Three-Dimensional Quantitative Structure−Activity Relationship Analyses of Peptide Substrates of the Mammalian H^+/Peptide Cotransporter PEPT1

Sabine Gebauer,[†] Ilka Knütter,[†,‡] Bianka Hartrodt,[†] Matthias Brandsch,[‡] Klaus Neubert,[†] and Iris Thondorf*[,†]

Institute of Biochemistry, Department of Biochemistry/Biotechnology, and Biozentrum, Martin-Luther-University Halle-Wittenberg, D-06099 Halle, Germany

Received July 28, 2003

The utilization of the carrier protein PEPT1 for the absorption of peptidomimetic drug molecules is a promising strategy for oral drug administration and increasing bioavailability. In the absence of structural information on the binding mode of substrates to PEPT1, a computational study was conducted to explore the structural requirements for substrates and to derive a predictive model that may be used for the design of novel orally active drugs. A comparative molecular field analysis (CoMFA) and a comparative molecular similarity indices analysis (CoMSIA) were performed on a series of 79 dipeptide-type substrates for which affinity data had been collected in a single test system under the same conditions. These studies produced models with conventional r^2 and cross-validated coefficient (q^2) values of 0.901 and 0.642 for CoMFA and 0.913 and 0.776 for CoMSIA. The models were validated by an external test set of 19 dipeptides and dipeptide derivatives. CoMSIA contour maps were used to identify the recognition elements that are relevant for the binding of PEPT1 substrates. The 3D QSAR models provide an insight in the interactions between substrates and PEPT1 on the molecular level and allow the prediction of affinity constants of new compounds.

Introduction

The oral bioavailability of drugs is often limited due to the barrier formed by the intestinal epithelium. A promising strategy to overcome this barrier for drug delivery lies in the utilization of existing transport systems located in the intestinal cell membranes. The PEPT1 protein expressed in the apical brush border membrane of the small intestine mediates the transport of di- and tripeptides from the intestinal lumen driven by a transmembrane proton gradient.[1−3] Given the theoretically >8000 endogenous di- and tripeptides, it is not surprising that the substrate specificity of PEPT1 is very broad, although the affinities of structurally diverse peptides may differ by some orders of magnitude. In addition, it is well-known that also certain drugs such as β-lactam antibiotics, angiotensin-converting enzyme (ACE) inhibitors, valacyclovir, and bestatin are recognized by PEPT1.[4−8] These findings make the PEPT1 pathway attractive for the design of orally active drugs either by adaptation of the drug's structure to the requirements of the transporter or by conjugation of a drug with a small and rigid substrate of PEPT1. Both strategies require a detailed knowledge of the binding site geometry or at least of the pharmacophoric pattern of substrates.

Due to its membrane-spanning character, the three-dimensional structure of PEPT1 is hitherto unknown and will not be available soon. Hence, the main source of information about structure−affinity/transport relationships is extensive binding and transport assays using a variety of structurally diverse compounds followed in some cases by computational investigations.[9−14] The only other available data come from experiments with chimeric mammalian peptide transporters or site-directed mutagenesis studies.[15,16] Since the current knowledge about the substrate specificity of PEPT1 is based on data collected from different cells, from different species and organs, it is, however, difficult to condense the results in a single model. The presence of a free N-terminal α-amino group and a peptide bond seem to be essential for substrate recognition, although the peptide bond may be replaced by a ketomethylene group.[17] At the C-terminus, derivatization of the free carboxyl group is less tolerated by the carrier than its replacement by hydrophobic groups. With regard to the stereochemistry of substrates, the carrier protein seems to be entirely specific for a trans arrangement of the peptide bond, and dipeptides consisting of two L-amino acids show in general the highest affinities, although also dipeptides with one D-amino acid, especially the DL-isomers, are often recognized by PEPT1.[18,19] A distance of 5.0−6.3 Å between the N- and C-terminus has been suggested as a further recognition element.[11] Considering not only the nutritional importance of PEPT1 but also its pharmacological relevance for the oral bioavailability of drugs, structural information about substrate−protein interactions are highly relevant for a rational design of orally active peptide mimetics. Therefore, we performed a computational study to explore the structural requirements for substrates and to derive a predictive model that may be used for the design of novel orally active drugs.

Here, we report the results of conformational analysis, pharmacophore identification, and three-dimensional

* To whom correspondence should be addressed. Tel: +49-345-5524862. Fax: +49-345-5527011. E-mail: thondorf@biochemtech.uni-halle.de.
[†] Institute of Biochemistry, Department of Biochemistry/Biotechnology.
[‡] Biozentrum.

10.1021/jm030976x CCC: $25.00
Published on Web 11/14/2003

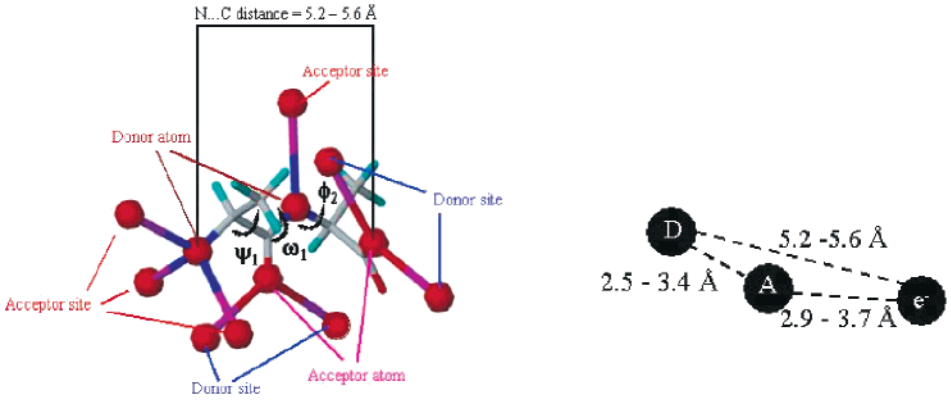


Figure 1. The 3D structure of **14** and the three-point recognition model for PEPT1 (D = Donor, A = Acceptor and e^- = high electron density).

quantitative structure−activity relationship (3D QSAR) studies using the comparative molecular field analysis (CoMFA) and the more recently introduced comparative molecular similarity indices analysis (CoMSIA) on a variety of structurally diverse dipeptides and dipeptide derivatives that exhibit a broad range of binding affinities for PEPT1.[20,21]

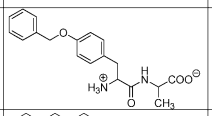
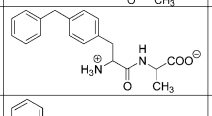
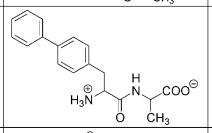
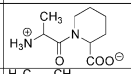
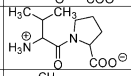
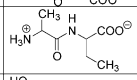
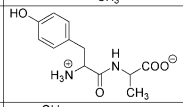
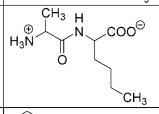
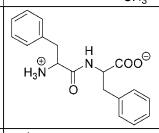
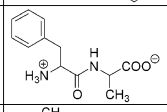
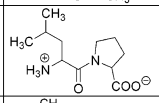
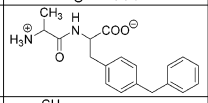
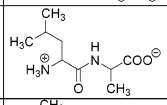
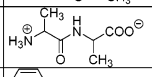
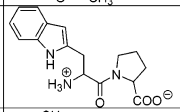
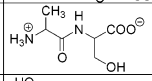
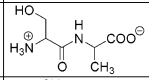
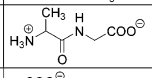
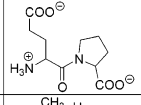
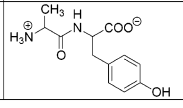
Results and Discussion

Conformational Analysis and Molecular Alignment. Since the substrates of PEPT1 are in general very flexible compounds that may exist in numerous local energy minima, it is a priori not clear whether the ligand interacts with the receptor in its global minimum structure or another similarly low energy conformation. Thus, a systematic conformational search and subsequent energy minimization was performed to identify the relevant conformers of each substrate from which individual conformational databases were generated. The number of conformers in these databases was regulated by considering only those in an energy range of 5 kcal mol^{-1} above that of the lowest energy conformer found and by filtering the dataset for unique minima using a rms deviation of 0.3 Å as identity criterion to eliminate redundant conformers. Depending on the number of rotatable bonds present in the molecule, the conformational databases contained between two (**30**) and 227 (**65**) conformers.

3D QSAR analyses require a suitable alignment of conformers based on a template conformation. This pharmacophore model or bioactive conformation is defined by a specific geometric arrangement of pharmacophoric features, such as hydrogen-bond donor and acceptor sites and aromatic and aliphatic ring centroids, which represent an image of the macromolecular binding site. The derivation of the pharmacophore model of PEPT1 substrates was done using the DISCO module implemented in the Sybyl program package, which compares the intramolecular distances between features across all active compounds. The reference for the DISCO run should possess, for efficiency, few conformers. This is realized in the relatively high affinity substrate Ala-ψ[CS−N]-Pro (**30**, K_i = 0.33 mM), for which only two conformers were found to exist with a trans arrangement of the peptide bond. However, since Ala-ψ[CS−N]-Pro is not a natural substrate of PEPT1 and its flexibility is restricted by the pyrrolidine ring, we decided to choose the simplest natural, chiral substrate of PEPT1, L-Ala-L-Ala (**14**, K_i = 0.12 mM) as the reference compound. For **14**, six unique conformers were obtained from the systematic search. Comparison of these six conformers with those of **30** revealed that only a single spatial arrangement is shared by both compounds (characterized by $\psi_1 = 165°$, $\omega_1 = 178°$, $\phi_2 = -65°$, $d_{N\cdots C} = 5.3$ Å). In the subsequent DISCO run with the reference conformer of **14** we considered (1) site hydrogen-bond-acceptors of the protonated basic nitrogen atom, (2) site hydrogen-bond-acceptors of the nitrogen atom of the peptide bond, (3) site hydrogen-bond-donors to the carbonyl oxygen and the deprotonated carboxylic group, (4) hydrogen-bond-donor atoms and (5) hydrogen-bond-acceptor atoms as pharmacophoric features (Figure 1). The DISCO models were required to match between three and 12 features.

To derive a pharmacophore model we used DISCO in a somewhat different way as initially described.[22,23] Due to the high structural diversity of the molecules, the model was built up stepwise by comparing each conformer database separately to the reference conformer of **14**. Then, the classical procedure was followed: (i) combining conformer databases to create a molecular spreadsheet, (ii) assigning features to the rows automatically, (iii) moving features to the conformers, (iv) scanning the reference compound **14**, (v) admitting 3−12 points of feature requirements, (vi) computing the model, and (vii) examination and analysis of the results spreadsheet. The best matches were combined in a conformational database providing the source for subsequent 3D QSAR analyses. However, not all DISCO runs yielded a model suitable for these analyses. Therefore, for compounds **15**, **19**, **29**, **44**, **61**, **73**, and **74**, a manual alignment by fitting of the peptide backbone onto the template was performed. For the conformers with defined backbone torsion angles and N···C distances (exceptions are, for example, compounds **25**, **32**, **33**, **35**, **36**, and **42**, Table 1), the following geometrical data characterize the model: $\psi_1 = 165°$ to 142°, $\omega_1 \approx 180°$, $\phi_2 = -57°$ to −98° and $d_{N\cdots C} = 5.2-5.6$ Å (Figure 1). From the pharmacophore generation a very simple three-point recognition model can be derived comprising a single hydrogen-bond-donor atom (D), one hydrogen-bond-acceptor atom (A), and an area of high electron densitiy (e^-) as shown in Figure 1 (cf., Table 1). A second hydrogen-bond-donor, such as the peptide nitrogen atom, does not seem to be essential (which is in

Table 1. Biological Data and Results of the Conformational Search and of the DISCO Analysis for the Compounds Used in the Training Set

No.	Compound	K_i [a] [mM]	lg 1/K_i	No. of conf.[b]	ΔE^c [kcal× mol^{-1}]	D...A[d] [Å]	A...e$^-$[e] [Å]	No. of Feat.[f]	RMS[g] [Å]
1		0.01	1.85	81	3.78	2.64	3.28	11	0.06
2		0.02	1.68	55	4.01	2.64	3.25	11	0.08
3		0.02	1.68	32	2.27	2.64	3.28	11	0.07
4		0.06	1.22	4	0.00	2.54	3.42	6	0.18
5		0.06	1.22	4	0.09	2.60	2.96	8	0.27
6		0.07	1.15	18	0.22	2.64	3.51	10	0.36
7		0.09	1.05	31	1.72	2.64	3.26	11	0.07
8		0.09	1.05	112	0.19	2.64	3.47	10	0.21
9		0.11	0.96	37	3.39	2.64	3.32	11	0.25
10		0.11	0.96	16	1.63	2.64	3.26	11	0.07
11		0.11	0.96	5	0.00	2.65	3.19	5	0.05
12		0.12	0.92	39	3.72	2.64	3.31	11	0.16
13		0.12	0.92	40	0.64	2.63	3.26	8	0.11
14		0.12	0.92	6	0.00	2.64	3.26	12[h]	
15		0.13	0.89	12	1.56	2.65	3.21	manu[i]	0.14[j]
16		0.14	0.85	24	0.00	2.64	3.28	11	0.10
17		0.14	0.85	27	0.20	2.65	3.26	11	0.05
18		0.14	0.85	9	0.08	2.64	3.25	11	0.18
19		0.14	0.85	12	0.00	2.65	3.20	manu[i]	0.14[j]
20		0.15	0.82	38	1.39	2.64	3.32	11	0.18

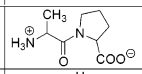
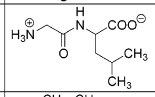
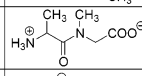
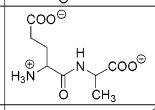
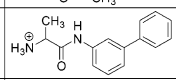
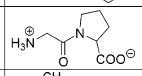
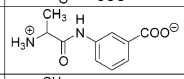
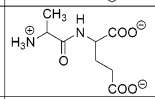
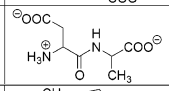
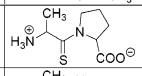
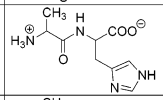
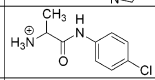
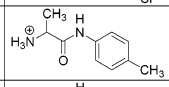
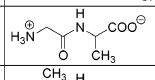
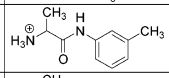
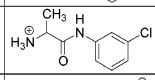
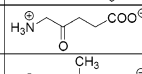
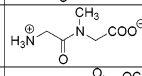
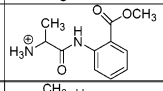
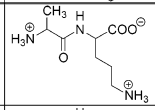
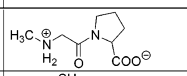
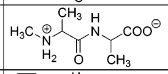
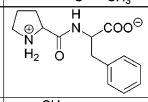
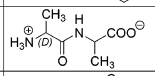
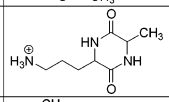
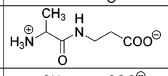
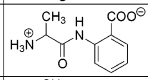
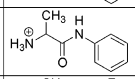
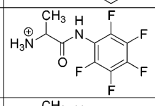
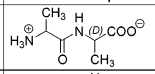
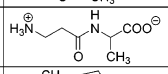
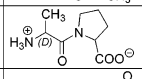
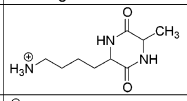
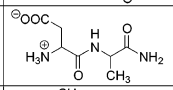
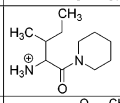
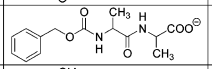
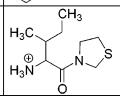
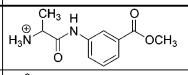
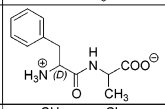
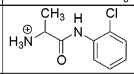
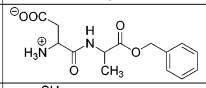
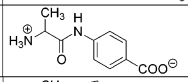
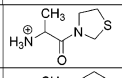
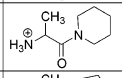
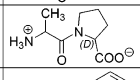
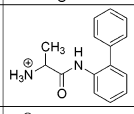
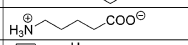
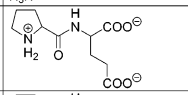
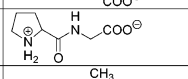
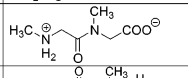
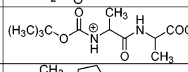
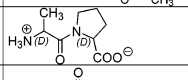
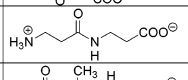
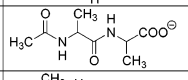
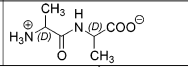
No.	Compound	K_i [a] [mM]	lg 1/K_i	No. of conf.[b]	ΔE^c [kcal× mol^{-1}]	D...A[d] [Å]	A...e$^-$[e] [Å]	No. of Feat.[f]	RMS[g] [Å]
21		0.15	0.82	2	0.00	2.65	3.21	5	0.15
22		0.16	0.80	45	0.33	2.65	3.44	8	0.14
23		0.25	0.60	4	0.20	2.55	3.24	6	0.19
24		0.25	0.60	43	0.51	2.64	3.25	11	0.17
25		0.29	0.54	24	0.00	2.63	3.78	8	0.08
26		0.30	0.52	3	0.00	2.63	3.21	6	0.20
27		0.31	0.51	12	0.00	2.63	3.78	8	0.08
28		0.32	0.49	67	1.77	2.64	3.45	10	0.19
29		0.32	0.49	24	0.15	2.64	3.39	manu[i]	0.01[j]
30		0.33	0.48	2	0.00	3.02	3.36	4	0.03
31		0.33	0.48	66	0.90	2.64	3.26	8	0.00
32		0.33	0.48	6	0.00	2.63	3.79	8	0.08
33		0.34	0.47	9	0.00	2.63	3.79	8	0.08
34		0.38	0.42	6	0.00	2.65	3.29	8	0.11
35		0.41	0.39	12	0.00	2.63	3.78	8	0.08
36		0.46	0.34	12	0.00	2.63	3.78	8	0.08
37		0.79	0.10	36	0.63	2.69	2.87	6	0.30
38		0.83	0.08	14	0.00	2.65	3.49	8	0.36
39		0.89	0.05	23	0.25	2.64	3.69	8	0.04
40		0.98	0.01	115	0.54	2.64	3.39	10	0.12
41		1.0	-0.01	10	0.00	2.65	3.29	9	0.11
42		1.1	-0.03	12	0.00	2.63	3.79	8	0.08
43		1.2	-0.06	3	0.00	2.63	2.95	8	0.03
44		1.4	-0.15	101	4.65	3.39	3.28	manu[i]	0.01[k]

Table 1. (Continued)

No.	Compound	K_i [a] [mM]	lg 1/K_i	No. of conf.[b]	ΔE[c] [kcal× mol^{-1}]	D...A[d] [Å]	A...e$^-$[e] [Å]	No. of Feat.[f]	RMS[g] [Å]
45		1.6	-0.20	8	0.74	2.66	3.25	5	0.24
46		1.8	-0.26	24	0.03	2.69	3.28	8	0.02
47		1.9	-0.28	23	1.23	2.68	3.32	8	0.14
48		2.0	-0.31	6	0.00	2.64	3.29	7	0.06
49		2.3	-0.37	13	0.44	n/a[l]	n/a[l]	5	0.26
50		2.7	-0.43	27	0.00	2.64	3.21	8	0.07
51		2.8	-0.44	9	0.00	2.64	3.67	8	0.07
52		2.9	-0.46	6	0.00	2.63	3.79	8	0.08
53		3.1	-0.49	6	0.00	2.64	3.77	8	0.07
54		4.2	-0.62	9	0.00	2.64	2.95	8	0.02
55		4.8	-0.68	17	0.00	3.35	3.20	6	0.00
56		4.9	-0.69	2	0.00	2.63	3.23	3	0.00
57		5.4	-0.73	52	1.36	n/a[l]	n/a[l]	5	0.27
58		5.4	-0.73	24	0.35	2.64	3.13	10	0.10
59		5.62	-0.75	9	0.60	2.54	n/a[l]	6	0.20
60		5.6	-0.75	19	2.34	2.63	3.30	8	0.07
61		6.3	-0.80	17	0.51	2.64	n/a[l]	manu[i]	0.13[m]

No.	Compound	K_i [a] [mM]	lg 1/K_i	No. of conf.[b]	ΔE[c] [kcal× mol^{-1}]	D...A[d] [Å]	A...e$^-$[e] [Å]	No. of Feat.[f]	RMS[g] [Å]
62		6.9	-0.84	24	0.00	2.64	4.53	8	0.08
63		7.1	-0.85	27	2.10	3.39	3.29	7	0.06
64		7.8	-0.89	12	0.00	2.64	3.72	8	0.06
65		8.3	-0.92	227	1.38	2.64	3.51	11	0.13
66		9.1	-0.96	6	0.00	2.63	3.79	8	0.08
67		11.0	-1.04	2	0.00	2.65	n/a[l]	5	0.16
68		12.9	-1.11	2	0.00	2.56	n/a[l]	6	0.21
69		15.1	-1.18	2	0.00	2.65	2.79	3	0.08
70		16.6	-1.22	12	0.00	2.64	3.73	8	0.07
71		17.4	-1.24	33	0.00	n/a[l]	n/a[l]	5	0.38
72		20	-1.30	51	1.66	2.68	3.46	8	0.45
73		22	-1.34	9	0.00	2.68	3.25	manu[i]	0.09[j]
74		≈ 32	-1.50	10	0.00	2.61	3.04	manu[i]	0.16[m]
75		>30 (≈ 62)	-1.79	8	0.77	2.63	3.29	8	0.05
76		>30 (≈ 65)	-1.81	2	0.00	2.61	2.77	3	0.03
77		>30 (≈ 77)	-1.89	40	1.46	2.61	4.86	5	0.32
78		>30 (≈ 90)	-1.96	22	0.63	2.64	3.30	8	0.05
79		>30 (≈ 100)	-2.00	9	0.00	2.64	2.95	4	0.00

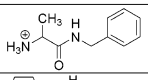
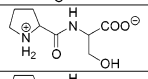
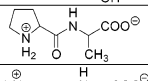
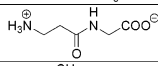
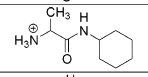
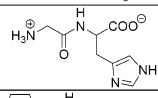
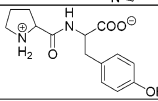
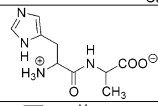
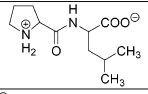
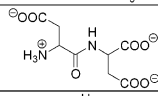
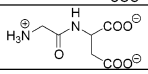
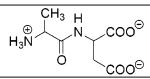
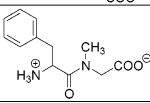
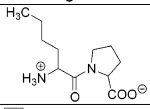
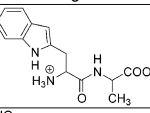
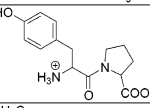
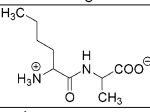
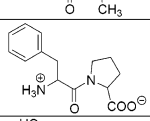
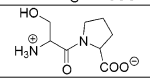
[a] Affinity constant. [b] Number of conformers of the systematic search using an energy cutoff of 5 kcal mol^{-1}. [c] Energy difference between the selected conformer and the lowest energy conformer found by the systematic search. [d] Distance between donor (D) and acceptor (A). [e] Distance between donor (D) and regions of high electron density (e$^-$), where e$^-$ is defined as the centroid of the two C-terminal carboxyl oxygen atoms or the aromatic moiety of aryl amides. [f] Number of features for a DISCO run. [g] Measure for the deviation of the distance between the features in the ligand and in the reference compound **14**. [h] Reference compound. [i] Manual fitting. [j] Root mean square measured along the backbone. [k] Root mean square measured along the backbone without the ammonium nitrogen atom. [l] Not available. [m] Root mean square measured along the backbone without the C-terminal carbon atom.

accordance with the high affinity of **4**, **5**, **11**, **15**, **19**, **21**, **23**, **26**, **37** and **38**).

CoMFA and CoMSIA Studies. Starting from the structural alignment of all conformers that is depicted in Figure 2, a comprehensive CoMFA and CoMSIA analysis was initiated. A training set comprising 79 dipeptide-type substrates of PEPT1 allowed the derivation of 3D QSAR models of statistical significance (Table 3). They provide a satisfying prediction of binding affinities for PEPT1. Fitted predictions versus measured binding affinities are plotted in Figure 3. In general, CoMSIA results yield slightly better correlations with regard to higher q^2 values, whereas the numbers of PLS components is equal. Models were evaluated for a 2 Å grid spacing. For CoMSIA, an additional lattice using an 1 Å grid spacing was evaluated to illustrate physicochemical features in smoother contour maps.

As the predictive ability of our models can be overestimated, an external validation was performed on a test set comprising 19 dipeptide-type substrates (Table 2). Plots of the predictive versus measured affinity constants are shown in Figure 4. The predicted sub-

Table 2. Biological Data and Results of the Conformational Search and of the DISCO Analysis for the Compounds Used in the Test Set

No.	Compound	K_i[a] [mM]	lg 1/K_i	No. of conf.[b]	ΔE[c] [kcal × mol^{-1}]	D…A[d] [Å]	A…e$^-$[e] [Å]	No of Feat.[f]	RMS[g] [Å]
80		14.1	-1.15	13	0.26	2.64	3.60[h]	8	0.10
81		14.0	-1.15	20	0.00	2.65	3.28	8	0.13
82		9.5	-0.98	8	0.00	2.65	3.26	8	0.10
83		7.3	-0.86	13	0.00	3.35	3.17	4	0.01
84		1.5	-0.17	9	0.00	2.64	n/a	8	0.02
85		1.0	-0.01	62	1.03	2.65	3.28	7	0.24
86		0.73	0.14	30	1.23	2.65	3.30	9	0.20
87		0.48	0.32	65	0.00	2.68	3.83	manu.[j]	0.08[k]
88		0.47	0.33	40	0.44	2.68	3.41	5	0.03
89		0.41	0.39	83	0.26	2.65	3.44	11	0.39
90		0.35	0.46	36	0.51	2.65	3.36	8	0.14
91		0.26	0.59	27	0.28	2.64	3.32	11	0.09
92		0.21	0.68	11	2.43	2.55	3.23	6	0.19
93		0.16	0.80	24	0.35	2.65	3.20	6	0.22
94		0.16	0.80	34	2.91	2.64	3.26	11	0.01
95		0.16	0.80	4	1.52	2.65	3.20	3	0.16
96		0.14	0.85	55	0.61	2.64	3.26	11	0.10
97		0.13	0.85	4	1.40	2.65	3.20	manu.[j]	0.14[k]
98		0.13	0.85	6	0.03	2.66	3.20	manu.[j]	0.14[j]

[a] Affinity constant. [b] Number of conformers of the systematic search using an energy cutoff of 5 kcal mol^{-1}. [c] Energy difference between the selected conformer and the lowest energy conformer found by the systematic search. [d] Distance between donor (D) and acceptor (A). [e] Distance between donor (D) and regions of high electron density (e$^-$), where e$^-$ is defined as centroid of the two C-terminal carboxyl oxygen atoms or the aromatic moiety of aryl amides. [f] Number of features for a DISCO run. [g] Measure for the deviation of the distance between the features in the ligand and in the reference compound **14**. [h] Aromatic moiety is situated at the opposite site. [i] Not available. [j] Manual fitting. [k] Root mean square measured along the backbone.

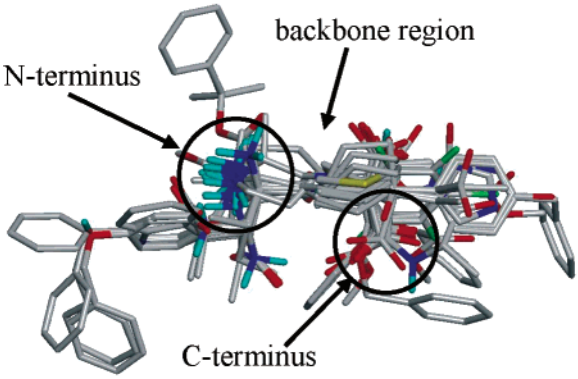


Figure 2. Structural alignment of the training set (hydrogen atoms bonded to carbon atoms have been omitted for clarity).

Table 3. Summary of CoMFA and CoMSIA Results

	CoMFA	CoMSIA
q^2	0.642	0.776
s_{press}	0.580	0.459
r^2	0.901	0.913
S	0.305	0.285
F	168	195
components	4	4
fraction		
steric		11.7
electrostatic		22.9
lipophilic		24.5
donor		24.2
acceptor		16.7
grid spacing (Å)	2	2

strate affinities for the test set are in almost all cases in reasonable agreement with the experimental values, which supports the predictive ability of the statistical models generated by both 3D QSAR methods. In CoMSIA, one outlier, **88** (Pro-Leu), was observed. This is probably due to the anomalous gradual decrease of measured affinity constants compared to physicochemical properties of molecules of the Pro-Xaa series. Affinity constants for Pro-Xaa dipeptides range from 0.47 mM for Pro-Leu up to 22 mM for Pro-Gly. While Pro-Xaa dipeptides bearing hydrophobic side chains show reasonable affinities (e.g., **47**, **86**, **88**, and Pro-Pro, 0.61 mM), substrates carrying charged and hydrophilic amino acids in the second position yield lower affinities (e.g., Pro-Lys, 3.2 mM; Pro-Asp, 9.8 mM; Pro-Ser, 14 mM; and Pro-Glu, 20 mM).[10] It seems that this huge variance of affinity constants is difficult to fit into our 3D QSAR models.

Graphical Interpretation of the Results. Klebe has recently pointed out that fields used in CoMFA imply some inherent deficits arising from Lennard-Jones and Coulomb potentials.[24] Due to these problems, CoMFA maps are often not contiguously connected and

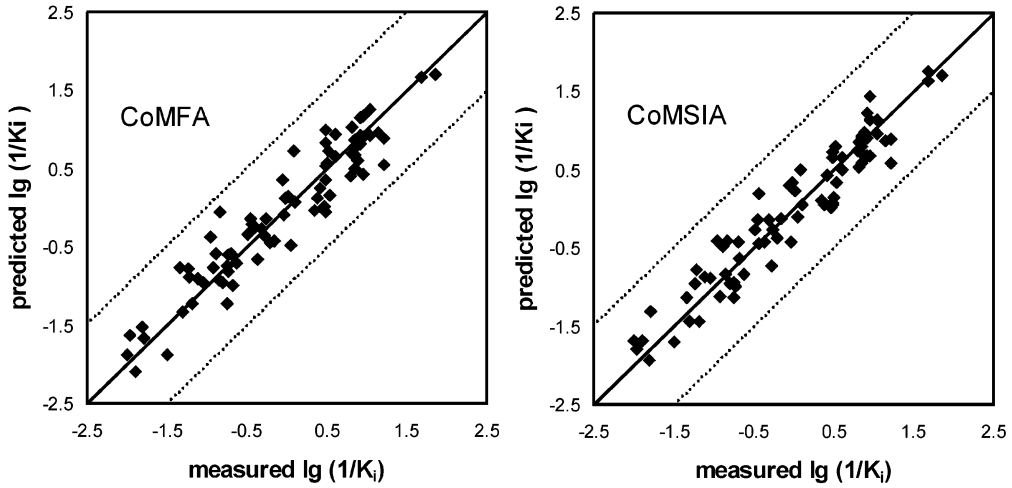


Figure 3. Predicted versus measured affinity constants for the 79 dipeptides and dipeptide derivatives of the training set. The predicted values were obtained by PLS analyses using CoMFA and CoMSIA with 2 Å grid spacing. Dashed lines denote deviations of one logarithmic unit.

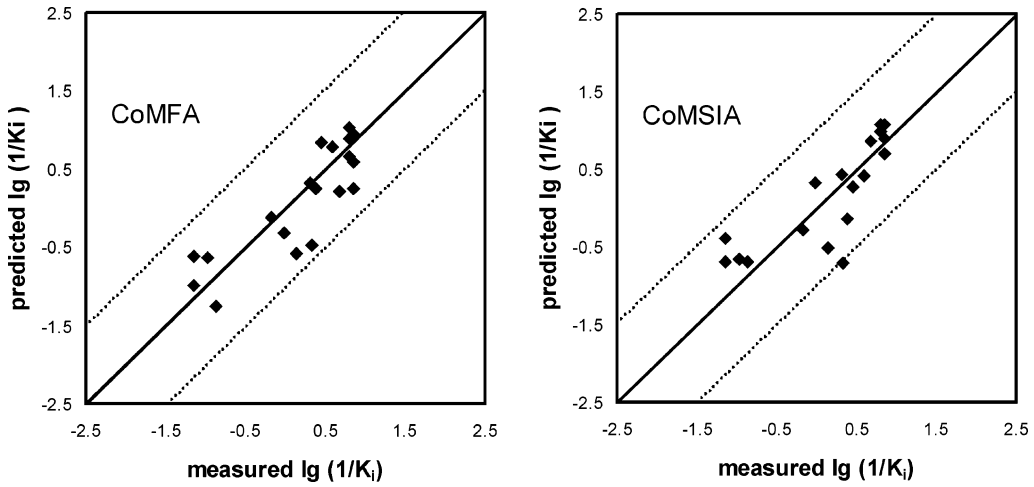


Figure 4. Predicted versus measured affinity constants for the test set of 19 compounds obtained by PLS analyses using CoMFA and CoMSIA with 2 Å grid spacing. Dashed lines denote deviations of one logarithmic unit.

are accordingly difficult to describe. For this reason, we will restrict the graphical interpretation of the models to CoMSIA field contributions to different physicochemical properties. However, it should be mentioned that the CoMFA model derived for the training set produced similar features for steric and electrostatic map properties compared to CoMSIA.

In Figure 5, the *steric field properties* are shown as contour maps. Green regions (denoted as **S1** and **S2**) correspond to areas where bulky group substitution is likely to enhance the affinity, while yellow contours (labeled as **S3**−**S5**) indicate regions where steric bulk is detrimental to binding affinity. The green contours imply the presence of two large cavities in the receptor binding site that accommodate the side chains of dipeptides. This finding is in agreement with our previous observation that bulky side chains of dipeptides (e.g., **1**−**3**, **7**−**12**, Table 1) increase affinity. In Figure 5a, this is demonstrated for two high-affinity substrates, i.e., **2** (gray) and **12** (purple), which fill those regions with their aromatic side chains.

Disfavored steric field contributions in the region **S3** may hold for the moderate to low affinities of Pro-Xaa dipeptides (**47**, **72**, **73**) and of ortho-substituted alanyl aryl amides (**39**, **51**, **53**, **64**, **70**). As shown exemplarily for **47** and **64** in Figure 5b, a portion of the pyrrolidine ring and the ortho-substituents lie in this region (marked by circles). A second smaller disfavored region (labeled as **S4**) is mainly occupied by the side chains or the N-terminal amino group of DL- and DD-dipeptides, as demonstrated in Figure 5b for D-Ala-Ala **48**. The third disfavored area marked with **S5** is occupied by the side chains of D-configured C-terminal amino acids (**54**, **69**, **76**, **79**), which show diminished binding affinity as compared to their L-analogues.

The contribution maps of *electrostatic fields* are depicted in Figure 6. They indicate regions where positive charges are favorable (or negative unfavorable, **E1**, blue) and where negative charges are favorable (or positive unfavorable, **E2**, red). The blue contour, **E1**, is related to the protonated amino group present in all natural dipeptides at physiological pH, while the red contour, **E2**, corresponds to the area where increased electron density results in a higher affinity.

Replacement of the N-terminal ammonium group by amide groups as in **60**, **75**, and **78** or its misplaced orientation for *β*-Ala-Ala **55** (Figure 6) or *β*-Ala-*β*-Ala **77** results in a notable to dramatic fall off in affinity. This indicates that a positively charged residue in this region is important for the PEPT1 affinity.

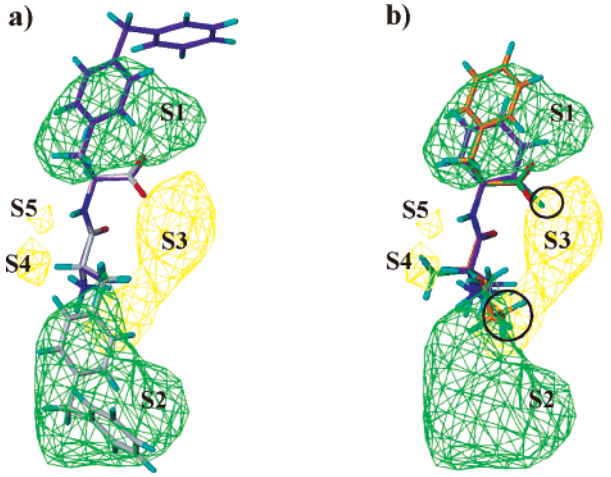


Figure 5. CoMSIA stdev*coeff contour plots for steric properties. Green isopleths indicate regions where bulky groups enhance affinity. Yellow isopleths are regions that should be kept unoccupied. (a) The high-affinity substrates **2** (gray) and **12** (purple) and (b) the low-affinity substrates **47** (orange), **48** (green) and **64** (purple) are displayed for demonstration purposes. Circles in **S3** denote substituents pointing into this disfavored area.

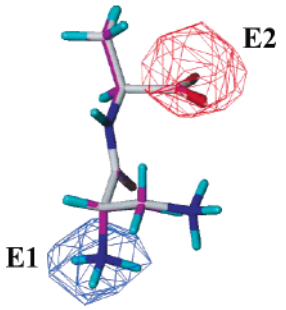


Figure 6. CoMSIA stdev*coeff contour plots for electrostatic properties. The blue contour denotes a region where positively charged groups enhance affinity, whereas the red contour indicates a region where negative charges increase affinity. The comparison of **14** (magenta) and **55** (gray) demonstrates the differences in the position of the positively charged N-terminus that may hold for the difference in binding affinity by a factor of 40.

The comparison of alanyl pipecolinic acid **4** and alanyl piperidide **68**, differing in their affinities by a factor of 215, illustrates the tremendous effect of a missing negative charge at the C-terminus. Interestingly, alanyl aryl amides show also reasonable affinities (**25**, **27**, **32**, **33**, **35**, **36**, **39**, **42**, and **51**−**53**). Here, the presence of the aromatic moiety increases the electron density in this region, which leads to appreciable affinities.

On the basis of electrostatic contour maps, it is possible to explain why only cyclic dipeptides carrying positively charged side chains (like lysine or ornithine) still show an appreciable affinity to PEPT1.[25] The positively charged amino group of the side chain coincides with the position of the N-terminal ammonium group of linear dipeptides. Moreover, one of the carbonyl oxygens of the diketopiperazine ring is situated in the same region where one oxygen atom of the carboxyl group at the C-terminus is placed.

The maps for *hydrophobic field contributions* are shown in Figure 7a−c. Lipophilic favorable regions are colored in orange (Figure 7a, labeled as **L1** and **L2**), and hydrophilic favorable regions are colored in cyan (Figure 7b, marked as **L3** and **L4**). The main hydrophilic contributions arise from the two polar regions at both tips of the molecule, which is in agreement with the electrostatic fields (Figure 7b,c). The lipophilic area **L1** in part coincides with the sterically favored region **S2**, thus suggesting that a lipophilic bulky side chain at the C-terminal end is advantageous for the binding affinity. Interestingly, there is no hydrophobic field contribution at the N-terminal amino acid, which indicates that both hydrophobic and hydrophilic side chains are comparably tolerated in this position.

In particular, lipophilic areas **L1** and **L2** enclose parts of the aromatic moieties of alanyl aryl amides, mainly their substituents such as phenyl, methyl, or chloride groups. This might explain why alanyl aryl amides carrying lipophilic meta and para substituents show noteworthy affinities (see Figure 7a, **25** and **33**), while their analogues carrying hydrophilic para substituents such as the carboxyl group in **66** possess lower affinities (Figure 7b,c). In contrast, ortho substitution by hydrophobic residues is unfavorable for affinity (e.g., **70**), whereas hydrophilic ortho substitution is advantageous (**39**, **51**).

In Figure 8, contribution maps of *hydrogen-bond-donor* fields are shown. Cyan areas belong to regions (denoted with **D1**−**D3**) where donor groups enhance binding affinity. In the case of the compounds under study, these regions are occupied by the hydrogen atoms of the N-terminal ammonium group or the ϵ-ammonium groups of diketopiperazines (vide supra) as shown exemplarily for compound **18** in Figure 8a. Replacement of one ammonium hydrogen by a methyl group is tolerated by PEPT1 (**46**). Contours colored in purple and labeled as **D4**−**D6** denote regions where hydrogen donors should not be present. Its occupancy by the imine hydrogens of Pro-Xaa dipeptides (**47**, **72**, **73**), amide hydrogens (**60**, **75**, **78**), or the ammonium groups of DL- and DD-isomers (**44**, **56**, **63**, **76**, **79**) generally results in low affinity. This is shown for compounds **73** and **78** in parts a and b of Figure 8, respectively.

In Figure 9, the contour maps of *hydrogen-bond-acceptor fields* are shown exemplarily for **14** and **54**. The region colored in magenta (labeled as **A1**) denotes the presence of hydrogen-bond-donors at the receptor site, while the region colored in red (marked for one example by **A2**) denotes that acceptor groups in the substrates have no counterparts in the receptor site. The acceptor fields **A1** and **A2** may explain the different affinity of LL-, LD-, and DD-isomers.[19,26] The C-terminal carboxylic group of LD- and DD-isomers is placed in the region **A2**, while the favorable region **A1** is occupied by the carboxyl group of LL- and DL-dipeptides, which is necessarily closely related to the contour field of high electron density (**E2** in Figure 6).

To proof the predictive power of our 3D QSAR model, we tried to predict affinities of substrates that were neither included in the training nor in the test set (Table 4). Compounds were first predicted and afterward measured in our laboratory. Deviations of affinities were not higher than ±0.57 logarithmic units. The highest deviation was observed for the dipeptide **100**, which bears valine at the N-terminal and the less hydrophobic amino acid alanine at the C-terminal position, while

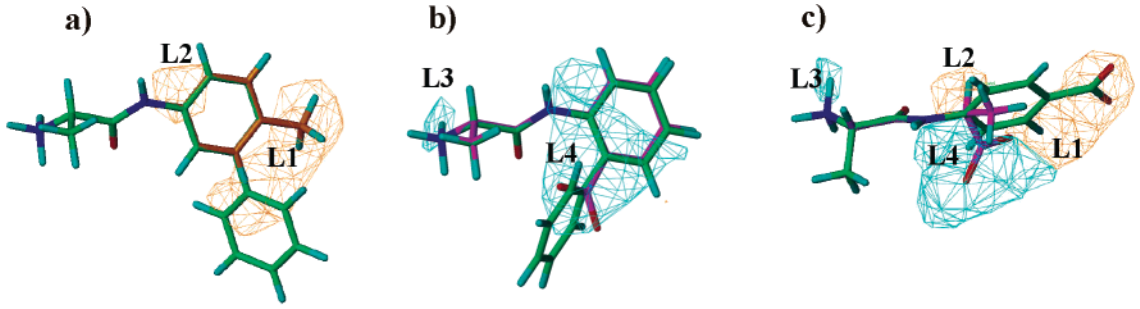


Figure 7. CoMSIA stdev*coeff contour plots showing lipophilic and hydrophilic properties. The orange contours (**L1** and **L2**) indicate regions where lipophilic groups enhance affinity, whereas cyan contours (**L3** and **L4**) indicate regions where hydrophilic groups increase affinity. (a) **25** (green) and **33** (orange), (b) **51** (purple) and **70** (green), (c) **14** (magenta) and **66** (green). In parts a and b, the fields of lipophilic and hydrophilic properties were separated for clarity.

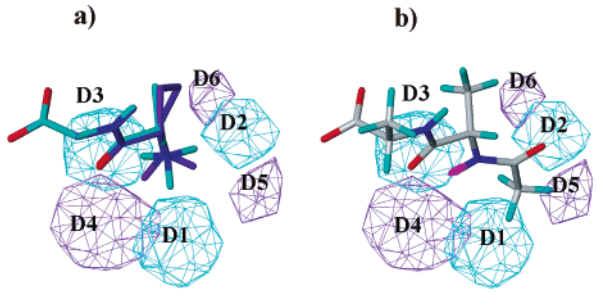


Figure 8. CoMSIA stdev*coeff contour plots for hydrogen-bond-donor fields. Cyan isopleths (**D1**−**D3**) enclose regions where hydrogen-bond-donor groups should be directed toward acceptors at the receptor site, while purple areas (**D4**−**D6**) indicate regions where donor groups decrease affinity. (a) **18** (cyan) and **73** (purple). (b) **78**. Hydrogen atoms directed toward disfavored regions are colored in purple.

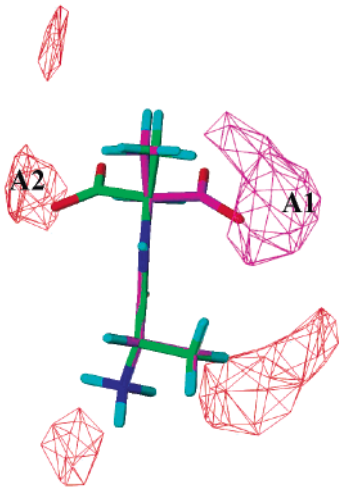


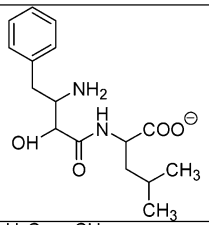
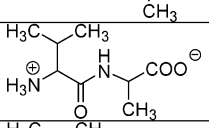
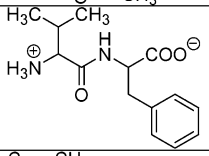
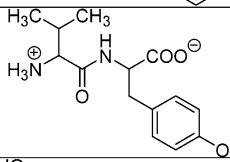
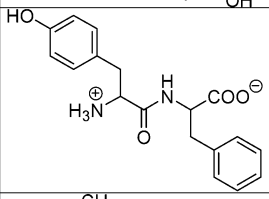
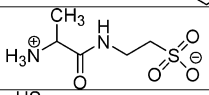
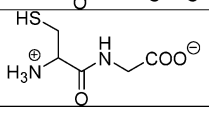
Figure 9. CoMSIA stdev*coeff contour plots for the hydrogen-bond-acceptor properties. The magenta isopleth marks the region favored by acceptor groups, while the red isopleths denote areas in which acceptor groups have no counterparts at the receptor site. Compounds **14** (magenta) and **54** (green) have been chosen to illustrate where differences in the position of the carboxyl group in the substrate affect the binding affinity.

predictions of hydrophobic, valine-containing dipeptides as **101** and **102** yield better correlations. The deviation might be explained by some special peculiarities of the natural system, e.g., hydrophobicity, that receive much greater influence as in silico methods can propose.

In general, using CoMSIA contour maps the affinity differences of substrates of PEPT1 can be reasonably explained. However, we are not able to explain all binding differences.

Finally, it should be pointed out that the CoMSIA method correlates and compares only similarities of molecules to each other and no absolute values. In this respect, one has to be aware that all the features derived from the QSAR studies are only a mirror of the structural variation of the data set.

Table 4. Predicted and Measured Affinity Constants of Substrates Not Included in the Training or the Test Set

No.	Compound	$K_{i\,pred}$ [a] [mM]	$K_{i\,meas}$ [b] [mM]	residuals [c]
99		0.55	0.40	0.14
100		0.33	0.09	0.57
101		0.08	0.05	0.20
102		0.10	0.10	0.00
103		0.04	0.09	-0.35
104		1.0	1.2	0.08
105		0.22	0.20	0.04

[a] Affinity constant predicted by CoMSIA using a 2 Å grid space. [b] Measured affinity constant. [c] Residuals are given as deviations from the measured value (log $1/K_{i\,meas}$ − log $1/K_{i\,pred}$).

Conclusions

In the present study, it has been demonstrated that 3D QSAR models are useful tools to explain different affinities of dipeptides and their derivatives to the intestinal peptide transporter PEPT1. Both methods, CoMFA and CoMSIA, yield similar results. By combining the five CoMSIA contour maps, i.e., steric, electrostatic, hydrophobic/hydrophilic, hydrogen-bond-donor, and hydrogen-bond-acceptor, we have identified six recognition elements that are favorable for binding to PEPT1: (i) the presence of bulky side chains, (ii) a positively charged N-terminus and a region of high electron density at the C-terminus, which are commensurate with (iii) two hydrophilic regions, (iv) a preferred hydrophobic region at the C-terminal part, (v) a hydrogen-bond-donor region at the N-terminus, and (vi) a hydrogen-bond-acceptor region crucial for differentiation between LL-, DL-, LD-, and DD-isomers. The present 3D QSAR models allow the prediction of K_i-values of new compounds.

Methods

Data Sets. For the computational investigations we selected a training set of 79 diverse dipeptides and dipeptide derivatives covering a range of K_i values from 0.02 to 100 mM (Table 1) whose affinity constants for PEPT1 were measured in our laboratory. The validation of the model was carried out using a test set of 19 compounds (Table 2), again tested in our laboratory for their affinity toward PEPT1.

Biological Data. The biological data were obtained using the intestinal Caco-2 cell line expressing PEPT1. Caco-2 cells were obtained from the German Collection of Microorganisms and Cell Cultures and cultured as described previously.[27,28] Uptake of [^{14}C]Gly-Sar (specific radioactivity 53 mCi mmol^{-1}, Amersham International, UK) was measured on the 6th to 7th day after cells reached confluence.[27,28] The uptake buffer (1 mL) contained 25 mM Mes/Tris (pH 6.0), 140 mM NaCl, 5.4 mM KCl, 1.8 mM $CaCl_2$, 0.8 mM $MgSO_4$, 5 mM glucose, 10 μM [^{14}C]Gly-Sar, and increasing concentrations (0−31.6 mM) of unlabeled substrates (Sigma-Aldrich Chemie, Bachem, or synthesized according to standard procedures in peptide chemistry[29]). Results are given as the mean ($n = 4$). Nonlinear regression analysis, calculation of inhibition constants (K_i) from IC_{50} values, and statistical analysis were carried out as described earlier.[27,28]

Conformational Search and Molecular Alignment. All molecular modeling studies were performed using the Sybyl program package[30] running on a SGI Octane2 R12000 workstation.

The starting structures were built with charged C- and N-termini as well as trans peptide bonds. Partial atomic charges were calculated according to the Gasteiger−Marsili method and energies were computed with the standard Tripos force field using a dielectric constant of $\epsilon = 80$ for the scaling of electrostatic interactions to mimic the aqueous environment.[31,32] The structures were submitted to a systematic search in which all rotatable bonds were allowed to vary in 30° steps. The energy cutoff was set to 10 kcal mol^{-1} above the lowest energy structure. The outcome of the conformational search was filtered with a search algorithm for local minima in torsional space.[33] The remaining structures were subsequently minimized using the Powell minimizer until the rms (root-mean-square) gradient was less than 10^{-3} kcal mol^{-1} Å^{-1}. Structures that converged to the same minimum were omitted on the basis of their energies and by a rms fitting procedure in which all non-hydrogen atoms were superimposed using 0.3 Å as similarity cutoff. The remaining conformers were stored in conformational databases.

The derivation of the pharmacophore model was performed with the DISCO[23] module of Sybyl using **14** as reference structure (see the Results and Discussion), with a maximum tolerance up to 0.5 Å, and requiring the models to have 3−12 matched features. In cases where more than a single conformer fitted to the specific distances, the lower energy conformer was chosen. The set of superimposed structures was stored in a molecule database for subsequent 3D QSAR analyses.

3D QSAR Analyses. All CoMFA calculations were performed using the Advanced CoMFA module implemented in Sybyl. Steric and electrostatic fields were calculated using Lennard-Jones and Coulomb potentials with a dielectric constant $\epsilon = 1/r$ and a nonbonded cutoff value of 30 kcal mol^{-1}. Hydrogen-bonding fields were created by assigning energies equal to the steric cutoff energy to lattice points that are close to H-bond-accepting or -donating atoms.[34] Indicator fields were used to convert continuous data to discrete results.[35] A sp^3 carbon atom was used as a probe atom with a charge of +1.

In CoMSIA analyses five physicochemical properties (steric, electrostatic, hydrophobic, and hydrogen-bond donor and acceptor) were examined, using a sp^3 carbon atom with an 1 Å radius and a charge, hydrophobicity, and hydrogen-bond property of +1. A Gaussian-type distance dependence was considered between the grid point q and each atom i of the molecule. The value of the attenuation factor α was set to 0.3, as described by Böhm et al.[36] For both methods, lattices with a grid spacing of 1 and 2 Å and equal box dimensions with a sufficiently large margin (>4 Å) were considered.

PLS analyses were performed by following the standard implementation in Sybyl. The descriptors were scaled to each other by the standard scaling option. The statistical significance of the 3D QSAR models was checked using the cross-validation procedure "leave-one-out" (LOO) and the enhanced version of PLS, the SAMPLS method.[37] The optimum number of components was determined by selecting the smallest standard error of prediction and the highest q^2 value. Subsequently, this value was used in the final PLS run without cross-validation to estimate r^2. The minimum σ value (column filter) for CoMSIA was set by considering about 10% of the variables in the PLS analyses (usually between 1.5 and 2).

Acknowledgment. This work was supported by a grant from the Land Sachsen-Anhalt (Grant 2880A/0028G) and by the Fonds der Chemischen Industrie.

Supporting Information Available: A table showing the residuals of the predictions of the training and test set by CoMFA and CoMSIA at 2 Å grid spacing. This material is available free of charge via Internet at http://pubs.acs.org.

References

(1) Meredith, D.; Boyd, C. A. Structure and function of eukaryotic peptide transporters. *Cell. Mol. Life Sci.* **2000**, *57*, 754−758.
(2) Fei, Y. J.; Ganapathy, V.; Leibach, F. H. Molecular and structural features of the proton-coupled oligopeptide transporter superfamily. *Prog. Nucleic Acid Res. Mol. Biol.* **1998**, *58*, 239−261.
(3) Rubio-Aliaga, I.; Daniel, H. Mammalian peptide transporters as targets for drug delivery. *Trends Pharmacol. Sci.* **2002**, *23*, 434−440.
(4) Bretschneider, B.; Brandsch, M.; Neubert, R. Intestinal transport of beta-lactam antibiotics: Analysis of the affinity at the H$^+$/peptide symporter (PEPT1), the uptake into Caco-2 cell monolayers and the transepithelial flux. *Pharm. Res.* **1999**, *16*, 55−61.
(5) Moore, V. A.; Irwin, W. J.; Timmins, P.; Lambert, P. A.; Chong, S.; Dando, S. A.; Morrison, R. A. A rapid screening system to determine drug affinities for the intestinal dipeptide transporter 2: Affinities of ACE inhibitors. *Int. J. Pharm.* **2000**, *210*, 29−44.
(6) Ganapathy, M. E.; Huang, W.; Wang, H.; Ganapathy, V.; Leibach, F. H. Valacyclovir: A substrate for the intestinal and renal peptide transporters PEPT1 and PEPT2. *Biochem. Biophys. Res. Commun.* **1998**, *246*, 470−475.
(7) Nakanishi, T.; Tamai, I.; Takaki, A.; Tsuji, A. Cancer cell-targeted drug delivery utilizing oligopeptide transport activity. *Int. J. Cancer* **2000**, *88*, 274−280.
(8) Lee, H. J.; Amidon, G. L. Oral peptide delivery: Improving the systemic availability of small peptides and enkephalin analogues. *NIDA Res. Monogr.* **1995**, *154*, 86−106.
(9) Meredith, D.; Temple, C. S.; Guha, N.; Sword, C. J.; Boyd, C. A.; Collier, I. D.; Morgan, K. M.; Bailey, P. D. Modified amino acids and peptides as substrates for the intestinal peptide transporter PEPT1. *Eur. J. Biochem.* **2000**, *267*, 3723−3728.

(10) Brandsch, M.; Knütter, I.; Thunecke, F.; Hartrodt, B.; Born, I.; Borner, V.; Hirche, F.; Fischer, G.; Neubert, K. Decisive structural determinants for the interaction of proline derivatives with the intestinal H+/peptide symporter. *Eur. J. Biochem.* **1999**, *266*, 502−508.
(11) Döring, F.; Will, J.; Amasheh, S.; Clauss, W.; Ahlbrecht, H.; Daniel, H. Minimal molecular determinants of substrates for recognition by the intestinal peptide transporter. *J. Biol. Chem.* **1998**, *273*, 23211−23218.
(12) Meredith, D.; Boyd, C. A.; Bronk, J. R.; Bailey, P. D.; Morgan, K. M.; Collier, I. D.; Temple, C. S. 4-aminomethylbenzoic acid is a nontranslocated competitive inhibitor of the epithelial peptide transporter PEPT1. *J. Physiol.* **1998**, *512*, 629−634.
(13) Bailey, P. D.; Boyd, C. A.; Bronk, J. R.; Collier, I. D.; Meredith, D.; Morgan, K. M.; Temple, C. S. How to Make Drugs Orally Active: A Substrate Template for Peptide Transporter PEPT1. *Angew. Chem., Int. Ed. Engl.* **2000**, *39*, 505−508.
(14) Swaan, P. W.; Koops, B. C.; Moret, E. E.; Tukker, J. J. Mapping the binding site of the small intestinal peptide carrier (PEPT1) using comparative molecular field analysis. *Receptors Channels* **1998**, *6*, 189−200.
(15) Fei, Y. J.; Liu, W.; Prasad, P. D.; Kekuda, R.; Oblak, T. G.; Ganapathy, V.; Leibach, F. H. Identification of the histidyl residue obligatory for the catalytic activity of the human H^+/peptide cotransporters PEPT1 and PEPT2. *Biochemistry* **1997**, *36*, 452−460.
(16) Döring, F.; Dorn D.; Bachfischer, U.; Amasheh, S.; Herget, M.; Daniel, H. Functional analysis of a chimeric mammalian peptide transporter derived from the intestinal and renal isoforms. *J. Physiol.* **1996**, *497*, 773−779.
(17) Döring, F.; Walter, J.; Will, J.; Focking, M.; Boll, M.; Amasheh, S.; Clauss, W.; Daniel H. Delta-aminolevulinic acid transport by intestinal and renal peptide transporters and its physiological and clinical implications. *J. Clin. Invest.* **1998**, *101*, 2761−2767.
(18) Brandsch, M.; Thunecke, F.; Küllertz, G.; Schutkowski, M.; Fischer, G.; Neubert, K. Evidence for the absolute conformational specificity of the intestinal H+/peptide symporter, PEPT1. *J. Biol. Chem.* **1998**, *273*, 3861−3864.
(19) Hartrodt, B.; Theis, S.; Knütter, I.; Börner, V.; Born, I.; Brandsch, M.; Daniel, H.; Neubert, K. New insights into stereospecificity of the intestinal H+/peptide symporter. In *Peptides 2000, Proceedings 26th Europ. Peptide Symposium*; Martinez, J., Fehrentz, J.-A., Eds.; EDK Paris, France, 2001; pp 983−984.
(20) Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959−5967.
(21) Klebe, G.; Abraham, U.; Mietzner, T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *J. Med. Chem.* **1994**, *37*, 4130−4146.
(22) Nicolotti, O.; Pellegrini-Calace, M.; Carrieri, A.; Altomare, C.; Centeno, N. B.; Sanz, F.; Carotti, A. Neuronal nicotinic receptor agonists: A multi-approach development of the pharmacophore. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 859−872.
(23) Martin, Y. C.; Bures, M. G.; Danaher, E. A.; DeLazzer, J.; Lico, I.; Pavlik, P. A fast new approach to pharmacophore mapping and its application to dopaminergic and benzodiazepine agonists. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 83−102.
(24) Klebe, G. Comparative molecular similarity indices analysis: CoMSIA. *Perspect. Drug Discovery Des.* **1998**, *12*, 87−104.
(25) Other diketopiperazines carrying either positively charged side chains or neutral groups do not show any appreciable binding affinity. Nevertheless, they have been included in the DISCO studies; however, they did not yield any features and have, therefore, been omitted from the 3D QSAR analyses.
(26) Li, J.; Tamura, K.; Lee, C. P.; Smith, P. L.; Borchardt, R. T.; Hidalgo, I. J. Structure-affinity relationships of Val-Val and Val-Val-Val stereoisomers with the apical oligopeptide transporter in human intestinal Caco-2 cells. *J. Drug Target* **1998**, *5*, 317−327.
(27) Brandsch, M.; Miyamoto, Y.; Ganapathy, V.; Leibach, F. H. Expression and protein kinase C-dependent regulation of peptide/H^+ co-transport system in the Caco-2 human colon carcinoma cell line. *Biochem. J.* **1994**, *299*, 253−260.
(28) Börner, V.; Fei, Y. J.; Hartrodt, B.; Ganapathy, V.; Leibach, F. H.; Neubert, K.; Brandsch, M. Transport of amino acid aryl amides by the intestinal H^+/peptide cotransport system, PEPT1. *Eur. J. Biochem.* **1998**, *255*, 698−702.
(29) Goodman, M.; Felix, A.; Moroder, L.; Toniolo, C. Synthesis of Peptides and Peptidomimetics. In *Houben-Weyl Methods of Organic Chemistry*; Georg Thieme: Stuttgart, New York, 2002.
(30) SYBYL 6.7 and 6.8; Tripos Associates, Inc., 1699 South Hanley Rd., St. Louis, MO 63144.
(31) Gasteiger, J.; Marsili, M. Iterative partial equalization of orbital electronegativity: A rapid access to atomic charges. *Tetrahedron* **1980**, *36*, 3219−3222.
(32) Clark, M.; Cramer, R. D., III; Van Opdenbosch, N. The Tripos Field, *J. Comput. Chem.* **1989**, *10*, 982−1012.
(33) Schinke, H. Ph.D. Dissertation, Martin-Luther-University Halle-Wittenberg, Halle, Germany, 1995.
(34) Bohacek, R. S.; McMartin, C. Definition and display of steric, hydrophobic, and hydrogen-bonding properties of ligand binding sites in proteins using Lee and Richards accessible surfaces: Validation of a high-resolution tool for drug design. *J. Med. Chem.* **1992**, *35*, 1671−1684.
(35) Krömer, R. T.; Hecht, P. Replacement of steric 6−12 potential-derived interaction energies by atom-based indicator variables in CoMFA leads to models of higher consistency. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 205−212.
(36) Böhm, M.; Stürzebecher, J.; Klebe, G. Three-dimensional quantitative structure−activity relationship analyses using comparative molecular field analysis and comparative molecular similarity indices analysis to elucidate selectivity differences of inhibitors binding to trypsin, thrombin, and factor Xa. *J. Med. Chem.* **1999**, *42*, 458−477.
(37) Bush, B. L.; Nachbar, R. B., Jr. Sample-distance partial least squares: PLS optimized for many variables, with application to CoMFA. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 587−619.

JM030976X