PNAS
www.pnas.org

Supplementary Information for

Structural basis for substrate specificity of heteromeric transporters of neutral amino acids

Carlos F Rodriguez, Paloma Escudero-Bravo, Lucía Díaz, Paola Bartoccioni, Carmen García-Martín, Joan G Gilabert, Jasminka Boskovic, Víctor Guallar, Ekaitz Errasti-Murugarren, Oscar Llorca and Manuel Palacín

Ekaitz Errasti-Murugarren; Oscar Llorca; Manuel Palacín

Email: ekaitz.errasti@irbbarcelona.org, ollorca@cnio.es and manuel.palacin@irbbarcelona.org

This PDF file includes:

SI Materials and Methods

Subcloning of hLAT2 and CD98hc proteins. cDNA encoding for the human LAT2 transporter protein was subcloned by amplification from human Strep-TagII-LAT2 (N-terminally tagged) in pcDNA3.1+ (1) with KOD polymerase (Toyobo, Osaka, Japan), following manufacturer's instructions, using the following primer pair: Fw: AGGAGATATACCATGgaagaaggagccaggcaccgaaacaacaccg; Rv: CTTCCAGACCGCTTGAgggctggggctgccccgccac (lower case indicates gene-specific sequence, upper case indicates plasmid-specific sequence). The resulting 1605 bp PCR product was then treated with DpnI to remove the template and purified using AmPure magnetic beads as per the manufacturer's instructions (Beckman Coulter, Brea, USA). The purified PCR product was then cloned into the KpnI and BmtI-cut pPEU24TT plasmid containing a C-terminal 3C site-eGFP-10xHis fusion by InFusion (Takara, Kyoto, Japan) following manufacturer's instructions. Similarly, human CD98hc (isoform f) cDNA was amplified from the His-CD98hc (N-terminally tagged) in pcDNA4-His-MaxC (2) by KOD polymerase using the following primer pair: Fw: TCGAAAAAAGCAGCGGCatgagccaggacaccgaggtggatatg; Rv: GTGATGGTGATGTTTAggccgcgtaggggaagcggag (lower case indicates gene-specific sequence, upper case indicates plasmid-specific sequence). The 1600bp product was then purified (as hLAT2 PCR product) and cloned into KpnI and PmeI-cut pPEU22TT plasmid by InFusion. pPEU22TT contains an N-terminal OneStrep tag (StrepII-spacer-StrepII) and in this instance the resident C-terminal His-tag was not used as a stop codon was inserted into the reverse primer. Both pPEUTT plasmid variants are derived from the pTT plasmid containing the OriP origin of replication to allow episomal replication in the 293-6E cells (3) (please see https://www.irbbarcelona.org/en/research/protein-expression for pPEUTT plasmid details). All DNA constructs were fully verified by sequencing before use.

Transient transfection and production of hLAT2/CD98hc heterodimer in HEK293-6E cells. Transient production of hLAT2/CD98hc heterodimer in a suspension of HEK293-6E cells was performed as previously described (3). Briefly, cells were grown to $1.5x10^6$ cells/ml in 2 l Erlenmeyer flasks with ventilation membrane caps (Triforest Plasticware Irvine, USA) with a working volume of 600 ml culture/flask at 37°C and 5% CO_2 with shaking at 120 rpm. PEI:DNA polyplexes were prepared by mixing a total of 1 µg (1:1 w/w hLAT2 and CD98hc) plasmid-DNA and 4 µg PEI-MAX 40000 (Polysciences Europe GmbH, Germany) per ml culture in approximately $1/10^{th}$ of total culture volume of fresh media. Polyplexes were allowed to form for 3 min at room temperature with intermittent mixing before addition to the cells. After addition of the DNA:PEI complexes, the cells were incubated for a further 48 h with shaking at 37°C and 5% CO_2. They were then harvested by centrifugation at 500 x *g* for 15 min, and the cell pellets were washed twice with 50 ml of PBS and then stored at -80ºC until use.

hLAT2/CD98hc purification for cryo-EM. All subsequent steps were carried out at 4°C. Whole HEK293-6E cells expressing hLAT2-3C-GFP-10His/OneStrep-CD98hc heterodimer were solubilized using 2% (w/v) digitonin (Merck, Darmstadt, Germany) for 2 h in solubilization buffer (20 mM Tris-Base, 150 mM NaCl, pH 7.4). Following ultracentrifugation (200,000 × *g* for 1 h), the soluble fraction was incubated for 2 h with Strep-Tactin Superflow resin (IBA Lifesciences, Göttingen, Germany) equilibrated with purification buffer (20 mM Tris-Base, 150 mM NaCl, 0.1% digitonin, pH 7.4). Protein-bound resin was washed twice with 12 column volumes (CVs) of purification buffer. Protein was eluted in Strep elution buffer (20 mM Tris-Base, 150 mM NaCl, 2.5 mM D-desthiobiotine, pH 7.4), and eluted protein was incubated for 2 h with Ni^{2+}-NTA Superflow beads (Qiagen, Hilden, Germany) equilibrated in Nickel washing buffer (20 mM Tris-Base, 150 mM NaCl, 0.1% digitonin, 20 mM imidazole, pH 7.4). Protein-bound beads were washed twice with 12 CVs of washing buffer supplemented with 20 mM and 40 mM imidazole. The beads were then washed once with 12 CVs of HRV-3C buffer (20 mM Tris-Base, 150 mM NaCl, 0.1% digitonin, 20 mM imidazole, 0.5 mM EDTA, pH 7.4) before on-column cleavage with HRV-3C protease (IRB Barcelona Protein Expression Core

Facility, Barcelona, Spain) for 16 h. Column flowthrough containing cleaved hLAT2/CD98hc was concentrated by centrifugation in an Amicon Ultra-15 filter unit (100,000 kDa molecular weight cut-off; Millipore, Temecula, CA) at 3,220 × *g* until reaching 1.2 mg/ml. The heterodimer was subjected to size exclusion chromatography (SEC) on a Superdex 200 5/150 GL column (GE Healthcare, Chicago, USA) equilibrated with 20 mM Tris-Base, 150 mM NaCl, and 0.1% digitonin, pH 7.4. The peak fraction was collected, concentrated and used for grid preparation.

hLAT2/CD98hc purification for amino acid transport assays. All subsequent steps were performed at 4°C. Whole HEK293-6E cells expressing the hLAT2-3C-GFP-10His/OneStrep-CD98hc heterodimer were solubilized using 2% (w/v) digitonin (Merck, Darmstadt, Germany) for 2 h in solubilization buffer (20 mM Tris-Base, 150 mM NaCl, pH 7.4). Following ultracentrifugation (200,000 × *g* for 1 h), the soluble fraction was incubated for 2 h with Strep-Tactin Superflow resin (IBA Lifesciences, Göttingen, Germany) equilibrated with purification buffer (20 mM Tris-Base, 150 mM NaCl, 0.1% digitonin, pH 7.4). Protein-bound resin was washed twice with 12 CVs of purification buffer. Protein was eluted in Strep elution buffer (20 mM Tris-Base, 150 mM NaCl, 2.5 mM D-desthiobiotine, pH 7.4), and eluted protein was incubated for 2 h with Ni^{2+}-NTA Superflow beads (Qiagen, Hilden, Germany) equilibrated in Nickel washing buffer (20 mM Tris-Base, 150 mM NaCl, 0.1% digitonin, 20 mM imidazole, pH 7.4). Protein-bound beads were washed three times with 20 CVs of washing buffer before elution with washing buffer supplemented with 350 mM imidazole. The purified protein was desalted on a PD minitrap G-25 desalting column (GE Healthcare, Chicago, USA) and centrifuged in an Amicon Ultra (100,000 kDa molecular weight cut-off; Millipore) at 3,220 × *g* until the desired concentration was reached.

Reconstitution into proteoliposomes. The liposomes, composed of a 5:1 ratio of L-α-phosphatidylcholine type II-S (Sigma-Aldrich, Sant Louis, USA) to brain total lipid extract, were prepared as previously described (4). The lipids were dried under N_2 and resuspended in reconstitution buffer (20 mM Na_3PO_4, 150 mM NaCl, pH 7) at a concentration of 20 mg/ml. After 10 rounds of freezing and thawing by liquid nitrogen, the liposomes were extruded through 0.4 µm polycarbonate membranes (Sigma-Aldrich, Sant Louis, USA) up to 21 times. The purified hLAT2-GFP/CD98hc protein was added to reach the desired protein to lipid ratio of 1:20 (w:w). To destabilize the liposomes, 1.25% ß-D-octylglucoside was added for 30 min, and the mixture was then incubated on ice for 5 min with occasional agitation. The protein was incubated with liposomes for 60 min at a concentration of 0.8 µg of protein per mg of lipid. Digitonin and ß-D-octylglucoside were removed by overnight dialysis at 4°C against 100 volumes of reconstitution buffer. Proteoliposome (PL) suspensions were frozen in liquid nitrogen and stored at -80°C until use.

Amino acid transport assays in proteoliposomes. For uptake experiments, liposomes were filled with 5 mM L-valine by adding the amino acid to the PL suspension, which was then subjected to three freeze/thaw cycles. The extraliposomal amino acid content was then removed by ultracentrifugation (100,000 × *g* for 1 h at 4°C) and PLs were resuspended to one-third of the initial volume with reconstitution buffer. Amino acid uptake assays were initiated after mixing 10 µl of cold PLs with 180 µl of transport buffer (20 mM Na_3PO_4, 150 mM NaCl, pH 7) plus 0.5–1 µCi/180 µl of radiolabeled L-valine (Perkin Elmer, Waltham, USA) supplemented with 10 µM unlabeled L-valine. This mixture was then incubated at room temperature for the indicated periods. Transport experiments were stopped by the addition of 2 ml of ice-cold stop buffer (reconstitution buffer containing 5 mM L-valine) and filtration through 0.45 µm pore-size membrane filters (Sartorius Stedim Biotech, Cedex, France). Filters were then washed twice with 2 ml of stop buffer and dried, and the trapped radioactivity was counted. For transport assays we used hLAT2-GFP/CD98hc since GFP-untagged hLAT2/CD98hc was unstable during our protocol to reconstitute it in proteoliposomes. Transport values are expressed in pmol of L-valine per µg of protein and for the indicated time. hLAT2-GFP/CD98hc protein in PLs was determined by silver staining (Pierce, Rockford, USA) in SDS-PAGE gels compared with known amounts of hLAT2-GFP/CD98hc in digitonin micelles, determined by nanodrop and loaded in the

same gel. hLAT2-GFP/CD98hc overexpressed in HeLa cells shows the same transport activity and substrate selectivity as the one shown by the non-GFP-tagged version.

1. A. Rosell *et al.*, Structural bases for the interaction and stabilization of the human amino acid transporter LAT2 with its ancillary protein 4F2hc. *Proc Natl Acad Sci U S A* **111**, 2966-2971 (2014).
2. J. Fort et al., The structure of human 4F2hc ectodomain provides a model for homodimerization and electrostatic interaction with plasma membrane. J Biol Chem 282, 31444-31452 (2007).
3. M. Loignon et al., Stable high volumetric production of glycosylated human recombinant IFNalpha2b in HEK293 cells. BMC Biotechnol 8, 65 (2008).
4. Y. Lee *et al.*, Cryo-EM structure of the human L-type amino acid transporter 1 in complex with glycoprotein CD98hc. *Nat Struct Mol Biol* **26**, 510-517 (2019).

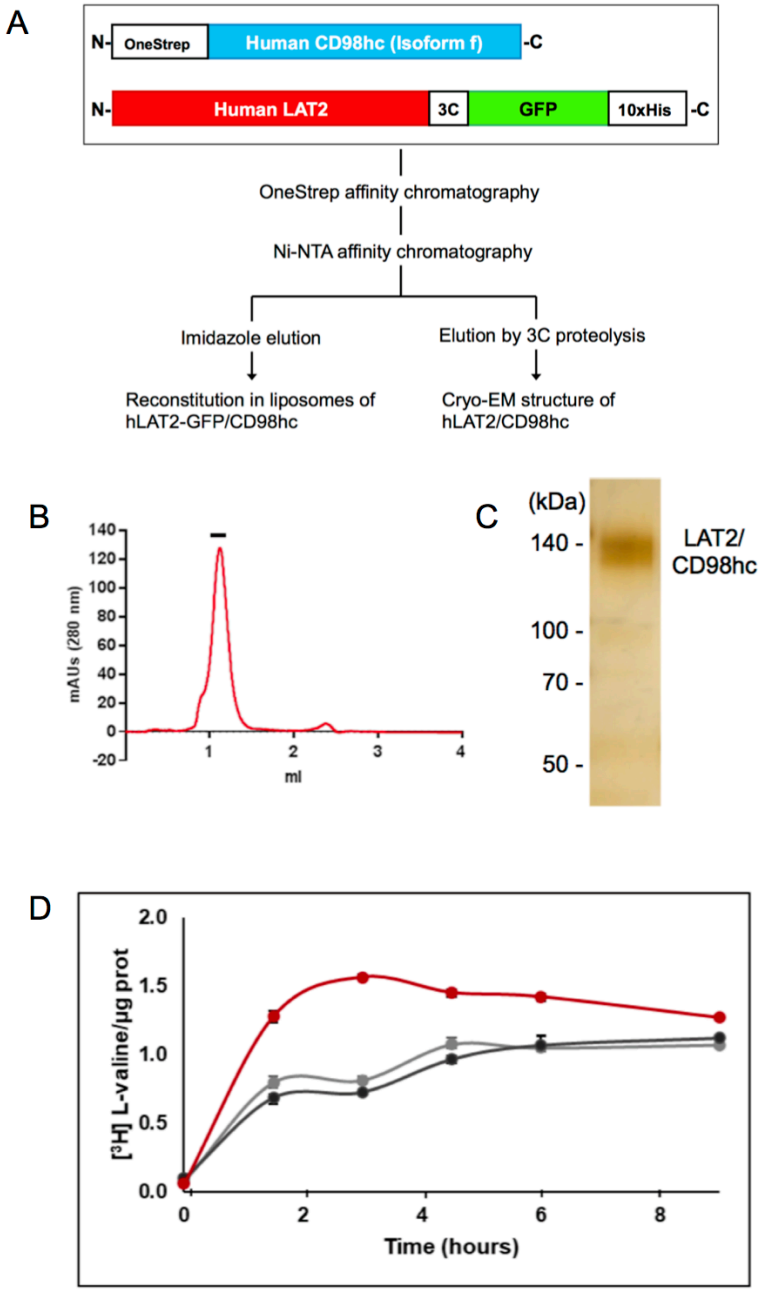

Figure S1. hLAT2/CD98hc purification and reconstitution in liposomes. (**A**) Schematic representation of the DNA constructions used for protein expression and purification protocols indicating key steps and further uses of the purified protein. (B) SEC of digitonin-solubilized and purified hLAT2/CD98hc protein. mAUs, milli absorbance units. (C) Silver staining of the SDS-PAGE gel of purified hLAT2/CD98hc protein. Heterodimeric hLAT2/CD98hc protein migrates as a prominent band at ~140 kDa. (D) Time course (0–9 h) of 10 µM [^{3}H] L-valine influx (pmol/µg protein) into hLAT2/CD98hc-GFP-proteoliposomes containing 5 mM L-valine (red circles) or no amino acid (black circles). Uptake in liposomes containing no protein is also shown (gray circles). Data correspond to a representative experiment, performed using four replicates. Two additional experiments gave similar results.

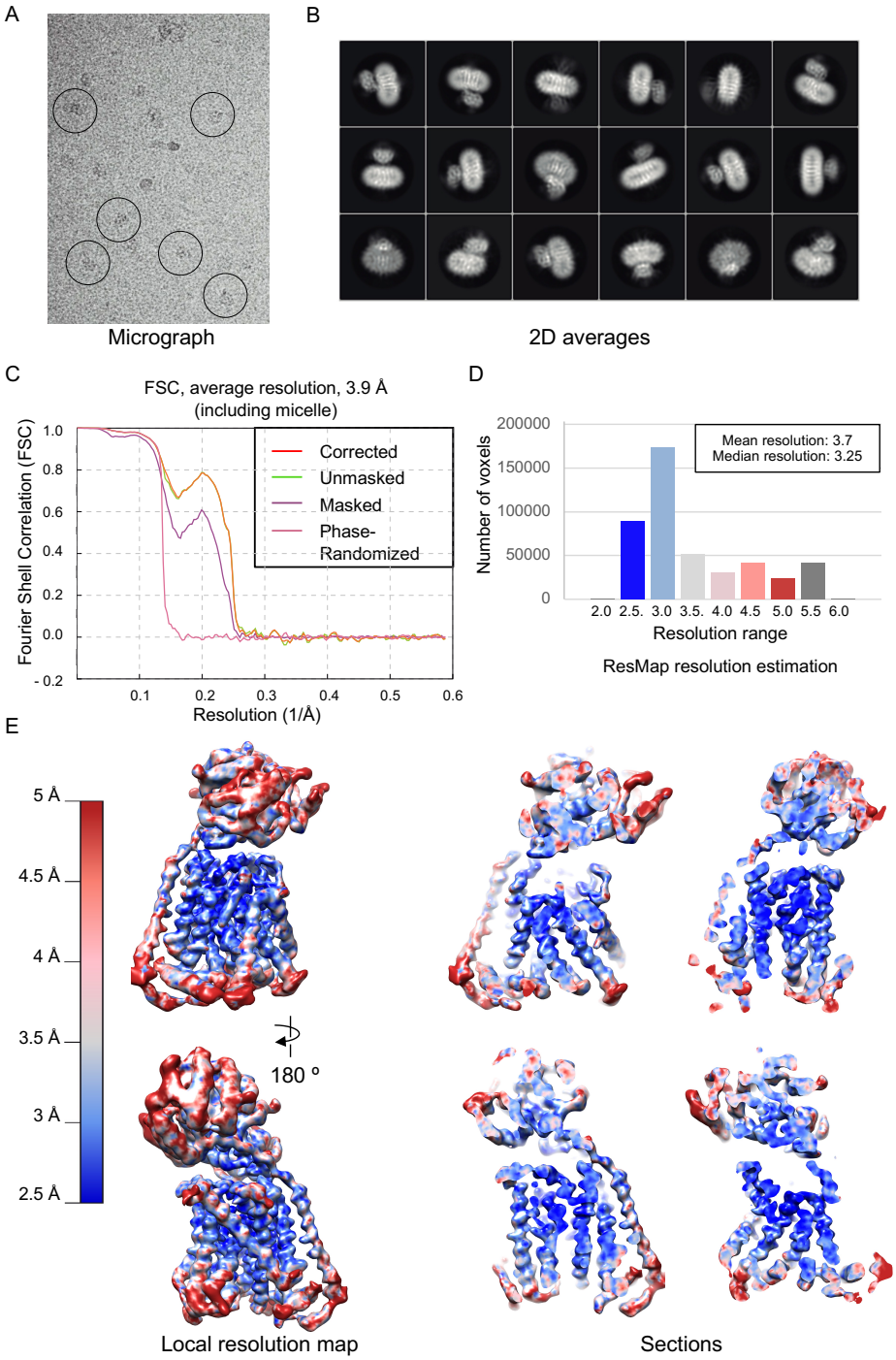

Figure S2. Cryo-EM of hLAT2/CD98hc and resolution estimation. (A) Representative cryo-EM micrograph of hLAT2/CD98hc. Some particles are highlighted within circles. (B) 2D averages of hLAT2/CD98hc showing a visible ectodomain and transmembrane helices embedded within the micelle. hLAT2/CD98hc was visualized in several orientations. (C) Fourier Shell Correlation (FSC) curves estimating the average resolution of the cryo-EM volume of hLAT2/CD98hc using a mask that includes the whole protein and the density of the micelle. (D) Plot of the number of voxels at each resolution range in the cryo-EM reconstruction of hLAT2/CD98hc. Average and median resolution estimates calculated using ResMap as indicates as inserted text. (E) Local resolution estimates of hLAT2/CD98hc cryo-EM map as provided by RELION. Color scale is shown on the left.

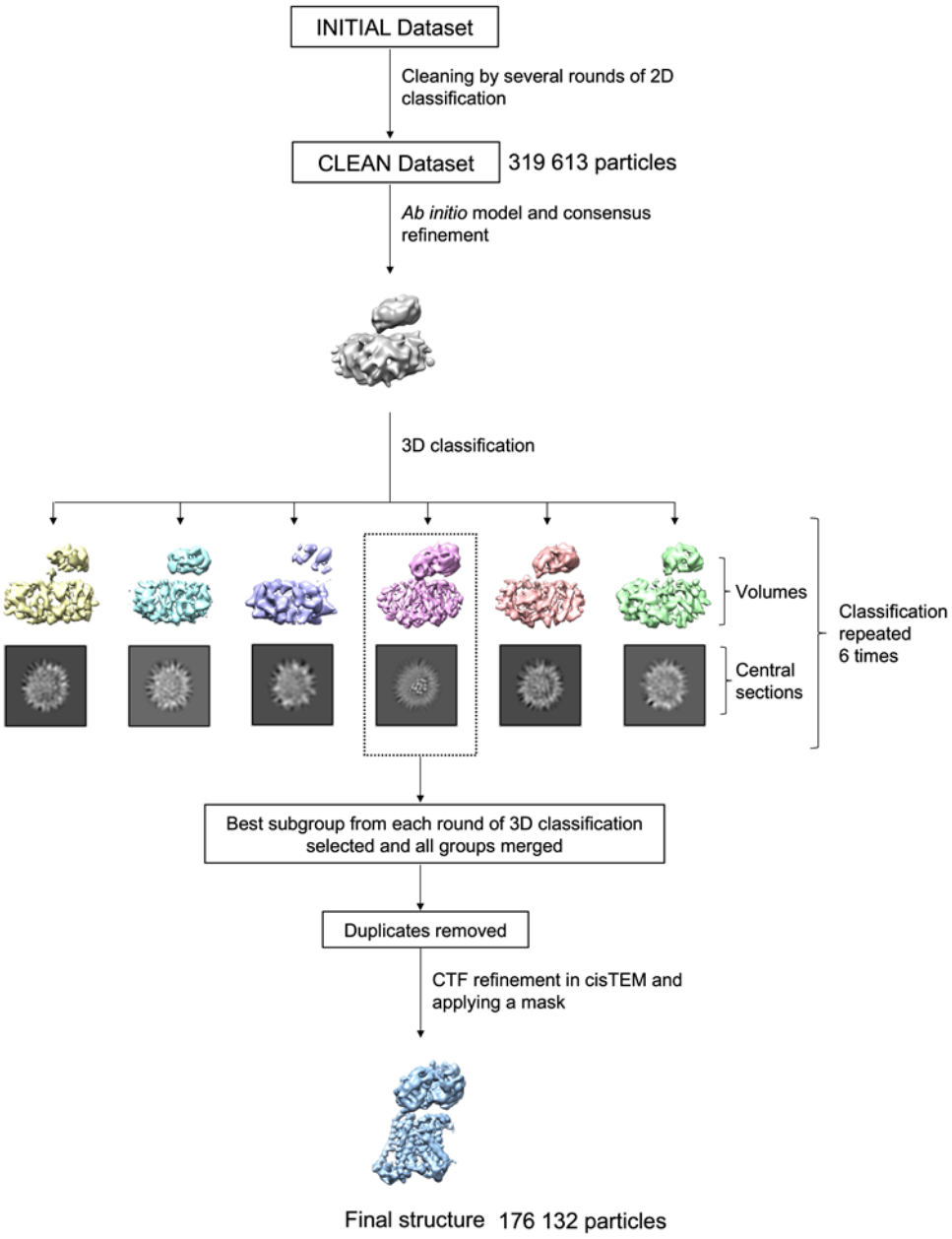

Figure S3. Image processing workflow of the cryo-EM images of hLAT2/CD98hc. Workflow of the image processing strategy used in this study. The initial data set was first subjected to multiple rounds of 2D classification and averaging to remove low quality particles. A clean dataset containing 319,613 particles was then used to generate an ab initial model and a consensus refined volume. This was the starting point for several rounds of 3D classification and the best subgroup was selected based on the quality of the volumes obtained during classification. The best subgroup was then further classified by multiple rounds of 3D classification of the same images, pooling the best groups and removing duplicates. A subset of 176,132 particles refined to high resolution and was selected as the final cryo-EM map.

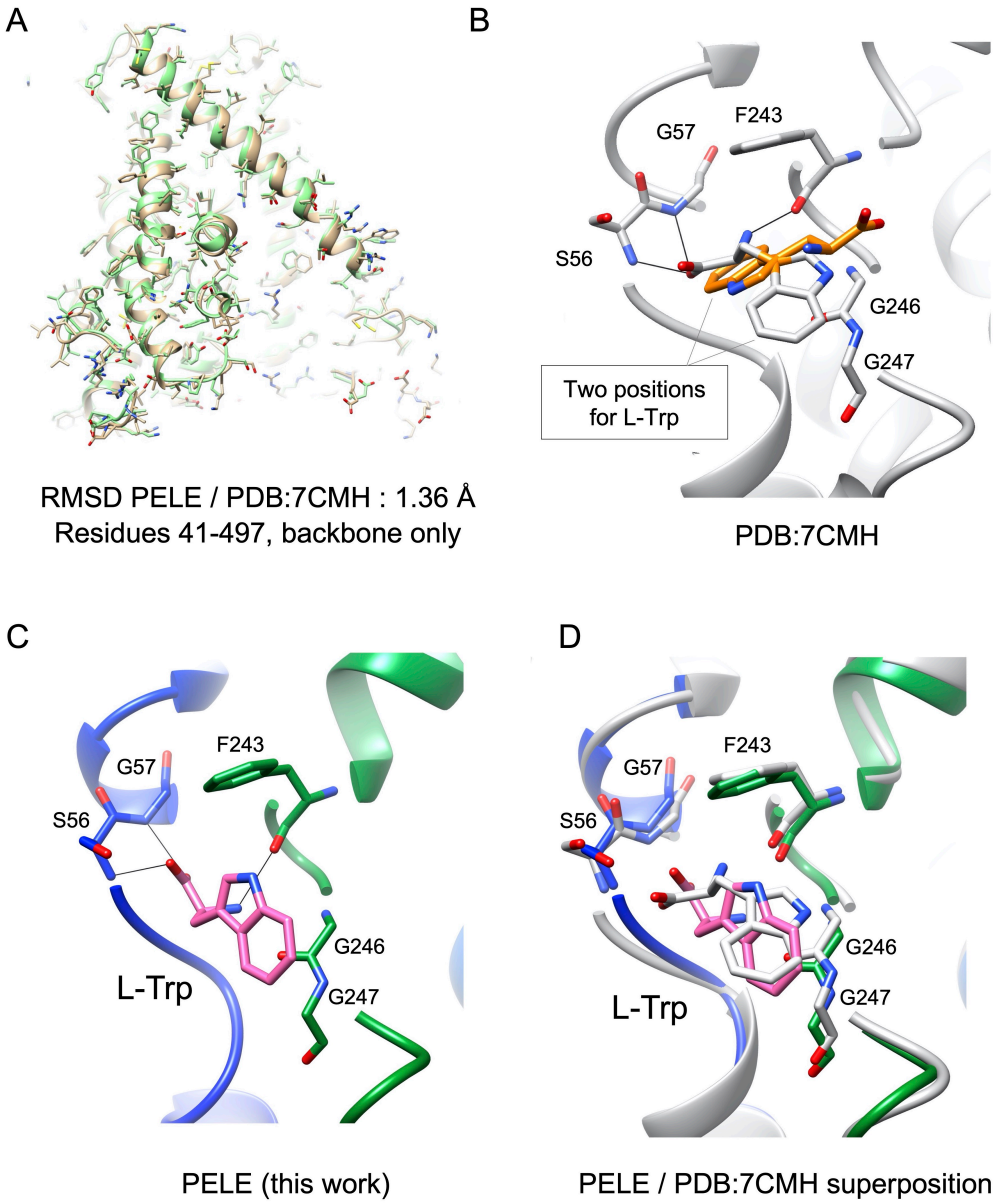

Figure S4. **Comparison between our predictions for hLAT2/CD98hc bound to L-tryptophan obtained using molecular docking and PELE studies (a representative pose selected), and the cryo-EM structure (PDB 7CMH).** (A) Root Mean Square Deviation (RMSD) between the two structures. One view of a region from both structures superimposed. Our prediction (green color), 7CMH (brown color). (B) View of the substrate-binding site for L-tryptophan in the cryo-EM structure (PDB 7CMH). The density for tryptophan in the cryo-EM map was insufficient to define a unique position for the substrate and two possible orientations for L-tryptophan were suggested, here colored in gray and orange. O and N atoms of substrates and residues are shown in red and blue, respectively. Black lines connect atoms located at H bond distance. (C) View of the substrate binding from the PELE analysis for L-tryptophan. Carbon atoms of the indicated residues follow the color codes for hLAT2 helices and residues as used in Figure 1. C atoms of the substrates are shown in pink, whereas O and N atoms of substrates and residues are shown in red and blue, respectively. Black lines connect atoms located at H bond distance. (D) Superimposition of the structures for the substrate binding site for L-tryptophan obtained in our work (in rainbow colors as in Figure 1) and in the cryo-EM structure (PDB 7CMH) in grey, in this later one showing only the L-Trp position that matches the position found in our PELE analysis.

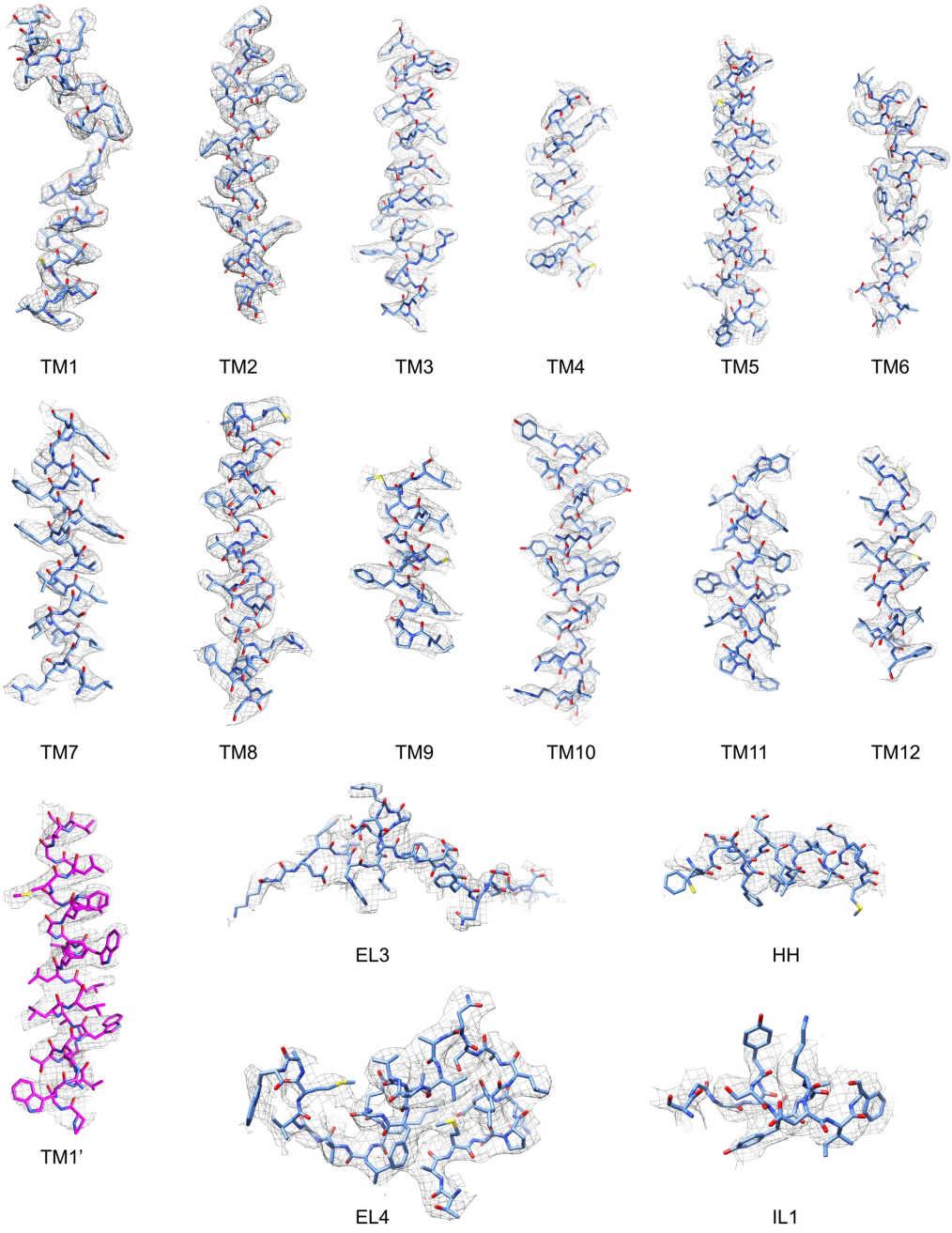

Figure S5. High-resolution features in the cryo-EM map of hLAT2/CD98hc. High-resolution structural features, including clear densities for side chains are shown as a mesh for the cryo-EM density with the structural model fitted. Figure shows representative region for each of the transmembrane helices (TM), external loops (EL), internal loops (IL) and the horizontal α-helix (HH) at the C-terminus of hLAT2.

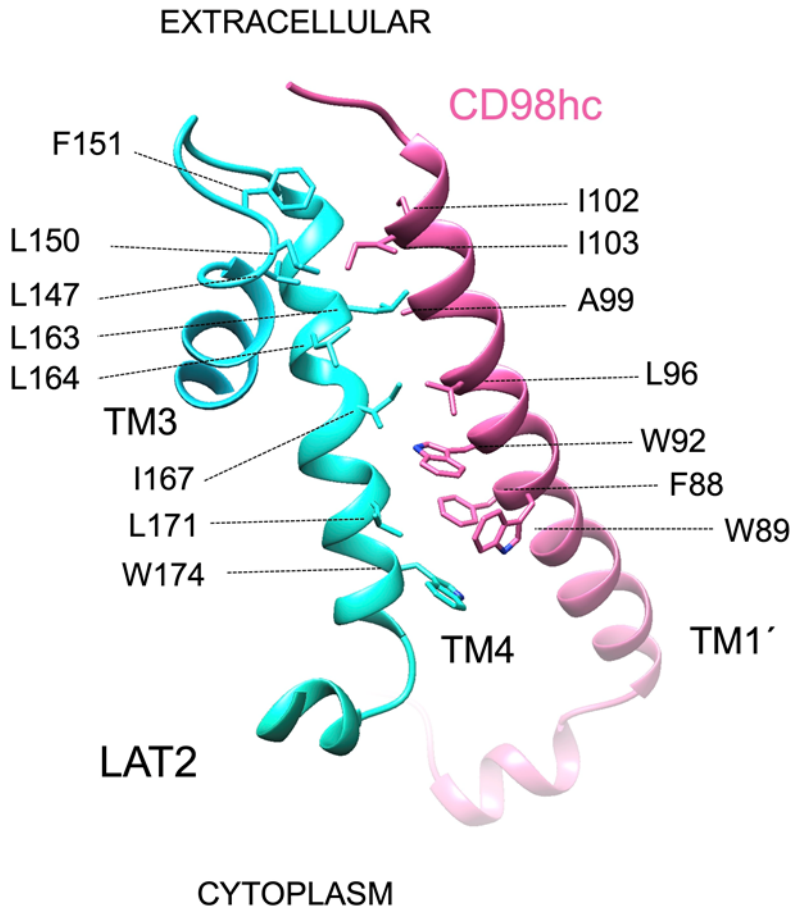

Figure S6. Contacts between CD98hc and LAT2 in hLAT2/CD98hc heterodimer. Figure shows the contacts between the transmembrane helix TM1´ in CD98hc with TM3 and TM4 of hLAT2. TM3, TM4 and TM1´are shown as ribbons and residues involved in interactions between the two subunits are highlighted. Color codes for hLAT2 helices and residues in all the panels are as used in Figure 1. Nitrogen atoms are shown in blue.

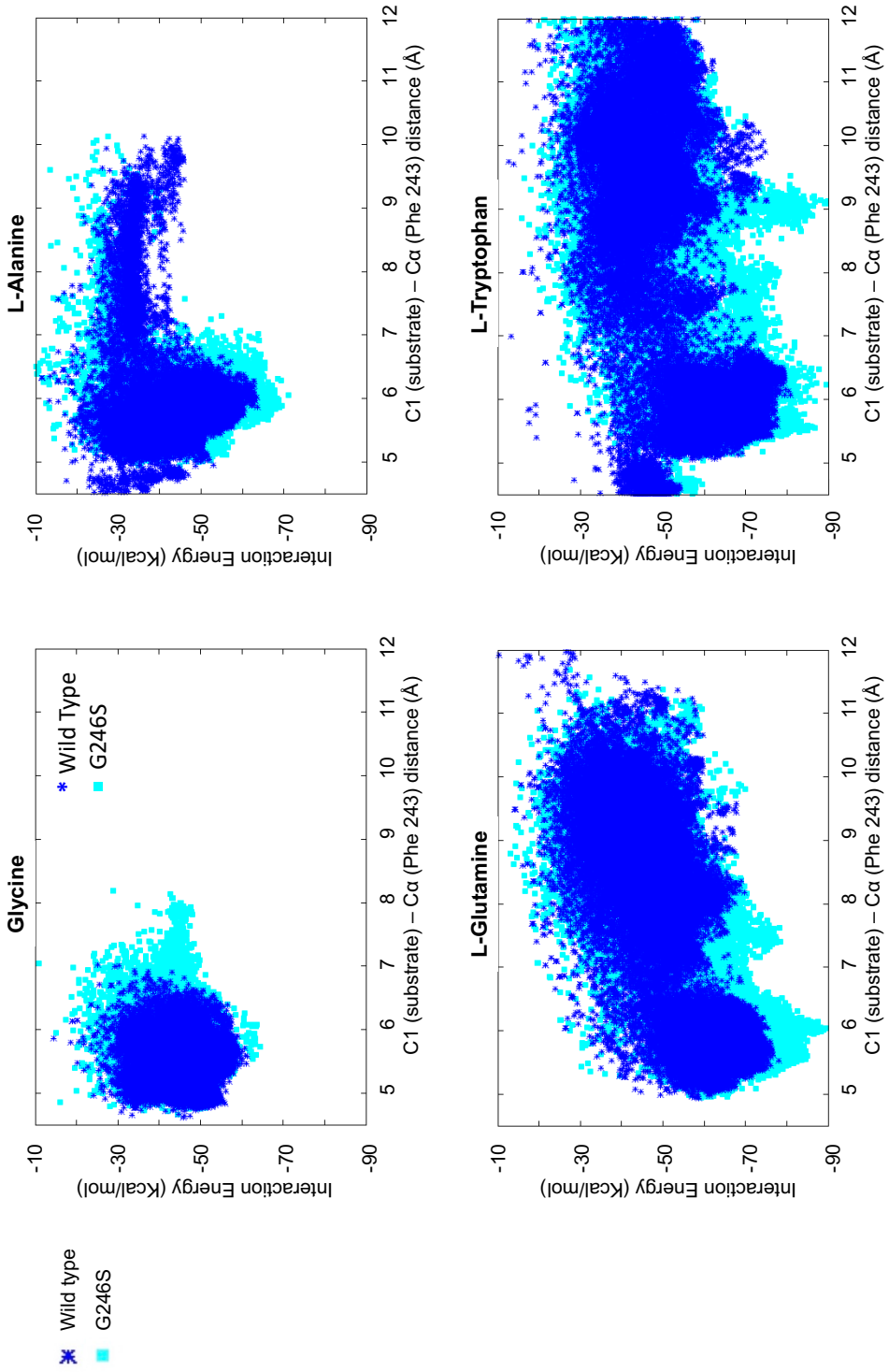

Figure S7. Energy landscape of the substrate binding to wild type hLAT2/CD98hc and G246S mutant in inward-facing conformation. Results from PELE simulations displaying interaction energies for the indicated amino acid substrates against the distance between alpha carbons of the substrate and of residue Phe 243. The minimal binding energy within ~6 Å of the indicated distance corresponds to the canonical binding site. Simulations for wild type hLAT2 (dark blue asterisk) and G246S mutant (cyan squares) are shown; 160,000 Monte Carlo PELE steps were attempted (~55,000 sampled binding modes) per condition.

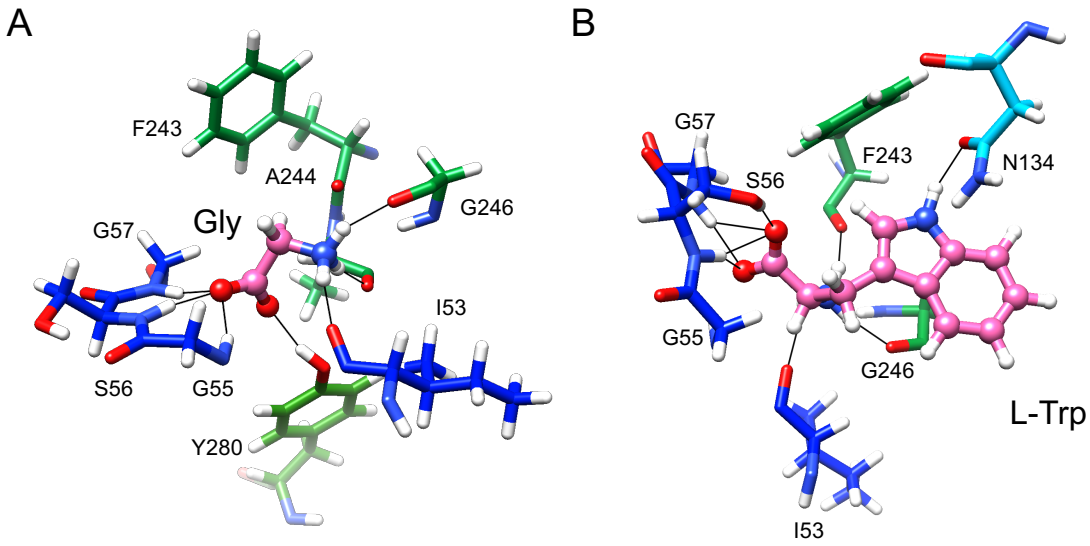

Figure S8. Selected poses in the canonical binding site from the PELE analysis for glycine and L-tryptophan substrates, respectively. Carbon atoms of the indicated residues follow the color codes for hLAT2 helices and residues as used in Figure 1. C atoms of the substrates are shown in pink, whereas O and N atoms of substrates and residues are shown in red and blue, respectively. Black lines connect atoms located at H bond distance.

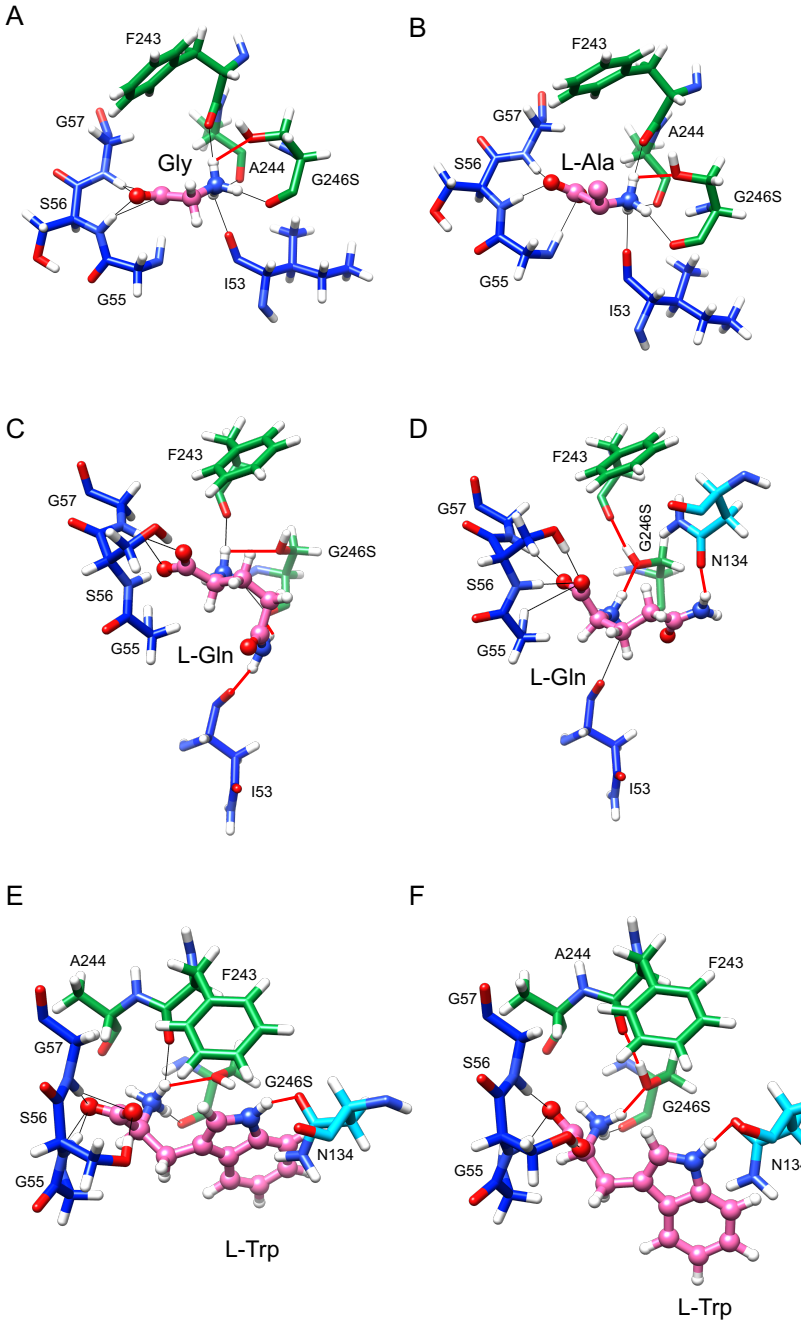

Figure S9. Selected poses in the substrate-binding site of G246S hLAT2 mutant from the PELE analysis for the indicated amino acid substrates. (A), (B) Canonical poses for glycine and L-alanine substrates. (C), (D) Selected poses within the substrate-binding site for L-glutamine before and after shifting of the substrate, respectively. (E), (F) Selected poses within the substrate-binding site for L-tryptophan before and after shifting of the substrate, respectively. In the shifted position (D and F), the hydroxyl group of G243S mutant bridges the connection of the α amino nitrogen atom of L-glutamine and L-tryptophan with residue Phe 243. In these shifted poses, the lateral chain of L-glutamine and L-tryptophan establishes H bonds with residue N134. H bonds mediated by the hydroxyl oxygen atom of G246S and the lateral chains of the substrates are highlighted with a red line. Other H bonds are indicated with a black line. Carbon atoms of the indicated residues follow the color codes for hLAT2 helices and residues as used in Figure 1. C atoms of the substrates are shown in pink, whereas O and N atoms of substrates and residues are shown in red and blue, respectively.

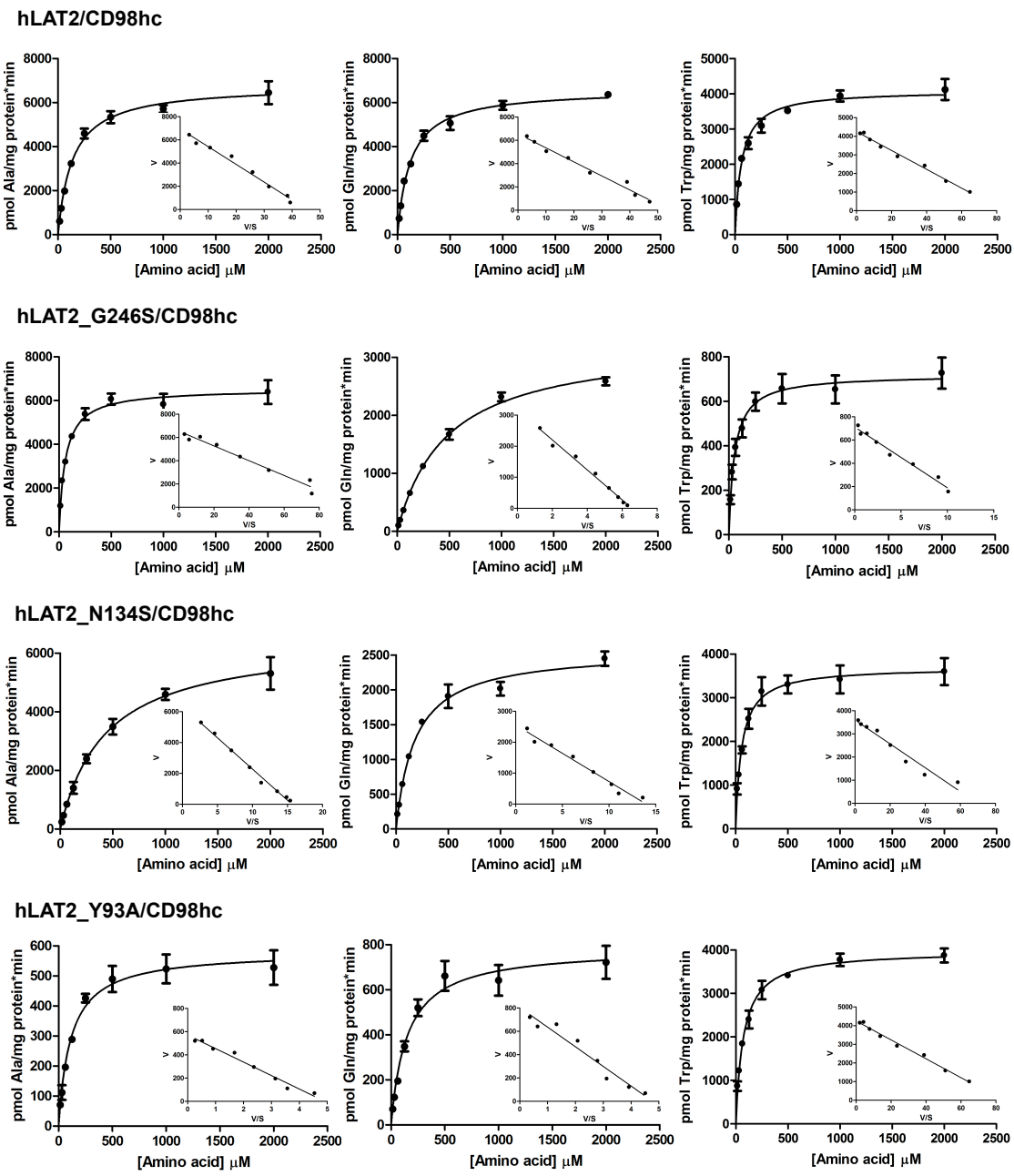

Figure S10. Kinetics of the hLAT2/CD98hc-induced uptake of the indicated L-amino acids. L-alanine (left), L-glutamine (middle) and L-tryptophan (right) kinetic analysis of wild type (upper), G246S (second line), N134S (third line) and Y93A (lower) versions of hLAT2. Data (mean±SEM) from representative experiments run in triplicates are shown. Inset: Eadie-Hofstee transformation.

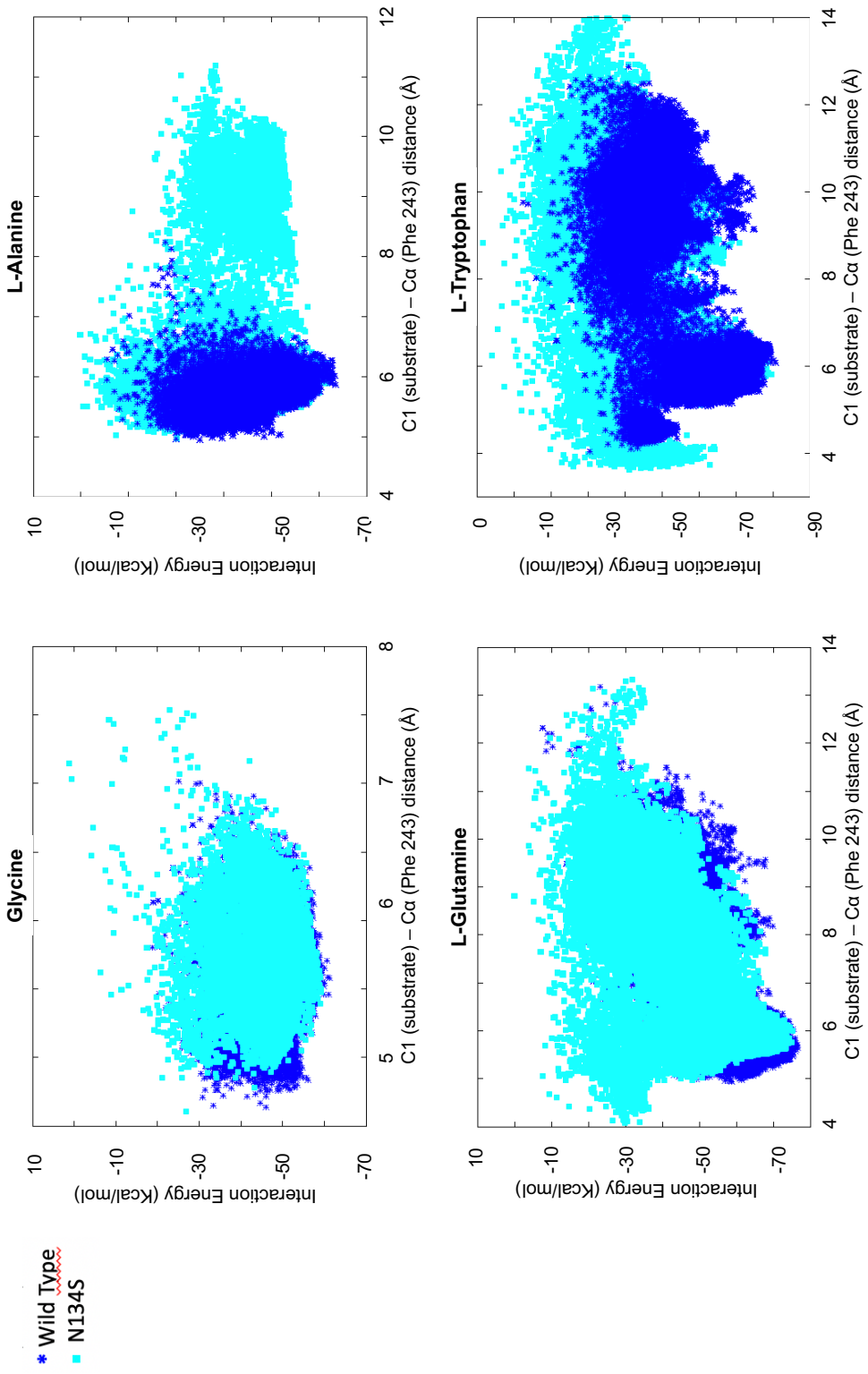

Figure S11. Energy landscape of the substrate binding to wild type hLAT2/CD98hc and N134S mutant in inward-facing conformation. Results from PELE simulations displaying interaction energies of the substrate binding for the indicated amino acid substrates against the distance between alpha carbon atoms of the substrate and of residue Phe 243. The minimal binding energy within ~6 Å of the indicated distance corresponds to the canonical binding site. Simulations for wild type hLAT2 (dark blue asterisk) and N134S mutant (cyan squares) are shown; 160,000 Monte Carlo PELE steps were attempted (~55,000 sampled binding modes) per condition.

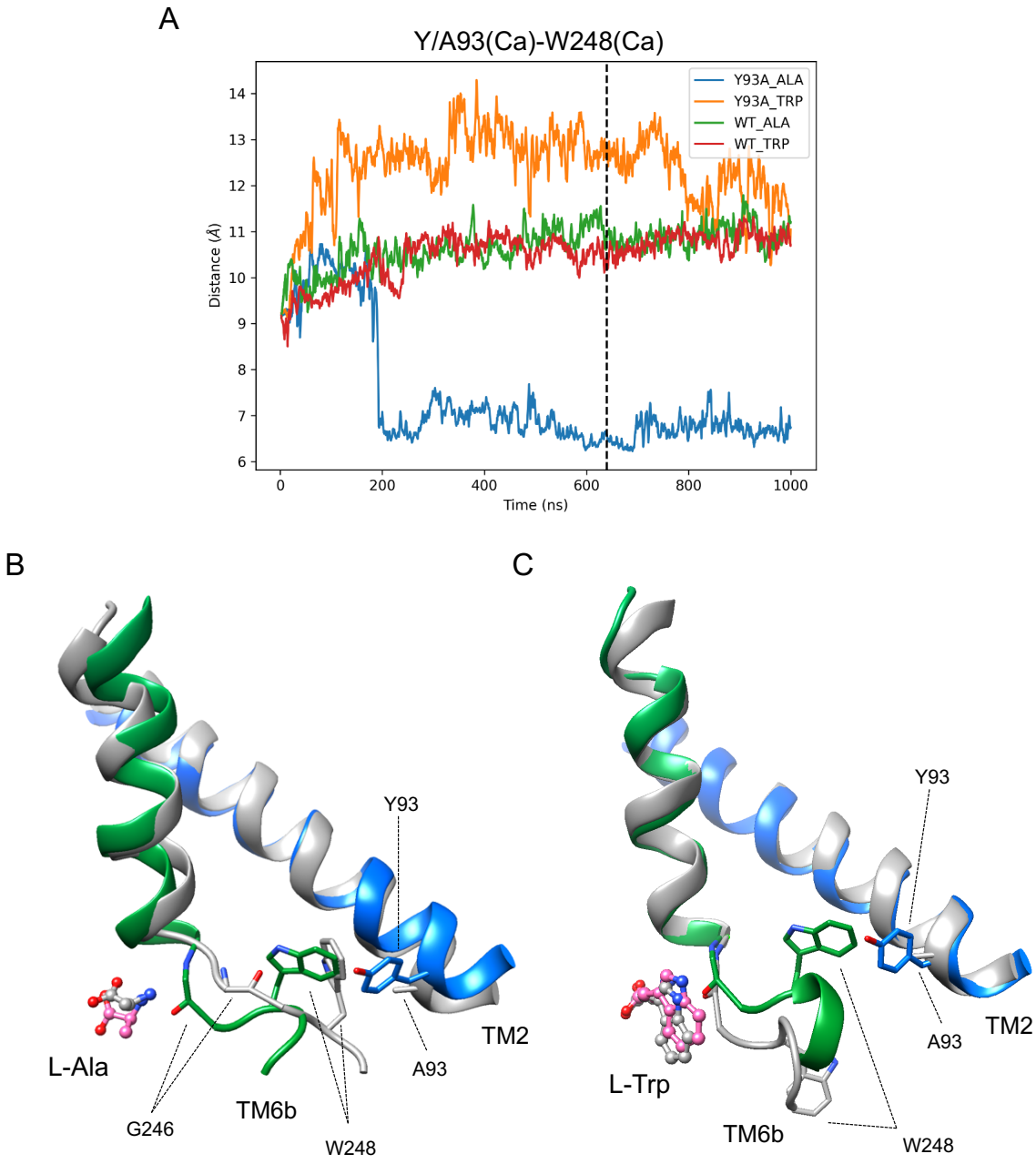

Figure S12. Molecular dynamics analysis of wild type hLAT2/CD98hc and mutant Y93A bound to L-alanine or L-tryptophan. (A) Evolution of the atom Cα of Tyr/Ala 93 - Cα atom of Trp 248 distance for wild type (WT) and Y93A mutant in the presence of substrate. The line in panel (A) indicates the time (640 ns) corresponding to the snapshots shown in panels B and C. Snapshots of L-alanine (B) and of L-tryptophan (C) bound to wild type hLAT2 and Y93A mutant. Cartoons and residues are colored as in Figure 1, blue for TM1 and green for TM6 in wild type hLAT2, whereas Y93A cartoons and residues are shown in gray. The substrates are depicted with C atoms in pink (wild type) or gray (Y93A mutant), and O and N atoms in red and blue, respectively.

Cryo-EM and modeling of hLAT2/CD98hc	(EMD-11952) (PDB 7B00)
Data collection and processing	
Microscope	FEI Titan Krios
Detector	Gatan K3 (counting mode)
Calibrated magnification	105,000
Voltage (kV)	300
Electron exposure (e–/Å^2)	45
Defocus range (μm)	-1.2 to -2.2
Pixel size (Å)	0.85
Symmetry imposed	C1
Initial particle images (no.)	319,613
Final particle images (no.)	176,132
FSC threshold	0.143
Map resolution (Å)	3.91
Map resolution range (Å)	2.5-5.5
Refinement	
Refinement software	phenix.real_space_refine
Initial model used (PDB code)	I-tasser homology model
Map sharpening *B* factor (Å^2)	-112
Model composition	
Non-hydrogen atoms	7379
Protein residues	926
Ligands	1 Digitonin (AJP)
R.m.s. deviations	
Bond lengths (Å)	0.005
Bond angles (°)	0.962
Validation	
MolProbity score	1.91
Clashscore	6.8
Poor rotamers (%)	0.39
Favorable rotamers (%)	99.61
Ramachandran plot	
Favored (%)	90.46
Allowed (%)	9.54
Disallowed (%)	0.00
Mask CC	0.78

Table S1. Cryo-EM data collection, image processing and modeling parameters.

	Substrate						
	Glycine	L-Alanine	L-Valine	L-Isoleucine	L-Glutamine	L-Histidine	L-Tryptophan
hLAT2 wild type	100±2.7	100±9.3	100±2.6	100±5.3	100±3.4	100±7.7	100±4.7
hLAT2 G246S	249.5±5.6	227.1±5.2	124.7±1.7	78.7±4.4	15.9±0.6	50.9±3.4	14.5±2.6
hLAT2 N134S	31.3±5.3	35.9±0.5	34.5±3.9	81.8±2.1	32.7±1.6	102.3±4.9	98.5±1.1
hLAT2 T402A	109.6±3.0	115.7±10.8	93.7±2.4	104.2±5.5	106.2±8.2	109.5±3.4	134.4±6.4
hLAT2 Y93A	14.3±2.0	12.7±2.6	33.3±0.9	65.3±3.0	19.7±0.3	91.4±3.3	76.7±5.4
hAsc1 wild type	100±3.7	100±2.7	100±1.8	100±6.9	100±11.7	100±13.7	100±1.1
hAsc1 S246G	27.1±2.6	26.8±1.5	234.2±16.6	158.9±17.6	92.1±15.0	633.5±60.5	286.1±6.7

Table S2. Amino acid uptake data from Figures 3D, E and 4C normalized as percentage (mean±S.E.M.) of the wild type transporter activity.

	L-Alanine		L-Glutamine		L-Tryptophan	
	K_m	V_{max}	K_m	V_{max}	K_m	V_{max}
hLAT2 wild type	134.7±9.7	7376±288	110.3±5.6	6706±256	52.7±2.5	4122±66
hLAT2 G246S	54.7±5.5 ***	6720±143 (ns)	375.7±70.9 *	3140±76 ***	47.3±4.1 (ns)	1083.3±382 ***
hLAT2 N134S	403.1±13.0 ***	7107±483 (ns)	169.1±22.5 *	2679±200 ***	55.0±2.9 (ns)	3701±136 (ns)
hLAT2 Y93A	120.7±4.2 (ns)	707±110 ***	102.7±9.1 (ns)	830±18 ***	58.1±3 (ns)	3555±166 *

Table S3. Kinetic characterization of hLAT2 wild type and mutants in HeLa cells. Transporter-mediated uptake of [^{3}H] L-alanine, [^{3}H] L-glutamine and [^{3}H] L-tryptophan (1 μCi/ml, 1 min) by hLAT2/CD98hc wild type, G246S, N134S and Y93A mutants was measured in transport medium containing 0–2000 μM of the corresponding cold amino acid. Mediated transport was calculated as uptake in transporter-transfected cells minus mock-transfected cells. K_m is expressed in μM and V_{max} in pmols [^{3}H] amino acid / mg protein * min. Data are expressed as the mean±S.E.M. of three experiments run in triplicates and carried out on different days on different batches of cells. Student t-test comparison versus hLAT2 wild type (*, p≤ 0.05; ***, p≤ 0.001).