Molecular Properties That Influence the Oral Bioavailability of Drug Candidates

Daniel F. Veber,*[,†] Stephen R. Johnson,[‡,§] Hung-Yuan Cheng,[||] Brian R. Smith,[⊥] Keith W. Ward,[⊥] and Kenneth D. Kopple[‡,#]

Departments of Medicinal Chemistry, Cheminformatics, Computational Analytical and Structural Sciences, and Drug Metabolism and Pharmacokinetics, GlaxoSmithKline, 709 Swedeland Road, P. O. Box 1539, King of Prussia, Pennsylvania 19406-0939

Received January 9, 2002

Oral bioavailability measurements in rats for over 1100 drug candidates studied at SmithKline Beecham Pharmaceuticals (now GlaxoSmithKline) have allowed us to analyze the relative importance of molecular properties considered to influence that drug property. Reduced molecular flexibility, as measured by the number of rotatable bonds, and low polar surface area or total hydrogen bond count (sum of donors and acceptors) are found to be important predictors of good oral bioavailability, independent of molecular weight. That on average both the number of rotatable bonds and polar surface area or hydrogen bond count tend to increase with molecular weight may in part explain the success of the molecular weight parameter in predicting oral bioavailability. The commonly applied molecular weight cutoff at 500 does not itself significantly separate compounds with poor oral bioavailability from those with acceptable values in this extensive data set. Our observations suggest that compounds which meet only the two criteria of (1) 10 or fewer rotatable bonds and (2) polar surface area equal to or less than 140 Å^2 (or 12 or fewer H-bond donors and acceptors) will have a high probability of good oral bioavailability in the rat. Data sets for the artificial membrane permeation rate and for clearance in the rat were also examined. Reduced polar surface area correlates better with increased permeation rate than does lipophilicity (*C* log *P*), and increased rotatable bond count has a negative effect on the permeation rate. A threshold permeation rate is a prerequisite of oral bioavailability. The rotatable bond count does not correlate with the data examined here for the in vivo clearance rate in the rat.

Introduction

High oral bioavailability is often an important consideration for the development of bioactive molecules as therapeutic agents. Thus, an important goal for drug research is to gain sufficient understanding of the molecular properties that limit oral bioavailability to facilitate the design of viable new drug candidates. Poor oral bioavailability can result in variable exposure to active drug, especially when the factors that limit it are compromised in a specific individual. Polymorphic variability in drug-metabolizing enzymes, or co-administration with drugs that may inhibit such enzymes, may reduce the first pass clearance and increase drug exposure to undesired levels, for example. A general lack of systematically obtained oral bioavailability data on compounds with diverse molecular properties in a single animal species has limited the ability to make correlations with structure and physical properties.

Analysis of the structures of orally administered drugs, and of drug candidates, as pioneered by Lipinski, has so far been the primary guide to correlating physical properties with successful drug development.[1,2] This analysis has been very useful and has led to a set of rules relating to the importance of lipophilicity (octanol−water partition), molecular weight (MW), and the number of hydrogen bond donors and acceptors. Nonetheless, there are limitations on it as it relates to oral bioavailability. These include the lack of quantitative assessment of oral bioavailability in the data analyzed, a need to assume that all orally administered drugs are intended to be absorbed, the assumption that oral bioavailability is generally high for orally administered drugs, and the fact that many properties other than oral bioavailability, such as crystallinity, ease of formulation, chemical stability, and practical availability by synthesis or isolation, enter into the choice of a compound for drug development.

In addition to the molecular properties discussed by Lipinski, other properties have been discussed in regard to oral bioavailability. Navia has postulated the desirability of molecular flexibility for membrane permeation.[3] Hirschmann has focused on the undesirable property of water complexation by amide bonds as a negative factor for oral bioavailability.[4] The negative impact of a high polar surface area on intestinal absorption is recognized.[5,6] Membrane permeation is recognized as a common requirement for oral bioavailability in the absence of active transport, and failure to achieve this usually results in poor oral bioavailability.

* To whom correspondence should be addressed at 290 Batleson Rd., Ambler, PA 19002. Phone/Fax: (215) 542 7831. E-mail: veberdf@comcast.net.
† Department of Medicinal Chemistry.
‡ Department of Cheminformatics.
|| Department of Computational Analytical and Structural Sciences.
⊥ Department of Drug Metabolism and Pharmacokinetics.
§ Present address: Bristol-Myers Squibb PR1, P.O. Box 4000, Princeton, NJ 08453-4000.
# Present address: 720 Glengarry Rd., Philadelphia, PA 19118. E-mail: kdkopple@bellatlantic.net.

10.1021/jm020017n CCC: $22.00
Published on Web 05/11/2002

Passive membrane permeation, which can be measured using artificial lipid bilayers[7] or hydrocarbon films,[8] is not structure specific but is dependent on the kind of macroscopic properties used in Lipinski's evaluation.

Many other factors are now also recognized as limiting oral bioavailability.[9−11] These include energy-driven export from the blood to the gut by transporter enzymes of intestinal or liver cells (e.g., P-glycoproteins) and first pass metabolism by enzymes of intestinal or liver cells, including oxidation by cytochromes P-450, glycosylation, sulfation, glucuronidation, etc. In contrast to membrane permeation, these enzymatic clearance mechanisms, both metabolic and transporter-mediated, are most likely to be subject to structure-specific recognition. Structural features involved in metabolic clearance have been made the basis of a classifier for oral bioavailability.[12] Structure-dependent oral bioavailability has been observed in the course of recently reported studies from these laboratories on inhibitors of the osteoclast specific protease cathepsin K.[13,14] In these studies, molecular rigidity was proposed to have played a role in locking out access to clearance enzymes while inhibitory potency against the target was retained. Specific structural requirements of transport mechanisms using Caco-2 cells have also been noted.[15] Such observations led us to the search for structure−oral bioavailability correlations in the SmithKline Beecham compound oral bioavailability database that is described here.

With the increased application of pharmacokinetic evaluation during the course of drug discovery, quantitative pharmacokinetic data in a single animal species have only recently become available in a quantity sufficient to allow analysis of the relationship of structural properties and oral bioavailability. The rat is the only species for which we have access to a database of the needed size and diverse composition of compounds. We have examined data for over 1100 compounds from diverse projects at the former SmithKline Beecham Pharmaceuticals with respect to the dependence on some of the molecular properties known to affect oral bioavailability. Molecular diversity of this set is limited by its members having been products of therapeutically directed medicinal chemistry efforts, and the group may thus be biased by analogue series. Nonetheless, a significant number of projects and compound classes are represented.

Rat whole body clearance rates were also recorded for those compounds for which rat oral bioavailability was measured. In addition, we had access to artificial membrane (lipid bilayer) permeation rate measurements for over 3000 compounds and rat liver microsomal clearance for over 4000. For comparison with oral bioavailability we explored the dependence of these observations on the same molecular properties.

Lipophilicity, commonly estimated by a calculated log *P*, and molecular mass as descriptors do not offer insight regarding the structure-specific properties that influence oral bioavailability. No precedent exists for finding simple descriptors to account for the consequences of cyclization and for nonbonded, intramolecular interactions as they limit molecular flexibility. We have taken a simple approach involving the rotatable bond count and found it to be unexpectedly useful. We have also found the use of the rotatable bond count in conjunction with the polar surface area to give an effective rule for prediction of oral bioavailability.

Table 1. Diversity As Measured by Fingerprint-Based Clusters per 1000 Compounds for the Compound Sets Examined in This Work[a]

Tanimoto similarity	0.6	0.7	0.8
oral bioavailability/ in vivo clearance set	229	332	500
permeation set	131	206	334

[a] To facilitate comparison, the number of clusters for each data set has been divided by the number of compounds and multiplied by 1000.

Methods

Measurements. Rat Oral Bioavailability and Clearance. The following protocol is typical of those used for ca. 1100 compounds that were submitted from a variety of SmithKline Beecham drug discovery programs, although variations, inconsequential in the context of this analysis, were employed.

Data were generated in male Sprague−Dawley rats using a crossover design on two separate study days. Intravenous dosages of test molecules were administered via femoral vein catheters which had been surgically implanted at least 3 days prior to the start of the study. On study day 1, the animals received test compounds at pharmacologically relevant dosages as an iv infusion of 0.5−1 h (4.0 mL/kg). On study day 2, the animals received test molecules, at similar doses, by oral gavage formulated as suspensions in methylcellulose or solution. In some cases formulation aids, such as Encapsin, were used to enhance solubilization. In these cases, the formulation aid was at a level that we have found to have a minimal effect on pharmacokinetics. Circulating concentrations of test compounds were determined using LC/MS/MS methods with demonstrated specificity and error over a concentration range of 10.0 ng/mL (LLQ) to 2500 ng/mL (1 day validation). Pharmacokinetic parameters were calculated from concentration versus time data using noncompartmental pharmacokinetic methods with the pharmacokinetic analysis software WinNonlin Professional Version 2.1.

Artificial membrane permeation rates were available for 3061 compounds. This set included 341 compounds also present in the oral bioavailability set. The membrane permeation assay used, developed at SmithKline Beecham,[16] is based on permeation through planar polycarbonate filter supported lipid bilayers[17] and employed the high-throughput parallel chemical analysis approach.[7,18] Rates measured ranged from 0.1 to 2000 nm/s. The maximum rate was limited by the filter pore size. Permeation rates measured in our laboratory by this technology correlated well with those reported in the literature for 11 marketed drugs through Caco2 cell monolayers. Over a rate range in either system of 1−500 nm/s, the correlation coefficient was 0.90, or 0.99 on a log(rate) basis.

Human oral bioavailability data for 277 drugs were collected from Appendix II of Goodman & Gilman's *The Pharmacological Basis of Therapeutics*, 8th and 10th eds.[19] Structures were extracted from the World Drug Index (WDI). The entries were selected on the criteria that they contained experimental oral bioavailability data (not based on the administration of a prodrug), were amenable to our property calculators, and did not contain metals. These data and the calculated properties are tabulated in the Supporting Information.

Calculated Molecular Properties. The degree of structural diversity of the two sets of compounds may be judged from the results, shown in Table 1, of group average clustering by Tanimoto similarity of Daylight fingerprints.[20] The oral bioavailability data span a somewhat wider range of structural classes than do the permeation rate data. As will be noted from Tables 2 and 3, the global averages of the measures related to molecular size are smaller in the artificial membrane permeation compound set.

Table 2. Oral Bioavailability Quartile Property Averages for the Rat Oral Bioavailability Data Set and for Subsets Divided on the Basis of Molecular Weight[a]

quartile	%*F* range (rat)	MW	no. of rotatable bonds	*C* log *P*[b]	H-bond donor count	H-bond acceptor count	H-bond total count	polar surface area (Å^2)
			Full Data Set, MW range 220−770, $n = 1117$					
4	>42.7−100	*431.6*	*6.17*	4.45	*1.75*	*5.81*	*7.56*	*87.8*
3	>15.5−42.7	483.9	8.15	*4.78*	2.03	6.60	*8.63*	*94.0*
2	>4.3−15.5	492.3	*9.00*	4.38	2.40	7.12	9.52	103.6
1	<4.3	*511.1*	*10.22*	*3.49*	*2.99*	*8.06*	*11.05*	*123.3*
all	av 28.7	479.8	8.39	4.27	2.29	6.90	9.20	102.2
	r with %*F*	−0.35	−0.39	0.10	−0.32	−0.35	−0.39	−0.30
			MW >550, $n = 237$					
4	23−100	600.1	*10.34*	*6.21*	*1.59*	*7.44*	*9.03*	*97.0*
3	6.2−23	608.8	11.02	5.91	*1.64*	*7.73*	*9.37*	99.9
2	2.7−6.1	623.4	12.98	5.22	2.47	*9.46*	*11.93*	129.3
1	0−2.7	627.8	*13.60*	*4.13*	*2.83*	*10.00*	*12.83*	*141.2*
all	av 15.3	615.1	11.99	5.36	2.14	8.66	10.80	117.0
	r with %*F*	−0.20	−0.40	0.27	−0.27	−0.33	−0.34	−0.27
			MW 400 to ≤550, $n = 668$					
4	39.1−100	*462.4*	*6.50*	*4.76*	*1.86*	*6.20*	*8.06*	*90.5*
3	15.3−39.1	477.3	8.07	*4.72*	2.30	6.78	9.08	98.2
2	4.7−15.2	476.4	*8.69*	4.23	2.57	7.01	9.58	103.4
1	0−4.7	473.1	*9.08*	*3.05*	*3.15*	*7.47*	*10.62*	*119.1*
all	av 27.3	472.3	8.08	4.19	2.47	6.87	9.33	102.7
	r with %*F*	−0.17	−0.34	0.23	−0.33	−0.22	−0.30	−0.24
			MW < 400, $n = 212$					
4	93−100	349.3	5.57	*3.94*	*1.55*	*4.64*	*6.19*	*77.2*
3	41−92.7	353.9	4.75	3.43	1.60	4.92	6.53	80.1
2	11−40.6	347.7	4.98	3.12	1.92	4.77	6.70	81.6
1	1−10	358.5	6.09	*2.79*	*2.57*	*5.87*	*8.43*	*97.2*
all	av 48.3	352.3	5.35	3.32	1.91	5.05	6.96	84.0
	r with %*F*	−0.05	−0.05	0.28	−0.33	−0.21	−0.28	−0.20

[a] Italicized values for the quartiles differ from the overall average of the set or subset at the 1% significance level or better. [b] Overall range −0.62 to +9.39.

Table 3. Property Averages for Artifical Membrane Permeation Rate Ranges[a,b]

permeation rate range [log(nm/s)]	*n*	MW	no. of rotatable bonds	*C* log *P*	H-bond donor count	H-bond acceptor count	H-bond total count	polar surface area (Å^2)
>3	90	*397.0*	*5.72*	*4.53*	*1.34*	*3.92*	*5.27*	*57.7*
>2.5 to ≤3	1541	*453.0*	*7.28*	*4.80*	*1.94*	*5.48*	*7.42*	*82.1*
>2 to ≤2.5	570	*482.2*	7.82	*4.70*	2.03	*5.90*	7.93	93.4
>1 to ≤2	393	*481.9*	*8.49*	*4.13*	*2.17*	*6.18*	*8.35*	*103.0*
≤1	467	*476.6*	*9.02*	*3.45*	*2.37*	*6.35*	*8.72*	*113.9*
all		464.1	7.76	4.48	2.04	5.74	7.78	91.0
r with log permeation rate		−0.12	−0.20	0.26	−0.21	−0.20	−0.25	−0.40

[a] Italicized values for the ranges differ from the overall average of the set or subset at the 1% significance level or better. [b] Overall molecular weight range 146−771, *C* log *P* range −1.59 to +10.98.

Rotatable bonds were defined as any single bond, not in a ring, bound to a nonterminal heavy (i.e., non-hydrogen) atom. Excluded from the count were amide C−N bonds because of their high rotational energy barrier.

Octanol−water partition coefficients, *C* log *P*, were calculated using the BioByte ClogP 4.0 estimator, as implemented in the Daylight Chemical Information System software, v. 4.71.[20]

Hydrogen bond donors were taken as any heteroatom with at least one bonded hydrogen.

Hydrogen bond acceptors were taken as any heteroatom without a formal positive charge, excluding halogens, pyrrole nitrogen, heteroaromatic oxygen and sulfur, and higher oxidation states of nitrogen, phosphorus, and sulfur but including the oxygens bonded to them.

The polar surface area was calculated by the atom-based method of Ertl, Rohde, and Selzer.[21] This calculated polar surface area correlated closely with the total hydrogen bond count, the sum of hydrogen bond donors and acceptors. For the oral bioavailability data set $r = 0.93$, and for the permeation set $r = 0.87$.

Supporting Information that includes the distributions of the calculated properties and the bioavailability values for the 1117-compound rat oral bioavailability data set and for the 279-drug human set is available.

Results and Discussion

To examine the contribution of molecular rigidity to oral bioavailability, we chose the rotatable bond count as the simplest measure. In general, the rotatable bond count increases with molecular weight.[22,23] To separate molecular weight from the rotatable bond count, we divided our oral bioavailability data at molecular weight 500, a currently popular delimiter. The two groups were further divided according to the number of rotatable bonds, 7 or fewer, 8−10, and more than 10. The fraction of compounds with rat oral bioavailability (%*F*) of 20% or greater for the six groups is shown in Figure 1. Although the cutoff for acceptable oral bioavailability remains a matter of considerable discussion throughout the pharmaceutical industry, the value of 20% chosen for Figure 1 is intended to be at the low end of what might currently be acceptable for drug development, yet inclusive of potentially interesting leads. It is recognized that many successful drugs fall below this cutoff and that studies in rats are an imprecise guide to the results that might be found in humans. Given these caveats,

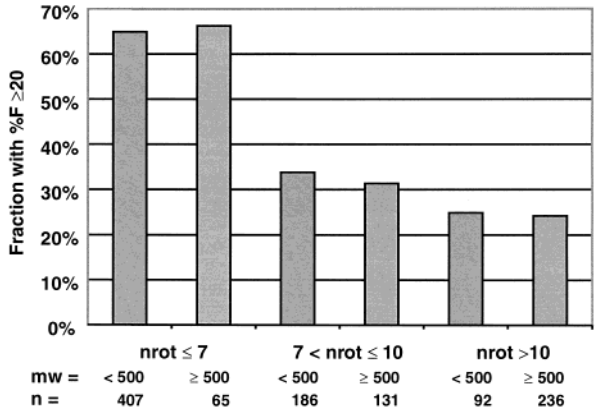

Figure 1. Fraction of compounds with a rat oral bioavailability of 20% or greater as a function of molecular weight and rotatable bond count (nrot). The effect of molecular rigidity on oral bioavailability is independent of molecular weight.

the apparent dependence of oral bioavailability on rotatable bonds, both above and below an MW of 500, is striking. About 65% of the compounds with seven or fewer rotatable bonds meet the %$F \geq 20$ criterion irrespective of the molecular weight range. In contrast, more than 75% of the compounds with more than 10 rotatable bonds, in either molecular weight group, show oral bioavailability of less than 20% and would most often be excluded for consideration as development candidates. The two intermediate groups showed intermediate and nearly identical distributions. These intermediate groups contain a significant fraction of compounds that would be deemed to have satisfactory oral bioavailability.

In a more detailed analysis, we divided the data into oral bioavailability quartiles and examined the quartile averages of the rotatable bond count and other simple calculable molecular properties recognized as relevant to oral bioavailability. We also separated the set into three subsets spanning approximately equal molecular weight ranges and calculated quartile averages for these. Values are given in Table 2. Correlation coefficients between %*F* and the calculated properties are also reported. Over the full MW 220−770 data set, the quartile averages confirm that higher oral bioavailability is indeed associated with lower molecular weight, but demonstrate that higher oral bioavailability is also associated with lower rotatable bond counts, lower hydrogen bond counts, and lower polar surface area. The averages of the calculated octanol−water partition coefficient, *C* log *P*, on the other hand, appear only to indicate that a certain minimum lipophilicity is required for oral bioavailability.[24]

When the narrower molecular weight ranges are considered, the correlation of the oral bioavailability quartile average with the molecular weight average is, of course, reduced. A clear relationship of %*F* with the rotatable bond count is apparent for the MW 400−550 and MW greater than 550 subsets, but for the MW < 400 set there is no correlation. In that MW < 400 set, 90% of the compounds have 10 or fewer rotatable bonds, and 70% have seven or fewer, suggesting that there is a molecular weight threshold below which flexibility is sufficiently limited to have little impact on %*F*. The negative correlations of the average hydrogen bond count and average polar surface area with oral bioavailability are retained in all three of the molecular weight subdivisions.

Increasing average oral bioavailability is paralleled in each subset with increasing hydrophobicity. In fact, in the high molecular weight range high oral bioavailability is associated with an average *C* log *P* in excess of 5, the cutoff suggested by Lipinski. Solubility, which might generally be reduced in more highly hydrophobic compounds, is less an issue in the rat studies reported here, because compounds are normally administered to the rats using formulation aids such as the cyclodextrin derivative Encapsin. Low solubility can be a more limiting factor in drug development, however, and may play a role in the contribution of *C* log *P* to Lipinski's analysis of compounds destined for drug development or used as drugs, as he has recognized.

The relation of *C* log *P* to oral bioavailability is complicated by the failure of the *C* log *P* calculation to consider the charge state of ionizable groups, which surely plays a role in membrane permeation rates. Because of the added uncertainty of the prediction of pK_a, we followed common practice in not calculating the partition coefficient at a fixed pH.

Overall, these results suggest that the success of molecular weight as a predictor of oral bioavailability is at least partially a result of the correlation of higher molecular weight with increased molecular flexibility and with higher polar surface area or hydrogen bond count.

The lack of influence of molecular weight on oral bioavailability when it is separated from a measure of molecular flexibility, the number of rotatable bonds, which is clearly seen in Figure 1, is at odds with much current thinking, and the impact of increasing molecular rigidity on oral bioavailability is not generally recognized. Correlation of increasing rotatable bond count with increasing molecular weight is intuitively obvious and is supported by analysis of a number of compound databases.[22,23] Thus, the perception that molecular weight is a determinant of oral bioavailability probably arises from the dependence of oral bioavailability on a molecular weight correlated property. An apparent molecular weight cutoff could be perceived because the number of rotatable bonds and corresponding molecular flexibility are reduced below molecular weight 500, and especially below molecular weight 400, to the point that oral availability can reach high levels even in groups of randomly chosen structures. For higher molecular weight compounds, the process of finding constraints that are compatible with target binding can be complex and often time-consuming, both in the molecular design and in the chemical synthesis. Thus, we see a relatively low population of compounds over molecular weight 500 with a low number of rotatable bonds in our data, 45% with 10 or fewer, and 15% with seven or fewer.

Molecular weight is a surrogate for other properties in addition to molecular rigidity. The polar surface area and hydrogen bond count also tend to increase with increasing molecular weight, though not as strongly in our data as do rotatable bonds.[25] That information relevant to drug absorption is sufficiently encoded in lipophilicity plus polar surface area, without explicit reference to molecular weight, has already been suggested by Egan et al.[26] Below, we suggest guidelines

derived from our data for rat oral bioavailability based on the rotatable bond count and polar surface area (or hydrogen bond count) alone.

Possible Sources of Correlations with Oral Bioavailability. Many factors control the entry of drugs from the intestine to the systemic blood. In vitro models of at least two of the processes involved are available in our laboratories. The passive membrane permeation rate of individual compounds is modeled by passage through a synthetic lipid bilayer.[7,16] This model mimics the cell membrane lacking the specific transporters that can facilitate entry or catalyze export into or out of a cell. Numerous enzymatic processes involving both metabolism and transporter-mediated expulsion (apical recycling) play a role in limiting the oral bioavailability of drugs. These enzymes are present in the cells of both liver and intestinal mucosa. Renal processes can also play a significant role in the clearance of drugs. Clearance as measured in the iv leg of our oral bioavailability studies is expected to offer some overall measure of these processes. An in vitro model using rat liver microsomes[27] measures an approximation of the rate of metabolism by many of the enzymes present in these tissues. It is thus a measure of some of the enzyme-mediated processes responsible for the first pass liver clearance by metabolism that can limit oral bioavailability. In vitro rates of metabolism were available from this assay for 4300 compounds. A detailed analysis of much of the same data has been reported elsewhere.[28] We have additionally searched these data for correlations with molecular properties.

Artificial Membrane Permeation Rate. Artificial membrane permeation rates available for 3061 compounds ranged from 0.1 to 2000 nm/s. The compounds were divided into five rate ranges on a logarithmic basis, and average values of molecular properties were calculated for each range. The correlation coefficients for permeation and the calculated properties over the full set were also computed. The averages, shown in Table 3, reveal that, in this data set as in other permeation data that have been reported, increasing permeation rate is paralleled by consistently decreasing polar surface area and hydrogen bond count. Increasing permeation rate is also paralleled by decreasing rotatable bond count, however, and it should be noted that again this is not solely a matter of size. For the three least permeable compound sets, spanning a rate range of at least 300-fold, the trend in rotatable bonds persists although the average molecular weight does not vary significantly. Further observations are that the compounds with the lowest rates are significantly less lipophilic on the average, and those few with the very high rates are of significantly lower average molecular weight.

The apparent dependence of the membrane permeation rate on the polar surface area or hydrogen bond count is consistent with the literature,[6,29,30] including the conclusions of Burton et al. for a series of peptide derivatives across Caco-2 cell monolayers.[31] In that work, it was found that the passive transmembrane permeation rate better correlated with measures of the desolvation necessary on entering the lipid region of the membrane, as measured by heptane−ethylene glycol partition (or an effective partition between water-saturated isooctane and water-saturated octanol), than with lipophilicity measured by partition between water and water-saturated octanol. Hydrogen-bonding potential also appears, inter alia, as an important factor in a multiple regression analysis study of intestinal absorption of 42 nonpeptide drugs.[32] Since most of the major contributors to the polar surface area are hydrogen bond donors or acceptors, it is not surprising that the polar surface area in our membrane permeation compound set correlates well with the sum of hydrogen bond donors and acceptors ($r = 0.87$). Both the polar surface area and hydrogen bond count correlate less well with *C* log *P* (−0.43 and −0.30, respectively.). In 21 peptide derivatives studied by Burton et al., although the properties were computed somewhat differently, the total hydrogen bond count also correlated well ($r = 0.84$) with the polar surface area while the measured log *P* correlated poorly with both ($r = 0.25$ and -0.02, respectively).[33]

Secondary amides, as in the peptide bond, contribute significantly to the polar surface area. The dependence of permeation on the polar surface area is thus consistent with the arguments of Hirschmann[4] regarding the limiting role of amide hydration on membrane permeation.

It has been suggested that while desolvation of polar groups is required for permeation, lipophilicity is necessary to drive solutes as far as the interfacial region of a membrane.[34,35] Our observations that the average hydrophobicity (*C* log *P*) is constant in the higher three permeation rate groups but shifts significantly lower (one *C* log *P* unit) for the lowest permeation rate quartile while the polar surface area decreases with increasing permeation rate are consistent with a change in the rate-limiting step in permeation through the lipid bilayer.

Navia and Chaturvedi[3] proposed that flexibility-allowed changes in surface properties from aqueous-compatible to lipid-compatible are important for a good permeation rate. The data here do not support this reasonable hypothesis, which may be valid in the specific classes of compounds to which they refer. Rather, the correlation we observe between the average membrane permeation rate and average rotational bond count (which is clearly molecular weight independent in the permeation rate ranges below 300 nm/s) may reflect a possible entropic cost of changes in conformation required to present an appropriate exterior to the hydrocarbon interior of the membrane.

Correlation of Observed Oral Bioavailability with Observed Artificial Membrane Permeation. For the 341 compounds common to the oral bioavailability and permeation rate data sets the correlation coefficient between oral bioavailability and log permeation rate is 0.33 (0.31 vs permeation rate). A scatter plot of %*F* vs log permeation rate is shown as Figure 2. Roughly speaking, it would appear that, above a certain threshold, perhaps 100 nm/s, the permeation rate is not the limiting factor for oral bioavailability. The average oral bioavailability for those compounds with permeation rates below 100 nm/s is 13%, and for those above 100 nm/s it is 35%. This is also consistent with the relationship between the Caco-2 monolayer permeation

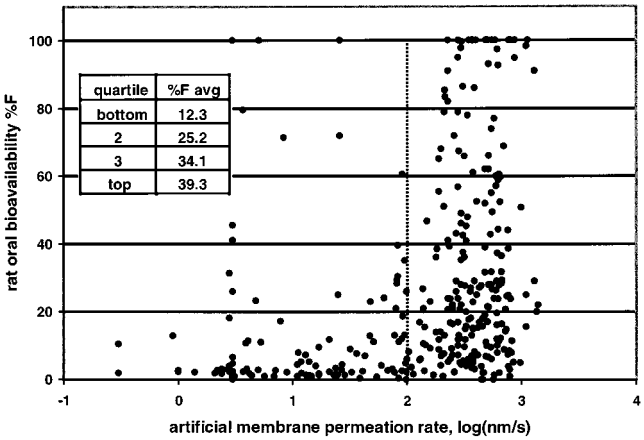

Figure 2. Scatter plot of rat oral bioavailability vs artificial membrane permeation rate. The inset gives the average oral bioavailability values for log (permeability) quartiles.

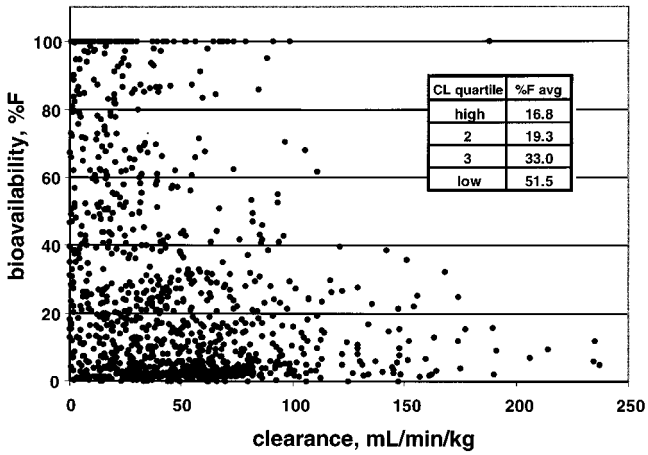

Figure 3. Scatter plot of rat oral bioavailability vs rat in vivo clearance rate. The data are those of Table 4.

rate and fraction absorbed after oral administration to humans for 21 drugs reported by Stenberg et al.[29]

Clearance as a Factor in Determining Oral Bioavailability. Clearance rates [(mL/min)/kg] were available for virtually all (1090 of 1117) compounds for which an oral bioavailability number had been reported. The compounds were grouped into quartiles by clearance rate, and the averages for the properties examined are given in Table 4, together with the overall correlation coefficients of in vivo clearance with the property values. According to these results, in vivo clearance does not correlate with molecular weight, rotatable bond count, lipophilicity, hydrogen bond count, or polar surface area.

A similar analysis of the rat liver microsomal clearance data,[28] not reported in detail here, suggests only questionable correlations between lipophilicity (positive) or polar surface area (negative) and microsomal clearance. We are therefore unable to associate the effects of the rotatable bond count and polar surface area on oral bioavailability with this in vitro measure. This probably reflects the fact that liver microsomes contribute to only a portion of the limits on oral bioavailability.

The correlation coefficient of oral bioavailability with in vivo clearance is −0.28. A scatter plot of %*F* vs in vivo clearance is given in Figure 3. Although the average oral bioavailability depends on the average clearance rate, the correlation is about the same as that with the artificial membrane permeation rate. Thus, both of these factors play a contributing role in oral bioavailability but even taken together do not fully account for the molecular properties that do influence oral bioavailability.

Evaluation of Potential New Rules for Prediction of Oral Bioavailability. It is often necessary to work with compounds of high molecular weight to achieve desired drug features such as high potency and target selectivity. We have therefore tried to develop rules related to properties independent of molecular weight that can be altered through structure modification to optimize oral bioavailability. Following from the observations on correlation of molecular properties with oral bioavailability, we explored several rules based on rotatable bonds and either the hydrogen bond donor/acceptor sum or polar surface area. The success of each new rule as applied to our oral availability database has been evaluated by plotting the cumulative fraction meeting the chosen rule as a function of increasing oral bioavailability. The most successful rules should have a high fraction of the compounds with an oral bioavailability of >20−40% while at the same time having a low fraction with 0−20% oral bioavailability.

If a rat oral bioavailability of >20−40% is taken as a level of acceptability, it appears that with our data either (i) polar surface area (PSA) ≤ 140 Å^2 and number of rotatable bonds ≤10 or (ii) sum of H-bond donors and acceptors ≤ 12 and number of rotatable bonds ≤10 is an efficient and selective criterion. This rule gives selection over the range 20−100% oral bioavailability similar to selection based on meeting any 3 of the 4 criteria of Lipinski's "rule of 5", but supplies somewhat better selection against compounds that have low oral bioavailability. The bias of our data set toward compounds that meet the rule of 5 is apparent in the fact that nearly 80% of its compounds meet 3 out of 4 of the Lipinski criteria.

These results would suggest that candidate design directed at reduced flexibility and polar surface area without specific reference to molecular weight limits would increase success in achieving high oral bioavailability. The upper limit of molecular weight is actually not predictable from our studies, nor are we aware of

Table 4. Property Averages for Rat Whole Body Clearance Data and for Quartiles Based on Clearance Rate (n = 1090)

quartile	CL [(mL/min)/kg]	%*F*	MW	no. of rotatable bonds	*C* log *P*	H-bond donor count	H-bond acceptor count	H-bond total count	polar surface area (Å^2)
4	103.2	16.8	478.6	8.59	4.22	2.12	6.85	8.97	98.9
3	51.4	19.3	491.5	9.07	3.99	2.37	7.32	9.69	107.4
2	28.1	33.0	489.4	8.37	4.26	2.26	6.92	9.18	100.7
1	8.4	51.5	461.2	7.63	4.58	2.43	6.63	9.06	102.8
total	47.8	30.3	480.2	8.42	4.26	2.29	6.93	9.22	102.5
r with in vivo clearance		−0.28	−0.02	0.05	−0.06	−0.09	−0.04	−0.06	−0.07

any studies that address this question uniquely. Examples of higher molecular weight compounds with unexpectedly high oral bioavailability include phalloidin (mushroom toxin), antamanide (antitoxin), cyclosporine (immunosuppressant), and vinorelbine (antitumor agent). Borchardt has demonstrated an effect of peptide cyclization in improving the transport of cyclic peptides compared to linear peptides in Caco-2 cells.[36] He has also demonstrated the importance of transporter-mediated expulsion in limiting the permeability of cyclic peptides.[37] Although the orally available molecules listed above are cyclic, nonbonded intramolecular interactions alone in larger compounds may serve to reduce molecular flexibility and thereby give unexpectedly high oral bioavailability. The simple rotatable bond count would be irrelevant in such cases.

We applied our proposed rule to a set of nonproprietary data on human oral bioavailability of drugs, taken from the compilations in Goodman & Gilman's *The Pharmacological Basis of Therapeutics* (see Methods), although the property distributions of this set, consisting of relatively low molecular weight compounds of relatively high oral bioavailability, are very different from those of the experimental compounds for which we had rat data. The average molecular weight is 336 vs 480, the average bioavailability is 59 vs 29%, the average number of rotatable bonds is 4.7 vs 8.4, and the average *C* log *P* is 1.8 vs 4.3 (see the Supporting Information). Only 20 of the 277 molecules have molecular weight above 500, so that it is not possible to demonstrate an effect of rotatable bond count at high molecular weight, and only 40 have human oral bioavailability below 20%. Not surprisingly, since it was derived from a larger subset of the World Drug Index,[1] application of the rule of 5 to this drug data set selects 93% of the compounds with an oral bioavailability of 20% or greater. It selects 84% of them in our rat data set; see Figure 4. Application of the rotatable bond, polar surface area rule selects 86% of the drugs with an oral bioavailability of 20% or greater, and 80% for the rat data set.

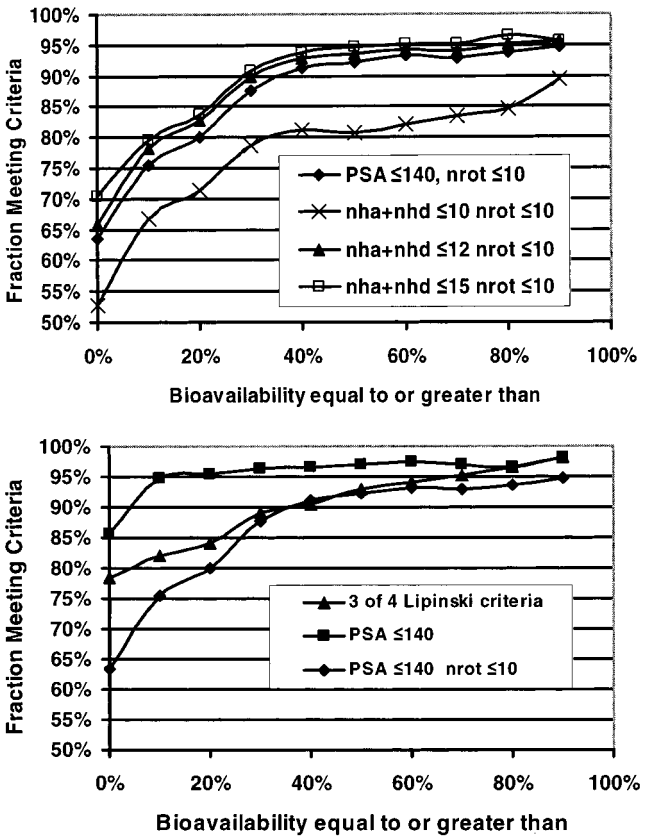

Figure 4. Effects of various criteria for selecting orally bioavailable compounds. Of compounds with an oral bioavailability of ≥20%, ~80% had PSA ≤ 140 Å^2 and (nrot) ≤ 10, while ~85% satisfy Lipinski's rule of 5. Thus, while the rule of 5 and PSA/nrot rule give similar results for compounds with %*F* > 20, the PSA/nrot guideline passes fewer compounds with %*F* < 20. (The line for a hydrogen bond count of ≤12 alone lies within 1% of the line for PSA ≤ 140 alone at all points.)

Conclusions

An analysis of the measured oral availability in rats for over 1100 compounds studied at SmithKline Beecham has revealed the unexpected positive influence of increasing molecular rigidity as measured by the rotatable bond count and the more expected negative impact of increasing polar surface area. A portion of these effects are discernible in the underlying process of passive membrane transport as measured using a synthetic membrane model. A dominant portion, which must relate to the processes involved in metabolic and transporter-mediated first pass clearance (intestinal membrane, portal vein, and liver), is not modeled by total body clearance or by in vitro metabolism by rat liver microsomes.

The key role that structure-based differentiation can play in oral bioavailability highlights the difficulties and paradoxes of finding potent and selective drugs with high oral bioavailability. One must freeze in an overall molecular shape and functional group presentation compatible with optimal target interaction. At the same time one must freeze out the shapes and functional group presentations responsible for substrate or binding interactions with proteins responsible for clearance. Both of these must be accomplished while properties compatible with membrane permeation are retained. Failure to avoid even one of the many factors can lead to poor oral bioavailability. This may reflect an evolutionary selection based on exposure to otherwise toxic chemicals. Thus, high molecular rigidity, although important, cannot guarantee high oral bioavailability. Nonetheless, our data make it clear that any medicinal chemistry effort needs to place a high priority on the introduction of conformational constraints when pharmacokinetics is an issue to be resolved. Furthermore, we know that molecular rigidity is a much more complex issue than the simple counting of rotatable bonds. Ring systems are not rigid as assumed by our count. The conformational consequences of ring substitution or even branching in acyclic systems is very difficult to quantitate or even predict qualitatively by the present theoretical methods. The problem is even further confounded by species differences in the nature and specificity of the first pass clearance, making the projection from animal models to humans a risky exercise at best. Until all of the enzymes involved are fully defined in both the model species and humans, success in medicinal chemistry will continue to be a matter of optimizing probabilities. It will also continue to require the in vitro and in vivo evaluation of oral bioavailability in multiple

animal species and man to speculate on the in vivo human result.

Despite the complexities of the underlying processes of oral bioavailability, the new methods for more rapidly measuring pharmacokinetic parameters have produced enough rat oral bioavailability data to allow for the development of proposed new prediction rules. Importantly, these rules suggest a path to the discovery of orally available drugs of higher molecular weight, a class of compounds only poorly exploited until now. We also see the likelihood of important improvement in these rules when theoretical methods can be developed that more adequately define molecular flexibility. We hope that analyses similar to ours can now be reported on other proprietary databases that may have been built with differing inherent bias and in species other than the rat.

Acknowledgment. We thank Dr. Robert W. Marquis, Dr. Dennis S. Yamashita, Dr. Stephen E. Clarke, and Dr. Peter D. Gorycki for valuable discussions and helpful editorial comment.

Supporting Information Available: Table giving the human oral bioavailability data from Goodman and Gilman and calculated molecular properties and figures showing the property distribution data for the 277-compound Goodman and Gilman human oral bioavailability data set and the 1117-compound propietary rat oral bioavailability data set. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estmate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **1997**, *23*, 4−25.

(2) Lipinski, C. A. Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 235−249.

(3) Navia, M. A.; Chaturvedi, P. R. Design Principles for Orally Bioavailable Drugs. *Drug Discovery Today* **1996**, *1*, 179−189.

(4) Hirschmann, R. Peptide research a means to further biological and chemical understanding. In *Peptides 1996, Proceedings of the 24th European Peptide Symposium*; Ramage, R., Epton, R., Eds.; Mayflower Scientific Ltd.: Kingswinford, UK, 1998; pp 3−17. Smith, A. B., III; Hirschmann, R.; Pasternak, A.; Yao, W.; Sprengler, P. A.; Halloway, M. K.; Kuo, L. C.; Chen, Z.; Darke, P. L.; Schleif, W. A. *J. Med. Chem.* **1997**, *40*, 2440−2444.

(5) 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.

(6) 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.

(7) Kansy, M.; Senner, F.; Gubernator, K. Physicochemical High Throughput Screening−Parallel Artificial Membrane Permeation Assay in the Description of Passive Absorption Processes. *J. Med. Chem.* **1998**, *41*, 1007−1010.

(8) Wohnsland, F.; Faller, B. High-throughput permeability pH profile and high-throughput alkane/water log P with artificial membranes. *J. Med. Chem.* **2001**, *44*, 923−930.

(9) Lin, J. H.; Chiba, M.; Baillie, T. A. Is the role of the small intestine in first-pass metabolism overemphasized? *Pharmacol. Rev.* **1999**, *51*, 135−157.

(10) Matheny, C. J.; Lamb, M. W.; Brouwer, K. L. R.; Pollack, G. M. Pharmacokinetic and pharmacodynamic implications of P-glycoprotein modulation. *Pharmacotherapy* **2001**, *21*, 778−796.

(11) Chaturvedi, P. R.; Decker, C. J.; Odinecs, A. Prediction of pharmacokinetic properties using experimental approaches during early drug discovery. *Curr. Opin. Chem. Biol.* **2001**, *5*, 452−463.

(12) Yoshida, F.; Topliss, J. G. QSAR model for drug human oral bioavailability. *J. Med. Chem.* **2000**, *43*, 2575−2585.

(13) Veber, D. F.; Marquis, R. W.; Yamashita, D. S.; Ru, Y.; Oh, H.-J.; Ward, K.; Smith, B. R. The Role of Conformational Constraint in Improved Oral Bioavailability of Cathepsin K Inhibitors. In *Peptides 2000: Proceedings of the Twenty-Sixth European Peptide Symposium*; Martinez, J., Fehrentz, J.-A., Eds.; Editions EDK: Paris, 2000; pp 113−114.

(14) Marquis, R. W.; Ru, Y.; LoCastro, S. M.; Zeng, J.; Yamashita, D. S.; Oh, H.-J.; Erhard, K. F.; Davis, L. D.; Tomaszek, T. A.; Tew, D.; Salyers, K.; Proksch, J.; Ward, K.; Smith, B.; Levy, M.; Cummings, M. D.; Haltiwanger, R. C.; Trescher, G.; Wang, B.; Hemling, M. E.; Quinn, C. J.; Cheng, H.-Y.; Lin, F.; Smith, W. W.; Janson, C. A.; Zhao, B.; McQueney, M. S.; D'Alessio, K.; Lee, C.-P.; Marzulli, A.; Dodds, R. A.; Blake, S.; Hwang, S.-M.; James, I. E.; Gress, C. J.; Bradley, B. R.; Lark, M. W.; Gowen, M.; Veber, D. F. Azepanone-Based Inhibitors of Human and Rat Cathepsin K. *J. Med. Chem.* **2001**, *44*, 1380−1395.

(15) Gao, J.; Sudoh, M.; Aube, J.; Borchardt, R. T. Transport characteristics of peptides and peptidomimetics: I. N-methylated peptides as substrates for the oligopeptide transporter and P-glycoprotein in the intestinal mucosa. *J. Pept. Res.* **2001**, *57*, 316−329.

(16) Cheng, H.-Y. Unpublished results.

(17) Thompson, M.; Lennox, R. B.; McClelland, R. A. Structure and Electrochemical Properties of Microfiltration Filter-Lipid Membrane Systems. *Anal. Chem.* **1982**, *54*, 76−81.

(18) Sugano, K.; Hamada, H.; Machida, M.; Ushio, H. High throughput prediction of oral absorption: Improvement of the composition of the lipid solution used in parallel artificial membrane permeation assay. *J. Biomol. Screening* **2001**, *6*, 189−196.

(19) (a) Hardman, J. G.; Limbird, L. E.; Gilman, A. G. Goodman & Gilman's *The Pharmacological Basis of Therapeutics*, 10th ed.; Medical Publishing Division, McGraw-Hill: New York, 2001. (b) Gilman, A. G.; Rall, T. W.; Niews, A. S.; Taylor, P. Goodman & Gilman's *The Pharmacological Basis of Therapeutics*, 8th ed.; Pergamon Press: New York, 1990.

(20) Daylight Chemical Information Systems, Mission Vieho, CA.

(21) Ertl, P.; Rohde, B.; Selzer, P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J. Med. Chem.* **2000**, *43*, 3714−3717.

(22) Ajay; Walters, W. P.; Murcko, M. A. Can we learn to distinguish between "drug-like" and "nondrug-like" molecules? *J. Med. Chem.* **1998**, *41*, 3314−3324.

(23) The correlation coefficient for rotatable bonds and molecular weight for both the bioavailability and permeation rate data sets is 0.68. For the former SmithKline Beecham historical database it is 0.72, and for 2851 drugs with U.S. Adopted Names (USAN) in the World Drug Index it is 0.61.

(24) *C* log *P* is not expected to parallel the polar surface area or hydrogen bond count. In this data set it correlates better with the ratio of polar surface area to molecular weight ($r = -0.69$) than with polar surface area alone ($r = -0.46$).

(25) In the oral bioavailability data set the correlation coefficients of the polar surface area and hydrogen bond count with molecular weight are 0.33 and 0.45, respectively. In the permeation rate data set the values are 0.25 and 0.39.

(26) Egan, W. J.; Merz, K. M.; Baldwin, J. J. Prediction of drug absorption using multivariate statistics. *J. Med. Chem.* **2000**, *43*, 3867−3877.

(27) Clarke, S. E.; Baldwin, S. J.; Bloomer, J. C.; Ayrton, A. D.; Sozio, R. S.; Chenery, R. J. Lauric acid as a model substrate for the simultaneous determination of cytochrome P450 2E1 and 4A in hepatic microsomes. *Chem. Res. Toxicol.* **1994**, *7*, 836−842.

(28) Clarke, S. E.; Jeffrey, P. Utility of metabolic stability screening: comparison of in vitro and in vivo clearance. *Xenobiotica* **2001**, *31*, 591−598.

(29) Stenberg, P.; Norinder, U.; Luthman, K.; Artursson, P. Experimental and computational screening models for the prediction of intestinal drug absorption. *J. Med. Chem.* **2001**, *44*, 1927−1937.

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

(31) Burton, P. S.; Conradi, R. A.; Ho, N. F. H.; Hilgers, A. R.; Borchardt, R. T. How Structural Features Influence the Biomembrane Permeability of Peptides. *J. Pharm. Sci.* **1996**, *85*, 1336−1340.

(32) Sugawara, M.; Takekuma, Y.; Yamada, H.; Kobayashi, M.; Iseki, K., Miyazaki, K. A General Approach For the Prediction of the Intestinal Absorption of Drugs−Regression Analysis Using the Physicochemical Properties and Drug-Membrane Electrostatic Interaction. *J. Pharm. Sci.* **1998**, *87*, 960−966.

(33) Goodwin, J. T.; Mao, B.; Vidmar, T. J.; Conradi, R. A.; Burton, P. S. Strategies toward predicting peptide cellular permeability from computed molecular descriptors. *J. Pept. Res.* **1999**, *53*, 355−369.

(34) Stenberg, P.; Luthman, K.; Artursson, P. Prediction of membrane permeability to peptides from calculated dynamic molecular surface properties. *Pharm. Res.* **1999**, *16*, 205−212.
(35) Stenberg, P.; Luthman, K.; Artursson, P. Virtual screening of intestinal drug permeability. *J. Controlled Release* **2000**, *65*, 231−243.
(36) Okumu, F. W.; Pauletti, G. M.; Vandervelde, D. G.; Siahaan, T. J.; Borchardt, R. T. Effect of Restricted Conformational Flexibility On the Permeation of Model Hexapeptides Across Caco-2 Cell Monolayers. *Pharm. Res.* **1997**, *14*, 169−175.
(37) Sudoh, M.; Pauletti, G. M.; Akimoto, K.; Yau, W.; Sprengler, P.; Smith, A. B., III; Hirschmann, R.; Borchardt, R. *Pharm. Res.* **1997**, S-26 (Abstr. 1081), 14.

JM020017N