and the Entropy of Vaporization for Pure Organic Compounds at 298.15 K and at Normal Boiling Temperature by a Group Contribution Method. Ind. Eng. Chem. Res. 2005, 44, 8436-8454.

- (15) Myrdal, P. B.; Yalkowsky, S. H. Estimating Pure Component Vapor Pressures of Complex Organic Molecules. Ind. Eng. Chem. Res. 1997, 36,
- (16) Sanghvi, R.; Yalkowsky S. H. Estimation of Heat Capacity of Boiling for Organic Compounds. Ind. Eng. Chem. Res. 2006, 45, 451-453.
- (17) Dannenfelser, R. M.; Yalkowsky, S. H. Predicting the Total Entropy of Melting: Application to Pharmaceutical and Environmentally Relevant Compounds. J. Pharm. Sci. 1999, 88, 722-724.
- (18) Chickos, J.; Acree, W. E. Enthalpies of Vaporization of Organic and Organometallic Compounds 1880-2002. J. Phys. Chem. Ref. Data **2003**. *32*. 519-878.
 - (19) Jain, A. Ph.D. Thesis, University of Arizona, Tucson, AZ, 2005.

Received for review November 17, 2005 Revised manuscript received February 15, 2006 Accepted February 16, 2006

IE051277G