Quantitative Structure-Property Relationships for Normal Saturated and Unsaturated Fatty Acids

N.V.K. Dutt*, Y.V.L. Ravi Kumar and **H. Sumathi Vedanayagam**
Indian Institute of Chemical Technology, Hyderabad - 500 007, India

Quantitative structure-property relationships, a technique of deriving properties of compounds from knowledge of their structures, has been applied to aliphatic carboxylic acids, an industrially important group of compounds. Equations have been developed from the molecular connectivity of the first order and the molar refraction as inputs which, on testing with data of seven properties, yielded average absolute deviations ranging between 0.9 and 6.7%. The compounds studied ranged between C_1 and C_{30} for saturated and C_3 to C_{22} for unsaturated fatty acids. The properties studied are normal melting and boiling points, critical temperature and pressure, and heats of fusion, combustion and formation.

KEY WORDS: Fatty acids, molar refraction, molecular connectivity, property, structure.

Methods for relating the structure of molecules to their properties is known as Quantitative Structure-Property Relationships (QSPR). Randic (1) initially devised a connectivity index, which showed such a relationship for hydrocarbon properties. He studied a series of molecules with different degrees of branching and showed that this index has a linear relationship to certain physical properties. Kier *et al.* (2) used this QSPR technique to study the relationship to nonspecific local anesthesia, which included hydrocarbons and alcohols. Murray *et al.* (3) showed the molecular connectivity index to be well correlated to water solubility and boiling point of branched cyclic and straight-chain alcohols and hydrocarbons. The same authors (4) also studied the relationship to partition coefficients, which included a variety of monofunctional chemical classes such as ethers, alcohols, ketones, esters, carboxylic acids, amines and hydrocarbons. In another study, Kier *et al.* (5) correlated the connectivity function with density. Structure-activity studies on odor molecules (6) have included the sensation threshold concentration of fatty acids to overcome anosmia. Gombar (7) has studied the correlation between molecular connectivity and the boiling points of saturated aliphatic carboxylic acids. In the present work, a simple method of estimating some properties of aliphatic carboxylic acids from the molecular connectivity of the first order, $^1\chi$ (chi 1, a sum of the reciprocal square roots of the product of two adjacent hydrogen suppressed valencies), and from molar refraction, R_M (a constitutive and additive property determined as the sum of the atomic refractive indices), is reported. The calculation procedure is explained further in the text. The properties studied are normal melting and boiling points, critical temperature and pressure, and heats of combustion, formation and fusion.

METHODS AND MATERIALS

The method followed in the present study consists of estimating $^1\chi$ and R_M of the saturated and unsaturated fatty acids, following the procedures reported by Kier and Hall (8) and the Eisenlohr atomic and group contributions toward atomic refractions reported by Dreisbach (9), respectively. Data on several properties of the compounds, collected from the literature (10-13), were then correlated to $^1\chi$ and R_M as inputs, by means of mathematical expressions. The degree of agreement between the literature and calculated values for any property was characterized as the percent deviation.

Calculation of $^1\chi$ and R_M. Following the procedure of calculating the molecular connectivities for compounds reported by Kier and Hall (8), we give the example of hexanoic acid, a typical fatty acid, as follows.

First, the molecular skeleton of the compound is drawn, as shown in Figure 1. Each carbon atom in the skeleton is then assigned a number ranging between 1 and 4, depending on the number of hydrogen atoms present at vertex i, h_i: $\delta_i = 4 - h_i$. The number 4 may represent the valence or the number of valence electrons for the carbon atom. Thus, the terminal methyl carbon, CH_3, is assigned the number 1 (4 − 3), and the methylene carbons

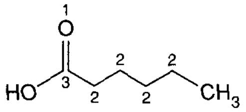


FIG. 1. Molecular skeleton of hexanoic acid with valence-suppressed indices shown at the vertices.

TABLE 1

Atomic Refractive Indices Used for Computing Molecular Refractive Index[a]

Carbon singly bound and alone	2.592
Carbon singly bound	2.418
Carbon double bond	1.733
Hydrogen	1.100
Oxygen-hydroxyl	1.525
Oxygen-ketonic	2.211

[a]All values are for the sodium line.

TABLE 2

Example of R_M Calculation for Hexanoic Acid

Atom/bond	Number	Contribution
Carbon singly bound and alone	6	6 × 2.418 = 14.508
Hydrogen	12	12 × 1.1 = 13.200
Oxygen-ketonic	1	1 × 2.211 = 2.211
Oxygen-hydroxyl	1	1 × 1.525 = 1.525
Total		31.444

---

*To whom correspondence should be addressed.

TABLE 3

Molecular Connectivities and Molar Refractions of Fatty Acids

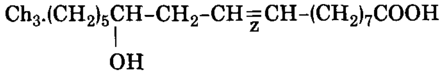
Compound no.	Fatty acid	Formula	$^1\chi$	R_M
1.	Formic	$HCOOH$	1.15	8.35
2.	Acetic	CH_3COOH	1.73	12.97
3.	Acrylic	$CH_2=CH.COOH$	1.82	14.02
4.	Propionic	CH_3-CH_2-COOH	2.27	17.59
5.	Isocrotonic	$CH_3-CH=CH-COOH$	2.39	18.64
6.	Vinylacetic	$CH_2=CH-CH_2-COOH$	2.30	18.64
7.	Butyric	$CH_3-(CH_2)_2COOH$	2.77	22.21
8.	Allylacetic	$CH_2=CH-(CH_2)_2COOH$	2.80	23.26
9.	Valeric	$CH_3-(CH_2)_3COOH$	3.27	26.83
10.	Hydrosorbic	$CH_3-CH_2-CH=CH-CH_2COOH$	3.42	27.87
11.	Caproic	$CH_3-(CH_2)_4COOH$	3.77	31.44
12.	Enanthic	$CH_3-(CH_2)_5COOH$	4.27	36.06
13.	Caprylic	$CH_3(CH_2)_6COOH$	4.77	40.68
14.	Pelargonic	$CH_3(CH_2)_7COOH$	5.27	45.29
15.	Capric	$CH_3(CH_2)_8COOH$	5.77	49.92
16.	Undecylenic	$CH_2=CH-(CH_2)_8COOH$	5.80	50.96
17.	Undecanoic	$CH_3-(CH_2)_9COOH$	6.27	54.53
18.	Lauroleic	$CH_3-CH_2-CH\underset{Z}{=}CH(CH_2)_7COOH$	6.42	55.58
19.	Lauric	$CH_3-(CH_2)_{10}COOH$	6.77	59.15
20.	Tridecanoic	$CH_3-(CH_2)_{11}-COOH$	7.27	63.77
21.	Myristoleic	$CH_3-(CH_2)_3CH\underset{Z}{=}CH-(CH_2)_7COOH$	7.42	64.82
22.	Myristic	$CH_3-(CH_2)_{12}COOH$	7.77	68.39
23.	Pentadecanoic	$CH_3(CH_2)_{13}COOH$	8.27	73.01
24.	Palmitoleic	$CH_3(CH_2)_5CH\underset{Z}{=}CH.(CH_2)_7COOH$	8.42	74.05
25.	Palmitic	$CH_3(CH_2)_{14}-COOH$	8.77	77.62
26.	Heptadecenoic	$CH_3(CH_2)_{15}COOH$	9.27	82.24
27.	Eleostearic	$CH_3-(CH_2)_3CH\underset{E}{=}CH-CH\underset{E}{=}CH-CH\underset{Z}{=}CH-(CH_2)_7COOH$	8.75	76.15
28.	Linolenic	$CH_3-CH_2-CH\underset{Z}{=}CH-CH_2-CH\underset{Z}{=}CH-CH_2-CH\underset{Z}{=}CH-(CH_2)_7COOH$	8.72	76.15
29.	Linoleic	$CH_3(CH_2)_4CH\underset{Z}{=}CH-CH_2-CH\underset{Z}{=}CH-(CH_2)_7COOH$	9.07	79.72
30.	Petroselenic	$CH_3-(CH_2)_{10}-CH\underset{Z}{=}CH-(CH_2)_4COOH$	9.42	83.29
31.	Oleic	$CH_3(CH_2)_7-CH\underset{Z}{=}CH-(CH_2)_7-COOH$	9.42	83.29
32.	Elaidic	$CH_3-(CH_2)_7CH\underset{E}{=}CH-(CH_2)_7COOH$	9.42	83.29
33.	Vaccenic	$CH_3-(CH_2)_5CH\underset{Z}{=}CH(CH_2)_9COOH$	9.42	83.29
34.	Stearic	$CH_3-(CH_2)_{16}COOH$	9.77	86.86
35.	Ricinoleic	 $Ch_3.(CH_2)_5\underset{\underset{OH}{\vert}}{CH}-CH_2-CH\underset{Z}{=}CH-(CH_2)_7COOH$	9.81	84.82
36.	Nonadecanoic	$CH_3-(CH_2)_{17}COOH$	10.27	91.48
37.	Gadoleic	$CH_3(CH_2)_9CH\underset{Z}{=}CH-(CH_2)_7COOH$	10.42	92.53
38.	Erucic	$CH_3(CH_2)_7CH\underset{Z}{=}CH.(CH_2)_{11}COOH$	11.42	101.76
39.	Brassidic	$CH_3(CH_2)_7CH\underset{E}{=}CH(CH_2)_{11}COOH$	11.4	101
40.	Eicosanoic (arachidic)	$CH_3-(CH_2)_{18}-COOH$	10.77	96.10
41.	Arachidonic	$CH_3-(CH_2)_4-CH\underset{Z}{=}CH-CH_2-CH\underset{Z}{=}CH-CH_2-CH\underset{Z}{=}CH-CH_2-CH\underset{Z}{=}CH-(CH_2)_3COOH$	9.36	81.82
42.	Heneicosanoic	$CH_3-(CH_2)_{19}COOH$	11.27	100.71
43.	Docosanoic (behenic)	$CH_3-(CH_2)_{20}-COOH$	11.77	105.33
44.	Tricosanoic	$CH_3(CH_2)_{21}COOH$	12.27	109.95
45.	Lignoceric	$CH_3(CH_2)_{22}COOH$	12.77	114.57
46.	Pentacosanoic	$CH_3(CH_2)_{23}COOH$	13.27	119.19
47.	Hexacosanoic	$CH_3(CH_2)_{24}COOH$	13.77	123.80
48.	Heptacosanoic	$CH_3(CH_2)_{25}COOH$	14.27	128.42
49.	Octacosanoic	$CH_3(CH_2)_{26}COOH$	14.77	133.04
50.	Nonacosanoic	$CH_3(CH_2)_{27}COOH$	15.27	137.66
51.	Triacontanoic	$CH_3(CH_2)_{28}COOH$	15.77	142.28

TABLE 4

Summary of Results

Property studied	No. of compounds	Correlating parameter	Type of equation	Coefficients in Equations 3-5 A	B	C	D	% Average absolute deviation ($\bar{e}$)
Normal melting point, K	37	R_M	Linear	255.7	0.7710	—	—	5.7
Normal boiling point, K	27	R_M	Linear	371.6	3.27	—	—	1.7
Critical temperature, K	20	R_M	Linear	575.5	2.614	—	—	1.4
Critical pressure, MPa	20	$1/{}^1\chi$	Parabolic	0.1278	12.79	−5.125	—	1.7
Heat of formation, KJ/mole	15	${}^1\chi$	Parabolic	415.80	36.72	2.073	—	0.9
Heat of fusion, KJ/mole	18	${}^1\chi$	Parabolic	13.60	−2.554	0.7939	—	6.7
Heat of combustion, KJ/mole	15	R_M	Cubic	−1605	193.9	−1.257	0.00931	1.4

2 each (4 − 2). According to Kier and Hall (8), the heteroatoms OH (hydroxyl) and O of the carboxyl group are assigned a valence number, $\delta = 1$ each. The carbonyl carbon atom of the carboxyl group (>C=O) is assigned a δ value of 3 (delta is a count of only sigma electrons). Molecular connectivity of the first order, ${}^1\chi$ is then calculated from the formula:

$$ {}^1\chi = \Sigma\,(\delta_i\delta_j)^{-1/2} \qquad [1] $$

where δ_i and δ_j are the valencies at the vertices i and j. Thus, ${}^1\chi$ for the molecule as a whole becomes:

$$ {}^1\chi = \sqrt{1/3} + \sqrt{1/3} + \sqrt{1/3\cdot 2} + \sqrt{1/2\cdot 2} + \sqrt{1/2\cdot 2} + \sqrt{1/2\cdot 2} + \sqrt{1/2\cdot 1} = 3\cdot 77 \qquad [2] $$

Molar refraction, R_M, for hexanoic acid is calculated as the sum of atomic and group contributions of Eisenlohr as reported by Dreisbach (9). According to this method, R_M is computed as the sum of various contributions (atomic refractive indices) as tabulated by Dreisbach (9). Some relevant atomic refractive indices are reproduced in Table 1 to explain the calculation of R_M.

Thus, the molar refraction R_M for hexanoic acid is calculated as exemplified in Table 2. By using the methods of calculation described, χ and R_M values for other compounds studied were calculated and are reported in Table 3.

RESULTS AND DISCUSSION

Data collected from the literature (10–12) on several properties for the compounds (as numbered in Table 3) are correlated with ${}^1\chi$ or R_M by means of the linear (L), parabolic (P) or cubic (C) equations.

$$ y = A + B\chi \qquad (L) \qquad [3] $$

$$ y = A + B\chi + C\chi^2 \qquad (P) \qquad [4] $$

$$ y = A + B\chi + C\chi^2 + D\chi^3 \qquad (C) \qquad [5] $$

where y denotes the property studied and χ denotes ${}^1\chi$, R_M or the reciprocal of ${}^1\chi$, as the case may be.

Results obtained on the estimation of several properties are summarized and reported in Table 4. In Table 4 the percent deviation, e, calculated for any compound is defined by the equation:

$$ e = 100\cdot\left(\frac{Y_{exp} - Y_{cal}}{Y_{exp}}\right) \qquad [6] $$

where Y_{exp} and Y_{cal} represent the experimental value of the property reported in the literature and the calculated value of the property from Equations 3–5, respectively, best fitted to the data. The percent average absolute deviation, $\bar{e}$, for a set of N number of compounds is defined by the equation:

$$ \bar{e} = \frac{\Sigma\,|e|}{N} \qquad [7] $$

The values of $\bar{e}$ reported in Table 4 are those obtained from the appropriate type of equation selected for representation on the basis of simplicity combined with reasonable accuracy. In the case of the critical temperature and normal boiling and melting points, no advantage was found in using higher-order equations other than the linear equation. Similarly, parabolic equations gave adequate representations of the data on critical pressure and heats of fusion and formation. A cubic equation gave the best results for the heat of combustion. As seen from Table 4, the information derived from knowledge of the structures of carboxylic acids studied (in the form of ${}^1\chi$ and R_M) can be correlated with physical properties.

REFERENCES

1. Randic, M., *J. Chem. Phys. 97*:6609 (1975).
2. Kier, L.B., L.H. Hall, W.J. Murray and M. Randic, *J. Pharm. Sci. 64*:1971 (1975).
3. Hall, L.H., L.B. Kier and W.J. Murray, *Ibid. 64*:19 (1975).
4. Murray, W.J., L.H. Hall and L.B. Kier, *Ibid. 64*:1978 (1975).
5. Kier, L.B., L.H. Hall, W.J. Murray and M. Randic, *Ibid. 65*:1226 (1976).
6. Kier, L.B., T. Di Paolo and L.H. Hall, *J. Theor. Biol. 67*:585 (1977).
7. Gombar, V.K., *Pharmacos 25*:46 (1981).
8. Kier, L.B., and Hall, L.H., *Molecular Connectivity in Structure Activity Analysis,* Research Studies Press, Ltd., Letchworth, Hertfordshire, and John Wiley & Sons Inc., New York, 1986.
9. Dreisbach, R.R., *Physical Properties of Chemical Compounds—III,* American Chemical Society, Washington, D.C., 1961.
10. Grayson, M. (ed.), *Kirk-Othmer Encyclopedia of Chemical Technology,* Vol. 4, John Wiley & Sons, New York, 1980.
11. Gunstone, F.D., *An Introduction to the Chemistry and Biochemistry of Fatty Acids and Their Glycerides,* 2nd edn., Chapman and Hall Ltd., England, 1967, p. 14.
12. Fischer, C.H., *J. Am. Oil Chem. Soc. 66*:1158 (1987).
13. Daubert, T.E., and R.P. Danner, *Physical and Thermodynamic Properties of Pure Chemicals,* Hemisphere Publishing Corp., New York, 1989.

[Received May 6, 1992; accepted September 18, 1992]