

Chemical Meaning of Octane Number Analyzed by Topological Indices*

Haruo Hosoya

*Department of Information Sciences, Ochanomizu University,
Bunkyo-ku, Tokyo 112–8610, Japan
(E-mail: hosoya@is.ocha.ac.jp)*

Received June 4, 2001; revised August 20, 2001; accepted August 21, 2001

Correlation of the octane number (ON) of heptane and octane isomers with various topological indices, Z , w , p , and B , was studied. For single parameter correlation the index B by Balaban and w by Wiener show good results, whereas for two-parameter correlation almost any combination among the above four indices was found to give relatively high r value. Further, by adding the third parameter the correlation is a little improved. One of the best results of regression analysis is: $\text{ON} = 76.389 + 8.179(0.3B + 1.1p - 0.6Z)$ with $r = 0.982$ and $S = 6.56$ for octane isomers. From these analyses one can conclude that in order for a gasoline isomer to have high ON, or to be burnt without knocking, it should be highly branched (small Z or low boiling point), its statistically averaged shape should be spherical (large B), and it should have a compact structure (large p or high liquid density). The weight of the contributions of these factors decreases in this order. From these results several candidates of nonane isomers whose ON is expected to be higher than 100 are suggested.

Key words: gasoline, octane number, Wiener number, centric index, topological index, QSPR, anti-knocking, boiling point.

INTRODUCTION

Anti-knock ability of gasoline of a variety of hydrocarbons and their mixture is rated in terms of »octane number (ON)«, which is determined by measuring the combustion efficiency compared with that of the prescribed

* Dedicated to Professor Milan Randić on the occasion of his 70th birthday.

binary mixture of isooctane (2,2,4-trimethylpentane, ON = 100) and *n*-heptane (ON = 0) under the standard test conditions.¹ Balaban and Motoc performed extensive QSPR analyses on the alkanes and cycloalkanes together with alkenes by using a number of topological indices.^{2,3} If the correlation of a single parameter with ON is concerned, the centric index *B* proposed by Balaban⁴ is the best with a relatively large correlation coefficient *r* and small standard error, while the Wiener's *w*⁵ is shown to have a fairly large – *r* value. Although Hosoya's topological index, *Z*,⁶ has a relatively large negative *r* value, the magnitude of standard error is rather large. On the other hand, the connectivity index, *X*,⁷ of Randić has a small *r* value in this case. Randić tried to correlate ON with other kinds of molecular descriptors and found that his graphical bond order *P*²/*P* gives almost the same result as *w*.⁸ However, further enlargement of the number of linearly combined parameters (up to three) for QSPR correlation by Balaban *et al.* was unrewarding.⁹ Gutman *et al.* defined the π index which is obtained by multiplying the distances between all pairs of vertices, and got a little better correlation coefficient for ON.¹⁰ Very recently the research group of Trinajstić tried to extend the idea of Balaban's average distance-sum connectivity index, *J*,¹¹ and obtained the highest correlation coefficient around 0.99.¹²

In spite of these attempts to improve the *r* value for ON correlation there has not been suggested any physico-chemical discussion on the mechanism of gasoline combustion or on the meaning of ON in the molecular level except for the following empirical rules.²

That is, i) ON increases with the number of tertiary and quaternary carbon atoms, while ii) it decreases with the total number of carbon atoms. Further, iii) as the branching point is moving toward the center of the longest chain in the alkane molecule, ON gets larger, as reflected by the fairly high *r* value for the correlation between ON and the centric index *B*.

Since ON is not a quantity rigorously defined in thermodynamics, one cannot »measure« its accurate value like boiling point (bp) and liquid density (*d*). Our aim here is not to search the best descriptor or topological index giving the highest *r* value for the correlation with the available ON's, but rather to induce physico-chemical interpretation or thermodynamic meaning of this useful index, ON. In this paper in order to find a crucial key for clarifying the chemical meaning of ON we will be concerned only with alkanes, or saturated acyclic hydrocarbon molecules.

OCTANE NUMBER

As has been explained in detail by Balaban *et al.*,^{2,9} two different kinds of ON scales, *i.e.*, research ON (RON) and motor ON (MON), are available.

Since these two scales are roughly parallel, the results of the former scale will mainly be shown, and supplemented by the latter, if necessary, in this paper. However, the ON value of 0.00 in both the scales for *n*-octane is found to be outlying in several different plots. Then in this paper we will choose the newly revised value of -19.00 for *n*-octane as shown in Table I,^{10,13} where the values of boiling point and liquid density are also given for the relevant discussion.¹⁴

TOPOLOGICAL INDICES

Since our purpose is to clarify the physico-chemical features of ON, only the selected number of topological indices (TIs) were chosen, *i.e.*, *Z*, *w*, *p*, and *B*. The values of these TIs for heptane and octane isomers are given in Table I. The *Z* index of Hosoya is known to reflect the degree of rotational freedom around the CC bonds of alkane molecules.¹⁵⁻¹⁸ The good correlation of *Z* with boiling point and entropy can be explained by using this property. Wiener's path number *w*⁵ and Balaban's centric index *B*⁴ have already been shown to be correlated well with ON.^{2,9} Contrary to the popularity of *w*¹⁹⁻²¹ the polarity number *p* has not so frequently been discussed. However, it is closely related to the density and refractivity of liquid, and its importance and physico-chemical meaning have gradually been getting recognized.¹⁵⁻¹⁸

On the other hand, topological features of *w* are known to change with the size of alkane molecules. Namely, for smaller alkanes (*N* ≤ 7) the *r* value between *w* and *Z* is greater than the absolute value of *r* between *w* and *p*,

TABLE IA
Various topological indices and octane numbers of heptane isomers

Isomer	<i>Z</i>	<i>w</i>	<i>p</i>	<i>B</i>	RON ^a	MON ^a	(bp/°C) ^b	<i>d</i> ^b
<i>n</i> -Hept	21	56	4	13	0.0	0.0	98.4	0.678
2-Me-Hex	18	52	4	17	42.4	46.4	90.0	0.674
3-Me-Hex	19	50	5	17	52.0	55.0	91.9	0.683
3-Et-Pent	20	48	6	19	65.0	69.3	93.4	0.694
2,2-Me-Pent	14	46	4	21	92.8	95.6	79.2	0.670
2,3-Me-Pent	17	46	6	21	91.1	88.5	89.7	0.691
2,4-Me-Pent	15	48	4	21	83.1	83.8	80.5	0.668
3,3-Me-Pent	16	44	6	21	80.8	86.6	86.0	0.699
2,2,3-Me-But	13	42	6	29	112.1	101.3	80.9	0.686

^a Ref. 9.

^b Ref. 14.

TABLE IB
 Various topological indices and octane numbers of octane isomers

Isomer	Z	w	p	B	RON ^a	MON ^a	(bp/°C) ^b	d^b
<i>n</i> -Oct	34	84	5	16	-19.0	-19.0	125.7	0.699
2-Me-Hept	29	79	5	18	21.7	23.8	117.7	0.694
3-Me-Hept	31	76	6	18	26.8	35.0	119.0	0.702
4-Me-Hept	30	75	6	18	26.7	39.0	117.7	0.701
3-Et-Hex	32	72	7	22	33.5	52.4	118.6	0.710
2,2-Me-Hex	23	71	5	24	72.5	77.4	106.9	0.691
2,3-Me-Hex	27	70	7	24	71.3	78.9	115.6	0.708
2,4-Me-Hex	26	71	6	24	65.2	69.9	109.5	0.696
2,5-Me-Hex	25	74	5	24	55.5	55.7	109.1	0.689
3,3-Me-Hex	25	67	7	24	75.5	83.4	112.0	0.706
3,4-Me-Hex	29	68	8	24	76.3	81.7	117.7	0.715
2-Me-3-Et-Pent	28	67	8	26	87.3	88.1	115.7	0.715
3-Me-3-Et-Pent	28	64	9	26	80.8	88.7	118.3	0.723
2,2,3-Me-Pent	22	63	8	30	109.6	99.9	109.9	0.712
2,2,4-Me-Pent	19	66	5	30	100.0	100.0	99.3	0.688
2,3,3-Me-Pent	23	62	9	30	106.1	99.4	114.8	0.722
2,3,4-Me-Pent	24	65	8	30	102.7	95.9	113.5	0.715
2,2,3,3-Me-But	17	58	9	40	-	-	106.3	-

^a Ref. 9.

^b Ref. 14.

but this order is reversed and enlarged with the size of molecules.¹⁸ This means that the topological character of w changes from Z -type to p -type with the increase of the size of molecules. In this sense the physico-chemical meaning of w is rather elusive.

In order to realize the important role of B in this paper, let us restate its definition.⁴ First count the number of endpoints, or the vertices of degree one, v_1 , of the original graph and prune them out. Then repeat this pruning process until the graph becomes vacant. The quantity B is defined to be the sum of the squares of all the v_i 's. From this definition B seems to reflect the magnitude of the surface area of an alkane molecule.

Before looking into the correlation pattern between these topological indices and ON we need to know the correlation among these indices, which is summarized in Table II. Remarkably good correlation (but with a negative r) is found between w and B , while Z is correlated modestly well with w and B . On the other hand, the most orthogonal pair of descriptors is Z and p as

has already been pointed out.¹⁵⁻¹⁸ The second next orthogonal pair is p and B . These informations are helpful for following the discussion below.

TABLE IIA

The r values for the correlation among various topological indices of heptane isomers

	Z	w	p	B
Z	1.000			
w	0.796	1.000		
p	-0.089	-0.648	1.000	
B	-0.898	-0.896	0.447	1.000

TABLE IIB

The r values for the correlation among various topological indices of octane isomers

	Z	w	p	B
Z	1.000			
w	0.761	1.000		
p	-0.257	-0.787	1.000	
B	-0.875	-0.915	0.625	1.000

CORRELATION OF ON WITH A SINGLE TOPOLOGICAL INDEX

The correlation coefficients r 's for the correlation of ON with various topological indices, Z , w , p , and B were calculated as shown in Table III. As has already been pointed out by Balaban *et al.*^{2,9} both B and w show the best correlation with ON. Note, however, that the sign of the r value of w is minus as is the case with Z . Although p shows the most inferior result as a single parameter correlation, it will later be shown that p has an auxiliary but rather important role in the correlation problem of gasoline combustion.

Compare n -alkane and a highly branched and spherical isomer of onion-like shape. It is easily inferred from its definition that the value of B of the latter is much greater than the former, as shown in Table I. In other words B may more or less reflect the dynamical surface area of an alkane molecule. One can infer that the larger the effective surface of alkane molecules the more smoothly their explosive oxidation may take place.

TABLE III

The r values and standard error (S) for the correlation with the octane number of heptane and octane isomers

	Heptane		Octane		
	r	S	r	S	
Single parameter					
Z	-0.860	16.19	Z	-0.828	19.42
w	-0.956	9.27	w	-0.957	9.53
p	0.487	27.67	p	0.594	27.81
B	0.917	12.62	B	0.960	9.64
Two-parameter					
$w+0.7Z$	-0.970	7.77	$w+0.7Z$	-0.984	6.15
$p-0.9Z$	0.950	9.93	$p-0.6Z$	0.978	7.24
$B-0.7Z$	0.928	11.77	$B-0.15Z$	0.962	9.44
$w+0.9p$	-0.972	7.43	$w+p$	-0.975	12.14
$w-0.4B$	-0.962	9.09	$B-0.7w$	0.979	7.10
$B+0.1p$	0.918	12.61	$B+0.3p$	0.963	9.27
Three-parameter					
$0.1B+p-0.7Z$	0.951	7.56	$0.3B+1.1p-0.6Z$	0.982	6.56
$w+0.9p-0.2B$	-0.974	7.21	$w+p-0.7B$	-0.983	6.33
$w+0.6Z-0.1B$	-0.970	7.70	$w+0.7Z-0.8B$	-0.986	5.36

Although the p index, which is the number of 3-step pairs of vertices, also seems to increase with branching, its behavior is a little more delicate. Very recently¹⁸ the present author found that within a group of alkane isomers the value of p is parallel to the number of gauche^{22,23} or staggered structures in the statistically most popular conformation. Then one can explain why p is an extremely good indicator of liquid density and refractive index,¹⁸ but not that of ON.

CORRELATION OF ON WITH BOILING POINT

From chemical reasoning the first stage of combustion of gasoline might be its vaporization followed by the oxidation of the outermost atoms of the molecules. In order to get some clue for clarifying this mechanism the r values between ON and bp were calculated for heptane and octane isomers to be 0.828 and 0.632, respectively, which are much worse than expected (see

Figure 1a and b). Although 2,3-dimethylhexane, 2-methyl-3-ethylpentane, and 2,3,3-trimethylpentane boil almost at the same temperature, 115 ± 0.6 °C,¹⁴ their ON values widely range from 73.4 to 108.6. This ordering was found to be the same as that of liquid density, d . In both groups of heptane and octane isomers we can find several pairs and trios whose bp and d behave similarly. Since d of alkane isomers is correlated fairly well with the

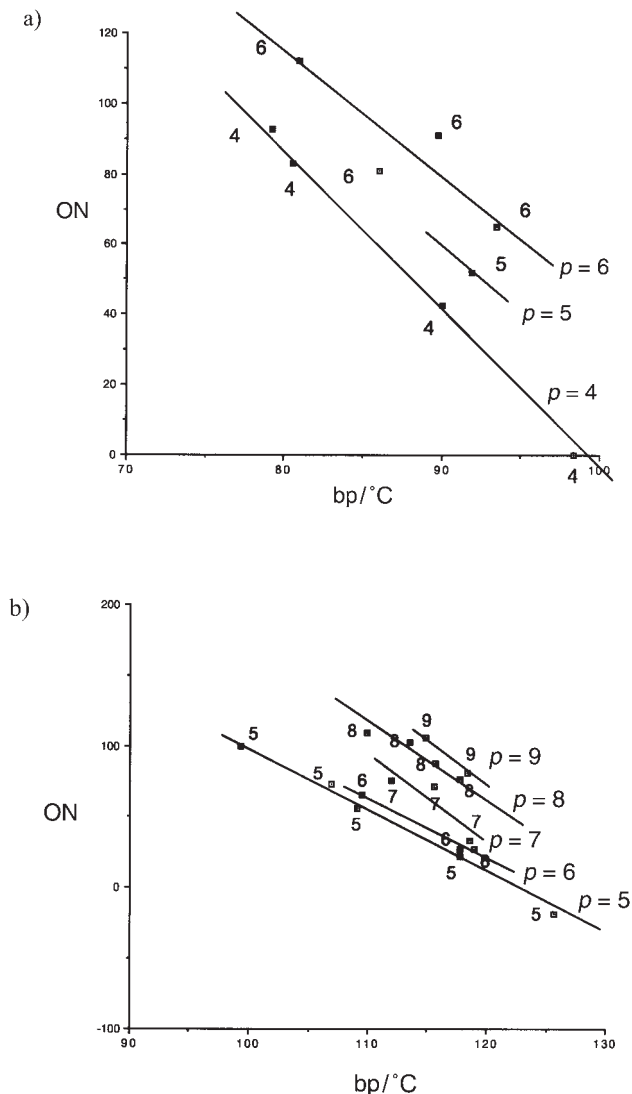


Figure 1. Correlation between the boiling point (°C) and octane number of (a) heptane and (b) octane isomers.

Wiener's polarity number, p ,¹⁵⁻¹⁸ the above discussion on the relation between bp and d is also valid between bp and p .

Then the values of p were written down near the corresponding points in the plot of bp -ON as shown in Figure 1a and b. In both the cases most of the points with the same p were lined up nicely. If similar modification is worked out with bp on Figure 1a and b, one can obtain contour maps of a rather smooth slope, suggesting that ON of alkanes is linearly dependent on bp and d .

At this stage one can summarize these results as follows. i) If isomers of almost the same bp are compared, the one with larger p number (or of higher liquid density) has a larger ON, and ii) if isomers of the same p number (or of similar liquid density) are compared, the one of lower bp has a larger ON.

TWO-PARAMETER CORRELATION OF ON

Since the Z index is known to be correlated fairly well with bp ,^{6,15-18} the $Z(p)$ -ON plots were similarly drawn for heptane and octane isomers separately as shown in Figure 2a and b, from which more clear QSPR features came out. Namely, among each of heptane and octane isomers all the points with the same p value are more neatly lined up as if they are radiating from the common origin. The only exception is the case with $p = 8$ and 9 of octane isomers, whose points are loosely scattered along a single line. Note that both in Figures 1b and 2b the points for n -octane seem to be located at the right position, supporting the revised ON value of -19.00 . Also the value of the motor ON for n -octane is suggested to be around -20 .

Then by changing the relative weights of Z and p the best linear coupling of these two TI's was searched without taking into account the difference in slopes in Figure 2. The obtained regression equations are as follows:

$$\begin{aligned} \text{ON} &= 188.55 + 11.63(p - 0.9Z) \\ r &= 0.950 \quad S = 9.93 \quad n = 9 \quad \text{for heptanes} \end{aligned} \quad (1)$$

$$\begin{aligned} \text{ON} &= 177.66 + 12.12(p - 0.6Z) \\ r &= 0.978 \quad S = 7.24 \quad n = 18 \quad \text{for octanes} \end{aligned} \quad (2)$$

Notice the dramatic improvement of r and S values from the single parameter values of the component TI's.

Similarly the best combination of two-parameter correlation with ON for each of all the possible pairs among the four TI's was searched to give the

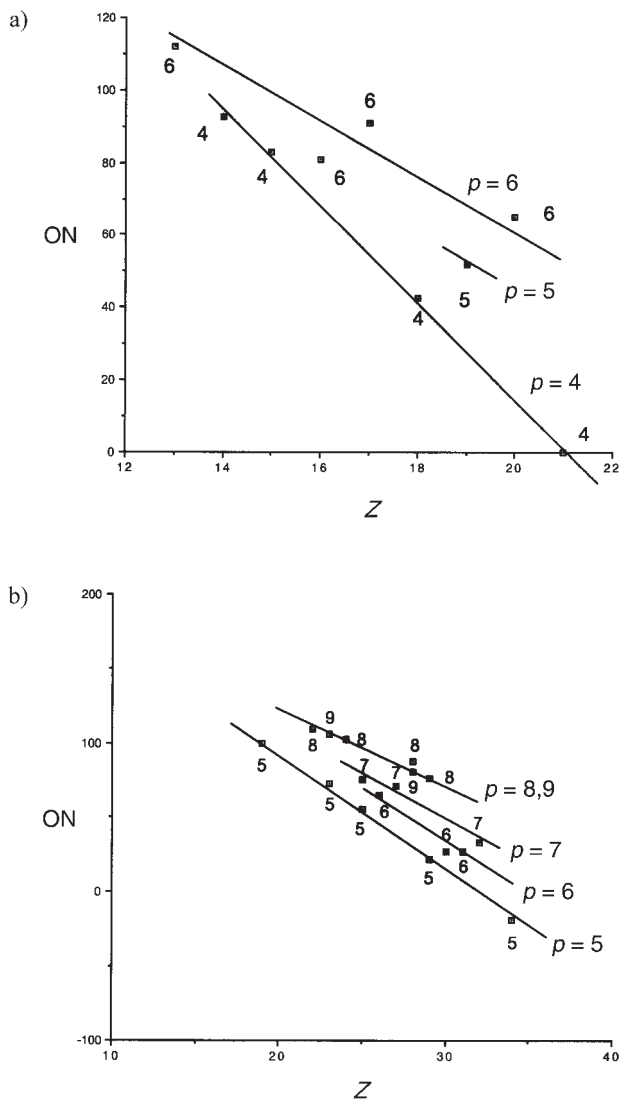


Figure 2. Correlation between the Z index and octane number of (a) heptane and (b) octane isomers.

largest r value. The results are summarized in Table III. Note that the absolute values of r 's for all the possible pairs chosen from the four TI's are as high as 0.96~0.98. Namely, as long as the statistical parameters are concerned, any combination of these four TI's can give more or less similar results. Although the value of r for the pair of Z and p is a little smaller than those for the two pairs, w and Z , and w and B , we preferred to focus our at-

tention on the pair of nearly orthogonal descriptors, Z and p , from the following observation.

The correlation plots like Figures 1 and 2 were drawn for all the two-parameter pairs to find other complementary pair as Z and p . However, we could not find any such pair of TIs suggesting some physico-chemical features of the combustion of gasoline.

THREE-PARAMETER CORRELATION OF ON

For further elaboration one of the rest TIs, B index, was added into the regression calculation. The obtained results are as follows:

$$\begin{aligned} \text{ON} &= 127.69 + 11.99 (0.1B + p - 0.7Z) \\ r &= 0.951 \quad S = 9.76 \quad n = 9 \quad \text{for heptanes} \end{aligned} \quad (3)$$

$$\begin{aligned} \text{ON} &= 76.39 + 8.179 (0.3B + 1.1p - 0.6Z) \\ r &= 0.982 \quad S = 6.56 \quad n = 18 \quad \text{for octanes} \end{aligned} \quad (4)$$

The weights of the contribution from the three TIs to the ON value were found to be in the descending order of Z , B , and p , by putting the TI values for these isomers into equations (3) and (4). The increment of the r value caused by the addition of B is not so large in this case, and other combination of TIs, such as the trio of w , Z , and B , gives a little higher correlation as seen in Table III. However, the relative signs of the coefficients of Z , B , and p are all consistent with those for the two-parameter pairs, contrary to the case with the other two combinations of three-parameter correlation. By taking into consideration of the relative signs and magnitudes of the contributions of Z , B , and p one can derive the following conclusion to the determining factors of ON.

In order for a gasoline isomer to have a high ON, or to be burnt without knocking, i) it should be highly branched (small Z or low bp), ii) its statistically averaged shape should be spherical (large B), and iii) it should have a compact structure (large p or high d). The weight of the contribution from these factors decreases in this order.

Although the above conclusion might have been inferred from chemical reasoning without sophisticated calculation, the present study can afford its quantitative verification. It is also to be mentioned here that the results obtained in this study are an outcome of careful scrutinization of the regression analysis by using a selected number of physico-chemically meaningful TIs.

PREDICTION

Among the eighteen octane isomers 2,2,3,3-tetramethylbutane (**1**) is estimated to have the largest ON = 135 from the extrapolation in Figure 2b (but not from equation (4)). This substance is known to have very peculiar properties called plastic crystal.²⁴ Since this molecule has such a high symmetry and fits into the stable crystal structure that the melting point is as high as +101 °C, just 6 °C below the boiling point!¹⁴ Thus this substance cannot be used as an ordinary fuel.

2,2,3,3-Tetramethylpentane, a highly branched isomer of nonane similar to **1**, is suggested to have a very high ON, and actually it was measured to be 116.8.⁹ Since there are only five available ON's for nonane isomers,⁹ it is difficult to get a reliable regression formula. However, from the regression formulas obtained for heptane and octane isomers one can roughly estimate the relative magnitude of ON of a nonane isomer by the following combination of TI's as $B+4p-2Z$. In Table IV are given several candidates of nonane isomers whose ON is nearly equal to or higher than 100, and compared with those which are already found to have higher ON.

TABLE IV
Predicted and observed nonane isomers with high ON

Isomer	<i>B</i>	<i>p</i>	<i>Z</i>	$B+4p-2Z$	RON ^a
2,2,3,3-Me-Pent	41	12	30	29	116.8
2,2,3,4-Me-Pent	41	12	31	27	
2,3,3,4-Me-Pent	41	12	33	23	
2,2,4,4-Me-Pent	41	6	24	17	
2,2-Me-3-Et-Pent	35	10	36	3	112.1
2,3-Me-3-Et-Pent	35	12	40	3	
2,3,3-Me-Hex	33	10	36	1	
2,2,3-Me-Hex	33	9	35	-1	
3,3,4-Me-Hex	33	11	39	-1	
2,4-Me-3-Et-Pent	35	10	39	-3	105.3
2,4,4-Me-Hex	33	8	34	-3	

^a Ref. 9.

REFERENCES

1. Phillips Petroleum Company, *Phillips 66 Reference Data for Hydrocarbons and Petro-Sulfur Compounds*, Bulletin No. 521, 1962 (Revised 1974).
2. A. T. Balaban and I. Motoc, *MATCH (Commun. Math. Chem.)* **5** (1979) 197–218.
3. A. T. Balaban, *Pure Appl. Chem.* **55** (1983) 199–206.
4. A. T. Balaban, *Theor. Chim. Acta* **5** (1979) 239–261.
5. H. Wiener, *J. Am. Chem. Soc.* **69** (1947) 17–20.
6. H. Hosoya, *Bull. Chem. Soc. Jpn.* **44** (1971) 2332–2339.
7. M. Randić, *J. Am. Chem. Soc.* **97** (1975) 6609–6615.
8. M. Randić, *J. Math. Chem.* **7** (1991) 155–168.
9. A. T. Balaban, L. B. Kier, and N. Josh, *MATCH (Commun. Math. Chem.)* **28** (1992) 13–27.
10. I. Gutman, W. Linert, I. Lukovits, and Z. Tomović, *J. Chem. Inf. Comput. Sci.* **40** (2000) 113–116.
11. A. T. Balaban, *Chem. Phys. Lett.* **89** (1982) 399–404.
12. S. Nikolić, D. Plavšić, and N. Trinajstić, *MATCH (Commun. Math. Computer Chem.)* **44** (2001) 361–386.
13. F. Jürgen (Ed.), *CD Römpp-Chemielexikon*, 9th ed., Georg Thieme Verlag, New York, 1995, pp. 310.
14. American Petroleum Institute Research Project 44 at the National Bureau of Standards, *Selected Values of Physical and Thermodynamic Properties of Hydrocarbons*, Texas A & M Univ., 1945.
15. H. Narumi and H. Hosoya, *Bull. Chem. Soc. Jpn.* **58** (1985) 1778–1786.
16. Y.-D. Gao and H. Hosoya, *Bull. Chem. Soc. Jpn.* **61** (1988) 3093–3102.
17. H. Hosoya, M. Gotoh, M. Murakami, and S. Ikeda, *J. Chem. Inf. Comput. Sci.* **39** (1999) 192–196.
18. H. Hosoya and Y.-D. Gao, *Mathematical and Chemical Analysis of the Wiener's Polarity Number*, in: R. B. King and D. Rouvray (Eds.), *Topology in Chemistry*, Ellis Horwood, New York, 2001, in press.
19. L. B. Kier and L. H. Hall, *Molecular Connectivity in Chemistry and Drug Design*, Academic Press, New York, 1976.
20. N. Trinajstić, *Chemical Graph Theory*, 2nd revised ed., CRC Press, Boca Raton, FL, USA, 1992.
21. J. Devillers and A. T. Balaban, *Topological Indices and Related Descriptors in QSAR and QSPR*, Gordon & Breach, Amsterdam, 1999.
22. T. Simanouti and S. Mizushima, *Sci. Rept. Inst. Phys. Chem. Res.* **40** (1943) 467–471.
23. S. Mizushima, *Structure of Molecules and Internal Rotation*, Academic Press, New York, 1954.
24. J. G. Aston, in: D. Fox, M. M. Labes, and A. Weissberger (Eds.), *Physics and Chemistry of the Organic Solid State*, Vol.1, Interscience Publ., New York, 1963, pp. 543.

SAŽETAK**Kemijsko značenje oktanskog broja analiziranog s pomoću topologijskih indeksa***Haruo Hosoya*

Proučavana je korelacija oktanskog broja (ON) heptana i oktanskih izomera s različitim topologijskim indeksima, Z (Hosoyin indeks), w (Wienerov indeks), p (broj parova atoma udaljenih za tri veze) i B (Balabanov indeks). Za korelaciju s jednim parametrom Balabanov indeks, B , i Wienerov indeks, w , pokazali su dobre rezultate, dok za korelaciju s dva parametra gotovo svaka kombinacija navedenih četiriju indeksa daje relativno veliku vrijednost koeficijenta korelacije (r). Nadalje, dodavanjem trećeg parametra korelacija se neznatno poboljšava. Jedan od najboljih rezultata regresijske analize jest $ON = 76,389 + 8,179 (0,3 B + 1,1p - 0,6 Z)$ s $r = 0,982$ i $S = 6,56$ za oktanske izomere. Iz tih analiza može se zaključiti da bi benzinski izomer imao veliki ON, ili da izgara bez detonacije u motoru, treba biti jako razgranano (što znači da ima mali Z ili nisko vrelište), njegov statistički prosječan oblik treba biti sferičan (veliki B), i treba imati kompaktnu strukturu (veliki p , ili veliku gustoću tekućine). Težina doprinosa tih faktora opada navedenim redoslijedom. Na temelju tih rezultata preporuča se nekoliko izomera nonana za koje se očekuje ON veći od 100.