Available online at www.sciencedirect.com
SCIENCE DIRECT®
European Journal of Pharmaceutical Sciences 20 (2003) 115–123
EUROPEAN JOURNAL OF PHARMACEUTICAL SCIENCES
www.elsevier.com/locate/ejps

Quantitative structure–pharmacokinetic relationships for disposition parameters of cephalosporins

Vangelis Karalis[a], Anna Tsantili-Kakoulidou[b], Panos Macheras[a,*]

[a]*Laboratory of Biopharmaceutics-Pharmacokinetics, School of Pharmacy, University of Athens, Athens, Greece*
[b]*Division of Pharmaceutical Chemistry, School of Pharmacy, University of Athens, Athens, Greece*

Received 12 February 2003; received in revised form 9 June 2003; accepted 24 June 2003

Abstract

This study explores the utility of quantitative structure–pharmacokinetic relationship models of the disposition parameters: clearance (CL), apparent volume of drug distribution (V_{ap}), fractal clearance (CL_f), and fractal volume (v_f), for a series of 23 cephalosporins used in therapeutics. Data for CL, V_{ap} and elimination half-life were obtained from literature, whereas CL_f and v_f were calculated from the literature data for CL and V_{ap}, respectively. A variety of descriptors expressing acidity/basicity, lipophilicity, molecular size and hydrogen bonding properties were estimated using computer packages. For each pharmacokinetic parameter, projection to latent structures (PLS) was applied to the total dataset. Adequate PLS models, with one principal component, were derived for CL, CL_f, V_{ap} and v_f. Identical descriptors were found to be significant for the two clearance as well as for the two volume of distribution terms. CL and CL_f expressed similar performance while the predictive performance of v_f was much higher than that of V_{ap}. Multiple linear and non-linear regression models were developed. The regression results were in agreement with the PLS models. The non-linear models were superior to the relevant linear relationships. The worst models found were for V_{ap} ($R^2=0.523$ and $R^2=0.571$ for the linear and non-linear model, respectively) and the best models found were for v_f ($R^2=0.729$ and $R^2=0.824$ for the linear and non-linear model, respectively).

Keywords: Quantitative structure–pharmacokinetic relationships; Cephalosporins; Multivariate data analysis; Regression analysis

1. Introduction

In recent years, the advent of combinatorial chemistry has increased the number of compounds entering drug discovery process. However, this enormous amount of candidate drugs forces for an early selection of the compounds which have the greatest possibility of success. In this aspect, the focus is not only on achieving the best pharmacological efficacy, but also on seeking desirable ADME (absorption, distribution, metabolism, excretion) characteristics (Boobis et al., 2002; Kretz and Probst, 2001). A variety of high throughput experimental and theoretical methods have arisen for screening candidate molecules. The development of quantitative structure–pharmacokinetic relationships (QSPR) using 'in-silico' procedures is of special interest. QSPR models focus on the association of structural features of compounds to pharmacokinetic (PK) parameters (Fouchécourt et al., 2001). The aim of such relationships is the prediction of the pharmacokinetic behavior from easily measured/estimated physicochemical or molecular properties.

Several successful attempts have been reported to establish QSPR models for intestinal absorption within congeneric series of drug molecules (Betageri and Rogers, 1989; Esaki, 1987; Markin et al., 1988; Toma, 1989). Moreover, QSPR models for human intestinal absorption of structurally unrelated compounds have been developed (Klopman et al., 2002; Zhao et al., 2001a,b). However, this task becomes complicated for disposition pharmacokinetic parameters and fewer articles have appeared in literature (Gobburu and Shelver, 1995; Poulin and Theil, 2000; Poulin et al., 2001). The difficulties in the development of successful QSPR models for disposition parameters should be attributed to the composite and interrelated nature of the distribution and elimination processes. Besides, the frequently encountered fictitious character of the numerical values of the apparent volume of drug distribution, V_{ap}, makes the QSPR modeling even more questionable.

*Corresponding author. Tel.: +30-210-727-4026; fax: +30-210-727-4027.
E-mail address: macheras@pharm.uoa.gr (P. Macheras).

0928-0987/03/$ – see front matter
doi:10.1016/S0928-0987(03)00177-5

Recently, a more physiologically relevant description for drug volume of distribution and clearance was reported; fractal volume of drug distribution, v_f, and fractal clearance, CL_f, were proposed as substitutes for V_{ap} and clearance, CL, respectively (Karalis et al., 2001; Karalis and Macheras, 2002). These two newly proposed parameters were found to exhibit better properties in interspecies allometric scaling (Karalis et al., 2001; Karalis and Macheras, 2002). QSPR modeling of the traditional V_{ap} and CL as well as of the corresponding fractal parameters, v_f and CL_f for a large number of structurally unrelated drugs has already been performed using multivariate statistics (Karalis et al., 2002). An adequate model was developed only for v_f but not for CL_f and the corresponding traditional parameters. Next to the complexity of the disposition parameters, the high degree of structural diversity of the dataset was considered as an additional difficulty to establish reliable models. In this context, we tried to apply the same methodology in a congeneric set of drugs. Cephalosporins comprise a widely used family of therapeutic agents and consequently the necessary pharmacokinetic properties were available for 23 drugs in this category. Both the conventional and fractal disposition parameters of these drugs were analyzed by multivariate data analysis (MVDA) (Eriksson et al., 1995; Eriksson and Johansson, 1996; Franke and Gruska, 1995; Wold, 1995) and regression analysis using a variety of molecular descriptors. The aim of this study was to explore the potential differentiation in information content between conventional and traditional disposition parameters and especially the suitability of CL_f and v_f in QSPR studies.

2. Methods

Pharmacokinetic data (CL, V_{ap} and elimination half-life, $t_{1/2}$) for the 23 cephalosporins used in this study were obtained from a classic textbook (Hardman et al., 1996). Values of v_f were estimated from the reported V_{ap} values (Hardman et al., 1996) using Eq. (1) (Karalis et al., 2001):

$$v_f = V_{pl} + (v - V_{pl})\left(1 - \frac{V_{pl}}{V_{ap}}\right) \quad (1)$$

where v is the total volume of the species body (equivalent to the total mass assuming a uniform density 1 g/ml), and V_{pl} is the plasma volume of the species. In our study, the typical human values for v and V_{pl} were used (70 and 3 l, respectively). The clearance analog of v_f, called for reasons of uniformity fractal clearance, CL_f, refers to the portion of v_f which is cleared per unit of time (Karalis and Macheras, 2002). CL_f estimates were derived from Eq. (2) using the reported elimination half-life ($t_{1/2}$) values (Hardman et al., 1996), and the calculated v_f values:

$$CL_f = \frac{\ln 2}{t_{1/2}} v_f \quad (2)$$

Fraction of drug bound to plasma proteins, f_b, was obtained from the same bibliographic source (Hardman et al., 1996) and was included as an important variant in the analysis of the disposition parameters. Table 1 summarizes the 23 cephalosporins used in this study along with the utilized PK parameters.

A variety of physicochemical and molecular descriptors (Table 2) expressing lipophilicity, ionization, molecular size and hydrogen bonding capacity, were calculated using HyperChem v.5.0/ChemPlus v.1.6 (Hypercube Inc.) and Pallas 2.1 (Compudrug Chemistry Ltd).

Molecular size was expressed by a variety of descriptors: molecular weight (*MW*), molar refractivity (*refr*), molecular polarizability (*polrz*), Van der Waals surface area (*SVdW*) or volume (*VVdW*), solvent accessible surface area (*Ssol*) or volume (*Vsol*). *Polarity* was expressed by molecular polar surface area based on solvent accessible surface area (*PSsol*) or Van der Waals surface area (*PSVdW*). For this purpose, as polar atoms were considered all O and N atoms, as well as all H atoms bound to O and N atoms. The corresponding *non-polar surface areas*, *nPSsol* and *nPSVdW*, were obtained by subtracting *PSsol* and *PSVdW* from *Ssol* and *SVdW*, respectively. All the above molecular size and polarity descriptors were calculated using the ChemPlus v.1.6 module implemented in Hyperchem v.5.0 after 3-D optimization. The geometry of a given molecule was first optimized at the empirical level using an MM+ molecular mechanics force field, followed by unrestricted geometric optimization at the semi-empirical level using an SCF calculation with convergence limit set at 0.1 Kcal/mol.

Hydrogen bonding capacity was expressed with two descriptors; the number of hydrogen bond donors (*HDO*) and the number of hydrogen bond acceptors (*HAC*). *HDO* represents the number of all O–H and N–H fragments, but excluding hydrogens belonging to all kind of acids and thiols (Oprea, 2000). Likewise, *HAC* counts all oxygen and nitrogen atoms, whereas exceptions were the nitrogen in carbamides, sulfonamides, and the nitrogen atoms which are bound with three alkyl groups.

Intrinsic lipophilicity was expressed by log *P* of the neutral species. The ChemPlus module implemented in Hyperchem v.5.0 was used for the estimation of log *P*-values of the various compounds according to the original Ghose–Crippen system (log *PG*) (Viswanadhan et al., 1989).

Dissociation constants, expressed as acidic and basic pK_a, were estimated using the pK_{alc} module of Pallas 2.1 and were used to estimate the fraction ionized. For compounds with more than one acidic center, only the pK_a-value for the most potent acidic group was considered.

The descriptors used to derive the final models are presented in Table 2.

Multivariate data analysis was performed using SIMCA-P v.8.0 (Umetri AB, Umea, Sweden). Principal component analysis (PCA) (Franke and Gruska, 1995), which is a

Table 1
Values of the pharmacokinetic parameters for the cephalosporins under study (Hardman et al., 1996)

	Drug	CL (l/min)	V_{ap} (l)	$t_{1/2}$ (min)	v_f [a] (l)	CL_f [b] (l/min)	f_b
1	Cefaclor	0.427	25.2	40.2	62.02	1.07	0.25
2	Cefadroxil	0.203	16.8	72	58.04	0.56	0.20
3	Cefamandole	0.196	11.2	46.8	52.05	0.77	0.74
4	Cefazolin	0.067	9.8	108	49.49	0.32	0.89
5	Cefixime	0.091	21.0	180	60.43	0.23	0.67
6	Cefmetazole	0.102	12.6	90	54.05	0.42	0.70
7	Cefonicid	0.022	7.7	264	43.90	0.12	0.98
8	Cefoperazone	0.084	9.8	132	49.49	0.26	0.91
9	Ceforanide	0.037	9.8	156	49.49	0.22	0.81
10	Cefotaxime	0.259	16.1	66	57.52	0.60	0.36
11	Cefotetan	0.039	9.8	216	49.49	0.16	0.85
12	Cefoxitine	0.426	17.5	45	58.51	0.90	0.73
13	Cefpodoxime	0.168	32.2	138	63.76	0.32	0.27
14	Ceftazidime	0.140	16.1	96	57.52	0.42	0.21
15	Ceftizoxime	0.142	25.2	108	62.02	0.40	0.28
16	Ceftriaxone	0.017	11.2	438	52.05	0.08	0.93
17	Cefuroxime	0.137	14.0	102	55.64	0.38	0.33
18	Cephalexin	0.301	18.2	54	58.96	0.76	0.14
19	Cephalothin	0.469	18.2	34.2	58.96	1.19	0.71
20	Cephapirin	0.483	14.7	43.2	56.33	0.90	0.62
21	Cephradine	0.336	32.2	54	63.76	0.82	0.14
22	Loracarbef	0.216	22.4	72	61.03	0.59	0.25
23	Moxalactame	0.130	17.5	126	58.51	0.32	0.60

[a] Calculated from Eq. (1).
[b] Calculated from Eq. (2).

multivariate projection method to extract the systemic variables in the data matrix, was applied to the total dataset of the PK parameters and the molecular descriptors. Each separate PK parameter (CL, CL_f, V_{ap}, v_f) was further explored by partial least squares analysis (PLS) (Eriksson et al., 1995; Eriksson and Johansson, 1996; Wold, 1995).

Table 2
Cephalosporins and the corresponding calculated values of descriptors entering the final models

	Drug	Descriptors						
		log P	*Pssol*	*HDO*	*HAC*	*MW*	*Vsol*	*VVdW*
1	Cefaclor	−0.92	101.3	3	5	367.8	926.1	289.7
2	Cefadroxil	−0.83	122.4	4	6	363.4	956.0	298.2
3	Cefamandole	1.22	175.3	2	8	462.5	1168.5	368.3
4	Cefazolin	0.26	209.8	1	9	454.5	1096.9	344.0
5	Cefixime	−1.43	180.2	3	10	453.4	1105.7	342.4
6	Cefmetazole	0.90	223.0	1	9	471.5	1149.2	366.2
7	Cefonicid	1.16	260.2	2	11	542.6	1277.6	404.1
8	Cefoperazone	−0.21	249.3	3	11	645.7	1558.3	514.6
9	Ceforanide	0.30	221.4	3	10	519.6	1298.6	414.4
10	Cefotaxime	−2.03	153.7	3	10	455.5	1134.0	349.9
11	Cefotetan	0.45	240.9	3	11	575.6	1297.4	420.2
12	Cefoxitine	−2.26	136.9	3	7	427.5	1036.4	327.9
13	Cefpodoxime	−1.88	123.1	3	9	427.5	1077.6	332.2
14	Ceftazidime	0.03	148.5	4	11	547.6	1359.2	437.1
15	Ceftizoxime	−1.31	131.5	3	8	383.4	957.7	291.1
16	Ceftriaxone	−0.79	208.6	4	12	554.6	1315.9	415.2
17	Cefuroxime	−2.18	161.9	3	9	424.4	1054.9	324.5
18	Cephalexin	−0.54	99.3	3	5	347.4	935.3	292.2
19	Cephalothin	−2.79	132.1	1	6	396.4	999.6	309.8
20	Cephapirin	−3.30	155.9	1	7	423.5	1006.5	335.5
21	Cephradine	−1.05	111.6	3	5	349.4	962.2	298.9
22	Loracarbef	−0.92	99.9	3	5	367.8	926.1	289.7
23	Moxalactame	0.01	244.7	2	11	520.5	1189.1	402.6

PLS is a regression extension of PCA applied to connect the information in the two blocks of variables X (descriptor matrix) and Y (PK parameters). The predictive ability of the derived PLS models was assessed in three different ways (Oprea, 2000; Wold, 1995): (i) using cross-validation with the default values of SIMCA-P, (ii) randomly reordering the responses (Y-data) and evaluation of the properties of the derived models, and (iii) dividing each parent set into a training and a validation set for the assessment each training's set ability to accurately predict the values of the validation set.

Multiple linear and non-linear regression analysis was performed using Mathematica v.4.0 (Wolfram Research, Inc.). The development of multiple linear and non-linear regression relationships was based on the best models obtained from the PLS analysis of each PK parameter. All descriptors prior to their application to multiple regression analysis were checked for linear interdependence assigning as a limit $R^2 < 0.4$. Ceftazidime was not included in the regression analysis. This drug includes a permanently charged nitrogen atom in its molecule, which renders the calculation of molecular descriptors by Hyperchem/ChemPlus disputable or erroneous. MVDA can handle inaccurate and missing values; however, this is not the case for regression analysis.

3. Results

According to the initial PCA (results not shown) applied to the total dataset, no strong outliers were identified since all drugs were lying inside the 95% confidence ellipse (Hotelling T^2). In addition, no separate groups were identified despite of the fact that the 23 cephalosporins can be divided into two large groups in regard to their percentage of urinary excretion (less than 55% for seven compounds and much higher than this value for the remaining). Hence, all drugs were used to develop PLS models for each PK parameter separately. The final models derived after variable selection were based on one principal component with R^2 values between 0.592 and 0.775 and Q^2 values between 0.554 and 0.731 (Table 3). The regression coefficients for each model describing the response of CL, CL_f, V_{ap}, and υ_f are shown in Fig. 1A–D. The two clearance expressions CL, and CL_f, were reflected by identical molecular properties and exhibited approximately the same Q^2-values, namely 0.731 and 0.709, respectively. In contrast, the PLS models obtained for the volume expressions (V_{ap} and υ_f) differ considerably in their statistics although they include similar descriptors. An adequate model was obtained for υ_f ($Q^2 = 0.717$), while the model for V_{ap} was significantly inferior ($Q^2 = 0.554$). Fig. 2A–D shows the u_1 vs. t_1 plots for CL, CL_f, V_{ap}, and υ_f, respectively. These plots reveal that adequately linear PLS inner relations for CL, CL_f, and υ_f were derived, while a higher degree of scattering is observed for V_{ap}.

Table 3
Correlation coefficients (R^2, Q^2) and number of components (A) for each PLS model

No.	PLS models	n[a]	R^{2}[b]	Q^{2}[c]	A
Total set					
1	CL	23	0.775	0.731	1
2	CL_f	23	0.753	0.709	1
3	V_{ap}	23	0.592	0.554	1
4	υ_f	23	0.754	0.717	1
Training set					
	CL	18	0.753	0.685	1
	CL_f	18	0.718	0.656	1
	V_{ap}	18	0.597	0.499	1
	υ_f	18	0.753	0.691	1

[a] Number of compounds analyzed.
[b] Coefficient of determination.
[c] Cross-validated coefficient of determination.

The validity of the derived PLS models was further examined by applying two additional statistical tools using permutation tests and division of parent data into training and validation sets. In permutation testing, which is based on the randomization of responses (i.e. the original data are permuted to appear in a different order), the R^2 and Q^2 estimates of the scrambled data are plotted against the R^2-value of the Y-vector itself. A good behavior for the PLS models of CL, CL_f and υ_f was observed, since the Y-intercepts of the R^2 and Q^2 estimates were very close to zero (not shown). The splitting of the parent set into a training (18 drugs) and a test set (five drugs) was based on a random generator program developed in Mathematica v.4.0. PLS analysis was applied to each training set separately and the derived models were similar to those obtained with the total data set (Table 3). In Fig. 3A–D the predicted values are plotted versus the observed values of CL, CL_f, V_{ap}, and υ_f for both the training and the test set.

The PLS models quoted above were used as the basis for the development of regression models in order to have a simple and interpretable relationship between the PK parameters and the molecular properties. Table 4 summarizes the results of the regression models, while the graphical representation of each PK parameter with the relevant descriptors for the non-linear relationships, is shown in Fig. 4A–D. In all cases, the derived non-linear relationships were superior to the corresponding linear models. The plots of the predicted values of the disposition parameters based on models B1–B4 quoted in Table 4 versus the observed values are shown in Fig. 5.

4. Discussion

One component PLS models were derived for CL, CL_f, V_{ap} and υ_f. Their validity was verified using the statistical

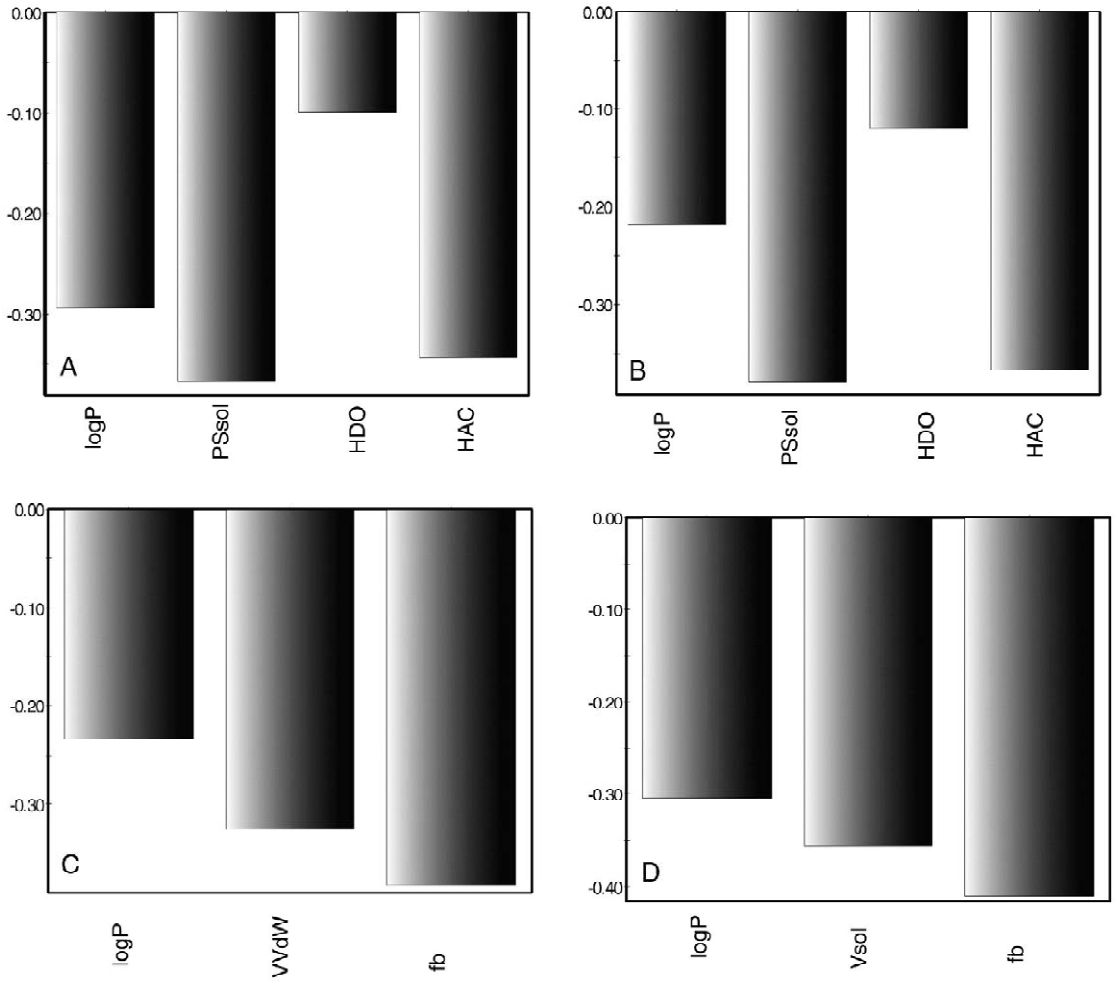

Fig. 1. Regression coefficients for each PLS model describing the response of *CL* (A), CL_f (B), V_{ap} (C), v_f (D) for the utilized cephalosporins.

approaches described previously. For *CL* and CL_f, both the descriptive and the predictive ability were found to be adequately high (models 1, 2; Table 3). Inspection of the regression coefficient plots (Fig. 1) reveals that similar descriptors, entering with the same sign, are responsible for the two clearance terms. Lipophilicity (log *P*), polar surface area (*PSsol*) and hydrogen bonding properties (*HDO*, *HAC*) contributed negatively to *CL* and CL_f.

In the case of the volume parameters an adequate model was derived for v_f, while a significantly inferior model was found for V_{ap} (models 3, 4; Table 3). For both V_{ap} and v_f, the dominating descriptor was the fraction of drug bound to plasma proteins (f_b). Besides, volume parameters (Van der Waals molecular volume in the case of V_{ap}, solvent accessible volume in the case of v_f), and lipophilicity were also found to be important for both V_{ap} and v_f. The negative sign of the contribution of f_b to the V_{ap} and v_f models is a reasonable finding since drugs with a high degree of protein binding exhibit smaller values of volume of distribution (Qie, 1996; Urien et al., 2001). The difference in the quality of the models 3 and 4 of Table 3 could be explained on the basis of the different information content of the two volume parameters although they both lie within the physiological range (<70 l) i.e. the V_{ap} values range from 7.7 to 32.2 l whereas the v_f values range from 43.9 to 63.8 l (Table 1).

Multiple regression analysis was based on the PLS models considering only the independent descriptors among those which had been found to be significant for the PLS models. As mentioned in the Methods section, the parameters were considered to be linearly independent when $R^2<0.4$. Thus, *HAC* could not be used in combination with *PSsol* and *PSVdW*. The relationships established are listed in Table 4. The non-linear models B1–B4 of Table 4 were found to describe better the relationship between the molecular properties and the pharmacokinetic parameters. Such complex relationships are frequently encountered in quantitative modeling (Palm et al., 1997). *PSsol* was found to be incorporated with its inverse square value in the relationships for both clearance parameters, while an additional parabolic log *P* term was found significant for *CL*.

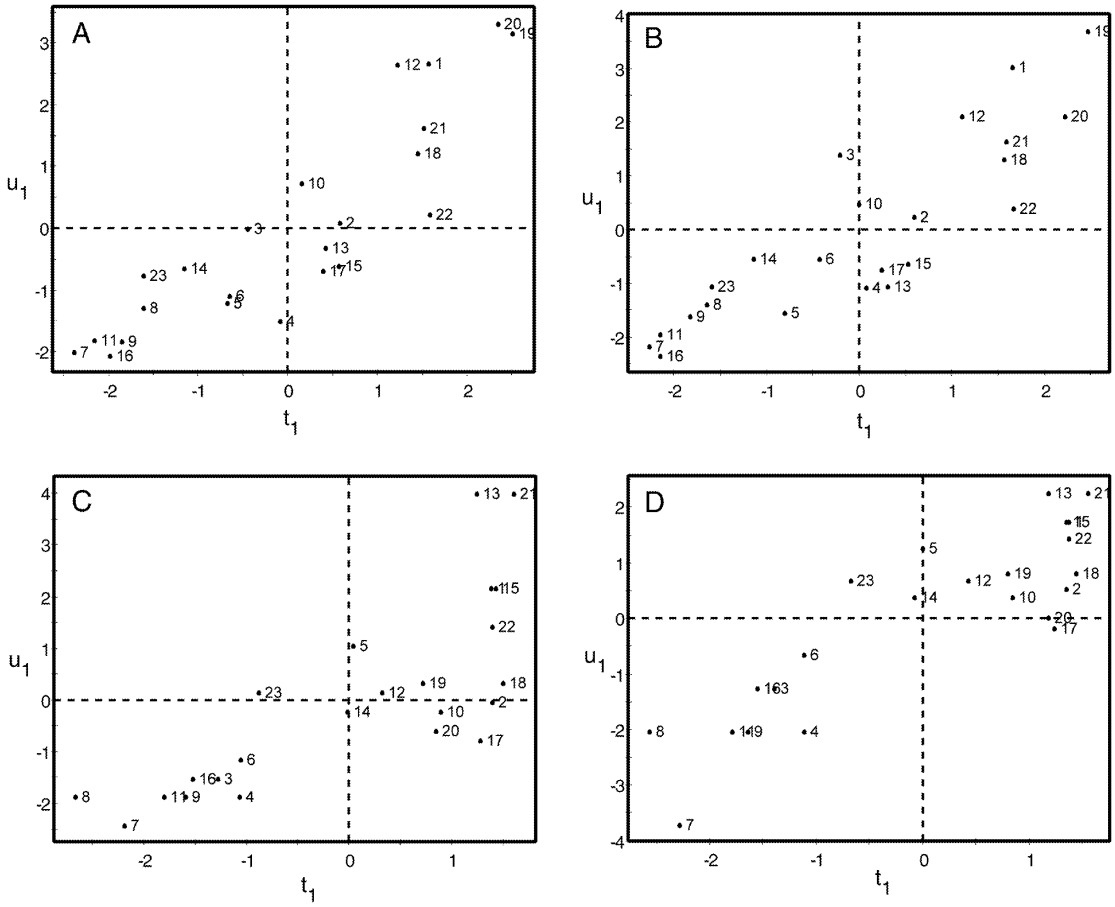

Fig. 2. The u_1 vs. t_1 plots showing the PLS inner relation for the models of: *CL*, (A), CL_f, (B), V_{ap}, (C), v_f, (D). u_1 and t_1 are the coordinates of the PK parameter and the descriptor matrix after PLS, respectively. The numbers represent the cephalosporins (see Table 1).

Replacement of *PSsol* with the number of hydrogen acceptors *HAC* led to similar models both for *CL* and CL_f with R^2 values equal to 0.724 and 0.641, respectively. This is a reasonable finding because *PSsol* is a polarity term encoding information of hydrogen bonding capability and a high degree of linear dependence was observed between *PSsol* and *HAC*.

Parabolic expressions were found for log *P* and f_b in the non-linear regression equations of V_{ap} and v_f (models B3, B4; Table 4). V_{ap} was found to be exclusively dependent on f_b. Although the negative sign of f_b in model B3 is a reasonable finding, the poor statistical properties ($R^2 =$ 0.571) make the validity of the V_{ap} model questionable. In contrast, satisfactory regression results were obtained for v_f, $R^2 = 0.824$ (model B4; Table 4). In this model apart from the negative contribution of f_b, lipophilicity was found to be an additional physicochemical descriptor contributing also in a negative way. The fitting results of the non-linear models to the experimental data are shown in Fig. 4.

The plots of the predicted values by models B1–B4 versus the observed values are shown in Fig. 5. A larger scattering is observed for the V_{ap} plot, Fig. 5C. It is worthy to mention that the data points corresponding to the highest V_{ap} values lie well below the line of complete concordance. This means that model B3 significantly underpredicts the higher V_{ap} values. The last observation should be considered in conjunction with the remarks regarding the validity of experimental V_{ap} values when the latter considerably exceed the plasma volume (Karalis et al., 2001). In the same vein, one should also note that the scattering of data points in the relevant v_f plot, Fig. 4D, is much less.

Some clarifications regarding the use of f_b in models A3, A4, and B3, B4 are required since f_b is not a molecular property but a pharmacokinetic parameter strongly related to drug's physicochemical characteristics. For this reason, attempts were made to find any possible relationship between f_b and the utilized descriptors. PLS analysis, using f_b as a dependent variable, showed that the best descriptive and predictive ability could be achieved when *MW*, *PSsol*, *HDO*, $fi_{(ac)}$, and $fi_{(ba)}$ were used as descriptors; the derived model was based on one principal component with R^2 and Q^2 values equal to 0.734 and 0.615, respectively. *MW* and *PSsol* contributed in a positive way, while the remaining parameters entered with a negative sign. In the case of multiple regression analysis, the best-significant models for f_b were observed when only *HDO* and *MW* were non-linearly expressed:

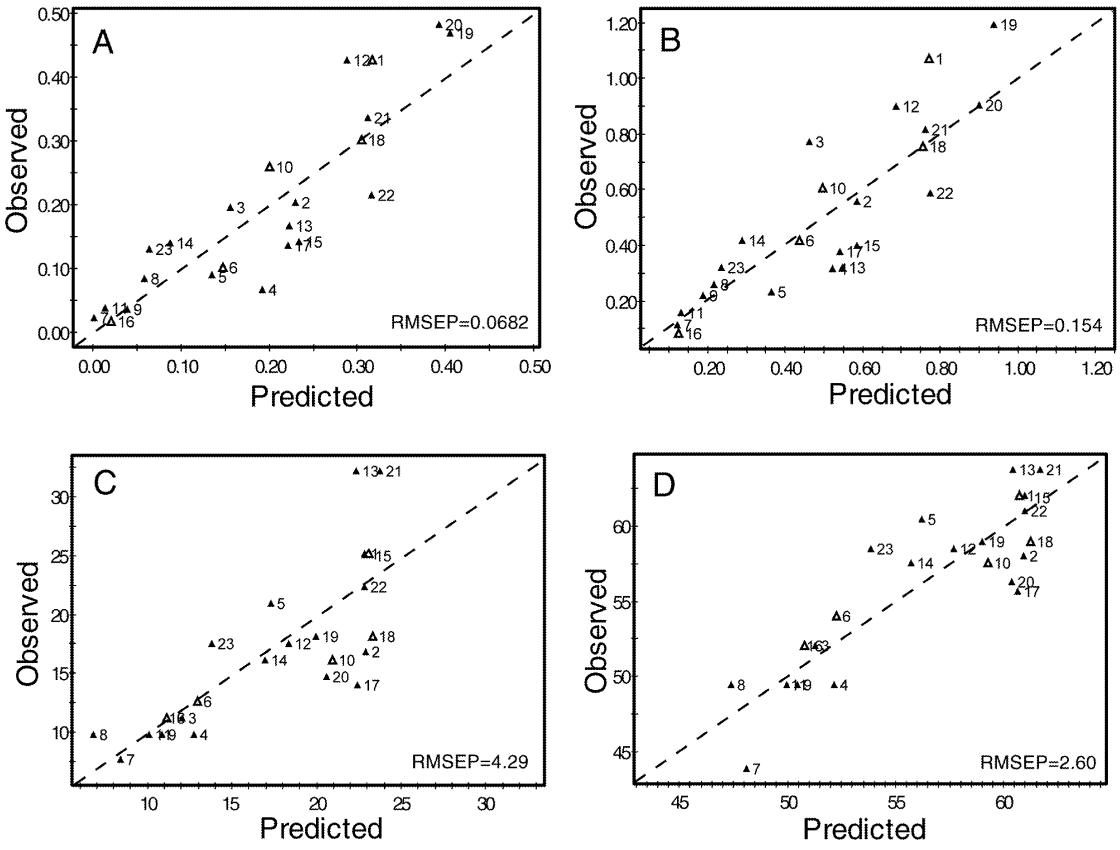

Fig. 3. Predicted versus observed values of *CL* (A), CL_f (B), V_{ap} (C), and v_f (D). Dashed line indicates complete concordance. Compounds belonging to training set are shown with solid triangles, while those of test set with open triangles. The term RMSEP denotes the root mean squared value for the prediction (Sheiner and Beal, 1981).

$$f_b = -109.8(\pm 13.3) + 0.4(\pm 0.1)\cdot\frac{1}{HDO} + 102.0(\pm 12.3)\cdot MW^{0.013} \quad (3)$$

$(R^2 = 0.789, \text{S.D.} = 0.132)$

Subsequently, regression models for V_{ap} and v_f were developed by replacing f_b with its relevant molecular descriptors. In both cases *MW* was found to contribute significantly. A linear combination of log *P* and *MW* led to a simple model for v_f with moderate statistics expressed with Eq. (4)

Table 4
Models derived after multiple linear and non-linear regression analysis for each PK parameter using the 22 cephalosporins (ceftazidime not included)

No.	Regression model	R^{2a}	S.D.[b]
Linear models (A)			
A1	$CL = 0.57(\pm 0.098) - 1.65\cdot 10^{-3}(\pm 0.00)\ (PSsol) - 3.63\cdot 10^{-2}(\pm 0.02)\ (\log P)$	0.703	0.080
A2	$CL_f = 1.65(\pm 0.17) - 4.60\cdot 10^{-3}(\pm 0.01)\ (PSsol)$	0.685	0.176
A3	$V_{ap} = 26.47(\pm 2.15) - 17.46(\pm 3.48)\ (f_b)$	0.523	4.752
A4	$v_f = 61.78(\pm 1.49) - 1.55(\pm 0.51)\ (\log P) - 12.41(\pm 2.17)\ (f_b)$	0.729	2.780
Non-linear models (B)			
B1	$CL = -0.079(\pm 0.036) + \frac{12332.1(\pm 1912.8)}{(PSsol)^2} + 0.0185(\pm 0.0058)\ (\log P)^2$	0.804	0.066
B2	$CL_f = \frac{32819.5(\pm 4254.6)}{(PSsol)^2}$	0.736	0.165
B3	$V_{ap} = 23.70(\pm 1.59) - 17.07(\pm 3.17)\ (f_b)^2$	0.571	4.609
B4	$v_f = 59.02(\pm 0.91) - 12.92(\pm 2.14)\ (f_b)^4 - 0.76(\pm 0.31)\ (\log P)^2 - 2.50(\pm 0.78)\ (\log P)$	0.824	2.291

[a] Coefficient of determination.
[b] Standard deviation.

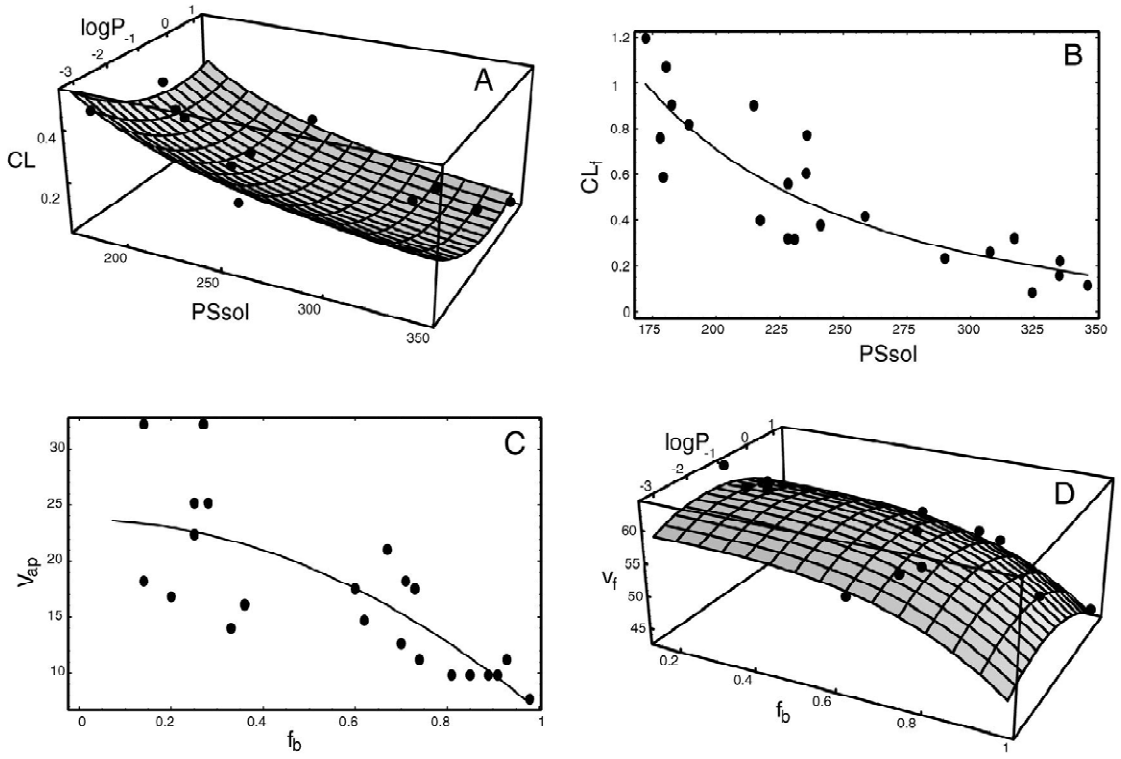

Fig. 4. Graphical representation of the non-linear regression models B1–B4 for *CL* (A), CL_f (B), V_{ap} (C), and v_f (D). Dots are the experimental data; curves or 3D surfaces correspond to the models.

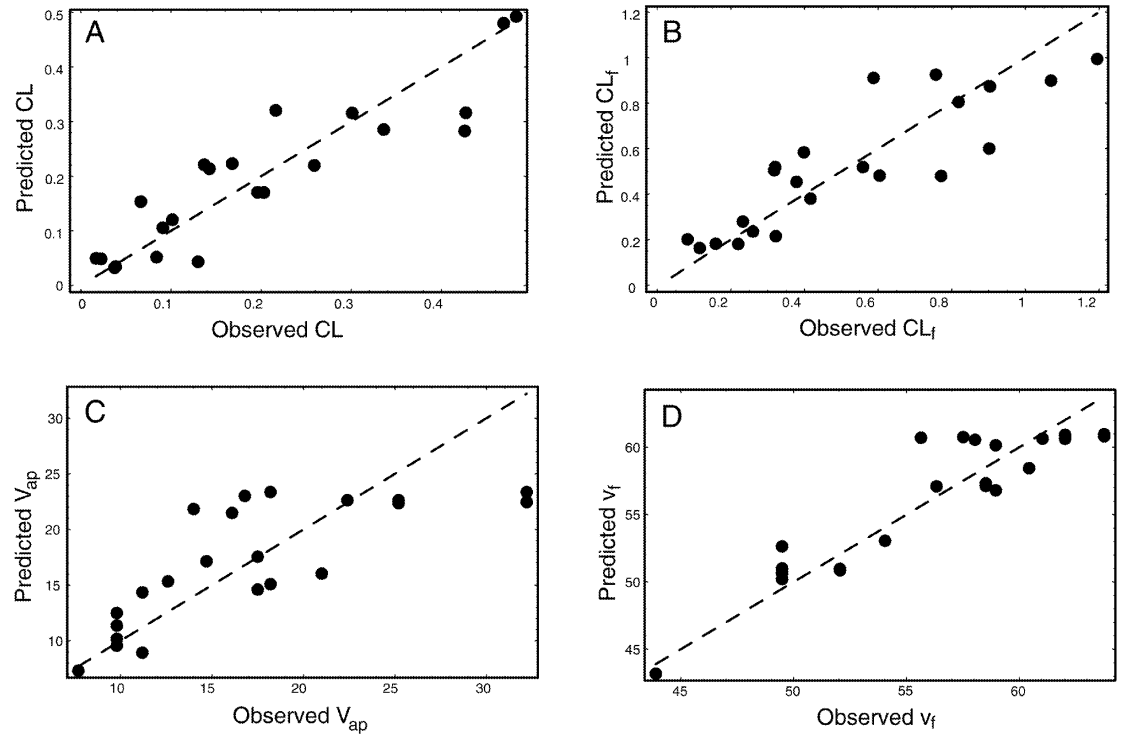

Fig. 5. Predicted versus observed values for *CL* (A), CL_f (B), V_{ap} (C), and v_f (D). The dashed line indicates complete concordance.

$$v_f = 73.6(\pm 4.9) - 1.42(\pm 0.65) \cdot \log P - 0.04(\pm 0.01) \cdot MW \quad (4)$$

($R^2 = 0.623$, S.D. = 3.35)

A significantly inferior model was obtained for V_{ap}, while the best model for V_{ap} was exclusively dependent nonlinearly on MW:

$$V_{ap} = 1236.5(\pm 288.1) - 1060.0(\pm 250.4) \cdot MW^{0.023} \quad (5)$$

($R^2 = 0.446$, S.D. = 5.24)

5. Conclusions

Within the congeneric series of the cephalosporins studied, the conventional and fractal clearance parameters led to satisfactory and analogous results with PLS as well as with multiple regression analysis. Regarding the volume parameters however, it was not possible to find adequate models for V_{ap}, although the volume values in this series of drugs lie within the physiological range (<70 l). In contrast, v_f could be successfully modeled by both PLS and regression analysis. This finding may further support the assumption that v_f constitutes a more suitable distribution parameter for QSPR studies.

Acknowledgements

This work was supported by the General Secretariat of Research and Technology of Greece (PENED Grant 70/3/6508).

References

Betageri, G., Rogers, J., 1989. Correlation of partitioning of nitroimidazoles in the *n*-octanol/saline and liposome systems with pharmacokinetic parameters and quantitative structure–activity relationships (QSAR). Pharm. Res. 6, 399–403.

Boobis, A., Gundert-Remy, U., Kremers, P., Macheras, P., Pelkonen, O., 2002. In silico prediction of ADME and pharmacokinetics. Report of an expert meeting organized by COST B15. Eur. J. Pharm. Sci. 17, 183–193.

Eriksson, L., Hermens, J., Johansson, E., Verhaar, H., Wold, S., 1995. Multivariate analysis of aquatic toxicity data with PLS. Aquat. Sci. 57, 217–241.

Eriksson, L., Johansson, E., 1996. Multivariate design and modeling in QSAR. Chemometr. Intell. Lab. Syst. 34, 1–19.

Esaki, T., 1987. Quantitative drug design studies. Quantitative structure–activity relationships of ionizable substances: antibacterial activities of phenols. Chem. Pharm. Bull. 35, 3105–3111.

Fouchécourt, M., Béliveau, M., Krishnan, K., 2001. Quantitative structure–pharmacokinetic relationship modeling. Sci. Total Environ. 274, 125–135.

Franke, R., Gruska, A., 1995. Principal component and factor analysis. In: van de Waterbeemd, H. (Ed.), Chemometric Methods in Molecular Design. VCH, Weinheim, pp. 113–158.

Gobburu, J., Shelver, W., 1995. Quantitative structure–pharmacokinetic relationships (QSPR) of beta blockers derived using neural networks. J. Pharm. Sci. 84, 862–865.

Hardman, J., Limbird, L., Molinoff, P., Ruddon, R., Gilman, A., 1996. Goodman and Gilman's the Pharmacological Basis of Therapeutics, 9th ed.. McGraw–Hill, New York.

Karalis, V., Claret, L., Iliadis, A., Macheras, P., 2001. Fractal volume of drug distribution: it scales proportionally to body mass. Pharm. Res. 18, 1056–1060.

Karalis, V., Macheras, P., 2002. Drug disposition viewed in terms of the fractal volume of distribution. Pharm. Res. 19, 697–704.

Karalis, V., Tsantili-Kakoulidou, A., Macheras, P., 2002. Multivariate statistics of disposition pharmacokinetic parameters for structurally unrelated drugs used in therapeutics. Pharm. Res. 19, 1829–1836.

Klopman, G., Stefan, L., Saiakhov, R., 2002. ADME evaluation. 2. A computer model for the prediction of intestinal absorption in humans. Eur. J. Pharm. Sci. 17, 253–263.

Kretz, O., Probst, A., 2001. High-throughput ADE screening. In: Testa, B., van de Waterbeemd, H., Folkers, G., Guy, R. (Eds.), Pharmacokinetic Optimization in Drug Research. Wiley, Weinheim, pp. 199–215.

Markin, R., Murray, W., Boxenbaum, H., 1988. Quantitative structure–activity study on human pharmacokinetic parameters of benzodiazepines using the graph theoretical approach. Pharm. Res. 5, 201–208.

Oprea, T., 2000. Property distribution of drug-related chemical databases. J. Comput. Aided Mol. Des. 14, 251–264.

Palm, K., Stenberg, P., Luthman, K., Artursson, P., 1997. Polar molecular surface properties predict the intestinal absorption of drugs in humans. Pharm. Res. 14, 569–571.

Poulin, P., Theil, F., 2000. A priori prediction of tissue:plasma partition coefficients of drugs to facilitate the use of physiologically-based pharmacokinetic models in drug discovery. J. Pharm. Sci. 89, 16–35.

Poulin, P., Schoenlein, K., Theil, F., 2001. Prediction of adipose tissue: plasma partition coefficients for structurally unrelated drugs. J. Pharm. Sci. 90, 436–447.

Qie, S., 1996. Drug distribution and binding. J. Clin. Pharmacol. 26, 583–586.

Sheiner, L., Beal, S., 1981. Some suggestions for measuring predictive performance. J. Pharmacokin. Biopharm. 9, 503–512.

Toma, E., 1989. Structure–activity relationship of quinolones. Clin. Invest. Med. 12, 7–9.

Urien, S., Tillement, J., Barré, J., 2001. The significance of plasma-protein binding in drug research. In: Testa, B., van de Waterbeemd, H., Folkers, G., Guy, R. (Eds.), Pharmacokinetic Optimization in Drug Research. Wiley, Weinheim, pp. 189–197.

Viswanadhan, V., Ghose, A., Revankar, G., Robins, R., 1989. Atomic physicochemical parameters for three-dimensional structure-directed quantitative structure–activity relationships. Additional parameters for hydrophobic and dispersive interactions and their application. J. Chem. Inf. Comput. Sci. 29, 163–172.

Wold, S., 1995. PLS for multivariate linear modeling. In: van de Waterbeemd, H. (Ed.), Chemometric Methods in Molecular Design. VCH, Weinheim, pp. 195–218.

Zhao, Y., Le, J., Abraham, M., Hersey, A., Eddershaw, P., Luscombe, C., Boutina, D., Beck, G., Sherborne, B., Cooper, I., Platts, J., 2001a. Evaluation of human intestinal absorption data and subsequent derivation of a quantitative structure–activity relationship (QSAR) with the Abraham descriptors. J. Pharm. Sci. 90, 749–784.

Zhao, Y., Abraham, M., Le, J., Hersey, A., Luscombe, C., Beck, G., Sherborne, B., Cooper, I., 2001b. Rate-limited steps of human oral absorption and QSAR studies. Pharm. Res. 19, 1446–1457.