Russian Journal of General Chemistry, Vol. 73, No. 3, 2003, pp. 371–374. Translated from Zhurnal Obshchei Khimii, Vol. 73, No. 3, 2003, pp. 400–403.
Original Russian Text Copyright © 2003 by Golovanov, Zhenodarova.

Structure–Property Correlation Relation: XII.[1] Simple Expression for a Property of Any Organic Compound Containing Unbranched Alkyl Substituents

I. B. Golovanov and S. M. Zhenodarova

Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences, Pushchino, Moscow oblast, Russia

Received January 29, 2001

Abstract—A simple expression that allows accurate calculation of physicochemical properties of organic compounds like RX (R is an unbranched alkyl substituent C_nH_{2n+1} and X is a functional group) was proposed. The potential of the proposed approach is demonstrated by the estimation of the boiling points and heat capacities at constant pressure of alkyl benzenes and carboxylic esters.

We earlier described [2] a general correlation equation that allows quite an accurate calculation of properties of unsaturated hydrocarbons and other organic compounds. Properties of linear molecules are described by an equation including five empirical parameters, whereas properties of branched and cyclic molecules are determined via properties of linear molecules of the same composition with corrections for branching or ring formation, whose value depends on three additional parameters. Thus, the first step in the estimation of properties of unsaturated hydrocarbons of any structure reduces to the estimation of properties of linear molecules of the same composition. In this connection it is desirable to find a simpler approach to estimating properties of linear molecules, using less parameters, since determination of parameters is frequently frustrated by the shortage of experimental evidence.

In estimating various properties of linear saturated hydrocarbons we noted that the properties smoothly vary with hydrocarbon chain length. Morever, the tendency to smoothly vary with hydrocarbon chain length within one class was revealed for many properties of aliphatic and aromatic compounds RX (R is an alkyl substituent C_nH_{2n+1} and X is a functional group). Therefore, not posing the problem of estimating a property from general considerations relying on fundamental concepts of quantum mechanics and quantum chemistry, one can approximate a property of a linear molecule by a simple function of *n*, including less parameters. In the present work we propose one of approaches to tackling this methodically very important problem.

The structure–property correlation equation proposed in [2] takes the following form for linear molecules.

$$P_n = n\alpha + (n - 1)\beta + (n - 2)\gamma + (n - 3)\delta_{14} + (n - 4)\delta_{15} + \ldots \qquad (1)$$

Here P_n is a property of a molecule, *n* is the number of C atoms in the chain, α is a one-particle contribution into the property, and β, γ, etc. are two-particle contributions into the property, relating to interactions like 1...2, 1...3, etc.

The two-particle contributions attenuate with distance, and, therefore, one should accurately estimate the most important contributions (α, β, γ, and δ_{14}), whereas the contributions of longer range interactions can be less accurately estimated under condition (2).

$$\delta_{1n} = \delta_{14}k^{(n-4)} \ (k < 1). \qquad (2)$$

Five parameters (α, β, γ, δ_{14}, and *k*) are determined from test examples. Provided α, β, and γ are estimated accurately and longer range interaction parameters are estimated by an equation like (2), namely, $\delta_{1n} = \gamma k^{(n-3)}$, the number of parameters will reduce to four, but the resulting estimates will be slightly less accurate.

Properties of RX molecules still smoother vary with chain length compared with saturated hydrocarbons [3–5]. Therefore, approach (1) can be changed, taking into account that a property of an RX molecule can be estimated from data for the “ancestor” of the series CH_3X by adding slowly changing corrections for additional CH_2 groups.

---

[1] For communication XI, see [1].

Table 1. Boiling points (°C) of normal saturated hydrocarbons

Alkane no.[a]	Experiment [3–5]	Present work				
		b	c	d	e	f
1	–161.8	–161.8	–155.9	–163.3	–161.8	–164.9
2	–88.7	–90.5	–95.4	–91.3	–92.0	–88.4
3	–42.1	–39.6	–44.9	–39.4	–40.3	–36.9
4	–0.6	2.0	–1.8	2.7	1.9	4.2
5	36.1	38.0	36.0	38.7	38.2	39.2
6	68.8	70.0	69.6	70.5	70.3	70.2
7	98.5	99.0	99.8	99.2	99.2	98.3
8	125.7	125.7	127.2	125.5	125.7	124.1
9	150.8	150.4	152.4	149.9	150.1	148.2
10	174.2	173.3	175.6	172.6	173.0	170.8
11	196.9	194.8	197.1	193.9	194.3	192.2
12	216.2	215.1	217.2	214.1	214.6	212.6
13	235.5	234.2	236.1	233.3	233.7	232.1
14	253.6	252.3	253.8	251.6	252.0	250.7
15	270.7	269.5	270.6	269.1	269.4	268.7
16	286.9	285.8	286.5	286.0	286.1	286.0
17	302.2	301.3	301.5	302.1	302.1	302.8
18	316.8	316.2	315.9	317.7	317.6	318.9
19	330.6	330.3	329.6	332.8	332.5	334.8
20	343.8	343.8	342.7	347.4	346.9	350.1
r		1.0000	0.9999	0.9999	0.9999	0.9999
s		1.3396	2.4586	2.2329	1.8874	3.5257

[a] Simpified formulas of hydrocarbons are given [6]: The numerals denote the number of carbon atoms in the linear chain.

[b] Calculated by Eq. (3) with use of Eq. (4) with α^* –161.7500±1.3396, k_1 6.4580±3.6296, k_2 66.6441±3.7762, and k_3 1.8967± ±0.7415. [c] Calculated by Eq. (5) with k_1 –335.5045±2.1320, k_2 301.3898±2.0367, and k_3 3.4909±0.1308. [d] Calculated by Eq. (6) with k_1 –125.1725±3.7124, k_2 235.0345±0.6653, and k_3 0.3854±0.0264. [e] Calculated by Eq. (7) with k_1 –341.2704±50.2007, k_2 –50.2737±17.1868, and k_3 452.6387±65.6078; k_4 was set at 0.3. [f] Calculated by Eq. (8) with k_1 51.7169±9.4295, k_2 –75.0050± 9.9682, and k_3 131.5452±1.8889.

$$P_n = \sum_1^n \alpha_i = \alpha^* + \sum_1^{n-1} \beta_m^*. \tag{3}$$

Here R_n is a property of an RX molecule (R = C_nH_{2n+1}), α^* is a property of the “ancestor” of the series, β_m^* is the correction for the *m*th CH_2 group. Therewith, β_m^* can be chosen as $\beta_1^* k^{m-1}$ or another function that decreases with increasing *m*.

$$\beta_m^* = k_1 + k_2 n^{1/2} + k_3 n^{-1/2} \quad (m = n - 1). \tag{4}$$

In this case, we can gat β^* to vary with *n* at a rate a close to desired, while the total number of parameters for molecules like RX will equal four.

Riding ourselves of the intention to impart to various contributions into a property *P* at least arbitrary physicochemical meaning, but simply considering *R* as a function of *n*, we can present this function by Eqs. (5)–(8), etc.

$$P_n = k_1 + k_2 \ln (n + k_3) \tag{5}$$

or

$$P_n = k_1 + k_2 n^{k_3} \tag{6}$$

or

$$P_n = k_1 + k_2 \ln n + k_3 n^{k_4} \tag{7}$$

or

$$P_n = k_1 + k_2 n^{-1/2} + k_3 n^{1/2}. \tag{8}$$

Values of parameters k_i can be determined from test examples. For *n* we can chose an “effective” *n* value equal to $(n + k_3)$; in the last of the above equations, $n^{1/2}$ can be replaced by n^k, etc.

Using Eqs. (3)–(8) we estimated some properties of normal saturated hydrocarbons: boiling point, heat capacity at constant pressure, entropy, and standard

Table 2. Properties of alkylbenzenes

Alkyl benzene no.[a]	bp, °C				Heat capacity, C_p, J mol^{-1} K^{-1}			
	experiment [3–5]	present work			experiment [3–5]	present work		
		b	c	d		b	c	d
1	80.1	80.1	80.3	79.7	81.7	81.7	91.0	90.0
2	110.6	109.9	108.5	109.2	103.7	104.5	96.0	94.5
3	136.2	136.6	136.2	137.0	128.5	127.7	118.9	118.5
4	159.2	161.3	161.7	162.2	152.4	151.0	145.4	145.7
5	183.3	184.3	185.0	185.2	175.2	174.3	172.3	173.0
6	205.5	205.9	206.6	206.5	198.1	197.6	198.8	199.6
7	226.1	226.2	226.7	226.4	221.0	220.8	224.5	225.2
8	246.1	245.4	246.6	245.1	243.9	244.0	249.3	250.0
9	264.4	263.6	263.4	262.8	266.8	267.1	273.3	273.7
10	282.0	280.8	280.3	279.7	289.6	290.2	296.4	296.6
11	297.9	297.1	296.3	295.8	312.5	313.2	318.8	318.7
12	313.2	312.5	311.7	311.3	335.4	336.2	340.4	340.2
13	327.6	327.1	326.4	326.2	358.3	359.0	361.3	360.9
14	341.3	340.9	340.6	340.6	381.2	381.8	381.6	381.1
15	353.9	353.9	354.2	354.5	404.1	404.6	401.4	400.7
16	365.9	366.2	367.4	368.0	427.0	427.3	420.5	419.7
17	377.9	377.9	380.1	381.1	449.9	449.9	439.2	438.3
r		1.0000	0.9999	0.9998		1.0000	0.9985	0.9983
s		0.8117	1.5463	1.9127		0.6709	6.8735	7.1694

[a] Simplified formulas of alkylbenzenes are given [6]: The numerals denote the number of carbon atoms in the normal alkyl substituent. [b] Calculated by Eq. (3) with use of Eq. (4) with α^* 353.2500±0.8117, k_1 32.2133±2.5404, k_2 2.9133±2.5071, and k_3 5.3313±0.5537 for boiling points and α^* 81.6717±0.6709, k_1 –25.7977±1.9508, k_2 –2.4000±1.9254, and k_3 –0.6492±0.4253 for C_p. [c] Calculated by Eq. (7) with k_1 –114.1104±51.8776, k_2 –115.2510±15.7252, and k_3 467.5626±64.6444 for boiling points and k_1 –738.6457±32.0175, k 269.4566±17.6277, k_3 829.6313±35.7407 for C_p; in both cses, k_4 0.3. [d] Calculated by Eq. (8) with k_1 188.6544±5.9912, k_2 54.5108±6.0231, and k_3 109.7076±1.2766 for boiling points and k_1 –291.2488±22.4564, k_2 217.1850±22.5761, and k_3 164.1753±4.7851 for C_p.

enthalpy and standard Gibbs energy of formation. As an example, Table 1 lists the boiling points.

As seen, the four-parameter equation (3) provides an excellent fit to experiment, and the three-parameter equations (7) and (8), a fairly good fit. In calculations by Eq. (7), the k_4 value was varied from 0.2 to 0.4, and, therefore, k_4 was set invariant at 0.3, which converted Eq. (7) into three-parameter. Similar results were obtained with other properties of saturated hydrocarbons; for this reason, in what follows we will operate by these equations only.

To validate the proposed approach, we estimated such properties as boiling point, heat capacity at constant pressure, entropy, and standard enthalpy and Gibbs energy of formation for organic compounds of different classes: 1-alkenes (R = CH_3 to C_8H_{17}, X = $CH_2{=}CH$), 1-alkynes (R = CH_3 to C_8H_{17}, X = $HC{\equiv}C$), alkylbenzenes (R = CH_3 to $C_{17}H_{35}$, X = C_6H_5), and carboxylic esters [R = CH_3 to $C_{10}H_{21}$, X = $CH_3C(O)O$ or $C(O)OCH_3$]. Qualitatively similar results were obtained in all the cases, and we exemplify them by the results for alkylbenzenes (Table 2) and carboxylic esters (Table 3), since, on the one hand, these compounds have much different structures of substituent X and, on the other, for them the greatest body of experimental data is available. For the "ancestors" of the alkylbenzene and carboxylic ester series we chose toluene and methyl acetate, respectively.

The results in Tables 2 and 3 suggest that the four-parameter equation (3) provides a very good fit of the calculated properties of organic compounds like RX (R is a normal alkyl substituent) to experiment, while

Table 3. Boiling points (°C) of carboxylic esters

Molecule[a]	Experiment [3–5, 7]	Present work		
		b	c	d
$CH_3C(O)OR$				
1	57.0	57.0	56.7	56.8
2	77.1	77.2	77.4	77.2
3	101.8	101.4	102.1	102.1
4	125.6	125.7	126.2	126.2
5	148.0	149.0	149.0	148.9
6	171.5	171.0	170.4	170.3
7	192.4	191.6	190.6	190.5
8	210.0	210.6	209.8	209.7
9	228.0	228.1	228.0	228.0
10	244.0	244.0	245.3	245.5
r		1.0000	0.9999	0.9999
s		0.5230	1.0535	1.1017
$RC(O)OCH_3$				
1	57.0	57.0	56.9	56.9
2	79.8	79.4	78.9	78.9
3	102.5	103.6	104.0	104.0
4	126.5	127.3	128.0	128.0
5	151.0	150.0	150.5	150.5
6	172.0	171.5	171.6	171.5
7	193.0	191.7	191.5	191.4
8	213.8	210.7	210.2	210.2
9	224.0	228.5	228.0	228.1
10	245.0	245.0	245.0	245.2
r		0.9995	0.9995	0.9995
s		1.9587	2.2969	2.3408

[a] Simplified denotation of alkyl radical in esters is given: The numerals denote the number of carbon atoms in normal alkyl radical R. [b] Calculated by Eq. (3) with use of Eq. (4) with α^* 57.0000 ± 0.5230, k_1 64.4138 ± 6.2241, k_2 –31.5794 ± 5.1967, and k_3 –12.6463 ± 1.6681 for $CH_3C(O)OR$ and α^* 57.0000 ± 1.9587, k_1 50.8011 ± 23.7113, k_2 –19.0618 ± 19.7972, and k_3 –9.2991 ± 6.3548 for $RC(O)OCH_3$. [c] Calculated by Eq. (7) with k_1 –198.1485 ± 10.1667, k_2 –146.2881 ± 4.7157, and k_3 527.9936 ± 10.7252 for $CH_3C(O)OR$ and k_1 –175.1523 ± 22.1667, k_2 –136.6674 ± 10.2817, and k_3 505.1950 ± 23.3843 for $RC(O)OCH_3$; in both cases, k_4 0.3. [d] Calculated by Eq. (8) with k_1 115.2528 ± 6.0724, k_2 96.7948 ± 5.2218, and k_3 117.8713 ± 1.5691 for $CH_3C(O)OR$ and k_1 128.6432 ± 12.9023, k_2 86.9014 ± 11.0950, and k_3 114.5320 ± 3.3340 for $RC(O)OCH_3$.

the three-parameter equations (7) and (8), quite satisfactory. Since no assumptions as to the molecular moiety denoted X we made, were believe that even the simplest variant of the proposed approach (8) provides fairly accurate estimates for various properties of any molecules comprising an alkyl radical. This variant is preferred from the practical viewpoint, because its parameters can be determined with a minimal set of experimental data.

REFERENCES

1. Golovanov, I.B. and Zhenodarova, S.M., *Zh. Obshch. Khim.*, 2003, vol. 73, no. 2, p. 260.
2. Golovanov, I.B. and Zhenodarova, S.M., *Zh. Obshch. Khim.*, 2003, vol. 73, no. 1, p. 90.
3. Stull, R. ., Westrum, E.F., Jr., and Sinke, G.C., *The Chemical Thermodynamics of Organic Compounds*, New York: Wiley, 1969.
4. Reid, R.C., Prausnitz, J.-M., and Sherwood, T.K., *The Properties of Gases and Liquids*, New York: McGraw–Hill, 1977, 3rd ed.
5. *CRC Handbook of Chemistry and Physics*, Lide, D.R., Ed., Boca Raton: CRC, 1993–1994, Sect. 3, 5.
6. Kier, L.B. and Hall, L.H., *Molecular Connectivity in Chemistry and Drug Research*, New York: Academic, 1976.
7. *Svoistva organicheskikh soedinenii: Spravochnik* (Properties of Organic Compounds: Handbook), Potekhin, A.A., Ed., Leningrad: Khimiya, 1984.