Journal of Chromatography, 235 (1982) 61–74
Elsevier Scientific Publishing Company, Amsterdam — Printed in The Netherlands

CHROM. 14,270

INVESTIGATION OF N-ALKYLBENZAMIDES BY REVERSED-PHASE LIQUID CHROMATOGRAPHY

III. CORRELATION OF CHROMATOGRAPHIC PARAMETERS WITH MOLECULAR CONNECTIVITY INDICES FOR THE C_1–C_5 N-ALKYLBENZAMIDES

MARTHA J. M. WELLS* and C. RANDALL CLARK*
Division of Medicinal Chemistry, Department of Pharmacal Sciences, School of Pharmacy, Auburn University, Auburn University, AL 36849 (U.S.A.)
and
RICHARD M. PATTERSON
Research Data Analysis, Auburn University, Auburn University, AL 36849 (U.S.A.)
(Received May 7th, 1981)

SUMMARY

Molecular connectivity indices were compared with measured and calculated reversed-phase liquid chromatographic retention data for the C_1–C_5 N-alkylbenzamides. Molecular connectivity is a topological index which encodes fundamental structural information about a molecule. The calculated indices are presented as bar-graph spectra and in tabular form. The various parameters of the solvophobic theory used to calculate log k'_w values for the amides were also compared to connectivity data. Highest correlations with k' were obtained for connectivity data which describes molecular bulk, branching, and site of branching in the hydrocarbon portion of the molecule.

INTRODUCTION

High-performance liquid chromatography (HPLC) is currently the most rapidly growing technique of the separation sciences[1]. Of all of the HPLC techniques, those involving reversed-phase application dominate the literature. Understanding of the retention process in reversed-phase liquid chromatography (RPLC) is growing, but much of the work published is still qualitative in nature and there is a need for increased quantitation of the obtained retention data.

One method of accounting for hydrophobic selectivity in RPLC is to correlate

* Present Address: USDA Forest Service, Southern Forest Experiment Station, George W. Andrews Forest Sciences Laboratory, Auburn University, AL 36849, U.S.A.

0021-9673/82/0000–0000/$02.75

retention values with a topological index which encodes structural information about the solute. Topological features of molecules can be quantitated by an easily computable set of molecular connectivity indices. These indices were first introduced by Randic[2], and further developed by Kier and Hall[3]. The connectivity index, χ, is a very fundamental parameter which reflects the shape and interatomic connections of a molecule. Many studies[4–8] have demonstrated that physicochemical and biological properties which depend upon the topology of a molecule may be related to the connectivity index.

Molecular connectivity calculations have been compared with gas–liquid chromatographic retention data[9–11]. The simple connectivity index as originally defined has been used by Karger *et al.*[12] and more recently by Colin and Guiochon[13] to evaluate hydrophobic effects in RPLC. The index has since been refined and expanded by Kier and Hall[3] to allow for differences in the identities of atoms and in bond orders[14]. In this study, relationships between predicted and measured chromatographic data and the expanded molecular connectivity indices will be pursued.

EXPERIMENTAL

The chromatographic data used in this study has been previously reported[15] as has the application of the solvophobic theory to the chromatographic data[16]. Additionally, some data obtained using a 25 cm × 4.60 mm I.D. Partisil ODS (10 μm) column purchased from Whatman (Clifton, NJ, U.S.A.) is also included in this study. Chromatographic procedures employed with this column were identical to those described previously[15] with the exception that the Partisil ODS column was operated at ambient temperature without a guard column.

RESULTS AND DISCUSSION

The reversed-phase liquid chromatographic retention of the C_1–C_5 N-alkylbenzamides has been investigated and the physical characteristics of these compounds described[15]. In Part II of this series[16] chromatographic parameters for the C_1–C_5 N-alkylbenzamides were predicted by application of the solvophobic theory[17–19] to measured retention data. The purpose of this article is to associate the predicted parameters as well as the actually measured chromatographic capacity factors with structural features of the molecules.

Understanding of the structural meaning of connectivity indices is facilitated by bar graph spectra[3] of the ${}^m\chi_t$ values for the C_1–C_5 N-alkylbenzamides involved in this study (Fig. 1). The horizontal axes in Fig. 1 represent the order, m, and the vertical axes the magnitude of ${}^m\chi_t$ values; t = type. These graphs can be visually compared to find trends in structure. A smooth falloff in the spectrum is observed for those compounds having long-chain portions. (Compare the methyl, ethyl, *n*-propyl, *n*-butyl, and *n*-pentyl derivatives.) The spectrum becomes more jagged (due to greater ${}^2\chi$ values) with an increase in the amount of branching in the N-alkyl chain. The graphs of the *tert.*-butyl, *tert.*-pentyl, and neopentyl compounds have high ${}^2\chi$ and ${}^3\chi_c$ values and non-zero values of ${}^4\chi_c$ due to the presence of carbon atoms assigned a δ value of four (*i.e.*, a quaternary carbon).

The number of terms for each subgraph-type of molecular connectivity index

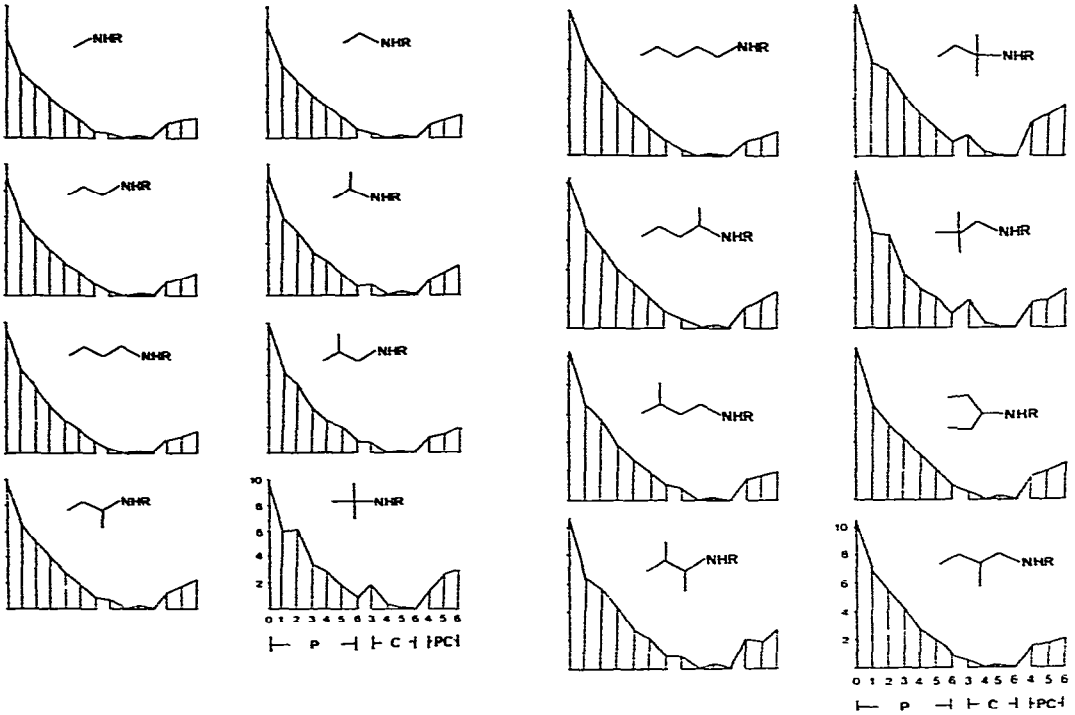


Fig. 1. Bar-graph spectra of the $^{m}\chi_t$ values for the C_1–C_5 N-alkylbenzamides: P = path; C = cluster; PC = path/cluster; R = benzoyl.

(path, P, cluster, C, and path/cluster, PC) is presented for the C_1–C_5 N-alkylbenzamides in Table I. The chain terms, CH (arising from an enclosed figure), were omitted since there are no 3rd, 4th, or 5th order chain terms possible for these compounds. Also, every compound has one 6th order chain term due to the presence of the aromatic ring.

By the use of the Statistical Analysis System procedure RSQUARE (SAS Institute, Raleigh, NC, U.S.A.) various parameters were screened against all possible two-variable combinations of the connectivity level (χ) and valence level(χ^{v}) indices, their inverses and their squares through sixth order. When the best combinations were selected, procedure SYSREG was used to evaluate the coefficients.

Comparison of log k' values with molecular connectivity indices

The method of predicting the log k'_w values for 0% acetonitrile on both the Partisil ODS-2 and Ultrasphere ODS columns was discussed previously[16]. The molecular connectivity indices were regressed on the predicted log k'_w values for the sixteen C_1–C_5 N-alkylbenzamides (Table IV, ref. 16), and the best two variable combinations chosen are given in eqns. 1 and 2:

Partisil ODS-2

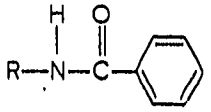
$$\log k'_w = 0.070\ (0.013)\ [^5\chi_{PC}]^2 + 20.930\ (0.917)\ [^6\chi_P^v]^2 + 2.154\ (0.082) \qquad r = 0.9885 \qquad (1)$$

Ultrasphere ODS

$$\log k'_w = 0.834\ (0.026)\ [^0\chi^v] + 0.326\ (0.031)\ [^6\chi_{PC}^v]^{-1} - 3.991\ (0.267) \qquad r = 0.995 \qquad (2)$$

TABLE I

NUMBER OF SUBGRAPH TERMS FOR PATH, CLUSTER, AND PATH/CLUSTER TYPES



R	$^{1}\chi$	$^{2}\chi$	$^{3}\chi$		$^{4}\chi$			$^{5}\chi$			$^{6}\chi$			*Total*		
	P	*P*	*P*	*C*	*P*	*C*	*PC*	*P*	*C*	*PC*	*P*	*C*	*PC*	*P*	*C*	*PC*
CH_3	10	12	14	2	14	0	7	14	1	12	8	0	19	72	3	38
CH_3CH_2	11	13	15	2	16	0	7	16	1	13	10	0	22	81	3	42
$CH_3CH_2CH_2$	12	14	16	2	17	0	7	18	1	13	12	0	23	89	3	43
$(CH_3)_2CH$	12	15	16	3	18	0	8	18	1	16	12	0	28	91	4	52
$CH_3(CH_2)_3$	13	15	17	2	18	0	7	19	1	13	14	0	23	96	3	43
$(CH_3)_2CHCH_2$	13	16	17	3	18	0	8	20	1	14	14	0	26	98	4	48
$CH_3CH_2CH(CH_3)$	13	16	18	3	19	0	9	20	1	17	14	0	31	100	4	57
$(CH_3)_3C$	13	18	17	6	20	1	10	20	1	22	14	0	39	102	8	71
$CH_3(CH_2)_4$	14	16	18	2	19	0	7	20	1	13	15	0	23	102	3	43
$CH_3CH_2C(CH_3)_2$	14	19	20	6	21	1	13	22	1	25	16	0	45	112	8	83
$CH_3CH_2CH_2CH(CH_3)$	14	17	19	3	21	0	9	21	1	18	16	0	32	108	4	59
$(CH_3)_3CCH_2$	14	19	18	6	19	1	10	22	1	17	16	0	32	108	8	59
$(CH_3)_2CHCH_2CH_2$	14	17	18	3	19	0	8	20	1	14	16	0	24	104	4	46
$(CH_3CH_2)_2CH$	14	17	20	3	21	0	10	22	1	19	16	0	35	110	4	64
$(CH_3)_2CHCH(CH_3)$	14	18	20	4	20	0	12	22	2	19	16	0	37	110	6	68
$CH_3CH_2CH(CH_3)CH_2$	14	17	19	3	19	0	9	21	1	15	16	0	27	106	4	51

The numbers in parentheses in all equations are the standard error of the regression coefficients. The values of the chosen connectivity indices are given in Table II. The results (eqns. 1 and 2) should be compared and contrasted with the best regression equations obtained for the sixteen compounds at other solvent compositions and on other columns (eqns. 3–9; MeOH = methanol, ACN = acetonitrile):

Partisil ODS

$$\log k'_{40\%\ \mathrm{MeOH}} = 0.024\ (0.005)\ [{}^5\chi_{\mathrm{PC}}]^2 + 6.879\ (0.349)\ [{}^6\chi^{\mathrm{v}}_{\mathrm{P}}]^2 + 0.200\ (0.031) \qquad r = 0.9847 \qquad (3)$$

Partisil ODS-2

$$\log k'_{35\%\ \mathrm{MeOH}} = 0.053\ (0.006)\ [{}^5\chi_{\mathrm{PC}}]^2 + 12.625\ (0.463)\ [{}^6\chi^{\mathrm{v}}_{\mathrm{P}}]^2 + 0.500\ (0.041) \qquad r = 0.9922 \qquad (4)$$

$$\log k'_{55\%\ \mathrm{MeOH}} = 0.041\ (0.004)\ [{}^5\chi_{\mathrm{PC}}]^2 + 8.439\ (0.303)\ [{}^6\chi^{\mathrm{v}}_{\mathrm{P}}]^2 + 0.055\ (0.027) \qquad r = 0.9927 \qquad (5)$$

$$\log k'_{20\%\ \mathrm{ACN}} = 0.068\ (0.006)\ [{}^5\chi_{\mathrm{PC}}]^2 + 13.500\ (0.416)\ [{}^6\chi^{\mathrm{v}}_{\mathrm{P}}]^2 + 0.490\ (0.037) \qquad r = 0.9946 \qquad (6)$$

$$\log k'_{40\%\ \mathrm{ACN}} = 0.054\ (0.004)\ [{}^5\chi_{\mathrm{PC}}]^2 + 8.697\ (0.304)\ [{}^6\chi^{\mathrm{v}}_{\mathrm{P}}]^2 + 0.001\ (0.027) \qquad r = 0.9935 \qquad (7)$$

Ultrasphere ODS

$$\log k'_{35\%\ \mathrm{MeOH}} = 0.056\ (0.007)\ [{}^5\chi_{\mathrm{PC}}]^2 + 13.334\ (0.519)\ [{}^6\chi^{\mathrm{v}}_{\mathrm{P}}]^2 + 0.162\ (0.046) \qquad r = 0.9912 \qquad (8)$$

$$\log k'_{20\%\ \mathrm{ACN}} = 0.077\ (0.006)\ [{}^5\chi_{\mathrm{PC}}]^2 + 14.520\ (0.450)\ [{}^6\chi^{\mathrm{v}}_{\mathrm{P}}]^2 + 0.123\ (0.040) \qquad r = 0.9946 \qquad (9)$$

The log k' values predicted by eqns. 1–9 are listed in Tables III and IV. When a connectivity level (χ) value is chosen in a regression of this type, it implies that the nature of the atom itself is not important. However, if a valence level (χ^{v}) value is selected, the identity of the atom (carbon *vs.* oxygen *vs.* nitrogen, etc.) is important to correlation with a given property. In eqns. 1 and 3–9, the same connectivity parameters were chosen as the best descriptors of retention in these systems. This implies that the same structural features were important to the chromatographic retention process at various eluent compositions on these columns. However, at 0% acetonitrile on the Ultrasphere ODS column different structural parameters were selected. The most apparent anomalies in the predicted log k' values (Tables III and IV) appear to be: (1) a consistently higher predicted value for the 1-ethylpropyl derivative, and (2) in all cases except one a reverse order of elution to that observed is predicted for the isopentyl and *n*-pentyl derivatives.

TABLE II

SELECTED MOLECULAR CONNECTIVITY INDICES FOR THE C_1 C_5 N-ALKYLBENZAMIDES

For general structure, see Table I.

R	$^0\chi$	$^0\chi^v$	$^1\chi$	$^2\chi^v$	$^5\chi_{pc}$	$^6\chi^v_p$	$^6\chi_{pc}$	$^6\chi^v_{pc}$
CH_3	7.39734	5.79500	4.84253	1.94515	1.33925	0.13065	1.51603	0.27826
CH_3CH_2	8.10444	6.50210	5.34253	2.22547	1.37377	0.17548	1.68269	0.32036
$CH_3CH_2CH_2$	8.81155	7.20921	5.84253	2.62192	1.32496	0.23039	1.70710	0.32504
$(CH_3)_2CH$	8.97469	7.37235	5.69837	2.99418	1.79251	0.20654	2.14567	0.46531
$CH_3(CH_2)_3$	9.51866	7.91632	6.34253	2.97547	1.32496	0.28953	1.67258	0.31756
$(CH_3)_2CHCH_2$	9.68180	8.07945	6.19837	3.46989	1.47000	0.26843	2.00320	0.42466
$CH_3CH_2CH(CH_3)$	9.68180	8.07945	6.23638	3.14878	1.81043	0.25137	2.33227	0.52879
$(CH_3)_3C$	9.89734	8.29500	5.98898	4.07015	2.63726	0.23159	3.10821	0.78616
$CH_3(CH_2)_4$	10.22576	8.62342	6.84253	3.32903	1.32496	0.32852	1.67258	0.31756
$(CH_3)_2CHCH_2CH_2$	10.38890	8.78656	6.69837	3.80444	1.52908	0.33098	1.77514	0.38642
$CH_3CH_2CH(CH_3)CH_2$	10.38890	8.78656	6.73638	3.58946	1.62530	0.31671	2.01586	0.46357
$CH_3CH_2CH_2CH(CH_3)$	10.38890	8.78656	6.73638	3.52920	1.96573	0.29966	2.34493	0.56771
$(CH_3CH_2)_2CH$	10.38890	8.78656	6.77438	3.35290	1.96673	0.29620	2.51089	0.61982
$(CH_3)_2CHCH(CH_3)$	10.55204	8.94970	6.60906	3.87217	1.98611	0.28243	2.83388	0.73751
$(CH_3)_3CCH_2$	10.60444	9.00210	6.48898	4.60936	1.97345	0.29911	2.60052	0.66814
$CH_3CH_2C(CH_3)_2$	10.60444	9.00210	6.54964	4.05758	2.89302	0.27042	3.46418	0.94794

A subset (Table V) of the pertinent connectivity values from Table II illustrates the relationships observed. Branching increases the value of $^5\chi_{PC}$ in the order: *n*-butyl < isobutyl < *sec.*-butyl < *tert.*-butyl. *sec.*-Butyl has a larger $^5\chi_{PC}$ value than isobutyl because the branch occurs on the carbon atom attached to the amide nitrogen atom. The $^6\chi_P^v$ values are decreased by branching in the same order to the increase of the $^5\chi_{PC}$ indices.

Comparison of solvophobic regression coefficients with molecular connectivity indices

The solvophobic theory[17–19] was used in Part II[16] of this series to determine values for the chromatographic parameters B, $(A + E)$, and C. The values obtained for $(A + E)$ and B (Tables II and III, ref. 16) were screened against the molecular connectivity indices for the sixteen C_1–C_5 N-alkylbenzamides and the results are given in eqns. 10 and 11 for the Partisil ODS-2 column and in eqns. 12 and 13 for the Ultrasphere ODS column:

Partisil ODS-2

$$B = 10.559\,(1.572)\,[^6\chi_{PC}]^2 + 2653.980\,(162.014)\,[^6\chi_P^v]^2 - 409.965\,(14.350) \qquad r = 0.9810 \qquad (10)$$

$$(A + E) = -10.206\,(1.525)\,[^6\chi_{PC}]^2 - 2584.280\,(157.215)\,[^6\chi_P^v]^2 + 383.320\,(13.925) \qquad r = 0.9811 \qquad (11)$$

Ultrasphere ODS

$$B = 101.329\,(21.228)\,[^1\chi]^2 - 1195.480\,(252.722)\,{}^1\chi + 3312.993\,(746.060) \qquad r = 0.8014 \qquad (12)$$

$$(A + E) = -99.020\,(20.660)\,[^1\chi]^2 + 1168.690\,(245.974)\,{}^1\chi - 3256.500\,(726.140) \qquad r = 0.8022 \qquad (13)$$

The best two-variable combination chosen for B and $(A + E)$ was the same for each column. This is not unreasonable since a high degree of correlation between the variables B and $(A + E)$ has already been established (see eqns. 10 and 11, ref. 16). The correlation between the connectivity indices and the values of B and $(A + E)$ derived from the Ultrasphere ODS column is unexpectedly poor. Eqns. 10 and 11 involve a path/cluster term, which accounts for branching in these molecules, as well as a path term which is greater for those compounds that are less branched. However, the first-order connectivity level index (quadratic relationships of eqns. 12 and 13) is not usually considered to account best for molecular branching. The values of the molecular connectivity indices which appear in eqns. 10–13 are reported in Table II.

Table VI presents the calculated values for $^6\chi_{PC}/^6\chi_P^v$ for the C_1–C_5 N-alkylbenzamides. This ratio describes the relative amount of branched to unbranched portions in these compounds. In the straight-chain homologs the ratio decreases as the number of methylene groups increases. Not unexpectedly, the *tert.*-butyl- and *tert.*-pentylbenzamides have the largest values for the ratio.

TABLE III

EXPERIMENTAL AND PREDICTED LOG k' VALUES OF SOME N-ALKYLBENZAMIDES ON PARTISIL ODS-2

For general structure, see Table I.

R	*0% ACN*		*20% ACN*		*40% ACN*		*35% MeOH*		*55% MeOH*	
	*log k' (calc.)**	*log k' (calc.)***	*log k' (obs.)*	*log k' (calc.)****	*log k' (obs.)*	*log k' (calc.)‡*	*log k' (obs.)*	*log k' (calc.)‡‡*	*log k' (obs.)*	*log k' (calc.)‡‡‡*
CH_3	2.612	2.637	0.804	0.842	0.228	0.246	0.805	0.810	0.268	0.272
CH_3CH_2	2.987	2.931	1.039	1.034	0.377	0.371	1.000	0.989	0.399	0.392
$(CH_3)_2CH$	3.305	3.272	1.303	1.284	0.552	0.546	1.201	1.209	0.537	0.547
$CH_3CH_2CH_2$	3.452	3.388	1.347	1.326	0.581	0.557	1.274	1.263	0.576	0.575
$CH_3CH_2CH(CH_3)$	3.638	3.706	1.570	1.566	0.742	0.728	1.434	1.471	0.709	0.723
$(CH_3)_3C$	3.768	3.763	1.718	1.687	0.874	0.843	1.569	1.546	0.813	0.793
$(CH_3)_2CHCH_2$	3.786	3.813	1.636	1.610	0.776	0.744	1.552	1.524	0.770	0.752
$CH_3(CH_2)_3$	3.946	4.031	1.717	1.741	0.818	0.825	1.638	1.651	0.826	0.834
$(CH_3CH_2)_2CH$	4.039	4.261	1.843	1.937	0.919	0.973	1.680	1.813	0.869	0.954
$(CH_3)_2CHCH(CH_3)$	4.098	4.100	1.880	1.835	0.944	0.908	1.743	1.716	0.906	0.890
$(CH_3)_3CCH_2$	4.248	4.299	1.956	1.963	0.984	0.989	1.855	1.836	0.986	0.970
$CH_3CH_2CH_2CH(CH_3)$	4.337	4.304	1.974	1.965	0.984	0.991	1.816	1.838	0.954	0.971
$CH_3CH_2C(CH_3)_2$	4.353	4.270	2.043	2.046	1.069	1.089	1.887	1.867	1.021	1.015
$CH_3CH_2CH(CH_3)CH_2$	4.490	4.438	2.028	2.024	1.014	1.016	1.924	1.906	1.014	1.010
$(CH_3)_2CHCH_2CH_2$	4.589	4.610	2.076	2.128	1.034	1.080	1.978	2.007	1.053	1.075
$CH_3(CH_2)_4$	4.722	4.536	2.146	2.066	1.079	1.034	2.036	1.956	1.095	1.038

* See ref. 16, Table III.

** Eqn. 1.

*** Eqn. 6.

‡ Eqn. 7.

‡‡ Eqn. 4.

‡‡‡ Eqn. 5.

TABLE IV

EXPERIMENTAL AND PREDICTED LOG k' VALUES OF SOME N-ALKYLBENZAMIDES

For general structure, see Table I.

R	*Partisil ODS*		*Ultrasphere ODS*					
	40% MeOH		*0% ACN*		*20% ACN*		*35% MeOH*	
	log k' (obs.)	*log k' (calc.)**	*log k' (calc.)***	*log k' (calc.)****	*log k' (obs.)*	*log k' (calc.)§*	*log k' (obs.)*	*log k' (calc.)§§*
CH_3	0.382	0.360	2.034	2.014	0.458	0.509	0.479	0.490
CH_3CH_2	0.455	0.457	2.427	2.449	0.713	0.715	0.677	0.678
$(CH_3)_2CH$	0.558	0.570	2.850	2.858	1.000	0.990	0.907	0.911
$CH_3CH_2CH_2$	0.608	0.607	2.972	3.024	1.050	1.029	0.981	0.968
$CH_3CH_2CH(CH_3)$	0.683	0.713	3.373	3.364	1.311	1.293	1.171	1.188
$(CH_3)_3C$	0.762	0.736	3.456	3.342	1.485	1.437	1.320	1.267
$(CH_3)_2CHCH_2$	0.759	0.748	3.522	3.515	1.385	1.336	1.289	1.244
$CH_3(CH_2)_3$	0.813	0.819	3.642	3.638	1.456	1.475	1.360	1.378
$(CH_3CH_2)_2CH$	0.805	0.896	3.692	3.863	1.600	1.695	1.405	1.548
$(CH_3)_2CHCH(CH_3)$	0.841	0.843	3.842	3.915	1.635	1.585	1.475	1.446
$(CH_3)_3CCH_2$	0.928	0.909	3.995	4.005	1.724	1.722	1.617	1.573
$CH_3CH_2CH_2CH(CH_3)$	0.889	0.910	3.937	3.911	1.725	1.724	1.553	1.576
$CH_3CH_2C(CH_3)_2$	0.924	0.904	3.920	3.860	1.809	1.829	1.610	1.606
$CH_3CH_2CH(CH_3)CH_2$	0.959	0.953	4.065	4.040	1.780	1.783	1.661	1.647
$(CH_3)_2CHCH_2CH_2$	1.003	1.010	4.177	4.181	1.834	1.894	1.720	1.754
$CH_3(CH_2)_4$	1.057	0.984	4.289	4.228	1.894	1.825	1.785	1.699

* Eqn. 3.

** See ref. 16, Table III.

*** Eqn. 2.

§ Eqn. 9.

§§ Eqn. 8.

TABLE V

CONNECTIVITY DATA COMPARED TO LOG k' FOR THE BUTYLBENZAMIDES

For general structure, see Table I.

R	*Partisil ODS-2* log k'	$^5\chi_{PC}$	$^6\chi^v_P$
$CH_3(CH_2)_3$	3.946	1.32495	0.28953
$(CH_3)_2CHCH_2$	3.786	1.46999	0.26843
$CH_3CH_2CH(CH_3)$	3.638	1.81043	0.25137
$(CH_3)_3C$	3.768	2.63726	0.23159

The third regression coefficient derived through application of the solvophobic theory was the parameter C. The value of C was determined for the C_1–C_5 N-alkylbenzamides (Tables II and III, ref. 16) and used to evaluate ΔA, the contact surface area of the associated solute-bonded ligand complex (Table V, ref. 16). Results of the best two-variable regression of connectivity indices on ΔA (in Å^2) are given in eqns. 14 and 15:

Partisil ODS-2

$$\Delta A = 0.337\ (0.039)\ [^0\chi]^2 + 36.429\ (6.130)\ [^6\chi_{PC}]^{-1} + 42.878\ (6.112) \qquad r = 0.9220 \qquad (14)$$

Ultrasphere ODS

$$\Delta A = -703.697\ (36.660)\ [^0\chi]^{-1} + 24.030\ (3.191)\ [^5\chi_{PC}]^{-1} + 147.197\ (3.187) \qquad r = 0.9835 \qquad (15)$$

TABLE VI

COMPARISON OF BRANCHED TO UNBRANCHED PORTIONS IN THE C_1–C_5 N-ALKYLBENZAMIDES

For general structure, see Table I.

R	$^6\chi_{PC}$ $^6\chi^v_P$
CH_3	11.604
CH_3CH_2	9.589
$CH_3CH_2CH_2$	7.410
$(CH_3)_2CH$	10.389
$CH_3(CH_2)_3$	5.777
$(CH_3)_2CHCH_2$	7.463
$CH_3CH_2CH(CH_3)$	9.278
$(CH_3)_3C$	13.421
$CH_3(CH_2)_4$	5.091
$(CH_3)_2CHCH_2CH_2$	5.363
$CH_3CH_2CH(CH_3)CH_2$	6.365
$CH_3CH_2CH_2CH(CH_3)$	7.825
$(CH_3CH_2)_2CH$	8.477
$(CH_3)_3CCH_2$	8.694
$(CH_3)_2CHCH(CH_3)$	10.034
$CH_3CH_2C(CH_3)_2$	12.810

The correlation of molecular connectivity with ΔA was much better for the Ultrasphere ODS column than for the Partisil ODS-2 column. Solute structure alone may not account for all of the variation in the contact surface area, since the solvent as well as the bonded ligand itself must play a part in determining the contact surface area between the solute and the ligand. Because the configurations of the two bonded phases may not be alike due to differences in bonding conditions, the fact is understandable that different connectivity indices were chosen to correlate with the contact surface area, ΔA, on these two columns.

The molecular connectivity indices were also regressed on the calculated total molecular surface areas (TSA) for the C_1–C_5 N-alkylbenzamides and the results are described by eqn. 16:

$$\begin{aligned} TSA = {} & 37.212\,(1.649)\,{}^{1}\chi - 58.877\,(12.108)\,[{}^{2}\chi']^{-1} + 15.980\,(13.620) \\ & r = 0.9971 \end{aligned} \tag{16}$$

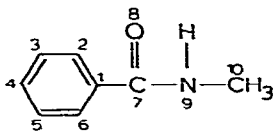
The ${}^{2}\chi'$ index does reflect some degree of branching because the value of the index increases with increases in the number of short branch points. The connectivity indices correlated with TSA are not the same as the indices chosen as the best variables for ΔA on either column.

CONCLUSIONS

A high degree of correlation was observed between molecular connectivity data and measured and calculated reversed-phase chromatographic retention data. The results suggest that bulk, branching, and site of hydrocarbon branching are the controlling factors in the retention of the C_1–C_5 N-alkylbenzamides and molecular connectivity allows for the quantitation of these molecular features. The same connectivity indices were chosen upon regression of retention data obtained on different columns. Inconsistent correlations were observed between the solute-bonded ligand contact surface area and molecular connectivity indices.

APPENDIX

Calculation of molecular connectivity indices



I

The calculations of the molecular connectivity indices through sixth order for N-methylbenzamide, I, are outlined. The atoms of the molecular structure are referred to as vertices, and the bonds are called edges. Each non-hydrogen atom (ten in

this example) is assigned a vertex value (δ) equal to the number of non-hydrogen atoms connected to it, and a vertex valence value (δ^v) calculated by subtracting the number of hydrogen atoms attached to that atom from the number of valence electrons in the atom. These values are given in Table A1 for N-methylbenzamide.

TABLE A1

VERTEX AND VERTEX VALENCE VALUES FOR N-METHYLBENZAMIDE

	Atom number									
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
δ	3	2	2	2	2	2	3	1	2	1
δ^v	4	3	3	3	3	3	4	6	4	1

The vertex values, δ, are used to calculate the connectivity level indices, χ, in which the nature of the atom is not taken into consideration. These indices were expanded by allowing for differences in the identity of atoms and differences in bond order by introducing a valence level index, χ^v, which is calculated by using the vertex valence values, δ^v.

The molecular structure is broken down into parts called subgraphs. If a graph is taken as being composed of vertices alone (*i.e.*, no edges), the order of this index is zero order. The number of edges, or bonds, in a subgraph is the order of that subgraph and appears as the leftside superscript to the symbol, χ, or χ^v. There are four types of subgraphs which may be present in a molecule: paths, clusters, path/clusters, and chains which appear as the subscripts P, C, PC, and CH respectively. If no subscript appears it is understood to be a path index.

The actual number of subgraphs of each order, m (through sixth order), and type, t, present in N-methylbenzamide are listed in Table A2. The dashes in this table indicate that for a given order it is not possible to have a subgraph of that type in any molecule. Some of the representative subgraphs present in N-methylbenzamide are illustrated in Fig. A1.

TABLE A2

NUMBER OF TERMS (t) FOR ORDER (m) OF $^m\chi_t$ INDICES FOR N-METHYLBENZAMIDE

Order	*Type*			
	Path	*Cluster*	*Path/cluster*	*Chain*
0	–	–	–	–
1	10	–	–	–
2	12	–	–	–
3	14	2	–	0
4	14	0	7	0
5	14	1	12	0
6	8	0	19	1

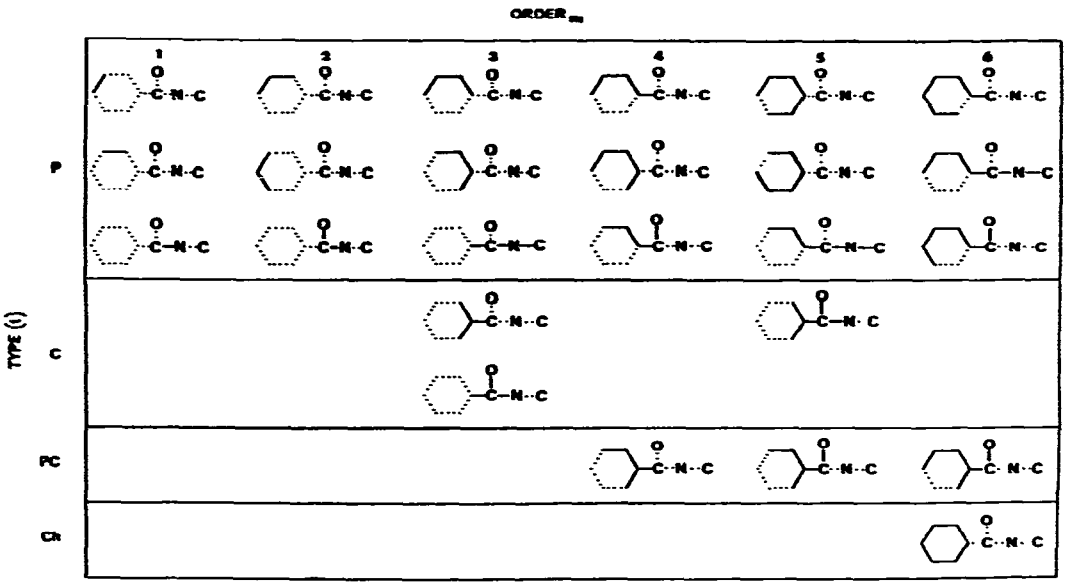


Fig. A1. Representative subgraphs (solid lines) of N-methylbenzamide.

Zero order indices are calculated by summing the reciprocal square roots of the vertex values, δ, or of the vertex valence values, δ^v. For subgraphs of order higher than zero, the reciprocal square root of the products of the δ or δ^v values assigned to the atoms participating in a bond or bonds is the value for that subgraph. All of the subgraph values are summed to give the χ or χ^v indices for the molecule. Molecular connectivity calculations were performed on the IBM 370/158, Computer Services, Auburn University. The program (CFUNC) used to calculate molecular connectivity indices was obtained from Dr. Lowell Hall (Eastern Nazarene College, Quincy, MA, U.S.A.). More detailed discussion of these calculations can be found in the literature[3,14].

ACKNOWLEDGEMENT

We sincerely appreciate the support of the American Foundation for Pharmaceutical Education Silas M. Burroughs Memorial Fellowship for M.J.M.W.

REFERENCES

1 W. Worthy, *Chem. Eng. News*, 58 (33) (1980) 30.
2 M. Randic, *J. Amer. Chem. Soc.*, 97 (1975) 6609.
3 L. B. Kier and L. H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York, 1976.
4 D. R. Henry and J. R. Block, *J. Med. Chem.*, 22 (1979) 465.
5 T. DiPaolo, L. B. Kier and L. H. Hall, *J. Pharm. Sci.*, 68 (1979) 39.
6 B. K. Evans, K. C. James and D. K. Luscombe, *J. Pharm. Sci.*, 68 (1979) 370.

7 R. A. Glennon, L. B. Kier and A. T. Shulgin, *J. Pharm. Sci.*, 68 (1979) 906.
8 L. B. Kier, *J. Pharm. Sci.*, 69 (1980) 1034.
9 Y. Michotte and D. L. Massart, *J. Pharm. Sci.*, 66 (1977) 1630.
10 J. S. Millership and A. D. Woolfson, *J. Pharm. Pharmacol.*, 30 (1978) 483.
11 L. B. Kier and L. H. Hall, *J. Pharm. Sci.*, 68 (1979) 120.
12 B. L. Karger, J. R. Gant, A. Hartkopf and P. H. Weiner, *J. Chromatogr.*, 128 (1976) 65.
13 H. Colin and G. Guiochon, *J. Chromatogr. Sci.*, 18 (1980) 54.
14 L. B. Kier and L. H. Hall, *J. Med. Chem.*, 20 (1977) 1631.
15 M. J. M. Wells and C. R. Clark, *J. Chromatogr.*, 235 (1982) 31.
16 M. J. M. Wells, C. R. Clark and R. M. Patterson, *J. Chromatogr.*, 235 (1982) 43.
17 O. Sinanoglu and S. Abdulnur, *Fed. Proc. Fed. Amer. Soc. Exp. Biol.*, 24, No. 2, Part III (1965) S-12.
18 O. Sinanoglu, in B. Pullman (Editor), *Molecular Associations in Biology*, Academic Press, New York, 1968.
19 Cs. Horváth, W. Melander and I. Molnár, *J. Chromatogr.*, 125 (1976) 129.