

# Quantitative Insight into the Design of Compounds Recognized by the L-Type Amino Acid Transporter 1 (LAT1)

Henna Ylikangas,<sup>[a]</sup> Kalle Malmioja,<sup>[a]</sup> Lauri Peura,<sup>[a]</sup> Mikko Gynther,<sup>[a]</sup> Emmanuel O. Nwachukwu,<sup>[a]</sup> Jukka Leppänen,<sup>[a]</sup> Krista Laine,<sup>[a]</sup> Jarkko Rautio,<sup>[a]</sup> Maija Lahtela-Kakkonen,<sup>[a]</sup> Kristiina M. Huttunen,<sup>[a]</sup> and Antti Poso<sup>\*,[a, b]</sup>

L-Type amino acid transporter 1 (LAT1) is a transmembrane protein expressed abundantly at the blood–brain barrier (BBB), where it ensures the transport of hydrophobic acids from the blood to the brain. Due to its unique substrate specificity and high expression at the BBB, LAT1 is an intriguing target for carrier-mediated transport of drugs into the brain. In this study, a comparative molecular field analysis (CoMFA) model with considerable statistical quality ( $Q^2=0.53$ ,  $R^2=0.75$ ,  $Q^2SE=0.77$ ,  $R^2SE=0.57$ ) and good external predictivity ( $CCC=0.91$ ) was

generated. The model was used to guide the synthesis of eight new prodrugs whose affinity for LAT1 was tested by using an in situ rat brain perfusion technique. This resulted in the creation of a novel LAT1 prodrug with L-tryptophan as the promoiety; it also provided a better understanding of the molecular features of LAT1-targeted high-affinity prodrugs, as well as their promoiety and parent drug. The results obtained will be beneficial in the rational design of novel LAT1-binding prodrugs and other compounds that bind to LAT1.

## Introduction

L-Type amino acid transporter 1 (LAT1) is a sodium-independent heterodimeric transmembrane protein. LAT1 is found in several peripheral tissues, but is expressed at 100-fold higher levels in the blood–brain barrier (BBB) than in peripheral tissues (e.g., placenta, retina, and gut).<sup>[1,2]</sup> The natural substrates of LAT1 are large neutral amino acids such as L-leucine, L-tryptophan, and L-phenylalanine. These compounds have excellent affinity for the LAT1 transporter protein and thus rapidly pass through the BBB.<sup>[3]</sup> However, LAT1 also transports thyroid hormones and amino acid derived drug molecules such as levodopa,<sup>[4]</sup> gabapentin,<sup>[5]</sup> and baclofen.<sup>[6]</sup> Its distinctive substrate specificity and relatively high expression at the BBB make LAT1 an interesting target for drug delivery to the central nervous system (CNS).

Natural substrates of LAT1 can be used as potential promoieties for LAT1-targeted prodrugs in order to improve drug delivery to the brain. In general, the substrates should have unsubstituted carboxylic acid and amine functionalities, and the parent drug should be attached to the side chain of amino acids in order to maintain efficient LAT1 binding.<sup>[7–9]</sup> These kinds of prodrugs have been reported earlier.<sup>[8,10–14]</sup> However,

previous studies have focused mostly on a few amino acids and their derivatives in clinical use.<sup>[3,9,15]</sup> Little information is available on the structural and chemical features that influence the molecular size and flexibility of the prodrug or the preferred position of the parent drug with respect to the promoiety essential for achieving LAT1-binding prodrugs with high affinities.

To design better LAT1 prodrugs, it is of the utmost importance to gain a good understanding of the three-dimensional (3D) structure–activity relationships of LAT1-binding compounds. For this reason, we were interested in analyzing the 3D quantitative structure–activity relationships (3D QSAR) of LAT1-binding compounds.

In this study, the first 3D QSAR study of LAT1-binding compounds was carried out by using classical and topomer comparative molecular field analysis (CoMFA).<sup>[16,17]</sup> The models are based on biological data determined by an in situ rat brain perfusion technique, which reveals the ability of the compounds to inhibit the brain uptake of the LAT1-selective substrate, [<sup>14</sup>C]L-leucine.<sup>[18]</sup> By using the information obtained from the CoMFA models, it was possible to design and synthesize eight new prodrugs. The model predicted the affinities of the synthesized compounds with reasonable evidence that the topomer CoMFA model can be used in the rational design of new LAT1 prodrugs. Herein we discuss the quantitative insights obtained from the CoMFA model and propose new concepts for LAT1-binding compounds and prodrugs, which can be useful for improved drug delivery to the brain. As far as we are aware, these novel 3D QSAR models are the first to be published for compounds that target LAT1.

[a] H. Ylikangas, K. Malmioja, Dr. L. Peura, Dr. M. Gynther, E. O. Nwachukwu, Dr. J. Leppänen, Dr. K. Laine, Prof. J. Rautio, Dr. M. Lahtela-Kakkonen, Dr. K. M. Huttunen, Prof. A. Poso  
School of Pharmacy, University of Eastern Finland  
P.O. Box 1627, 70211 Kuopio (Finland)  
E-mail: antti.poso@uef.fi

[b] Prof. A. Poso  
University Hospital Tübingen, Department of Internal Medicine 1  
Division of Translational Gastrointestinal Oncology  
Otfried-Müller-Strasse 10, 72076 Tübingen (Germany)

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cmdc.201402281>.

## Results and Discussion

## CoMFA models

In general, prodrugs can be divided into two main sections: the parent drug and the promoiety, which are connected to each other by a cleavable prodrug bond. Following this nomenclature, these sections in the CoMFA coefficient maps were analyzed based on the natural substrates of LAT1 and previously reported LAT1 prodrugs (Table 1).

The data series used in CoMFA models consist of 47 compounds which were selected based on their relatively diverse structures in order to obtain as much information on 3D QSAR as possible. The training set (Table 1) consists of 39 and an external test set of eight compounds with varying affinities for LAT1 (0–100% L-leucine uptake inhibition). Their ability to inhibit L-leucine uptake into the brain was used as a measure of biological activity, as it describes the ability of the investigated compounds to replace competing substrates, that is, their affinity for LAT1.<sup>[8]</sup> High affinity is a prerequisite for efficient

uptake into the brain; this makes it a reliable method for determining biological activity, although as such, it does not directly guarantee uptake into the brain.

In addition to our previously published compounds<sup>[7,8,11,19]</sup> and commercial compounds (Table 1), the training set includes one novel prodrug (compound 26), and an amino acid derivative (39) which were designed as a continuum to our previously published prodrugs in order to broaden the chemical space presented so far. The detailed synthesis procedures of these compounds are described in the Supporting Information.

Six 3D QSAR models were generated, and from those, one model with two components was selected and used in guiding the synthesis of the new LAT1 substrates. The model was selected based on results obtained from partial least squares (PLS) and progressive scrambling stability analyses (topomer CoMFA model E: Table 2). The topomer CoMFA model revealed that the steric component explains most of the variance in biological activity of LAT1-binding compounds. The contribution of steric interactions was stronger than the effect of electrostatics for both topomers, R1: amino acid terminal and R2: side chain, prodrug bond, and parent drug, respectively (Figure 1).

**Table 1.** Training set used in CoMFA models with percent L-leucine uptake inhibition determined by the in situ rat brain perfusion technique.

Compd (Inh. [%]) <sup>[a]</sup>	Structure	Compd (Inh. [%]) <sup>[a]</sup>	Structure
1 (99±4)		2 (94±2)	
3 (93±2)		4 (93±3)	
5 (92±1)		6 (88±3)	
7 (86±2)		8 (86±4)	
9 (83±2)		10 (81±1)	
11 (79±2)		12 <sup>[b]</sup> (79±6)	
13 (75±2)		14 (66±1)	
15 (66±5)		16 (64±3)	
17 (64±4)		18 (63±7)	
19 (59±4)		20 (57±20)	
21 (56±1)		22 (39±6)	
23 (38±5)		24 (36±7)	

Table 1. (Continued)

Compd (Inh. [%]) <sup>[a]</sup>	Structure	Compd (Inh. [%]) <sup>[a]</sup>	Structure
<b>25</b> (35±11)		<b>26</b> <sup>[c]</sup> (33±6)	
<b>27</b> (31±1)		<b>28</b> (25±7)	
<b>29</b> (23±3)		<b>30</b> (21±7)	
<b>31</b> (8±5)		<b>32</b> (5±10)	
<b>33</b> <sup>[d]</sup> (0±0)		<b>34</b> <sup>[d]</sup> (0±0)	
<b>35</b> <sup>[d]</sup> (0±0)		<b>36</b> <sup>[d]</sup> (0±0)	
<b>37</b> <sup>[d]</sup> (0±0)		<b>38</b> <sup>[d]</sup> (0±0)	
<b>39</b> <sup>[d]</sup> (0±0)			

[a] [<sup>14</sup>C]-leucine brain capillary surface area (PA) percent inhibition data are the average ± SD of *n* = 3 experiments conducted at a compound concentration of 100 μM. [b] Measured at 70 μM compound concentration. [c] *n* = 4. [d] *n* = 2.

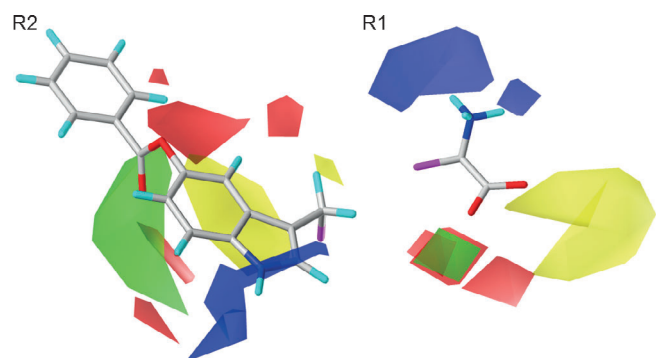
## Model analysis

Our CoMFA model indicates that adding steric moieties that occupy regions beyond the amino acid terminal decreases affinity for LAT1 (Figure 1: yellow contour in R1). These findings are in agreement with our previous studies.<sup>[20]</sup> The model also revealed that increasing the positive charge near the amine function (blue contours at R1) can be beneficial for affinity. In addition, adding a negative charge near the carboxylic acid can increase LAT1 affinity (red cubic-shaped contours at R1). So far, all the tested compounds that fulfill the electrostatic features of the model but do not contain an amino acid functionality (**28**, **30**, and **32**) have been rather poor inhibitors (i.e., < 25 % of L-leucine uptake into the brain; Table 1).

The model shows that the promoiety for LAT1 prodrugs should be relatively planar, because large, branched substituents can reach areas that can decrease affinity (yellow contour in R2). Moreover, the addition of steric substitutions near the 5- and 6-positions of L-tryptophan and its derivatives can be beneficial for efficient binding (Figure 1: green crescent-shaped

contour at R2). These regions correspond roughly to the 3- and 4-positions of L-phenylalanine and indicate positions where one can attach the parent drug in the promoiety. This is in agreement with results of previous studies.<sup>[7,15]</sup>

The most efficient valproic acid prodrugs in the training set (**3** and **4**; > 90 % L-leucine uptake inhibition) contain the parent drug substituted at the 3-position of L-phenylalanine.<sup>[8]</sup> These correspond better with steric features revealed by the CoMFA model (green crescent-shaped contour at R2) than prodrugs **16** and **17** (64 % L-leucine uptake inhibition). Prodrugs **16** and **17** contain parent drugs at the 4-position and are less efficient in inhibiting L-leucine uptake than **3** and **4** (Table 1). The CoMFA model shows that prodrugs with parent drug at position 4 (e.g., **1**, **16**, and **17**) do not occupy sterically disfavored areas. However, their orientations do not reach the sterically favored regions described by the topomer CoMFA model either (green contours in R2). In addition to the position of the parent drug, the CoMFA model indicates that relatively rigid and large prodrugs bind more efficiently to LAT1 than more



**Figure 1.** Topomer CoMFA model presented with L-tryptophan prodrug **40** (99 % L-leucine uptake inhibition). Green and red represent areas where adding steric features and negative charge or hydrogen bond acceptors are favored, respectively. Blue represents areas where more positive charge or hydrogen bond donors are favored, whereas yellow contours designate sterically disfavored regions. R1 and R2 indicate the common core; amino acid function and variable topomer; side chain and parent drug.

**Table 2.** Progressive scrambling stability of the six CoMFA models generated for LAT1-binding compounds represented as mean  $\pm$  SD of 20 independent scrambling tests.

Model	$n^{[a]}$	$N^{[b]}$	$Q_s^{*2[c]}$	$Q_0^{*2[d]}$	$SDEP_s^{*[e]}$	$SDEP_0^{*[f]}$	$dq^2/dr_{yy}^{2[g]}$	$CCC^{[h]}$
A	34	1	0.415 $\pm$ 0.013	0.488 $\pm$ 0.01	0.872 $\pm$ 0.01	0.836 $\pm$ 0.01	0.237 $\pm$ 0.08	0.813
		2	0.460 $\pm$ 0.02	0.541 $\pm$ 0.03	0.847 $\pm$ 0.02	0.803 $\pm$ 0.03	0.408 $\pm$ 0.12	
		3	0.421 $\pm$ 0.03	0.494 $\pm$ 0.03	0.892 $\pm$ 0.02	0.853 $\pm$ 0.03	0.648 $\pm$ 0.18	
B	32	2	0.523 $\pm$ 0.02	0.616 $\pm$ 0.02	0.747 $\pm$ 0.01	0.668 $\pm$ 0.01	0.819 $\pm$ 0.13	0.696
		3	0.534 $\pm$ 0.02	0.628 $\pm$ 0.02	0.753 $\pm$ 0.01	0.961 $\pm$ 0.08	0.961 $\pm$ 0.08	
		4	0.465 $\pm$ 0.03	0.547 $\pm$ 0.03	0.820 $\pm$ 0.02	0.751 $\pm$ 0.02	1.135 $\pm$ 0.18	
C	33	2	0.450 $\pm$ 0.01	0.529 $\pm$ 0.01	0.833 $\pm$ 0.01	0.781 $\pm$ 0.01	0.868 $\pm$ 0.12	0.569
		3	0.450 $\pm$ 0.02	0.529 $\pm$ 0.02	0.845 $\pm$ 0.01	0.792 $\pm$ 0.01	1.139 $\pm$ 0.15	
		4	0.349 $\pm$ 0.02	0.411 $\pm$ 0.03	0.935 $\pm$ 0.02	0.896 $\pm$ 0.02	1.108 $\pm$ 0.23	
D	36	1	0.304 $\pm$ 0.01	0.358 $\pm$ 0.02	0.925 $\pm$ 0.01	0.904 $\pm$ 0.01	0.517 $\pm$ 0.61	1.00
		2	0.342 $\pm$ 0.01	0.402 $\pm$ 0.02	0.912 $\pm$ 0.01	0.886 $\pm$ 0.01	0.726 $\pm$ 0.12	
		3	0.321 $\pm$ 0.02	0.377 $\pm$ 0.02	0.939 $\pm$ 0.01	0.770 $\pm$ 0.20	0.770 $\pm$ 0.20	
E	33	1	0.382 $\pm$ 0.01	0.450 $\pm$ 0.01	0.868 $\pm$ 0.01	0.783 $\pm$ 0.10	0.783 $\pm$ 0.10	0.907
		2	0.436 $\pm$ 0.01	0.513 $\pm$ 0.01	0.843 $\pm$ 0.01	0.803 $\pm$ 0.01	0.922 $\pm$ 0.11	
		3	0.386 $\pm$ 0.02	0.454 $\pm$ 0.03	0.892 $\pm$ 0.02	0.846 $\pm$ 0.02	1.11 $\pm$ 0.12	
F	35	2	0.377 $\pm$ 0.02	0.443 $\pm$ 0.02	0.901 $\pm$ 0.01	0.870 $\pm$ 0.02	0.642 $\pm$ 0.11	0.823
		3	0.380 $\pm$ 0.03	0.447 $\pm$ 0.03	0.912 $\pm$ 0.02	0.880 $\pm$ 0.03	0.672 $\pm$ 0.19	
		4	0.313 $\pm$ 0.04	0.368 $\pm$ 0.04	0.976 $\pm$ 0.02	0.902 $\pm$ 0.03	0.765 $\pm$ 0.23	

[a] Number of ligands in the training set. [b] Number of PLS components. [c] Predictivity at the critical threshold level of perturbation  $s$  (0.85); maximum value of  $Q_s^{*2}$ . [d]  $Q_0^{*2} = Q_s^{*2}s$ ; adjusted  $Q_s^{*2}$ , corresponding to the value expected for an unperturbed, non-redundant  $Q^2$ . [e] Standard error of prediction at the critical threshold level of perturbation  $s$ . [f]  $SDEP_0^* = \{[(2-s)(n-N-1)(SDEP_s)^2 - (1-s)(n-1)SD_y^2]/(n-N-1)\}^{1/2}$ ;  $N$  = number of compounds in the training set;  $SD_y = 1.1$  (response standard deviation); adjusted  $SDEP_s^*$ , corresponding to the value expected for an unperturbed, non-redundant model. [g] Sensitivity to perturbation. [h] Concordance correlation coefficient.<sup>[20,21]</sup>

flexible prodrugs. For example, the ketoprofen prodrug with a derivative of L-phenylalanine as a promoiety (compound **1**) inhibits >90% of L-leucine uptake, although it is substituted at position 4. Prodrug **12**, which has a more flexible promoiety (L-lysine), inhibits L-leucine uptake by almost 20 percentage units less than prodrug **1** (79% L-leucine uptake inhibition). It is possible that the large aromatic portion of **1** allows the prodrug to adopt a conformation beneficial for LAT1 binding. Similarly, dopamine prodrug **8** inhibits at 86%, whereas the more flexible prodrugs **22** and **25** inhibit <40% of L-leucine uptake into the brain. Based on the CoMFA model, it seems that the flexible prodrugs protrude into sterically disfavored areas (yellow contours at R2) which can decrease their affinity for LAT1. This finding was further studied with the newly synthesized compounds.

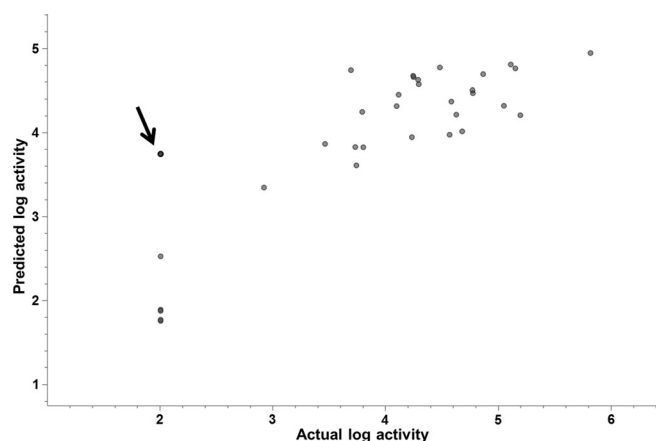
In summary, the 3D QSAR model created for LAT1-binding compounds indicates that the following chemical modifications to currently known substrates can increase their affinity for LAT1: 1) increasing the negative charge over the aromatic ring of L-phenylalanine and L-tryptophan; 2) adding substituents that increase the negative charge in proximity to the 5- and 6-positions of L-tryptophan, and similarly to the 3- and 4-positions of L-phenylalanine; 3) adding substituents to L-phenylalanine and L-tryptophan at the same plane with their aromatic side chains at the 5- and 6-positions of L-tryptophan, and similarly to the 3- and 4-positions of L-phenylalanine to

fulfill sterically favored areas. In addition, the following characteristics can decrease affinity for LAT1: 4) adding steric features above the aromatic plane of the amino acid side chains; 5) inserting substituents beyond the amino acid terminal which reach sterically disfavored regions and disturb the amino acid functionality of the promoiety.

### Model validation

The CoMFA model showed acceptable internal predictivity ( $Q^2 > 0.5$ ) when this was determined with leave-one-out (LOO) cross-validation. The experimental and predicted activities for the compounds showed good accuracy with the exception of two compounds: valproic acid prodrug **26** and aliphatic compound **35**. The predictions of activity for them (72 and 43% L-leucine uptake inhibition, respectively) differ by >1.2 logarithmic units from measured activities, which translate into 39 and 43 percentage units difference in L-leucine uptake inhibition (Supporting Information).

This might be due to the fact that the simultaneously *meta*- and *para*-substituted prodrug **26** and non-aromatic compound **35** are located outside the chemical space presented by the CoMFA model. Therefore, activity predictions for **26** and **35** are less reliable due to the large extrapolation. Another modest outlier, compound **39**, is observed in the plot of predicted versus measured activities (Figure 2). It is an amino acid derivative which is unable to inhibit the uptake of L-leucine. Excluding **39** from the training set decreases the statistical quality of the CoMFA model which is suggested to result from the strong electrostatic interactions of the compound (i.e., the nitro substituent at position 3 withdrawing electrons from the aromatic ring). However, compound **39** provides important information on electrostatics in the CoMFA model which is missed if the compound is omitted from the training set. Nevertheless, the predicted activity values of CoMFA model for the training set are in agreement with the experimental data within a statistically tolerable error range. Subsequently, the CoMFA model was used to guide the synthesis of eight novel prodrugs (**40–47**) and then used to predict their activities, which were compared with [ $^{14}$ C]L-leucine uptake inhibition determined by using the in situ rat brain perfusion method. These compounds (*prediction set*) were further used to determine the external predictivity of the CoMFA model.

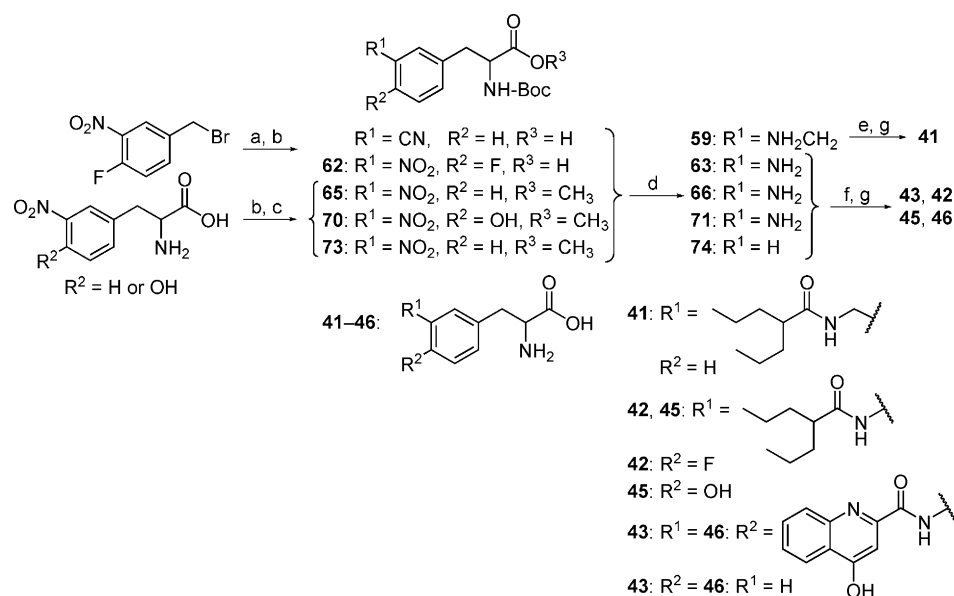


**Figure 2.** Predicted versus measured binding affinities of compounds in the training set shown on a logarithmic scale. Those compounds with the highest percent [ $^{14}\text{C}$ ]L-leucine uptake inhibition are located at the upper right-hand side of the coordinates. The arrow indicates compound **39**.

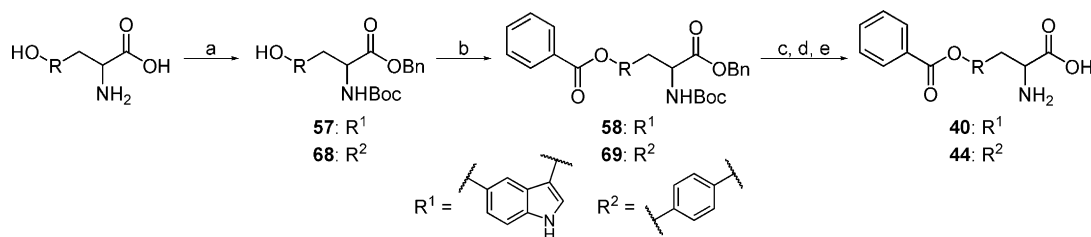
### Chemistry of the prediction set

The synthesized prodrugs include two benzoic acid prodrugs (**40** and **44**), four valproic acid prodrugs (**41**, **42**, **45**, and **47**), and two prodrugs of kynurenic acid (**43** and **46**). Prodrugs **40** and **44** were synthesized by protecting amino and carboxylic acid groups of L-5-hydroxytryptophan and L-tyrosine with *tert*-butoxycarbonyl (Boc) and benzyl (Bn) groups, respectively, and coupling the obtained compounds with benzoic acid using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) (Scheme 1). The Bn and Boc protecting groups were cleaved sequentially by hydrogenolysis on palladium and trifluoroacetic acid (TFA), respectively. Finally, the products were treated with hydrochloric acid to remove TFA

residues. Prodrugs **41**, **42**, **43**, **45**, and **46** were prepared by starting from commercially available modified amino acids (Scheme 2). Amino acid **63** was prepared by reacting 4-nitro-3-fluorobenzylbromide with diethyl acetamidomalonate in the presence of sodium hydride. The amino groups of all amino acid derivatives were protected with Boc if the commercial starting compounds were not protected already. The carboxylic acid groups were converted into their methyl esters before palladium-catalyzed reduction of nitro or cyano groups. The parent drugs and amino acids **63**, **66**, **71**, and **74** were then coupled by EDC. The carboxylic acid of the amino acid derivative with a cyano group was left unprotected, as compound **59** was reacted with valproic acid chloride after reduction of the cyano group. Finally, all the protecting groups were removed through base-catalyzed hydrolysis for the methyl ester and by acid-catalyzed hydrolysis for the Boc group, respectively. The products were treated either with  $\text{HCl}_{(\text{g})}$  or HCl in dioxane to remove any remaining TFA. Valproic acid prodrug **47** (Table 3) was prepared by coupling 2-propylpentanoyl chloride



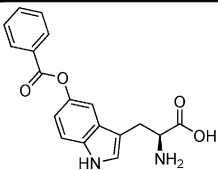
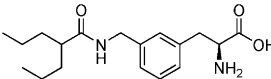
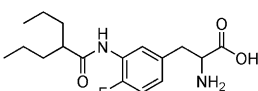
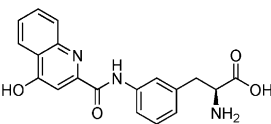
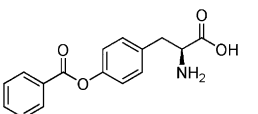
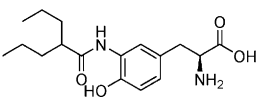
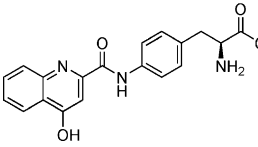
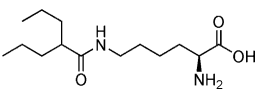
**Scheme 2.** Synthesis of **41–43** and **45** and **46**. *Reagents and conditions:* a) 1. NaH, acetamidomalonate, DMF, RT, 4 h; 2. HCl, reflux 18 h; b)  $\text{Boc}_2\text{O}$ , TEA, THF, RT, 18 h; c) dimethyl sulfate,  $\text{K}_2\text{CO}_3$ , acetone, RT, 18 h, or  $\text{SOCl}_2$  in MeOH, reflux, 4 h, 69–99%; d) 10% Pd/C, MeOH, RT, 4 h, 95–99%; e) 2-propylpentanoyl chloride, NaOH,  $\text{CH}_2\text{Cl}_2$ , RT, 3 h, 72%; f) valproic or kynurenic acid, EDC, DMAP, DMF, RT, 18 h, 31–66%; g) 1. NaOH or  $\text{LiOH}_{(\text{aq})}$ , MeOH, RT, 30 min; 2. TFA,  $\text{CH}_2\text{Cl}_2$ , RT, 30 min, 81–91% or  $\text{HCl}_{(\text{g})}$ ,  $\text{CH}_3\text{CN}$ , RT, 30 min, 99%.



**Scheme 1.** Synthesis of **40** and **44**. *Reagents and conditions:* a) 1.  $\text{Boc}_2\text{O}$ , TEA, THF, RT, 18 h; 2. benzyl alcohol, EDC, DMAP,  $\text{CH}_2\text{Cl}_2$ , RT, 18 h, 58–70%; b) benzoic acid, EDC, DMAP,  $\text{CH}_2\text{Cl}_2$ , RT, 18 h, 35–67%; c) Pd (10% on activated charcoal), MeOH, RT, 45 min; d) TFA,  $\text{CH}_2\text{Cl}_2$ , RT, 30 min; e) 4 M HCl in dioxane, RT, 10 min, 48–72%.



**Table 3.** Prediction set synthesized based on the statistically best CoMFA model with the values of percent L-leucine uptake inhibition determined by the in situ rat brain perfusion technique.

Compd (Inh. [%]) <sup>[a]</sup>	Structure	Compd (Inh. [%]) <sup>[a]</sup>	Structure
<b>40</b> (99±1)		<b>41</b> (92±2)	
<b>42</b> (87±4)		<b>43</b> (85±4)	
<b>44</b> (73±6)		<b>45</b> (65±3)	
<b>46</b> (25±3)		<b>47</b> (3±0) <sup>[b]</sup>	

[a] [<sup>14</sup>C]L-leucine brain capillary surface area (PA) percent inhibition data are the average±SD of *n*=3 experiments conducted at a compound concentration of 100 μM.  
[b] *n*=2.

with Boc-L-lysine and subsequently removing the Boc group by acid-catalyzed hydrolysis. Detailed synthesis procedures are described in the Supporting Information.

### External predictivity

The CoMFA model predicted compounds **40–45** to be active (>50% L-leucine uptake inhibition; >4.00 in logit transformation), although the measured and predicted L-leucine uptake inhibition values of compounds **40**, **44**, and **45** differ from each other by >10 percentage units (Table 4). Benzoic acid prodrug **40** is the only L-tryptophan derivative substituted at position 6 of the indole ring (Figure 1). This was suggested by the CoMFA model, and it is a singularity in the chemical space of the data set. The model also anticipated the *meta*- and *para*-substituted prodrug **46** to be threefold more active (77% L-leucine uptake inhibition) than determined by the in situ rat brain perfusion (25% L-leucine uptake inhibition) (Table 4). Prodrug **46** can be considered to locate slightly outside the chemical space of the data set, which complicates the correct prediction of its activity. The model concluded that aliphatic compound **47** would be inactive (<4.00 logit transformation), which is in agreement with its affinity (3% L-leucine uptake inhibition). Altogether, the CoMFA model was shown to possess statistically significant external predictivity (concordance correlation coefficient, CCC=0.91; Table 2).<sup>[20,21]</sup> CCC was selected as an indicator for this study, as it has been proposed to be the most stable and restrictive in determining real predictivity of QSAR models.<sup>[20]</sup> These results show that the CoMFA model

can be used to prioritize compounds for synthesis and further optimization of the LAT1 prodrugs.

### QSAR of synthesized compounds

Topomer CoMFA suggested that the steric and electrostatic properties of L-tryptophan have a positive effect on affinity for LAT1. Consequently, we were interested in introducing L-tryptophan and its analogues as promoieties for new LAT1 prodrugs. As far as we are aware, L-tryptophan has not been previously used as a promoiety for LAT1. The CoMFA model indicated that adding steric features at position 5 of L-tryptophan could be beneficial for affinity (Figure 1; green crescent-shaped contour at R2). For this reason, benzoic acid was substituted into position 5 of L-tryptophan (compound **40**), which is a representative prodrug with the highest affinity for LAT1 in our data set (99% L-leucine uptake inhibition). Relative to natural L-tryptophan (**6**; 88% L-leucine uptake inhibition) and 5-methoxy-D,L-tryptophan (**12**; 75% L-leucine uptake inhibition), prodrug **40** expresses significantly higher affinity for LAT1. This indicates that L-tryptophan is a valuable promoiety for novel LAT1-targeted prodrugs.

**Table 4.** Measured versus predicted percent L-leucine uptake inhibition by the prediction set.

Compd	Inhibition [%] <sup>[a]</sup>		log (Inh. [%]) <sup>[b]</sup>	
	Measured	Predicted	Measured	Predicted
<b>40</b>	99±1	86	6.00	4.78
<b>41</b>	92±2	87	5.06	4.83
<b>42</b>	87±4	84	4.83	4.73
<b>43</b>	85±4	84	4.75	4.71
<b>44</b>	73±6	80	4.44	4.60
<b>45</b>	65±3	85	4.26	4.74
<b>46</b>	25±3	77	3.51	4.52
<b>47</b>	3±0 <sup>[b]</sup>	43	2.43	3.87

[a] [<sup>14</sup>C]L-leucine brain capillary surface area percent inhibition data are the average±SD of *n*=3 experiments conducted at a compound concentration of 100 μM. [b] Logit transformation of L-leucine uptake inhibition =  $\text{Sc} + \log(I/(100-I))$ , in which *I*=percent L-leucine uptake inhibition;  $I > 99 \rightarrow I = 99$ ,  $I < 1 \rightarrow I = 1$ ; Sc (scale factor) =  $-\log(c)$ , for which *c*=standard molar concentration ( $10^{-4}$ ).<sup>[33]</sup>

Similar to **40**, prodrug **44** was substituted at position 4 of L-phenylalanine which was also observed to increase affinity for LAT1 by the CoMFA model (green crescent-shaped contour at R2). Both representative prodrugs inhibit L-leucine uptake by >50%, but the results show that substitution at position 5 of L-tryptophan is preferred over position 4 of L-phenylalanine (99 and 73% uptake inhibition, respectively).

In our previous studies we also demonstrated that the 3-position of L-phenylalanine is preferred over position 4.<sup>[11]</sup> This was further studied with kynurenic acid prodrugs **43** and **46**,

as the CoMFA model suggests that adding steric features at positions 3 and 4 of L-phenylalanine can enhance affinity for LAT1 (Figure 1). Similarly to previous studies, kynurenic acid at position 3 (**43**) is considerably more efficient than the parent drug at position 4 (**46**) (85 and 25 % L-leucine uptake inhibition, respectively). Adding kynurenic acid successfully as the parent drug (**43** and **46**) shows that in addition to flexible prodrugs, LAT1 can accommodate prodrugs with larger and more rigid parent drugs than ketoprofen, which was the parent drug used in our earlier work.<sup>[8]</sup>

Inserting steric features at position 3 was also studied by introducing a methylene group between the amide and aromatic ring of L-phenylalanine (compound **41**). The modification did not disrupt binding to LAT1, as prodrug **41** inhibits L-leucine uptake inhibition to the same extent as high-affinity prodrugs **3** and **4** (> 90 % L-leucine uptake inhibition). This finding indicates that the distance between the parent drug and promoiety can be modified, and that prodrugs are relatively tolerant in the sense of geometry of the parent drugs. This is encouraging when one considers its implications for the future design of larger LAT1 prodrugs.

The steric and electrostatic properties of L-tryptophan are similar to those of L-phenylalanine, which has been successfully applied as a promoiety in several previously reported LAT1 prodrugs (**1**, **3**, **4**, **16**, and **17**).<sup>[8,11,19]</sup> Based on the CoMFA model, electrostatic and steric modifications were made to the aromatic ring of L-phenylalanine to increase their affinity for LAT1. Following from this step, fluoro and hydroxy substituents were placed at position 4 of L-phenylalanine (**42** and **45**). Prodrugs **42** and **45** are close analogues of the methoxy-substituted valproic acid prodrug **26** in the training set (Tables 1 and 3). The substituents used in **42** and **45** are similar in size, and both can activate the aromatic ring. Both prodrugs exhibit high affinity for LAT1 (87 and 65 % L-leucine uptake inhibition, respectively), whereas prodrug **26** inhibits only 33 % of L-leucine uptake. The molecular size of the methoxy substituent in **26** is slightly larger than either fluoride or hydroxy and it can rotate more freely. This could affect the conformation of the prodrug, directing the parent drug into an disfavored region, thereby decreasing affinity for LAT1. Nevertheless, compounds **26**, **42**, and **45** show that it is possible to employ *para* and *meta* substitutions simultaneously and still have efficient LAT1 prodrugs. Successful simultaneous substitution of **42** and **45** also provides an interesting possibility to modify the chemical stability of the prodrug bond by a suitable substituent at the *meta* position which could affect the bioconversion rate of LAT1 prodrugs. However, this needs to be further studied before any definitive conclusions can be made.

Lastly, compound **47** was designed to examine the flexibility of LAT1 prodrugs in greater detail. In **47**, exchanging the rigid promoiety for a flexible L-lysine resulted in an almost complete loss of inhibitory activity (3 % L-leucine uptake inhibition). However, poor affinity does not necessarily result from increased flexibility alone, as replacing aromaticity with aliphatic features has been shown in previous studies to decrease affinity.<sup>[7]</sup> Non-aromatic prodrug **47** inhibits L-leucine uptake over 70 percentage units less efficiently than prodrug **12** with an

aliphatic promoiety and aromatic parent drug (Tables 1 and 2). This further shows that complete removal of aromaticity has a significant effect on LAT1 affinity.

## Conclusions

The topomer CoMFA model generated in this study provides novel insight into the molecular features required for LAT1-binding compounds with high affinity. It clarifies the molecular size, flexibility, and preferred position of the parent drug of LAT1-targeted prodrugs. The CoMFA model was used successfully in the design of new prodrugs, which had high affinity for LAT1; it resulted in the discovery of the first representative L-tryptophan prodrug. This indicates that the CoMFA model can be used in the rational design of novel prodrugs and other compounds that bind efficiently to LAT1.

## Experimental Section

### Computational design

**Ligands:** The training set contains substrates of LAT1 and their derivatives, ketoprofen, valproic acid, and dopamine prodrugs with activities spanning three log units. The prediction set contains close structural analogues with training set compounds and includes new valproic acid and kynurenic acid prodrugs.

**Training set ( $n=39$ ):** O-[2-(3-benzoylphenyl)-1-oxopropyl]-L-tyrosine (**1**),<sup>[8]</sup> L-1-naphthylalanine (**2**), 3-(2-propylpentanamido)-L-phenylalanine (**3**),<sup>[11]</sup> 3-((2-propylpentanoyl)oxy)-L-phenylalanine (**4**), 5-fluoro-L-tryptophan (**5**), L-tryptophan (**6**), L-phenylalanine (**7**), 3-((3,4-dihydroxyphenethyl)carbamoyl)-L-phenylalanine (**8**),<sup>[19]</sup> D,L-cyclopentanealanine (**9**), 3-cyano-L-phenylalanine (**10**), L-(+)- $\alpha$ -phenylglycine (**11**), N6-[2-(3-benzoylphenyl)-1-oxopropyl]-L-lysine (**12**),<sup>[6]</sup> 5-methoxy-D,L-tryptophan (**13**),<sup>[7]</sup> levodopa (**14**), 3-methoxy-L-tyrosine (**15**), 4-((2-propylpentanoyl)oxy)-L-phenylalanine (**16**),<sup>[11]</sup> 4-(2-propylpentanamido)-L-phenylalanine (**17**),<sup>[11]</sup> D,L-homophenylalanine (**18**), gabapentin (**19**), N-phenyl-L-glutamine (**20**), 5-hydroxy-L-tryptophan (**21**), N6-[2-(3,4-dihydroxyphenyl)ethyl]-6-oxolysine (**22**),<sup>[19]</sup>  $\alpha$ -methyl-L-tyrosine (**23**), D,L- $\beta$ -phenylalanine (**24**), 4-((3,4-dihydroxyphenethyl)amino)-4-oxo-L-asparagine (**25**),<sup>[19]</sup> 4-methoxy-3-(2-propylpentanamido)-L-phenylalanine (**26**), L-alanyl-L-tyrosine (**27**), 2-amino-5-methoxybenzoic acid (**28**), L-2-amino-3-phenylpropanol (**29**), saclofen (**30**), L-alanyl-L-tyrosyl-L-alanine (**31**), sulfanilamide (**32**), 2-[2-(3-benzoylphenyl)-1-oxopropoxy]ethyl ester-L-leucine (**33**), 2-[2-(3-benzoylphenyl)-1-oxopropoxy]ethyl ester-L-phenylalanine (**34**), diaminopimelic acid (**35**), N-[2-(3-benzoylphenyl)-1-oxopropyl]-L-leucine (**36**), N-[2-(3-benzoylphenyl)-1-oxopropyl]-L-phenylalanine (**37**), N2-benzoyl-L-lysine (**38**), 2-Amino-4-((3-nitrophenyl)amino)-4-oxobutanoic acid (**39**).

**Prediction set ( $n=8$ ):** 5-benzoyloxy-L-tryptophan (**40**), 3-((2-propylpentanamido)methyl)-L-phenylalanine (**41**), 4-fluoro-3-(2-propylpentanamido)-D,L-phenylalanine (**42**), 3-(4-hydroxyquinoline-2-carboxamido)-L-phenylalanine (**43**), 4-benzoyloxy-L-phenylalanine (**44**), (4-hydroxy-3-(2-propylpentanamido)-L-phenylalanine (**45**), 4-(4-hydroxyquinoline-2-carboxamido)-L-phenylalanine (**46**), 2-propylpentanamido-L-lysine (**47**).

**Virtual screening:** Brutus<sup>[22,23]</sup> was used to identify novel compounds with affinity for LAT1 by considering 3D molecular interaction fields of a compound database. The databases from Asinex Merged Libraries ( $n=398\,926$ ), Bachem ( $n=4874$ ), Acros ( $n=$

16545), and Sigma–Aldrich ( $n=41\,390$ ) were searched by using **3**, **13**, and **40** as queries. The query compounds and databases were pre-processed with the LigPrep package of Schrödinger Suite 2011.<sup>[24]</sup> Tautomers and stereoisomers for the database compounds and query compounds were generated at  $\text{pH } 7.4 \pm 1$  using Epik.<sup>[25]</sup> Only *S* configurations were selected for the query compounds. Partial charge calculation and minimization of the molecular structures in 500 steps were performed using the OPLS\_2005 force field.<sup>[26]</sup> For Brutus screening, partial charges were assigned with modified MMFF94 force field as implemented in MOE 2012.17.<sup>[27]</sup> Default settings for minimum hydrogen bond (0.5), van der Waals (0.7), hydrophobic (0.2), and steric (0.5) similarity indices in Brutus were employed.<sup>[22,23]</sup> The top 5% of the hit compounds ( $n=172$ ) were selected for further analysis based on Brutus total similarity and hydrophobic, van der Waals, steric, and hydrogen bond similarities, respectively. After visual inspection, a small selection ( $n=5$ ) of hit compounds was selected based on their novelty with respect to known LAT1-binding compounds and synthetic feasibility in pro-drug design. None of them had affinity for LAT1.

**CoMFA models:** The structural features of a series of LAT1-binding compounds ( $n=47$ ) were correlated with experimentally determined LAT1 affinity using topomer and classical CoMFA methodologies.<sup>[16,28]</sup> Both methods focus on 3D molecular interaction fields of molecules instead of 2D structures. Consequently, they can mimic the way a protein binding site recognizes its substrates. Contrary to classical CoMFA, topomer CoMFA does not require pre-determined alignment of compounds. This is beneficial for compound series with varying molecular sizes such as our dataset, which contains small moiety-sized compounds ( $< 250$  Da) and larger prodrugs (270–420 Da). Topomer CoMFA is also able to elucidate the effect of small structural changes on a series of structurally similar compounds more specifically than classical CoMFA. This results in identifying a common core in the data series and dividing the rest of the molecule into varying parts, i.e., topomers. For this reason, the effect of different topomers and their effect on biological activity can be studied in greater detail. In classical CoMFA the effect of small structural changes on biological activity can be overlooked more easily because the method perceives the entire molecule and its molecular interaction fields due to variations in alignment instead.

With the topomer approach the compounds were fragmented into two variable R groups between the single bond connecting C $\alpha$  and C $\beta$ . The majority ( $n=31$ ) were automatically fragmented, whereas four compounds had to be manually fragmented (**11**, **23**, **24**, **35**) (Table 1). 3D representations of the compounds solely based on their 2D topology were created using deterministic rules.<sup>[29]</sup> Alignment was performed by defining a common core for the compounds which was followed by a single fragment conformation being generated for each compound. As a result, self-consistent and absolute configurations and 3D conformations were generated for the compound data set.<sup>[29]</sup> Four compounds (**19**, **28**, **30**, and **32**; Table 1) could not be used in topomer CoMFA modeling due to the lack of a common core, that is, an amino acid functionality. In total, six models were generated with slightly different compounds included in the training set (model A:  $n=34$ ; model B:  $n=32$ ; model C:  $n=33$ ; model D:  $n=36$ ; model E:  $n=35$ ; model F:  $n=35$ ; Table 2). The best model was selected based on statistics and external predictivity.

In classical CoMFA, the complete series of compounds ( $n=52$ ) was aligned with Brutus.<sup>[22,23]</sup> Brutus considers compounds based on their 3D molecular interaction fields instead of simple 2D topological similarities, which allows it to align structurally diverse com-

pounds based on their interaction potential. Conformational sampling of the dataset was performed using LigPrep packages of Schrödinger suite 2011.<sup>[24]</sup> Tautomers and stereoisomers were generated at pH 7.4 using Ionizer and MMFFs charges. Conformations were generated using ConfGen and Fast CF protocol,<sup>[30]</sup> which resulted in a total of 7325 conformations. The lowest-energy conformation of compound **40**, which possesses high affinity for LAT1 and is relatively rigid, was selected as the template compound. Default similarity filters were disabled to guarantee the whole database to be aligned with the template. Alternative alignments for each compound with the template were then ranked according to total similarity, and the top solution for each compound was selected for further analysis. The alignment of the compound series was visually checked to identify compounds with poorly aligned conformations. Alignment was followed by generating electrostatic and steric interaction fields using Coulomb and Lennard–Jones potentials, respectively.<sup>[16]</sup> The compounds were positioned in a regular 3D cubic grid with 2 Å spacing. The interaction energies between the compound and probe atoms were calculated at each grid point. Subsequently, the relationships of the intermolecular interaction fields with biological activity ( $[^1\text{C}]$ -leucine uptake inhibition) were analyzed by multivariate statistical technique, PLS.<sup>[31]</sup>

**Validation of CoMFA models:** LOO cross-validation was used to evaluate the optimal number of components for non-cross-validated analysis. Each compound of the training was omitted in turn, and the derived model was used to predict the biological activity of the omitted compound. The optimal number of components for non-cross-validated analysis was determined based on the “5% rule” (Table 2). If adding one component increased  $Q^2$  by 5%, the additional component was considered justified. In general,  $Q^2$  values  $> 0.5$  were considered statistically significant.<sup>[33]</sup> Predicted versus measured affinities were plotted to identify outliers. Compounds **27** and **38** were omitted from the model which increased the  $Q^2$  value and decreased the standard error of prediction to  $> 5\%$ . The plot of predicted versus measured affinities in logarithmic scale (Figure 2) suggested compound **39** as being a potential outlier, but its inclusion in the training set was needed in order to produce a model with reasonable internal predictivity. The regression equation with thousands of coefficients is represented as CoMFA contour maps, which show regions in space where adding negative or positive charge or steric bulk either increases or decreases ligand activity (Figure 1). Information obtained from the best model (Model E) was used in further lead optimization to suggest structural modifications that could increase the compounds' affinity for LAT1. Detailed information on the measured and predicted activities for training set can be found in the Supporting Information.

The sensitivity of the CoMFA models to chance correlations was determined by using progressive scrambling<sup>[32]</sup> (Table 2) as implemented in Sybyl-X 2.0.<sup>[33]</sup> Scramblings were performed in 20 individual runs to obtain statistically valid results. The model consisting of two components with  $dq^2/dr^2_{yy}$  slope closer to the optimal unity and smaller error margin than the model with three components was considered valid and employed in guiding the synthesis of new compounds whose activity (affinity) was predicted with the CoMFA models.

The set of synthesized prodrugs ( $n=8$ ) (Table 3) was used as a prediction set for analyzing the external predictivity of the model. CCC was determined for the prediction set based on measured affinity versus predicted affinity.<sup>[20,21]</sup> Topomer CoMFA models E and D with  $\text{CCC} > 0.85$  were considered significant.



## Synthesis

All of the purchased commercial compounds were of analytical grade or better. The synthesized prodrugs possess >95% purity, which was determined by elemental analysis (C, H, N). Characterization of the prodrugs was carried out with  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectroscopy and mass spectrometry. Reactions were monitored by thin-layer chromatography using aluminum sheets coated with silica gel 60  $F_{254}$  (0.24 mm) with suitable visualization. Purifications by flash chromatography were performed using silica.  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra were recorded on a Bruker Avance 500 spectrometer (Bruker Biospin, Fallanden, Switzerland) operating at 500.13 and 125.75 MHz. The products were also characterized by MS with a Finnigan LCQ quadrupole ion trap mass spectrometer (Finnigan MAT, San Jose, CA, USA) equipped with an electrospray ionization source. Purity was determined by elemental analysis (C, H, N) with a ThermoQuest CE Instruments EA 1110-CHNS-O elemental analyzer (CE Instruments, Milan, Italy). Detailed synthetic procedures and characterization data can be found in the Supporting Information.

## Animal testing

**Determination of LAT1 affinity:** The ability of the investigated compounds to bind LAT1 was studied by the in situ rat brain perfusion technique, which has been described and validated in detail previously.<sup>[8, 11, 18]</sup> The 100% permeability surface area (PA) product of [ $^{14}\text{C}$ ]L-leucine, a known substrate of LAT1, was determined by 30 s perfusion of  $0.2\ \mu\text{Ci mL}^{-1}$  [ $^{14}\text{C}$ ]L-leucine ( $0.64\ \mu\text{M}$ ). To determine the interactions of the investigated compounds with LAT1 at the BBB, compounds at  $100\ \mu\text{M}$  were co-perfused for 30 s with  $0.2\ \mu\text{Ci mL}^{-1}$  [ $^{14}\text{C}$ ]L-leucine. The PA product of  $0.2\ \mu\text{Ci mL}^{-1}$  [ $^{14}\text{C}$ ]L-leucine co-perfused with the investigated compounds was then compared with the 100% PA product of  $0.2\ \mu\text{Ci mL}^{-1}$  [ $^{14}\text{C}$ ]L-leucine. The LAT1 binding of the investigated compounds is shown by the decrease in PA product of [ $^{14}\text{C}$ ]L-leucine caused by competitive binding to LAT1.

## Acknowledgements

This work was supported by funding from the Academy of Finland (grant no. 256837), the FinPharma Doctoral Program, the Magnus Ehrnrooth Foundation, the Jenny and Antti Wihuri Foundation, the Kuopio University Foundation, the Finnish Pharmaceutical Society, the Finnish Cultural Foundation, the Orion Research Foundation, the Alfred Kordelin Foundation and the Emil Aaltonen Foundation. The CSC-IT Center for Science Limited is acknowledged for providing computing resources, as is Dr. Ewen MacDonald (University of Eastern Finland) for reviewing the language of this text.

**Keywords:** 3D QSAR • CoMFA • drug design • LAT1 • prodrugs

- [1] F. Verrey, *Pfluegers Arch.* **2003**, *445*, 529–533.
- [2] R. J. Boado, J. Y. Li, M. Nagaya, C. Zhang, W. M. Pardridge, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 12079–12084.
- [3] Y. Kanai, Y. Fukasawa, S. H. Cha, H. Segawa, A. Chairoungdua, D. K. Kim, H. Matsuo, J. Y. Kim, K. Miyamoto, E. Takeda, H. Endo, *J. Biol. Chem.* **2000**, *275*, 20787–20793.
- [4] P. Gomes, P. Soares-da-Silva, *Brain Res.* **1999**, *829*, 143–150.
- [5] K. C. Cundy, R. Branch, T. Chernov-Rogan, T. Dias, T. Estrada, K. Hold, K. Koller, X. Liu, A. Mann, M. Panuwat, S. P. Raillard, S. Upadhyay, Q. Q. Wu, J.-N. Xiang, H. Yan, N. Zerangue, C. X. Zhou, R. W. Barrett, M. A. Gallop, *J. Pharmacol. Exp. Ther.* **2004**, *311*, 315–323.
- [6] J. B. van Bree, K. L. Audus, R. T. Borchardt, *Pharm. Res.* **1988**, *5*, 369–371.
- [7] H. Ylikangas, L. Peura, K. Malmioja, J. Leppänen, K. Laine, A. Poso, M. Lahtela-Kakkonen, J. Rautio, *Eur. J. Pharm. Sci.* **2013**, *48*, 523–531.
- [8] M. Gynther, K. Laine, J. Ropponen, J. Leppänen, A. Mannila, T. Nevalainen, J. Savolainen, T. Jarvinen, J. Rautio, *J. Med. Chem.* **2008**, *51*, 932–936.
- [9] H. Uchino, Y. Kanai, D. K. Kim, M. F. Wempe, A. Chairoungdua, E. Morimoto, M. W. Anders, H. Endou, *Mol. Pharmacol.* **2002**, *61*, 729–737.
- [10] M. Gynther, A. Jalkanen, M. Lehtonen, M. Forsberg, K. Laine, J. Ropponen, J. Leppänen, J. Knuuti, J. Rautio, *Int. J. Pharm.* **2010**, *399*, 121–128.
- [11] L. Peura, K. Malmioja, K. Laine, J. Leppänen, M. Gynther, A. Isotalo, J. Rautio, *Mol. Pharm.* **2011**, *8*, 1857–1866.
- [12] I. Walker, D. Nicholls, J. W. Irwin, S. Freeman, *Int. J. Pharm.* **1994**, *104*, 157–167.
- [13] F. P. Bonina, L. Arenare, F. Palagiano, A. Saija, F. Nava, D. Trombetta, P. de Caprariis, *J. Pharm. Sci.* **1999**, *88*, 561–567.
- [14] D. M. Killian, S. Hermeling, P. J. Chikhale, *Drug Delivery* **2007**, *14*, 25–31.
- [15] Q. R. Smith, International Congress Series 1277, **2005**, 63–74.
- [16] R. D. Cramer, D. E. Patterson, J. D. Bunce, *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.
- [17] R. D. Cramer, *J. Med. Chem.* **2003**, *46*, 374–388.
- [18] Y. Takasato, S. I. Rapoport, Q. R. Smith, *Am. J. Physiol.* **1984**, *247*, H484–H493.
- [19] L. Peura, K. Malmioja, K. Huttunen, J. Leppänen, M. Hämäläinen, M. M. Forsberg, M. Gynther, J. Rautio, K. Laine, *Pharm. Res.* **2013**, *30*, 2523–2537.
- [20] N. Chirico, P. Gramatica, *J. Chem. Inf. Model.* **2011**, *51*, 2320–2335.
- [21] N. Chirico, P. Gramatica, *J. Chem. Inf. Model.* **2012**, *52*, 2044–2058.
- [22] A. J. Tervo, T. Rönkkö, T. H. Nyronen, A. Poso, *J. Med. Chem.* **2005**, *48*, 4076–4086.
- [23] T. Rönkkö, A. J. Tervo, J. Parkkinen, A. Poso, *J. Comput. Aid. Mol. Des.* **2006**, *20*, 227–236.
- [24] LigPrep ver. 2.5, Schrödinger LLC, New York, NY (USA), **2012**.
- [25] Epik ver. 2.2, Schrödinger LLC, New York, NY (USA), **2011**.
- [26] W. L. Jorgensen, J. Tirado-Rives, *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
- [27] Chemical Computing Group Inc., 1010 Sherbooke Street West, Suite 910, Montreal, QC, H3A 2R7 (Canada) **2009**.
- [28] R. D. Cramer, *J. Comput. Aid. Mol. Des.* **2012**, *26*, 805–819.
- [29] R. J. Jilek, R. D. Cramer, *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1221–1227.
- [30] I. J. Chen, N. Foloppe, *J. Chem. Inf. Model.* **2010**, *50*, 822–839.
- [31] B. L. Bush, R. B. Nachbar, Jr., *J. Comput. Aided Mol. Des.* **1993**, *7*, 587–619.
- [32] R. D. Clark, P. C. Fox, *J. Comput. Aid. Mol. Des.* **2004**, *18*, 563–576.
- [33] Sybyl-X ver. 2.0, Tripos International, 1699 South Hanley Road, St. Louis, MO, 63144 (USA).

Received: July 31, 2014

Published online on September 9, 2014