# Estimation of CACO-2 Cell Permeability Using Calculated Molecular Descriptors

**ARTICLE** *in* QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIPS · JANUARY 1996

DOI: 10.1002/qsar.19960150604

# Estimation of Caco-2 Cell Permeability using Calculated Molecular Descriptors

**Han van de Waterbeemd*, Gian Camenisch**

F. Hoffmann-La Roche Ltd., Pharma Research New Technologies, CH-4070 Basel, Switzerland.

**Gerd Folkers**

ETHZ, Department Pharmacy, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland.

**Oleg A. Raevsky**

The Institute of Physiologically Active Compounds of the Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia.

## Abstract

Permeability of a compound by passive diffusion through a biological membrane such as Caco-2 cell monolayers depends mainly on lipophilicity, molecular size and hydrogen bonding capacity of the solute. Since these properties are often used in structure-permeability correlations, we have evaluated different calculated descriptors for size and hydrogen bonding, as well as their intercorrelations. A new descriptor for hydrogen bonding potential is introduced. It is demonstrated that a combination of appropriate size and H-bonding descriptors using graphical or Eq.-based approaches may be of potential use for membrane permeation estimation.

**Key words**: Lipophilicity, Permeability, Caco-2, Hydrogen Bonding, Drug Absorption, Molecular Size, Molecular Weight

## Abbreviations and Symbols

| | |
|---|---|
| Å | Angström |
| Ca | free energy hydrogen bond acceptor factor |
| Cd | free energy hydrogen bond donor factor |
| Cad | sum of absolute values of free energy H-bond factors, characterizing the total H-bond ability of a compound |
| Caco-2 | monolayer of cultured colon carcinoma cells |
| $\Delta$log P | difference between octanol/water and alkane/water partition coefficient, reflecting hydrogen bonding capacity of a solute |
| HA | number of hydrogen bond acceptors (lone-pair count) in neutral form ($HA_i$ for state at pH7.4) |
| HB | total number of atoms capable of forming hydrogen bonds |
| HD | number of hydrogen bond donors in neutral form ($HB_i$ for state at pH7.4) |
| HT | total number of potential hydrogen bonds for neutral form ($HT_i$ for state at pH7.4) |
| $A$ | polarity term, mainly hydrogen bonding |
| log D | octanol/water distribution coefficient at pH 7.4 |
| log P | octanol/water partition coefficient |
| log $P_{erm}$ | permeability constant |
| MW | molecular weight |
| O | ovality |
| pKa | ionization constant |
| S | surface area |
| SNP | nonpolar part of the surface area |
| SP | polar part of the surface area |
| SW | water accessible surface area |
| V | molar volume |
| VNP | nonpolar part of the molecular volume |
| VP | polar part of the molecular volume |
| VW | water accessible molecular volume |
| X,Y,Z | principal axes (length, width1, width2) |
| X/Y, X/Z | ratios of principal axes (length/width1, length/width2) |

## 1 Introduction

Much interest is currently focused on cell culture models for the prediction of oral drug absorption [1,2]. The human intestinal epithelial cell line Caco-2 has been particularly recommended for such studies, since these cells express various biological membrane properties, including enzymatic and transporter systems [3,4]. It is assumed that a large majority of drugs crosses biological barriers by means of a passive diffusion mechanism. Some compounds, including di- and tripeptides and related peptidomimetics, may use active transport systems. Permeation may also be hindered by efflux systems involving P-glycoproteins [5,6]. Two possible pathways exist for permeation, the transcellular and paracellular routes. The paracellular route is the tight-junctional pathway between cells, which can be considered as pores filled with water [7].

In a number of studies it has been demonstrated that permeability coefficients measured for transport through Caco-2 monolayer cell cultures are correlated with lipophilicity [8], while others discussed the role of hydrogen bonding [9–11] or charge [12].

It is commonly assumed that the permeability-lipophilicity relationship is sigmoidal [13–15]. Outliers from a general sigmoidal rela-

* to receive all correspondence
FAX +41-61-688 1075

0931-8771/96/0612-0480 $5.00+.25/0

tionship have been explained in terms of active transport, molecular size, diffusion limitation through aqueous stagnant layers at the membrane, or solubility of the drug [13,15,16]. Nevertheless, many authors still assume incorrectly that the permeability-lipophilicity relationship is linear [17–20].

Lipophilicity is often expressed by the experimental or calculated octanol/water partition coefficient log P. In order to account for the degree of ionization, correlations with permeation data should be performed using distribution coefficients log D, usually measured at physiological pH 7.4.

It has recently been demonstrated that log P, or log D, values are composed of two components, namely size and polarity [21–24]. Such approaches go back to early work on the dissection of amino acid lipophilicities [25]. Written in a mathematical form as:

$$\log P = a.V - A \quad (1)$$

The molar volume V is just one simple example of a size or bulk descriptor. Others may be used, since there is a high collinearity among these descriptors, such as molecular weight, polarizability, surface area, molar refractivities, parachor [22,25]. Particularly polarizability of the molecule is assumed to be of great importance, but will not be further discussed here since it is highly collinear in molecular weight [26,27].

For uncharged compounds these $A$ values (Eq. 1) appear to encode mainly, but not exclusively, the hydrogen bonding capacity of a substance [24,28]. More extended Eqs. of this type have been derived to express the individual contributions of hydrogen bond acceptors and donors, using thermodynamic or solvatochromic parameters [29–32]. For a restricted data set of compounds containing one hydroxy, aldehyde or ketone group, Raevsky and coworkers showed that for octanol/water partitioning there is an essential contribution of the free energy H-bond acceptor factor (Ca) and a very small contribution of the free energy H-bond donor factor (Cd) [31]. The solvation Eq. developed by Abraham [29,32], used to predict partitioning coefficients but also brain penetration [32] or skin permeability [17,33], suffers from the serious drawback that the required molecular descriptors for H-bonding, namely H-bond acceptor basicity ($\beta$) and H-bond donor acidity ($\alpha$) are very difficult to estimate.

From a different approach it was found that the difference between octanol/water and alkane/water log P values also encodes for H-bonding capacity [34–36]. It has been argued that experimental $\Delta$log P values provide better descriptors for H-bonding than the calculated ones [37]. However, this descriptor has the obvious disadvantage of requiring tedious measurements. Furthermore, experimental descriptors are limited to synthesized compounds. Often, an *a priori* property estimation is of interest, in order to filter structural proposals.

Hydrogen-bonding appears to be an important factor in transport processes through the blood-brain barrier (BBB) [35,38,39] and the skin [33,40,41]. Differences in hydrogen bonding also rationalize peptide transport through Caco-2 cell culture monolayers [9]. Therefore, probably hydrogen bonding plays a role in membrane permeability in general.

One of the simplest theoretical approaches to account for hydrogen bonding of a drug is to count the number of hydrogen bond forming groups, which goes back to an early proposal by Stein [42]. An extension was adopted by Conradi and colleagues, taking into account the potential number of hydrogen bonds each functional group can make [10]. This method treats all hydrogen bonds as energetically equivalent. Another problem concerns the conformational aspect. Some potential hydrogen bonds may be sterically hindered. Therefore these bonds should not be considered in an estimation of H-bonding potential. Furthermore, the parametrization of ionization was not addressed.

On the basis of the literature, the current paradigm of structure-permeability correlations can be expressed as:

$$\text{permeability} = f(\text{lipophilicity, molecular size, H-bonding capacity, charge}) \quad (2)$$

However, charge is included in lipophilicity when distribution coeffcients (log D) instead of partition coefficients (log P) are used. Furthermore, as seen in Eq. 1, molecular size and H-bonding are components of lipophilicity. Thus, one can also write:

$$\text{permeability} = f(\text{molecular size, H-bonding capacity}) \quad (3)$$

This simpler Eq. eliminates the knowledge of lipophilicity for the estimation of membrane permeation. These properties, molecular size and H-bonding, besides solubility and dissolution, are believed to play an important role for the optimization of oral bioavailability of drugs [19,43,44]. Our goal is to study in more detail the optimal range and combination of these properties in order to optimize passive drug absorption through biological membranes.

In the present paper we have investigated a number of different calculated molecular size and hydrogen-bonding descriptors and studied their relationship to membrane permeation. In particular, we introduce a new descriptor for hydrogen bonding potential of a compound, Cad, as the sum of the previously defined free energy H-bond donor (Cd) and acceptor (Ca) factors [30,31]. Our objective is to find a set of simple descriptors to estimate membrane permeation from molecular structure and which can be used in high-throughput screening and in the design of new orally active compounds.

## 2 Materials and Methods

### *2.1 Data set*

The structurally heterogeneous data set used covers a relatively wide range of molecular size and lipophilicity [8]. All compounds are regarded as being transported by passive diffusion. The compounds are, at pH 7.4, either neutral or in their fully ionized form, allowing us to evaluate the influence of ionization on permeation. The following seventeen compounds are considered: corticosterone (Co), testosterone (Te), propranolol (Pr), alprenolol (Al), warfarin (Wa), metoprolol (Me), felodipine (Fe), hydrocortisone (Hy), dexamethasone (De), salicylate (Sa), acetylsalicylate (Ac), practolol (Pa), terbutaline (Tb), atenolol (At), mannitol (Ma), sulphasalazine (Su) and olsalazine (Ol).

### 2.2 *Experimental Data*

Permeability data through Caco-2 cells (log $P_{erm}$) and 1-octanol/water distribution coefficients (log D, for molecules considered in their ionization state at pH 7.4) were taken from the literature [8]. These permeability studies were performed at a pH of 7.4. 1-octanol/water partition coefficients (log P, for molecules considered in their neutral form) and pKa values were taken from the MedChem94 database [42]. The experimental data are presented in Table 1.

### 2.3 *Calculated Molecular Descriptors*

An overview of the calculated molecular descriptors is presented in Table 2. The following programs have been used for the calculations: PCMODELS [45], HYBOT 5.0 [30,31,46], MOLOC (Roche in-house modeling system) [47], TSAR [48], HBOND [49].

Different **molecular size descriptors** have been compared (Table 3). A number of these are conformation-dependent. In all cases energy-minimized structures using the united-atom approach have been considered [47]. In preliminary calculations using a set of low-energy conformations it was found that the standard deviations of the surface and volume properties are in the order of 10%, which corresponds to recent findings by others [20]. We do not consider this as critical for the present analysis. Furthermore, for conformationally flexible compounds the actual conformation inside a membrane is difficult to estimate. The computed values have been compared to volume and surface data from the literature [54] and those calculated by the TSAR program [48]. Due to differences in parametrization, each of these programs produces different, but strongly intercorrelated results. **Molecular shape** has been evaluated using ovality $O = S/[4\pi(3V/4\pi)^{2/3}]$ defined as the ratio between the actual surface (S) and the minimum ideally spheric surface [51], and finally using ratios of the length to the smallest and largest width of the molecule (X/Y and X/Z) given by its principal axes.

**Hydrogen bonding descriptors** are compared in Table 4. According to Eq. 1, $A$ values have been calculated taking the alkanes as reference [38]. The free energy-based hydrogen bond factors Ca and Cd have been computed with the HYBOT 5.0 program and database using experimental data of 12,000 H-bonded complexes [30,31]. Cd appears to be highly correlated to Abraham's free energy factor for H-donors $\alpha_2^H$, while Ca reflects $\beta_2^H$ for H-acceptors [30,31]. A new descriptor Cad was defined as the sum of the absolute values of Ca and Cd, with the intention to reflect total hydrogen bonding capacity. By contrast to the original work of Raevsky *et al.* [30,31], in the present study Cd values are taken as positive values (see below).

Since part of the compounds is ionized at the pH of permeability measurement, both the neutral form (HA, HD, HT) and ionized form ($HA_i$, $HD_i$, $HT_i$) have been considered in the computation of H-bonding potential. Probably the effect of ionization on the total H-bonding capacity is small. For a carboxylic acid in its neutral form we may define three acceptors and one donor, and in its

**Table 1.** Experimental values and related calculated descriptors

| Symbol a) | log $P_{erm}$ b) | log D c) | log P d) | CLOGP e) | pKa | Netto-Charge c) |
|---|---|---|---|---|---|---|
| Sa | –4.924 | –2.14 | 2.24 | 2.187 | 3.0 | – |
| Ac | –5.620 | –2.57 | 1.19 | 1.023 | 3.5 | – |
| Ma | –6.745 | –3.10 | –3.10 | –4.670 | – | 0 |
| Tb | –6.420 | –1.40 | – f) | – f) | 8.8 | + |
| Al | –4.393 | 1.00 | 2.89 | 2.652 | 9.7 | + |
| Pr | –4.378 | 1.54 | 3.30 | 2.735 | 9.5 | + |
| Pa | –6.046 | –1.40 | 0.79 | 0.755 | 9.5 | + |
| At | –6.700 | –2.14 | 0.17 | –0.108 | 9.6 | + |
| Me | –4.569 | 0.07 | 2.04 | 1.196 | 9.7 | + |
| Te | –4.286 | 3.31 | 3.31 | 3.219 | – | 0 |
| Ol | –6.959 | –4.50 | n.a.[g] | 4.50 | ~3[h] | – – |
| Wa | –4.417 | 0.12 | 2.52 | 2.785 | 5.1 | – |
| Co | –4.263 | 1.89 | 1.53 | 1.163 | – | 0 |
| Hy | –4.668 | 1.53 | 1.53 | 0.537 | – | 0 |
| Fe | –4.644 | 3.48 | 3.48 | 4.525 | – | 0 |
| De | –4.903 | 1.74 | 1.74 | 1.505 | – | 0 |
| Su | –6.886 | –0.13 | n.a.[g] | 3.831 | 2.4 | – |

a) Compound names; see Data Set under Materials and Methods
b) Permeability [cm/s] through cultured Caco-2 monolayers
c) At pH 7.4
d) Data from the MedChem94 database
e) Calculated with PCMODELS.
f) Over the whole pH range terbutaline is charged.
g) Not available by shake flask lipophilicity measurement method.
h) Two carboxylic groups.

**Table 2.** Calculated molecular descriptors evaluated in the present study.

| Descriptor | Symbol | Reference |
|---|---|---|
| ***Lipophilicity*** | | |
| octanol/water partition coefficient | CLOGP | 47 |
| ***Molecular size and shape***[a] | | |
| Van der Waals volume [Å$^3$] | V | |
| surface area [Å$^2$] | S | |
| water accessible volume [Å$^3$] | VW | |
| water accessible surface [Å$^2$] | SW | |
| ovality | $O = S/[4\pi(3V/4\pi)^{2/3}]$ | 48 |
| polar part of the volume [Å$^3$] | VP | 35 |
| nonpolar part of the volume [Å$^3$] | VNP | 35 |
| nonpolar part of the surface area [Å$^2$] | SNP | 35 |
| molecular weight [g/mol] | MW | |
| principal axes [Å] | X, Y, Z (X > Y > Z) | |
| ratio length/width1 | X/Y | |
| ratio length/width2 | X/Z | |
| ***Hydrogen bonding***[b] | | |
| number of H-bond acceptors | HA, $HA_i$ | 9–11, 35 |
| number of H-bond donors | HD, $HD_i$ | 9–11, 35 |
| total number of H-bonds | HT (= HA + HD), $HT_i$ | 9–11, 35 |
| total number of atoms capable of H-bonding | HB | 49, 50 |
| total number of possible H-bonds | HBOND | 46 |
| free energy of H-bond acceptor factor | Ca | 27,28 |
| free energy of H-bond donor factor | Cd | 27,28 |
| total H-bonding potential | Cad (= Ca + Cd)[c] | 27,28 |
| interactive polar parameter | $A$ | 25 |
| polar part of the surface [Å$^2$] | SP | 35 |

[a] Energy-minimized structures using MOLOC

[b] Molecules considered in their neutral state; $HA_i$, $HD_i$, $HT_i$, refer to the ionized state

[c] Absolute values of Ca and Cd (see text)

**Table 3.** Molecular size and shape descriptors for the selected compounds

| Symbol | MW | V a) | VW b) | VNP a) | VP a) | S a) | SW b) | O a) | X a) | Y a) | Z a) | X/Y | X/Z |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sa | 138 | 125 | 1023 | 101 | 24 | 151 | 302 | 1.25 | 4.00 | 2.46 | 0.00 | 1.63 | – |
| Ac | 180 | 163 | 827 | 133 | 30 | 194 | 361 | 1.34 | 3.94 | 3.12 | 0.61 | 1.26 | 6.44 |
| Ma | 182 | 165 | 990 | 117 | 49 | 206 | 355 | 1.41 | 6.31 | 1.20 | 0.97 | 5.24 | 6.49 |
| Tb | 225 | 229 | 1825 | 197 | 32 | 265 | 460 | 1.46 | 10.99 | 2.14 | 0.73 | 5.14 | 15.08 |
| Al | 249 | 270 | 1497 | 248 | 22 | 311 | 531 | 1.54 | 14.56 | 3.24 | 0.58 | 4.49 | 24.93 |
| Pr | 259 | 293 | 2323 | 271 | 22 | 340 | 556 | 1.59 | 17.71 | 2.98 | 0.58 | 5.95 | 30.70 |
| Pa | 266 | 272 | 1339 | 235 | 37 | 317 | 546 | 1.56 | 20.87 | 1.74 | 0.86 | 11.97 | 24.29 |
| At | 266 | 272 | 1707 | 231 | 41 | 317 | 543 | 1.56 | 20.35 | 1.82 | 0.93 | 11.16 | 21.98 |
| Me | 267 | 283 | 2323 | 256 | 28 | 329 | 567 | 1.58 | 20.11 | 3.10 | 0.86 | 6.49 | 23.39 |
| Te | 288 | 298 | 1536 | 281 | 17 | 309 | 483 | 1.43 | 8.08 | 2.34 | 1.63 | 3.45 | 4.95 |
| Ol | 302 | 258 | 2333 | 192 | 66 | 298 | 525 | 1.52 | 15.52 | 3.83 | 0.40 | 4.05 | 38.80 |
| Wa | 308 | 290 | 1811 | 261 | 30 | 317 | 531 | 1.50 | 10.55 | 3.42 | 1.73 | 3.08 | 6.09 |
| Co | 346 | 341 | 1936 | 307 | 33 | 350 | 538 | 1.49 | 9.00 | 2.68 | 1.73 | 3.36 | 5.21 |
| Hy | 362 | 350 | 1962 | 308 | 42 | 363 | 546 | 1.51 | 11.40 | 2.49 | 1.30 | 4.58 | 8.74 |
| Fe | 384 | 335 | 1739 | 300 | 34 | 369 | 582 | 1.58 | 8.42 | 5.27 | 1.60 | 1.60 | 5.25 |
| De | 392 | 365 | 2620 | 324 | 41 | 371 | 569 | 1.50 | 10.70 | 2.91 | 1.43 | 3.68 | 5.25 |
| Su | 398 | 340 | 3414 | 266 | 74 | 381 | 651 | 1.62 | 30.65 | 2.46 | 0.67 | 12.48 | 45.48 |

[a] Calculated with MOLOC.

[b] Calculated with MOLOC using a water radius of 1.45 Å.

ionized form four acceptors. In both cases four potential hydrogen bonds may be formed. More experimental data for larger data sets are needed to determine how H-bonding of (partially) charged compounds is best accounted for. The present study intends to unravel trends.

### *2.4 Regression Analysis*

It is commonly assumed that the general form of the relationship between permeability and lipophilicity is sigmoidal. However, a sigmoidal function needs at least four parameters to be estimated, which with the present limited data set is statistically not acceptable. Therefore, we chose three approximate approaches. First a principal component analysis (PCA) was performed on the full set of generated size and H-bonding descriptors in order to extract the key information for permeability prediction. Second, the first two principal components (Table 5) were used as new descriptors in a simple multiple linear regression (MLR) approach. Third, we explored linear combinations of the original variables as approximate predictors for permeability. The statistical analyses were performed using Systat 5.0 for Windows [55] and Scan 1.0 for Windows [56].

Principal components can be found in Table 5. Regression Eqs. are reported in Table 8. A correlation matrix of the hydrogen bonding descriptors is given in Table 6, and a partial correlation matrix for the molecular weight is seen in Table 9. Linear correlations between hydrogen bonding descriptors and permeability as well as lipophilicity are considered in Table 7.

## 3. Results and Discussion

### *3.1 Correlations among the Descriptors*

An elegant means to visualize the relationships between the descriptors included in this study is principal component analysis (PCA). The loadings plot is presented in Figure 1. Descriptors related to H-bonding are seen in the upper part, while the size- and

**Table 5.** Principal Component Analysis

| Variable | PC1 (43.4%) | PC2 (34.2%) | PC3 (9.3%) |
|---|---|---|---|
| MW | –0.192 | –0.218 | 0.232 |
| V | –0.154 | –0.270 | 0.098 |
| VP | –0.270 | 0.116 | 0.104 |
| VNP | –0.096 | –0.300 | 0.076 |
| VW | –0.222 | –0.154 | –0.012 |
| VWP | –0.199 | 0.111 | 0.162 |
| VNPW | –0.180 | –0.187 | –0.098 |
| S | –0.170 | –0.270 | 0.016 |
| SP | –0.255 | 0.116 | 0.211 |
| SNP | –0.086 | –0.313 | –0.056 |
| SW | –0.189 | –0.254 | –0.062 |
| SPW | –0.214 | 0.189 | 0.174 |
| SNPW | –0.014 | –0.315 | –0.157 |
| O | –0.183 | –0.215 | –0.197 |
| Z | –0.019 | –0.185 | 0.276 |
| Y | 0.021 | –0.163 | 0.361 |
| X | –0.203 | –0.097 | –0.361 |
| X/Y | –0.181 | –0.010 | –0.450 |
| HA | –0.237 | 0.104 | 0.289 |
| HD | –0.158 | 0.210 | –0.209 |
| HT | –0.241 | 0.180 | 0.085 |
| HB | –0.273 | 0.093 | 0.118 |
| *HBOND* | –0.243 | 0.175 | –0.057 |
| Ca | –0.285 | –0.006 | –0.014 |
| Cd | –0.160 | 0.243 | –0.183 |
| Cad | –0.256 | 0.142 | –0.117 |

X/Z and $\Lambda$ excluded due to missing values

**Table 4.** Hydrogen bonding descriptors for the selected compounds

| Symbol | HA | $HA_i$ | HD | $HD_i$ | HT | $HT_i$ | HB | HBOND a) | Ca b) | Cd b) | Cad b) | SP c) | $\Lambda$ d) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sa | 3 | 4 | 2 | 1 | 5 | 5 | 3 | 5 | 3.7 | 4.8 | 8.5 | 37 | – 1.9 |
| Ac | 4 | 5 | 1 | 0 | 5 | 5 | 4 | 5 | 4.3 | 3.1 | 7.4 | 52 | – 4.2 |
| Ma | 6 | 6 | 6 | 6 | 12 | 12 | 6 | 12 | 9.9 | 11.4 | 21.3 | 70 | – 8.6 |
| Tb | 3 | 3 | 4 | 5 | 7 | 8 | 4 | 8 | 6.3 | 8.1 | 14.5 | 43 | – |
| Al | 2 | 2 | 2 | 3 | 4 | 5 | 3 | 5 | 6.1 | 3.8 | 9.9 | 27 | – 6.0 |
| Pr | 2 | 2 | 2 | 3 | 4 | 5 | 3 | 5 | 5.9 | 3.8 | 9.7 | 27 | – 6.4 |
| Pa | 4 | 4 | 3 | 4 | 7 | 8 | 5 | 8 | 9.3 | 5.7 | 15.0 | 49 | – 8.2 |
| At | 4 | 4 | 4 | 5 | 8 | 9 | 5 | 9 | 9.3 | 8.0 | 17.3 | 55 | – 8.8 |
| Me | 3 | 3 | 2 | 3 | 5 | 6 | 4 | 6 | 7.6 | 3.8 | 11.4 | 36 | – 7.3 |
| Te | 3 | 3 | 1 | 1 | 4 | 4 | 2 | 3 | 4.6 | 2.2 | 6.8 | 29 | – 6.5 |
| Ol | 8 | 10 | 4 | 2 | 12 | 12 | 8 | 11 | 9.4 | 9.7 | 19.1 | 94 | – |
| Wa | 5 | 7 | 1 | 0 | 6 | 7 | 4 | 5 | 6.7 | 2.2 | 8.9 | 50 | – 7.1 |
| Co | 5 | 6 | 2 | 1 | 7 | 7 | 4 | 6 | 7.6 | 3.1 | 10.7 | 55 | – 9.4 |
| Hy | 5 | 6 | 3 | 3 | 8 | 9 | 5 | 8 | 9.1 | 5.3 | 14.4 | 67 | –10.0 |
| Fe | 5 | 5 | 1 | 1 | 6 | 6 | 5 | 6 | 8.7 | 1.7 | 10.4 | 49 | – 7.6 |
| De | 5 | 5 | 3 | 3 | 8 | 8 | 6 | 8 | 10.3 | 5.3 | 15.6 | 64 | –10.3 |
| Su | 8 | 9 | 3 | 2 | 11 | 11 | 9 | 11 | 14.1 | 7.0 | 21.1 | 106 | – |

a) Calculated with HBOND.
b) Calculated with HYBOT. The sign of Cd has been reversed (see text).
c) Calculated with MOLOC.
d) Calculated based on Eq. (1) from the relationship of 1-octanol/water log P values versus Van de Waals volume V of different n-alkanes as the distance on the ordinate [25,35].

**Table 6.** Correlation matrix (r) of the evaluated H-bonding descriptors

| | HA | $HA_i$ | HD | $HD_i$ | HT | $HT_i$ | HB | HBOND | Ca | Cd | Cad | SP | Λ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 14 |
| HA | 1.000 | 0.964 | 0.399 | 0.074 | 0.879 | 0.814 | 0.922 | 0.755 | 0.749 | 0.487 | 0.713 | 0.961 | 0.530 |
| $HA_i$ | | 1.000 | 0.255 | 0.263 | 0.784 | 0.728 | 0.816 | 0.612 | 0.891 | 0.356 | 0.548 | 0.912 | 0.365 |
| HD | | | 1.000 | 0.841 | 0.789 | 0.824 | 0.532 | 0.873 | 0.520 | 0.967 | 0.870 | 0.487 | 0.450 |
| $HD_i$ | | | | 1.000 | 0.390 | 0.490 | 0.190 | 0.595 | 0.366 | 0.736 | 0.646 | 0.060 | 0.470 |
| HT | | | | | 1.000 | 0.981 | 0.895 | 0.961 | 0.773 | 0.830 | 0.931 | 0.898 | 0.576 |
| $HT_i$ | | | | | | 1.000 | 0.868 | 0.971 | 0.791 | 0.844 | 0.949 | 0.862 | 0.655 |
| HB | | | | | | | 1.000 | 0.870 | 0.879 | 0.602 | 0.855 | 0.957 | 0.658 |
| HBOND | | | | | | | | 1.000 | 0.803 | 0.890 | 0.984 | 0.821 | 0.573 |
| Ca | | | | | | | | | 1.000 | 0.485 | 0.853 | 0.799 | 0.881 |
| Cd | | | | | | | | | | 1.000 | 0.870 | 0.567 | 0.294 |
| Cad | | | | | | | | | | | 1.000 | 0.789 | 0.644 |
| SP | | | | | | | | | | | | 1.000 | 0.610 |
| Λ | | | | | | | | | | | | | 1.000 |

**Table 7.** Correlations between permeability or lipophilicity and hydrogen bonding descriptors.

| H-bonding Descriptor | log $P_{erm}$ | log P | log D |
|---|---|---|---|
| SP | 0.668[a)] | 0.680 | 0.381 |
| HA | 0.563 | 0.549 | 0.304 |
| HD | 0.753 | 0.882 | 0.584 |
| HT | 0.774 | 0.866 | 0.511 |
| HB | 0.709 | 0.647 | 0.379 |
| $HA_i$ | 0.470 | 0.391 | 0.345 |
| $HD_i$ | 0.509 | 0.681 | 0.309 |
| $HT_i$ | 0.783 | 0.647 | 0.529 |
| HBOND | 0.841 | 0.877 | 0.556 |
| Ca | 0.527 | 0.494 | 0.061 |
| Cd | 0.860 | 0.912 | 0.724 |
| Cad | 0.810 | 0.822 | 0.467 |

[a)] Correlation coefficient r.

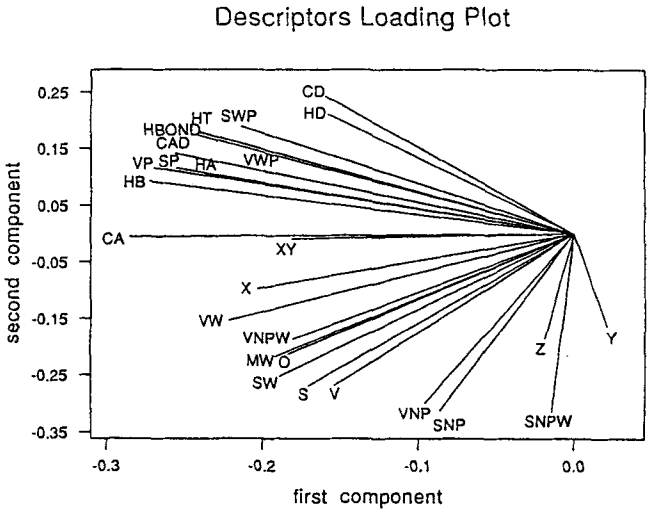


**Figure 1.** Loadings plot of the calculated descriptors

shape-related descriptors are found in the lower part of the plot. The first group is more clustered (more similar information in each of the variables) than the second group, since the latter encodes both for size and shape. In a preliminary analysis we noticed that Cd pointed away from the other H-bonding properties. This relates to the original definition in which Cd has negative values [30,31]. Therefore we decided to change the sign of the Cd values and to consider only positive values.

The eigenvectors are presented in Table 5. The first three principal components already explain 86.9% of the variance. The first component (43.4%) to a large extent contains information on the H-bonding potential, the second component (34.2%) encodes for molecular size, and the third component (9.3%) has shape information. Using these components as new variables, Eq. 4 was derived.

$$\log P_{erm} = -0.179(\pm 0.039)\ PC1 - 0.235(\pm 0.044)\ PC2 - 5.342(\pm 0.126)$$

$$n = 17 \quad r = 0.884 \quad r_{cv} = 0.836 \quad s = 0.520 \quad F = 25.2 \qquad (4)$$

where r is the correlation coefficient, $r_{cv}$ the cross-validated (leave-one-out) correlation coefficient, s the standard deviation of the regression, and F a measure for the statistical significance of the regression model. A scores plot of the compounds is interesting (Figure 2). Poorly permeating compounds are found in the upper right corner, as is tentatively indicated with two dividing lines. A very similar plot is obtained (Figure 3) taking instead of the two principal components a representative of each, e.g., the molecular weight (MW) and the number of potential hydrogen bonds (HT). This simple graphical approach might be used as a first estimate of permeability. However, this concept should be tested and may be extended with larger data sets.

Molecular size and H-bonding are often intercorrelated. For example adding a polar group to a molecule capable of H-bonding may enlarge H-bonding capacity and MW at the same time. However, numerically one effect might be larger than the other, giving rise to a modification of the lipophilicity, *cf.* Eq. 1. For the steroids used in this study (Te, Co, De and Hy) such an interrelationship can be recognized in Figure 3 (r = 0.969; n = 4). Therefore, in a series of congeneric compounds apparent linear relationships between permeability and lipophilicity or H-bonding or MW might be observed [9–11].

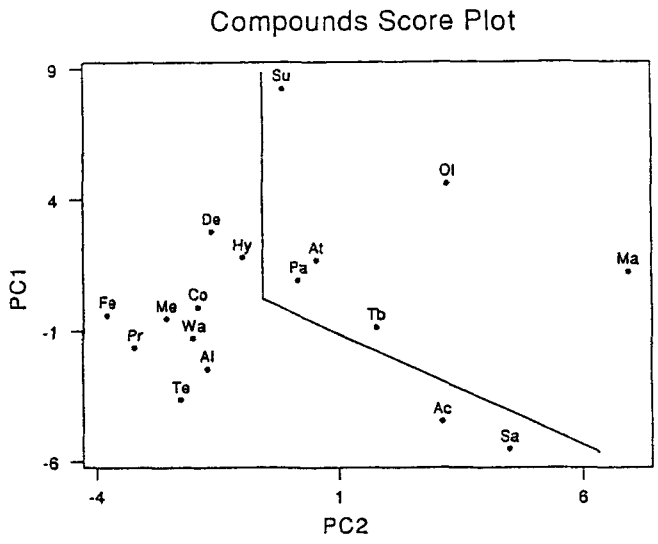


**Figure 2.** Scores plot of the first two PC's

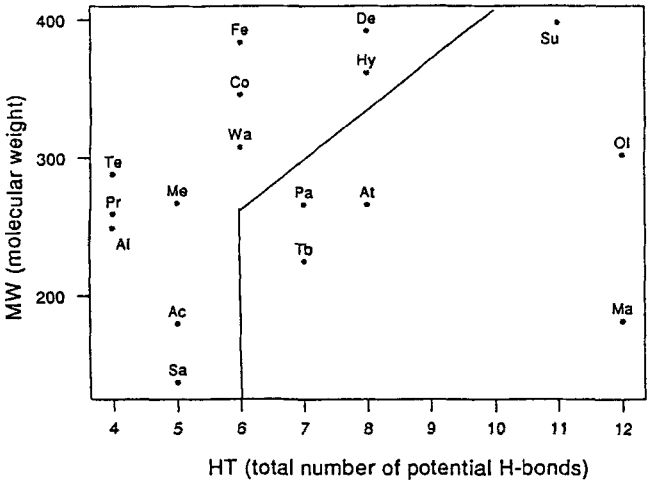


**Figure 3.** Scatter plot of a molecular size (MW) against a hydrogen bonding descriptor (HT)

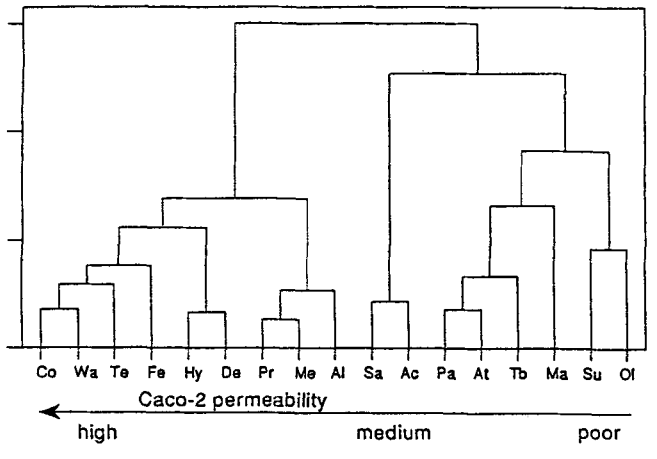


**Figure 4.** Dendrogram from a cluster analysis using Ward's method.

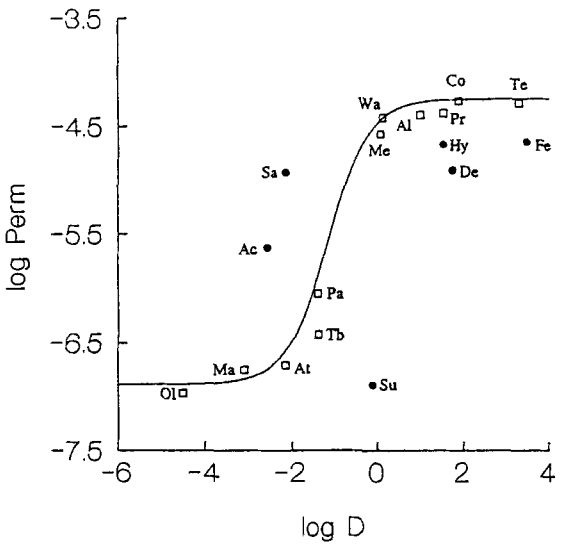


**Figure 5.** Sigmoidal relationship between permeability and experimental log D values fitted with an empirical Eq.

Finally, we applied hierarchical clustering using Ward's algorithm to the present descriptor set (Figure 4). The dendrogram shows a good classification with respect to the permeability data.

### *3.2 Relationship between Permeability and Lipophilicity*

The plot of the measured Caco-2 cell permeability data versus experimental log D values is shown in Figure 5. The distribution coefficients refer to the same pH as the permeability coefficients and therefore some of the compounds are charged under these conditions. As supported by theoretical models [57,58] and accurate experimental data for various types of biological membranes [13,59–61], a sigmoidal relationship between permeability and lipophilicity is assumed [14]. In Figure 5 an empirical sigmoidal function is drawn. Six compounds diverge from this curve, namely the two *smallest* compounds within this data set (Ac, Sa) lie above and the four *largest* ones (Su, Hy, De and Fe) below the sigmoid. Compounds with a molecular weight below 200 may use the paracellular pore pathway [62], while the size becomes the diffusion-limiting factor for larger compounds [13]. It has also been suggested that salicylic acid (Sa) is absorbed by carrier-mediated processes with proton-cotransport and/or pH-dependent anion exchange mechanisms [63] or a carboxylic acid transporter [64].

In order to understand the individual contributions of molecular size and H-bonding potential to the lipophilicity and thus their effect on permeation, both descriptors will now be considered separately.

### *3.3 Relationship between Permeability and Molecular Size*

In Table 3 different molecular size descriptors for the data set are presented. The calculation of these descriptors, with the exception of the molecular weight MW, requires computational chemistry techniques. Therefore the molecular weight is often used as a simple accessible molecular size descriptor in the literature. In Table 9

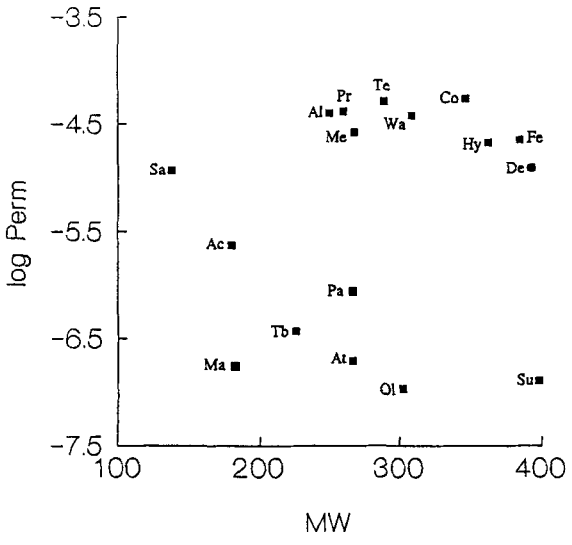


**Figure 6.** *Relationship between permeability and molecular weight.*

the correlation of these descriptors compared to the molecular weight is given. The Van der Waals volume V and the surface area S show the best correlations with MW. These three descriptors refer to the entire molecule (in contrast to VP or VNP, see below). V and S take into account the three-dimensional structure of the compounds. The molar volume of a liquid or crystal is defined by the ratio of its molecular weight and density. Since the density of most organic compounds atoms used as drugs, containing mainly C,H,N,O, do not vary greatly (range from 1 to 2), it is obvious that MW reasonably substitutes for V. The surface area S is roughly proportional to $V^{2/3}$. Thus it can be expected, that the linear correlation of S with MW for very large molecules flattens to a hyperbole.

The plot of the measured Caco-2 permeability values versus MW is shown in Figure 6. It can be seen that permeability is not simply correlated to the molecular weight of the compounds ($r = 0.155$). Similarly poor correlations are obtained with any of the other size descriptors. This is not unexpected considering the complex mechanism of permeation. Transcellular lipid permeation depends both on molecular size via lipophilicity (see Eq. 1) and the diffusion coefficient through the membrane, while paracellular pore permeation depends on molecular size via the sieving effect [13,15] and on diffusion in water.

### *3.4 Relationship between Permeability and H-bonding Capability*

In Table 4 a set of different H-bonding descriptors is presented. Again most of these descriptors require computational chemistry techniques. Exceptions are the number of H-bonds (HA, HD, HT), which are obtained by counting the number of potential H-bonds on electronegative atoms (N, O, S), and the number of all atoms capable of H-bonding (HB). The program HBOND does the same in computerized form. However, simple counting can be rather problematic and arbitrary, since bulky groups or conformational flexibility may shield the hydrogen bond strength [10]. Simple counting assumes all kinds of H-bonds to be energetically equivalent. In particular, effects of ionization are ignored. At present, it is not clear how to deal exactly with partly ionized groups.

By contrast, experimental and computational techniques may take into consideration the three-dimensional structure of the compounds and different energy values for the different kinds of H-bonds. For example the Ca and Cd parameters proposed by Raevsky and coworkers [30] come from the analysis of experimental thermodynamic H-bonding data. In Table 6 the correlation matrix of the hydrogen bonding descriptors is given. The comparison shows that descriptors for the total H-bonding capability (HT, HBOND and Cad) on the one side, and the descriptors for H-bonding donor capability (HD and Cd) on the other hand are more or less correlated ($r > 0.9$). By contrast, the descriptors for H-bonding acceptor capability (HA and Ca) show a poorer correlation ($r = 0.749$), underlining the problem of counting hydrogen bond acceptors. In addition it shows that HB and SP especially correlate with the H-bonding acceptor capability. For $\Lambda$ poor correlations are found. It has been shown earlier that $\Lambda$ accounts not only for H-bonding, but also for other interactive forces as dipolarity or polarizability [28].

In a number of papers linear correlations between permeability and hydrogen bonding have been suggested, particularly involving peptidic structures [37]. In Table 7 linear correlations between permeability, as well as lipophilicity, and hydrogen bonding descriptors are considered. However, at present it is not clear whether this relationship should be linear or not. In Figure 7 the present Caco-2 permeability values are plotted against calculated Cad values. Obviously, the correlation is not linear ($r = 0.81$). Rather, the plot suggests an interesting sigmoidal relationship for the majority of compounds, which cannot currently be rationalized. The two smallest compounds (Ac, Sa) diverge again.

### *3.5 Relationship between Permeability and Charge*

To account for repulsion and attraction phenomena at the membrane surface, some authors proposed, for a series of peptidic RGD-ana-

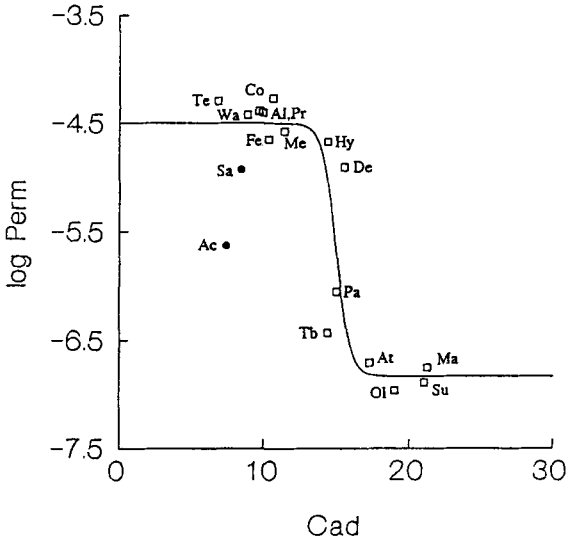


**Figure 7.** Correlation of permeability and total hydrogen bonding capacity (Cad).

logues a non-linear relationship between permeability (normalized for differences in molecular size with $MW^{1/3}$) and net charge of a compound, showing a charge optimum for single or doubly charged anions [12]. Since biological membranes are negatively charged, this is rather astonishing [2,12]. However, peptides and peptidomimetics, as in the study above, may behave differently compared to small organic molecules [44]. Using the present permeability data from Table 1, no such relationship could be found (graph not shown). However, charge effects influencing lipophilicity of the compounds are already taken into account by the use of the distribution coefficient log D.

### *3.6 Estimation of Membrane Permeation using MLR or PLS*

The principal component analysis presented above allowed us to reduce the descriptor space to a generalized hydrogen bonding and generalized molecular size descriptor. The two first principal components are complex descriptors to use for the design of new compounds, are not intuitive and require computational effort. Therefore we have tried linear (two-dimensional) combinations of the original physicochemical descriptors using multiple linear regression (MLR). A representative set is the combination of MW and various H-bonding descriptors presented in Table 8. Using molar volume V or surface area S instead of MW (see Table 9 for their intercorrelation), Eqs. of similar statistical relevance are obtained ($r = ca\ 0.9$). Slightly worse are Eqs. using ovality ($r = 0.85$) or principal axes ($r = ca\ 0.8$) as a size descriptor. However, considering the limited size of the data set no final conclusions can be drawn on the selection of the best combination. The small limited data set does not allow us to use three descriptors in the Eq.. Several compounds are of the same structural class, lowering the degrees of freedom. We conclude that a combination of very simple descriptors can be used to make a first estimate of potential absorption problems. The most simple Eq. is based on the molecular weight and the number of atoms capable of forming H-bonds. However, the limited data set does not permit a final recommendation to be made.

A partial least squares (PLS) analysis yielded one significant component ($r = 0.894$ and $r_{cv} = 0.852$) with H-bonding descriptors as the most dominant ones and size descriptors as a secondary important component. In this case PLS did not contribute to further understanding.

## 4 Conclusions

It is well recognized that the permeation of drugs through biological membranes by a passive diffusion mechanism is governed by a set of strongly intercorrelated physicochemical properties, namely lipophilicity, molecular size, hydrogen bonding capability, and the degree of ionization at the pH of interest. Since in the past congeneric series of compounds have often been studied [64,65], such interrelationships were not always visible at first sight. Generally experimental descriptors provide more reliable data than calculated parameters. Therefore there is a clear need for high-throughput methods for measuring properties related to membrane permeation for the optimization of oral bioavailability.

However, in many instances approximate estimations of the drug absorption potential may be a useful contribution to the optimization of oral bioavailability. In such cases calculated properties may be of practical use. Often attention is mainly focused on lipophilicity. However, log P calculations may be unreliable, or with some algorithms even impossible because of missing fragment values. Furthermore computed log P values refer to unionized compounds, while in fact distribution coefficients should be considered. An alternative may be the computation of the two main constituents of log P, namely molecular size and hydrogen bonding [27,31].

**Table 8.** Estimation of permeability constants (n = 17).

| $\log P_{erm}$ = | r |
|---|---|
| 0.008(±0.002) MW – 0.043(±0.008) SP – 5.165(±0.605) | 0.833 |
| 0.007(±0.003) MW – 0.487(±0.120) HA – 5.313(±0.729) | 0.742 |
| 0.001(±0.002) MW – 0.568(±0.135) HD – 4.030(±0.833) | 0.754 NS[b] |
| 0.005(±0.002) MW – 0.343(±0.058) HT – 4.344(±0.619) | 0.851 |
| 0.008(±0.002) MW – 0.573(±0.080) HB – 4.974(±0.503) | 0.890 |
| 0.006(±0.003) MW – 0.298(±0.108) $HA_i$ – 5.485(±0.861) | 0.607 ns[a] |
| 0.001(±0.003) MW – 0.293(±0.138) $HD_i$ – 4.831(±1.051) | 0.513 NS |
| 0.005(±0.002) MW – 0.361(±0.059) $HT_i$ – 4.021(±0.631) | 0.857 |
| 0.004(±0.002) MW – 0.364(±0.050) HBOND – 3.898(±0.558) | 0.892 ns |
| 0.011(±0.003) MW – 0.424(±0.078) Ca – 5.239(±0.606) | 0.830 |
| 0.000(±0.002) MW – 0.328(±0.053) Cd – 3.483(±0.670) | 0.861 NS |
| 0.005(±0.002) MW – 0.201(±0.029) Cad – 4.100(±0.566) | 0.883 |

[a] ns = non-significant Eq. at the 95% confidence interval level [64].
[b] NS = non-significant Eq.

**Table 9.** Correlation matrix (r) of MW versus other evaluated molecular size descriptors (n = 17)

| | V | VW | VNP | VP | S | SW | O | X | Y | Z | X/Y | X/Z |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MW | 0.946 | 0.742 | 0.871 | 0.401 | 0.918 | 0.854 | 0.669 | 0.392 | 0.397 | 0.611 | 0.176 | 0.078 |

We have demonstrated that using a linear combination of a suitable molecular size and hydrogen bonding descriptor, thus without knowledge of the distribution coefficient, reasonable estimates of permeability can be made. This may be done either via a graphical approach (Figs. 2 and 3, or the method suggested by Leahy *et al.* [13] using separate permeability-lipophilicity curves for each molecular weight), or by using a set of descriptors and multiple linear or partial least squares regression.

As recently repeatedly demonstrated, hydrogen bonding descriptors are of paramount importance in describing membrane permeation. The advantage of computer-calculated H-bonding descriptors over *ad hoc* ones is that larger series of compounds, and potentially whole databases, can easily be computed. We have presented here a new descriptor Cad, which can be readily calculated from molecular structure. It is strongly correlated to the HBOND descriptor. Simple alternatives may be a count of all atoms capable of H-bonding or the computation of the polar surface area. Further studies with larger data sets should allow the optimal choice and should clarify how ionization affects the hydrogen bonding descriptors.

**Acknowledgements**

We thank Dr. Paul Gerber and Dr. Peter Straub (F. Hoffmann-La Roche Ltd.) for their assistance in molecular size calculations and Dr. Subash Basak (University of Minnesota) for the calculation of HBOND descriptors. Discussions with Dr. Werner Rubas (Genentech Inc.), Prof. Per Artursson (University of Uppsala), Prof. Klaus Müller and Dr. Klaus Gubernator (F. Hoffmann-La Roche Ltd ) have been highly appreciated.

## 5 References

[1] Gan, S., Eads, C., Niederer, T., Bridgers, A., Yanni, S., Hsyu, P.H., Pritchard, F.J. and Thakker, D., *Drug Develop. Ind. Pharm.* 20 615–631 (1994).

[2] Conradi, R.A., Burton, P.S. and Borchardt, R.T., in: Pliska, V., Testa, B, van de Waterbeemd, H. (Eds.), *Lipophilicity in Drug Action and Toxicology*, VCH, Weinheim 1996, pp. 233–252.

[3] Hillgren, K.M., Kato, A. and Borchardt, R.T., *Med. Res. Revs.* 15, 83–109 (1995).

[4] Artursson, P., *Crit. Rev. Ther. Drug Car. 8*, 305–330 (1991).

[5] Hunter, J., Hirst, B.H. and Simmons, N.L., *Pharm. Res. 10*, 743–749 (1993).

[6] Schinkel, A.H., Wagenaar, E., Van Deemter, L., Mol, C.A.A., and Borst, P., *J.Clin.Invest.* 96, 1698–1705 (1995).

[7] Tanaka, Y., Taki, Y., Sakane, T., Nadai, T., Sezaki, H. and Yamashita, S., *Pharm. Res. 12*, 523–528 (1995).

[8] Artursson, P. and Karlsson, J., *Biochem. Biophys. Res. Comm. 175*, 880–885 (1991).

[9] Conradi. R.A., Hilgers, A.R., Ho, N.F.H. and Burton, P.S., *Pharm. Res. 8*, 1453–1460 (1991).

[10] Conradi, R.A., Hilgers, A.R., Burton, P.S. and Hester, J.B., *J. Drug Target. 2*, 167–171 (1994).

[11] Chikdale, E.G., Ng, K.-Y., Burton, P.S. and Borchardt, R.T., *Pharm. Res.* 11, 412–419 (1994).

[12] Rubas, W., Cromwell, M., Gadek, T., Narindray, D. and Mrsny, R., *Mat. Res. Soc. Symp. Proc.* 331, 179–185 (1994).

[13] Leahy, D.E., Lynch, J. and Taylor, D.C., in: Prescott, L.F. and Nimmo, W.S., (Eds.), *Novel Drug Delivery and Its Therapeutic Application*, John Wiley, New York 1989, pp. 33–44.

[14] Macheras, P., Reppas, C. and Dressman, J.B., *Biopharmaceutics of Orally Administered Drugs*, Ellis Horwood, Hemel Hempstead 1995.

[15] Camenisch, G., Folkers, G. and Van de Waterbeemd, H., *Pharm. Acta Helv.*, (1996) in press.

[16] Yalkowski, S.H., Slunick, T.G. and Flynn, G.L., *J. Pharm. Sci. 63*, 691–695 (1974).

[17] Potts, R.O. and Guy, R.H., *Pharm.Res. 12*, 1628–1633 (1995).

[18] Pidgeon, Ch., Ong, S., Liu, H., Qiu, X., Pidgeon, M., Dantzig, A.H., Munroe, J., Hornback, W.J., Kasher, J.S., Glunz, L. and Szczerba, T., *J.Med.Chem. 38*, 590–594 (1995).

[19] 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. and Hicks, J., *J.Med.Chem. 38*, 1446–1455 (1995).

[20] Palm, K., Luthman, K., Ungell, A.-L., Strandlund, G. and Artursson, P., *J.Pharm.Sci. 85*, 32–39 (1996).

[21] Testa, B. and Seiler, P., *Arzneim. Forsch. (Drug Res.) 31*, 1053–1058 (1981).

[22] Van de Waterbeemd, H. and Testa, B., *Adv. Drug Res. 16*, 85–225 (1987).

[23] El Tayar, N., Tsai, R.-S., Testa, B., Carrupt, P.-A., Leo, A., *J.Pharm. Sci. 80*, 590–598 (1991).

[24] Testa, B., Carrupt, P.-A., Gaillard, P. and Tsai, R.-S., in: Pliska, V., Testa, B, van de Waterbeemd, H. (Eds.), *Lipophilicity in Drug Action and Toxicology*, VCH, Weinheim 1996, pp. 49–71.

[25] Charton, M. and Charton, B.I., *J.Theoret.Biol. 99*, 629–644 (1982)

[26] Charton, M. and Charton, B.I., *J.Org.Chem. 44*, 2284–2288 (1979)

[27] Raevsky, O.A., in: Van de Waterbeemd, H., Testa, B., Folkers, B. (Eds.), *Computer-Assisted Lead Finding and Optimization*, VCH, Weinheim 1997, in press.

[28] El Tayar, N., Testa, B. and Carrupt, P.-A., *J. Phys. Chem.* 96, 1455–1459 (1992).

[29] Abraham, M., *Chem.Soc.Rev. 22*, 73–83 (1993).

[30] Raevsky, O.A., Grigor'ev, V., Kireev, D. and Zefirov, N., *Quant. Struct. Act. Relat. 11*, 49–63 (1992).

[31] Raevsky, O.A., Schaper, K.-J. and Seydel, J.K., *Quant. Struct. Act. Relat. 14*, 433–436 (1995).

[32] Abraham, M.H. and Chadha, H.S. in: Pliska, V., Testa, B. and Van de Waterbeemd, H. (Eds.) *Lipophilicity in Drug Action and Toxicology*, VCH, Weinheim. 1996, pp. 311–337.

[33] Abraham, M., Chadha, H.S. and Mitchell, R.C., *J.Pharm.Pharmacol. 47*, 8–16 (1995).

[34] Seiler, P., *Eur.J.Med.Chem. 9*, 473–479 (1974).

[35] Young, R.C., Mitchell, R.C., Brown, Th.H., Ganellin, C.R., Griffiths, R., Jones, M., Rana, K.K., Saunders, D., Smith, I.R., Sore, N.E. and Wilks, T.J., *J. Med. Chem. 31*, 656–671 (1988).

[36] Ganellin, C.R., Brown, Th.H., Griffiths, R., Jones, M., Mitchell, R.C., Rana, K., Saunders, D., Smith, I.R., Sore, N.E., Wiley, T.J. and Young, R.C., in: Silipo, C. and Vittoria, A. (Eds.), *QSAR: Rational Approaches to the Design of Bioactive Compounds*, Elsevier, Amsterdam 1991, pp. 103.

[37] Burton, P.S., Ho, N.F.H., Conradi, R.A. and Hilgers, A.R., *Pharm. Res. 10*, 636–637 (1993).

[38] Van de Waterbeemd, H. and Kansy, M., *Chimia 46*, 299–303 (1992).

[39] Ter Laak, A.M., Tsai, R.S., Donné-Op den Kelder, G.M., Carrupt, P.A.,Testa, B. and Timmerman, H., *Eur. J. Pharm. Sci. 2*, 373–384 (1994).

[40] El Tayar, N., Tsai, R.-S., Testa, B., Carrupt, P.-A., Hansch, C. and Leo, A., *J.Pharm.Sci. 80*, 744–749 (1991).

[41] Potts, R.O. and Guy, R.H., *Pharm. Res. 10*, 635–636 (1993).

[42] Stein, W.D., in: *The Movement of Molecules Across Cell Membranes*, Academic Press, London 1967, pp. 65–125.

[43] Kempf, D.J., *Perspect. Drug Discov. Des. 2*, 427–436 (1994).

[44] Stewart, B.H., Chan, O.H., Lu, R.H., Reyner, E.L., Schmid, H.L., Hamilton, H.W., Steinbaugh, B.A. and Taylor, M.D., *Pharm.Res.* 12, 693–699 (1995)
[45] Daylight Chemical Information Systems Inc., Irvine, USA, MedChem94 database.
[46] HYBOT 5.0 is commercially available from Prof. O.A. Raevsky.
[47] Gerber, P. and Müller, K., *J. Comput. Aided Mol. Des. 9*, 251–268 (1995).
[48] TSAR, Oxford Molecular Group, The Magdalen Centre, Oxford Science Park, Oxford OX4 4GA, UK (FAX +44–1865–784601)
[49] Basak, S., *Pharm. Res. 13*, 775–778 (1996).
[50] Leo, A.J., *Chem. Revs. 93*, 1281–1306 (1993).
[51] Bodor, N., Gabanyi, Z. and Wong, C.K., *J. Amer. Chem. Soc.* 111, 3783- 3786 (1989).
[52] Yang, G., Lien, E.J. and Guo, Z., *Quant.Struct.-Act.Relat. 5*, 12–18 (1986)
[53] Lien, E.J., Lien, L.L. and Gao, H., in: Sanz, F. (Ed.), *Trends in QSAR and Molecular Modeling 94*, Prous, Barcelona 1996, pp 94.
[54] Pearlman, R.S., in: Dunn, W.J., Block, J.H., Pearlman, R.S. (Eds.), *Partition Coefficient. Determination and Estimation.*, Pergamon Press, New York 1986, pp. 3–20.
[55] SYSTAT 5.0 from Systat Intelligent Software Inc., Evanston, IL, USA, (1995).
[56] SCAN 1.0 for Windows from Minitab Inc., 3081 Enterprise Drive, State College, PA 16801, USA.
[57] Ho, N.F.H., Park, J.Y., Morozowich, W. and Higuchi, W.I. in: Roche, E.B. (Ed.), *Design of Biopharmaceutical Properties through Prodrugs and Analogs* APhA/APS, Washington D.C. 1977, pp. 136–227.
[58] Wilschut, A., Ten Berge, W.F., Robinson, P.J. and McKone, T.E., *Chemosphere* 30, 1275–1296 (1995).
[59] Oldendorf, W.H., *Proc.Soc.Exp.Biol.Med.* 147, 813–816 (1971)
[60] Hilgers, A.R., Conradi, R.A. and Burton, P.S. *Pharm. Res. 7* (1990) 902–910.
[61] Potts, R.O. and Guy, R.H., in: *Drug Permeation Enhancement*, Hsieh, D.S. (Ed.), Marcel Dekker, New York 1994, pp. 148–155.
[62] Lennernäs, H. *Pharm.Res. 12*, 1573–1582 (1995).
[63] Takanaga, H., Tamai, I. and Tsuji, A., *J.Pharm.Pharmacol. 46*, 567–570 (1994).
[64] Tsuji, A., Takanaga, H., Tamai, I. and Terasaki, T., *Pharm.Res. 11*, 30–37 (1994).
[65] Houston, J.B., Upshall, D.G. and Bridges, J.W., *J. Pharmacol. Exp. Ther. 189*, 244–254 (1974).
[66] Potts, R.O. and Guy, R.H. *Pharm. Res. 9*, 663–669 (1992).
[67] Kubinyi, H., *QSAR: Hansch Analysis and Related Approaches*, VCH, Weinheim, 1993.

Received on July 5th, 1996;
accepted on September 23rd, 1996.