

On Topological Indices, Boiling Points, and Cycloalkanes

Gerta Rücker and Christoph Rücker*

Institut für Organische Chemie und Biochemie, Universität Freiburg, Albertstrasse 21,
D-79104 Freiburg, FRG

Received March 4, 1999

The experimental boiling points (bp) of saturated hydrocarbons (acyclic through polycyclic) up to decanes are systematically compiled. The bp values are classified into groups of lower or higher reliability according to the accuracy and frequency with which they were reproduced by independent researchers. For each hydrocarbon structure the values of several simple topological indices (TI) of widely varying origin are given, including the values of molecular walk counts of various lengths and their sum. The sensitivity of the TIs for structural changes within comprehensive groups of cyclic saturated hydrocarbons is evaluated, and the total walk count is found to be most sensitive. By multilinear regression structure–bp correlations are obtained for various comprehensive compound samples. Both the detour index and the walk counts are found to play a major role in the best models. Comparison of the bp models obtained with those from the recent literature reveals significant improvements for both cyclic and acyclic alkanes, which is attributed in part to the higher quality of experimental data, in part to the use of novel descriptors, and in part to the use of a more diverse pool of descriptors to select from. Despite this a descriptor combination allowing to accurately model cycloalkane bps is not yet found.

INTRODUCTION

The boiling point at normal pressure (bp) has become a benchmark property for the evaluation of descriptors for structure–property relationship studies.¹ Numerous models for the bp of saturated acyclic hydrocarbons (the alkanes) were established in terms of more or less simple graph invariants, the so-called topological indices (TIs), a few of which from the past decade are given in a later section (Table 7).² For (poly)cyclic saturated hydrocarbons, on the other hand, structure–bp studies, though highly desirable, are scarce. Such studies suffered from problems both with the data and with the TIs:

First, the diversity in the structures of (poly)cyclic saturated hydrocarbons is overwhelming (fused, bridged, spiro ring systems, bearing as substituents open-chain or again (poly)cyclic attachments), so as to render the include/exclude decision everything but clear-cut. Second, a small fraction of all possible (poly)cycloalkanes even of small size has been synthesized, and their bps, if obtained at all, were determined mostly during purification (distillation) and as such are of low precision. Third, the problem of stereochemistry, existent but neglected in most studies of acyclic alkanes, has to be dealt with in the case of cycloalkanes. Most substituted or polycyclic cycloalkanes exist as diastereomers differing in their bps, whereas experimental data are often from mixtures of unknown composition of stereoisomers. On the other hand, TIs by definition do not include stereochemical information. Fourth, it is sometimes not straightforward to extend the definition of a TI, first given for acyclics, to cyclic compounds. This is particularly so for a TI based on paths, since in (poly)cyclics there may exist several paths of equal length as well as different lengths between a pair of vertices.³

One result of Hosoya's pioneering study (as late as 1972⁴) was that (poly)cycloalkane bps exhibit linear dependence on

$\log(Z)$ and possibly on the number of rings, n_r . Even far later others included in their alkane-centered investigations some cycloalkanes, such as Lukovits⁵ and Zinn.^{6,7} Lukovits,⁸ Trinajstić,⁹ and ourselves¹⁰ studied the detour index, a TI recently introduced especially for cyclic compounds, and its correlation with bps for mixed samples of cyclic and acyclic compounds. Plavšić et al. devoted a study to the monocycloalkanes of up to six carbon atoms, using an index Z^* derived from the Z matrix.¹¹ Das et al. found for a sample of monocycloalkanes nonlinear correlations of the bps with the Wiener, Schultz, and Szeged indices and a high linear interrelation among these indices.¹² Diudea regressed the bps of some monocycloalkanes against indices derived from the Wiener, Szeged, and Cluj matrices as well as from the corresponding reciprocal matrices.¹³ Estrada regressed the bps of some mono- and bicyclic alkanes against descriptors derived from the edge adjacency matrix.¹⁴ Finally, Lukovits and Linert compared combinations of several Wiener-type indices for the bps of a mixed sample of some monocyclic and all acyclic octanes.^{3c}

In contrast to all the work done in constructing and comparing molecular descriptors for structure–bp correlations, far less attention has been paid to the data, that is, to the compounds included in such studies and to their bps. The samples of cyclic compounds used were ill-defined in most of the above studies, in that typically a particular compound was included if it happened to appear with its bp in one or another handbook, or it simply was selected without any comment given, resulting in no two of these studies dealing with the same sample. Moreover, several bp values used in the above studies are more or less incorrect, so that in certain cases the values for a compound included in two studies differed considerably.¹⁵ For these reasons the qualities of the above structure–bp relationships for cycloalkanes or

mixed alkane/cycloalkane samples cannot be compared.

We therefore undertook to compile a list as comprehensive as possible of all saturated hydrocarbons with normal bps known as reliably as possible up to a certain molecular size limit. Earlier we had used for the first time a comprehensive sample of acyclic and cyclic alkanes with known bps at 760 Torr up to size $n = 8$.¹⁰ In the present work first we extend that compilation up to $n = 10$ (including bps measured at ≥ 720 Torr, extracting the data from the most comprehensive base for this kind of data, the Beilstein database) and add the values for a few important TIs, in the hope to provide a list which may be useful to others in future studies. Second, we subject a few simple TIs of potential or demonstrated value for bp studies to a test for structural sensitivity within two comprehensive samples of cyclic alkanes. Third, we propose best (within our limited pool of descriptors) 1-descriptor, 2-descriptor, etc., multilinear models for the bp within several comprehensive samples of saturated hydrocarbons. We compare our results to those found in the recent literature. Fourth, we demonstrate how improper use of literature data and use of incomplete compound samples can result in structure–property correlations of deceptively high quality.

RESULTS AND DISCUSSION

Experimental Bps of Saturated Hydrocarbons up to $n = 10$. A data file of all saturated hydrocarbon structures up to $n = 10$ with known normal bps was compiled by the procedure detailed in the Experimental Section (531 structures). The file contains for each structure a shorthand name (composed in an obvious, semisystematic manner), the size (number of C atoms, n), the number of rings (n_r), for monocyclics the size of the ring (s_r), and what we think is the best available bp value, based on a literature survey only, that is, before undertaking any structure–bp correlations.

A classification of bps is made as follows: Bps are regarded as “reproduced” ($\text{rel} = 1$) if at least two independent research groups reported values differing by no more than 4 °C, otherwise a bp is “not reproduced” ($\text{rel} = 0$). As an exception, a bp value resulting from the careful measurements of Rossini (National Bureau of Standards) is regarded “reproduced” even if it is the only one available.¹⁶ This qualification enables us to exclude completely unchecked data from structure–bp correlations; it had proved useful in the earlier work.¹⁰ Since the quality even of the reproduced experimental bps may still be too low—after all, the bp values result from measurements done all over the world and during a time period of well over a century—we went a step further in this work. Bps which were reported by at least four independent research groups within a span of no more than 2 °C are classified as “known beyond reasonable doubt” ($\text{rel} = 2$).

Also included in the data file are the values of the following simple TIs: Wiener index (W), Hosoya index (Z ; calculated in this work in a novel manner, see Experimental Section), Randic index (χ), detour indices (ω and ω'),¹⁰ diameter (longest distance, dia), molecular walk counts of different lengths k (mwck), total walk count (twc),¹⁷ and path counts of length k including and excluding cyclic paths (pk and apk). These descriptors are defined for cyclic as well as acyclic graphs without any problems. Descriptors mwck and

twc were introduced by us earlier: They are conceptually extremely simple, extremely easily calculated by a Morgan type summation procedure, and highly discriminant within acyclic alkanes (first degeneracies of twc in three pairs of dodecanes). In (cyclo)alkanes twc is a good measure of structural complexity.¹⁷ Walk counts had scarcely been used in structure–bp studies when the present work was begun;¹⁸ for their fundamental nature we included them here for a test.

The data file is too voluminous to be printed here; it is available from the authors in SAS-readable electronic form. The 531 structures along with their respective shorthand names as an identifier are shown in Figure 1, where they appear in the order of increasing n , n_r , s_r (monocyclics only), and complexity as measured by twc .¹⁷ By this ordering it should be easy to verify a particular structure (or its absence) in Figure 1. Table 1 (an extract from the data file) gives, in the same order, the bp values, their reliability status ($\text{rel} = 0/1/2$), and the twc values from the file.

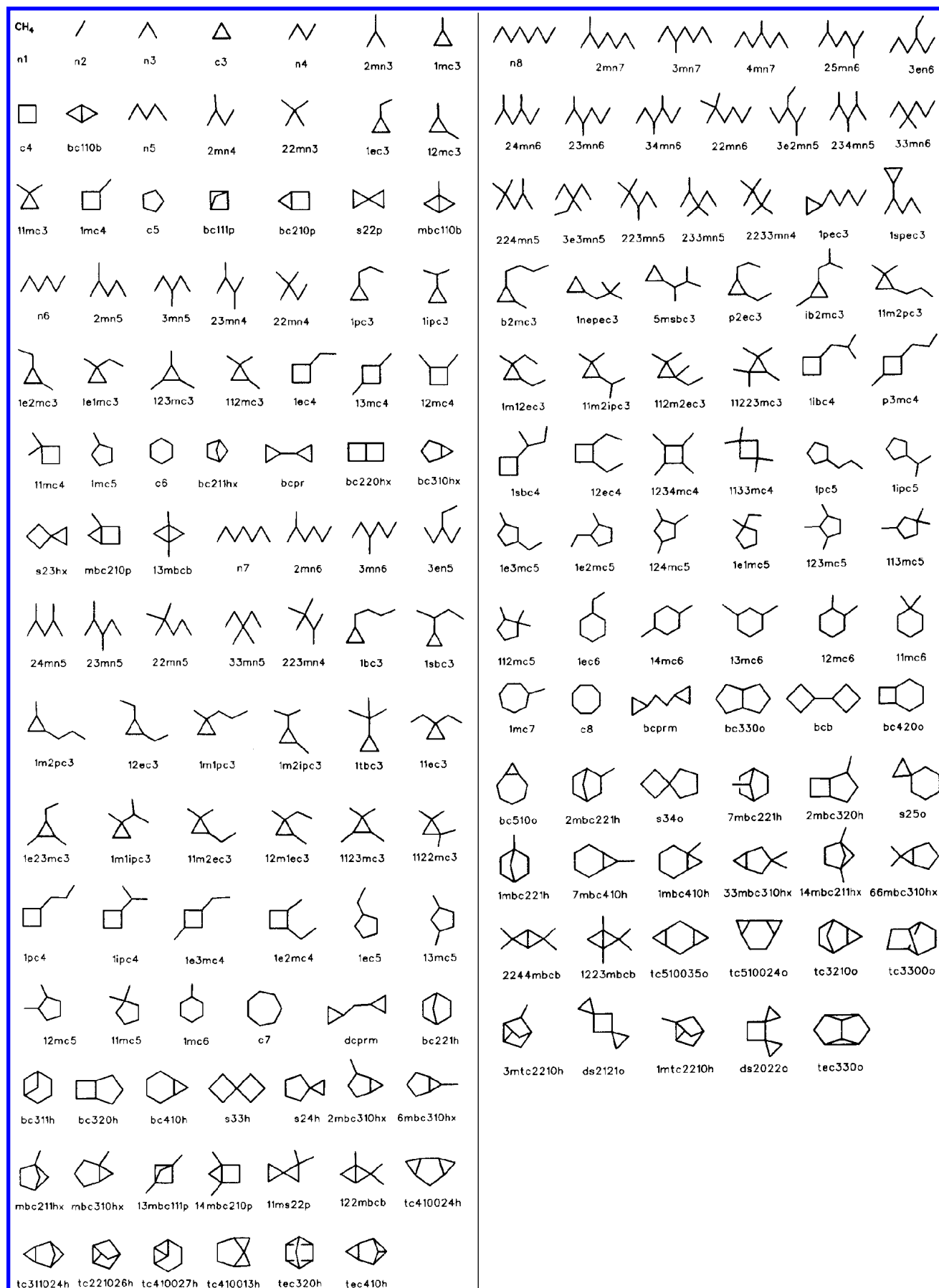
Table 2 gives an overview of the number of structures contained in Table 1/Figure 1 and of all possible (polycyclo)-alkane structures, ordered by size n and number of rings n_r . In each cell of Table 2 in the upper line the number of structures of a particular n and n_r in Table 1 is given and in the lower line the number of all (mathematically) possible cycloalkane structures (neglecting stereoisomerism) of that n and n_r , as generated by the program MOLGEN.¹⁹ Though MOLGEN for high n_r includes many structures which would be considered “impossible” for hydrocarbons by most chemists, at least for low n_r it gives the correct number of chemically reasonable structures. Comparison of the upper and lower lines in Table 2 demonstrates how small a fraction of all cycloalkanes is actually known with bps at normal pressure. Comparison of the upper line entries of Table 2 with the corresponding counts of structures included in refs 3c, 5–9, and 12–14 demonstrates that these studies were far from comprehensive within their respective sets of structures.

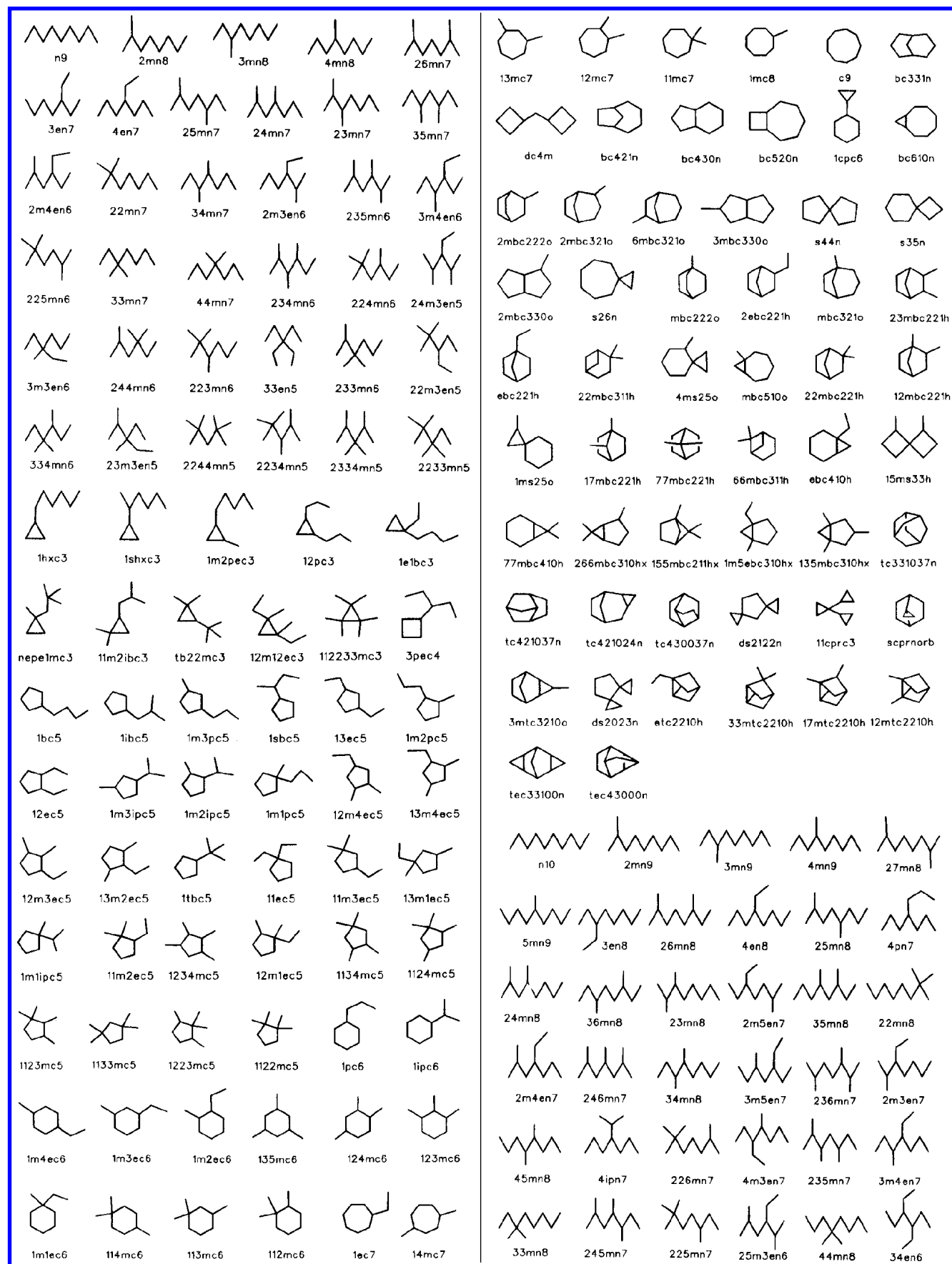
Sensitivity of TIs within Comprehensive Cycloalkane Samples. We decided to evaluate the sensitivity of some TIs for structure changes in cycloalkanes in two comprehensive samples thought to be typical.

The challenge in the first test sample should be in a large number of monocyclic isomers differing in ring size and thus in number, size, and structure of side chains. Table 3 (an expansion of one column of Table 2) gives the number of structures of monocycloalkanes from cyclopropane to the monocyclodecanes, ordered by size n and ring size s_r , where the meaning of the upper and lower entries is as in Table 2. It is seen from Table 3 that of the small-ring monocycloalkanes most (or at least most bps) are unknown. The monocyclic decanes (475 structures) were chosen for the test (sample 1).

The second test sample was chosen to challenge the TIs due to a large variety of polycyclic systems within a narrow size range. It consists of all mathematically possible saturated hydrocarbon structures including (poly)cycloalkanes of $n = 4$ –6, 105 structures according to Table 2 (sample 2).

All structures in samples 1 and 2 were generated by MOLGEN, and those not already present in the data file (i.e., 402 monocyclodecanes and 64 polycyclic butanes through hexanes) were additionally included as a new section in the





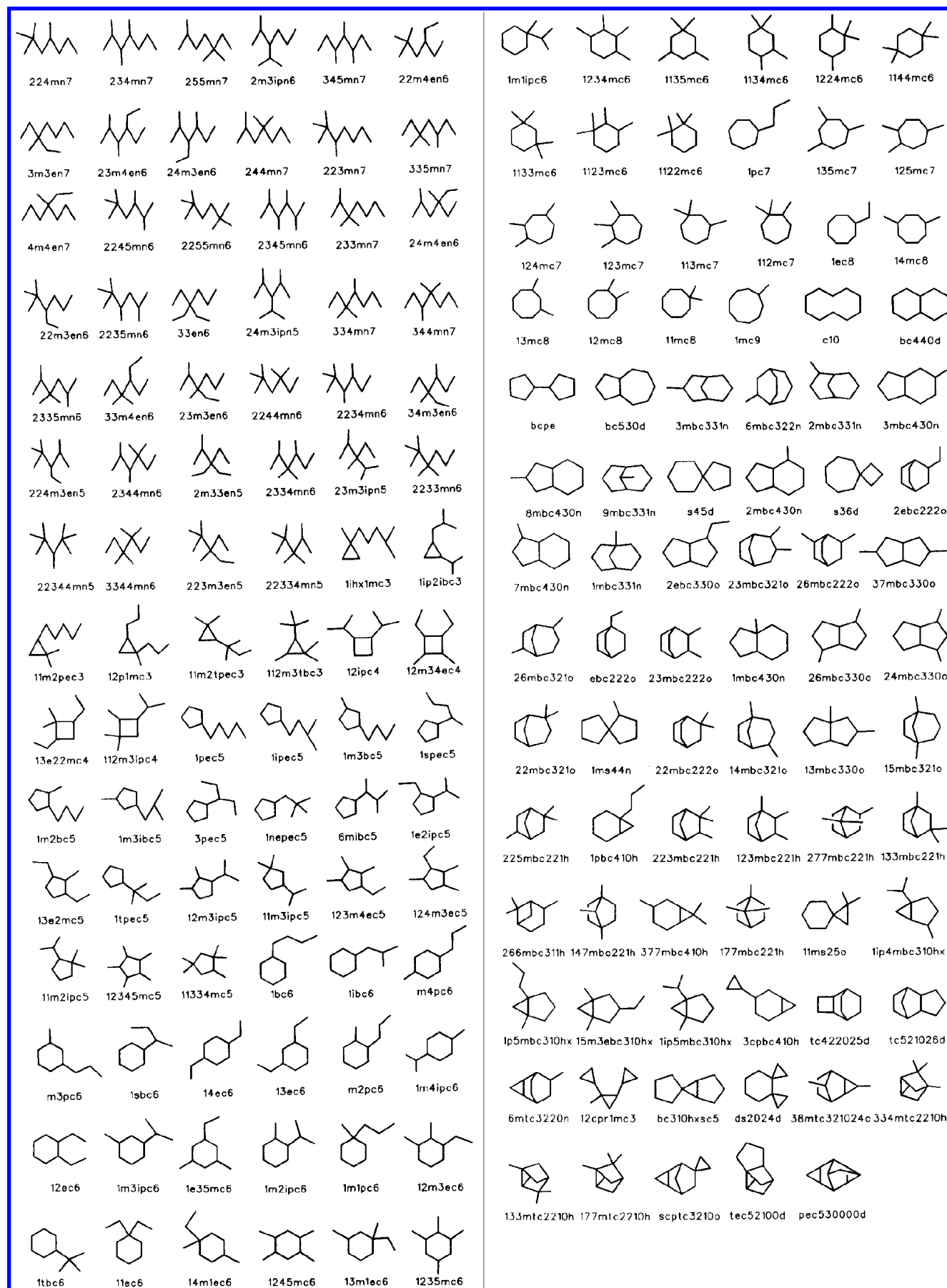


Figure 1. Saturated hydrocarbon structures ($n = 1-10$) for which bps are available, with their respective shorthand names. The structures appear in the order of increasing n , n_r , s_r (monocyclics only), and two-letter codes for bicyclic compounds.

name	bp (°C)	rel	twc	name	bp (°C)	rel	twc	name	bp (°C)	rel	twc	name	bp (°C)	rel	twc
n1	−161.5	2	0	1mc6	101	2	547	bc420o	133	1	2593	1m1pc5	145	0	5317
n2	−88.6	2	1	c7	118.4	2	441	bc510o	141	1	2692	12m4ec5	142.5	1	5335
n3	−42.1	2	5	dcprm	102	1	987	2mbc221h	125	1	2962	13m4ec5	138	1	5399
c3	−32.8	1	9	bc221h	105.5	2	987	s34o	128	1	3043	12m3ec5	147	0	5565
				bc311h	110	0	1003	7mbc221h	128	0	3068	13m2ec5	151	0	5629
n4	−0.5	2	16	bc320h	110.5	1	1037	2mbc320h	130.5	0	3122	1tbc5	145	1	5630
2mn3	−11.7	2	18	bc410h	116	2	1063	s25o	125	1	3123	1lec5	151	2	5775
1mc3	0.7	1	32	s33h	96.5	0	1204	1mbc221h	117	1	3280	11m3ec5	133	1	5887
c4	12.6	2	28	s24h	98.5	2	1216	7mbc410h	138	1	3452	13m1ec5	135.5	1	6139
bc110b	8	0	51	2mbc310hx	100	0	1253	1mbc410h	125	2	3772	1m1ipc5	143	1	6230
				6mbc310hx	103	1	1320	33mbc310hx	115	1	3864	11m2ec5	138	1	6413
n5	36.0	2	44	mbc211hx	81.5	0	1327	14mbc211hx	91	0	4414	1234mc5	135.9	1	6424
2mn4	27.8	2	53	mbc310hx	92	2	1433	66mbc310hx	126.1	0	4656	12m1ec5	142.5	1	6591
22mn3	9.5	2	70	13mbc111p	71.5	0	1767	2244mbcb	104	0	7613	1134mc5	127	1	6825
1ec3	35.9	2	93	14mbc210p	74	0	1894	1223mbcb	105	1	8886	1124mc5	129.5	1	7047
12mc3	32.6	2	107	11ms22p	78	0	1913	tc510035o	142	1	4670	1123mc5	134	1	7228
11mc3	20.6	2	116	122mbcb	84	1	2330	tc510024o	149	0	4938	1133mc5	118.2	2	7409
1mc4	36.3	2	89	tc410024h	105	1	1867	tc3210o	136	2	4943	1223mc5	138	1	7450
c5	49.3	2	75	tc311024h	107	0	1868	tc3300o	125	1	5162	1122mc5	135	1	8309
bc111p	36	0	147	tc221026h	106	1	1890	3mtc2210h	120.5	0	6006	1pc6	156.7	2	3396
bc210p	46	2	150	tc410027h	110	1	2113	ds2121o	103	0	6014	lipc6	154.8	2	4020
s22p	39	2	166	tc410013h	107.5	0	2234	1mtc2210h	111	1	6679	1m4ec6	150.8	2	4020
mbc110b	33.5	0	188	tec320h	108.5	1	3131	ds2022o	115	0	6682	1m3ec6	150	2	4092
				tec410h	104	0	3138	tec330o	137.5	0	8457	1m2ec6	154.3	2	4325
n6	68.7	2	111									135mc6	139.5	1	4680
2mn5	60.3	2	134	n8	125.7	2	627	n9	150.8	2	1439	124mc6	144.8	1	4779
3mn5	63.3	2	142	2mn7	117.6	2	764	2mn8	142.8	2	1763	123mc6	149.4	1	5016
23mn4	58.0	2	165	3mn7	118.9	2	838	3mn8	144	2	1942	1m1ec6	152	2	5037
22mn4	49.7	2	185	4mn7	117.7	2	856	4mn8	142.4	2	2019	114mc6	136	1	5247
1pc3	69	2	245	25mn6	109.1	2	911	26mn7	134	2	2098	113mc6	136.6		

Table 1. (Continued)

name	bp (°C)	rel	twc	name	bp (°C)	rel	twc	name	bp (°C)	rel	twc	name	bp (°C)	rel	twc
scprnorb	140.7	0	15500	223mn7	158	1	9133	124m3ec5	171	0	17268	8mbc430n	173.5	0	18097
3mtc3210o	161	0	15959	335mn7	155.7	1	9140	11m2ipc5	163	0	18379	9mbc331n	189.9	0	18197
ds2023n	146	0	18029	4m4en7	166	1	9149	12345mc5	164	0	20290	s45d	185	2	18493
etc2210h	137.5	0	20122	2245mn6	147.9	1	9509	11334mc5	141.5	1	22887	2mbc430n	182	0	18663
33mtc2210h	137.5	1	20325	2255mn6	137	1	9624	1bc6	180.9	1	7954	s36d	183	0	18761
17mtc2210h	131	0	21804	2345mn6	158	1	9713	1ibc6	171.3	2	8978	2ebc222o	183	0	18792
12mtc2210h	128	0	24329	233mn7	160	1	9869	m4pc6	173.4	1	9725	7mbc430n	182	1	18850
tec33100n	168.3	0	21609	24m4en6	158	1	10014	m3pc6	169	1	9882	1mbc331n	178	1	19598
tec43000n	153	1	23294	22m3en6	159	1	10215	1sbc6	179.3	1	10240	2ebc330o	174	1	20560
				2235mn6	148.7	1	10249	14ec6	175.5	1	10240	23mbc321o	174.5	0	21174
n10	174.1	2	3250	33en6	166.3	1	10263	13ec6	172	1	10458	26mbc222o	172.5	0	21356
2mn9	166.8	2	3948	24m3ipn5	157	0	10305	m2pc6	174.5	1	10611	37mbc330o	166	0	21403
3mn9	167.8	2	4394	334mn7	164	1	10544	1m4ipc6	170	2	11322	26mbc321o	164.5	0	21452
4mn9	165.7	2	4597	344mn7	164	1	10712	12ec6	176	1	11338	ebc222o	179.5	0	21791
27mn8	160	2	4658	2335mn6	153	1	11117	1m3ipc6	167	1	11553	23mbc222o	171.5	1	21971
5mn9	165.1	2	4665	33m4en6	165	1	11351	1e35mc6	168.5	2	11900	1mbc430n	175.8	1	22230
3en8	168	1	5026	23m3en6	169	1	11412	1m2ipc6	171	1	12417	26mbc330o	160.5	0	23101
26mn8	158.5	2	5118	2244mn6	153	1	11524	1m1pc6	174.3	2	12554	24mbc330o	165	0	23393
4en8	164	1	5285	2234mn6	155	1	11603	12m3ec6	175	0	13122	22mbc321o	172.5	0	23524
25mn8	157	2	5393	34m3en6	170	1	11890	1tbc6	171.5	2	13333	1ms44n	186.2	0	23611
4pn7	161.8	1	5475	224m3en5	155.3	1	12053	11ec6	179.5	2	13826	22mbc222o	174.5	0	23633
24mn8	153	1	5557	2344mn6	162	1	12305	14m1ec6	168	1	14126	14mbc321o	164	0	23947
36mn8	159.5	2	5649	2m33en5	174	1	12678	1245mc6	167	1	14344	13mbc330o	160.5	0	26214
23mn8	164	2	5808	2334mn6	164	1	13053	13m1ec6	166.6	1	14405	15mbc321o	159.5	0	27046
2m5en7	159.7	1	5835	23m3ipn5	169.4	1	13889	1235mc6	166.5	1	14471	225mbc221h	162	1	29493
35mn8	159	1	6010	2233mn6	160	1	14580	1m1ipc6	177.5	0	15021	1pbc410h	172.5	0	29677
22mn8	155	1	6195	22344mn5	159.3	1	14829	1234mc6	172.5	1	15031	223mbc221h	165.5	1	31361
2m4en7	160	1	6252	3344mn6	170.5	1	15816	1135mc6	153	1	15328	123mbc221h	159.5	0	31780
246mn7	145	1	6465	223m3en5	168	1	15958	1134mc6	160.3	0	15644	277mbc221h	163	1	32172
34mn8	166	1	6496	22334mn5	166	0	17193	1224mc6	158	0	16562	133mbc221h	150	2	32785
3m5en7	160	1	6514	1ihx1mc3	155.1	0	14800	1144mc6	153	1	16712	266mbc311h	168	1	32851
236mn7	155.7	1	6621	1ip2ibc3	148.3	0	17674	1133mc6	155	1	17188	147mbc221h	153	0	36023
2m3en7	166	1	6671	11m2pec3	157.8	0	18862	1123mc6	167	0	17256	377mbc410h	168.5	1	36067
45mn8	162.4	1	6693	12p1mc3	153.5	0	22225	1122mc6	161.5	0	20476	177mbc221h	160	2	37026
4ipn7	159.5	1	6924	11m2tpec3	146.7	0	29301	1pc7	182.8	1	7793	11ms25o	152	0	37138
226mn7	148.2	1	6937	112m3tbc3	146	1	35781	135mc7	164.2	0	10569	1ip4mbc310hx	157.5	2	39776
4m3en7	167	1	7286	12ipc4	158	1	17040	125mc7	173	0	10909	1p5mbc310hx	158.5	0	42873
235mn7	159.7	1	7326	12m34ec4	155.5	0	21097	124mc7	170.7	0	11080	15m3ebc310hx	159.5	0	44140
3m4en7	167	1	7343	13e22mc4	154	0	23393	123mc7	177.5	0	11791	1ip5mbc310hx	152	0	47509
33mn8	161.2	1	7387	112m3ipc4	145.5	1	23812	113mc7	161	0	12496	3cpbc410h	175.5	0	29041
245mn7	157	1	7492	1pec5	180	2	8030	112mc7	174	0	13803	tc422025d	219	0	30968
225mn7	147	1	7569	1ipe5	172	1	8871	1ec8	191.4	0	7431	tc521026d	188	1	31680
25m3en6	157	1	7729	1m3bc5	170.7	0	10232	14mc8	183.5	0	8323	6mtc3220n	189.5	0	38037
44mn8	160	1	7832	1spec5	176.5	1	10667	13mc8	181.5	1	8492	12cpr1mc3	158.3	0	45624
34en6	162	1	7933	1m2bc5	172	1	10892	12mc8	185.5	1	8997	bc310hxxc5	192.7	0	47196
224mn7	147.7	1	8073	1m3ibc5	153	0	11303	11mc8	170.5	0	10175	ds2024d	160	0	48556
234mn7	163	1	8150	3pec5	175.7	0	11390	1mc9	193.6	0	6637	38mtc321024o	152.5	0	50296
255mn7	152.8	1	8301	1nepec5	173.9	0	11883	c10	202	1	5110	334mtc2210h	151	1	70170
2m3ipn6	163	1	8533	6mibc5	173.5	0	12294	bc440d	191.5	1	14943	133mtc2210h	143.5	2	73139
345mn7	164	1	8650	1e2ipc5	160	0	13903	bcpe	190	2	14943	177mtc2210h	153	2	75358
22m4en6	147	1	8667	13e2mc5	172	1	14745	bc530d	193	1	15024	scptc3210o	174	0	62817
3m3en7	163.8	1	8794	1tpec5	173.9	0	15167	3mbc331n	182	1	16487	tec52100d	155	0	76738
23m4en6	164	1	8821	12m3ipc5	160.5	1	15928	6mbc322n	190	0	17001	pec530000d	171	0	95615
24m3en6	164	1	8967	11m3ipc5	148.5	0	16422	2mbc331n	187	0	17126				
244mn7	152	1	8973	123m4ec5	172	0	17113	3mbc430n	178	0	17607				

^a Numbers in italics are duplicate twc values.

data file with their TIs. The results of the sensitivity search are given in Table 4, where the distinction between upper and lower lines is as before (compounds with known bps and all possible structures, respectively). For the TIs W , Z , χ , ω , ω' , ($W\omega$), and twc the number of distinct values and the sensitivity (number of distinct values divided by number of distinct structures) are shown. As seen from Table 4, lower lines, W , Z , χ , ω , and ω' are not very discriminant TIs, ($W\omega$) is intermediate, while twc despite its simplicity is very sensitive to structure variations in both samples. In fact there were found only 14 pairs of monocyclodecanes with identical twc in sample 1, and four pairs of cyclic hexanes with identical twc in sample 2 (to be compared with 15 doublets and 1 triplet found earlier for a complexity measure in sample 2²⁰). For comparison, Diudea had constructed a putatively comprehensive sample of 439 monocyclodecanes and found

for W a sensitivity of 0.116, for Szeged and Cluj indices sensitivities between 0.223 and 0.483, and for the more complex corresponding Harary indices values between 0.355 and 1.000 in that sample.^{13a}

Interestingly, in the subsamples of compounds with known bps (Table 4, upper lines), all sensitivity values are considerably higher than in the complete samples (lower lines). Obviously neither the known monocyclodecanes nor the known cyclic $n = 4-6$ saturated hydrocarbons are representative for the respective complete populations, no surprise in the light of Tables 2 and 3. The ranking of TIs by sensitivity agrees for the two full samples; in the subsamples slight differences are seen, but twc ranks first in all cases.

As a further comparison consider the known sensitivity values in the sample of all 75 acyclic decanes: W 0.53, Z 0.53, χ 0.65, Balaban's J 0.99, twc 1.00.^{17,21} Obviously

Table 2. Number of Saturated Hydrocarbon Structures, Ordered by Increasing Size n and Number of Rings n_r ^a

	n_r												
n	0	1	2	3	4	5	6	7	8	9	10	11	Σ
1	1												1
	1												1
2	1												1
	1												1
3	1	1											2
	1	1											2
4	2	2	1	—									5
	2	2	1	1									6
5	3	5	4	—	—	—	—						12
	3	5	5	4	2	1	1						21
6	5	12	7	—	—	—	—	—					24
	5	12	17	18	14	8	3	1					78
7	9	24	15	5	2	—	—	—	—				55
	9	29	56	79	79	59	31	9	2				353
8	18	34	18	8	1	—	—	—	—	—			79
	18	73	182	326	430	427	298	134	35	6			1929
9	35	58	36	13	2	—	—	—	—	—	—		144
	35	185	573	1278	2161	2768	2616	1714	707	154	16		12207
10	75	73	46	11	2	1	—	—	—	—	—	—	208
	75	475	1792	4875	10162	16461	20346	18436	11477	4399	845	59	89402
Σ	150	209	127	37	7	1	—	—	—	—	—	—	531
	150	782	2626	6581	12848	19724	23295	20294	12221	4559	861	59	104000

^a Upper rows, structures with known bps; lower rows, all mathematically possible structures.**Table 3.** Number of Monocyclic Alkane Structures, Ordered by Increasing Size n and Increasing Ring Size s_r ^a

n	s_r								Σ
	3	4	5	6	7	8	9	10	
3	1								1
	1								1
4	1	1							2
	1	1							2
5	3	1	1						5
	3	1	1						5
6	6	4	1	1					12
	6	4	1	1					12
7	14	4	4	1	1				24
	15	8	4	1	1				29
8	12	6	9	5	1	1			34
	33	24	9	5	1	1			73
9	10	1	28	12	5	1	1		58
	83	55	28	12	5	1	1		185
10	6	4	19	30	7	5	1	1	73
	196	147	71	40	13	6	1	1	475
Σ	53	21	62	49	14	7	2	1	209
	338	240	114	59	20	8	2	1	782

^a The meaning of upper and lower rows is as in Table 2.

complete discrimination between isomers is far more difficult to achieve in cyclic than in acyclic alkanes.

Some Structure—Bp Correlations. For each sample of compounds of constant size n considered, the following descriptors were included in a pool from which to select descriptors for multilinear regressions: W , Z , $\log(Z)$ (abbreviated lZ), χ , ω , $mwc2$ — $mwc6$, twc , $\log(twc)$ (abbreviated $ltwc$), $p3$ — $p5$, $ap2$ — $ap5$, dia , $W\omega$, $(W\omega)^{0.125}$. For samples of variable n additionally n , n^2 , and $n^{0.5}$, for monocyclics s_r , and for samples of variable number of rings n_r were included. For acyclic structures ω is identical to W ; $(W\omega)$ therefore reduces to W^2 , and $(W\omega)^{0.125}$ reduces to $W^{0.25}$, a quantity

Table 4. Test Results for the Structure Sensitivity of TIs within Two Comprehensive Samples of Cycloalkanes^a

	W	Z	χ	ω	ω'	$W\omega$	twc
Sample 1							
distinct	30	45	43	51	52	71	72
values	50	62	136	102	132	321	461
sensitivity	0.411	0.616	0.589	0.699	0.712	0.973	0.986
	0.105	0.131	0.286	0.215	0.278	0.676	0.971
Sample 2							
distinct	20	18	34	28	36	39	40
values	27	37	79	49	62	83	101
sensitivity	0.488	0.439	0.829	0.683	0.878	0.951	0.976
	0.257	0.352	0.752	0.467	0.590	0.790	0.932

^a The meaning of upper and lower rows is as in Table 2.

found earlier to be of some value.²² This selection of TIs was guided by the criteria proven usefulness, availability, and simplicity as an indication of fundamentality.

Using the SAS procedures $rsquare$ and reg we obtained the few best one-variable, two-variable, etc., (multi)linear regressions for each sample. Comprehensive samples considered were the acyclic, monocyclic, and bicyclic heptanes, octanes, nonanes, decanes, combinations $n = 7$ –10, and combinations $n = 2$ –10 (all such compounds, those with bp values of $rel > 0$ only, and those with values of $rel = 2$ only) and the heptanes, octanes, nonanes, decanes, and combinations containing any number of rings, including zero. Structure—bp correlations for such a diverse set of saturated hydrocarbons were never obtained before. Some of the best correlations are given in Table 5 in the format descriptors, r^2 (squared coefficient of correlation), s (standard deviation).

From Table 5 the following conclusions can be drawn: (1) In the samples of high-quality bps, the s values for an increasing number of descriptors level off to a lower plateau than in the corresponding samples containing low-quality or

Table 5. Best Multilinear Structure–Bp Correlations Including 1, 2, 3, etc., Descriptors, in the Format Descriptor(s), r^2 , s

A. Acyclics									
all			rel > 0 only			rel = 2 only			
Octanes $n = 8$ $N = 18$			Octanes $n = 8$ $N = 18$			Octanes $n = 8$ $N = 18$			
Z	0.7868	2.91	(see left)			(see left)			
W p3	0.9785	0.96							
1twc $W^{0.25}$ p3	0.9842	0.85							
Nonanes $n = 9$ $N = 35$			Nonanes $n = 9$ $N = 35$			Nonanes $n = 9$ $N = 15$			
Z	0.5849	3.90	(see left)			Z	0.9442	1.98	
mwc4 mwc6	0.9414	1.49				mwc4 mwc6	0.9734	1.42	
Z p4 p5	0.9587	1.27				W mwc4 twc	0.9908	0.88	
mwc5 twc p3 $W^{0.25}$	0.9727	1.05							
W^2 mwc3 mwc4 mwc6 twc	0.9810	0.89							
Decanes $n = 10$ $N = 75$			Decanes $n = 10$ $N = 73$			Decanes $n = 10$ $N = 10$			
Z	0.3360	5.75	1Z	0.3855	5.57	p2	0.8677	2.01	
W p3	0.8588	2.67	W p3	0.8583	2.70	p3 $W^{0.25}$	0.9621	1.15	
W^2 mwc2 mwc5	0.8881	2.39	W^2 mwc2 mwc5	0.8879	2.40				
mwc3 mwc6 p3 $W^{0.25}$	0.9047	2.22	W mwc2 mwc3 mwc6	0.9046	2.24				
Acyclic Alkanes $n = 7-10$ $N = 137$			Acyclic Alkanes $n = 7-10$ $N = 135$			Acyclic Alkanes $n = 7-10$ $N = 52$			
$n^{0.5}$	0.9189	6.57	χ	0.9236	6.40	χ	0.9801	3.68	
1Z mwc5	0.9796	3.31	1Z mwc5	0.9803	3.26	1Z mwc6	0.9900	2.64	
mwc2 mwc5 $W^{0.25}$	0.9919	2.10	mwc2 mwc5 $W^{0.25}$	0.9918	2.11	mwc2 mwc5 $W^{0.25}$	0.9967	1.53	
mwc4 mwc6 p3 $W^{0.25}$	0.9929	1.97	mwc4 mwc6 p3 $W^{0.25}$	0.9930	1.96	mwc2 mwc5 $W^{0.25}$ p5	0.9978	1.27	
						mwc5 mwc6 $W^{0.25}$ p3 n	0.9987	1.00	
Acyclic Alkanes $n = 2-10$ $N = 149$			Acyclic Alkanes $n = 2-10$ $N = 147$			Acyclic Alkanes $n = 2-10$ $N = 64$			
$n^{0.5}$	0.9756	6.67	$n^{0.5}$	0.9756	6.70	$W^{0.25}$	0.9849	6.44	
p3 $W^{0.25}$	0.9891	4.47	p3 $W^{0.25}$	0.9891	4.49	p2 $n^{0.5}$	0.9944	3.97	
p3 $W^{0.25}$ n^2	0.9964	2.56	p3 n^2 $W^{0.25}$	0.9964	2.58	$W^{0.25}$ p3 n^2	0.9986	2.01	
mwc5 p2 $W^{0.25}$ n^2	0.9976	2.12	mwc5 p2 n^2 $W^{0.25}$	0.9976	2.13	$W^{0.25}$ p3 n^2 dia	0.9991	1.57	
mwc2 mwc3 mwc6 $W^{0.25}$ n^2	0.9980	1.92	mwc2 mwc3 mwc6 n^2 $W^{0.25}$	0.9980	1.93	$W^{0.25}$ mwc4 mwc5 p3 n	0.9994	1.34	
						W mwc5 mwc6 $W^{0.25}$ p3 n^2	0.9997	1.01	
B. Monocyclics									
all			rel > 0 only			rel = 2 only			
Octanes $n = 8$ $N = 34$			Octanes $n = 8$ $N = 24$			Octanes $n = 8$ $N = 13$			
Z	0.8188	5.21	Z	0.8751	4.47	Z	0.8852	3.76	
p3 1twc	0.9272	3.35	p3 1twc	0.9455	3.02	χ dia	0.9656	2.16	
1twc p4 ap3	0.9461	2.93	mwc4 twc dia	0.9710	2.26				
1twc p4 ap3 ap5	0.9576	2.65	1twc p4 ap3 ap5	0.9821	1.82				
Nonanes $n = 9$ $N = 58$			Nonanes $n = 9$ $N = 46$			Nonanes $n = 9$ $N = 15$			
Z	0.8149	4.86	Z	0.8438	4.41	1Z	0.9017	3.34	
Z ω	0.9038	3.53	Z s_r	0.9142	3.31	1Z s_r	0.9662	2.09	
Z ap5 s_r	0.9367	2.89	1twc p5 ap3	0.9470	2.63				
1Z p4 ap5 s_r	0.9512	2.56	Z p3 p5 ap3	0.9600	2.31				
			Z p3 p5 ap3 ap4	0.9713	1.98				
Decanes $n = 10$ $N = 73$			Decanes $n = 10$ $N = 38$			Decanes $n = 10$ $N = 7$			
Z	0.6919	6.56	Z	0.8291	4.97	Z	0.6958	2.76	
Z s_r	0.8401	4.76	Z s_r	0.9219	3.41				
W Z p4	0.8619	4.45	1Z p4 dia	0.9471	2.85				
mwc5 mwc6 p4 ap3	0.8876	4.05	W 1Z p4 ap5	0.9695	2.20				
mwc5 mwc6 1twc p4 ap3	0.9018	3.81	χ p4 ap3 ap5 s_r	0.9769	1.94				
mwc4 mwc5 1twc p4 ap2 ap3	0.9113	3.65							
Monocycloalkanes $n = 7-10$ $N = 189$			Monocycloalkanes $n = 7-10$ $N = 123$			Monocycloalkanes $n = 7-10$ $N = 42$			
1Z	0.9616	5.72	1Z	0.9695	4.90	1Z	0.9751	4.12	
1Z s_r	0.9786	4.28	1Z s_r	0.9841	3.56	1Z s_r	0.9871	3.01	
1Z dia s_r	0.9802	4.12	χ p4 s_r	0.9860	3.35	1Z p5 s_r	0.9916	2.45	
1Z p4 dia s_r	0.9834	3.79	χ p4 dia s_r	0.9886	3.04	1Z p4 dia s_r	0.9939	2.12	
Z 1Z p4 dia s_r	0.9847	3.65	W 1Z mwc2 mwc4 mwc6	0.9904	2.80	1twc ap2 ap4 ap5 s_r	0.9953	1.88	
1Z p4 ap5 ($W\omega$) s_r n^2	0.9864	3.45	Z 1Z twc p4 ap5 s_r	0.9924	2.50	Z 1Z twc p4 ap5 s_r	0.9968	1.57	
W 1Z twc 1twc p4 ap5 $n^{0.5}$	0.9873	3.33	W 1Z twc 1twc p4 ap5 $n^{0.5}$	0.9941	2.22				
Monocycloalkanes $n = 3-10$ $N = 209$			Monocycloalkanes $n = 3-10$ $N = 143$			Monocycloalkanes $n = 3-10$ $N = 54$			
1Z	0.9796	5.80	1Z	0.9859	5.16	1Z	0.9902	4.46	
1Z s_r	0.9858	4.84	W χ	0.9904	4.28	W 1Z	0.9955	3.04	
Z 1Z s_r	0.9893	4.22	Z 1Z s_r	0.9935	3.53	1Z ($W\omega$) s_r	0.9965	2.74	
Z 1Z ap5 s_r	0.9908	3.91	Z 1Z p5 s_r	0.9948	3.15	W χ ap4 s_r	0.9975	2.31	
W 1Z p4 dia ($W\omega$) ^{0.125}	0.9921	3.64	1Z p4 dia ($W\omega$) ^{0.125} n^2	0.9958	2.85	1twc p4 ap2 ap5 s_r	0.9982	1.99	
1Z p4 ap5 ($W\omega$) n s_r	0.9930	3.43	Z 1Z twc p4 ap5 s_r	0.9969	2.44	Z 1Z twc p4 ap5 s_r	0.9989	1.54	
W 1Z twc 1twc p4 ap5 n	0.9936	3.29	1Z twc p4 ap5 ($W\omega$) ($W\omega$) ^{0.125} s_r	0.9974	2.26				

Table 5. (Continued)

C. Bicyclics									
all			rel > 0 only			rel = 2 only			
Octanes $n = 8$ $N = 18$			Octanes $n = 8$ $N = 13$			Octanes $n = 8$ $N = 2$			
Z	0.7925	6.20	Z	0.8370	4.35				
mwc2 1twc	0.9235	3.89	1twc ap2	0.9340	2.90				
mwc3 mwc5 mwc6	0.9557	3.06							
Nonanes $n = 9$ $N = 36$			Nonanes $n = 9$ $N = 14$			Nonanes $n = 9$ $N = 1$			
mwc2	0.8011	4.53	mwc3	0.8433	3.77				
mwc2 ($W\omega$) ^{0.125}	0.8259	4.30	1twc p5	0.8641	3.66				
ω mwc4 ($W\omega$) ^{0.125}	0.8332	4.27							
χ mwc4 mwc6 twc	0.8540	4.06							
ω ap2 ap5 dia ($W\omega$) ^{0.125}	0.8789	3.76							
Decanes $n = 10$ $N = 46$			Decanes $n = 10$ $N = 18$			Decanes $n = 10$ $N = 5$			
mwc2	0.7718	5.78	mwc2	0.8962	4.13				
Z 1twc	0.8060	5.39	mwc4 ap4	0.9355	3.36				
mwc2 ap4 dia	0.8471	4.85	mwc3 ap3 ap4	0.9595	2.76				
Z 1twc ap4 dia	0.8582	4.72							
Bicycloalkanes $n = 7-10$ $N = 115$			Bicycloalkanes $n = 7-10$ $N = 53$			Bicycloalkanes $n = 7-10$ $N = 12$			
1Z	0.9517	6.34	1Z	0.9728	4.63	1Z	0.9685	5.97	
1Z twc	0.9622	5.63	1Z ap4	0.9783	4.18	1Z mwc6	0.9876	3.95	
mwc4 mwc5 $n^{0.5}$	0.9672	5.27	1Z 1twc ($W\omega$) ^{0.125}	0.9811	3.94				
mwc4 mwc6 dia $n^{0.5}$	0.9702	5.05	1Z 1twc p5 ($W\omega$) ^{0.125}	0.9834	3.73				
mwc4 mwc6 ap5 dia $n^{0.5}$	0.9728	4.85	mwc2 p2 ap3 ap4 ($W\omega$) ^{0.125}	0.9847	3.62				
mwc6 ap2 ap4 ap5 dia $n^{0.5}$	0.9753	4.64	ω 1twc p4 ap3 ($W\omega$) ^{0.125} n^2	0.9857	3.54				
Bicycloalkanes $n = 4-10$ $N = 127$			Bicycloalkanes $n = 4-10$ $N = 61$			Bicycloalkanes $n = 4-10$ $N = 15$			
1Z	0.9741	6.33	1Z	0.9851	4.70	1Z	0.9842	5.92	
1Z twc	0.9802	5.55	1Z twc	0.9884	4.18	1Z n^2	0.9941	3.76	
mwc4 mwc5 $n^{0.5}$	0.9828	5.20	1Z ($W\omega$) ^{0.125} n^2	0.9903	3.85	mwc3 mwc4 ($W\omega$) ^{0.125}	0.9966	2.97	
W mwc4 mwc6 n	0.9838	5.07	ap4 1Z ($W\omega$) ^{0.125} n^2	0.9912	3.70				
mwc4 mwc5 p5 ap5 $n^{0.5}$	0.9852	4.85							
mwc6 ap2 ap4 ap5 dia $n^{0.5}$	0.9871	4.56							
D. Saturated Hydrocarbons (Acyclic–Polycyclic)									
all			rel > 0 only			rel = 2 only			
Octanes $n = 8$ $N = 79$			Octanes $n = 8$ $N = 59$			Octanes $n = 8$ $N = 34$			
χ	0.6586	7.27	χ	0.7363	5.96	1Z	0.7689	5.10	
Z p4	0.7655	6.06	Z p4	0.8584	4.41	Z p4	0.9308	2.83	
Z ω p4	0.8329	5.15	Z ω p4	0.8941	3.84	Z p4 dia	0.9547	2.33	
Z ω p4 n_r	0.8601	4.75	Z ω p4 n_r	0.9133	3.51	W Z 1twc p4	0.9711	1.89	
Z ω mwc6 twc ap4	0.8829	4.37	Z ω 1twc p4 ap3	0.9330	3.11				
Nonanes $n = 9$ $N = 144$			Nonanes $n = 9$ $N = 98$			Nonanes $n = 9$ $N = 31$			
χ	0.6725	6.45	χ	0.7191	5.72	Z	0.8540	4.36	
Z twc	0.7891	5.19	Z ap4	0.8461	4.26	Z twc	0.9310	3.05	
Z ω ap4	0.8280	4.71	W Z ap4	0.8911	3.60	W Z mwc5	0.9519	2.59	
1Z twc ($W\omega$) ($W\omega$) ^{0.125}	0.8592	4.27	W Z 1twc ap4	0.9157	3.18	Z p3 p5 ap3	0.9614	2.37	
1Z mwc3 mwc4 ($W\omega$) ($W\omega$) ^{0.125}	0.8742	4.05	W Z 1twc ap3 ap4	0.9236	3.05				
Decanes $n = 10$ $N = 208$			Decanes $n = 10$ $N = 133$			Decanes $n = 10$ $N = 24$			
χ	0.5090	8.50	χ	0.6119	7.00	χ	0.7284	5.87	
Z twc	0.6622	7.07	Z p5	0.8200	4.78	1Z twc	0.9074	3.51	
W Z ap4	0.7440	6.17	W Z p4	0.8698	4.08	1Z ω ap4	0.9464	2.74	
Z p3 p4 ap3	0.7933	5.56	W Z p4 ap3	0.8841	3.87				
Z p3 p4 ap3 dia	0.8144	5.28	W 1Z p4 ($W\omega$) ($W\omega$) ^{0.125}	0.9084	3.45				
Saturated Hydrocarbons $n = 7-10$ $N = 486$			Saturated Hydrocarbons $n = 7-10$ $N = 326$			Saturated Hydrocarbons $n = 7-10$ $N = 109$			
χ	0.9245	7.50	χ	0.9406	6.31	χ	0.9570	5.57	
1Z n_r	0.9320	7.12	χ dia	0.9496	5.82	1Z p5	0.9668	4.92	
χ mwc6 dia	0.9440	6.47	1Z p4 $n^{0.5}$	0.9571	5.38	χ twc dia	0.9773	4.09	
χ ω ap4 ($W\omega$)	0.9496	6.14	W 1Z p4 n	0.9650	4.86	W χ ω p4	0.9829	3.57	
χ ω ap4 dia ($W\omega$)	0.9551	5.80	W 1Z 1twc ap4 n	0.9696	4.54	χ ω p4 dia ($W\omega$)	0.9859	3.25	
1Z ω 1twc ap4 ($W\omega$) n^2	0.9573	5.67	W 1Z mwc3 1twc ap4 n	0.9723	4.34				
Saturated Hydrocarbons $n = 2-10$ $N = 530$			Saturated Hydrocarbons $n = 2-10$ $N = 366$			Saturated Hydrocarbons $n = 2-10$ $N = 136$			
χ	0.9601	7.97	χ	0.9692	7.31	($W\omega$) ^{0.125}	0.9787	7.05	
W χ	0.9645	7.53	W χ	0.9765	6.39	W χ	0.9840	6.14	
W χ ap4	0.9726	6.63	1Z p4 $n^{0.5}$	0.9820	5.61	W χ ap4	0.9909	4.66	
W 1Z mwc4 $n^{0.5}$	0.9761	6.19	W 1Z p4 $n^{0.5}$	0.9860	4.95	W 1Z mwc4 $n^{0.5}$	0.9942	3.72	
W 1Z 1twc ap4 n	0.9794	5.75	W 1Z 1twc ap4 n	0.9882	4.56	W 1Z mwc4 ap4 $n^{0.5}$	0.9949	3.52	
W 1Z 1twc ap4 dia $n^{0.5}$	0.9803	5.64	W 1Z 1twc ap4 n^2 $n^{0.5}$	0.9886	4.47	χ ω ap4 dia n^2 $n^{0.5}$	0.9956	3.28	

unchecked bps. This justifies the quality classification of bps introduced in this work. This moreover is an independent

justification for the use of TIs in QSPR studies: If the high-quality bps were less well-described by TIs than all bps, then

Table 6. Comparison of Structure–Bp Relationships for (Poly)Cyclic Alkanes from the Recent Literature (left, mostly incomplete samples), from This Work for the Same Samples (middle), and for the Corresponding Complete Samples, as Far as $\text{rel} = 2$ (right)

no. of descr	<i>s</i>	no. of descr	descriptors	<i>s</i>	no. of descr	descriptors	<i>s</i>
Monocyclic Alkanes $n = 3-6$ $N = 20^a$					Monocyclic Alkanes $n = 3-6$ $\text{rel} = 2$ $N = 12$		
1	3.69 ^b	1	1Z	2.84	1	1Z	2.71
2	2.87 ^b	2	1Z W	2.48	2	p2 ($W\omega$) ^{0.125}	2.05
Some Monocyclic Alkanes $n = 4-10$ $N = 45$					Monocyclic Alkanes $n = 4-10$ $\text{rel} = 2$ $N = 54$		
1	5.93 ^c	1	1Z	5.13	1	1Z	4.46
Some Monocyclic Alkanes $n = 7-10$ $N = 30$					Monocyclic Alkanes $n = 7-10$ $\text{rel} = 2$ $N = 42$		
1	7.10 ^d	1	1Z	5.40	1	1Z	4.12
2	5.30 ^d	2	1Z s_r	2.88	2	1Z s_r	3.01
Acyclic and Some Monocyclic Octanes $N = 32$					Acyclic and Monocyclic Octanes $\text{rel} = 2$ $N = 31$		
		2	Z p4	2.29	2	Z p4	2.86
3	2.98 ^e	3	Z p4 dia	1.70	3	Z p4 dia	2.26
4	2.54 ^e	4	Z p4 ap5 s_r	1.26	4	Z p4 ω ap5	1.78
Acyclic and Some Monocyclic Octanes $N = 29$					Acyclic and Monocyclic Octanes $\text{rel} = 2$ $N = 31$		
2	(4.78) ^{f,g}	2	Z p5	2.92	(see above)		
2	3.00 ^f	2	Z p4	2.38			
Some Acyclic and Monocyclic Alkanes $n = 4-9$ $N = 69$					Acyclic and Monocyclic Alkanes $n = 4-9$ $\text{rel} = 2$ $N = 99$		
7	3.43 ^h	3	1Z $n^{0.5}$ p5	3.16	6	mwc4 mwc6 ltwc ap4 ap5 s_r	2.46
8	3.60 ^h	7	W mwc4 mwc6 ap4 ap5 n s_r	1.70			
Some Mono- and Bicyclic Alkanes $n = 3-9$ $N = 80$					Mono- and Bicyclic Alkanes $n = 3-9$ $\text{rel} = 2$ $N = 57$		
		2	W χ	4.23	1	χ	5.00
6	4.80 ⁱ	6	n χ p4 ap3 dia ($W\omega$) ^{0.125}	2.34	4	1Z mwc2 dia ($W\omega$) ^{0.125}	2.54
Some Acyclic–Bicyclic Alkanes $n = 4-10$ $N = 130$					Acyclic–Bicyclic Alkanes $n = 4-10$ $\text{rel} = 2$ $N = 131$		
7	5.04 ^j	3	1Z ap5 ($W\omega$) ^{0.125}	4.89	7	W mwc3 mwc6 p4 ap3 n s_r	3.08
9	3.65 ^j	5	mwc5 p4 ap3 ap5 ($W\omega$) ^{0.125}	3.63			
12	3.03 ^j	7	W 1Z mwc4 mwc6 ap2 ap5 n^2	2.85			

^a Complete sample. ^b Reference 11. ^c Reference 13b. ^d Reference 13a. ^e Reference 13c. ^f Reference 3c. ^g *s* value calculated in this work. ^h Reference 6. ⁱ Reference 14. ^j Reference 7.

the better correlations would be suspect to be due to mere coincidence. (2) In going from the octanes to the decanes, the correlations with a given number of descriptors become less and less satisfactory. (3) In many samples the best single descriptor for bps is Z or $\log(Z)$. (4) The detour index ω or its variants ($W\omega$) or ($W\omega$)^{0.125}, used earlier for samples of cyclic compounds, is apparently very useful. (5) The molecular walk counts of length k (mwck), their half-sum (twc), and the logarithm thereof are useful descriptors, taking part in most of the more precise models. (6) There is no descriptor or combination of few descriptors found which would describe the bps in all or most of the various samples with reasonable precision (say, $s < 1^\circ\text{C}$). On the contrary, with the possible exception of the acyclic alkanes, we are far from that goal, even in the samples with high-quality bp values.

For this latter disappointing result at least two possible reasons are obvious: First, of course, there is no warranty that the good descriptor (combination) which may exist for bps is included in our limited descriptor pool. Second, there is no reason why the actual dependence of bps on any descriptor should be linear.

In Table 6 we compare our structure–bp correlations for cycloalkanes with results from the recent literature. In each case we first use the compound sample from the respective reference; then we consider the corresponding comprehensive sample with highly reliable bp values ($\text{rel} = 2$), which is assumed to be more significant than the corresponding complete sample including low-quality bps.

Acyclic alkanes are not the focus of the present study. Nevertheless, having compiled the data and having obtained the best correlations within our pool of descriptors for acyclics as well, we compare in Table 7 our models to the best ones (as far as we are aware) from the recent literature. The data may be useful to other workers in the field.

It is seen from Tables 6 and 7 that for most samples considered, our pool of diverse TIs allowed to construct better regression models for the bp than were available hitherto. Of course, in much of the work referenced in the tables, the primary objective was to test the members of a family of descriptors, rather than to obtain the very best model. In contrast, the better performance of our models built from diverse descriptors emphasizes the usefulness of nonrelated descriptors, a few of which may span the relevant topological space more efficiently than highly intercorrelated descriptors from one family.

Deceptively Impressive Regressions Obtained from Incomplete or Biased Data. The best two-variable regression we were able to find for monocyclic octanes includes as independent variables p3 and ltwc (see Table 5B), for all such compounds with known bps:

$$\text{bp} = (4.78 \pm 0.55)\text{p3} - (159.47 \pm 10.48)\text{ltwc} + (586.70 \pm 28.72)$$

$$N = 34, r^2 = 0.9272, s = 3.35, F = 197.5 \quad (1)$$

and for the compounds with reproduced bps ($\text{rel} > 0$):

$$\text{bp} = (5.55 \pm 0.71)\text{p3} - (165.95 \pm 11.63)\text{ltwc} + (599.85 \pm 30.86)$$

$$N = 24, r^2 = 0.9455, s = 3.02, F = 182 \quad (2)$$

The experimental bps used in these regressions are shown once more in column 2 of Table 8. We now want as a caveat to demonstrate the effect of selective perception of data.

First trap: noninclusion of some structures. If 1-ethyl-2-propylcyclopropane (1e2pc3), 1-butyl-2-methylcyclopropane (b2mc3), 1,1-dimethyl-2-isopropylcyclopropane (11m2ipc3), and 1,1-dimethyl-2-propylcyclopropane (11m2pc3) are simply excluded from the regression, most readers will not miss them. In fact these compounds have never before been included in structure–bp correlations (though all their bp data known today are available since 1973), and there are even reasons for excluding them, in that these are “exotic” or “unimportant” substances whose bps were measured only once or twice and as such are not known with some certainty. The result of this exclusion is the following correlation:

$$\text{bp} = (4.84 \pm 0.45)\text{p3} - (159.10 \pm 8.84)\text{ltwc} + (584.92 \pm 24.36)$$

$$N = 30, r^2 = 0.9499, s = 2.71, F = 256 \quad (3)$$

which obviously is somewhat better than eqs 1 and 2. Note that the four structures excluded were not the worst cases as measured by the residuals from regressions 1 and 2, columns 6 and 7 in Table 8.

Second trap: selective perception of bp records. Table 8 also contains the bp range and the number of bp records in Beilstein for the given compounds (columns 4 and 5). We next changed all bp values, shifting them into the direction of better fit, as suggested by the residuals, but not beyond the range of the observations documented in Beilstein. The new fictitious bps (columns 8 and 9 in Table 8) therefore look innocent: they are manipulated but not out of the range of true experimental observations. Now using these fictitious “experimental” bp values, we obtained the following regressions for all monocyclic octanes with known bps:

$$\text{bp} = (4.79 \pm 0.28)\text{p3} - (158.09 \pm 5.34)\text{ltwc} + (582.09 \pm 14.63)$$

$$N = 34, r^2 = 0.9794, s = 1.71, F = 737 \quad (4)$$

and for the monocyclic octanes with reproduced bps (rel > 0):

$$\text{bp} = (5.73 \pm 0.20)\text{p3} - (167.38 \pm 3.34)\text{ltwc} + (602.56 \pm 8.86)$$

$$N = 24, r^2 = 0.9952, s = 0.87, F = 2177 \quad (5)$$

These latter regressions (to be compared to eqs 1 and 2) seem impressive and even seem to allow predictions, yet in fact they are the result of severe data manipulation. It is needless to detail what may happen if both erroneous procedures are unintentionally combined.

CONCLUDING REMARK

It is obvious that the problems discussed here in the context of multilinear regression of bp data of cycloalkanes using TIs pertain to other molecular properties, to other sets of

Table 7. Comparison of Structure–Bp Relationships for Samples of Acyclic Alkanes from the Literature (left) and from the Present Work (right)^z

no. of descr	s	no. of descr	descriptors	s
Acyclic Octanes $N = 18^a$				
2	1.11 ^{b-d}	2	W p3	0.96
2	0.866 ^d	3	W ^{0.25} p3 ltwc	0.85
Acyclic Nonanes $N = 35^e$				
1	3.56 ^{f,g}	1	XMT4R	3.42
2	2.70 ^h	2	mwc4 mwc6	1.49
3	1.60 ⁱ	3	Z p4 p5	1.27
5	1.69 ^j	4	mwc5 twc p3 W ^{0.25}	1.05
5	0.98 ^k	5	W ² mwc3 mwc4 mwc6 twc	0.89
6	1.687 ^l			
Acyclic Alkanes $n = 6-8$ $N = 32$				
		2	1Z mwc5	2.26
3	2.37 ^l	3	W ^{0.25} p3 n	0.87
Acyclic Alkanes $n = 3-8$ $N = 38$				
2	4.92 ^m	2	$\chi^{0.33}$ mwc5	2.21
3	3.75 ^m	3	p3 W ^{0.25} n	1.38
4	2.37 ^m	4	p3 W ^{0.25} dia n ²	0.79
5	1.75 ^m	5	p3 W ^{0.25} W ² dia n ²	0.72
4	0.91 ^m			
5	0.64 ^m			
Acyclic Alkanes $n = 2-8$ $N = 39$				
1	5.79 ⁿ	1	$\chi^{0.33}$	4.13
Acyclic Alkanes $n = 3-9$ $N = 73$				
1	5.97 ^o	1	$\chi^{0.25}$	4.96
2	3.63 ^o	2	mwc5 $\chi^{0.25}$	2.75
3	2.17 ^o	3	p3 W ^{0.25} n ²	1.75
4	1.41 ^o	4	ltwc p3 W ^{0.25} n	1.40
		5	mwc2 mwc3 mwc6 W ^{0.25} n ²	1.19
Acyclic Alkanes $n = 2-9$ $N = 74^{e,p}$				
2	4.74 ^q	1	$\chi^{0.25}$	4.94
2	3.83 ^r	2	mwc5 $\chi^{0.25}$	2.74
3	3.52 ^q	3	p3 W ^{0.25} n ²	1.73
3	1.84 ^k	4	ltwc p3 W ^{0.25} n	1.43
5	2.21 ^s	5	mwc2 mwc3 mwc6 W ^{0.25} n ²	1.19
5	1.38 ^k			
9	3.39 ^t			
12	1.64 ^t			
Acyclic Alkanes $n = 6-10$ $N = 109^u$				
3	6.59 ^v	3	mwc2 mwc5 W ^{0.25}	1.98
6	5.25 ^v	4	mwc4 mwc6 p3 W ^{0.25}	1.90
Acyclic Alkanes $n = 6-10$ $N = 142$				
		3	mwc2 mwc5 W ^{0.25}	2.15
		4	mwc2 mwc3 mwc6 W ^{0.25}	2.02
Acyclic Alkanes $n = 6-10$ rel = 2 $N = 57$				
		3	mwc2 mwc5 W ^{0.25}	1.57
		4	mwc2 mwc5 p5 W ^{0.25}	1.31
		5	mwc5 mwc6 p3 W ^{0.25} n	1.06
Acyclic Alkanes $n = 2$ $N = 149$				
2	9.46 ^w	2	$\chi^{0.25}$ mwc5	3.55
10	3.34 ^{x,y}	5	mwc2 mwc3 mwc6 W ^{0.25} n ²	1.92

^a See also ref 23. ^b Descriptors W and p3, bp values different from ours were used. ^c Reference 24. ^d Reference 25. ^e For this sample XMT4R was included as an additional descriptor; see ref 26. ^f Descriptor XMT4R, bp values different from ours were used. ^g Reference 26. ^h Reference 27. ⁱ Reference 18b. ^j Reference 28. ^k Reference 29. ^l Reference 23b. ^m Reference 30. ⁿ Reference 31. ^o Reference 32. ^p See also refs 26 and 33. ^q Reference 34. ^r Reference 35. ^s Reference 36a. ^t Reference 36b. ^u Incomplete sample. ^v Reference 37. ^w Reference 38. ^x Two outliers excluded. ^y Reference 39. ^z In the calculations given here $\chi^{0.25}$, $\chi^{0.33}$, and $\chi^{0.5}$ were included as additional descriptors; see ref 22.

compounds, to other types of descriptors, and to other statistical methods as well. We stress the importance of obtaining a comprehensive unbiased data set before under-

Table 8. Bp Data for Monocyclic Octanes

name	bp (°C)	rel	in Beilstein		residual		fictitious bp used	
			bp range	no. of bp rec	eq 1	eq 2	eq 4	eq 5
c8	149	2	120.3–156	23	3.58	3.75	145.4	145.2
1mc7	134	2	113–136	10	−0.26	−0.21	134.3	134.2
1ec6	131.8	2	129–136	28	−0.93	−1.39	132.7	133.2
1pc5	131	2	127–131.5	12	0.61	0.23	130.4	130.8
1pec3	128	2	124.1–129	11	−0.32	−0.61	128.3	128.6
14mc6	121.8	2	119–124.9	29	−2.95	−3.09	124.8	124.9
13mc6	122.3	2	105–125	40	−1.97	−2.09	124.3	124.4
1ibc4	120.1	0	119.9–120.4	1	−1.47	—	120.4	—
12mc6	126.6	1	122–130	26	0.72	−0.05	125.9	126.7
1ipc5	126.4	2	118–129.3	9	2.33	1.64	124.1	124.8
1e3mc5	121	2	117–121.4	8	−1.66	−2.29	121.4	121.4
1e2mc5	124.7	1	121.2–128.4	13	0.54	−0.74	124.2	125.4
p3mc4	117.4	1	116.1–118.7	2	−1.64	−2.13	118.7	118.7
11mc6	119.5	2	115–125	19	5.23	5.52	115	115
1sbc4	123	0	123	1	−0.52	—	123	—
1spec3	117.7	0	117.74	1	−3.05	—	117.7	—
b2mc3	124	1	121.1–128	2	6.26	5.25	121.1	121.1
12ec4	119	1	115.5–123	2	−2.79	−4.55	121.8	123
124mc5	115	1	109–120	19	−1.50	−2.46	116.5	117.5
1e1mc5	121.5	2	98–121.5	5	6.28	5.37	115.2	116.1
123mc5	117	1	109–123	24	−2.03	−3.67	119	120.7
1nepec3	106	0	106–106.2	1	1.83	—	106	—
5msbc3	115.5	0	115.5	1	−1.20	—	115.5	—
1e2pc3	108	0	99–117.5	2	−8.57	—	116.6	—
ib2mc3	110	0	109.5–110.5	1	−1.35	—	110.5	—
113mc5	104.9	2	104–116	10	−1.01	−0.96	105.9	105.9
112mc5	114	2	113–114.5	9	3.29	1.98	113	113
1234mc4	114.5	0	106.5–122.5	2	0.56	—	113.9	—
11m2pc3	105.9	1	105.5–106.3	2	3.48	1.94	105.5	105.5
1133mc4	86	1	84–87.6	3	−3.14	−2.99	87.6	87.6
1m12ec3	108.9	1	108.3–109.9	2	2.50	−0.35	108.3	109.3
11m2ipc3	94.4	0	93.9–94.9	1	4.82	—	94.9	—
112m2ec3	104.5	1	103.5–104.5	2	5.63	1.91	103.5	103.5
11223mc3	100.5	0	100–101	1	−1.65	—	101	—

taking a structure–property relationship study and of not restricting one's attention to descriptors from one class of related TIs.

EXPERIMENTAL SECTION

All saturated hydrocarbons up to $n = 10$ with known bps were identified in the Beilstein database Crossfire Plus Reactions, as of the third update for 1998 (BS9803PR). Whenever possible, data were checked in the primary literature. Unfortunately, in many cases this was not possible for lack of access to older, in particular to Russian, journals. Only those structures were included in the data file for which at least one bp record at a pressure ≥ 720 Torr was present. Bp values measured between 720 and 800 Torr were adjusted to 760 Torr using an increment of 0.05 °C/Torr. Multiple entries of bps in Beilstein, resulting from multiple measurements at the same or different pressures or from stereoisomers or isotopomers, were averaged. All bp and s values in this paper are given in °C.

For a few compounds bp values differing by an unreasonable margin or bps unreasonable on comparison with those of structurally related compounds were reported; such structures were excluded from the data file. For example, for 1,1,4-trimethylcycloheptane bp values as discrepant as 152, 158, and 164 °C are reported. For 1,4-dimethylbicyclo[2.2.1]heptane a bp of 124.5 °C is given, far too low compared to the reliable value of 117 °C for 1-methylbicyclo[2.2.1]heptane. The bp given for 1-*tert*-butylbicyclo[1.1.1]-

pentane, 90 °C 760 Torr, is far too low compared with the 100 °C 70 Torr for 1-*tert*-butyl-3-methylbicyclo[1.1.1]-pentane and in fact is far lower than the bp of any other nonane. For similar reasons, the bps recorded in Beilstein for 1-methyl-2-isopentylcyclopropane (114 °C), 1-ethyl-2-propylcyclopentane (72.9 and 77.4 °C), 1,3,3,5-tetramethylbicyclo[3.1.0]hexane (137.2 °C), and 3-propyltricyclo[2.2.1.0^{2,6}]-heptane (132 °C) are regarded as unreasonable.

While this study was underway, a rather old book was brought to our attention (thanks go to Dr. E. Estrada) in which a few other saturated hydrocarbons with bps are collected, which for reasons unknown to us do not appear in Beilstein.⁴⁰ These were also included in the data file but regarded as “not reproduced”, $rel = 0$ (unless reported by Rossini), since the values could not be checked in the primary literature. The bp value of scptc3210o was found in a paper referenced by Beilstein in another context.⁴¹ There are probably some more saturated hydrocarbons of $n \leq 10$ with known bps buried in the literature, but we do not see a systematic way to uncover them.

Hosoya index (Z) values in this work were obtained in a novel way by our unpublished program ZPATH. Program ZPATH reduces the problem of generating all possible combinations of nonadjacent edges from the graph to a path-tracing problem. A selection of nonadjacent edges is a string of labels (in increasing order) of nonadjacent edges, which can be found by means of the edge-adjacency matrix in the same way as a string of labels of adjacent vertices (a path) is found using the usual vertex-adjacency matrix. The resulting $p(G,k)$ values and their sums, the Z values, were checked against and found identical with those published for acyclic alkanes (up to $n = 8$),^{42a} mono- and bicyclic alkanes (up to $n = 8$),^{42b} and (perhydro)fullerenes (up to C_{42}).⁴³ Program ZPATH is thus a means to calculate the matching polynomial of a graph in a very elementary manner.

The Wiener index, detour index, and diameter values were obtained using our program DETOUR.¹⁰ The walk counts $mwck$ and the total walk count twc were obtained using our program MORGAN.¹⁷ The path counts pk (including cyclic paths (circles)) and the acyclic path counts apk (without circles) were obtained by our unpublished program PATHS; the apk were checked against and found identical to published values.^{6,44}

In the comparisons with published structure–bp relationships we used the experimental bp values as contained in our data file, that is in Table 1, not the values used in the original work, since we tried to obtain comparable s values and since we think our carefully checked bp values, as a rule, are more correct. In several cases we employed, as a double-check, additionally the bp values as given in the respective reference, without significant changes in the obtained s values, except for refs 30 (where very uncommon bp values were used), 13c, and 14.

Where the original authors treated *cis/trans* isomeric cycloalkanes as two independent cases (exhibiting identical TIs but different bp values), we (not distinguishing stereoisomers) included the same structure (with its averaged bp value) twice in order to obtain a comparable sample.

Zinn et al.^{6,7} use a convention for the designation of paths that differs from the one used by most other authors including us. They characterize a path by the number of nodes on it, while we use the number of edges on it, so that their pk is

our $pk-1$ for acyclic structures. Further, for cyclic structures they count acyclic paths only, so that their pk corresponds to our $apk-1$.

In ref 7 the descriptor and bp values given for spiro[4.5]-decane actually correspond to bicyclo[5.3.0]decane, so that we included bc530d rather than s45d in our comparison sample. For one compound in ref 7, 1-butyl-1-methylcyclopentane, we were not able to find a bp value in the primary literature. This compound therefore is not contained in our data file; it was however included in the sample for comparison with Zinn's results with the bp value given in ref 7.

REFERENCES AND NOTES

- (1) (a) Katritzky, A. R.; Lobanov, V. S.; Karelson, M. Normal Boiling Points for Organic Compounds: Correlation and Prediction by a Quantitative Structure-Property Relationship. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 28–41. (b) Ivanciuc, O.; Ivanciuc, T.; Balaban, A. T. Quantitative Structure-Property Relationship Study of Normal Boiling Points for Halogen-/Oxygen-/Sulfur-Containing Organic Compounds Using the CODESSA Program. *Tetrahedron* **1998**, *54*, 9129–9142.
- (2) Randić, M.; Mihalić, Z.; Nikolić, S.; Trinajstić, N. Graph-Theoretical Correlations – Artifacts or Facts? *Croat. Chem. Acta* **1993**, *66*, 411–434.
- (3) (a) Diudea, M. V.; Gutman, I. Wiener-Type Topological Indices. *Croat. Chem. Acta* **1998**, *71*, 21–51. (b) Gutman, I.; Diudea, M. V. Correcting the Definition of Cluj Matrices. *MATCH* **1998**, *37*, 195–201. (c) Lukovits, I.; Linert, W. Polarity-Numbers of Cycle-Containing Structures. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 715–719.
- (4) Hosoya, H.; Kawasaki, K.; Mizutani, K. Topological Index and Thermodynamic Properties. I. Empirical Rules on the Boiling Point of Saturated Hydrocarbons. *Bull. Chem. Soc. Jpn.* **1972**, *45*, 3415–3421.
- (5) Lukovits, I.; Linert, W. A Novel Definition of the Hyper-Wiener Index for Cycles. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 899–902.
- (6) Gautzsch, R.; Zinn, P. List Operations on Chemical Graphs. 5. Implementation of Breadth-First Molecular Path Generation and Application in the Estimation of Retention Index Data and Boiling Points. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 791–800.
- (7) Hermann, A.; Zinn, P. List Operations on Chemical Graphs. 6. Comparative Study of Combinatorial Topological Indexes of the Hosoya Type. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 551–560.
- (8) Lukovits, I. The Detour Index. *Croat. Chem. Acta* **1996**, *69*, 873–882.
- (9) Trinajstić, N.; Nikolić, S.; Lucić, B.; Amić, D.; Mihalić, Z. The Detour Matrix in Chemistry. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 631–638.
- (10) Rücker, G.; Rücker, C. Symmetry-Aided Computation of the Detour Matrix and the Detour Index. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 710–714.
- (11) Plavšić, D.; Soskić, M.; Daković, Z.; Gutman, I.; Graovac, A. Extension of the Z Matrix to Cycle-Containing and Edge-Weighted Molecular Graphs. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 529–534.
- (12) Das, A.; Dömötör, G.; Gutman, I.; Joshi, S.; Karmarkar, S.; Khaddar, D.; Khaddar, T.; Khadikar, P. V.; Popović, L.; Sapre, N. S.; Sapre, N.; Shirhatti, A. A Comparative Study of the Wiener, Schultz and Szeged Indices of Cycloalkanes. *J. Serb. Chem. Soc.* **1997**, *62*, 235–239.
- (13) (a) Diudea, M. V. Indices of Reciprocal Properties or Hary Indices. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 292–299. (b) Diudea, M. V. Cluj Matrix Invariants. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 300–305. (c) Diudea, M.; Katona, G.; Lukovits, I.; Trinajstić, N. Detour and Cluj-Detour Indices. *Croat. Chem. Acta* **1998**, *71*, 459–471.
- (14) Estrada, E. Spectral Moments of the Edge Adjacency Matrix in Molecular Graphs. 3. Molecules Containing Cycles. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 23–27.
- (15) For example: propylcyclobutane, ref 13b used 110, ref 14 100.6 °C; isopropylcyclohexane, ref 13b used 146, ref 7 154.6, refs 8 and 9 171.3 °C; 1,3-dimethylcyclohexane, ref 13a used 120, ref 13b 124.5 °C.
- (16) Rossini, F. D.; Pitzer, K. S.; Arnett, R. L.; Braun, R. M.; Pimentel, G. C. *Selected Values of Physical and Thermodynamic Properties of Hydrocarbons and Related Compounds, Comprising the Tables of the American Petroleum Institute Research Project 44, Extant as of December 31, 1952*; Carnegie Press: Pittsburgh, PA, 1953. We were not able to locate a more recent edition of this compilation (as far as to bp values). However, a more recent reference work, citing as source of data the findings of API Research Project 44, gives somewhat different bp values: *Physical Constants of Hydrocarbons C₁ to C₁₀*; prepared by ASTM Committee D-2 on Petroleum and Lubricants and API Research Project 44 on Hydrocarbons and Related Compounds; Philadelphia, PA, 1971. Many of the latter bp values were used and printed (citing as data source the API Research Project 44, 1947–1991) in the following: Gakh, A. A.; Gakh, E. G.; Sumpter, B. G.; Noid, D. W. Neural Network-Graph Theory Approach to the Prediction of the Physical Properties of Organic Compounds. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 832–839. In the absence of evidence for the latter values being better than the former, we decided to use the original values from Rossini (1953).
- (17) Rücker, G.; Rücker, C. Counts of All Walks as Atomic and Molecular Descriptors. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 683–695.
- (18) (a) Diudea, M. V.; Minailiuc, O. M.; Katona, G. Novel Connectivity Descriptors Based on Walk Degrees. *Croat. Chem. Acta* **1996**, *69*, 857–871. (b) Ivanciuc, O.; Diudea, M.; Khadikar, P. V. New Topological Matrices and Their Polynomials. *Indian J. Chem., Sect. A* **1998**, *37A*, 574–585.
- (19) Wieland, T.; Kerber, A.; Laue, R. Principles of Generation of Constitutional and Configurational Isomers. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 413–419.
- (20) Hendrickson, J. B.; Huang, P.; Toczko, A. G. Molecular Complexity: A Simplified Formula Adapted to Individual Atoms. *J. Chem. Inf. Comput. Sci.* **1987**, *27*, 63–67.
- (21) Razingier, M.; Chretien, J. R.; Dubois, J. E. Structural Selectivity of Topological Indexes in Alkane Series. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 23–27.
- (22) Randić, M.; Hansen, P. J.; Jurs, P. C. Search for Useful Graph Theoretical Invariants of Molecular Structure. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 60–68.
- (23) (a) Randić, M. Comparative Regression Analysis: Regressions Based on a Single Descriptor. *Croat. Chem. Acta* **1993**, *66*, 289–312. (b) Bhattacharjee, S.; Dasgupta, P. Molecular Property Correlation in Alkanes With Geometric Volume. *Comput. Chem.* **1994**, *18*, 61–71. (c) Nikolić, S.; Trinajstić, N.; Mihalić, Z. The Wiener Index: Development and Applications. *Croat. Chem. Acta* **1995**, *68*, 105–129.
- (24) (a) Wiener, H. Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc.* **1947**, *69*, 17–20. (b) Kirby, E. C. Sensitivity of Topological Indices to Methyl Group Branching in Octanes and Azulenes, or What Does a Topological Index Index? *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1030–1035.
- (25) Randić, M. Linear Combinations of Path Numbers as Molecular Descriptors. *New J. Chem.* **1997**, *21*, 945–951.
- (26) Medeleau, M.; Balaban, A. T. Real-Number Vertex Invariants and Schultz-Type Indices Based on Eigenvectors of Adjacency and Distance Matrices. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 1038–1047.
- (27) Plavšić, D.; Soskić, M.; Landeka, I.; Gutman, I.; Graovac, A. On the Relation between the Path Numbers $\sum Z$, $\sum^2 Z$ and the Hosoya Z Index. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1118–1122.
- (28) Randić, M.; Trinajstić, N. Isomeric Variations in Alkanes: Boiling Points of Nonanes. *New J. Chem.* **1994**, *18*, 179–189.
- (29) Lucić, B.; Trinajstić, N. New Developments in QSPR/QSAR Modeling Based on Topological Indices. *SAR QSAR Environ. Res.* **1997**, *7*, 45–62.
- (30) Bonchev, D. Novel Indices for the Topological Complexity of Molecules. *SAR QSAR Environ. Res.* **1997**, *7*, 23–43.
- (31) Ren, B. A New Topological Index for QSPR of Alkanes. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 139–143.
- (32) Toropov, A.; Toropova, A.; Ismailov, T.; Bonchev, D. 3D Weighting of Molecular Descriptors for QSPR/QSAR by the Method of Ideal Symmetry (MIS). 1. Application to Boiling Points of Alkanes. *J. Mol. Struct. (Theochim)* **1998**, *424*, 237–247.
- (33) (a) Needham, D. E.; Wei, I.-C.; Seybold, P. G. Molecular Modeling of the Physical Properties of the Alkanes. *J. Am. Chem. Soc.* **1988**, *110*, 4186–4194. (b) Basak, S. C.; Niemi, G. J.; Veith, G. D. Predicting Properties of Molecules Using Graph Invariants. *J. Math. Chem.* **1991**, *7*, 243–272.
- (34) Toropov, A. A.; Toropova, A. P.; Ismailov, T. T.; Voropaeva, N. L.; Ruban, I. N. Extended Molecular Connectivity: Prediction of Boiling Points of Alkanes. *J. Struct. Chem. (Engl. Transl.)* **1997**, *38*, 965–969.
- (35) Basak, S. C.; Grunwald, G. D.; Niemi, G. J. In *From Chemical Topology to Three-Dimensional Geometry*; Balaban, A. T., Ed.; Plenum: New York, 1997; Chapter 4, p 78.
- (36) (a) Estrada, E.; Guevara, N.; Gutman, I. Extension of Edge Connectivity Index. Relationship to Line Graph Indices and QSPR Applications. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 428–431. (b) Estrada, E. Generalized Spectral Moments of the Iterated Line Graphs Sequence. A Novel Approach to QSPR Studies. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 90–95.

- (37) Estrada, E.; Ivanciuc, O.; Gutman, I.; Gutierrez, A.; Rodriguez, L. Extended Wiener Indices. A New Set of Descriptors for Quantitative Structure–Property Studies. *New J. Chem.* **1998**, 819–822.
- (38) Yang, Y.-Q.; Xu, L.; Hu, C.-Y. Extended Adjacency Matrix Indices and Their Applications. *J. Chem. Inf. Comput. Sci.* **1994**, 34, 1140–1145.
- (39) Liu, S.; Cao, C.; Li, Z. Approach to Estimation and Prediction for Normal Boiling Point (NBP) of Alkanes Based on a Novel Molecular Distance-Edge (MDE) Vector, λ . *J. Chem. Inf. Comput. Sci.* **1998**, 38, 387–394.
- (40) Ferris, S. W. *Handbook of Hydrocarbons*; Academic Press: New York, 1955.
- (41) Varushchenko, R. M.; Pashchenko, L. L.; Druzhinina, A. I. Enthalpies of Vaporization of Some Mono- and Polycyclic Hydrocarbons. *Russ. J. Phys. Chem. (Engl. Transl.)* **1996**, 70, 208–211.
- (42) (a) Mizutani, K.; Kawasaki, K.; Hosoya, H. Tables of Non-Adjacent Numbers, Characteristic Polynomials and Topological Indices. I. Tree Graphs. *Nat. Sci. Rep. Ochanomizu Univ.* **1971**, 22, 39–58. (b) Kawasaki, K.; Mizutani, K.; Hosoya, H. Tables of Non-Adjacent Numbers, Characteristic Polynomials and Topological Indices. II. Mono- and Bicyclic Graphs. *Nat. Sci. Rep. Ochanomizu Univ.* **1971**, 22, 181–214.
- (43) (a) Manoharan, M.; Balakrishnarajan, M. M.; Venuvanalingam, P.; Balasubramanian, K. Topological Resonance Energy Predictions of the Stability of Fullerene Clusters. *Chem. Phys. Lett.* **1994**, 222, 95–100. (b) Balasubramanian, K. Exhaustive Generation and Analytical Expressions of Matching Polynomials of Fullerenes C_{20} – C_{50} . *J. Chem. Inf. Comput. Sci.* **1994**, 34, 421–427.
- (44) Randić, M.; Wilkins, C. L. Graph Theoretical Ordering of Structures as a Basis for Systematic Searches for Regularities in Molecular Data. *J. Phys. Chem.* **1979**, 83, 1525–1540.

CI9900175