ELSEVIER

International Journal of Pharmaceutics 181 (1999) 219–226

international journal of pharmaceutics

Are calculated log *P* values for some guanine derivatives by different computer programs reliable?

Albin Kristl *, Slavko Pečar, Vojko Kmetec

University of Ljubljana, Faculty of Pharmacy, Aškerčeva 7, 1000 Ljubljana, Slovenia

Received 15 October 1998; received in revised form 22 December 1998; accepted 18 January 1999

Abstract

The log *P* values of *n*-octanol/water for some guanine derivatives, acyclovir, deoxyaciclovir and their acetyl congeners, were calculated by some commercially available computer programs for log *P* calculation. These values were compared with those obtained by the conventional shake-flask method. It was established that the calculations of log *P* values for examined guanine derivatives by these computation programs do not give reliable results.

Keywords: Log *P* computation; Guanine derivatives; Acyclovir; Deoxyacyclovir

1. Introduction

Besides the classical determination of *n*-octanol/water partition coefficient (*P*) by the shake-flask method, other approaches can be used: filter probes, the generator column method, different reversed-phase high-performance liquid chromatography (HPLC) retention parameters and others (Dearden and Bresnen, 1988). In addition to determining *P* by direct measurement, it can be predicted by using one of several estimation techniques (Hansch and Leo, 1979; Rekker and Mannhold, 1992). The estimation of log *P* for complex structures may have a restricted importance because, so far, none of the available methods can include all the effects of molecular conformation, proximity and hydrogen bonds into the calculation procedure.

There are many different commercially available computer programs which have simplified such computations. They are arranged into three major groups: programs based on fragmental methods, those based on atomic contributions and those based on molecular (conformation dependent) properties.

Different reports are available in the literature where the authors compared experimentally determined log *P* values with those estimated by differ-

* Corresponding author.

0378-5173/99/$ - see front matter
PII: S0378-5173(99)00028-9

ent computer programs. It was shown that log *P* values, calculated by the CLOGP program (Biobyte) for some phenyl β-D-glucopyranosides (Kim and Martin, 1986), steroids (Alvarez Nunez and Yalkowsky, 1997) and for some large and flexible peptidomimetics (a series of renin inhibitors) (Karajiannis and van de Waterbeemd, 1994) were underestimated, i.e. they were lower than the measured ones. It is important to notice that in these studies (Kim and Martin, 1986; Alvarez Nunez and Yalkowsky, 1997), measured log *P* values were taken from the literature (it can be thus questionable whether these values are reliable *n*-octanol/water partition coefficients) or determined by reverse-phase HPLC (RP-HPLC) (Karajiannis and van de Waterbeemd, 1994). It was also established (Karajiannis and van de Waterbeemd, 1994) that PROLOGP program (Compu Drug Chemistry) usually overestimates the lipophilicity of the studied peptidomimetics and that for all molecules investigated, the CLOGP and PROLOGP programs give neither identical nor correlating log *P* values. Besides, log *P* values of a series of 23 propafenone-type multidrug resistance modulators were estimated using different software packages, i.e. MOLGEN (CHERS, Slovak Republic), SYBYL Version 6.2 (Tripos, Germany), PROLOGP of the PALLAS system (CompuDrug Chemistry) (Prets et al., 1996). For many compounds, large differences (even two log units) between calculated log *P* values, computed by these programs, were observed. The best correlation between lipophilicity indices obtained by HPLC and log *P* values calculated by computer programs was established for MOLGEN ($r = 0.990$) and SYBYL ($r = 0.980$) software packages. Mannhold et al. (1990) compared experimentally determined log *P* values (by shake flask method) with calculated ones, i.e. Σ*f*-values (Rekker and Mannhold, 1992) and CLOGP (Hansch and Leo, 1979), for four groups of drugs (15 antiarrhythmics, 11 β-blockers, 13 phenothiazines and nine benzamides), and rather large differences (even larger than one log unit) between experimental and calculated log *P* values were observed. The results also showed that both calculative procedures (Σ*f*-values and CLOGP) exhibit similar qualities in calculating log *P*. Additionally, Mannhold and Dross (1996) have checked the predictive power of 14 calculation procedures for molecular lipophilicity using the database of 138 test compounds (which comprises 90 simple organic structures and 48 chemically heterogeneous drug molecules). Their analysis demonstrated a significantly higher quality of the fragmental methods as compared with the programs based on atomic contributions and on molecular properties. They also concluded that the predictive power of the calculation procedures is significantly better for the simple organic molecules than for chemically heterogeneous drug structures.

The aim of this work was to compare log *P* values for some guanine derivatives, calculated by different computer programs, with those determined by the experiment and to establish whether these values are trustworthy.

2. Materials and methods

The antivirus guanine derivatives examined in this study were acyclovir (ACV; 9-(2-hydroxyethoxymethyl)guanine) and deoxyacyclovir (DCV; 2-amino-9-(2-hydroxyethoxymethyl)-9H-purine) with their *O*-acetyl (OAcACV, OAcDCV), *N*-acetyl (NAcACV, NAcDCV) and *N*,*O*-diacetyl (diAcACV, diAcDCV) congeners and were synthesised at the National Institute of Chemistry, Ljubljana, Slovenia (Štimac and Kobe, 1990). The structures of ACV and DCV are given in Fig. 1. The log *P* values of these

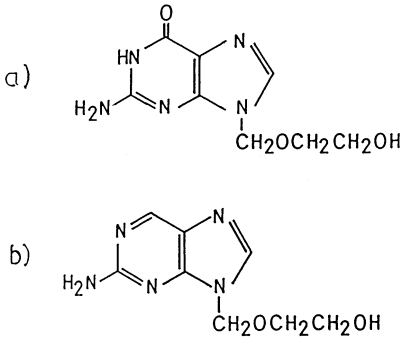

Fig. 1. Structural formulae of (a) keto form of ACV and (b) DCV.

Table 1
Experimental and calculated log P values by different programs (HYPERCHEM, PACO, CLOGP, KOWWIN, MICROQSAR and PROLOGP: ATOMIC, ATOMIC5, CDR and combined)

	log P exp.	log P HYPERCHEM keto	log P HYPERCHEM enol	log P PACO keto	log P PACO enol	log P CLOGP keto	log P CLOGP enol	log P KOWWIN keto	log P KOWWINENOL	log P MICROQ keto	log P MICROQ enol	log P ATOMIC keto	log P ATOMIC enol	log P ATOMIC5 keto	log P ATOMIC5 enol	log P CDR keto	log P CDR enol	log P combined keto	log P combined enol
ACV	−1.56	−0.98	−0.41	−0.35	−1.76	−2.30	−0.69	−1.70	−1.60	−1.48	−0.79	−1.98	−2.08	−2.31	−0.96	−1.47	−2.05	−2.08	−1.25
NAcACV	−1.30	−0.68	−0.12	−0.77	−2.37	−2.14	−0.54	−1.75	−1.64	−1.65	−0.96	−1.50	−1.86	−2.74	−0.98	−1.79	−3.12	−2.49	−1.55
OAcACV	−1.08	−0.85	−0.29	+0.48	−0.93	−1.40	0.21	−0.69	−0.60	−1.45	−0.76	−1.39	−1.49	−1.92	−0.57	−0.57	−1.15	−1.56	−0.72
diAcACV	−0.83	−0.55	+0.01	+0.07	−1.54	−1.25	0.36	−0.75	−0.64	−1.62	−0.93	−0.91	−1.26	−2.35	−0.60	−0.89	−2.22	−1.96	−1.03
DCV	−1.08	−0.83		−0.38		−1.32		−1.52		−0.39		−1.74		−0.97		−1.53		−1.12	
NAcDCV	−1.33	−0.53		−0.99		−1.18		−1.56		−0.56		−1.51		−1.00		−2.59		−1.42	
OAcDCV	−0.61	−0.70		+0.45		−0.43		−0.52		−0.36		−1.15		−0.58		−0.63		−0.59	
diAcDCV	−1.05	−0.40		−0.16		−0.28		−0.56		−0.53		−0.92		−0.61		−1.69		−0.90	

Table 2
The Pearson correlation coefficients obtained by the least-squares linear regression method for experimentally determined log P values and those calculated[a]

	HYPERCH	PACO	CLOGP	KOWWIN	MICROQ	ATOMIC	ATOMIC5	CDR	Combined
Keto[b]	0.339	0.696	0.736*	0.806*	0.343	0.740*	0.428	0.636	0.602
Enol[c]	−0.110	0.648	0.425	0.788*	0.373	0.749*	0.759*	0.628	0.751*

[a] Some correlation equations are also included (i.e. PACO, KOWWIN and combined for keto and enol forms).
[b] $\log P_{PACO} = 1.161 + 1.237 \log P_{exp}$; $\log P_{KOWWIN} = 0.500 + 1.476 \log P_{exp}$; $\log P_{combined} = -0.084 + 1.295 \log P_{exp}$.
[c] $\log P_{PACO} = 1.254 + 2.004 \log P_{exp}$; $\log P_{KOWWIN} = 0.490 + 1.420 \log P_{exp;}$ $\log P_{combined} = -0.147 + 0.837 \log P_{exp}$.
* The correlation is significant at the 0.05 level.

substances were determined previously by the shake flask method (Kristl and Vesnaver, 1995). In this study, we used some commercially available computer programs for log P computations, i.e. HYPERCHEM Release 5.0 (HyperChem, 1996) which uses atomic parameters; PACO, Version 2.9 (PACO, 1990) where the calculation principles are based on an additivity scheme; CLOGP (Hansch and Leo, 1979) based on Hansch's fragmental approach; KOWWIN Version 1.60 (Meylan and Howard, 1995) which applies atom/fragment contributions; MICROQSAR Version 2.0 using the Ghose and Crippen method with atomic physicochemical parameters (Ghose and Crippen, 1986); and PROLOGP module of the Pallas system (Pallas, 1995) which offers different methods: the CDR database is based on unrevised Rekker's hydrophobic fragmental constants, the ATOMIC database contains the atomic fragments according to Broto et al. (1984) and the ATOMIC5 database including the atomic fragments from Ghose et al. (Viswanadhan et al., 1989) are used for calculation. The module PROLOGP offers also the possibility to calculate a combined log P value, whereby the default values are set to $\log P_{combined} = 0.733 \log P_{ATOMIC5} + 0.267 \log P_{CDR}$.

3. Results and discussion

Experimentally determined and calculated values for log P are given in Table 1. One can observe rather large discrepancies between calculated and experimentally determined values in almost all cases; a large scatter can be observed in Fig. 2. In some cases (i.e. CLOGP, ATOMIC5), experimentally determined log P values for ACV derivatives lie between the calculated log P values for keto and enol forms, but in other calculations (i.e. KOWWIN, ATOMIC, etc.) this is not the case. This supports the idea that keto–enol tautomerism cannot play the major role in log P calculations; it is more the matter of log P computational programs.

The Pearson correlation coefficients (Table 2) obtained by the least-squares linear regression method for experimentally determined log P values and the calculated ones are relatively low and, in general, show very weak correlation. A similar observation can be made when one considers the slope and the intercept of the correlation equations given in Table 2. There are only a few correlation coefficients which exhibit significant correlation at the 0.05 level (there is no significant correlation at the 0.01 level): CLOGP, KOWWIN and ATOMIC for keto forms; KOWWIN, ATOMIC, ATOMIC5 and combined for enol forms.

The differences between calculated and measured log P values (Δ), given in Table 3, show that log P values calculated by HYPERCHEM, PACO, CLOGP (enol) and MICROQSAR are generally overestimated, on the other hand CLOGP (keto), ATOMIC, CDR and combined give underestimated values, while ATOMIC5 gives mixed values (for ACV keto derivatives, the log P values are underestimated, and for ACV enol and DCV derivatives, they are overestimated). Mixed values are calculated also by KOWWIN, which in fact exhibits the smallest differences between calculated and experimentally determined log P values (Δ_{av} is the low-

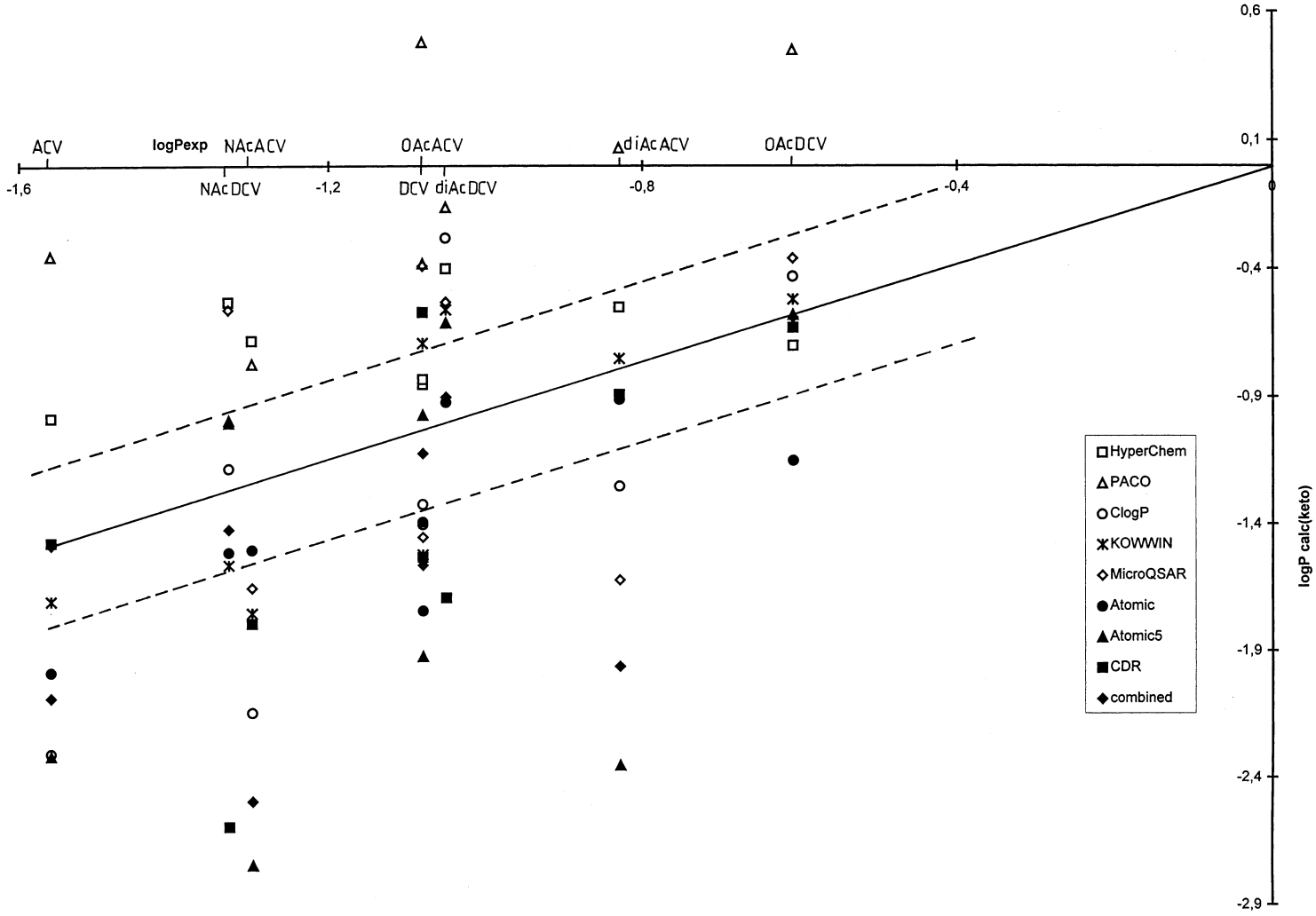

Fig. 2. The dependence of log P calculated by different programs on log P_{exp} is given. The solid line represents the values where log P_{exp} = log P_{calc}, while the dotted lines represent the area of (± 0.3 log P units).

Table 3
The differences between calculated and experimentally determined data ($\Delta = \log P_{calc} - \log P_{exp}$) with averages of their absolute values ($n = 8$), are given[a]

	$\Delta_{HYPERCH}$ keto	$\Delta_{HYPERCH}$ enol	Δ_{PACO} keto	Δ_{PACO} enol	Δ_{CLOGP} keto	Δ_{CLOGP} enol	Δ_{KOWWIN} keto	Δ_{KOWWIN} enol	$\Delta_{MICROQSAR}$ keto	$\Delta_{MICROQSAR}$ enol	Δ_{ATOMIC} keto	Δ_{ATOMIC} enol	$\Delta_{ATOMIC5}$ keto	$\Delta_{ATOMIC5}$ enol	Δ_{CDR} keto	Δ_{CDR} enol	$\Delta_{combined}$ keto	$\Delta_{combined}$ enol
ACV	+0.58	+1.15	+1.21	−0.20	−0.74	+0.87	−0.14	−0.04	+0.08	+0.77	−0.42	−0.52	−0.75	+0.60	+0.09	−0.49	−0.52	+0.31
NAcACV	+0.62	+1.18	+0.53	−1.07	−0.84	+0.76	−0.45	−0.34	−0.35	+0.34	−0.20	−0.56	−1.44	+0.32	−0.49	−1.82	−1.19	−0.25
OAcACV	+0.23	+0.79	+1.56	+0.15	−0.32	+1.29	+0.39	+0.48	−0.37	+0.32	−0.31	−0.41	−0.84	+0.51	+0.51	−0.07	−0.48	+0.36
diAcACV	+0.28	+0.84	+0.90	−0.71	−0.42	+1.19	+0.08	+0.19	−0.79	−0.10	−0.08	−0.43	−1.52	+0.23	−0.06	−1.39	−1.13	−0.20
DCV	+0.25		+0.70		−0.24		−0.44		+0.69		−0.66		+0.11		−0.45		−0.04	
NAcDCV	+0.80		+0.34		+0.15		−0.23		+0.77		−0.18		+0.33		−1.26		−0.09	
OAcDCV	−0.09		+1.06		+0.18		+0.09		+0.25		−0.54		+0.03		−0.02		+0.02	
diAcDCV	+0.65		+0.89		+0.77		+0.49		+0.52		+0.13		+0.44		−0.64		+0.15	
Δ_{av}	0.44	0.99	0.90	0.53	0.46	0.68	0.29	0.29	0.48	0.47	0.32	0.43	0.68	0.32	0.44	0.77	0.45	0.28

[a] $\Delta_{av} = \dfrac{\sum_{i=1}^{n} |\Delta_i|}{n}$.

est and none of differences between log P_{exp} and log P_{calc} is higher than 0.5). Additionally, in KOWWIN computations there are also the highest values for Pearson coefficients (Table 2).

The differences between calculated and experimentally determined data (Δ) are in many cases even larger than 1; only in the case of HYPERCHEM (keto), CLOGP (keto), KOWWIN, MICROQSAR, ATOMIC, ATOMIC5 (enol) and combined (enol) all the calculated differences are smaller than 1. There is also a considerably large amount of differences in the range $0.5 < \Delta < 1$ for almost all computational programs; the exceptions are only KOWWIN and combined (enol). If one takes into account more rigorous conditions (some authors have reported that log P values determined for the same compound are acceptable if the differences among them are lower than 0.3 log unit, i.e. Dearden and Bresnen, 1988; Rekker and Mannhold, 1992) than among 80 log P computations for keto only 29 values are acceptable (Fig. 2).

Careful comparison of log P values for ACV and its derivatives with those for DCV and its derivatives shows that log P differences between the parent molecule (i.e. ACV or DCV) and its acetylated derivatives, calculated by any particular method, are irrespective of keto–enol tautomerism almost the same. This indicates that the methods used apply different methodologies, although all are based on the same principle of the additivity; every particular method itself uses a similar group contribution (i.e. the contribution of *O*-acetylation or *N*-acetylation is, in the case of ACV and DCV, the same).

It is evident that for the examined eight relatively complicated test substances, rather unreliable values were calculated by these computational programs. On the basis of the results obtained, one cannot conclude whether the computations based on fragmental methods or on atomic contributions give more reliable log P estimations. However, regarding the Pearson correlation coefficients (Table 2) and the differences between calculated and experimentally determined values (Table 3), the best calculation of log P values are obtained by KOWWIN. These findings are in agreement with Mannhold and Dross (1996), where the best overall results showed KOWWIN computations. Relatively good results are also obtained by ATOMIC, combined and CLOGP (keto), while log P values obtained by MICROQSAR exhibit relatively low Δ_{av}, but the Pearson correlation coefficient is very low.

4. Conclusions

Overall one can conclude that the calculations of log P values for complex molecules such as tested guanine derivatives by different computer programs do not give reliable results. These results support our previous findings that tested guanine derivatives exhibit rather unexpected hydrolipophilic properties; neither the calculation of log P by Rekker's approach nor the determination of hydrolipophilic properties with HPLC gives reliable results (Kristl and Pečar, 1997). This also indicates that the validity of calculated log P values should be established, i.e. by comparing the observed and calculated values for structurally similar compounds, before using them in different QSAR studies.

References

Alvarez Nunez, F.A., Yalkowsky, S.H., 1997. Correlation between log P and C log P for some steroids. J. Pharm. Sci. 86, 1187–1189.

Broto, P., Moreau, G., Vandycke, C., 1984. Molecular structures perception, auto correlation descriptor and SAR studies—system of atomic contributions for the calculation of the normal octanol water partition coefficients. Eur. J. Med. Chem.-Chim. Ther. 19, 71–78.

Dearden, J.C., Bresnen, G.M., 1988. The measurement of partition coefficients. Quant. Struct.-Act. Relatsh. 7, 133–144.

Ghose, A.K., Crippen, G.M., 1986. Atomic physicochemical parameters for 3-dimensional structure-directed quantitative structure–activity relationships. l. Partition coefficients as a measure of hydrophobicity. J. Computat. Chem. 7, 565–577.

Hansch, C., Leo, A., 1979. Substituent Constants for Correlation Analysis in Chemistry and Biology. Wiley, New York.

HyperChem Release 5.0, 1996. HyperCube, Canada.

Karajiannis, H., van de Waterbeemd, H., 1994. The prediction of the lipophilicity of peptidomimetics. A comparison between experimental and theoretical lipophilicity values of

renin inhibitors and their building blocs. Pharm. Acta. Helv. 70, 67–77.

Kim, K.H., Martin, Y.C., 1986. Comparison of calculated versus measured partition coefficients of some phenyl β-D-glucopyranosides. J. Pharm. Sci. 75, 637–638.

Kristl, A., Pečar, S., 1997. Hydrolipophilic anomalies of some guanine derivatives. Eur. J. Med. Chem. 32, 3–8.

Kristl, A., Vesnaver, G., 1995. Thermodynamic investigation of the effect of octanol–water mutual miscibility on the partitioning and solubility of some guanine derivatives. J. Chem. Soc. Faraday. Trans. 91, 995–998.

Mannhold, R., Dross, K., 1996. Calculation procedures for molecular lipophilicity: a comparative study. Quant. Struct.-Act. Relatsh. 15, 403–409.

Mannhold, R., Dross, K.P., Rekker, R.F., 1990. Drug lipophilicity in QSAR practice: I. A comparison of experimental with calculative approaches. Quant. Struct.-Act. Relatsh. 9, 21–28.

Meylan, W.M., Howard, P.H., 1995. Atom/fragment contribution method for estimating octanol–water partition coefficients. J. Pharm. Sci. 84, 83–92.

PACO, Version 2.9, 1990. Chemodata-Computer Chemie GmbH, Germany.

Pallas, PrologP 5.1, 1995. CompuDrug Chemistry Ltd., Hungary.

Prets, S., Jungreithmair, A., Chiba, P., Ecker, G., 1996. Comparison of different methods for estimation of lipophilicity of propafenone-type modulators of multidrug resistance. Sci. Pharm. 64, 627–636.

Rekker, R.F., Mannhold, R., 1992. Calculation of Drug Lipophilicity (The Hydrophobic Fragmental Constant Approach). VCH, Weinheim.

Štimac, A., Kobe, J., 1990. A new synthesis of acyclovir prodrugs. N2 acetylacyclovir and 6-deoxyacyclovir, Synthesis, 461–464.

Viswanadhan, V.N., Robins, R.K., Revankar, G.R., Ghose, A.K., 1989. Atomic physicochemical parameters for 3-dimensional structure directed quantitative structure activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics. J. Chem. Inf. Comput. Sci. 29, 163–172.