J|A|C|S
ARTICLES
Published on Web 10/06/2006

Conformational Flexibility, Internal Hydrogen Bonding, and Passive Membrane Permeability: Successful in Silico Prediction of the Relative Permeabilities of Cyclic Peptides

Taha Rezai,[†] Jonathan E. Bock,[†] Mai V. Zhou,[†] Chakrapani Kalyanaraman,[‡] R. Scott Lokey,*[,†] and Matthew P. Jacobson*[,‡]

Contribution from the Department of Chemistry and Biochemistry, University of California at Santa Cruz, Santa Cruz, California 95064, and Department of Pharmaceutical Chemistry, University of California at San Francisco, 600 16th Street, San Francisco, California 94143-2240

Received May 2, 2006; E-mail: matt.jacobson@ucsf.edu

Abstract: We report an atomistic physical model for the passive membrane permeability of cyclic peptides. The computational modeling was performed in advance of the experiments and did not involve the use of "training data". The model explicitly treats the conformational flexibility of the peptides by extensive conformational sampling in low (membrane) and high (water) dielectric environments. The passive membrane permeabilities of 11 cyclic peptides were obtained experimentally using a parallel artificial membrane permeability assay (PAMPA) and showed a linear correlation with the computational results with $R^2 = 0.96$. In general, the results support the hypothesis, already well established in the literature, that the ability to form internal hydrogen bonds is critical for passive membrane permeability and can be the distinguishing factor among closely related compounds, such as those studied here. However, we have found that the number of internal hydrogen bonds that can form in the membrane and the solvent-exposed polar surface area correlate more poorly with PAMPA permeability than our model, which quantitatively estimates the solvation free energy losses upon moving from high-dielectric water to the low-dielectric interior of a membrane.

Introduction

Many previous studies have explored correlations between various properties of small molecules and their membrane permeability, a critical property underlying the bioavailability of drugs and one determinant of success or failure in preclinical development.[1−10] Predictive models of membrane permeability have been developed,[2−9] usually by the application of multivariable regression to a "training set", followed by evaluation of the model on a separate "test set". The molecular descriptors most commonly employed in these studies are compound size and the solvent-exposed polar surface area. It is physically reasonable to expect both of these properties to affect membrane permeability. Other approaches have used quantum mechanical calculations to provide descriptors.[11] Beyond the reasonability of the descriptors used, however, "knowledge-based" methods of this type do not attempt to predict membrane permeability on the basis of an atomistic physical model. The reliability of existing models of membrane permeability is a subject of debate, and failures in practical application have been reported[12] that have been attributed, in part, to inadequate training sets.

We and others have used cyclic peptides as model systems for studying membrane permeability.[13−16] The advantages of cyclic peptides for this purpose include their relative ease of synthesis and the ability to precisely modulate chemical functionality and stereochemistry. The ability to rationally engineer cyclic peptides with druglike passive membrane permeability would also open new possibilities for using cyclic

[†] University of California at Santa Cruz.
[‡] University of California at San Francisco.

(1) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Delivery Rev.* **2001**, *46*, 3−26.
(2) Ekins, S.; Rose, J. *J. Mol. Graphics Modell.* **2002**, *20*, 305−309.
(3) Refsgaard, H. H.; Jensen, B. F.; Brockhoff, P. B.; Padkjaer, S. B.; Guldbrandt, M.; Christensen, M. S. *J. Med. Chem.* **2005**, *48*, 805−811.
(4) Krarup, L. H.; Christensen, I. T.; Hovgaard, L.; Frokjaer, S. *Pharm. Res.* **1998**, *15*, 972−978.
(5) Ekins, S.; Waller, C. L.; Swaan, P. W.; Cruciani, G.; Wrighton, S. A.; Wikel, J. H. *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 251−272.
(6) Ekins, S.; Durst, G. L.; Stratford, R. E.; Thorner, D. A.; Lewis, R.; Loncharich, R. J.; Wikel, J. H. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1578−1586.
(7) Yamashita, F.; Wanchana, S.; Hashida, M. *J. Pharm. Sci.* **2002**, *91*, 2230−2239.
(8) Malkia, A.; Murtomaki, L.; Urtti, A.; Kontturi, K. *Eur. J. Pharm. Sci.* **2004**, *23*, 13−47.
(9) Fujikawa, M.; Ano, R.; Nakao, K.; Shimizu, R.; Akamatsu, M. *Bioorg. Med. Chem.* **2005**, *13*, 4721−4732.
(10) Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. *J. Med. Chem.* **2002**, *45*, 2615−2623.
(11) Fujiwara, S.; Yamashita, F.; Hashida, M. *Int. J. Pharm.* **2002**, *237*, 95−105.
(12) Stouch, T. R.; Kenyon, J. R.; Johnson, S. R.; Chen, X. Q.; Doweyko, A.; Li, Y. *J. Comput.-Aided Mol. Des.* **2003**, *17*, 83−92.
(13) Rezai, T.; Yu, B.; Millhauser, G. L.; Jacobson, M. P.; Lokey, R. S. *J. Am. Chem. Soc.* **2006**, *128*, 2510−2511.
(14) Gangwar, S.; Jois, S. D.; Siahaan, T. J.; Vander Velde, D. G.; Stella, V. J.; Borchardt, R. T. *Pharm. Res.* **1996**, *13*, 1657−1662.
(15) Gudmundsson, O. S.; Vander Velde, D. G.; Jois, S. D.; Bak, A.; Siahaan, T. J.; Borchardt, R. T. *J. Pept. Res.* **1999**, *53*, 403−413.
(16) Gudmundsson, O. S.; Jois, S. D.; Vander Velde, D. G.; Siahaan, T. J.; Wang, B.; Borchardt, R. T. *J. Pept. Res.* **1999**, *53*, 383−392.

peptides to inhibit intracellular proteins. Most cyclic peptides have poor membrane permeability, but there are exceptions, including cyclosporine A, a cyclic undecapeptide that is used as an orally active immunosuppressive drug.[17] In addition, in a previous paper we identified a cyclic hexapeptide, cyclo(D-Leu-D-Leu-Leu-D-Leu-Pro-Tyr), with passive membrane permeability that slightly exceeds that of cyclosporine A.[13]

We report here a physics-based model for the passive membrane permeability, as measured in a parallel artificial membrane permeability assay (PAMPA), of cyclic peptides that shows good predictive ability. The computational modeling was performed in advance of the experiments and did not involve the use of "training data". The model involves extensive conformational sampling of the peptides, represented in all-atom detail, in low (membrane) and high (water) dielectric environments. The conformational sampling is needed to understand the role of conformational flexibility, which is not explicitly considered in many models of membrane permeability. Water and the membrane are not treated in atomistic detail. Rather, they are both treated as a dielectric continuum. Thus, our model neglects atomic-level details of the membrane and water, including the membrane−water interface. Implicit solvent models have also been used in prior studies of membrane permeability.[18,19] Our model is thus intermediate in both computational speed and level of physical detail between molecular dynamics simulations of membrane permeability using explicit water and membranes[20−24] and knowledge-based methods based on simple chemical descriptors.[2−9]

The passive membrane permeabilities of 11 cyclic peptides were obtained experimentally using a PAMPA and showed a linear correlation with the computational predictions with $R^2 = 0.96$. The success of the model suggests that, although it does not treat all aspects of the physics of membrane permeability, it captures two critical elements: the ability of the peptides to adopt multiple conformations with different populations in low- and high-dielectric environments and the free energy cost of desolvating the peptides upon entering the membrane. The physical underpinnings of our model are considered in some detail below, including possible improvements to the model. In general, the results support the hypothesis, already well-established in the literature,[8,14−16,18,25] that the ability to form internal hydrogen bonds is critical for passive membrane permeability and can be the distinguishing factor among closely related compounds, such as those studied here. However, we have found that the number of internal hydrogen bonds that can form in the membrane and the solvent-exposed polar surface area each correlate somewhat weakly with membrane permeability. The implicit solvent model used in our approach reflects the internal hydrogen bonding and polar surface area but provides a quantitative free energy difference that can be thought of as a measure of the overall lipophilicity/hydrophilicity of a particular conformation of the compound (i.e., the one that we postulate is adopted in the low-dielectric membrane).

Finally, we also discuss limitations of our model, the most important of which is that it currently can only be used to predict relative PAMPA permeabilities in a series of compounds and not absolute permeabilities. One important reason for this is that the model does not account for translational, rotational, and internal entropy losses of compounds upon entering the membrane.

Results and Discussion

Overview of the Computational Model. Figure 1 schematically depicts the conceptual underpinnings of our model. Note that, although we depict a bilayer in this schematic, the artificial membrane in the PAMPA is considerably thicker (approximately 30 Å) and the interior contains a hydrocarbon, dodecane in this work. It is also important to emphasize that this work does not consider the effects of changing the composition of the membrane on permeability, but only the effects of changing the chemical structure of the permeant, using an artificial membrane of constant composition.

We assume that the peptides adopt multiple, rapidly interconverting conformations in water, but that membrane permeation is dominated by a single conformation. (Conformational entropy losses upon entering the membrane have been observed in molecular dynamics simulations.[21,22]) In practice, we identify this conformation as the lowest energy state found by extensive sampling of the peptide in a low-dielectric solvent (see the Methods). In the following, we will refer to this conformation as the "low-dielectric conformation", LDC. The LDC generally has maximal internal hydrogen bonding, but we identify the LDC by the internal energy, which includes intramolecular strain, electrostatics, and dielectric screening treated by an implicit solvent model. The relationship between hydrogen bonding and permeability is considered in greater detail below.

We further assume that the LDC is contained within the ensemble of peptide conformations in water; i.e., it is at least transiently populated in water (this assumption is supported by computational evidence, as discussed below). The flux across the membrane is assumed to be limited by the rate of the LDC passing from the high-dielectric water into the low-dielectric membrane. More precisely, we postulate that the measured PAMPA permeability is proportional to the partition coefficient of the LDC between water (high dielectric) and the membrane interior (low dielectric) or, equivalently, that the log of the permeability coefficient (log P_e) is proportional to the free energy of transferring the LDC from a high-dielectric environment to a low-dielectric environment (ΔG_I, the free energy of "insertion"). In the language of transition-state theory, this implies that the LDC in the membrane can be considered the "transition state" for diffusion through the membrane.

This model bears some similarity to the classical solubility−diffusion model of passive membrane transport;[24,26] i.e., $P_e = K_p D/d$, where P_e is the permeabililty (cm/s), K_p is the unitless partition coefficient of the compound in the membrane, D is the diffusion coefficient in the membrane (cm^2/s), and d is the membrane thickness (cm). Our model has two important

(17) Schreiber, S. L.; Crabtree, G. R. *Immunol. Today* **1992**, *13*, 136−142.
(18) Goodwin, J. T.; Mao, B.; Vidmar, T. J.; Conradi, R. A.; Burton, P. S. *J. Pept. Res.* **1999**, *53*, 355−369.
(19) Parsegian, A. *Nature* **1969**, *221*, 844−846.
(20) Marrink, S. J.; Berendsen, H. J. C. *J. Phys. Chem.* **1996**, *100*, 16729−16738.
(21) Bemporad, D.; Luttmann, C.; Essex, J. W. *Biophys. J.* **2004**, *87*, 1−13.
(22) Bemporad, D.; Luttmann, C.; Essex, J. W. *Biochim. Biophys. Acta* **2005**, *1718*, 1−21.
(23) Wilson, M. A.; Pohorille, A. *J. Am. Chem. Soc.* **1996**, *118*, 6580−6587.
(24) Pohorille, A.; New, M. H.; Schweighofer, K.; Wilson, M. A. *Curr. Top. Membr.* **1999**, *48*, 49−76.
(25) Goodwin, J. T.; Conradi, R. A.; Ho, N. F.; Burton, P. S. *J. Med. Chem.* **2001**, *44*, 3721−3729.
(26) Walter, A.; Gutknecht, J. *J. Membr. Biol.* **1986**, *90*, 207−217.

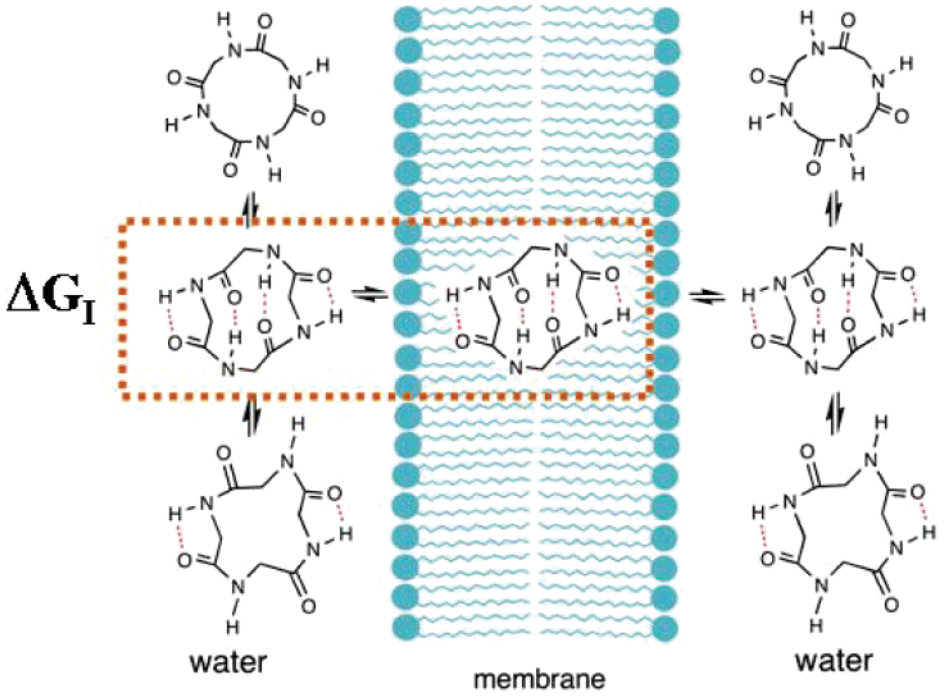

Figure 1. Qualitative overview of the model. The peptides in water are assumed to adopt multiple, rapidly interconverting conformations. The low dielectric of the membrane promotes the formation of internal hydrogen bonds, and the conformation in the membrane (the low-dielectric conformation, LDC) is assumed to be populated in the ensemble of water conformations. The key predictive quantity is ΔG_I, the free energy for transferring the LDC from water to the low-dielectric environment.

differences. First, we assume that differences in the rate of intramembrane diffusion are unimportant for the relatively similar compounds we consider here. Second, we calculate the partition coefficient for *one member* of the ensemble of possible peptide conformations, rather than the overall partition coefficient, which would reflect the overall ensembles in the two different environments. Indeed, experimentally measured phase partitioning with bulk solvents sometimes correlates weakly with the membrane permeability, especially for flexible compounds capable of forming internal hydrogen bonds.[8]

Finally, we point out one important limitation of our model, which is that it does not account for differences in the size or shape of the cyclic peptides. Molecular size is frequently included as a descriptor in parametrized models of membrane permeability, and both computational and experimental studies have suggested the importance of size and shape effects in passive membrane permeability.[20,24,26−29] Evidently, among the relatively similar compounds tested here, this limitation does not substantially degrade the predictive ability of the model when the data obtained are compared to PAMPA data. However, it is clear that we will need to address this limitation of our model, and probably others, in extending this work to more diverse compounds.

Cyclic Peptide Virtual Libraries. As discussed in the Methods, all of the peptides contained one L-tyrosine, which was radiolabeled by methylating the tyrosine −OH with $^{14}CH_3I$. The cyclic peptides used in the computational study were generated as a set of sublibraries categorized by ring size, the number of prolines, and proline spacing (Figure 2). The hexapeptides contained either one or two prolines, and the

1. cyclo[D,LLeu - D,LLeu - D,LLeu - D,LLeu - D,L**Pro** - LTyr(OMe)]
2. cyclo[D,LLeu - D,LLeu - D,LLeu - D,L**Pro** - D,L**Pro** - LTyr(OMe)]
3. cyclo[D,LLeu - D,LLeu - D,L**Pro** - D,LLeu - D,L**Pro** - LTyr(OMe)]
4. cyclo[D,LLeu - D,L**Pro** - D,LLeu - D,LLeu - D,L**Pro** - LTyr(OMe)]

5. cyclo[D,LLeu - D,LLeu - D,LLeu - D,LLeu - D,LLeu - D,L**Pro** - LTyr(OMe)]
6. cyclo[D,LLeu - D,LLeu - D,LLeu - D,LLeu - D,L**Pro** - D,L**Pro** - LTyr(OMe)]
7. cyclo[D,LLeu - D,LLeu - D,LLeu - D,L**Pro** - D,LLeu - D,L**Pro** - LTyr(OMe)]
8. cyclo[D,LLeu - D,LLeu - D,L**Pro** - D,LLeu - D,LLeu - D,L**Pro** - LTyr(OMe)]
9. cyclo[D,LLeu - D,L**Pro** - D,LLeu - D,L**Pro** - D,LLeu - D,L**Pro** - LTyr(OMe)]

Figure 2. Cyclic peptides included in the computational study. Sublibraries 1−4 contain 128 cyclic hexapeptides with one or two prolines. Sublibraries 5−9 contain 320 cyclic heptapeptides with one, two, or three prolines. All indicated stereochemical combinations were included within each sublibrary.

heptapeptides contained one, two, or three prolines. Within each sublibrary, the stereochemistry of the L-Tyr(OMe) was held constant, while the stereochemistries of the remaining residues were permuted to give all possible diastereomers. Permutation of ring size and stereochemistry thus provided access to a diverse array of conformations, while inclusion of proline residues was expected to induce turn structures and limit the number of low-energy conformers available to a given scaffold. Leucine was chosen for the remaining positions since it is one of the most hydrophobic natural amino acids that does not have a branching β-carbon, ensuring efficient couplings during the synthesis. A total of 128 hexapeptides and 320 heptapeptides were included in the study. Detailed data on the entire libraries are provided in the Supporting Information.

The results on the virtual libraries allow us to investigate correlations between predicted permeability and various structural properties of the peptides. The good correlation of ΔG_I with the experimental PAMPA permeability, for the subset of cyclic peptides that were characterized experimentally, suggests that these predictions may be valid, but the relatively small

(27) Deyoung, L. R.; Dill, K. A. *Biochemistry* **1988**, *27*, 5281−5289.
(28) Deyoung, L. R.; Dill, K. A. *J. Phys. Chem.* **1990**, *94*, 801−809.
(29) Xiang, T. X.; Anderson, B. D. *J. Membr. Biol.* **1994**, *140*, 111−122.

Table 1. Correlation of ΔG_I with the Number of Internal Hydrogen Bonds, Obtained from the Computational Analysis of the Virtual Cyclic Peptide Library[a]

	hexapeptides				heptapeptides			
no. of H bonds	no. of peptides	av ΔG_I	std dev	std error	no. of peptides	av ΔG_I	std dev	std error
0	29	3.1	1.8	0.3	18	2.1	1.9	0.4
1	34	2.1	1.4	0.2	93	0.6	1.9	0.2
2	47	1.7	1.3	0.2	132	−0.5	1.6	0.1
3	19	−1.4	1.1	0.3	62	−1.1	1.6	0.2
4	0				14	−3.2	0.9	0.2

[a] The column "no. of H bonds" refers to the number of internal hydrogen bonds, "no. of peptides" is the number of cyclic peptides with that number of hydrogen bonds, "av ΔG_I" is the value of the insertion free energy quantity (kcal/mol) averaged over all peptides with the specified number of hydrogen bonds, and "std dev" and "std error" are the standard deviations and standard errors of these values (kcal/mol).

number of cyclic peptides tested experimentally does not permit direct confirmation of all of these predictions.

First, we investigated the relationships between the predicted passive permeability and the number of hydrogen bonds in the low-dielectric environment. Specifically, we computed the number of hydrogen bonds in the LDC using a simple measure based on the distance between any two electronegative atoms (this distance must be less than the sum of the van der Waals radii). The averages values of ΔG_I for the cyclic peptides with different numbers of hydrogen bonds in the LDC are shown in Table 1. There is clearly a correlation, but it is somewhat weak, and the standard deviation of the averages is quite large. As discussed further below, some compounds with multiple internal hydrogen bonds have relatively poor PAMPA permeability (predicted or measured) and vice versa.

The relationship between permeability and the number of proline residues is also of interest. On one hand, the Pro residues do not contain a hydrogen bond donor group, while the remaining amino acids have a free NH group. This would suggest that increasing numbers of prolines might increase permeability. On the other hand, the prolines have greatly reduced backbone flexibility, which could reduce the ability of the peptides to adopt conformations with good internal hydrogen bonds. In addition, the Leu side chain has a somewhat larger hydrophobic surface area than Pro, and thus, increasing the number of prolines could arguably decrease the overall hydrophobicity of the molecule. However, it is not clear that this difference in the side chain hydrophobicity would outweigh the elimination of the backbone NH group. The computational results are unambiguous that increasing numbers of prolines reduce the predicted passive permeability, on average. For the hexapeptides, the average values of ΔG_I are 0.2 and 2.2 kcal/mol for the peptides with one and two Pro residues, respectively. For the heptapeptides, the trend is similar: the average values of ΔG_I are −1.4, −0.7, and 1.6 kcal/mol for the peptides with one, two, and three Pro residues, respectively. On the basis of the present results, we cannot distinguish whether this effect is due to rigidification or decreased hydrophobicity caused by increasing numbers of prolines.

Overall, the heptapeptides were predicted to have a larger range of permeabilities and on average to be somewhat more permeable than the hexapeptides with the same number of Pro residues (data in the Supporting Information). We attribute this computational result in part to the greater flexibility of the heptapeptides, on average, making it easier for them to form internal hydrogen bonds. However, when hexa- and heptapeptides with the same number of internal hydrogen bonds are considered, the heptapeptides are still predicted to have higher permeability, possibly due to the larger number of hydrophobic side chains. As discussed above, we do not explicitly take molecular size or weight into account in our model, and thus, this prediction may be incorrect, given the generally observed inverse correlation between molecular weight and permeability. The size of the experimental data set is not sufficient to confirm or refute this prediction, although the hexa- and heptapeptides tested experimentally do fall on the same line in the plot of PAMPA permeability vs ΔG_I.

Finally, the computational results can be used to directly assess the validity of our assumption that the LDC is contained among the ensemble of rapidly interconverting conformations in water. The relevant quantity for assessing this assumption is the free energy difference between the energy of the LDC in water and the lowest energy state found in water (ΔG_W). With adequate sampling, i.e., if the lowest energy state found in water is truly the global minimum, this value should always be positive. The average value of this quantity for the 128 hexapeptides is 1.1 kcal/mol, and for the 320 heptapeptides it is 1.0 kcal/mol. Since thermal energy at room temperature (RT) corresponds to 0.6 kcal/mol, the computational data largely support our simple model, subject to the limitations of the sampling and the energy model. That is, for most of the cyclic peptides, the calculations suggest that the free energy of the LDC in water is in fact low enough that it can be at least transiently populated at room temperature. The rate of interconversion among the different low-energy states of the peptide in water is much more difficult to evaluate, and we do not attempt to do so. However, it is highly likely that the different conformations can interconvert on a time scale much faster than that of the PAMPA measurements.

Choice of Peptides for Experimental Investigation. A total of eight cyclic hexapeptides and eight cyclic heptapeptides from the virtual library were chosen for synthesis and experimental testing. These were chosen to span the range of calculated ΔG_I values and to have a variety of different predicted properties, especially the number of hydrogen bonds predicted for the LDC (see Table 2). Due to low cyclization yields, however, four cyclic heptapeptides and seven cyclic hexapeptides were successfully made.

Experimental Results and Correlation with Computational Predictions. The experimental PAMPA data are summarized in Table 2. The correlation between the log P_e determined by the PAMPA assay and the computationally determined values of ΔG_I is shown in Figure 3. Clearly, there is a strong, linear correlation ($R^2 = 0.96$), which is highly encouraging because the computations were performed in advance of the experiments and did not use a training set. The

Table 2. Cyclic Peptides Studied Experimentally[a]

compd	cyclic peptide	ΔG_I	no. of H bonds	PSA	log P_e
3	Leu-D-Leu-D-Leu-D-Leu-Pro-Tyr	−2.87	3	112.2	−6.47 ± 0.07
25	Leu-Leu-D-Leu-D-Pro-Pro-Tyr	1.63	2	150.5	−7.05 ± 0.03
42	D-Leu-D-Pro-Leu-Leu-Pro-Tyr	3.48	2	119.8	−7.27 ± 0.06
46	D-Leu-D-Leu-D-Leu-Leu-Pro-Tyr	3.88	0	170.4	−7.16 ± 0.05
69	D-Leu-D-Leu-D-Leu-D-Leu-Pro-Tyr	−3.04	3	118.1	−6.52 ± 0.01
100	D-Leu-D-Leu-Leu-D-Leu-Pro-Tyr	−2.31	3	116.1	−6.68 ± 0.08
112	D-Leu-Pro-D-Leu-D-Pro-Leu-D-Pro-Tyr	5.36	1	144.5	−7.29 ± 0.08
114	Leu-D-Leu-Leu-Leu-D-Pro-Tyr	−1.17	3	120.5	−6.75 ± 0.13
218	Leu-Leu-Pro-D-Leu-Leu-D-Pro-Tyr	1.47	1	148.4	−6.95 ± 0.20
225	Leu-D-Leu-D-Leu-D-Pro-Leu-D-Pro-Tyr	3.36	1	134.4	−7.12 ± 0.01
294	D-Leu-Leu-D-Leu-Leu-Leu-D-Pro-Tyr	−3.70	3	130.1	−6.41 ± 0.08

[a] The cyclic peptides were chosen to have a wide range of predicted properties, on the basis of the computational results; see the text for details. The values of ΔG_I (see Figure 1) are in kilocalories per mole. The cyclic peptides chosen span nearly the entire range of predicted values of ΔG_I. The column "compd" lists the cyclic peptide identification number, "cyclic peptide" specifies the amino acid sequence, "no. of H bonds" refers to the number of hydrogen bonds observed in the lowest energy structure in the low-dielectric simulation, "PSA" is the polar surface area (Å^2), and "log P_e" is the PAMPA permeability (cm/s). The uncertainties are calculated from the standard deviation of the three replicate measurements.

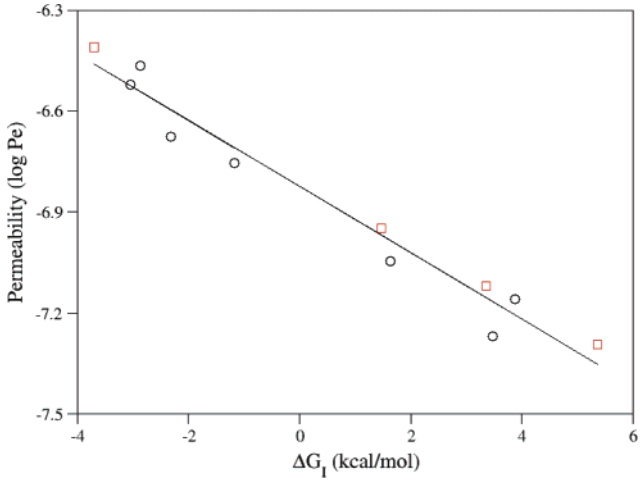

Figure 3. The calculated free energy of insertion (ΔG_I) shows a strong correlation with the permeability determined in the PAMPA (log P_e). The units of P_e are centimeters per second. The correlation coefficient R^2 is 0.96. The cyclic hexapeptides are represented by black circles and the cyclic heptapeptides by red squares.

only adjustable parameters are thus the slope and intercept of the linear fit.

The slope of the linear fit is −0.098. This observed slope deviates significantly from the value predicted by our simple model that the passive permeability is directly proportional to the partition coefficient of the LDC, which implies that log $P_e \propto -\Delta G_I/2.3RT$. The slope of the line would thus be predicted to be −0.71, using energy units of kilocalories per mole, which is obviously much larger in magnitude than the observed slope.

One possible contribution to this discrepancy is our choice of dielectric to represent the interior of the membrane. Largely due to the availability of parameters for the implicit solvent model, we chose chloroform to represent the low-dielectric environment, with a dielectric constant of 4.8. This is lower than many estimates of the dielectric constant of the interior of the membrane and also neglects electrostatic screening from high-dielectric water outside the membrane.[19] The values of ΔG_I are highly sensitive to the choice of the membrane dielectric and scale approximately as $1/\epsilon_{hi} - 1/\epsilon_{lo}$, where ϵ_{hi} and ϵ_{lo} are the high and low dielectrics, because the transfer free energies are dominated by the difference in the solvation free energy computed by the generalized Born solvent model. However, using a membrane dielectric of 10 would only change the slope by a factor of 2. Clearly, other factors must also play a role. A related possibility is that the free energy cost of partitioning into the membrane is reduced by the peptide "dragging" water with it as it passes through the membrane, as has been observed in molecular dynamics simulations of passive membrane permeation.[22]

Finally, it is important to note that we do not account for entropic losses of the peptides upon moving from water to the membrane. The loss of translational and rotational entropy will be determined largely by the size (and shape) of the molecules. Changes in internal entropy are difficult to estimate but would certainly oppose permeation, as has been seen in molecular dynamics simulations.[22] The loss of internal entropy would likely vary among the compounds, with larger entropic losses for those compounds forming the best internal hydrogen bonds in the membrane. The loss of internal entropy is thus likely to reduce the observed slope.

Hydrogen Bonding and Polar Surface Area. Our results suggest that the differing PAMPA permeabilities of the cyclic peptides are related to their ability to form internal hydrogen bonds.[8,14−16,18,25] That is, in general, the ability to form internal hydrogen bonds can promote passive membrane permeability by reducing the free energy cost of desolvating the peptides upon insertion into the membrane. However, as can be seen in Figure 4, the number of hydrogen bonds, by itself, provides only a relatively weak correlation with the PAMPA data (see also Table 1). Figures5 and 6 depict the computationally predicted conformations of three of the experimentally studied cyclic peptides in the low-dielectric medium. The two cyclic heptapeptides in Figure 5 represent the most and least permeable cyclic peptides, as predicted by computation and confirmed by the PAMPA results. The most permeable cyclic peptide shows three internal hydrogen bonds, while the least permeable cyclic peptide has only one. For the most part, the relative permeabilities of the other cyclic peptides investigated correlate loosely with the number of hydrogen bonds, but there are exceptions. Notably, Figure 6 depicts the predicted conformation of a cyclic hexapeptide with two internal hydrogen bonds, which nonetheless has one of the poorest permeabilities according to both the predicted ΔG_I and the experimental PAMPA results. One key feature underlying this discrepancy appears to be that all four carbonyl groups not involved in hydrogen bonds point outward and are not buried by the Leu side chains. These carbonyls

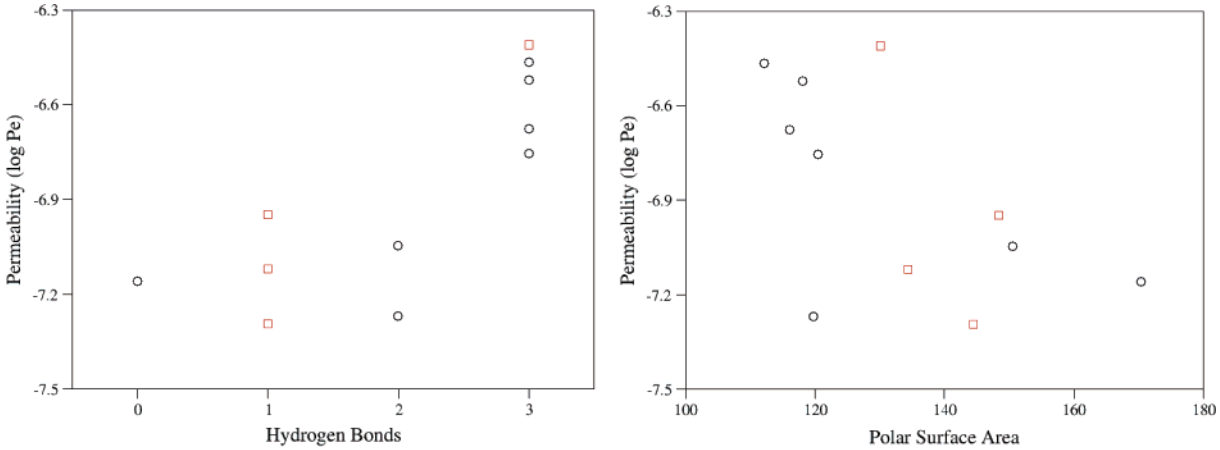

Figure 4. Correlations of permeability with the number of internal hydrogen bonds (left) and the polar surface area (Å^2) (right) in the LDC.

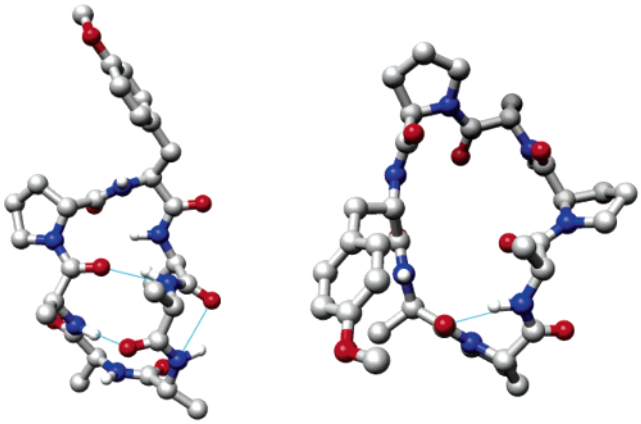

Figure 5. Predicted lowest energy conformations, in a low-dielectric environment, for the most and least permeable cyclic peptides. Only the Cβ atoms of the Leu/D-Leu side chains are shown for clarity. Hydrogen bonds are shown as cyan lines. Left: Cyclic heptapeptide no. 294, which is the most permeable of those tested experimentally and is also predicted to be the most permeable according to ΔG_I. This conformation contains three hydrogen bonds. Right: Cyclic heptapeptide no. 112, which is the least permeable of those tested experimentally and is also predicted to be the least permeable according to ΔG_I. Only one, somewhat strained, hydrogen bond is formed in this conformation.

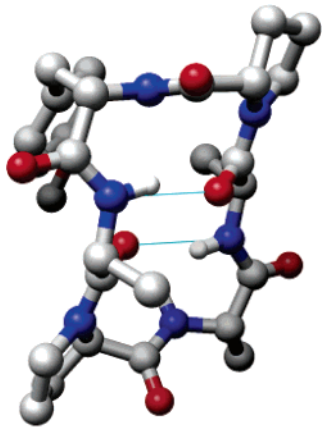

Figure 6. Predicted low-dielectric conformation of cyclic hexapeptide no. 42. Only the Cβ atoms of the Leu/D-Leu side chains are shown for clarity. Hydrogen bonds are shown as cyan lines.

would thus incur a significant desolvation cost upon moving from water to the membrane interior.

The exposed polar surface area is also frequently used as a descriptor in models of membrane permeability. Conformations with greater polar surface area are generally expected to be less permeable because they will incur a larger free energy cost of desolvating the compounds as they pass into the membrane. However, as can be seen in Figure 4, the polar surface area does not correlate nearly as well ($R^2 = 0.34$) with the PAMPA data as the calculated free energy of insertion (ΔG_I).

Initial Testing with Small-Molecule Inhibitors. Significant additional work will be required to demonstrate the utility of our model for small-molecule drug discovery. However, as an initial proof-of-concept, we have applied the same model that we used for the cyclic peptides to a series of fluoroquinolones, for which PAMPA data are available in the literature.[50]

These compounds have significantly fewer rotatable bonds than the cyclic peptides, and for this reason the only sampling we performed was direct energy minimization in the low-dielectric environment, i.e., to generate the LDC. On the other hand, the fluoroquinolone compounds contain a more diverse set of functional groups than the cyclic peptides. For this reason, it is encouraging that a strong linear correlation, $R^2 = 0.83$, is observed between ΔG_I and the intrinsic PAMPA permeability for this series (Figure 7), albeit not as strong as for the cyclic peptides.

In vivo absorption data for these compounds are also available, specifically the results of rat gut in situ perfusion assays (Table 3 in ref 50). A linear correlation between the logarithm of the measured absorption and ΔG_I is observed, with $R^2 = 0.81$.

These results provide additional support for the utility of the free energy of insertion (ΔG_I) as a descriptor for predicting relative permeabilities within a series of relatively closely related compounds. Because the slope and intercept of the linear fit are different between the cyclic peptides and the fluoroquinolones, the results also highlight that additional work will be needed to develop a model capable of predicting absolute PAMPA permeability that is valid across highly diverse sets of compounds. We believe that accurate estimates of translational, rotational, and internal entropy losses upon entering the membrane will be needed to achieve this goal.

Methods

Computational Methods for Predicting Cyclic Peptide Conformations. Some prior studies have used molecular dynamics methods to sample molecular conformations in the context of studying fundamental aspects of passive permeability[20−24] or predicting permeation

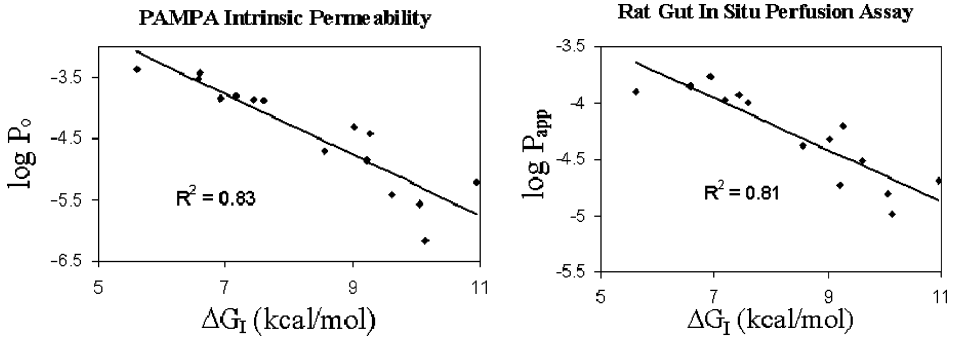

Figure 7. Correlations of the calculated free energy of insertion (ΔG_I) for a series of fluoroquinolones with PAMPA intrinsic permeability (P_0, cm/s) and rat in situ absorption (P_{app}, cm/s). Experimental data were taken from ref 50, Tables 1 and 3.

rates (e.g., by averaging the polar surface area[4]). Here we use a different sampling technology based on dihedral angle sampling,[18] adapted from a reported method for protein loop sampling. The advantage of this method is that it samples the conformational space in an unbiased manner and the time scales for converting between different conformations are irrelevant. Several other computational methods for predicting cyclic peptide structures have also been reported in the literature.[30−36]

All calculations were performed with version "1-9e" of the Protein Local Optimization Program (PLOP), the software platform developed in the Jacobson group. The program is freely available to academic researchers and is part of the commercial Prime package (Schrödinger, Inc.).

Cyclic peptide conformations were generated using the loop prediction algorithm of Jacobson et al.[37] In brief, candidate backbone conformations were generated by splitting the cyclic peptide in half and using a buildup procedure for each side, eventually joining in the middle. The buildup algorithm employs a rotamer-like library of acceptable backbone dihedral angle combinations (ϕ, ψ); Pro and Gly were treated separately from other amino acids, and D-amino acids were treated by inverting the standard Ramachandran preferences. During the buildup procedure, conformations were rejected primarily on the basis of steric clashes and also if there was insufficient space for the side chain to fit adequately or if the peptide could not close, on the basis of geometric criteria. When all acceptable conformations of the C- and N-terminal halves of the peptide had been identified, at a given sampling resolution, closed peptides were generated by pairing halves whose end points were close in space. The resultant closed peptides were filtered on the basis of whether the backbone dihedrals of the closure residue were within Ramachandran-allowed regions, whether steric clashes existed between the two halves, and whether side chains could fit acceptably given the backbone conformation.

Rather than choose in advance the sampling resolution for the backbone dihedral angles, the resolution was adaptively controlled to produce an acceptable number of peptide candidate structures. In this case we generated a minimum of 100 closed peptide backbones. To reduce redundancy among the candidate structures, we employed a K-means clustering algorithm[38,39] to select a representative subset for scoring (52 in this work, a number that was obtained empirically; i.e., doubling the number of clusters did not significantly change the results). Each of the peptides among the subset was then subjected to side chain optimization,[37,40] followed by complete energy minimization.

Side chain sampling was accomplished primarily by using a highly detailed (10° resolution) rotamer library constructed by Xiang and Honig from a database of 297 proteins.[41] This library contains, for example, 2086 rotamers for lysine. The computational expense of such a detailed library was mitigated by prescreening the rotamers using only hard-sphere overlap as a criterion (using a cell list for computational efficiency), allowing many rotamers to be excluded before any energy evaluations were performed. The method we used for the combinatorial optimization was also adapted from the method of Xiang and Honig,[41] which is similar in spirit to earlier work.[42] In brief, all side chains were initially built onto the fixed backbone in a random rotamer state, and then each side chain in the protein was optimized one at a time, holding the others fixed. The procedure was iterated until no side chains changed rotamer states. After convergence was achieved, all side chains were simultaneously energy minimized in Cartesian coordinates.

The all-atom OPLS force field[43,44] was used to describe the peptide intramolecular energetics. The OPLS torsional energy parameters have recently been refined using high-level quantum chemical calculations[44] and validated using protein side chain prediction;[40] the updated parameters were used here. The solvation free energy was estimated using an implicit solvent model consisting of the surface generalized Born (SGB) model of polar solvation.[45] Correction terms have also been developed to improve the agreement between SGB and Poisson−Boltzmann solvation free energy calculations.[45,46] The simulations in water used a dielectric of 80. To treat the low-dielectric environment of the membrane, we modified the generalized Born implicit solvent model to treat chloroform by changing the solvent dielectric constant to 4.8. The "nonpolar" term, which represents the free energy of cavity formation, was treated with a standard surface-area-dependent term, i.e., $a + b$(SASA), where SASA stands for solvent-accessible surface area, and the values of the constants a and b were taken from previous work by Gilson and co-workers.[47] The GB/SA solvent model, for water and chloroform, has also been used in prior studies of membrane permeability.[18]

The experimental strategy for measuring the membrane permeability, described below, required the use of radiolabeled *O*-methyltyrosine. The force field parameters for this modified amino acid were obtained using an automated "atom-typing" program implemented in the Impact

(30) Riemann, R. N.; Zacharias, M. *J. Pept. Res.* **2004**, *63*, 354−364.
(31) Baysal, C.; Meirovitch, H. *Biopolymers* **1999**, *50*, 329−344.
(32) Baysal, C.; Meirovitch, H. *Biopolymers* **2000**, *53*, 423−433.
(33) Baysal, C.; Meirovitch, H. *Biopolymers* **2000**, *54*, 416−428.
(34) Deem, M. W.; Bader, J. S. *Mol. Phys.* **1996**, *87*, 1245−1260.
(35) Nikiforovich, G. V.; Kolodziej, S. A.; Nock, B.; Bernad, N.; Martinez, J.; Marshall, G. R. *Biopolymers* **1995**, *36*, 439−452.
(36) Che, Y.; Marshall, G. R. *J. Med. Chem.* **2006**, *49*, 111−124.
(37) Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J.; Honig, B.; Shaw, D. E.; Friesner, R. A. *Proteins* **2004**, *55*, 351−367.
(38) Hartigan, J. A.; Wong, M. A. *Appl. Stat.* **1979**, *28*, 100−108.
(39) Hartigan, J. A. *Clustering Algorithms*; Wiley: New York, 1975.
(40) Jacobson, M. P.; Kaminski, G. A.; Friesner, R. A.; Rapp, C. S. *J. Phys. Chem. B* **2002**, *106*, 11673−11680.
(41) Xiang, Z. X.; Honig, B. *J. Mol. Biol.* **2001**, *311*, 421−430.
(42) Bruccoleri, R. E.; Karplus, M. *Biopolymers* **1987**, *26*, 137−168.
(43) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(44) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.
(45) Ghosh, A.; Rapp, C. S.; Friesner, R. A. *J. Phys. Chem. B* **1998**, *102*, 10983−10990.
(46) Gallicchio, E.; Zhang, L. Y.; Levy, R. M. *J. Comput. Chem.* **2002**, *23*, 517−529.
(47) Luo, R.; Head, M. S.; Given, J. A.; Gilson, M. K. *Biophys. Chem.* **1999**, *78*, 183−193.

software package. The partial charges on the *O*-methyl group were assigned to be −0.35 (O), 0.02 (C), and 0.06 (H). The remaining partial charges on the residue were retained from the standard force field parameters for Tyr.

The overall computational expense for sampling one cyclic peptide in either a high-dielectric environment or a low-dielectric environment was approximately 6 min for the hexapeptides and approximately 15 min for the heptapeptides, using a single 2.8 GHz Xeon processor.

Computational Methods for Predicting Membrane Permeability. Having identified the lowest energy conformations of the cyclic peptides in the low- and high-dielectric media, the values of ΔG_I were computed as defined in Figure 1. As discussed in the Results and Discussion, we also computed ΔG_W, the free energy difference between the energy of the LDC in water and the lowest energy state found in water. This value of ΔG_W directly assesses the validity of our assumption that the LDC is at least transiently populated in water. It also permits an assessment of the effectiveness of the sampling method. If the computational method succeeds in identifying the lowest energy conformations in the two dielectrics, then ΔG_W must be a positive value. In practice, we used a cutoff of $\Delta G_W > -1.0$ kcal/mol to account for the inherent precision of the calculations (vide infra). Two of the 128 cyclic hexapeptides and 25 of the 320 cyclic heptapeptides showed evidence of incomplete sampling according to this criterion and were excluded from further analysis. Overall, this measure of internal consistency suggests that the conformational sampling is satisfactory for the cyclic hexapeptides and generally satisfactory for the cyclic heptapeptides. Larger cyclic peptides would likely require qualitatively more sampling than we have performed here.

We also performed a series of calculations aimed at estimating the inherent precision of the computational values. Specifically, we computed the values of the transfer free energies for pairs of enantiomers, specifically for all cyclic hexapeptides with the sequence cyclo(Leu-Leu-Leu-Leu-Leu-O-Me-Tyr), with all combinations of stereochemistries for the Leu residues (32 total, 16 pairs). The pairs of enantiomers should, of course, have identical physical properties including permeability and phase partitioning. However, due to technical aspects of the calculations, the results can be obtained completely independently and are not identical. The RMS values of ΔG_I were ∼1.0 kcal/mol over the 16 pairs of enantiomers, which we consider a rough estimate for the precision of the computations.

Synthesis. All cyclic peptides were synthesized utilizing Fmoc solid-phase peptide synthesis,with linkage to the solid phase via the tyrosine side chain as the silyl ether.[13,48] Each reaction was performed on 100 mg of resin. After cyclization and cleavage from the resin, each peptide was purified by preparative HPLC and lyophilized to dryness. The purified peptides were then methylated by dissolving the peptide in DMF (6 mM) and adding K_2CO_3 (2 equiv) and $^{14}CH_3I$ (0.12 mCi, 10 equiv). When the methylation had gone to completion (∼24 h), the reaction mixtures were filtered through disposable C18 cartridges and the resin was rinsed with 5 column volumes of water. The peptides were then eluted with acetonitrile, and the solvent was evaporated under a stream of air. HPLC analysis of the resulting products showed >98% purity for all products except one, which showed double methylation, likely the result of a small amount of *N*-methylation. This peptide was discarded from the study.

PAMPA. DMSO stocks (1 mM) of the radiolabeled cyclic peptides were made, and each peptide was analyzed by liquid scintillation counting to determine its specific activity. Next, 2.5 μM solutions of each peptide were made in PBS (containing 0.25% DMSO) as starting donor well solutions for the PAMPA (MultiScreen-IP hydrophobic plate, cat. no. MAIPN4510/Millipore). A 1% solution of lecithin in dodecane was then applied to each filter well at 5 μL per well. Immediately after application of the lipid membrane, donor solutions were added to the wells. Incubation times for all peptides were 19 h, after which the acceptor well radioactivity was measured by adding 200 μL from the acceptor well to 10 mL of scintillation fluid and submitting it to liquid scintillation counting. Final log P_e values for each cyclic peptide were obtained as averages of the values from three wells (calculated according to ref 49).

Acknowledgment. This work was supported in part by NIH Grants AI035707 and GM56531 (to M.P.J.) and CA104569-03 (to R.S.L.). We thank the anonymous reviewers for their helpful suggestions. M.P.J. thanks Ken Dill (UCSF) for illuminating conversations.

Supporting Information Available: Detailed experimental procedures on compound synthesis and PAMPA, computationally predicted low-dielectric structures for all cyclic peptides tested experimentally, and detailed computational data for the virtual cyclic peptide library. This material is available free of charge via the Internet at http://pubs.acs.org.

JA063076P

(48) Tallarico, J. A.; Depew, K. M.; Pelish, H. E.; Westwood, N. J.; Lindsley, C. W.; Shair, M. D.; Schreiber, S. L.; Foley, M. A. *J. Comb. Chem.* **2001**, *3*, 312−318.
(49) Schmidt, D.; Lynch, J. Application Note AN1725EN00; Millipore Corp.: Billerica, MA.
(50) Bermejo, M.; Avdeef, A.; Ruiz, A.; Nalda, R.; Ruell, J. A.; Tsinman, O.; Gonzalez, I.; Fernandez, C.; Sanchez, G.; Garrigues, T. M.; Merino, V. *Eur. J. Pharm. Sci.* **2004**, *21*, 429−441.