

The nature of topological parameters. II. The composition of topological parameters

Marvin Charton^{1,*} & Barbara I. Charton²

¹*Chemistry Department, Pratt Institute, Brooklyn, NY 11205, U.S.A.*

²*Cole Science Library, New York University, New York, NY, U.S.A.*

MS received 28 December 2002; accepted for publication 15 January 2003

Summary

The composition of thirty topological parameters of various types including the χ , κ and τ parameters and the Wiener indices was determined. Those derived from alkyl and substituted alkyl groups are a function of branching and steric effects; those derived from haloalkanes and haloalkenes are a function of some combination of polarizability, dipole moment, the number of halogen atoms of each type, and the ration of the number of branches to the number of atoms in the longest chain. It follows that topological parameters are actually counts of the number of atoms of each type, bonds of each type, valence, nonbonding and core electrons, branches at each position, OH and NH bonds, and nonbonding electron pairs on O and/or N atoms. These counts can be related to polarizability, dipole moment, hydrogen bonding, and steric effects, which are the true determinants of properties and biological activities. Topological parameters themselves do not have a cause and effect relationship with the quantities they are used to model. They function because:

1. They are counts of quantities that are related to the fundamental quantities that determine properties.
2. They are composite parameters which for the purpose of obtaining a predictive empirical relationship are able to represent a data set as well as pure parameters.

Introduction

In the previous paper of this series [1] we have shown that topological parameters are not ‘fundamental’, that is, they do not have a direct cause and effect relationship with molecular properties. We have proposed that ‘in those data sets with variable topology the topological parameters are representing polarizability and dipole moment, and probably steric effects as well.’ That they are functions of polarizability is further shown by their correlation of molar refraction [2]. They are composite parameters. Topological parameters actually are counts of numbers of atoms, of bonds, of branching, and of electrons.’ In this work we provide further evidence concerning the composition of topological parameters. We have examined χ and κ parameters, τ parameters, and Wiener indices. We have also examined various types of substituents and skeletal groups.

*Correspondence. E-mail: mcharton@pratt.edu

Alkyl and substituted alkyl substituents

When the structural features were alkyl groups only steric effects and polarizability need be considered as parameters. As the alkyl groups vary in length and branching the correlation equation was based on the simple branching Equation (3).

$$S\psi = \sum_{i=1}^m a_i n_i + a_b n_b \quad (1)$$

where $S\psi$ represents the steric effect parameterization, the a_i and a_b are coefficients, n_i is the number of branches attached to the i -th atom, and n_b is the number of bonds between the first and last atoms of the group skeleton. It follows that n_b is a measure of group length. Unfortunately it is frequently highly collinear in group polarizability, which greatly limits its utility. For saturated cyclic substituents it is necessary to determine values of n_i from an appropriate regression equation. For planar π -bonded groups n_i is taken to be 1 for each atom in the group skeleton.

For other groups n_i is obtained simply by counting branches. The model makes the assumption that all of the branches attached to a skeleton atom are equivalent. This is at best only a rough approximation. Distinguishing between branches results in an improved model called the expanded branching equation:

$$S\psi = \sum_{i=1}^m \sum_{j=1}^3 a_{ij} n_{ij} + a_b n_b \quad (2)$$

which allows for the difference in steric effect that results from the order of branching [3]. This difference follows from the MSI principle which states that the preferred conformation is that which results in the minimum steric interaction. The first branch has the smallest steric effect because a conformation in which it is rotated out of the way of the active site is preferred. In this conformation the active site is in contact with two hydrogen atoms. The preferred conformation in the case of a second branch has the larger of the two branches directed out of the way. The smaller branch and a hydrogen atom are in contact with the active site. When there are three branches the largest will be directed out of the way and the other two will be in contact with the active site. Regrettably, the expanded branching method requires a large number of parameters. Data sets large enough to permit its use are rarely found.

To the correlation equation we then add a term in n_C , the number of carbon atoms in the alkyl group. The term in n_b is dropped from the correlation as in the general as it cannot be distinguished from n_C . We have also in some cases made use of the parameter $r_{b/cl}$ which is defined as the number of branches on the longest chain of a substituent divided by the number of atoms in the longest chain in the hydrogen suppressed structure. It is a measure of:

1. Substituent or compound shape.
2. Substituent or compound flexibility.

and is given by the equation:

$$r_{b/cl} = \frac{\sum n_i}{n_{A,cl}} \quad (3)$$

where:

n_i is the number of branches at the i -th atom of the longest chain.

$n_{A,cl}$ is the number of atoms in the longest chain.

Examples of values for alkyl and substituted alkyl groups are given in Table 1.

When the alkyl group is substituted by other groups it is necessary to replace n_C by some other

Table 1. Values of $r_{b/cl}$ for some alkyl and substituted alkyl groups

Alkyl or substituted alkyl group	$\sum n_i$	$n_{A,cl}$	$r_{b/cl}$
-CHMe-CMe ₂ -CHMe-CH ₃	4	4	1
Cl-CCl ₂ -CCl ₂ -Cl	4	4	1
-CH ₂ -CMe ₂ -CH ₂ -CH ₂ -CH ₃	2	5	0.400
-CH ₂ -CMe ₂ -CH ₂ -CHCl-CH ₃	3	5	0.600

measure of polarizability and to parameterize the electrical effect. As a polarizability parameter we have used α , given by:

$$\alpha = \frac{MR_X - MR_H}{100} = \frac{MR_X}{100} - 0.0103 \quad (4)$$

where MR_X and MR_H are the group molar refractivities of X and H [4]. The factor 1/100 is introduced to scale the α parameter so that its coefficients in the regression equation are roughly comparable to those obtained for the other parameters used. The α parameters used in this work are from Charton [] or were estimated by additivity.

As the only electrical effect generally observed for a group bonded to an sp^3 hybridized C atom is the localized (field) effect we have used the σ_1 constants as electrical effect parameters. The σ_1 values used in this work are from Charton [5].

Haloalkanes, haloalkenes, and halobenzenes

In the case of haloalkanes and haloalkenes the simplest objects available for counting are the halogen atoms of each type. Thus the correlation equation contains the variables n_F , n_{Cl} , and n_{Br} . The use of these variables is justified by the success of the group number (GN) Equation (6) which is given by:

$$Q_{Xi} = \sum_{i=1}^l b_i n_{Xi} + b_C n_C + b_o \quad (5)$$

where n_{Xi} is the number of the i -th X groups attached in equivalent positions to the skeletal group and n_C is the number of carbon atoms in the skeletal group. The GN equation can be obtained from a simplified form of the intermolecular force (IMF) Equation (7):

$$Q_x = \rho \sigma_x + S \psi_x + A \alpha_x + M \mu_x + H_1 n_{HX} + H_2 n_{nX} + a_o \quad (6)$$

where σ is a composite electrical effect parameter such as a Hammett constant, ψ is a steric effect parameterization, α is the polarizability parameter defined in

Table 2. Data used in the correlations

Set TP 1. X, $^0\chi^v$: Me, 1.447; Et, 2.154; BrCH₂CHBrCH₂, 6.361; ClCH₂CHMe, 3.864; Pr, 2.861; H₂NCH₂CHMe, 3.309; Bu, 3.569; iBu, 3.732; Et₂NCH₂CH₂, 4.069; cHx, 4.560; PhOCH₂CH₂, 5.656; Oc, 6.397; No, 7.104; Dc, 7.811; Cl₃CCH₂, 5.051; BuEtCHCH₂, 6.560; iPr, 3.025; Udc, 8.518; ClCH₂CH₂CH₂, 3.701. [8]

Sets TP 2, 3, 4, 5. X, $^1\chi$, $^2\chi$, $^3\chi$, $^4\chi$: Me, 0.707, 0, 0, 0; Et, 1.207, 0.500, 0, 0; iPr, 1.563, 1.394, 0, 0; tBu, 1.854, 2.561, 0, 0; Pr, 1.707, 0.854, 0.354, 0; Bu, 2.207, 1.207, 0.604, 0.250; Et₂CH, 2.639, 1.683, 0.866, 0.289; Pe, 2.707, 1.561, 0.854, 0.427; iPrCH₂CH₂, 2.563, 2.036, 0.781, 0.408; tBuCH₂, 2.354, 2.811, 0.750, 0; Pr₂CH, 3.639, 2.442, 1.394, 0.961; iBuMeCH, 3.247, 3.462, 0.961, 1.045; iBu₂CH, 4.625, 4.072, 1.971, 1.633; tBuCH₂(Me₂)C, 3.561, 5.109, 1.362, 1.436; (tBuCH₂)₂CH, 4.932, 6.030, 1.802, 1.804; Et₃C, 3.536, 2.561, 2.250, 0.750; (tBuCH₂)(tBu)MeC, 4.811, 6.548, 2.731, 1.436; tBuMe₂C, 3.104, 4.134, 2.030, 0. [9]

Sets TP 6, 7. X, $^2\kappa$, $^3\kappa$: Me, 1.000, 1.450; Et, 2.000, 2.000, Pr, 3.000, 3.378; Bu, 4.000, 4.000, Pe, 5.000, 5.333; Hx, 6.000, 6.000; Hp, 7.000, 7.200; iBu, 2.250, 4.000, CH₂sBu, 3.200, 3.000; CH₂iBu, 3.200, 5.333; CH₂tBu, 1.633, 5.333; iPr, 1.333, 0; sBu, 2.250, 4.000; 2-Pe, 3.200, 5.333; 2-Hx, 4.167, 6.000; 2-Hp, 5.143, 7.200; 3-Pe, 3.200, 3.000, 3-Hx, 4.167, 3.840; iPrEtCH, 3.061, 2.666; iPr₂CH, 3.928, 2.812. [10]

Sets TP 8, 9, 10, 11. X, T_R, F, T, B: Me, 1.000, 1.816, -0.816, 0; Et, 1.414, 1.284, 0.130, 0; Pr, 1.914, 1.284, 0.630, 0; iPr, 1.732, 1.049, 0.683, 0.182; Bu, 2.414, 1.284, 1.130, 0; Pe, 2.914, 1.284, 1.163, 0; sBu, 2.270, 1.049, 1.221, 0.144; iPe, 2.770, 1.284, 1.486, 0.144; tPe, 2.561, 0.908, 1.653, 0.353; 3-Pe, 2.808, 1.049, 1.757, 0.106; 2-Pe, 2.770, 1.049, 1.721, 0.144; PrMeCHCH₂, 2.808, 1.284, 1.524, 0.106; iBu, 2.270, 1.284, 0.986, 0.144; tBu, 2.000, 0.908, 1.092, 0.414; Hx, 3.414, 1.284, 2.130, 0; Hp, 3.914; 1.284, 3.130, 0. [11]

Sets TP 12, 12R, 13, 13R, 14, 14R, 15, 15R, 16, 17, 18. Cmpd, $^1\chi$, $1/^1\chi$, $^0\chi^v$, $1/^0\chi^v$, $^1\chi^v$, $1/^1\chi^v$, χ_s^v , $1/\chi_s^v$, $^3\chi_c^v$, $^4\chi_c^v$, $^4\chi_{pc}^v$: MeCl, 1.0000, 1.0000, 2.1339, 0.4686, 1.1339, 0.8819, 1.1339, 0.8819, 0, 0, 0; CH₂Cl₂, 1.4142, 0.7071, 2.9749, 0.3361, 1.6036, 0.6236, 0.9091, 1.1000, 0, 0, 0; CHCl₃, 1.7321, 0.5773, 3.9791, 0.2513, 1.9640, 0.5092, 0.8417, 1.1881, 0.8417, 0, 0; CCl₄, 2.0000, 0.5000, 5.0356, 0.1986, 2.2648, 0.4415; 1.4638, 0.6832, 2.9157, 1.4638, 0; ViCl, 1.4142, 0.7071, 2.4184, 0.4135, 1.0629, 0.9408, 0.4629, 2.1603, 0, 0, 0; Cl₂C=CH₂, 1.7321, 0.5773, 3.4749, 0.2878, 1.4874, 0.6723, 0.4546, 2.1997, 0.4546, 0, 0; Z-1,2-VnCl₂, 1.9142, 0.6224, 3.4225, 0.2922, 1.6426, 0.6088, 0.4286, 2.3332, 0, 0, 0; E-1,2-VnCl₂, 1.9142, 0.6224, 3.4225, 0.2922, 1.6426, 0.6088, 0.4286, 2.3332, 0, 0, 0; Cl₂C=CHCl, 2.2701, 0.4405, 4.4790, 0.2233, 2.0772, 0.4814, 0.4208, 2.3764, 0.3712, 0, 0.4208; Cl₂C=CCl₂, 2.6429, 0.3784, 5.5356, 0.1806, 2.5178, 0.3972, 0.4133, 2.4195, 0.6429, 0, 1.4579; EtCl, 1.4142, 0.7071, 2.8410, 0.3520, 1.5089, 0.6627, 0.8018, 1.2472, 0, 0, 0; CHCl₂Me, 1.7321, 0.5773, 2.8451, 0.3515, 1.8867, 0.5300, 0.7423, 1.3472, 0.7423, 0, 0; ClCH₂CH₂Cl, 1.9142, 0.5224, 3.6820, 0.2716, 2.1036, 0.4753, 0.6429, 1.5555, 0, 0, 0; CCl₃Me, 2.2701, 0.4405, 4.6861, 0.2134, 2.5193, 0.3969, 0.5952, 1.6801, 0.5249, 0, 0.5952; Cl₂HC-CH₂Cl, 2.0000, 0.5000, 4.9017, 0.2040, 2.2008, 0.4544, 0.7284, 1.3719, 7.6575, 0.7289, 0; CHCl₂CHCl₂, 2.6127, 0.3784, 5.6903, 0.1757, 2.9519, 0.3388, 0.5510, 1.8149, 0.8571, 0, 1.9438; CCl₃CHCl₂, 2.5607, 0.3905, 5.7427, 0.1741, 2.8562, 0.3501, 0.5844, 1.7112, 2.0926, 0.5154, 1.0309; CH₂F₂, 1.4142, 0.7071, 1.4630, 0.6835, 0.5345, 1.8709, 0.1010, 9.9010, 0, 0, 0; CH₂FCl, 1.4142, 0.7071, 2.2190, 0.4507, 1.0690, 0.9335, 0.3030, 3.3003, 0, 0, 0; CH₂BrCl, 1.4142, 0.7071, 3.8050, 0.2628, 2.1905, 0.4565, 1.5747, 0.6350, 0, 0, 0; CH₂Br₂, 1.4142, 0.7071, 4.6350, 0.2157, 2.7775, 0.3600, 2.7274, 0.3666, 0, 0, 0; CF₃CHBrCl, 2.9434, 0.3397, 5.0309, 0.1988, 2.6442, 0.3782, 0.0347, 28.8184, 0.7936, 0.0156, 1.1122; ViBr, 1.4142, 0.7071, 3.2484, 0.3078, 1.5421, 0.6484, 0.8018, 1.2472, 0, 0, 0; BrCH₂CH₂Cl, 1.9142, 0.5224, 4.5121, 0.2216, 2.6905, 0.3717, 1.1135, 0.8981, 0, 0, 0.

Sets TP 19, 20, 21, 22, 23, 24. Cmpd, w, w_r, w_r^v, w_r^v, p, f: MeCl, 1, 0.2500, 1.000, 0.2500, 0, 0; CH₂Cl₂, 4, 0.4444, 2.706, 0.3007, 0, 2; CHCl₃, 9, 0.5625, 5.118, 0.3199, 0, 6; CCl₄, 16, 0.6400, 8.236, 0.3294, 0, 12; ViCl, 4, 0.4444, 2.353, 0.2614, 0, 2; Cl₂C=CH₂, 9, 0.5325, 4.912, 0.3070, 0, 6; Z-1,2-VnCl₂, 10, 0.6250, 5.412, 0.3383, 1, 4; E-1,2-VnCl₂, 10, 0.6250, 5.412, 0.3383, 1, 4; Cl₂C=CHCl, 18, 0.7200, 9.177, 0.3671, 2, 8; Cl₂C=CCl₂, 29, 0.8056, 14.148, 0.3930, 4, 12; EtCl, 4, 0.4444, 3.353, 0.3726, 0, 2; CHCl₂Me, 9, 0.5625, 6.412, 0.4008, 0, 6; ClCH₂CH₂Cl, 10, 0.6250, 7.412, 0.4633, 1, 4; CCl₃Me, 18, 0.7200, 12.177, 0.4871, 2, 8; CHCl₂CH₂Cl, 16, 0.6400, 10.177, 0.4071, 0, 12; CHCl₂CHCl₂, 29, 0.8056, 18.648, 0.5180, 4, 12; CCl₃CHCl₂, 28, 0.7778, 17.648, 0.4902, 3, 14; CH₂F₂, 4, 0.4444, 3.334, 0.3704, 0, 2; CH₂FCl, 4, 0.4444, 3.020, 0.3356, 0, 2; CH₂BrCl, 4, 0.4444, 2.524, 0.2804, 0, 2; CH₂Br₂, 4, 0.4444, 2.342, 0.2602, 0, 2; CF₃CHBrCl, 42, 0.8571, 29.643, 0.6050, 6, 18; ViBr, 4, 0.4444, 2.171, 0.2412, 0, 2; BrCH₂CH₂Cl, 10, 0.6250, 7.048, 0.4405, 1, 4. [12]

TP 25, 26. X, 1B , 0B : 2-Me, 3.1884, 2.4510; 3-Me, 3.1284, 2.4333; 4-Me, 3.0808, 2.4187; 2,3-Me₂, 3.2931, 2.4699; 2,4-Me₂, 3.2144, 2.4468; 2,5-Me₂, 3.2153, 2.4471; 2,6-Me₂, 3.2940, 2.4703; 3,4-Me₂, 3.2139, 2.4465; 2-Cl, 3.4668, 2.5691; 3-Cl, 3.3908, 2.5468; 4-Cl, 3.3316, 2.5279; 2,3-Cl₂, 3.7980, 2.6773; 2,4-Cl₂, 3.6833, 2.6432; 2,5-Cl₂, 3.6792, 2.6413; 3,4-Cl₂, 3.6869, 2.6449; 3,5-Cl₂, 3.6738; 2,6409. [13]

Abbreviations: Pe, pentyl; Hx, hexyl; Hp, heptyl; Oc, octyl; No, nonyl; Dc, decyl; Udc, undecyl; cHx, cyclohexyl; Vi, vinyl; Vn, vinylene, Cmpd, compound.

Equation (3), μ is the dipole moment, n_H , the number of OH and NH bonds, and n_n , the number of lone pairs on O or N, are hydrogen bonding parameters. The correlation equation actually used was of the form:

$$Q_X = \sum_{i=1}^l b_i n_{Xi} + A\alpha + M\mu + \dots + b_o \quad (7)$$

where α is the polarizability of the entire molecule and μ is its dipole moment.

Method

Statistics

Values of the parameters taken from the literature were correlated with equations that represented counts of structural features of various types by means of multiple linear regression analysis. The best regression equations are reported, they are obtained by dropping independent variables whose coefficients were not significant.

The quantitative description of independent variable composition

It is helpful in discussing results to use C_i , the percent contribution of each independent variable in the regression equation, defined as: [1]

$$C_i = (100 a_i x_i) / \left(\sum_1^m a_i x_i \right) \quad (8)$$

where a_i is the regression coefficient of the i -th independent variable and x_i is its value for some reference group. We can define a hypothetical or choose an actual reference group.

The C_i values are useful in comparing the relative importance of the contribution of the different independent variables to the regression equation both for a given data set and between data sets.

Results

Alkyl and substituted alkyl groups

Protic and Sablic [8] have reported the correlation of LC_{50} , values for alcohols and other compounds in the fathead minnow with the $^{\circ}\chi^v$, parameters,

Table 3. Correlation matrix for Equation (10)

n_1	n_2	n_3	n_4	σ_1	α	
1	0.056	0.220	0.210	0.131	0.051	n_1
	1	0.169	0.014	0.613	0.306	n_2
		1	0.698	0.183	0.668	n_3
			1	0.136	0.713	n_4
				1	0.054	σ_1
					1	α

Table 4. Correlation matrix for Equation (12)

n_1	n_2	n_3	n_c	
1	0.451	0.219	0.635	n_1
	1	0.252	0.720	n_2
		1	0.794	n_3
			1	n_c

The parameters for alcohols were correlated with the equation:

$$Q_X = s_1 n_{1X} + s_2 n_{2X} + s_3 n_{3X} + s_4 n_{4X} + L \sigma_{1X} + A \alpha_X + a_o \quad (9)$$

The data used in the correlations are reported in Table 2. The correlation matrix is set forth in Table 3.

The regression equation is:

$$^{\circ}\chi^v = -0.838 (\pm 0.172) n_{4Ak} + 18.2 (\pm 0.956) \alpha_{Ak} + 0.419 (\pm 0.213) \quad (10)$$

$100R^2$, 96.80; $A100R^2$, 96.63; F , 257.5; S_{est} , 0.362; S° , 0.194; N_{dp} , 20; Correlation of the LC_{50} values for the alcohols with the variables n_{4X} and α_X gave the equation:

$$LC_{50} = 0.795 (\pm 0.177) n_{4X} - 13.9 (\pm 0.983) \alpha_X + 3.75 (\pm 0.219) \quad (11)$$

$100R^2$, 93.99; $A100R^2$, 93.66; F , 132.9; S_{est} , 0.372; S° , 0.266; N_{dp} , 20; C_{n4} , 22.2; C_{α} , 77.8

Equation (11) is in reasonable agreement with Equation (10).

Murray [9] has studied the relationship between the Taft E_S steric parameter for alkyl groups and the $^n\chi$ parameters. We have correlated the latter with the equation:

$$Q_{Ak} = s_1 n_1 + s_2 n_2 + s_3 n_3 + a_c n_c + a_o \quad (12)$$

Composition of the $n^{\wedge}\chi$ parameters

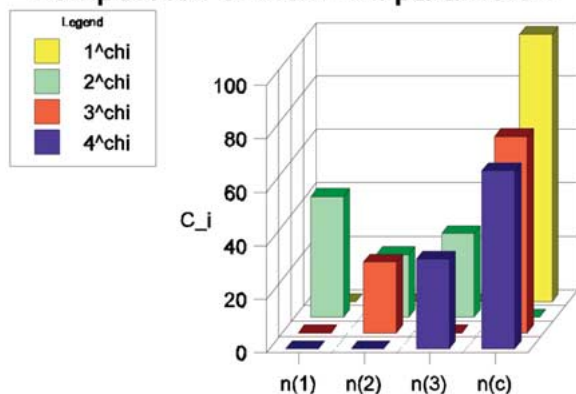


Figure 1.

The data (sets TP 2 through 5) is given in Table 2, the correlation matrix in Table 4. The regression equations are:

for $^1\chi$;

$$^1\chi = 0.425 (\pm 0.0185)n_c + 0.401 (\pm 0.117) \quad (13)$$

$100R^2$, 97.05; F, 526.9; S_{est} , 0.215; S^0 , 0.182; N_{dp} , 18.

for $^2\chi$;

$$^2\chi = 0.870 (\pm 0.186)n_1 + 0.453 (\pm 0.144)n_2 + 0.606 (\pm 0.906)n_3 - 0.362 (\pm 0.344) \quad (14)$$

$100R^2$, 90.59; $A100R^2$, 89.34; F, 44.93; S_{est} , 0.627; S^0 , 0.348; N_{dp} , 18.

for $^3\chi$;

$$^3\chi = 0.315 (\pm 0.0793)n_2 + 0.174 (\pm 0.0337)n_c - 0.428 (\pm 0.150) \quad (15)$$

$100R^2$, 90.86; $A100R^2$, 90.29; F, 74.57; S_{est} , 0.273; S^0 , 0.331; N_{dp} , 18.

and for $^4\chi$;

$$^4\chi = 0.224 (\pm 0.0507)n_3 + 0.0887 (\pm 0.0314)n_c - 0.226 (\pm 0.142) \quad (16)$$

$100R^2$, 89.57; $A100R^2$, 88.86; F, 64.00; S_{est} , 0.222; S^0 , 0.355; N_{dp} , 18.

The results are interesting. $^1\chi$ is a function only of polarizability, $^2\chi$ is a function only of branching, while $^3\chi$ and $^4\chi$ are a function of both. The results are compared in Figure 1.

Kier [10] has reported a correlation of pC_{50} values for the inhibition of aniline hydroxylase by alkanols

Table 5. Correlation matrix for Equation (17)

n_1	n_2	n_3	n_4	n_5	n_c	
1	0.32	0.058	0	0.05	0.51	n_1
	1	0.2	0.23	0.17	0.57	n_2
		1	0.44	0.32	0.43	n_3
			1	0.73	0.56	n_4
				1	0.49	n_5
					1	n_c

with the $^2\kappa$ and $^3\kappa$ parameters. these parameters were correlated with the equation:

$$Q = s_1n_1 + s_2n_2 + s_3n_3 + s_4n_4 + s_5n_5 + a_c n_c + a^o \quad (17)$$

The correlation matrix is given in Table 5. The regression equations obtained were:

for $^2\kappa$;

$$^2\kappa = -0.916 (\pm 0.298)n_1 - 0.760 (\pm 0.178)n_2 + 1.15 (\pm 0.125)n_c + 0.247 (\pm 0.498) \quad (18)$$

$100R^2$, 84.36; $A100R^2$, 82.52; F, 28.76; S_{est} , 0.666; S^0 , 0.442; N_{dp} , 20.

and for $^3\kappa$;

$$^3\kappa = 0.950 (\pm 0.501)n_3 + 1.75 (\pm 0.711)n_4 + 0.365 (\pm 0.0991)n_c + 1.38 (\pm 0.852) \quad (19)$$

$100R^2$, 69.24; $A100R^2$, 65.62; F, 12.00; S_{est} , 1.14; S^0 , 0.620; N_{dp} , 20.

The κ parameters depend on branching and polarizability as expected for alkyl derivatives.

Correlation of the pC_{50} values with equation 17 gave on exclusion of the value for Me the regression equation:

$$pIC_{50} = -0.356 (\pm 0.135)n_1 - 0.436 (\pm 0.0747)n_2 + 0.320 (\pm 0.0557)n_c - 0.702 (\pm 0.276) \quad (20)$$

$100R^2$, 76.63; $A100R^2$, 73.70; F, 16.39; S_{est} , 0.280; S^0 , 0.544; N_{dp} , 19.

On further exclusion of the value for Et the results were much improved. The regression equation is:

$$pIC_{50} = -0.387 (\pm 0.0910)n_1 - 0.463 (\pm 0.0506)n_2 + 0.250 (\pm 0.0406)n_c - 0.214 (\pm 0.215) \quad (21)$$

Table 6. Correlation matrix for Equation (17) with τ values

n_1	n_2	n_3	n_4	n_5	n_C	
1	0.070	0.341	0.296	0.247	0.104	n_1
	1	0.232	0.051	0.043	0.521	n_2
		1	0.431	0.360	0.590	n_3
			1	0.835	0.649	n_4
				1	0.679	n_5
					1	n_C

$100R^2$, 88.50; $A100R^2$, 86.97; F , 35.92; S_{est} , 0.188; S^o , 0.384; N_{dp} , 18.

Pal, Sengupta and De [11] have reported τ values for alkanols. The data sets (sets TP8–TP11) are given in Table 2. They were correlated with Equation (17) giving the regression equations:
for T_R ;

$$T_R = 0.365 (\pm 0.104)n_4 + 0.421 (\pm 0.0262)n_C + 0.557 (\pm 0.110) \quad (22)$$

$100R^2$, 97.70; $A100R^2$, 97.55; F , 297.6; S_{est} , 0.139; S^o , 0.167; N_{dp} , 17.

for F ;

$$F = -0.250 (\pm 0.0224)n_1 + 0.0535 (\pm 0.0271)n_2 + 1.62 (\pm 0.0454) \quad (23)$$

$100R^2$, 89.99; $A100R^2$, 89.32; F , 62.91; S_{est} , 0.0711; S^o , 0.347; N_{dp} , 17.

for T

$$T = 0.194 (\pm 0.0617)n_1 - 0.374 (\pm 0.169)n_5 + 0.444 (\pm 0.0373)n_C - 1.04 (\pm 0.150) \quad (24)$$

$100R^2$, 96.98; $A100R^2$, 96.55; F , 139.3; S_{est} , 0.176; S^o , 0.199; N_{dp} , 17.

and for B

$$B = 0.127 (\pm 0.0202)n_1 - 0.0651 (\pm 0.0368)n_4 - 0.0622 (\pm 0.0355) \quad (25)$$

$100R^2$, 79.40; $A100R^2$, 78.02; F , 26.97; S_{est} , 0.0615; S^o , 0.500; N_{dp} , 17.

The correlation matrix is given in Table 6.

T_R and T are functions of both polarizability and branching while F and B are functions only of branching.

Haloalkanes and haloalkenes

Gorgas, Seybold, and Anderson [12] have correlated solvent : air and tissue : air partition coefficients and

Composition of parameters

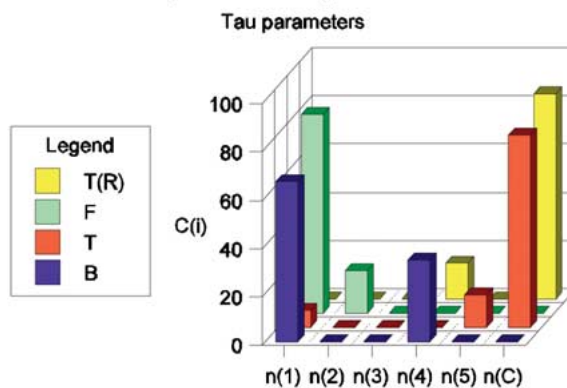


Figure 2.

Composition of parameters

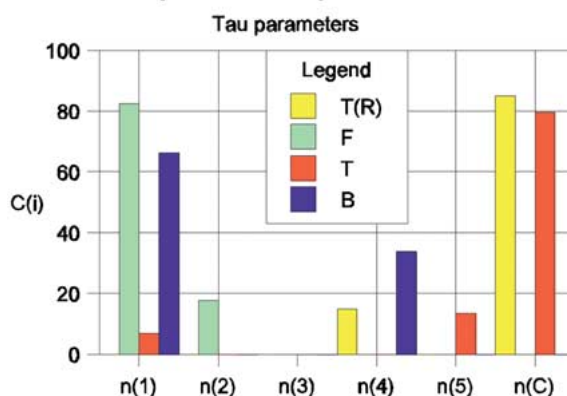


Figure 3.

V_{max} values for haloethanes, haloethanes, and haloethenes with various types of χ topological parameters and Wiener indices. Their topological parameters were correlated with the equation:

$$Q = M\mu + A\alpha + b_{b/c}/r_{b/c} + b_{Cl}n_{Cl} + b_{Br}n_{Br} + b_F n_F + b^o \quad (26)$$

Table 7. Correlation matrix for Equation (26)

μ	α	$r_{b/c}$	n_{Cl}	n_{Br}	n_F	
1	0.591	0.408	0.634	0.145	0.328	μ
	1	0.75	0.854	0.03	0.339	α
		1	0.752	0.216	0.185	$r_{b/c}$
			1	0.489	0.354	n_{Cl}
				1	0.087	n_{Br}
					1	n_F

which is a form of equation 7. The correlation matrix is given in Table 7. The regression equations obtained with the χ parameters are:

for $^1\chi$

$$^1\chi = 8.12 (\pm 0.391)\alpha - 0.221 (\pm 0.0520)n_{Br} + 0.385 (\pm 0.0319)n_F + 0.265 (\pm 0.0872) \quad (27)$$

$100R^2$, 95.38; $A100R^2$, 94.99; F , 158.2; S_{est} , 0.134; S^0 , 0.233; N_{dp} , 27.

for $^0\chi^v$

$$^0\chi^v = 11.8 (\pm 2.19)\alpha + 0.500 (\pm 0.115)n_{Cl} + 0.971 (\pm 0.185)n_{Br} + 0.291 (\pm 0.0520)n_F + 0.301 (\pm 0.192) \quad (28)$$

$100R^2$, 98.15; $A100R^2$, 97.91; F , 292.5; S_{est} , 0.216; S^0 , 0.151; N_{dp} , 27.

for $^1\chi^v$

$$^1\chi^v = 0.163 (\pm 0.0787)\mu + 6.44 (\pm 1.89)\alpha + 0.236 (\pm 0.105)n_{Cl} + 0.684 (\pm 0.163)n_{Br} + 0.0907 (\pm 0.0444)n_F - 0.144 (\pm 0.207) \quad (29)$$

$100R^2$, 94.96; $A100R^2$, 94.05; F , 79.22; S_{est} , 0.183; S^0 , 0.254; N_{dp} , 27.

for $^3\chi^v$

$$^3\chi^v = 1.21 (\pm 0.573)r_{b/cl} + 0.355 (\pm 0.113)n_{Cl} - 0.398 (\pm 0.261) \quad (30)$$

$100R^2$, 69.98; $A100R^2$, 68.78; F , 27.97; S_{est} , 0.595; S^0 , 0.581; N_{dp} , 27.

for $^4\chi_c^v$

$$^4\chi_c^v = -0.210 (\pm 0.0969)\mu + 0.401 (\pm 0.196)r_{b/cl} + 0.308 (\pm 0.164) \quad (31)$$

$100R^2$, 38.51; $A100R^2$, 36.11; F , 7.535; S_{est} , 0.281; S^0 , 0.831; N_{dp} , 27.

for $^4\chi_{pc}^v$

$$^4\chi_{pc}^v = -16.2 (\pm 2.41)\alpha + 0.367 (\pm 0.196)n_F - 2.77 (\pm 0.533) \quad (32)$$

$100R^2$, 65.44; $A100R^2$, 64.06; F , 22.72; S_{est} , 0.825; S^0 , 0.624; N_{dp} , 27.

for χ_s^v

$$\chi_s^v = 10.3 (\pm 2.51)\alpha + 0.566 (\pm 0.135)n_{Cl} + 1.47 (\pm 0.217)n_{Br} - 0.325 (\pm 0.0609)n_F - 1.45 (\pm 0.225) \quad (33)$$

Parameter Composition

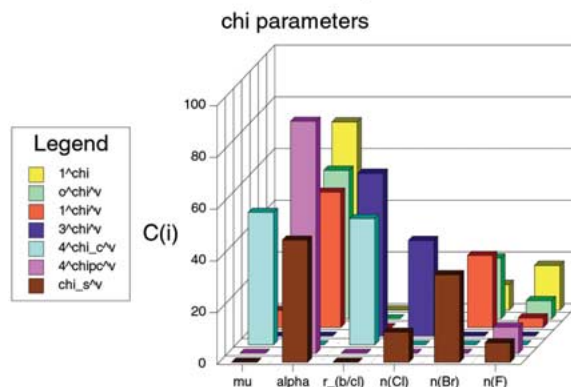


Figure 4.

$100R^2$, 82.01; $A100R^2$, 79.67; F , 25.08; S_{est} , 0.253; S^0 , 0.470; N_{dp} , 27.

for $1/^1\chi$ (set TP12R),

$$1/^1\chi = -2.86 (\pm 0.458)\alpha + 0.183 (\pm 0.104)r_{b/cl} + 0.0886 (\pm 0.0300)n_{Br} - 0.128 (\pm 0.0249)n_F + 1.10 (\pm 0.0724) \quad (34)$$

$100R^2$, 85.73; $A100R^2$, 83.86; F , 33.03; S_{est} , 0.0670; S^0 , 0.419; N_{dp} , 27.

for $1/^\circ\chi^v$ (set TP13R);

$$1/^\circ\chi^v = -1.46 (\pm 0.153)\alpha - 0.0459 (\pm 0.0216)n_{Br} + 0.585 (\pm 0.0331) \quad (35)$$

$100R^2$, 79.70; $A100R^2$, 78.88; F , 47.10; S_{est} , 0.0558; S^0 , 0.478; N_{dp} , 27.

for $1/^1\chi^v$ (TP14R);

$$1/^1\chi^v = -3.43 (\pm 0.539)\alpha - 0.147 (\pm 0.0760)n_{Br} + 1.30 (\pm 0.116) \quad (36)$$

$100R^2$, 64.45; $A100R^2$, 63.03; F , 21.75; S_{est} , 0.196; S^0 , 0.632; N_{dp} , 27.

and for $1/\chi_s^v$;

$$1/\chi_s^v = 10.4 (\pm 0.929)n_F + 1.12 (\pm 0.858) \quad (37)$$

$100R^2$, 83.26; F , 124.4; S_{est} , 4.16; S^0 , 0.425; N_{dp} , 27.

The χ parameters are a function of the polarizability, dipole moment, and the number of halogen atoms of type. In some cases there is also a dependence on $r_{b/cl}$. The results are compared in Figures 4 and 5.

Correlation of the Wiener indices with equation 26 gave the regression equations:

Table 8. Correlation matrix for Equation (44)

n_{26Me}	n_{35Me}	n_{4Me}	n_{26Cl}	n_{35Cl}	n_{4Cl}	
1	0.120	0.033	0.361	0.391	0.301	n_{26Me}
	1	0.092	0.333	0.361	0.277	n_{35Me}
		1	0.277	0.301	0.230	n_{4Me}
			1	0.120	0.092	n_{26Cl}
				1	0.033	n_{35Cl}
					1	n_{4Cl}

Table 9. Composition of TPs for alkyl groups

TP	n_1	n_2	n_3	n_4	n_5	n_C/α
$^0\chi$	0	0	0	18.7	nd	81.3
$^1\chi$	0	0	0	0	nd	100
$^2\chi$	45.1	23.5	31.4	0	nd	0
$^3\chi$	0	26.6	0	0	nd	73.4
$^4\chi$	0	0	32.5	0	nd	67.5
$^1\kappa$	12.3	10.2	0	0	0	77.5
$^3\kappa$	0	0	21.0	38.7	0	40.3
T_R	0	0	0	14.9	0	85.1
F	82.4	17.6	0	0	0	0
T	7.0	0	0	0	13.4	79.6
B	69.2	0	0	33.8	0	0

nd, not determined.

where the variable n_{ijX} is the number of X groups in the positions i and j. The correlation matrix is reported in Table 8. The regression equations obtained are: for **I**B

$$I_B = 0.0987 (\pm 0.0222)n_{26Me} + 0.0676 (\pm 0.0297)n_{35Me} + 0.348 (\pm 0.0279)n_{26Cl} + 0.297 (\pm 0.0214)n_{35Cl} + 0.236 (\pm 0.0307)n_{4Cl} + 3.10 (\pm 0.0273) \quad (45)$$

$100R^2$, 97.89; $A100R^2$, 92.13; F, 92.88; Sest, 0.0418; S^o , 0.184; N_{dp} , 16.

and for **o**B

$$B = 0.0176 (\pm 0.00652)n_{26Me} + 0.121 (\pm 0.00791)n_{26Cl} + 0.105 (\pm 0.00587)n_{35Cl} + 0.0942 (\pm 0.00870)n_{4Cl} + 2.44 (\pm 0.00628) \quad (46)$$

$100R^2$, 98.62; $A100R^2$, 98.28; F, 196.9; S_{est} , 0.0128; S^o , 0.142; N_{dp} , 16.

Table 10. Composition of TPs for haloalkanes and haloalkenes

TP	α	μ	n_{Cl}	n_{Br}	n_F	$r_{b/l}$
w_r	66.2	0	0	10.2	15.0	8.52
w^v	71.7	0	0	7.52	20.7	0
w_r^v	59.7	10.7	0	11.6	18.0	0
w	73.6	0	0	8.39	18.0	0
f	65.6	0	0	8.99	16.3	9.14
p	80.6	0	0	0	0	19.4
$1^l\chi$	65.0	0	0	10.1	14.5	10.4
$1^o\chi^v$	86.4	0	0	13.6	0	0
TP	α	μ	n_{Cl}	n_{Br}	n_F	$r_{b/l}$
$1^l\chi^v$	82.4	0	0	17.6	0	0
$1^l\chi_s^v$	0	0	0	0	100	0
$^1\chi$	72.8	0	0	9.92	17.3	0
$^o\chi^v$	57.4	0	12.1	23.5	7.04	0
$^1\chi^v$	52.3	6.63	9.58	27.8	3.69	0
$^3\chi^v$	0	0	37.0	0	0	63.0
$^4\chi_c^v$	0	51.2	0	0	0	48.8
$^4\chi_{pc}^v$	89.8	0	0	0	0	10.2
χ_s^v	47.2	0	11.6	33.8	7.46	0

Discussion

Composition of topological parameters

C_i values for topological parameters of alkyl and substituted alkyl groups are reported in Table 9. All of the parameters studied are a function of n_i (branching), n_C or α (polarizability), or both. C_i values for topological parameters of haloalkanes, haloalkenes, and haloanilines are reported in Table 10. They are a function of n_{Xi} , α , μ , and $r_{b/Cl}$. They are therefore a function of steric effects, flexibility, bond moment, and polarizability. We have not studied systems in which significant hydrogen bonding occurs.

In most cases topological parameters are composite parameters.

How topological parameters work

We have shown in the previous paper of this series [1] that topological parameters are not fundamental because they cannot account for the properties of isotopological data sets [1]. We have further shown in this work that they are counts of the number of atoms of each kind in a structural feature as well as the number of branches, and that they are generally composite parameters. How does this relate to the modeling of properties and bioactivities? It has been shown that

Table 11. Intermolecular forces and the quantities upon which they depend [15]

Intermolecular force	Quantity
<i>molecule–molecule</i>	
Hydrogen bonding (hb)	q _{MH} , q _{ME} , orbital type
Dipole–dipole (dd)	dipole moment
Dipole–induced dipole (di)	dipole moment, polarizability
Induced dipole–induced dipole (ii)	polarizability
Charge transfer (ct)	ionization potential, electron affinity
<i>ion–molecule</i>	
ion–dipole (Id)	ionic charge, dipole moment
ion–induced dipole (Ii)	ionic charge, polarizability

Abbreviations are in parentheses. dd interactions are also known as Keesom interactions; di interactions are also known as Debye interactions; ii interactions are also known as London or dispersion interactions. Collectively, dd, di, and ii interactions are known as Van der Waals interactions. Charge transfer interactions are also known as donor-acceptor interactions.

the variation of both the properties and the biological activities as a function of structure is dependent on the difference in intermolecular forces between an initial and a final state and on steric effects [7,14]. The intermolecular forces involved are set forth in Table 11. The fundamental quantities they depend on are polarizability, bond moment, and when appropriate a measure of hydrogen bonding.

Polarizability is directly related to the number of atoms of each type in the variable structural feature or alternatively to the number of core, valence and non-bonding electrons in the structural feature [15]. Thus, the advantage of α is that it is easily estimated either by additivity from the values for fragments or from group molar refractivities calculated from the equation

$$MR_X = 0.320n_c + 0.682n_b - 0.0825n_n + 0.991 \quad (47)$$

where n_c , n_b , and n_n are the number of core, bonding, and nonbonding electrons respectively in the group X. The relation between MR_X and α , the polarizability parameter used in this work, is given by equation 4. The bond moment is directly related to the difference in electronegativity between the bonded atoms. The Allred–Rochow electronegativity [16] for the elements with atomic numbers Z equal to 3–9, 11–27, and 30–56 were correlated with the Equation (17):

$$\chi_{AR} = a_1 1/r_C^2 + a_2 Z + a_3 n_{e,hopo} + a_4 n_{e,hodo} + a_o \quad (48)$$

where χ_{AR} is the Allred - Rochow electronegativity, $1/r_C^2$ is the reciprocal of the square of the covalent radius of the atom with atomic number Z , $n_{e,hopo}$ and $n_{e,hodo}$ are the numbers of electrons in the outermost

Table 12. Correlation matrix for Equation (48)

$1/r_C^2$	Z	$n_{e,hopo}$	$n_{e,hodo}$	
1	0.536	0.319	0.006	$1/r_C^2$
	1	0.019	0.262	Z
		1	0.454	$n_{e,hopo}$
			1	$n_{e,hodo}$

p and d orbitals respectively. The correlation matrix is given in Table 12. The regression equation is:

$$\begin{aligned} \chi_{AR} = & 0.858 (\pm 0.0733) 1/r_C^2 \\ & - 0.00204 (\pm 0.00176) Z \\ & + 0.272 (\pm 0.0132) n_{e,hopo} \\ & + 0.0263 (\pm 0.00881) n_{e,hodo} \\ & + 0.740 (\pm 0.0850) \end{aligned} \quad (49)$$

100R², 96.52; F, 311.7; S_{est}, 0.143; N_{dp}, 50.

Clearly, electronegativities, and therefore bond moments, will be encompassed in counts of atoms of each type present in the structural feature. As the molecular dipole moment is the vector sum of the bond moments it too may be accounted for. Finally, the number of OH and NH bonds is a rough measure of hydrogen donor hydrogen bonding capability, and a count of the number of O or N atoms is a rough measure of the hydrogen acceptor hydrogen bonding capacity [18]. We have shown elsewhere that composite parameters may be used in place of pure parameters to obtain a predictive model, the only disadvantage being the nonexplicative nature of the resulting model.

Obviously, a mix of pure and composite parameters would work as well.

Authors have often modified topological parameters by taking the reciprocal. A comparison of the composition for the four parameters for which the reciprocal composition was also determined shows that this is indeed a method of varying parameter composition.

Conclusions

Topological parameters are generally composite parameters. Their composition depends on the algorithm used to generate them and any further mathematical operations to which they may be subjected such as taking the reciprocal. They represent counts of the numbers of atoms of each kind, the numbers of bonds of each type, of core, valence (bonding), and of non-bonding electrons, of branches, of OH and NH bonds and of lone pairs (full nonbonding orbitals) on O and N in some variable structural feature. **They are not fundamental parameters and have no direct relationship to any property.** The quantities of which they are counts are related to polarizability, dipole moment, hydrogen bonding and steric effects which in turn are the quantities that actually determine properties and biological activities. Topological parameters are useful; in the determination of empirical relationships that can be used for predictive purposes. They have the advantage that they can be generated by algorithms for any structural variation. They have the disadvantage that they are not directly explicative.

Generally, in order to be used successfully, two or more topological parameters are required. There is no *a priori* method for predicting which topological parameters are required in a particular case.

References

1. Charton, M., Perspectives in Drug Design, 3 (2003) in press.
2. Kier, L.B. and Hall, L.H., Molecular Connectivity in Chemistry and Drug Research, Academic Press, New York, 1976.
3. Charton, M., Stud. Org. Chem., 42 (1992) 629.
4. Charton, M., in The Chemistry of Dienes and Polyenes, Z. Rapoport, Ed., Wiley, Chichester, (1997) 683.
5. Charton, M., Prog. Phys. Org. Chem., 13 (1981) 119.
6. Charton, M., Environ. Health Perspect., 61 (1985) 229.
7. Charton, M., in Classical and 3-D QSAR in Agrochemistry and Toxicology, Hansch, C. and Fujita, T. (eds.), American Chemical Society, Washington, D.C., (1995) 75.
8. Protic, M. and Sablic, A., Aquat. Toxicol., 14 (1989) 47.
9. Murray, W., J. Pharm. Sci., 66 (1977) 1352.
10. Kier, L., Med. Res. Rev., 7 (1987) 417.
11. Pal, D.K., Sengupta, C. and De, A.U., Ind. J. Chem., 27b (1988) 734.
12. Gorgas, M.L., Seybold, P.G. and Anderson, M.E., Toxicol. Lett., 43 (1988) 235.
13. Pyka, A., J. Planar Chromatogr., 11 (1998) 61.
14. Charton, M., Advances in Quantitative Structure Property Relationships, 2 (1999) 177.
15. Charton, M., in The Chemistry of Organic Germanium, Tin and Lead Compounds, Vol. 2, Z. Rapoport, Ed., Wiley, Chichester, (2002) 537.
16. Allred, A.L. and Rochow, E.G., J. Inorg. Nucl. Chem., 5 (1958) 264; Little, E.J. and Jones, M.M., J. Chem. Ed., 37 (1960) 231; Allen, L.C. and Huheey, J.E., J. Inorg. Nucl. Chem., 42 (1980) 1523.
17. This work.
18. Charton, M., Prog. Phys. Org. Chem., 18 (1990) 163.