High selectivity of the γ-Aminobutyric acid transporter (GABA) transporter 2 (GAT-2, SLC6A13) revealed by structure-based approach

Avner Schlessinger[1,2,4], Matthias B. Wittwer[1,4], Amber Dahlin[1,3], Natalia Khuri[1,2], Massimiliano Bonomi[1,2], Hao Fan[1,2], Kathleen M. Giacomini[1]*, Andrej Sali[1,2]*

[1]Department of Bioengineering and Therapeutic Sciences; [2]California Institute for Quantitative Biosciences, University of California, San Francisco, CA 94158; [3]Present address: Channing Laboratory for Network Medicine, Brigham and Women's Hospital and Harvard Medical School.

[4]Equal contribution

*To whom correspondence should be addressed: Andrej Sali, E-mail sali@salilab.org; Kathleen M. Giacomini, E-mail: kathy.giacomini@ucsf.edu

Template selection. The structure of LeuT was determined by X-ray crystallography in different conformations, with a variety of ligands (1,2). We selected the template structures (PDB identifier 2A65 (3) and 3F3A (4) for the occluded and outward-facing conformations, respectively) based on the following criteria: (i) highest resolution (1.65 Å and 2 Å), (ii) substrate (leucine)- and inhibitor (tryptophan)-bound, and (iii) an occluded and an outward-facing conformations.

GAT-2-LeuT alignment. An initial GAT-2-LeuT alignment was extracted from a comprehensive comparison of the SLC6 family including eukaryotic and prokaryotic members (5). The alignment was corrected with the updated sequence of the active isoform of GAT-2 (April 2011) and was subsequently refined; four long segments distant from the primary binding site (extracellular loop 2, the loop between transmembrane helices 11 and 12, and the N- and C- termini) were excluded from modeling (Fig. S1). The sequence identity between the modeled fraction of GAT-2 and LeuT is 23%.

Sidechain refinement. For each one of the initial models generated by MODELLER, the sidechains of the binding site residues were repacked on a fixed backbone using SCWRL4 (6). The coordinates of the sodium ions from these initial models were used as steric constraints for the sidechains. In particular, we ran SCWRL4 on different combinations of residues. For both the occluded and outward-facing models, we ran SCWRL4 on E48 alone, four binding site residues (*ie*, E48, F288, L294, and Q391), or residues located within 4 Å from the coordinates of the ligand in the initial model (*ie*, leucine and tryptophan for the occluded and outward-facing models, respectively). For the outward facing model, we also ran SCWRL4 on an additional residue separately (*ie*, D447), three binding site residues (E48, Q391, and D447), or the whole protein.

MD simulations. All MD simulations were performed with GROMACS4 MD code (7). Each model, including pre- or post-sidechain optimization models, was refined using the following MD protocol. The

model was subjected to 10000 steps of conjugate gradient minimization under the Amber99SB-ILDN force field (8,9). To account for the membrane hydrophobic environment, an implicit model for the solvent based on a generalized Born formalism was used. A dielectric constant equal to 2 was used to model the membrane interior. Whenever ions were present in the model, the system was simulated "in vacuum" and coulomb interactions were screened by using a dielectric constant equal to 2. All bond lengths were constrained to their equilibrium values using the LINCS algorithm (10). A time step of 2 fs was adopted. A cutoff of 1.0 nm was used for the Lennard-Jones and the electrostatic interactions.

Enrichment. The enrichment factor was defined as:

$$EF_n = \frac{(\text{ligand}_{selected} / N_n)}{(\text{ligand}_{total} / N_{total})} \tag{1}$$

where ligand_{total} is the number of known ligands in a database containing N_{total} compounds, and $\text{ligand}_{selected}$ is the number of ligands found in a given subset of N_n compounds (11-13). Additionally, we used the area under the enrichment curve as a measure of virtual screening accuracy (11-13). The enrichment curve can be obtained by plotting the percentage of actual ligands correctly predicted (y-axis) within the top ranked subset of all database compounds (x-axis on logarithmic scale). We calculated the area under the curve (logAUC) of the enrichment plot for $\Delta x = 0.1$:

$$\log AUC = \frac{1}{\log_{10} 100 / 0.1} \sum_{0.1}^{100} \frac{\text{ligand}_{selected}(x)}{\text{ligand}_{total}} \Delta x \quad \text{where} \quad x = \log_{10} \frac{N_n}{N_{total}} \tag{2}$$
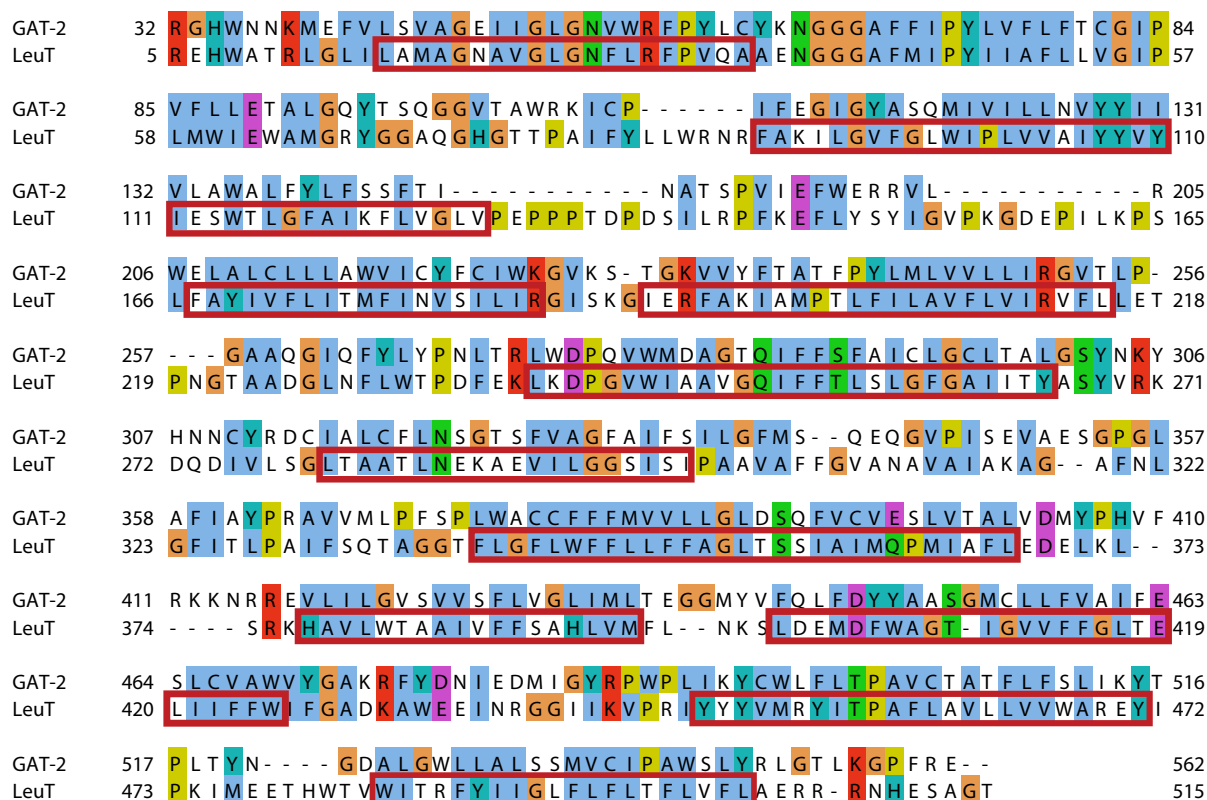
Fig S1. GAT-2-LeuT alignment. Putative transmembrane regions (in red rectangles) were obtained from the comprehensive SLC6 study (5), and were originally based on the PDB_TM database (14). The alignment was corrected and refined with an updated GAT-2 sequence. The LeuT sequence was derived from occluded X-ray structure (*ie*, PDB id 2A65). The sequence alignment was visualized using Jalview (15). The aligned residues were colored based on their type using the "Clustlx" color scheme.
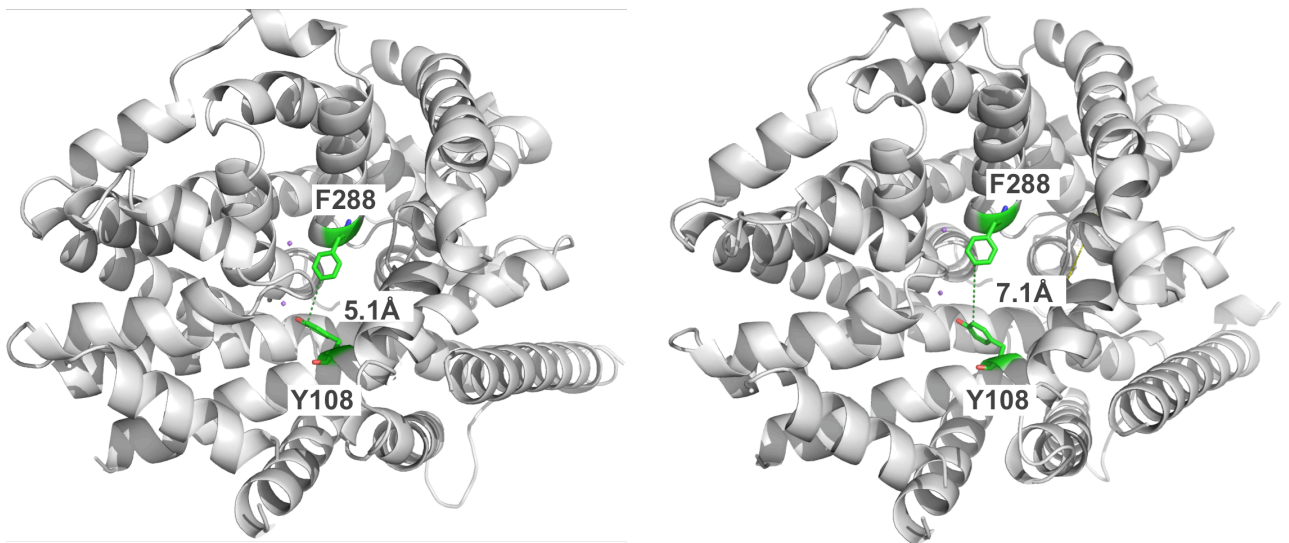
Fig S2. The extracellular gate. Top view of the extracellular gate formed by F288 and Y108 (green sticks) in the occluded (left panel) and outward-facing (right panel) conformations. All other GAT-2 residues are illustrated by white ribbons.
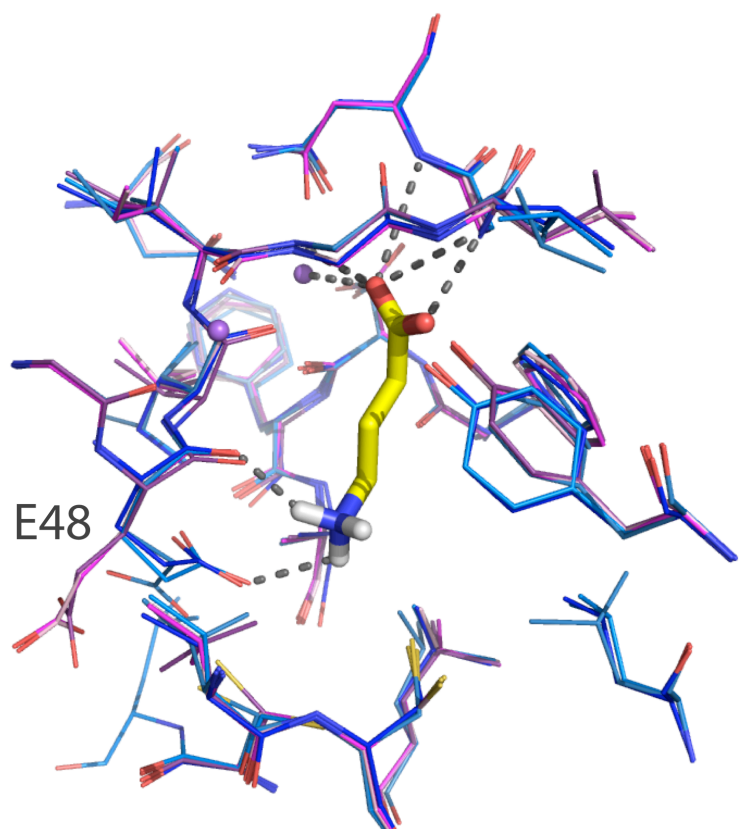
Fig S3. E48 in various models of GAT-2. The three occluded models with the best (blue) and worst (pink) enrichment scores are shown. Atoms are illustrated by lines, with oxygen, nitrogen, and hydrogen atoms in red, blue, and white, respectively. The sodium ions Na1 and Na2 are visualized with purple spheres. GABA is depicted in yellow sticks and its hydrogen bonds with GAT-2 in the final refined model (involving E48, G51, G53, N54, and Na1) are shown as dotted gray lines.
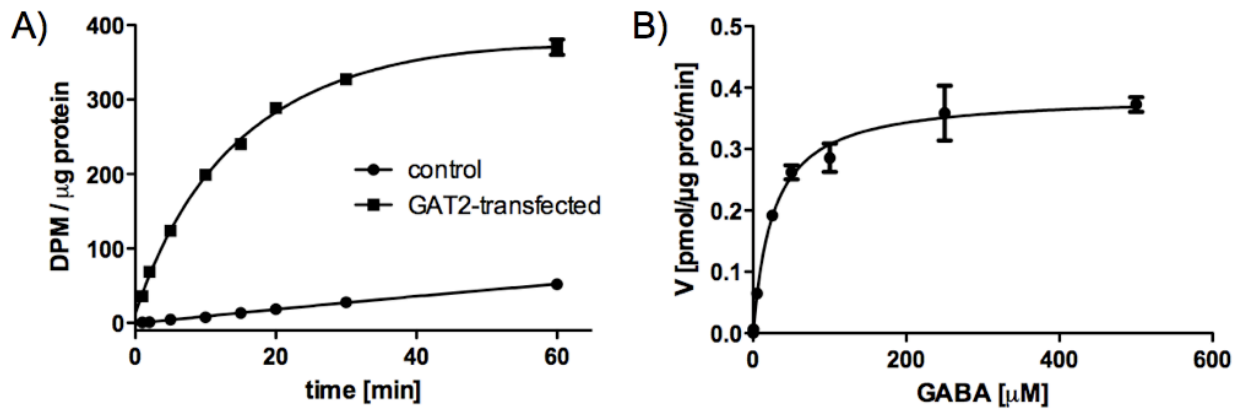
Fig S4. Time course and uptake kinetics of HEK293-cells transiently transfected with GAT-2. A) Time course of ^{3}H-GABA uptake into HEK293-cells transiently transfected with pcDNA5/FRT-GAT-2 or empty vector (pcDNA5/FRT, control). Uptake was determined after 1, 2, 5, 10, 15, 20, 30, and 60 minutes. The uptake was linear up to 2 minutes. B) Uptake kinetics of GABA into HEK293-cells transiently transfected with pcDNA5/FRT-GAT-2. The K_m was determined to be 26.2 µM, V_{max} was 0.39 pmol/µg protein/minute.
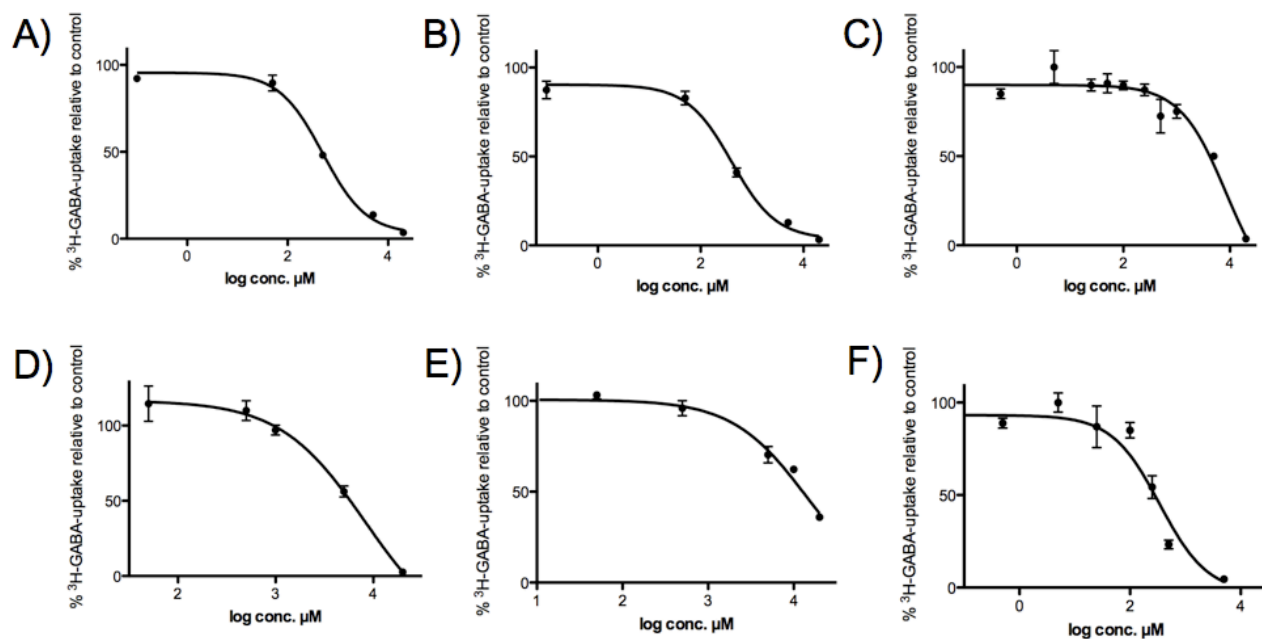
Fig S5. IC$_{50}$ curves for selected compounds. IC$_{50}$ values (in brackets) were determined for selected hits. A) 3-Aminobutanoic acid (499 µM), B) GABOB (402 µM), C) Pyridoxal-5'-phosphate (8.5 mM), D) 5-Aminolevulinic acid (8.2 mM), E) Baclofen (ca. 14 mM), F) Desipramine (344 µM). For baclofen, the solubility limit in the assay buffer was reached at a concentration of 20 mM.
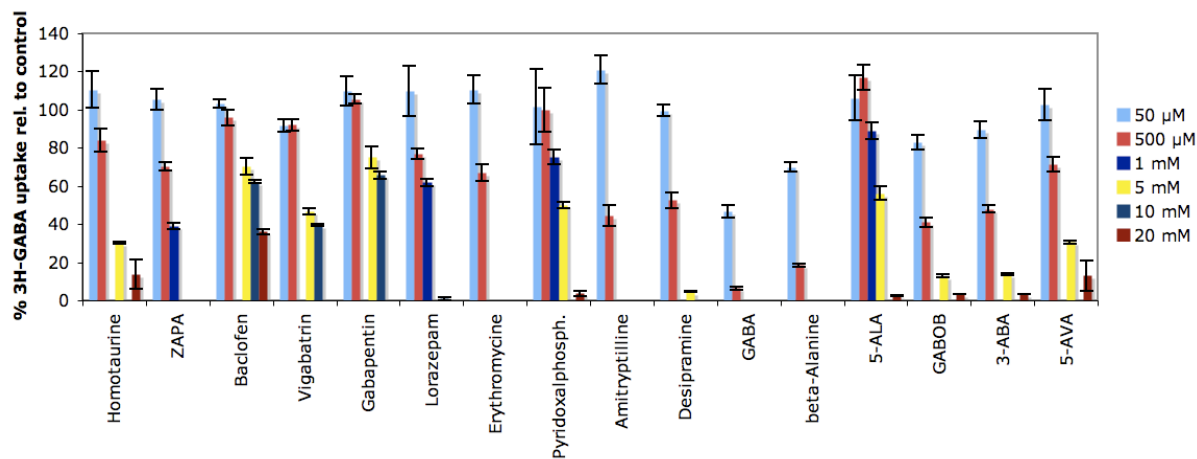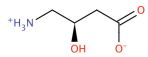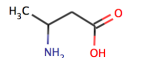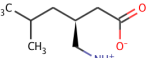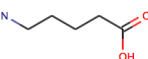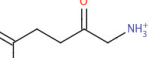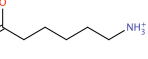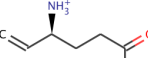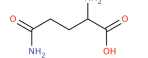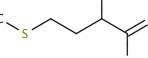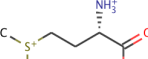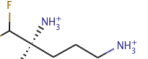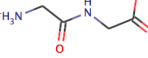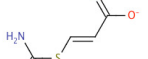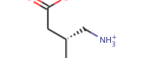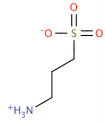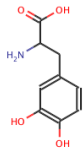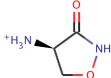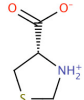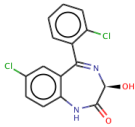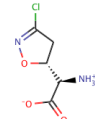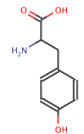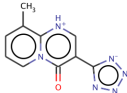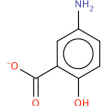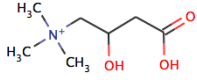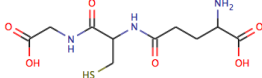
Fig S6. Additional data points for GAT-2 inhibitors. The inhibition of ^{3}H-GABA uptake was assessed at additional concentrations for compounds that were inhibitors of GAT-2 at 50, 500 or 5000 µM. ^{3}H-GABA uptake consistently decreased at higher concentrations up to 20 mM. All data are shown with standard error of the mean (SEM) bars.

TABLES

Name[a]	Function[b]	Tc[c]	Sketch[d]
Docking screen			
γ-Amino-β-hydroxybutyric acid (GABOB)	Anticonvulsant	0.93	
3-Aminobutanoic acid (3-ABA)	Fragment	0.87	
Pregabalin	Anticonvulsant; side effects related to renal excretion	0.84	
5-Aminovaleric acid (5-AVA)	Metabolite (Lysine degradation)	0.84	
Aminolevulinic acid (5-ALA)	Antineoplastic	0.79	
Amicar	Hemostatic	0.77	
Vigabatrin	Anticonvulsant	0.75	
L-Glutamine	Dietary supplement	0.72	
L-Methionine	Amino Acid	0.61	
Vitamin U	Gastrointestinal ulceration	0.60	
Ornidyl	Antineoplastic; Antiprotozoal	0.55	
Glycylglycine	Metabolite	0.50	
ZAPA	GABA analog for research	0.40	
Baclofen	Relaxant	0.36	

9

Tramiprosate (Homotaurine)	Alzheimer's disease,	0.30	
L-DOPA	Antiparkinsonian	0.30	
Cycloserine	Antibacterial; side effects include seizures	0.29	
Timonacic	Hepatic protectant	0.28	
Lorazepam	Relaxant	0.26	
Acivicin	Antineoplastic	0.20	
L-Tyrosine	Precursor of epinephrine, thyroid hormones, and melanin; antidepressant	0.19	
Gabapentin	Anticonvulsant	0.17	
Pemirolast	Antipruritics; anti-allergic agent	0.11	
Mesalazine	Anti-inflammatory; Crohn's disease	0.11	

Other tested molecules			
Name[a]	Function[b]	Tc[c]	Sketch[d]
Carnitine	Metabolite lysine and methionine and nutrition supplement	0.73	
Glutatione	Antioxidant	0.36	
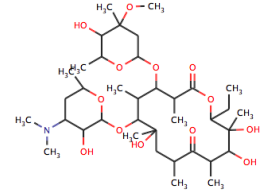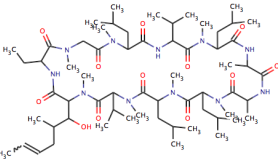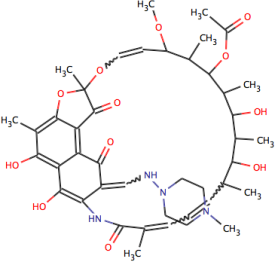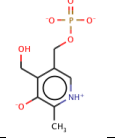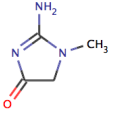
Name	Function	Tc	Sketch
Erythromycin	Antibiotic	0.25	
Cyclosporin A	Immunosuppressant	0.22	
Rifampicin	Antibiotic	0.20	
Pyridoxine phosphate	Vitamin B6 metabolism	0.16	
Creatinine	Metabolite and indicator for renal function	0.14	

Table S1: Small molecules tested in uptake kinetic assays.

a   *Name* is the generic or chemical name of the molecule; names of experimentally confirmed hits are marked in bold font.

b   *Function* gives the pharmacological function of the drug or the physiological function of the metabolites, when applicable.

c   *Tc* is the Tanimoto coefficient calculated relying on the Daylight fingerprints. Tc values of < 0.5 suggest that the molecule is chemically different from all known GAT-2 ligands.

d   *Sketch* provides the 2D sketch of the molecule.

LeuT	SLC6A13 (GAT-2)	SLC6A12 (BGT1)	SLC6A11 (GAT-3)	SLC6A1 (GAT-1)
S355	S390	S395	S410	S396
G24	**G51**	G55	G69	G63
G26	G53	G57	G71	G65
V104	L125	L129	L143	L136
Y108	Y129	Y133	Y147	Y140
F253	F288	F293	F308	F294
T254	S289	S294	S309	S295
N21	**E48**	E52	E66	Y60
S256	A291	A296	A311	G297
F259	L294	Q299	L314	L300
A261	C296	C301	C316	S302
I359	C394	C399	C414	T400
-	S353	S358	S373	S359
L400	F443	F448	F463	F447
Y107	Y128	Y132	Y146	Y139
L29	W56	W60	W74	W68
I111	**V132**	I136	I150	I143
L25	L52	L56	L70	L64
D404	D447	D452	D467	D451
T409	S452	S457	S472	S456

Table S2: Comparison between the four GATs and LeuT

Residues in close proximity to the S1 binding site are listed. Where residues that were experimentally tested using site-directed mutagenesis are marked in bold font.

REFERENCES

1.   Nyola, A., Karpowich, N. K., Zhen, J., Marden, J., Reith, M. E., and Wang, D. N. (2010) *Curr Opin Struct Biol* **20**, 415-422
2.   Krishnamurthy, H., and Gouaux, E. (2012) *Nature* **481**, 469-474
3.   Yamashita, A., Singh, S. K., Kawate, T., Jin, Y., and Gouaux, E. (2005) *Nature* **437**, 215-223
4.   Singh, S. K., Piscitelli, C. L., Yamashita, A., and Gouaux, E. (2008) *Science* **322**, 1655-1661
5.   Beuming, T., Shi, L., Javitch, J. A., and Weinstein, H. (2006) *Mol Pharmacol* **70**, 1630-1642
6.   Krivov, G. G., Shapovalov, M. V., and Dunbrack, R. L., Jr. (2009) *Proteins* **77**, 778-795
7.   Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) *Journal of chemical theory and computation* **4**, 435-447
8.   Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., and Shaw, D. E. (2010) *Proteins* **78**, 1950-1958
9.   Shaw, D. E., Maragakis, P., Lindorff-Larsen, K., Piana, S., Dror, R. O., Eastwood, M. P., Bank, J. A., Jumper, J. M., Salmon, J. K., Shan, Y., and Wriggers, W. (2010) *Science* **330**, 341-346
10.   Hess, B. (2007) *Journal of chemical theory and computation* **4**, 116-122
11.   Huang, N., Shoichet, B. K., and Irwin, J. J. (2006) *J Med Chem* **49**, 6789-6801
12.   Irwin, J. J., and Shoichet, B. K. (2005) *J Chem Inf Model* **45**, 177-182
13.   Fan, H., Irwin, J. J., Webb, B. M., Klebe, G., Shoichet, B. K., and Sali, A. (2009) *J Chem Inf Model* **49**, 2512-2527
14.   Tusnady, G. E., Kalmar, L., and Simon, I. (2008) *Nucleic acids research* **36**, D234-239
15.   Clamp, M., Cuff, J., Searle, S. M., and Barton, G. J. (2004) *Bioinformatics* **20**, 426-427