The nature of topological parameters. I. Are topological parameters 'fundamental properties'?

Marvin Charton
Chemistry Department, Pratt Institute, Brooklyn, NY 11205, USA (E-mail: mcharton@pratt.edu)

Summary

Topological parameters are often used descriptors in QSAR and QSPR studies. Though easily generated from an appropriate algorithm useful in developing predictive models they are not readily interpretable. Though every molecule has an intrinsic topology it is not clear whether topological parameters are 'fundamental' properties or not. The results of correlations of various properties of sets of molecules with constant topology shows that are a function of the polarizability α alone in the case of nonpolar molecules and of both dipole moment μ and polarizability α in the case of polar molecules. Properties studied include melting point, boiling point, critical temperature, heat of melting, heat of vaporization, solubilties of gases in various solvents, and the ratio of the Van der Waals constants (a/b). Types of molecules studied include monatomic (rare gases), diatomic, trigonal pyramidal, and tetrahedral. Our results show that these properties are successfully modeled by α and μ. As they are isotopological in each data set they cannot be modeled by topological parameters. It follows then that topological parameters do not have a direct cause and effect relationship with these properties and therefore are not fundamental parameters. The conclusion of Kier and Hall that they represent electron accessibility is incorrect. Topological parameters are composite parameters representing counts of the numbers of atoms, bonds, electrons, and branching.

Abbreviations: T_M – melting point; T_B – boiling point; T_C – critical temperature; ΔH_m – heat of melting; ΔH_v – heat of vaporation; TP – topological parameters; μ – dipole moment; σ; α – polarizability; P_M – atomic polarizability; $100R^2$ and $100r^2$ – correlation coefficient X 100; $A100R^2$ – adjusted correlation coefficient X 100; r_{ij} – zeroth order correlation coefficients between independent variables; $r_{df/iv}$ – ration of the degrees of freedom to the number of independent variables; C – the percent contribution of each independent variable in the regression equation; ω – indicator variable; S_{est} – standard error of estimate; S° – standard error of estimate corrected for the nature of the data; N_{dp} – number of data points in the regression; A – coefficient of α in the equation; M – coefficient of μ in the equation.

Introduction

Though frequently used as descriptors in QSAR and QSPR studies topological parameters (TPs) [1–3] are not well understood. They belong to a category best described as black box parameters, that is, the meaning of the relationship between the parameters and the quantities they model is unclear. Here we present results which provide insight into the nature of TPs and the way in which they work.

If topological parameters are 'fundamental properties' they must account for molecular properties in all molecules. If the molecular properties of a set of molecules **with constant topology** can be accounted for by some group of parameters then we may reach the following conclusions:

1. As the topology is constant the properties cannot be accounted for by topological parameters.
2. As topological parameters cannot model the data while other parameters can topological parameters are not fundamental.

This **isotopological method** provides a test for determining the nature of topological parameters.

We have applied this method to the phase change properties T_b, T_m, T_c, ΔH_m, and ΔH_v;[1] the solubilities of the rare gases and of diatomic molecules in

Table 1. Intermolecular forces and the quantities upon which they depend

Intermolecular force	Quantity
molecule–molecule	
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
ion–molecule	
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.

various solvents, and the ratio a/b of the constants of the Van der Waals equation of state [4–6].

Method

Parameters

Many properties are a function of the difference in intermolecuar forces (imf) between compounds in an initial and a final phase. The change in the free energy associated with the contribution of these forces to the property is given by the expression:

$$Q = f(E_{imf_f} - E_{imf_i}) = f(\Delta E_{imf}) \qquad (1)$$

These properties, and many others, are a function of molecular structure. Such properties can be modeled by parameterizing the imf involved as functions of structural variation and using the resulting correlation equation to obtain a quantitative model of the property of interest by means of linear regression analysis. Commonly encountered imf are summarized in Table 1 [7].

In the data sets studied in this work only van der Waals interactions are involved. These interactions (dd, di, ii) are a function of dipole moment and polarizability. It has been shown that the dipole moment cannot be replaced entirely by the use of electrical effect substituent constants as parameters [8, 9]. This is because the dipole moment has no sign. Either an overall electron donor group or an overall electron acceptor group may have the same value of μ. It has also been shown that the bond moment rather than the molecular

dipole moment is the parameter of choice. The dipole moments of MeX and PhX were taken as measures of the bond moments of substituents bonded to sp^3 and sp^2 hybridized carbon atoms respectively of a skeletal group. Application to substituents bonded to sp hybridized carbon atoms should require a set of dipole moments for substituted ethynes. Dipole moments for MeX and PhX can be calculated from the equations [8]:

$$\mu_{MeX} = 5.11(\pm0.497)\,\sigma_{lX} + 1.99(\pm0.541)\,\sigma_{dX} \\ + 0.0129(\pm0.205) \qquad (2)$$

and:

$$\mu_{PhX} = 5.47(\pm0.150)\,\sigma_{lX} + 4.30(\pm0.446)\,\sigma_{dX} \\ + 6.94(\pm1.91)\,\sigma_{eX} + 0.420(\pm0.172)\,(3)$$

The polarizability parameter generally used in this work, α, is given by the expression

$$\alpha \equiv \frac{M\,R_X - M\,R_H}{100} = \frac{M\,R_X}{100} - 0.0103 \qquad (4)$$

where MR_X and MR_H are the group molar refractivities of X and H [10–12] 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. If all of the carbon atoms in a compound have the same hybridization then n_C, the number of these atoms, can be used as a polarizability parameter when it varies. There are many other polarizability parameters including parachor, group molar volumes of various kinds, van der Waals volumes, and accessible surface areas any of which will do as well as they are all highly collinear

Table 2. ΔP_M and α values of the atoms M

M, ΔP, α: H, 0, 0; F, -0.11, -0.001; Cl, 1.513, 0.0510; Br , 2.383, 0.079; I, 4.033, 0.129; O, 0.135, 0.015; S, 2.233, 0.082; Se, 3.103; 0.114; N, 0.433, 0.038; P, 2.963, 0.107; C, 1.093, 0.039; He, -0.462, -0.017; Ne, -0.271, -0.010; Ar, 0.974, 0.029; Kr, 1.818, 0.056; Xe, 3.377, 0.106; Rn, 4.633, 0.146.

α values for the rare gases are calculated from equation 10.

in each other [13]. Proposing polarizability parameters remains a popular sport. Values of α can be estimated by additivity from the values for fragments or from group molar refractivities calculated from the equation

$$MR_X = 0.320\,n_c + 0.682\,n_b - 0.0825\,n_n + 0.991 \qquad (5)$$

where n_c, n_b, and n_n are the number of core, bonding, and nonbonding electrons respectively in the group X [13].

Atomic polarizabilities P_M for atoms M were tabulated by Miller [14]. While they may be used as polarizability parameters they are not directly comparable to α values. Coefficients obtained from correlations with α are not directly comparable to those obtained with P_M. In order to obtain α values for some atoms we have first defined:

$$\Delta P_M \equiv P_M - P_H \qquad (6)$$

where M is the atom of interest and H is hydrogen. This is directly comparable to the definition of α in equation 4. ΔP_M values for the set H, F, Cl, Br, I (Table 2) were then correlated with the equation:

$$\Delta P_M = A\,\alpha_M + a_o \qquad (7)$$

to give the regression equation:

$$\Delta P_{at,M} = 31.4(\pm 0.428)\alpha - 0.0507(\pm 0.0305) \qquad (8)$$

$100r^2$, 99.94; F, 5396; S_{est}, 0.0472; S°, 0.0304; N_{dp}, 5.

Inclusion in the data set of the $\Delta P_{at,M}$ values for N, P, O, S, and Se (Table 2) gave the regression equation:

$$\Delta P_M = 30.6(\pm 1.70)\,\alpha_M - 0.197(\pm 0.126) \qquad (9)$$

$100r^2$, 97.26; F, 320.0; S_{est}, 0.248; S°, 0.183; N_{dp}, 11.

Though the goodness of fit of equation 9 is decidedly poorer than that of equation 8, the values of A are not significantly different and the intercepts in both equations are not significant, furthermore, both equations are significant at the 99.9% confidence level. $\Delta P_{at,M}$ values for the Group 14 elements seem to lie on their own line. We have used equation 8 in the form:

$$\alpha_M = 0.0318\Delta P_M - 0.00183 \qquad (10)$$

Table 3. Parameter values for sets TB1 to TB9, TM1 to TM5, and TC1 to TC3

Sets TB1, TM1, TC1, DHM1, DHV1, S11–S16, VdW1. M, α; He, -0.017; Ne, -0.010; Ar, 0.029; Kr, 0.056; Xe, 0.106; Rn, 0.146.
Sets TB2, TM2, TC2, DHM2, DHV2. S21, S22, VdW2. M^1M^2, α, μ, ω; F_2, -0.002, 0, 1; Cl_2, 0.100, 0, 1; Br_2, 0.158, 0, 1; I_2, 0.258, 0, 1, H_2, 0, 0, 1; N_2, 0.031, 0, 1; O_2, 0.028, 0, 1; CO, 0.031, 0.128, 0.5; NO, 0.030, 0.158, 0.5; ICl, 0.179, 0.65, 0.5; IBr, 0.208, 1.21, 0.5; ClF, 0.049, 0.881, 0.5; BrCl, 0.129, 0.57; 0.5; BrF, 0.078, 1.29, 0.5; HCl, 0.060, 1.0929, 0.5; HBr, 0.089, 0.826, 0.5; HI, 0.139, 0.4475, 0.5.
Sets TB4, TM4, DHM4, DHV4, VdW4. Trigonal pyramidal MZ_3, α, μ; NF_3, 0.020, 0.234; NMe_3, 0.191, 0.612; PH_3, 0.122, 0.574; PF_3, 0.088, 1.03; PCl_3, 0.241, 0.56; PI_3, 0.478, 0; PMe_3, 0.248, 1.192; AsH_3, 0.156, 0.22; AsF_3, 0.141, 2.65; $AsCl_3$, 0.294, 2.19; $AsBr_3$, 0.281, 1.66; AsI_3, 0.531, 0.96; $AsMe_3$, 0.282, 1.10; SbH_3, 0.189, 0.116; $SbCl_3$, 0.308, 3.93; $SbBr_3$, 0.395, 2.48; SbI_3, .545, 1.59.
Sets TB5, TM5, TC5, DHM5, DHV5, VdW5. Tetrahedral $MX^1X^2X^3X^4$, α, μ, ω; SiH_4, 0.111, 0, 1; SiF_4, 0.097, 0, 1; $SiCl_4$, 0.301, 0, 1; $SiBr_4$, 0.417, 0, 1; SiI_4, 0.587, 0, 1; $SiMe_4$, 0.285, 0, 1; GeH_4, 0.126, 0, 1; $GeCl_4$, 0.316, 0, 1; $GeBr_4$, 0.432, 0, 1; GeI_4, 0.632, 0, 1; $GeMe_4$, 0.300, 0, 1; SnH_4, 0.179, 0, 1; $SnCl_4$, 0.369, 0, 1; $SnBr_4$, 0.485, 0, 1; SnI_4, 0.685, 0, 1; $SnMe_4$, 0.353, 0, 1; $PbCl_4$, 0.386, 0, 1; $PbMe_4$, 0.370, 0, 1; SiH_3Br; 0.180, 1.32, 0.75; SiH_3Cl, 0.151, 1.292, 0.75; $SiBr_3H$, 0.348, 0.79, 0.25; SiH_2Cl_2, 0.191, 1.181, 0.50; $SiCl_3H$, 0.261, 0.855, 0.25; SiF_3Cl 0.148, 0.49, 0.75; SiH_3I, 0.230, 1.62, 0.75; SiF_3H, 0.077, 1.26, 0.75; $SiCl_3Me$, 0.296, 1.91, 1; $SiMe_2H_2$, 0.183, 0.75, 0.50; SiH_3Me, 0.147, 0.73, 0.75; $SiClMe_3$, 0.289, 2.02, 1; GeH_3Br, 0.205, 1.31, 0.75; GeH_3Cl, 0.176, 2.124, 0.75; GeH_2Cl_2, 0.216, 2.22, 0.50; CH_4, 0.046, 0, 1; CF_4, 0.042, 0, 1; CCl_4, 0.237, 0, 1; CMe_4, 0.232, 0, 1, $CHCl_3$, 0.191, 1.03, 0.25; CH_2Cl_2, 0.146, 1.62,0.50; CH_3Cl, 0.096, 1.94, 0.75; CH_3F, 0.45, 1.822, 1; $CClF_3$, 0.093, 0.50, 0.75, $CClMe_3$, 0.234, 2.14, 1; CF_2Br_2, 0.202, 0.66, 0.50; $CHFCl_2$, 0.145, 1.30, 0.50.

to calculate α values for the rare gas elements. These values are also given in Table 2.

The parameters used in this work are set forth in Table 3.

Statistics

The data sets were correlated by simple or multiple linear regression analysis depending on the number of independent variables in the correlation equation. The statistics reported for multiple linear regression analysis include the $100R^2$ statistic which represents the per cent of the variance of the data accounted for by the regression equation. The $A100R^2$ statistic is the adjusted value of $100R^2$, it takes into account the number of independent variables in the regression

Table 4. Monatomic species (rare gases). Values of T_b, T_m, T_c, ΔH_m, ΔH_v used in correlations

Set Q	TB1 T_b	TM1 T_m	TC1 T_c	DHM1 ΔH_m	DHV1 ΔH_v
He	4.215	–	5.19	0.0138	0.0829
Ne	27.05	24.55	44.44	0.335	1.71
Ar	87.28	83.77	150.71	1.12	6.43
Kr	119.75	115.95	209.4	1.37	9.08
Xe	165.05	161.35	289.74	1.81	12.64
Rn	211.15	202.15	376.99	3.24	18.10

Temperatures in degrees Kelvin, ΔH values in Kj/mole.

Table 5. Results of correlations for monatomic species M (rare gases)

Set Q	TB1 T_b	TM1 T_m	TC1 T_c	DHM1 ΔH_m	DHV1 ΔH_v
A	1219	1106	2186	17.4	104
S_A	82.1	71.5	138	2.14	5.72
a_o	−39.5	45.2	872	0.417	2.63
S_{ao}	6.44	6.12	28.0	0.160	0.449
$100r^2$	98.22	98.76	98.44	94.79	98.81
F	220.2	239.4	252.5	72.81	331.2
S_{est}	11.9	17.3	19.9	0.295	0.827
S^o	0.164	0.135	0.153	0.279	0.134
N_{dp}	6	5	6	6	6
$r_{df/iv}$	4.00	3.00	4.00	4.00	4.00

equation. The F statistic is a measure of the significance of the correlation, its confidence level is given in parentheses. The S_{est} statistic is the standard error of the estimate, it is useful in detecting outliers. It is of no use as a measure of the goodness of fit of different

Table 6. Homonuclear diatomic species. Values of T_b, T_m, T_c, ΔH_m, ΔH_v used in correlations

Set Q	TB2a T_b	TM2a T_m	TC2a T_c	DHM2a ΔH_m	DHV2a ΔH_v
H_2	20.28	14.01	33.3	0.117	0.904
N_2	77.15	63.29	126	0.720	5.577
O_2	90.11	54.75	155	0.444	6.820
F_2	85.01	53.53	144	0.510	6.62
Cl_2	238.55	172.17	417	6.406	20.41
Br_2	331.93	265.95	584	10.57	29.96
I_2	457.5	386.65	785	15.066	41.6

Temperatures in degrees Kelvin, ΔH values in Kj/mole.

Table 7. Results of correlations for homonuclear diatomic molecules M_2

Set Q	TB2a T_b	TM2a T_m	TC2a T_c	DHM2a ΔH_m	DHV2a ΔH_v
A	1643	1413	2852	61.4	154
S_A	123	72.8	231	4.04	11.3
a_o	51.3	28.6	87.2	−0.194	3.55
S_{ao}	15.0	7.87	28.0	0.491	1.37
$100r^2$	97.26	98.69	96.83	97.88	97.38
F	177.2	377.3	152.7	238.7	185.9
S_{est}	29.3	17.3	54.8	0.960	2.69
S^o	0.196	0.135	0.211	0.172	0.192
N_{dp}	7	7	7	7	7

$r_{df/iv}$ is 5.00 for all of these data sets.

Table 8. Heteronuclear diatomic species. Values of T_b, T_m, T_c, ΔH_m, ΔH_v used in correlations

Set Q	TB2b T_b	TM2b T_m	TC2b T_c	DHM2b ΔH_m	DHV2b ΔH_v
NO	121.35	109.55	180	2.30	13.83
CO	81.65	74.15	133	0.837	6.04
BrCl	*278.15*	*207.15*	–	10.4	34.7
BrF	253.15	240.15d	–	–	25.1
ClF	172.35	119.15	–	–	24
IBr	389.15d	315.15	–	–	–
ICl	370.55	300.35(α), 287.07(β)	–	11.60	–
HCl	188.25	158.35	325	1.992	16.14
HBr	206.15	184.65	363	2.406	17.61
HI	237.8	222.35	423	2.87	19.77

Temperatures in degrees Kelvin, ΔH values in Kj/mole.

data sets to a model if they involve different types of data as it depends on the nature of the data correlated. The S° statistic may be thought of as a standard error of the estimate which has been corrected for the nature of the data, it is a useful measure of the goodness of fit of the data to the regression equation. With the exception of $100R^2$ which is replaced by $100r^2$ the same statistics are used for simple linear regression analysis. The regression equation reported is that which results after the elimination of variables whose coefficients are not significant. Also reported for MLRA are the zeroth order correlation coefficients r_{ij} that are a measure of colinearity among the independent variables. The quantity $r_{df/iv}$ is the ratio of the number of degrees of freedom, df, to the number of independent

Table 9. Results of correlations for homo- and heteronuclear diatomic molecules M^{1}M^2

Set Q	TB2′ T$_b$	TM2′ T$_m$	TC2′ T$_c$	DHM2′ ΔH_m	DHV2′ ΔH_v
A	1582	1333	2806	64.8	155
S$_A$	85.7	62.0	197	4.77	15.8
M	33.8	19.9	–	–	6.10
S$_M$	13.4	11.1	–	–	2.51
a$_o$	54.7	37.5	90.1	−0.461	3.39
S$_{ao}$	10.5	7.60	21.0	0.533	1.84
100R^2	96.74	97.59	95.30[a]	94.37[a,b]	90.26[c]
A100R^2	96.51	97.42	–	–	89.45
F	193.0	263.2	202.9	184.4	50.98
S$_{est}$	25.1	18.1	50.3	1.30	4.10
S°	0.200	0.172	0.237	0.258	0.352
r$_{\alpha\mu}$	0.220	0.273	0.020	0.100	0.015
r$_{\alpha\_}$	–	0.129	–	–	–
r$_{\mu\_}$	–	0.774	–	–	–
N$_{dp}$	16	16	12	13	14
C$_\alpha$	90.3	93.1	100	100	71.7
C$_\mu$	9.66	6.94	–	–	28.3
r$_{df/iv}$	6.50	6.50	4.50	5.00	5.50

[a] 100r^2 value.
[b] Value for HI excluded.
[c] Value for ClF excluded.

variables, iv. The Statistics are reported either beneath the regression equation or in the appropriate Table.

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: [7]

$$C_i = (100\, a_i\, x_i) / \left(\sum_1^m a_i\, x_i \right) \qquad (11)$$

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.

Table 10. Trigonal pyramidal tetratomic species MZ$_3$. Values of T$_b$, T$_m$, ΔH_m, and ΔH_v used in correlations

Set Q	TB4p T$_b$	TM4p T$_m$	DHM4p ΔH_m	DHV4p ΔH_v
NF$_3$	144.09	66.37	–	11.6
NCl$_3$	344.15^x	246.15^x	–	–
NMe$_3$	276.05	156.05	–	–
PH$_3$	185.37	139.34	1.130	14.60
PF$_3$	171.77	121.85	–	16.5
PCl$_3$	348.15	182.15	7.10	30.5
PBr$_3$	374.15^{d,x}	>375.15^{d,x}	–	38.8
PI$_3$	473.15^{d,x}	334.65	–	43.9
PMe$_3$	310.95	188.15	–	–
AsH$_3$	210.65	156.25	–	16.7
AsF$_3$	324.15	267.20	10.4	29.7
AsCl$_3$	403.15	257.15	10.1	35.0
AsBr$_3$	494.15	304.15	11.7	41.8
AsI$_3$	676.15	414.95	–	59.3
AsMe$_3$	324.15	185.85	–	–
SbH$_3$	254.75	181.65	–	21.3
SbCl$_3$	493.15	346.55	12.7	45.2
SbBr$_3$	562.15	368.15	14.6	59
SbI$_3$	673.15	444.15	22.8	68.6
SbMe$_3$	353.75	211.15	–	–
BiCl$_3$	720.15	505.15	10.9	72.6
BiBr$_3$	733.15	491.15	21.7	75.4
BiI$_3$	773.15^{d,x}	681.15	–	20.9^x
BiMe$_3$	383.15	165.45	–	–

Temperatures in degrees Kelvin, ΔH values in Kj/mole.
d = decomposes; x = excluded from the correlation

Results

Phase change properties

We present results for the correlation of values of T$_m$, T$_b$, T$_c$, ΔH_m, and ΔH_v for the monatomic rare gases, M; homonuclear diatomic molecules, M$_2$; and the pentatomic tetrahedral molecules MZ$_4$. The values used are given in Tables 4, 6, and 12. The data were correlated with the equation:

$$Q_{Mol} = A\, \alpha_{Mol} + a_0 \qquad (12)$$

where Q is the property to be correlated, Mol designates the type of molecule, α is a polarizability parameter, A and B$_o$ are the coefficient of α respectively. Results of the correlations are reported in Tables 5, 7, 13, and 14. The data have come from a number of different sources. The fit of the model to the data is very good.

We have also studied the correlation of T$_b$, T$_c$, ΔH_m, and ΔH_v values for combined data sets of

Table 11. Results of correlations for trigonal pyramidal tetraatomic molecules, MZ_3

Set Q	TB4p T_b (MZ_3)	TM4p T_m (MZ_3)	DHM4p ΔH_m	DHV4p ΔH_v
A	1028	618	37.2	95.2
S_A	84.2	53.5	7.73	6.76
M	29.7	34.0	1.31	4.36
S_M	11.6	7.59	0.893	0.901
a_o	72.7	36.4	−2.72	2.55
S_{ao}	23.6	17.3	2.87	2.32
$100R^2$	93.48	92.73	85.37	95.53
$A100R^2$	93.05	92.25	82.93	95.19
F	100.4	89.3	14.59	128.2
S_{est}	45.0	31.8	2.80	4.05
$S°$	0.281	0.297	0.484	0.236
$r_{\alpha\mu}$	0.344	0.185	0.193	0.189
N_{dp}	17	17	8	15
C_α	87.4	78.4	74.0	68.6
C_μ	12.6	21.6	26.0	31.4
$r_{df/iv}$	7.00	7.00	2.50	6.00

homo- and heteronuclear diatomic molecules, and of trigonal pyramidaltetratomic molecules MZ_3 as well as T_m, T_b, T_c, ΔH_m, and ΔH_v values with the correlation equation:

$$Q_{Mol} = A\,\alpha_{Mol} + M\,\mu_{Mol} + a_0 \qquad (13)$$

The data used in the correlations are reported in Tables 8, 10, and 12. The results of the correlatioms are given in Tables 9, 11, 13, and 14. Polarizability is the major descriptor in all of the sets studied. For the diatomic and pentatomic molecules the values of A in the combined sets are in excellent agreement with those in the M_2 and MZ_4 sets. The fit of the data to the model is very good.

The combination of T_m values for homo- and heteronuclear diatomic molecules into a single data set (set TM2') may require a term in ω, the probability that the molecule approaches the crystal surface in the proper direction for binding. ω takes the values 1 for a homonuclear molecule and 0.5 for a heteronuclear molecule [15]. Thus, the correlation equation becomes:

$$T_m = A\alpha + M\mu + P\omega + a_o \qquad (14)$$

Though no dependence on ω was observed (Table 9) $r_{\mu\omega}$ shows that μ and ω are significantly colinear. The most important descriptor again is polarizability.

Table 12. Tetrahedral pentatomic molecules MZ_4. Values of T_b

M/Z	C	Si	Ge	Sn	Pb
			T_b		
H	91.15	161.25	184.65	221.15	–
F	123.15	187.15	–	–	–
Cl	250.15	330.75	356.25	388.15	–
Br	365.15	422.15	459.65	480.15	–
I	–d	576.15	621.15	637.65	–
Me	256.65	299.65	316.55	351.15	383.15
			T_m		
H		88.45	107.25	123.15	–
F		152.95	–	–	–
Cl		203.15	223.65	239.15	258.15
Br		278.55	299.25	303.15	–
I	171d	393.65	417.15	417.65	–
Me		173.65	185.15	218.35	245.65
			T_c		
H	190.4	270			
F	227.6	259			
Cl	556.4	508.1	551	592	
Br	–	663			
I	–	944			
Me	433.8				
			ΔH_m		
H	0.94	0.67	–	–	–
F	0.7	–	–	–	–
Cl	3.28	7.60	–	9.20	–
Br	–	–	–	11.9	–
I	–	19.7	–	*18.1*	–
Me	3.10	6.88	–	–	10.86
			ΔH_v		
H	8.2	12.1	14.1	19.1	–
F	12.6	–	–	–	–
Cl	29.8	28.7	27.9	34.9	–
Br	48.7	37.9	41.4	43.5	–
I	–	56.9	–	–	–
Me	22.7	–	27.0x	30.6x	–

d = decomposes; x = excluded from the correlation.

Again, in the correlation of T_m for the combined set of MZ_4 and $MZ_n^1Z_{4-n}^2$ pentatomic molecules (set TM5'), it is necessary to consider that while all four surfaces of the tetrahedral molecule MZ_4 are equivalent those of $MZ_n^1Z_{4-n}^2$ molecules are not. Thus, $MZ_2^1Z_2^2$ has two sets of two equivalent faces while $MZ_3^1Z^2$ has a set of three equivalent faces and a set of one face. We assume that the face with the largest

Table 13. Results for tetrahedral pentatomic molecules, MZ_4 and $MZ_nZ'_{4-n}$. T_b, T_m, T_c

Q	T_b		T_m		T_c
Set	TB5	TB5$'$	TM5	TM5$'$	TC5
	MZ_4	$MZ_4 + MZ_nZ'_{4-n}$	MZ_4	$MZ_4 + MZ_nZ'_{4-n}$	MZ_4
A	836	831	613	594	1347
S_A	20.0	21.5	19.9	18.7	55.3
M	–	30.9	–	52.3	–
S_M	–	4.50	–	5.63	–
P	–	–	–	97.6	–
S_P	–	–	–	18.1	–
a_o	78.0	78.8	17.3	−72.2	121
S_{ao}	7.87	8.21	8.05	18.5	19.4
$100r^2$	99.15	98.32	98.44	97.97	99.16
$A100R^2$	–	98.26	–	97.80	–
F	1251	788.8	946.5	985.4	592.6
S_{est}	14.0	16.4	12.9	13.5	23.6
$S°$	0.0981		0.133		0.108
N_{dp}	17	30	17	28	7
$r_{\alpha\mu}$	–	0.137	–	0.154	–
$r_{\alpha\omega}$	–	–	–	0.402	–
$r_{\mu\omega}$	–	–	–	0.726	–
C_α	100	84.3	100	44.2	100
C_μ	–	15.7	–	19.5	–
C_ω	–	–	–	36.3	–
$r_{df/iv}$	15.0	13.5	15.0	12.5	2.50

value of $\Sigma\alpha_Z$ will preferentially bind to the crystal surface. Then we assign values of the probability ω as 1 for all MZ_4, 0.75 to $MZ_3^1Z^2$ when the set of three equivalent faces has the higher value of $\Sigma\alpha_X$, 0.25 when the reverse is the case, and 0.50 to $MZ_2^1Z_2^2$. The correlation equation is again Equation (14). Results of the correlation are given in Table 13. Though the value of $r_{\mu\omega}$ shows significant colinearity between μ and ω the term in the latter is significant and α remains the most important descriptor.

Solubility of gases in liquids

We now examine further evidence derived from the solubilities of the rare gases and of diatomic molecules in various solvents. The solubilities are reported as χ_G, the mole fraction of gas in solution. At a constant temperature, $\log \chi_G$ is proportional to ΔG^o, the free energy of solution of the gas. Thus:

$$\log \chi_g = \frac{\Delta G^o}{2.303RT} \tag{15}$$

where χ_g is the mole fraction of the gas in solution and ΔG^o is the Gibbs free energy of solution in J/mole. At 298.15 degrees Kelvin:

$$\log \chi_g = C \Delta G^o \tag{16}$$

For the rare gases polarizability is the only variable. The correlation equation for solution in a given solvent is Equation (12), with which $\log \chi$ values for n-alkanes and cyclohexane were correlated.

The data sets studied (sets S11a through S11m are given in Table 15. Results of the correlations are reported in Table 16.

The results for the n-alkanes show that when $N_{dp} = 4$ A lies between 20.3 and 21.6, $100r^2$ between 98.83 and 99.68; while when $N_{dp} > 4$ A lies between 15.0 and 18.5, $100r^2$ between 97.20 and 98.03. The value of the intercept, a_0, seems to fall off with increasing n_C, the number of carbon atoms in the alkane. Cyclohexane has an intercept which is very different from those of the n-alkanes. In order to combine all of the n-alkanes into the same data set (set 11$'$) it is necessary to account for the polarizability of the n-alkanes

Table 14. Results for tetrahedral pentatomic molecules, MZ_4. ΔH_m, ΔH_v

Q	ΔH_m		ΔH_v	
Set	DHM5	DHM5$'$	DHV5	DHV5$'$
	MZ_4	$MZ_4 + MZ_nZ'_{4-n}$	MZ_4	$MZ_4 + MZ_nZ'_{4-n}$
A	32.2	31.7	80.8	83.6
S_A	2.84	2.46	3.88	5.11
M		−2.36		−4.02
S_M		0.499		1.00
P		–		
S_P		–		
a_o	−2.25	−1.90	5.27	5.32
S_{ao}	0.933	0.830	1.26	1.67
$100r^2$	93.45	90.71	97.31	90.13
$A100R^2$		90.19	–	89.83
F	128.5	82.99	434.3	146.0
S_{est}	1.60	3.92	2.41	3.57
$S°$	0.283	0.331	0.177	0.329
N_{dp}	11	20	14	35
$r_{\alpha\mu}$	–	0.359	–	0.506
$r_{\alpha\omega}$	–		–	
$r_{\mu\omega}$	–		–	
C_α	100	57.2	100	67.5
C_μ	–	42.8	–	32.5
C_ω	–	–	–	–
$r_{df/iv}$	9.00	8.50	12.0	16.0

Table 15. Solubility of rare gases M in alkanes, $10^4\chi_M$

Set	S11a	S11b	S11c	S11d	S11e	S11f	
Solvent	PeH	HxH	HpH	OcH	NnH	DcH	
M							
He	2.6	2.60	2.49	2.363	2.41	2.387	
Ne	4.1	3.71	3.55	3.62	3.43	3.48	
Ar	28.4	25.2	25.0	24.45	24.8	24.8	
Kr	78.5	69.5	71.2	70.7	71.2	71.9	
Xe	273.1	259	260.2	na	na	na	
Rn	na	737	na	na	na	na	
Set	S11g	S11h	S11i	S11j	S11k	S11l	S11m
Solvent	UdcH	DdcH	TrdH	TedH	PedH	HxdH	cHxH
M							
He	1.9	2.22	1.9	2.28	1.8	1.8	1.22
Ne	3.7	3.13	3.6	3.24	3.5	3.2	1.81
Ar	25.6	25.7	24.7	25.9	24.7	24.7	14.9
Kr	71.0	75.7	69.6	78.1	70.4	69.6	47.0
Xe	na	307	na	na	na	na	209
Rn	na	na	na	na	na	na	na

na, not avalable; Pe, pentyl; Hx, hexyl; Hp, heptyl; Oc, octyl; Nn, nonyl; Dc, decyl; Udc, undecyl, Ddc, dodecyl; Trd, tridecyl; Ted, tetradecyl; Ped, pentadecyl; Hxd, hexadecyl; cHx, cyclohexyl.

Table 16. Results of correlations of log χ_g with Equation (3). Rare gases, M, in n-alkanes and cyclohexane

Set	S11a	S11b	S11c	S11d	S11e	S11f	
Solvent	PeH	HxH	HpH	OcH	NnH	DcH	
A	16.6	15.0	16.7	20.3	20.4	20.5	
S_A	1.63	1.13	1.50	0.839	0.830	0.815	
B^o	0.817	0.808	0.773	0.749	0.745	0.745	
S_{Bo}	0.0910	0.0887	0.0836	0.0277	0.0274	0.0269	
$100r^2$	97.20	97.79	97.66	99.66	99.67	99.68	
F	104.0	176.8	125.3	583.4	604.3	630.6	
S_{est}	0.165	0.163	0.151	0.0498	0.0493	0.0484	
S^o	0.216	0.182	0.197	0.0827	0.0812	0.0795	
N_{dp}	5	6	5	4	4	4	
Set	S11g	S11h	S11i	S11j	S11k	S11l	S11m
Solvent	UdcH	DdcH	TrdH	TedH	PedH	HxdH	cHx
A	21.1	17.8		21.3	21.4	21.6	18.5
S_A	1.63	1.58		0.863	1.58	1.44	1.51
B^o	0.720	0.740		0.734	0.699	0.685	0.496
S_{Bo}	0.0538	0.0880		0.0285	0.0523	0.0471	0.0844
$100r^2$	98.83	97.70		99.67	98.92	99.13	98.03
F	168.9	127.5		611.6	182.7	227.1	149.4
S_{est}	0.0966	0.159		0.0513	0.0941	0.0852	0.153
S^o	0.153	0.196		0.0807	0.147	0.132	0.181
N_{dp}	4	5		4	4	4	5

which in the combined set must vary. This results in the correlation equation:

$$\log \chi_G = A \alpha_M + B_C n_{C,Ak} + a_0 \tag{17}$$

where α_M is the polarizability of the rare gas M and n_C, the number of C atoms in AkH, is a polarizability parameter[2] for the solvent. The regression equation obtained for the combined set (set GS1′) is:

$$\log \chi_M = 18.5 \, (\pm 0.435) \, \alpha_M$$
$$+ \, 0.748 \, (\pm 0.0188) \tag{18}$$

$100r^2$, 97.30; F, 1805; S_{est}, 0.118; S^o, 0.167;

N_{dp}, 52.

No dependence on n_C is observed. Again, the fit of the data to the equation is excellent. Solubilities in cyclohexane cannot be included in this combined set because of the very large difference in the intercept between it and the n-alkanes.

For the rare gases in PhX (sets 12a through 12g) the correlation equation is equation 12. The results are reported in Table 17. The fit is excellent. Both A and a^o vary with X in PhX. We may combine the

individual data sets into a single data set (set S12′) by parameterizing the polarizability of X and the bond moment of the X-Ph [7, 8, 16] It is not necessary to parameterize the polarizability of the Ph group as it is constant throughout set 2. Thus, the correlation equation becomes either:

$$\log \chi_G = A_M \alpha_M + M_{PhX} \mu_{PhX} + A_X \alpha_X$$
$$+ \, a_0 \tag{19}$$

or:

$$\log \chi_G = A_M \alpha_M + L \sigma_{lX} + D \sigma_{dX} + R \sigma_{eX}$$
$$+ \, A_X \alpha_X + a_0 \tag{20}$$

depending on the choice of parameterization of the dipole moment of PhX. Correlation with equation 19 gave on the exclusion of the value for Rn in PhH the regression equation:

$$\log \chi_G = 18.3 \, (\pm 0.570) \, \alpha_M$$
$$- \, 0.0524 \, (\pm 0.0240) \, \mu_{PhX}$$
$$+ \, 1.98 \, (\pm 0.679) \, \alpha_X$$
$$+ \, 1.40 \, (\pm 0.0542) \tag{21}$$

Table 17. Solubility of rare gases M in PhX, $10^5 \chi_M$

Set	S12a	S12b	S12c	S12d	S12e	S12f	S12g
Solvent	PhH	PhMe	PhF	PhCl	PhBr	PhI	$PhNO_2$
M							
He	7.59	9.90	1.15	6.93	5.37	3.74	3.50
Ne	11.51	14.35	1.58	9.8	7.86	5.26	4.35
Ar	88.2	109.6	11.5	86.1	67.9	50.0	44.5
Kr	274	335	33.3	273	228	169	139
Xe	1160	1460	131	1230	1080	882	na
Rn	4020	4950	na	na	na	na	na

Table 18. Results of correlations of $\log \chi_g$ with Equation (3). Rare gases, M, in substituted benzenes

Solvent	PhH	PhMe	PhF	PhCl	PhBr	PhI	$PhNO_2$
Set	S12a	S12b	S12c	S12d	S12e	S12f	S12g
A	16.6	16.5	17.1	18.7	19.1	19.7	22.7
S_A	1.14	1.10	1.42	1.55	1.49	1.47	1.20
B^o	1.308	1.409	1.431	1.246	1.144	0.987	0.914
S_{Bo}	0.0893	0.0860	0.0793	0.0864	0.0834	0.0823	0.0397
$100r^2$	98.15	98.27	97.97	97.99	98.20	98.35	99.44
F	212.7	227.1	144.8	146.1	163.5	178.8	357.5
S_{est}	0.165	0.159	0.144	0.156	0.151	0.149	0.0713
S°	0.166	0.161	0.184	0.183	0.173	0.166	0.105
N_{dp}	6	6	5	5	5	5	4

$100R^2$, 97.29; $A100R^2$, 97.12; F, 370.4; S_{est}, 0.158; S^o, 0.175; N_{dp}, 35; $C_{\alpha M}$, 92.7; $C_{\mu X}$, 2.25; $C_{\alpha X}$, 5.02; $r_{\alpha M,\mu X}$, 0.166; $r_{\alpha M,\alpha X}$, 0.021; $r_{\mu X,\alpha X}$, 0.349.

while correlation with equation 20 gave the regression equation:

$$
\begin{aligned}
\log \chi_G = {} & 17.8\,(\pm 0.531)\,\alpha_M \\
& - 0.205\,(\pm 0.117)\,\sigma_{lX} \\
& - 0.481\,(\pm 0.145)\,\sigma_{dX} \\
& - 1.71\,(\pm 0.656)\,\alpha_X \\
& + 1.40\,(\pm 0.0542) \quad\quad (22)
\end{aligned}
$$

$100R^2$, 97.60; $A100R^2$, 97.38; F, 315.8; S_{est}, 0.156; S^o, 0.167; N_{dp}, 36; $r_{\alpha M,\sigma l}$, 0.224; $r_{\alpha M,\sigma d}$, 0.007; $r_{\alpha M,\sigma e}$, 0.085; $r_{\alpha M,\alpha X}$, 0.093; $r_{\sigma l,\sigma d}$, 0.314; $r_{\sigma l,\sigma e}$, 0.068; $r_{\sigma l,\alpha X}$, 0.715; $r_{\sigma d,\sigma e}$, 0.240; $r_{\sigma d,\alpha X}$, 0.124; $r_{\sigma e,\alpha X}$, 0.679; $C_{\alpha M}$, 77.6; $C_{\sigma l}$, 7.45; $C_{\sigma d}$, 4.47; $C_{\alpha X}$, 10.5.

Clearly, in both equations the dominant variable is α_M.

For alcohols, AkOH (sets S13b through S13f), and water set (3a) correlation with equation 12 gives very good results, they are set forth in Table 19. All of the available data for alcohols can be combined into a single set (set S13′) by including a term in n_C. The correlation equation used is equation 13, the regression equation is:

$$
\begin{aligned}
\log \chi_G = {} & 15.7\,(\pm 1.26)\,\alpha_M \\
& + 0.0419\,(\pm 0.0213)\,n_{C,Ak} \\
& + 2.08\,(\pm 0.102) \quad\quad (23)
\end{aligned}
$$

$100R^2$, 91.76; $A100R^2$, 91.21; F, 77.97; S_{est}, 0.207; S°, 0.316; N_{dp}, 17; $r_{\alpha M,nC}$, 0.171; $C_{\alpha M}$, 94.9; C_{nC}, 5.07.

It is possible to include water in the alcohol data set (set S13″) by means of an indicator variable n_W which takes the value 1 when the gases are dissolved in water and is 0 otherwise, the correlation equation is:

$$
\log \chi_G = A\,\alpha_M + B_C\,n_{C,Ak} + B_W\,n_W + a_0 \quad (24)
$$

The regression equation is:

$$
\begin{aligned}
\log \chi_G = {} & 12.3\,(\pm 1.19)\,\alpha_M \\
& - 1.43\,(\pm 0.133)\,n_W \\
& + 2.28\,(\pm 0.0655) \quad\quad (25)
\end{aligned}
$$

Table 19. Solubility of rare gases M in water and ROH, $10^5 \chi_M$

Set	S13a	S13b	S13c	S13d	S13e	S13f
Solvent	H$_2$O	MeOH	EtOH	iBuOH	OcOH	cHxOH
M						
He	0.6997	5.95	7.91	10.2	12.07	4.73
Ne	0.8152	8.07	10.90	15.3	16.93	7.18
Ar	2.519	44.7	62.57	95.2	126	53.0
Kr	4.512	na	na	255	na	na
Xe	7.890	na	na	na	na	na
Rn	16.71	na	1330	na	na	na

Table 20. Results of correlations of log χ_g with Equation (3). Rare gases, M, in various solvents

Set	S13a	S13c	S13d	S14a	S14b	S14c	S14d
Solvent	H$_2$O	EtOH	iBuOH	MeNO$_2$	MeAc	Me$_2$SO	CS$_2$
A	8.32	13.4	19.2	14.4	13.5	14.7	18.9
S$_A$	1.06	1.20	0.874	1.75	1.90	0.686	0.884
B°	1.059	2.218	2.366	0.937	1.433	0.735	1.023
S$_{Bo}$	0.0549	0.0901	0.0289	0.0975	0.136	0.0383	0.0693
100r^2	97.25	98.42	99.59	95.77	94.40	99.35	99.13
F	141.7	124.7	483.3	67.86	50.59	458.9	457.6
S$_{est}$	0.101	0.157	0.0519	0.176	0.250	0.0694	0.128
S°	0.203	0.178	0.0908	0.266	0.305	0.104	0.114
N$_{dp}$	6	4	4	5	5	5	6

100R^2, 89.17; A100R^2, 88.65; F, 82.32; S$_{est}$, 0.263; S°, 0.353; N$_{dp}$, 23; r$_{\alpha M,nW}$, 0.350; C$_{\alpha M}$, 63.2; C$_{nC}$, 36.8.

Correlation of solubilities of rare gases in MeNO$_2$, MeAc, Me$_2$SO and CS$_2$ (sets S14a through S14b) with equation 12 gave fair fit for the first two solvents and excellent fit for the last two. The results are reported in Table 20.

For diatomic molecules some of which have dipole moments, the correlation equation for solubilities in n-alkanes (sets S21a through S21f) is equation 13. Results are reported in Table 21. No dependence on the dipole moment was observed. The fit of the data to the equation is good. In order to combine all of the data in n-alkane solvents into a single data set (set S21′) we have added to equation 13 a term in n$_C$ of the solvent giving the correlation equation:

$$\log \chi_G = A\, \alpha_{M^1 M^2} + M\, \mu_{M^1 M^2} + B_C\, n_{C,Ak} + a_0 \qquad (26)$$

The regression equation is:

$$\log \chi_G = 26.2\, (\pm 1.26)\, \alpha_{M^1 M^2}$$

Table 21. Solubility of rare gases M in various solvents, $10^5 \chi_M$

Set	S14a	S14b	S14c	S14d
Solvent	MeNO$_2$	MeAc	Me$_2$SO	CS$_2$
M				
He	3.86	19.74	2.84	3.9
Ne	5.40	15.7	3.68	5.9
Ar	31.5	90.6	15.4	48.63
Kr	83.8	320	44.6	175.6
Xe	201	na	170	1042
Rn	na	1700	na	4870

$$+\, 0.0578\, (\pm 0.0291)\, \mu_{M^1 M^2}$$
$$-\, 0.482\, (\pm 0.0650) \qquad (27)$$

100R^2, 95.38; A100R^2, 95.17; F, 216.9; S$_{est}$, 0.171; S°, 0.230; N$_{dp}$, 24;

No dependence on n$_C$ is observed. The dependence on μ is borderline.

Table 22. Solubility of gases M^1M^2 in various solvents, $10^4 \chi_M$

Set Solvent M^1M^2	S21a HxH	S21b HpH	S21c OcH	S21d NnH	S21e DcH	S21f cHxH	S22a CHCl$_3$	S22b CCl$_4$
H$_2$	6.63	na	6.84	na	6.5	na	2.22	3.193
N$_2$	13.76	13.5	13.1	12.8	12.13	7.68	4.45	*6.407*
O$_2$	19.8	21.6	22.0	21.3	22.0	*12.48*	na	*12.09*
Cl$_2$	na	na	*1510*	na	*1720*	na	1140	*1620*
HCl	112	na	na	na	298	na	185	136
HBr	416	na	448	na	519	na	454	382
NO	*23.2*	na	na	na	na	na	na	*13.6*
CO	17.7	na	17.14	na	16.62	na	6.79	8.70

Values in italics were estimated from values at other temperatures by interpolation or extrapolation.

Table 23. Results of correlations of log χ_g for diatomic molecules with equation 3

Set Solvent	S21a HxH	S21c OcH	S21e DcH	S22a CHCl$_3$	S22b CCl$_4$
A	23.6	25.2	26.8	30.8	25.8
S$_A$	2.45	2.36	2.81	3.91	3.77
B^o	−0.368	−0.409	−0.469	−0.913	−0.512
S$_{Bo}$	0.107	0.133	0.160	0.235	0.211
100r^2	94.90	96.62	94.79	93.96	88.65
F	93.07	114.3	90.93	62.25	46.88
S$_{est}$	0.152	0.196	0.237	0.319	0.343
S^o	0.267	0.225	0.270	0.301	0.389
N$_{dp}$	7	6	7	6	8
r$_{\alpha\mu}$	0.843	0.431	0.403	0.360	0.369

Table 24. Values of $r_{a/b}$ used in the correlation

Set VdW1. M, $r_{a/b}$: He, 1.469; Ne, 12.17; Ar, 42.33; Kr, 58.71; Xe, 81.30; Rn, 197.4.

Set VdW2. M^1M^2, $r_{a/b}$: H$_2$, 9.370; N$_2$, 35.52; O$_2$, 43.38; F$_2$, 34.53; Cl$_2$, 118.8; Br$_2$, 165.0; CO, 37.28; NO, 50.52; Hcl, 91.11; Hbr, 101.9; HI, 119.0.

Set VdW3. MZ$_3$, $r_{a/b}$: NF$_3$, 66.74; NMe$_3$, 121.4; PH$_3$, 86.70; PF$_3$, 76.10; AsH$_3$, 104.6; AsCl$_3$, 165.8; SbBr$_3$, 253.8.

Set VdW5a. MZ$_4$, $r_{a/b}$: CH$_4$, 121.8; CF$_4$. 63.87; CCl$_4$, 156.2; CMe$_4$, 121.8; SiH$_4$, 74.19; SiF$_4$, 72.68; SiCl$_4$, 142.6; GeH$_4$, 87.61; GeCl$_4$; 155.5; SnCl$_4$, 166.1.

Set VdW5b. MZ$_n^1$Z$_{4-n}^2$, $r_{a/b}$: CHCl$_3$, 150.5; CH$_2$Cl$_2$, 143.2; CH$_3$Cl, 116.8; CH$_3$F, 89.18; CF$_3$Cl, 84.75; CClMe$_3$, 142.3; CF$_2$Br$_2$, 132.3; CHFCl$_2$, 126.7; SiH$_2$Cl$_2$, 126.0; SiClMe$_3$, 244.4.

The correlation of solubilities of diatomic molecules in CHCl$_3$ and CCl$_4$ (sets S22a and S22b) with equation 13 gave fair and good results respectively. They are given in Table 22.

Ratios of the van der Waals a and b constants

The van der Waals a constant has the dimensions of the product of energy and volume. The b constant has the dimensions of volume. Then a/b has the dimensions of energy. This quantity is designated $r_{a/b}$. The data sets studied are reported in Table 25.

For the rare gases (set VdW1) the correlation equation is Equation (12). On exclusion of the value for Rn the regression equation is:

$$r_{a/b,M} = 642\,(\pm 59.8)\,\alpha_M + 18.2\,(\pm 3.34) \qquad (28)$$

$100r^2$, 97.46; F, 115.1; S$_{est}$, 6.05; S^o, 0.206; N$_{dp}$, 5; r$_{df/iv}$, 3.00

For diatomic molecules (set VdW2) the correlation equation is Equation (13). The regression equation is:

$$r_{a/b,M^1M^2} = 864\,(\pm 89.4)\,\alpha_M + 19.6\,(\pm 11.1)\mu$$
$$+ 19.1\,(\pm 6.42) \qquad (29)$$

$100R^2$, 90.26; A$100R^2$, 89.45; F, 50.98; S$_{est}$, 4.10; S^o, 0.352; N$_{dp}$, 14; r$_{df/iv}$, 5.50; r$_{\alpha\mu}$, 0.015; C$_\alpha$, 71.7, C$_\mu$, 28.3.

As all trigonal pyramidal tetratomic molecules (set VdW4) have dipole moments the correlation equation is Equation (13). The regression equation is:

$$r_{a/b,MZ_3} = 463\,(\pm 133)\,\alpha_M - 7.37\,(\pm 14.0)\mu$$
$$+ 31.2\,(\pm 12.9) \qquad (30)$$

$100R^2$, 96.26; A$100R^2$, 95.51; F, 51.44; S$_{est}$, 15.6; S^o, 0.256; N$_{dp}$, 7; r$_{df/iv}$, 2.00; r$_{\alpha\mu}$, 0.920; C$_\alpha$, 86.3; C$_\mu$, 13.7.

The large value of $r_{\alpha\mu}$ shows that polarizability and dipole moment are inseparable in this data set.

For MZ_4 molecules the correlation equation is Equation (12). On exclusion of the value for CCl_4 the regression equation is:

$$r_{a/b,MZ_4} = 342\,(\pm 135)\,\alpha_M + 41.8\,(\pm 2.93) \quad (31)$$

$100r^2$, 98.92; F, 641.6; S_{est}, 4.74; S^o, 0.117; N_{dp}, 9; $r_{df/iv}$, 3.50.

For the combined data set of MZ_4 and $MZ^1Z^2Z^3Z^4$ (set VdW5) the correlation equation is Equation (13). The regression equation obtained on exclusion of the points for $SiClMe_3$ and CCl_4 is:

$$
\begin{aligned}
r_{a/b,MZ^1Z^2Z^3Z^4} =\ & 333\,(\pm 27.5)\,\alpha_M \\
& + 19.1\,(\pm 3.32)\mu \\
& + 45.7(\pm 5.98) \quad (32)
\end{aligned}
$$

$100R^2$, 91.47; $A100R^2$, 90.94; F, 80.42; S_{est}, 10.8; S^o, 0.320; N_{dp}, 18; $r_{df/iv}$, 7.50; $r_{\alpha\mu}$, 0.165; C_α, 63.5, C_μ, 36.5.

Discussion

We have shown that isotopological data sets of melting points, boiling points, critical temperatures, enthalpies of melting and of vaorization, solubilities of gases in liquids, and ratios of the Van der Waals a and b constants can all be accounted for by polarizability and dipole moment. The predominant independent variable is polarizability. a much smaller contribution from dipole moment is observed in the case of those data sets which include some polar molecules. Kier and Hall[5] have reported correlations of boiling points, solubilities, ΔH_v values, and the separate correlation of the a and b constants with topological parameters. As they are topologically constant they cannot be modeled by TPs. **There is no direct cause and effect relationship between these properties and topological parameters. It is polarizability and dipole moment that are fundamental paramenters.**

What, then do topological parameters represent? If there are data sets with variable topology for which topological parameters TP can model the quantity as well as the appropriate fundamental parameters FP then:

$$Q = f\left(\sum_{i=1}^{n} TP_i\right) \quad (33a)$$

$$Q = f\left(\sum_{i=1}^{n} FP_i\right) \quad (33b)$$

Equating the right sides of equations 33a and 33b gives:

$$f\left(\sum_{i=1}^{n} TP_i\right) = f\left(\sum_{i=1}^{n} FP_i\right) \quad (34)$$

Thus, 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 [1]. They are composite parameters. Topological parameters actually are counts of numbers of atoms, of bonds, of branching, and of electrons. The proposal that they are a measure of electron accessibility is incorrect [17]. In part II of this work we will discuss the composition of topological parameters and the way in which they work.

References

1. Kier, L.B. and Hall, L.H., Molecular Connectivity in Chemistry and Drug Research, Academic Press, New York, 1976.
2. Merrifield, R.E. and Simmons, H.E., Topological Methods in Chemistry, Wiley-Interscience, Wiley, New York, 1989.
3. Balaban, A.T. Ed., From Chemical Topology to Three-Dimensional Geometry, Plenum Press, New York, 1997.
4. Charton, M., Abstr. 9th Intl. Conf. Correlation Anal. Chem., Borowno, Poland, 2001, p. 0–10.
5. Charton, M., Abstr. 212th Mtg. Amer. Chem. Soc., Boston, 2002.
6. Charton, M., Abstr. 14th Eur. Symp. QSAR, Bournemouth, 78 (2002).
7. Charton, M., in The Chemistry of Germanium, Tin and Lead Compounds, Vol. 2, Z. Rapaport, Ed., Wiley, Chichester. (2002) 537.
8. 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.
9. Charton, M. and Charton, B.I., J. Phys. Org. Chem., 7 (1994) 196.
10. Charton, M. and Charton, B.I., J. Theoret. Biol., 99 (1982) 629.
11. Charton, M., Prog. Phys. Org. Chem., 18 (1990) 163.
12. Charton, M., in Trends in Medicinal Chemistry '88, van der Goot, H., Domany, G., Pallos, L. and Timmerman, H. (eds.), Elsevier, Amsterdam (1989) 89.
13. Charton, M. and Charton, B.I., J. Org. Chem., 44 (1979) 2284.
14. Miller, T.M., in Handbook of Chemistry and Physics, 67th Edn., Weast, R.C., Astle, M.J. and Beyer, W.H. (eds.), CRC Press, Boca Raton (1986) E66.
15. Charton, M., Adv. Quantitative Structure Property Relationships, 3 (2002) 137.
16. Charton, M., in The Chemistry of Dienes and Polyenes, Z. Rappoport, Ed., Wiley, Chichester (1997) 683.
17. Kier, L.B. and Hall, L.H., J. Chem. Inf. Comp. Sci., 40 (2000) 792.