Russian Journal of General Chemistry, Vol. 74, No. 5, 2004, pp. 701–705. Translated from Zhurnal Obshchei Khimii, Vol. 74, No. 5, 2004, pp. 765–769.
Original Russian Text Copyright © 2004 by Golovanov, Zhenodarova.

Structure–Property Correlation: XVI.[1] Linear Molecules

I. B. Golovanov and S. M. Zhenodarova

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

Received May 7, 2002

Abstract—The structure–property correlation for linear molecules of the A_n type is represented by a simple function of n with a restricted number of parameters. The correlation is similar for all linear molecules with similar variations of all contributing values, so that related properties of any linear molecules vary similarly. This was exemplified by various physicochemical properties (boiling point, melting point, heat of sublimation, heat of vaporization, diamagnetic susceptibility, ionization potential, long-wave transition energy, hyperpolarizability, logarithm of the octanol–water distribution coefficient) for all saturated linear hydrocarbons and corresponding perfluoro derivatives, unsaturated linear hydrocarbons and those with conjugated double bonds, oligothiophenes, linearly fused aromatic rings, RNA duplexes, and bilayer lipide membranes. The simplified relations obtained allow a sufficiently precise assessment of various properties of any linear molecule.

Proceeding from reasonable assumptions we proposed in [2] a structure–property correlation that allows the “property” to be defined in the most general form. This correlation allows precise assessment of various physicochemical properties of organic compounds belonging to different classes [1, 3–5]. For molecules of the A_n type, built of the same fragments, the correlation takes the following form:

$$P = n\alpha + (n - 1)\beta + (n - 2)\gamma + (n - 3)\delta_{14} + \ldots, \quad (1)$$

where α is a one-center contribution to the property of the A_i fragment, β is a two-center contribution corresponding to interaction of the ith and $(i + 1)$th fragments, γ is a two-center contribution corresponding to interaction of the ith and $(i + 2)$th fragments, etc.

As far as any property a set of molecules A_n (n = 1, 2, 3...) can be considered as a function of n [$P = f(n)$], in this work we will explore the possibility to represent correlation (1) in the form of a simple function of n, that comprises a restricted number of parameters but allows a rather precise assessment of various properties. Therewith, the simplification of Eq. (1) should not be an occasional choice but rely on the simplification of an equation already successfully applied previously [1, 3–5].

Insofar as similar properties of different molecules A_n, B_n, etc., are described by the same equation (1) and its contributions β, γ and others tend to decrease in a similar manner with increasing distance between fragments, the same properties of different linear molecules should similarly vary with n. Actually, various physicochemical properties such as boiling point, heat of vaporization, diamagnetic susceptibility, ionization potential, energy of long-wave transitions, hyperpolarizability, and octanol–water distribution coefficient vary similarly in series of saturated linear hydrocarbons (n-alkanes) [6,7,11], related perfluoro derivatives (perfluo-n-alkanes) [6], linear unsaturated hydrocarbons with conjugated double bonds (alkyloligoenes) [6, 9], oligothiophenes [10], and linearly fused aromatic rings (acenes) [6, 8, 11] (Tables 1 and 2). Moreover, the dependences for the melting points of RNA duplexes of different size [12, 13] or bilayer membranes formed by lipids of different size [14, 15] resemble the dependence of the boiling points of n-alkanes on the hydrocarbon chain length (n).

It follows from Tables 1 and 2 that the character of variation of different properties with the number of fragments n is not the same. Thus the boiling points increase slower as n increases, ionization potentials and long-wave transition energies decrease with n, and hyperpolarizability increases as n^a ($a > 1$). The latter property is described by an equation whose numerator slowly varies with n and denominator is the difference in the energies of the ground and excited states, which decreases with n; as a result, polarizability rapidly increases. The other properties may vary analogously. To describe them as a function of n we transform Eq. (1) to the following form:

---

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

Table 1. Boiling points of linear molecules A_n [6]

Molecule	T_b, K	Molecule	T_b, K
n-Alkanes		Perfluoro-*n*-alkanes	
CH_4	111.7	CF_4	145.1
C_2H_6	184.5	C_2F_6	195.0
C_3H_8	231.1	C_3F_8	236.5
C_4H_{10}	272.7	C_4F_{10}	271.2
C_5H_{12}	309.2	C_5F_{12}	302.4
C_6H_{14}	341.9	C_6F_{14}	329.8
$C_{10}H_{22}$	447.3	$C_{10}F_{22}$	417.4
Alkyloligoenes		Acenes	
Ethylene	169.4	Benzene	353.2
1,3-Butadiene	268.7	Naphthalene	471.1
1,3,5-Hexatriene	351.6	Anthracene	613.1
		Tetracene	616.1

$$\begin{aligned} P &= n\alpha + (n-1)\beta + (n-2)\gamma + (n-3)\delta_{14} \\ &+ \ldots \pm\beta\pm 2\gamma\pm 3\delta_{14}\pm\ldots = [n(\alpha+\beta+\gamma+\delta_{14}+\ldots)-\beta]-2\gamma \\ &- 3\delta_{14} - \ldots = [n(\alpha+\beta+\gamma+\delta_{14}+\ldots)-\beta-2\gamma]-3\delta_{14}-\ldots \\ &= k_1 + k_2 n + k_3 f(n). \qquad (2) \end{aligned}$$

Here k_2 is the sum $(\alpha + \beta + \gamma + \delta_{14} +\ldots)$, k_1 is β or β combined with other nearest contributions, $k_3f(n)$ comprises all contributions not included to k_1. The decomposition of $P(n)$ into separate contributions is represented schematically in Fig. 1. It is seen that $f(n)$ is nonlinear function descending with n, and it can be represented as $f(n) \cong e^{-n}$ or $f(n) \cong 1/n$. Thus, all problems associated with choosing the form of $P(n)$ we have transferred to choosing the form of $f(n)$. Obviously, the choice of $f(n)$ will not always be satisfactory. To go around this drawback, we assume that n in Eq. (2) is an "effective" value that differs from the real n.

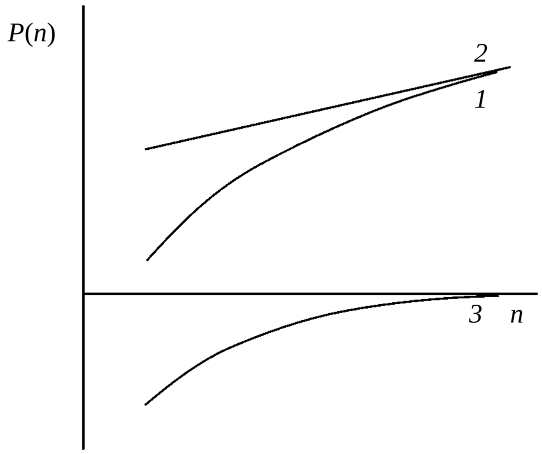

Fig. 1. Decomposition of $P(n)$ into separate contributions. (*1*) $P_{exp} \cong k_1 + k_2n + k_3f(n)$, (*2*) $k_1 + k_2n$, and (*3*) $k_3f(n)$.

Table 2. Selected properties of linear molecules A_n

Type of molecule	Numer of fragments *n*				
	1	2	3	4	5
Heat of sublimation, kcal mol^{-1} [7]					
n-Alkanes	2.20	4.90	6.81	8.57	10.03
Acenes	10.7	17.3	23.6	29.8	37.3
Diamagnetic susceptibility, $-\chi_m \times 10^6$ [6]					
n-Alkanes	12.2	27.2	40.0	50.0	63.0
Acenes	54.8	91.0	130.0	168.0	205.4
Ionization potential, eV [6]					
n-Alkanes	12.51	11.52	10.95	10.53	10.35
Alkyloligoenes	10.49	9.02	8.23	7.80	
Long-wave transition energy, eV					
Acenes [11]	3.59	2.18	1.66	1.10	0.79
Alkyloligoenes [9]	7.65	5.91	4.93	4.40	
3rd-order hyperpolarizability, $\times 10^{-36}$ au					
Alkyloligoenes [9]	0.76	2.13	7.53		
Oligothiophenes [10]	4.1	22	160	800	2600
Logarithm of the octanol–water distribution coefficient [11]					
n-Alkanes	1.09	1.81	2.36	2.89	3.39
Acenes	2.13	3.30	4.45	5.90	
Polyols	–0.77	–1.36	–1.76	–2.29	

Equation (1) can also be transformed in another way:

$$\begin{aligned} P &= [n\alpha + (n-1)\beta + (n-2)\gamma] + (n-3)\delta_{14} + \ldots \\ &= [(n-1)\alpha + (n-1)\beta + (n-1)\gamma + \alpha - \gamma] + (n-3)\delta_{14} + \ldots \\ &= [(n-2)\alpha + (n-2)\beta + (n-2)\gamma + 2\alpha - \beta] + (n-3)\delta_{14} \\ &= k_1 + k_2 n_{eff} + k_3 f(n_{eff}). \qquad (3) \end{aligned}$$

Here n_{eff} is $(n - i)$, k_2 is the sum $(\alpha + \beta + \gamma)$, k_1 is a combination of the α, β and γ contributions, $k_3f(n_{eff})$ is the sum of all other contributions. Like in the preceding case, correlation (3) is shown in Fig. 2. As seen, $f(n_{eff})$ is a nonlinear function growing with n_{eff}. It can be written down as $f(n_{eff}) = e^{n_{eff}}$. Equations (2) and (3) exhaust all possible simplifications of Eq. (1). They assume that a property P is a combination of linear and nonlinear functions. Generally, both these functions can be nonlinear, then $n_{eff} = n^a$, where $a \neq 1$.

One can easily find n_{eff} values providing the best estimates of various molecular properties. However, a

Table 3. Correlation coefficients of properties of saturated hydrocarbons with n^a

Property	n	$n^{0.5}$	$n^{0.25}$
Boiling point	0.977	0.999	0.998
Heat of vaporization, ΔH_{vap}	0.992	1	0.999
Heat capacity, C_p	1	0.984	0.962
Formation entropy, S^0	1	0.987	0.966
Formation enthalpy, ΔH_f^0	1	0.984	0.961
Formation Gibbs energy, ΔG_f^0	0.999	0.989	0.970
Logarithm of the gas–hexadecane distribution coefficient, log L^{16}	1	0.989	0.971

simpler way is to correlate various properties with assumed values of $n_{eff} \cong n^a$. These data are given in Table 3.

From Table 3 it follows that T_b and ΔH_{vap} best correlate with $n^{0.5}$, while C_p, S^0, ΔH_f^0, ΔH_f^0 and log L_{16}, with n. Therefore, it is reasonable to take $n_{eff} = n^{0.5}$ for the first two properties and $n_{eff} = n$ for all the others. The estimates for T_b and ΔH_{vap} obtained by Eqs. (2) and (3) with $n_{eff} = n^{0.5}$ are given in Table 4.

Using Eq. (2) with $n_{eff} = n$ we obtained estimates for the other properties with a very good precision. The statistical criteria for these estimates are listed in Table 5. The fit of calculation to experiment was always good.

Unlike the above-considered properties, ionization potentials, hyperpolarizabilities and long-wave transition energies are known for much less molecules. As noted above (Table 2), the ionization potential and long-wave transition energy, by contrast to the other characteristics considered, decrease with increasing n

Table 4. Boiling points and vaporization temperatures of linear saturated hydrocarbons

Molecule	bp, K			ΔH_{vap}, kcal mol^{-1}		
	experiment [6]	calculation by		experiment [7]	calculation by	
		Eq. (2)	Eq. (3)		Eq. (2)	Eq. (3)
CH_4	111.4	108.9	115.7	1.95	1.97	1.96
C_2H_6	184.5	182.9	180.9	3.52	3.41	3.39
C_3H_8	231.1	235.1	230.6	4.49	4.51	4.49
C_4H_{10}	272.6	276.9	272.3	5.35	5.43	5.42
C_5H_{12}	309.2	312.5	308.7	6.16	6.23	6.23
C_6H_{14}	341.9	343.8	341.3	6.90	6.96	6.96
C_7H_{16}	371.6	372.1	371.0	7.58	7.63	7.63
C_8H_{18}	398.8	398.0	398.3	8.22	8.25	8.26
C_9H_{20}	423.9	422.1	423.6	9.03	8.83	8.85
$C_{10}H_{22}$	447.3	444.7	447.3	9.39	9.38	9.40
$C_{11}H_{24}$	470.0	466.0	469.3	9.92	9.91	9.93
$C_{12}H_{26}$	489.3	486.2	490.1	10.43	10.41	10.43
$C_{13}H_{28}$	508.6	505.5	509.5	10.91	10.88	10.90
$C_{14}H_{30}$	526.7	524.0	527.8	11.38	11.34	11.36
$C_{15}H_{32}$	543.8	541.8	544.9	11.82	11.79	11.80
$C_{16}H_{34}$	560.0	559.0	561.0	12.24	12.22	12.22
$C_{17}H_{36}$	575.3	575.6	576.0	12.64	12.63	12.63
$C_{18}H_{38}$	589.9	591.6	590.0	13.02	13.04	13.02
$C_{19}H_{40}$	603.7	607.2	603.1	13.39	13.43	13.40
$C_{20}H_{42}$	616.9	622.4	615.1	13.74	13.81	13.77
k_1		33.8773 ± 6.0515	–42.0513 ± 1.7850		–1.2739 ± 0.1368	–1.5024 ± 0.0721
k_2		131.9840 ± 1.5515	159.5338 ± 0.8823		3.3748 ± 0.0351	3.4694 ± 0.0356
k_3		–154.7081 ± 17.2563	–0.6428 ± 0.0344		–0.3617 ± 0.3902	–0.0028 ± 0.0014
r		0.9998	0.9999		0.9998	0.9998
s		3.0655	1.5820		0.0693	0.0639

Table 5. Statistical criteria of estimates for selected properties of saturated hydrocarbons

Property[a]	r	s	k_1	k_2	k_3
C_p	1.0000	0.1176	1.2376 ± 0.0686	5.4773 ± 0.0054	4.6169 ± 0.3735
ΔG_f^0	0.9998	0.2487	–11.7855 ± 0.1451	1.9846 ± 0.0114	–5.5375 ± 0.7902
ΔH_f^0	0.9999	0.3381	–10.1804 ± 0.1972	–4.9349 ± 0.0156	–6.6704 ± 1.0743
S^0	1.0000	0.0662	36.9344 ± 0.0386	9.3179 ± 0.0030	–4.8294 ± 0.2103
$\log L^{16}$	1.0000	0.0113	–0.3677 ± 0.0066	0.5051 ± 0.0005	–1.2475 ± 0.0359

[a] Range of variation: C_p 8.5 to 110.7, ΔG^f –12.1 to 28.0, ΔH_f^0 –17.9 to –28.0; S^0 44.5 to 223.3; and $\log L^{16}$ –93.0 to 9.73.

and tend to a limit at high n values. The dependence of the ionization potential on n can be given as

$$I = k_1 + k_2 e^{-n_{\text{eff}}}.$$

This equation with $n_{\text{eff}} = n^{0.5}$ fits the experimental data with $r = 0.9984$ and $s = 0.0557$. The long-wave transition energies and reciprocal hyperpolarizabilities are estimated similarly with approximately the same precision.

The described approach is feasible for assessment of properties of R–X and R^1–X–R^2 linear molecules with only one functional group X in the chain (e.g., alcohols and ethers). The precision of the assessment is the same as in the case of saturated hydrocarbons.

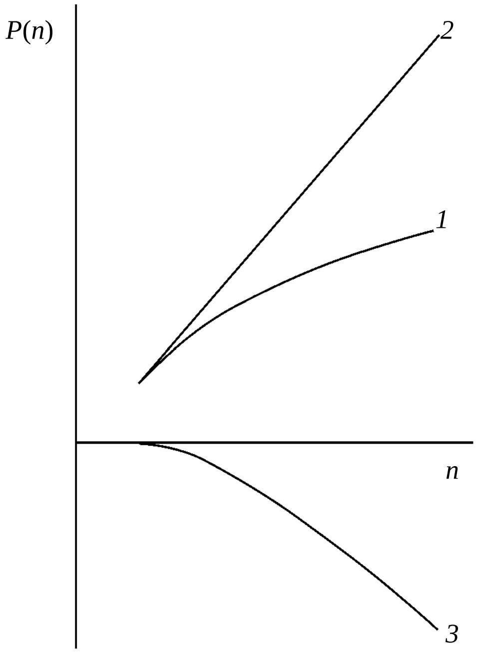

Fig. 2. Decomposition of $P(n)$ into separate contributions. (*1*) $P_{\text{exp}} \cong k_1 + k_2 n_{\text{eff}} + k_3 f(n_{\text{eff}})$, (*2*) $k_1 + k_2 n_{\text{eff}}$, and (*3*) $k_3 f(n_{\text{eff}})$.

Thus, Eq. (1) and the form of dependence of its terms β, γ and others on the distance between fragments provides evidence for a similar dependence on chain length of related properties of linear molecules belonging to different classes: *n*-alkanes, perfluoro-*n*-alkanes, alkylpolyenes, acenes, oligothiophenes, polyols, etc. Simplification of Eq. (1) leads to fairly simple expressions (2) and (3) whose application allows various properties for any linear molecules to be assesses to a good precision. This is useful not only methodically but also conceptually, as far as allows one to predict from properties of one group of linear molecules related properties of linear molecules belonging to other classes.

REFEREENCES

1. Golovanov, I.B. and Zhenodarova, S.M., *Zh. Obshch. Khim.*, 2003, vol. 73, no. 10, p. 1696.
2. Golovanov, I.B. and Zhenodarova, S.M., *Zh. Obshch. Khim.*, 2003, vol. 73, no. 1, p. 90.
3. Golovanov, I.B., and Zhenodarova, S.M., *Zh. Obshch. Khim.*, 2003, vol. 73, no. 2, p. 260.
4. Golovanov, I.B. and Zhenodarova, S.M., *Zh. Obshch. Khim.*, 2003, vol. 73, no. 3, p. 400.
5. Golovanov, I.B., Zhenodarova, S.M., and Smolyaninova, O.A., *Zh. Obshch. Khim.*, 2003, vol. 73, no. 4, p. 553.
6. *CRC Handbook of Chemistry and Physics,* Lide, D.R., Ed., Boca Raton: CRC, 1993–1994, Sect. 3, 9, 10.
7. Kitaigorodskii, A.I., *Molekulyarnye kristally* (Molecular Crystals), Moscow: Nauka, 1971.
8. Pariser, R.J., *Chem. Phys.,* 1956, vol. 24, no. 2, p. 250.
9. Pierce, B.J., *J. Chem. Phys.,* 1989, vol. 91, no. 2, p. 791.

10. Zhao, M.T., Singh, B.P., and Prasad, P.N., *J. Chem. Phys.*, 1988, vol. 89, no. 9, p. 5535.
11. Hansh, C., Leo, A., and Hoekman, D., *Exploring QSAR*, Washington: Am. Chem. Soc., 1995.
12. Martin, F.H., Uhlenbeck, O.C., and Doty, P., *J. Mol. Biol.*, 1971, vol. 57, no. 2, p. 201.
13. Uhlenbeck, O.C., Martin, F.H., and Doty, P., *J. Mol. Biol.*, 1971, vol. 57, no. 2, p. 217.
14. Phillips, M.C. and Chapman, D., *Biochim. Biophys. Acta*, 1969, vol. 163, no. 3, p. 301.
15. Phillips, M.C., Ledbrooke, B.D., and Chapman, D., *Biochim. Biophys. Acta,* 1970, vol. 196, no. 1, p. 35.