J. Chem. Inf. Comput. Sci. **2001,** *41,* 1521−1530

QSPR Analysis of Flash Points

Alan R. Katritzky,*[,||] Ruslan Petrukhin,[||] Ritu Jain,[||] and Mati Karelson[§]

Center for Heterocyclic Compounds, Department of Chemistry, University of Florida, P.O. Box 11720,
Gainesville, Florida 32611-7200, and Department of Chemistry, University of Tartu,
Jakobi St. 2, Tartu 51014, Estonia

Received April 28, 2001

A quantitative structure property relationship study of the flash point of a diverse set of 271 compounds provided a general three-parameter QSPR model ($R^2 = 0.9020$, $R^2_{\mathrm{cv}} = 0.8985$, $s = 16.1$). Use of the experimental boiling point as a descriptor gives a three-descriptor equation with $R^2 = 0.9529$. Use of the boiling point predicted by a four-parameter reported relationship gives a three-parameter flash point equation with a R^2 value of 0.9247.

INTRODUCTION

Flash point (FP) is one of the most widely used, important characteristics of the flammability properties of liquids and low-melting substances. It provides a simple, convenient index of the flammability and combustibility of substances and is of importance, since it gives the knowledge needed for the handling and transporting of the compound in bulk quantities. It is a subject of interest in terms of understanding the fundamental chemical and physical processes in combustion chemistry. It has also received attention in recent years from the point of view of safety.

The FP is defined as the lowest temperature at which a substance will ignite in air with an initiator (spark or flame). The combustible substance reacts with oxygen in the air in an exothermic oxidation reaction. The flash point occurs when the rate of heat evolved by these reactions is equal to or higher than the rate at which heat is lost to the surroundings. There are two standard methods for measuring FP: open cup and closed cup. The error in the measurement of FP by either method may be in the range of 5−8 °C.[1,2] The open-cup method gives a somewhat arbitrary value because of the unpredictable rate of mass transfer between the liquid and the surrounding atmosphere, but still provides the best match to reality. The closed-cup method produces the most consistent results, because the FP depends only on the vapor pressure and the heat effect of the initial oxidation. There is a very strong relationship between these values and the lower concentration limit of flammability.[2] Closed cup flash points are commonly used in determining the classification of liquids which flash in the ordinary temperature range, but for materials which have relatively high flash points, the open cup flash point is preferred.[3]

Experimental flash point data are desirable, but due to the advancement of technology in discovery or synthesis of new compounds, there is often a significant gap between the demand for such data and their availability. Moreover, for some toxic, explosive, or radioactive compounds the experimental determination of the flash point is extremely difficult. Hence a reliable theoretical method for estimating the flash point is desired.

Flash points are frequently not reported in the literature even for industrially important materials. When the flash points are listed, the reference source of the information is often not provided, leading to uncertainty as to whether the value was measured experimentally or estimated via one of the several prediction methods. The increasing importance of safety related issues in the chemical industry dictates that attention is devoted in reporting material safety data to the experimental method used and its accuracy and repeatability; standards are needed for all physical property data used in the design and construction of chemical processes and plants.[4]

Of the some 20 million known substances, FP values are only recorded for a few thousand. In attempts to solve problems in classifying substances on the basis of their flammability, several methods have previously been developed for the prediction of FP. Empirical group-additive methods with errors of approximately 10 °C,[2] suffer from the usual limitations of such methods. More recent methods are based on functions of boiling point. These methods are sometimes rather complicated[5−7] but can give small errors. The best published equations for the estimation of the flash point appear to contain an exponential function. For example eq 1 relates the flash point to the boiling point,[6] where T_{f} denotes the flash point temperature (K); T_{b} the normal boiling temperature (K); and a, b, c are constants. Equation 1 fits the flash point of 1200 organic compounds and 21 petroleum fractions as a function of normal boiling temperature with an average absolute error of less than 1%.

$$T_{\mathrm{f}} = a + b(c/T_{\mathrm{b}})e^{-c/T}/(1-e^{-c/T})^2 \qquad (1)$$

In the present work, we have developed a method for the prediction of FP based on the multi-parameter regression methodology of the CODESSA computer program,[8] which has eliminated the need for any experimental data input and therefore can be used for the prediction of the flash point of unavailable or unknown compounds.

---

* Corresponding author phone: (352)392-0554; fax: (352)392-9199; e-mail: katritzky@chem.ufl.edu.

[||] University of Florida.

[§] University of Tartu.

10.1021/ci010043e CCC: $20.00

METHODOLOGY

The experimental flash point data for 271 out of the dataset of 298 compounds previously used in our group for the prediction of boiling points[1] was collected from the Acros[9] and Aldrich[10] catalogs [out of these 298 compounds, 27 compounds have no literature flash point data available (many of these are halogenated compounds, which do not have flash points).] All structures were entered into the MDL ISIS[11] chemical database (Table 1). Structures were optimized using the MOPAC[12] quantum chemical package. The molecular descriptors were then calculated using CODESSA, and correlations were calculated, then analyzed, using the CODESSA software package.[8]

RESULTS AND DISCUSSION

To perform a complete search for the best multilinear correlations with a multitude of descriptors, the heuristic method was employed in the present study to pick out the best descriptor combinations to explain the flash point. The heuristic method of the descriptor selection (HM) proceeds with a preselection of descriptors by eliminating (i) those descriptors that are not available for each structure, (ii) descriptors having a small variation in magnitude for all structures, (iii) descriptors that give a *F*-test's value below 1.0 in the one-parameter correlation, and (iv) descriptors whose *t*-values are less than the user-specified value, etc.

This procedure orders the descriptors by decreasing correlation coefficient when used in one-parameter correlations. The next step involves correlation of the given property with (i) the top descriptor in the above list with each of the remaining descriptors and (ii) the next one with each of the remaining descriptors, etc.

The best pairs, as evidenced by the highest *F*-values in the two-parameter correlations, are chosen and used for further inclusion of descriptors in a similar manner.

The Heuristic method usually produces correlations 2−5 times faster than other methods, with comparable quality. The rapidity of calculations from the Heuristic method render it the first method of choice in practical research. Thus, we have used this method for our calculations.

Correlation of Experimental Flash Point with Theoretical Descriptors. Treatment of the dataset by a heuristic selection method[8] afforded the three-parameter correlation given in eq 2-3, Table 2 and Figure 1.

$$T_{FP} = (44.50 \pm 0.99)\sqrt[3]{G_b} + (16731 \pm 574.45)HDCA + (4.95 \pm 0.73)\,M_R - (117.70 \pm 9.14) \quad (2\text{-}3)$$

$$R^2 = 0.9020,\ R^2_{\mathrm{cv}} = 0.8985,\ F = 819.60,\ s = 16.1$$

Here $\sqrt[3]{G_b}$ is the gravitational index over all bonded atoms; *HDCA* is a hydrogen donor charged solvent accessible surface area; and M_R is the relative molecular weight. All descriptor values were obtained by the AM1 parametrization of the MOPAC computer program.

The gravitational index accounts both for the atomic masses (volumes) and for their distribution within the molecular space and seems to quantify effectively the bulk cohesiveness of compound arising from the dispersion and hydrophobic interactions. *HDCA* is a hydrogen donor charged solvent accessible surface area, and this descriptor represents the sum of solvent-accessible surface area of the H-bonding donor atoms. Both descriptors have explicit physical meaning; the first is connected with the dispersion and cavity formation effects in liquids (gravitation index) and the second with the hydrogen-bonding ability of compounds.

The gravitational index over all bonded atoms *i*, *j* in the molecule is defined as

$$G_I = \sum_{ij} (m_i m_j / r_{ij}^{\,2})$$

(where m_i and m_j are the atomic masses of the bonded atoms and r_{ij} denotes the respective bond lengths).[8] The range value of this descriptor for the present set is from 6.6 to 12.4.

HDCA accounts for the hydrogen bonding donor ability of the molecule and is calculated as the sum of the solvent accessible surface areas of the hydrogen atoms as the possible hydrogen donors.[13] The hydrogen atoms directly connected with the electronegative atom in the molecule are considered as possible hydrogen-bonding donors. For our dataset, the maximum value of *HDCA* is 11.78, and the minimum value is 0.

M_R represents the molecular weight divided by the number of atoms in the molecule.[13]

Correlation of Experimental Flash Point with Experimental Boiling Point as a Descriptor. Using the heuristic method, but including the experimental boiling point as one of the descriptors, we obtained a three-parameter correlation given in eq 3-3, Table 3, Figure 2, where T_{FP} is the temperature of a flash, T_{BP} is experimental temperature of boiling, *RNCG* is relative negative [surface] charge and characterizes the dispersion of partial charges in the molecule, *HDCA* is H-donors charged surface area and represents the sum of solvent-accessible surface area of H-bonding donor atoms.

$$T_{FP} = (0.72 \pm 0.014)T_{BP} + (76.99 \pm 11.97)RNCG + (2.05 \pm 0.33)HDCA - (8.40 \pm 7.63) \quad (3\text{-}3)$$

$$R^2 = 0.9529,\ R^2_{\mathrm{cv}} = 0.9513,\ F = 1798.73,\ s = 11.2$$

RNCG is defined as the ratio of maximum (by absolute value) atomic partial negative surface charge and the sum of similar negative charges in the molecule.[13] The *RNCG* values lie in the range of 0.06−0.50.

Correlation of Experimental Flash Point with Predicted Boiling Points as a Descriptor. Reported two-, four-,[1] and eight-parameter[14] boiling point equations for the correlations of the flash point are capable of allowing the prediction of the boiling point of structurally diverse organic compounds, with average errors of 16.2, 12.4, and 15.5 K. The data set used for two- and four-parameter equations consisted of 298 structurally diverse compounds and contains saturated and unsaturated hydrocarbons, halogenated compounds, and hydroxyl, cyano, amino, ester, ether, carbonyl, and carboxyl functionalities. However, the eight-parameter equation was calculated using 584 datapoints and is a representative for all major classes of compounds containing C, H, O, N, S, F, Cl, Br, and I. These models are theoretically justified and provide significant additional insight into the relationship between the structure and the boiling point of the compounds.

Table 1. Database Listing

no.	name	T_{FP} (exp)	T_{FP} (pred eq 2-3)	T_{FP} (pred, eq 3-3)	T_{FP} (pred, eq 4-3)	T_{FP} (pred, eq 5-3)	T_{FP} (pred, eq 6-3)	T_{FP} (pred, eq 7-3)
1	acrylonitrile	268	282.8	273.7	265.4	272.1	284.9	297.6
2	acrolein	247	280.1	258.2	246.5	262.8	275.5	269
3	acrylic acid	321	348.5	331.4	342.9	334	344.8	335.6
4	3-chloropropene	244	233.2	246.5	251.2	231.6	238.8	246.8
5	allyl alcohol	294	285.4	304.9	288.9	283.9	293.5	280.7
6	propionaldehyde	233	252.2	253.3	242	258.5	251.9	246.2
7	ethyl formate	253	255.4	254.8	258.3	242.3	252	245.2
8	methyl acetate	263	255.8	259	255.5	237.7	249.4	245.6
9	propionic acid	324	335.4	329.5	331.2	329.9	336.2	328.6
10	1-bromopropane	298	279.3	265.4	248.3	264.7	260.2	247.1
11	2-chloropropane	241	225	241	245.3	228.7	235.4	245.7
12	1-chloropropane	255	225.9	248.1	245.8	230.1	236.8	246.6
13	allylamine	245	259.7	258.2	242.3	250.8	254	257.1
14	methyl acetate	264	214.8	230.1	224.4	214.7	229.5	211.3
15	1-propanol	288	279.2	304.9	278.4	281.3	289.5	278.4
16	dimethoxymethane	255	254.6	248.6	269.2	249.1	263.3	249
17	1,2-propanediol	380	379.3	371.7	369.6	369.3	381.3	378.9
18	isopropylamine	236	250	242.5	231.7	245.8	250	250.3
19	propylamine	247	256.1	256.7	234.6	249.2	259.2	256.7
20	trimethylamine	267	239.7	223.3	236.9	259.5	225.3	239.6
21	1,3-butadiene	197	203.8	208.4	202.5	209	204.3	197.6
22	isobutenal	273	284	264.9	267	273.8	279.6	276.7
23	crotonic acid	361	333.4	350.8	342.4	347.7	336.6	324.5
24	methyl acrylate[a]	270	277.9	273	277.9	256.7	267.9	266.6
25	vinyl acetate[b]	265	278.4	264.5	278.1	258.4	266	267.2
26	*n*-butyronitrile	289	268.3	299	284.9	274.6	272.6	292.4
27	butyraldehyde	267	268.8	267.2	262.6	269.2	268.4	265.1
28	isobutyraldehyde	249	272.3	259.7	261.9	274	263	268.7
29	2-butanone	266	272.8	268.4	258.2	277	266.3	269.3
30	tetrahydrofuran	252	251.1	266.2	242	244.3	260.1	246.8
31	*n*-butyric acid	342	325.5	333.9	332.8	344.8	335.1	320.7
32	ethyl acetate	269	273.2	268.4	270.4	256.9	264.8	266.3
33	isobutyric acid	328	340	334.9	331	339.7	346.5	335.9
34	methyl propionate	279	273.2	270.9	270.5	257.4	265.9	266.3
35	propyl formate	273	273	272.8	282.3	257.2	267.2	266.1
36	1-bromobutane	296	291.9	282.1	268.6	284.2	275	267.7
37	1-chlorobutane	267	246.7	266	270.8	252.5	256.5	271.3
38	1-butanol	308	297.4	314.2	302.8	304.1	305.1	298.1
39	*sec*-butanol	299	291.8	297	286.1	292.3	292.4	292.2
40	*tert*-butyl alcohol	277	287.8	281	288.3	290.3	286.5	288.1
41	ether	233	238	241.2	243.8	237.2	244	236.2
42	2-methyl-1-propanol	301	298.6	306.6	289.2	297	302.6	299.3
43	1,3-butanediol	394	376.6	381.2	390.3	389.1	378.5	377.5
44	*n*-butylamine	266	274.6	272	259.3	266.6	273.4	276.2
45	diethylamine	234	271.1	254.8	253.8	265.2	273	272.5
46	isobutylamine	264	264.6	263.2	255.4	260.1	272.4	265.8
47	pyridine	290	277.8	290.2	298.6	270	274.9	266.4
48	cyclopentene	273	238.6	237.2	242.4	244.1	239.6	235.2
49	2-methyl-1,3-butadiene	225	228.8	232.9	230	232.9	228.4	224.8
50	2,4-pentanedione	307	343.1	305.5	325.8	336.2	329	337.4
51	allyl acetate	279	293.3	285.1	290.2	274.8	283.3	285.5
52	ethyl acrylate	289	293	282.6	292.2	274.4	282.2	285.3
53	cyclopentane	236	235.4	239.2	236.7	245.9	239	234.8
54	1-pentene	273	225.2	229.6	227.1	231.6	225.9	224
55	3-pentanone	286	283.2	281.5	275.1	287.3	277.1	281.2
56	2-pentanone	280	288.9	281.8	278.5	291.2	281.7	287.2
57	valeraldehyde	285	286.7	283.8	283.2	287	282.4	284.9
58	ethyl propionate	285	289.4	281.8	284.3	274.9	281.5	284.9
59	isobutyl formate	283	289.1	279	284.9	279.5	283.3	284.6
60	*n*-propyl acetate	283	289.4	283.3	285.6	275	283.9	284.9
61	1-butanecarboxylic acid	359	352.9	356	350.3	357.4	354.7	351.1
62	1-chloropentane	314	265.6	283.7	288.5	273	274.5	293.1
63	2-methylbutane	222	222.1	226.7	219.6	233	224.8	223.1
64	pentane	224	222.4	232.3	219.3	233.2	224.6	223.4
65	2-methyl-1-butanol	316	290.5	308.9	303.9	294.5	292.9	291.4
66	*tert*-amyl alcohol (2-methyl-2-butanol)	293	305.1	292.1	302.8	306.6	302	306.6
67	3-methyl-1-butanol	316	310.2	318.6	306.9	311.8	314	311.9
68	DL-3-methyl-2-butanol	312	307.4	301.4	306.2	309.5	307.8	309
69	1-pentanol	305	312.6	324	317	318.9	315.7	314.5
70	DL-2-pentanol	313	305.9	306.3	308.9	309.7	306.8	307.5
71	3-pentanol	313	306	303.8	309.2	310.7	308.5	307.6
72	2,2-dimethyl-1,3-propanediol	380	368.2	373.3	382.9	377.6	373.1	369.8
73	1,5-pentanediol	402	387.4	403.1	407.6	405.8	399	389.8
74	1,2-dichlorobenzene	340	352.4	333.8	351.6	341.7	360.1	347.8
75	bromobenzene	324	354.5	317.8	323	330.7	327.9	316.6

Table 1. (Continued)

no.	name	T_{FP} (exp)	T_{FP} (pred eq 2-3)	T_{FP} (pred, eq 3-3)	T_{FP} (pred, eq 4-3)	T_{FP} (pred, eq 5-3)	T_{FP} (pred, eq 6-3)	T_{FP} (pred, eq 7-3)
76	chlorobenzene	302	312.5	297.3	327.9	306.9	314.7	323.2
77	benzene	262	268.8	258.8	268.3	269.8	265.7	260.5
78	phenol	352	357	349.4	352.3	347	356.5	349
79	hydroquinone	438	434.1	431.6	429.2	428.4	433.9	426.5
80	aniline	343	344.8	344.7	320	312.4	336.2	339.3
81	2-picoline	301	322.7	298.5	314.5	327	320.2	316.3
82	1,4-dicyanobutane	432	354.3	415.2	408.7	360.7	402.2	406.6
83	cyclohexene	253	260.2	262.3	265.5	266.3	261.6	258.6
84	cyclohexanone	319	319	319.8	312.3	315.6	310.6	316.9
85	mesityl oxide	305	306.3	300.9	301.9	297.8	303.7	303.6
86	cyclohexane	255	256.8	259.3	258.8	267.1	260.1	257.5
87	2,3-dimethyl-1-butene	255	247.1	246.1	248.4	252.7	247.8	247.3
88	2,3-dimethyl-2-butene	257	248.4	255.4	249.2	258.3	250	248.5
89	2-ethyl-1-butene	247	247.4	252.2	249.1	253.2	247.6	247.5
90	1-hexene	247	247.2	250.6	248.9	253.1	247	247.3
91	2-hexene	252	247.8	253.8	249.7	254.9	247.9	247.9
92	*trans*-2-hexene	253	247.8	253.1	249.8	254.9	248	247.9
93	methylcyclopentane	263	256.1	256.8	257.5	262.1	259.3	256.7
94	2-methyl-1-pentene	247	247.4	250.4	248.9	253	247.4	247.4
95	4-methyl-1-pentene	242	246.9	244	249.3	252.7	246.7	247
96	*n*-butyl vinyl ether	264	279.7	275.1	285.9	276.1	278.5	278.5
97	cyclohexanol	340	335.2	335.3	337.7	332.2	343.4	336.2
98	2-hexanone	296	301.8	297.6	297	304.1	292.6	301.5
99	4-methyl-2-pentanone	287	298.9	289.7	296.9	299.3	292.2	298.5
100	ethyl butyrate	299	304.6	295.4	298.9	291.6	296.5	302.1
101	2-ethylbutyric acid	360	365.3	360.1	355.9	369.2	365.3	365.3
102	hexanoic acid	377	363	368.7	362.4	371.7	364.9	362.9
103	isobutyl acetate	291	304.5	292.1	300	291.4	296.3	301.9
104	cyclohexylamine	300	318	306.8	308.7	308.4	314.9	320.1
105	2,2-dimethylbutane	225	244	239.9	242.4	254	245.8	245.9
106	2,3-dimethylbutane	244	244.3	245.6	242.5	254.3	245.8	246.1
107	*n*-hexane	251	244.8	253.3	243.3	254.5	245.9	246.6
108	2-methylpentane	250	244.5	247.2	242.9	254.2	245.6	246.3
109	3-methylpentane	267	244.5	249.6	243.2	254.4	246.3	246.4
110	isopropyl ether	244	276.1	255.8	278.6	275.2	274.7	276.8
111	*n*-propyl ether	245	276.7	273.2	279.4	275.7	278.3	277.5
112	acetal	252	305.1	280.1	307.4	299.5	302	304.9
113	diisopropylamine	266	301.8	267.3	289.5	296.6	297.6	305
114	dipropylamine	276	304.8	286.3	293.3	299.2	302.8	308.1
115	benzaldehyde	335	337.8	339	340.8	329.7	330.6	328.5
116	benzoic acid	394	403	402.4	399.7	393.5	402.2	392.8
117	toluene	277	285.7	283.3	286.9	284.2	285.1	280.2
118	benzyl alcohol	367	364.3	371	366.1	352.9	366.5	359.3
119	*m*-cresol	359	368.7	362.7	370	364.8	369.7	364
120	*o*-cresol	354	369.1	354.8	364.3	361.5	368.3	364.3
121	m-toluidine	359	355	355.9	337.8	326.3	347.6	352
122	o-toluidine	358	346.5	353.9	334.6	320.2	339.6	343.2
123	butyl acrylate	312	321.4	312.8	319	306.9	313	318.2
124	ethyl cyclopentane	288	274.7	277.6	277.3	280.3	277.3	276.3
125	2,3-dimethylpentane	267	264.1	266.7	263.3	273.6	265.3	266.8
126	heptane	269	264.6	272.6	263.5	273.9	265.4	267.3
127	2-methylhexane	270	264.3	266.6	262.9	273.7	265.1	267
128	3-methylhexane	270	264.4	268.1	263.5	273.8	265.3	267.1
129	styrene	305	304.4	308.8	309	298.3	302.1	298.1
130	acetophenone	350	354.6	351.7	349.9	344.4	345.9	347.8
131	ethylbenzene	288	301	301.8	304.4	298.5	300.8	297.7
132	*m*-xylene	298	301.4	302	304.5	301.6	302	298.1
133	*p*-xylene	300	301.4	301.4	304	301.7	302.1	298.1
134	2,6-dimethylphenol	346	367.7	356.4	365.9	360.7	367.7	364.8
135	*n*,*n*-dimethylaniline	336	344.9	344.3	347.3	347.8	344.3	342.9
136	1,2-dimethylcyclohexane	288	291.9	294.9	294.1	297.9	294.5	294.5
137	*trans*-1,2-dimethylcyclohexane	279	291.8	290.2	294	297.8	294.3	294.3
138	*cis*-1,3-dimethylcyclohexane	278	292	287.7	293.8	298	294.4	294.5
139	*cis*-1,2-dimethylcyclohexane	285	291.7	291.1	294.2	297.8	294.4	294.3
140	1,4-dimethylcyclohexane	283	292	290.9	293.9	298	294.4	294.5
141	*cis*-1,4-dimethylcyclohexane	279	291.8	287.3	294	297.9	294.3	294.4
142	ethylcyclohexane	291	291.9	296.4	294.5	297.7	294.2	294.5
143	*n*-propylcyclohexane	308	291.6	295.6	294	297.4	293.7	294.2
144	2,4,4-trimethyl-1-pentene	267	283.8	274.9	287.4	289.6	283.8	286
145	2,4,4-trimethyl-2-pentene	275	284.2	276.3	288.1	292.1	285.1	286.3
146	isobutyl isobutyrate	313	332	310.7	322.7	322.1	323.8	332.3
147	2,3-dimethylhexane	270	282	283.6	280.9	291.5	283.1	285.5
148	octane	286	282.5	290.7	281.9	291.9	318.8	286.1
149	2,2,4-trimethylpentane	261	281.5	272	282	291.3	283.2	285
150	di-*n*-butyl ether	298	308.8	305.1	309.8	309.2	308.9	311.6

Table 1. (Continued)

no.	name	T_{FP} (exp)	T_{FP} (pred eq 2-3)	T_{FP} (pred, eq 3-3)	T_{FP} (pred, eq 4-3)	T_{FP} (pred, eq 5-3)	T_{FP} (pred, eq 6-3)	T_{FP} (pred, eq 7-3)
151	2-ethyl-1-hexanol	350	352	350.5	356.6	361.1	351.8	356.6
152	quinoline	374	351.1	373.7	376	340.2	349.4	341.7
153	alpha-methylstyrene	318	318.4	322	325.9	313.9	317.3	314.4
154	ethyl benzoate	357	361.7	360.1	358.9	340.8	350.8	355.2
155	cumene	304	315.2	311.3	320.2	314.7	314.9	313.6
156	*o*-ethyltoluene	312	315.7	320.9	320.8	315	316.3	314.1
157	*p*-ethyltoluene	309	315.8	318.7	320.7	315	316.2	314.2
158	mesitylene	317	316.1	319.1	319.9	317.9	317.7	314.6
159	propylbenzene	320	315.4	316.4	320	314.4	315.4	313.9
160	1,2,3-trimethylbenzene	321	316	327.4	320	317.6	318.6	314.4
161	1,2,4-trimethylbenzene	321	316.1	322.4	319.4	317.9	317.5	314.6
162	isophorone	357	362.3	357.9	360	351.3	356.1	362.8
163	*n*-propylcyclohexane	308	307.5	312.9	309.9	313.9	310	311
164	*n*-nonane	304	298.7	307.5	298.3	308.6	299.2	303
165	1,2,3,4-tetrahydronaphthalene	350	338.3	347.7	346.1	339.8	340.8	336.5
166	*n*-butylbenzene	332	329	332.2	335.4	329.6	329.5	328.9
167	*tert*-butylbenzene	317	328.3	321.9	334.5	329.6	329.2	328.1
168	*p*-cymene	320	329.1	327.9	335.4	330.2	329.8	329
169	*m*-diethylbenzene	329	329.3	331	335.1	329.9	330.3	329.2
170	*o*-diethyl benzene	322	329.2	332.6	334.3	329.8	330.2	329.1
171	isobutylbenzene	328	328.7	324.5	334.5	329.6	329.1	328.6
172	*n*-butylcyclohexane	314	321.9	329.1	324.3	329.3	324.6	326.2
173	decane	319	313.8	323.3	313.1	324.3	311.5	318.8
174	2-ethylhexyl acrylate	359	369.6	356.8	366.7	362.7	362.4	371.8
175	phenyl ether	388	389.7	383.9	399.2	383.1	387.7	383.5
176	diphenylamine	425	419.8	424.6	417.1	406.6	413.7	416.6
177	diethyl phthalate	433	426.3	412.1	414.9	412.2	414.9	422.2
178	1,3-diisopropylbenzene	349	353.8	344.5	361.6	358.2	356.3	356
179	bicyclohexyl	365	355	367.4	360.6	366.9	362	360
180	1-dodecene	350	342.2	350.7	346.1	351.8	343.4	347.6
181	dodecane	344	340.9	352.2	340.1	353.5	344.2	347.3
182	di-*n*-hexyl ether	351	361	361.6	360.3	366.4	361.7	367
183	diphenylmethane	400	380	385.4	393	382.1	382.2	376.8
184	benzyl benzoate	420	424.6	434.8	428.4	406.8	414.1	419.5
185	bibenzyl	402	389.9	396.4	403.7	394.6	393.5	388.3
186	1-tetradecene	388	366.2	376.6	369.9	378.8	369.6	373
187	1-tetradecane	372	366.2	378.4	370.1	378.7	369.2	373
188	dibutyl phthalate	444	457.6	442.5	442.6	455.7	453.7	459.3
189	1-hexadecene	405	387.9	400.1	391.2	403.7	392.7	395.9
190	dibutyl sebacate	440	463.1	446.1	441.1	477.9	464.6	470.5
191	1-octadecene	421	407.8	420.9	410.6	427.3	414.2	416.9
192	stearic acid	469	407.8	464.4	410.6	427.3	454.7	416.9
193	octadecane	438	406.8	421.9	405.1	428.9	414.9	416.6
194	nonadecane	441	416.3	431.1	413.9	440.3	425.2	426.6
195	methyl isopropenyl ketone	286	296.8	279.4	282.9	292.2	286.5	292.2
196	1,1-dichloropropane	280	276.4	279.8	271.7	271.2	292.2	276.5
197	1,3-dichloropropane	293	277.1	296.6	288.8	275.3	288.1	283.7
198	1,3-propanediol	404	380.3	396	388	387.3	388.6	379.9
199	isobutyronitrile	281	269.9	288.7	283	276.4	274.1	294
200	1,4-dichlorobutane	313	290	314.9	305.3	294.9	302.7	304
201	2-bromobutane	294	290.3	274.6	265.2	283	273.7	266.1
202	2-chlorobutane	258	245.9	258.8	267.3	251.3	255.2	270.4
203	1,4-butanediol	394	387.5	401.7	385.5	391.3	391.6	388.8
204	2,3-butanediol	358	367.5	356.8	369.2	370.7	378.2	368
205	*sec*-butylamine	254	263.6	259.6	254.5	258.4	262.3	264.7
206	*tert*-butylamine	264	266.2	246.1	251.1	262.2	265.2	267.4
207	valeronitrile	313	284.8	310.8	307.4	290.9	285.2	310.9
208	2-methyl-1-butene	239	225.4	231.3	225.6	231.9	226.8	224.3
209	2-methyl-2-butene	228	226.2	233	226.3	236.4	227.8	225.1
210	3-methyl-1-butene	217	224.8	222.6	226	231.5	226.1	223.6
211	1,5-dichloropentane	299	303.7	331.5	319.5	312.4	319.3	322.7
212	3-methyl-2-butanone	279	288.6	276.1	276.9	292.4	281.4	286.9
213	isopropyl acetate	284	289.4	272.4	285.5	274.5	278.8	284.9
214	dl-2-methylbutyric acid	350	354.4	350.2	345.2	357.3	357.2	352.7
215	methyl butyrate	285	289.6	284.4	287.3	275.4	281.8	285.1
216	piperidine hydrochloride	338	304.8	286.3	293.5	298.6	304.5	305.9
217	2,2-dimethyl-1-propanol	309	312.5	306.7	301.7	312.1	314.3	314.3
218	*tert*-butyl methyl ether	245	257.8	248.9	262.1	257.1	257.9	257.3
219	methyl *sec*-butyl ether	243	258.3	254.4	264.4	257.4	262.6	257.9
220	*n*-amylamine	277	292.9	287.6	281.3	284.5	289.3	295.4
221	3-picoline	313	319.9	308.1	316.7	328.4	317.5	313.4
222	1,3-cyclohexadiene	267	263.9	260.2	269.5	267.4	263.5	259.4
223	methylcyclopentadiene	299	261.8	256.4	270.3	263	260.2	257.2
224	2,3-dimethyl-1,3-butadiene	251	250.4	255.4	254.4	254.1	250.1	248.3
225	1,5-hexadiene	246	249.9	248.1	256.4	252.9	247.5	247.8

Table 1. (Continued)

no.	name	T_{FP} (exp)	T_{FP} (pred eq 2-3)	T_{FP} (pred, eq 3-3)	T_{FP} (pred, eq 4-3)	T_{FP} (pred, eq 5-3)	T_{FP} (pred, eq 6-3)	T_{FP} (pred, eq 7-3)
226	hexanenitrile	316	300.1	323	326.2	306.7	299.8	328
227	3,3-dimethyl-1-butene	245	246.3	235.1	249.5	252.5	246.5	246.3
228	2-methyl-2-pentene	250	248.1	253	248.9	255.4	249	248.2
229	3-methyl-1-pentene	245	246.9	244.3	249.1	252.7	247.2	246.9
230	hexanal	305	300.4	299.2	304.6	300.9	294.9	300
231	3-hexanone	287	303.3	294.6	295.4	305.7	295.3	303
232	*sec*-butyl acetate	289	304.4	287.7	298.2	291.2	294.7	301.8
233	*tert*-butyl acetate	288	304.1	275.3	299.5	290.6	293	301.5
234	ethyl isobutyrate	286	304.4	287.3	297.6	291.4	296.2	301.8
235	ethyl *n*-butyrate	292	304.7	296.1	298	291.7	296.3	302.1
236	butyl ethyl ether	268	276.8	274.2	279.9	276	277.5	277.5
237	2-ethyl-1-butanol	330	302.4	317.7	316	305	309	304.2
238	hexyl alcohol	333	325.5	334.5	334.9	334.9	328.1	328.4
239	2-hexanol	314	319.7	318.6	327.4	326.8	319.7	322.3
240	2-methyl-1-pentanol	323	328.7	328.9	320.5	329.4	327.3	331.6
241	4-methyl-2-pentanol	314	320.3	312.8	314.4	321.1	318.8	322.9
242	methyl tert-pentyl ether	262	276.1	268.8	279.8	275.3	275	276.9
243	1,6-hexanediol	420	393.3	406.2	412.7	413.7	397.3	396.8
244	hexylamine	281	309.9	304	298.4	301.7	304.4	313.4
245	acetone	253	254.5	255.8	236.2	258.9	249.4	248.5
246	2-propanol	285	274.5	289.3	274.7	276.5	280	273.5
247	ethyl vinyl ether	228	242.1	240.5	250.3	237.4	237	237.3
248	butyl formate	291	289.2	285.2	289.1	279.9	284.7	284.7
249	isovaleric acid	343	354.6	347.5	349	352.7	354.4	352.9
250	*n*-butyl acetate	295	304.7	298.5	300.2	291.7	296.2	302.1
251	hexamethyleneimine	291	318.7	302.7	310.6	311.2	317.7	320.8
252	*p*-cresol	362	364.7	361.9	365.3	360.2	365.2	359.8
253	*p*-toluidine	360	354.1	353.9	335.7	325.8	346.7	351
254	1-heptene	265	266.6	270.3	269.3	272.3	264.5	267.8
255	methylcyclohexane	270	275.2	275.8	277.8	280.9	278.2	276.9
256	2,2,3-trimethylbutane	267	263.5	260.2	262.8	273.3	264.8	266.2
257	*o*-xylene	305	301.4	305.8	304.2	301.7	301.8	298.1
258	1-octene	286	284.3	288.4	287.4	290.4	283.9	286.5
259	benzyl acetate	375	361.8	359.6	360.5	341.2	350.6	355.4
260	*m*-ethyltoluene	311	315.8	318.2	320.6	315	316.3	314.2
261	dimethyl phthalate	419	410.4	406.6	402.3	389.8	399.2	401.8
262	*sec*-butylbenzene	318	328.8	325.2	335.6	329.3	329.3	328.6
263	1,4-diethylbenzene	329	329.3	332.9	335	329.9	330.3	329.2
264	1-decene	329	315.3	321.5	318.8	322.7	312.9	319.2
265	decyl alcohol	355	372.1	379.9	386.8	389.5	373.4	378.1
266	1,4-diisopropylbenzene	349	353.8	349.8	361.5	358.3	356.5	356
267	hexadecane	408	386.9	401	385.6	405.4	393.2	395.8
268	propionitrile	279	253.9	292.3	260.1	262.1	263.8	275
269	2-bromopropane	274	277.7	256.9	245.5	263.6	258.4	245.4
270	2-chloro-2-methylpropane	291	245.1	245.8	267.6	251	254	269.5
271	4-picoline	313	319.89	309.5	315.8	326.3	317.4	313.3

[a] Stabilized with methyl ethyl hydroquinone. [b]1 stabilized with 8 to 12 ppm hydroquinone.

Table 2. Theoretical Flash Point Correlations [Eqs 2-1, 2-2, and 2-3]

	R^2	R^2_{cv}	F	s^2		X	DX	*t*-test	descriptor
1	0.5902	0.5792	387.45	1076.34	0	197.91	5.76	34.34	1
					1	0.148	0.008	19.68	2
2	0.8853	0.8826	1034.17	302.42	0	−95.38	9.20	−10.4	1
					1	45.19	1.06	42.63	3
					2	16905	619.86	27.27	4
3	0.9020	0.8985	819.60	259.21	0	−117.7	9.14	−12.9	1
					1	4.45	0.99	45.09	3
					2	16731	574.45	29.13	4
					3	4.95	0.73	6.76	5

The predicted value of boiling point from these equations was further used for the correlation of the flash point.

$$T_{BP} = (64.87 \pm 1.46)\sqrt[3]{G_b} + (44.92 \pm 4.46)HACA + (0.26 \pm 0.035)H_f + (563.3 \pm 74.55)FHBCA + (45.95 \pm 7.70)N_b - (152.40 \pm 12.30) \quad (4)$$

$$R^2 = 0.8918,\ R^2_{cv} = 0.8832,\ F = 436.87,\ s = 23.03$$

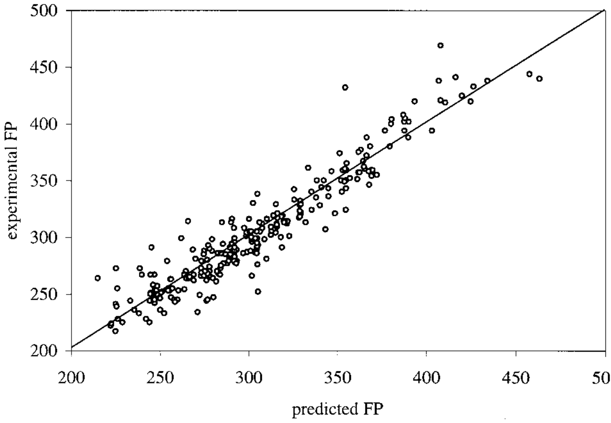

Figure 1. Plot of observed T_{FP} versus predicted T_{FP} using three-parameter correlation [eq 2-3].

The correlation of the experimental boiling point for the present set of compounds (used as a descriptor in eq 3-3) using the heuristic method afforded a five-parameter equation given as eq 4, where $\sqrt[3]{G_b}$ is the cubic root of gravitational

Table 3. FP Correlations Using Experimental BP [Eqs 3-1, 3-2, and 3-3]

	R^2	R^2_{cv}	F	s^2		X	DX	t-test	descriptor
1	0.9218	0.9205	3169.67	205.48	0	23.74	5.06	4.69	1
					1	0.69	0.012	56.30	6
2	0.9465	0.9453	2371.21	141.02	0	28.15	4.21	6.69	1
					1	0.67	0.01	64.37	6
					2	11654	1046.8	11.13	4
3	0.9529	0.9513	1798.73	124.76	0	−8.40	7.63	−1.10	1
					1	0.72	0.013	52.69	6
					2	76.99	11.97	6.43	7
					3	2.05	0.33	6.17	8

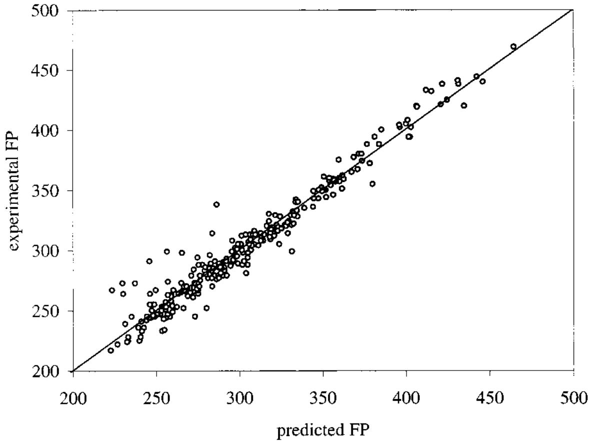

Figure 2. Plot of observed T_{FP} versus predicted T_{FP} using three-parameter correlation [eq 3-3].

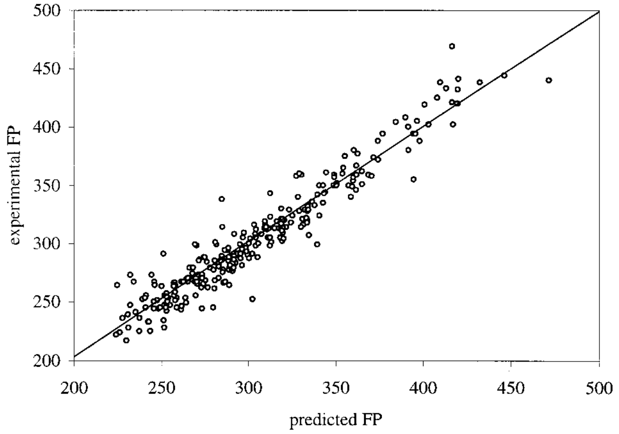

Figure 3. Plot of observed T_{FP} versus predicted T_{FP} using three-parameter correlation [eq 4-3].

index, *HACA* is the hydrogen acceptor charged area, H_f is the final heat of formation, *FHBCA* is the fractional hydrogen bonding surface area relative to the total molecular surface area and is responsible for polar interactions between molecules, and N_b is the number of triple bonds.

HACA is defined as the sum of solvent accessible surface area of hydrogen bonding acceptor atoms in the molecule.[13] The minimum value of *HACA* is 0, and the maximum value is 2.1 for our dataset. The fractional hydrogen bonding charged surface area accounts for the property of the pairs in all components in a mixture.[13] The descriptor values lie in the range of 0−0.14.

$$T_{FP} = (0.67 \pm 0.014)T_{BP,pred} + (3.45 \pm 0.27)DPSA + (0.95 \pm 0.17)E_{e-n,C} - (161.8 \pm 31.07) \quad (4\text{-}3)$$

Table 4. FP Correlations Using BP Predicted by QSPR (Five-Parameter) for the Present Data Set [Eqs 4-1, 4-2, and 4-3]

	R^2	R^2_{cv}	F	s^2		X	DX	t-test	descriptor
1	0.8780	0.8760	1935.83	320.46	0	9.77	6.78	1.44	1
					1	0.73	0.016	44.00	9
2	0.9155	0.9132	1451.43	222.83	0	8.46	5.66	1.50	1
					1	0.67	0.015	45.42	9
					2	2.64	0.24	10.90	10
3	0.9243	0.9217	1086.59	200.34	0	−161.8	31.07	−5.21	1
					1	0.67	0.014	47.72	9
					2	3.45	0.27	12.69	10
					3	0.95	0.17	5.56	11

Table 5. FP Correlations Using Predicted BP (Two-Parameter Equation[1]) [Eqs 5-1, 5-2, and 5-3]

	R^2	R^2_{cv}	F	s^2		X	DX	t-test	descriptor
1	0.8432	0.8407	1446.11	411.96	0	−10.78	8.38	−1.28	1
					1	0.79	0.021	38.03	12
2	0.8946	0.8922	1137.31	277.89	0	0.48	6.95	0.07	1
					1	0.75	0.017	42.95	12
					2	0.39	0.034	11.43	13
3	0.9073	0.9033	871.24	245.27	0	32.85	8.44	3.89	1
					1	0.59	0.03	18.92	12
					2	0.51	0.037	13.57	13
					3	0.28	0.0046	6.05	14

$$R^2 = 0.9243,\ R^2_{cv} = 0.9217,\ F = 1086.59,\ s = 14.15$$

Using the predicted boiling point from eq 4 as a descriptor gives a three-parameter equation [eq 4-3, Table 4 and Figure 3], where $T_{BP,Pred}$ is the predicted boiling point based on the eq 4, and *DPSA* is the difference in charge partial positive surface area of the positive partial charged atoms and for the negative charged atoms in the molecule. The range of *DPSA* is from 3.9 to 24.6. $E_{e-n,C}$ is the minimum electron attraction for a C atom. $E_{e-n,C}$ describes the nuclear-electron attraction-driven processes in the molecule and is related to the conformational (rotational, inversional) changes or atomic reactivity in the molecule. The nuclear-electron attraction energy for a given atomic species (atom C) in the molecule is defined as

$$E_{e-n,C} = \sum_{B} \sum_{\mu,\nu \Sigma A} P_{\mu\nu} \left\langle \mu \middle| \frac{Z_B}{R_{iB}} \middle| \nu \right\rangle$$

Here the first summation is performed over all atomic nuclei in the molecule (B), whereas the second summation is carried out over all atomic orbitals at a given atom (C). The term $\langle \mu | Z_B/R_{iB} | \nu \rangle$ denotes the nuclear-electron attraction integrals on the given atomic basis.[13] The minimum value of $E_{e-n,C}$ is 159.6, and the maximum value is 179.9. *DPSA* is responsible for polar interactions between molecules, whereas the quantum-chemical descriptor $E_{e-n,C}$ can be related to the reactivity of a carbon atom in the molecule in combustion reaction.

$$T_{FP} = (0.59 \pm 0.03)T_{BP,pred} + (0.51 \pm 0.04)HBSA + (0.28 \pm 0.05)M_w + (32.85 \pm 8.44) \quad (5\text{-}3)$$

$$R^2 = 0.9073,\ R^2_{cv} = 0.9033,\ F = 871.24,\ s = 15.66$$

Using the value of boiling point predicted on the basis of the two-parameter equation[1] gives a three-parameter equation [eq 5-3, Table 5 and Figure 4], where $T_{BP,Pred}$ is the predicted boiling point based on the reference,[1] *HBSA* is the hydrogen

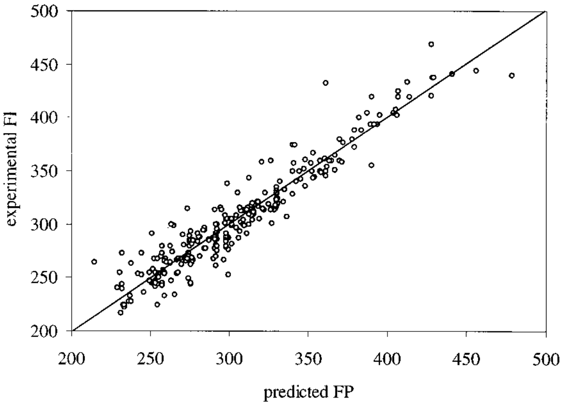

Figure 4. Plot of observed T_{FP} versus predicted T_{FP} using three-parameter correlation [eq 5-3].

Table 6. FP Correlations Using Predicted BP (Four-Parameter Equation[1]) [Eqs 6-1, 6-2, and 6-3]

	R^2	R^2_{cv}	F	s^2		X	DX	*t*-test	descriptor
1	0.8891	0.8873	2156.87	291.26	0	22.36	6.16	3.63	1
					1	0.70	0.001	46.44	15
2	0.9200	0.9181	1540.23	211.01	0	−46.9	8.60	−5.46	1
					1	0.79	0.016	50.57	15
					2	134.9	13.27	10.16	7
3	0.9247	0.9222	1092.99	199.26	0	−26.1	9.77	−2.67	1
					1	0.75	0.018	42.58	15
					2	103.65	14.98	6.92	7
					3	0.23	0.055	4.10	16

bonding surface area and is defined as the difference between the respective hydrogen-bonding donor and hydrogen-accepting descriptors, and M_w is the molecular weight.

$$T_{FP} = (0.75 \pm 0.018)T_{BP} + (0.23 \pm 0.06)HASA + (103.65 \pm 14.98)RNCG - (26.11 \pm 9.77) \quad (6\text{-}3)$$

$$R^2 = 0.9247,\ R^2_{\mathrm{cv}} = 0.9222,\ F = 1092.99,\ s = 14.12$$

Using the predicted value of boiling point from the four-parameter equation[1] as one of the descriptor for the correlation of the flash point gives a three-parameter equation [eq 6-3, Table 6, Figure 5], where T_{FP} is the temperature of a flash, T_{BP} is the predicted value of boiling point on the basis of predicted value of the four-parameter boiling point equation,[1] *HASA* is hydrogen acceptor surface area and is calculated as a sum of exposed surface areas of all possible hydrogen acceptor atoms in the molecule, and *RNCG* is the relative negative charge.

$$T_{FP} = (46.85 \pm 2.27)\sqrt[3]{G_b} + (17451 \pm 563.85)HDCA + (0.003 \pm 0.004)T_{BP} - (310.02 \pm 41.15)\delta_{Cl} + (28.79 \pm 5.04)N_b - (113.6 \pm 8.53) \quad (7\text{-}5)$$

$$R^2 = 0.9140,\ R^2_{\mathrm{cv}} = 0.9079,\ F = 563.53,\ s = 15.14$$

Using the predicted value of boiling point from the eight-parameter equation[14] as one of the descriptors for the correlation of the flash point gives a five-parameter equation [eq 7-5, Table 7, Figure 6], where T_{FP} is the temperature of flash, $\sqrt[3]{G_b}$ is the cubic root of gravitational index, *HDCA* is the hydrogen donor charged area relative to the total molecular surface area, T_{BP} is the predicted boiling point,[14] δ_{Cl} is the maximum partial charge for a chlorine atom, and N_b is the number of triple bonds. Whereas the first two descriptors describe the intermolecular interactions in the bulk, the number of triple bonds and δ_{Cl} can be related to the reactivity of compounds.

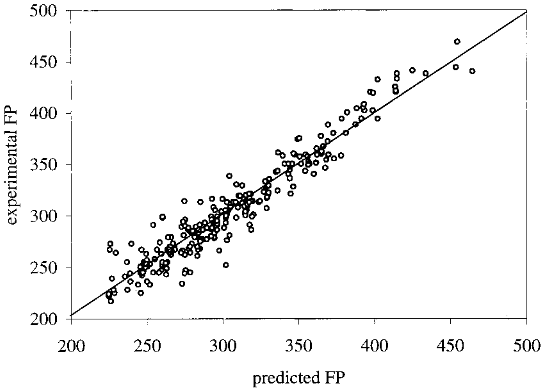

Figure 5. Plot of observed T_{FP} versus predicted T_{FP} using three-parameter correlation [eq 6-3].

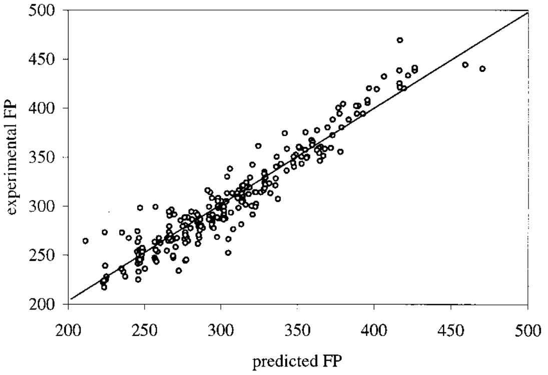

Figure 6. Plot of observed T_{FP} versus predicted T_{FP} using five-parameter correlation [eq 7-5].

Table 7. FP Correlations Using Predicted BP (Eight-Parameter Equation[14]) [Eqs 7-1, 7-2, 7-3, 7-4, and 7-5][a]

	R^2	R^2_{cv}	F	s^2		X	DX	*t*-test	descriptor
1	0.5669	0.5617	352.15	1137.52	0	−6.40	16.69	−0.38	1
					1	37.00	1.97	18.77	17
2	0.8853	0.8826	1034.17	302.42	0	−95.38	9.20	−10.4	1
					1	45.19	1.06	42.63	17
					2	16905	619.86	27.27	4
3	0.8860	0.8799	691.68	301.69	0	−92.63	9.44	−9.82	1
					1	42.26	2.51	16.81	17
					2	16729	634.08	26.38	4
					3	0.06	0.04	1.29	18
4	0.9035	0.8974	622.27	256.46	0	−106.4	8.92	−11.9	1
					1	43.60	2.33	18.75	17
					2	17540	596.2	29.42	4
					3	0.06	0.04	1.39	18
					4	−301.7	43.50	−6.93	19
5	0.9140	0.9079	563.53	229.21	0	−113.6	8.53	−13.3	1
					1	46.85	2.27	20.63	17
					2	17451	563.85	30.95	4
					3	0.003	0.004	0.075	18
					4	−310.0	41.15	−7.53	19
					5	28.79	5.04	5.71	20

[a] The correlation tables include *n*-1 old descriptors (of the previous correlation) and a new descriptor.

HDCA represents the sum of solvent-accessible surface area of H-bonding donor atoms and δ_{Cl} reflect characteristics of the charge distribution of the molecule.

The correlations of descriptors for the eqs 2-3 to 7-5 are given in Table 8(a−f).

Table 8. Intercorrelation of Descriptors

		(a) For Eq 2-3			
	D1[a]	D2[b]	D3[c]		
R^2	0.0891	0.0812	0.0111		
		(b) For Eq 3-3			
	D1[d]	D2[e]	D3[f]		
R^2	0.5043	0.5103	0.3525		
		(c) For Eq 4-3			
	D1[g]	D2[h]	D3[i]		
R^2	0.1276	0.3779	0.3097		
		(d) For Eq 5-3			
	D1[j]	D2[k]	D3[l]		
R^2	0.7347	0.2849	0.7278		
		(e) For Eq 6-3			
	D1[m]	D2[n]	D3[o]		
R^2	0.5038	0.5011	0.3097		
		(f) For Eq 7-5			
	D1[p]	D2[q]	D3[r]	D4[s]	D5[t]
R^2	0.8481	0.1576	0.8373	0.0605	0.0759

[a] D1 = cubic root of gravitational energy. [b] D2 = HA dependent HDCA-2/TMSA. [c] D3 = relative molecular weight. [d] D1 = experimental boiling point. [e] D2 = relative negative charge. [f] D3 = H-donors charged surface area. [g] D1 = predicted boiling point. [h] D2 = difference in charged partial surface area. [i] D3 = minimum electron attraction for a C atom. [j] D1 = predicted boiling point on the basis of two-parameter equation.[6] [k] D2 = H-bonding surface area. [l] D3 = molecular weight. [m] D1 = predicted boiling point on the basis of four-parameter equation.[6] [n] D2 = relative negative charge. [o] D3 = H-acceptors surface area. [p] D1 = cubic root of gravitational index. [q] D2 = hydrogen donor charged area relative to total molecular surface area. [r] D3 = predicted boiling point on the basis of eight-parameter equation.[14] [s] D4 = maximum partial charge for a Cl atom. [t] D5 = number of triple bonds.

Table 9. Validation of Correlations for Eq 2-3

subsets	R^2	group referred
A+B	0.9091	C
A+C	0.8975	B
B+C	0.9019	A
whole set	0.8982	

Table 10. Allocation of Descriptors for the Tables 2−7

name of descriptor	no.
intercept	1
total molecular one-center e−e repulsion	2
cubic root of gravitational index over all bonds	3
HA dependent HDCA-2/TMSA	4
relative molecular weight	5
experimental BP	6
relative negative charge	7
H-donors charged surface area	8
predicted BP (for the present data set)	9
difference in charged partial surface area	10
minimum electron attraction for a C atom	11
predicted BP (two-parameter equation)	12
HBSA H-bonding surface area	13
molecular weight	14
predicted BP (four-parameter equation)	15
H-acceptors surface area	16
cubic root of gravitational index	17
predicted boiling point (eight-parameter equation)	18
maximum partial charge for a Cl atom	19
number of triple bonds	20

VALIDATION OF RESULTS

The following procedure was used for the cross-validation of the three-parameter correlation eq 2-3. The whole data set of 271 compounds was randomly divided into three groups A (1,4,7,...), B (2,5,8,....), and C (3,6,9,....) (91, 90, and 90 points, respectively). The descriptors that were obtained in the best three-parameter correlation eq 2-3 for the whole set were selected. Using these descriptors, the flash point of subsets A+B, A+C, and B+C were predicted. The coefficients of the descriptors for subsets A+B, A+C, and B+C were recorded. The descriptors from the subset A+B were then used for the prediction of the flash point of group C. Similarly, group B and group A flash points were predicted using the descriptor coefficients from subsets A+C and B+C, respectively. The correlation values were given in Table 9.

The predicted flash points were then correlated with the experimental values. The predicted flash point for group A, B, and C were then combined and then correlated with the experimental flash point of the whole set. This cross-validation afforded a squared correlation coefficient of 0.8982 compared to that found in the original three-parameter correlation (eq 2-3, 0.9020). The predictions based on cross-validation are important since several experimental methodologies were employed in the determination of the flash point.

In a recent paper, Tetteh et al.[7] developed radial basis function (RBF) neural network models for the estimation of flash point and boiling point, based on Free-Wilson parameters for 25 molecular functional groups and their first-order molecular connectivity index. After dividing the total database of 400 compounds into training (134), validation (133), and test (133) sets, the average absolute error obtained for the validation set and testing sets ranges from 10 °C to 12 °C for FP and are in agreement with the experimental errors of about 10 °C. Single output 26-36-2 RBF prediction for the training set gives a R^2 value of 0.9553 and s (standard error) = 10.8 °C, the validation set gives $R^2 = 0.9271$ and $s = 14.1$ °C, and the test set gives $R^2 = 0.9155$ and $s = 14.3$ °C. Double output RBF prediction for the training set gives an $R^2 = 0.9614$ and $s = 10.1$ °C, the validation set gives an $R^2 = 0.9278$ and $s = 14.6$ °C, and the test set gives a R^2 value of 0.9220 and $s = 14.0$ °C. The quality of predictions by eq 2-3 is similar to this result, although with somewhat lower R^2_{cv} values. However, only three theoretical molecular descriptors are involved in eq 2-3, which underlines the robustness of the QSPR model.

Notably, Tetteh et al.[7] advocate that the neural network model developed by them accounts for the intrinsic nonlinearity between the flash point and molecular descriptors employed. This is based on the observation that the predictions by their neural network model do not exhibit systematic positive deviations at lower flash point values inherent to the earlier developed PLS model.[5] Such deviations can be observed in the case of our MLR models (Figures 1−6) but to a much lesser extent. Nevertheless, the application of nonlinear neural network approach with the theoretical descriptors employed in this work may improve the performance of the QSPR model.

CONCLUSION

The development of significant QSAR or QSPR equations by the extraction of molecular descriptors from large descriptor spaces has been successful for the prediction of many physical properties and biological activities of chemical compounds. The present work clearly demonstrates that analogous QSPR equations can be developed for the description of complex combustion physicochemical processes. Importantly, the descriptors employed in the best correlation equations are indicative of the nature of the interactions. We obtained a correlation with $R^2 = 0.9020$ for the descriptors calculated purely on the basis of theoretical descriptors. The model based on chemical structure and the experimental T_{BP} has $R^2 = 0.9529$. However, when the predicted values of the boiling point (two-parameter equation and four-parameter equation, respectively) is used for the correlation of the flash point, it gives three-parameter equations with $R^2 = 0.9073$ and 0.9247, respectively. When the predicted value of the boiling point on the basis of the eight-parameter equation is used as one of the descriptors, it gives a five-parameter flash point equation with $R^2 = 0.9140$. These correlations allow the prediction of the flash points of toxic or explosive compounds, where experimental determination is extremely difficult.

Many of the outliers are very small molecules, which have low flash points and are difficult to determine accurately. Our method correlates flash points within the limits of ±11 K. A similar error range is found for measured flash points.

REFERENCES AND NOTES

(1) Katritzky, A. R.; Mu, L.; Lobanov, V. S.; Karelson, M. Correlation of Boiling Points with Molecular Structure. 1. A Training Set of 298 Diverse Organics and a Test Set of 9 Simple Inorganics. *J. Phys. Chem*. **1996**, *100*, 10400−10407.

(2) Elvanov, S. F. Open-Cup Flash Point and Lower Flammability Temperature of Liquids. *Zh. Prikl. Khim. (English Translation)* **1991**, *64*, 747−752.

(3) Fire Protection Guide on Hazardous Materials. National Fire Protection Association, 1973.

(4) Mitchell, J. W.; Vratsanos, M. S.; Hanley, B. F.; Parekh, V. S. Experimental Flash Points of Industrial Amines. *J. Chem. Eng. Data*. **1999**, *44*, 209−211.

(5) Suzuki, T.; Ohtaguchi, K.; Koide, K. A Method for Estimating Flash Points of Organic Compounds From Molecular Structures. *J. Chem. Eng. Jpn*. **1991**, *24*, 258−261.

(6) Satyanarayana, K.; Rao, P. G. Improved Equation to Estimate Flash Points of Organic Compounds *J. Hazard. Mater*. **1992**, *32*, 81−85.

(7) Tetteh, J.; Suzuki, T.; Metcalfe, E.; Howells, S. Quantitative Structure−Property Relationships for the Estimation of Boiling Point and Flash Point Using a Radial Basis Function Neural Network. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 491−507.

(8) Katritzky, A. R.; Lobanov, V. S.; Karelson, M. CODESSA Reference Manual Version 2.0; Gainesville, 1996.

(9) 98/99 Catalog of fine chemicals. Fisher Scientific, Inc., 1997.

(10) Catalog Handbook of Fine Chemicals 1996−1997. Aldrich Chemical Co., Inc., 1996.

(11) Integrated Scientific Information System (documentation package), MDL Information Systems, Inc., 1997.

(12) Dewar, M. J. S.; et al. MOPAC 7.0 for UNIX; Quantum Chemistry Program Exchange, Project 688.

(13) Karelson, M. In *Molecular Descriptors in QSAR/QSPR*; Wiley-Interscience: New York, 2000.

(14) 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.

CI010043E