Rate-Limited Steps of Human Oral Absorption and QSAR Studies

Yuan H. Zhao,¹ Michael H. Abraham,^{1,4} Joelle Le.¹ **Anne Hersey,2 Chris N. Luscombe,2 Gordon Beck,3 Brad Sherborne,3 and Ian Cooper3**

Received May 30, 2002; accepted July 1, 2002

Purpose. To classify the dissolution and diffusion rate-limited drugs and establish quantitative relationships between absorption and molecular descriptors.

Methods. Absorption consists of kinetic transit processes in which dissolution, diffusion, or perfusion processes can become the ratelimited step. The absorption data of 238 drugs have been classified into either dissolution or diffusion rate-limited based on an equilibrium method developed from solubility, dose, and percentage of absorption. A nonlinear absorption model derived from first-order kinetics has been developed to identify the relationship between percentage of drug absorption and molecular descriptors.

Results. Regression analysis was performed between percentage of absorption and molecular descriptors. The descriptors used were ClogP, molecular polar surface area, the number of hydrogenbonding acceptors and donors, and Abraham descriptors. Good relationships were found between absorption and Abraham descriptors or ClogP.

Conclusions. The absorption models can predict the following three BCS (Biopharmaceutics Classification Scheme) classes of compounds: class I, high solubility and high permeability; class III, high solubility and low permeability; class IV, low solubility and low permeability. The absorption models overpredict the absorption of class II, low solubility and high permeability compounds because dissolution is the rate-limited step of absorption.

KEY WORDS: human intestinal absorption; rate-limited step; hydrogen bonding; solubility.

INTRODUCTION

Absorption may be determined by the rate of dissolution in gastrointestinal fluid, passive diffusion across intestinal membrane, or perfusion in portal vein (1–4). Many physicochemical properties of drugs and physiologic factors that influence the rate-limited process can affect the extent of absorption. The methods of determining these kinetic rate constants and identification of the rate-limited step within the environmental conditions of the gastrointestinal tract are the key factors in modeling absorption.

Dissolution rate is obviously important because the drug must be in solution for uptake to occur (5). Based on the Noyes-Whitney model (6,7), dissolution rate is governed by solubility, as well as by the volume of the lumen, motility (hydrodynamics), diffusivity, particle size, density, wettability, *etc*. (5). Formulation scientists use dissolution tests to choose between candidate formulations (8). For those drugs that are diffusion rate limited (i.e., diffusion into and across the membrane is the slowest step), blood flow and dissolution will have little effect on gastrointestinal absorption. One of the methods used to measure the diffusion rate of a compound is intestinal permeability in rats, developed by Schanker *et al.* (9) and Dolusio *et al.* (10). A non-animal procedure, diffusion through Caco-2 cell monolayers, has been used to screen permeability and is especially valuable for examining a large number of compounds (11).

It is generally assumed that physicochemical descriptors of drug molecules can be useful for predicting absorption for passive diffusion of drugs. Several recent studies have shown their importance to the prediction of human intestinal absorption (12–17). A solute requires a certain affinity to lipid structures to enter the cell membrane (18). Many attempts have been made to correlate *in vivo* absorption with drug lipophilicity (7,19,20). However, most recent studies showed that both H-bonding acceptors and donors play very important roles in gastrointestinal absorption for passive diffusion drugs (13,15–17). Clark and Palm (15,16) recently developed theoretical methods, based on the determination of dynamic surface properties—polar molecular surface area (PSA), to predict human intestinal absorption. However, when the model was applied to a larger data set, the fit was not very good (15). One of the reasons is that absorption of these drugs may not be controlled by the same kinetic process. For instance, dissolution is the rate-determining step of absorption for some poorly soluble drugs, whereas intestinal wall permeability becomes rate-controlling if the drug is polar (5). The Lipinski "rule of 5" has proved very popular as a rapid screen for compounds likely to be poorly absorbed (13,15,18). Lipinski *et al.* thought that these descriptors were globally associated with solubility and permeability. Because solubility is the key parameter determining dissolution rate (5), it is reasonable to believe that drugs alerted by the "rule of 5" can be either dissolution or diffusion rate-limited drugs.

Several workers have not only tried to establish absorption models for passive diffusion drugs but also for dissolution rate-limited drugs. To correct for low solubility, Dressman *et al.* (5) introduced the absorption potential (AP). With this approach, log P is corrected for the molar fraction of nonionized species at pH 6.5 (F_{non}) , the solubility of the nonionized species in water (S_w) , the volume of the luminal contents (V_L) , and the dose administered (X_O) .

$$
AP = \log\left(P \times F_{\text{non}} \times \frac{S_{\text{w}} \times V_{\text{L}}}{X_{\text{O}}}\right) \tag{1}
$$

In their study, they found a sigmoidal relationship between the fraction absorbed in humans and the AP for seven chemically different drug compounds. It is important to note that Eq. (1) may only be suitable for drugs for which the dissolution is the rate-limited step (21). Balon *et al.* (22) suggested that a high solubility to dose ratio may outweigh a low lipophilicity, resulting in a high percentage of absorption. Even highly lipophilic compounds can be poorly absorbed when the solubility to dose ratio is small, such as miconazole

¹ Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, United Kingdom.

² DMPK, Mechanism and Extrapolation Technologies, GlaxoSmith-Kline, Park Road, Ware SG12 0DP, United Kingdom.

³ Roche Products Ltd, Welwyn Garden city, Herts AL7 3AY, United Kingdom.

⁴ To whom correspondence should be addressed. (e-mail: m.h.abraham@ucl.ac.uk)

Human Intestinal Absorption and QSAR 1447

and rifabutine. It is believed that $S_w \times V_L/X_O$ is a function of the dissolution rate (5); after the solubility to dose ratio reaches a certain limit, the dissolution rate is relatively rapid. Hence, it is the permeability characteristics of the drug that determine the rate of absorption (8).

Scales of hydrogen-bonding acceptors and donors form a major part of the Linear Free Energy Relation (LFER) method of Abraham and coworkers (20,23).

$$
SP = c + eE + sS + aA + bB + vV
$$
 (2)

where SP is a property of a series of compounds in a given system, \bf{E} is an excess molar refraction in units of $\rm{cm³}$ mol−1)/10, **S** is the dipolarity/polarizability, **A** and **B** are the hydrogen-bonding acceptors and donors, respectively, and **V** is the McGowan characteristic volume in units of $\text{cm}^3 \text{ mol}^{-1}$)/ 100. Equation (2) has been applied to numerous physicochemical and biochemical processes. We applied the LFER equation by use of calculated descriptors to human intestinal absorption of 241 drug and druglike compounds with a wide range of physicochemical properties (12). Based on the analysis of the solubility and dose, 19 drugs considered to be dissolution rate-limited drugs were removed from the data set. The regression analysis of the remaining compounds showed that the dominant descriptors are hydrogen-bonding acceptors and donors, which is quite similar to previously reported results (13,15,16,17,22).

Following on from this work (12), the aims of the present study are as follows: (i) to give a theoretical background based on the rate-limited step; (ii) to classify the dissolution and diffusion rate-limited drugs based on the solubility, dose, and percentage of absorption; and (iii) to establish quantitative relationships between absorption and molecular descriptors, such as ClogP, PSA, and Abraham descriptors and then apply the absorption models to a large data set to test the models established.

MATERIALS AND METHODS

Human Intestinal Absorption Data

The human intestinal absorption dosed orally was collected and evaluated from 244 sources of literature; details were published previously (12).

Physicochemical Descriptors

Abraham molecular descriptors were calculated by the method of Platts *et al.* (23). The program was written to read molecular structures as SMILES strings. After calculation of the solvation descriptors, an error code was given by the program for each drug as an indication of the quality of the parameter calculations.

The logarithm of the octanol-water partition coefficient (ClogP) was calculated with use of the ClogP for Windows software (Biobyte version 2.0.0b, Claremont, CA, USA). The experimental octanol-water partition coefficient was also obtained from the program database.

Polar molecular surface area was calculated by using the SAVOL program (Tripos Inc., St. Louis, MO, USA). The three-dimensional structure of a drug was transferred from SMILES by using CONCORD. After energy minimization, the polar surface area for each drug was calculated.

Statistical Analysis

The linear regression analysis was performed by using Excel 97, and nonlinear regression was performed by using JMP program (version 3.2.5, 1989-1999 SAS Institute Inc., Cary, NC, USA). Stepwise regression analysis was used to determine the most significant descriptors. The regression coefficients were obtained by least-squares regression analysis. For each regression, the following descriptive information is provided: number of observations used in the analysis (n), square of the correlation coefficient (r^2) , and standard error of the estimate (S).

RESULTS AND DISCUSSION

Theoretical Background—Rate-Limited Step

Oral absorption refers to the movement of a drug from its site of administration into the blood. Sietsema (24) defined absorption as "the drug passing from the lumen of the gastrointestinal (GI) tract into the tissue of the gastrointestinal tract. Once in the tissue, the drug is considered absorbed."

A number of factors affect the movement of a drug from a site of administration into the blood. Drug solubility is important. Drugs in solution are generally absorbed faster than undissolved suspensions (25). The rate of passive diffusion of active substance passing through biologic membranes into the bloodstream is directly proportional to the concentration of dissolved substance. Thus, insufficient solubility of drugs can lead to poor absorption. The steps involved in the absorption of an orally administered drug can be simply depicted below.

Solid drug
$$
\xrightarrow{\text{Disolution}}
$$
 Drug in solution $\xrightarrow{\text{Difusion}}$

\nAbsorbed drug $\xrightarrow{\text{Perfusion}}$

\nOutput

\nDescription:

Dissolution Rate-Limited Step

Very often the dosage form does not contain the drug as a solution, but rather as solid particles as in tablets, in capsules, or in suspensions. Because solid particles cannot pass through membranes, a drug has to dissolve to be absorbed. Many drugs administered orally in solid dosage forms, such as tablets or capsules, must dissolve in the aqueous digestive fluids and diffuse across the intestinal wall into the hepatic portal vein for transport through the liver before they reach the systemic circulation (2). Dissolution may limit the rate of absorption of drugs, which disintegrate slowly and readily diffuse across the gut wall (4). Dissolution is considered the rate-limited step in the absorption of some drugs from solid dosage forms (3). This step is bypassed by the administration of a solution, but the rate of absorption of a drug from solution may also be limited by perfusion at the site of absorption.

A particularly important factor is the absolute solubility of the drug at the pH of the gut. Discoumarol has a high oil-water partition ratio and is rapidly absorbed from solution, but the absorption of the drug is slow and erratic because it precipitates in the gut, making the dissolution rate of the

drug crystals the limiting factor in the absorption rate (26). Solubility is important in limiting the rate of drug absorption of a number of phenylbutazone analogues. Several of these compounds given parentally to rats are highly active in protecting against formaldehyde-induced edema but are virtually devoid of action when given orally to man (26).

The Noyes-Whitney equation (6,7) gives the relationship between the rate of dissolution, solubility (C_s) and surface area (A) of the solid.

$$
\frac{dX_d}{dt} = \frac{AD}{\delta}(C_s - X_d/V) \tag{3}
$$

where *A* is the effective surface area of the solid drug, *D* is the diffusion coefficient of the drug, δ is the effective diffusion boundary layer thickness adjacent to the dissolving surface, C_s is the saturation solubility of the drug under lumenal conditions, X_d is the amount of drug already in solution, and *V* is the volume of the dissolution medium.

Diffusion Rate-Limited Step

If dissolution is relatively rapid, it is the permeability characteristics of the drug that determine the rate of absorption. Passive diffusion can be generally described by Fick's law (1): the rate of diffusion is a function of the concentration difference, the surface area and the distance involved, and characteristic factors of the nature of the biologic barrier and the diffusing substance. Drug concentration on the receiving site (portal vein) is often negligible in relation to that on the driving site (intestinal tract). There must be a sufficient quantity of the substance dissolved in the small intestinal fluid. Then, the rate-determining step for absorption is the passive diffusion to and through the membrane and the percentage absorption is related to the diffusion rate (1). If the rate of diffusion follows first-order kinetics (2,14,21,27), the following relationships exist between percentage absorption (%*Abs*) and the diffusion rate constant (k_{dif}) .

$$
\frac{dC_I}{dt} = -k_{\text{dif}}C_I \tag{4}
$$

$$
\ln \frac{(C_l^o - C_p^t)}{C_l^o} = -K_{\text{dif}} \cdot t \tag{5}
$$

$$
\ln(1 - FA) = -k_{\text{dif}} \cdot t \tag{6}
$$

$$
\%Abs = 100 \times (1 - e^{-k \text{dif}t}) = 100 \times (1 - e^{-10 \text{log}k_{\text{dif}} + \text{log}t}) \tag{7}
$$

or

$$
\log\left(\ln\frac{1}{1 - FA}\right) = \log k_{\text{dif}} + \log_t\tag{8}
$$

where dC_I/dt is the diffusion rate through the gastrointestinal membrane, k_{dif} is the diffusion rate constant, C_{I} is the drug concentration in intestinal fluid, C_1° is the starting concentration in intestinal fluid, C_p^t is the concentration in the portal vein at time *t*, *FA* is the fraction absorbed; log *t* is a constant if it is assumed that the transit time is the same through gastrointestinal tract for all drugs.

Although Eqs. (4)–(8) may be reasonable approximations of the absorption kinetics at one site, they may not always apply to gastrointestinal absorption. The physiologic environment in the gastrointestinal tract is quite variable (27).

A pH value of 1–3 prevails in the stomach owing to the secretion of hydrochloric acid. Luminal pH values increase along the intestine from 5 or 6 up to pH 8 in the lower small intestine and the ascending colon. Transit time from the empty stomach to the colon is variable, from 3 to 8 h on average. With exposure of drug to widely different environments along the gastrointestinal tract, gastrointestinal absorption kinetics is sometimes complex and not easily defined by a simple equation.

Perfusion Rate-Limited Step

Blood flow or perfusion in the portal vein is important in determining absorption of drugs that readily diffuse through the mucosal membrane. Blood flow removes diffused drug from the absorption site and carries it to the site of activity. This removal establishes a concentration gradient across the gastrointestinal mucosa, which is the driving force behind passive diffusion (2).

Classification of Drugs Based on Rate-Limited Step

Any model for predicting oral drug absorption, which attempts to account for all the factors involved, will be very complex (5). However, the absorption model can be significantly simplified if a method is developed to identify the ratelimited step within the overall process for all the drugs studied. For example, if the absorption process of some drugs is determined by the dissolution rate and the rate is governed by the solubility (5), a quantitative structure-activity relationship (QSAR) model could be developed by use of solubility or other related descriptors for predicting drug dissolution (21). If the absorption process of some drugs is determined by the diffusion rate and the rate was governed by hydrogen bonding, QSAR models could be developed by using hydrogenbonding acceptors and donors or polar surface area descriptors for predicting passive diffusion drugs (12,15–17).

It may be assumed that a QSAR model developed for diffusion rate-limited drugs (permeability drugs) can only be used to predict absorption for diffusion rate-limited drugs and cannot be used to predict the absorption for dissolution or perfusion rate-limited drugs. However, this is not the case. To make the explanation clear, let us create a set of absorption data from 1 to 99% for diffusion rate-limited drugs. Suppose that diffusion follows a first-order process and a drug takes 4 h to pass through the gastrointestinal tract. We can then calculate the diffusion rate constants using Eq. (6). Figure 1 shows the plot of % absorption against $\log k_{\text{dif}}$. It can be seen that if the diffusion rate constant of a drug is higher than a certain value (i.e., $k_{\text{dif}} = 1 \text{ h}^{-1}$), these drugs should have similar absorption values, near 100%. This means that if the rate constant of rate-limited step of a drug is higher than a certain value (i.e., >1 h^{-1}), the absorption value (i.e., near 100%) can be predicted from the QSAR model developed for diffusion rate-limited drugs, irrespective of whether dissolution or perfusion is the rate-limited step. However, if the dissolution or perfusion rate of a drug is the rate-limited step and is lower than a certain limit (i.e., <1 h⁻¹), absorption will be overpredicted by the QSAR model developed for diffusion rate-limited drugs.

Suppose that the perfusion rate is very high (i.e., the rate >1 h⁻¹, the absorption value should be near 100%) for those

Fig. 1. Plots of absorption percent against log k_{dif} using Eq. (8).

drugs that are perfusion rate limited (1). Then, the QSAR model developed only from diffusion drugs should be very similar to that developed from both diffusion and perfusion drugs. This is because all the perfusion rate-limited drugs have near 100% absorption, which is in agreement with the absorption predicted by the model for diffusion drugs. Furthermore, perfusion rate-limited drugs are usually small. Many synthesized drugs are usually large and absorbed via passive diffusion (2). The remaining question is how to identify dissolution and diffusion rate-limited drugs.

The ideal method of identifying dissolution and diffusion rate-limited drugs is to compare the dissolution and diffusion rates to see which one is the slowest. Unfortunately, the dissolution tests that best predict the *in vivo* performance have not been developed for drug compounds with wide physicochemical properties (6). In addition, it is very expensive and time consuming to obtain the diffusion rate from animal experiments, even with the assumption that the absorption process in animals (i.e., rat) is similar to that in humans (28).

The dissolution limitation may not only be kinetic but also equilibrium. In equilibrium, there is not enough fluid available in the gastrointestinal tract to dissolve the dose (6). This can be checked by use of the dose and solubility ratio, or percentage of insoluble drug (12,21). The small intestine volume is assumed to be 250 mL (12), and the percentage of undissolved drugs for a single dose in 250 mL water was calculated by Eq. (9) and listed in Table 2 of Reference 12.

$$
\% Undissolved = 100 \times \left(1 - \frac{0.25 \times S_w}{\text{Dose}}\right) \tag{9}
$$

The results showed that 78 drugs of the 238 data set were not completely dissolved in the gastrointestinal tract. There are 62 drugs for which the insoluble fraction was >50%. Equation (9) only gives the calculated value of % undissolved at the start of administration. However, absorption is a kinetic process, and some of these drugs can gradually dissolve in the gastrointestinal tract after an amount has been absorbed by the small intestine and removed from the site of absorption. Dissolution will also depend on the pH of the small intestinal environment.

Dissolution rate is not only governed by the solubility, dose, and volume of dissolution medium but also by the diffusion rate coefficient. (The amount of drug already in gastrointestinal solution is very low if the diffusion rate is very high.) If we consider the absorption factor, Eq. (9) can be modified to Eq. (10).

$$
\% Undissolved = 100 \times \frac{\text{Dose} \times (1 - FA) - 0.25 \times S_w}{\text{Dose}} \tag{10}
$$

where *FA* is the fraction absorbed. Table 2 of Reference 12 lists the ratio calculated by Eq. (10) between the amount of insoluble drug and dose in 250 mL of water after correction of absorption. The results show that there are 33 drugs that are still not completely dissolved after the above correction applied. Dissolution would be the rate-limited step if a drug is not fully dissolved during its transit through the gastrointestinal tract.

However, the experimental error in obtaining the absorption value of a drug is sometimes quite high (12). Furthermore, if the diffusion rate constant of a dissolution ratelimited drug reaches a certain value, then the absorption reaches 100% (Fig. 1), and the absorption values can be correctly predicted by the model developed from diffusion drugs. If a 20% estimation error is assumed for % absorption or if it is believed that the diffusion rate constant is quite high for dissolution rate-limited drugs with 80% absorption values, then the definition of dissolution rate-limited drugs can be modified to Eq. (11).

$$
\% Undissolved = 100 \times \frac{\text{Dose} \times (1 - FA) - 0.25 \times S_w}{\text{Dose}} > 20 \tag{11}
$$

In a previous article (12), 238 drugs were classified into the following groups

- Diffusion rate-limited drugs: 169 drugs.
- Diffusion rate-limited and zwitterionic drugs: 20 drugs.
- Diffusion rate-limited drugs with missing fragments from Platts method: 9 drugs.
- Dissolution rate-limited drugs (or called dose-limited drugs) (DL): 19 drugs.
- Dose-dependent drugs (DP): 7 drugs.
- Drugs with unreliable absorption: 13 drugs.

The names of these drugs are listed in Table I. The details of the classification were given in Reference 12.

SAR Studies—Rule of 5

If we classify the absorption of drugs 1–225 in Table I as high = $100-67\%$, medium = $66-33\%$, and low = $32-0\%$, and apply the Lipinski "rule of 5" to these drugs (13), 25 compounds are alerted (Table I). Of 33 poorly absorbed drugs, only 14 drugs (42%) are alerted; 6 of 37 drugs (16%) are alerted in the medium absorption group, whereas 5 of 156 drugs (3%) are alerted in the high absorption group. Obviously, most alerts are given to poorly absorbed drugs. However, 19 drugs (58%) are not alerted at all although they are poorly absorbed. The same result was found by Clark (15) in an analysis of 88 drug compounds. By examining the number of hydrogen-bonding acceptors and donors, we find that for most of these poorly absorbed drugs, either N_A is near to 10 or N_D is near to 5. Because the cutoff values have not been surpassed, an alert is not triggered. The results from the "rule of 5" are listed in Table II.

In effect, a discovery alert is a very coarse filter that identifies compounds lying in a region of property space where the probability of useful oral activity is very low (13). A compound that fails the computational alert will likely be poorly bioavailable because of poor absorption. Most compounds failing the alert will prove troublesome if they progress far enough to be studied experimentally (drugs 152–158 and 160–162 in Table I). In addition, compounds passing the alert can still prove troublesome in experimental studies. However, the simple alert has a primary value in identifying problem compounds (13).

QSAR Studies

As mentioned above, 238 drug and druglike compounds were classified as diffusion rate-limited drugs, dissolution rate-limited drugs, zwitterionic drugs, dose-dependent drugs, and drugs with unreliable absorption. To carry out QSAR studies for diffusion rate-limited drugs, ClogP, PSA, the Abraham descriptors, number of hydrogen-bonding acceptors and donors were calculated for these drugs and are listed in Table I. Drugs ($n = 189$) for which diffusion is the ratelimited step of absorption were used for regression analysis (drugs 1–189). Among them, 20 drugs (drugs 170–189) are zwitterions. It is possible that the equilibrium method Eq. (11) identified some drugs that are not dissolution ratelimited drugs because pH can greatly affect the solubility in the gastrointestinal tract. However, it is important to note that removing some drugs for which dissolution is not the rate-limited step is much better than leaving some drugs for which dissolution is the rate-limited step in the diffusion ratelimited drug set.

The absorption data can be expressed in different ways. Traditional methods use percentage to express the extent of absorption. This has the advantage of being very simple, direct, and easily understood. Models 1, 2, and 3 in Table III show the results of linear regression analysis between absorption percent and the descriptors, such as ClogP, PSA, and the Abraham descriptors. The relationship between the absorption percent and ClogP or PSA looks sigmoidal (Figs. 2 and 3). The best model is the one using the Abraham descriptors $(r^2 = 0.68)$. By examining the difference between observed and predicted absorption, it is found that most of the prediction errors arise due to zwitterions. Removing these zwitterions increases regression coefficients and results in models 4–8 (Figs. 2 and 3). Stepwise regression results (model 5) show that hydrogen bond acceptors and donors play the main roles in the absorption process; this finding agrees with previous work (13,15–17,22).

Linear regression analysis and stepwise analysis were also performed by using the parameters from the "rule of 5" in which hydrogen-bonding acceptors and donors were simply counted as the number of O and N groups present (model 8). The stepwise regression results show that the dominant descriptors are the number of hydrogen bond acceptors, donors, and molecular weight. These results are comparable to the result of stepwise regression using the Abraham descriptors (model 5). Absorption increases with decreasing hydrogenbonding acceptors and donors and increasing molecular volume or weight. However, the regression coefficient and standard error from N_A , N_D , and MW are not as good as that from the Abraham descriptors because the hydrogen bond acceptors and donors are simply based on the number of O and N groups. In reality, different O and N groups contribute differently to hydrogen bonding (20).

Although observed absorption is roughly linearly related with the absorption predicted by the Abraham descriptors (model 1), the model 1 predicts absorption $> 100\%$ for some of the drug compounds and overpredicts some poorly absorbed drugs. There is also negative prediction in the model 1 (12). The same case was found for the linear regression analysis between absorption percent and ClogP or PSA (Figs. 2 and 3). If diffusion is a first-order kinetic process, percent of absorption can be converted to the kinetic constant k_{dif} or log k - $_{\text{dif}}$ by Eq. (6). The relationship between percent of absorption and $\log k_{\text{dif}}$ is shown in Fig. 1. The relationship between percent of absorption and $log k_{\text{dif}}$ is sigmoidal. Because the same sigmoidal relationship between percent of absorption and ClogP or PSA (Figs. 2 and 3) is obtained, a linear relationship between $log k_{\text{dif}}$ and these descriptors is possible. Therefore, linear regression analysis between $log k_{dif}$ converted by Eq. (8) and descriptors, and nonlinear relationship between percent of absorption and descriptors, Eq. (7), were investigated. The problem of prediction of negative absorption percent or absorption percent > 100 can be overcome by this transformation.

Models 9–11 and Figs. 2, 3, and 4 show the regression results by using nonlinear regression analysis for 189 diffusion rate-limited drugs. (The lines in Figs. 2 and 3 are the regression lines, and the line in Fig. 4 is from 0 to 100%). The regression coefficients between the absorption percent and the Abraham descriptors or ClogP were improved. The standard errors of nonlinear regression analysis are smaller than that of linear regression analysis, especially for the absorption values near 100% and 0%. There is no negative prediction and absorption >100% in the nonlinear absorption models. Removing zwitterionic compounds improved regression results (models 12–15). In comparing these models, PSA does not seem as good as the Abraham descriptors or ClogP. This is because, although polar terms play very important roles in the absorption, nonpolar properties of a drug also contribute to the absorption.

Table III also lists the linear regression results between log_{k_{dif} and ClogP, PSA, and the Abraham descriptors (mod-} els 16–20). However, the 41 absorption values with 100 and 0% had to be removed from the data set because Eq. (8) collapses when the absorption percent is 0 or 100. We find regression coefficients of nonlinear (models 12–14) and linear regression analysis (models 17–19) are quite similar. It is important to note that compounds cannot have exactly 100% and 0% absorption if absorption is considered to be a firstorder kinetic process. All the 100% or 0% absorption values observed should be theoretically near 100 or 0, but not equal to 100 or 0.

Analysis between diffusion kinetic constant (k_{dif}) and ClogP or Abraham descriptors was performed. The result clearly shows that there are no linear relationships between k_{dif} and ClogP or Abraham descriptors. The same trend was observed if we draw a theoretical plot between $\log k_{\text{dif}}$ and k_{dif} using Eq. (6).

In addition to using the percentage and kinetic constant to express the extent of absorption, logit-transformed absorption data have been used previously in QSAR analysis (14).

$$
\log \frac{FA}{1 - FA} = \log \text{ it} FA = Z \tag{12}
$$

TABLE I. Percentage of Absorption, Molecular Weight, Octanol-Water Partition Coefficient, Sum of N and Hydrogen Bond Acceptors and Donors, Polar Surface Area, and Abraham Descriptors

No	Names	%Abs. ^a	MW	$\text{Mlog} \ensuremath{\mathbf{P}}^{\ensuremath{\mathbf{b}}}$	ClogP ^c	$N_A{}^e$	N_D^f	Rule of $5g$	PSA ^h	E	S	A	B	V
$\mathbf{1}$	aminopyrine	100	231		1.00	4	$\boldsymbol{0}$	Pass	25	1.78	1.78	$\boldsymbol{0}$	1.37	1.87
2	bornaprine	100	329		4.30	3	$\mathbf{0}$	Pass	27	1.29	1.38	$\mathbf{0}$	1.20	2.79
3	caffeine	100	194	-0.07	-0.06	6	$\boldsymbol{0}$	Pass	47	1.94	1.81	$\overline{0}$	1.47	1.36
4	camazepam	100	372		3.64	6	$\boldsymbol{0}$	Pass	52	2.63	2.56	$\boldsymbol{0}$	1.84	2.67
5	cicaprost	100	374		2.01	5	3	Pass	99	1.26	1.55	1.19	1.44	3.03
6	cisapride	100	466		3.43	7	\overline{c}	Pass	83	2.30	3.40	0.46	2.04	3.40
7	corticosterone	100	346	1.94	2.32	4	$\mathbf{2}$	Pass	73	1.90	2.98	0.53	1.71	2.74
8	cyproterone acetate	100	417		3.39	4	$\boldsymbol{0}$	Pass	49	2.07	3.17	$\overline{0}$	1.57	3.09
9	desipramine	100	266	4.90	4.09	\overline{c}	$\mathbf{1}$	Pass	20	1.99	1.57	0.09	1.04	2.26
10	diazepam	100	285	2.99	3.29	3	$\boldsymbol{0}$	Pass	28	2.38	2.11	$\overline{0}$	1.15	2.07
11	diclofenac	100	296	4.40	3.03	3	$\sqrt{2}$	Pass	40	1.97	1.88	0.78	0.87	2.03
12	ethinylestradiol	100	296	3.67	3.66	\overline{c}	$\sqrt{2}$	Pass	46	2.12	2.50	0.97	1.16	2.39
13	fenclofenac	100	297	4.80	4.96	3	$\mathbf{1}$	Pass	48	1.80	1.76	0.59	0.62	1.98
14	fluvastatin	100	411		3.19	5	3	Pass	76	2.39	2.45	1.28	1.60	3.13
15	gallopamil	100	485		3.14	7	$\boldsymbol{0}$	Pass	68	1.72	2.56	$\overline{0}$	2.28	3.99
16	glyburide	100	358		4.08	8	3	Pass	110	2.60	3.89	1.30	1.88	3.56
17	granisetron	100	312		1.79	5	$\mathbf{1}$	Pass	48	2.18	2.53	0.37	2.03	2.45
18	imipramine	100	280	4.80	4.41	\overline{c}	$\boldsymbol{0}$	Pass	8	1.97	1.56	$\overline{0}$	1.15	2.40
19	indomethacin	100	358	4.27	4.18	5	$\mathbf{1}$	Pass	68	2.39	2.72	0.59	1.19	2.53
20	isoxicam	100	335	2.83	2.40	8	$\sqrt{2}$	Pass	116	2.47	3.53	0.58	1.92	2.21
21	levonorgestrel	100	312		3.31	\overline{c}	$\mathbf{1}$	Pass	40	1.79	2.46	0.43	1.18	2.58
22	lormetazepam	100	335		2.60	$\overline{4}$	$\mathbf{1}$	Pass	53	2.69	2.37	0.10	1.39	2.26
23	lornoxicam	100	372		3.15	7	\overline{c}	Pass	100	2.95	3.60	0.58	2.04	2.30
24	mexiletine	100	179	2.15	2.57	\overline{c}	$\mathbf{2}$	Pass	34	0.97	0.81	0.03	0.84	1.58
25	nefazodone	100	470		5.00 ^d	7	$\boldsymbol{0}$	Pass	51	3.07	2.83	$\boldsymbol{0}$	2.08	3.59
26	nicotine	100	162	1.17	1.32	\overline{c}	$\boldsymbol{0}$	Pass	15	1.05	1.09	$\overline{0}$	1.11	1.37
27	ondansetron	100	293		2.64	$\overline{4}$	$\boldsymbol{0}$	Pass	31	2.13	2.15	$\overline{0}$	1.46	2.27
28	oxatomide	100	426 288		5.41	5	$\mathbf{1}$	Pass	44	3.43	2.83 1.89	0.33	2.25 1.59	3.40 2.39
29	phenglutarimide	100	331		1.54 2.70	$\overline{4}$	$\mathbf{1}$	Pass	49 99	1.61 2.84		0.32 0.58	2.06	
30 31	piroxicam	100 100	312	1.98	3.43	7 4	$\mathfrak{2}$ $\boldsymbol{0}$	Pass Pass	36	1.94	3.61 2.42	$\overline{0}$	1.60	2.25 2.45
32	praziquantel progesterone	100	314	3.87	3.78	\overline{c}	$\boldsymbol{0}$	Pass	30	1.58	2.47	$\overline{0}$	1.16	2.62
33	salicylicacid	100	138	2.26	2.19	3	$\mathfrak{2}$	Pass	55	1.05	0.89	0.72	0.38	0.99
34	stavudine	100	224	-0.81	-0.48	6	$\mathbf{2}$	Pass	86	1.91	2.06	0.49	1.77	1.56
35	sudoxicam	100	337	1.64	2.60	7	\overline{c}	Pass	101	2.87	3.60	0.58	1.91	2.17
36	tenoxicam	100	337		2.42	7	$\mathbf{2}$	Pass	100	2.82	3.51	0.58	2.08	2.17
37	testosterone	100	288	3.32	3.22	\overline{c}	$\mathbf{1}$	Pass	40	1.61	2.32	0.35	1.13	2.38
38	theophylline	100	180	-0.02	-0.06	6	$\mathbf{1}$	Pass	64	1.93	1.84	0.42	1.38	1.22
39	toremifene	100	406		6.35	$\mathfrak{2}$	$\boldsymbol{0}$	Pass	15	2.43	2.03	0.02	1.11	3.30
40	valproicacid	100	144	2.75	2.76	\overline{c}	$\mathbf{1}$	Pass	40	0.24	0.47	0.59	0.44	1.31
41	verapamil	100	455	3.79	3.71	6	$\boldsymbol{0}$	Pass	64	1.70	2.48	$\boldsymbol{0}$	2.07	3.79
42	carfecillin	99	454	2.96	3.12	8	\overline{c}	Pass	111	2.83	3.31	0.56	2.47	3.20
43	naproxen	99	230	3.34	2.82	3	$\mathbf{1}$	Pass	51	1.62	1.40	0.59	0.75	1.78
44	nordiazepam	99	270	2.93	3.01	3	$\mathbf{1}$	Pass	43	2.33	2.21	0.28	1.24	1.93
45	prenisolone	99	360	1.62	1.64	5	3	Pass	97	2.19	3.26	0.72	2.00	2.75
46	propranolol	99	259	2.98	2.75	3	$\mathfrak{2}$	Pass	43	1.85	1.36	0.10	1.29	2.15
47	atropine	98	289	1.83	1.32	4	$\mathbf{1}$	Pass	50	1.44	1.71	0.35	1.48	2.28
48	lamotrigine	98	256		3.24	5	4	Pass	97	2.79	2.81	0.50	1.09	1.65
49	minoxidilne	98	209	1.24	1.09	6	4	Pass	94	2.46	2.87	0.50	1.71	1.59
50	tolmesoxide	98	214		0.89	3	$\boldsymbol{0}$	Pass	37	1.19	2.21	$\overline{0}$	1.28	1.62
51	viloxazine	98	237		1.34	4	$\mathbf{1}$	Pass	45	1.15	1.42	0.32	1.47	1.87
52	warfarin	98	308	2.70	2.44	4	$\mathbf{1}$	Pass	51	2.30	2.43	0.55	1.26	2.31
53	antipyrine	97	188	0.38	0.41	3	$\boldsymbol{0}$	Pass	24	1.53	1.58	$\boldsymbol{0}$	1.05	1.48
54	clofibrate	97	243		3.68	3	$\boldsymbol{0}$	Pass	31	0.93	1.23	$\boldsymbol{0}$	0.69	1.82
55	disulfiram	97	296	3.88	3.88	$\sqrt{2}$	$\boldsymbol{0}$	Pass	5	2.12	1.25	$\boldsymbol{0}$	1.39	2.29
56	trimethoprim	97	290	0.91	0.95	7	4	Pass	107	2.52	2.81	0.50	1.76	2.18
57	venlafaxine	97	277		2.11	3	$\mathbf{1}$	Pass	26	1.24	1.32	0.35	1.36	2.37
58	bumetanide	96	364		3.90	7	$\overline{4}$	Pass	121	2.20	2.73	1.41	1.76	2.64
59	torasemide	96	348		3.34	7	3	Pass	95	2.14	2.95	1.12	1.90	2.58
60	trapidil	96	205		1.94	5	$\boldsymbol{0}$	Pass	43	1.68	1.52	$\overline{0}$	0.98	1.63
61	codeine	95	299	$1.14\,$	0.82	4	$\mathbf{1}$	Pass	48	2.02	1.78	0.26	1.75	2.21
62	fluconazole	95	306		-0.11	7	$\mathbf{1}$	Pass	61	1.69	2.30	0.35	1.62	2.01
63	flumazenil	95	303		1.06	6	$\boldsymbol{0}$	Pass	52	1.80	2.38	$\boldsymbol{0}$	1.76	2.09

TABLE I. Continued

No	Names	%Abs. ^a	MW	$\mathbf{M} \mathsf{log} \mathbf{P}^\mathsf{b}$	ClogP ^c	N_A^{e}	N_D^f	Rule of 5 ^g	PSA ^h	E	S	A	$\, {\bf B}$	$\mathbf V$
64	ibuprofen	95	206	3.50	3.68	\overline{c}	$\mathbf{1}$	Pass	40	0.86	0.84	0.59	0.50	1.78
65	labetalol	95	328		2.50	5	5	Pass	95	2.20	2.13	0.77	1.62	2.64
66	metoprolol	85	267	1.88	1.20	4	$\mathfrak{2}$	Pass	56	1.13	1.18	0.10	1.44	2.26
67 68	oxprenolol practolol	95 95	265 266	2.10 0.79	1.69 0.75	4 5	$\mathfrak{2}$ 3	Pass Pass	53 77	1.26 1.45	1.18 1.95	0.10 0.58	1.49 1.64	2.22 2.18
69	scopolamine	95	303		0.26	5	$\mathbf{1}$	Pass	61	1.64	1.96	0.35	1.84	2.23
70	sotalol	95	272	-0.44	0.23	5	3	Pass	85	1.54	1.98	0.74	1.74	2.10
71	timolol	95	316	1.83	1.61	7	$\mathfrak{2}$	Pass	76	1.47	1.81	0.10	2.03	2.38
72	alprenolol	93	249	2.89	2.65	3	$\mathfrak{2}$	Pass	43	1.25	1.03	0.10	1.25	2.16
73	amrinone	93	189		-0.59	$\overline{4}$	3	Pass	75	1.84	2.11	0.53	1.29	1.40
74	isradipine	92	371	4.18	3.57	8	$\mathbf{1}$	Pass	95	1.67	2.46	0.32	1.62	2.71
75	ketoprofen	92	254	3.12	2.76	3	$\mathbf{1}$	Pass	59	1.63	1.78	0.59	0.86	1.98
76 77	hydrocortisone naloxone	91 91	362 327	1.61 2.09	1.70 -0.04	5 5	3 2	Pass Pass	96 69	2.06 2.24	3.16 2.09	0.72 0.55	1.98 2.11	2.80 2.36
78	alprazolam	90	309	2.12	2.30	4	$\boldsymbol{0}$	Pass	39	2.58	2.22	$\boldsymbol{0}$	1.32	2.20
79	amphetamine	90	135	1.76	1.59	1	$\mathfrak{2}$	Pass	27	0.94	0.77	0.18	0.63	1.24
80	betaxolol	90	307	2.81	2.17	4	$\boldsymbol{2}$	Pass	55	1.33	1.29	0.10	1.44	2.57
81	chloramphenicol	90	323	1.14	0.69	7	3	Pass	118	1.86	2.46	0.66	1.62	2.07
82	felamate	90	283		-0.29	6	4	Pass	110	1.44	1.48	0.70	1.12	1.77
83	ketorolac	90	255		1.62	4	$\mathbf{1}$	Pass	62	1.69	2.02	0.59	1.23	1.87
84	meloxicam	90	351	3.01	3.10	7	$\mathfrak{2}$	Pass	101	2.88	3.57	0.58	1.91	2.32
85	nisoldipine	90	388	4.53	4.24	8	$\mathbf{1}$	Pass	82	1.71	2.43	0.32	1.54	2.92
86	nizatidine	90	331		0.50	7	$\mathfrak{2}$	Pass	83	1.87	2.55	0.20	2.41	2.46
87	phenytoin	90 90	252	2.47	2.08	4	$\mathfrak{2}$ $\mathbf{1}$	Pass	59	2.21	1.68 3.09	0.48 0.59	1.21	1.87
88 89	sulindac terazosin	90	356 387	3.05	2.81 2.71	3 9	$\mathfrak{2}$	Pass Pass	58 102	2.28 3.25	3.78	0.25	1.28 2.64	2.57 2.83
90	tramadol	90	263	2.63	2.31	3	$\mathbf{1}$	Pass	22	1.24	1.30	0.35	1.5	2.23
91	dihydrocodeine	89	301		1.30	4	$\mathbf{1}$	Pass	49	1.88	1.68	0.26	1.73	2.25
92	oxazepam	89	287	2.24	2.29	4	$\mathfrak{2}$	Pass	67	2.51	2.33	0.38	1.49	1.99
93	sultopride	89	354		1.93	6	$\mathbf{1}$	Pass	68	1.77	3.25	0.22	2.18	2.71
94	tenidap	89	321		0.63 ^d	5	3	Pass	77	2.70	2.63	0.68	1.04	2.07
95	felodipine	88	384	4.80	4.96	5	$\mathbf{1}$	Pass	60	1.75	2.17	0.32	1.37	2.71
96	moxonidine	88	242		1.02	6	$\mathfrak{2}$	Pass	69	1.62	2.06	0.61	2.34	1.67
97	nitrendipine	88	360	4.15	3.39	8	$\mathbf{1}$	Pass	105	1.72	2.47	0.32	1.53	2.64
98	saccharin	88	183	0.91	0.52	4	$\mathbf{1}$	Pass	71	1.59	2.14	0.68	0.95	1.15
99 100	bupropion lamivudine	87 87	240 229	-0.93	3.21 -1.54	$\mathfrak{2}$ 6	$\mathbf{1}$ 3	Pass Pass	24 93	1.14 2.34	1.31 2.36	0.09 0.51	1.07 1.92	1.94 1.53
101	pindolol	87	248	1.75	1.67	4	3	Pass	63	1.68	1.48	0.47	1.58	2.01
102	topiramate	86	339		-0.07	9	$\mathfrak{2}$	Pass	121	1.10	2.40	0.54	2.78	2.21
103	lansoprazole	85	369		3.07	5	$\mathbf{1}$	Pass	65	2.02	3.15	0.42	1.84	2.37
104	morphine	85	285	0.76	0.24	4	$\boldsymbol{2}$	Pass	61	2.10	1.68	0.55	1.76	2.06
105	oxyfedrine	85	313		2.84	4	$\mathfrak{2}$	Pass	57	1.85	1.94	0.19	1.74	2.54
106	tolbutamide	85	270	2.34	2.50	5	2	Pass	78	1.35	2.15	0.93	1.09	2.06
107	acetylsalicylicacid	84	180	1.19	1.02	4	$\mathbf{1}$	Pass	60	0.93	1.35	0.59	0.80	1.29
108	bromazepam	84	316	1.69	1.69	4	$\mathbf{1}$	Pass	53	2.48	2.46	0.28	1.54	1.94
109	captopril	84	217		1.19	4	$\mathbf{1}$	Pass	58	1.15	1.68	0.50	1.31	1.62
110 111	propiverine methylprednisolone	84 82	367 374		4.06 1.96	4 5	$\boldsymbol{0}$ 3	Pass Pass	28 95	1.73 2.18	1.88 3.23	$\overline{0}$ 0.72	1.45 2.02	3.00 2.90
112	mifobate	82	359		0.69	7	$\boldsymbol{0}$	Pass	70	0.76	2.38	$\boldsymbol{0}$	2.28	2.36
113	sorivudine	82	349		-1.66	8	4	Pass	127	2.55	2.71	0.93	2.39	2.00
114	digoxin	81	781	1.26	1.32	14	6	Alert	216	3.20	5.34	1.72	4.62	5.75
115	flecainide	81	414		4.43	5	2	Pass	55	0.82	1.80	0.54	1.41	2.60
116	piroximone	81	217		0.96	5	$\mathfrak{2}$	Pass	82	1.77	2.08	0.61	1.37	1.60
117	quinidine	81	324	2.64	2.93	4	$\mathbf{1}$	Pass	40	2.30	1.90	0.26	1.88	2.55
118	acebutolol	80	336	1.71	1.63	6	3	Pass	88	1.60	2.40	0.58	1.97	2.76
119	acetaminophen	80	151	0.51	0.49	3	$\mathfrak{2}$	Pass	56	1.27	1.81	1.02	0.85	1.17
120	dexamethasone	80	392	2.01	2.01	5	3	Pass	90	2.09	3.22	0.77	2.01	2.91
121	ethambutol	80	204		0.12	4	4	Pass	69	0.78	0.79	0.22	1.81	1.83
122 123	guanabenz isoniazid	80 80	231 137	-0.70	2.96 -0.71	4 4	4 3	Pass Pass	76 72	1.85 1.21	1.60 1.89	0.11 0.55	1.51 1.51	1.56 1.03
124	methadone	80	309	3.93	3.13	2	$\boldsymbol{0}$	Pass	16	1.61	1.59	$\boldsymbol{0}$	1.25	2.71
125	omeprazole	80	345	2.23	2.53	6	$\mathbf{1}$	Pass	72	2.35	3.41	0.42	2.16	2.52
126	urapidil	78	387		2.56	8	$\mathbf{1}$	Pass	65	2.98	2.99	0.20	2.81	3.02
127	famciclovir	77	321		-0.36	9	2	Pass	113	1.98	2.71	0.25	1.83	2.34
128	mercaptoethanesulfonicacid	77	142		-0.52	3	$\mathbf{1}$	Pass	59	1.13	1.60	0.35	0.99	0.89

Human Intestinal Absorption and QSAR 1453

TABLE I. Continued

No	Names	$%$ Abs. a	МW	$\text{Mlog} \ensuremath{\mathbf{P}}^{\ensuremath{\mathbf{b}}}$	ClogP ^c	N_A^e	N_D^f	Rule of $5g$	PSA ^h	Ε	S	A	B	V
191	ximoprofen	98	261		2.18	4	\overline{c}	Pass	77	1.31	1.34	0.94	0.79	2.07
192	clonidine	95	230	1.57	1.37	3	$\mathbf{2}$	Pass	42	1.60	1.49	0.64	1.16	1.53
193	viomycin	85	685		-8.03	23	19	Alert	379	4.31	6.77	4.02	7.65	4.91
194	ceftizoxime	72	382		$-4.30d$	10	$\overline{4}$	Pass	147	2.62	3.23	0.97	2.58	2.43
195	capreomycin	50	653		-7.25	21	18	Alert	357	3.96	6.12	3.18	7.20	4.84
196	AAFC	32	243		-3.91 ^d	τ	3	Pass	110	1.45	1.89	0.43	2.01	1.47
197	bretyliumtosylate	23	244		-1.25	$\mathbf{1}$	$\boldsymbol{0}$	Pass	$\boldsymbol{0}$	0.87	0.62	$\boldsymbol{0}$	0.13	1.72
198	distigminebromide	8	578			8	$\boldsymbol{0}$	Alert	80	1.75	2.23	$\boldsymbol{0}$	2.30	3.32
	Dissolution rate-limited drugs													
199	spironolactone	73	417	2.26	2.25	$\overline{4}$	$\boldsymbol{0}$	Pass	60	2.25	3.74	$\boldsymbol{0}$	1.82	3.17
200	etoposide	$50(25-75)$	588	0.60	-1.89	13	3	Alert	170	3.03	3.81	0.38	3.81	3.90
201	cefetamet pivoxil (globocef)	47	511		2.33	12	$\mathfrak{2}$	Alert	153	2.97	4.16	0.31	3.22	3.49
202	cefuroximeaxetil	$44(36-52)$	510	0.89	0.25	14	3	Alert	173	2.58	3.89	0.41	3.31	3.36
203	azithromycin	37	749		1.83	14	5	Alert	183	1.97	3.26	0.93	5.04	6.00
204	fosinopril	36	564		7.74	8	$\mathbf{1}$	Alert	89	1.61	3.25	0.50	2.92	4.47
205	pravastatin	34	425		0.57	7	$\overline{4}$	Pass	112	1.37	2.08	1.63	1.81	3.37
206	cyclosporin	$28(10-65)$	1202		3.80 ^d	23	5	Alert	324	3.97	6.84	1.54	8.65	10.02
207	bromocriptine	28	654		6.69	10	3	Alert	101	3.94	4.38	0.84	4.03	4.48
208	olsalazine	$24(17-31)$	302		4.50	8	$\overline{4}$	Pass	130	2.21	1.61	1.43	0.81	2.03
209	doxorubicin	$12(0.7-23)$	543	0.10	-1.45	12	7	Alert	204	3.51	2.91	0.81	2.93	3.73
210	cefuroxime	1	424	-0.16	-0.17	12	$\overline{4}$	Pass	168	2.66	3.38	0.91	2.93	2.73
211	iothalamatesodium	1.9	613		1.42	6	3	Pass	87	3.44	3.39	1.57	1.33	2.50
212	sulfasalazine	$59(56-61)$	398		3.83	9	3	Pass	136	3.18	3.10	1.27	1.49	2.70
213	benazepril	>37	425		1.82	7	$\overline{2}$	Pass	85	2.34	2.43	0.28	1.84	3.27
214	lisinopril	$28(25-50)$	405		-1.71	8	5	Pass	142	1.79	2.42	0.96	2.56	3.19
215	enalaprilat	$25(10-40)$	348		0.86	τ	3	Pass	112	1.60	2.18	0.78	2.08	2.66
216	amphotericin b	$3(2-5)$	924		$-2.46d$	18	13	Alert	297	3.70	5.37	3.37	5.70	7.12
217	aztreonam	1	435		$-3.46d$	13	$\overline{4}$	Pass	204	2.77	4.36	1.15	3.18	2.76
	Dose-dependent drugs													
218	mibefradil	$69(37-100)$	516		4.41	6	$\mathbf{1}$	Pass	59	2.28	2.57	0.43	1.91	3.89
219	ranitidine	$64(39-88)$	314	0.27	1.33	7	$\overline{2}$	Pass	82	1.60	2.29	0.20	2.28	2.40
220	chlorothiazide	49 (36–61)	296	-0.24	-0.31	7	3	Pass	128	2.18	3.12	1.21	1.97	1.69
221	acyclovir	$23(15-30)$	225	-1.56	-2.07	8	$\overline{4}$	Pass	125	2.34	2.67	0.83	1.87	1.52
222	norfloxacin	71	319	-1.03	1.57	6	\overline{c}	Pass	76	2.08	2.46	0.31	2.10	2.27
223	methotrexate	$70(57-83)$	454		-0.30	13	τ	Alert	211	3.91	4.73	1.80	2.77	3.22
224	gabapentin	$59(43-74)$	171		-1.18	3	3	Pass	66	0.63	0.83	0.77	0.93	1.44
	Formulation-dependent drugs													
225	prazosin	86 (77–95)	383		2.45	9	$\overline{2}$	Pass	103	3.40	3.81	0.25	2.38	2.74
	Drugs expected to have higher absorption													
226	ciprofloxacin	>69	331	-1.08	1.40	6	\overline{c}	Pass	77	2.27	2.57	0.31	2.10	2.30
227	ribavirin	>33	244	-1.85	-3.23	9	6	Pass	158	1.71	2.66	1.13	2.44	1.58
228	pafenolol	>29	337		1.67	6	$\overline{4}$	Pass	86	1.41	1.75	0.67	2.01	2.84
229	azosemide	>10	371		1.35	8	\overline{c}	Pass	140	2.84	3.17	1.46	1.50	2.34
230	xamoterol	$>\!\!5$	339	0.61	0.39	8	$\overline{4}$	Pass	109	1.80	2.37	0.74	2.66	2.57
231	enalapril	$>66(61-71)$	376		0.79	7	$\mathfrak{2}$	Pass	96	1.50	2.29	0.28	2.09	2.94
232	phenoxymethylpenicillin	$59(49-68)$	350	2.09	1.90	7	\overline{c}	Pass	100	2.20	2.58	0.40	2.28	2.44
233	gliclazide	>65	323		1.09	6	\overline{c}	Pass	83	1.85	2.52	0.77	1.76	2.36
234	benzylpenicillin	>30	334	1.83	1.70	6	\overline{c}	Pass	91	2.18	2.49	0.56	2.06	2.38
235	Thiacetazone	>20	236		$0.88\,$	5	$\overline{4}$	Pass	87	2.05	2.25	0.98	1.82	1.77
236	lovastatin	${>}10$	405	4.26	4.08	5	$\mathbf{1}$	Pass	64	1.29	2.22	0.35	1.32	3.29
237	cromolynsodium	>0.4	468	1.92	1.85	11	3	Pass	167	3.10	3.64	1.35	2.41	3.04
238	erythromycin	>35	734	2.54	0.65	14	5	Alert	198	1.97	3.55	1.02	4.71	5.77

^a Absorption data taken from Reference 12.

^b Experimental logP (MlogP) from (MlogP) from ClogP program.

^c Calculated logP (ClogP).

^d Calculated logP from Meylan method (24).

^e Sum of N and O H-bonding acceptors.

^f Sum of N and O H-bonding donors.

^g Computational alert according to the rule of 5; pass, no problem detected; alert, poor absorption more likely.

^h Polar surface area (PSA) from SAVOL program.

i The definition of zwitterionic compounds is based on the presence of both an ionizable acid group (either a carboxylic acid or an H-bearing tetrazole) and an ionizable base group (either a primary, secondary, or tertiary amine or a pyridine). They may not be zwitterions according to pKa values.

Table II. The Number and Percentage of Alerted Drugs with Poor, Medium, and High Absorption

		Low			Medium		High			
	Total	Alert	%Alert			Total Alert %Alert Total Alert			%Alert	
Rule of 5	-33	14	42	37	₆	16.	156			

$$
\%Abs = \frac{100}{1 + 10^{-Z}}\tag{13}
$$

If logit *FA* values are linearly related with the molecular descriptors, we can use either Eq. 12 to do linear regression analysis or Eq. 13 to do non-linear regression analysis. Table III lists the regression results (Models 21-26) both from linear and non-linear regression analysis by use of Eqs. 12 and 13. The results showed that the equation coefficients from the linear and non-linear regression are quite similar. Again, to apply Eq. 12, we had to remove the 100 and 0% absorption values from the linear regression analysis (models 24–26).

By examining the regression coefficients and standard errors of nonlinear models $12-14$ (using log k_{dif}) and models 21–23 (using logit *FA*) using the Abraham descriptors, Clog P and PSA, we find that the models, using $log k_{dif}$ and logit FA , respectively, gave the same regression coefficients and standard errors. The plot of logit *FA* against %*Abs* absorption by using Eq. (13) also showed a sigmoidal relationship (not shown here), which is quite similar to Fig.1, which uses Eq. (7). The similarity between Eqs. (13) and (7) is not surprising if we plot log k_{dif} against logit *FA*. The logarithm of the absorption rate constant (log k_{dif}) calculated from Eq. (6) is almost collinear with logit *FA* calculated from Eq. (12). It seems that logit *FA* is an approximation to the more correct $\log k_{\text{dif}}$.

The above analysis shows that there are two obvious outliers, ganciclovir and digoxin, when using the Abraham and PSA models. Ganciclovir shows a lower than expected absorption with Abraham descriptors (Fig. 4), and digoxin shows a greater than expected absorption with PSA (Fig. 3). For digoxin, it is possible that the descriptor is incorrect for digoxin or solubility was not correctly calculated in Reference 12 because it is a very large molecule. The reason is not clear for ganciclovir. It is possible that percent absorption value of ganciclovir is incorrect or ganciclovir is not diffusion ratelimited drug. Although the key parameter determining dissolution rate is solubility and the method based on equilibrium removes all low-solubility drugs, it is certainly possible that some dissolution rate-limited drugs are still present in the 189 compound data set. Dissolution rate is not only determined by the solubility and absorption but is also affected by effec-

Table III. Regression Results between Absorption and Molecular Descriptors

No	Method	Model	r^2	$\mathbf n$	S.
		%Absorption			
1	Abraham $+Z$	%Abs = $94 + 3.02$ E + 3.96 S – 21.7 A – 20.0 B + 9.18 V	0.68	189	16
2	$ClogP + Z$	%Abs = $72 + 7.32$ ClogP	0.55	189	118
3	$PSA + Z$	%Abs = $109 - 0.345$ PSA	0.52	189	19
4	Abraham	%Abs = $92 + 2.94$ E + 41.0 S – 21.7 A – 21.1 B + 10.6 V	0.74	169	14
5	Abraham $- S$	%Abs = 96 - 20.0 \bf{A} - 19.8 \bf{B} + 13.9 \bf{V}	0.72	169	15
6	ClogP	%Abs = $70 + 8.22$ ClogP	0.65	169	16
7	PSA	%Abs = $109 - 0.354$ PSA	0.56	169	18
8	Rule of 5	%Abs = 90 - 3.34 N_A - 5.56 N_D + 0.0716 MW	0.61	169	17
		$\log k_{\text{diff}} = \log{\ln[1/(1 - FA)]}\log t$			
9	Abraham $+Z$	%Abs = $100 \times [1 - EXP(-10^{0.423+0.135E-0.00398S-0.333A-0.370B+0.216V})]$	0.72	189	15
10	$ClogP + Z$	%Abs = $100 \times [1 - EXP(-10^{0.151+0.126} \text{ ClogP})]$	0.60	189	17
11	$PSA + Z$	%Abs = $100 \times [1 - EXP(-10^{0.712 - 0.0057} PSA)]$	0.53	189	19
12	Abraham	$\%$ Abs = 100 × [1 - EXP(-10 ^{0435+0.0848E+0.0405S-0.348A-0.403B+0.232V)]}	0.78	169	13
13	ClogP	%Abs = $100 \times [1 - EXP(-10^{0.100+0.166} \text{ ClogP})]$	0.72	169	15
14	PSA	%Abs = $100 \times [1 - EXP(-10^{0.777-0.00654} PSA)]$	0.59	169	18
15	Rule of 5	%Abs = $100 \times [1 - EXP(-10^{0.506 - 0.0754} N_A - 0.141 N_B + 0.00184 MW)]$	0.70	169	15
16	Abraham	$\log k_{\text{diff}} = 0.568 - 0.0363E + 0.141S - 0.507B + 0.232V$	0.77	128	0.31
17	Abraham $- S$	$\log k_{\text{dif}} = 0.600 - 0.347$ A - 0.498 B + 0.305 V	0.75	128	0.32
18	ClogP	$\log k_{\text{dif}} = -0.0532 + 0.195 \text{ ClogP}$	0.70	128	0.35
19	PSA	$\log k_{\text{dif}} = 0.841 - 0.00835 \text{ PSA}$	0.66	128	0.37
20	Rule of 5	$\log k_{\text{dif}} = 0.501 - 0.0929 \text{ N}_{\text{A}} - 0.115 \text{ N}_{\text{D}} + 0.00155 \text{ MW}$	0.70	128	0.36
		logit $FA = log[FA/(1 - FA)]$			
21	Abraham	%Abs = $100/[1 + 10^{-(1.02 + 0.0622E + 0.0977S - 0.599A - 0.681B + 0.445V)}]$	0.79	169	13
22	ClogP	%Abs = $100/[1 + 10^{-(0.453 + 0.283 \text{ ClogP})}]$	0.72	169	14
23	PSA	%Abs = $100/[1 + 10^{-(1.72 - 0.0124)}$ PSA)	0.59	169	18
24	Abraham	logit FA = $1.16 + 0.0451E + 0.129S - 0.625A - 0.685A - 0.685B + 0.353V$	0.71	128	0.50
25	ClogP	logit FA = $0.400 + 0.266$ ClogP	0.64	128	0.56
26	PSA	logit FA = $1.61 - 0.0113$ PSA	0.58	128	0.60

Note: S, stepwise regression; Z, zwitterionic compounds.

Fig. 2. Nonlinear fit of ClogP to absorption percent by model 13.

tive surface area of the solid drug and diffusion boundary layer thickness adjacent to the dissolving surface Eq. (1).

To examine absorption prediction for the dissolution rate-limited drugs identified by Eq. (11), absorption models (models 12–14), derived from diffusion rate-limited drugs, were used to predict the percent absorption value for dissolution rate-limited drugs (drugs 199–217). The results indicate that absorption prediction for some of the dissolution ratelimited drugs is in agreement with or is in the range of the observed absorption, whereas for others, the predicted values are higher than the observed absorption (12). This is in agreement with what we would expect; dissolution is the ratelimited step for most of these drugs and the observed absorption is equal to, or lower than the absorption predicted by the models derived from diffusion rate-limited drugs.

Absorption consists of very complex kinetic processes. Although absorption can be simplified into three transit processes (i.e., dissolution, diffusion, and perfusion), other processes can also be involved in absorption and can become rate limiting. Therefore, the identification of the rate-limiting step is very important in absorption modeling. On the basis of the dissolution rate equation, Eq. (3), it is reasonable to believe that dissolution rate-limited drugs could be classified from solubility and absorption [Eqs. (10) or (11)]. However, some of the calculated solubility values will not be reliable enough, and not all dissolution rate-limited drugs can be correctly classified from Eq. (10). This fact can be seen when comparing solubility values calculated from different models for the

Fig. 3. Nonlinear fit of PSA to absorption percent by model 14.

Fig. 4. Plot of absorption percent observed and predicted by model 12.

same drug compound. Large estimation errors were found for calculated solubility values depending on which method is used. Therefore, the use of experimental values in Eq. (10) or (11) is strongly suggested.

Recently, a Biopharmaceutics Classification Scheme (BCS) was proposed (6,29). Based on this scheme, drugs can be categorized into four basic groups according to their solubility and ability to penetrate the gastrointestinal mucosa: class I, high solubility and high permeability; class II, low solubility and high permeability; class III, high solubility and low permeability; and class IV, low solubility and low permeability. Dressman *et al.* (6) suggested that the rate of dissolution of the drug was almost certain to be the principal limitation to its oral absorption for class II drugs. If BCS is used to categorize the drugs (Table I) into the four BCS groups according to solubility and permeability (here we used predicted absorption instead of permeability), we find that most dissolution rate-limited drugs identified by Eq. (10) (drugs 199–217) belong to class II. The predicted percent of absorption is much higher than the observed percent of absorption because the dissolution rate is lower than the diffusion rate, and hence, dissolution is the rate-limited step for these drugs (12). Only two drugs, cyclosporin and amphotericin B, belong to class IV. Their predicted percent of absorption values (36 and 17, respectively) are low and close to observed percent of absorption (28 and 3, respectively).

Although it is easy to categorize the drugs into the four basic groups, it is very difficult to give a cutoff between low and high solubility because dissolution rate does not only depend on the solubility and dose but also on the permeability. However, if a drug is completely dissolved at the end of the gastrointestinal tract, then we can assume that dissolution rate will not affect the absorption. The minimum required solubility of a drug should satisfy the following equation.

Dose × (1 – Fraction absorbed) – $0.25 \times S_w(\text{min}) = 0$ (14)

where 0.25 (L) is the gastrointestinal volume (12), and S_w (min) is the minimum solubility in gastrointestinal fluid (mg/ L). If subjects were administered 100 mg as a single dose, the solubility values can be calculated; the relationship between percent of absorption and cutoff of solubility is linear. The cutoff of solubility increases when absorption decreases. If percent of absorption is zero, the solubility should be >400

Human Intestinal Absorption and QSAR 1457

mg/L, so that drugs can be completely dissolved in gastrointestinal tract. If percent of absorption is 99%, the solubility should be >4 mg/L, so that these drugs are calculated to be completely dissolved once absorption is finished. However, this is only the case at equilibrium. Absorption is a kinetic process. The cutoff of solubility values will be higher than that calculated by Eq. (14) if dissolution rate is very slow or the dose is very large. It is important to note that the absorption used to identify dissolution drugs in Eq. (10) is based on experimental values. In practice, the observed absorption may be unknown for some drugs. If we use the predicted absorption in Eq. (10), the cutoffs of solubility are lower than that based on the observed absorption for dissolution ratelimited drugs because the predicted absorption is higher than observed absorption for such drugs.

To test the models established and the method that is used to identify the dissolution rate-limited drugs, the absorption of 466 drug compounds, categorized as high, medium, and low absorption by GlaxoSmithKline, was predicted by using nonlinear model 12. The results show that the predicted absorption is in good agreement with the observed absorption for most of these compounds. However, the model overpredicts absorption for most of the poorly soluble compounds. This is because dissolution becomes the rate-limited step if the solubility of a drug is very low.

In conclusion, the present results show that nonlinear absorption models developed from a first-order kinetic process are better than the linear absorption model developed by our previous work (12). Identifying the rate-limited step and removing dissolution rate-limited drug compounds from the data set is the key factor in modeling human absorption of diffusion rate-limited drugs. Based on the absorption models developed from 189 diffusion rate-limited compounds in this article and application of the models to 466 drugs, our absorption models can predict the following three classes of compounds of BCS: class I, high solubility and high permeability; class III, high solubility and low permeability, and class IV, low solubility and low permeability. The absorption model overpredicts the absorption of class II, low solubility and high permeability. The following facts were found from absorption data for >500 drugs: (i) Permanently charged compounds have very low absorption. (ii) The absorption is usually very low if the calculated solubility is < 0.0001 mg/L. (iii) There are large prediction errors for zwitterions identified by Platts method (23).

REFERENCES

- 1. Y. C. Martin, E. Kutter, and V. Austel. *Modern Drug Research— Paths to Better and Safer Drugs*. Dekker, New York, 1989.
- 2. P. O. Gubbins and K. E. Bertch. Drug absorption in gastrointestinal disease and surgery. Clinical pharmacokinetic and therapeutic implications. *Clin. Pharmacokinet.* **21**:431–447 (1991).
- 3. M. Gibaldi. Limitations of classical theories of drug absorption. In: Prescott and Nimmo (eds.), *Drug Absorption: Proceeding of the Edinburgh International Conference*, ADIS Press, Auckland, 1979 pp. 1–5.
- 4. G. S. Banker and V. E. Sharma. Advances in controlled gastrointestinal absorption. In: Prescott and Nimmo (eds.), *Drug Absorption: Proceeding of the Edinburgh International Conference,* ADIS Press, Auckland, 1979 pp. 194–204.
- 5. J. B. Dressman, G. L. Amidon, and D. Fleisher. Absorption potential: estimating the fraction absorbed for orally-administered compounds. *J. Pharm. Sci.* **74**:588–589 (1985).
- 6. J. B. Dressman, G. L. Amidon, C. Reppas, and V. P. Shah. Dis-

solution testing as a prognostic tool for oral drug absorption: immediate dosage forms. *Pharm. Res.* **15**:11–22 (1998).

- 7. J. C. Dearden. *Molecular Structure and Drug Transport. Comprehensive Medicinal Chemistry*. C. Hansch (ed.). Pegamar, Oxford, 1990 pp. 375–411.
- 8. R. W. Foster. *Basic Pharmacology*. Reed Educational & Professional Publishing Ltd, Great Britain, 1996.
- L. S. Schanker, D. J. Tocco, B. B. Brodie, and C. A. M. Hogben. Absorption of drugs from the rat small intestine. *J. Pharmacol. Exp. Ther.* **123**:81–88 (1958).
- 10. J. T. Dolusio, N. F. Billups, L. W. Dittert, E. T. Sugita, and J. V. Swintosky. Drug absorption I: An *in situ* rat gut technique yielding realistic absorption rates. *J. Pharm. Sci.* **58**:1196–1200 (1969).
- 11. S. Yee. *In vitro* permeability across caco-2 cells (colonic) can predict *in vivo* (small intestinal) absorption in man—fact or myth. *Pharm. Res.* **14**:763–766 (1997).
- 12. Y. H. Zhao, J. Le, M. H. Abraham, A. Hersey, P. J. Eddershaw, C. N. Luscombe, D. Butina, G. Beck, B. Sherborne, I. Cooper, and J. A. Platts. Evaluation of human intestinal absorption data for use in QSAR studies and a quantitative relationship obtained with the Abraham descriptors. *J. Pharm. Sci.* **90**:749–784 (2001).
- 13. C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **23**:3–25 (1997).
- 14. O. A. Raevsky, V. I. Fetisov, E. P. Trepalina, J. W. McFarland, and K. J. Schaper. Quantitative estimation of drug absorption in humans for passively transported compounds on the basis of their physical-chemical parameters. *Quant. Struct. Act. Relat.* **19**:366– 374 (2000).
- 15. D. E. Clark. Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 1. Prediction of intestinal absorption*. J. Pharm. Sci.* **88**:807–814 (1999).
- 16. K. Palm, P. Stenberg, K. Luthman, and P. Artursson. Polar molecular surface properties predict and the intestinal absorption of drugs in humans. *Pharm. Res.* **14**:568–571 (1997).
- 17. M. D. Wessel, P. C. Jurs, J. W. Tolan, and S. M. Muskal. Prediction of human intestinal absorption of drugs from molecular structure. *J. Chem. Inf. Comput. Sci.* **38**:726–735 (1998).
- 18. S. D. Krämer. Absorption prediction from physicochemical parameters. *PSTT* **2**:36–42 (1999).
- 19. K. J. Schaper. Absorption of ionizable drugs: nonlinear dependence on logP, pKa and pH-quantitative relationships. *Quant. Struct. Act. Relat.* **1**:13–27 (1982).
- 20. M. H. Abraham, H. S. Chadha, G. S. Whiting, and R. C. Mitchell. Hydrogen bonding. 32. An analysis of water-octanol and wateralkane partitioning and the delta log P parameter of Seiler. *J. Pharm. Sci.* **83**:1085–1100 (1994).
- 21. J. B. Dressman and D. Fleisher. Mixing-tank model for predicting dissolution rate control of oral absorption. *J. Pharm. Sci.* **75**:109– 116 (1986).
- 22. B. Balon, B. U. Riebesehl, and B. W. Müller. Drug liposome partitioning as a tool for the prediction of human passive intestinal absorption. *Pharm. Res.* **16**:890–896 (1999).
- 23. J. A. Platts, M. H. Abraham, A. Hersey, and D. Butina. Estimation of molecular linear free energy relationship descriptors by a group contribution approach. 2. Prediction of partition coefficients. *J. Chem. Inf. Comp. Sci.* **40**:71–80 (2000).
- 24. W. K. Sietsema. The absorption oral bioavailability of selected drugs. *Int. J. Clin. Pharmacol. Ther. Toxicol.* **27**:179–211 (1989).
- 25. P. Michael Conn and G. F. Gebhart. *Essentials of Pharmacology*. F.A. Davis Company, Philadelphia, Pennsylvania, 1989.
- 26. T. B. Binns. *Absorption and Distribution of Drugs*. E. & S. Livingstone LTD., Edinburgh and London, 1964.
- 27. M. Rowland and T. N. Tozer. *Clinical Pharmacokinetics: concepts and Applications*. Lea & Febiger, Philadelphia, Pennsylvania, 1995.
- 28. W. L. Chiou and A. Barve. Linear correlation of the fraction of oral dose absorbed of 64 drugs between humans and rats. *Pharm. Res.* **15**:1792–1795 (1998).
- 29. G. L. Amidon, H. Lennernas, V. P. Shah, and J. R. Crison. A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability. *Pharm. Res.* **12**:413–420 (1995).