Chromatographic Hydrophobicity Index by Fast-Gradient RP-HPLC: A High-Throughput Alternative to log P/log D

ARTICLE *in* ANALYTICAL CHEMISTRY · JUNE 1997

Impact Factor: 5.64 · DOI: 10.1021/ac961242d · Source: PubMed

3 AUTHORS, INCLUDING:

Klara Valko
GlaxoSmithKline plc. Stevenage, Herts. UK
99 PUBLICATIONS **2,565** CITATIONS
SEE PROFILE

Derek Peter Reynolds
University of Cambridge
43 PUBLICATIONS **1,394** CITATIONS
SEE PROFILE

Anal. Chem. **1997,** *69,* 2022−2029

Chromatographic Hydrophobicity Index by Fast-Gradient RP-HPLC: A High-Throughput Alternative to log P/log D

Klára Valkó,* Chris Bevan, and Derek Reynolds

Department of Physical Sciences, GlaxoWellcome Medicines Research Centre, Stevenage, Herts, SG1 2NY, United Kingdom

A new chromatographic hydrophobicity index (CHI) is described which can be used as part of a protocol for high-throughput (50−100 compounds/day) physicochemical property profiling for rational drug design. The index is derived from retention times (t_R) observed in a fast gradient reversed-phase HPLC method. The isocratic retention factors (log k') were measured for a series of 76 structurally unrelated compounds by using various concentrations of acetonitrile in the mobile phase. By plotting the log k' as a function of the acetonitrile concentration, the slope (S) and the intercept (log k'_w) values were calculated. The previously validated index of hydrophobicity φ_0 was calculated as $-\log k'_w/S$. A good linear correlation was obtained between the gradient retention time values, t_R and the isocratically determined φ_0 values for the 76 compounds. The constants of this linear correlation can be used to calculate CHI. For most compounds, CHI is between 0 and 100 and in this range it approximates to the percentage (by volume) of acetonitrile required to achieve an equal distribution of compound between the mobile and the stationary phases. CHI values can be measured using acidic, neutral, or slightly basic eluents. Values corresponding to the neutral form of molecules could be measured for 52 of the compounds and showed good correlation ($r = 0.851$) to the calculated octanol/water partition coefficient (c log P) values.

Since the pioneering work of Meyer[1] and Overton,[2] the hydrophobic binding and the lipophilicity of drug molecules have been of great interest. Hansch and Fujita[3,4] suggested the use of the logarithmic value of the octanol/water partition coefficient (log P) for modeling the biological partition behavior of drug molecules.

The expressions "hydrophobicity" and "lipophilicity" are often used interchangeably, and their frequency of usage is different in various scientific fields. While chromatographers tend to use hydrophobicity, medicinal chemists prefer the lipophilicity term. According to the definitions suggested by IUPAC,[5,6] distribution behavior of compounds in a biphasic system, such as liquid/liquid or solid/liquid represents lipophilicity. The lipophilicity measure derived by our chromatographic method is not a simple binary solvent partition value, and we have therefore called it a "hydrophobicity" index.

The most recent monograph edited by Mannhold et al.[7] provides an overview of the role of lipophilicity in drug action and toxicology and explains the details of the various methods for the measurements and calculations of log P. A large data base of measured log P values is available[8] and several calculation methods[9−12] provide prediction of log P values. Due to the difficulties in the traditional shake flask method, there is a need for high-throughput methods for measurements of lipophilicity.

High-performance liquid chromatography in reversed-phase separation mode has long been recognized as a potential method for lipophilicity determination.[13−18] Various approaches have been described[17,18] which employ octanol in the chromatographic system or just use conventional octadecyl silica columns and hydroorganic mobile phases. When octanol-coated stationary phases are used with octanol-saturated aqueous mobile phases, only a narrow range of lipophilicity can be measured. These chromatographic systems usually show poor efficiency and offer measurement of only a limited range of lipophilicity. When highly efficient reversed-phase stationary phases are used with hydroorganic mobile phases, the correlation between the chromatographic partition data and the octanol/water partition data is weak when structurally unrelated compounds are investigated[19,20] due to the different nature of the partitioning solvents. To cover a wide range of lipophilicity, the mobile-phase composition should be adjusted and, in these cases, the log k' value extrapolated to the 0% organic-phase concentration (log k'_w or log k'_0) is used as

(1) Meyer, H. *Arch. Exp. Pathol. Pharmacol.* **1899**, *42*, 109−18.
(2) Overton, E. *Studien uber die Narkose.*; *Fischer*: Jena, 1901.
(3) Hansch, C.; Fujita, T. *J. Am. Chem. Soc.* **1964**, *86*, 1616−25.
(4) Leo, A. J.; Hansch, C.; Elkins, D. *Chem. Rev.* **1971**, *71*, 525−616.
(5) *Glossary of Terms Used in Medicinal Chemistry*, IUPAC, in press.
(6) *Glossary of Terms Used in Computational Drug Design*, IUPAC, in press.
(7) Mannhold, R.; Kubinyi, H.; Timmerman, H. In *Methods and Principles in Medicinal Chemistry*, Vol. 4, *Lipophilicity in Drug Action and Toxicology*; Pilska, V., Testa B., van de Waterbeemd H., Eds.; VCH: Weinheim, 1996.
(8) Hansch, C.; Leo, A. *MedChem Database*; Medicinal Chemistry Project, Pomona, CA, 1993; Leo A. J. *Chem. Revs.* **1993**, *91*, 1281.)
(9) Rekker, R. F. In *The Hydrophobic Fragmental Constant*; Pharmacochemistry Library I; Nauta, W. Th., Rekker, R. F., Eds.; Elsevier: Amsterdam, 1977.
(10) Hansch C.; Leo A. *c* log *P Program*; Medicinal Chemistry Project, Pomona, CA, 1993; (Leo, A. J. *Chem. Rev.* **1993**, *93*, 1281.)
(11) PALLAS for Windows, *ProLog P Module*; Compudrug Chemistry, Budapest, Hungary, 1994; Version 2.
(12) *ACD/log P*, Advanced Chemistry Development Inc.: Toronto, ON, Canada, 1996; Version 1.
(13) Haggerty, W. J.; Murrill, E. A. *Res. Dev.* **1974**, *25*, 30−42.
(14) Henry, D.; Block, J. H.; Anderson, J. L.; Carlson, G. R. *J. Med. Chem.* **1976**, *19*, 619−26.
(15) Mirrlees, M. S.; Moulton, S. J.; Murphy, C. T.; Taylor, P. J. *J. Med. Chem.* **1976**, *19*, 615−9.
(16) Unger, S. H.; Cook, J. R.; Hollenberg, J. S. *J. Pharm. Sci.* **1978**, *67*, 1364−6.
(17) Braumann, Th.; Weber, G.; Grimme L. H. *J. Chromatogr.* **1983**, *261*, 329−43.
(18) Lambert, W. J. *J. Chromatogr., A* **1993**, *656*, 469−84.
(19) Harnisch, M.; Mockel, H. J.; Schulze, G. J. *J. Chromatogr.* **1983**, 282, 315−28.
(20) Valkó, K. In *Chromatography, the State of the Art*; Kalász, H., Ettre, L., Eds.; Akadémia: Budapest, 1985; pp 739−50.

S0003-2700(96)01242-5 CCC: $14.00

an HPLC hydrophobicity parameter. With the use of standard reference compounds with known log P values, this method is suggested by OECD guidelines.[19] It has been shown that the correlation is poor between the log P values and the log k' values extrapolated to the 0% organic modifier (i.e., log P vs log k'_0 or log k'_w), when structurally unrelated compounds are investigated.[20] However, the correlation can be improved by using the slope value (S) of the log k' vs organic-phase concentration line as a second parameter ($\log P = a \log k'_w + bS + c$). This approach means that the log k' values are extrapolated backward to an optimum organic-phase concentration (rather than to 0%) in the mobile phase, which best models the octanol/water partition.[21] The regression coefficients (a, b, and c) are dependent on the properties of the reversed-phase column[22] and the investigated set of compounds. This approach has the advantage that it allows optimization of the HPLC partition system to model not only octanol/water partition but also biological partition directly.[23]

Another approach has been recently introduced[24,25] to overcome the low correlation between the chromatographic retention data and log P values when diverse sets of compounds are considered. The φ_0 value of a compound was defined as the percentage (by volume) of acetonitrile required to achieve an equal distribution of compound between the mobile and stationary phases (i.e., $\log k' = 0$). For most compounds this is a physically attainable volume percent of organic phase with a value between 0 and 100%. Sometimes compounds have values out of this range that can be obtained by extrapolation, i.e., when either the compound is highly hydrophobic and $\log k' > 0$ with 100% organic phase or when a compound is highly hydrophilic and $\log k' < 0$ with neat aqueous mobile phase. The method involves measurement of log k' values with various organic solvent concentrations in the mobile phase (preferably close to that concentration region where $\log k' = 0$). By plotting the log k' values as a function of the organic solvent concentration, the φ_0 value can be obtained from the slope and the intercept of the straight line ($\varphi_0 = -\text{intercept}/\text{slope}$). This method means that the estimation of the φ_0 values can be done from a bracketing range of values, unlike the extrapolation to log k'_w, which is usually outside of the measurable range and whose value is strongly dependant on the curve of the log k' vs φ plot. In practice, the values obtained for log k'_w by extrapolation back to pure aqueous mobile phase for the same compound may be different if different organic modifiers are used. In our method for the determination of φ_0 values, we prefer to use the term "intercept" rather than log k'_w, as it is used for expressing the coordinates of one point only of the log k' vs φ plot. The φ_0 is an index that characterizes the compound hydrophobicity, as a higher organic-phase concentration is needed for more hydrophobic compounds to achieve the equal distribution. A good correlation of φ_0 values to the log P values was shown for almost 500 drug molecules[25] under conditions where the mobile-phase pH suppresses the ionization of the compound. This relationship between the c log P and φ_0 values has been utilized successfully in an expert system for HPLC method development without preliminary experiments.[26]

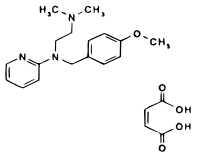
The experimental hydrophobicity index φ_0 shows acceptable correlation with the calculated log P values, and has the advantages of all chromatographic methods (i.e., it can be automated, only a small amount of sample is required, impurities do not disturb the measurements, etc.) However, the method cannot be regarded as strictly high throughput, as the retention times have to be measured at several isocratic mobile-phase compositions, which must be decided before the experiment.

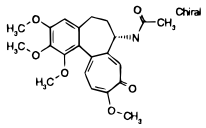
With the application of combinatorial synthesis of small molecules, the preparation of a large number of compounds is feasible and has consequently become an important additional source of molecular diversity for drug discovery screens.[27] High throughput in vitro pharmacological screening is also available in several pharmaceutical companies. The importance of physicochemical properties such as lipophilicity, solubility, pK, etc., is widely recognized in the design principles for orally bioavailable drugs,[28] for blood brain barrier distribution, formulation, etc. As a consequence, there is an urgent need for high throughput physicochemical property measurements to help understand structure - activity relationships and predict the bioavailability, pharmacokinetic properties, and in vivo distribution data of compounds.[29] It has been reported recently that retention parameters (log k'_w) obtained on high bonding density C-18 phases provide a better model of biopartitioning and bioavailability than the octanol/water partition data.[30]

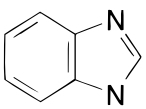
It has been reported that the prediction of precise isocratic retention data from two or more gradient elution runs is possible.[31] When the retention data are measured by using a slow and a fast gradient run, the S and log k'_w values can be estimated. A so-called $\bar{\varphi}$ value has also been defined as the organic-phase concentration at band center when the band is at the midpoint of column. We believe that this is basically an equivalent parameter to the φ_0 value. A generic gradient HPLC method has been developed[32] for the quality control of a wide variety of compounds without preliminary method development. By the application of third-generation reversed-phase columns, excellent separation efficiency and peak shape can be achieved for almost every compound. The method was speeded up, and with slightly modified mobile phase it allows an identification and purity check of compounds within 5−10 minutes by utilizing HPLC/electrospray Mass Spectrometry.[33] The question arose as to whether a lipophilicity parameter could be obtained at the same time by using only one fast gradient run.

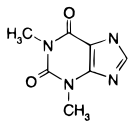
In principle, by applying a linear gradient increase of the organic-phase concentration, any point of the run time is equivalent to a certain mobile-phase composition. By knowing the void volumes of the column, mixing system, etc., it is possible to estimate the organic-phase concentration when the compound is eluting from the column. Considering a fast gradient run, the S parameter (i.e., the slope of the log k' vs organic modifier concentration straight line plot) will only have a small influence

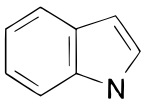
(21) Valkó, K. *J. Liq. Chromatogr.* **1984**, *7*, 1405−24.
(22) Valkó, K. In *Chromatography '84*; Kalász, H., Ettre, L., Eds.; Akadémia: Budapest, 1986; pp 73−82.
(23) Valkó, K.; Friedmann, T.; Báti, J.; Nagykáldi, A. *J. Liq. Chromatogr.* **1984**, *7*, 2073−92.
(24) Valkó, K. *TrAC, Trends Anal. Chem.* **1987**, *6*, 214−9.
(25) Valkó, K.; Slégel, P. *J. Chromatogr.* **1993**, *631*, 49−61.
(26) Csókán, P.; Darvas, F.; Csizmadia, F.; Valkó, K. *LC-GC Int.* **1993**, *6*, 361−9.
(27) Patel, D. V.; Gordon, E. M. *Res. Focus* **1996**, *1*, 134−44.
(28) Navia, M. A.; Chaturvedi, P. R. *Res. Focus* **1996**, *1*, 179−89.
(29) Ashton, M. J.; Jaye, M. C.; Mason, J. S. *Res. Focus* **1996**, *1*, 71−78.
(30) Hsieh, M.-M.; Dorsey, J. G. *Anal. Chem.* **1995**, *67*, 48−53.
(31) Quarry, M. A.; Grob, R. L.; Snyder, L. R. *Anal. Chem.* **1986**, *58*, 907−917.
(32) Mutton, I. M. *J. Chromatogr., A* **1995**, *697*, 191−201.
(33) Boughtflower, R. B.; Lane, S. J.; Underwood, T.; Davidson, I.; Cook, T.; Brinded, K.; Traube, S. *Compound Characterization and Isolation—The Hands on Automated Approach*; HPLC'96, San Francisco, **1996**; abstract.

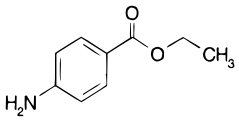
1  Tonsylamine maleate

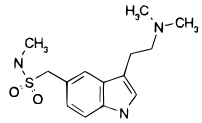
3  Colchicine

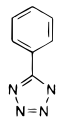
4  Benzimidazole

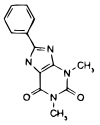
5  Theophylline

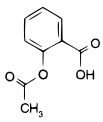
11  Indole

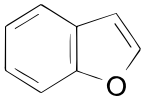
13  Benzocaine

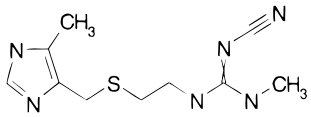
18  Sumatriptan

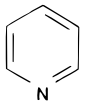
19  5-Phenyl-1H-tetrazole

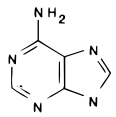
20  8-Phenyl-Theophylline

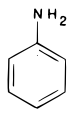
24  Acetyl salicylic acid

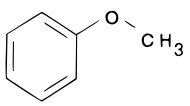
25  Benzofuran

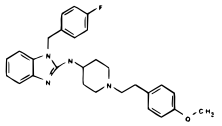
26  Cimetidine

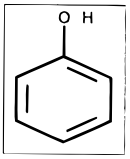
27  Pyridine

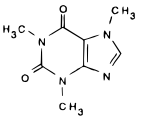
28  Adenine

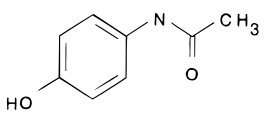
29  Aniline

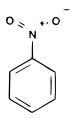
30  Anisole

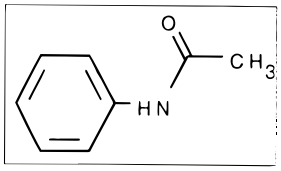
31  Astenizole

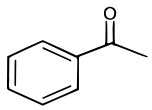
32  Phenol

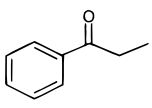
33  Caffein

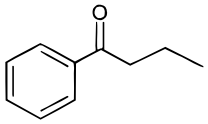
34  Paracetamol

35  Nitrobenzene

38  Acetanilide

73  Acetophenone

74  Propiophenone

75  Butyrophenone

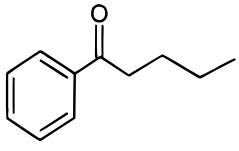
76  Valerophenone

Figure 1. Chemical structures of the commercially available and patented compounds.

Table 1. Measured and Calculated Data for the Investigated Compounds

					t_R			CHI			
no.	*c* log *P*	slope	log k'_w	φ_0 pH 7.3	pH 7.3	pH 2.6	pH 8.5	pH 7.3	pH 2.6	pH 8.5	CHIN[a]
1	2.77	−0.015	1.034	67	8.83	5.81	8.88	67.59	22.89	68.29	68.29
2	2.02	−0.017	0.779	47.1	7.81	6.25	7.68	52.48	29.37	50.57	52.48
3	0.92	−0.03	1.187	40	7.06	7.26	7.1	41.37	44.32	41.96	44.32
4	1.55	−0.016	0.561	35.2	6.34	4.69	6.39	30.71	6.18	31.49	31.49
5	-0.05	−0.04	0.667	17	5.33	5.56	5.38	15.76	19.17	16.43	19.17
6	3.7	−0.015	0.767	51	7.7	6.26	7.75	50.85	29.51	51.52	51.52
7	3.58	−0.013	0.702	52.2	7.64	6.21	7.69	49.93	28.77	50.70	50.70
8	3.46	−0.014	0.945	68.4	8.69	5.34	8.74	65.52	15.90	66.19	66.19
9	2.83	−0.015	0.757	50.3	7.59	5.55	7.64	49.27	18.92	49.92	49.92
10	4.52	−0.016	1.202	76	9.75	6.22	9.65	81.2	28.91	79.74	81.2
11	2.14	−0.016	1.244	75.7	8.94	9.24	8.99	69.15	73.71	69.96	73.71
12	1.83	−0.021	0.928	44.6	7.13	7.37	7.18	42.46	45.93	43.16	45.93
13	1.96	−0.021	1.275	61.8	8.26	8.44	8.3	59.16	61.79	59.74	61.79
14	3.16	−0.016	1.147	70.6	8.84	6.5	8.81	67.77	33.07	67.28	nm[b]
15	3.25	−0.015	0.983	65.4	8.71	6.36	8.7	65.83	31	65.67	65.83
16	5.21	−0.026	1.462	55.3	8.15	7.6	8.19	57.52	49.37	58.11	nm
17	3.42	−0.015	0.822	56.2	7.87	6.12	7.9	53.34	27.44	53.86	nm
18	0.74	−0.018	0.378	21.5	5.97	5.18	6.01	25.23	13.47	25.77	nm
19	1.42	−0.036	0.728	20	5.63	7.27	5.68	20.18	44.44	20.95	44.44
20	2.05	−0.012	0.692	57.5	7.83	7.78	7.68	52.78	52.04	50.54	52.78
21	4.31	−0.029	1.52	51.7	8	10.83	8.03	55.25	97.24	55.7	97.24
22	5.81	−0.029	1.537	53.9	8.11	10.97	10.91	56.96	99.35	98.46	99.35
23	3.2	−0.054	2.013	37.3	6.95	8.48	7	39.75	62.41	40.42	62.41
24	1.1	−0.04	0.798	20.1	5.51	8.39	5.56	18.37	61.09	19.17	61.09
25	2.7	−0.017	1.468	85.8	9.82	10.12	9.89	82.31	86.75	83.24	86.75
26	0.59	−0.036	0.812	22.3	6.17	4.73	6.02	28.18	6.87	26.02	28.18
27	0.67	−0.01	0.409	41	6.41	1.55	6.26	31.74	-40.26	29.53	31.74
28	-0.45	−0.037	0.198	5.2	4.88	2.06	4.75	9.04	-32.73	7.19	9.04
29	0.92	−0.012	0.756	61.2	7.49	2.61	7.31	47.68	-24.57	45.04	47.68
30	2.06	−0.014	1.207	86.2	9.62	9.47	9.46	79.3	77.08	76.88	79.3
31	6.06	−0.021	1.539	74.1	9.7	5.9	9.63	80.51	24.21	79.43	nm
32	3.15	−0.011	0.6157	57	7.58	7.54	7.47	49.00	48.48	47.44	49.00
33	0.06	−0.031	0.711	23.1	5.98	5.85	5.82	25.31	23.41	22.92	25.31
34	0.49	−0.029	0.583	19.9	5.66	5.54	5.5	20.59	18.77	18.22	20.59
35	1.89	−0.017	1.339	78.7	9.25	9.1	9.1	73.83	71.62	71.61	73.83
36	1.85	−0.018	0.838	46.3	7.46	5.82	7.35	47.35	22.92	45.66	nm
37	2.85	−0.018	1.529	84	9.88	9.74	9.69	83.12	81.11	80.31	83.12
38	1.16	−0.008	0.3682	44.1	7.13	7.06	7.03	42.33	41.32	40.85	42.33
39	2.67	−0.018	1.559	86.8	9.96	9.92	9.87	84.34	83.74	82.93	84.34
40	3.5	−0.021	1.8361	86.1	10.65	10.56	10.46	94.56	93.23	91.75	94.56
41	3.7	−0.019	1.62	84.2	10.28	10.22	10.2	89.08	88.19	87.89	89.08
42	3.15	−0.016	1.394	85	10.4	10.4	10.3	90.86	90.86	89.38	90.86
43	3.9	−0.016	0.978	61.02	7.99	6.31	8.25	55.15	30.26	58.99	nm
44	5.8	−0.018	1.386	78.75	9.55	7.3	9.56	78.22	44.95	78.41	nm
45	5.5	−0.014	1.088	76.87	9.01	7.17	9.26	70.32	42.97	74.01	nm
46	7.3	−0.021	2.038	95.43	10.99	8.95	10.98	99.6	69.4	99.51	nm
47	5.9	−0.023	2.1	91.74	10.71	10.72	10.7	95.39	95.52	95.3	95.52
48	4	−0.019	1.559	81.6	9.7	9.71	9.69	80.47	80.66	80.34	80.66
49	5.5	−0.02	1.749	85.89	8.81	7.17	9.03	67.33	42.98	70.54	nm
50	3	−0.013	0.837	63.03	8.11	6.67	8.31	56.99	35.6	59.86	nm
51	6.3	−0.024	1.984	81.16	9.89	7.63	9.88	83.31	49.84	83.2	nm
52	3.9	−0.012	0.946	78.76	8.97	7.1	9.26	69.7	41.93	73.98	nm
53	6.1	−0.014	1.127	80.37	9.17	7.27	9.42	72.62	44.42	76.38	nm
54	6.3	−0.015	1.261	86.1	9.5	7.4	9.78	77.54	46.35	81.73	nm
55	7.3	−0.023	2.191	94.86	10.97	10.09	10.95	99.27	86.25	99.05	nm
56	6.5	−0.027	2.48	93.77	10.99	10.99	10.98	99.61	99.66	99.5	nm
57	5.5	−0.017	1.622	93.03	10.46	10.29	10.45	91.79	89.21	91.64	91.79
58	3.9	−0.013	1.0199	77.12	8.92	7.02	9.14	68.96	40.75	72.23	72.23
59	6.1	−0.027	2.5595	96.6	11.26	11.26	11.26	103.63	103.55	103.54	103.63
60	6.9	−0.025	2.317	92.85	10.88	10.88	10.87	97.91	98.03	97.81	nm
61	3.1	−0.017	1.295	76.14	9.23	8.86	9.21	73.55	68.02	73.19	73.55
62	2.9	−0.017	1.259	74.92	9.15	9.17	9.12	72.37	72.66	71.94	72.66
63	4.9	−0.018	1.616	89.61	10.2	10.27	10.2	87.94	88.93	87.83	88.93
64	4.9	−0.022	2.085	93.73	10.8	10.81	10.81	96.83	96.89	96.93	96.93
65	4.4	−0.017	1.407	81.45	9.58	9.58	9.6	78.71	78.65	79	79
66	4.1	−0.014	0.711	49.75	8.05	6.62	8.17	56.04	34.86	57.81	nm
67	3.6	−0.014	0.711	49.77	7.68	5.51	7.78	50.51	18.37	52.01	nm
68	2.1	−0.011	0.551	48.9	7.96	7.83	7.78	54.63	52.72	52.05	54.63
69	1.26	−0.009	0.328	35.6	7.34	7.22	7.19	45.44	43.69	43.22	45.44
70	2.37	−0.011	0.706	62.7	8.86	8.77	8.7	68.1	66.63	65.7	68.1
71	1.38	−0.009	0.4218	48.5	7.77	7.48	7.64	51.89	47.59	50.01	51.89
72	2.68	−0.012	0.833	68.9	8.74	8.69	8.67	66.29	65.55	65.16	66.29
73	1.66	−0.011	0.746	66.01	8.65	8.65	8.65	64.93	64.93	64.93	64.93
74	2.2	−0.014	1.103	78.6	9.56	6.56	9.56	78.41	78.41	78.41	78.41
75	2.73	−0.017	1.457	86	10.24	10.24	10.24	88.49	88.49	88.49	88.49
76	3.26	−0.02	1.833	91.7	10.86	10.86	10.86	97.67	97.67	97.67	97.67

[a] CHIN = CHI value of the unionized form and is taken to be the highest of the three measured values. [b] nm, a value for the uncharged molecule could not be reliably determined between pH 2.6 and 8.6 because the estimated pK_a for the compound was <4 (for acids) or >7 (for bases).

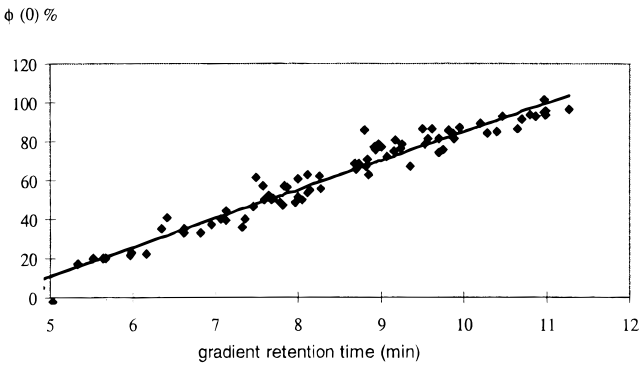


Figure 2. φ_0 values (obtained from isocratic retention measurements) as a function of gradient retention times (t_R).

on the retention time of the compounds; it can be considered as constant for each molecule. We can assume that each compound is running with the unretained peak volume when the appropriate organic-phase concentration reaches the top of the column. With these assumptions, the retention time in a fast gradient run should be linearly related to the chromatographic hydrophobicity index, φ_0.

In this paper, the application of fast generic gradient reversed-phase methods for the characterization of lipophilicity of potential drug molecules is described. A correlation study was carried out between the fast gradient retention time values and the φ_0 values obtained from isocratic runs for a large set of structurally unrelated compounds.

EXPERIMENTAL SECTION

A Hewlett-Packard 1090 series high-performance liquid chromatograph was used. Data acquisition and processing were performed on a Viglen IBM-compatible PC with HP Chemstation software (Hewlett-Packard Co., Amsterdam, The Netherlands).

The reversed-phase HPLC measurements were carried out on an ODS2-IK5 Inertsil column with the dimensions of 150 × 4.6 mm (Capital HPLC Ltd., Broxburn, Scotland). The mobile phase was 50 mM ammonium acetate (pH ranging from 7.0 to 7.3) obtained from Fisons (Loughborough, United Kingdom). For the measurements of the pH dependence of the retention times, the pH was adjusted by adding concentrated formic acid or ammonium hydroxide solutions (analytical grade, Fisons) keeping the ionic strength of buffer A constant. For the investigation of the acidic conditions, 0.2% formic acid solution (pH 2.6) was used without ammonium acetate. The mobile-phase flow rate was 1.00 mL/min. HPLC grade acetonitrile (Rathburn, Walkerburn, Scotland) was used as organic modifier. Both the gradient and isocratic mixing have been carried out by the low-pressure gradient mixer built into the HP 1090 controlled through the HP Chemstation program. For the fast gradient retention time measurements the following gradient program was applied: 0−1.5 min, 0% acetonitrile; 1.5 −10.5 min, 0−100% acetonitrile; 10.5−11.5 min, 100% acetonitrile; 11.5−12.0 min, 0% acetonitrile; 12.0−15.0 min, 0% acetonitrile. The dead time (t_0) was measured by injecting sodium nitrate together with the sample.

Theophylline, 5-phenyl-1*H*-tetrazole, benzimidazole, colchicine, 8-phenyltheophylline, indole, acetophenone, propiophenone, butyrophenone, and valerophenone were used as the standard mix for the quality control of the HPLC system that covers the log *P* range from −0.02 to 3.26. These compounds are not ionized at pH 7.4, so their distribution coefficient is equivalent to their log *P* values. These compounds were used for a reproducibility study of the gradient retention time measurements and for the quality control of the HPLC column and system. The gradient retention times of each component in the test mix were measured 14 times on two different HPLC systems and two different ODS2-IK5 Inertsil columns for the reproducibility study.

The chemical structures of the commercially available and patented compounds are shown in Figure 1. The additional compounds were synthesized at GlaxoWellcome recently. They are represented by a serial number in Table 1. The samples were dissolved at ∼1 mg/mL in a mixture of 50% acetonitrile and 50% aqueous ammonium acetate buffer. The average retention time (t_R) of two consecutive injections of 5 μL of sample was used to calculate the log k' values (log $k' = \log((t_R - t_0)/t_0)$). Then the log k' values were plotted against the applied acetonitrile concentration. The pH of the mobile phase was 7, and 50 mM ammonium acetate was mixed with the acetonitrile. The slope (S) and the intercept (log k'_w) values were calculated based on log k' values obtained by a minimum of five organic-phase concentrations. The correlation coefficients of the linear fit were always higher than 0.99. The isocratic hydrophobicity index (φ_0) was calculated as $-\log k'_w/S$, and it refers to the hydrophobicity of the compound at pH ∼7.

The octanol/water partition coefficients (*c* log *P*) of the compounds were calculated by using the MedChem database (Pomona College). The multiple regression analysis was carried out using the Drugidea software package (Chemicro Ltd., Budapest, Hungary).

RESULTS AND DISCUSSION

The 76 compounds included in this study were partially commercially available compounds and partially confidential research compounds. The set represents a wide range of chemical structures and lipophilicity ($-0.45 < c \log P < 7.3$). The isocratic retention times were measured by using various volume percents of acetonitrile in the mobile phase, preferably bracketing the retention when log $k' = 0$ (that is the retention time was close to the double of the dead time). The isocratic hydrophobicity index, φ_0, was calculated from the slope (S) and the intercept (log k'_w) values of the straight lines obtained by plotting log k' vs φ ($\varphi_0 =$ −intercept/slope), based on minimum three points and $r > 0.99$. The gradient retention time (t_R) was measured under the gradient profile described in the Experimental Section, which included a 9 min linear acetonitrile gradient from 0 to 100%.

Table 1 shows the calculated c log *P* values and the measured *S*, log k'_w, φ_o values as well as the gradient retention time values obtained at pH 7.3, 2.6 and 8.3.

Figure 2 shows the plot of the isocratic φ_0 (pH 7) values and the gradient retention time values (pH 7.3). The relationship is described by eq 1.

$$\varphi_0 = -\text{intercept}/\text{slope} = 14.34(\pm 0.39)t_R - 58.72(\pm 3.30) \quad (1)$$

$$n = 76 \qquad r = 0.974 \qquad s = 5.3 \qquad F = 1371$$

It can be seen that an excellent correlation was found and it provides experimental confirmation of the hypothesis that a linear gradient retention time can be used as a measure of compound hydrophobicity. The constants in eq 1 can be used to convert

Table 2. Chromatographic Data Obtained for the Compounds in the Calibration Mixture Including the Standard Deviation Obtained from 2 × 7 Measurements with Two HPLC Systems[a]

compound	φ_0	S	log k'_w	t_R	CHI
theophylline (5)	17.0	−0.0340	0.667	5.33 ± 0.08	15.76 ± 0.8
phenyltetrazole (19)	20.0	−0.0364	0.728	5.63 ± 0.05	20.18 ± 0.7
benzimidazole (4)	35.2	−0.0160	0.561	6.34 ± 0.01	30.71 ± 0.4
colchicine (3)	40.0	−0.0181	1.559	7.06 ± 0.01	41.37 ± 0.4
phenyltheophylline (20)	57.5	−0.0120	0.692	7.83 ± 0.01	52.04 ± 0.4
acetophenone (73)	66.0	−0.0137	0.976	8.65 ± 0.01	64.90 ± 0.2
indole (11)	75.7	−0.0164	1.244	8.94 ± 0.01	69.15 ± 0.2
propiophenone (74)	78.6	−0.0157	1.256	9.56 ± 0.01	78.41 ± 0.1
butyrophenone (75)	86.0	−0.0170	1.475	10.24 ± 0.01	88.49 ± 0.1
valerophenone (76)	91.2	−0.0186	1.700	10.86 ± 0.01	97.67 ± 0.3

[a] The CHI values were calculated by the constants given in eq 1 (derived from 76 compounds).

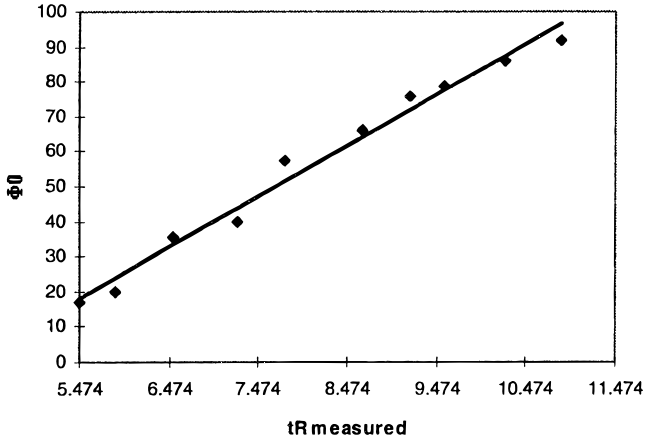


Figure 3. Gradient retention times against the isocratic φ_0 values for the test compounds: $\varphi_0 = 14.00(\pm 0.64)t_R - 55.9$ ($n = 10$, $r = 0.992$).

Table 3. Correlation Coefficients between the Individual Parameters Based on the Data from Table 1

	c log P	slope	log k'_w	φ_0	$CHI_{7.3}$	$CHI_{2.6}$	$CHI_{8.5}$
c log P	1.00	0.06	0.75	0.73	0.74	0.48	0.78
slope		1.00	−0.30	0.36	0.29	−0.07	0.26
log k'_w			1.00	0.74	0.79	0.72	0.79
φ_0				1.00	0.97	0.66	0.95
$CHI_{7.3}$					1.00	0.72	0.98
$CHI_{2.6}$						1.00	0.74
$CHI_{8.5}$							1.00

the gradient retention time values to a chromatographic hydrophobicity index parameter (CHI), and these are also shown in Table 1. This approach puts CHI and φ_0 on the same scale so that the CHI value for a compound approximates to the percentage (by volume) of acetonitrile required to achieve an equal distribution of compound between the mobile and the stationary phases. Several gradient profiles have been tested for a smaller subset of compounds, and it was found that a slightly better correlation coefficient between the gradient retention times and φ_0 values was obtained when the gradient speed was increased.

In order to make the method easier to apply, a representative calibration set of 10 compounds was chosen. Table 2 shows the gradient retention times, φ_0, S, log k'_w, and CHI values for the selected set which can be used to calibrate any reversed-phase column running a linear gradient profile on any HPLC instrument. The reproducibility of the gradient retention times was very good, based on 14 measurements on two different HPLC systems on two different days and columns. On the basis of these results, we expect CHI values to be repeatable within ±2 units.

Figure 3 shows the correlation of the isocratic φ_0 values with the gradient retention times for the calibration set of compounds. The correlation is described by eq 2. The constants of eqs 1 and 2 are very similar, which means that the 10 compounds in the test mixture are representative of the whole set (76 structurally unrelated compounds).

$$\varphi_0 = 14.00(\pm 0.64)t_R - 55.88(\pm 5.3) \qquad (2)$$

$$n = 10 \qquad r = 0.992$$

The following simplified procedure is suggested for setting up the CHI method on a new column or new instrument or with a different mobile phase. First measure the linear gradient retention time for the 10 compounds listed in Table 2 on a reversed-phase system. The linear range of the gradient from 0 to 100% organic phase could be between 2 and 10 min, depending on the dimensions of the column and the flow rate. Then perform a linear regression between the measured retention times and the CHI values of the test compounds (see Table 2) to give the coefficients A and B of eq 3. Keeping the conditions the same as those used for the test mixture, the CHI value of compound X can be calculated as $CHI_x = At_R + B$. In our experience, a column can be used for several days without the calibration changing significantly. A mixture containing all 10 of the calibration set should be injected at regular intervals, and recalibration was performed if the retention time differences of the standard compounds in our case was greater than 0.1 min.

$$CHI = At_R + B \qquad (3)$$

CHI is a high-throughput chromatographic hydrophobicity index, and the exact values obtained for a compound will depend on the type of stationary phase, type of organic phase (acetonitrile or methanol), and, for acidic or basic compounds, the pH. Any reversed-phase HPLC system can be calibrated by using the data from the test mixture. The given CHI values (see Table 2) of the calibration compounds can be correlated with the actual retention times of these compounds in any other system, and the constants can be used to convert any fast gradient retention times to CHI values. If a column other than ODS is used, then appropriate isocratic φ_0 values will need to be redetermined for the calibration compounds in order to align the CHI and φ_0 scales as closely as possible.

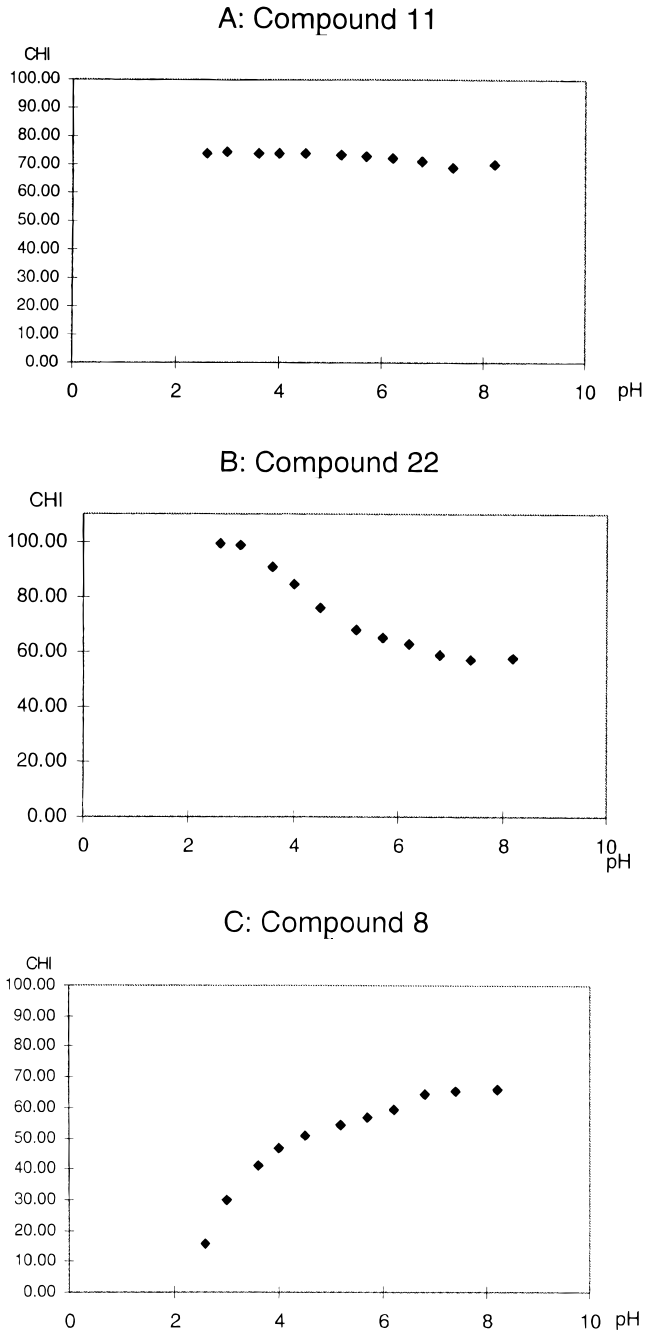


Figure 4. CHI values as a function of pH: (A) neutral, (B) acid, and (C) base.

Table 3 shows the correlation coefficients between the investigated parameters for 76 compounds in a correlation matrix. It was previously shown that φ_0 values are in a good correlation to the calculated log *P* values for several hundreds of structurally unrelated compounds.[25] It can be seen that the c log *P* values did not show high correlation with the φ_0 values and the CHI values in our case. The highest correlation coefficient is 0.78 with the CHI values obtained at pH 8.5. The explanation is that the *c* log *P* values refer to the partition of the neutral (uncharged species). At the applied pH, most of the basic and acidic compounds were fully or partially ionized, so the retention data reflect the distribution of both species. It is difficult to define precisely the pH of a mixed solvent that contains a high proportion of organic component (in our case acetonitrile), also it is difficult (practically impossible) to maintain constant ionic strength during a gradient run. Most buffers, other than ammonium acetate, cause high pressure and blockage on the column due to their precipitation at high organic solvent concentration. Only the pH of buffer A can be accurately measured and adjusted. For a subset of the compounds investigated, the CHI values were determined after stepwise variation of the pH in buffer A. Figure 4 shows typical plots for neutral (A), acidic (B), and basic (C) compounds. These plots are similar to those obtained from partitioning experiments, i.e., log *D*/pH profiles, in that an approximately constant lipophilicity difference can be observed between the charged and uncharged molecule.[34] As the operational pH range is limited to between 2 and 8 on a silica-based reversed-phase column, the CHIN (i.e., CHI value of the uncharged form of the molecule) could be measured only for a subset of the investigated molecules (i.e., neutral, acid, and weak bases). The correlation of the CHIN values to the *c* log *P* values can be described by eq 4.

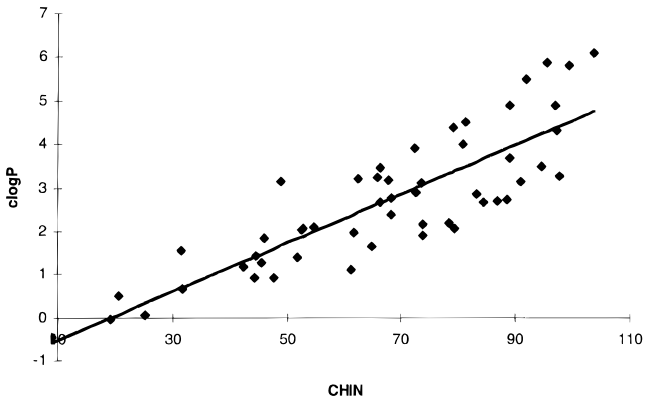


Figure 5. Relationship between the *c* log *P* values and the CHIN values (chromatographic hydrophobicity index of the neutral molecules).

$$c\log P = 0.0566(\pm 0.005)\mathrm{CHIN} - 1.107 \qquad (4)$$

$$n = 52 \qquad r = 0.851 \qquad s = 0.82 \qquad F = 131$$

The plot of *c* log *P* as a function of CHIN can be seen in Figure 5.

The *S* and log k'_w values did not show correlation to each other ($r = -0.296$) indicating that a structurally diverse set is present.[35]

Excellent correlation was found between the CHI values and the slope (S) and intercept (log k'_w) values, as described by eq 5.

$$\mathrm{CHI}_{7.3} = 1607.0(\pm 97.8)S + 41.1(\pm 1.5)\log k'_w + 46.8 \qquad (5)$$

$$n = 76 \qquad r = 0.959 \qquad s = 6.7 \qquad F = 412.8$$

It shows that the same principles are valid in the isocratic and the gradient methods. The correlation coefficient between $\mathrm{CHI}_{7.3}$ and log k'_w values was only 0.79.

Equation 5 supports the concept that the parameter CHI can be used as an alternative to log k' as a measure of lipophilicity of compounds.

CONCLUSIONS

A high-throughput fast gradient reversed-phase method has been introduced to characterize the partition behavior of a large

(34) Hersey, A.; Hill, A.; Hyde, R. M.; Livingstone, D. J. *Quant. Struct.—Act. Relat.* **1989**, *8*, 288−96.
(35) Valkó, K. *J. Liq. Chromatogr.* **1987**, *10*, 1663−86.

number of compounds in an automated experiment. The defined CHI can be obtained from the gradient retention time after calibrating the system with a test mixture. The CHI values show good correlation to the φ_0 values obtained from several isocratic runs. For most compounds, CHI is between 0 and 100 and in this range it approximates the percentage (by volume) of acetonitrile required to achieve an equal distribution of compound between the mobile and the stationary phases.

From the gradient retention time obtained by incorporating acidic and basic buffers in solvent A, the acid or base character of the compounds can be revealed. The CHI (un-ionized) value showed good correlation to the *c* log *P* values for a structurally unrelated set of compounds.

As shown previously for φ_0 values, the CHI values show good correlation to the *S* and log k'_w values (see eq 5).

The method is recommended as part of a protocol for high-throughput physicochemical property profiling for rational drug design.

DEFINITION OF SYMBOLS

log k'	retention factor, derived from isocratic retention time (t_R) and dead time (t_0) as log $k' =$ log $((t_R - t_0)/t_0)$
φ	volume percent of organic concentration in the mobile phase
φ_0	isocratic chromatographic hydrophobicity index, which is defined by the organic-phase concentration (volume percent) at which the log $k' = 0$; that is, the retention time is double of the dead time. Its values depend on the pH and the type of organic solvent
CHI	chromatographic hydrophobicity index, obtained from gradient retention times after calibrating the gradient system with test compounds. It is equivalent to the volume percent of organic phase concentration by which the compound is eluting from the column. Its value depends on pH and the type of organic solvent
t_R	retention time
S	the slope of the straight line obtained by plotting the isocratically determined log k' values in the function of the organic phase
log k'_w	intercept of the straight line obtained by plotting the isocratically determined log k' values as a function of organic-phase concentration. Its value depends on the type of organic phase and the range of organic-phase concentration from which it has been extrapolated
n	number of compounds the data of which was included in the regression equation
s	standard error of the dependent variable estimates
F	Fisher test value
r	correlation coefficient

Received for review December 6, 1996. Accepted March 21, 1997.[⊗]

AC961242D

[⊗] Abstract published in *Advance ACS Abstracts,* May 1, 1997.