Journal of Chromatography, 506 (1990) 119–131
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMSYMP. 1760

Partition coefficients and capacity factors of some nucleoside analogues

ANDREW P. CHEUNG* and DEBORAH KENNEY
Life Sciences Division, SRI International, Menlo Park, CA 94025 (U.S.A.)

SUMMARY

The AIDS epidemic has brought into focus the development of antiviral agents, of which nucleoside analogues are an important class. The single most important physico-chemical property of a chemotherapeutic agent is its hydrophobicity. This paper reports the hydrophobicity,determined as log partition coefficient (P) by the shake-flask method, of 25 nucleoside analogues. The capacity factors (k') of these analogues were also obtained by reversed-phase liquid chromatography. There is a very strong linear correlation between the log P and the log k' values of all the nucleoside analogues, indicating that hydrophobicity of nucleosides can be determined by liquid chromatography.

Examination of the P data indicates that an empirical mathematical relationship exists between the partition coefficient value and the molecular structure of the nucleoside analogues. A table of constants and an equation is proposed to estimate the P of nucleoside analogues.

INTRODUCTION

The AIDS (Acquired-Immune-Deficiency Syndrome) epidemic has brought into focus the development of antiviral agents. Analogues of nucleosides represent an important class of antiviral agents. For a chemotherapeutic agent to be effective, among other properties, it has to be able to be adsorbed and distributed to the target organ or tissue. This ability is dictated by the hydrophobicity of the drug. Hydrophobicity is one of the most important physico-chemical properties affecting the drug's biological activity. It is commonly expressed as the logarithm of the partition coefficient (log p) of the chemical between 1-octanol and water. This property is usually determined by the traditional shake-flask method.

The shake-flask method has many disadvantages. It is laborious and requires that pure chemicals be used. As alternatives to this shake-flask method, several investigators [1–3] have proposed the use of chromatography to determine the hydrophobicity of a chemical. Particularly, the logarithm of the capacity factor (log k') obtained from reversed-phase liquid chromatography (LC) has been shown to have good correlation to the log P for several classes of chemicals[4,5]. The LC method

0021-9673/90/$03.50

offers many advantages over the traditional shake-flask method. Among them, LC requires only a small amount of the chemical. Presence of impurities in the chemical does not interfere with the determination. Proper measurement of concentration is not required. LC gives better results when the solubility of the chemical in one phase is much lower than in the other phase. According to Henry *et al.*[6] and Baker *et al.*[7], chromatographic methods are more similar to the *in vivo* process than the shake-flask method.

Determination of hydrophobicity using LC is not free of limitations. The log k' values of the chemical and several model compounds are usually determined by LC with a single mobile phase—water or buffer with organic modifier. The log P value of the chemical is then extrapolated from linear regression analysis of the Collander equation[8] for the model compounds.

$$\log k' = a \log p + b \tag{1}$$

This monocratic approach requires model compounds of known log P. The major weakness of the method is that the Collander relationship has been shown valid for congeners only. Haky and Young[9] determined log k' for 68 compounds using a commercial ODS column and 55% methanol in aqueous buffer. They found good linear correlation ($r=0.966$) according to eqn. 1 between log P and log k' values. They found the correlation improved if they separated the compounds into congener classes. The hydrogen bonding of the residual silanol in the ODS column to different classes of compound interferes with the partition process differently. As a result, slopes (a) and intercepts (b) of eqn. 1 for different classes of compounds may be different. Miyake *et al.*[10] and Biasi and Lough[11] used polymer-based reversed-phase columns and obtained similar results—linear correlation between log P and log k' exists only among congeners. Several investigators[1,12–16] bypassed the congener restriction with a polycratic approach to determine hydrophobicity by LC. In this approach, the log k'_x of the chemical is determined with several mobile phases containing different % (x) methanol in water. They found that the log k'_0, the log k' when the mobile phase is 100% aqueous, follows the relationship of eqn. 2, if x is between 20 and 80.

$$\log k'_x = \log k'_0 - \mathrm{S}\, x \tag{2}$$

Garst and Wilson[13,14] and Minick *et al.*[15] reported that the log k'_0 thus obtained correlates linearly with the binding activity as well as log P. This polycratic approach offers several advantages over the monocratic one. In the monocratic approach, the log k' of the chemical and model compounds have to be simultaneously determined under a single LC condition. This is not so required for the polycratic approach. In addition, if log k'_0 is accepted as an independent hydrophobicity scale, even model compounds are not needed.

Because of the current interest in anti-AIDS agents, many derivatives of nucleosides have been synthesized as potential antiviral agents. The hydrophobicity of these chemicals would be of interest. The purpose of this paper is to report log P of 25 nucleosides and analogues determined by the shake-flask method and their log k' obtained by the monocratic LC approach. A discussion of the partition coefficients,

capacity factors and the correlation between the two is presented, and an empirical equation and a table of constants to calculate *P* of nucleoside analogues is proposed. The names and structures of these 25 chemicals are listed below and depicted in Fig. 1.

Compound	*Chemical name*	*Abbreviation*
1	Cytidine	C
2	2'-Deoxycytidine	DC
3	2',3'-Dideoxycytidine	DDC
4	2'.3'-Dideoxycytidine-2'-ene	DDCene
5	Cytarabine	AraC
6	5-Fluoro-2',3'-dideoxycytidine	5FDDC
7	5-Bromo-2',3'-dideoxycytidine	5BDDC
8	2',3'-Dideoxy-3'-azidocytidine	AZC
9	Carbocytidine	CarboC
10	Uridine	U
11	2',3'-Dideoxy-3'-azidouridine	AZU
12	Thymidine	T
13	3'-Deoxythymidine	DDT
14	3'-Deoxythymidine-2'-ene	DDTene
15	3'-Deoxy-3'-azidothymidine	AZT
16	Adenosine	A
17	2'-Deoxyadenosine	DA
18	2',3'-Dideoxyadenosine	DDA
19	2',3'-Dideoxyadenosine-2'-ene	DDAene
20	2'-Fluoro-2',3'-dideoxyadenosine	2'FDDA
21	Adeninearabinoside	AraA
22	Neplanosine	CarboA
23	Inosine	I
24	2',3'-Dideoxyinosine	DDI
25	2',3'-Dideoxyguanosine	DDG

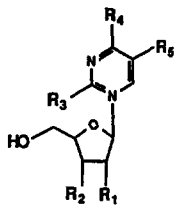
EXPERIMENTAL

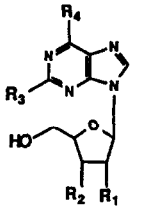
Reagents and materials

The nucleosides were received from the National Cancer Institute, National Institutes of Health. Their identities were confirmed by UV and mass spectra. The purities were checked by UV and LC. Due to their high purity, >98%, they were used without further purification. Buffers were prepared from KH_2PO_4 (Mallinckrodt, analytical-reagent grade) with Milli-Q quality water and the pH adjusted with dilute KOH or H_3PO_4. 1-Octanol (Aldrich) was HPLC grade. The buffer and 1-octanol were presaturated with each other before use.

Determination of partition coefficients (P)

Accurately weighed portions (0.2, 0.5, 1 and 1.5 mg) of each nucleoside were individually dissolved in 10.0 ml of pH 7.0, 0.05 *M* phosphate buffer in volumetric flasks. The UV of each solution was recorded from 300 to 200 nm with a Uvikon 810 UV–VIS spectrophotometer (Kontron, Switzerland). The absorbances of these solutions at the absorption maximum were plotted against their concentrations. If linear relationship was not obtained from the plot, the experiment was repeated with pro-





	R1	R2	R3	R4	R5
1.	OH	OH	OH	NH_2	H
2.	H	OH	OH	NH_2	H
3.	H	H	OH	NH_2	H
6.	H	H	OH	NH_2	F
7.	H	H	OH	NH_2	B_r
8.	H	N_3	OH	NH_2	H
10.	OH	OH	OH	OH	H
11.	H	N_3	OH	OH	H
12.	H	OH	OH	OH	CH_3
13.	H	H	OH	OH	CH_3
15.	H	N_3	OH	OH	CH_3

	R1	R2	R3	R4
16.	OH	OH	H	NH_2
17.	H	OH	H	NH_2
18.	H	H	H	NH_2
20.	F	H	H	NH_2
23.	OH	OH	H	OH
24.	H	H	H	OH
25.	H	H	NH_2	OH

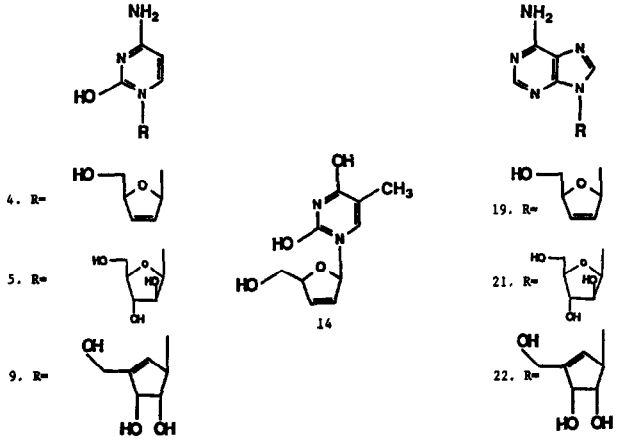


Fig. 1. Structural formula of the nucleoside analogues.

gressively more dilute solutions. This ensured that the molar absorptivity at the maximum (ε_{max}) was obtained from complete solutions and that the UV follows Beer's Law. The ε_{max} of each nucleoside at about 260 nm was calculated. Similarly, the ε_{max} of the nucleoside in 1-octanol was determined. Because some of the nucleosides have very low solubility in 1-octanol, they were first dissolved in 1 ml ethanol before diluting to mark with octanol.

From the freshly prepared buffer solutions which gave linear response of absorbance *vs.* concentration, 5.00 ml each were transferred to individual centrifuge cones (glass equipped with stopper). To each cone were added 5.00 ml octanol and the cone was stoppered. The mixtures were shaken 80 times followed by centrifugation at 1000 *g* for 1–2 h. The aqueous and the organic phases were separated and the UV of

each recorded. The nucleoside concentration in each phase was calculated from its respective absorbance. The *P* value of the nucleoside was determined by the following expression:

$$P = [c]_o/[c]_a$$

where $[c]$ = nucleoside concentration, o = octanol phase and a = aqueous phase.

Determination of capacity factor (k')

The *k'* values were determined isocratically (3% acetonitrile in pH 7.0, 0.05 *M* phosphate buffer, 1.0 ml/min) on a Chemcosorb-5-C_{18}-H bonded phase, 250 × 4.6 mm I.D. column (DyChrom, Sunnyvale, CA, U.S.A.) using a Model 600 solvent delivery system (Waters Assoc., Milford, MA, U.S.A.), a Model 712WISP Autosampler (Waters) and a Model 481 LC spectrophotometer (Waters). A Maxima 820 integration system (Dynamic Solution, Ventura, CA, U.S.A.) was used to collect chromatographic and retention data.

The *k'* value of each nucleoside was calculated as $(t_i - t_0)/t_0$, where t_i is the retention time of the nucleoside and t_0 is the retention time of an unretained compound. The unretained peak was determined by an injection of methanol. The *t* values of the nucleosides (except compounds, **10, 12, 21** and **25**) were obtained from one chromatogram. The *t* values for compounds **10, 12, 21** and **25** were extrapolated from a second chromatogram.

RESULTS AND DISCUSSION

Nucleosides are ionizable in aqueous solutions. In order to determine the hydrophobicity of their neutral species the experimental condition should suppress the ionization of the nucleosides to the minimum. The 25 chemicals studies are analogues of four classes of nucleosides: cytidines (**1–9**), uridines (**10–15**), adenosines (**16–22**), and inosines (**23–25**). They contain the pyrimidine and purine bases as their ionizable moiety. Their ionization constants and charge sites are depicted in Fig. 2. The basic

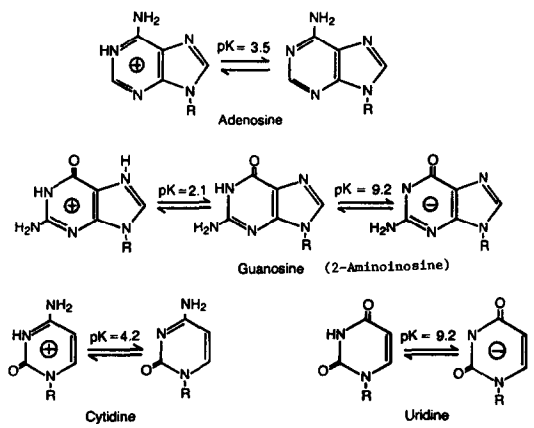


Fig. 2. Sites and p*K* values for protonation and ionization of representative nucleosides.

TABLE I

P VALUES OF DDA AND DDC BETWEEN 1-OCTANOL AND PHOSPHATE BUFFERS

The buffers were prepared from 0.05 *M* KH_2PO_4. The pH was adjusted with dilute KOH or H_3PO_4. *P* values were determined by the traditional shake-flask method.

pH of buffer	*P of DDA*	*P of DDC*
5	0.57	0.043
6	0.58	0.050
7	0.60	0.054
8	0.60	0.051

pK_a values are 4 or lower and the acidic pK_a values are about 9, therefore, a pH 6–7 buffer solution would keep the nucleosides in their unionized forms. Table I presents the *P* values of DDC (**3**) and DDA (**18**) determined between 1-octanol and pH 5, 6, 7 and 8 buffers. For DDC which has both basic and acidic pK_a values, *P* decreases on both sides of pH 7 due to ionization of the basic or acidic function. For DDA which has only the basic function, *P* decreases at pH less than 7. Based on this result, *P* values of the rest of the chemicals were determined with the 1-octanol–pH 7.0 phosphate buffer system. Several investigators[1,12–15] emphasized the advantages of extrapolating log k'_0 for each chemical with the polycratic solvent approach. They also pointed out that the extrapolation is valid only when the organic modifier (methanol) in the mobile phase is within a limited range (20 to 80%). For their model compounds with large (0–8) log *P* values, the polycratic solvent approach is applicable. For nucleosides, the log *P* values of which are often small (<0), the amount of organic modifier in the mobile phase is usually less than 10%. Because of this limitation, the log k' values of the nucleosides were obtained monocratically, using a 3% acetonitrile in pH 7.0 phosphate buffer. The buffer in the mobile phase ensured that the ionization of the nucleosides were suppressed to a minimum. The organic modifier in the mobile phase was kept to a deliberately small amount (3% acetonitrile) in order to keep the LC system as similar as possible to the octanol-buffer partition system. Table II lists the determined *P* and k' values and their log values of these nucleosides and analogues.

In order to see if the nucleosides' hydrophobicity can be predicted from the k' values obtained with LC, the linear plots and linear regression analyses between the log *P* and log k' values in Table II were performed. Results of the linear regression analysis are summarized in Table III. The linear plots are presented in Figs. 3–5. Except for the inosines (**23–25**), the log *P vs.* log k' plot of the remaining 22 nucleoside gives a straight line joining **9** and **15** (Fig. 3). The linear correlation coefficient (r) for these 22 compounds is 0.991. Fig. 4 is a log *P vs.* log k' plot of the 25 compounds after they are separated into the pyrimidine and purine nucleoside series. Both the pyrimidines (**1–15**) and the purines (**16–25**) give straight lines with r of 0.994 and 0.958, respectively. Fig. 5 is the log *P vs.* log k' plot of the nucleosides after they are further separated as cytidines (**1–9**), uridines (**10–15**), adenosines (**16–22**), and inosines (**23–25**). The r of the cytidines, the uridines, the adenosines and the inosines are respectively 0.991, 0.996, 0.981 and 1.000. Even when all twenty 25 nucleosides are considered as a whole, r between log *P* and log k' of these 25 compounds is still a respectable

TABLE II

PARTITION COEFFICIENTS AND CAPACITY FACTORS OF NUCLEOSIDE ANALOGUES

P Values were determined by the traditional shake-flask method; *k'* values were determined by LC (see text for details).

Compound	*P* ± *S.D.*	*log P*	*k'*	log *k'*	*n*
1	0.008 ± 0.001	−2.097	0.79	−0.102	4
2	0.013 ± 0.001	−1.886	1.87	0.272	3
3	0.055 ± 0.003	−1.260	4.49	0.652	4
4	0.028 ± 0.002	−1.553	2.53	0.402	4
5	0.008 ± 0.001	−2.097	1.28	0.107	4
6	0.090 ± 0.003	−1.046	6.47	0.811	3
7	0.320 ± 0.007	−0.495	20.49	1.312	3
8	0.208 ± 0.008	−0.682	16.14	1.208	4
9	0.005 ± 0.001	−2.301	0.73	−0.136	3
10	0.013 ± 0.001	−1.886	2.02	0.306	4
11	0.465 ± 0.006	−0.333	28.81	1.460	3
12	0.079 ± 0.002	−1.102	7.68	0.885	4
13	0.264 ± 0.012	−0.578	20.82	1.303	7
14	0.193 ± 0.011	−0.714	13.83	1.141	4
15	1.091 ± 0.002	0.038	73.37	1.866	3
16	0.105 ± 0.003	−0.979	13.33	1.125	4
17	0.245 ± 0.002	−0.611	16.14	1.208	4
18	0.602 ± 0.017	−0.220	38.90	1.590	5
19	0.311 ± 0.008	−0.507	24.22	1.384	3
20	0.801 ± 0.002	−0.096	42.31	1.626	4
21	0.111 ± 0.004	−0.955	9.63	0.984	3
22	0.050 ± 0.003	−1.301	6.47	0.811	4
23	0.010 ± 0.001	−2.000	3.60	0.556	3
24	0.068 ± 0.005	−1.167	13.33	1.125	4
25	0.085 ± 0.002	−1.071	15.24	1.183	4

TABLE III

LINEAR REGRESSION ANALYSIS OF LOG *P VS.* LOG *k'*

log *P* = *S* log *k'* + *I*; *S* = slope; *I* = intercept; S.D. = standard deviation; *r* = correlation coefficient.

Data base[a]	*S* ± *S.D.*	*I* ± *S.D.*	*r*	*n*
Compounds **1–25** (all)	1.203 ± 0.064	−2.187 ± 0.170	0.969	25
Compounds **1–15** (pyrimidines)	1.198 ± 0.038	−2.116 ± 0.086	0.994	15
Compounds **16–25** (purines)	1.631 ± 0.172	−2.782 ± 0.170	0.958	10
Compounds **1–9** (cytidines)	1.224 ± 0.062	−2.106 ± 0.092	0.991	9
Compounds **10–15** (uridines)	1.254 ± 0.057	−2.217 ± 0.067	0.996	6
Compounds **16–22** (adenosines)	1.401 ± 0.124	−2.414 ± 0.093	0.981	7
Compounds **23–25** (inosines)	1.476 ± 0.016	−1.822 ± 0.008	1.000	3
Compounds **1–22** (all except inosines)	1.199 ± 0.035	−2.132 ± 0.093	0.991	22

[a] From Table II.

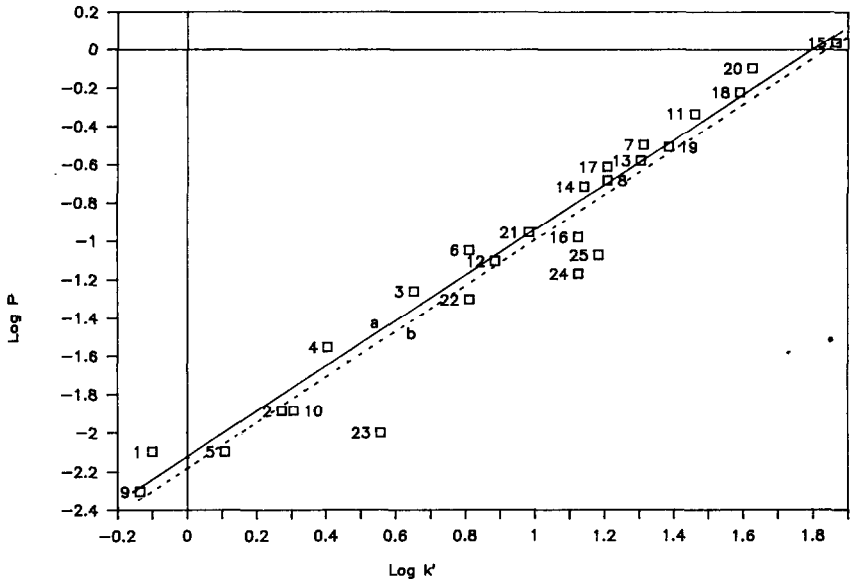


Fig. 3. Linear plot of log P *vs.* log k' (pH 7) of all analogues. The solid line is for compounds **1–22** (all nucleosides minus the inosines). It fits the linear equation of log $P = 1.199 \log k' - 2.132$. The dotted line is for all 25 compounds. It fits the linear equation of log $P = 1.203 \log k' - 2.187$.

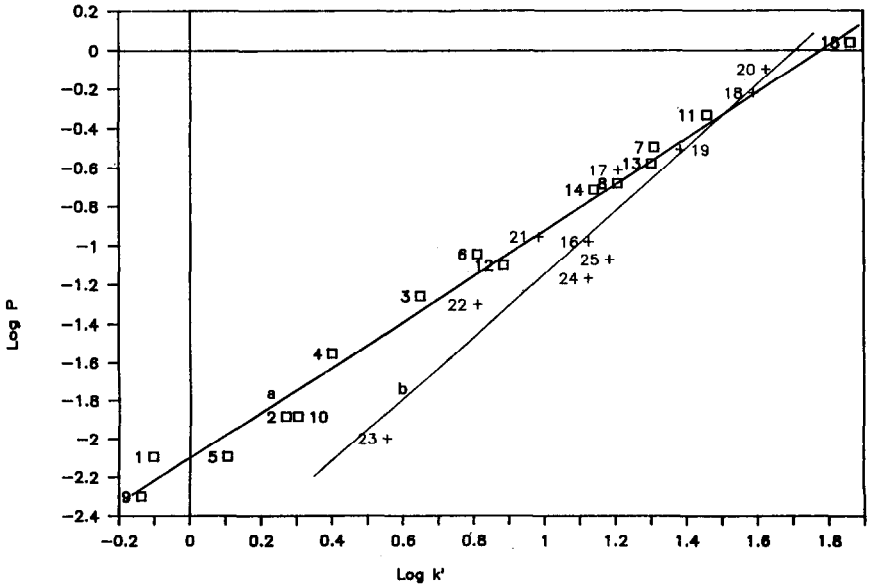


Fig. 4. Linear plot of log P *vs.* log k' (pH 7) of the analogues according to pyrimidine or purine derivatives. The pyrimidine (**1–15**) line (□, a) fits the equation of log $P = 1.198 \log k' - 2.116$. The purine (**16–25**) line (+, b) is represented by the equation of log $P = 1.631 \log k' - 2.782$.

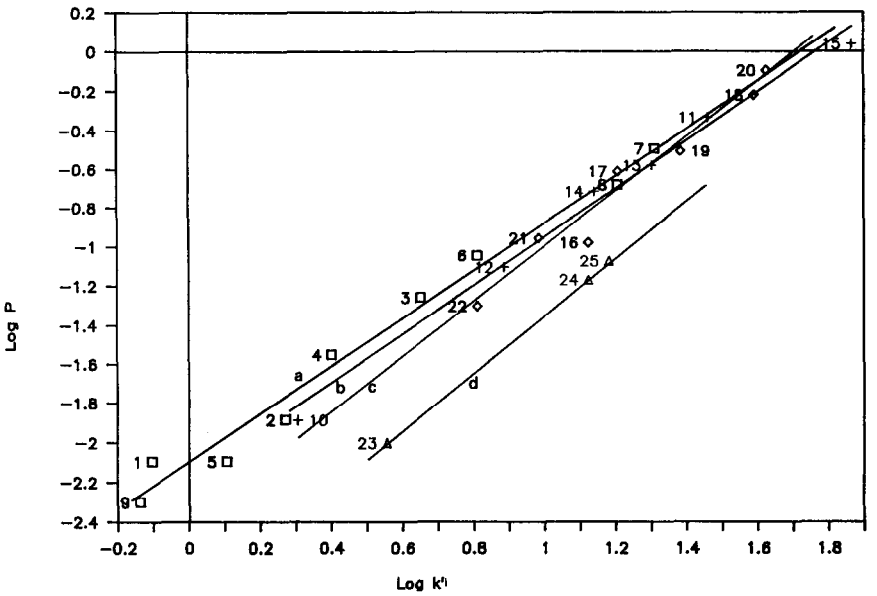


Fig. 5. Linear plot of log *P vs.* log *k'* (at pH 7) of the analogues according to cytidine, uridine, adenosine and inosine classes. Line a (□) for the cytidines (**1–9**) has an equation of log $P = 1.224 \log k' - 2.106$. Line b (+) for the uridines (**10–15**) has an equation of log $P = 1.254 \log k' - 2.217$. Line c (◇) for the adenosines (**16–22**) has an equation of log $P = 1.401 \log k' - 2.414$. Line d (△) for the inosines (**23–25**) has an equation of log $P = 1.476 \log k' - 2.822$.

0.969. These results are much better than those recently published by Balzarini *et al.*[17] These data indicate[17] that the linear correlation between log *P* and retention time (t_R) is good ($r \geqslant 0.96$) only among nucleosides having the identical base. When all of the fifteen purine derivatives are considered as one class, *r* is a fair 0.956. But when all fifteen pyrimidine derivatives are considered together, *r* is a disappointing 0.266[a]. *R* is only 0.713 when all thirty nucleoside derivatives are considered together[b]. Several factors may have contributed to this lack of linear correlation between the log *P* and t_R data in the paper by Balzarini *et al.*[17]. (a) The log *P* data in ref. 17 were obtained with an octanol–pH 7.5 buffer system. Under this condition the ionizaton of both the basic ($-NH_2$, $pK_a \leqslant 4$) and the acidic moiety (phenol, $pK_a \geqslant 9$) of the nucleosides were suppressed. The t_R data, on the other hand, were obtained from a pH 3.2 buffer environment. In this case, the $-NH_2$ group in the nucleoside bases would be partially ionized. For the $-NH_2$-containing nucleosides (cytidines, adenosines, guanosines and diaminopurines), the log *P* values were obtained from neutral species, but the t_R values were obtained from partially ionized species. For nucleosides without an $-NH_2$ group (uridines and thymidines), both the log *P* and t_R values were derived from neutral species. Therefore, the log *P* and t_R obtained by Balzarini *et al.*[17] would not correlate linearly among all nucleosides. (b) In ref. 17, the t_R values were obtained with a gradient mobile phase in which the acetonitrile concentration increased from 4

[a] Calculated from values in Table 1 in ref. 17.
[b] Calculated from values in Tables 1 and 2 in ref. 17.

to 25%, non-linearly. For compounds of different polarity, the effect of changing mobile phase on t_R would be quite different. Minick and co-workers[15,16] observed that the effect on LC retention time due to changes in acetonitrile concentration in the mobile phase was independent of chemical structures. Thus, the use of gradient mobile phase in LC may have contributed to the poor linear correlation of Balzarini *et al.*'s data between the log P and t_R of the entire nucleoside derivatives.

It is evident from our data presented in Table III and Figs. 3–5 that there is a strong linear correlation between log P and log k' values of diversified nucleoside analogues and that hydrophobicity of nucleoside analogues can be estimated by their log k' values. The LC system, however, should closely resemble the octanol–aqueous partition system and the analytes should be kept in the neutral form.

Examination of Figs. 4 and 5 reveals that the slopes for the cytidines and the uridines are identical ($S = 1.23$) but are smaller than those of the adenosines and inosines ($S = 1.45$). In LC, where the C_{18} stationary phase is coated on a solid support, the size of the nucleoside base rings probably contributed more to the nucleosides' affinity to the lipophilic phase than is in the case of the shake-flask method. Since the pyrimidine ring is smaller than the purine ring, similar structural modification in the pyrimidines will have larger impact than the purines on their k' values. In the shake-flask method P is less affected by the molecular size than by the polarity of the compound, similar structural change would have similar impact on the P value of both the purine and the pyrimidine analogues. It appears that different impact of the molecular size on the k' and P of the nucleosides resulted in different slopes for the log k' *vs.* log P plot.

Table II shows that the P values of the uridines (**10**, **11**), thymidines (**12–15**), adenosines (**16–19**, **21**, **22**) and inosines (**23** and **24**) are, respectively, about 2, 4, 10 and 1.2 times those of the corresponding cytidines (**1–5**, **8**, **9**). Substituting the 5-H of the pyrimidine base with F, CH_3 or Br increases its hydrophobicity by 1.6, 2 or 5 times, respectively. This substitution effect on P is in line with that observed by Garst[14]. Variations in the sugar moiety also have consistent effects on the hydrophobicity of the chemicals. The effect appears to be independent of the base moiety. Comparing the P values with those of the corresponding ribosides (**1**, **10**, **12**, **16**, **23**), the P values of the arabinosides (**5**, **21**) are identical. The 2′-deoxyribosides (**2**, **17**) are 2 times and the 2′,3′-dideoxyribosides (**3**, **13**, **18**, **24**) are 6 times more hydrophobic than the ribosides. Creation of a 2′-vinyl bond in the 2′,3′-dideoxyribosides reduces their hydrophobicity by half (**4**, **14**, **19** *vs.* **3**, **13**, **18**) while insertion of a 3-azido in the dideoxyribosides increases their P values by 4 times (**8**, **15** *vs.* **3**, **13**). The P values of the carbo analogues (**9**, **22**) are about half those of the corresponding ribosides (**1**, **16**). The above observation is expected. While the spatial orientation of the $-OH$ would have no significant effect, removal of $-OH$ groups and introduction of lipophilic groups would enhance the hydrophobicity of a chemical. These changes are consistent with the atomic hydrophobicity contribution of Ghose and Crippen[18]. Balzarini *et al.*[17] also observed similar correlation between the P values and the structural differences of their nucleoside derivatives.

In an effort to see if any empirical relationship exists between the P values and structural modifications on the nucleosides, we pooled our P data and those of Balzarini *et al.* together. We studied the combined P data from 48 nucleoside derivatives and found an empirical mathematical relationship existed between the P data and the

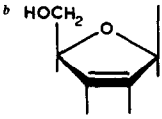
TABLE IV

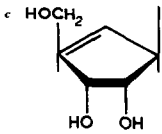
EFFECT OF STRUCTURAL CHANGE ON THE *P* OF 2',3'-DIDEOXYADENOSINE (DDA)

F_i values were derived by two-dimensional correlation between the combined *P* values from Table II of this paper and Tabels 1 and 2 of ref. 17 and the chemical structures.

Factors	F_i	*Structural modification from DDA*
F_1	0.08	Adenine changed to cytosine
F_2	0.19	Adenine changed to uracil
F_3	0.10	Adenine changed to hypoxanthine
F_4	0.18	Adenine changed to guanine
F_5	0.75	Aryl H substituted by $-NH_2$
F_6	2.05	Aryl H subsituted by $-CH_3$
F_7	1.73	Aryl H substituted by F
F_8	2.38	Aryl H substituted by Cl
F_9	4.13	Aryl H substituted by Br
F_{10}	6.61	Aryl H substituted by I
F_{11}	0.15	Sugar[a] changed to ribose or arabinose
F_{12}	0.35	Sugar[a] changed to 2'-deoxyribose
F_{13}	0.66	Sugar[a] changed to 2',3'-dideoxy-2'-vinylribose[b]
F_{14}	1.17	2'-H substituted by F
F_{15}	2.21	3'-H substituted by F
F_{16}	4.33	3'-H substituted by $-N_3$
F_{17}	0.09	Sugar[a] changed to carbovinyl derivative[c]

[a] 2',3'-Dideoxyribose.

[b] 

[c] 

structural differences of the nucleosides when they are considered as analogues of 2',3'-dideoxyadenosine (DDA). Each structural modification from DDA appears to affect the P of DDA (0.60) by a constant factor (F_i). These factors ($F_{i,j,k,\ldots}$), derived from the combined *P* data of Balzarini *et al.* and ours, are listed in Table IV. Most of the factors are derived from four or more data points. We also discovered that there is an empirical mathematical relationship between *P* of a nucleoside analogue (P_N) and that of DDA (P_A). That relationship can be expressed as eqn. 3, where P_A is taken as 0.6[a] and the *F* values are those listed in Table IV.

$$P_N = P_A F_i F_j F_k{}^* \ldots \tag{3}$$

[a] Average of 0.602 (this paper) and 0.605 (reference 16).

TABLE V
COMPARISON OF CALCULATED (EQN. 3) AND ACTUAL *P* OF NUCLEOSIDE ANALOGUES

Compound	*P calculated*	*P actual*	
		This paper	*Ref. 17*
1	0.007	0.008	
2	0.017	0.013	0.017
3	0.048	0.055	0.050
4	0.032	0.028	0.038
5	0.007	0.008	
6	0.083	0.090	
3′-Fluoro-2′,3′-dideoxycytidine	0.106		1.121
7	0.198	0.320	
8	0.208	0.208	0.231
9	0.004	0.005	
10	0.017	0.013	
2′-Deoxyuridine	0.040		0.031
5-Chloro-2′-deoxyuridine	0.095		0.078
5-Bromo-2′-deoxyuridine (5BDU)	0.165		0.102
5-Iodo-2′-deoxyuridine	0.264		0.244
2′,3′-Dideoxyuridine	0.114		0.129
3′-Fluoro-2′,3′-dideoxyuridine	0.252		0.303
3′-Fluoro-5-chloro-2,3′-dideoxyuridine	0.600		0.678
3′-Fluoro-5-bromo-2′,3′-dideoxyuridine	1.041		0.903
3′-Fluoro-5-iodo-2′,3′-dideoxyuridine	1.665		1.620
2,3′-Deoxy-2′-vinyluridine	0.075		0.085
11	0.494	0.465	0.480
12	0.082	0.079	0.067
3′-Fluoro-3′-deoxythymidine	0.516		0.529
13	0.234	0.264	0.233
14	0.154	0.193	0.154
15	1.012	1.091	0.964
16	0.090	0.105	
17	0.210	0.245	0.283
2-Amino-2′-deoxyadenosine (2ADA)	0.157		0.300
18	0.600	0.602	0.605
2-Amino-2′,3′-dideoxyadenosine (2ADDA)	0.450		0.344
3′-Fluoro-2-amino-2′,3′-dideoxyadenosine	0.994		1.128
3′-Fluoro-2′,3′-dideoxyadenosine	1.326		1.207
19	0.396	0.311	0.440
20	1.038	0.801	
Azidoadenosine	2.598		2.249
2-Amino-azidoadenosine	1.949		1.725
21	0.090	0.111	
22	0.054	0.050	
23	0.009	0.010	
24	0.060	0.068	
2′-Fluoro-2′,3′-dideoxyinosine	0.070	0.069[a]	
2′-Deoxyguanosine	0.038		0.050
25	0.108	0.085	0.098
3′-Fluoro-2′,3′-dideoxyguanosine	0.239		0.220
2′,3′-Dideoxy-2′-vinylguanosine	0.070		0.061
Azidoguanosine	0.468		0.472

[a] Ref. 19.

Thus, the *P* of AZT (**15**) is calculated to be $0.6 \cdot 0.19 \cdot 2.05 \cdot 4.33 = 1.012$. Table V compares the *P* values of the nucleoside analogues calculated according to eqn. 3 and to those determined by the shake-flask method. Except for 5BDDC, 5BDU, 2 ADA and 2ADDA, there is good agreement between the calculated and the actual *P* values.

Eqn. 3 and the factors *F* in Table IV are derived from experimentally determined *P* values of 48 nucleoside analogues. The data base is still relatively limited. However, it appears that the *P* or hydrophobicity of a nucleoside analogue can be calculated with eqn. 3 in conjunction with the factors in Table IV. As more data become available, the *F* values will be improved and refined and the validity of eqn. 3 will be further tested.

ACKNOWLEDGEMENTS

We wish to thank Mr. James Berg for technical assistance in obtaining data for four of the compounds, Ms. Gwen Wilkins for typing the manuscript and Dr. Karl Flora of the National Cancer Institute for encouragement in this work. This work is supported by the National Cancer Institute, NIH, PHS, under contract No. N01-CM-67864.

REFERENCES

1 N. Funasaki, S. Hada and S. Neya, *J. Chromatogr.*, 361 (1986) 33.
2 O. Papp, K. Volko, G. Szasz, I. Hermecz, J. Vamos, K. Hanko and Z. Ignath-Halasz, *J. Chromatogr.*, 252 (1982) 67.
3 N. El Tayer, H. van de Waterbeemd and B. Testa, *J. Chromatogr.*, 320 (1985) 305.
4 A. Negro, R. Mendez and F. Salto, *J. Liq. Chromatogr.*, 10 (1987) 2789.
5 A. Opperhuizen, *Toxicol. Environ. Chem.*, 15 (1987) 249.
6 D. Henry, J. H. Bock, J. L. Anderson and G. R. Carlson, *J. Med. Chem.*, 19 (1976) 619.
7 J. K. Baker, D. O. Rauls and R. F. Borne, *J. Med. Chem.*, 22 (1979) 1301.
8 R. Collander, *Acta Chem. Scand.*, 5 (1951) 774.
9 J. E. Haky and A. M. Young, *J. Liq. Chromatogr.*, 7 (1984) 675.
10 K. Miyake, F. Kitawa, N. Mizuno and H. Terada, *Chem. Pharm. Bull.*, 35 (1987) 377.
11 V. D. Biasi and W. J. Lough, *J. Chromatogr.*, 353 (1986) 279.
12 K. Valko, *J. Liq. Chromatogr.*, 7 (1984) 1405.
13 J. E. Garst and W. C. Wilson, *J. Pharm. Sci.*, 73 (1984) 1616.
14 J. E. Garst, *J. Pharm. Sci.*, 73 (1984) 1623.
15 D. J. Minick, J. J. Sabatka and D. A. Brent, *J. Med. Chem.*, 30 (1987) 2565.
16 D. J. Minick, J. H. Frenz, M. A. Patrick and D. A. Brent, *J. Med. Chem.*, 31 (1988) 1923.
17 J. Balzarini, M. Cools and E. De Clercq, *Biochem. Biophys. Res. Commun.*, 158 (1989) 413.
18 A. K. Ghose and G. M. Crippen, *J. Computational Chem.*, 7 (1986) 565.
19 A. P. Cheung, unpublished results.