Correlation of Human Jejunal Permeability (in Vivo) of Drugs with Experimentally and Theoretically Derived Parameters. A Multivariate Data Analysis Approach

Susanne Winiwarter,[†] Nicholas M. Bonham,[†] Fredrik Ax,[†] Anders Hallberg,[†] Hans Lennernäs,[‡] and Anders Karlén*,[†]

Department of Organic Pharmaceutical Chemistry and Department of Pharmacy, Uppsala Biomedical Centre, Uppsala University, SE-751 23 Uppsala, Sweden

Received March 2, 1998

The effective permeability (P_{eff}) in the human jejunum (in vivo) of 22 structurally diverse compounds was correlated with both experimentally determined lipophilicity values and calculated molecular descriptors. The permeability data were previously obtained by using a regional in vivo perfusion system in the proximal jejunum in humans as part of constructing a biopharmaceutical classification system for oral immediate-release products. pK_a, log P, and, where relevant, log P_{ion} values were determined using the pH-metric technique. On the basis of these experiments, log D values were calculated at pH 5.5, 6.5, and 7.4. Multivariate data analysis was used to derive models that correlate passive intestinal permeability to physicochemical descriptors. The best model obtained, based on 13 passively transcellularly absorbed compounds, used the variables HBD (number of hydrogen bond donors), PSA (polar surface area), and either log $D_{5.5}$ or log $D_{6.5}$ (octanol/water distribution coefficient at pH 5.5 and 6.5, respectively). Statistically good models for prediciting human in vivo P_{eff} values were also obtained by using only HBD and PSA or HBD, PSA, and CLOGP. These models can be used to predict passive intestinal membrane diffusion in humans for compounds that fit within the defined property space. We used one of the models obtained above to predict the log P_{eff} values for an external validation set consisting of 34 compounds. A good correlation with the absorption data of these compounds was found.

Introduction

Effective intestinal permeability (P_{eff}) is a fundamental parameter describing both rate and extent of intestinal drug absorption.[1,2] Due to the importance of predicting this property at an early stage of a drug development project[3,4] many investigations have been made to study drug permeability in cell cultures and animals and to correlate it with physicochemical properties of the drugs in question. It has been suggested that lipophilicity, molecular size, molecular shape, polar surface area, hydrogen bonding capacity, and similar parameters correlate to absorption or permeability.[5−12] Lipophilicity, often determined either as a pH-independent partition coefficient (P) or as a pH-dependent distribution coefficient (D), is probably the most frequently used parameter in correlations with membrane permeability.

Due to experimental difficulties, very few correlation studies have been performed using direct measurements of in vivo permeability of drugs and nutrients in the human intestine. However, determination of effective permeability in humans has recently become more readily accessible experimentally through the development of a regional jejunal perfusion system.[2,13] This experimentally validated approach gives a direct in vivo estimation of the local absorption rate P_{eff} (in cm/s) across the intestinal barrier. As part of constructing a biopharmaceutical classification system (BCS) for oral immediate-release products,[1] the human jejunal P_{eff} values for 22 compounds were determined using the above method. The aim of the BCS is to establish in vitro−in vivo correlations and to estimate oral drug absorption based on physiological relevant variables such as dissolution and permeability. This classification system is presently under consideration by the regulatory authorities.

The aim of the present study was to derive a quantitative structure−activity relationship (QSAR) equation which, based on previously determined human in vivo P_{eff} values and relevant physicochemical descriptors of the above set of compounds, will allow for the prediction of passive absorption of drugs in the human intestine. The lipophilicity variables log P and log D were carefully determined using a recently introduced pH-metric technique[14] and a number of theoretical molecular descriptors were calculated. To find relationships between P_{eff} and the physicochemical data, the PLS (projections to latent structures) method was applied.

Materials and Methods

Compound Data Sets. Two compound data sets were used in this study. Data set 1 consists of the 22 drugs shown in Figure 1 for which human permeability data (Table 1) have been determined as part of the BCS.[15−25] The in vivo P_{eff} values for the 22 drugs were obtained by using a technique based on single-pass perfusion of a human jejunum segment between two

[†] Department of Organic Pharmaceutical Chemistry.
[‡] Department of Pharmacy.

10.1021/jm9810102 CCC: $15.00
Published on Web 11/07/1998

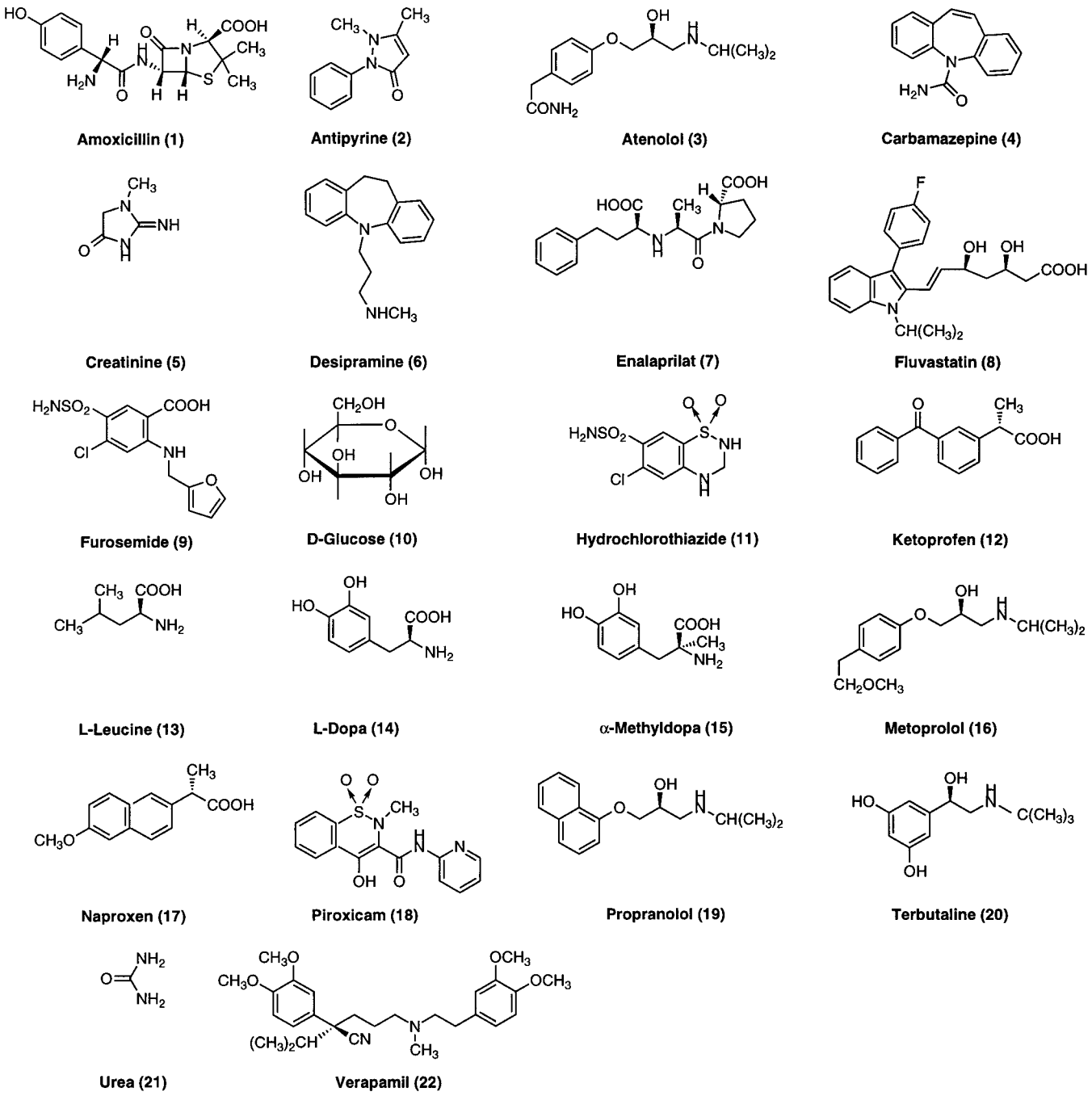

Figure 1. Chemical structures of drugs in data set 1.

inflated balloons.[2,13] For atenolol, fluvastatin, ketoprofen, metoprolol, propranolol, terbutaline, and verapamil the racemate was used in the experimental determination of P_{eff}. When the enantiomers of terbutaline[22] and verapamil[19] were studied, no difference in human permeability was observed.

Different routes of transportation across the intestinal membrane exist. It is generally assumed that sufficiently lipophilic compounds are transported by transcellular passive diffusion, while small (<200 Da), hydrophilic compounds are transported by diffusion and/or convection via the paracellular route,[26] if they do not use a carrier-mediated transport mechanism. The 22 compounds of data set 1 are transported across the epithelial barrier by at least one of these routes: amoxicillin, D-glucose, L-leucine, L-dopa, and α-methyldopa predominantly use various carrier transport systems across the enterocyte membrane.[15,27−29] Creatinine and urea, which have molecular weights much below 200, were assumed to be transported, at least to some extent, by the paracellular pathway.[26] The remaining 15 compounds, including verapamil, were considered to be passively transported across the intestinal membrane by the transcellular route. Verapamil, which is a well-known substrate for *P*-glycoprotein, was investigated at a high and clinically relevant luminal concentration, where the efflux transport was saturated and the passive transcellular diffusion is considered to be the predominant process.[19]

Data set 2 consists of the 22 compounds of the first data set combined with 136 compounds found in an Internet database of the Pomona College Medicinal

Table 1. Experimentally Determined Permeability Values, pK_a Values, Octanol/Water Partition, and Distribution Coefficients[a]

compound	P_{eff}[b]	log P_{eff}	pK_a value(s)[c]	log P[d]	log P_{ion}[e]	GOF[f]	log $D_{7.4}$[g]	log $D_{6.5}$[h]	log $D_{5.5}$[i]
amoxicillin·3H_2O	0.3 ± 0.4	−4.47	2.64/7.39/9.61	−1.71 ± 0.05	−1.22 ± 0.03 −1.56 ± 0.04 −1.98 ± 0.12	3.1	−1.6	−1.7	−1.7
antipyrine	4.5 ± 2.5	−3.35	1.44	0.56 ± 0.04		1.2	0.6	0.6	0.6
atenolol	0.2 ± 0.2	−4.70	9.60	0.18 ± 0.02		1.6	−2.0	<−2	<−2
carbamazepine	4.3 ± 2.7	−3.37		2.45[j]			2.45[j]	2.45[j]	2.45[j]
creatinine	0.3 ± 0.2	−4.52	4.85	<−2		0.5	<−2	<−2	<−2
desipramine·HCl	4.4 ± 1.8	−3.36	10.65	4.54 ± 0.01	1.05 ± 0.04	1.2	1.5	1.1	1.1
enalaprilat	0.2 ± 0.3	−4.70	1.25[k]/3.17/7.84	−0.13 ± 0.02	−1.07 ± 0.02 −0.99 ± 0.02	1.6	−1.0	−1.0	−1.0
fluvastatin Na	2.4 ± 1.8	−3.62	4.31	4.17 ± 0.01	1.12 ± 0.03	1.2	1.4	2.0	3.0
furosemide	0.05 ± 0.04	−5.30	3.34/10.46	2.53 ± 0.01	−1.03 ± 0.03	1.7	−0.9	−0.5	0.4
D-glucose[l]	10.0 ± 8.0	−3.00	nm[m]						
hydrochlorothiazide	0.04 ± 0.05	−5.40	8.84/10.07	−0.17 ± 0.01		1.6	−0.2	−0.2	−0.2
ketoprofen	8.4 ± 3.3	−3.08	3.89	3.37 ± 0.01	−0.34 ± 0.03	1.1	0.1	0.8	1.8
L-leucine	6.2 ± 2.9	−3.21	2.38/9.61	−1.55 ± 0.01	−2.07 ± 0.04 −1.58 ± 0.02	2.6	−1.6	−1.6	−1.6
L-dopa	3.4 ± 2.6	−3.47	2.21/8.77/9.81/12.73[k]	<−2	<−2	1.4	<−2	<−2	<−2
α-methyldopa	0.2 ± 0.06	−4.70	2.21/8.94/10.11/12.66[k]	<−2	<−2	1.3	<−2	<−2	<−2
metoprolol·$^1/_2$tartrate	1.3 ± 1.0	−3.89	9.60	2.07 ± 0.01	−0.63 ± 0.06	0.9	0.0	−0.5	−0.6
naproxen	8.3 ± 4.8	−3.08	4.01	3.58 ± 0.01	−0.22 ± 0.02	0.9	0.3	1.1	2.1
piroxicam[n]	7.8 ± 7.5	−3.11	nm[m]						
propranolol·HCl	2.9 ± 2.2	−3.54	9.52	3.43 ± 0.02	0.70 ± 0.05	0.9	1.4	0.9	0.7
terbutaline·$^1/_2$sulfate	0.3 ± 0.3	−4.52	8.72/10.00/11.10	−0.09 ± 0.01	−0.96 ± 0.06 −0.81 ± 0.05 −1.35 ± 0.04	3.3	−1.1	−1.3	−1.3
urea[l]	1.4 ± 0.4	−3.85	nm[m]						
verapamil·HCl	6.7 ± 2.9	−3.17	8.66	3.96 ± 0.01	0.96 ± 0.03	0.6	2.7	1.9	1.2

[a] The log P constants measured by the pH-metric technique are macroscopic; thus, where relevant, values obtained will relate to the partitioning of the sum of unionized and zwitterionic species. The ionic strength adjuster is 0.15 M KCl. The log P and log P_{ion} values are presented with the estimated standard deviation. [b] Effective permeability in 10^{-4} cm/s and estimated standard deviation between human subjects for each drug studied. It should be noted that some of the P_{eff} values given in this column differ from those originally reported (ref 15−25). This is due to a recent revision of the data set that included more human subjects. [c] Estimated standard deviations of the pK_a values derived from least-squares refinement are 0.01 for all compounds except furosemide (0.08), fluvastatine (0.02), and verapamil (0.04). [d] Logarithm of the partition coefficient measured in octanol/water. [e] Logarithm of the partition coefficient for the ionic species (if several ionic species partition several values are given). [f] GOF values refer to multisets of log P data. [g] Distribution coefficient in octanol/water at pH 7.4. [h] Distribution coefficient in octanol/water at pH 6.5. [i] Distribution coefficient in octanol/water at pH 5.5. [j] The log P value for carbamazepine was taken from literature (ref 43), and the log D values were given the same value, since the molecule lacks proteolytic groups. [k] pK_a value was extrapolated from less than 50% of the appropriate data points of the difference curve (Bjerrum plot). [l] The pK_a values of glucose (12.47) and urea (0.2) fall outside of the titration range, precluding lipophilicity measurements. [m] Not measured. [n] Piroxicam possesses very limited aqueous solubility, making pK_a measurement difficult even in water/cosolvent mixtures.

Chemistry Project[30] giving altogether 158 compounds. Four of the compounds in the database were already present in data set 1 and were not considered a second time. The database predominantly consists of drugs for which both structural (2D) information and calculated log P values (CLOGP) are supplied. The molecular weights of these compounds range from 100 to 550 Da and the CLOGP values from −2.7 to 6.7. We assumed that the compounds in this database are representative of the structural variation seen in the more common drugs. Thus, data set 2 was used in the molecular diversity study in order to ensure that the molecules in data set 1 are representative of drugs in general.

pK_a, log P, and log D Determination. The pH-metric technique was used in all pK_a, log P, and log D determinations.[14,31,32] Generally, when using this method only two types of titration are needed. The first titration takes place in aqueous solution and allows pK_a to be measured. For the second titration, octanol is added and an apparent pK_a (p$_o K_a$) is obtained (log P titration). The shift between the pK_a and p$_o K_a$ curves and the octanol-to-water ratio are used to calculate the log P value. Note: log P values below −2 measured by this method are not considered to be reliable, since the shift of the titration curves becomes too small to be measured accurately. The partitioning of ion pair species (described by log P_{ion}) may affect the magnitude of log D values. If ion pair partitioning is likely to occur, additional titrations are necessary. At least two titrations should be performed with different octanol−water ratios. Ideally, octanol-to-water ratios in each of the titrations should be as different as possible. Ion pair partitioning is dependent on the concentration of possible counterions, thus the ionic strength has to be considered. A 0.15 M KCl solution was used as ionic strength adjuster in all titrations. This is comparable to the ionic strength in humans. Partition coefficients obtained by using the pH-metric technique compare favorably to those obtained using the traditional shake-flask method.[33]

Distribution coefficients were calculated from pK_a, log P, and, where relevant, log P_{ion} values at three different pH values: (a) pH 6.5 which is the normal luminal pH in the proximal human jejunum[34] where the reference P_{eff} values were determined, (b) pH 5.5 which is found in the thin unstirred water/mucus layer adjacent to the intestinal wall,[35,36] and (c) pH 7.4 which was used for comparison since most log D values reported in the literature refer to this pH.

Calculated Molecular Descriptors. The 22 drugs in data set 1 were built in their neutral form in an extended conformation using SYBYL.[37] For chiral compounds, the stereoisomer drawn in Figure 1 was used in the theoretical calculations. For the molecules

of data set 2, the SMILES codes and CLOGP[38] values were extracted from the Medicinal Chemistry Project online database. The SMILES codes were processed using CONCORD within SYBYL to yield 3D coordinates. All structures were minimized with the AM1 method[39] as implemented in SYBYL using the keywords PRECISE, XYZ, and NOMM. From these calculations highest occupied molecular orbital and lowest unoccupied molecular orbital energies (E_HOMO and E_LUMO) and dipole moments (DM) were extracted. The size descriptors molecular weight (MW) and number of atoms (NATOM) were obtained using SYBYL spreadsheet functions. Molecular volume (V) and molecular surface area (S) were derived from a Connolly surface created with the MOLCAD module within SYBYL.[40] Hardness (H = (E_LUMO − E_HOMO)/2) and ovality ($O = S/[4\pi(3V/4\pi)^{2/3}]$) were computed using SYBYL spreadsheet functions.

Hydrogen-bonding capacity was estimated by counting the number of possible hydrogen bond donors and possible hydrogen bond acceptors.[6] This was done automatically with an SPL-script[40] using the SYBYL built-in set {*possible_hbond*}, which defines all atoms that can be part of a hydrogen bond and distinguishes between donors (HBD, number of hydrogens connected to N- and O-atoms) and acceptors (HBA, number of O- and N-atoms in an appropriate functional group). The sum of HBD and HBA was denoted HB. HBD can be determined easily while HBA is dependent on the atom type. We used CONCORD to assign SYBYL atom types for all molecules.

The polar surface area (PSA)[11,41,42] was calculated using the MOLCAD module within SYBYL (versions 6.1 and 6.2). The PSA was defined as the part of the surface area (as defined above) associated with oxygens, nitrogens, sulfur, and the hydrogens bonded to any of these atoms. An SPL script was written to automate these calculations.[40] However, for some molecules (amoxicillin, furosemide, and ketoprofen) the PSA could not be obtained using this script. In these cases the PSA was calculated by using MOLCAD interactively. CLOGP[38] values for the molecules in data set 1 were compiled from the drug compendium in Comprehensive Medicinal Chemistry,[43] except for fluvastatin.[7] The above values are based on calculations using CLOGP[44] versions 3.54 and 3.55.

Data Analysis. Multivariate data analyses were performed with SIMCA (version 6.1)[45] using default settings. Cross validation was used to quantify the predictive power of the PLS models.[46] The predictive capability of the PLS models was also ascertained by dividing the compounds into a training and a test set. Correlations between each variable and log P_{eff} (single correlations) were performed in SYBYL.

Results

The following strategy was applied to obtain models that can be used to predict passive absorption of drugs in humans from physicochemical data: (i) characterization of the physicochemical properties of the compounds in data set 1 with experimentally determined log *P* and log *D* values and theoretically calculated molecular descriptors; (ii) calculation of the theoretical molecular descriptors also for the compounds in data set 2 and performance of a principal component analysis (PCA) on all theoretical data in order to check the molecular diversity of the 22 compounds of data set 1; (iii) selection of a training and a test set of compounds from data set 1 according to statistical design principles based on the PCA above; (iv) investigation of the relationship between physicochemical variables and human in vivo permeability data of the training set compounds by PLS analysis; (v) evaluation of the resulting PLS models by use of the test set of compounds; (vi) calculating final models based on both test and training set compounds. The results of each of these steps are discussed below.

(i) Experimentally Determined Physicochemical Properties and Theoretically Calculated Molecular Descriptors of Data Set 1. Ionization constants were obtained from the potentiometric titrations in aqueous solution. These values were generally in good agreement with data previously reported.[43,47−49] Only creatinine, furosemide, and ketoprofen gave values which differed from the literature by more than 0.5 pK_a units. On individual titrations after refinement all data used gave good GOF ("goodness-of-fit")[14] of ca. 1 or less. After multiset refinement (where processed data from single titrations were grouped together and refined further) a GOF value of ca. 1 was obtained, suggesting that the pK_a values of the present study are reliable. The low solubility of furosemide in water made direct measurement of its pK_a difficult: the aqueous pK_a values were estimated from three measurements carried out in water/methanol mixtures by Yasuda−Shedlovsky extrapolation.[14,31] The very low solubility of piroxicam made measurements of its pK_a in water/cosolvent mixtures very difficult. No quality data was obtained for this compound, and consequently no data is presented.

The log *P* values were obtained after additional titrations in octanol/water mixtures. If log *P* was found to be above 1 or the compound could exist as a zwitterion, ion pair partitioning was determined with further titrations using different octanol/water ratios. The resulting multiset GOF values generally ranged from 0.5 to 1.7, indicating excellent overall log *P* refinements for each of the compounds (Table 1). Higher GOF values (up to 3.3) were obtained for only a few compounds, where ion pair partitioning was considered in the refinement (amoxicillin, L-leucine, and terbutaline). These high values are probably due to the high octanol/water ratios that were used in some of the titrations for these compounds.

Table 1 lists pK_a, log *P*, log P_{ion}, and log *D* values at pH 5.5, 6.5, and 7.4 for the substances in data set 1. pK_a and log *P* data could not be determined by the potentiometric method for carbamazepine, D-glucose, piroxicam, and urea. Carbamazepine possesses no ionizable groups so log *P* data was taken from literature[43] (see Table 1). pK_a values of D-glucose (12.47)[49] and urea (0.2)[43,47] fell outside of the practical titration range, precluding log *P* measurement with this method. For piroxicam, neither pK_a nor log *P* values were obtained due to its low aqueous solubility. Creatinine, L-dopa, and α-methyldopa were found to have log *P* and log *D* values below −2. For atenolol, log *D* values at

Table 2. Calculated Molecular Descriptors for Data Set 1

compound	CLOGP[a]	MW[b]	PSA[c]	HBA[d]	HBD[e]	HB[f]
amoxicillin	0.33	365.4	154.4	6	5	11
antipyrine	0.19	188.2	26.5	1	0	1
atenolol	−0.11	266.3	88.1	4	4	8
carbamazepine	1.98	236.3	41.6	1	2	3
creatinine	nd[g]	113.1	67.3	3	2	5
desipramine	4.09	266.4	16.2	1	1	2
enalaprilat	0.01	348.4	102.1	6	3	9
fluvastatin	3.24	411.5	81.4	5	3	8
furosemide	2.29	330.7	124.3	6	4	10
D-glucose	nd[g]	180.2	113.6	6	5	11
hydrochlorothiazide	−0.15	297.7	132.6	6	4	10
ketoprofen	2.79	254.3	54.2	3	1	4
L-leucine	−1.54	131.2	66.3	3	3	6
L-dopa	−2.92	197.2	105.5	5	5	10
α-methyldopa	−2.61	211.2	102.6	5	5	10
metoprolol	1.20	267.4	57.8	4	2	6
naproxen	2.82	230.3	48.2	3	1	4
piroxicam	nd[g]	331.3	90.9	6	2	8
propranolol	2.75	259.3	39.2	3	2	5
terbutaline	0.48	225.3	76.4	4	4	8
urea	−2.11	60.1	77.5	1	4	5
verapamil	3.53	454.6	64.4	6	0	6

[a] All CLOGP values were taken from the drug compendium in Comprehensive Medicinal Chemistry (ref 43) except the value for fluvastatin (ref 7). [b] Molecular weight. [c] Polar surface area. [d] HBA = number of possible hydrogen bond acceptor atoms. [e] HBD = number of possible hydrogen bond donor atoms. [f] HB = HBA + HBD. [g] No data (CLOGP was not obtained due to missing fragment values).

pH 5.5 and 6.5 were below −2. Since such values are not considered reliable, no exact values are given.

Ion pair partitioning was observed for seven of the monoprotic compounds, for one of the diprotic compounds, and for all compounds that can exist as zwitterions. The difference between the partition coefficients of the uncharged and ionic species of monoprotic compounds was about 3−4 log units, as was already observed in other studies.[50,51] Zwitterionic compounds did not have such a large difference between log *P* and the log P_{ion} values.[50] For amoxicillin two of the log P_{ion} values were even less negative than log *P* (see Table 1).

The most important calculated molecular descriptors for data set 1 can be found in Table 2 (the remaining data used in the multivariate data analysis is given in the Supporting Information). The molecular weight ranges from 60 to 455 Da (if only the 15 passively absorbed compounds are considered, the molecular weight range starts at 188 Da). CLOGP values ranged from −2.92 to 4.09. CLOGP values could not be obtained for creatinine, D-glucose, and piroxicam due to missing fragment values.

(ii) Molecular Diversity of the Compounds. The molecular diversity of data set 1 was analyzed by use of PCA. This analysis was performed using the 158 compounds of data set 2 and the 14 theoretical descriptors CLOGP, MW, NATOM, V, S, O, E_HOMO, E_LUMO, H, DM, HBA, HBD, HB, and PSA. The result of the analysis was a model with two principal components (PCs), which were found significant according to the eigenvalue test (a PC having an eigenvalue of more than 2 is considered significant).[45] These PCs describe 67% of the variance in the data set. The first PC described 40.4% and the second 26.6% of the variance. The third PC would account for only 11.7% of the variance and was not considered. The loading plot of the first two PCs (Figure 2a) shows that the size descriptors MW,

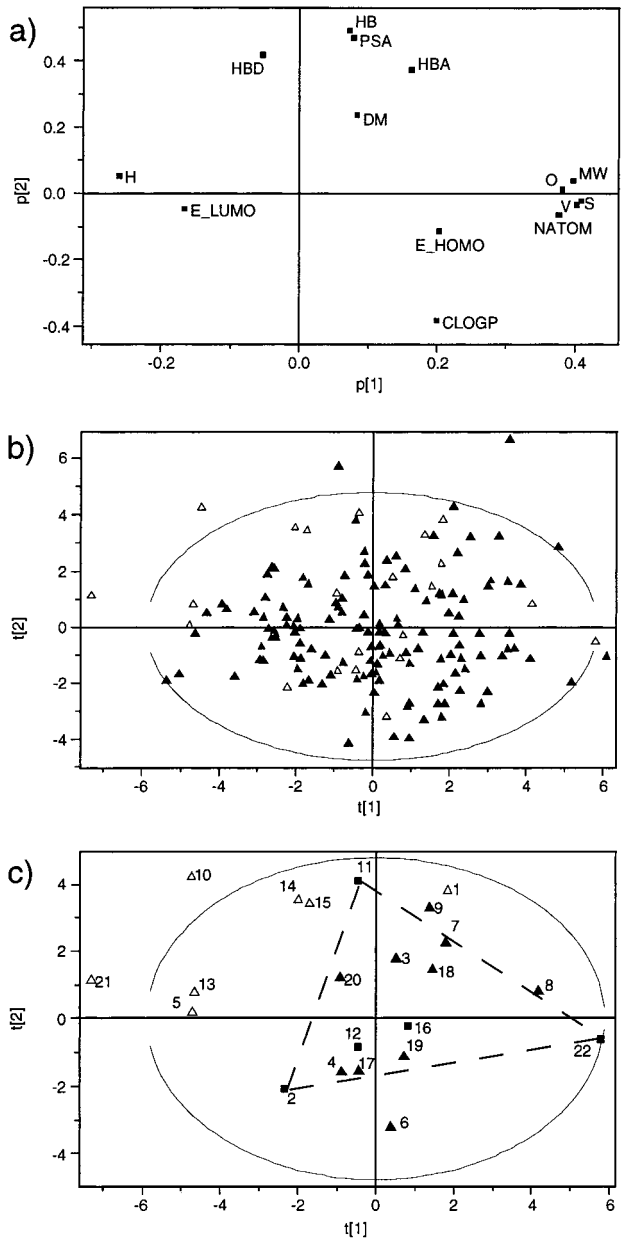

Figure 2. (a) Loading plot of p1 against p2 for data set 2. (b) Score plot showing t1 against t2 for data set 2. Compounds depicted with unfilled triangles correspond to data set 1. The ellipse corresponds to the confidence region based on Hotelling T^2 (0.05). (c) The same as in Figure 2b but only data set 1 is shown. Compounds depicted with unfilled triangles are actively transported or use, presumably, the paracellular pathway whereas compounds depicted with solid symbols are passively absorbed. Compounds depicted with a square correspond to the training set compounds. The ellipse corresponds to the confidence region based on Hotelling T^2 (0.05).

NATOM, V, and S contain similar information. The shape descriptor ovality can also be found in this cluster. The hydrogen-bonding descriptors HBA, HBD, and HB and the polar surface area (PSA) are found in the upper part of the loading plot, indicating that they contain similar information. CLOGP is found well separated from the other variables.

Figure 2b shows the score plot with all 158 compounds used for the PCA. The 22 compounds of data set 1 correspond to the unfilled triangles. These compounds are found to be reasonably well separated, implying that they are representative of drugs in general.

Table 3. PLS Models Derived from the Training Set Compounds[a]

variables[b]	no. of components[c]	Q^2	R^2	mean residuals[d]
18[e]	2	0.91	0.99	0.22
8[f]	2	0.95	1.0	0.33
HBD, PSA, HB	2	0.72	0.90	0.36
HBD, PSA[g]	1	0.80	0.88	0.26
HBD, PSA, log *P*	1	0.98	0.98	0.30
HBD, PSA, log $D_{5.5}$[h]	1	0.81	0.94	0.23
HBD, PSA, log $D_{6.5}$	1	0.75	0.90	0.23
HBD, PSA, CLOGP[i]	1	0.96	0.98	0.29

[a] Training set compounds: antipyrine, hydrochlorothiazide, ketoprofen, metoprolol, and verapamil. [b] Variables used in the PLS analyses. [c] Number of significant components obtained in the PLS analyses. [d] Calculated mean residual of the test set compounds (carbamazepine, desipramine, enalaprilat, fluvastatin, furosemide, naproxen, propranolol, and terbutaline). [e] MW, NATOM, S, V, O, PSA, HB, HBD, HBA, DM, E_HOMO, E_LUMO, H, CLOGP, log *P*, log $D_{7.4}$, log $D_{6.5}$, and log $D_{5.5}$. [f] HBD, PSA, HB, log *P*, log $D_{5.5}$, CLOGP, log $D_{6.5}$, and DM. [g] Model 2a. [h] Model 1a. [i] Model 3a.

(iii) Selection of Training and Test Sets. A training set of compounds was selected based on statistical design principles using the PCA described above. In Figure 2c the score plot of the 22 molecules of data set 1 (the molecules depicted as unfilled triangles in Figure 2b) is shown. When the actively transported compounds (**1**, **10**, **13**, **14**, and **15**) and the two compounds which may use the paracellular route (**5** and **21**) are omitted, the remaining compounds (depicted as filled squares and triangles) are distributed in a triangular fashion. We selected the three corner compounds of the triangle (antipyrine (**2**), hydrochlorothiazide (**11**), and verapamil (**22**)) and two compounds in the center (ketoprofen (**12**) and metoprolol (**16**)) for the training set (marked as squares in Figure 2c). Such a design is referred to as a truncated factorial design in two variables (geometrically a triangle).[52] All parameters were available for the compounds of the training set. The remaining passively transported compounds for which all data existed were used for the evaluation. Thus, we obtained a test set consisting of eight molecules (**4**, **6**, **7**, **8**, **9**, **17**, **19**, and **20**). We did not include the passively absorbed compounds **3** and **18** in the test set since data were lacking for these compounds.

(iv) PLS Analysis of Training Set Compounds. A PLS analysis using the 18 available descriptors (the 14 theoretical descriptors used in the PCA together with log *P*, log $D_{5.5}$, log $D_{6.5}$, and log $D_{7.4}$) and the five compounds of the training set resulted in a two-component model with $Q^2 = 0.91$ and $R^2 = 0.99$ (see Table 3). A $Q^2 > 0.5$ indicates that the PLS model has a predictive capability better than chance.[45] To obtain a more simple model we aimed to reduce the number of variables. We started by selecting the most significant variables according to the variable importance in the projection (VIP) plot available in SIMCA (Figure 3). The VIP value gives an indication of the relative importance of the variables (*x*-variables) for explaining log P_{eff} (*y*-value). Terms with VIP > 1 have an above average influence on log P_{eff}.[45] To evaluate the quality of the resulting models we compared both the Q^2 and R^2 values and calculated the mean residual of the test set compounds for each model.

Inspection of the VIP plot (Figure 3) shows that eight of the variables have a VIP value above 1. A PLS analysis using these eight variables gave a two-component model with $Q^2 = 0.95$ and $R^2 = 1.0$ (Table 3). Taking into account only the parameters with a VIP value above 1.2 (HBD, PSA, and HB), we obtained a two component model with $Q^2 = 0.72$ and $R^2 = 0.90$. By eliminating HB, a one component model was obtained with $Q^2 = 0.80$ and $R^2 = 0.88$. Including instead log *P*, the most important lipophilicity variable according to the VIP plot results in a one-component model with $Q^2 = 0.98$ and $R^2 = 0.98$. We also tried several of the other lipophilicity parameters together with PSA and HBD and found that CLOGP gave results similar to log *P*, whereas using log *D* values resulted in lower Q^2 and R^2 values. Several other combinations of variables were also tested but did not give better results (data not shown). We also observed that the PLS models did not improve if several closely related parameters were taken into account at the same time, e.g., log $D_{5.5}$ and log $D_{6.5}$.

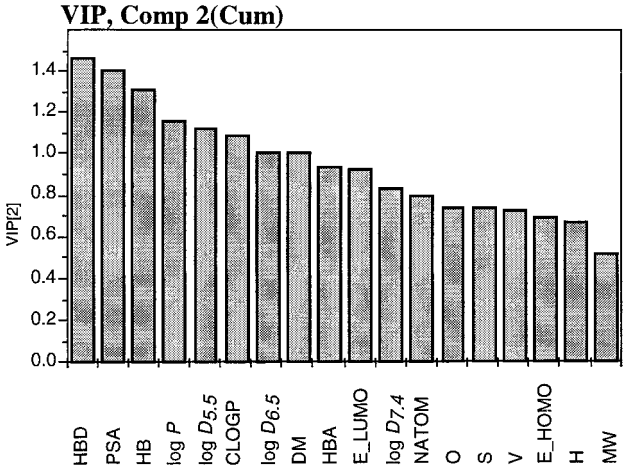

Figure 3. VIP plot from PLS analysis using the 18 available descriptors and the five compounds of the training set.

Table 4: Q^2 and R^2 Values Obtained from Correlations between log P_{eff} and Each Individual Variable[a]

variable	Q^2	R^2
MW	−0.47	0.02
V	−0.32	0.02
S	−0.33	0.02
NATOM	−0.26	0.06
O	−0.35	0.02
log *P*	0.23	0.42
log $D_{7.4}$	0.31	0.47
log $D_{6.5}$	0.36	0.52
log $D_{5.5}$	0.22	0.41
CLOGP	0.12	0.36
E_HOMO	−0.21	0.34
E_LUMO	−0.43	0.09
H	−0.19	0.07
DM	0.11	0.37
PSA	0.68	0.76
HBA	0.16	0.41
HBD	0.63	0.75
HB	0.55	0.70

[a] The passively transported compounds **2**, **4**, **6**, **7**, **8**, **9**, **11**, **12**, **16**, **17**, **19**, **20**, and **22** were considered.

To understand the importance of each variable we calculated the linear correlation coefficients between log P_{eff} and each of the 18 variables individually using the 13 passively transported compounds for which all data was available, i.e., the compounds of both the training and the test set (see Table 4). The best correlation was obtained with PSA ($Q^2 = 0.68$, $R^2 = 0.76$; see also Figure 4a). Good correlations were also found with HBD (Figure 4b) and HB ($Q^2 > 0.5$, $R^2 \geq 0.7$). Correlations with the experimentally derived lipophilicity descriptors

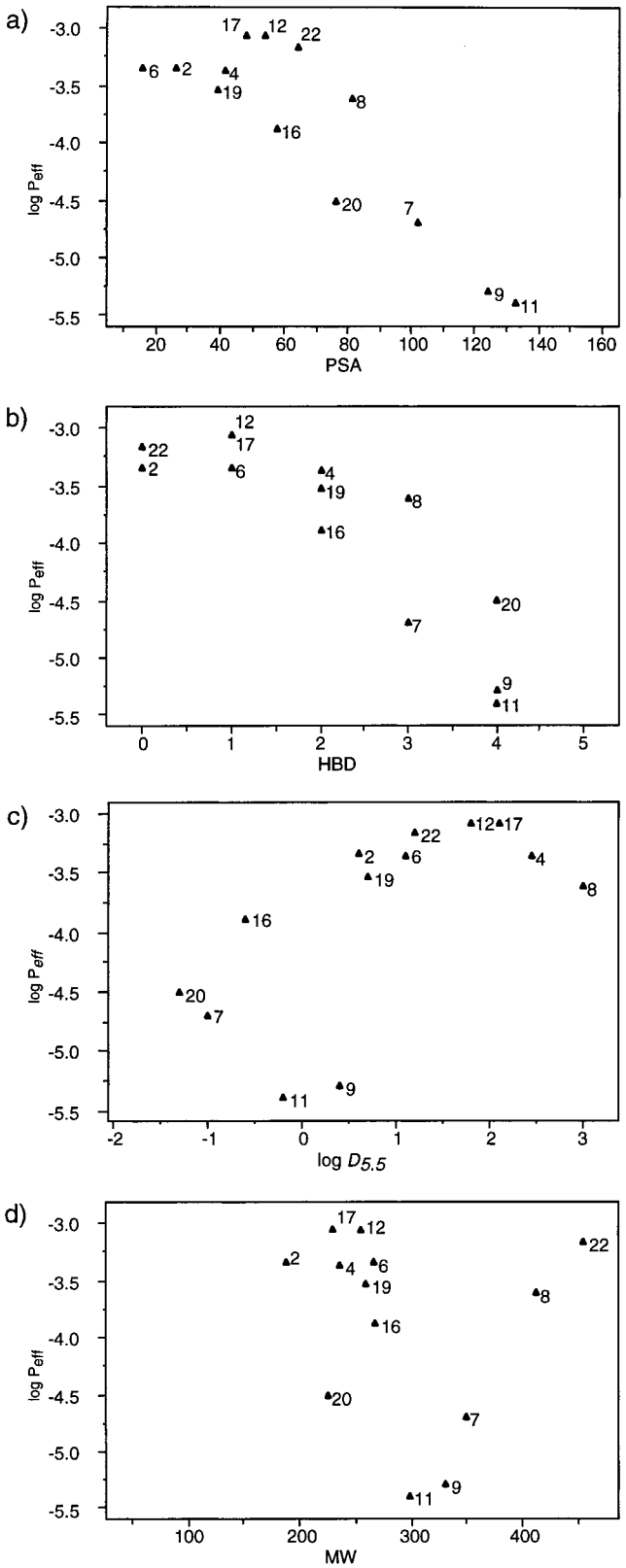

Figure 4. Scatter plots of log P_{eff} and four of the molecular descriptors (only the 13 passively absorbed compounds are shown; see also Table 4).

had markedly lower Q^2 values (about 0.3). In Figure 4c a near parabolic relationship between log P_{eff} and log $D_{5.5}$ can be discerned if the sulfonamides furosemide (**9**) and hydrochlorthiazide (**11**) are not considered.[53] No correlations (Q^2 negative and R^2 below 0.2) were found with H, E_LUMO and the molecular size descriptors (see Figure 4d for the plot of log P_{eff} against MW). Since the same parameters (PSA, HBD, and lipophilicity variables) that were found to be important in the PLS models also gave good correlations on their own, we assumed that we could exclude the other parameters from our models.

(v) Evaluation of Resulting PLS Models Using the Test Set Compounds. The most interesting models obtained above were evaluated by predicting log P_{eff} of the test set compounds and determining the mean residual for each model (Table 3). The PLS model using all 18 variables gave the lowest mean residual (0.22) of the evaluated models but was not considered further due to its complexity. The models using HBD, PSA, and either log $D_{5.5}$ or log $D_{6.5}$ yielded nearly as low mean residuals (0.23). These two models seem to be similar which can also be seen when the residuals of each compound of the test set are compared (data not shown). Due to its higher Q^2 value, we selected the model using HBD, PSA, and log $D_{5.5}$ as model 1a. The model using HBD and PSA which gives a reasonable mean residual (0.26), despite its relatively low Q^2 value, was chosen as model 2a. The models with the highest Q^2 values using HBD, PSA, and either log P or CLOGP gave surprisingly high mean residuals. Since both models were similar and CLOGP is an easily obtained quantity, we used it in model 3a.

The plots of predicted versus observed log P_{eff} values for the three selected models are shown in Figure 5a–c. Predictions were performed for all 22 compounds of data set 1 provided that the necessary data were available. Thus, molecules for which log $D_{5.5}$ values were missing were not predicted with model 1a, and for molecules lacking the CLOGP value predictions were not performed using model 3a. One can see that the actively transported molecules **1**, **10**, **13**, **14,** and **15** are transported faster than the models predict (i.e., they have a predicted log P_{eff} value which is lower than the observed value). Urea (**21**), which has a low molecular weight and is therefore believed to use the paracellular pathway (at least partially), is also transported faster than is predicted by the models. However, creatinine (**5**), which also has a low molecular weight, is predicted too high in model 2a. This difference might be due to a significant efflux of creatinine (**5**) back into the jejunum.

Most of the remaining molecules, which were transported by passive diffusion, were predicted well, with absolute residuals less than 0.5. One exception is fluvastatin (**8**), which has a large residual in model 2a (0.80), whereas a better prediction is obtained when including a lipophilicity value as in model 1a or 3a. Piroxicam (**18**), which could only be predicted in model 2a, also shows a high residual (1.0). Desipramine (**6**) has the highest residual in model 3a (−0.67), whereas for terbutaline (**20**) a residual slightly above 0.5 was found in model 1a.

(vi) Final Model Development. Final models were calculated using both the training and the test set molecules (n = 13) and the variables from models 1a-3a, respectively. The PLS analyses resulted in a higher Q^2 for model 1b (Q^2 = 0.90, R^2 = 0.93) than for model 1a (Q^2 = 0.81, R^2 =0.94). Q^2 and R^2 were about the same for model 2a (Q^2 = 0.80, R^2 = 0.88) and model 2b (Q^2 = 0.82, R^2 = 0.85). In model 3b both Q^2 = 0.85 and R^2 = 0.88 were found to be lower than in model 3a (Q^2 = 0.96, R^2 = 0.98). The following equations were

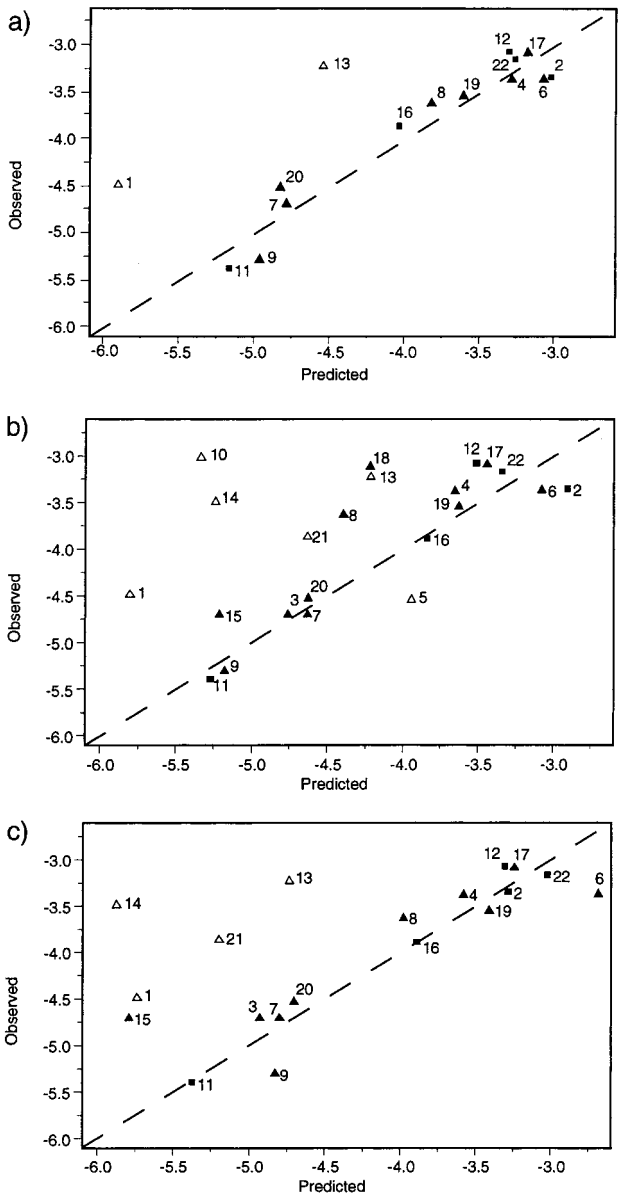

Figure 5. Correlations between observed and calculated log P_{eff} values. Predictions were only performed for compounds for which all necessary data was available. Compounds depicted with unfilled triangles are actively transported or use, presumably, the paracellular pathway whereas compounds depicted with solid symbols are passively absorbed. Compounds depicted with a square correspond to the training set compounds. (a) Model 1a (QSAR eq: log $P_{eff} = -2.980 - 0.009$ PSA $+ 0.299 \log D_{5.5} - 0.236$ HBD). (b) Model 2a (QSAR eq: log $P_{eff} = -2.607 - 0.011$ PSA $- 0.287$ HBD). (c) Model 3a (QSAR eq: log $P_{eff} = -3.061 + 0.190$ CLOGP $- 0.010$ PSA $- 0.246$ HBD).

obtained:

$$\log P_{eff} = -2.883 - 0.010 \text{ PSA} + 0.192 \log D_{5.5} - 0.239 \text{ HBD} \quad \text{(model 1b)}$$

$$\log P_{eff} = -2.546 - 0.011 \text{ PSA} - 0.278 \text{ HBD} \quad \text{(model 2b)}$$

$$\log P_{eff} = -3.067 + 0.162 \text{ CLOGP} - 0.010 \text{ PSA} - 0.235 \text{ HBD} \quad \text{(model 3b)}$$

Discussion

The extent of intestinal drug absorption, most often described by the fraction of drug absorbed (Fa), is governed by several different processes: dose/dissolution ratio, chemical degradation and/or metabolism in the lumen, complex binding in the lumen, intestinal transit, and effective permeability across the intestinal mucosa. In many cases P_{eff}, which was shown to be related to Fa by a sigmoidal function,[25] is considered to be the rate-limiting step in the overall absorption process. Therefore, it is of special interest to identify those chemical descriptors that might be used to predict P_{eff}.

We have calculated molecular descriptors for the 22 compounds in data set 1 and have also determined the pK_a values for 18 and log P values for 15 of the compounds by use of potentiometric titrations. Using PLS we found that a good estimate of passive intestinal permeability in human jejunum (in vivo) could be obtained using the theoretical descriptors HBD and PSA alone (model 2a) or combined with a lipophilicity descriptor. The best relationship was found using either log $D_{5.5}$ or log $D_{6.5}$ together with HBD and PSA. The model using log $D_{5.5}$ (model 1a) was only marginally better. Thus, we cannot conclude based on this study if pH 5.5, corresponding to the pH of the microenvironment in the unstirred water layer adjacent to the intestinal wall, or pH 6.5, corresponding to the pH of the lumen, is more appropriate for modeling the effective permeability. Nevertheless, using log D values gives better results than using the pH-independent log P or CLOGP. The two models combining one of the two latter descriptors with HBD and PSA were found to be remarkably similar. Since CLOGP values are more easily obtained, the model using this parameter seemed more attractive (model 3a).

Molecular weight is generally considered to be an important parameter when predicting membrane permeability. However, we did not find any correlation between MW and log P_{eff} for the molecules in this study ($R^2 = 0.0$; see Table 4 and Figure 4d), nor did the use of molecular weight in the multivariate analysis improve our models, in contrast to what has been observed in other studies.[12,54] One reason for the lack of correlation may be that the molecular weight range of the investigated molecules (188−455 Da) was too narrow.

Two parameters (HBD and PSA) were found to have R^2 values above 0.7 when correlated on their own to log P_{eff} (see Figure 4a,b). The dynamic PSA was previously shown to correlate with permeability determined in the Caco-2 cell model (in vitro) and with Fa in humans.[11,41] In the present study we did not perform a conformational analysis of the compounds and thus did not generate the Boltzmann-averaged (dynamic) PSA. However, even though we only used one conformation, PSA was one of the most important parameters for predicting permeability.

All our correlations and models assume that a linear relationship exists between permeability and the investigated parameters.[53] A sigmoidal relationship has been observed between different molecular descriptors and permeability measured in Caco-2 cells.[12] Inspection of Figure 4a,b indicates that a sigmoidal relationship can also be present in our data. However, no definite conclusions can be made concerning the type of relationship. A linear instead of a sigmoidal relationship can be due to the minor importance of the paracellular route in vivo for low permeability drugs with a size larger

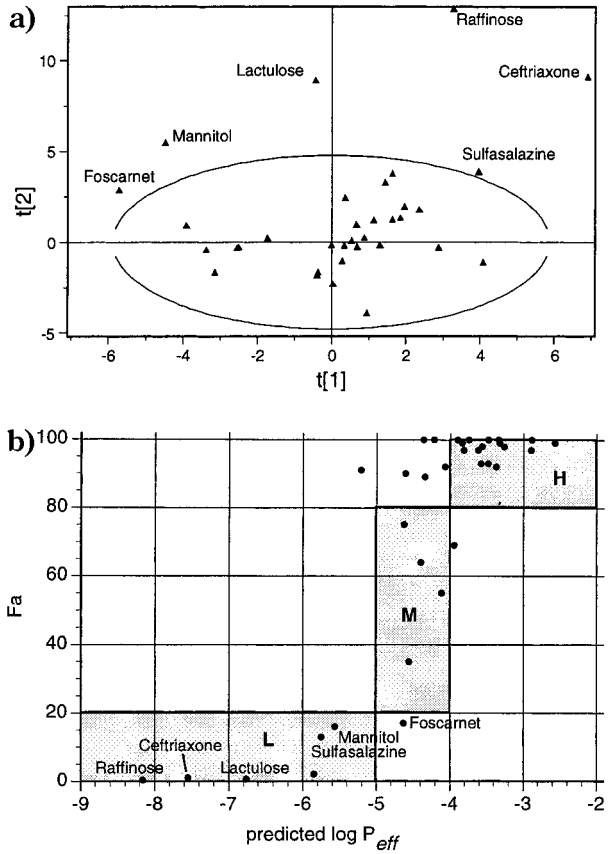

Figure 6. (a) Score plot showing t1 against t2 for the external validation set (see Figure 2 for comparison). The ellipse corresponds to the confidence region based on Hotelling T^2 (0.05). (b) Scatter plot of Fa and predicted log P_{eff} values. The gray rectangles correspond to low (L), medium (M), and high (H) absorption regions.

than 200 Da, and the unstirred water layer is not an absorption barrier for high permeability drugs in vivo.[26,55]

Due to its simplicity we believe model 2b (PSA and HBD) to be very interesting and we have therefore used it to predict the log P_{eff} values for a set of 34 compounds for which the Fa values after oral administration to humans have been published.[41,56,57] Ten of these compounds were also present in data set 2. To examine if the 34 compounds fit into the property space defined by data set 2 we calculated their molecular descriptors and fitted them into the previously described PCA model (Figure 6a). Six compounds were found outside the confidence limits of the model. We then calculated the log P_{eff} values for the 34 compounds based on model 2b. These compounds were classified into low (Fa < 20%, log P_{eff} < −5), medium (Fa = 20 to 80%, log P_{eff} = −5 to −4), and high (Fa > 80%, log P_{eff} > −4) absorption drugs according to both Fa and predicted log P_{eff}. Figure 6b shows that most compounds are classified correctly by log P_{eff}. This figure also shows the expected sigmoidal relationship between Fa and log P_{eff}.[25] The main outlier in this plot corresponds to cephalexin which is absorbed to 91% but was predicted to have a low log P_{eff} value (<−5). This compound is considered to be actively transported[57] which likely explains the poor prediction. It should be noted that other drugs in this data set may also use different transport routes.

In conclusion, three models with good statistics have been developed that can be used for predicting passive intestinal absorption in vivo in humans. The choice of which model to use depends on the availability of descriptor data. It should be remembered that even though the three models seem to have good predictive capabilities, they are based on only 13 compounds. If more in vivo data were available, the equations and combinations of variables might look different. It should also be kept in mind that the models are not applicable for predicting effective permeability for peptides, polysaccharides, or other compounds which do not fit into the defined property space.

Experimental Section

pK_a, log *P*, and log *D* Determination. The Sirius PCA101 (Sirius Analytical Instruments Ltd., Forest Row, U.K.) computerized titration instrument was used to carry out the pK_a and log *P* assays (25 ± 1 °C, under an argon stream to minimize CO_2 absorption at high pH). The operation of this instrument has been described comprehensively by Avdeef and co-workers.[14,31,32]

HCl and KOH solutions were made using standardized ampules (Merck, Darmstadt, Germany). Partition-coefficient grade *n*-octanol was obtained from Sigma (Poole, U.K.). General chemicals used in this study were obtained from Sigma (Poole, U.K.), and the compounds in data set 1 were obtained from various drug companies.

More than 200 titrations were performed. The majority of the titrations were alkalimetric, and sample concentrations ranged from 0.4 to 29.8 mM in 0.15 M KCl. The more concentrated sample solutions were used for compounds with high solubilities and where a greater titration range was required (pH < 3 and pH >11). Each titration was performed on a freshly prepared sample.

For furosemide, which was insufficiently water soluble for its aqueous pK_a values to be determined directly, the ionization constants were obtained by Yasuda−Shedlovsky analysis. Three titrations were performed in the presence of methanol (10%, 20%, and 30% v/v, respectively), and the aqueous pK_a values were extrapolated.[14,31]

For log *P* titrations the ratios of octanol/water were optimized and varied from 1:80 for lipophilic compounds to 25:1 for hydrophilic compounds. Most titrations were carried out in standard 20 mL vials. For the more hydrophilic compounds, large ratios of octanol/water were used. This was possible using the Sirius lipophilicity beaker apparatus: a special beaker with ca. 100 mL capacity and a separate special glass sleeve which allowed the temperature of the titration to be controlled.

Ionization constants were estimated from Bjerrum plots and were refined by a nonlinear least squares procedure.[14,32] The log *P* values were calculated from the shifts of the titration curves that occurred when octanol was present. All pK_a and log *P* values listed were based on multiset refinement of at least two high-quality titrations.

The partition coefficients for ion pairs (log P_{ion}) were determined for those substances in which the measured log *P* value for the uncharged species was above 1. Initially, two titrations were performed at markedly different octanol/water ratios (usually 3:1 and 1:80). These ratios and the apparent pK_a values (p$_o K_a$) at each ratio were input into an automated equation ("ion-pair") on the Sirius instrument, and the magnitude of log P_{ion} was calculated. If log P_{ion} was observed, up to four titrations were performed with a range of ratios. The value for log P_{ion} was introduced into the partitioning model as a fixed contribution. Refinement of the log *P* value was carried out on individual data sets and then in a multiset calculation. The log P_{ion} value was then refined in the multiset calculation.

Ion pair partitioning was also investigated for compounds that can exist as zwitterions (amoxicillin, enalaprilat, L-

leucine, L-dopa, α-methyldopa, and terbutaline). Up to four titrations were performed for each of the above compounds with large octanol/water ratios (up to 25:1) to maximize the shifts observed with the titration curves obtained in the presence of octanol, relative to the titrations carried out in aqueous solution alone. The mode of partitioning, i.e., which species partition, was elucidated using the Sirius "twelve-case chart",[58] and the appropriate partition equations were proposed in the refinement model. The data sets were refined individually, and the log P and log P_{ion} values were refined further in multiset calculations. In some cases, e.g., terbutaline, all possible species partition (four species) and thus there is one more partition equation than ionization equation. In such a case, one of the log P_{ion} values was added as a fixed contribution (fixed at an appropriate approximate value, e.g., −2), and the remaining log P and log P_{ion} values were refined in the individual sets. These log P and log P_{ion} values were refined further in a multiset calculation, and the last remaining log P_{ion} was obtained by refinement in this multiset.

The distribution coefficients at pH 5.5, 6.5, and 7.4 in octanol/water were computed from the pK_a, log P, and log P_{ion} values using eqs 1−3 for monoprotic, diprotic, and triprotic compounds, respectively.[50,59] All equations can be used for both acids and bases. P_X denotes the partition coefficient of the deprotonated species, i.e., in eq 1 it is equal to the partition coefficient P in the case of a base and equal to the ion pair partition coefficient P_{ion} for an acid, whereas P_{XH} denotes the partition coefficient of the protonated species (in eq 1 it is P_{ion} for a base and P for an acid).

$$\log D = \log(P_X + P_{XH}10^{pK_a-pH}) - \log(1 + 10^{pK_a-pH}) \quad (1)$$

$$\log D = \log[P_X/(1 + 10^{pK_{a2}-pH}(1 + 10^{pK_{a1}-pH})) + P_{XH}/(1 + 10^{pK_{a1}-pH} + 10^{pH-pK_{a2}}) + P_{XH_2}/(1 + 10^{pH-pK_{a1}}(1 + 10^{pH-pK_{a2}}))] \quad (2)$$

$$\log D = \log[P_X/(1 + 10^{pK_{a3}-pH}(1 + 10^{pK_{a2}-pH}(1 + 10^{pK_{a1}-pH}))) + P_{XH}/(1 + 10^{pH-pK_{a3}} + 10^{pK_{a2}-pH}(1 + 10^{pK_{a1}-pH})) + P_{XH_2}/(1 + 10^{pK_{a1}-pH} + 10^{pH-pK_{a2}}(1 + 10^{pH-pK_{a3}})) + P_{XH_3}/(1 + 10^{pH-pK_{a1}}(1 + 10^{pH-pK_{a2}}(1 + 10^{pH-pK_{a3}})))] \quad (3)$$

Acknowledgment. We thank The Swedish Research Council for Engineering Sciences (TFR) and the Swedish Medical Research Council (MFR) for financial support. The authors express their gratitude to Dr. Lennart Eriksson and Dr. Erik Johansson (UMETRI), Dr. John Comer (Sirius Analytical Instruments), and Dr. Anders Sokolowski (Pharmacia&UpJohn) for many valuable discussions. Mr. Wes Schaal is acknowledged for critical reading of the manuscript.

Supporting Information Available: Data used in the multivariate data analysis for data set 1 (1 page). Ordering information is given on any current masthead page.

References

(1) Amidon, G. L.; Lennernäs, H.; Shah, V. P.; Crison, J. R. A Theoretical Basis for a Biopharmaceutic Drug Classification: The Correlation of *in Vitro* Drug Product Dissolution and *in Vivo* Bioavailability. *Pharm. Res.* **1995**, *12*, 413−420.

(2) Lennernäs, H.; Ahrenstedt, Ö.; Hällgren, R.; Knutson, L.; Ryde, M.; Paalzow, L. K. Regional Jejunal Perfusion, a New *in Vivo* Approach to Study Oral Drug Absorption in Man. *Pharm. Res.* **1992**, *9*, 1243−1251.

(3) Smith, D. A.; Jones, B. C.; Walker, D. K. Design of Drugs Involving the Concepts and Theories of Drug Metabolism and Pharmacokinetics. *Med. Res. Rev.* **1996**, *16*, 243−266.

(4) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Delivery Rev.* **1997**, *23*, 3−25.

(5) Ren, S.; Das, A.; Lien, E. J. QSAR Analysis of Membrane Permeability to Organic Compounds. *J. Drug Targeting* **1996**, *4*, 103−107.

(6) Hamilton, H. W.; Steinbaugh, B. A.; Stewart, B. H.; Chan, O. H.; Schmid, H. L.; Schroeder, R.; Ryan, M. J.; Keiser, J.; Taylor, M. D.; Blankley, C. J.; Kaltenbronn, J. S.; Wright, J.; Hicks, J. Evaluation of Physicochemical Parameters Important to the Oral Bioavailability of Peptide-like Compounds: Implications for the Synthesis of Renin Inhibitors. *J. Med. Chem.* **1995**, *38*, 1446−1455.

(7) Roth, B. D.; Bocan, T. M. A.; Blankley, C. J.; Chucholowski, A. W.; Creger, P. L.; Creswell, M. W.; Ferguson, E.; Newton, R. S.; O'Brien, P.; Picard, J. A.; Roark, W. H.; Sekerke, C. S.; Sliskovic, D. R.; Wilson, M. W. Relationship between Tissue Selectivity and Lipophilicity for Inhibitors of HMG-CoA Reductase. *J. Med. Chem.* **1991**, *34*, 463−466.

(8) Hollander, D.; Ricketts, D.; Boyd, C. A. R. Importance of "Probe" Molecular Geometry in Determining Intestinal Permeability. *Can. J. Gastroenterol.* **1988**, *2*, 35A−38A.

(9) Nook, T.; Doelker, E.; Buri, P. Intestinal Absorbtion Kinetics of Various Model Drugs in Relation to Partition Coefficients. *Int. J. Pharmaceut.* **1988**, *43*, 119−129.

(10) Walter, A.; Gutknecht, J. Permeability of Small Nonelectrolytes through Lipid Bilayer Membranes. *J. Membr. Biol.* **1986**, *90*, 207−217.

(11) Palm, K.; Luthman, K.; Ungell, A.-L.; Strandlund, G.; Artursson, P. Correlation of Drug Absorption with Molecular Surface Properties. *J. Pharm. Sci.* **1996**, *85*, 32−39.

(12) van de Waterbeemd, H.; Camenisch, G.; Folkers, G.; Raevsky, O. A. Estimation of Caco-2 Cell Permeability using Calculated Molecular Descriptors. *Quant. Struct.-Act. Relat.* **1996**, *15*, 480−490.

(13) Knutson, L.; Odlind, B.; Hällgren, R. A New Technique for Segmental Jejunal Perfusion in Man. *Am. J. Gastroenterol.* **1989**, *84*, 1278−1284.

(14) Avdeef, A. pH-Metric log P. Part 1. Difference Plots for Determining Ion-Pair Octanol−Water Partition Coefficients of Multiprotic Substances. *Quant. Struct.-Act. Relat.* **1992**, *11*, 510−517.

(15) Lennernäs, H.; Nilsson, D.; Aquilonius, S.-M.; Ahenstedt, Ö.; Knutson, L.; Paalzow, L. K. The Effect of L-Leucine on the Absorption of Levodopa, Studied by Regional Jejunal Perfusion in Man. *Br. J. Clin. Pharmac.* **1993**, *35*, 243−250.

(16) Lennernäs, H.; Knutson, L.; Knutson, T.; Hussain, A.; Lesko, L.; Salmonson, T.; Amidon, G. L. Human Effective Permeability for Five Cardiovascular Drugs: Atenolol, Metoprolol, Hydrochlortiazide, Furosemide and R/S−Verapamil. *Pharm. Res.*, submitted.

(17) Lennernäs, H.; Knutson, L.; Knutson, T.; Hussain, A.; Lesko, L.; Salmonson, T.; Amidon, G. L. Human Effective Permeability of Two NSAID's and Two CNS Drugs. *J. Pharm. Sci.*, submitted.

(18) Lennernäs, H.; Knutson, L.; Hussain, A.; Lesko, L.; Salmonson, T.; Amidon, G. L. The Effect of Amiloride on the Effective Permeability of Amoxicillin in Human Jejunum. *Br. J. Clin. Pharmac.*, submitted.

(19) Sandström, R.; Karlsson, A.; Knutson, L.; Lennernäs, H. Jejunal Absorption and Metabolism of R/S−Verapamil in Humans. *Pharm. Res.* **1998**, *15*, 856−862.

(20) Lennernäs, H.; Knutson, L.; Lesko, L.; Salmonson, T.; Amidon, G. L. Human Jejunal Permeabilities Comparison to α-Methyldopa and L-dopa: Two Carrier Mediated Drugs. *Pharm. Res.* **1996**, *13*, S329.

(21) Lindahl, A.; Sandström, R.; Ungell, A.-L.; Abrahamsson, B.; Knutson, T. W.; Knutson, L.; Lennernäs, H. Jejunal Permeability and Hepatic Extraction of Fluvastatin in Humans. *Clin. Pharmacol. Ther.* **1996**, *60*, 493−503.

(22) Fagerholm, U.; Borgström, L.; Ahrenstedt, Ö.; Lennernäs, H. The Lack of Effect of Induced Net Fluid Absorption on the *in vivo* Permeability of Terbutaline in the Human Jejunum. *J. Drug Targeting* **1995**, *3*, 191−200.

(23) Fagerholm, U.; Nilsson, D.; Knutson, L.; Lennernäs, H. Jejunal Permeability in Humans *in vivo* and Rats *in situ*: Investigation of Molecular Size Selectivity and Solvent Drag. *Acta Physiol. Scand.*, in press.

(24) Takamatsu, N.; Welage, L. S.; Idkaidek, N. M.; Liu, D. Y.; Lee, P. I.; Hayashi, Y.; Rhie, J. K.; Lennernäs, H.; Barnett, J. L.; Shah, V. P.; Lesko, L.; Amidon, G. L. Human Intestinal Permeability of Piroxicam, Propranolol, Phenylalanine, and PEG 400 Determined by Jejunal Perfusion. *Pharm. Res.* **1997**, *14*, 1127−1132.

(25) Fagerholm, U.; Johansson, M.; Lennernäs, H. Comparison Between Permeability Coefficients in Rat and Human Jejunum. *Pharm. Res.* **1996**, *13*, 1336−1342.

(26) Lennernäs, H. Does Fluid Flow Across the Intestinal Mucosa Affect Quantitative Oral Drug Absorption? Is It Time for a Reevaluation? *Pharm. Res.* **1995**, *12*, 1573−1582.
(27) Paintaud, G.; Alván, G.; Dahl, M. L.; Grahnén, A.; Sjövall, J.; Svensson, J. O. Nonlinearity of Amoxicillin Absorption Kinetics in Human. *Eur. J. Clin. Pharmacol.* **1992**, *43*, 283−288.
(28) Nakashima, E.; Tsuji, A.; Kagatani, S.; Yamana, T. Intestinal Absorption Mechanism of Amino-Beta-Lactam Antibiotics. III. Kinetics of Carrier-Mediated Transport across the Rat Small Intestine in situ. *J. Pharmacobiodyn.* **1984**, *7*, 452−464.
(29) Hidalgo, I. J.; Borchardt, R. T. Transport of a Large Neutral Amino Acid (Phenylalanine) in a Human Intestinal Epithelial Cell Line: Caco-2. *Biochim. Biophys. Acta* **1990**, *1028*, 25−30.
(30) The Pomona database can be reached on the World Wide Web (http://clogP.pomona.edu/medchem/chem/clogp/).
(31) Avdeef, A.; Comer, J. E. A.; Thomson, S. J. pH-Metric log *P*. 3. Glass Electrode Calibration in Methanol−Water, Applied to pK_a Determination of Water-Insoluble Substances. *Anal. Chem.* **1993**, *65*, 42−49.
(32) Avdeef, A. pH-Metric log *P*. II. Refinement of Partition Coefficients and Ionization Constants of Multiprotic Substances. *J. Pharm. Sci.* **1993**, *82*, 183−190.
(33) Takács-Novák, K.; Avdeef, A. Interlaboratory Study of Log *P* Determination by Shake-Flask and Potentiometric Methods. *J. Pharm. Biomed. Anal.* **1996**, *14*, 1405−1413.
(34) Meldrum, S. J.; Watson, B. W.; Riddle, H. C.; Bown, R. L.; Sladen, G. E. pH Profile of Gut as Measured by Radiotelemetry Capsule. *Br. Med. J.* **1972**, *2*, 104−106.
(35) Thomson, A. B. R.; Dietschy, J. M. The Role of the Unstirred Water Layer in Intestinal Permeation. In *Pharmacology of Intestinal Permeation II*; Csaky, T. Z., Ed.; Springer-Verlag: Berlin, 1984.
(36) Lucas, M. L.; Schneider, W.; Haberich, F. J.; Blair, J. A. Direct Measurement by pH-Microelectrode of the pH Microclimate in Rat Proximal Jejunum. *Proc. R. Soc. London B.* **1975**, *192*, 39−48.
(37) SYBYL: Molecular Modeling Software, Tripos Associates, Inc.: St. Louis, MO 63144, 1996.
(38) Leo, A. J. Calculating log P_{oct} from Structures. *Chem. Rev.* **1993**, *93*, 1281−1306.
(39) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(40) The SPL-scripts performing these calculations can be obtained from the authors upon request.
(41) Palm, K.; Stenberg, P.; Luthman, K.; Artursson, P. Polar Molecular Surface Properties Predict the Intestinal Absorption of Drugs in Humans. *Pharm. Res.* **1997**, *14*, 568−571.
(42) van de Waterbeemd, H.; Kansy, M. Hydrogen-Bonding Capacity and Brain Penetration. *Chimia* **1992**, *46*, 299−303.
(43) *Cumulative Subject Index and Drug Compendium;* Hansch, C., Sammes, P. G., Tailor, J. B., Ed.; Pergamon Press: Oxford, 1991; Vol. 6.
(44) CLOGP, LogP Calculation Algorithm, Pomona College Medicinal Chemistry Project: Claremont, CA 91711, 1989.
(45) SIMCA, Umetri AB, Box 7960: SE-90719 Umeå, Sweden, 1996.
(46) Wold, S. Validation of QSAR's. *Quant. Struct.-Act. Relat.* **1991**, *10*, 191−193.
(47) *Clark's Isolation and Identification of Drugs;* Moffat, A. C., Ed.; The Pharmaceutical Press: London, 1986.
(48) *Therapeutic Drugs;* Dollery, C., Ed.; Churchill Livingstone: Edinburgh, 1991.
(49) *Lange's Handbook of Chemistry;* Dean, J. E., Ed.; McGraw-Hill Book Co.: 1979.
(50) Avdeef, A. Assessment of Distribution-pH Profiles. In *Lipophilicity in Drug Action and Toxicology*; Pliska, V., Testa, B., van de Waterbeemd, H., Ed.; VCH: Weinheim, 1996; pp 109−139.
(51) Chamberlain, K.; Evans, A. A.; Bromilow, R. H. 1-Octanol/Water Partition Coefficient (K_{OW}) and pK_a for Ionisable Pesticides Measured by a pH-Metric Method. *Pestic. Sci.* **1996**, *47*, 265−271.
(52) Eriksson, L.; Jonsson, J.; Sjöström, M.; Wold, S. A Strategy for Ranking Environmentally Occurring Chemicals. Part II. An Illustration with Two Data Sets of Chlorinated Aliphatics and Aliphatic Alcohols. *Chemom. Intell. Lab. Syst.* **1989**, *7*, 131−141.
(53) In the multivariate analysis we also included squared lipophilicity values. However, according to the VIP plot these variables were not found to be important.
(54) Shah, M. V.; Audus, K. L.; Borchardt, R. T. The Application of Bovine Brain Microvessel Endothelial-Cell Monolayers Grown onto Polycarbonate Membranes *in Vitro* to Estimate the Potential Permeability of Solutes Through the Blood-Brain Barrier. *Pharm. Res.* **1989**, *6*, 624−627.
(55) Fagerholm, U.; Lennernäs, H. Experimental Estimation of the Effective Unstirred Water Layer Thickness in the Human Jejunum, and its Importance in Oral Drug Absorption. *Eur. J. Pharm. Sci.* **1995**, *3*, 247−253.
(56) Artursson, P.; Karlsson, J. Correlation between Oral Drug Absorption in Humans and Apparent Drug Permeability Coefficients in Human Intestinal Epithelial (Caco-2) Cells. *Biochem. Biophys. Res. Comm.* **1991**, *175*, 880−885.
(57) Kansy, M.; Senner, F.; Gubernator, K. Physicochemical High Throughput Screening: Parallel Artificial Membrane Permeation Assay in the Description of Passive Absorption Processes. *J. Med. Chem.* **1998**, *41*, 1007−1010.
(58) Theory Guide, issued as standard with the Sirius instrument.
(59) McFarland, J. W.; Berger, C. M.; Froshauer, S. A.; Hayashi, S. F.; Hecker, S. J.; Jaynes, B. H.; Jefson, M. R.; Kamicker, B. J.; Lipinski, C. A.; Lundy, K. M.; Reese, C. P.; Vu, C. B. Quantitative Structure−Activity Relationships among Macrolide Antibacterial Agents: In Vitro and in Vivo Potency against *Pasteurella multocida*. *J. Med. Chem.* **1997**, *40*, 1340−1346.

JM9810102