Topological Index and Thermodynamic Properties. II.[†] Analysis of the Topological Factors on the Absolute Entropy of Acyclic Saturated Hydrocarbons

Hideyuki NARUMI and Haruo HOSOYA*,[††]
Department of Chemistry, Hokkaido University, Sapporo 060
[††] *Department of Chemistry, Ochanomizu University, Bunkyo-ku, Tokyo 112*

(Received September 11, 1979)

The absolute entropy of acyclic saturated hydrocarbons is found to be well correlated linearly with the logarithm of the topological index (Z_G) of the graph of the carbon atom skeleton as, $S°=a\log Z_G+b$, excpet for sterically overcrowded and highly symmetrical molecules. This relation is improved by taking the symmetry factor σ into account as, $S°=a\log Z_G+b+c\log\sigma$. The physical meaning of these relations is clarified by analyzing the mathematical properties of Z_G and the rotational partition function proposed by Pitzer.

Recently importance of the topological or graph-theoretical aspects of molecules has been recognized and detailed analyses have been developed by many authors.[1-4] The problems discussed are, for example, enumeration of the number of the structural isomers,[1] analysis of the stability and the electronic distribution on π-electron networks,[5-7] coding and classification of the structures of molecules,[8,9] and interpretation of empirical rules on the thermodynamic properties of organic molecules.[10-12] Concerning the last problem one of the present authors has proposed the topological index[13] (Z_G) for the carbon atom skeleton of a saturated hydrocarbon and has shown that several empirical rules on the boiling point (bp) are interrelated to each other by assuming the relation

$$\text{bp} = a\log Z_G + b.^{12)} \tag{1}$$

However, physical interpretation of Eq. 1 is not yet fully given.

It is practically very difficult to estimate the value of the boiling point of a compound from purely theoretical calculations based on an *a priori* model for the motions of molecular aggregates. Then a clue for solving this problem might be obtained by analyzing the topological aspects of other quantities, such as the entropy, critical pressure and volume, *etc.*, whose behaviors are similar to that of the boiling point but are more interpretable by the molecular theory.[14]

According to the theory developed by Pitzer,[15] the observed values of the absolute entropy $S°$ of small acyclic saturated hydrocarbons can accurately be reproduced from the partition function obtained solely by the atomic and molecular constants and temperature.[15-19] However, from this theory and numerical calculations, it is difficult to draw generalized rules and their physical interpretation for the topological dependence of $S°$. Namely, so many different factors especially on the rotational conformations are entangled in the Pitzer's theory that practically the calculation of $S°$ values is limited to small alkanes, say, with fewer than ten carbon atoms.

On the contrary, the topological index has been known to reflect the topological characteristics of the zig-zag chain especially on the degree of rotational freedom and to have a number of mathematically interesting properties. Namely, we have found that $S°$ of saturated hydrocarbons is also related to the topological index as

$$S° = a\log Z_G + b. \tag{2}$$

The purposes of the present paper are to obtain logical bases for the empirical relation 2 in terms of the Pitzer's partition function for clarifying the topological nature of the absolute entropy of van der Waals gases.

In Section I the mathematical properties of the topological index for tree graphs are reviewed with particular emphasis on the effect of branching including some additivity relationship. In Section II parallelism between the observed and calculated (by the topological index, Pitzer's partition function, and an empirical formula) entropies of small alkanes is presented and several important features of the branching effect on $S°$ are exposed. It is also demonstrated that as the increase of the size of the molecule, to the extent where the Pitzer's theory cannot follow, the relation between $\log Z_G$ and observed $S°$ gets better. In Section III by analyzing the relation between Z_G and the partition function the empirical relation 2 is justified.

I. Topological Index

Definition.[13] A graph G is composed of points and lines, or carbon atoms and C–C bonds of the carbon atom skeleton of a saturated hydrocarbon molecule. The non-adjacent number $p(\mathrm{G},k)$ for G is defined as the number of ways for choosing k ($0\sim m$) disjoint lines, with $p(\mathrm{G},0)$ being defined as unity. The topological index Z_G is the sum of the $p(\mathrm{G},k)$'s,

$$Z_G = \sum_{k=0}^{m} p(\mathrm{G},k). \tag{3}$$

For a tree graph, or an acyclic hydrocarbon, $p(\mathrm{G},k)$ is nothing else but the coefficient of the characteristic polynomial of G

$$\begin{aligned} P_G(X) &= (-1)^n \det(\boldsymbol{A}-X\boldsymbol{E}) \\ &= \sum_{k=0}^{n} a_k X^{n-k} \\ &= \sum_{k=0}^{[n/2]} (-1)^k p(\mathrm{G},k) X^{n-2k}, \end{aligned} \tag{4}$$

where $\boldsymbol{A}$ is the adjacency matrix of G with elements

---

[†] Part I, *Bull. Chem. Soc. Jpn.*, **45**, 3415 (1972).

$$A_{rs} = \begin{cases} 1 & \text{if points r and s are neighbors} \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

$\boldsymbol{E}$ is the unit matrix of the same order n (the number of points in G) as $\boldsymbol{A}$, and det stands for a determinant. From Eqs. 3 and 4 one gets

$$Z_G = (-i)^n P_G(i) \tag{6}$$

for a tree graph.

Recursion Formulas.[13,20] The Z_G values for small tree graphs are given in Table 1, where the observed and calculated absolute entropy of the corresponding alkanes are compared for the later discussion.[21] For larger graphs, Z_G can easily be obtained from the values of smaller subgraphs by using the recursive relations described below.

Suppose that a graph G is obtained from graphs A and B by joining them with line l as in Fig. 1. The topological index Z_G of G can be given as

$$Z_G = Z_L \cdot Z_M + Z_A \cdot Z_B, \tag{7}$$

where L and M are the subgraphs obtained by deleting line l from G.

Suppose that a graph G is obtained from graphs A and B by joining them with point p as in Fig. 2. Graphs L and M which are obtained by adding point p to A and B, respectively, are the same as in Fig. 1. The topological index Z_G of G can be obtained by

$$Z_G = Z_L \cdot Z_B + Z_A \cdot Z_M - Z_A \cdot Z_B. \tag{8}$$

Branching Properties.[12] The topological indices of the various series of branched alkanes have a number of interesting mathematical properties which are in parallel with the experimentally known features of the absolute entropy. It is remarkable that these properties are maintained up to an infinitely large system. In the following discussion the carbon atom skeleton of an acyclic alkane is represented by a tree graph.

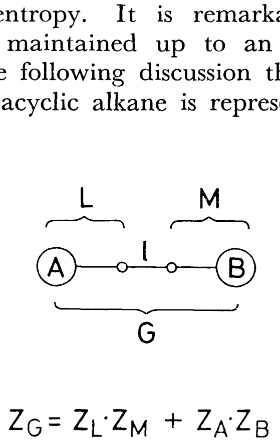


Fig. 1. Recursion formula 1.

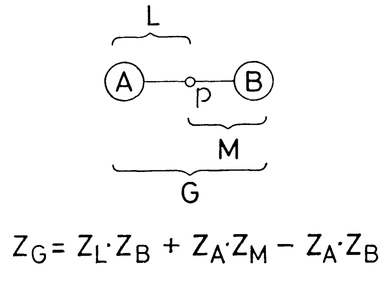


Fig. 2. Recursion formula 2.

TABLE 1. TOPOLOGICAL INDEX AND OBSERVED AND CALCULATED ABSOLUTE ENTROPY OF SMALL ALKANES

Alkane[b]	Z_G	Calculated S° [a]			Observed S° [f]
		$a \log Z_G + b$[c]	Pitzer[d]	P. & S.[e]	
4	5	74.41	74.10	74.1	74.07
2m3	4	70.12	70.48	69.9	70.60
5	8	83.44	83.18	83.1	83.52
2m4	7	80.87	82.27	81.2	82.13
22m3	5	74.41	73.04	71.8	73.11
6	13	92.76	92.32	92.4	92.91
3m5	12	91.23	91.48	90.3	91.52
2m5	11	89.55	90.80	90.3	91.03
23m4	10	87.72	87.32	85.4	87.43
22m4	9	85.70	85.66	84.6	85.62
7	21	101.98	101.45	101.5	102.29
3e5	20	101.04	97.87	97.2	98.36
3m6	19	100.05	101.38	100.8	101.80
2m6	18	99.02	100.04	99.4	100.48
23m5	17	97.92	98.79	97.3	99.08
33m5	16	96.75	95.83	94.5	95.16
24m5	15	95.51	94.64	94.5	94.86
22m5	14	94.19	93.34	93.7	93.83
223m4	13	92.76	92.27	90.2	91.62
s[g]	—	1.160	0.380	0.539	—
r[h]	—	0.9927	0.9993	0.9985	—

a) At 298.15 K in cal·K^{-1}·mol^{-1}. b) Abbreviated notation. For example, 24m5 stands for 2,4-dimethylpentane, and e for ethyl. See also Fig. 4. c) Equation 2 with $a=43.483$ and $b=44.240$. d) Equation 25 with $C_0=43.012$, $C_1=7.679$, $C_2=-1.068$, $C_3=-3.425$. See Table 3 for σ values ($a=500$ cal·mol^{-1}). e) Equation 30. See Table 3 for σ. f) Ref. 22. g) Standard deviation. h) Correlation coefficient.

TABLE 2. TOPOLOGICAL CHARACTERISTICS OF SEVERAL SERIES OF GRAPHS

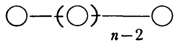
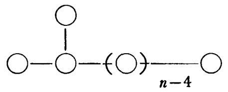
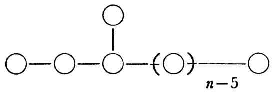
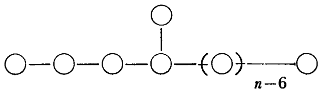
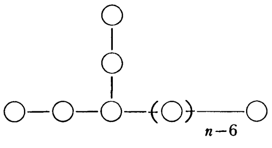
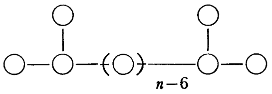
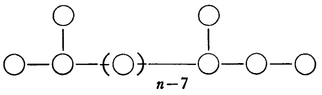
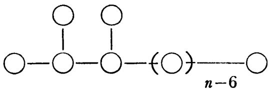
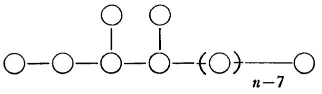
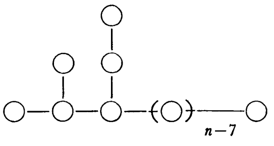
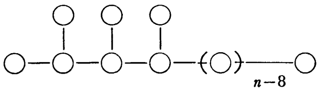
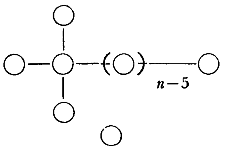
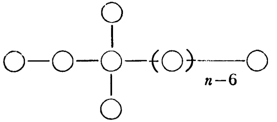
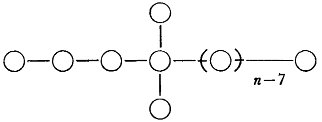
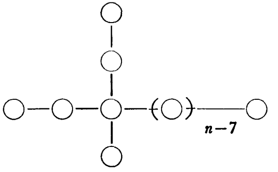
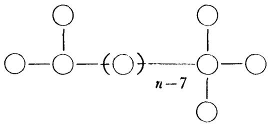
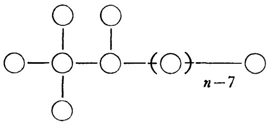
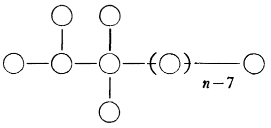
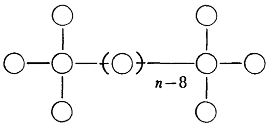
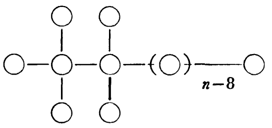
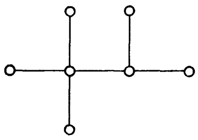
Type	G_n	n_Y	n_X	Z_1	Z_2	a_G	b_G	$1000\delta_G = 1000\log(a_N/a_G)$	
N		0	0	1	2	0.7236	0.2763	0	
A		1	0	2	1	0.6180	−1.6180	68.5	→ 46
B				−1	3	0.6583	3.3417	41.0	
C				7	−2	0.6429	−9.6429	51.3	
D				4	0	0.6833	−4.6833	24.9	
E		2	0	−5	5	0.5279	9.4721	137.0	→101
F				13	−6	0.5623	−19.5623	109.5	
H				5	−1	0.5777	−6.5777	97.8	
I				−11	9	0.6085	19.3915	75.2	
J				−3	4	0.5931	6.4069	86.4	
K		3	0	29	−16	0.5315	−45.5315	134.0	→ 134
L		0	1	−2	3	0.4876	4.5124	171.5	→135
M				8	−3	0.5374	−11.5374	129.2	
O				−18	13	0.5183	30.4817	144.9	
P				−16	12	0.5836	27.4164	93.4	

Table 2. (Continued)

Type	G_n	n_Y	n_X	Z_1	Z_2	a_G	b_G	$1000\delta_G = 1000 \log (a_N/a_G)$	
Q		1	1	17	−9	0.4164	−26.4164	240.0	→205
R				−7	6	0.4626	12.5374	194.3	
S				−15	11	0.4780	25.5220	180.1	
T		0	2	−45	29	0.3285	73.6715	343.0	→308
U				33	−19	0.3856	−52.3856	273.4	



$Q_7 = R_7 = S_7$

Given a tree graph and try to enlarge it to form a series of graphs with similar mode of branching by inserting points one by one in between certain prescribed lines. As in Fig. 3 there are three possible ways for doing so, since Q_7 has three topologically different types of lines. In this way one gets three series of graphs, Q, R, and S. The original graph is turned out to be the smallest and the common member of the series Q, R, and S. As expected from this procedure, an infinite number of topologically different series of graphs are derived. In Table 2 are given twenty fundamental series of graphs with different modes of branching. They are subdivided into several classes with respect to the number of tertiary (n_Y) and quarternary (n_X) carbon atoms.

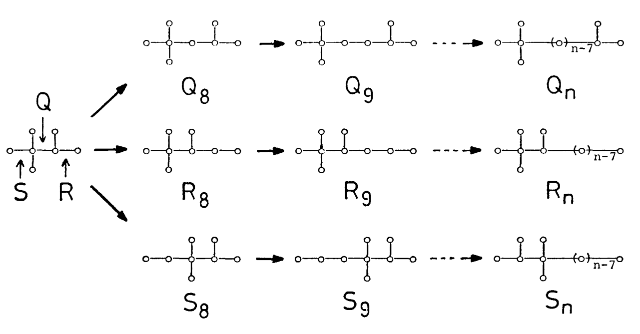
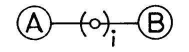


Fig. 3. Generation of different series of graphs from a parent graph.

Consider a series of graphs A-*i*-B composed of subgraphs A and B and *i* consecutive line of points. Successive application of Eqs. 7 and 8 gives the following recursive relation for the topological indices of these graphs.



A-*i*-B ($i \geq 0$)

$$Z_{A-(i+2)-B} = Z_{A-(i+1)-B} + Z_{A-i-B}. \tag{9}$$

Therefore, if the topological index of any series of graph G_n with n points as given in Table 2 is denoted by Z_n, a recursive relation similar to Eq. 9 can be written as

$$Z_{n+2} = Z_{n+1} + Z_n. \tag{10}$$

Although graph A-*i*-B has been defined for non-negative integers of *i*, the recursive relation 10 can formally be extended down to as small integers as one wishes irrespective of the existence of the corresponding graph. In Table 2 are given the Z_1 and Z_2 values for all the series of molecules studied. Although from these values the topological index Z_n for graph G_n can be obtained by the successive application of Eq. 10, one can obtain the closed and asymptotic forms of Z_n as in the following discussion.

It is known that if a series of numbers Z_n obey a recursive relation as 10, Z_n can be expressed explicitly by

$$Z_n = a\alpha^n + b\beta^n,^{23,21)} \tag{11}$$

where α and β are the roots of the equation

$$x^2 - x - 1 = 0,$$

namely,

$$\alpha = (1+\sqrt{5})/2 \text{ and } \beta = (1-\sqrt{5})/2. \tag{12}$$

The constants a and b are obtained if two of the members of the series $\{Z_n\}$ are given. Here by putting $n=1$ and 2 one gets

$$\begin{aligned} a &= (Z_1+\alpha Z_2)/(1+3\alpha) \\ b &= (Z_1+\beta Z_2)/(1+3\beta), \end{aligned} \tag{13}$$

which enables the calculation of any member of Z_n with Eq. 11. The values of a and b differ from series to series as in Table 2 and may be called as the structural factors.

For larger n the second term of Eq. 11 vanishes because of the relation $\alpha>1>|\beta|$. Then one gets

$$Z_n \rightarrow a\alpha^n \tag{14}$$

or

$$\log Z_n \rightarrow \log a + n \log \alpha. \tag{15}$$

The ratio Z_{n+1}/Z_n approaches to a constant value α for all the series of graphs,

$$Z_{n+1}/Z_n \rightarrow \alpha. \tag{16}$$

It is found from Table 2 that the value of a_G (a for a series G) is the largest for normal alkanes (series N) and is roughly determined by the number and mode of branching, whereas the value of b_G is scattered from series to series. This feature will be clarified if the value $\delta_G=\log(a_N/a_G)$ is compared with each other. Note that

$$\begin{aligned} \log Z_N - \log Z_G &\rightarrow \log a_N - \log a_G \\ &= \delta_G. \end{aligned} \tag{17}$$

As is evident from Table 2 the $\overline{\delta}_G$ averaged over a group of series with the same values of n_Y and n_X is roughly given by

$$100\,\overline{\delta}_G \fallingdotseq 5(n_Y+3n_X). \tag{18}$$

This means that the decrement of the logarithm of the topological index Z_G of a given graph from that of the linear graph with the same number of points is roughly determined by the number of the tertiary and quarternary carbon atoms in a simple manner as in Eq. 18. Recall that the effect of n_X on the lowering of the $p(\mathrm{G}, 2)$ number is three times as large as that of n_Y. See Eq. 12 of Ref. 13.

The effect of the quarternary carbon atom δ_X is three times larger than that of the tertiary carbon atom δ_Y. Namely,

$$\delta_L = 2.5\delta_A, \quad \delta_M = 3.2\delta_B, \quad \delta_O = 2.8\delta_C \tag{19}$$

or simply as

$$\delta_X \fallingdotseq 3\delta_Y. \tag{20}$$

Further inspection of the δ_X values reveals interesting additivity relationships. Namely, we have rigorous equalities as

$$\begin{aligned} \delta_E &= 2\delta_A \\ \delta_T &= 2\delta_L \\ \delta_Q &= \delta_A + \delta_L \;. \end{aligned} \tag{21}$$

That is, the value of δ for a graph with two branches is the sum of the δ values for the two graphs each with a single component branch, if the two roots of the branching are separated at least by one point, *i.e.*, one CH_2 unit. In Appendix I a general proof of this property of the δ value is given.

If two branches are adjacent the δ value is always smaller than the sum of the δ values for the component structures. For example, we can see

$$\begin{aligned} \delta_H &< \delta_A + \delta_B, \\ \delta_I &< \delta_B + \delta_C, \\ \delta_J &< \delta_A + \delta_D. \end{aligned} \tag{22}$$

Longer branching causes much smaller decrement in δ value than a shorter branching at the same root, for example,

$$\delta_D < \delta_B. \tag{23}$$

Finally the relations

$$\begin{aligned} &\delta_A > \delta_C > \delta_B, \\ \text{and} \quad &\delta_L > \delta_O > \delta_M \end{aligned} \tag{24}$$

show a zig-zag effect of the branching in going from the terminal to the center. All these properties can be deduced mathematically as in the case of the relation 20.

II. Relation between S° and Topological Quantities**

In Fig. 4 are shown the relations between the Z_G values and the observed absolute entropy S° of smaller ($n\leq 7$) acyclic saturated hydrocarbons. Similar plots are extended as in Fig. 5 for larger members of several series (normal (●), 3-methyl (△), 2,2-dimethyl (○), 2,3-dimethyl (+), and 2,4-dimethyl (□) alkanes) of saturated hydrocarbons. These two plots indicate several important aspects of the relation 2. Namely, Fig. 4 shows that the more the branching the smaller the value of S° as does Z_G. Introduction of a quarternary carbon atom causes much larger effect than that of two tertiary carbon atoms. It should also be noted here that as in the case of the boiling point 2-methyl substitution causes larger decrease in S° than does 3-methyl substitution in parallel with the behavior of Z_G. Inspection of the correlation

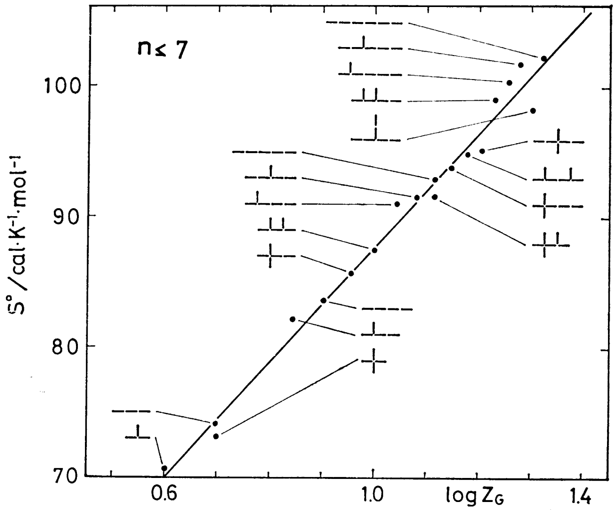


Fig. 4. $\log Z_G$—S° plot for smaller alkanes ($n\leq 7$).

** 1 cal=4.184 J.

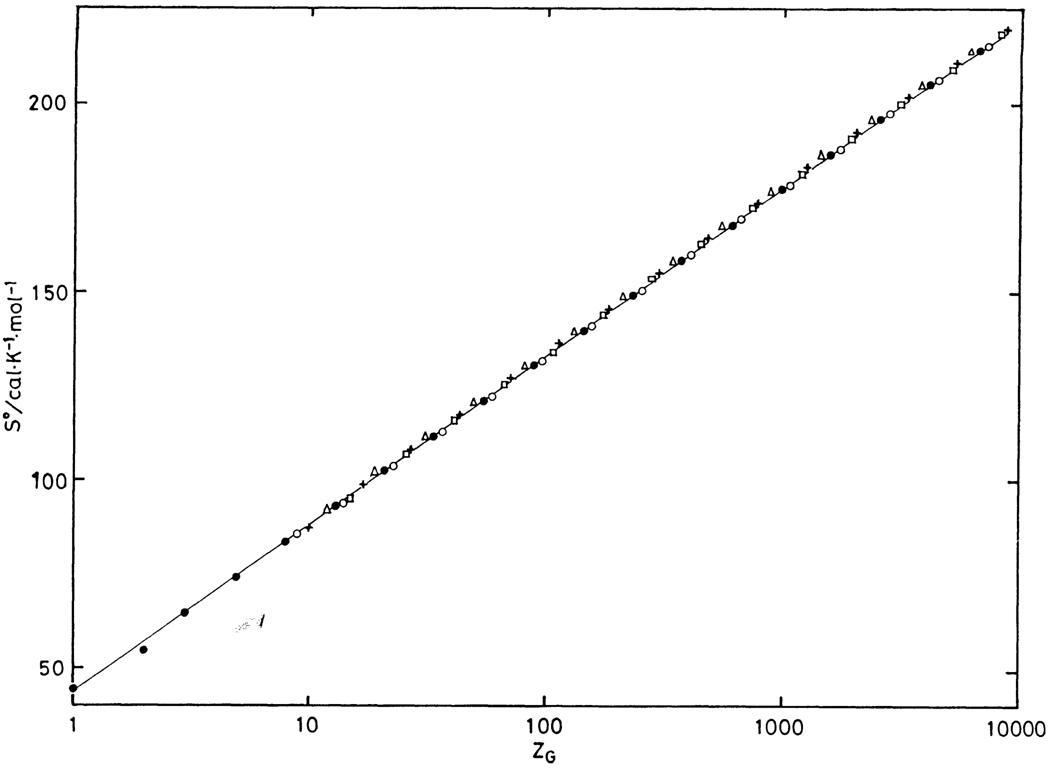


Fig. 5. log Z_G—S° plot for larger members ($n \leq 20$) of several series of alkanes.
●: Normal, △: 3-methyl, ○: 2,2-dimethyl, +: 2,3-dimethyl, □: 2,4-dimethyl.

between the topological structure of alkane and entropy reveals that the effect of the tertiary and/or quarternary carbon atoms in lowering the S° value is additive if there is at least a secondary (CH_2) carbon atom between these branching sites. Further, it is observed that the higher the symmetry, the smaller the S° value. However, Z_G has nothing to do with the symmetry of a graph. The remarkably good linear relation between S° and $\log Z_G$, especially within each series, in Fig. 5 demonstrates that the growth of the chain length by one CC bond in eash series contributes the same amount (9.38 cal·K^{-1}·mol^{-1}) of increase in the entropy and that $\log Z_G$ has just the same property as has been shown above.

Although a number of empirical relations have been proposed for the topological dependence of the absolute entropy or the related quantities of branched compounds, especially of the saturated hydrocarbons, in this paper, we will discuss only three of them, namely, the relations proposed by the present author (topological index[12,13]), by Pitzer (partition function[15]), and by Pitzer and Scott (purely empirical one[18]), since the parameters involved in other theories do not seem to be easily interpretable.[4,10,11]

Pitzer has derived the expression of S° of acyclic saturated hydrocarbons in terms of the molecular constants by enumerating the partition functions for the molecular motions. His original expression can be transformed into as

$$S^\circ = C_0 + C_1 n + C_2 n_Y + C_3 n_X + \Delta, \qquad (25)$$

$$\Delta = R(\ln Q + Q'/Q - \ln \sigma), \qquad (26)$$

$$Q = \sum_{i}^{\text{confor-mations}} \exp(-\Delta\varepsilon_i/kT) = \sum_{j} g_j \exp(-\Delta\varepsilon_j/kT), \qquad (27)$$

$$Q' = \sum_{i}^{\text{confor-mations}} \frac{\Delta\varepsilon_i}{kT} \exp(-\Delta\varepsilon_i/kT)$$

$$= \frac{a}{kT} \sum_{j} g_j \exp(-\Delta\varepsilon_j/kT), \qquad (28)$$

$\Delta\varepsilon_i$=relative energy of i-th conformation
g_j=number of conformations with j excess *gauche* pairs relative to the conformation with the lowest energy, $\Delta\varepsilon_j$ being approximated as ja (a is the energy of a *gauche* form relative to the *trans* form)
σ=symmetry number=$\sigma_e \cdot \sigma_i / I$
σ_e=symmetry number for external rotation
σ_i=symmetry number for internal rotation
I=number of racemic isomers (2 for racemic mixture, 1 for otherwise)
n=number of carbon atoms
n_Y=number of tertiary carbon atoms
n_X=number of quarternary carbon atoms.

The derivation of Eq. 25 will be given in Appendix II.

The numerical values of C_0, C_1, C_2, and C_3 for S° at 298.15 K given in Table 1 are obtained from the paper by Person and Pimentel[16] and the book by Janz,[17] both of whom refined the Pitzer's original values.[25] Then Eq. 25 is expressed in the unit of cal·K^{-1}·mol^{-1} as

$$S^\circ_{\text{thm}} = 43.012 + 7.679n - R \ln \sigma$$
$$- 1.068\, n_Y - 3.425\, n_X + R(\ln Q + Q'/Q). \qquad (29)$$

The data for calculating the values of σ, Q, and Q'

TABLE 3. STERIC AND SYMMETRY FACTORS OF SMALL ALKANES

Alkane[a]	Type[b]	g_j[c]						σ[c]			$R(\ln Q+Q'/Q)$[d]
		$j=0$	1	2	3	4	∞	σ_e	σ_i	I	
4	N	1	2				0	2	1	1	1.745
2m3	A	1					0	3	1	1	0
5	N	1	4	2			2	2	1	1	3.153
2m4	A	2	1				0	1	1	1	1.927
22m3	L	1					0	12	1	1	0
6	N	1	6	8	2		10	2	1	1	4.610
3m5	B			3	4		2	1	1	1	3.462
2m5	A		2	3			4	1	1	1	2.781
23m4	H			1	2		0	2	1	1	1.745
22m4	L			3			0	1	3	1	2.183
7	N	1	8	18	12	2	40	2	1	1	6.060
3e5	D				2	9	16	3	1	1	4.358
3m6	B			3	7	2	15	1	1	2	4.307
2m6	A		2	7	4		14	1	1	1	4.340
23m5	H				1	4	4	1	1	2	2.779
33m5	M				7		2	2	1	1	3.867
24m5	E			2			7	2	1	1	1.377
22m5	L			3			6	1	3	1	2.183
223m4	R					3	0	1	3	1	2.183

a) See Table 1. b) See Table 2. c) See Text for notations. d) $R=1.9872\ \mathrm{cal\cdot K^{-1}\cdot mol^{-1}}$, $T=298.15\ \mathrm{K}$, and $a=500\ \mathrm{cal\cdot mol^{-1}}$.

are given in Table 3, where the g_j numbers have been obtained by the method of Balaban with 4-digit codes representing the direction of the C–C bond in the diamond lattice.[26,27]

Although Eq. 29 gives fairly good results (See Table 1) for the lower members of saturated hydrocarbons, it is a formidable task to get the g_j numbers for larger molecules. Then Pitzer and Scott have proposed the empirical formula

$$S^\circ_{\mathrm{emp}} = 39.00 + 9.13n - R\ln\sigma - 3.5B \qquad (30)$$

where the value B is the number of branches.[18]

If one compares Eqs. 29 and 30 by noting that $B=n_Y+2n_X$, it is found that Eq. 30 does not properly take into account the effect of branching and also the rotational partition function. The effect of the site of the branching as discussed in Section I is completely neglected in Eq. 30.

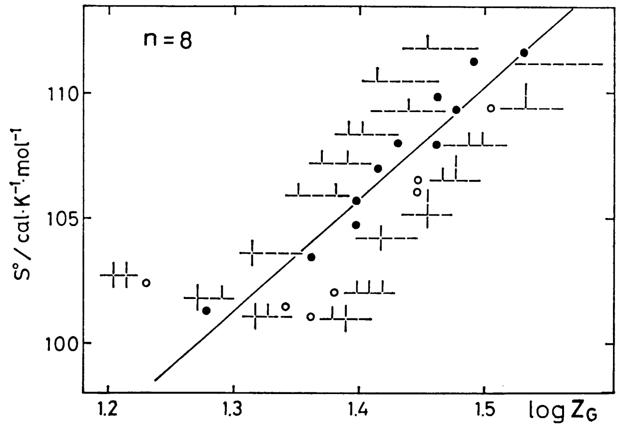


Fig. 6. log Z_G—S° plot for octane isomers.

The three different sets of calculated entropies, Eqs. 2, 29, and 30 are compared with the observed values in Table 1. As judged from the standard deviations (s) and the correlation coefficients (r) for the three different sets of calculations, the thermodynamic formula 29 gives the best and the topological formula 2 the worst agreement with experiments. Deviation of the plot of Z_G—S°_{top} from linearity in Fig. 4 is seen mostly for molecules with high symmetry.

In Figs. 6 and 7 the log Z_G—S°_{top} plots are also given for hydrocarbons with $n=8$ and 9, respectively,

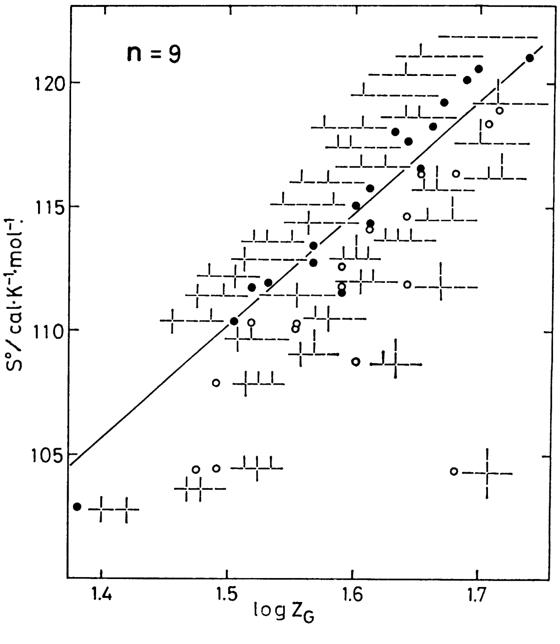
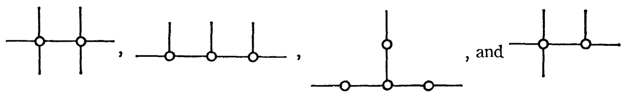


Fig. 7. log Z_G—S° plot for nonane isomers,

where those hydrocarbons containing skeletons



are marked with open circles, which tend to be scattered beneath the correlation line given in Fig. 5. Note that most of the rest of the molecules marked with filled circles are still on the straight line except for several melecules. Then as long as we are concerned with the fundamental analysis of the effect of branching and extension of the chain unit on the entropy, the topological index is found to be fairly effective, whereas the Pitzer's theory fails to be applied to larger molecules owing to the infeasibility of the enumeration of the steric factors.

III. Analysis of the Topological Dependence of the Entropy

As is found above, sterically overcrowded and/or highly symmetrical acyclic hydrocarbons are beyond the limit of the estimation of the absolute entropy by the use of only the topological index. If the symmetry factor σ is taken into consideration by adding the correction term to Eq. 2 for the set of data in Table 1, one gets

$$S^{\circ}_{\mathrm{top}} = 46.453 + 42.100 \log Z_{G} - 2.9630 \log \sigma \tag{31}$$

where the coefficients were obtained by the least squares calculation. The values of $s=0.744$ and $r=0.9970$ show fair improvement from Eq. 2. In both of Eqs. 29 and 30, the contribution of the symmetry factor σ is given by $-R\ln\sigma=-4.576\log\sigma$, which is 50 per cent larger than the corresponding term in Eq. 31. This means that the higher the symmetry of the molecule the smaller the topological index as the result of branching, but that this tendency of Z_G is not large enough for explaining the symmetry dependence of entropy completely.

With the above discussion in mind let us make term-to-term comparison of Eqs. 29 and 31 for larger alkanes. According to Relation 15, n can be substituted by

$$n = (\log Z_G - \log a)/\log \alpha. \tag{32}$$

Equation 25 or 29 is then transformed as

$$\begin{aligned} S^{\circ}_{\mathrm{thm}} = {} & 43.012 + 36.744 \log Z_G - 4.5757 \log \sigma \\ & - (36.744 \log a + 1.068\, n_{Y} + 3.425\, n_{X}) \\ & + 1.9872\, (\ln Q + Q'/Q). \end{aligned} \tag{33}$$

There are observed small discrepancy between the corresponding coefficients of the first three terms of S°_{thm} and S°_{top}. Let us examine the relation between these discrepancies and the last two terms in S°_{thm}, which are missing from the expression of S°_{top}. As seen from Table 2 the fourth term is found to be constant as

$$-(36.744 \log a + 1.068\, n_{Y} + 3.425 n_{X}) = 6.9 \pm 1.4 \tag{34}$$

throughout all the series studied. On the other hand, the last term seems to increase with n or $\log Z_G$ but with a few exceptions at least for small members of alkanes (See Table 3). Person and Pimentel[16] have noticed the linear increase of this term with respect to n for n-alkanes. One may then expect a rough estimate of this term as

$$R(\ln Q + Q'/Q) = 6.5 \log Z_G - 3.9 \pm 1.0 \tag{35}$$

for larger alkanes. Now substitution of these two relations into Eq. 33 yields

$$S^{\circ}_{\mathrm{thm}} = 46.0 + 43.2 \log Z_G - 4.6 \log \sigma \pm 2.4 \tag{36}$$

which is turned out to be very close to the round-off experssion of Eq. 31 as

$$S^{\circ}_{\mathrm{top}} = 46.5 + 42.1 \log Z_G - 3.0 \log \sigma\,. \tag{37}$$

Large deviation from Eq. 34 is not expected, since the components in Eq. 34 are related with each other as in Eqs. 17 and 18. Although at the present stage we cannot draw any quantitative conclusion on Relation 35, this is qualitatively approved by the following discussion. The branching of the carbon atom skeleton diminishes the rotational degree of freedom mainly and additional decrement in the partition function is brought about by the increase of the overcrowded conformations. The above analysis shows that both the effects are reflected well in (the logarithm of) the topological index; the latter effect contributing by the amount of as large as fifteen per cent. The symmetry factor for a molecule with higher symmetry lowers further the entropy as much as a few per cent for each of additional symmetry operations.

One may well say that the above argument is based on superficial comparison of figures between experiments and models. However, the parallelism between the mathematical properties of S° and Z_G as discussed in Sections I and II ensures that these relations obtained in this paper can be used as the start for further analysis of the thermodynamic properties of van der Waals fluids.

Appendix I

Proof of Eq. 21

Consider four series of graphs, N, X, Y, and XY, as shown in Fig. A. Graph N_k is a path progression with k $(=l+m+n)$ points. Graph X_k, which is a member of the series X with k points, is composed of the root Ⓧ with l points and a path progression N_{m+n} $(m+n>0)$. Graph Y_k is composed of the root Ⓨ with m points and a path progression N_{l+m} $(l+m>0)$. The roots Ⓧ and Ⓨ characterize graphs X_k and Y_k, respectively, by their modes of branching. Graph XY_k is composed of the roots Ⓧ and Ⓨ and a path progression

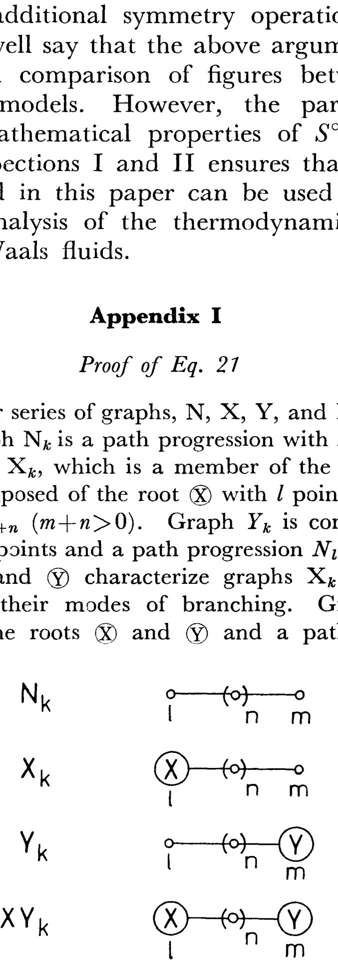


Fig. A. Four different series of graphs. The total number of points is $k=l+m+n$.

N_n ($n>0$).

If the topological indices of these graphs are denoted respectively as N_k, X_k, Y_k, and XY_k, these values can be expressed by using Eq. 8 as follows:

$$\begin{aligned}
\mathrm{N}_k &= \mathrm{N}_{l+n}\cdot\mathrm{N}_m + \mathrm{N}_{l+n-1}\cdot\mathrm{N}_{m+1} - \mathrm{N}_{l+n-1}\cdot\mathrm{N}_m \\
\mathrm{X}_k &= \mathrm{X}_{l+n}\cdot\mathrm{N}_m + \mathrm{X}_{l+n-1}\cdot\mathrm{N}_{m+1} - \mathrm{X}_{l+n-1}\cdot\mathrm{N}_m \\
\mathrm{Y}_k &= \mathrm{N}_{l+n}\cdot\mathrm{Y}_m + \mathrm{N}_{l+n-1}\cdot\mathrm{Y}_{m+1} - \mathrm{N}_{l+n-1}\cdot\mathrm{Y}_m \\
\mathrm{XY}_k &= \mathrm{X}_{l+n}\cdot\mathrm{Y}_m + \mathrm{X}_{l+n-1}\cdot\mathrm{Y}_{m+1} - \mathrm{X}_{l+n-1}\cdot\mathrm{Y}_m.
\end{aligned} \tag{A1}$$

Note that these relations are valid for the case with $n>0$, since graphs X_k and Y_k are defined, respectively, for the cases with $k\geq l$ and $k\geq m$. By combining these equations one gets

$$\begin{aligned}
\mathrm{XY}_k\cdot\mathrm{N}_k - \mathrm{X}_k\cdot\mathrm{Y}_k = {} & (\mathrm{X}_{l+n}\cdot\mathrm{N}_{l+n-1}-\mathrm{X}_{l+n-1}\cdot\mathrm{N}_{l+n}) \\
& \times (\mathrm{Y}_m\cdot\mathrm{N}_{m+1}-\mathrm{Y}_{m+1}\cdot\mathrm{N}_m).
\end{aligned} \tag{A2}$$

Division of the both sides of Eq. A2 by $\mathrm{N}_k{}^2$ gives

$$\begin{aligned}
\frac{\mathrm{XY}_k}{\mathrm{N}_k} - \frac{\mathrm{X}_k}{\mathrm{N}_k}\cdot\frac{\mathrm{Y}_k}{\mathrm{N}_k} &= \frac{\mathrm{X}_{l+n}\cdot\mathrm{Y}_m\cdot\mathrm{N}_{l+n}\cdot\mathrm{N}_m}{\mathrm{N}_k{}^2} \\
&\quad \times\left(\frac{\mathrm{N}_{l+n-1}}{\mathrm{N}_{l+n}}-\frac{\mathrm{X}_{l+n-1}}{\mathrm{X}_{l+n}}\right)\left(\frac{\mathrm{N}_{m+1}}{\mathrm{N}_m}-\frac{\mathrm{Y}_{m+1}}{\mathrm{Y}_m}\right) \\
&< \frac{\mathrm{X}_{l+n}}{\mathrm{N}_{l+n}}\cdot\frac{\mathrm{Y}_m}{\mathrm{N}_m}\left(\frac{\mathrm{N}_{l+n-1}}{\mathrm{N}_{l+n}}-\frac{\mathrm{X}_{l+n-1}}{\mathrm{X}_{l+n}}\right) \\
&\quad \times\left(\frac{\mathrm{N}_{m+1}}{\mathrm{N}_m}-\frac{\mathrm{Y}_{m+1}}{\mathrm{Y}_m}\right).
\end{aligned} \tag{A3}$$

The last inequality is obtained by using the relation

$$\mathrm{N}_k > \mathrm{N}_{l+n}\cdot\mathrm{N}_m \qquad \text{(from Eq. 7).} \tag{A4}$$

Since the topological index of the path progression N_k is the largest among the isomeric graphs with the same number of points, we have

$$\frac{\mathrm{X}_{l+n}}{\mathrm{N}_{l+n}},\ \frac{\mathrm{Y}_m}{\mathrm{N}_m} < 1.$$

As k increases the remaining two terms in the parentheses individually approach to zero because of the asymptotic relation 14. Now Eq. A3 sharply converges to zero with increasing the size of the graph, or

$$\frac{\mathrm{XY}_k}{\mathrm{N}_k} \rightarrow \frac{\mathrm{X}_k}{\mathrm{N}_k}\cdot\frac{\mathrm{Y}_k}{\mathrm{N}_k}. \tag{A5}$$

With respect to the coefficient a defined in Eq. 11, a rigorous relation

$$\frac{a_{\mathrm{XY}}}{a_{\mathrm{N}}} = \frac{a_{\mathrm{X}}}{a_{\mathrm{N}}}\cdot\frac{a_{\mathrm{Y}}}{a_{\mathrm{N}}} \tag{A6}$$

holds, which completes the proof of Eq. 21.

Appendix II

Derivation of Eq. 25

According to Pitzer[15)] the absolute entropy of an alkane is expressed as

$$S^\circ = f_0 + \sum_{i=1}^{7} f_i n_i + \Delta \tag{A7}$$

with

n_1=number of C–C stretching modes
n_2=number of C–C–C bending modes
n_3=number of internal rotations (C–C–C–C)
n_4=number of $\mathrm{CH_3}$ groups
n_5=number of $-\mathrm{CH_2}-$ groups
n_6=number of $-\overset{|}{\mathrm{C}}\mathrm{H}-$ groups$=n_{\mathrm{Y}}$
n_7=number of $-\underset{|}{\overset{|}{\mathrm{C}}}-$ atoms$=n_{\mathrm{X}}$.

Δ is defined as Eq. 26. Person and Pimentel[16)] give the coefficients f_i ($i=0\sim7$) for various temperatures. The values for $T=298.15$ K are:

$$\begin{aligned}
&f_0=57.473,\ f_1=0.792,\ f_2=2.173,\\
&f_3=4.523,\ f_4=2.314,\ f_5=0.191,\\
&f_6=-0.65,\ \text{and}\ f_7=-2.78
\end{aligned}$$

in units of $\mathrm{cal\cdot K^{-1}\cdot mol^{-1}}$.

A simple combinatorial consideration on tree graphs gives the following relations among $n_1\sim n_7$ and n (the number of the carbon atoms),

$$\begin{aligned}
n_1 &= n - 1 \\
n_2 &= n - 2 + n_6 + 2n_7 \\
n_3 &= n - 3 - n_6 - 2n_7 \\
n_4 &= 2 + n_6 + 2n_7 \\
n_5 &= n - 2 - 2n_6 - 3n_7.
\end{aligned} \tag{A8}$$

By putting Relations A8 into Eq. A7 one gets Eq. 25 with

$$\begin{aligned}
C_0&=f_0-f_1-2f_2-3f_3+2f_4-2f_5=43.012\\
C_1&=f_1+f_2+f_3+f_5=7.679\\
C_2&=f_2-f_3+f_4-2f_5+f_6=-1.068\\
C_3&=2f_2-2f_3+2f_4-3f_5+f_7=-3.425.
\end{aligned}$$

References

1) A. T. Balaban, Ed., "Chemical Application of Graph Theory," Academic Press, London (1976).

2) A. Graovac, I. Gutman, and N. Trinajstić, "Topological Approach to the Chemistry of Conjugated Molecules (Lecture Notes in Chemistry, No. 4)," Springer-Verlag, Berlin (1977).

3) H. Hosoya, *Kagaku No Ryoiki*, **26**, 989 (1972).

4) L. Kihr and L. H. Hall, "Molecular Connectivity in Chemistry and Drug Research," Academic Press, New York (1976).

5) H. Hosoya, K. Hosoi, and I. Gutman, *Theor. Chim. Acta*, **38**, 37 (1975); H, Hosoya and K. Hosoi, *J. Chem. Phys.*, **64**, 1065 (1976).

6) I. Gutman, M. Milan, and N. Trinajstić, *J. Am. Chem. Soc.*, **99**, 1692 (1977); I. Gutman and H. Hosoya, *Theor. Chim. Acta*, **48**, 279 (1978).

7) J. Aihara, *J. Am. Chem. Soc.*, **98**, 2750 (1978); **99**, 2048 (1977).

8) H. Hosoya, *J. Chem. Doc.*, **12**, 181 (1972).

9) M. Randić, *J. Chem. Phys.*, **60**, 3920 (1974); *J. Am. Chem. Soc.*, **97**, 6609 (1975).

10) H. Wiener, *J. Am. Chem. Soc.*, **69**, 17, 2636 (1947); *J. Chem. Phys.*, **15**, 766 (1947); *J. Phys. Chem.*, **52**, 425 (1948).

11) J. R. Platt, *J. Phys. Chem.*, **56**, 328 (1952).

12) H. Hosoya, K. Kawasaki, and K. Mizutani, *Bull. Chem. Soc. Jpn.*, **45**, 3415 (1972).

13) H. Hosoya, *Bull. Chem. Soc. Jpn.*, **44**, 2332 (1971).

14) H. Hosoya and Y. Maruyama, "Kōzō to Bussei," Asakura, Tokyo (1975).

15) K. S. Pitzer, *J. Chem. Phys.*, **8**, 711 (1940).

16) W. B. Person and G. C. Pimental, *J. Am. Chem. Soc.*, **75**, 532 (1953).

17) G. J. Janz, "Thermodynamic Properties of Organic Compounds—Estimation Methods, Principles and Practice," Academic Press, New York (1967).

18) K. S. Pitzer and D. W. Scott, *J. Am. Chem. Soc.*, **63**, 2419 (1941).

19) D. W. Scott, *J. Chem. Phys.*, **60**, 3144 (1974).

20) H. Hosoya, *Fibonacci Quart.*, **11**, 255 (1973).

21) Observed values of the absolute entropy of alkanes were taken from Ref. 22.

22) "Selected Values of Properties of Hydrocarbons and Related Compounds (API Report 44)," Thermodynamic Research Center, Texas A & M Univ. (1975).

23) N. N. Vorob'ev, "Fibonacci Numbers," Nauka Publ., Moskow (1964).

24) V. E. Hoggatt, Jr., "Fibonacci and Lucas Numbers," Houghton Mifflin Co., Boston (1969).

25) For several cases Pitzer gave wrong numbers for g_n. Although Person and Pimentel adjusted the parameters C_0, C_1, and C_2 for smaller normal alkanes, we did not use this sophisticated correction. Discrepancy of the $S°$ values from those of the previous papers might have arisen from the above alternatives.

26) A. T. Balaban, *Rev. Roum. Chim.*, **21**, 1049 (1976).

27) A number of conflicting theories have been proposed on the conformations of alkane skeletons. For example, Sheraga(28) claims that more than three potential minima exist around a C-C axis for larger alkanes, Scott(19) proposes a sophisticated method for estimating the steric repulsion energy in branched alkanes, Allinger(29) proposes another criterion for determining the conformations, *etc.*(30,31) However, according to some trial calculations of the steric factors for the various models, the topological dependence of the steric factor as a whole is not strongly dependent on the choice of these models.

28) R. A. Scott and H. A. Scheraga, *J. Chem. Phys.*, **44**, 3054 (1966).

29) N. L. Allinger, M. T. Tribble, M. A. Miller, and D. H. Wertz, *J. Am. Chem. Soc.*, **93**, 1637 (1971).

30) A. Abe, R. L. Jernigan, and P. J. Flory, *J. Am. Chem. Soc.*, **88**, 631 (1966).

31) A. L. Verma, W. F. Murphy, and H. J. Bernstein, *J. Chem. Phys.*, **60**, 1540 (1974).