ELSEVIER

Journal of Chromatography A, 697 (1995) 59–69

JOURNAL OF CHROMATOGRAPHY A

Effect of various organic modifiers on the determination of the hydrophobicity parameters of non-homologous series of anticancer drugs

Esther Forgács*, Tibor Cserháti

Central Research Institute for Chemistry, Hungarian Academy of Sciences, P.O. Box 17, 1525 Budapest, Hungary

Abstract

The lipophilicities and specific hydrophobic surface areas of 21 commercial anticancer drugs were determined by means of reversed-phase thin-layer chromatography using methanol, ethanol, 1-propanol, 2-propanol, acetonitrile, dioxane and tetrahydrofuran as organic modifiers at various concentrations. The data were evaluated by various multivariate mathematical–statistical methods such as the spectral mapping technique and principal component analysis followed by two-dimensional non-linear mapping, varimax rotation and cluster analysis. The results indicate that the solvent strength and selectivity of organic modifiers are strongly related to the steric characteristics of the solvent molecule, suggesting competition between the anticancer drugs and solvent molecules for the hydrophobic surface of the stationary phase. It was established that for the evaluation of large retention data matrices the use of principal component analysis followed by two-dimensional non-linear mapping provides more information than one-dimensional cluster analysis.

1. Introduction

Quantative structure–activity relationship (QSAR) methods have been widely accepted and applied in the design of new bioactive compounds [1,2]. To find the best relationship between chemical structure and biological activity, many molecular parameters have been introduced and tested in QSAR studies [3]. Most of these parameters can easily be determined by various chromatographic methods [4]. Chromatographic techniques have some advantages: they are rapid and relatively simple; very small amounts of the substances are required and the compounds need not be very pure as the impurities separate during the chromatographic processes. Lipophilicity is one of the physicochemical parameters frequently applied in QSAR studies [5,6]. Lipophilicity can be determined by the classical partition method between water and *n*-octanol [7], by reversed-phase high-performance liquid chromatography (RP-HPLC) [8–10] and by reversed-phase thin-layer chromatography (RP-TLC) [11]. The lipophilicity values determined with RP-HPLC or RP-TLC generally show excellent correlation [12]. RP-TLC has been extensively applied to determine the lipophilicity of bioactive compounds [13]. To increase the accuracy of the lipophilicity determination, linear correlations have been calculated between the R_M values and the concentration of organic component in the mobile phase; the R_M value extrapolated to zero organic

* Corresponding author.

0021-9673/95/$09.50
SSDI 0021-9673(94)00807-8

phase concentration (R_{M0}) was regarded as the most accurate estimate of lipophilicity [14]. However, chromatographic methods have some drawbacks. The supports may partially retain their original adsorptive characteristics even after impregnation [15] and the R_M value changes with the amount [16]) and quality [17] of coating substance.

Much effort has been devoted to the elucidation of the mode of action of various anticancer drugs. They can bind to different biomolecules such as model and native membranes [18], DNAs [19,20] and various proteins [21]. The binding of anticancer drugs to proteins may modify protein structures [22] and can increase or decrease enzyme activity [23,24], resulting in modified biological efficiency of the drugs [25]. It is reasonable to assume that the hydrophobicity parameters (lipophilicity and specific hydrophobic surface area) of drugs may have a significant impact on their various biological effects.

Multivariate mathematical–statistical methods such as principal component analysis (PCA) [26], factor analysis [27] and the spectral mapping technique [28] have been successfully used for the extraction of maximum information from large retention data matrices.

The objectives of our work were the determination of the lipophilicity and specific hydrophobic surface area of non-homogeneous series of anticancer drugs under various RP-TLC conditions, to find the relationships between retention characteristics and physico-chemical parameters by the use of multivariate mathematical–statistical methods and to compare the information content of the methods. The drugs studied are listed in Table 1.

2. Experimental

2.1. Reversed-phase thin-layer chromatography

Polygram UV 254 plates (Macherey–Nagel, Düren, Germany) were impregnated by overnight predevelopment in *n*-hexane–paraffin oil (95:5, v/v). The IUPAC nomenclature for the anticancer drugs used is shown in Table 1. Anticancer drugs were separately dissolved in methanol at a concentration of 3 mg/ml and 2 μl of the solutions were spotted on the plates. Methanol–water, ethanol–water, 1-propanol–water, 2-propanol–water, acetonitrile–water, dioxane–water and tetrahydrofurane–water mixtures were used as eluents, the concentration of organic modifiers ranging from 10 to 75 vol.-% in steps of 5 vol.-%. The application of this wide range of organic modifier concentrations was motivated by the highly different hydrophobicities of anticancer drugs. Developments were carried out in sandwich chambers ($22 \times 22 \times 3$ cm) at room temperature, the distance of development being about 16 cm. After development the plates were dried at 105°C and the spots of anticancer drugs were revealed by their visible and UV spectra, with iodine vapour and with phosphomolybdic acid reagent. Each experiment was run in quadruplicate. The R_M values of the anticancer drugs were determined by

$$R_M = \log (1/R_F - 1) \quad (1)$$

Linear correlations between the R_M values of anticancer drugs and the concentration of organic modifiers in the eluent were calculated separately for each drug:

$$R_M = R_{M0} + bC \quad (2)$$

The intercept and slope values of Eq. 2 were considered as the best estimation of the hydrophobicity and specific hydrophobic surface area [29] of the drugs, respectively. It has been stated [29] that this slope perhaps reflect hydrocarbonaceous surface area but only for a set of solutes of closely similar electrostatic properties (dipole moments, polarizabilities). Obviously this set of compounds did not comply with the requirements mentioned above. However, it has recently been reported that the assumption also holds for non-homologous series of solutes [30,31]. In the case of a given drug the slope values are also related to the solvent strength of the organic modifiers (decrease in the R_M value caused by a 1% (v/v) increase in the concentration of the organic modifier [32]).

Table 1
Anticancer drugs studied

No.	Common name	IUPAC name	Source
1	Ftorafur	N-(2-Furanidyl)-5-fluorouracil	Medexport (Russia)
2	Bicnu	N,N-Bis(2-chloroethyl)-N-nitrosourea	Laboratoire Bristol (France)
3	Vincristin	22-Oxo-(3α,14β,16α)-14,15-dihydro-14-hydroxy-eburnamenine-14-carbocyclic acid methyl ester	Gedeon Richter (Hungary)
4	Vinblastine	(3α,14β,16α)-14,15-Dihydro-14-hydroxyeburnamenine-14-carbocyclic acid methyl ester	Gedeon Richter (Hungary)
5	Vumon	4′-O-Demethyl-1-O-(4,6-O-2-thenylidene-β-D-glucopyranosyl)epipodophyllotoxin	Bristol-Arzneimittel (Germany)
6	Provera	17-α-Acetoxy-6-α-(methyl)progesterone	Upjohn (UK)
7	Bleogin	N^1-[3-Dimethyl(sulfonio)propyl]bleomycinamide	Nippon Kayaku (Japan)
8	Paraplatin	9,11,15-Trihydroxy-15-methylprosta-5,13-dienoic acid	Bristol-Arzneimettel (Germany)
9	Zitazonium	2-[4-(2-Chloro-1,2-diphenylethynyl)phenoxy]-N,N-diethylethamine citrate	EGIS Pharmaceutical Works (Hungary)
10	Farmorubicin	(8*S*-*cis*)-10-[(3-Amino-2,3,6-trideoxy-α-L-*arabino*-hexopyranosyl)oxy]7,8,9,10-tetrahydro-6,8,11-trihydroxy-8-(hydroxyacetyl)-1-methoxy-5,12-naphthacenedione	Farmitalia (Italy)
11	Adriblastine (Doxorubicin)	10-[3-(Amino-2,3,6-trideoxy-α-L-hexapyranosyl)oxy]-7,8,9-tetrahydro-6,8,11-trihydroxy-8-(hydroxyacetyl)-1-methoxy-5,12-naphthacenedione	Farmitalia (Italy)
12	Natulan	*N*-(1-Methylethyl)4-[(2-methylhydrazino)methyl]-benzamide	Roche (Switzerland)
13	Alexan	4-Amino-1-β-D-arabifuranosyl-2(1*H*)-pyrimidine	Mack (Germany)
14	Mitomycin C Kyowa	[1-a*R*]-6-Amino-8-[(aminocarbonyl)oxymethyl]-1,1a,2,8,8a,8b-hexahydro-8a-methoxy-5-methyl azirino-[2′,3′:3,4]pyrrolo[1,1*a*]indole-4,7-dione	Kyowa (Japan)
15	Cytoxan	2-[Bis(2-chloroethyl)amino]tetrahydro-2*H*-1,3,2-oxazaphosphorine 2-oxide monohydrate	Bristol-Myers (Germany)
16	Estracyt	Estra-1,3,5-(10)-triene-3,17-diol-3-[bischloroethyl)carbamate	Aktiebolaget (Sweden)
17	Deticene	5-(3,3-Dimethyl-1-triazenyl)-1*H*-imidazole-4-carboxamide	Rhone-Poulenc (France)
18	Methotrexate	2,4,-Diamino-10-methylpteroylglutamic acid	Lachema (Czech Republic)
19	Myelobromol	1,6-Dibromo-1,6-bis(desoxy)-D-mannitol	Chinoin (Hungary)
20	Zitostop	1,2,5,6-Tetramethylsulfon-D-mannitol	EGIS Pharmaceutical Works (Hungary)
21	Elobromol	1,6-Dibromo-1,6-bis(desoxy)-D-galactitol	Chinoin (Hungary)
22	Taxol	{2a*R*-[2aα,4β,4aβ,6β,9α(a*R**,β*S**),11α,12α,12aα,12bα}-β-(Benzoylamino)-α-hydroxybenzenepropanoic acid 6,12b-bis(acetyoxy)-12-(benzoyloxy)-2a,3,4,4a,5,6,9,10,11,12,12a,12b-dodecahydro-4,11-dihydroxy-4a,8,13,13-tetramethyl-5-oxo-7,11-methano-1*H*-cyclodeca[3,4]benz[1,2-*b*]oxet-9-yl ester	Sigma Chemie (Germany)

2.2. Multivariate mathematical–statistical methods

Spectral mapping technique [33]

In order to separate the elution strength and elution selectivity of organic modifiers, the spectral mapping technique was applied. The *b* values of Eq. 2 and the anticancer drugs were the variables and observations, respectively. The dimensions of the spectral map were reduced to two by the non-linear mapping technique [34]. Stepwise regression analysis [35] was used for the

elucidation of the impact of the physico-chemical parameters of organic modifiers on their elution strength and selectivity. We should stress that the elution strength probably depends on the solutes specifically used in the experiments, therefore the conclusions drawn are valid only for this set of solutes and any extrapolation may lead to severe misinterpretation of the results. Stepwise regression analysis was used three times, the elution strength and the two coordinates of the spectral map being the dependent variables. The independent variables were in each instance the following physico-chemical parameters of organic modifiers:

π	Hansch–Fujita substituent constant characterizing hydrophobicity [36,37];
H-Ac and *H-Do*	indicator variables for proton acceptor and proton donor properties, respectively [38];
M-RE	molar refractivity [39];
F and *R*	electronic parameters characterizing the inductive and resonance effect, respectively [40];
σ	Hammett's constant, characterizing the electron-withdrawing power of the substituent [41];
Es	Taft's constant, characterizing steric effects of the substituent [42];
B_1, B_4 and B_1/B_4	Sterimol width parameters determined by the distance of substituents at their maximum point perpendicular to attachment [43,44].

The two equations used in stepwise regression analysis were the following:

Elution strength

$$= a + b_1\pi + b_2\text{H-Do} + b_3\text{M-RE} + b_4F + b_5R + b_6\sigma + b_7Es + b_8B_1 + b_9B_4 + b_{10}B_1/B_4 \quad (3)$$

First coordinate of the spectral map

$$= a + b_1\pi + b_2\text{H-Do} + b_3\text{M-RE} + b_4F + b_5R + b_6\sigma + b_7Es + b_8B_1 + b_9B_4 + b_{10}B_1/B_4 \quad (4)$$

The acceptance level for the individual independent variables was set to the 95% significance level.

Principal component analysis [45] followed by two-dimensional non-linear mapping, cluster analysis and varimax rotation [46]

The data matrix consisted of the parameters of Eq. 2 (intercept = R_{M0} and slope = b values determined with the seven organic modifiers) and various physico-chemical characteristics of drugs (altogether 23 variables) were considered as variables and the anticancer drugs were the observations. The data matrix is compiled in Table 2. As compounds **19–21** were very near to the front these data were omitted from the calculations. The physico-chemical parameters included in the calculation were as listed above. The limit of the variance explained was set to 99.9%. To facilitate the evaluation of the results of PCA, both two-dimensional non-linear mapping and cluster analysis were carried out on the principal component loadings and variables. To compare the information content of the two-dimensional non-linear map and that of varimax rotation, varimax rotation around two axes was carried out on the principal component loadings. As it was assumed that PCA can cause some data distortion, cluster analysis was also applied to the original data matrix. Cluster analysis, the non-linear mapping technique and varimax rotation are theoretically similar: each method calculates and visualizes the relative distances between the members of data matrix (in our cases physico-chemical and chromatographic parameters of drugs). To compare their information content, linear correlations were calculated between the corresponding distances on the non-linear map and cluster dendogram:

$$Y = a + bX_{1,2} \quad (5)$$

Table 2
Retention characteristics and physico-chemical parameters of anticancer drugs

Parameter	Anticancer drug																		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**22**
$R_{M(MeOH)}$	0.33	1.45	2.20	1.97	1.73	2.33	1.35	0.32	3.81	2.13	1.90	1.24	0.99	1.17	1.33	4.55	0.94	0.88	4.08
$b_{(MeOH)}$	1.57	2.14	2.47	2.24	1.81	2.85	1.31	1.26	3.18	3.00	2.39	2.83	6.27	3.00	2.59	6.39	1.82	2.53	6.29
$R_{M(EtOH)}$	0.32	1.04	2.26	2.14	2.20	3.42	1.30	−0.36	3.57	1.95	1.71	1.18	0.58	1.23	1.25	3.16	0.84	1.00	2.72
$b_{(EtOH)}$	2.87	2.19	3.91	3.71	4.88	6.63	4.70	1.94	6.14	4.28	3.70	3.92	10.00	5.01	3.48	5.11	2.46	6.12	5.16
$R_{M(1\text{-}PROP)}$	0.09	1.01	1.48	1.77	1.81	2.28	1.29	−0.49	2.34	1.38	1.20	0.98	0.48	0.70	0.92	1.86	0.63	1.06	1.94
$b_{(1\text{-}PROP)}$	2.06	3.46	4.41	5.22	6.28	6.96	2.19	1.91	6.21	4.80	4.24	4.12	10.59	3.61	3.60	5.03	2.22	10.02	5.27
$R_{M(2\text{-}PROP)}$	0.27	1.03	1.50	1.67	1.77	2.90	1.21	0.52	3.18	1.51	1.38	0.96	0.43	0.72	0.98	2.70	0.61	0.77	2.41
$b_{(2\text{-}PROP)}$	4.06	2.61	3.22	3.52	4.70	6.95	4.85	1.57	7.39	4.27	3.96	3.67	9.81	3.49	3.26	5.59	2.14	6.28	5.56
$R_{M(MeCN)}$	0.33	1.16	2.15	2.20	2.33	3.44	1.44	0.59	2.71	1.75	1.55	1.16	1.24	1.24	1.24	2.86	0.72	0.99	2.98
$b_{(MeCN)}$	2.85	2.90	4.23	4.30	5.59	7.16	6.54	0.93	5.34	4.57	3.89	3.55	15.53	5.62	3.36	6.52	2.32	6.36	6.56
$R_{M(DIOX)}$	0.19	0.96	1.63	1.82	2.01	2.64	1.26	−0.90	2.38	1.47	1.41	0.76	0.42	0.46	1.02	2.20	0.50	0.48	2.65
$b_{(DIOX)}$	2.91	2.24	2.86	3.09	4.71	4.86	6.19	0.50	3.76	3.93	4.30	2.79	10.70	2.70	3.69	3.66	1.91	3.77	5.41
$R_{M(THF)}$	0.24	1.11	1.43	1.53	1.71	2.39	1.35	0.04	1.89	1.02	0.99	0.31	−0.89	0.18	0.87	2.90	0.12	−0.12	2.33
$b_{(THF)}$	5.83	1.31	2.59	2.51	3.19	4.41	8.24	3.83	2.98	2.45	2.45	1.47	1.17	1.64	2.63	6.40	0.71	2.72	3.64
π	−0.81	0.37	16.72	17.93	7.44	8.76	−16.29	7.98	3.84	−1.56	−1.56	−1.07	−6.11	1.43	0.38	13.76	2.28	−5.23	9.11
H-Do	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
M-Re	103.24	52.63	267.87	266.14	248.32	158.54	487.04	117.98	127.44	213.16	213.16	72.25	99.21	154.62	43.88	152.82	41.67	146.51	865.81
F	1.01	0.50	1.25	0.90	1.96	0.06	4.01	0.51	1.37	1.52	1.52	0.06	0.05	1.77	0.63	−0.56	0.68	−0.69	1.88
R	2.13	−0.72	−0.13	−0.39	−4.54	−0.63	−15.28	3.93	−1.29	−2.91	−2.91	−2.08	−4.38	−3.58	0.27	−0.89	−1.77	−6.97	−5.44
σ	1.44	1.43	0.86	0.44	2.38	0.76	3.94	−0.26	0.78	−0.69	0.69	−0.47	0.34	−3.18	0.74	−0.59	0.56	1.35	3.33
Es	−7.16	7.24	−30.09	−29.80	−26.19	19.97	−82.19	−28.80	21.83	−27.46	−27.46	−9.30	−10.84	23.21	−6.21	−17.82	−8.02	−34.05	−5.34
B_1	10.00	9.90	27.67	27.59	27.10	16.20	79.43	30.73	13.66	29.61	29.61	9.24	16.78	22.45	7.42	19.32	6.18	60.56	62.49
B_4	13.24	14.68	56.59	56.27	40.89	26.30	92.28	41.23	21.10	39.58	39.58	16.86	23.74	31.94	10.92	34.16	12.07	43.76	79.08

where Y = relative distances between anticancer drugs on the non-linear map X_1 = relative distances between anticancer drugs on the cluster dendogram calculated from the original data matrix and X_2 = relative distances between anticancer drugs on the cluster dendogram after PCA. To facilitate the calculations only the distances between the nearest neighbour drugs on the maps were included in the equations. The comparison of distances was hampered by the fact that their absolute values depend on the dimensions of the plots. We overcame this difficulty by data normalization: the greatest distances on each map were considered to be 100% and the other distance were calculated accordingly.

To compare the information content of non-linear mapping and varimax rotation techniques, linear correlations were calculated between their corresponding coordinates:

$$Y_{1-2} = a + bX_{1-2} \quad (6)$$

where Y_{1-2} = first and second coordinates of the varimax rotation and X_{1-2} = first and second coordinates of the non-linear map.

3. Results and discussion

Compounds **19–21** were near to the front in each eluent system, which means that these drugs are highly hydrophilic and their hydrophobicity parameters cannot be determined under the experimental conditions used. Eq. 2 was significant for each anticancer drug in each eluent system, that is, these solutes do not show anomalous retention behaviour which can inhibit the calculation of their hydrophobicity and specific hydrophobic surface area (see retention data in Table 2).

The solvent strength and solvent selectivity of organic modifiers are compiled in Table 3. The solvent strength was calculated taking into consideration simultaneously the slope value of Eq. 2 for each anticancer drug. It is a dimensionless number without a concrete physico-chemical meaning, giving the order of solvent strengths of the organic modifiers relative to each other. The data clearly show that the solvent strengths of organic modifiers differ considerably; for alcohols they increase with increasing number of methylene groups in the molecule. The results further indicate that not only the solvent strengths but also the solvent selectivities are markedly different. Stepwise regression analysis selected two equations describing the relationship between the chromatographic parameters of organic modifiers and their physico-chemical parameters. The chromatographic parameters of each anticancer drug except compounds **19–21** were included in the calculations, and therefore the terms solvent strength and solvent selectivity (coordinates of the spectral map) refer to the solutes mentioned above:

Solvent strength

$$= -286.9 + 1003.6B_1/B_4 - 814.1(B_1/B_4)^2$$

$$F_{calc.} = 22.65\ ;\ r^2 = 0.9189$$

Table 3
Solvent strength and selectivity of organic modifiers: results of spectral mapping technique

Organic modifier	Solvent strength (arbitrary units)	Selectivity map	
		First coordinate	Second coordinate
Methanol	12.84	101.52	114.90
Ethanol	19.78	154.12	99.75
1-Propanol	21.15	113.44	63.09
2-Propanol	19.93	164.83	80.87
Acetonitrile	22.51	174.23	148.17
Dioxane	16.97	199.07	108.26
Tetrahydrofuran	13.80	212.79	4.42

First coordinate of spectral map

$$= 457.0 - 487.3B_1/B_4$$

$$F_{calc.} = 21.73\ ;\ r^2 = 0.8129$$

Good relationships were found between the derived retention parameters (solvent strength and the first coordinate of the spectral map) and the physico-chemical parameters of the anticancer drugs, the significance level being over 99% (see $F_{calc.}$ values). The steric parameter (molecular shape) accounts for 91 and 81% of the change in elution strength and elution selectivity of organic modifiers, respectively.

The results of PCA are summarized in Table 4. Six principal components explain the majority (91.67%) of the total variance. This result indicates that the 23 physico-chemical and chromatographic parameters can be substituted by six background variables without a substantial loss of information. Unfortunately, PCA does not prove the existence of such background variables

Table 4
Relationship between retention characteristics and physico-chemical parameters for anticancer drugs: results of principal component analysis

No. of principal component	Eigenvalue	Variance explained (%)	Total variance explained (%)
1	7.69	33.42	33.42
2	4.89	21.24	54.66
3	4.27	18.57	73.23
4	2.18	9.49	82.72
5	1.14	4.97	87.69
6	0.92	3.98	91.67

Parameter	Principal component loadings: No. of principal component				
	1	2	3	4	5
$R_{M0(MeOH)}$	−0.18	0.01	0.13	0.87	0.28
$b_{(MeOH)}$	0.55	0.36	−0.38	0.05	0.50
$R_{M0(EtOH)}$	0.83	0.36	0.37	−0.01	−0.16
$b_{(EtOH)}$	0.61	0.17	−0.73	0.06	−0.08
$R_{M0(1\text{-}PROP)}$	0.86	0.23	0.32	−0.06	−0.23
$b_{(1\text{-}PROP)}$	0.51	0.30	−0.66	−0.08	−0.08
$R_{M0(2\text{-}PROP)}$	0.85	0.32	0.35	0.07	−0.16
$b_{(2\text{-}PROP)}$	0.67	0.16	−0.63	0.29	−0.08
$R_{M0(MeCN)}$	0.90	0.32	0.23	−0.04	−0.05
$b_{(MeCN)}$	0.56	0.07	0.78	0.03	0.06
$R_{M0(DIOX)}$	0.88	0.23	0.37	0.01	−0.06
$b_{(DIOX)}$	0.58	0.11	−0.71	0.13	0.05
$R_{M0(THF)}$	0.71	0.13	0.64	0.01	0.01
$b_{(THF)}$	0.34	−0.53	0.25	0.42	−0.01
π	0.18	0.50	0.54	−0.20	0.34
H-Do	0.04	0.12	−0.12	0.88	0.11
M-Re	0.61	0.49	0.18	0.11	0.49
F	0.30	0.66	0.30	0.23	−0.05
R	−0.20	0.89	0.29	0.01	0.10
σ	0.40	0.56	0.12	0.28	0.15
Es	−0.23	0.77	−0.03	0.14	0.45
B_1	0.42	0.80	−0.06	−0.20	0.11
B_4	0.51	0.72	0.14	−0.28	0.20

as concrete physico-chemical entities but only indicates its mathematical possibility. The majority of retention parameters (hydrophobicity = R_{M0} and specific hydrophobic surface area = b values) have high loadings in the first PC, indicating that this principal component can be regarded as related to the molecular hydrophobicity. The second PC contains only the calculated physico-chemical parameters of drugs. The specific hydrophobic surface area values and the calculated hydrophobicity (π) values have high loadings in the third PC, suggesting a more or less strong relationship between these parameters.

Neither the retention characteristics nor the physico-chemical parameters of anticancer drugs

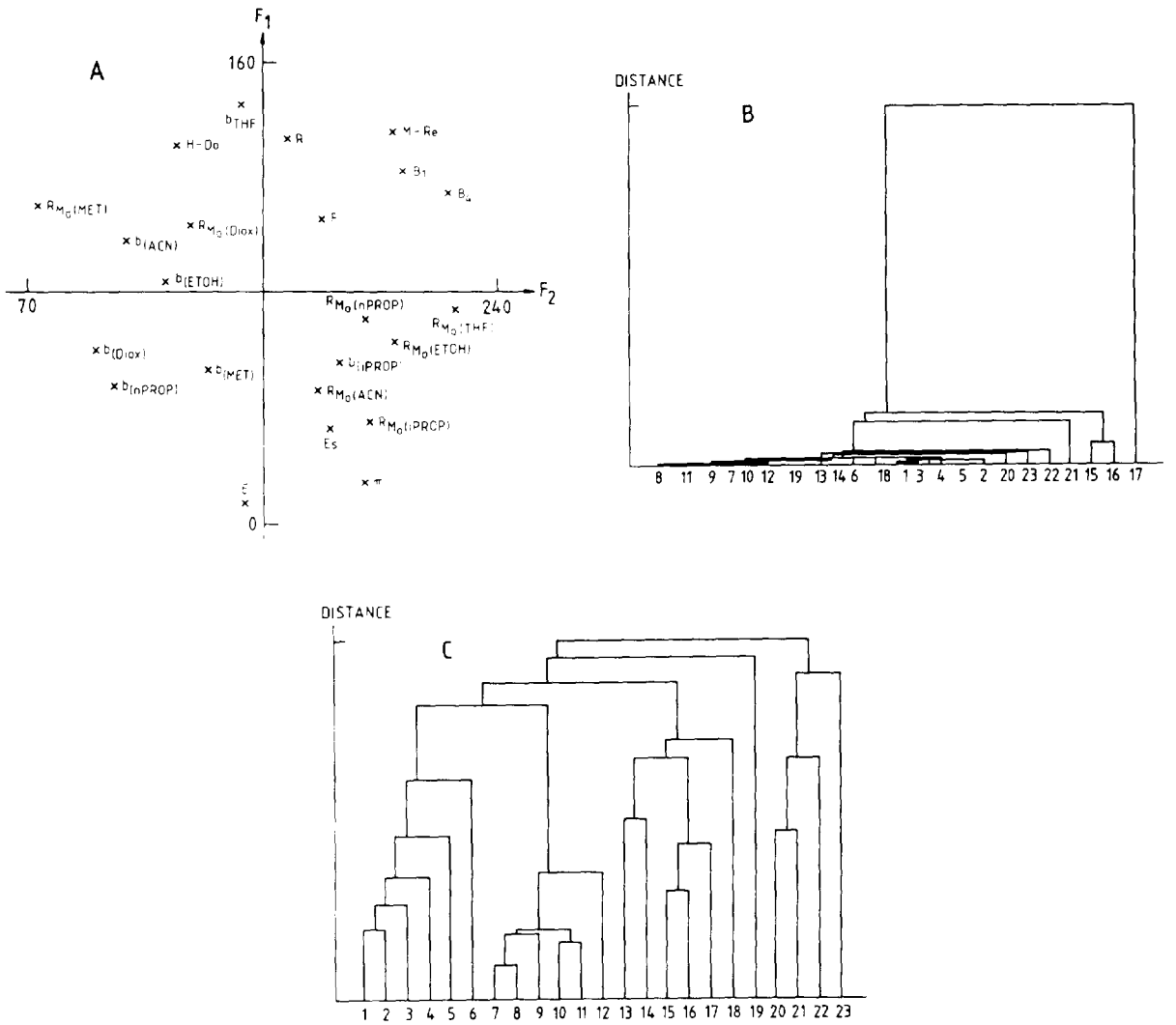

Fig. 1. Similarities and dissimilarities between the retention characteristics and physico-chemical parameters of anticancer drugs. (A) Two-dimensional non-linear map of PC loadings. Number of iterations, 90; maximum error $2.2 \cdot 10^{-3}$. (B) Cluster dendogram calculated from the original data matrix. (C) Cluster dendogram calculated from the PC loadings. Numbers on abscissa: 1 = $R_{M0(MeOH)}$; 2 = $b_{(MeOH)}$; 3 = $R_{M0(EtOH)}$; 4 = $b_{(EtOH)}$; 5 = $R_{M0(1\text{-}PROP)}$; 6 = $b_{(1\text{-}PROP)}$; 7 = $R_{M0(2\text{-}PROP)}$; 8 = $b_{(2\text{-}PROP)}$; 9 = $R_{M0(MeCN)}$; 10 = $b_{(MeCN)}$; 11 = $R_{M0(DIOX)}$; 12 = $b_{(DIOX)}$; 13 = $R_{M0(THF)}$; 14 = $b_{(THF)}$; 15 = π; 16 = *H-Do*; 17 = *M-RE*; 18 = *F*; 19 = *R*; 20 = σ; 21 = *Es*; 22 = B_1; 23 = B_4.

form distinct clusters on the two-dimensional non-linear map of PCA loadings and on the cluster dendograms (Fig. 1). These results indicate that each measured and calculated parameter has a different information content, they cannot be replaced by each other and they can be used separately in quantitative structure–activity relationship calculations. The findings discussed above are supported by the fact that only one significant linear correlation was found between the distances on the various maps (two-dimensional non-linear map and cluster analysis carried out on PC loadings: $r_{calc.} = 0.9884$, $r_{99\%} = 0.6524$). Although cluster analysis and non-linear mapping give theoretically similar results, we strongly advocate the application of the two-dimensional non-linear mapping technique instead of cluster analysis because on the two-dimensional non-linear map the observations or variables are distributed in a plane whereas they are along a line in the cluster analysis.

Significant linear correlations were found between the rotated PC loadings and the coordi-

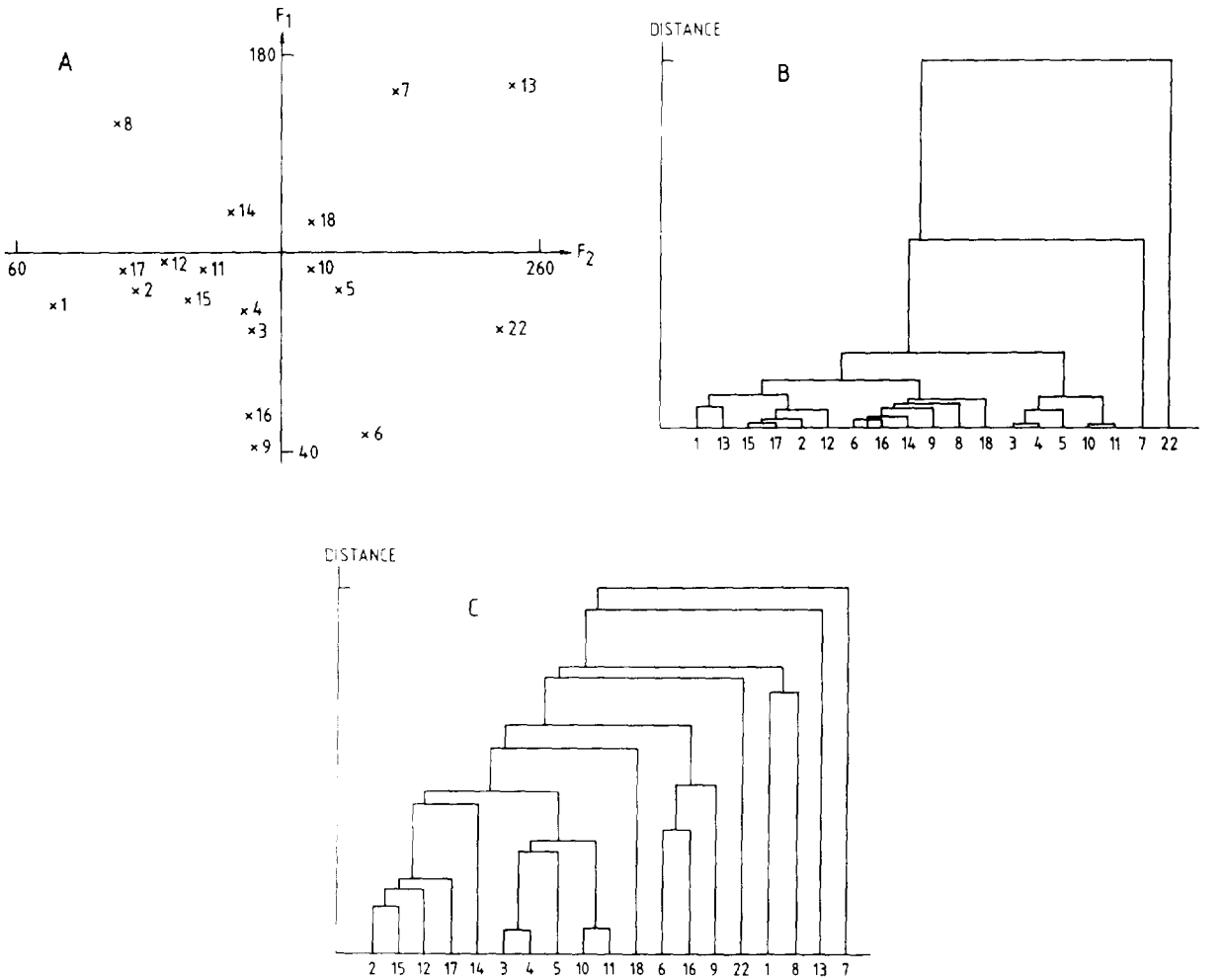

Fig. 2. Distribution of anticancer drugs according to their retention characteristics and physico-chemical parameters. (A) Two-dimensional non-linear map of PC variables. Number of iterations, 178; maximum error, $5.3 \cdot 10^{-3}$. (B) Cluster dendogram calculated from the original data matrix. (C) Cluster dendogram calculated from the PC variables. Numbers refer to anticancer drugs in Table 1.

nates of two-dimensional non-linear map of PC loadings ($n = 22$):

$$\text{varimax}_1 = 2.45 - (2.22 \pm 0.28) \cdot 10^{-3} \cdot \text{nlmap}_1$$

$$r_{\text{calc.}} = 0.9765 \; ; \; r_{99.9\%} = 0.6524$$

$$\text{varimax}_2 = 1.43 - (9.21 \pm 5.89) \cdot 10^{-3} \cdot \text{nlmap}_2$$

$$r_{\text{calc.}} = 0.9431 \; ; \; r_{99.9\%} = 0.6524$$

These data indicate that the information contents of varimax rotation and two-dimensional non-linear mapping (nlmap) are similar but not identical and both methods can be used to decrease the dimensionality of complicated data matrices.

The distribution of anticancer drugs according to their retention characteristics and physico-chemical parameters on the various maps is shown in Fig. 2. Compounds with similar retention characteristics and physico-chemical parameters are near to each other on the map whereas compounds with different characteristics are far from each other.

Acknowledgement

This work was supported by grant OTKA 7340 of the Hungarian Academy of Sciences.

References

[1] J. Andrew, W.E. Stuper, P. Brugger and S. Jurs, *Computer Assisted Studies of Chemical Structure and Biological Function*, Wiley, New York, 1979.
[2] C.J. Cavallito, *Structure–Activity Relationships*, Vol. 1, Pergamon Press, New York, 1973.
[3] R. Franke, in J.K. Seydel (Editor), *QSAR and Strategies in the Design of Bioactive Compounds*, VCH, Weinheim, 1985, p. 59.
[4] R. Kaliszan, *Quantitative Structure–Chromatographic Retention Relationships*, Wiley, New York, 1987.
[5] A. Fini, A. Roda, A.M. Bellini, E. Mencini and M. Guarneri, *J. Pharm. Sci.*, 79 (1990) 603.
[6] M. Vaara, W.Z. Plachya and H. Nikaido, *Biochim. Biophys. Acta*, 1042 (1990) 152.
[7] C. Hansch and S.M. Anderson, *J. Org. Chem.*, 32 (1967) 2583.
[8] A. Siwek and J. Sliwiok, *J. Chromatogr.*, 506 (1990) 109.
[9] W.J. Lambert, L.A. Wright and J.K. Stevens, *J. Pharm. Res.*, 7 (1990) 577.
[10] C. Yamagami, T. Ogura and N. Takao, *J. Chromatogr.*, 514 (1990) 123.
[11] C.B.C. Boyce and B.V. Milborrow, *Nature*, 208 (1965) 537.
[12] G.L. Biagi, M.C. Guerra, A.M. Barbaro, S. Barbieri, M. Recanatini and P.A. Borea, *J. Liq. Chromatogr.*, 13 (1990) 913.
[13] T. Cserháti and K. Magyar, *J. Biomed. Biophys. Methods*, 24 (1992) 249.
[14] J. Draffehn, B. Schonecker and K. Ponsold, *J. Chromatogr.*, 205 (1980) 113.
[15] W.F. Giesen and L.H.M. Janssen, *J. Chromatogr.*, 237 (1982) 199.
[16] T. Cserháti, Y.M. Darwish and Gy. Matolcsy, *J. Chromatogr.*, 270 (1983) 97.
[17] L. Ogierman and A. Silovecki, *J. High Resolut. Chromatogr. Chromatogr. Commun.*, 4 (1981) 357.
[18] J.B.A. Custodio, L.M. Almeida and V.M.C. Madeira, *Biochim. Biophys. Acta*, 1150 (1993) 123.
[19] T. Araka, T. Kusakabe, J. Kuwahara, M. Otsuka and Y. Sugiura, *Biochem. Biophys. Res. Commun.*, 190 (1993) 362.
[20] G.J. Finlay, E. Marshall, J.H.L. Matthews, K.D. Paull and B.C. Baguley, *Cancer Chemother. Pharmacol.*, 31 (1993) 401.
[21] M. Kimura, *Yakugaku Zasshi*, 112 (1992) 914.
[22] J.A. Broomhead, L.M. Rendina and L.K. Webster, *J. Inorg. Biochem.*, 49 (1993) 221.
[23] D.W. Fry, T.J. Boritzki, R.C. Jackson, P.D. Cook and W.R. Leopold, Mol. Pharmacol., 44 (1993) 479.
[24] C. Leguellec, B. Lacarelle, J. Catalin and A. Durand, *Cancer Chemother. Pharmacol.*, 32 (1993) 491.
[25] G.N. Kumar, U.K. Walle, K.N. Bhalla and T. Walle, *Res. Commun. Chem. Pathol. Pharmacol.*, 80 (1993) 337.
[26] R. Kaliszan and K. Osmialowski, *J. Chromatogr.*, 506 (1990) 3.
[27] B. Walczak, M. Dreux, J.R. Chretien, L. Morín-Allory, M. Lafosse and G. Felix, *J. Chromatogr.*, 464 (1989) 237.
[28] T. Cserháti, *J. Chromatogr. Sci.*, 29 (1991) 210.
[29] C. Horváth, W. Melander and I. Molnár, *J. Chromatogr.*, 125 (1976) 129.
[30] T. Cserháti and E. Forgács, *J. Chromatogr. A*, 665 (1994) 17.
[31] Y. Darwish, T. Cserháti and E. Forgács, *J. Chromatogr. A*, 668 (1994) 485.
[32] T. Cserháti, *J. Chromatogr. Sci.*, 31 (1993) 220.
[33] P.J. Lewi, *Arzneim.-Forsch.*, 26 (1976) 1295.
[34] J.W. Sammon, Jr., *IEEE Trans. Comput.*, C18 (1969) 401.
[35] H. Mager, *Moderne Regressionsanalyse*, Salle, Säuerlander, Frankfurt am Main, 1982, p. 135.
[36] T. Fujita, J. Iwasa and C. Hansch, *J. Am. Chem. Soc.*, 86 (1964) 5175.
[37] A. Leo, C. Hansch and M. Ames, *J. Pharm. Sci.*, 64 (1975) 559.

[38] C. Hasch and A. Leo, *Substituent Constants for Correlation Analysis in Chemistry and Biology*, Wiley, New York, 1979, p. 1.
[39] L. Pauling and D. Pressman, *J. Am. Chem. Soc.*, 67 (1945) 1003.
[40] R.W. Taft and I.C. Lewis, *J. Am. Chem. Soc.*, 80 (1958) 2436.
[41] L.P. Hammett, *Chem. Rev.*, 17 (1935) 125.
[42] R.W. Taft, *J. Am. Chem. Soc.*, 74 (1952) 3120.
[43] A. Verloop and J. Tipker, *Pestic. Sci.*, 7 (1976) 379.
[44] A. Verloop, W. Hoogenstraaten and J. Tipker, in J. Ariens (Editor), *Drug Design*, Vol. VII, Academic Press, New York, 1976, p. 165.
[45] K.V. Mardia, J.T. Kent and J.M. Bibby, *Multivariate Analysis*, Academic Press, London, 1979, p. 213.
[46] P. Willett, *Similarity and Clustering in Chemical Information Systems*, Research Studies Press, New York, 1987.