Articles

Presentation of a Structurally Diverse and Commercially Available Drug Data Set for Correlation and Benchmarking Studies

Christian Sköld,[†] Susanne Winiwarter,[‡] Johan Wernevik,[§] Fredrik Bergström,[§] Leif Engström,[‡] Ruth Allen,[||] Karl Box,[||] John Comer,[||] Jon Mole,[||] Anders Hallberg,[†] Hans Lennernäs,[⊥] Torbjörn Lundstedt,[†] Anna-Lena Ungell,[‡] and Anders Karlén*[,†]

Division of Organic Pharmaceutical Chemistry, Department of Medicinal Chemistry, and Department of Pharmacy, BMC, Uppsala University, Sweden, Drug Metabolism and Pharmacokinetics & Bioanalytical Chemistry and Lead Generation; DMPK & Physical Chemistry, AstraZeneca R&D Mölndal, Sweden, and Sirius Analytical Instruments Ltd., Riverside, East Sussex, United Kingdom

Received June 30, 2005

A multivariate analysis of drugs on the Swedish market was the basis for the selection of a small, physicochemically diverse set of 24 drug compounds. Factors such as structural diversity, commercial availability, price, and a suitable analytical technique for quantification were considered in the selection. Lipophilicity, pK_a, solubility, and permeability across human Caco-2 cell monolayers were measured for the compiled data set. The results show that, by use of a physicochemically diverse data set, experimental responses over a wide range were obtained. The paper also shows how experimental difficulties due to the diversity of the data set can be overcome. We anticipate that this data set can serve as a benchmark set for validation of new experimental techniques or in silico models. It can also be used as a diverse starting data set for the development of new computational models.

Introduction

In drug discovery there is a strong demand for early information on compound characteristics such as absorption, distribution, metabolism, excretion, and toxicity (referred to as ADME/Tox) and properties related to these phenomena. Various in vitro ADME/Tox screens have significantly increased the amount of experimental data generated in the early drug discovery process. In addition to these experimental techniques, there is a strong need for different in silico methods that can predict such properties. For example, ADME/Tox filters that can sort out compounds with undesirable ADME/Tox properties can be applied in virtual screening or compound design so as to reduce compound attrition rates.[1] The basis for constructing these predictive models is the existence of high-quality experimental data. Several data sets exist in the literature and have been used for building models that predict ADME/Tox-related properties such as permeability,[2] fraction absorbed (FA%),[3] solubility,[4] CYP inhibition,[5] plasma protein binding,[6] etc. However, many of these data sets have been compiled from experimental results from different labs, which may not have used the same experimental conditions. Another possible problem with the existing data sets is that different sets of compounds have been used for studying and modeling different properties. This makes it difficult if the aim is to construct robust and more complex quantitative models of, for example, drug absorption that would be based on log *P*/log *D*, solubility, dissolution rate, intestinal permeability, metabolic stability, and carrier-mediated membrane transport data.[7] In this case a relevant and common data set that represents the available ADME/Tox space is essential. New experimental tools are also constantly being developed that aim to complement in vitro and/or in vivo experiments. For example, surface plasmon resonance (SPR) technology and PAMPA have been introduced to model intestinal permeability.[8,9] To validate these novel techniques, a data set with experimental permeability or ADME/Tox data is often compiled and correlated to the measured data. The compounds included in these data sets most often differ, which makes an independent comparison between the techniques difficult. It would therefore be valuable if a small, commercially available, and structurally diverse data set existed, for which high-quality in vitro and/or in vivo ADME/Tox data can be determined in the same lab. This would make interlaboratory comparisons of data easier and more relevant.

The aim of the present study is to derive such a "benchmark" data set of reasonable size having the following characteristics: (a) druglike, (b) physicochemically diverse, (c) commercially available and inexpensive, and (d) amenable to analytical measurements. Herein we describe how such a data set consisting of 24 compounds was compiled by use of multivariate analysis and experimental design. Also, highly toxic compounds (as stated in the material safety data sheet from the respective supplier) were avoided when possible. To start the characterization of these compounds, we have (a) determined their pK_a values, by a pH-metric or pH-UV-based technique; (b) determined their log *P* values by a pH-metric log *P* technique (ionizable compounds) or by liquid−liquid chromatography with UV detection (nonionizable compounds); (c) determined the aqueous intrinsic solubility values by a pH-metric technique (ionizable compounds) or a shake-plate method with UV detection (nonionizable compounds); (d) determined lipophilicity,

* To whom correspondence should be addressed: phone +46-18-471-4293; fax +46-18-471-4474; e-mail anders.karlen@orgfarm.uu.se.
[†] Division of Organic Pharmaceutical Chemistry, Department of Medicinal Chemistry, Uppsala University.
[‡] Drug Metabolism and Pharmacokinetics & Bioanalytical Chemistry, AstraZeneca R&D.
[§] Lead Generation, DMPK & Physical Chemistry, AstraZeneca R&D.
[||] Sirius Analytical Instruments Ltd.
[⊥] Department of Pharmacy, Uppsala University.

10.1021/jm0506219 CCC: $33.50
Published on Web 10/14/2006

expressed as chromatographic capacity factors, by reversed-phase high-performance liquid chromatography (RP-HPLC); and (e) determined the permeability across Caco-2 cell monolayers. We hope that other groups are willing to supplement the herein-started characterization of these 24 pharmaceutical compounds and that this data set is used as a benchmark data set for different types of studies in the future.

Materials and Methods

Preparation of the Data Set. In a previous study we constructed a database of 673 pharmaceutical compounds used on the Swedish market in the year 2000.[10] As the source for this database we used the Swedish list of drugs (FASS 2000).[11] For the present study we have used the same database but also included new compounds from the list of drugs in the year 2001 to give 691 pharmaceutical compounds (data set 1). We used the same protocol for database creation as in the previous study but with the difference that all zwitterions in the database were added in the neutral form. Shortly, compounds with a molecular weight (MW) above 900, polymer structures, polypeptides (more than nine amino acids), inorganic compounds, and metal-containing complexes were removed. Structures were drawn with ChemDraw[12] or Sybyl sketch, saved in SMILES[13] format, and as such read into Sybyl.[14] Acids and bases were drawn in their neutral form, and only permanently charged compounds were considered as charged species. Chirality was used as given in FASS. If no chirality was specified, either the more active enantiomer was used, if such information was available, or one chiral form was selected randomly. Three-dimensional starting structures were generated with Concord.[15] All structures were optimized by the semiempirical AM1 method[16] by use of the MOPAC[17] program within Sybyl, with the keywords PRECISE, NOMM, and XYZ added. For the charged compounds, the keyword CHARGE was also used.

Calculation of Theoretical Molecular Descriptors. The 28 molecular descriptors used in this study were in large part the same as those used in our previous study.[10] They can tentatively be divided into general descriptors, hydrogen-bond donor and hydrogen-bond acceptor descriptors, and general hydrogen-bonding descriptors.

(A) General Descriptors. MW = molecular weight, V = molecular volume, S = molecular surface area, O = ovality, E_{LUMO} = energy of lowest unoccupied molecular orbital, E_{HOMO} = energy of highest occupied molecular orbital, H = hardness, DM = dipole moment, NPSA = nonpolar surface area, log P_{Cr} = log P calculated according to the method developed by Ghose and Crippen,[18] log P_{Mor} = log P calculated according to the method developed by Moriguchi et al.,[19,20] log P_{ACD} = log P calculated with ACD/Labs,[21] log $D_{ACD\_6.5}$ = log D at pH 6.5 calculated with ACD/Labs, log $D_{ACD\_7.4}$ = log D at pH 7.4 calculated with ACD/Labs.

(B) Hydrogen-Bond Donor Descriptors. Q_H = highest partial charge on a hydrogen atom, HBD = number of hydrogen-bond donors, ΣQ_H = sum of the partial charge of all hydrogen atoms attached to an O, N, or S atom, $CWPSA_{HBD}$ = sum of absolute charge-weighted surface area of all hydrogen atoms attached to an O, N, or S atom.

(C) Hydrogen-Bond Acceptor Descriptors. Q_{MN} = absolute value of lowest partial charge of a non-hydrogen atom, HBA = number of hydrogen-bond acceptors, ΣQ_{MN} = sum of absolute partial charge of all negatively charged O, N, S, and F atoms, $CWPSA_{HBA}$ = sum of absolute charge-weighted surface area of all negatively charged N, O, S, F, Cl, Br, and I atoms.

(D) General Hydrogen-Bonding Descriptors. HB = sum of HBD and HBA, PSA = molecular surface area of all N, O, S, and attached H atoms, PSA_{NO} = molecular surface area of all N, O, and attached H atoms, PSA_{rel} = PSA/S × 100, PSA/NPSA, PSA_{cw} = sum of (absolute) charge-weighted surface area of all N, O, halogen, and attached H atoms.

Data Analysis. Multivariate data analysis was performed with SIMCA-P+[22] at default settings. The selection of the commercial data set was accomplished by using a multivariate design approach in the first three principal components. Multivariate design, MVD,[23,24] is a method for selecting a set of representative observations among a large set of data. First, a number of descriptors are calculated for each compound (multivariate characterization). Next, a PCA (principal component analysis) model is fitted to the data to find the principal properties (principal components) that best describe the data. Then, a representative choice of observations according to the principal properties can be made. The principal components in a PCA model are mathematically independent (orthogonal) and limited in number, properties that make them well-suited for statistical experimental design. There are a number of approaches for a multivariate design, all with the aim of maximizing the information content of the selected observations. In this case the selection from principal properties was based on a face-centered central composite design.

Measurements of pK_a. Before we attempted to experimentally measure the 24 molecules, their pK_a values were calculated by use of ACD software.[25] Having these values aids the experimental design. The program also identifies the ionizable groups and denotes whether they are acidic or basic, which helps when the ionization is assessed and later also for the lipophilicity models. The structures were then examined for the likelihood that the molecule would have good UV absorbance and whether the ionizable group(s) was(were) close enough to the chromophores to change the UV spectra as a function of pH.

pH-Metric pK_a. The pK_a values of molecules with weak chromophores or with at least one ionizable group remote from the chromophore were analyzed by the pH-metric technique. In this technique, a solution of the compound is acid−base-titrated over a wide pH range and the pK_a is calculated from the shape of the titration curve, which is governed by pK_a values of the sample.[26] Compounds with log P_{ACD} values below 2.5 were assumed to be water-soluble across the entire pH range at pH-metric concentrations (at least 0.5 mM for pK_a values between 3 and 10), and their pK_a values were measured pH-metrically in aqueous solution. It should be noted that higher concentrations of sample are required to measure pK_a values outside the range 3−10. This is because measured pH changes less in response to unit additions of acid or base titrant at low and high pH, so changes in the shape of the titration curve caused by ionization of the sample appear smaller in relation to the background signal.[27] It was assumed that molecules with log P_{ACD} greater than 2.5 might precipitate over part of the pH range, leading to poor-quality data for aqueous pK_a measurement. Their p$_s K_a$ values were measured with the pH-metric technique in water−solvent mixtures, and their aqueous pK_a values were determined by Yasuda−Shedlovsky extrapolation.[28]

pH-UV pK_a. The pK_a values of molecules with ionizable groups that were capable of interacting with UV-absorbing chromophores because of structural proximity or conformational flexibility were selected for pK_a measurement by the pH-UV technique.[29] Because the analytical concentrations required for pH-UV methods (typically 0.05 mM over the entire pH range) are significantly lower than for pH-metric methods, it was assumed that molecules with log P_{ACD} below 4.5 would be sufficiently water-soluble for measurement in aqueous solution.

Measurements of Lipophilicity: *(A) pH-Metric Log P.* Log P values were determined pH-metrically by titrating the sample in a two-phase water−octanol mixture and calculating the result from the shape of the titration curve, which is governed by the pK_a and log P values of the sample. In effect, the presence of octanol shifts the pK_a to a higher value (known as the p$_o K_a$) for acids and a lower value for bases.[30] As well as measuring log P for the un-ionized species, this technique also measures log P for the ionized species of the sample. With these log P and pK_a values, a profile of log D versus pH can be drawn, and log D values at specified pH values can be determined by interpolation.[31]

(B) Capacity Factor by RP-HPLC. A reversed-phase HPLC method with a linear wide-polarity mobile phase gradient was used as a high-throughput method for determining lipophilicity.[32] By use of a calibration set of five well-studied standard compounds, the retention time was related to a standardized capacity factor, k. The

retention time for each studied compound was then transformed to standardized *k*, which is used as a measure of the lipophilicity.

Measurements of Solubility: *(A) pH-Metric Solubility.* Aqueous intrinsic solubility values for the 20 ionizable compounds were measured by the pH-metric CheqSol technique.[33] Intrinsic solubility is calculated from the pK_a values along with information about pH changes that occur when the molecule dissolves and precipitates in response to additions of acid and base titrant while close to equilibrium in the presence of excess precipitate. The procedure also provides a measurement of the kinetic solubility. Comparing the kinetic to the intrinsic solubility gives an estimation of the level of supersaturation that could be expected. Intrinsic solubility is the equilibrium solubility of the un-ionized form of an ionizable molecule, where equilibrium solubility is the concentration of compound in a saturated solution when excess solid is present and the solution and solid are at equilibrium. Intrinsic solubility thus has a clear definition, and the measured results should be independent of the measurement technique used.[34] Kinetic solubility is the concentration of a compound in solution at the time when an induced precipitate first appears. In this work, precipitation was induced by pH change and detected by the appearance of turbidity; the values reported here for kinetic solubility may differ from values obtained by other techniques such as inducing precipitation by adding water to a solution in a solvent.

(B) Shake-Plate Solubility. Nonionizable compounds were measured by a shake-plate method with UV detection.[35]

Measurements of Permeability across Caco-2 Cells. To check for possible solubility problems during the transport experiments, compound solubility was investigated in HBSS, pH 7.4. Permeability of the Caco-2 cell monolayers for all the compounds was measured in the absorptive direction, that is, from the apical to the basolateral side of the membrane (A to B), by an automated procedure. The measurements were performed with a pH 6.5/7.4 gradient, thereby mimicking the situation in the intestine. A donor concentration of 10 μM was utilized for all compounds to balance low solubility and analytical detection problems during the experiments. For analysis, two different short RP-HPLC columns, one appropriate for extremely hydrophilic compounds, were employed before detection by mass spectrometry.

Results and Discussion

Definition of a Druglike, Physicochemically Diverse, Commercially Available and Inexpensive Data Set. Many different databases of drugs and drug-like compounds exist, and these have been profiled to enable extraction of descriptors that portray, for example, drug likeness. The most well-known of these are Comprehensive Medicinal Chemistry,[36] World Drug Index,[37] and MDL Drug Data Report.[38] In this study we used 691 pharmaceutical compounds, which are used on the Swedish market and which are taken from the Swedish list of drugs (FASS 2000 and FASS 2001), as the drug database (data set 1). To define a data set that meets our criteria, we first checked for commercial availability in the Available Chemicals Directory (ACD).[39] Out of the 691 drugs, 450 were commercially available and had a catalog price that ranged from a few cents to more than $100,000 per gram. It was decided to include only those compounds with a price of less than $800 per gram, which gave 370 compounds. These compounds were further screened to remove compounds that were included only in topically administered preparations (ointment, cream, gel, etc.) or those with an unusual administration route (patch, rinsing fluid, nail polish, etc.). The following ATC classes (Anatomical Therapeutic Chemical Classification System)[40] of drugs were also removed: A01, stomatological preparations; A05, bile and liver therapy; A06, laxatives; A09, digestives, including enzymes; A11, vitamins; A16A, other alimentary tract and metabolism products; B05, blood substitutes and perfusion solutions; C05A, antihemorrhoidals for topical use; D, dermatologicals; M02, topical products for joint and muscular pain; N01AB, halogenated hydrocarbons (general anesthetic); R02, throat preparations; R07A, other respiratory system products; S01J, diagnostic agents (eye); and V, various. After this procedure, 284 compounds remained.

The next task was to select a subset that still covered the physicochemical space of data set 1 by statistical molecular design. Principal component analysis was performed in Simca with the 691 compounds and the 28 molecular descriptors. Thirteen principal components were significant according to the standard criteria in Simca. The first three principal components were used for further analysis. These explained 74% of the variance in the X-block and thus represent a good summary of the properties of the 28 variables. The first principal component (PC1) mainly encodes size-related descriptors; the second (PC2), lipophilicity descriptors; and the third (PC3), hydrogen-bonding and charge descriptors. These new variables (principal components) were used in the experimental design to select the data set. A face-centered central composite design (Figure 1) was used that consisted of the corners of the cube (red squares), the axis points (blue squares), and the center point (yellow square). In this design each principal component is explored at three levels: low (minus sign), center (zero sign), and high (plus sign).

By inspection of the three principal components and choosing from the 284 filtered and commercially available compounds, two were selected from each of the corner regions of the box. Two compounds were also selected near the center and one compound from each of the axis points. Selection of substances with high absolute values in the principal components (guided by the Euclidean distance to origin) was preferred to maximize the span of the compound space. The compounds in the corners of the box were selected from those 25% having the highest (plus sign in the principal component) or lowest (minus sign in the principal component) values in each of the three components. However, for PC1 the percentage had to be adjusted (lower 35% and upper 15%) since very few compounds were identified with the 25% cutoff. If a compound was identified in the corner or center points that was structurally very similar to one already selected, it was substituted for a less similar compound. We selected two compounds from each corner because these compounds have more extreme properties, which may be more challenging in the experimental work (very lipophilic or very polar, etc). Even though the two compounds in each corner have the same sign, their values differ somewhat, which makes them physicochemically slightly different. Very toxic compounds, as defined in the material safety data sheet (e.g., antineoplastic agents, carcinogens, potent neurotoxins, etc.), were avoided in the selection process. We also avoided drugs with a very low bioavailability, since pharmacokinetic data may be difficult to obtain for such drugs. An adjustment of the data set was also performed to allow for the analytical procedure. This gave 24 compounds altogether (data set 2, Figure 2) that have good chemical diversity. The total cost for buying at least 1 g of all 24 compounds corresponds to less than $1500, although the price fluctuates somewhat over time.

Characteristics of the Compiled Data Set. A comparison of the minimum, maximum, and mean values of some of the key descriptors in data sets 1 and 2 are shown in Table 1. To a large extent the values in data set 2 reflect those in data set 1 except for the more extreme max and min values. For example, the maximum number of HBD atoms in data set 1 corresponds to 19, while in data set 2 it only corresponds to 8. However, inspection of the mean values shows that these are similar for

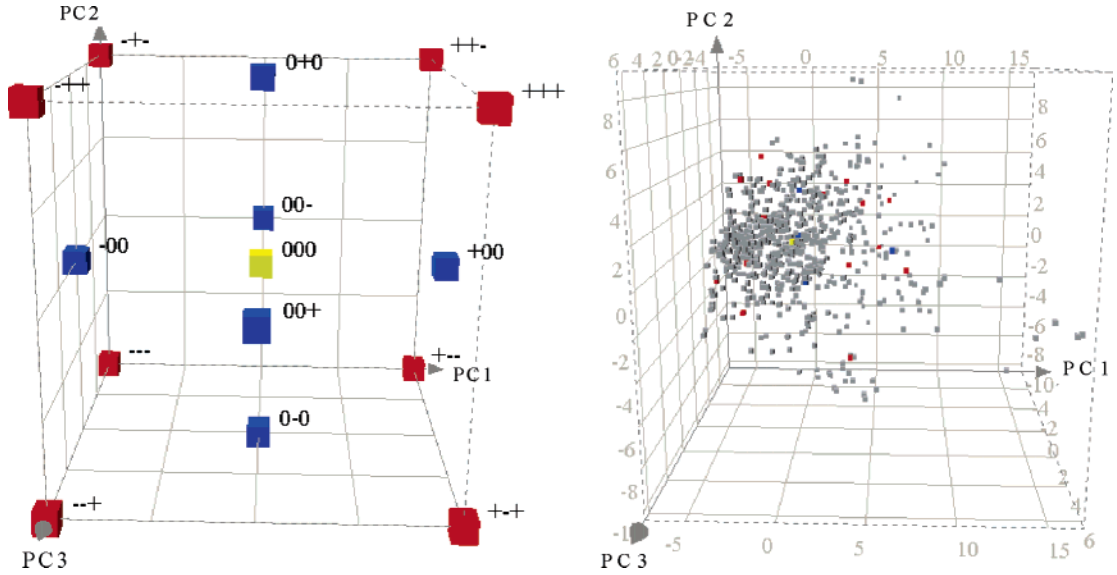

Figure 1. Illustration of a face-centered central composite design with red squares representing the experiments at the corners of the cube, blue squares at the axis points, and a yellow square at the center point (left) and a score plot showing scores in PC1−PC3 for data set 1 with compounds of data set 2 colored according to these codes (right).

Table 1. Comparison of Data Sets 1 and 2 with Respect to Some Common Molecular Descriptors

	data set 1, 691 compounds			data set 2, 24 compounds		
	min	max	mean	min	max	mean
MW	60	854	347	114	777	349
PSA	0	373	93	8	246	99
log P_{Mor}	−6.4	7.6	1.9	−2.0	5.3	1.9
log $D_{ACD\_6.5}$	−10.6	12.3	0.74	−5.0	4.8	0.94
HBD	0	19	2.4	0	8	2.7
HBA	0	19	4.9	1	14	4.7

the two data sets. This can partly be explained by the fact that only nine compounds can be found in data set 1 that have nine or more HBD atoms.

Measurements of pK_a. The structures were first examined to ascertain whether they had groups that would ionize in aqueous solution. Four molecules did not − carbamazepine, carisoprodol, fenofibrate, and prednisone − and they were excluded from the pK_a measurements. pK_a values were successfully measured for the remaining 20 molecules by pH-metric or pH-UV techniques. The results for the compounds are shown in Table 2.

pH-Metric pK_a. pK_a analysis determined with the pH-metric technique of compounds that have a relatively low log P_{ACD} (amantadine, captopril, erythromycin, L-dopa, and thiamazole) was easy. Tetracycline, which also belongs to this group, was measured pH-metrically in aqueous solution in case the absorbance change associated with the ionization of the tertiary amine would be too small for an accurate pH-UV result (even though it has excellent UV absorbance in general). For molecules with log P_{ACD} greater than 2.5, p_sK_a values were measured in water−solvent mixtures, and their aqueous pK_a values were determined by Yasuda−Shedlovsky extrapolation.[28] Flupenthixol and sulindac were in this category, and their pK_a values were easily determined from p_sK_a values measured in water−methanol mixtures. The pK_a values of meclizine and terfenadine and one of the pK_a values of folic acid were also measured pH-metrically, but the measurements were challenging (see Supporting Information for details).

pH-UV pK_a. pK_a analysis, by the pH-UV technique, of amiloride, bendroflumethiazide, chlorzoxazone, hydrochlorothiazide, metoclopramide, and tinidazole was easy. However, the pK_a values of glipizide, levothyroxine, and folic acid (three pK_a values) were more challenging (see Supporting Information for details). The log P_{ACD} of chlorprothixene was greater than 4.5, and its pK_a was determined from p_sK_a values measured in water−methanol mixtures, although its weak UV absorbance called for high-quality spectral data to avoid calculation difficulties.

Measurements of Lipophilicity. The most widely used measure of lipophilicity in modeling biological partition/distribution is log *P*, describing the partitioning between *n*-octanol and water. Log *D*, a related lipophilicity descriptor, describes the distribution of a compound between the two phases at a given pH. A multitude of methods are utilized for assessing lipophilicity.[41] In the present study, both a pH-metric technique to obtain log *P* and log *D* and a fast RP-HPLC method that results in capacity values *k* were used.

Measurement of Log P. Log *P* octanol/water values for the 20 ionizable compounds were measured by the pH-metric technique, and results were obtained for all except L-dopa (below the limit of detection of −2) and folic acid (too insoluble in octanol; see Supporting Information for details). The pH-metric log *P* technique is generally easy at analytical concentrations of 0.5 mM and above for compounds with log *P* values between about −1 and 5 and whose acidic pK_a values are below about 10 and basic pK_a values are above about 4. Log *P* analysis for molecules in this category included amantadine, amiloride, captopril, chlorzoxazone, erythromycin, hydrochlorothiazide, metoclopramide, and sulindac. Of the remaining molecules, bendroflumethiazide and thiamazole have high acidic pK_a values and their p_oK_a values in the two-phase titration shift to higher pH, where pH measurement is less sensitive to changes caused by ionization of the samples. Similarly, flupenthixol and tinidazole have basic groups with low pK_a values whose p_oK_a values in the two-phase titration shift to lower pH where pH measurement is less sensitive. High concentrations of these compounds (up to 7 mM for tinidazole) were required to obtain good signal-to-noise ratios at extreme pH for successful log *P* measurement. While their pK_a values were at the extremes for pH-metric log *P*, thiamazole and tinidazole have low log *P* values, and the shifts in pK_a were small enough to remain within the measurable pH range. Tetracycline has two acidic and one

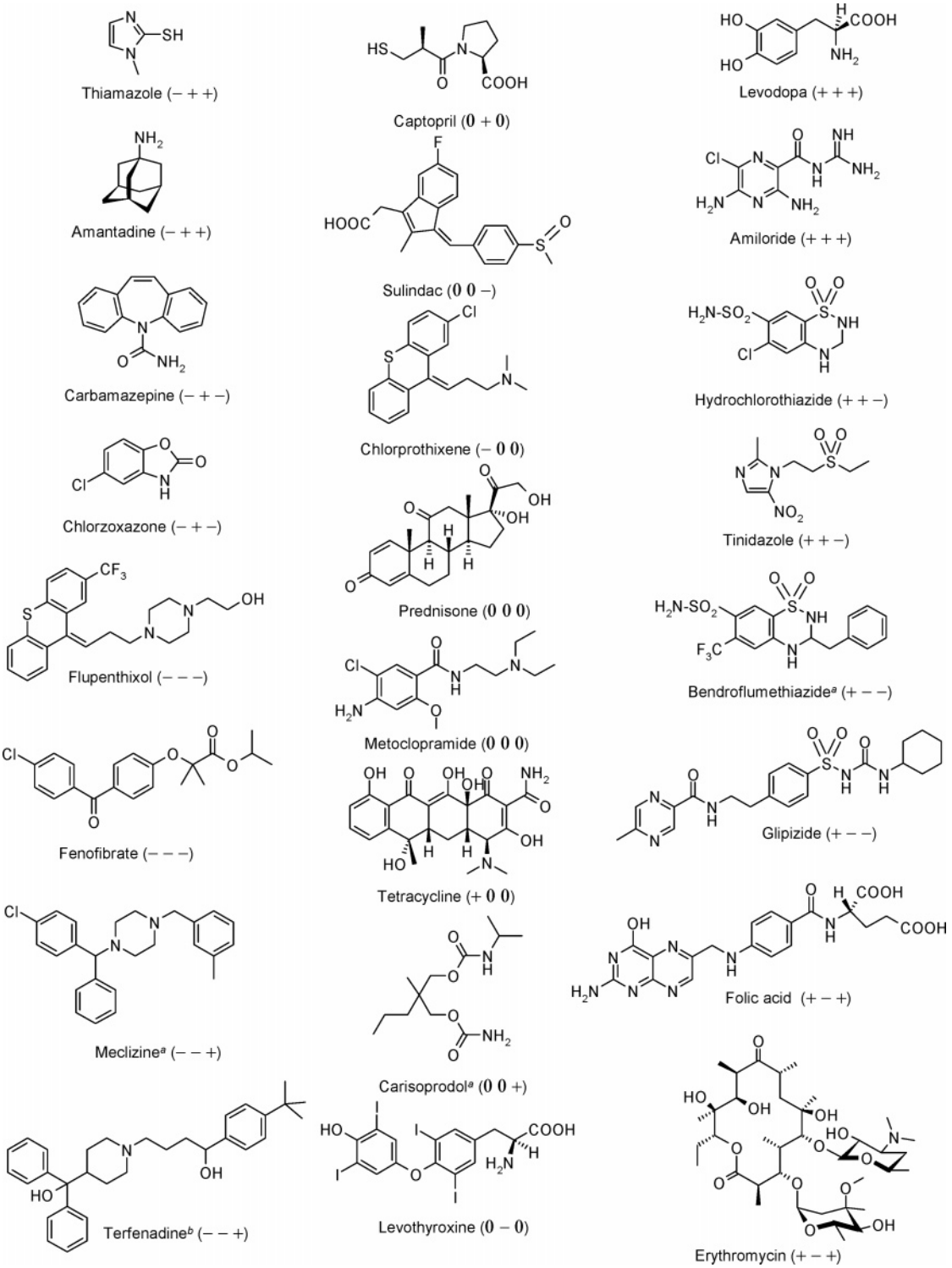

Figure 2. Chemical structures of data set 2. The level of the principal components PC1, PC2, and PC3 for each compound is shown in parentheses. The compounds are ordered according to the sign in PC1−PC3. [a]The *R*-configuration is used in calculations. [b]The *S*-configuration is used in calculations.

basic pK_a and is highly water-soluble. The shifts in pK_a were small in the two-phase titration, and given the complicated ionization and partition model, the experimental log *P* can only be reported with confidence as being less than zero. Of the four nonionizable compounds, fenofibrate was measured by liquid−liquid chromatography with UV detection.[42] This technique was also used to measure carbamazepine, but the results from two different columns differed significantly and were not considered reliable. Log *P* measurement of prednisone was not done during this study. Carisoprodol has no UV absorption and its log *P* could not be measured by this technique. Obtained log *P* and log *P*(ion) values are shown in Table 2. Knowledge about log *P*, log *P*(ion), and pK_a, as obtained by the pH-metric method, can be used to calculate log *D* at any given pH. Log *D* values at pH 7.4 by this method are shown in Table 2.

Measurement of Capacity Factors by RP-HPLC. Lipophilicity was also measured by chromatographic separation, resulting in the capacity factor *k*. Values could be obtained for all but three compounds: Folic acid had too low a signal to be detected after

Table 2. Experimental Caco-2 Permeability, pK_a, Log *P*, Log *D*, HPLC Capacity Factor, Kinetic and Intrinsic Solubility, and Supersaturation Ratio for Data Set 2

compd	CAS registry no.	$P_{app-Caco-2}$ (10^{-6} cm s^{-1})	pK_a	log *P* (neutral)	log *P* (ion)	log *D* (pH 7.4)	*k*[a]	kinetic sol. (μg/ mL)	intrinsic sol. (μg/mL)	supersatn ratio
amantadine	768-94-5	22	10.49	2.41	−0.59[b]	−0.33	7.2	2622.0	2119.0	1.2
amiloride	2609-46-3	0.71	8.63	−0.32		−1.57	4.4			
bendroflumethiazide	73-48-3	1.3	8.46, 9.74	1.95	−0.63[c]	1.91	10.9	159.0	19.6	8.1
captopril	62571-86-2	0.07	3.45, 9.80	0.35		<−2	6.4			
carbamazepine	298-46-4	40					9.8		110.3[d]	
carisoprodol	78-44-4	22					10.3			
chlorprothixene	113-59-7	18	9.06[e]	5.48	1.51[b]	3.81	11.9	1.4	0.16	8.8
chlorzoxazone	95-25-0	65	8.16	2.11	−0.44[d]	2.04	9.3	1208.0	416.0	2.9
erythromycin	114-07-8	0.06	8.87	2.88	−0.81[b]	1.4	9.7			
fenofibrate	49562-28-9	nd[f]		5.1[g]		5.1	15.0		0.7[c]	
flupenthixol	2709-56 -0	13	3.29,[e] 7.57[e]	4.68	1.79[b]	4.29	12.2	35.2	41.8	0.8
folic acid[h]	59-30-3	0.87	2.16, 3.79, 4.47, 7.90			<−2	nd[i]	30.0	2.3	12.8
glipizide	29094-61-9	19	5.13[j]	2.58	−0.46[c]	0.38	9.1	59.5	1.4	42.5
hydrochlorothiazide	58-93-5	0.71	8.77, 9.79	−0.10		−0.12	8.3	2501.0	629.0	4.0
levodopa	59-92-7	0.21	2.38, 8.75, 9.76, >12	<−2		<−2	<2			
levothyroxine	51-48-9	3.0	6.84,[j] 8.72[j]	3.21[k]	0.35[l]	2.53	10.1	93.4	42.8	2.2
meclizine	569-65-3	lr[m]	2.23,[e,n] 7.24[e,n]	6.20[o]	2.18[b,o]	5.98	16.5	0.09[p]	0.13[p]	0.8
metoclopramide	364-62-5	14	9.25	2.74	−0.11[b]	0.93	7.3	379.0	75.5	5.0
prednisone	53-03-2	23					9.2		206.6[d]	
sulindac	38194-50-2	6.2	4.08[e]	3.42	0.05[c]	0.38	8.3	70.7	10.8	6.5
terfenadine	50679-08-8	9.3	9.25[e]	5.42	2.24[b]	3.58	12.5	0.15[q]	0.009[q]	17.2
tetracycline	60-54-8	0.53	3.31, 7.50, 9.14	<0		<0	8.1	9662.0	440.0	22.0
thiamazole	60-56-0	48	11.45	−0.22		−0.22	<2			
tinidazole	19387-91-8	30	1.80	−0.94		−0.94	6.5			

[a] HPLC capacity factor used as a measure of the lipophilicity of the compounds based on two or more independent measurements. Errors are within 0.1 as judged from the standard deviation of repeated measurements. [b] Cation. [c] Anion. [d] Solubility by shake-plate. [e] By Yasuda−Shedlovsky extrapolation. [f] Not detectable and low recovery. [g] By Sirius ProfilerLDA. [h] Combination of pH-metric and pH-UV used in pK_a measurements. [i] Not detectable. [j] Fast D-PAS. [k] Species XH2 (+1−1). [l] Species XH (+1−1−1). [m] Not reliable due to low recovery. [n] MDM cosolvent. [o] 0.01M ionic strength [p] CheqSol MDM cosolvent. [q] CheqSol cosolvent extrapolation.

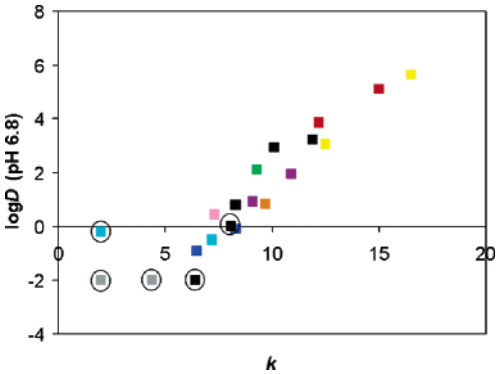

Figure 3. Correlation between the measured HPLC capacity factor (*k*) and pH-metric log *D* (pH 6.8). The compounds from the eight corner points have different colors and symbols, but the two compounds at each corner point have the same color (in the case when both compounds are present). The axis points are colored black and the center point is colored pink (data were obtained for only one compound). Encircled data points have not fully quantified values in one or both of the descriptors but are plotted for completeness.

the chromatographic separation, and both levodopa and thiamazole were found to have extremely low values, below 2. The determined capacity factors are shown in Table 2.

Comparison of the Two Methods. A comparison of the lipophilicity values obtained by the pH-metric method and RP-HPLC method is shown in Figure 3. Since the *k* values were obtained in a buffer at pH 6.8, pH-metric log *D* values at pH 6.8 was also used for this comparison. The relationship between *k* and log *D* in Figure 3 is linear ($R^2 = 0.92$, excluding the five encircled nonquantified data points), which has also been shown previously when gradient elution was used in RP−HLPC.[43] In the figure, the two compounds corresponding to each of the corners of the box are labeled with the same color (when data for both compounds could be determined). The six axis compounds are colored black and the compound in the center is colored pink. For the most part, the corner compounds, which should have similar physicochemical properties, are grouped, even though only lipophilicity is considered.

The obtained *k* values span the region from 4.4 (<2 for levodopa and thiamazole) to 16.5, and log *P* ranged between <−2 and 6.2, which illustrates the physicochemical diversity of data set 2.

Measurement of Solubility. Provided a compound dissolves in ionized form at low or high pH, measurements of intrinsic solubility by CheqSol in aqueous solution are generally easy, as was the case for amantadine, bendroflumethiazide, chlorzoxazone, hydrochlorothiazide, metoclopramide, and tetracycline. Once their pK_a values had been determined, CheqSol aqueous solubility measurements were also quite straightforward for folic acid, glipizide, and levothyroxine. Of these compounds, amantadine was a nonchaser while the other compounds were chasers. A chaser is a compound that readily forms a supersaturated solution and a nonchaser is a compound with no tendency to from a supersaturated solution.[34] The highest supersaturation ratio occurred for glipizide, whose kinetic solubility was 42.5 times higher than its intrinsic solubility. Chlorprothixene and sulindac were also easy to measure, but they precipitated at first as nonchasers that converted after 10−20 min to chasers with much lower solubilities; these lower values are reported in Table 2. Three basic molecules, flupenthixol, meclizine, and terfenadine, did not dissolve in aqueous solution at low pH because their salts were insoluble. Their aqueous intrinsic solubilities were determined by extrapolation from solubilities measured in solvent−water mixtures. This procedure was straightforward for flupenthixol (a nonchaser), but was challenging for meclizine and terfenadine (see Supporting Information for details). Results were obtained for all except captopril, L-dopa, thiamazole, and tinidazole, which were too soluble to measure, and amiloride and erythromycin, which behaved anomalously (see Supporting Information for

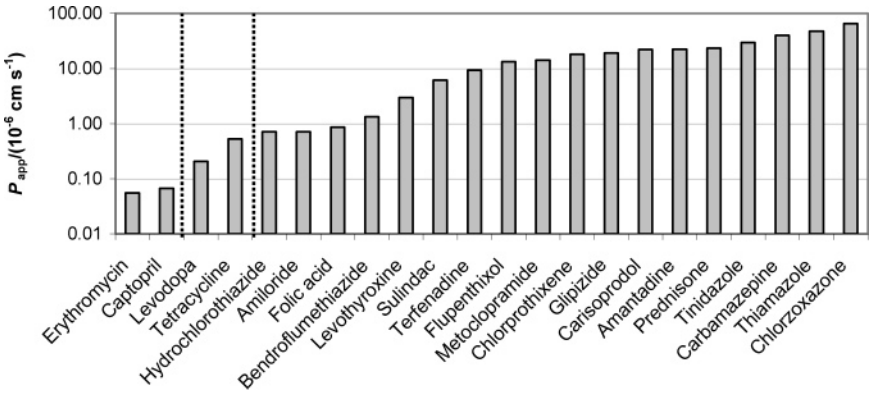

Figure 4. Plot of the obtained in vitro P_{app} values in human Caco-2 cells for data set 2, divided into groups with low, medium, and high permeability values.

Table 3. Compiled Pharmacokinetic Literature Data[a] for Data Set 2

compound	FA% (%)	protein binding (%)	V_{ss} (L kg^{-1})	clearance (mL min^{-1} kg^{-1})	$t_{1/2}$ (h)	dose[b] (mg)	transport route[c]	abs class[d]	perm class[e]
amantadine[49,50]	50−90	67	6.6 ± 1.5	4.8 ± 0.8	16 ± 3.4	100	P	m−h	h
amiloride[51]	80−90	40	17 ± 4	9.7	21 ± 3	5−10	P	h	h
bendroflumethiazide[52−54]	>90	95	1−2	4−7	3−9	10−20	P	h	h
captopril[55−57]	70−80	25	0.01	12.8	1−2	75	C	m	l
carbamazepine[58]	>90	25−50	0.6−0.9	1.2 ± 0.4	5−7	400−600	P	h	h
carisoprodol[59]		40−67			1.7 ± 0.8	700	P		h
chlorprothixene[60,61]	>90		10−20	5−20	26 ± 14	100−150	P	h	h
chlorzoxazone[62,63]	>90			3.5 ± 1.5	1.1 ± 0.34	250	P	h	h
erythromycin[64,65]	>80	84 ± 3	0.8 ± 0.4	9 ± 4	1.6 ± 0.7	1000	C	h	l
fenofibrate[66]	>90	>99	0.9	0.3−0.5	19−27	200	P	h	
flupenthixol[67,68]	>90		12−17	6−9	14−36	5	P	h	h
folic acid[69,70]	>90				2	10	C	h	h
glipizide[71−73]	>95	98	0.2 ± 0.02	0.5 ± 0.2	3.4 ± 0.7	7.5−15	P	h	h
hydrochlorothiazide[74−79]	55	58 ± 17	0.8 ± 0.3	4.9 ± 1.1	2.5 ± 0.2	50	P	m	h
levodopa[80−85]	100	>10	0.9−1.6	5	1−2	200	C	h	m
levothyroxine[86]	80					0.1−0.2	P	m	h
meclizine						25			
metoclopramide[87−89]	>95	40 ± 4	3.4 ± 1.3	6.2 ± 1.3	5.4 ± 1.4	20	P	h	h
prednisone[90−92]	>95	75 ± 2	1 ± 0.11		3.3 ± 1.3	50	P	h	h
sulindac[93,94]	90	94	3−13			200−400	P	h	h
terfenadine[95−97]	>95					60	P	h	h
tetracycline[98,99]	>80	25−65	1−2	2−3	6−9	500	P	h	m
thiamazole[100]	>90	<10	0.7	2−3	3−5	10−30	P	h	h
tinidazole[101−103]	>95		7.2	0.73	9−14	650	P	h	h

[a] Fraction absorbed, plasma protein binding, volume of distribution, clearance, elimination half-life, dose, transport route, and classification regarding fraction absorbed and permeability values are given. [b] Common dose per dosing event. [c] P = mainly passive absorption, C = mainly carrier-mediated transport. [d] Absorption classified according to FA% as high (h, FA% > 80%), medium (m, FA% = 20−80%), or low (l, FA% < 20%) absorption. [e] Permeability classified according to $P_{app-Caco-2}$ as high (h, $P_{app} > 0.67 \times 10^{-6}$ cm s^{-1}), medium (m, $P_{app} = 0.11 \times 10^{-6}$ to 0.67×10^{-6} cm s^{-1}), or low (l, $P_{app} < 0.11 \times 10^{-6}$ cm s^{-1}) permeability.

details). Of the four nonionizable compounds, carbamazepine, fenofibrate, and prednisone were measured by a shake-plate method with UV detection.[35] Carisoprodol has no UV absorption and its solubility could not be measured by this technique.

Measurement of Permeability across Human Caco-2 Cells. Many studies have demonstrated a good correlation between in vitro human Caco-2 cell permeability and extent of absorption in humans.[44,45] Especially when passive permeability is the dominating mechanism, this prediction has high accuracy.[46,47] The technique has for a long time been considered the golden standard for in vitro assessment of intestinal permeability in humans and is routinely used in drug discovery screening. In the present study, permeability measurements over Caco-2 cell monolayers at a physiological pH gradient of pH 6.5/7.4 were performed in an automated robot system, which was optimized for rapid screen analysis. The solubility of the 24 compounds was checked to ensure that the subsequent permeability measurement could be performed in the Caco-2 cell system. Three of the compounds (fenofibrate, levothyroxine, and meclizine) had an HBSS solubility below 10 μmol/L at pH 7.4 (data not shown).

By use of an automated procedure, reliable permeability values could be obtained for 22 of the 24 compounds (Table 2), showing that the robot system is well-suited for investigating a variety of drug or druglike compounds. Two compounds (fenofibrate and meclizine) had a very low recovery ($\ll$80%) and P_{app} values could therefore not be determined accurately. The low recovery of these compounds was most likely due to their high lipophilicity (RP-HPLC k values were above 15 and pH-metric log D values were above 5), which can lead to adsorption to plastics, cell monolayers, or filter supports. Two compounds had low P_{app} values ($<0.11 \times 10^{-6}$ cm s^{-1} = FA% < 20%), two had moderate values, and 18 had high permeability values ($P_{app} > 0.67 \times 10^{-6}$ cm s^{-1} = FA% > 80%) according to a ranking defined for the robotic system used,[48] which is based on the ranking originally proposed by Ungell and Karlsson[44] (Figure 4). In drug data sets used for modeling

intestinal permeability, compounds with high permeability are often more frequent than compounds with low permeability. Inspection of data set 2 shows that all compounds are given orally and it is therefore not surprising to find that many of the compounds have high permeability. However, by selecting physicochemically diverse compounds, both low- and medium-permeability compounds could be identified. There is also a noticeable diversity among the highly permeable compounds (P_{app} values between 0.70 and 65, which is a factor of about 100).

Before we attempt to make predictive models of intestinal permeability by use of this data set, further studies are needed to delineate the possible transport routes and uptake or efflux systems relevant for these compounds. A tentative classification of the transport route for each compound (Table 3) indicates that 19 compounds can be considered to be mainly passively transported across the intestinal membrane, while four compounds predominantly use various carrier-mediated transport systems.

In Table 3 the compounds have been classified into low, medium, and high absorption on the basis of the literature FA% values. The compounds were also classified on the basis of the P_{app} values as defined above. For the passively absorbed compounds, a good correlation was observed between the two systems, which has also been noted previously.

Compilation of Literature Pharmacokinetic Data. To create predictive ADME models, basic clinical pharmacokinetic data is necessary. We have therefore gathered this information from the literature for data set 2. The pharmacokinetic variables presented in Table 3 (plasma protein binding, V_{ss}, $t_{1/2}$, and clearance) have been identified in clinical reports of healthy adults. In the table, a combination of values for each ADME variable has been selected from the literature and FA% has been calculated according to eq 1. The definition of fraction absorbed is the fraction of the drug that reaches the inside of the intestinal side. This means that the bioavailability (F) is dependent on three major parameters: the fraction absorbed (FA%) and the first-pass extraction of the drug in the gut wall (E_G) and by the liver (E_H):

$$F = \mathrm{FA\%}(1 - E_G)(1 - E_H) \quad (1)$$

Suggestions on Data Set Usage. The 24 compounds can be used as a test set for testing already-derived models of FA%, Caco-2 cell permeability, lipophilicity, and solubility as well as a validation set for new experimental methods. However, we hope that more experimental data will be compiled in the future. The data set can also be used on its own for building and validating models by dividing it into a training set and a test set. For this purpose one compound from each of the corners, one compound from the center, and the compounds at the axis points (see Figures 1 and 2) will make a well-covered training set (for example, amantadine, amiloride, bendroflumethiazide, captopril, carisoprodol, chlorprothixene, chlorzoxazone, erythromycin, fenofibrate, hydrochlorothiazide, levothyroxine, meclizine, metoclopramide, sulindac, and tetracycline). The remaining compounds could be used as an external test set.

Conclusions

Many data sets in the literature contain high-quality experimental ADME-related data. Unfortunately, many of these include compounds that are not commercially available. Others are not structurally diverse, which is an important property when attempting to build general predictive models of ADME-related properties. We propose here a physicochemically diverse data set consisting of 24 drugs, where the cost of buying the entire data set (at least 1 g of each compound) is less than $1500. To meet the requirements of a "high-quality" data set, it is preferable that the same laboratory measures each property. Due to the rather small size of the data set, this should be possible for most properties that can be measured by in vitro experiments. However, some types of more complicated data will still most likely be gathered from different laboratories (e.g., fraction absorbed, bioavailability, volume of distribution, and clearance).

To start the characterization of this data set, we have determined pK_a, log P, and solubility. Furthermore, lipophilicity has been determined by reversed-phase HPLC and permeability across human Caco-2 cell monolayers (apical to basolateral direction) at a pH gradient of 6.5/7.4 in a robotic system. The results show that, by using a physicochemically diverse data set, experimental responses over a wide range were obtained. The measurement of these properties also illustrated some of the experimental difficulties that may occur due to the physicochemical diversity of the data set. An explanation of the difficulties and a discussion of how these were overcome can be found in the Supporting Information. Nevertheless, the compounds can be used in routine screening methods and help define the usability of new experimental methods.

Similar to the measurements in this study, missing data for some compounds will also inevitably occur in future characterizations, but this information may be valuable when investigating the scope and limitations of experimental procedures. However, due to the redundancy in each corner group, a problem with one of the compounds is not critical with respect to diversity since a similar compound is available. For example, intestinal permeability values could not be obtained for fenofibrate ($- - -$) and meclizine ($- - +$), which belong to two corner groups in the multivariate design. However, since permeability values for flupenthixol ($- - -$) and terfenadine ($- - +$) (the other compounds in these groups) could be measured, the diversity of the data set is still maintained. Concerning theoretical modeling with this data set, some missing values should not be a problem since statistical tools such as PCA and PLS can handle missing data.

We have determined properties such as pK_a, lipophilicity, solubility, and permeability for this data set and suggest that it now can serve as a benchmark data set used in, for example, validation of new experimental tools or for building or validating theoretical models of these properties. We hope that other groups are willing to help us to supplement the herein-started characterization of these compounds and that this data set could serve as a benchmark data set for many types of studies in the future.[104]

Experimental Section

Calculated Molecular Descriptors. The molecular descriptors were mainly calculated in Sybyl[14] by internal or external scripts. AM1 calculations were used for calculation of E_{HOMO}, E_{LUMO}, DM, and partial atom charges. HBD and HBA were calculated according to definitions in Sybyl. The descriptors surface area (S), volume (V), polar surface area (PSA = molecular surface area of all nitrogen, oxygen, sulfur, and attached H-atoms, PSA_{NO} = molecular surface area of all nitrogen, oxygen, and attached H-atoms) were calculated by use of Savol[105,106] with probe radius set to 0.0 in order to obtain the van der Waals surface. Ovality ($O = S/[4\pi(3V/4\pi)^{2/3}]$), relative PSA ($PSA_{rel} = PSA/S \times 100$), nonpolar surface area (NPSA $= S -$ PSA), PSA divided by NPSA, sum of hydrogen-bond donor and acceptor atoms (HB = HBD + HBA), and hardness [$H = (E_{LUMO} - E_{HOMO})/2$] were calculated from the respective functions. The charge-weighted polar surface areas (PSA_{CW}, $CWPSA_{HBD}$, and $CWPSA_{HBA}$) were calculated by summing up the charge-weighted

surface areas of defined atoms. The charge-weighted surface area of a given atom is the surface area of that atom multiplied by the absolute value of its partial atomic charge. Log P_{Mor} was calculated by the method of Moriguchi et al.[19,20] as implemented in a Sybyl Programming Language (SPL) script included in the Sybyl distribution. Log P_{Cr} was calculated by the method by Ghose and Crippen.[18] Log P_{ACD}, log $D_{ACD6.5}$, and log $D_{ACD7.4}$ were calculated with ACD/Labs.[21] ACD/ Labs could not provide data for all structures (28 missing values for log P_{ACD}, 48 for log $D_{ACD6.5}$, and 48 for log $D_{ACD7.4}$), and in those cases the descriptor values were treated as missing data.

Compounds: *(A) Compounds Used in Measurements of pK_a, Log P, and Solubility.* Amantadine hydrochloride, amiloride hydrochloride, bendroflumethiazide, carbamazepine, carisoprodol, chlorprothixene hydrochloride, chlorzoxazone, erythromycin, fenofibrate, flupenthixol dihydrochloride, folic acid, glipizide, hydrochlorothiazide, levodopa, levothyroxine, meclizine dihydrochloride, metoclopramide hydrochloride, prednisone, sulindac, tetracycline, thiamazole, and tinidazole were obtained from Sigma−Aldrich (Stockholm, Sweden). Captopril and terfenadine were obtained from ICN Biomedicals (Aurora, Ohio).

(B) Compounds Used in Measurements of Capacity Factors and Permeability across Caco-2 Cells. Levodopa and glipizide were obtained from Acros Organics (Geel, Belgium); captopril, sulindac, carbamazepine, prednisone and meclizine were from Göteborgs Termometerfabrik (Göteborg, Sweden); amiloride hydrochloride hydrate, erythromycin, folic acid, metoclopramide, tetracycline free base, chlorprothixene hydrochloride, and tinidazole were from Sigma−Aldrich (Stockholm, Sweden); and levothyroxine was obtained from VWR International AB (Stockholm, Sweden). Amantadine, bendroflumethiazide, carisoprodol, chlorzoxazone, fenofibrate, flupenthixol dihydrochloride, hydrochlorothiazide, terfenadine, and thiamazole were obtained internally from AstraZeneca's Compound Collection.

Apparatus Used for Measuring pK_a, Log *P*, and Solubility. Titrations of ionizable compounds were done on a GLpK_a instrument (Sirius Ltd, Forest Row, Sussex, U.K.) fitted with a double-junction pH electrode, together with a D-PAS attachment with fiber optic UV dip probe for pH-UV assays. Results were calculated from experimental data by use of RefinementPro 2, Fast D-PAS, and CheqSol software.[107] Log *P* octanol−water of nonionizable compounds was measured by liquid−liquid chromatography on a Sirius ProfilerLDA instrument (Sirius Ltd, Forest Row, Sussex, U.K.). Solubility of nonionizable compounds was measured on a μSOL Evolution instrument (pION Inc., Woburn MA).

Measurements of Capacity Factors by RP-HPLC. The lipophilicity was investigated by use of a Waters Xterra C_{18} column, 100 × 2.1 mm, with 3.5 μm particle size (Waters Corp.). The mobile phase consisted of acetonitrile (HPLC grade; Rathburn, Scotland) in 10 mM ammonium acetate (p.a. quality, Merck, Germany), pH 6.8. For the first 2 min 5% acetonitrile was used, followed by a linear gradient for 15 min to end up with 95% acetonitrile, which was used for 3 min to wash the column. The initial conditions were then used for 3 min to equilibrate the column before the next injection. The flow rate was 0.3 mL/min and the injection volume was 10 μL. Samples were analyzed on a Micromass LCT (Micromass UK Limited).

A mixture of warfarin, verapamil, metoprolol, propranolol, and felodipine at concentrations between 10 and 50 μM in acetonitrile/water (1/1) was used as calibration compounds for lipophilicity determinations.

Measurements of Permeability across Human Caco-2 Cells: *(A) Solubility in the Experimental Setup*. Solubility in HBSS (Hank's balanced salt solution), pH 7.4, was examined as follows: The compound was dissolved to 10 mM in dimethyl sulfoxide (DMSO), followed by dilution to 100 μM in HBSS, thus having 1% DMSO (99.9% anhydrous, Sigma−Aldrich, Germany) as cosolvent. After incubation at room temperature for 24 h, the solution was filtered through a glass-filled polypropylene filter (Whatman Inc.) and quantified by RP-HPLC-UV with a calibration curve. The same system as described above for measurement of the capacity factors was used for the analysis. The mobile phase was, however, different, with a gradient from 5% to 95% acetonitrile in 0.1% formic acid (p.a. quality, Merck, Germany). UV absorption monitored between 200 and 500 nm on a diode array detector was used for quantification.

(B) Cell Culture. Caco-2 cell line was obtained from the ATCC (American Type Culture Collection, Rockville, MD). All tissue culture medium was purchased from Gibco (Paisley, Scotland, U.K.). The cells were cultured according to Neuhoff et al.[108] and were seeded at a density of 3.7×10^5 cells/cm^2 on polycarbonate filters on a 24-transwell plate (3396 HTS Transwell, Costar, Cambridge, MA). Passages between 28 and 43 were used after 15 days on filters.

(C) Caco-2 Cell Monolayer Permeability Experiments. The Caco-2 cell monolayers were washed once with HBSS prior to start. TEER (transepithelial electrical resistance) was measured both before and after performing all the transport experiments. Experiments were made in the A to B direction and in duplicate. TEER measurement, sampling, dilution, and incubation were performed by a laboratory robot: Freedom EVO 200 equipped with a Te-MO 96 and a TEER Station (Tecan, Männerdorf, Switzerland). Transport buffer, 800 μL (HBSS, pH 7.4)< was first dispensed to the basal side of the monolayer. The assay was then initiated when 200 μL of each compound was added to the apical side at 10 μM (all test compounds were diluted in HBSS, pH 6.5, with 1% DMSO as cosolvent). Samples were withdrawn before addition and at 45 and 120 min postaddition of test compound. Aliquots (2 μL and 200 μL) were withdrawn from the apical donor compartment and the basolateral receiver compartment, respectively. During incubation the transwell plates were placed in a shaking incubator at 480 rpm and 37 °C between the sampling up to 120 min. All samples were analyzed subsequently.

(D) Permeability Calculations. The apparent permeability coefficient (P_{app}, centimeters per second) was calculated according to

$$P_{app} = (\Delta Q/\Delta t)/(AC_D) \quad (2)$$

where $\Delta Q/\Delta t$ is the cumulative amount of test compound transported over time, A is the surface area of the monolayer membrane (square centimeters), and C_D is the average drug concentration in the donor chamber over the period which $\Delta Q/\Delta t$ was determined.

Recovery during the experiment was calculated as

$$\text{recovery (\%)} = [(QD_{end} + QR_{end})/QD_{start}] \times 100 \quad (3)$$

where $QD_{end} + QR_{end}$ is the final amount of compound found on both donor and receiver sides and QD_{start} is the initial amount on the donor side. Recovery between 80% and 120% was considered to define a high-quality experiment.

(E) Analysis. The analytical method described in this study represents a standardized approach. The Caco-2 matrix, containing MES and HEPES buffers and acetonitrile, is a relatively uncomplex matrix for LC-MS/MS analysis. No suppression patterns were observed.

Separation was made on a Thermo HyPURITY Elite C18 column (5 μm and 3 × 30 mm). Polar analytes (levodopa, captopril, amiloride, hydrochlorothiazide, methimazole, and tetracycline) were separated on a Waters Atlantis C18 (3 μm and 4.6 × 50 mm). The mobile phase consisted of two phases and was used for both columns: water phase A contained 2% acetonitrile (HPLC grade; Rathburn, Scotland) and 0.2% formic acid (p.a. quality; Merck, Germany); and phase B consisted of acetonitrile and 0.2% formic acid. The gradient was linear during the first 1.8 min, with the percentage of phase B in phase A increased from 4% to 90%. Between 1.8 and 2.6 min the phase consisted of 10% A and 90% B. After the elution of the compounds, the columns were reequilibrated with initial conditions for 0.9 min to be ready for the next sample. Elution was performed at a flow rate of 0.55 mL/min by use of a gradient pump (Agilent 1100), and the injection volume was 5−20 μL (autosampler HTS PAL, CTC, Switzerland). All

compounds were then analyzed quantitatively by use of Micromass Quattro Ultima (Waters Corp.) with a generic tune file.

Most of the analytes were detected by SRM (selected reaction monitoring). Cone voltage and product ion and collision energy were optimized automatically by the QuanOptimize software. Folic acid was detected by SIR (single ion reaction) in negative mode, and carisoprodol was detected by SIR in the positive mode measuring the sodium adduct ions. All samples were acquired by use of a generic MS protocol into which the masses for the corresponding compounds and the internal standard, that is, verapamil, was quantified. Verapamil was used as an internal volume marker and the CV (coefficient of variation) was calculated for each of the different compound analyses. The CVs did not exceed 10% for any of the compounds for the total procedure. All chromatograms were integrated by use of the QuanLynx software and a generic quantitation method to obtain drug concentrations in the samples.

Acknowledgment. We gratefully acknowledge support from the Swedish Foundation for Strategic Research.

Supporting Information Available: Discussion of experimental difficulties seen in some of the compounds (amiloride, erythromycin, folic acid, glipizide, levothyroxine, meclizine, and terfenadine) and descriptor table for data set 1 (in pdf format) and for data set 2 (as Excel file, .xls format). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Kola, I.; Landis, J. Opinion: Can the pharmaceutical industry reduce attrition rates? *Nat. Rev. Drug Discovery* **2004**, *3*, 711−716.
(2) Winiwarter, S.; Bonham, N. M.; Ax, F.; Hallberg, A.; Lennernaes, H.; Karlen, A. Correlation of Human Jejunal Permeability (in Vivo) of Drugs with Experimentally and Theoretically Derived Parameters. A Multivariate Data Analysis Approach. *J. Med. Chem.* **1998**, *41*, 4939−4949.
(3) 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.
(4) Bergstroem, C. A. S.; Wassvik, C. M.; Norinder, U.; Luthman, K.; Artursson, P. Global and local computational models for aqueous solubility prediction of drug-like molecules. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1477−1488.
(5) Afzelius, L.; Zamora, I.; Masimirembwa, C. M.; Karlen, A.; Andersson, T. B.; Mecucci, S.; Baroni, M.; Cruciani, G. Conformer- and Alignment-Independent Model for Predicting Structurally Diverse Competitive CYP2C9 Inhibitors. *J. Med. Chem.* **2004**, *47*, 907−914.
(6) Kratochwil, N. A.; Huber, W.; Muller, F.; Kansy, M.; Gerber, P. R. Predicting plasma protein binding of drugs: a new approach. *Biochem. Pharmacol.* **2002**, *64*, 1355−1374.
(7) van de Waterbeemd, H.; Gifford, E. ADMET in silico modelling: Towards prediction paradise? *Nat. Rev. Drug Discovery* **2003**, *2*, 192−204.
(8) Kim, K.; Cho, S.; Park, J. H.; Byun, Y.; Chung, H.; Kwon, I. C.; Jeong, S. Y. Surface Plasmon Resonance Studies of the Direct Interaction Between a Drug/Intestinal Brush Border Membrane. *Pharm. Res.* **2004**, *21*, 1233−1238.
(9) 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.
(10) Winiwarter, S.; Ax, F.; Lennernas, H.; Hallberg, A.; Pettersson, C.; Karlen, A. Hydrogen bonding descriptors in the prediction of human in vivo intestinal permeability. *J. Mol. Graph. Model.* **2003**, *21*, 273−287.
(11) FASS, Elanders, Stockholm, Sweden.
(12) CS ChemDraw Std 6.0, CambridgeSoft Corporation, 100 Cambridge Park Drive, Cambridge, MA.
(13) Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31−36.
(14) Sybyl 6.7, Tripos Inc., 1699 South Hanley Rd., St. Louis, MO, 63144.
(15) Pearlman, R. S. Concord, distributed by Tripos Inc., St. Louis, MO, 63144.
(16) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(17) Stewart, J. J. P. MOPAC: a semiempirical molecular orbital program. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1−105.
(18) Ghose, A. K.; Crippen, G. M. Atomic physicochemical parameters for three-dimensional structure-directed quantitative structure−activity relationships. I. Partition coefficients as a measure of hydrophobicity. *J. Comput. Chem.* **1986**, *7*, 565−577.
(19) Moriguchi, I.; Hirono, S.; Liu, Q.; Nakagome, I.; Matsushita, Y. Simple method of calculating octanol/water partition coefficient. *Chem. Pharm. Bull.* **1992**, *40*, 127−130.
(20) Moriguchi, I.; Hirono, S.; Nakagome, I.; Hirano, H. Comparison of reliability of log *P* values for drugs calculated by several methods. *Chem. Pharm. Bull.* **1994**, *42*, 976−978.
(21) ACD/Labs version 4.5, Advanced Chemistry Development, Inc., Toronto, Ontario, Canada.
(22) Simca-P+ 10.0, Umetrics AB, SE−-90719 Umeå, Sweden.
(23) Wold, S.; Sjoestroem, M.; Carlson, R.; Lundstedt, T.; Hellberg, S.; Skagerberg, B.; Wikstroem, C.; Oehman, J. Multivariate design. *Anal. Chim. Acta* **1986**, *191*, 17−32.
(24) Carlson, R.; Lundstedt, T.; Albano, C. Screening of suitable solvents in organic synthesis. Strategies for solvent selection. *Acta Chem. Scand. B* **1985**, *B39*, 79−91.
(25) ACD pK_a DB version 4.56, Advanced Chemistry Development, Inc., Toronto, Ontario, Canada.
(26) Takacs-Novak, K.; Box, K. J.; Avdeef, A. Potentiometric pK_a determination of water-insoluble compounds: validation study in methanol/water mixtures. *Int. J. Pharm.* **1997**, *151*, 235−248.
(27) Comer, J. E. A.; Avdeef, A.; Box, K. J. Limits for successful measurement of pK_a and log *P* by pH-metric titration. *Am. Lab.* **1995**, *27*, 36C−−36D, 36T, 36H−−36I.
(28) 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.
(29) Allen, R. I.; Box, K. J.; Comer, J. E. A.; Peake, C.; Tam, K. Y. Multiwavelength spectrophotometric determination of acid dissociation constants of ionizable drugs. *J. Pharm. Biomed. Anal.* **1998**, *17*, 699−712.
(30) Slater, B.; McCormack, A.; Avdeef, A.; Comer, J. E. A. pH-Metric log *P*. 4. Comparison of Partition Coefficients Determined by HPLC and Potentiometric Methods to Literature Values. *J. Pharm. Sci.* **1994**, *83*, 1280−1283.
(31) Avdeef, A. Assessment of distribution−pH profiles. In *Lipophilicity in Drug Action and Toxicology*; Pliska, V., Testa, B., van de Waterbeemd, H., Eds.; Wiley−VCH: Weinheim, Germany, 1996; Vol. 4, pp 109−139.
(32) Valko, K.; Bevan, C.; Reynolds, D. Chromatographic Hydrophobicity Index by Fast-Gradient RP-HPLC: A High-Throughput Alternative to Log *P*/log *D*. *Anal. Chem.* **1997**, *69*, 2022−2029.
(33) Stuart, M.; Box, K. Chasing Equilibrium: Measuring the Intrinsic Solubility of Weak Acids and Bases. *Anal. Chem.* **2005**, *77*, 983−990.
(34) Box, K. J.; Volgyi, G.; Baka, E.; Stuart, M.; Takacs-Novak, K.; Comer, J. E. A. Equilibrium versus kinetic measurements of aqueous solubility, and the ability of compounds to supersaturate in solution−a validation study. *J. Pharm. Sci.* **2006**, *95*, 1298−1307.
(35) Avdeef, A. High-throughput measurements of solubility profiles. In *Pharmacokinetic Optimization in Drug Research: Biological, Physicochemical, and Computational Strategies* [Proceedings of the 2nd LogP2000, Lipophilicity Symposium Lausanne, Switzerland, Mar. 5−9, 2000], 2001, pp 305−325.
(36) Comprehensive Medicinal Chemistry; MDL Information Systems, Inc., 14600 Catalina St., San Leandro, CA.
(37) World Drug Index; Derwent Publications Ltd.
(38) MDL Drug Data Report; MDL Information Systems, Inc., 14600 Catalina St., San Leandro, CA.
(39) Avaliable Chemicals Directory v2000.1; MDL Information Systems, Inc., 14600 Catalina St., San Leandro, CA.
(40) The ATC system provides a global standard for classifying medical substances. It is maintained by the WHO Collaborating Centre for Drug Statistics Methodology in Oslo, Norway: http://www.whocc.no/atcddd/.
(41) Poole, S. K.; Poole, C. F. Separation methods for estimating octanol−water partition coefficients. *J. Chromatogr. B* **2003**, *797*, 3−19.
(42) Gocan, S.; Cimpan, G.; Comer, J. Lipophilicity measurements by liquid chromatography. *Adv. Chromatogr.* **2006**, *44*, 79−176.
(43) Kerns, E. H.; Di, L.; Petusky, S.; Kleintop, T.; Huryn, D.; McConnell, O.; Carter, G. Pharmaceutical profiling method for lipophilicity and integrity using liquid chromatography−mass spectrometry. *J. Chromatogr. B* **2003**, *791*, 381−388.
(44) Ungell, A.-L.; Karlsson, J. Cell cultures in drug discovery: an industrial perspective. In *Drug Bioavailability*; van de Waterbeemd, H., Lennernaes, H., Artursson, P., Eds.; Wiley−VCH: Weinheim, Germany, 2003; Vol. 18, pp 90−131.

(45) Artursson, P.; Palm, K.; Luthman, K. Caco-2 monolayers in experimental and theoretical predictions of drug transport. *Adv. Drug Delivery Rev.* **2001**, *46*, 27−43.
(46) Sun, D.; Lennernas, H.; Welage, L. S.; Barnett, J. L.; Landowski, C. P.; Foster, D.; Fleisher, D.; Lee, K.-D.; Amidon, G. L. Comparison of Human Duodenum and Caco-2 Gene Expression Profiles for 12,000 Gene Sequences Tags and Correlation with Permeability of 26 Drugs. *Pharm. Res.* **2002**, *19*, 1400−1416.
(47) Lennernaes, H.; Palm, K.; Fagerholm, U.; Artursson, P. Comparison between active and passive drug transport in human intestinal epithelial (Caco-2) cells in vitro and human jejunum in vivo. *Int. J. Pharm.* **1996**, *127*, 103−107.
(48) Twenty-two drugs with known FA% values (four compounds with FA% < 20, six compounds with FA% between 20 and 80, and 12 compounds with FA% > 80) were run four times through the robotic assay. Resulting permeability values were correlated to the FA% values as suggested by Ungell and Karlsson and resulted in the presented limits.
(49) Aoki, F. Y.; Sitar, D. S. Clinical pharmacokinetics of amantadine hydrochloride. *Clin. Pharmacokinet.* **1988**, *14*, 35−51.
(50) Aoki, F. Y.; Sitar, D. S.; Ogilvie, R. I. Amantadine kinetics in healthy young subjects after long-term dosing. *Clin. Pharmacol. Ther.* **1979**, *26*, 729−736.
(51) Spahn, H.; Reuter, K.; Mutschler, E.; Gerok, W.; Knauf, H. Pharmacokinetics of amiloride in renal and hepatic disease. *Eur. J. Clin. Pharmacol.* **1987**, *33*, 493−498.
(52) Beermann, B. Aspects on pharmacokinetics of some diuretics. *Acta Pharmacol. Toxicol.* **1984**, *54* (Suppl. 1), 17−29.
(53) Beermann, B.; Groschinsky-Grind, M. Clinical pharmacokinetics of diuretics. *Clin. Pharmacokinet.* **1980**, *5*, 221−245.
(54) Vergin, H.; Nuss, U.; Strobel, K. Studies on the pharmacokinetics and bioavailability of bendroflumethiazide in combination with spironolactone. *Arzneim.-Forsch.* **1986**, *36*, 490−495.
(55) Song, J. C.; White, C. M. Clinical pharmacokinetics and selective pharmacodynamics of new angiotensin-converting enzyme inhibitors. An update. *Clin. Pharmacokinet.* **2002**, *41*, 207−224.
(56) Kubo, S. H.; Cody, R. J. Clinical pharmacokinetics of the angiotensin converting enzyme inhibitors. A review. *Clin. Pharmacokinet.* **1985**, *10*, 377−391.
(57) Duchin, K. L.; McKinstry, D. N.; Cohen, A. I.; Migdalof, B. H. Pharmacokinetics of captopril in healthy subjects and in patients with cardiovascular diseases. *Clin. Pharmacokinet.* **1988**, *14*, 241−259.
(58) Bertilsson, L.; Tomson, T. Clinical pharmacokinetics and pharmacological effects of carbamazepine and carbamazepine-10,11-epoxide. An update. *Clin. Pharmacokinet.* **1986**, *11*, 177−198.
(59) Olsen, H.; Koppang, E.; Alvan, G.; Moerland, J. Carisoprodol elimination in humans. *Ther. Drug Monit.* **1994**, *16*, 337−340.
(60) Poklis, A.; Maginn, D.; Mackell, M. A. Chlorprothixene and chlorprothixene-sulfoxide in body fluids from a case of drug overdose. *J. Anal. Toxicol.* **1983**, *7*, 29−32.
(61) Bagli, M.; Rao, M. L.; Hoeflich, G.; Kasper, S.; Langer, M.; Barlage, U.; Beneke, M.; Sueverkruep, R.; Moeller, H.-J. Pharmacokinetics of chlorprothixene after single intravenous and oral administration of three galenic preparations. *Arzneim.-Forsch.* **1996**, *46*, 247−250.
(62) Mishin, V. M.; Rosman, A. S.; Basu, P.; Kessova, I.; Oneta, C. M.; Lieber, C. S. Chlorzoxazone pharmacokinetics as a marker of hepatic cytochrome P4502E1 in humans. *Am. J. Gastroenterol.* **1998**, *93*, 2154−2161.
(63) Desiraju, R. K.; Renzi, N. L., Jr.; Nayak, R. K.; Ng, K. T. Pharmacokinetics of chlorzoxazone in humans. *J. Pharm. Sci.* **1983**, *72*, 991−994.
(64) Periti, P.; Mazzei, T.; Mini, E.; Novelli, A. Clinical pharmacokinetic properties of the macrolide antibiotics. Effects of age and various pathophysiological states (Part II). *Clin. Pharmacokinet.* **1989**, *16*, 261−282.
(65) Kanazawa, S.; Ohkubo, T.; Sugawara, K. The effects of grapefruit juice on the pharmacokinetics of erythromycin. *Eur. J. Clin. Pharmacol.* **2001**, *56*, 799−803.
(66) Miller, D. B.; Spence, J. D. Clinical pharmacokinetics of fibric acid derivatives (fibrates). *Clin. Pharmacokinet.* **1998**, *34*, 155−162.
(67) Balant-Gorgia, A. E.; Balant, L. Antipsychotic drugs. Clinical pharmacokinetics of potential candidates for plasma concentration monitoring. *Clin. Pharmacokinet.* **1987**, *13*, 65−90.
(68) Jann, M. W.; Ereshefsky, L.; Saklad, S. R. Clinical pharmacokinetics of the depot antipsychotics. *Clin. Pharmacokinet.* **1985**, *10*, 315−333.
(69) Hannon-Fletcher, M. P.; Armstrong, N. C.; Scott, J. M.; Pentieva, K.; Bradbury, I.; Ward, M.; Strain, J. J.; Dunn, A. A.; Molloy, A. M.; Kerr, M. A.; McNulty, H. Determining bioavailability of food folates in a controlled intervention study. *Am. J. Clin. Nutr.* **2004**, *80*, 911−918.
(70) Ahn, E.; Kapur, B.; Koren, G. Study on circadian variation in folate pharmacokinetics. *Can. J. Clin. Pharmacol.* **2005**, *12*, e4−9.
(71) Kobayashi, K. A.; Bauer, L. A.; Horn, J. R.; Opheim, K.; Wood, F., Jr.; Kradjan, W. A. Glipizide pharmacokinetics in young and elderly volunteers. *Clin. Pharmacy* **1988**, *7*, 224−228.
(72) Kradjan, W. A.; Kobayashi, K. A.; Bauer, L. A.; Horn, J. R.; Opheim, K. E.; Wood, F. J., Jr. Glipizide pharmacokinetics: effects of age, diabetes, and multiple dosing. *J. Clin. Pharmacol.* **1989**, *29*, 1121−1127.
(73) Brogden, R. N.; Heel, R. C.; Pakes, G. E.; Speight, T. M.; Avery, G. S. Glipizide: a review of its pharmacological properties and therapeutic use. *Drugs* **1979**, *18*, 329−353.
(74) Beermann, B.; Groschinsky-Grind, M. Enhancement of the gastrointestinal absorption of hydrochlorothiazide by propantheline. *Eur. J. Clin. Pharmacol.* **1978**, *13*, 385−387.
(75) Beermann, B.; Groschinsky-Grind, M. Gastrointestinal absorption of hydrochlorothiazide enhanced by concomitant intake of food. *Eur. J. Clin. Pharmacol.* **1978**, *13*, 125−128.
(76) Beermann, B.; Groschinsky-Grind, M. Pharmacokinetics of hydrochlorothiazide in man. *Eur. J. Clin. Pharmacol.* **1977**, *12*, 297−303.
(77) Beermann, B.; Groschinsky-Grind, M.; Rosen, A. Absorption, metabolism, and excretion of hydrochlorothiazide. *Clin. Pharmacol. Ther.* **1976**, *19*, 531−537.
(78) Jordo, L.; Johnsson, G.; Lundborg, P.; Persson, B. A.; Regardh, C. G.; Ronn, O. Bioavailability and disposition of metoprolol and hydrochlorothiazide combined in one tablet and of separate doses of hydrochlorothiazide. *Br. J. Clin. Pharmacol.* **1979**, *7*, 563−567.
(79) O'Grady, P.; Yee, K.-F.; Lins, R.; Mangold, B. Fosinopril/hydrochlorothiazide: single dose and steady-state pharmacokinetics and pharmacodynamics. *Br. J. Clin. Pharmacol.* **1999**, *48*, 375−381.
(80) Lennernas, H.; Nilsson, D.; Aquilonius, S. M.; Ahrenstedt, O.; 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. Pharmacol.* **1993**, *35*, 243−250.
(81) Granerus, A. K.; Jagenburg, R.; Svanborg, A. Intestinal decarboxylation of L-dopa in relation to dose requirement in Parkinson's disease. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **1973**, *280*, 429−439.
(82) Sasahara, K.; Nitanai, T.; Habara, T.; Morioka, T.; Nakajima, E. Dosage form design for improvement of bioavailability of levodopa II: bioavailability of marketed levodopa preparations in dogs and parkinsonian patients. *J. Pharm. Sci.* **1980**, *69*, 261−265.
(83) Hinterberger, H.; Andrews, C. J. Catecholamine metabolism during oral administration of levodopa. *Arch. Neurol.* **1972**, *26*, 245−252.
(84) Nutt, J. G.; Woodward, W. R.; Anderson, J. L. The effect of carbidopa on the pharmacokinetics of intravenously administered levodopa: the mechanism of action in the treatment of parkinsonism. *Ann. Neurol.* **1985**, *18*, 537−543.
(85) Cedarbaum, J. M.; Breck, L.; Kutt, H.; McDowell, F. H. Controlled-release levodopa/carbidopa. I. Sinemet CR3 treatment of response fluctuations in Parkinson's disease. *Neurology* **1987**, *37*, 233−241.
(86) Wenzel, K. W. Bioavailability of levothyroxine preparations. *Thyroid* **2003**, *13*, 665.
(87) Kearns, G. L.; Van Den Anker, J. N.; Reed, M. D.; Blumer, J. L. Pharmacokinetics of metoclopramide in neonates. *J. Clin. Pharmacol.* **1998**, *38*, 122−128.
(88) Lauritsen, K.; Laursen, L. S.; Rask-Madsen, J. Clinical pharmacokinetics of drugs used in the treatment of gastrointestinal diseases (Part I). *Clin. Pharmacokinet.* **1990**, *19*, 11−31.
(89) Rotmensch, H. H.; Mould, G. P.; Sutton, J. A.; Kilminster, S.; Moller, C.; Pero, R. W. Comparative central nervous system effects and pharmacokinetics of Neumetoclopramide and metoclopramide in healthy volunteers. *J. Clin. Pharmacol.* **1997**, *37*, 222−228.
(90) Gustavson, L. E.; Benet, L. Z. The macromolecular binding of prednisone in plasma of healthy volunteers including pregnant women and oral contraceptive users. *J. Pharmaceut. Biomed.* **1985**, *13*, 561−569.
(91) Pickup, M. E. Clinical pharmacokinetics of prednisone and prednisolone. *Clin. Pharmacokinet.* **1979**, *4*, 111−128.
(92) Sullivan, T. J.; Hallmark, M. R.; Sakmar, E.; Weidler, D. J.; Earhart, R. H.; Wagner, J. G. Comparative bioavailability: eight commercial prednisone tablets. *J. Pharmaceut. Biomed.* **1976**, *4*, 157−172.
(93) Ravis, W. R.; Diskin, C. J.; Campagna, K. D.; Clark, C. R.; McMillian, C. L. Pharmacokinetics and dialyzability of sulindac and metabolites in patients with end-stage renal failure. *J. Clin. Pharmacol.* **1993**, *33*, 527−534.
(94) Malloy, M. J.; Ravis, W. R.; Pennell, A. T.; Hagan, D. R.; Betagari, S.; Doshi, D. H. Influence of cholestyramine resin administration on single dose sulindac pharmacokinetics. *Int. J. Clin. Pharm. Ther.* **1994**, *32*, 286−289.
(95) Simons, F. E.; Simons, K. J. Pharmacokinetic optimisation of histamine H1-receptor antagonist therapy. *Clin. Pharmacokinet.* **1991**, *21*, 372−393.
(96) Garteiz, D. A.; Hook, R. H.; Walker, B. J.; Okerholm, R. A. Pharmacokinetics and biotransformation studies of terfenadine in man. *Arzneim.-Forsch.* **1982**, *32*, 1185−1190.

(97) Okerholm, R. A.; Weiner, D. L.; Hook, R. H.; Walker, B. J.; Leeson, G. A.; Biedenbach, S. A.; Cawein, M. J.; Dusebout, T. D.; Wright, G. J.; et al. Bioavailability of terfenadine in man. *Biopharm. Drug Dispos.* **1981**, *2*, 185−190.
(98) Garty, M.; Hurwitz, A. Effect of cimetidine and antacids on gastrointestinal absorption of tetracycline. *Clin. Pharmacol. Ther.* **1980**, *28*, 203−207.
(99) Raghuram, T. C.; Krishnaswamy, K. Pharmacokinetics of tetracycline in nutritional edema. *Chemotherapy* **1982**, *28*, 428−433.
(100) Jack, D. B. *Handbook of clinical pharmacokinetic data*; MacMillan Publishers Ltd.: Basingstoke, U.K., 1992.
(101) Lamp, K. C.; Freeman, C. D.; Klutman, N. E.; Lacy, M. K. Pharmacokinetics and pharmacodynamics of the nitroimidazole antimicrobials. *Clin. Pharmacokinet.* **1999**, *36*, 353−373.
(102) Wood, S. G.; John, B. A.; Chasseaud, L. F.; Brodie, R. R.; Baker, J. M.; Faulkner, J. K.; Wood, B. A.; Darragh, A.; Lambe, R. F. Pharmacokinetics and metabolism of ^{14}C-tinidazole in humans. *J. Antimicrob. Chemother.* **1986**, *17*, 801−809.
(103) Bergan, T.; Solhaug, J. H.; Soreide, O.; Leinebo, O. Comparative pharmacokinetics of metronidazole and tinidazole and their tissue penetration. *Scand. J. Gastroenterol.* **1985**, *20*, 945−950.
(104) We will continuously gather information from the literature concerning this data set and list it on our web site (http://www.orgfarm.uu.se/compchem/fass/).
(105) Pearlman, R. S. Sarea, QCPE 413.
(106) Pearlman, R. S. In *Partition coefficient: determination and estimation*; Dunn, W. J., Block, J. H., Pearlman, R. S., Eds.; Pergamon Press: New York, 1986; p 165.
(107) Fast D-PAS and CheqSol module in Refinement Pro 2, Sirius Ltd., Forest Row: Sussex, U.K.
(108) Neuhoff, S.; Ungell, A.-L.; Zamora, I.; Artursson, P. pH-Dependent Bidirectional Transport of Weakly Basic Drugs Across Caco-2 Monolayers: Implications for Drug−Drug Interactions. *Pharm. Res.* **2003**, *20*, 1141−1148.

JM0506219