Use of Surface Charges from DFT Calculations To Predict Intestinal Absorption

Ron Jones,*,† Paul C. Connolly,‡ Andreas Klamt,§ and Michael Diedenhofen§

Strategic Technologies, Pharmacy Division, GlaxoSmithKline Research and Development, South Eden Park Road, Beckenham, Kent BR3 3BS, U.K., Strategic Technologies, Pharmacy Division, GlaxoSmithKline Research and Development, Park Road, Ware SG12 ODP, U.K., and COSMOlogic GmbH&Co.KG, Burscheider Strasse 515, 51381 Leverkusen, Germany

Received November 17, 2004

Ⓦ **This paper contains enhanced objects available on the Internet at http://pubs.acs.org/journals/jcisd8.**

A model for prediction of percent intestinal absorption (%Abs) of neutral molecules was developed based upon surface charges of the molecule calculated by density functional theory (DFT). The surface charges are decomposed into σ moments which are correlated to a partition coefficient representing transfer of the molecule between water and the epithelial membrane. The model was built and tested using a data set of 241 drugs. It achieved an RMS deviation of 13% on a training set of 38 compounds as well as on a test set of 107 drugs for which the experimental data were classified as high quality. Property maps of the molecule, depicting which atoms contribute to or hinder absorption, are produced to aid drug design.

1. INTRODUCTION

Predicting the interactions of drug molecules in biological systems provides a continuing challenge to the computational chemist. Quantum mechanical descriptions of the drug coupled with continuum models of the solvent are effective for describing the interaction in a homogeneous solution.[1,2] The COSMO-RS theory goes beyond simple continuum models by treating the solvent as an ensemble of neutral molecules for which the surface polarization charges have been calculated using quantum mechanics.[2] The purpose of this work is to extend the approach to intestinal absorption, where the intestinal epithelium cannot be completely described, and to derive a general absorption equation which can be applied to drug candidates based only on a calculation of the surface charge using standard quantum mechanics.

Oral absorption of drugs is a complex process including disintegration of the formulation, dissolution of the drug, passive absorption across the intestinal epithelium, and in some cases active transport, efflux (e.g. by P-glycoprotein), and gut wall metabolism. As detailed information on active transport or efflux is not available on an individual drug at the early development stage, the preferred approach is to model passive transcellular absorption without explicitly considering other factors, while recognizing that it will introduce inaccuracies for some molecules.

Previous approaches for predicting intestinal absorption include a nonlinear neural network model using 2d and 3d molecular descriptors,[3] methods using dynamic polar surface area[4] and polar surface area,[5] a fragment based method which was developed using Abraham's descriptors,[6] and a combination of a fragment based method with hydrogen bond donor ability.[7] Recent work has also been directed toward predicting intestinal permeability directly.[8,9] However the expense of permeability studies in human subjects limits the size of the data set that is available to be modeled.

Theoretical Framework. The theoretical basis for interactions of a molecule with the intestinal membrane is derived from COSMO-RS theory which addresses the interaction of molecular surfaces as computed by quantum chemical methods (QM). COSMO-RS combines an electrostatic theory of locally interacting molecular surface descriptors (which are available from QM calculations) with a statistical thermodynamics methodology. QM-COSMO calculations provide a discrete surface around a molecule embedded in a virtual conductor. Each segment i of this surface is characterized by its area a_i and the screening charge density (SCD) σ_i on the segment which takes into account the electrostatic screening (reduction of electrostatic forces) on the solute molecule by its surrounding (which in a virtual conductor is perfect screening) and the back-polarization of the solute molecule which induces additional surface charges. In addition, the total energy of the ideally screened molecule, E_{COSMO}, is provided. Within COSMO-RS theory a liquid is now considered an ensemble of closely packed ideally screened molecules. To achieve this close packing, the system has to be compressed, and thus the cavities of the molecules become slightly deformed although the volume of individual cavities does not change significantly.

Each piece of the molecular surface is in close contact with another molecular surface. Assuming that there is still a conducting surface between the molecules, i.e., that each molecule is still enclosed by a virtual conductor, in any contact area, the surface segments of both molecules have net SCDs σ and σ'. In reality there is no conductor between the surface contact areas. Thus an electrostatic interaction arises from the contact of two different SCDs. The specific interaction energy per unit area resulting from this "misfit" of SCDs is given by

$$E_{misfit}(\sigma,\sigma') = a_{eff}\frac{\alpha'}{2}(\sigma+\sigma')^2 \quad (1)$$

* Corresponding author phone: +44 (0)20-8639-6447; fax: +44 (0)-1920-88-3873; e-mail: ron.6.jones@gsk.com.
† GlaxoSmithKline Research and Development, Beckenham.
‡ GlaxoSmithKline Research and Development, Ware.
§ COSMOlogic GmbH&Co.KG.

10.1021/ci049653f CCC: $30.25
Published on Web 08/20/2005

where a_{eff} is the effective contact area between two surface segments and α' is an adjustable parameter. The basic assumption of eq 1 (which is the same as in other surface pair models, e.g. UNIQUAC[10]) is that residual nonsteric interactions can be described by pairs of geometrically independent surface segments. Hydrogen bonding (HB) can also be described by the two adjacent SCDs. HB donors have a strongly negative SCD, whereas HB acceptors have strongly positive SCDs. Generally, an HB interaction can be expected if two sufficiently polar pieces of surface of opposite polarity are in contact. Such a behavior can be described by a functional of the form

$$E_{hb}(\sigma,\sigma') = a_{eff}\, c_{hb} \min(\sigma\sigma' + \sigma_{hb}{}^2, 0) \quad (2)$$

where c_{hb} and σ_{hb} are adjustable parameters. In addition to electrostatic misfit and HB interaction, COSMO-RS also takes into account van der Waals (vdW) interactions between surface segments via

$$E_{vdW}(e,e') = a_{eff}(\tau(e) + \tau(e')) \quad (3)$$

where e and e' indicate the elements of the atoms below the surface segments, and $\tau(e)$ is an adjustable element specific parameter.

The link between the microscopic surface interaction energies and the macroscopic thermodynamic properties is provided by statistical thermodynamics. As COSMO-RS treats all molecular interactions as local pairwise interactions of surface segments, the statistical averaging can be done on the ensemble of interacting surface pieces. A knowledge of the probability distribution of σ for all compounds X_I is sufficient to describe the ensemble. The probability distributions $p_x(\sigma)$ are called σ-profiles. The σ-profile of the mixture $p_s(\sigma)$ is the sum of the σ-profiles of the components weighted with their mole fraction.

The chemical potential of a surface segment with SCD σ in an ensemble described by $p_s(\sigma)$ is

$$\mu_S(\sigma) = -kT \ln\left[\int \frac{p_S(\sigma')}{A_S} \exp\left\{-\frac{E(\sigma,\sigma') - \mu_S(\sigma')}{kT}\right\} d\sigma'\right] \quad (4)$$

where $E(\sigma,\sigma')$ denotes all types of interactions of the segments, and A_S is the mean surface area of the mixture. In the case of transfer across a membrane, the membrane molecules are constrained, and this theoretical approach cannot be used to determine chemical potential directly. Instead, the σ-potential of the drug is expanded in a similar way to a Taylor series to give σ-moments.

$$\mu_S(\sigma) \cong \sum_{i=-2}^{m} c_S^i f_i(\sigma) \quad (5)$$

with

$$f_i(\sigma) = \sigma^i \quad \text{for } i \geq 0 \quad (6)$$

and

$$f_{-2/-1}(\sigma) = f_{acc/don}(\sigma) \cong \begin{cases} 0 \text{ if } \pm\sigma < \sigma_{hb} \\ \mp\sigma + \sigma_{hb} \text{ if } \pm\sigma > \sigma_{hb} \end{cases} \quad (7)$$

Any logarithmic partition coefficient between 2 phases (e.g. aqueous solution and membrane) should now be a linear combination of the drug σ-moments:

$$\log K^X_{S,S'} = c_{S,S'} + c^{acc}_{S,S'} M^X_{acc} + c^{don}_{S,S'} M^X_{don} + c^0_{S,S'} M^X_0 + c^2_{S,S'} M^X_2 + c^3_{S,S'} M^X_3 + \ldots \quad (8)$$

where all coefficients $c_{SS'}$ depend only on the two solvent phases, and the σ-moments M^X_i of the solute X are defined by

$$M^X_i = \int p^X(\sigma) f_i(\sigma) d\sigma \quad (9)$$

with $p^X(\sigma)$ being the composition function of the molecular surface with respect to the polarization charge density σ (σ-profile). Note that the first σ-moment M^X_1 has been left out in eq 7, because it is the total charge of the solute molecule and it thus vanishes for neutral compounds. This approach is extremely versatile and has been successfully applied to such different problems as adsorption onto activated carbon,[11] soil sorption coefficients,[12] and to various organic solvent−water partition coefficients. A more detailed description of the σ-moment approach is given in a recent review.[13]

2. METHODS

Experimental Data. The experimental data for the percentage of human intestinal absorption of 241 drugs has been taken from the study of Zhao et al.[6] Compound names and experimental data are listed in Table 1 of the Supporting Information. A total of 241 drugs were in the original paper, of which 169 were deemed to have reliable absorption data.

The model was built using drugs which do not exhibit dose or formulation dependent absorption.

Calculation of σ-Moment Descriptors. σ-Moments are not commonly used within the pharmaceutical industry, though much of the basic software for their calculation is used within the industry for drug discovery. A typical sequence to calculate σ-moments is given below. It should be noted that all of the software is commercially available.

1. If the molecule is flexible, select a low-energy conformation by use of a force field such as MMFF[14] or a program such as CORINA[15] which predicts conformation.

2. Optimize the geometry using MOPAC 2000 or a later version, using the AMI method and setting the van der Waals radii to 1.17 times the default values.[16]

3. Using the MOPAC geometry as input, carry out a single point density functional calculation to produce a COSMO file holding the surface polarization charges. Suitable software includes Turbomole[17] (as used in the current work) and Gaussian.[18] For Turbomole use the SVP basis set, the BP functional, and the RIDFT method. A single calculation using a PC with a 2 GHz processor running Turbomole under LINUX takes 1 h for a typical drug molecule e.g. naloxone.

4. Calculate the σ-moments using the COSMOtherm software.[19]

In our experience with other systems, a partition equilibrium can be characterized using moments up to the third order in eq 1. Thus we need 5 σ-moments to describe arbitrary partition coefficients.

Most of the σ-moments have a well-defined meaning. M^X_0 is the surface area of the molecule, M^X_2 is an integral

measure of electrostatic interaction energy, M_{acc}^{X} is a descriptor for hydrogen bond acceptor interactions of the compound, and M_{don}^{X} is the equivalent descriptor for hydrogen bond donor interactions. The third-order moment is related to the asymmetry of the σ-profile but has no simple interpretation. M_1^X is the molecular charge which is zero for a neutral molecule and is therefore not included.

Statistical Analysis. A number of alternative approaches have been used to convert percent intestinal absorption (%Abs) into an absorption rate. Zhao et al. in their original work modeled %Abs directly as a free energy function.[6] Lennernas has suggested that intestinal absorption is membrane controlled for both high and low permeability compounds.[20] Abraham et al. proposed that absorption is controlled by a combination of diffusion through the stagnant mucosal layer together with transfer across the mucosal−membrane interface.[22] For this study we compared Abraham's diffusion and transfer (kinetic) model with a simple partition model which simulates transfer across the epithelial membrane. Although the kinetic model should reproduce the in vivo situation, in practice a partition model fitted the data better and is described below.

The σ-moments are correlated to free energy changes and are suitable for linear regression analysis of logarithmic partition coefficients. However the experimental data of percent intestinal absorption (%Abs) are not themselves partition coefficients. They can be converted into partition coefficients, $K_{ia,}$ by the relationship

$$K_{ia} = \frac{\%\text{Abs}}{100 - \%\text{Abs}} \tag{10}$$

The relationship is analogous to the distribution of 100% of solute between a phase containing a% and a second phase containing $100 - a$% and produces a sigmoid relationship similar to that between %Abs and intestinal permeability. However this leads to a problem for the experimental values reported as 0% or 100%, respectively, for which K_{ia} cannot be estimated accurately from eq 10. To use the full data set we define two thresholds %Abs_min and %Abs_max = 100 − %Abs_min and values K_{ia}_max and K_{ia}_min which correspond to these thresholds via eq 10.

For the subsequent regression analysis of log K_{ia} with respect to the σ-moments, we consider initially the range K_{ia}_min to K_{ia}_max. Outside this range, log K_{ia} is taken as smaller than log K_{ia}_min if %Abs < %Abs_min, and log K_{ia} as greater than log K_{ia}_max, if %Abs > Abs_min. The error for a compound is zero, if the predicted value of log K_{ia} is less than log K_{ia}_min and %Abs < %Abs_min, or if the predicted value of log K_{ia} is greater than log K_{ia}_max and %Abs > %Abs_max.

This treatment of errors for compounds outside the range means that standard linear regression analysis can no longer be used to determine the optimal coefficients of the σ-moments in the linear model of log K_{ia}. Therefore we applied a self-written quasi-Newton nonlinear optimization procedure to minimize the sum of the squared residual errors with respect to the coefficients $c_{IA}{}^{i}$ in eq 8. This procedure converges rapidly, and the results are independent of the initialization of the variables. The resulting linear eq 8 using these coefficients is then applied to drug molecules to calculate log K_{ia} (see eq 11 below).

Finally the results for log K_{ia} are converted back into values for %Abs using eq 10, and the residuals in %Abs are used to calculate a root-mean-squared error (RMSE) which can be compared with other methods.

3. RESULTS AND DISCUSSION

The data set was divided into training and test sets in order to test the predictive power of the model. Application of the procedure described in the previous section to the 38 compound training set 1 described by Zhao[6] yields the regression equation

$$\log K_{ia} = 0.0040M_0 - 0.0053M_2 - 0.0024M_3 - 0.113M_{acc} - 0.11656M_{don} + 1.37 \tag{11}$$

On the training set this yields an RMSE of 12.5 for %Abs. A value of %Abs _min = 3 was chosen in the regression analysis, since this corresponds to a shallow minimum of the RMSE (RMSE = 13.8 for %Abs_min = 0.1, RMSE = 13.9 for %Abs_min = 12). Hence %Abs cutoffs from 3 to 97% were used in deriving the equation. On the test set of 131 compounds[6] the RMSE is 15. If we reduce the test set to the 107 experimental values classified as GOOD or OK in the source data,[6] the RMSE reduces to 12.8. Hence the model is very stable when applied to the test set.

Comparison with Other Models. Most of the reported methods have not used this large data set and cannot be compared directly. The dynamic polar surface area performs somewhat better (RMSE 9.2) on a smaller calibration set of 20 molecules but has not been applied to an independent test set.[4] The single conformer polar surface area intended as a rapid screening method performs less well on a larger data set taken from Wessel et al.[3,5]

A neural net method gave an RMSE of 16.0 for an external test set.[3]

The same data set has been modeled previously using Abraham's molecular descriptors.[6,23] A fragment-based method (ABSOLV) is used to decompose the molecule into fragments for which data are available from previous measurements.

In comparison to Abraham's descriptors, COSMO gives slightly better results for the training set (12.5 vs 14) and slightly worse for a test set of 131 drugs (15 vs 14) which included both high quality and questionable data. Figure 1 shows a plot of predicted vs experimental %Abs using the COSMO approach for the calibration set and the high quality data set.

It can be seen from Figure 1 that the data are clustered close to 100% absorption. This is a feature of the data available in the literature.

The largest outlier within the good-quality data is lamivudine, for which an experimental value of 87% is reported, while COSMO-%Abs yields 42%. Absorption of the drug may be increased by a nucleoside specific transporter.[24] Within the questionable data from the source data set, the errors for mannitol (exp. 16%, pred 69%) and lincomycin (exp. 28%, pred 77%) are even larger. Both would have been classified as well absorbed by COSMO-%Abs, although the experimental results indicate poor absorption.

The data set of 20 zwitterionic compounds (compounds 173−192) was treated in nonionic form. Applying eq 6 to

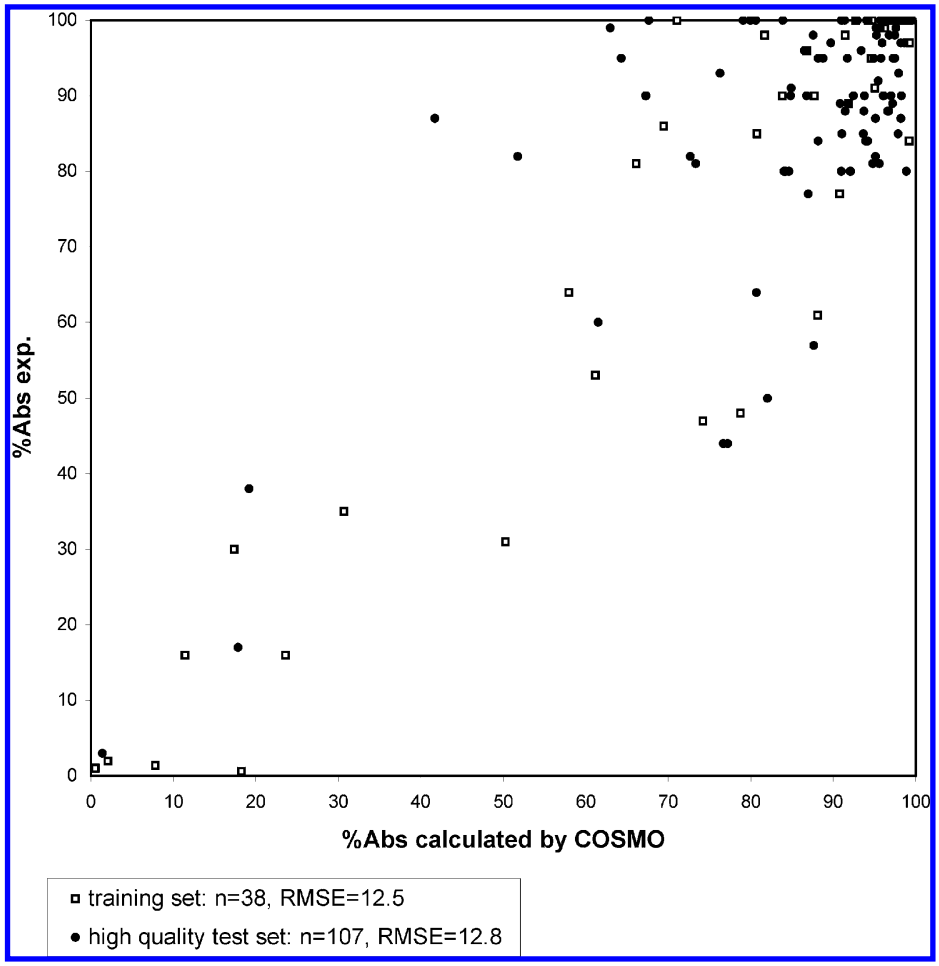

Figure 1. Experimental vs predicted intestinal absorption.

this subset yields an RMSE of about 22%. The current model appears to underestimate the absorption in most cases. It may be that approach to the membrane surface is more energetically favorable for the charge separated form relative to the neutral molecule The current model applies only to uncharged molecules but will be modified in the future to treat permanent charges or charge separation.

Compounds 193−201 are considered to give unreliable predictions by the Abraham model due to missing fragments in the calculation of the parameters. Apart from compounds 193 and 201, which are permanently ionized across the physiological pH range, these compounds have been successfully treated by COSMO-%Abs. Five of them give very close agreement with the experimental data. For viomycin and capreomycin we predict the lowest absorption values of the entire data set (0.01% and 0.02%), while experimental values of 85% and 50% are reported for these compounds. Viomycin and capreomycin are both peptides which are unlikely to be well absorbed by passive diffusion but may be absorbed via a peptide transporter.

For most of the 27 compounds (202−228) classified as dose-limited or as dose or formulation dependent, the predictions of COSMO-%Abs consistently overestimate the percentage absorption. This is as expected for a model which assumes that dissolution is not rate-limiting. Only for methotrexate is the predicted value significantly lower than the reported experimental range of 57−83%. This may be due to active transport via the folate transporter. Finally, for 12 of the 13 compounds (229−241) for which the reported

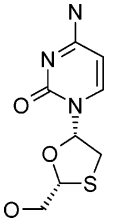

Figure 2. Structure of lamivudine.

experimental value is believed to be underestimated, we correctly predict higher values. In one case we have an underestimation of 5% which is not thought to be significant.

The COSMO descriptors are intended to provide a complete description for the surface properties of molecules. By using the full set of descriptors it is hoped that, despite the limited size of the training set, the method will prove robust when challenged with novel structures.

Prediction of Structural Effects on Absorption. Since the σ-moments are surface integrals, log K_{ia} determined by eq 11 can be color coded on the molecular surface. This visualization is a useful tool for molecular design, allowing the identification of features which aid or hinder absorption. As an example, the structure of lamivudine is shown in Figure 2, and the color coded surface is shown in Figure 3.

Surface regions with a positive effect on absorption are shown in red, regions of strong negative influence are indicated by blue and violet, and regions of negligible influence are colored green. It can be seen that the carbonyl and the adjacent unsubstituted pyrimidine nitrogen have a

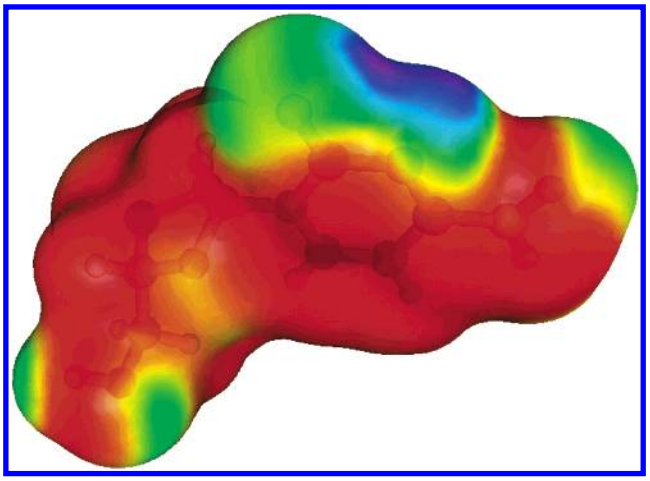

Figure 3. Surface of lamivudine colored by log Ki_a.

Ⓦ A 3D rotatable image in wrl format is available.

negative influence on intestinal absorption, while the rest of the pyrimidine ring and the oxathiole have a positive influence. The hydroxy and amino ring substituents have little effect.

It is possible using group contribution methods to classify functional groups as helping or hindering absorption. However the use of whole molecule calculations is expected to be more reliable for heterocyclic ring systems which are common in drug structures and where the entire molecule must be treated to obtain accurate charge distributions. The DFT calculations are slower than group contribution methods, which must be set against the extra information produced.

Recent application of this model via the high-throughput fragmentation based COSMO-RS shortcut COSMOfrag[21] resulted in an increase in the rms error of only 0.4% compared to the nonfragment based calculation procedure described here. This shows that the results of the intestinal absorption model are not very sensitive to the details of the conformations, since the conformations of the molecules used for the fragments are different in many cases from those used in the present study.

4. SUMMARY

COSMO-%Abs is a new, broadly applicable method for the a priori prediction of percentage intestinal absorption of drugs. It is based on σ-moments as molecular descriptors, which are derived from quantum chemical density functional calculations combined with the continuum solvation model, COSMO. The underlying σ-moment approach is theoretically well justified and has been successfully validated for other partition properties. The rms-deviation of COSMO-%Abs from experimental data is about 13%. As expected, the results are low for drugs where absorption is solubility or formulation dependent. The predictive quality of COSMO-%Abs is comparable to previous models. Potential advantages of this approach are the calculation from molecular structure without the need for a fragment database and the production of color coded property maps of the molecular surface to aid drug design. A possible disadvantage is that the speed of calculation is more suited to comparing a limited set of molecules than for high-throughput screening. However rapid calculation of σ-profiles via the COSMOfrag method only slightly decreases the accuracy and may offer an approach to rapid screening.

ACKNOWLEDGMENT

We thank Mike Abraham, Chris Luscombe, and Anne Hersey for supplying the structures and tables of data in electronic format and Per Artursson for valuable advice.

Supporting Information Available: Experimental and calculated data used in the paper (Table 1). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Miertus, S.; Scrocco, E.; Tomasi, J. Electrostatic Interaction of a Solute with a Continuum. A Direct Utilization of ab initio Molecular Potentials for the Prediction of Solvent Effects. *Chem. Phys.* **1981**, *55*, 117−129.
(2) Klamt, A.; Eckert, F.; Hornig, M. COSMO-RS: A novel view to physiological solvation and partition questions. *J. Comput.-Aid. Mol. Des.* **2001**, *15*, 355−365.
(3) Wessel, M. D.; Jurs, P. C.; Tolan, J. W.; Muskal, S. M. Prediction of Human Intestinal Absorption of Drug Compounds from Molecular Structure. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 726−735.
(4) 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.
(5) Clark, D. E. Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 1. Prediction of intestinal absorption. *J. Pharm. Sci.* **1999**, *88*, 807−814.
(6) Zhao, Y. H.; Le, J.; Abraham, M. H.; Hersey, A.; Eddershaw, P. J.; Luscombe, C. N.; Boutina, D.; Beck, G.; Sherborne, B.; Cooper, I.; Platts, J. A. Evaluation of human intestinal absorption data and subsequent derivation of a quantitative structure−activity relationship (QSAR) with the Abraham descriptors. *J. Pharm. Sci.* **2001**, *90*, 749−784.
(7) Klopman, G.; Stefna, L. R.; Saiakov, R. D. ADME evaluation: 2. A computer model for the prediction of intestinal absorption in humans. *Eur. J. Pharm. Sci.* **2002**, *17*, 253−263.
(8) Winiwarter, S.; Bonham, N. M.; Ax, F.; Hallberg, A.; Lennernas, H.; Karlen, A. Correlation of human jejunal permeability (in vivo) of drugs with experimentally and theoretically derivied parameters. A multivariate data analysis approach. *J. Med. Chem.* **1998**, *41*, 4939−4949.
(9) Egan, W.; Lauri, G. Prediction of intestinal permeability. *Adv. Drug Deliv. Rev.* **2002**, *54*, 273−289.
(10) Abrams, D. S.; Prausnitz, J. M. Statistical thermodynamics of liquid mixtures: A new expression for the excess Gibbs energy of partly or completely miscible systems. *AICHE J.* **1975**, *21*, 116−128.
(11) Mehler, C.; Klamt, A.; Peukert, W. Use of COSMO-RS for the prediction of adsorption equilibria. *AICHE J.* **2002**, *48*, 1093−1099.
(12) Klamt, A.; Eckert, F.; Diedenhofen, M. Prediction of Soil Sorption Coefficients with COSMO-RS. *Environ. Toxicol. Chem.* **2002**, *21*, 2562−2566.
(13) Klamt, A.; Eckert, F.; Hornig, M. COSMO-RS: A novel view to physiological solvation and partition questions. *J. Comput.-Aid. Mol. Des.* **2001**, *15*, 355−365.
(14) Halgren, T. A.; Nachbar, R. B. Merck molecular force field. IV. conformational energies and geometries for MMFF94. *J. Comput. Chem.* **1996**, *17*, 587−615.
(15) Sadowski, J.; Gasteiger, J. From atoms and bonds to three-dimensional atomic coordinates: automatic model builders. *Chem. Rev.* **1993**, *93*, 2567−2581.
(16) Stewart, J. J. P. MOPAC 2000.00 manual, Fujitsu: Tokyo, 1999.
(17) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. Electronic structure calculations on workstation computers: The program system turbomole. *Chem. Phys. Lett.* **1989**, *162*, 165−169.
(18) Gaussian 03, Revision C.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck,

A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Wallingford, CT, 2004.

(19) Klamt, A.; Eckert, F. COSMO*therm*, Version C1.1-Revision 01.01; COSMO*logic* GmbH&CoKG, Leverkusen, Germany, 2001.

(20) Lennernas, H. Human intestinal permeability. *J. Pharm. Sci.* **1998**, *87*, 403−410.

(21) Hornig, M.; Klamt, A. COSMO*frag*: A novel tool for high throughput ADME property prediction and similarity screening based on quantum chemistry, submitted for publication in *J. Chem. Inf. Model.* **2005**.

(22) Abraham, M. H.; Zhao, Y. H.; Le, J.; Hersey, A.; Luscombe, C. N.; Reynolds, D. P.; Beck, G.; Sherborne, B.; Cooper, I. On the mechanism of human intestinal absorption. *Eur. J. Med. Chem.* **2002**, *37*, 595−605.

(23) Platts, J. A.; Butina, D.; Abraham, M. H.; Hersey, A. Estimation of Molecular Linear Free Energy Relation Descriptors Using a Group Contribution Approach. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 835−845.

(24) Yao, S. Y.; Cass, C. E.; Young, J. D. Transport of the antiviral nucleoside analogues 3′-azido-3′-deoxythymidine and 2′,3′-dideoxycytidine by a recombinant nucleoside transporter (rCNT) expressed in *Xenopus laevis* oocytes. *Mol. Pharmacol.* **1996**, *50*, 388−393.

CI049653F