Structural studies of Heteromeric Amino acid Transporters (HATs): Validation of the first 3D structural model of a HAT (human 4F2hc/LAT2) and identification of new HAT targets for 3D-crystallization

Mª Elena Álvarez Marimón

Aquesta tesis ha estat realitzada al Institut de Recerca Biomèdica (IRB Barcelona)

Dins el programa de Doctorat de Biotecnologia de la Universitat de Barcelona (UB)

Facultat de Biologia, Departament de Bioquímica i Biologia Molecular (UB)

Structural studies of Heteromeric Amino acid Transporters (HATs): Validation of the first 3D structural model of a HAT (human 4F2hc/LAT2) and identification of new HAT targets for 3D-crystallization

Memòria per optar al grau de Doctora per la Universitat de Barcelona presentada per:

Mª ELENA ÁLVAREZ MARIMON

Barcelona, 2014

Dedico aquesta tesi a les persones més importants:
als meus pares, pel seu gran carinyo i recolzament
incondicional que ha permés que arribés fins aquí,
per ser els millors pares que es pugui tenir!

així com a Rubén per fer-me tant feliç

Contents

List of Figures

List of Tables

Abbreviations

4F2hc	4F2 heavy chain
4F2hc-ED	ectodomain of 4F2hc
AA	amino acid
Ab	antibody
APA	amino acid/polyamine antiporter
APC	amino acid, polyamine and organocation
asc	sytem asc amino acid transporter
ATP	adenosine-5'-triphosphate
BCA	bicinchoninic acid protein assay
BMOE	bis(maleimido)ethane
bp	base pairs
BSA	bovine serum albumin
cAMP	3'-5'-cyclic adenosine monophosphate
cDNA	complementary DNA
CHS	cholesterol hemisuccinate Tris salt
CMC	critical micellar concentration
CoIP	coimmunoprecipitation
CRAC	cholesterol recognition / interaction amino acid consensus
CRB	cell Resuspension Buffer
Cymal-6	6-cyclohexyl-1-hexyl-β-D-maltoside
Cymal-7	7-cyclohexyl-1-heptyl-β-D-maltoside
DAT	Drosophila Dopamine transporter
DDAO	N,N-dimethyldodecylamine N-oxide
DDM	n-Dodecyl-b-D-maltopyranoside
dH2O	deionized H2O
DM	n-Decyl-b-D-maltopyranoside
DMEM	Dulbecco’s modified Eagle's medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNAse	deoxyribonuclease
dNTP	deoxyribonucleotide
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EL	extracellular loop
ER	endoplasmic reticulum
FACS	fluorescence activated cell sorting
FBS	fetal bovine serum
Fos-9	foscholine-9
Fos-12	foscholine-12
FPLC	fast protein liquid chromatography
FSEC	fluorescent size exclusion chromatography
g	gravitational-force
GFP	green fluorescent protein

GPCR	G-protein-coupled receptors
HA	human influenza hemagglutinin tag
HAT	heteromeric amino acid transporter
HEK 293T	human embryonic kidney 293T
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	horseradish peroxidase
HUGO	Human Genome Organization
Ig	immunoglobulin
IMAC	immobilized metal ion affinity chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside
kb	kilobase
kDa	kiloDalton
Km	Michaelis–Menten constant
kpsi	kilopound per square inch
KSHV	Kaposi's sarcoma herpesvirus
LAT	system L amino acid transporter
LAPAO	3-dodecylamido-N,N'-dimethylpropyl amine oxide
LDAO	N,N-Dimethyldodecylamine N-oxide
LiAc	Litic acid
LPI	lysinuric protein intolerance
LSB	Laemmli sample buffer
mAb	monoclonal antibody
mAb	monoclonal antibody
MRB	membrane Resuspension Buffer
mRNA	messenger RNA
NG	nonyl glucoside
NM	nonyl maltoside
OD	optical density
OG	octyl glucoside
OMIM	online Mendelian inheritance in man
PBS	phosphate buffer saline
PCR	polymerase chain reaction
PDB	protein data bank
PEG	polyethylene glycol
PEI	polyethylenimine
rBAT	related to basic amino acid transporter, heavy subunit
RNA	ribonucleic acid
RT	room temperature
SB	solubilization buffer
SDS	sodium dodecyl sulfate
SDS-PAGE	sds-polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SLC	Solute carrier (gene)
TCDB	the transporter classification database
TEV	Tobacco etch virus
TM	transmembrane

TMD	Transmembrane domain
Tris	tris(hydroxymethyl)aminomethane
UDS	Ultracentrifugation Dispersity Sedimentation Assay
-URA	Yeast synthetic drop-out medium supplement without Ura
WB	Washing buffer
WT	wild-type
xCT	system xC- amino acid transporter
yEGFP	yeast enhanced green fluorescent protein
y+LAT	system y+L amino acid transporter
YPD	Yeast peptone dextrose media
YSB	Yeast suspension buffer

Amino acid abbreviations

A	**Ala**	alanine	**M**	**Met**	methionine
C	**Cys**	cysteine	**N**	**Asn**	asparagine
D	**Asp**	aspartate	**P**	**Pro**	proline
E	**Glu**	glutamate	**Q**	**Gln**	glutamine
F	**Phe**	phenylalanine	**R**	**Arg**	arginine
G	**Gly**	glycine	**S**	**Ser**	serine
H	**His**	histidine	**T**	**Thr**	threonine
I	**Ile**	isoleucine	**V**	**Val**	valine
K	**Lys**	lysine	**W**	**Trp**	tryptophan
L	**Leu**	leucine	**Y**	**Tyr**	tyrosine

Introduction

1. Membrane protein transporters

1.1. Transport across the plasma membrane

Mammalian cells are delimited by a plasma membrane formed by a phospholipid bilayer, cholesterol, and proteins that protects mechanically the cell and contributes to the selective transport of molecules; furthermore, it also participates in cell adhesion and communication with nearby and distant cells. In addition, phospholipids and proteins in eukaryotic membranes are usually glycosylated on the external face of the cell membrane, conferring a "cell identity" that allows the immune system to recognize the cells and detect pathogens or cancer cells that have changed the glycosylation pattern from the healthy ones.

The selective permeability of the plasma membrane allows controlling the transport of specific molecules and conferring equilibrium between the extracellular and the intracellular medium. Only a limited number of molecules can cross plasmatic membranes without transmembrane proteins, like O_2 or CO_2 by diffusion. The uptake and efflux of sugars, amino acids, nucleotides, ions, and drugs and even water is controlled by transporters or channels. Depending on the energy expenditure transporters can be classified in passive, active transporters or channels. The passive transporters, also known as facilitated transporters, allow diffusion of solutes to cross the cell membrane using already existing ion/solute gradients.

In the active transport, molecules move against a concentration gradient utilizing diverse energy-coupling mechanisms that classify them into primary or secondary active transporters. Primary active transporters use the free energy provided by ATP hydrolysis to transport against an electrochemical gradient. These pumps enable the generation and storage of electrochemical ion gradients across membranes that afterwards secondary-active ion-coupled transporters use to drive uphill transport of solutes. Cotransport or symport happens when the ion forming the electrochemical gradient and the driven molecule go in the same direction, and exchange or antiport in the case that they move in opposite direction.

Membrane transporters have high degree of specificity for the transported substrate and bind only one (or a few) substrates at the same time and the transporter undergoes conformational changes to allow the transport of the bounded substrate (Hediger *et al.,* 2013). Channels allow as well the movement of solutes down their electrochemical gradients but are not so specific and do not translocate during transport (Armstrong, 2003; Decoursey, 2003; Jiang *et al.,* 2003; Yu and Catterall, 2003).

1.2. Solute Carrier Transporters (SLC)

In mammalian, genes encoding passive transporters, ion coupled transporters and exchangers are grouped into Solute Carrier transporters families (Figure 1).

SLC classification proposed by the Human Genome Organization (HUGO) Nomenclature Committee Database (http://www.gene.ucl.ac.uk/nomenclature/) (Hediger *et al.* 2004) classifies the amino acid transporters in 11 different families (Table 2) from the 52 total families belonging to SLC nomenclature (Table 1). They are classified based on the number of predicted or observed transmembrane (TM) α-helices (usually 10–14) and sequence similarity, of at least 20% with at least one other family member (Schlessinger *et al.* 2013). Members within the same SLC family can have substrates with different physicochemical properties despite their common evolutionary origin. Conversely, SLC families can be unrelated evolutionarily but still have chemically similar substrates such as the amino acid transporter families SLC1 and SLC7 (Schlessinger *et al.* 2013). Although the majority of amino acid transporters have been identified and characterized, a significant number of orphan transporters remain, for instance in the SLC16 and SLC38 families (Bröer & Palacín 2011).

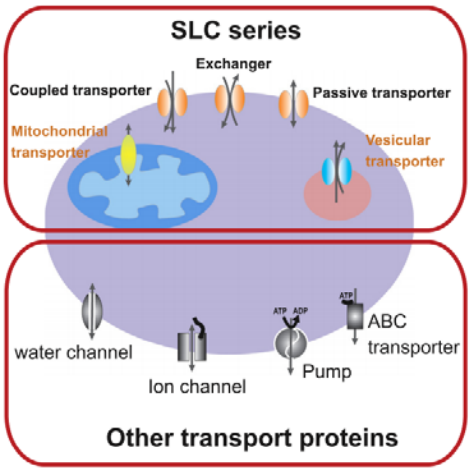


Figure 1. Cartoon of membrane transporters in a cell. SLC transporters are shown on the top and include passive transporters, secondary active transporters (cotransporters and exchangers), mitochondrial transporters and vesicular transporters. Other transport proteins do not belong to SLC (at the bottom) like water channels, ion channels, pumps, and ABC transporters (from Hediger *et al.,* 2013).

The HGNC Solute Carrier Family Series	Total 2013
SLC1: The high affinity glutamate and neutral amino acid transporter family	7
SLC2: The facilitative GLUT transporter family	14
SLC3: The heavy subunits of the heteromeric amino acid transporters	2
SLC4: The bicarbonate transporter family	10
SLC5: The sodium glucose cotransporter family	12
SLC6: The sodium- and chloride-dependent neurotransmitter transporter family	21
SLC7: The cationic amino acid transporter/glycoprotein-associated amino-acid transporter family	14
SLC8: The Na^+/Ca^{2+} exchanger family	3
SLC9: The Na^+/ H^+ exchanger family	13
SLC10: The sodium bile salt cotransport family	7
SLC11: The proton coupled metal ion transporter family	2
SLC12: The electroneutral cation-Cl cotransporter family	9
SLC13: The human Na^+-sulfate/carboxylate cotransporter family	5
SLC14: The urea transporter family	2
SLC15: The proton oligopeptide cotransporter family	5
SLC16: The monocarboxylate transporter family	14
SLC17: The vesicular glutamate transporter family	9
SLC18: The vesicular amine transporter family	4
SLC19: The folate/thiamine transporter family	3
SLC20: The type III Na^+-phosphate cotransporter family	2
SLC21/SLCO: The organic anion transporting family	12
SLC22: The organic cation/anion/zwitterion transporter family	23
SLC23: The Na^+-dependent ascorbic acid transporter family	4
SLC24: The $Na^+/(Ca^{2+}–K^+)$ exchanger family	6
SLC25: The mitochondrial carrier family	53
SLC26: The multifunctional anion exchanger family	11
SLC27: The fatty acid transport protein family	6
SLC28: The Na^+-coupled nucleoside transport family	3
SLC29: The facilitative nucleoside transporter family	4
SLC30: The zinc efflux family	10
SLC31: The copper transporter family	2
SLC32: The vesicular inhibitory amino acid transporter family	1
SLC33: The Acetyl-CoA transporter family	1
SLC34: The type II Na^+-phosphate cotransporter family	3
SLC35: The nucleoside-sugar transporter family	30
SLC36: The proton-coupled amino acid transporter family	4
SLC37: The sugar-phosphate/phosphate exchanger family	4
SLC38: The System A & N, sodium-coupled neutral amino acid transporter family	11
SLC39: The metal ion transporter family	14
SLC40: The basolateral iron transporter family	1
SLC41: The MgtE-like magnesium transporter family	3
SLC42: The Rh ammonium transporter family (pending)	3
SLC43: Na^+-independent, system-L like amino acid transporter family	3
SLC44: Choline-like transporter family	5
SLC45: Putative sugar transporter family	4
SLC46: Folate transporter family	3
SLC47: Multidrug and Toxin Extrusion (MATE) family	2
SLC48: Heme transporter family	1
SLC49: FLVCR-related transporter family	4
SLC50: Sugar efflux transporters	1
SLC51: Transporters of steroid-derived molecules	2
SLC52: Riboflavin transporter family	3
Total	395

Table 1. List of current SLC families and the total number of members in each family. (Table adapted from Hediger *et al.,* 2013).

1.3. Importance of getting the 3D-structure of membrane proteins

Total membrane proteins in eukaryotic cells represent one third of the whole proteome (Krogh *et al.,* 2001), and a majority of 75% in specialized organelle like mitochondria (Alberts B, Molecular Biology of the Cell. 4th edition, 2002). The number and type of membrane proteins are highly variable and give the different functional properties to each membrane. Therefore, dysregulation and malfunction of their biological activity often leads to severe diseases (Sanders & Myers, 2004).

Due to its physiologic relevance, membrane proteins represent two thirds of all current therapeutic targets (Hediger *et al.,* 2013). Nevertheless, so far, only 466 unique structures (mainly prokaryotic) have been deposited in the Protein Data Bank (PDB, April 2014) representing less than 2% of all available protein crystal structures (Figure 2). Furthermore, very few of these structures were solved at high resolution (better than 2 Å). The small amount of solved membrane protein structures is mainly due to the hydrophobic segments of integral membrane proteins that are embedded in phospholipid bilayers, making difficulty to be overexpressed, purified and crystallized (Arinaminpathy *et al.,* 2009; Bill *et al.,* 2011).

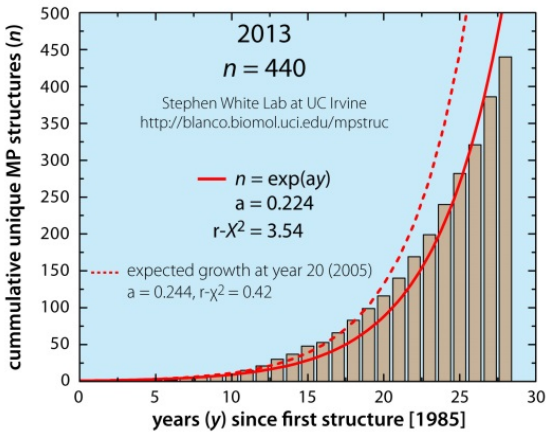


Figure 2. Representation of membrane protein solved structures relative to each year. Blanco membrane protein database (http://blanco.biomol.uci.edu/mpstruc/) statistics of cumulative unique membrane proteins structures published since 1985 shows an n=440 of solved membrane proteins in 2013.

For a long time it has been considered a difficult or even an impossible task, but last years the number of high-resolution structures solved has increased, including many homologues of SLC family members (Figure 3) (Hediger *et al.,* 2013). Most of these proteins were obtained from lower organisms homologues due to the facility of expression, purification and crystallization compared with eukaryote systems. These structures have been used to model the putative structures of human homologues that can guide biochemical studies to study the structure of the eukaryotic homologues and give the start point for structure-based drug design. Nevertheless, work towards the elucidation of human transporters is still

required to learn the structural bases of these transporters. At 2012, from almost 400 membrane protein solved, 25 were SLC family members.

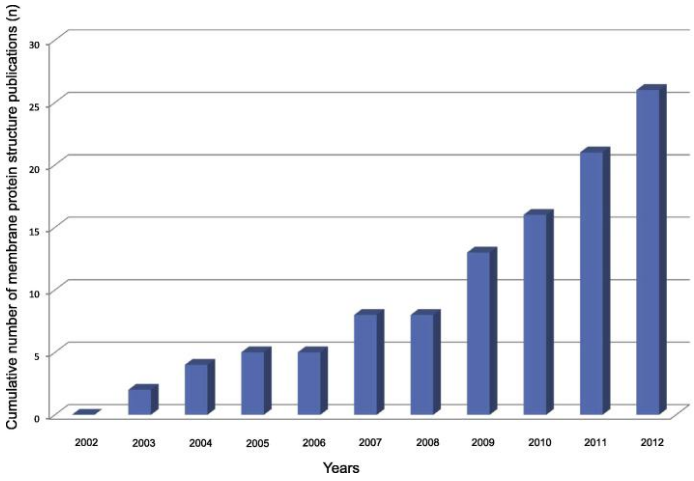


Figure 3. Representation of membrane protein solved structures belonging to the SLC family relative to each year. Cumulative number of membrane protein structure publications since 2002 show a exponential evolution of the number of SLC family members crystal structures during the 2002-2012 (Hediger *et al.,* 2013).

1.4. Amino acid transporters and their physiological relevance

Amino acid transporters mediate amino acids transport across the cell membrane providing cells of one of the building blocks of live. Amino acids play key roles in all cells, such as protein synthesis, metabolism regulation, cell growth, regulation of cell volume, metabolic energy production, neural transmission, and cell signaling molecules. Amino acid transporters are essential for the absorption of amino acids from nutrition, mediating the transfer of amino acids between organs, between different cells, and between compartments in cells. Together with other membrane proteins the amino acid transporters help to the maintenance of ionic concentration, flowing nutrients, providing amino acids to all tissues of the body (Christensen *et al.,* 1990).

Our diet is composed of more than 30% protein, which after digestion, the resulting peptides and amino acids are absorbed by enterocytes in the small intestine (Adibi *et al.,* 1967), there, peptides are hydrolyzed and released to blood by amino acid transporters and delivered to all tissues. Kidney filters blood and reabsorbs those valuable metabolites including amino acids. Amino acid transport across cellular membranes is mediated in most of the cases by multiple transporters with overlapping specificities (Palacín *et al*., 1998; Bröer, 2008). Moreover, they play a role in the reuptake of the nervous synapsis, where some amino acids like glutamate, glycine or aspartate, are neurotransmitters. Finally, in the placenta, amino acid transporters provide the nutrients from mother to fetal blood and help to detoxify it.

Absence, overexpression and malfunction of certain amino acid transporters can affect whole-body homeostasis giving raise to diverse human diseases (Table 2).

SLC	Acronym	Substrate(s)	Function	Disease/phenotype
SLC1A1	EAAT3	D,E,Cn	System X^-_{AG}	Dicarboxylic aminoaciduria, OCD
SLC1A2	EAAT2	D,E	System X^-_{AG}	
SLC1A3	EAAT1	D,E	System X^-_{AG}	Episodic ataxia?
SLC1A4	ASCT1	A,S,C	System ASC	
SLC1A5	ASCT2	A,S,C,T,Q	System ASC	Tumour growth
SLC1A6	EAAT4	D,E	System X^-_{AG}	
SLC1A7	EAAT5	D,E	System X^-_{AG}	
SLC3A1	rBAT	Trafficking subunits	Heavy chains of heteromeric AAT	Cystinuria
SLC3A2	4F2hc	Trafficking subunits	Heavy chains of heteromeric AAT	Tumour growth
SLC6A5	GlyT2	G	System Gly	Hyperekplexia
SLC6A7	PROT	P	Proline transporter	
SLC6A9	GlyT1	G	System Gly	
SLC6A14	$ATB^{0,+}$	All neutral and cationic amino acids	System $B^{0,+}$	Obesity?
SLC6A15	B^0AT2	P,L,V,I,M	System B^0	
SLC6A17	$NTT4/B^0AT3$	L,M,P,C,A,Q,S,H,G	System B^0	
SLC6A18	$XT2/B^0AT3$	G, A	System Gly	Hyperglycinuria? Hypertension?
SLC6A19	B^0AT1	All neutral amino acids	System B^0	Hartnup disorder, hypertension?
SLC6A20	IMINO	P	System IMINO	Iminoglycinuria
SLC7A1	CAT-1	K,R,O	System y^+	
SLC7A2	CAT-2	K,R,O	System y^+	
SLC7A3	CAT-3	K,R,O	System y^+	
SLC7A5	LAT1/4F2hc	H,M,L,I,V,F,Y,W	System L	Tumour growth
SLC7A6	$y^+LAT2/4F2hc$	K,R,Q,H,M,L	System y^+L	
SLC7A7	$y^+LAT1/4F2hc$	K,R,Q,H,M,L,A,C	System y^+L	Lysinuric protein intolerance
SLC7A8	LAT2/4F2hc	All neutral amino acids, except P	System L	
SLC7A9	$b^{0,+}AT/rBAT$	R,K,O,Cn	System $b^{0,+}$	Cystinuria
SLC7A10	Asc-1/4F2hc	G,A,S,C,T	System asc	
SLC7A11	xCT/4F2hc	D,E,Cn	Sytem x^-_c	
SLC7A12	Asc-2	G,A,S,C,T	System asc	
SLC7A13	AGT1	D,E	Asp, Glu transporter	
SLC16A10	TAT1	W,Y,F	System T	Blue diaper syndrome?
SLC17A6	VGLUT2	E	Vesicular Glu transporter	
SLC17A7	VGLUT1	E	Vesicular Glu transporter	
SLC17A8	VGLUT3	E	Vesicular Glu transporter	Non-syndromic deafness
SLC25A2	ORC2	K,R,H,O,Cit	Orn/Cit carrier	
SLC25A12	AGC1	D,E	Asp/Glu carrier	Global cerebral hypomyelination
SLC25A13	AGC2	D,E	Asp/Glu carrier	Type II citrullinaemia, neonatal intrahepatic cholestasis
SLC25A15	ORC1	K,R,H,O,Cit	Orn/Cit carrier	HHH syndrome
SLC25A18	GC2	E	Glu carrier	
SLC25A22	GC1	E	Glu carrier	Neonatal myoclonic epilepsy
SLC32A1	VIAAT	G,GABA	Vesicular Gly/GABA transporter	
SLC36A1	PAT1	G,P,A	Proton AAT	Hair colour (horses)
SLC36A2	PAT2	G,P,A	Proton AAT	Iminoglycinuria
SLC36A4	PAT4	P,W	Amino acid sensor	
SLC38A1	SNAT1	G,A,N,C,Q, H,M	System A	
SLC38A2	SNAT2	G,P,A,S,C,Q,N,H,M	System A	
SLC38A3	SNAT3	Q,N,H	System N	
SLC38A4	SNAT4	G,A,S,C,Q,N,M	System A	
SLC38A5	SNAT5	Q,N,H,A	System N	
SLC43A1	LAT3	L,I,M,F,V	System L	
SLC43A2	LAT4	L,I,M,F,V	System L	
Not assigned	Cystinosin	Cn	Lysosomal Cys transporter	Cystinosis

Table 2. List of the 11 SLC amino acid transporters family members, nomenclature, transport activity and involvement in diseases. Amino acid substrates indicated with their letter code and Cit, citrulline; Cn, cystine; O, ornithine. In the transport function column, the acronyms of the substrate specificity of the transporter follow a nomenclature: upper-case letters indicate Na^+-dependent transporters (with the exception of system L, system T and the proton amino acid transporters); lower case is used for Na^+-independent transporters (for example asc, y^+ and x^-_c). X^- or x^- indicate transporters for anionic amino acids (as in X^-_{AG} and x^-_c); AG means that it transports aspartate and glutamate, and "c" indicates that the transporter also accepts cystine. Y^+ or y^+ to cationic amino acids transporters (still there is not an Na^+-dependent cationic amino acid transporter well defined, thus, Y^+ is not used); B or b accounts for amino acid transporters of broad specificity; superscript "0" indicates neutral amino acids transport and "+" cationic amino acid transporter. "T" indicates aromatic amino acid transport, and "N", selectivity for amino acids with nitrogen atoms in the side chain. In the rest of the cases, the amino acid letter code represents the preferred substrate. For example, system L are transporters that prefer leucine and system ASC, alanine, serine and cysteine.

Proline and hydroxyproline are referred to as imino acids. Due to a historical use of this nomenclature it is still widely extended even though is not completely consistent (Table from Bröer and Palacín 2011).

Disorders with defective amino acid transport activities reflect the variety of roles that they play in human physiology. Mutations of brain amino acid transporters affect neuronal excitability. Mutations of renal and intestinal amino acid transporters produce malabsorption and renal problems that affect the whole-body homoeostasis, resulting in malabsorption and renal problems. Amino acid transporters that are integral parts of metabolic pathways reduce the function of these pathways. Finally, amino acid uptake is essential for cell growth, thereby explaining their role in tumor progression (Bröer and Palacín, 2011). Due to their important role in physiology, a deeper study of the structure-function of amino acid transporters would have a great contribution in pharmacology and medicine to understand, prevent and treat the diseases related with the amino acid transporters.

2. Heteromeric Amino acid Transporters (HATs)

Heteromeric Amino acid Transporters (HATs) are the only transporter family composed by two different subunits: a light and a heavy subunit belonging to SLC7 and SLC3 families, respectively, linked by a disulphide bridge (Table 3). Interestingly, they are covalently linked by a disulphide bridge between the cysteines located a few residues away from the transmembrane domain of the heavy subunit and in the putative extracellular loop 2 of the light subunit (Figure 4), having found the cysteines involved in the disulphide bridge highly conserved in all the family members (Pfeiffer *et al.*, 1998; Verrey *et al.,* 1999; Chillarón *et al.,* 2001; Palacín & Kanai*,* 2004). It is suggested that additional interactions between the heavy and the light subunit may occur, non-covalent ones, since mutation on cysteines responsible for the disulphide bond do not break the interaction (Estévez *et al.*, 1998).

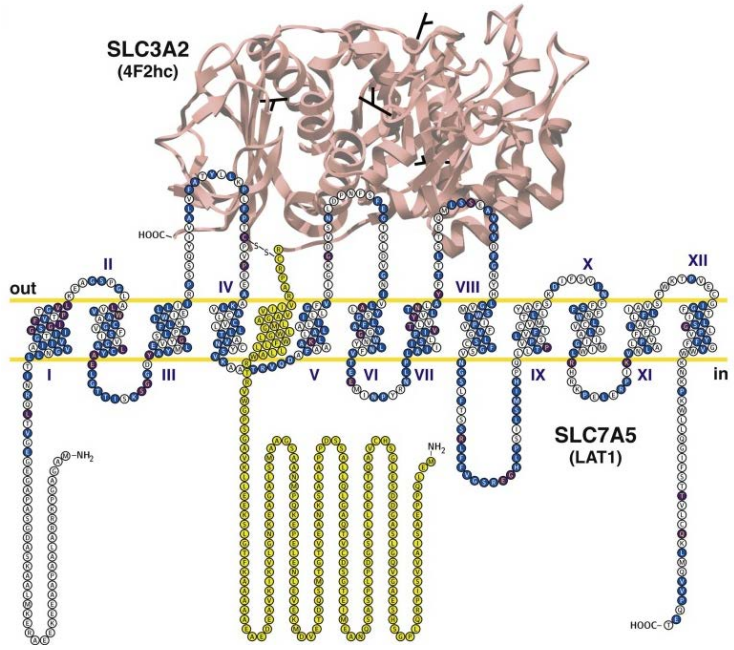


Figure 4. Topology model of human 4F2hc/LAT1 (SLC3A2/SLC7A5) as a representative member of HAT proteins. The crystal structure of human 4F2hc-ED was solved and is displayed (Fort *et al.,* 2007), and the 4 glycosylation sites of 4F2hc are indicated on the structure. Topology of the single TMD and N-term of 4F2hc is illustrated in yellow. Identical and conserved residues within the corresponding families, i.e. CATs and the light subunits of HATs, are colored in violet and blue, respectively. Transmembrane topologies were predicted using the HMMTOP server (Tusnady and Simon, 2001) and topology data were plotted using T(E)Xtopo (Beitz *et al.,* 2000). Adapted from (Fotiadis, Kanai & Palacín, 2013).

The minimal functional unit of the transporter is formed by one heavy subunit and one light subunit, from which, the light subunit is the catalytic part of the transporter conferring the amino acid transport specificity to the HAT (Reig et al., 2002; Fernández *et al.,* 2006), and the heavy subunit is the responsible of the trafficking of the holotransporter to the plasma membrane (Estévez *et al.,* 1998; Pfeiffer *et al.,* 1998; Deves *et al.,* 2000; Palacín & Kanai, 2004). Ten different light subunits (asc1, LAT1, LAT2, xCT, y^+LAT1, y^+LAT2, $b^{0,+}$AT, AGT1, asc2 and arpAT) and two heavy subunits (4F2hc and rBAT) have been

identified until now as members of HATs (Table 3). Structurally, light subunits are polytopic non-glycosylated (Kanai *et al.*, 1998; Torrents *et al.*, 1998) proteins composed by 12 predicted transmembrane domains (TMD), with N and C-term intracellular (Gasol *et al.,* 2004) (Figure 4). Their molecular weight is~50 kDa; however, when visualizing them in SDS-PAGE, they run anomalously at 35-40 kDa due to their hydrophobicity (Verrey *et al.*, 2000; Chillarón *et al.*, 2001).

Light subunits belong to the L-type amino acid transporter (LAT) family within the Amino acid, Polyamine and Organocation (APC) superfamily described by the Transporter Classification Database (TCDB; http://www.tcdb.org). LATs, together with CATs are members of the mammalian solute carrier SLC7 family, but this second ones have 14 TMD described, are glycosylated proteins and do not form heterodimers (Fotiadis *et al.*, 2013).

Heavy subunits are type II N-glycoproteins with a single transmembrane domain, a cytosolic N-term, and a bulky extracellular ectodomain (ED) that shows structural homology with insect maltases and bacterial α-glucosidases (Estévez *et al.,* 1992, Torrents *et al.,* 1998; Fort *et al.,* 2007; Gabrisko and Janecek, 2009) (Figure 3). 4F2hc and rBAT share 22 % amino acid identity and 50 % similarity (Bertran *et al.,* 1992a; Bertran *et al.,* 1992b; Tates *et al.,* 1992; Wells *et al.,* 1992) and whereas 4F2hc is ubiquitously expressed in all the cells, rBAT is expressed in the renal proximal tubule in kidney and small intestine (Fernandez *et al.,* 2006). They localize differently in polarized epithelial cells (intestine and kidney): rBAT, in the apical membrane and 4F2hc in the basolateral membrane (Furriols *et al.*, 1993; Pfeiffer *et al.,* 1998; Palacín *et al.* 1998), and are visualized by SDS-PAGE as ~94 KDa and ~85 kDa band, respectively.

4F2hc forms heterodimer with six of the known light subunits: LAT1, LAT2, y^+LAT1, y^+LAT2, asc1 and xCT (Fotiadis, Kanai & Palacín, 2013); and rBAT associates with $b^{0,+}$AT (Reig *et al.* 2002) or AGT1 (Unpublished results, Kanai, Nunes and Palacín laboratories) (Table 3). The heavy subunit corresponding to asc2 and arpAT remains unknown (Table 3). Light subunits when combined with their corresponding heavy subunit constitute different amino acid transporters systems which are mostly obligatory exchangers with the exception of 4F2hc/asc1 and 4F2hc/asc2 that also work as uniporters (Busch *et al.,* 1994; Chillarón *et al.,* 1996; Fukasawa *et al.,* 2000).

They transport a broad spectrum of substrates ranging from neutral amino acids to aromatic amino acids, negatively charged amino acids and cationic amino acids plus neutral amino acids. Transporters y^+L are the only sodium-depended co-transporter HAT members, and only with neutral amino acids as co-substrates (Fotiadis *et al.,* 2013).

Heavy subunit	Light subunit	Gene Family	Human Chromosome	Transporter system	Predominant Substrates/ function	Disease linkage
4F2hc		SLC3A2	11q13		Traffic of the holotransporter to the membrane	Overexpressed in tumors
	LAT1	SLC7A5	16q24.3	L	Large neutral L-AA, T_3, T_4, L-DOPA, BCH	Overexpressed in tumors
	y+LAT2	SLC7A6	16q22.1	y^+L	Na+ indep.: cationic AA Na+/large neutral AA	–
	y+LAT1	SLC7A7	14q11.2	y^+L	Na+ indep.: cationic AA Na+/large neutral L-AA	Lysinuric protein intolerance (LPI)
	LAT2	SLC7A8	14q11.2	L	Neutral L-AA, T_3, T_4, BCH	–
	asc1	SLC7A10	19q13.1	Asc	small neutral AA	–
	xCT	SLC7A11	4q28.3	x_c^-	Cystine (anionic form), L-glutamate	Overexpressed in tumors Kaposi Sarcoma Herpes Virus entry
rBAT		SLC3A1				Cystinuria Type A
	$b^{0,+}$AT	SLC7A9	19q13.1	$b^{0,+}$	neutral/dibasic amino acids	Cystinuria Type B
	AGT-1	SLCA13	8q21.3		L-Aspartate and L-glutamate	
?						
	asc-2	Slc7a12		Asc		
	ArpAT	Slc7a15	(silenced in primates)		Aromatic-preferring AA	

Table 3. List of Heteromeric amino acid transporters (HATs). SLC gene member, human chromosome, transport system, predominant substrate/function and disease linkage are indicated for each subunit. 4F2hc forms heterodimer with LAT1, LAT2, y^+LAT1, y^+LAT2, asc1 or xCT and rBAT with $b^{0,+}$AT or AGT1. (AA) amino acid(s); (?) not identified; (-) not described.

2.1. rBAT and its heterodimers

This was the first subunit cloned and characterized in *Xenopus laevis* oocytes, which expression induced a high Na^+-independent uptake of dibasic and neutral amino acids (Bertran *et al.,* 1992; Tate *et al.,* 1992; Wells and Hediger, 1992). The rBAT/$b^{0,+}$AT heterodimer oligomerizes into stable heterotetramers and it is expressed mainly in the brush border membranes of the epithelial cells from renal proximal tubule (from S1 to S3 segments rBAT expression increases) (Fernández *et al.,* 2006), also in the small intestine, localizing in the apical membrane (Kanai *et al.,* 1992; Furriols *et al.,* 1993; Pickel *et al.,* 1993; Pfeiffer *et al.,* 1999), and lower expressed in brain (Bertran *et al.,* 1992, Yan *et al.,* 1992).

2.1.1. rBAT/$b^{0,+}$AT (system $b^{0,+}$)

Transport system $b^{0,+}$, molecularly represented by rBAT/$b^{0,+}$AT, induces a Na^+-independent electrogenic exchange of extracellular cationic amino acids and L-cystine for intracellular neutral amino acids, except imino acids (Bertran *et al.,* 1992; Busch *et al.,* 1994; Chillaron *et al.,* 1996; Pfeiffer *et al.,* 1999). Transport direction of substrate exchange is determined by the high intracellular concentration of neutral amino acids, the intracellular reduction of cystine to cysteine and the electric potential across the plasma membrane which is negative inside.

rBAT/$b^{0,+}$AT is a dimer of heterodimers ([rBAT/$b^{0,+}$AT]$_2$), where the single rBAT/$b^{0,+}$AT heterodimer is the transport functional unit (Fernández *et al.* 2006). Chimerical fusion proteins of rBAT with 4F2hc-light subunits also show functional heterotetramers, suggesting that the heavy subunit dictates the oligomerization state (Fernández *et al.* 2006). It has been proved that complete folding and maturation of rBAT needs interaction with $b^{0,+}$AT (Rius *et al.,* 2012). When expressed alone, rBAT remains endoglycosidase H sensitive and is degraded, whereas $b^{0,+}$AT is stable (Bauch *et al.,* 2002; Pineda *et al.,* 2004; Bartoccioni *et al.,* 2008). The assembly with $b^{0,+}$AT within the endoplasmic reticulum (ER) blocks rBAT degradation (Bartoccioni *et al.,* 2008). Thus, intracellular assembly of rBAT with $b^{0,+}$AT to form a disulfide-linked heterodimer is required for functional expression. Reconstitution into liposomes of membranes of transfected cells with only $b^{0,+}$AT or with both rBAT and $b^{0,+}$AT resulted in the same transport activity induction. This result suggested that the light subunit is the catalytic part of the holotransporter and that it folds independently of rBAT (Reig *et al.,* 2002).

The rBAT/$b^{0,+}$AT heterodimer is expressed in the brush border membranes of the epithelial cells from renal proximal tubule (where rBAT expression increases from S1 to S3 segments) (Fernández *et al.,* 2006); it was demonstrated by immunocytochemistry to be expressed mainly in kidney and small intestine (Kanai *et al.,* 1992; Furriols *et al.,* 1993; Pickel *et al.,* 1993; Pfeiffer *et al.,* 1999). In addition, it is also detected in brain (Bertran *et al.,* 1992, Yan *et al.,* 1992). Transport activity performed by the rBAT/$b^{0,+}$AT

heterodimer constitutes the main apical reabsorption system for cystine in kidney (Fernández *et al.,* 2002). On the extracellular side, rBAT/$b^{0,+}$AT has a high affinity for cationic amino acids and cystine (Km~100 µM within the physiological range in plasma), while on the intracellular side the affinity for neutral amino acids is in the milimolar range (e.g., ~ 2.5 mM for L-leucine) (Reig *et al.,* 2002).

Mutations in either of the transporter subunits rBAT or $b^{0,+}$AT can cause cystinuria (Online Mendelian Inheritance in Man (OMIM) database, entry 220100) (Calonge *et al.,* 1994; Feliubadaló *et al.,* 1999). 133 and 95 mutations in rBAT and $b^{0,+}$AT, respectively, have been identified so far, causing cystinuria type A, B or AB for patients with mutations in SLC3A1 gene, SLC7A9 gene or both, respectively (Dello Strologo *et al.*, 2002; Chillarón *et al.,* 2010). This autosomal recessive disorder is the most common primary inherited aminoaciduria. Cystinuria causes a failure in the reabsorption of filtered cystine and dibasic amino acids in the proximal tubule producing hyperexcretion of dibasic amino acids (lysine, arginine and ornithine) and cystine in urine. High concentration of the hyperexcreted amino acid in the urinary tract and its poor solubility of cystine results most often times in the formation of renal calculi (urolithiasis), which can cause obstruction, infection and ultimately chronic kidney disease (Chillarón *et al.,* 2010). Cystinuria represents 1–2% of all cases of renal stone formation and even 6–8% in pediatric patients (Milliner, *et al.,* 1993).

Cystinuria specific mutations in SLC3A1 and SLC7A9 cause loss of function. In the case of rBAT mutants, they showed problems in the trafficking of the heterodimer to the membrane (Bartoccioni *et al.,* 2008). Two different mechanisms underlie the trafficking defects. One type of mutation like L89P, located in the TMD, makes the assembly of rBAT with $b^{0,+}$AT really slow, resulting in very small amount of functional and mature glycosylated (rBAT/ $b^{0,+}AT)_2$ in the membrane. In contrast, mutations in the ectodomain of rBAT allow efficient assembly with $b^{0,+}$AT to form disulfide-linked heterodimers, but HATs remain core-glycosylated, fail to oligomerize and are degraded, most probably via the proteasome. Cystinuria-specific mutations in $b^{0,+}$AT studied so far cause lack of protein expression (mutation G105R), defective trafficking to the plasma membrane (mutation A182T) and defective transport function (mutations A354T and P482L) (Font *et al.,* 2001; Reig *et al.,* 2002; Shigeta *et al.,* 2006).

2.1.2. rBAT/AGT1

AGT1 (aspartate/glutamate transporter 1) was firstly described when, searching for new light subunits, it was cloned from a cDNA library of mouse kidney successfully expressed in the plasma membrane of *Xenopus* oocytes as fusion proteins of the transporter with 4F2hc or rBAT and functionally characterized (Matsuo *et al.,* 2002). Actually, this transporter has the cysteine residue responsible for the disulphide bridge conserved in all light subunits and it is 35-37 % identical to the rest of light subunits, presenting a 48 % identity to asc2. It transports in a sodium-independent manner anionic amino acids with

high affinities for L-aspartate and L-glutamate. AGT1, predominantly localized in kidney at the basolateral membrane of the proximal straight tubules and distal convoluted tubules is most probably an artifact (Matsuo *et al.,* 2002). Previous studies reported that in western blot analysis AGT1 shows mobility of 40 kDa in reducing conditions, whereas in the absence of reducing agents the protein appears forming part of a 250 kDa complex, indicating the formation of heterotetramers as $b^{0,+}AT$. Recently, it has been discovered that AGT1 associates with rBAT (unpublished results, Kanai, Nunes and Palacín laboratories).

2.2. 4F2hc and its heterodimers

4F2hc is an almost ubiquitous transmembrane protein in mammalian cells (Parmacek *et al.,* 1989; Nakamura *et al.,* 1999;), with an exclusively basolateral localization in intestine and kidney epithelia (Quackenbush *et al.,* 1986; Rossier *et al.,* 1999). Moreover, 4F2hc is present in all cell lines and tumor cells (Quackenbush *et al.*, 1986; Parmacek *et al.*, 1989; Nakamura *et al.*, 1999; Rossier *et al.*, 1999). Experimental evidence suggests that in native tissues 4F2hc-associated transporters are single heterodimers, whereas $rBAT/b^{0,+}AT$ is a dimer of heterodimers$[rBAT/b^{0,+}AT]_2$ (Fernández *et al.*, 2006).

No mutations have been identified in 4F2hc causing any disease, since its defect is believed to be lethal at the early embryonic phases. Silencing of 4F2hc caused a defective fetal implantation in the placenta (Domínguez *et al.*, 2010) leading to embryonic lethality (Tsumura *et al.*, 2003) (as observed in 4F2hc KO mice).

Apart from targeting the light subunit to the plasma membrane, 4F2hc plays multiple functions which are tightly related with β-integrin signaling regulation. Specifically, 4F2hc interacts with cytoplasmic tails of integrin-β chains through its transmembrane domain, which is needed and suffices integrin association (Zent *et al.,* 2000; Henderson *et al.,* 2004). This interaction mediates β1- and β3-integrin signaling (Feral *et al.,* 2005) that induces cell proliferation (Cantor *et al.,* 2009), migration (Fogelstrand *et al.,* 2009), and adhesion (Prager *et al.,* 2007) all key roles in cancer and metastatic cells.

4F2hc attaches pathogens, like enteropathogenic *E.coli* (EPEC) and *Citrobacter rodentium* producing loss of barrier function and polarity in epithelial cells, resulting in redistribution of basolateral proteins such as 4F2hc and β1-integrin and developing inflammatory bowel disease (Charania *et al.,* 2013).

This suggests that 4F2hc integrates integrin signaling functions and amino acid transport in order to be coupled processes. Still the impact of integrins modulating amino acid transport via 4F2hc has not been experimentally addressed.

4F2hc is also called CD98 and FRP-1, since it was firstly discovered, not yet as a transporter, but as a surface antigen of activated lymphocytes and identified as 4F2 antigen heavy chain or CD98 (following the systematic CD designation for antigens) (Haynes *et al.,* 1981; Quackkenbush *et al.,* 1987). And 4F2hc was designated 'fusion regulatory protein' (FRP-1) to reflect its function in cell fusion events that lead to multinucleated giant cells such as osteoclasts (Ohgimoto *et al.,* 1996; Suga *et al.,* 1997). Transport impact of 4F2hc was tested in *Xenopus* oocytes due to its sequence similarity with rBAT (previously discovered) and characterized as part of transport system y^+L (Bertran *et al.,* 1992; Wells and Hediger, 1992).

2.2.1. System L

System L mediates the transport of large branched and aromatic neutral amino acids in a Na^+-independent way. It is characteristic for being able to transport the non-metabolizable analog BCH (2-aminobicyclo-(2,2,1)-heptane-2-carboxylic acid) (Mastroberardino *et al.,* 1998; Meier *et al.,* 2002). There are identified in total four proteins that induce system L activity: LAT1-4, from which only LAT1 and LAT2 belong to SLC7 family and form heterodimer with 4F2hc and are exchangers, whereas the last two belong to SLC43 family and are uniporters (Babu et al., 2003; Bodoy *et al.,* 2005). LAT1 and LAT2 display similarly much lower intracellular (in the millimolar range) than extracellular (in the physiological micromolar range) apparent substrate affinities (Meier *et al.,* 2002).

2.2.1.1. 4F2hc/ LAT1

4F2hc/LAT1 is a sodium-independent obligatory exchanger (with 1:1 stoichiometry) that does not mediate any measurable amino acid efflux in the absence of extracellular amino acids. They transport mainly large neutral and aromatic neutral amino acids, including leucine, isoleucine and tyrosine (the majority are essential amino acids) and BCH (Mastroberardino *et al.,* 1998; Meier *et al.,* 2002); it can also transport thyroid hormone (Kinne *et al.,* 2011). The selectivity of the efflux function resembles that of uptake, but with some differences: L-leucine, L-isoleucine and L-methionine are relatively better efflux than influx substrates (Meier et al., 2002). This indicates that LAT1 equilibrate the relative concentrations of different amino acids, that exchanges them against intracellular amino acids accumulated by some other transporters (Meier et al., 2002; Verrey et al., 2003).

When analyzing tissue distribution of SLC7A5 mRNA in human tissues, this is found expressed in placenta > brain > spleen > testes, ovary and colon (Kanai *et al.,* 1998). In placental membranes it supplies thyroid hormones and amino acids to the developing fetus (Ritchie and Taylor, 2001). In brain, LAT1 is found expressed in microvessels participating in the transport of L-3,4-dihydroxyphenylalanine (L-DOPA) across the blood–brain barrier (Kageyama *et al.,* 2000). It has also been demonstrated that LAT1 is involved in blood-to-retina L-leucine transport at the inner blood-retinal barrier and seems to be closely involved in visual functions by supplying neurotransmitter precursors (Tomi *et al.,* 2005).

- 4F2hc/LAT1 overexpresses in cancer

4F2hc/LAT1 is upregulated in proliferative cells, cancer cell lines and most of types of human cancer tissues, including lung, colon, breast, prostate, head, neck, ovarian cancers, as well as in gliomas (Wolf *et al.,* 1996; Yanagida *et al.,* 2001; Fuchs and Bode, 2005; Kaira *et al.,* 2008; Kobayashi *et al.,* 2008), suggesting that LAT1 is essential for tumor cell survival and progression. In addition, the overexpression of 4F2hc/LAT1 correlates positively with the biological aggressiveness of tumors, cell proliferation and angiogenesis, and metastases (Fuchs and Bode, 2005; Ichinoe *et al.,* 2011; Kaira *et al.,* 2009a,b,c, 2008, 2011; Yoon *et al.,* 2005).

It has been described that 4F2hc/LAT1 forms a complex with other amino acid transporters and accessory proteins which activation is regulated by metabolism. The different transporters that have been described to be assiated there are: the neutral amino acid transporter ASCT2 (SLC1A5); the lactate/H^+ symporters (monocarboxylate transporter 1 (MCT1) (SLC16A1) and MCT4 (SLC16A3). As accessory protein it has been found CD147 (also called Basigin or Emmprin), which has been described to be responsible for the trafficking of MCT1 and MCT4 to the plasma membrane (Halestrap and Meredith, 2004). In the complex, also participate cell adhesion molecules as the epithelial cell adhesion molecule (EpCAM), which, in turn, is a regulator of cell proliferation (Xu and Hemler, 2005), and integrines that, in fact, interact with intracellular domain of 4F2hc. At the same time, CD147 induces the expression of matrix metalloproteinases and hyaluronic acid in the cell surface. Thus, this protein is also a prognostic marker for a wide range of tumors (Weidle *et al.,* 2010).

- 4F2hc/LAT1 activator of mTOR signaling

Actually, transport of neutral branched amino acids, and specifically leucine via LAT1 stimulates mTOR (mammalian target of rapamycin) (Nicklin *et al.,* 2009), which is a serine / threonine protein kinase that controls essential cellular processes, such as cell growth, proliferation and survival (Sengupta *et al.,* 2010). mTOR is a nutrient/energy/redox sensor that controls protein synthesis by integrating signaling from growth factors (via PI3K and Akt), energy metabolism (via AMP-dependent protein kinase; AMPK) and amino acids (particularly leucine) (Hay *et al.,* 2004).

In the uptake of leucine in tumor cells, needs the coordinated action of two transporters, LAT1 and ASCT2, to adjust the essential amino acid concentrations for metabolic demands and signaling to mTOR (Fuchs *et al.,* 2005). It was demonstrated in cultured cells that leucine transport required the efflux of glutamine, taken into the cell by ASCT2, via the exchanger LAT1 (Nicklin *et al.,* 2009; Wullschleger *et al.,* 2006). When cells are starved for glutamine, leucine cannot be transported into cells. As a result, mTORC1 is not activated, even in the presence of growth factors (Cohen *et al.,* 2009).

- 4F2hc/LAT1, a good target for cancer therapy

4F2hc/LAT1 is considered oncogenic since its inhibition decreases the viability of cancer cells and induces cell death. This inhibition can be achieved by the system L transport inhibitor BCH (Yamauchi *et al.,* 2009; Fan *et al.,* 2010; Imai *et al.,* 2010; Kaira *et al.,* 2013), or by antibodies against human LAT1, or knockdown of LAT1, which inhibited in vivo and in vitro the tumor growth, respectively (Kim *et al.,* 2006; Ohkawa *et al.,* 2011). Thus, LAT1 is considered to be both a prognosis marker and a target for both imaging and therapy for cancer.

In this regard, efforts have been focused to develop different anticancer therapies for LAT1: anticancer drugs that could be transported and internalized (del Amo *et al.,* 2008), or inhibitors of these transporters with the aim to reduce in that way tumor proliferation and progression. Still the number of specific inhibitors is scarce. In this direction compound KYT-0353 was develop as specific LAT1 inhibitor, which inhibits leucine uptake and cell growth in colon cancer cell line HT-29 with IC50s of 60 nM and 4 μM, respectively (Oda *et al.,* 2010); it also diminished growth of the HT-29 tumors transplanted into nude mice (Oda *et al.,* 2010).

Also different LAT1 positron emission tomography (PET) tracers have been developed with different success to image, evaluate and stage tumors in vivo. Among radiolabeled amino acids that are transport substrates for LAT1, tyrosine (L-[3-^{18}F]-α-methyl tyrosine) (Kaira *et al.,* 2007, 2009c, 2010) is the most specific for LAT1 due to its α-methyl moiety (Wiriyasermkul *et al.,* 2012). Others are: phenylalanine (L-p-(2-[^{18}F]fluoroethyl)-phenylalanine) (Wang *et al.,* 2011), tryptophan (5-(2-^{18}F-fluoroethoxy)-L-tryptophan) (Kramer *et al.,* 2012) and methionine (S-(3-[^{18}F]fluoropropyl)-D-homocysteine) (Denoyer *et al.,* 2012). Similarly, the LAT1 selective α-methyl amino acid probe L-3-[123I]-α-methyl tyrosine has been used for tumor imaging by SPECT (Wiriyasermkul *et al.,* 2012).

Nevertheless, this strategy has met limited success for oncologic imaging outside of brain (Oluwatayo *et al.,* 2013). The limited sensitivity is a consequence of the bi-directional nature of LAT1 and other system L transporters, which mediate both influx and efflux of substrates from cells through an exchange mechanism of transport. Thus, LAT1 cannot directly concentrate substrates, leading to relatively low tumor to tissue ratios. Additionally, the selectivity for radiolabeled amino acids targeting LAT1 over other system L transporters (LAT2, LAT3 and LAT4) is not typically complete. For that it has been develop a zirconium-89 labeled monoclonal antibody targeting the extracellular domain of LAT1, that shows specificity and high affinity for LAT1 in vitro and in vivo (Oluwatayo *et al.,* 2013).

2.2.1.2. 4F2hc/ LAT2

4F2hc/LAT2 is a sodium-independent, high-affinity obligatory exchanger, that transports a wide variety of neutral amino acids, including small ones, like alanine, glycine, cysteine and serine, all them poor substrates for LAT1 (Meier *et al.,* 2002; Pineda *et al.,* 1999; Rossier *et al.,* 1999).

Its expression is more ubiquitous than LAT1 in quiescent cells. It is found in many tissues in different levels (kidney >>> placenta >> brain, liver > spleen, skeletal muscle, heart, small intestine and lung) (Pineda *et al.,* 1999); also found expressed at high levels in the prostate, the testes, ovaries and in the fetal liver (Park *et al.,* 2005). Its localization in polarized cells is basolateral both in renal and intestinal epithelial cells (Rossier *et al.,* 1999).

In heterologous expression systems and cell culture models of renal epithelia 4F2hc/LAT2 efficiently exports L-cysteine (Bauch *et al.,* 2003; Fernandez *et al.,* 2003). In contrast, the null knockout of Slc7a8 in mice did not show renal reabsorption defects of neutral amino acids and cystine in particular (Braun *et al.,* 2011), and presents normal development and growth suggesting a functional compensation by other amino acid transporters which would explain the lack of a severe phenotype.

2.2.2. System y^+L

System y^+L (Torrents *et al*., 1998; Deves and Boyd, 1998), molecularly represented by 4F2hc/y^+LAT1 and 4F2hc/y^+LAT2, mediates the electro neutral exchange of cationic amino acids (y^+) (sodium-independent) and large neutral amino acids (L) (sodium-dependent) (Chillarón *et al.,* 2001). Under physiological conditions (i.e., 150 mM Na^+ gradient) system y^+L transporters would maintain a ~5-fold gradient of cationic (lower in the cytosol) and neutral (higher in the cytosol) amino acids. Thus, physiologically, the transporter mediates the efflux of a dibasic amino acid in exchange for an extracellular neutral amino acid plus Na^+ (Torrents *et al.,* 1998; Pfeiffer *et al.,* 1999; Bröer *et al.,* 2000; Kanai *et al.,* 2000). It is speculated that the Na^+ ion replaces the positive charge of the side chain of dibasic amino acids when neutral amino acids are transported.

2.2.2.1. 4F2hc/y$^+$LAT1

4F2hc/y$^+$LAT1 is expressed in the small intestine, the kidney, the placenta, the lung, the spleen and in circulating monocytes and macrophages (Barilli *et al.,* 2010; Pfeiffer *et al.,* 1999; Torrents *et al.,* 1998; Yeramian *et al.,* 2006). In the epithelial cells of the small intestine and in the renal proximal tubule, y$^+$LAT1 has a basolateral location (Torrents 1998; Pfeiffer, 1999b). It is expected that in cells with active influx of cationic amino acids (e.g., cells expressing rBAT/b$^{0,+}$AT or CATs) 4F2hc/y$^+$LAT1 will mediate export of cationic amino acids. Similar to LAT2 and b$^{0,+}$AT, y$^+$LAT1 expression along the proximal kidney tubule follows a decreasing axial gradient (Bauch *et al.,* 2003).

Mutations in the SLC7A7 gene encoding y$^+$LAT1 cause Lysinuric Protein Intolerance (LPI; OMIM database, entry 222700) (Borsani *et al.,* 1999; Torrents *et al.,* 1999). LPI is a very rare primary inherited aminoaciduria of cationic amino acids (mainly lysine, but also arginine and ornithine) with an autosomal recessive mode of inheritance (Palacin *et al.,* 2001; Simell, 2001). There are different clinical manifestations of LPI that appear after weaning and may include diarrhea, vomiting, failure to thrive, hepatosplenomegaly, lung involvement, osteoporosis, bone-marrow abnormalities, mental retardation, episodes of coma, altered immune response and chronic renal disease. LPI is highly prevalent in Finland (1/60,000), Southern Italy and Japan (1/50,000) (Palacin *et al.,* 2005). LPI leads to urine hyperexcretion and intestinal malabsorption of cationic amino acids (Palacin *et al.,* 2005). As a result, a metabolic derangement is caused, characterized by low plasma levels of dibasic amino acids, raised urinary excretion, dysfuntion of the urea cycle, leading to hyperammonaemia, orotic aciduria and protein aversion (Mykkänen, *et al.,* 2000; Torrents 1998; Borsani, 1999).

Frameshift mutations in y$^+$LAT1 have been found to be retained in the cytoplasm (Toivonen *et al*., 2013; Mykkänen, *et al.,* 2000), and some point mutations have been identified to inactivate the transporter. L334R and G54V y$^+$LAT1 mutants reached the *Xenopus* oocyte plasma membrane when co-expressed with 4F2hc but showed inactivated the transport (Mykkänen, *et al.,* 2000). No LPI mutations have been identified in the heavy subunit SLC3A2 (4F2hc).

2.2.2.2. 4F2hc/y$^+$LAT2

The transport characteristics of 4F2hc/y$^+$LAT2 are very similar to those of 4F2hc/y$^+$LAT1 (Kanai *et al.,* 2000; Pfeiffer *et al.,* 1999; Torrents *et al.,* 1998), but 4F2hc/y$^+$LAT2 preferentially mediates the efflux of L-arginine in exchange of L-glutamine plus Na$^+$ (Bröer *et al.,* 2000). y$^+$LAT2 has a broader tissue distribution than y$^+$LAT1 and is expressed in many organs like brain, heart, testis, kidney, small intestine and parotids (Bröer *et al.,* 2000). In fibroblasts of some LPI patients 4F2hc/y+LAT2 is responsible for system y$^+$L activity (Dall'Asta *et al.,* 2000).

2.2.3. System asc

System asc is a Na^+-independent transport system for short chain neutral amino acids such as glycine, alanine, serine, threonine and cystine. 4F2hc/asc1 and asc2 (which is linked to a yet unknown heavy subunit) are the two isoforms of system asc belonging to de SLC7 family.

2.2.3.1 4F2hc/asc1

4F2hc/asc1 transports, Na^+-independently, glycine, L-alanine, L-serine, L-threonine, L-cysteine, α-aminoisobutyric acid and β-alanine (Fukasawa *et al.,* 2000; Nakauchi *et al.,* 2000). This transporter has two peculiarities among light subunits. Firstly, it is able to transport D-isomers of the small neutral amino acids, in particular D-serine, with high apparent affinity (Km values of ~50 μM), and second, 4F2hc/asc1 also works as a facilitate diffuser, although it preferentially works as an exchanger.

It is expressed in the brain, lung, skeletal muscle, heart, kidney, small intestine and placenta (Fukasawa *et al.,* 2000; Nakauchi *et al.,* 2000). It is a neuronal transporter distributed all over the CNS (Helboe *et al.,* 2003), with the main physiological role in the CNS to control the synaptic concentration of the glutamate co-agonists that activate the N-methyl-D-aspartate (NMDA) receptors, like D-serine and glycine (Hashimoto and Oka, 1997). A Slc7a10 knockout mice model shows tremors, ataxia, seizures and early postnatal death (Xie *et al.,* 2005), probably due to over-activation of NMDA receptors for the elevated extracellular D-serine which causes hyper-excitability of Slc7a10 mice.

2.2.4. System X_c^-

System X_c^- is a Na^+-independent and electroneutral exchanger of extracellular anionic cystine (for which it has a high-affinity) and intracellular glutamate with a 1:1 stoichiometry (Bannai *et al.,* 1980; Makowske *et al.,* 1982; Sato *et al.,* 1999; Bassi *et al.,* 2001). This system plays an important antioxidant role because it provides cystine for the synthesis of the major endogenous antioxidant responsible for the maintenance of intracellular redox homeostasis (reduced tripeptide glutathione (GSH)), and it also supports the non-vesicular release of glutamate in brain (Bridges *et al.,* 2012; Conrad and Sato, 2012).

2.2.4.1. 4F2hc/xCT

System X_c^- is molecularly represented by 4F2hc/xCT (Sato *et al.,* 1999). The amino acid direction of this exchange is driven by the fact that at physiological conditions there are high intracellular levels of glutamate and low levels of cystine, because intracellular cystine becomes rapidly reduced to cysteine and incorporated into glutathion and protein biosynthesis pathways (Bannai *et al.,* 1984; Bannai *et al.,* 1986).

Importantly, cysteine is the rate-limiting substrate for GSH synthesis, and GSH in particular is required for proliferation, redox cycling and antioxidative defense (Reddy et al., 2008; Seiler et al., 2008).

xCT is expressed in neuronal cells, macrophages, fibroblast, pancreas, hepatocytes, kidney and cell culture lines (Sato *et al.,* 1999; Bassi *et al.,* 2001). Expression of xCT is up-regulated in cells requiring GSH, like activated macrophages and in various cancers including primary malignant brain tumors (gliomas), leukemias, lymphomas, Karposi's sarcoma or pancreatic cancer (Sato *et al.,* 1999; Bridges *et al.,* 2001; Kim *et al.,* 2001; Huang *et al.,* 2005; Lo *et al.,* 2008; Seib *et al.,* 2011). Cancer cells adjust their increased demand for nutrients due to their excessive glucose metabolism via anaerobic glycolysis (called Warburg effect; Warburg, 1956; Ganapathy *et al.,* 2009). To keep up with increased energy demand due to this permanent respiratory defect that bypasses the Pasteur effect, cancer cells express nutrient transporters for lactate and amino acids, which increase the reactive oxygen species (ROS) and redox cycle regulation (Vafa *et al.,* 2002; Semenza, 2003; Kroemer and Pouyssegur, 2008). Thus, overexpression of 4F2hc/xCT may protect tumor cells against oxidative stress by increasing glutathione synthesis (Takeuchi et al., 2013), which indeed, would protect from the oxidative stress induced by radiotherapy and chemotherapy. On the other hand, xCT loss-of-function in cellular and mouse models presents defective protection against oxidative stress in activated macrophages (Nabeyama *et al.,* 2010) and activation of the ROS/autophagy pathway in hepatocellular carcinoma (Guo *et al.,* 2011). Furthermore, malfunctioning of system X_c^-, can cause oxidative stress and excitotoxicity, both important phenomena in the pathogenesis of Parkinson's disease.

Expression of xCT in brain seems to be more related to the homeostasis of glutamate, and not so much to the important role contributing to the balance of glutathione levels, protecting cells from oxidative stress as it was though before (Sato *et al.,* 2002). Nowadays it is known that system X_c^-, due to its role of transporting glutamate, can affect excitability and have behavioral consequences; it is required for optimal spatial working memory, and its inactivation decreases susceptibility for limbic epileptic seizures (De Bundel *et al.,* 2011). When overexpressed in gliomas, it causes excess of glutamate secretion that produces excitotoxicity which ends up in neurodegeneration and brain edema in the periphery of glioma tumors and the invasion of glioma cells (when xc^- system was silenced this effects were alleviated) (Lyons *et al.,* 2007; Savaskan, *et al.,* 2008). Contrary to what happens in gliomas, in the case of chronic cocaine consumption, system xc- is downregulated, and it results in reduced extracellular levels of glutamate and diminished stimulation of extrasynaptic group II metabotropic glutamate receptors in the nucleus accumbens and prefrontal cortex (Xi *et al.,* 2002a,b; Baker *et al.,* 2003; Kau *et al.,* 2008;) causing the relapse of cocaine addiction. In these cases glutamate concentration can be restored by N-acetylcysteine and attenuates the reinstatement of cocaine-seeking behavior (Moussawi *et al.,* 2009).

In addition to its transport function, xCT serves as a fusion-entry receptor for the Kaposi's sarcoma-associated herpesvirus (KSHV; human herpesvirus 8) (Kaleeba and Berger, 2006), causative agent of Kaposi's sarcoma, since overexpression of xCT increases severely the effectiveness of fusion of the KSHV in different cell lines (Kaleeba and Berger, 2006; Qin *et al.,* 2010).

2.3. Light subunits with unknown heavy subunits

2.3.2. asc2

asc-2 (asc-type amino acid transporter-2) has a relatively low, but significant, amino acid sequence identity and similarity (~30%) to other HATs. A human ortholog of mouse Slc7a12 is not present in the human genome, and it is still not known the heavy subunit associated with mouse asc2. Mouse asc2 is expressed in kidney collecting duct cells, placenta, spleen, lung and skeletal muscle (Chairoungdua *et al.,* 2001). There are some evidences of a possible heavy subunit associated with asc2. Firstly, mouse asc-2 has the cysteine residue in the 2nd extracellular loop (between TMDs III and IV) that is conserved in all light subunits. Secondly, western blot analysis of mouse erythrocytes and kidney detected asc2 as multiple bands under non-reducing conditions and as a single band with a lower molecular mass under reducing conditions (Chairoungdua *et al.,* 2001), suggesting its association with other protein(s) via a disulfide bond.

Although co-expression of asc2 and 4F2hc or rBAT does not induce any transport activity at the cell plasma membrane, mouse asc2 was possible to be functionally characterized using fusion proteins of asc2 and 4F2hc or rBAT that were sorted to the plasma membrane. These fusion proteins exhibited a transport function corresponding to the Na^+-independent transport system asc (Chairoungdua *et al.,* 2001). Unlike the other system asc isoform (4F2hc/asc1), asc2 does not accept some of the high-affinity substrates of asc-1, such as α-aminoisobutyric acid and β-alanine.

2.3.3. arpAT

The Slc7a15 cDNA encoding arpAT (aromatic-preferring amino acid transporter) was the last member identified belonging to SLC7 family in mammals. It was found by sequence similarity search in vertebrate genomes using known human and mouse light subunit sequences (Fernandez *et al.,* 2005). Slc7a15 mRNA is expressed mainly in the small intestine and in some areas of the brain. In the intestine it shows a decreasing expression gradient in the enterocytes from the crypts to the tip of the villi. SLC7A15 is inactivated in the primate lineage (Casals *et al.,* 2008).

Like the other SLC7 members ArpAT has the conserved cysteine residue responsible for the disulfide bridge formation. It was tried the co-expression of mouse arpAT with the human heavy subunits rBAT and 4F2hc and both formed disulfide-linked heterodimers that allowed the functional characterization of the transporter. arpAT transported, sodium-independently, L-alanine, L-tyrosine, L-DOPA, L-glutamine, L-serine, L-cystine and L-arginine (Fernandez *et al.*, 2005). Kinetic and cis-inhibition studies showed a Km of 1–2 mM for L-alanine and an IC50 in the millimolar range for most amino acids, except L-proline, glycine, anionic and D-amino acids, which were not inhibitory. L-DOPA and L-tyrosine were the most effective competitive inhibitors of L-alanine transport, with IC50 values in the physiological micromolar range (Fernandez *et al.,* 2005). Nevertheless, the associated heavy subunit (rBAT and/or 4F2hc) in the original tissue remains unknown (Fotiadis *et al.,* 2013). So far, any new member of SLC3 has been identified by homology search through the genome databases, nevertheless everything points toward the existence of (a) yet unknown heavy subunit(s).

3. Structural information about HATs

Elucidating the atomic structures of the heteromeric amino acid transporters may provide valuable information to understand the translocation mechanism of the transporter and the substrate binding, which will be a tool to drive the design of specific therapeutic drugs (Kim *et al.,* 2006). Unfortunately, due to the nature of these proteins, structural data of HATs remains very scarce. At present, it has only been possible to solve at atomic level the structure of the human 4F2hc ectodomain (Fort *et al.,* 2007). Therefore, there are no available atomic structures of eukaryotic light subunits and the corresponding heterodimers.

3.1. Structure of heavy subunits

The atomis structure of human 4F2hc ectodomain was solved in 2007 in our laboratory using monoclinic (PDB 2DH2) and orthorhombic (PDB 2DH3) crystal forma at 2.1 Å and 2.8 Å of resolution, respectively (Fort *et al.,* 2007). 4F2hc-ED (Glu111-Ala529) shows similar structure as α-amylases, having in common domains A and C, but lacking of the B domain (Janacek *et al.,* 1997; Chillarón *et al.,* 2001; Fort *et al.,* 2007) (Figure 5). The domain A is a TIM barrel (or $(\beta\alpha)_8$ barrel) containing a pseudo-active site; domain B is a long loop region inserted between the third β strand and the third α-helix of domain A that contains calcium-binding site(s) (present in rBAT-ED, but not in 4F2hc-ED); finally, domain C is a C-terminal β-sheet domain which consists in 8 antiparallel β sheets. Although presenting some variability in sequence and length between amylases, domain C is always present in such proteins (Pujadas and Palau, 2001).

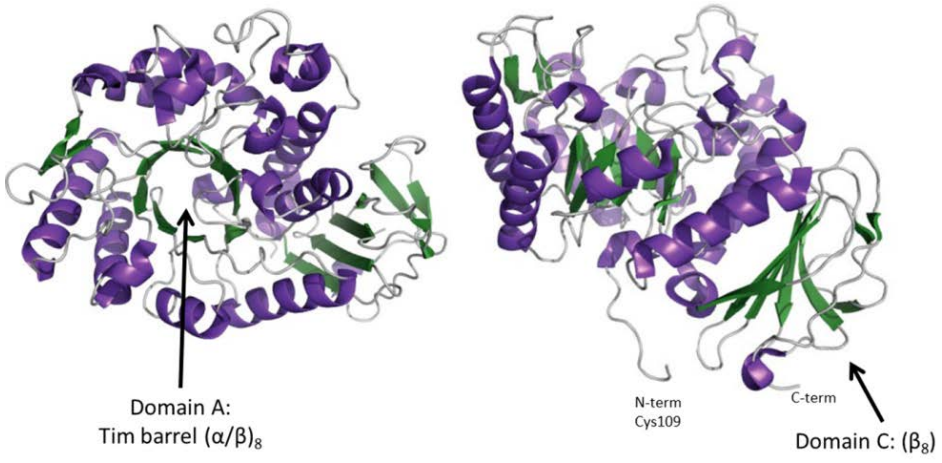


Figure 5. 4F2hc-ED structure. Top view on the left and lateral views on the right of the structure of 4F2hc-ED, that has in common with α-amylases the A domain (TIM barrel $(\beta / \alpha)_8$) and C domain (8 antiparallel β-sheets), but lacks the B domain (alpha-helixs are colured in purple and beta-sheets in green) (From Fort *et al.*, 2007).

Despite this structural similarity, 4F2hc lacks the key catalytic residues corresponding to the active site of amylases necessary for glycosidase activity. Until now, no enzymatic activity was detected in purified 4F2hc-ED when tested D-glucose, D-galactose, or D-manose derivatives of 4-methylumbelliferone as substrates (Fort *et al.,* 2007).

Interestingly, rBAT ectodomain (rBAT-ED) presents all (A, B (α2-β3 loop) and C domains), suggesting a possible glycosidase-like activity, which has not yet been proved. Unfortunately, rBAT-ED was neither possible to be expressed in *E.coli* nor purified as it was 4F2hc-ED. Nevertheless, it was possible to build a model structure using as templates the 4F2hc-ED and oligo-1,6-glucosidase from *Bacillus cereus* (PDB: 1UOK), that has 49% of amino acid sequence identity with domain β of rBAT (Joana Fort, unpublished).

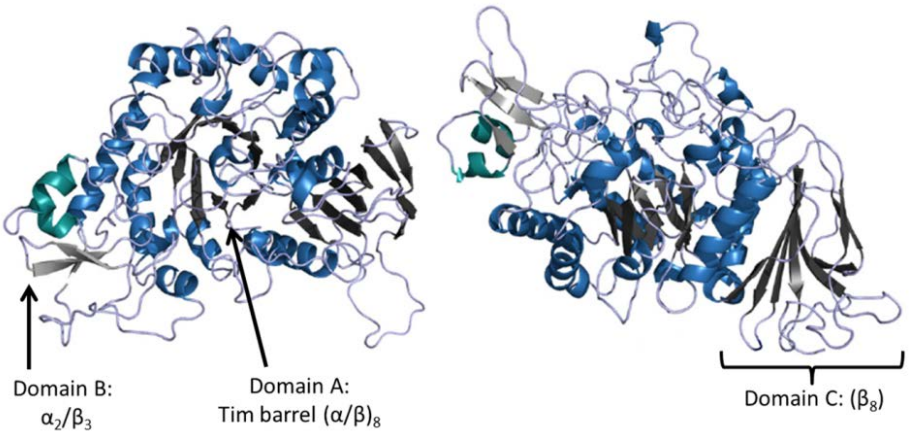


Figure 6. rBAT-ED 3D-model. Top view on the left and lateral views on the right of the 3D model of rBAT-ED obtained from 4F2hc-ED and oligo-1,6-glucosidase (PDB: 1UOK), that has all the domains characteristic of α-amylases: the A and C- domains present also in 4F2hc-ED and even the α2β3 B-domain, absent in 4F2hc (Adapted from Laura R. de la Ballina thesis) (alpha-helixs are colured in blue and beta-sheets in green).

When studying the surface charge distribution in 4F2hc-ED, it was shown a macromolecular dipole of the protein, with two charged faces: with a predominantly negative charge residues located on top (at the C termini of the Beta-sheets of the A domain) and lateral sides, and predominant positively charge residues on the bottom side (facing at the N termini of the beta-strands of the A-domain) and where the N-terminal of the ectodomain is located (Fort *et al.,* 2007). This two differentiated large positive and hydrophobic patches suggest two types of interactions.

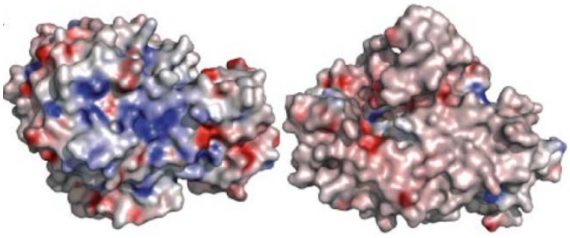


Figure 7. 4F2hc-ED electrostatic surface monomer in the monoclinic structure (2DH2). The electrostatic surface of the 4F2hc-ED monomer has a dipolar distribution of charges with a positive patch in N-terminal region (on the left) and negative surface residues on the other side of the protein (on the right).

The homodimer structure of 4F2hc opened the question about its orientation on the membrane, which was modeled by a collaborating group (Dr. Modesto Orozco) with our laboratory. The proposed model showed an electrostatic interaction between 4F2hc-ED (the positive charge region near the Cys109) and the phospholipid polar heads of the plasma membrane, which is characteristic of several peripheral and integral membrane proteins (McLaughlin and Aderem, 1995; Bhatnagar and Gordon, 1997; Conte and Matthews, 1998; Mulgrew-Nesbitt *et al.,* 2006).

4F2hc tends to homodimerize when is overexpressed in different cell types presenting a 190 KDa band in SDS-PAGE cleavable by DTT. Both monomers are linked by a disulphide bridge between cysteine residues. The same cysteine responsible for the disulfide bridge in the heterodimer formation (Fort *et al.,* 2007). Human 4F2hc homodimers showed sensitivity to endo H treatment, demonstrating they are formed in the endoplasmic reticulum (ER), and surface biotin labeling revealed that these homodimers arrived to the plasma membrane (Fort *et al.,* 2007). Human 4F2-ED has been crystallized as a monomer (PDB 2DH2) and as a homodimer (PDB 2DH3) (Fort *et al.,* 2007). A Zn^{++} atom was found at the interface of the two subunits presenting a tetrahedral coordination (Figure 8). Nevertheless, Zn^{++} does not seem to play a role in the homodimerization since mutation of the coordinating residues did not affect 4F2hc-4F2hc interaction even in the absence of the Cys109-intervening disulfide bridge (Fort *et al.,* 2007).

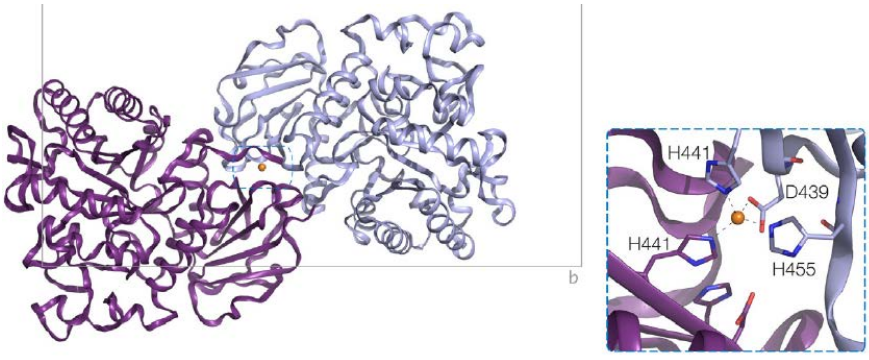


Figure 8. Homodimer of 4F2hc ectodomain (PDB 2DH3) with a zinc atom (orange sphere) coordinated in the interface zone between the two 4F2hc-ED. And on the right, detail of the residues involved in the coordination of the zinc atom: D439, H441, H455 (Fort *et al.,* 2007, adapted from R. de la Ballina thesis). This coordination patch is not conserved in all mammals.

Surprisingly, when Cys109 was mutated to serine to abolish disulfide bridge formation, 4F2hc still was able to form homodimers, suggesting that there is an important interaction area between monomers (Estévez *et al.*, 1998; Fort *et al.,* 2007). In fact, the crystallized 4F2hc-ED homodimer does not present Cys109. These results suggest that this residue would be relevant in the stabilization of the interaction but is not essential for the homodimerization process. Homo and heterodimerization both require the Cys109 to form the disulfide bridge between 4F2hc and its light subunits (3 Å distant with the cysteine residue of the subunit partner), and apparently it has preference for the light subunit *in vivo*. Nevertheless, it seems that the homodimer of 4F2hc is an artefact as it has not been found in normal tissues or tumours. Moreover, homodimerization of the 4F2hc-ED in solution has yet not been demonstrated (Turnay *et al.,* 2011).

Analyzing possible helix-helix interaction motifs, it was found in 4F2hc, but not in rBAT, the Engelman's motif, GxxxG, which can be considered as a framework for homodimerization of transmembrane α-helices (Russ *et al.,* 2000; Senes *et al.,* 2000; Senes *et al.,* 2004), and was first described in Glycophorine-transmembrane homodimerization (Lemmon *et al*., 1994). 4F2hc GxxxG residues are Gly94 and Gly98 which are located in the transmembrane domain of the protein and are highly conserved in all vertebrates. This motif seems to be sufficient to drive homodimerization of transmembrane α-helices. Nevertheless, this motif is not found in any of the light subunits associated with 4F2hc. In contrast, it is found in integrins (Berger *et al.,* 2009). So, this motif may be important in the 4F2hc homodimerization and in the 4F2hc-integrin interaction.

Apart from Cys109 responsible for the disulfide bridge, 4F2hc has another cysteine residue, Cys330 which is conserved in all mammals but not in the rest of vertebrates. By checking the structure of 4F2hc-ED it is possible to see that Cys330 is buried into the TIM barrel and it is not accessible to the outside (Figure 9). Nevertheless, it has been reported to participate in processes of cell fusion (Okamoto *et al.,* 1997). These results suggest that even Cys330 appears hidden in the solved structure; there might be a conformational movement that would make Cys330 accessible and able to interact with other proteins like β1-integrins. Actually, when Dr. Rodríguez de la Ballina and co-workers (unpublished results) analyzed the temperature factor (B factor) to analyze if Cys330 position was determined in the structure with accuracy (low B factor) or poorly defined area (high B factor) it was shown that helixes Aα5 and Aα6 have a higher B factor values than the rest of the helixes of the TIM barrel (Figure 9), meaning that they are more mobile, and that a movement of Aα5 and Aα6 could render an accessible Cys330.

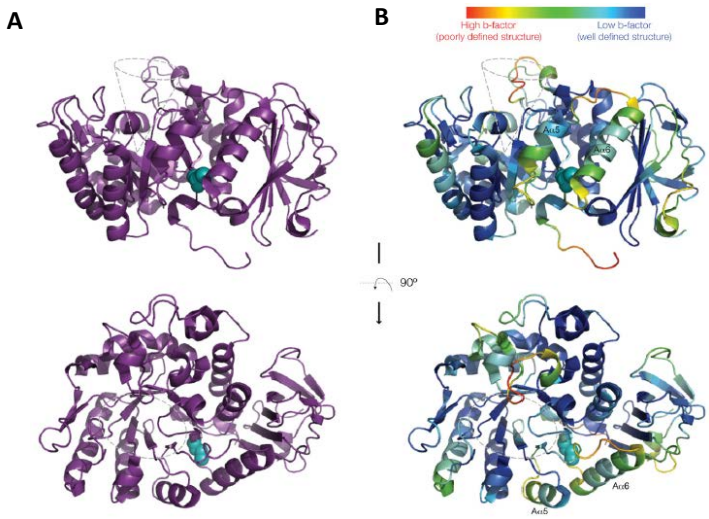


Figure 9. Cys330 location in 4F2hc-ED and analysis of the mobile areas by B factor data. A) Cys330 colored in cyan is highlighted on the 4F2hc-ED structure (lateral and top views are represented, upper and bottom parts of the figure). **B)** On the right panel the B factor parameter is represented in a gradient color to visualize the parts of the protein that are well defined (low B factor, colored in blue) from those that are not so well (high B factor, colored in red) (adapted from R. de la Ballina thesis)

3.2. Light subunit paradigm structures

Still there are not 3D structures of HAT light subunits (LHAT) available. The closest structural models are several crystal structures of prokaryotic membrane transporters homologues of the SLC7 family (Table 4). Although presenting low amino acid sequence identity with the eukaryotic light subunits all APC superfamily members structures have in common (≤ 20% amino acid sequence identity) the number of TMD as eukaryotic light subunits (Gasol *et al.,* 2004) (or, when additional transmembrane domains they appear on the periphery of this common helix core) (Faham *et al.,* 2008; Weyand *et al.,* 2008; Ressl *et al.,* 2009; Jeschke *et al.,* 2012); and a structural fold, called "5+5 inverted repeat fold" or LeuT fold (see 3.b.1. The Leu T fold), that is also expected to appear in eukaryotic light subunit structures. In SLC7 transporters the 5+5 inverted TMDs would correspond to the first 10 TMDs (Bartoccioni *et al.,* 2010).

LAT homologues	Oligomeric state	Co-crystallization	Resolution	Protein Data Bank (PDB)	References
AdiC	Tetramer	Fab Fragment	3.2 Å	3NCY	Fang *et al.,* 2009
	Dimer	-	4 Å	3LRC	Gao *et al.,* 2009
		L-arginine	3 Å	3L1L	Gao *et al.,* 2010
		L-arginine	3 Å	3OB6	Kowalczyk *et al.,* 2011
ApcT	-	-	2.32 Å	3GIA	Shaffer *et al.,* 2009
		Fab Fragment	2.48 Å	3GI9	
			2.59 Å	3GI8	
GadC	Dimer	-	3.1 Å	4DJK	Ma *et al.,* 2012

Table 4. List of the APC superfamily crystal structures of the closests homologues to LAT (SLC7). This table lists the available structures in the PDB of the prokaryotic homologues of LAT family which are the present structural paradigms of LATs. The oligomeric state, if co-crystallized, its resolution (Å) and protein data bank accession number are provided (Hediger *et al.,* 2013).

3.2.1. The LeuT fold

The LeuT fold groups distant families of secondary transporters (sequence-unrelated symporters and antiporters) into a single structural family which their 3D structure elucidation revealed the common structure fold. It was discovered for first time in the bacterial LeuT (Na^+-dependent and Cl^--dependent amino acid transporter (Yamashita *et al.,* 2005). Other crystallized transporters showing the LeuT fold were the APC superfamily members (Table 4): AdiC (arginine/agmatine antiporter) (Fang *et al.,* 2009; Gao *et al.,* 2009, 2010; Kowalczyk *et al.,* 2011); ApcT (H^+-dependent amino acid transporter) (Shaffer *et al.,* 2009); GadC (glutamate/ GABA antiporter) (Ma *et al.,* 2012) as well as transporters from other families; Mhp1 (benzyl hydantoin transporter) (Weyand *et al.,* 2008; Shimamura *et al.,* 2010); CaiT (carnitine/γ-butyrobetaine antiporter) (Shulze *et al.,* 2010; Tang *et al.,* 2010); BetP (betaine transporter) (Ressl *et al.,* 2009; Pérez *et al.,* 2012) and vSGLT (Na^+-coupled glucose transporter) (Faham *et al,* 2008). The LeuT fold is characterized by a pseudosymmetry where TMDs 1 to 5 (1st repeat) are related with 6 to 10 TMDs (2nd repeat) by a pseudo two-fold axis of symmetry located in the plane of the membrane, thus, the second repeat has an inverted topology in the membrane respect the first repeat (Figure 10) (Yamashita *et al.,* 2005).

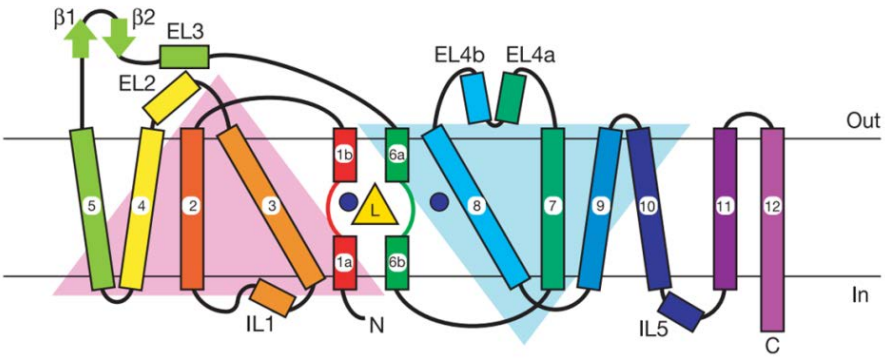


Figure 10. Topology of LeuTAa. The 12 transmembrane domains (TMD) are numbered and colored in gradient. In triangles, are represented the TMD that form the "Leu T fold" or "5+5 inverted repeat fold" between TMD1-5 and TMD 6-10. Substrate (leucine) is depicted as a yellow triangle, and the two sodium ions as two blue circles. Figure adapted from (Yamashita *et al.,* 2005).

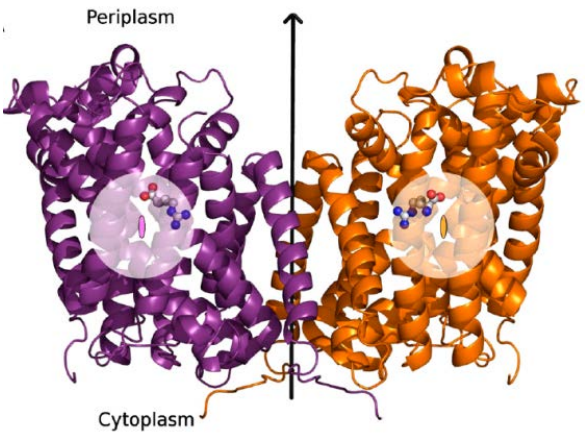
3.2.2. The closest paradigm to HAT light subunits: AdiC (L-Arginine^{+}/L-Agmatine^{++} antiporter)

AdiC belongs to the Amino acid/Polyamine Antiporter (APA) subfamily, and, together with the LAT subfamily, they form part of the APC superfamily of transporters (Casagrande *et al.*, 2008). AdiC is an arginine/agmatine antiporter expressed as a dimer in the plasma membrane of enteric bacteria, like *E. coli*, to survive in extreme acidic environments of the stomach (pH 1.5-4) and cause gastrointestinal infections (Donnenberg, 2000). AdiC forms a multiprotein complex with the cytoplasmatic decarboxylase enzyme AdiA that works together in the acid resistance system. When AdiC transports arginine into the cytoplasm, AdiA descarboxilates arginine^{+} to form a C-H bond on agmatine^{2+}, which is afterwards, exported out by AdiC, and a new arginine is importet in a one-to-one exchange stoichiometry, maintaining in this way the cytoplasmic pH in 4.7 (Foster, 2004).

The atomic structure of AdiC is the closest paradigm to HAT light subunits (with ~15-17% amino acid sequence identity with LAT subfamily), and represents an excellent model for understanding the molecular architecture of transporters from the APC superfamily. The first structure of AdiC, obtained in 2008 at 6.5 Å by cryo-TEM, showed a dimeric protein assembly (Casagrande *et al.,* 2008). In the last years it has been solved the atomic structure of AdiC in three conformational states (Table 4; Figure 12): the outward-open substrate free conformation (PDB: 3LRB and 3NCY) (Gao *et al.,* 2009; Fang *et al.,* 2009); the open-to-out Arg^{+} bound conformation (PDB: 3OB6) (Kowalczyk *et al.,* 2011); and the occluded substrate-bound conformation (PDB: 3L1L) (Gao *et al.,* 2010), but remain to be elucidated the rest of conformers (i.e., the inward-facing structures) of the LeuT fold transport cycle for this protein (Figure 12).

In our laboratory, the crystal structure of the N101A mutant of AdiC transporter in the open-to-out Arg^{+} bound conformation was solved at 3.0 Å (Kowalczyk *et al.,* 2011) (Figure 11).

A



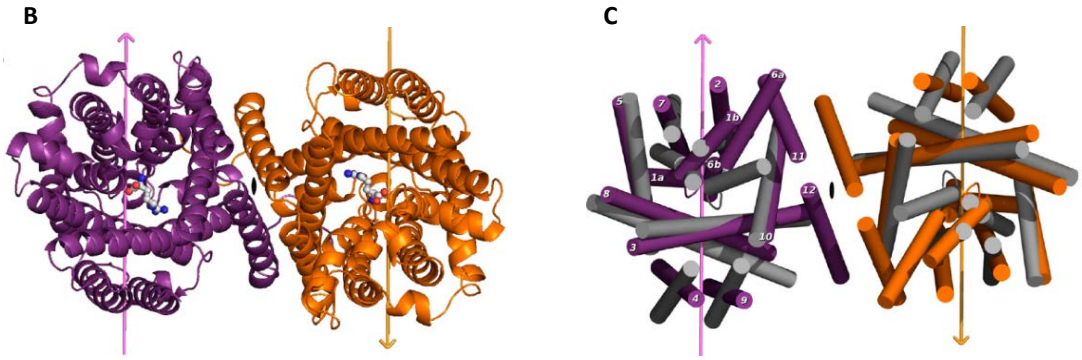


Figure 11. Structure of AdiC-N101A bound to Arg^+. A) Lateral and **B)** periplasmic views of the AdiC-N101A-Arg^+ complex homodimer (in purple, protomer 1 and in orange, protomer 2). The bound substrate is depicted with a ball-and-stick model. The two-fold subunit axes (ovals or arrows) are colored as their corresponding protomers. The dimer axis (arrow or ovals) is indicated in black. **C)** Model of the open-to-in conformation (gray cylinders) generated by the application of the 5 + 5 inverted repeat symmetry to the first 10 TMs of the open-to-out conformation (colored as in A and B). The conformational changes occur away from the dimer interface and along each subunit axis (Kowalczyk *et al.,* 2011).

The combination of structural, functional, substrate docking, and molecular dynamics (MD) studies reveal the coordination of the substrate guanidinium group with AdiC residues Trp293 (in TMD8) and Asn101 (in TMD101) during the transit to the occluded state (Figure 12). In both conformations, open-to-out apo and occluded conformations, the guanidinium group interacts with the side chain of Trp293 through π-cation interaction, and with the side chain of Asn101 to form a hydrogen bond (Kowalczyk *et al.,* 2011).

Interestingly, molecular dynamics also showed the high mobility of L-arginine in the AdiC N101A structure between residues Trp293 and Trp202, whereas the position of L-arginine docked in the open-to-out apo structure of AdiC (PDB 3LRB) (Gao *et al.,* 2009) is very stable (Kowalczyk *et al*, 2011). In the later structure, Trp202 (in TMD6) and Phe350 (loop between TMD9-10) have the orientation and distance of a π-π interaction, contributing to hold Trp202 in a rotamer closer to the substrate. Indeed, the presence of L-arginine freeze Trp202-Phe350 interaction (molecular dynamic studies), which in turn might contribute to hold L-arginine (guanidinium group)-Trp293 (π-cation interaction) (i.e. semi-occluded state) (Kowalczyk *et al,* 2011). This represents the second L-arginine-induced fitting mechanism (or conformational sampling) in the interaction of periplasmic L-arginine with AdiC. Transition from the semioccluded state to the outward-facing L-arginine occluded state (PDB 3L1L) (Gao *et al.,* 2010) occurs through 40º tilting of TMD6 (i.e. the N-terminal part of TMD6) and 10º tilting of TMD10 towards the substrate. In conclusion, Trp293 and Asn101, two high conserved residues among APA family members, play a dual role in the transport cycle: first, they recognize the substrate in the open-to-out state, and second, they stabilize the occluded conformation, which is needed for the transition to the inward-facing conformations.

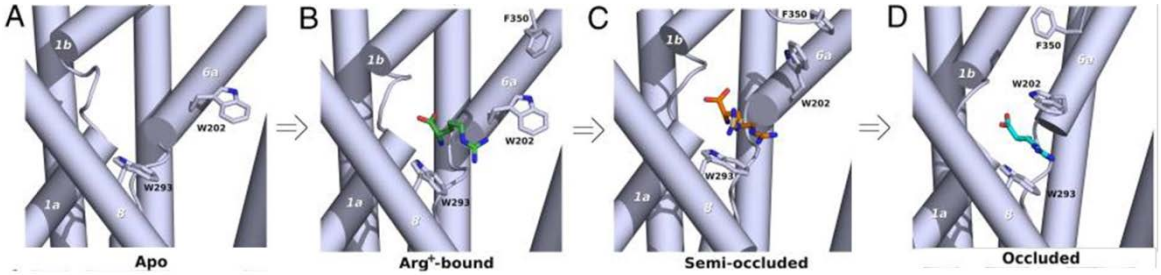


Figure 12. Proposed mechanism of substrate (Arginine) recognition and induced fit by AdiC. In this scheme it is represented the transition from the Apo to the occluded conformations where: a) Arg is recognized by the apo conformation of AdiC (3NCY), b) bound (3OB6), c) semi-occluded and finally d) occluded (Kowalczyk *et al.,* 2011).

3.2.3. The alternative access mechanism of prokaryotic transporters with LeuT fold

When the substrate binds, the transporters undergo changes not only at the level of internal residues, but also, there is a general conformational change that makes the protein transit from outward to inward conformation. It has been described that transporters belonging to the LeuT structure family translocate the substrate transiting through different conformations in a common way (Figure 13).

The appearance of more atomic structures of functionally different yet structurally related transporter families has provided insight into the principle of alternating-access mechanism which describe this transport cycle (Figure 13): it starts in an open-to-out conformation (or apo outward-facing) in which the substrate binding pocket is accessible from the extracellular side. The binding of the substrate to the apo transporter (substrate-bound open) triggers the transition to the occluded state, where first, a so called thin gate encloses the substrate (occluded outward-facing conformation). At this step, the protein undergoes a conformational change that makes a fully occluded state of the protein; the substrate is blocked by a thin (usually the side chain of a single residue) and a thick (usually several TMDs) gate preventing the diffusion of the substrate to either side of the membrane). Transition to the open-to-in state starts by an occluding inward-facing conformation in which the substrate remains occluded by a thin gate in the inner side. This state is followed by the substrate-bound inward-facing conformation, to the apo inward facing conformation in which substrate is finally released to the cytoplasmatic side of the membrane. At a given time, the substrate can only be accessed from one side of the membrane (Kowalscyk *et al.,* 2011).

Only one of these gates is open at a time, allowing substrates and ions to reach the primary binding sites without opening up a continuous transmembrane pore. This is consistent with the alternating access mechanisms depicted in the 1950s (Mitchell, 1957; Patlak, 1957).

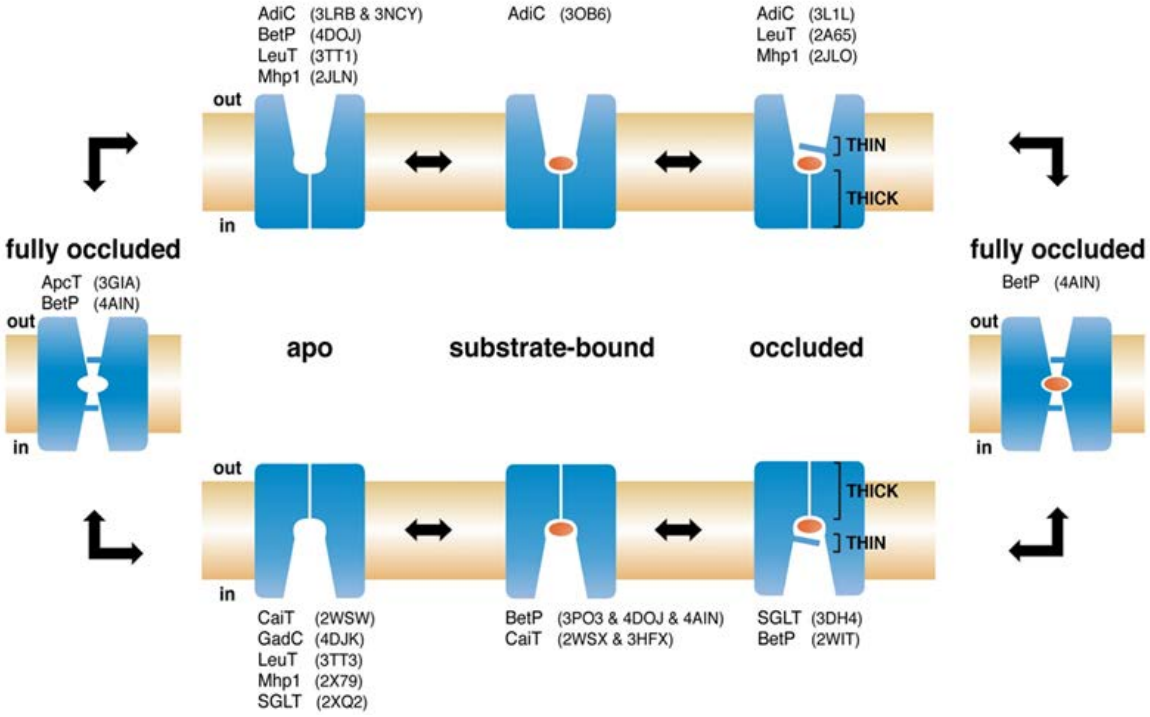


Figure 13. Conformational states of the alternative access mechanism of prokaryotic transporters with LeuT fold. Substrate is represented by a red ellipse. Binding to the open-to-out apo state, the substrate-bound state evolves to an occluded state, where two gates (thick and thin) prevent the diffusion of the substrate to either side of the membrane. Occlusion of the substrate by a thin gate is a common mechanism in the transport cycle of these transporters, in spite of involving different molecular events, as described for LeuT, vSGLT, Mhp1, BetP and AdiC. The inward-facing states are symmetrically related to the outward-facing ones. Transition to the inward-facing states requires a transient fully occluded symmetrical intermediate. In ion-coupled symporters (LeuT, vSGLT, Mhp1, ApcT and BetP) a free transition between the apo structures (outward- and inward-facing) is required to close the transport cycle. The apo occluded structure of ApcT is close to this state. In antiporters (AdiC and CaiT), the return to the outward-facing states requires the binding and translocation of a new intracellular substrate that will move the transporter back through all the states but in the opposite direction. PDB access codes of the corresponding transporter structures/conformational states are indicated in parentheses. (Adapted from (Fotiadis *et al.,* 2013).

3.2.4. Closest bacterial SLC7 transporter characterized: SteT (L-serine/L-threonine antiporter)

The first prokaryotic member of the SLC7 family (~30% amino acid identity with eukaryotic light subunits) cloned and functionally characterized in our laboratory was L-serine/L-threonine antiporter (SteT) of *Bacillus subtilis* (Reig *et al.,* 2007). Differently to all vertebrate SLC7 members, SteT lacks of the conserved cysteine between TMD3 and TMD4 responsible for the disulfide bridge formation. Nowadays, is still not known the existence of any HAT heavy subunit homologue in prokaryotic organisms. The single-particle negative-staining electron microscopy (TEM) of detergent-purified SteT showed that this protein expresses as a monomer and has elliptical shape according to its dimensions (diameters 6x7 nm), and presents a central vestibule (Figure 14) (Reig *et al.,* 2007).

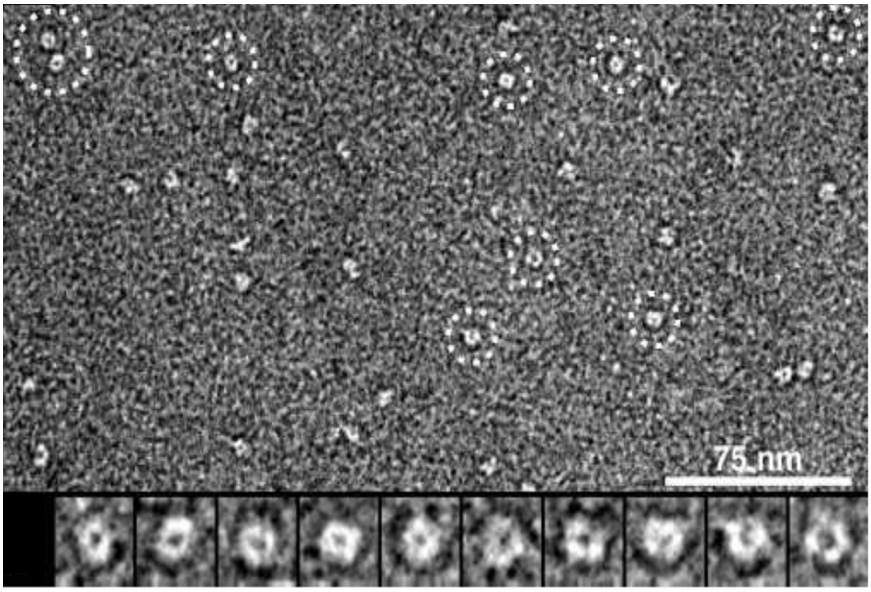


Figure 14. Single particles of detergent-solubilized and purified SteT were visualized by negative staining transmission electron microscopy (TEM). Detergent-solubilized SteT was purified and analyzed by SP-NS TEM showing an homogeneous population of monomers, with a diameter around 7 nm, corresponding to the protein and the additional mass arising from the detergent belt and endogenous lipid bound to the protein that have to be considered.

Functionally, SteT is an electroneutral obligatory exchanger that transports preferentially L-serine and L-threonine (1:1 stoichiometry), although it can also recognize with less affinity aromatic amino acids (Reig *et al.,* 2007). The binding pocket of SteT has been studied by testing cysteine reactivity of single-cysteine mutants of TMD 8 introduced in a functional cysteine-less SteT mutant combined with functional assays (Bartoccioni *et al.,* 2010). This work revealed that Cys 291 and Lys 295 are implicated in the substrate recognition and specificity (Bartoccioni *et al.,* 2010). Due to the lack of atomic structures

available of SteT, it was built a structural model of this transporter based on AdiC. Substitution of the residue Lys295 for a smaller residue opers in a wider space in the substrate binding site that results in acquisition of a broader substrate profile (Bartoccioni *et al.,* 2010). At the moment there are no atomic structures of SteT wild type available, since it behaves as a very unstable protein and mutants have diffracted not at high resolution after several trials performed in our laboratory. Nevertheless, two mutant forms of SteT, I134V-A377T and L210Q-229V have been found to be better candidates for crystallization studies (Rodríguez-Banqueri *et al.,* 2012). These mutants were generated by random mutagenesis and selected using the GFP-split technology adapted for membrane protein (Rodríguez-Banqueri *et al.,* 2012) and the adapted high-through put pipeline to screen membrane proteins for crystallization from E. coli (Drew et al., 2005; Kawate and Gouaux, 2006; Rodríguez-Banqueri *et al.,* 2012).

3.3. Interaction between heavy and light subunits

The first evident interaction that exists between heavy and light subunits is the disulphide bridge that is conserved in all amino acid transporters in mammals (HATs) (Palacín *et al.,* 2005). In the heavy subunits this Cys is located a few amino acids away from the transmembrane domain, in the extracellular neck, located in the Cys 114 in human rBAT and Cys 109 in human 4F2hc. Light subunits have also a conserved cysteine between transmembrane domains 3 and 4 (Mastrobernardino *et al.* 1998; Pfeifer *et al.* 1998; Torrents *et al.* 1998; Nakamura *et al.* 1999,).

The mutation of the cysteines responsible for the disulfide bond reduces, but does not abolish, trafficking of the heterodimer to the cell plasma membrane, which is observed by reduced transport activity (Pfeiffer *et al.* 1998; Nakamura *et al.* 1999). This observation suggests that other types of interactions (non-covalent) occur between the heavy and the light subunit (Estévez *et al.*, 1998; Pfeiffer *et al.*, 1998).

Several studies have been conducted to address which region of heavy subunit is needed to interact with light subunits:

Chimaeric and truncated forms of rBAT and 4F2hc co-expressed in *Xenopus* oocytes with $b^{0,+}$AT or LAT1 were analyzed by immunoprecipitation and transport function, and resulted that cytoplasmatic tail and transmembrane domain of rBAT play a dominant role in selective functional interaction with $b^{0,+}$AT; and the extracellular domain of rBAT appears to facilitate specifically L-cystine uptake (Franca *et al.*, 2005). In the case of 4F2hc, it is dispensable either the cytoplasmatic or extracellular glycosidase-like domains are for the functional interaction with LAT1.

A different group demonstrated that 4F2hc truncations in the C-terminal slowed down the traffic of the light subunit LAT1 to the plasma membrane but did not kill the transport, behaving as efficiently as the wild type (Bröer *et al.,* 2001). Conversely, the same truncations produced a complete loss of function in LAT2 and y^+LAT2, indicating that LAT1 might have different characteristics that the rest of light subunits interacting with 4F2hc. In these experiments, deletion of 4F2hc-ED abolished the heterodimer formation and transport at the plasma membrane with LAT2, y+LAT2 subunits (Bröer *et al.,* 2001).

Interestingly, chimeras of 4F2hc and CD69 (type II glycoprotein) showed that the relevant part of 4F2hc interaction with LAT1 was the extracellular domain (-COOH) (not the transmembrane or the cytoplasmatic domain) to keep the transport activity of LAT1 (Fenczic *et al*. 2001) (Figure 15). It was shown that the lack of transport was due to a defect in the trafficking of the holotransporter to the plasma membrane (L.R. de la Ballina PhD thesis, 2011). These results suggest that the different light subunits recognition would require non-covalent interactions of the ectodomain to reach the plasma membrane. In addition, they demonstrated that the cytoplasmatic and the transmembrane domains of 4F2hc are needed for its effects on integrin function. And, that transport activity and integrin interaction functions of 4F2hc are independent and are mediated by different domains in 4F2hc (Fenczik *et al.*, 2001; Henderson, 2004; Cai *et al.*, 2005).

Thus, different group's results are in concordance about the importance of the heavy subunit ectodomain for the light subunit trafficking to the plasma membrane and transport function. In addition, it is also shown that additional interactions are present between both subunits and needed, but still the specific way of interaction has not been cleared.

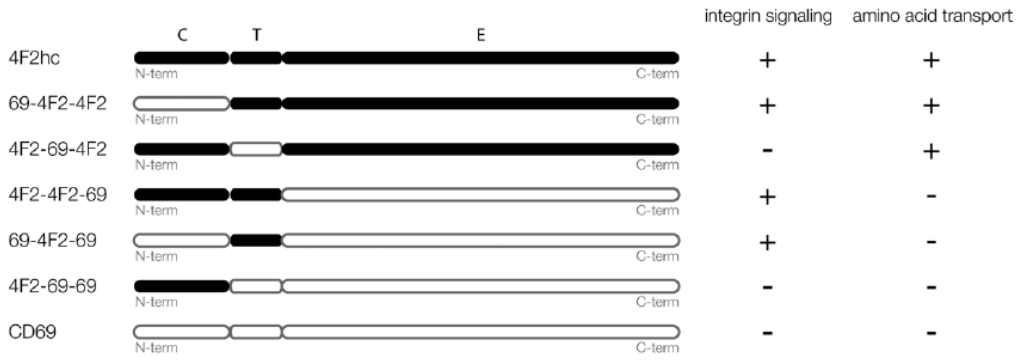


Figure 15. Scheme of the results obtained from testing the interaction between the different chimeras generated of 4F2hc and CD69 (another type II glycoprotein not implicated in amino acid transport or integrin signaling) **with integrins or with amino acid transporters indicated by transport function.** Chimeras were constructed by combinations of the ectodomain, transmembrane and cytoplasmic domains. (Fenczik *et al.*, 2001; Henderson, 2004; Cai *et al.*, 2005 and L.R. de la Ballina, PhD thesis, 2011).

In addition, 4F2hc interacts and stabilizes other transporters like GLUT1 without the need of a disulphide bridge. Endogenous association was observed in brain (mouse) and HeLa cells, and also by coimmunoprecipitation (CoIP). Over 4F2hc overexpression, GLUT1 increased and subsequently, the

glucose transport. In addition, GLUT1 degradation by cycloheximide was significantly diminished by 4F2hc overexpression and increased by 4F2hc siRNA treatment, indicating that 4F2hc is involved in the stabilization of GLUT1 and contribut to the regulation of also glucose metabolism (Ohno *et al.,* 2011).

On the other hand, when light subunits do not reach the plasma membrane due to a lack of heavy subunit or a lack of the interacting domain of heavy subunit, it does not mean that they are not functional, since it has been demonstrated that reconstituted light subunits in proteoliposomes are functional (Reig *et al.,* 2002). All these data point out that non-covalent interaction between heavy subunit and light subunit are present, and needed to keep trafficking to the plasma membrane the holotransporter, but still the specific way of interaction has not been cleared.

Objectives

Objectives

HATs play central roles in human physiology and are involved in diseases such as inherited aminoacidurias, viral infection and cancer. The low amino acid sequence identity of the eukaryotic light subunits with transporters of known structure preclude the generation of robust human transporter models. Moreover, these solved structures give no clues about the supramolecular organization of HATs, i.e., the architecture and interactions between the heavy and the light subunits. When the atomic structures of HATs be available, these structures will facilitate the comprehension of their molecular mechanisms of transport, the generation of HAT-specific ligands by structure-based drug design, and the understanding of the mechanisms involved in the functional interaction between integrins and 4F2hc-associated transporters.

One of the research interests of my group is the obtainment of structural information of Heteromeric amino acid transporters. In this context, my thesis had two main objectives:

I **Validation of the first low resolution 3D model of a human HAT (4F2hc/LAT2) by intersubunit crosslinking and study of 4F2hc-ED impact on LAT2 stability**

II **Identification of new HAT targets for 3D-crystallization**

The thesis is structured in two differentiated chapters containing each an introduction, materials and methods, results and discussion.

Chapter I

Validation of the first low resolution 3D model of human 4F2hc/LAT2 by intersubunit crosslinking

I. Introduction

Chapter I: Introduction

Previous work in the laboratory carried out by Dr. Rosell and Dr. Costa culminated with the selection of human 4F2hc/LAT2 as the best candidate for structural studies after having tested all human light and heavy subunits, and the heterodimers of those with high overexpression yields in the yeast *Pichia pastoris* and efficiently solubilized in detergent. n-Dodecyl-β-D-maltopyranoside (DDM) resulted the most efficient solubilizing the heterodimer and 4F2hc and LAT2 monomers (Costa *et al.,* 2013). Functionality of the overexpressed heterodimer was then tested in vivo in *Pichia* cells by analyzing the [3H]L-leucine uptake, which showed transport function compared with the untransformed cells (Figure 16).

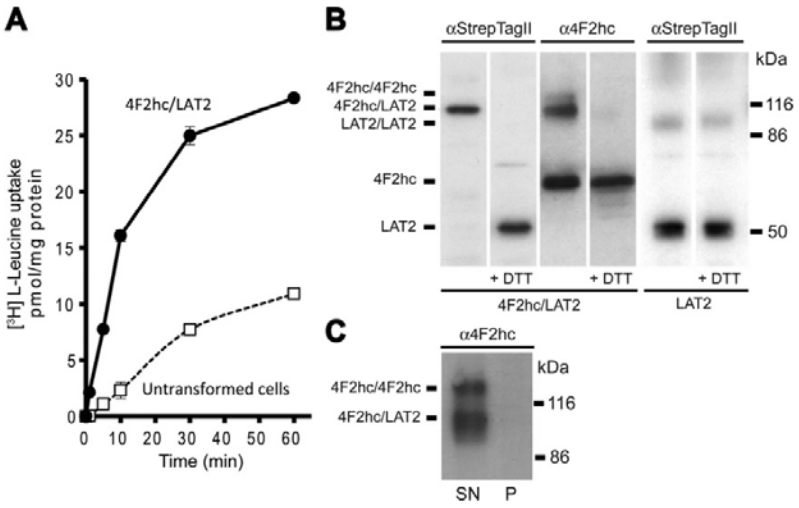


Figure 16. Expression, function and solubilization of the human heterodimer 4F2hc/LAT2. A) In vivo uptake of radiolabeled [3H]L-leucine (10 µM) into *P. pastoris* cells overexpressing human 4F2hc/LAT2 is observed when compared with control (untransformed) cells (Costa *et al.,* 2013). **B)** Western blot analysis of membranes isolated from *Pichia* cells expressing 4F2hc and LAT2. The anti-StrepTagII antibody indicated the presence of heterodimers only and no LAT2 monomers under non-reducing conditions. Under reducing conditions (+ DTT), the heterodimer disappeared yielding the LAT2 monomer band. The anti-4F2hc antibody indicated the presence of 4F2hc/4F2hc homodimer, 4F2hc/LAT2 heterodimer and 4F2hc monomer in the absence of DTT. Under reducing conditions (+ DTT), only 4F2hc monomers were visible. Western blot analysis using anti-StrepTagII antibody of isolated membranes from LAT2 expressing *Pichia* cells indicated the presence of LAT2 (major fraction; at ~50 kDa) and an artefactual LAT2 homodimer (minor fraction; at ~100 kDa), both under reducing and non-reducing conditions. Such dimerizations of light subunits in PAGE have been reported (Fernández *et al.,* 2006). **C)** The detergent DDM completely solubilized 4F2hc/LAT2 and 4F2hc/4F2hc from isolated membranes: all protein is found in the supernatant (SN) and none in the pellet (P) after ultracentrifugation.

This heterodimer was then purified by Co^{2+} and Strep-Tactin affinity chromatographies and its behavior analyzed by size exclusion chromatography. Human 4F2hc/LAT2 eluted in a peak corresponding to the heterodimer size, even though, the heterodimer was not stable enough for 3D crystallization studies, since it tended to aggregate during the purification process. Fortunately, the amount and quality of the purified heterodimer was good enough to determine the structure and oligomeric state of detergent-solubilized human 4F2hc/LAT2 from projection images obtained by

Transmission Electron Microscopy and Single Particle Analysis (TEM-SPA). In collaboration with Dr. Fotiadis (University of Bern), the first 3D low resolution model at 21Å of a human HAT, the 4F2hc/LAT2 was obtained (Rosell *et al.*, 2014).

The electron micrographs of the purified 4F2hc/LAT2 contained different particles composed of two globular domains of different sizes (Figure 17).

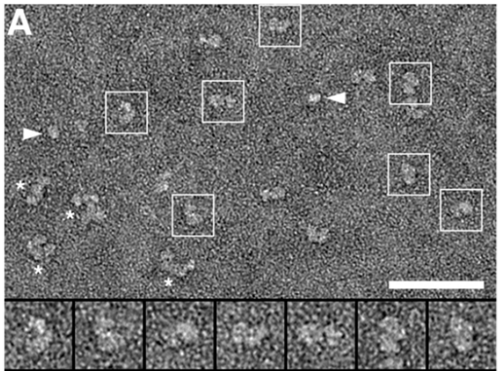


Figure 17. Transmission Electron Microscopy and Single Particle Analysis (TEM-SPA). Overview electron micrograph of purified, negatively stained 4F2hc/LAT2 heterodimers. The 4F2hc/LAT2 heterodimers are boxed, magnified and are displayed in the gallery bellow. Disrupted heterodimers lead to some monomeric forms of 4F2hc or LAT2 that are indicated with arrowheads, and asterisks indicate small protein aggregates. (Scale bar: 50 nm.) The frame sizes of the magnified particles in the gallery are 21.8 nm (Rosell *et al.*, 2014).

The different particle projections were image refined, filtered and contrasted to increase the signal to noise ratio (Figure 18A) and a gallery of 15,000 images was built from which single particle projections class averages were calculated (Figure 18B). Most class averages indicate a clear difference in size between the two domains.

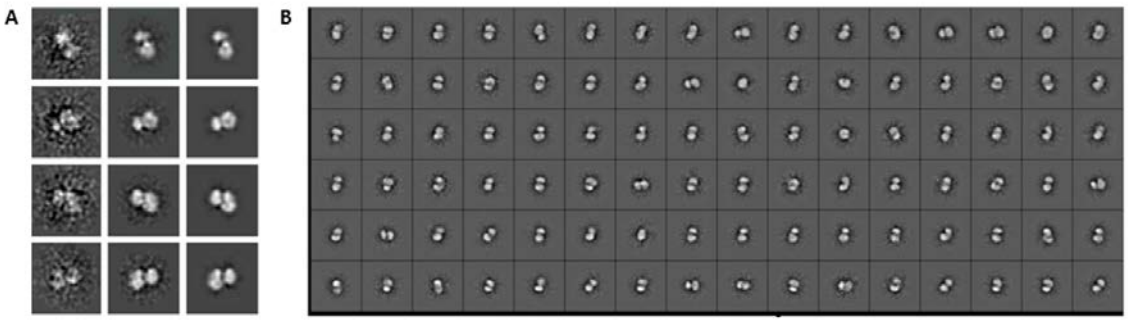


Figure 18. TEM image refinement (A) and gallery of the different single particle projections class average (B) (Rosell *et al.*, 2014)**.**

The structure of 4F2hc/LAT2 was determined by 3D reconstruction from negatively stained complexes (Fig. 17). The resolution of the 3D map obtained from the final refinement run was 21 Å (see Figure S1B in Rosell *et al.,* 2014) and had a homogeneous angular distribution of projections (see Figure S1C in Rosell *et al.,* 2014). The calculated 3D map reflects a structure composed of a smaller and a larger

density (Figure 19; see also Figure S1D for additional views in Rosell *et al.,* 2014). The smaller domain lies tilted (not flat) on the larger domain (Figure 19A, see black, dotted line indicating tilt). Consequently, the complex has on one side a cleft (Figure 19A, arrowhead) and on the opposite side a seal (Figure 19A, white, dotted curve). Importantly, the crystal structure of 4F2hc-ED (Fort *et al*., 2007) nicely fits into the smaller density (Figure 19B) making the assignment of 4F2hc and LAT2 in the 3D map possible.

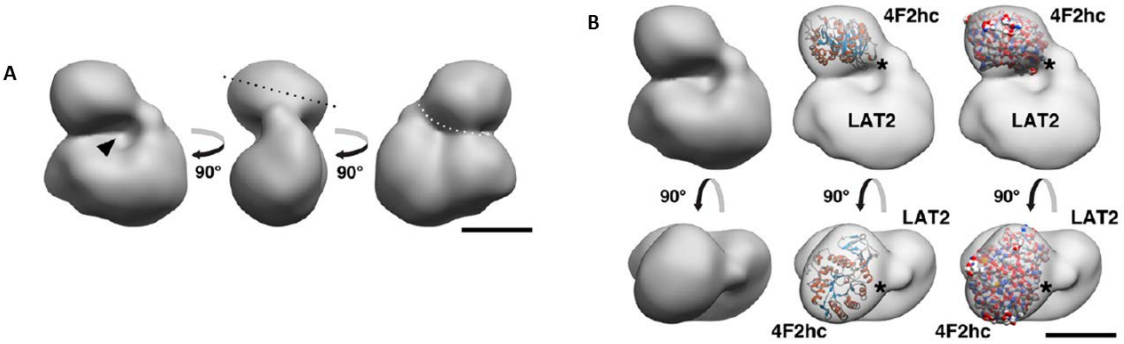


Figure 19. 3D reconstruction of 4F2hc/LAT2 calculated from projections of negatively stained heterodimer particles. A) Different side views of the 3D model are shown. 4F2hc/LAT2 is composed of a large and a small density. The small density is located on top of the large density and is tilted, as indicated by the black dotted line. As indicated by an arrowhead, the 3D model features a distinct cavity. On the opposite side, both subunits are in close contact, as marked by the white, dotted curve. (Scale bar: 5 nm.). **B)** Side (Upper) and top (Lower) views of the 4F2hc/LAT2 3D reconstruction without and with the fitted crystal structure of the 4F2hc-ED (Protein Data Bank ID code: 2DH2). The fitting assigns the small and large subunits to 4F2hc and LAT2, respectively. The structure of 4F2hc-ED is represented as a cartoon and surface model. Asterisks indicate the location of the N terminus in the 4F2hc-ED crystal structure. (Scale bar: 5 nm.) (Rosell *et al.*, 2014).

As determined by scanning transmission electron microscopy mass measurement, unstained n-dodecyl-β-D-maltoside (DDM) purified membrane proteins prepared in a conventional way for TEM are associated with >55 kDa of co-purified endogenous lipids and DDM molecules (Jastrzebska *et al.,* 2011). Thus, the total mass of the LAT2/lipid/DDM ternary complex plus the single TMD and the cytoplasmic N-terminal domain of 4F2hc containing His-tag and protease cleavage site (>120 kDa) is significantly larger than that of the 4F2hc-ED (~50 kDa). This, together with the fitting of the 4F2hc crystal structure into the smaller density, supports the assignment of LAT2 to the larger domain of the 3D map. In figure I.4.B the location of the N-terminus in the 4F2hc-ED crystal structure is marked by an asterisk. This location is close to an additional density connecting small and large domain that possibly arises from the N-terminal TMD of 4F2hc and extracellular loops of LAT2. This low resolution 3D model represents the first structural information on the supramolecular organization of the heteromeric amino acid transporters, and it suggests that 4F2hc ectodomain (4F2hc-ED) covers the extracellular surface of the transporter subunit LAT2.

Due to the low resolution of the 4F2hc/LAT2 structure obtained from TEM-SPA, part of this thesis focuses on the validation of this model. To that aim, intersubunit crosslinking between sulfhydryl groups of cysteine residues in both subunits was performed. Pairs of cysteines, site-specific introduced into each subunit, were tested for crosslinking in purified 4F2hc/LAT2 heterodimers, with the molecular

rulers, bismaleimidoethane (BMOE) and/or 1,8-Bismaleimido-diethyleneglycol ($BM(PEG)_2$) which are homobifunctional thiol-reactive probes with a dynamic simulated average length of 6.27-10.52Å and 3.51-14.26Å respectively (Green *et al.,* 2001). Chemical crosslinking is found to be a ideal strategy for unambiguous demonstration of protein-protein interactions, in vitro. The formation of crosslinks between two distinct proteins is a direct and convincing evidence of their close proximity. In addition to information on the identity of the interacting proteins, crosslinking experiments can reveal the regions of contact between them.

Interestingly, whereas the purified human 4F2/LAT2 heterodimer was soluble, it was impossible the purification of LAT2 alone because this light subunit aggregated fast during the purification process (Rosell *et al*., 2014). Moreover, isolated *Pichia* membranes expressing 4F2hc/LAT2 heterodimers or LAT2 monomers were solubilized in DDM and reconstituted into proteoliposomes. In addition, only 4F2hc/LAT2 heterodimers but not LAT2 monomers could be successfully reconstituted as functional proteins into proteoliposomes because LAT2 monomers aggregated before inserting into liposomes (Figure 20). These results indicated a possible role of 4F2hc stabilizing the light subunit.

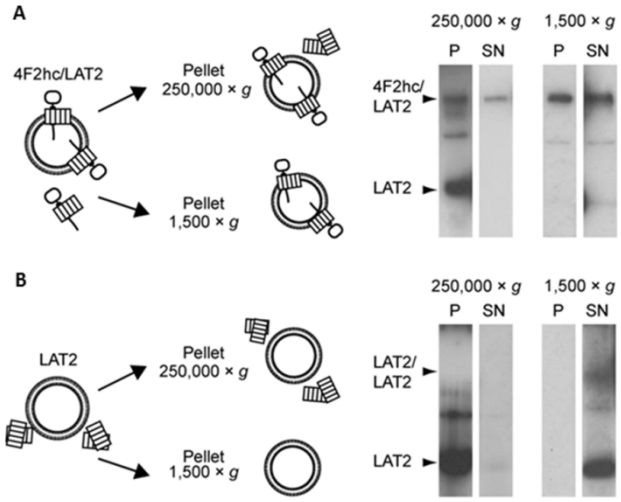


Figure 20. Reconstitution into liposomes of 4F2hc/LAT2 and LAT2. Reconstitution of heterodimers and LAT2 into liposomes. Membranes from *Pichia pastoris* co-expressing His-4F2hc and Strep-LAT2, or expressing Strep-LAT2 alone were solubilized with 1 % DDM. Detergent was then removed to induce reconstitution of the membrane proteins into proteoliposomes. To distinguish between reconstituted and aggregated proteins, preparations were centrifuged (1,500 × g, 1 h, 4 ºC) to precipitate proteoliposomes or ultracentrifuged (250,000 × g, 1 h, 4 ºC) to precipitate proteoliposomes and protein aggregates (scheme). Western blot analysis under non-reducing conditions with αStrep was used to detect Strep-LAT2, as monomers (LAT2) or heterodimers (4F2hc/LAT2). P, pellet. SN, supernatant (Results and image from Dr. Costa thesis).

These results opened a new question, whether the ectodomain alone could suffice this stabilizing effect on LAT2. To that aim, the second part of this chapter is focused in the study of 4F2hc-ED impact on LAT2.

I. Results

I.1. Crosslinking of 4F2hc-ED and LAT2 and docking analysis

To confirm the interaction between 4F2hc-ED and LAT2 that has been observed in the 3D model of 4F2hc/LAT2, different cysteine mutants on each subunit were generated to test intersubunit crosslinking between cysteine (Cys) residues. Crosslinking experiments were performed in mammalian cells (HEK 293T cell line) (co-transfected with both subunits), because of the convenience in generating mutants compared to *Pichia pastoris*. To mimic the most the conditions of the 3D model, crosslinking was assessed in purified 4F2hc/LAT2 heterodimers using the same detergent conditions (solubilized with n-dodecyl β-D-maltoside (DDM) and purified by Ni^{2+}-affinity chromatography) using the same protein tag design (N-terminally His-tag 4F2hc (His-4F2hc) and N-terminally Strep-TagII-tag LAT2 (Strep-LAT2)).

Endogenous Cys forming the disulphide bridge were maintained in order to restrict the heterodimer mobility, through the original fixed position between C109 in 4F2hc and C154 in LAT2. In fact, these Cys were not expected to interfere in the crosslinking reaction since they are not reactive. For that reason, crosslinkers resistant to DTT were chosen, so that, remaining heterodimers after crosslinking reaction and DTT addition (to quench reagent excess and cleave endogenous disulphide bridge) would correspond only to crosslinked subunits. Finally, western blotting under reducing conditions and using an anti-Strep antibody would reveal intersubunit crosslinking as DTT-resistant heterodimers. Interestingly, coimmunoprecipitated Strep-tag-LAT2 will also reveal percentage of crosslinking. Thus, CoIP Strep-tag-LAT2 monomers would correspond to immunoprecipitated 4F2hc/LAT2 heterodimers not crosslinked. In contrast, CoIP Strep-tag-LAT2 heterodimers would represent CoIP of crosslinked 4F2hc/LAT2 heterodimers.

Since several new Cys site-directed mutants were generated, they were validated previous crosslinking experiments: checking the heterodimer formation of new Cys mutants and their functionality (i.e., transport activity).

I.1.1. Residues selected in LAT2 and 4F2hc to perform intersubunit cysteine crosslinking assays

At the beginning of this project, there was no human LAT2 3D model built and no docking of 4F2hc-ED on LAT2 was still calculated. The different cysteine mutants generated were located in the region of 4F2hc most probable to be interacting with LAT2 extracellular loops. That means, that 4F2hc-ED region with glycosylated residues was orientated facing the extracellular medium.

In the case of LAT2, the extracellular loops are the expected region to be interacting with 4F2hc-ED. Analyzing the endogenous LAT2 Cys residues predicted to be located at the extracellular face, we found that it has two: Cys 154, which is the cysteine responsible to form the disulphide bridge with the heavy subunit (colored in blue, Figure 21); and Cys 210, not conserved cysteine present at the beginning of the extracellular loop 2 of LAT2 (between TMD5 and TMD6) (colored in purple, Figure 21).

This endogenous cysteine is localized far enough of the disulphide bridge cysteine at the external surface of LAT2 to be a good candidate residue to test crosslinking with 4F2hc-ED. Then, LAT2 wild-type was selected to be assayed for crosslinking with the different 4F2hc cysteine mutants through its endogenous Cys 210.

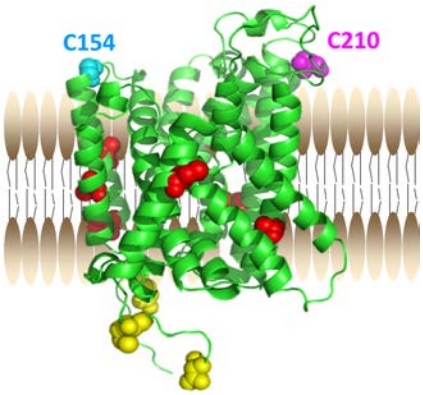


Figure 21. Reactive cysteines of human LAT2 located on SteT 3D model based on AdiC structure (Bartoccioni *et al.,* 2010). C154 and C210 are the only Cys located at the extracellular side. Cys 154 colored in blue corresponds to the Cys responsible for the disulphide bridge with 4F2hc. Cys 210 colored in purple is an endogenous Cys on LAT2 that is located in a desired area for crosslinking with 4F2hc, and kept for crosslinking studies. In red the Cys located in the intermembrane area are depicted. And in yellow, those cysteines located intracellularly.

Regarding the endogenous Cys present in human 4F2hc, it has only two: Cys109, participating in the intersubunit disulphide bridge (located in the "neck" connecting TMD and the ectodomain) and Cys330, a partially hidden residue (located in the A-subdomain of the ectodomain). To avoid doubtful results, residue Cys330 was mutated into serine (Ser) to avoid any kind of interference during crosslinking reaction. Thus, all mutants were generated on a C330S background, and residue Cys109 was maintained to hold the disulfide intersubunit bridge (Figure 22 A), as for LAT2 was kept the Cys154. Then, nine Ser residues spread over the surface of 4F2hc-ED potentially facing LAT2 (Fort *et al.,* 2007) were selected and mutated individually to Cys for crosslinking experiments (see locations in Figure 22 B).

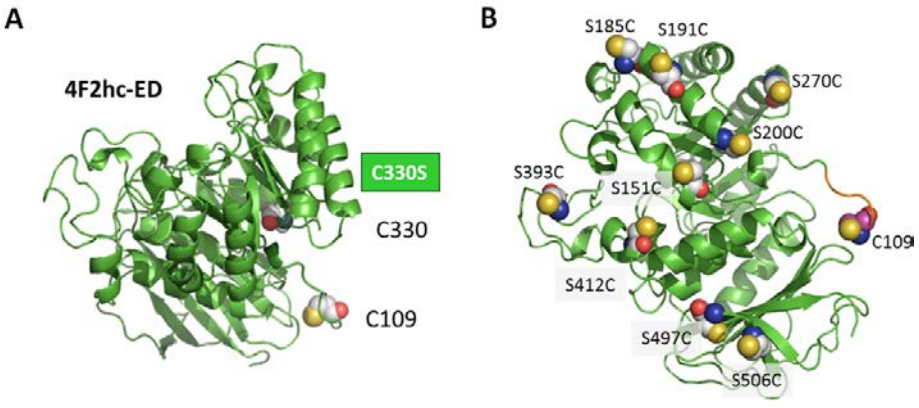


Figure 22. Representation of the different serine to cysteine mutants of 4F2hc-ED generated on a background C330S conserving the Cys 109 responsible for the disulphide bridge. A) 4F2hc-ED endogenous cysteines: C330 and C109 responsible for the disulphide bridge. Cys 330 was mutated to Ser to avoid background signal. C atoms for Cys residues are colored in grey, O atoms in red, N atoms in blue and S atoms in yellow. **B)** 4F2hc-ED most probable area to be interacting with the extracellular loops of LAT2 is represented. The nine different Cys mutants generated on C330S background are colored as Cys in figure (A). Only Cys109, responsible for the disulphide bridge is

differentiated by having C atom colored in purple. The N-terminal end of 4F2hc-ED (C109-S-E-L-P113; one letter code amino acid sequence) is colored in orange to denote its high mobility (Fort *et al.,* 2007). (C atoms, gray; O atoms, red; N atoms, blue; S atoms, yellow).

I.1.2. Validation of new mutants, previous crosslinking assays, by checking heterodimerization and functionality

Previous crosslinking, the new mutants generated on 4F2hc were validated by demonstrating i) its heterodimerization by western blot (from total cell extract) (Figure 23 and 24) and ii) the induction of L-Ala transport in HEK 293T cells. These parameters were first assessed for the control conditions corresponding to transfected: a) empty vector, b) 4F2hc C330S, c) LAT2 WT, and d) co-transfected 4F2hc C330S and LAT2 WT (Figure 23).

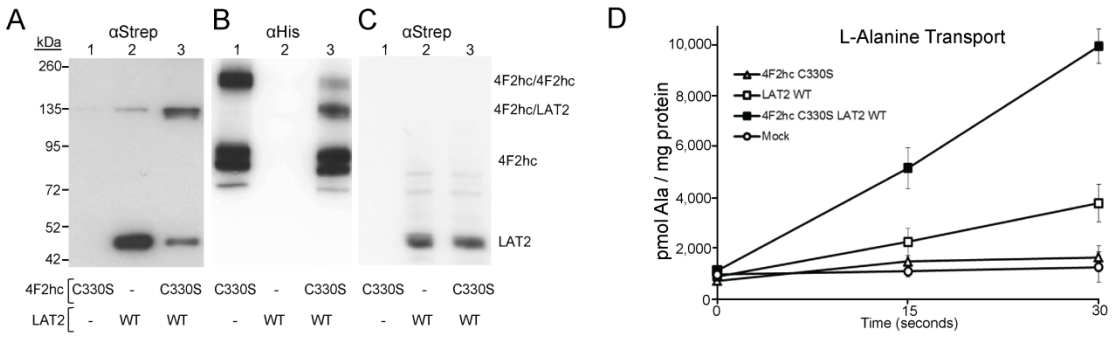


Figure 23. Heterodimer formation and amino acid transport in HEK 293T cells co-transfected with 4F2hc C330S and LAT2 wild-type. Western blot analysis of 1 % DDM-solubilized homogenates of cells transfected with the indicated versions of His-4F2hc and Strep-LAT2 using αStrep **(A, C)** and αHis **(B)** antibodies. Bands corresponding to heterodimers (4F2hc/LAT2), 4F2hc homodimers (4F2hc/4F2hc) and monomers (4F2hc or LAT2) are indicated under non-reducing conditions **(A, B)**. In contrast, under reducing conditions only LAT2 monomers are visible (C). The low expression of 4F2hc/LAT2 heterodimers in cells transfected only with LAT2 WT (wild-type) (panel A, lane 2) is most probably due to heterodimerization with endogenous 4F2hc. (-): Transfection with the corresponding empty vector. 4F2hc monomers were revealed as a doublets of bands **(B)** that correspond to core and mature N-glycosylation forms, as previously described (Fort *et al*., 2007). **D)** Transport of L-Alanine. Cells were analyzed for transport of 10 µM L-[^{3}H] Alanine for the indicated times. Co-transfection with His-4F2hc (C330S) and Strep-LAT2 wild-type (closed squares) resulted in the induction of L-Ala transport over background (mock transfected or His-4F2hc (C330S) transfected cells). Transfection with Strep-LAT2 alone induced L-Ala transport over background, but less than co-transfection with both proteins ($p \leq 0.001$; Student T-test). Induction of L-Ala transport by LAT2 alone could be the result of heterodimerization with endogenous 4F2hc and thus traffic to the plasma membrane or LAT2 monomers reaching the cell surface due to overexpression. Data (mean ± S.E.M) are from 3 experiments run in quadruplicates.

From the nine 4F2hc mutants tested to form heterodimer with LAT2 wild type, only four were selected as good candidates for crosslinking experiments: S151C, S185C, S393C and S412C (Figure 24A and C, lanes 1, 3, 5 and 6) and induced L-Ala transport to values similar than 4F2hc C330S (Figure 24D).

The remaining five from nine mutants were not considered for further studies because they produced proteolyzed proteins (S191C and S200C) (Figure 24B, lanes 2 and 3), or lacked (S270C) (Figure 24B, lane 4) or inefficiently formed intersubunit disulphide bridges (S497C and S506C) (Figures 24A and C, lanes 7 and 8) when co-transfected with Strep-LAT2.

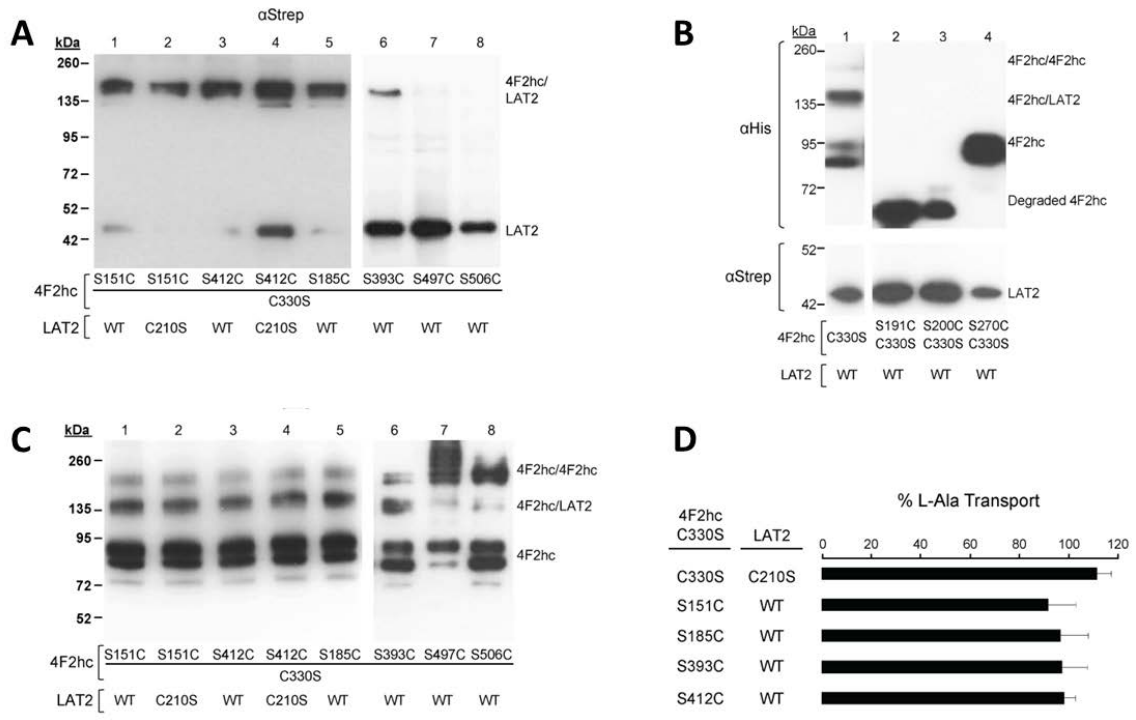


Figure 24. Heterodimer formation and amino acid transport in HEK 293T cells co-transfected with 4F2hc and LAT2 mutants. A and C) Expression of 4F2hc/LAT2 heterodimers by mutants of 4F2hc and LAT2 generated before the docking studies. Western blot analysis of 1 % DDM-solubilized homogenates of HEK 293T cells transfected with the indicated versions of His-4F2hc band Strep-LAT2 using αStrep (A) and αHis (C) antibodies under non-reducing conditions. All mutants in 4F2hc were generated in C330S background. Bands corresponding to heterodimers (4F2hc/LAT2), 4F2hc homodimers (4F2hc/4F2hc) and monomers (4F2hc or LAT2) are indicated under non-reducing conditions (A, C). WT, wild-type version of LAT2. **B)** Lack of formation of 4F2hc/LAT2 heterodimers by some mutants of 4F2hc in HEK 293T cells. Western blot analysis of cells transfected with the indicated versions of His-4F2hc and Strep-LAT2 using αHis under non-reducing conditions (top) or αStrep under reducing conditions (bottom). Bands corresponding to heterodimers (4F2hc/LAT2), 4F2hc homodimers (4F2hc/4F2hc) and monomers (4F2hc or LAT2) are indicated. Degraded 4F2hc, bands corresponding to proteolytic fragments of the protein. WT, wild-type version of LAT2. **D)** Induction of L-Alanine transport by the 4F2hc and LAT2 mutants generated before the docking analysis. HEK 293T cells were transfected with WT or C210S mutant of Strep-LAT2 alone or together with the indicated mutants of His-4F2hc. Cells were analyzed for transport of 10 µM L-[^{3}H] Alanine for 15 seconds. Transport due to 4F2hc and LAT2 co-expression was calculated by subtraction of that elicited by transfecting Strep-LAT2 alone. For 4F2hc C330S and LAT2 WT these transport values were 2463 ± 163 L-Ala uptake (pmol/mg protein) during 15 seconds (mean ± S.E.M. from 4 experiments run in quadruplicates). For the mutants shown, transport is expressed as percentage of that of 4F2hc C330S and LAT2 WT. Data (mean± S.E.M.) are from two experiments run in quadruplicates. The co-expression of all the indicated mutants induced L-Ala transport over the background due to expression of the corresponding LAT2 version alone. Thus, all LAT2 mutants studied were functional.

I.1.3. Crosslinking assays with the validated mutants

Mutants S151C and S412C in 4F2hc were crosslinked with wild-type LAT2, through its endogenous C210, using BMOE and BM(POE)$_2$ crosslinkers (Figure 25, lanes 1 and 7 (BMOE) and lanes 9 and 12 (BM(POE)$_2$)). Interestingly, crosslinking with 4F2hc mutants S151C and S412C was totally abolished when LAT2 Cys210 was mutated to Ser (C210S) (Figure 25A, lanes 2 and 6 (BMOE) and lanes 8 and 11 (BM(POE)$_2$)). It is worth mentioning that C210S (LAT2) mutant is fully functional (it heterodimerized (Figure 24A and C, lanes 2 and 4) and induced L-Ala transport with 4F2hc (Figure 24D)).

In contrast, mutant C330S (Figure 25, lanes 5 and 10), and mutants S185C (Figure 25, lane 3) and S393C (Figure 25, lane 4) incorporated in this background did not show crosslinking with wild-type LAT2. These results demonstrated specific crosslinking between Cys residues at positions 151 or 412, which are located in the loop Aα1-Aβ2 and helix Aα8''', respectively, of subdomain A of 4F2hc-ED, with Cys210, which is located in the extracellular loop TMD5-6 in LAT2.

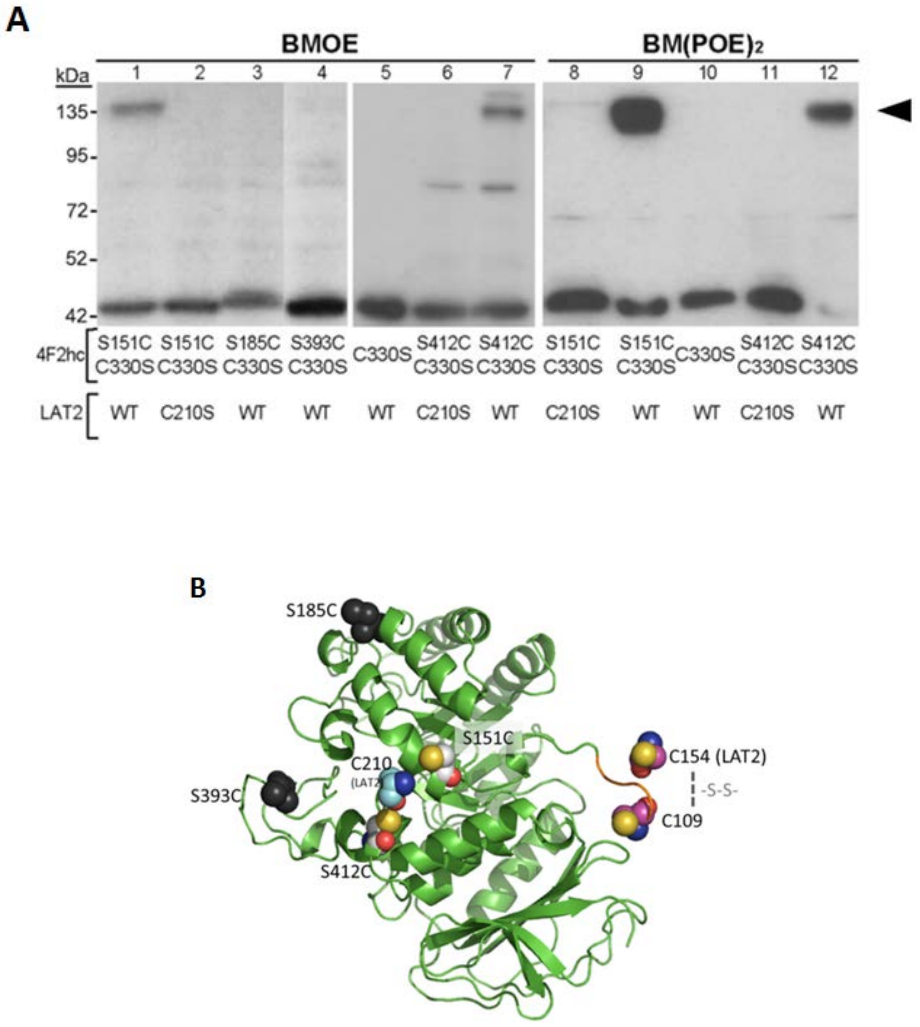


Figure 25. Intersubunit crosslinking between 4F2hc and LAT2. A) Versions of His-4F2hc/Strep-LAT2 heterodimers expressed in 293T HEK cells and purified by His-affinity chromatography were treated with the DTT uncleavable crosslinkers BMOE or $BM(POE)_2$. Crosslinking was detected as DTT-resistant 4F2hc/LAT2 heterodimers (arrowhead) by Western blotting using αStrep antibodies under reducing conditions. The band of LAT2 monomer corresponds to LAT2 forming heterodimers with 4F2hc (i.e., co-purified) but not crosslinked. Single Cys mutants were introduced in 4F2hc (C330S) and LAT2 wild-type (WT). In all cases, crosslinking was abolished when 4F2hc (C330S) or LAT2 (C210S) was used, indicating specificity of the crosslinked sites. **B)** Schematic representation of 4F2hc-ED visualizing the interacting region with LAT2. Crosslinked 4F2hc-ED residues, S412C and S151C are colored as C atoms in gray; O atoms in red; N atoms in blue; S atoms in yellow. Both are crosslinked with the endogenous Cys 210 in LAT2, colored as C atoms in cyan; O atoms in red; N atoms in blue; S atoms in yellow, that is represented at the most probable region between both residues. Non-crosslinked residues, S393C, S185C, are colored in black. And disulphide bridge between C109 in 4F2hc and C154 in LAT2 is illustrated (-S-S-), and Cys residues colored as C atoms in pink; O atoms in red; N atoms in blue; S atoms in yellow.

I.1.4. Human LAT2 model from AdiC and multialignment

Due to the difficulty to work without a 3D structure of human LAT2 and/or the heterodimer, a collaboration was started with the group of Dr. Fernández-Recio and Dr. Pérez (from the Joint Barcelona Supercomputing Center-IRB Program in Computational Biology) who built 20 homology models of human LAT2 and, a docking of 4F2hc-ED with LAT2. To that aim, they used as a template the atomic structure of the bacterial homologue L-arginine/agmatine exchanger AdiC in the open-to-out Arg^+ bound conformation (PDB 3OB6) (Kowalczyk *et al.,* 2011) and an extensive multiple sequence alignment between AdiC and 250 close homologues including bacterial (part of the alignment is showed in the Figure 26), simple eukaryotic and metazoan LATs. They used Modeler 8v1 (Eswar *et al.,* 2008) to build the 20 homology models of human LAT2 (Figure 27A).

```
                                                                    TMD1a
LAT2    ----------MEEGARHRNNT-----EKKHPGGGESDASPEAGSGGGGVALKKEIGLVSA 45
asc1    ----------MAGHTQQPSGR-----GNPRPAPSPSPVPGTVPGASERVALKKEIGLLSA 45
LAT1    MAGAGPKRRALAAPAAEEKEE-----AREKMLAAKSADGSAPAGEGEGVTLQRNITLLNG 55
y+LAT2  ----------MEAREPGRPTPTYHLVPNTSQSQVEEDVSSPPQRSSETMQLKKEISLLNG 50
y+LAT1  ---------MVDSTEYEVASQ-----PEVETSPLGDGASPGP----EQVKLKKEISLLNG 42
xCT     ---------MVRKPVVSTISKGG--YLQGNVNGRLPSLGNKEPPGQEKVQLKRKVTLLRG 49
b0,+AT  ----------MGDTGLRKRRE-----DEKSIQSQEP----------KTTSLQKELGLISG 35
Stet    ----------------------------------------MHTEDNGLKKEIGLLFA 17
AdiC    ----------------------------------------MSSDADAHKVGLIPV 15

        TMD1a       TMD1b  ooooo          TMD2
LAT2    CGIIVGNIIGSGIFVSPKGVLENAGSVGLALIVWIVTGFITVVGALCYAELGVTIPKSGG 105
asc1    CTIIIGNIIGSGIFISPKGVLEHSGSVGLALFVWVLGGGVTALGSLCYAELGVAIPKSGG 105
LAT1    VAIIVGTIIGSGIFVTPTGVLKEAGSPGLALVVWAACGVFSIVGALCYAELGTTISKSGG 115
y+LAT2  VSLVVGNMIGSGIFVSPKGVLVHTASYGMSLIVWAIGGLFSVVGALCYAELGTTITKSGA 110
y+LAT1  VCLIVGNMIGSGIFVSPKGVLIYSASFGLSLVIWAVGGLFSVFGALCYAELGTTIKKSGA 102
xCT     VSIIIGTIIGAGIFISPKGVLQNTGSVGMSLTIWTVCGVLSLFGALSYAELGTTIKKSGG 109
b0,+AT  ISIIVGTIIGSGIFVSPKSVLSNTEAVGPCLIIWAACGVLATLGALCFAELGTMITKSGG 95
Stet    LTLVIGTIIGSGVFMKPGAVLAYSGDSKMALFAWLLGGILTLAGGLTIAEIGTQIPKTGG 77
AdiC    TLMVSGNIMGSGVFLLPANLASTGG--GIAIYGWLVTIIGALGLSMVYAKMSFLDPSPGG 73

                          TMD3                    ooooooooooo TMD4
LAT2    DYSYVKDIFGGLAGFLRLWIAVLVIYPTNQAVIALTFSNYVLQPLFPTCFPPESGLRLLA 165
asc1    DYAYVTEIFGGLAGFLLLWSAVLIMYPTSLAVISMTFSNYVLQPVFPNCIPPTTASRVLS 165
LAT1    DYAYMLEVYGSLPAFLKLWIELLIIRPSSQYIVALVFATYLLKPLFPTCPVPEEAAKLVA 175
y+LAT2  SYAYILEAFGGFIAFIRLWVSLLVVEPTGQAIIAITFANYIIQPSFPSCDPPYLACRLLA 170
y+LAT1  SYAYILEAFGGFLAFIRLWTSLLIIEPTSQAIIAITFANYMVQPLFPSCFAPYAASRLLA 162
xCT     HYTYILEVFGPLPAFVRVWVELLIIRPAATAVISLAFGRYILEPFFIQCEIPELAIKLIT 169
b0,+AT  EYPYLMEAYGPIPAYLFSWASLIVIKPTSFAIICLSFSEYVCAPFYVGCKPPQIVVKCLA 155
Stet    LYTYLEEVYGEFWGFLCGWVQIIIYGPAIIGALGLYFGSLMAN-LFGW---GSGLSKVIG 133
AdiC    SYAYARRCFGPFLGYQTNVLYWLACWIGNIAMVVIGVGYLSY--FFPILKDP-LVLTITC 130

        TMD4          TMD5                    ooooooooooooooooo
LAT2    AICLLLLTWVNCSSVRWATRVQDIFTAGKLLALALIIIMGIVQICKGEYFWLEPKNAFEN 225
asc1    MACLMLLTWVNSSSVRWATRIQDMFTGGKLLALSLIIGVGLLQIFQGHFEELRPSNAFAF 225
LAT1    CLCVLLLTAVNCYSVKAATRVQDAFAAAKLLALALIILLGFVQIGKGDVSNLDPNFSFEG 235
y+LAT2  AACICLLTFVNCAYVKWGTRVQDTFTYAKVVALIAIIVMGLVKLCQGHSEHF--QDAFEG 228
y+LAT1  AACICLLTFINCAYVKWGTLVQDIFTYAKVLALIAVIVAGIVRLGQGASTHF--ENSFEG 220
xCT     AVGITVVMVLNSMSVSWSARIQIFLTFCKLTAILIIIVPGVMQLIKGQTQNF--KDAFSG 227
b0,+AT  AAAILFISTVNSLSVRLGSYVQNIFTAAKLVIVAIIIISGLVLLAQGNTKNF--DNSFEG 213
Stet    IIAVLFLCVINIIGTKYGGFVQTLTTIGKLIPIACIIVFGLWKGDQ--HIFTAVNESISD 191
AdiC    VVVLWIFVLLNIVGPKMITRVQAVATVLALIPIVGIAVFGWFWFRGETYMAAW-NVSGLG 189
```

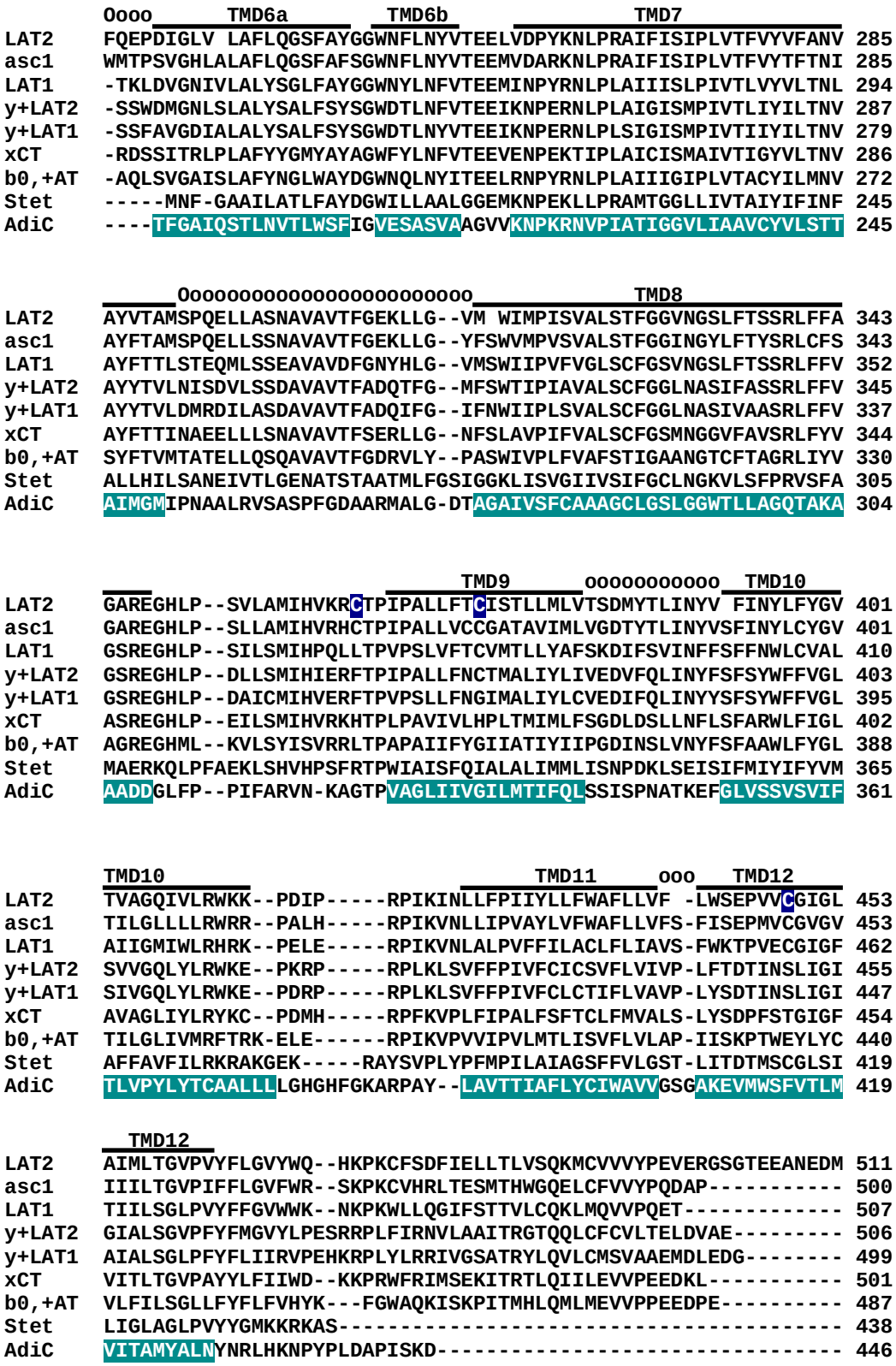


Figure 26. **Human LAT2 homology models.** Part of the extensive multiple amino acid sequence alignment between AdiC from *E. coli* and 250 close homologs (including bacterial, simple eukaryotic and metazoan LATs) (Bartoccioni *et al.,* 2010) used for homology modeling of human LAT2 is shown. The TMDs in human light subunits of HATs and the prokaryotic LAT SteT (Serine/Threonine exchanger) from *Bacillus subtilis* based on the crystal structure of AdiC are indicated over the alignment and the TMD residues in AdiC are highlighted in green. Residues in external loops are shown as open circles. Conserved cysteine residues involved in the intersubunit disulphide bridge are boxed (in

black). Other endogenous Cys residues in human LAT2 are boxed in blue. Small adjustments have been done by hand in loops TMD5-6, TMD7-8, TMD8-9, TMD10-11 and TMD11-12 of SteT and human LATs. The C-terminal ends (intracellular) of LAT2 (24 residues), asc1 (23 residues), y^+LAT2 (9 residues) and y^+LAT1 (12 residues) are truncated.

The loops between TMD5-6 and TMD7-8 are shorter in AdiC. Due to the structural uncertainty in these regions it was decided to select 20 models of LAT2 for the docking calculation. As shown in Figure 27, the 20 superposed models differ mainly in the extracellular loops mentioned before. In addition, the calculated LAT2 model was situated into a predicted membrane and the Cys residues C210 and C154 were confirmed to be located at the extracellular face.

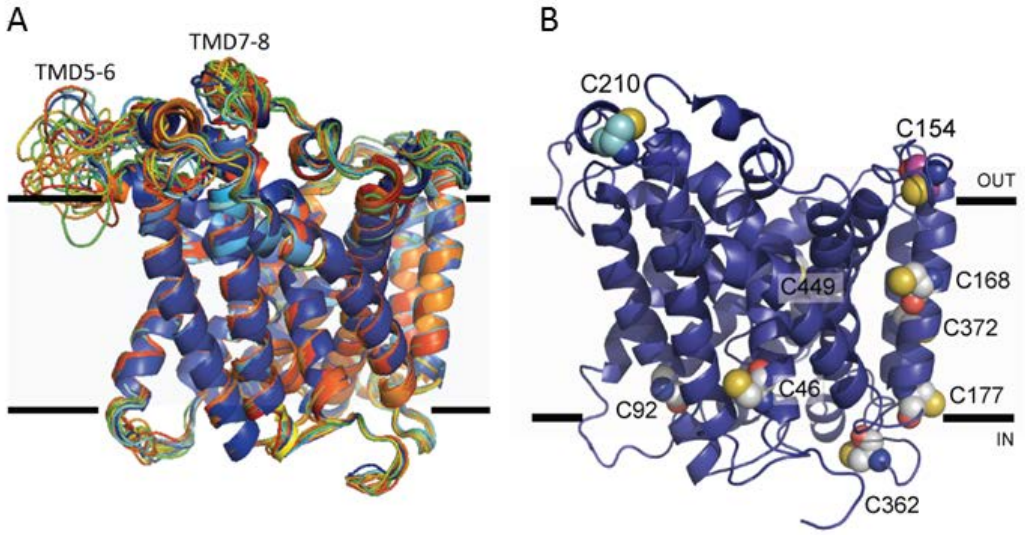


Figure 27. **Human LAT2 homology 3D models. A)** Superposition of the 20 homology models of human LAT2 built from AdiC monomer atomic structure (PDB ID: 3OB6; (Kowalczyk *et al.*, 2011)). Maximal divergence between the selected models occurs in the more flexible loops TMD5-6 and TMD7-8. **B)** Cys residues in the best human LAT2 homology model. Lateral view with atoms of Cys residues shown as spheres. The orientation of LAT2 in the membrane (marked by horizontal lines) was predicted using OPM (Lomize *et al.,* 2006) (A and B). There are only two external Cys residues: C154 (C atoms in purple), involved in the intersubunit disulphide bridge and C210 (C atoms in cyan), which is located almost diametrically opposite.

To characterize the interactions between 4F2hc-ED and LAT2, Dr. Fernández-Recio and Dr. Pérez (from BSC) used each of the 20 models (Figure 27A) and the atomic structure of 4F2hc-ED (PDB ID code 2DH2) (Fort *et al.*, 2007), and allowing only two restrictions:

i) Allowing the formation of the disulphide bridge between residues Cys109 (4F2hc-ED) and Cys154 (LAT2),

ii) Avoiding clashes between 4F2hc-ED residues and the plasma membrane. Out of 3,145 possible docking poses, the model with the lowest-energy is depicted in Figure 28.

Interestingly, the lowest-energy 4F2hc-ED-LAT2 model complies with the above mentioned crosslinking analysis (Figure 28).

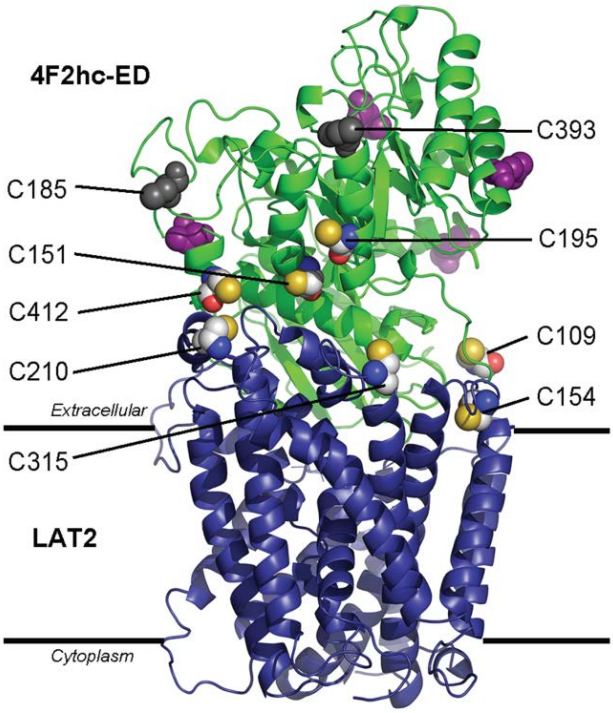


Figure 28. Docking model of the 4F2hc-ED and LAT2 complex. Lowest energy model of the 4F2hc-ED-LAT2 complex. Highlighted Cys residues (endogenous or added through mutation) are shown as sphere models (C atoms, gray; O atoms, red; N atoms, blue; S atoms, yellow). The putative N-glycosylation sites (N264, N280, N323 and N405 indicated in purple) and the not crosslinked residues (C185 and C393 in gray) of 4F2hc are located in the most external face.

Thus, crosslinked residues S151C (4F2hc) – Cys210 (LAT2) and S412C (4F2hc) – Cys210 (LAT2) were located at minimal interatomic distances of 13.4 Å and 11.5 Å, respectively (Table 5). These distances are in good agreement with the results obtained from crosslinking with BMOE and $BM(POE)_2$, which have a range of distances separating both reactive maleimide ends of 6.3 Å to 10.5 Å, and 3.5 Å to 14.3 Å, respectively (Green *et al.,* 2001). In contrast, 4F2hc residues that did not crosslink, S393C and S185C, are located at 18.8 Å and 28.2 Å distance from LAT2 Cys210, respectively, also corroborating the docking model (Table 5).

I.1.5. Docking confirmation and guidance for new crosslinking pair positions

Then, guided by this docking model new Cys mutants in protein secondary structure elements and subdomains not studied previously were considered for crosslinking: S195C in helix 2 of subdomain A, and S487C and G505C in subdomain C of 4F2hc-ED, and A235C, A315C, G392C and S441C in four external loops of LAT2 (Table 5). To ascribe unequivocally the crosslinked residues, 4F2hc mutants were generated in C330S background as previously done and those of LAT2 in C210S background taking advantage of the fact that LAT2 C210S heterodimerized with the new considered mutants of His-4F2hc (Figure 29A and B, lanes 7, 12 and 17) in HEK 293T cells. Moreover, the rest of the new 4F2hc and LAT2 mutants heterodimerized and induced L-Ala transport in the transfected cells (Figure 30).

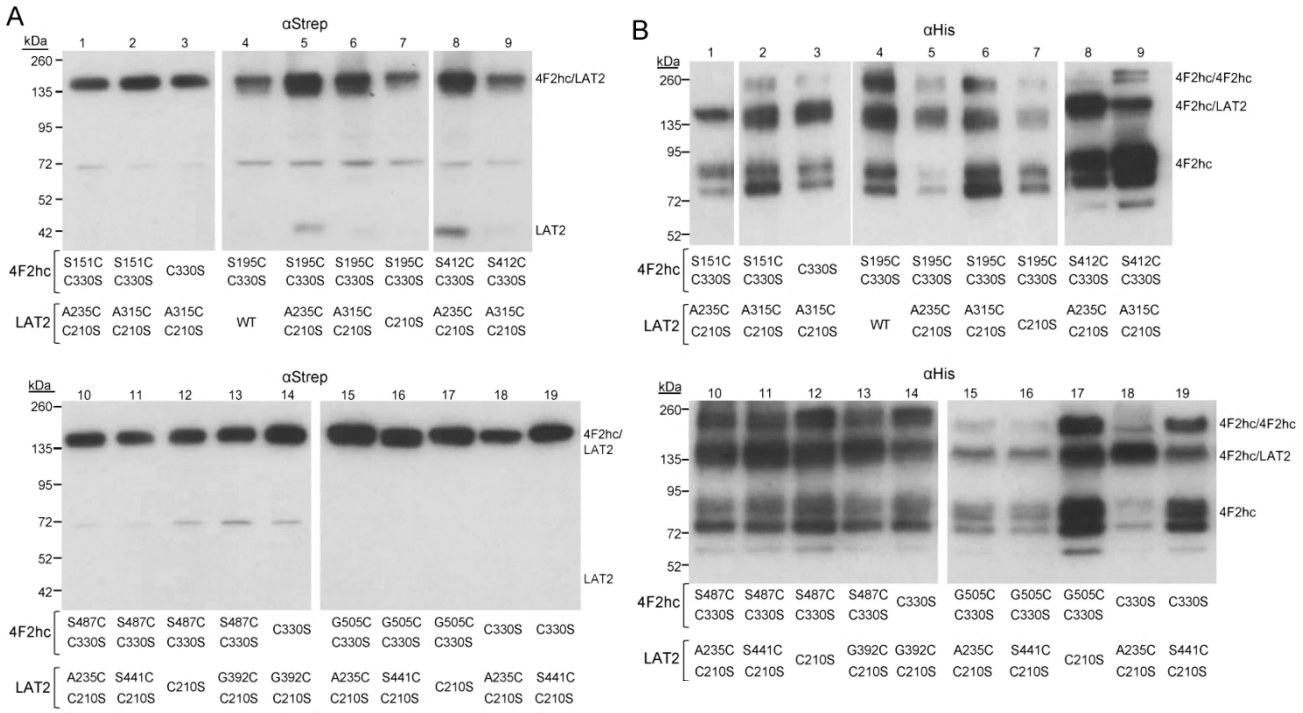


Figure 29. Expression of 4F2hc/LAT2 heterodimers of the new mutants of 4F2hc and LAT2 generated after the docking studies. Western blot analysis of 1 % DDM-solubilized homogenates of HEK 293T cells transfected with the indicated versions of His-4F2hc and Strep-LAT2 using αStrep (**A)** and αHis (**B)** antibodies under non-reducing conditions. Bands corresponding to heterodimers (4F2hc/LAT2), 4F2hc homodimers (4F2hc/4F2hc) and monomers (4F2hc or LAT2) are indicated. WT, wild-type version of LAT2.

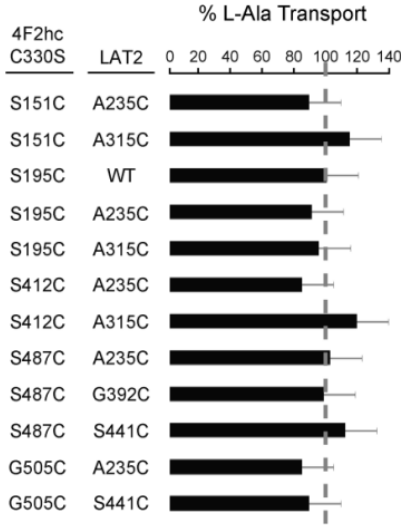


Figure 30. Induction of amino acid transport by mutants of 4F2hc and LAT2 generated after the docking studies. Transport of L-Alanine by the generated 4F2hc and LAT2 mutants. HEK 293T cells were transfected with WT or the indicated mutants of Strep-LAT2 alone or together with the indicated mutants of His-4F2hc. One day after, cells were analyzed for transport of 10μM L-[^{3}H] Alanine for 15 seconds. Transport is expressed as percentage of that elicited by the co-expression of 4F2hc C330S and LAT2 WT and has been calculated as indicated in the legend to Figure 18E. Data (mean± S.E.M.) from 1-2 experiments run in quadruplicates. The co-expression of all the indicated mutants induced L-Ala transport over the background due to expression of the corresponding LAT2 version alone. Thus, all LAT2 mutants studied were functional.

I.1.6. Crosslinking results corresponding to the mutants generated by docking guidance

Six new paired positions in 4F2hc and LAT2 were crosslinked by BMOE and/or $BM(POE)_2$, (Figure 31). The crosslinked heterodimer bands of the western blot results were quantified by densitometry using ImageJ. Interestingly, the new paired positions G505C (4F2hc) with A235C and S441C (LAT2) (Figure 31A, lanes 6 and 7), and the already mentioned S151C and S412C (4F2hc) with Cys210 (LAT2) (Figure 28) crosslinked almost completely (80-95%) (Table 5).

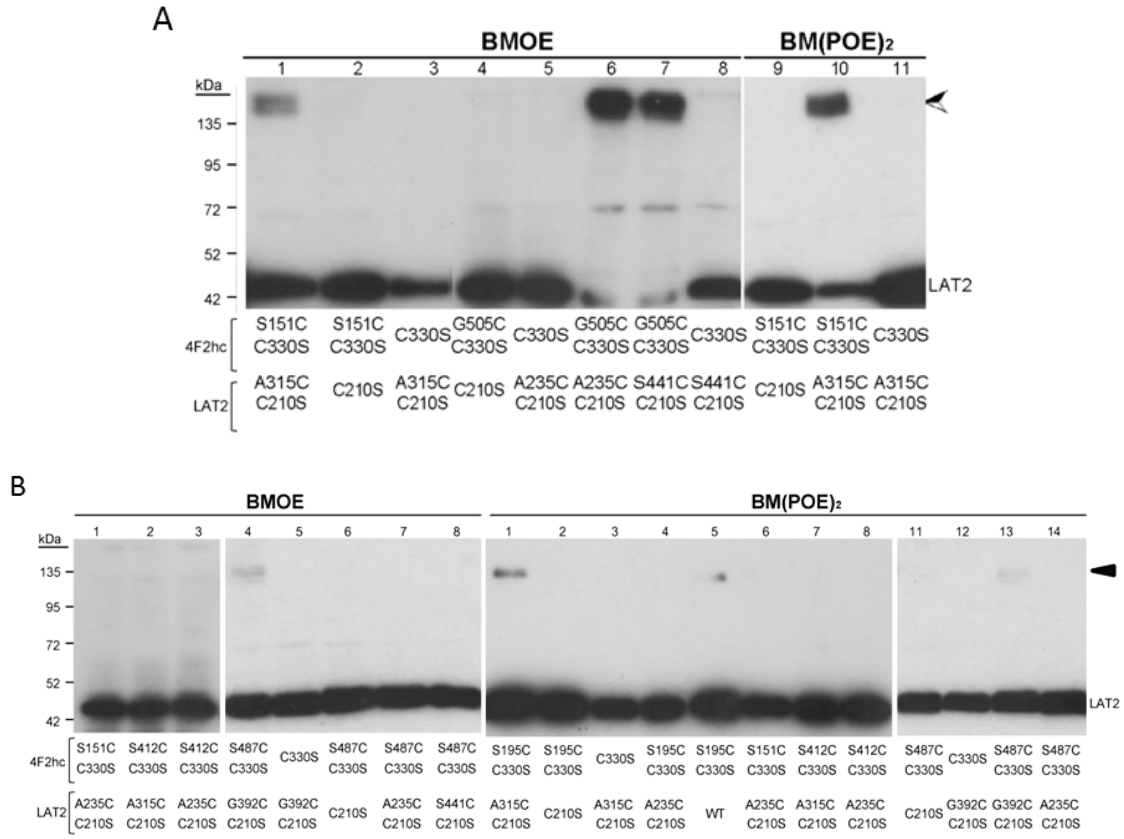


Figure 31. Additional intersubunit crosslinking analysis confirming the lowest energy docking model of 4F2hc-ED and LAT2 complex. A) Strong crosslinking was detected between the following positions (4F2hc-LAT2): S151C-A315C (lane 1), G505C-A235C (lane 6) and G505C-S441C (lane 7) by BMOE, and between position S151C-A315C (lane 10) by $BM(POE)_2$. **B)** Weak but significant crosslinking was detected between the following positions (4F2hc-LAT2): S487C-G392C by BMOE (lane 4), and between positions S195C-A315C (lane 1), S195C-Cys210 (LAT2 WT) (lane 5) and S487C-G392C (lane 13) by $BM(POE)_2$. These crosslinking products were specific of these positions because crosslinking could be abolished when using the single mutants C330S (4F2hc) or C210S (LAT2). Crosslinking by BMOE and by $BM(POE)_2$ was negative for the following positions (4F2hc-LAT2): S151C-A235C, S412C-A235C, S412C-A315C and S487C-A235C. Crosslinking by BMOE was negative for the paired positions S487C-S441C.

Together with the intersubunit disulphide bridge between Cys109 (4F2hc) and Cys154 (LAT2), which links to 4F2hc >95% of the expressed LAT2 mutants (Figure 24C, lanes 1 and 3, and Figure 29A, lanes 15 and 16), these results triangulate the location of 4F2hc-ED over the external face of LAT2 (Figure 32). Indeed, our crosslinking studies showed that four of the six external loops of LAT2 (Cys210 and Ala235 in TMD5-6, Ala315 in TMD7-8, Gly392 in TMD9-10 and Ser441 in TMD11-12) (Figure 27,

asterisks) are in the vicinity of residues in subdomains A and C of the 4F2hc-ED. Moreover, the intersubunit disulphide bridge connects Cys154 in TMD3-4 of LAT2 and Cys109 located five residues apart from the C-terminal end of the 4F2hc TMD (Figure 32 and Table 5).

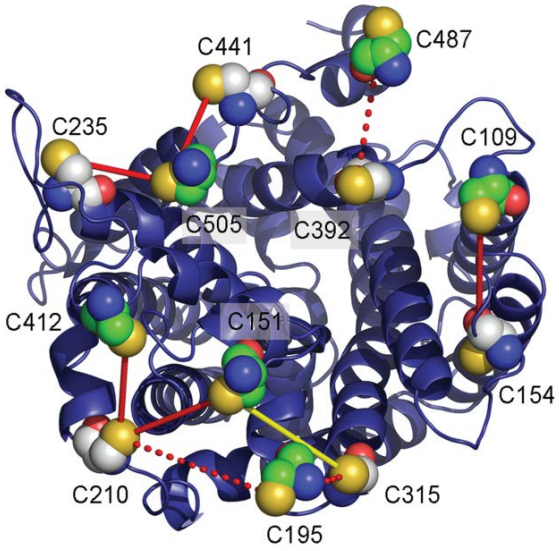


Figure 32. Summary of the crosslinked residues. The 4F2hc-ED-LAT2 model in (Figure 20) was rotated by 90º, i.e., view on the 4F2hc-ED. Of the 4F2hc-ED only the Cys residues (endogenous or added through mutation) are shown as sphere models with C atoms in green. In LAT2 (cartoon in blue) the Cys residues (endogenous or added through mutation) are shown as sphere models with C atoms in gray. Crosslinked residues are connected by a line denoting the percentage of crosslinking (solid red line: 80-95%, solid yellow line: ~60% and red dotted line: 5-10%). Similarly, the intersubunit disulphide bridge between C109 (4F2hc) and C154 (LAT2) is represented by a solid red line because it is present in >95% of the expressed LAT2 (see text for details).

According to the lowest-energy 4F2hc-ED-LAT2 model, all the crosslinked positions in the present work are between 8.1 Å and 17.5 Å apart (Table 5). In contrast, paired positions separated >15 Å or >18 Å were not crosslinked by BMOE or $BM(POE)_2$, respectively (Figures 25A and 31B, and Table 5).

					% crosslinking*		
Position	4F2hc residue	Location†	LAT2 residue	Location†	BMOE (6.3–10.5 Å)‡	$BM(POE)_2$ (3.5–14.3 Å)‡	Distance, Å§
Crosslinked	S151	Aα1-Aβ2	WT (C210)	TMD5–6	50	80	13.4
			A315	TMD7–8	25	60	14.9
	S195	Aα2	WT (C210)	TMD5–6	NA	5	17.5
			A315	TMD7–8	NA	10	15.5
	S412	Aα8′′′	WT (C210)	TMD5–6	35	95	11.5
	S487	Cβ5	G392	TMD9–10	5	10	12.9
	G505	Cβ6	A235	TMD5–6	90	NA	8.4
			S441	TMD11–12	90	NA	8.1
Not crosslinked	S151	Aα1-Aβ2	A235	TMD5–6	No	No	27.0
	S185	Aα2	WT (C210)	TMD5–6	No	No	28.2
	S195		A235	TMD5–6	NA	No	37.9
	S393	Aα6-Aβ7	WT (C210)	TMD5–6	No	No	18.8
	S412	Aα8′′′	A315	TMD7–8	No	No	27.1
			A235	TMD5–6	No	No	19.9
	S487	Cβ5	A235	TMD5–6	No	No	27.2
			S441	TMD11–12	No	NA	14.2

Table 5. Crosslinking analysis summary. [a] The structural element where the residues are located is shown. [b] The range of crosslinking distance of BMOE and $BM(POE)_2$ is indicated in brackets. Percentage of LAT2 crosslinked was calculated as the amount of LAT2 crosslinked from the total LAT2 co-immunoprecipitated with His-4F2hc. NO, residues not crosslinked. na, not analyzed for crosslinking. [c] Minimal interatomic distance between crosslinking residues in the lowest-energy docking model.

I.1.7. Discarding that negative crosslinking Cys were not forming a disulphide bridge

Blocking Cys residues by direct disulphide bridge formation, due to very close proximity, was not the cause of the lack of crosslinking in the indicated paired positions (Figure 33).

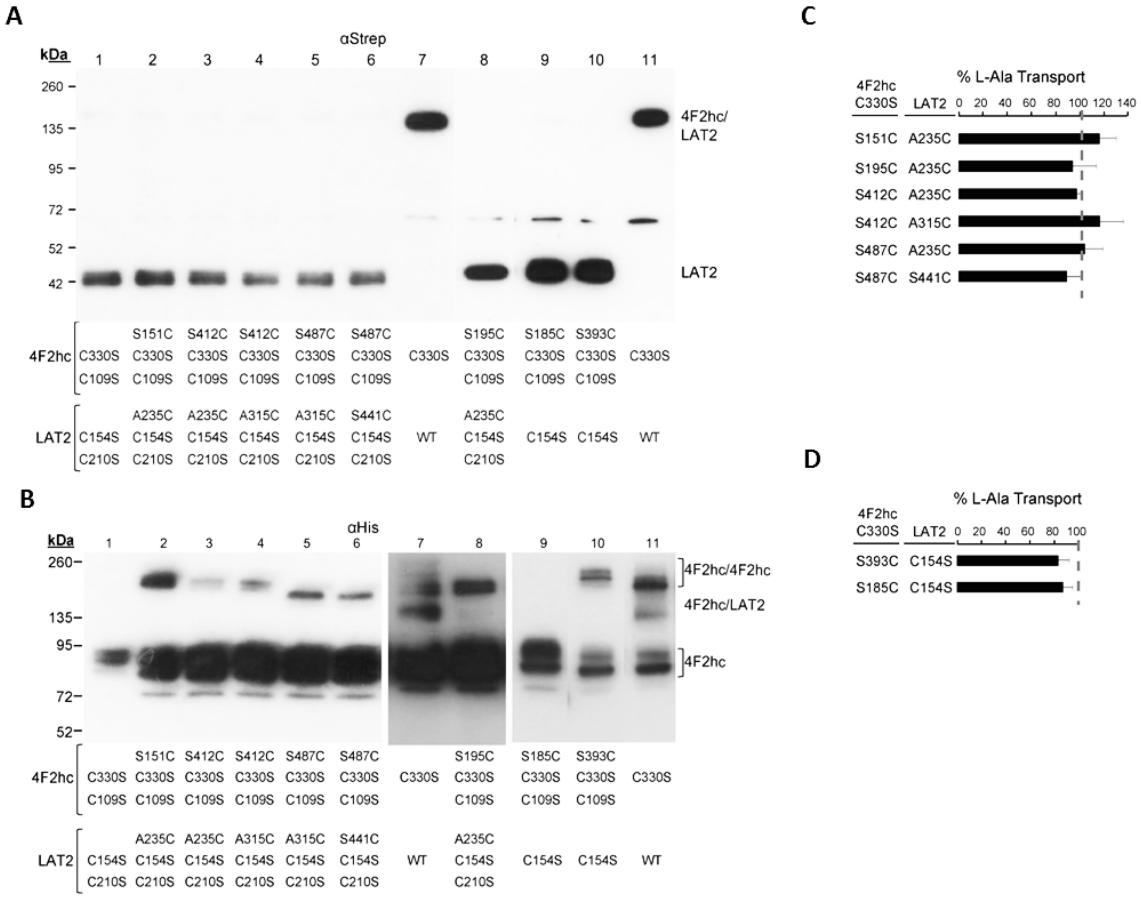


Figure 33. Cysteine pairs negative for crosslinking do not form direct intersubunit disulphide bridges. HEK 293T cells were transfected with the indicated mutants of His-4F2hc and Strep-LAT2. Two days after transfection Western Blot analyses using αStrep (A) or αHis (B) under non-reducing conditions were performed to detect the formation of intersubunit disulphide bridges. As indicated, analyzed mutants were introduced in 4F2hc (C330S/C109S) and LAT2 (C154S/C210S) or LAT2 (C154S) where the endogenous Cys residues involved in the intersubunit disulphide bridge have been eliminated. In this way, 4F2hc/LAT2 heterodimers connected by a possible new disulphide bridge will be revealed under non-reducing conditions. With the exception of the positive control (co-transfection of 4F2hc C330S and LAT2 WT (lanes 7 and 11)) none of the studied mutants showed the formation of 4F2hc/LAT2 heterodimers. Cysteine residues in several positions (S151C, S195C, S393C, S412C and S487C), but not in position S185C, produced disulphide bridge-linked homodimers of 4F2hc (D). Induction of L-Ala transport in transfected HEK 293T cells (C and D). To validate our strategy, one day after transfection, the mutants studied in (A) and (B) were assayed for L-Ala transport during 15 seconds (C and D). Transport is presented as percentage of that induced by co-transfection of 4F2hc (C109S/C330S) and LAT2 (C154S/C210S) (C), and 4F2hc (C109S/C330S) and LAT2 (C154S) (D) as described in the legend to Figure 30. Transport in these control conditions were 3062 ± 167 and 3017 ± 280 pmol L-Ala during 15 seconds (mean ± S.E.M. from 2 experiments run in quadruplicates), respectively. Induction of L-Ala transport over the background elicited by LAT2 alone demonstrates the functional interaction of the mutants of 4F2hc and LAT2 studied.

When analyzing the crosslinking results with the different docking calculated, it is observed that the experimental data is coincident with the docking model of lowest energy. Moreover, the five strongest (≥60% efficiency) crosslinking experiments are in concordance with the three best lowest-energy-docking models of 4F2hc-ED-LAT2 (Figure 34).

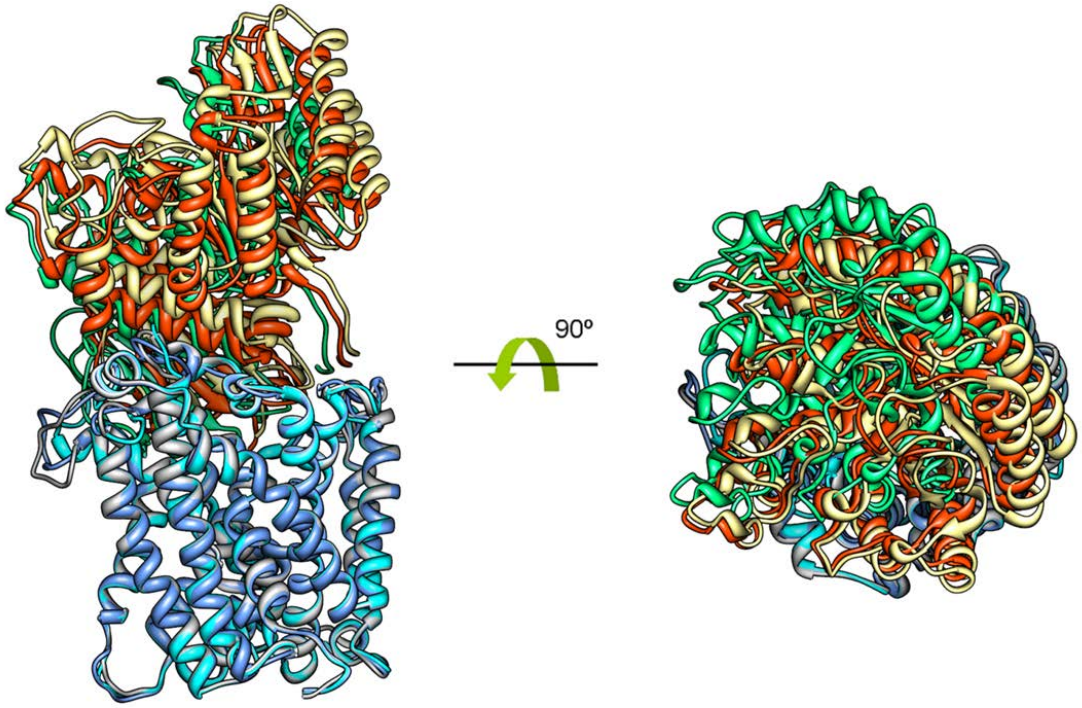


Figure 34. 4F2hc-ED-LAT2 docking models that satisfy the crosslinking experiments. The three 4F2hc-ED-LAT2 docking models that satisfy the crosslinking experiments (see Table 5) are displayed. The lowest-energy docking model (4F2hc-ED in red; LAT2 in gray) is the one that satisfies best the crosslinking experiments. There are two additional docking models (ranked 3 and 15) that satisfy the five strongest (≥ 60 %) crosslinking experiments (considering an additional distance of 6 Å to account for small conformational changes). Docking models ranked 3 and 15 are shown in (4F2hc-ED-LAT2) green-blue and yellow-cyan, respectively.

Interestingly, all these docking models, as well as the TEM-SPA 3D model position the 4F2hc-ED tilted on top of LAT2 (Figure 34 and 19). Even though the atomic structure of the ectodomain of 4F2hc (Fort *et al.*, 2007) can be fitted in the small density of the 3D map, the 'larger domain' of the map, which contains LAT2, the TMD and N-terminal domain of 4F2hc, lipid and detergent and the low resolution of the 3D model of 21 Å preclude a reliable fitting of the 4F2hc-ED-LAT2 docking model into the 3D 4F2hc/LAT2 map.

In summary, TEM, SPA, docking and crosslinking studies strongly indicate that 4F2hc-ED is positioned tilted on LAT2 covering almost completely the extracellular face of the transporter.

I.2. 4F2hc increases the stability of LAT2

According to this model, 4F2hc-ED interacts with all the extracellular loops or ends of TMDs of LAT2. 4F2hc is N-glycosylated in *Pichia* (Costa *et al.,* 2013) and mammalian cells (Fort, *et al.,* 2007), and the four putative sites are located in the model in the most external face of 4F2hc-ED, out of the interface with LAT2 (Figure 35). There is a large interaction surface between the two proteins, the LAT2 buried surface area upon 4F2hc binding being 1,735 $Å^2$ (Figure 35). As much as 60% of the LAT2 interaction surface is formed by hydrophobic residues, and 40% is actually formed by aromatic residues. This is consistent with the major contribution of desolvation to the binding energy as calculated by our scoring function. The most important contacting residues are shown in Figure 35B.

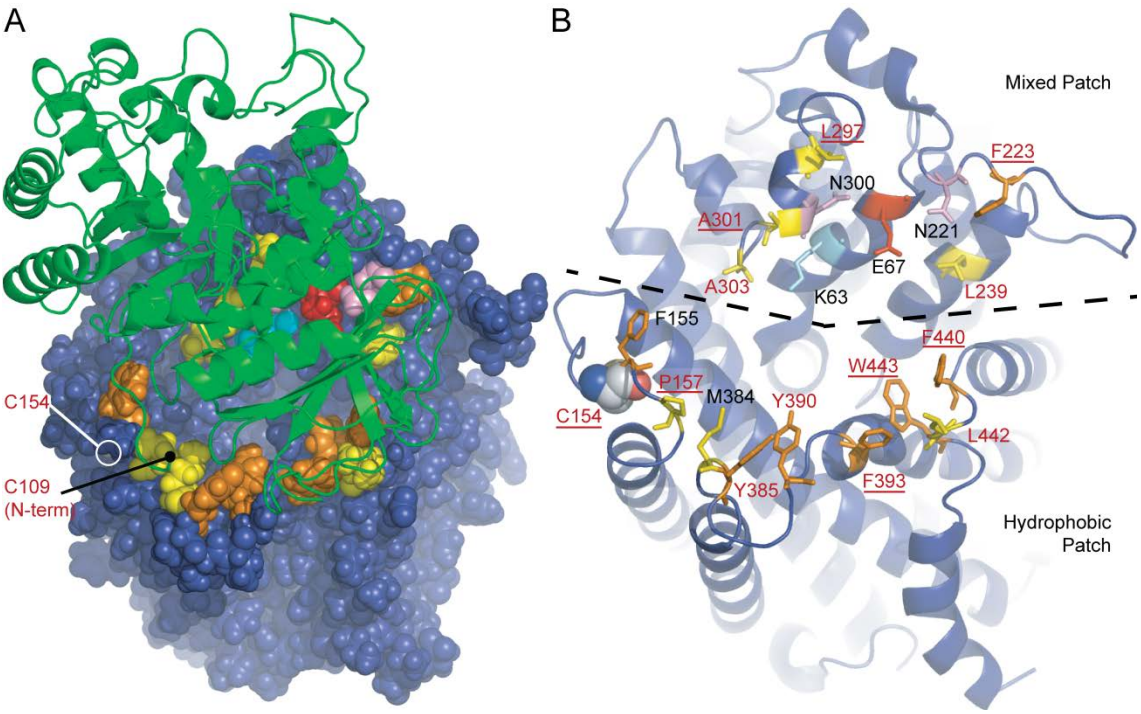


Figure 35. LAT2 residues interacting with 4F2hc-ED. Most important residues of LAT2 for binding to 4F2hc-ED according to docking-derived NIP values **(A and B)**. Residues with NIP >0.4 are highlighted in the representation of LAT2 (blue, atoms as spheres) with the following color code: aromatic residues (orange), aliphatic residues (yellow), Asn (light-pink), Glu (red) and Lys (cyan). The cartoon representation of 4F2hc-ED on top of LAT2 is shown in green. The hydrophobic patch of LAT2 interacts with subdomain C (antiparallel β8) of 4F2hc-ED. The Cys residues involved in the disulphide bridge are indicated: LAT2 C154 is partially hidden and one of its C atoms is indicated by a circle, and 4F2hc C109 corresponds to the N-terminal end of the structure of 4F2hc-ED (indicated by black line). Location of NIP-based interacting residues with cartoon representation of LAT2. **B)** Two defined patches (separated by the broken line) are evident: one hydrophobic and one mixed patch. The residues are represented by sticks with the same color code as in (**A**). C154 is highlighted with its atoms represented in spheres (C, grey; O, red; N, blue and S, yellow). Totally or highly conserved residues are written in red and non-conserved residues in black. Residues not conserved in cyanobacteria, yeast and fungi homologs of LAT2 are underlined. Loop TMD5-6, containing residues N221 and F223, is larger in LAT2 than in AdiC and other bacterial LATs, precluding the generation of a robust homology model. The fact that a significant part of this loop is not at the interface with 4F2hc-ED might be the consequence of the poor modeling of this loop.

The interaction between 4F2hc-ED and the overall external face of LAT2 suggests a possible effect of the heavy subunit on the stability of the light subunit.

To answer this question the behavior of purified 4F2hc/LAT2 and LAT2 was studied. Solubilization with DDM of isolated membranes from yeast cells co-expressing His-4F2hc and Strep-LAT2, and sequential purification by Co^{2+}- and Strep-Tactin affinity chromatographies yielded 4F2hc/LAT2 heterodimers (Costa *et al.,* 2013). In contrast, DDM solubilization of yeast membranes expressing Strep-LAT2 alone and subsequent Strep-Tactin affinity chromatography and size exclusion chromatography yielded only aggregates of LAT2 (results obtained by Dr. Costa) (Figure 36A).

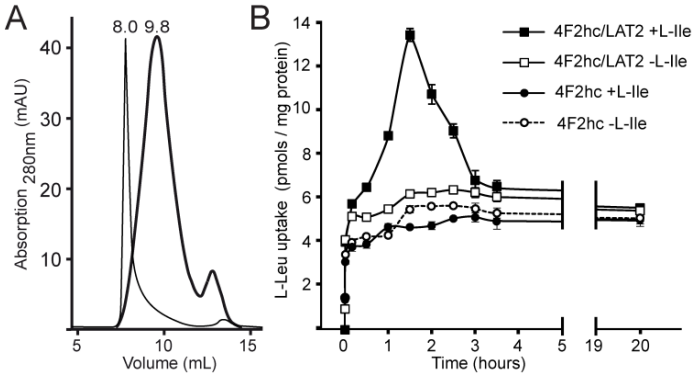


Figure 36. A) Size exclusion chromatography profiles of purified 4F2hc/LAT2 (thick curve) and purified LAT2 (thin curve). 4F2hc/LAT2 elutes at 9.8 mL whereas LAT2 is eluted in the void volume (8 mL). **B) Time-course of L-Leucine uptake in 4F2hc/LAT2 and 4F2hc proteoliposomes.** Transport of 10 µM L-[^{3}H]Leucine into proteoliposomes was measured during 10, 30, 60, 90, 120, 150, 180 and 210 min and 20 h. Proteoliposomes were loaded or not with 4 mM cold L-Isoleucine. Data (mean ± S.E.M.) corresponds to triplicates of a representative experiment. Transport of 10 µM L-[^{3}H]Leucine into 4F2hc/LAT2 proteoliposomes preloaded with 4 mM L-Isoleucine resulted in an overshoot. In contrast, L-Leucine transport into 4F2hc/LAT2 proteoliposomes with no amino acids inside showed passive diffusion similar to that shown by 4F2hc proteoliposomes filled or not with L-Isoleucine. This behavior is characteristic of coupled transporters like the H^+/lactose cotransporter LacY (Newman *et al.,* 1981) or exchangers like the LAT transporter SteT, an L-Serine/L-Threonine exchanger from *Bacillus subtilis* (Reig *et al.,* 2007).

To study if 4F2hc was able to modulate the functionality of LAT2, we tried the reconstitution of purified 4F2hc/LAT2 heterodimers into liposomes but our attempts failed. As an alternative approach to test the functionality of 4F2hc/LAT2 heterodimers and LAT2 monomers, isolated *Pichia* membranes were solubilized in DDM and reconstituted into proteoliposomes. Interestingly, only 4F2hc/LAT2 heterodimers but not LAT2 monomers could be successfully reconstituted as functional proteins into proteoliposomes (Figure 36B) because LAT2 monomers aggregated before inserting into liposomes (results obtained by Dr. Costa) (Rosell *et al.,* 2014). Surprisingly, expression of Strep-LAT2 alone, as well as co-expression of His-4F2hc and Strep-LAT2, resulted in a similar transport of 10 µM L-Leucine over background in *Pichia pastoris* cells (results obtained by Dr. Costa) (Rosell *et al.,* 2014). These results suggest that LAT2 is properly folded and trafficked to the plasma membrane in yeast cells. However, the interactions with 4F2hc would be essential to maintain the correct folding of the light subunit in detergent.

I.2.2. The ectodomain of the heavy subunit 4F2hc increases the stability of LAT2

Next, we asked whether the ectodomain of 4F2hc is able to stabilize LAT2 without the generation of the conserved disulphide bridge between both HAT subunits and without the TMD and N-terminus of 4F2hc. *Pichia* membranes expressing Strep-LAT2 were incubated with purified His-4F2hc-ED, and then LAT2 was solubilized with different concentrations of DDM (Figure 37A). Interestingly, pre-incubation with 4F2hc-ED, but not with same concentration of BSA, increased the solubilization of LAT2 at the three concentrations (1%, 0.5% and 0.25% DDM) of detergent tested. These results suggested that 4F2hc-ED stabilizes LAT2 in solution. To test this possibility, 1% DDM-solubilized LAT2 in the absence and presence of 4F2hc-ED was ultracentrifuged 2 h, 24 h and 72 h after solubilization. Remarkably, the presence of 4F2hc-ED prevented dramatically the precipitation of LAT2 (Figure 37B and C).

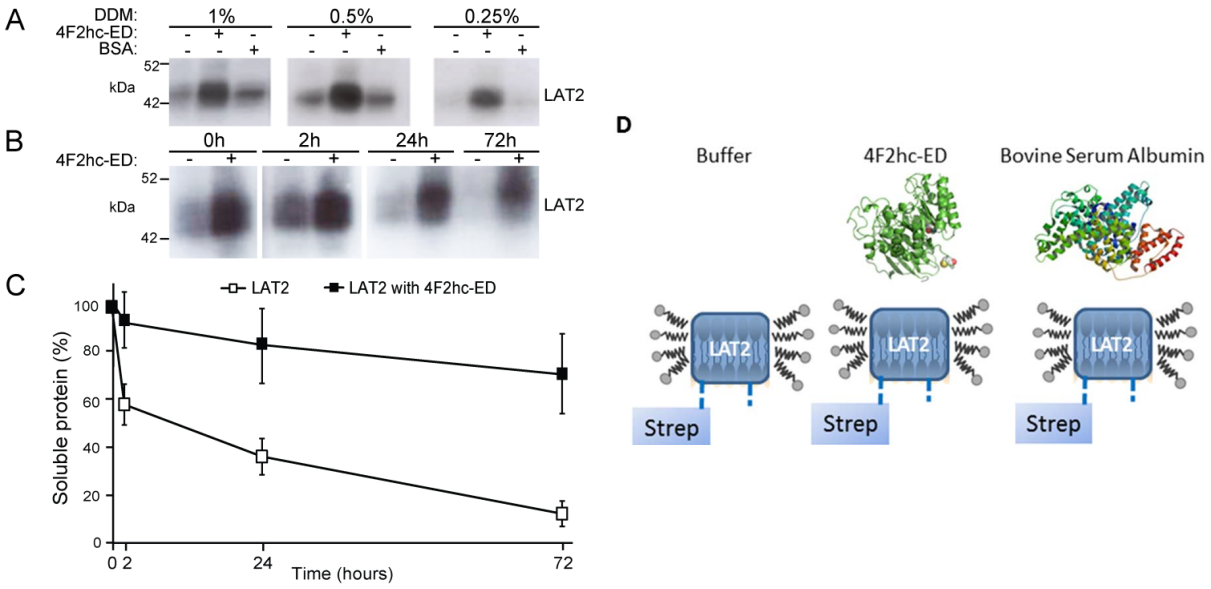


Figure 37. 4F2hc-ED increases LAT2 solubility and stability. A) Western blot analysis of DDM-solubilized LAT2 in the presence of 1 mg/mL of 4F2hc-ED or BSA, or plain buffer as indicated. **B)** Western blot analysis of solubilized LAT2 (1% DDM) in the presence or absence of 1 mg/mL of 4F2hc-ED after different times. **C)** Quantification of the soluble LAT2 after the indicated time. Data (mean ± SEM) corresponds to three independent experiments as the one shown in (B). Soluble LAT2 was considered as the remaining in solution after ultracentrifugation. LAT2 was detected using αStrep antibody. The presence of 4F2hc-ED increased LAT2 stability at all analyzed times. **D)** Scheme representing the *Pichia pastoris* membranes overexpressing Strep-LAT2 and the addition of buffer, purified 4F2hc-ED or BSA to test its interaction were kept overnight at 4ºC and next day solubilized with DDM.

In summary, these results indicate that the ectodomain of 4F2hc suffices to increase the stability of DDM-solubilized LAT2.

I. Discussion

I. Discussion

Three lines of evidence (TEM and SPA, docking and crosslinking analyses) strongly support the idea that 4F2hc-ED covers, almost completely, the extracellular surface of the transporter subunit LAT2. Indeed, we identified specific residues that nearly completely (>80%) crosslinked 4F2hc and LAT2 (Figure 25 and 31) and triangulate with the endogenous intersubunit disulphide bridge (Figure 24 and 32). Importantly, the proposed docking model, as well as two additional and highly similar poses, fulfills the steric restrictions defined by the results from crosslinking (Figure 34).

The docking analysis used the atomic structure of human 4F2hc-ED and homology models of human LAT2 based on the atomic structure of the amino acid transporter AdiC in an outward-facing conformation. Because of the low sequence conservation in the loops of AdiC and LAT2 (Figure 27), the docking models do not intend to represent high-resolution structures of the 4F2hc-ED and LAT2 complex. Nevertheless, docking analysis revealed desolvation as the main contributor to the binding energy of 4F2hc-ED and LAT2. Indeed, the LAT2 residues that are predicted to be more relevant for the interactions are mainly hydrophobic, and they could be grouped clearly in two defined patches (see Normalized Interface Propensity (NIP) residues in Table 6 and Figures 35A and B).

	NIP		Conservation				
Loop	residue	Location	LAT2	Other 4F2hc LS		$b^{0,+}AT$	Cyanobacteria, yeast and fungi
				human	all		
TMD1-2	K63	TMD1	YES	NO[1]	NO[1]	YES	NO
	G64	TMD1	YES	YES	YES	NO	NO
	E67	TMD1 (C-term)	YES	NO	NO	NO	NO
TMD3-4	F155	Loop	YES	NO	NO	NO	NO
	P157	Loop	YES	YES	YES	YES	NO
TMD5-6	N221	Loop	YES[2]	Polar	Polar	Polar	NO
	F223	Loop	YES	YES	YES	YES	NO
	L239	TMD6	YES	L/Y	L/Y	L/Y	V/Y/I
TMD7-8	L297	Loop	YES	YES	L/I/V[3]	YES	NO
	N300	Loop	YES	Polar	Polar	Polar	Polar
	A301	Loop	YES	YES	YES	YES	NO
	A303	Loop	YES	YES	YES	YES	YES
TMD9-10	M384	Loop	YES[4]	NO	NO	NO	NO
	Y385	Loop	YES	Y/F[5]	Y/F[5]	NO	Y/F
	Y390	Loop	YES	Y/F	Y/F	Y/F	Y/L/F
	F393	TMD11 (N-term)	YES	YES	YES	YES	NO
TMD11-12	F440	Loop	YES	F/V/L	F/V/L	NO	NO
	L442	TMD12	YES	L/F	L/F/Y	L/Y/I	L/V/F/Y
	W443	TMD12	W/Y	W/Y	W/Y	I	NO

Table 6. Conservation of the NIP residues placed in the human 4F2hc/LAT2 interface. The amino acid sequence of 72 proteins belonging to the LAT family (SLC7) and grouped in 8 categories (LAT2, LAT1, y^+LAT1, y^+LAT2, asc1, xCT, $b^{0,+}AT$ (8 vertebrate proteins per category) and cyanobacterial, yeast and fungal LATs (16 proteins per this category)) were aligned by the ClustalW2 program. The conservation of the NIP residues was analyzed in vertebrate LAT2, in other vertebrate light subunits associated to 4F2hc (L.S light subunit), in vertebrate $b^{0,+}AT$,in cyanobacterial LATs and in fungal LATs. The protein sequences were retrieved from the UniProt ([a]) and GenBank databases ([b]): **LAT2:** *Homo sapiens* (Q9UHI5)[a], *Mus musculus* (Q9QXW9)[a], *Ornithorhynchus anatinus*

(XM_003429011)[b], *Orcinus orca* (XM_004283215)[b], *Anolis carolinensis* (H9GCK9)[a], *Xenopus tropicallis* (Q6P891)[a], *Carassius carassius* (B8YMU3) [a], *Danio rerio* (NM_001271897.1)[b]. **LAT1:** *Homo sapiens* (Q01650)[a], *Mus musculus* (Q9Z127)[a], *Ornithorhynchus anatinus* (XM_001508884)[b], *Orcinus orca* (XM_004280070)[b], *Gallus gallus* (XM_414194)[b], *Xenopus tropicallis* (F6QFZ9)[a], *Takifugu rubripes* (H2SUN0)[a], *Danio rerio* (NM_001128358)[b]. **y^+LAT1:** *Homo sapiens* (NP_001119577.1)[b], *Rattus norvergicus* (AAF07216.1)[b], *Ornithorhynchus anatinus* (XP_003430084.1)[b], *Orcinus orca* (XP_004283244.1)[b], *Anolis carolinensis* (H9G4U3)[a], *Xenopus tropicallis* (F6S1W4)[a], Latimeria chalumnae (H3A009)[a], *Danio rerio* (NP_001032648.1)[b]. **y^+LAT2:** *Homo sapiens* (Q92536)[a], *Mus musculus* (Q8BGK6)[a], *Orcinus orca* (XM_004273192)[b], *Gallus gallus* (F1ND99)[a], *Anolis carolinensis* (XM_003222960)[b], *Xenopus tropicallis* (Q28I80)[a], *Pelodiscus sinensis* (K7G1V4)[a], *Danio rerio* (NM_001020474)[b]. **asc1:** *Homo sapiens* (Q9NS82)[a], *Mus musculus* (P63115)[a], *Orcinus orca* (XM_004284140)[b], *Gallus gallus* (XM_414136)[b], *Pelodiscus sinensis* (K7GB90)[a], *Xenopus tropicallis* (XM_002935793)[b], *Xiphophorus maculatus* (M4AC17)[a], *Danio rerio* (XM_690716)[b]. **xCT:** *Homo sapiens* (Q9UPY5)[a], *Mus musculus* (Q9WTR6)[a], *Orcinus orca* (XM_004265135)[b], *Gallus gallus* (E1C734)[a], *Anolis carolinensis* (XM_003221678)[b], *Pelodiscus sinensis* (K7FII9)[a], *Xenopus tropicallis* (XM_002932048)[b], *Danio rerio* (XM_002664447)[b]. **b^{0+}AT:** *Homo sapiens* (P82251)[a], *Rattus norvergicus* (NM_053929)[b], *Orcinus orca* (XM_004284134)[b], *Gallus gallus* (E1BT87)[a], *Anolis carolinensis* (H9GCM6)[a], *Xenopus laevis* (A9JS28)[a], *Danio rerio* (F1QGJ2)[a], Salmo salar (NM_001173819)[b]. **Cyanobacterial LATs**: *Chroococcidiopsis sp.* (WP_015152622.1)[b], *Cyanothece sp.* (WP_012630008)[b], *Fischerella sp.* (WP_009454021.1)[b], *Gloeobacter violaceus* (NP_924598)[b], *Oscillatoria sp.*(WP_007357767)[b], *Pleurocapsa minor* (WP_015143465)[b], *Microcoleus sp.* (WP_015184929.1)[b], *Rivularia sp.* (WP_015120455.1)[b]. **Yeast and Fungal LATs**: *Saccharomyces cerevisiae* (NM_001181184)[b], *Arthroderma gypseum* (XM_003171846)[b], *Schyzophyllum commune* (XM_003036114)[b], *Yarrowia lipolytica* (XM_505103)[b], *Aspergillus niger* (XM_001396535)[b], *Coprinopsis cinerea* (XM_001833155)[b], *Nectria haematococca* (XM_003050065)[b], *Ajellomyces dermatitidis* (XM_002620161)[b]. [1] K not present only in LAT1; [2] H in *Anolis carolinensis* LAT2; [3] M in *Xenopus tropicallis* and H in *Mus musculus* y^+LAT2; [4] I in *Anolis carolinensis* LAT2; [5] D in human xCT and H in *Xiphophorus maculatus* and *Danio rerio* asc1; [6] I in human asc1.

One of the patches, which is totally hydrophobic, is composed of nine residues distributed along three external loops or TMD ends of LAT2 (TMD3-4, which harbours Cys154 residue involved in the disulphide bridge, TMD9-10 and TMD11-12). The other is a mixed patch composed of four polar and five hydrophobic residues located in loops TMD1-2, TMD5-6 and TMD7-8.

The interaction between 4F2hc-ED and LAT2 brings the question of whether the heavy subunit might have an impact on LAT2 transport cycle. The structural paradigms of HAT light subunits, the bacterial APC transporters, present the LeuT-fold (Shi *et al.,* 2013). Secondary transporters, upon binding of an external substrate, transit to inward-facing conformations to release the substrate inside the cell. Besides transporter-specific changes, atomic structures of LeuT-fold transporters showed a commonality of features in this transition (Shimamura *et al.,* 2010 and Krishnamurthy *et al.*, 2012): the "hash" domain (TMDs 3, 4, 8 and 9) and the extracellular ends of the "bundle" domain (TMDs 1, 2, 6 and 7) get closer and the loop TMD7-8 occludes the external vestibule of the transporter. Our results indicate that these conformational changes are compatible with a direct contact of 4F2hc-ED with the external loops or ends of TMDs of LAT2. The highly conserved hydrophobic patch glues to 4F2hc-ED the "hash" domain, with the intersubunit disulphide bridge, and TMDs 11 and 12, the most static domain of LAT2 according to the structural homologue AdiC (Kowalczyk *et al.,* 2011).

In contrast, the mixed patch is less conserved and involves the "bundle" domain and the occluding loop (TMD7-8). This architecture suggests that 4F2hc/LAT2, and probably other HATs, have evolved to bind firmly the ectodomain of the heavy subunit to the "hash" domain. In this scenario,

energetically similar interactions of 4F2hc-ED with the "bundle" domain and loop TMD7-8 would be broken and replaced with the different conformations that the light subunit undergoes during the transport cycle.

According to amino acid transporters with LeuT-fold, the substrate binding site in LAT2 is expected to be located at the bottom of the extracellular vestibule (Fotiadis *et al.*, 2013 and Bartoccioni *et al.*, 2010), and therefore far away from the interacting 4F2hc-ED residues (Figures 38A and B). Moreover, our docking analysis located 4F2hc-ED tilted with subdomain C interacting with LAT2 TMD11-12 and leaving open a window in the opposite location flanked by LAT2 TMD7-8 for the entrance of the substrate to the binding site (Figure 38A).

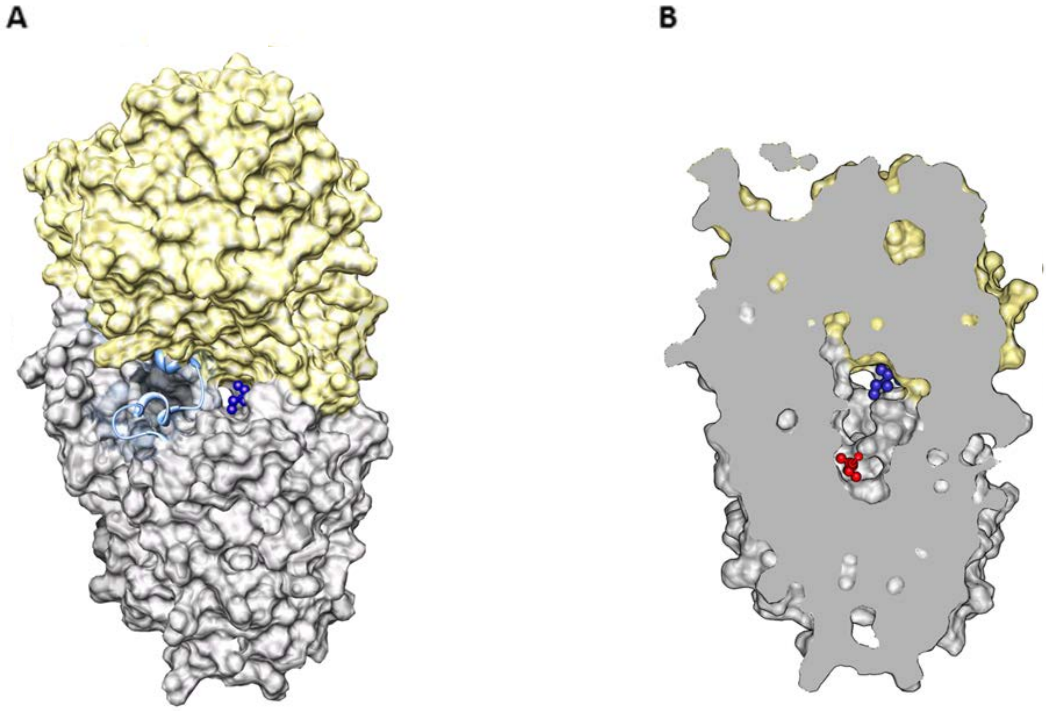


Figure 38. The LAT2 external substrate vestibule is accessible in the lowest-energy 4F2hc-ED-LAT2 model. To visualize the accessibility of the LAT2 substrate vestibule from the exterior, we have docked L-Isoleucine substrate with AutoDock and show here the most populated docking poses (in blue and red). **A)** We can see that the entrance of the substrate vestibule, where one of the most populated L-Isoleucine positions is located, is accessible from the exterior (**A**). TMD7-8 loop (cartoon in cyan) flanks the entrance to this vestibule. The interior of the substrate vestibule is visualized showing a second putative binding site of L-isoleucine (red) as determined by docking (**B**).

In spite of the different structures of the LAT2 monomer and the 4F2hc/LAT2 heterodimer, the measured K_M values for L-Leucine transport in *Pichia* cells were comparable (178±15 µM and 120±7 µM, respectively) suggesting that 4F2hc would have no significant impact on the substrate affinity of LAT2 (Figure 39). Thus, the interaction between 4F2hc and the extracellular surface of LAT2 increases the stability of the detergent-solubilized light subunit but did not alter dramatically the substrate apparent affinity of this transporter *in vivo*.

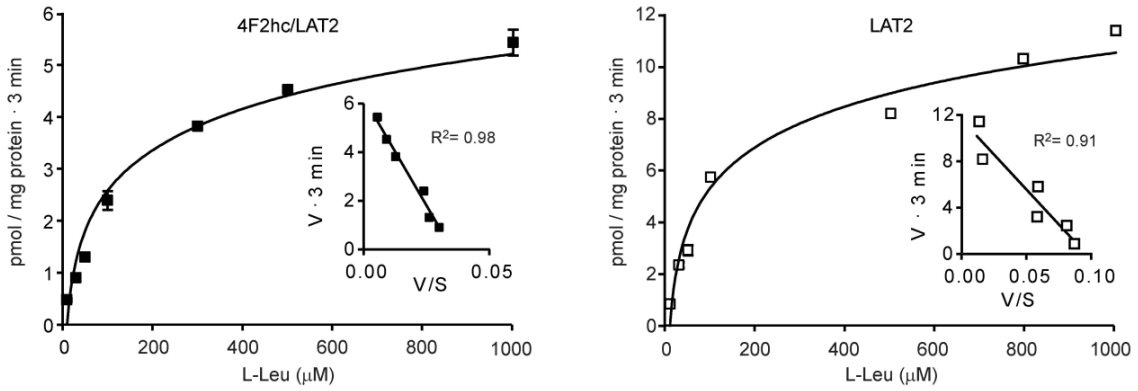


Figure 39. Kinetic analysis of L-Leucine transport in *Pichia pastoris* cells. Kinetic analysis of L-Leucine transport in *Pichia pastoris* cells transformed with His-4F2hc and Strep-LAT2 (left) or Strep-LAT2 alone (right) yielded comparable K_Ms of 178±15 μM and 120±7 μM, respectively. Transport was measured in linear conditions during 3 min. Data points were corrected for basal activity in untransformed yeast cells. The concentrations used were 10, 30, 50, 100, 300, 500 and 1000 μM. Eadie-Hofstee plots are shown as insets. Data (mean ± S.E.M.) corresponds to triplicates from representative experiments. When not visible, error bars are smaller than symbol. R^2, is the correlation coefficient of the linear regression.

The hydrophobic character of the interaction between 4F2hc-ED and LAT2 might be at the basis of the stabilization of this light subunit by 4F2hc. We show that 4F2hc allows functional reconstitution into proteoliposomes and purification of the transporter after solubilization with DDM. Moreover, we show here that 4F2hc-ED suffices to increase the stability of DDM-solubilized LAT2. The binding of 4F2hc-ED to the extracellular surface of LAT2 would bury a large hydrophobic patch that would be otherwise exposed to solvent, thus reducing aggregation and increasing stability. In addition, the compact and thermally stable structure (T_m~59º C) (Turnay *et al.*, 2011) of 4F2hc-ED may have additional effects in the improvement of LAT2 stability.

HATs appeared in metazoans with an ancient heavy subunit that evolved to differentiated 4F2hc and rBAT in vertebrates (Fotiadis *et al.* 2013). The fact that LAT2 monomers elicited amino acid transport in *Pichia* cells, which do not express 4F2hc or related proteins, indicates that 4F2hc is not required for proper folding of LAT2, as it has been demonstrated for the light subunit $b^{0,+}$AT in the absence of rBAT (Reig *et al.*, 2002). In contrast, 4F2hc and rBAT are necessary for heterodimerization and functional expression at the cell surface of its light subunits when expressed in *Xenopus oocytes* and mammalian cells (Fotiadis *et al.* 2013). Specifically, 4F2hc-ED is necessary for these functions (Fenczik *et al.*, 2001 and Bröer *et al.*, 2001). Desolvation of hydrophobic residues might be extended to the recognition of other light subunits by 4F2hc. The hydrophobic NIP residues in our docking analysis are highly conserved in 4F2hc-associated light subunits, but poorly conserved in cyanobacterial, yeast and fungal LATs, which are the closest homologs to vertebrate LATs (up to 37% amino acid sequence identity) not interacting with a heavy subunit (Table 6). Hydrophobic NIP residues are largely conserved in $b^{0,+}$AT also (Table 6). Thus, desolvation of hydrophobic residues might also contribute to the binding of the ectodomain of rBAT and its light subunit. Indeed, co-transfection of 4F2hc and $b^{0,+}$AT (i.e., the

rBAT-associated light subunit) in mammalian cells yielded functional transporter at the cell surface (Rajan *et al.*, 2000). Interaction through a hydrophobic patch looks like a clever solution to glue 4F2hc to six different light subunits (LAT1, LAT2, y^+LAT1, y^+LAT2, asc1 and xCT) that share as little as ~43% amino acid sequence identity (Fotiadis *et al.* 2013). Our work revealed that 4F2hc-ED covers the external face of LAT2 increasing transporter stability. The idea of a common hydrophobic patch for subunit recognition by 4F2hc is susceptible to future experimental examination.

An experiment where the NIP residues corresponding to the hydrophobic patch would be mutated one by one and in groups with the aim to change the charge properties of the LAT2 surface and hence lose the interaction with 4F2hc-ectodomain in a non-disulfide bridge background would be the aim for the next steps to confirm the existence of this hydrophobic patch. To demonstrate it, the same strategy to check heterodimer formation and functionality would be followed. In addition, heterodimer presence in the plasma membrane would be checked by biotin labeling.

I. Materials & Methods

I. Materials & Methods

I.1. Cloning of human 4F2hc and human LAT2 for mammalian cell expression

Human His-4F2hc (N-terminally tagged) cloned in pcDNA4His-MaxC (Fort *et al.*, 2007) was used. The gene encoding human LAT2 was cloned into the *Eco*RI and *Xba*I sites of pcDNA3.1+ (Invitrogen) using the same fusion protein design (N-terminal Strep II tag-Factor X protease site) as it was done for the overexpressed human LAT2 in *Pichia pastoris* (Costa *et al.,* 2013), since the objective was to perform complementing studies in mammalian cells to those obtained in *P.pastoris*. Human Strep-TagII-LAT2 (N-terminally tagged) cloned in pPICZ was subcloned adapting to the mammalian Kozak.

A

Human LAT2 primers to clone into pcDNA3.1+	
oligo	**5'-3' sequence**
LAT2 (EcoRI) F	CCG**GAATTC**CGGATGGCCTGGTCACATCCTCAATTTGAAAAA*ATCGAAGGTCGT*GAAGAAGGAGCCA
	GGCACCGAAAC (in bold: *Eco*RI site; underlined: Strep II tag; in italics: Factor X protease target)
LAT2 (XbaI) R	CTAG**TCTAGA**CTAG TCAGGGCTGGGGCTGCCC (in bold: *Xba*I site; underlined: the stop codon)

B

Reaction mix for insert amplification	(µL)
Template plasmid (20 ng)	X
Reaction buffer with $MgCl_2$ (10x)	5
dNTP mix	1
Forward primer (10 µM)	2
Reverse primer (10 µM)	2
Expand High Fidelity enzyme*	1
Milli-Q water, adjust until final volume	50

C

PCR Insert amplification	Ta	Time	Cycles
Initial denaturation	95°C	5 min	1x
Denaturation	95 °C	1 min	30x
Annealing	54 °C	1 min	
Elongation	72 °C	2 min	
Final Elongation	72 °C	7 min	1x
Cooling	4 °C	∞	

Table 7. Oligonucleotides and PCR conditions to clone human LAT2 into pcDNA3.1+ (Invitrogen). A) Oligonucleotides to clone human LAT2 into pcDNA3.1+ are indicated on top (F: forward primer; R: reverse primer). *Eco*RI and *Xba*I sequences sites were designed according to NEB recommendations to achieve higher digestion rate: https://www.neb.com/~/media/NebUs/Files/Chart%20image/cleavage_olignucleotides_old.pdf. StrepIITag, factor X protease were exact as for pPICZ (Costa *et al*., 2013) and Kozack sequence was adapted for human LAT2. **B) C)** Reaction mix and PCR for insert amplification are detailed in the tables below.

Human LAT2 cDNA was amplified by PCR (Polymerase Chain Reaction) using the proofreading enzyme Expand High Fidelity Taq polymerase (Roche). The sequence of the corresponding oligonucleotides and the PCR conditions are showed in Table 6. To analyze and separate the PCR product by size, DNA samples with 6x loading buffer were loaded to agarose gels of 1% (SeaKem LE, Lonza) prepared with TAE buffer (40 mM Tris base, 20 mM acetic acid, 1 mM EDTA) and run at 80-100 V for ~1 hour. SyberSafe (Invitrogen) at a 1/10,000 dilution was used as intercalating DNA agent to visualize the DNA. 1 kb ladder from Fermentas were used to reference our samples. Gels were visualized with the GeneGenius Gel Imaging system (Syngene) or with a UV-transilluminator. The PCR product was purified from agarose gel using the Gel Band Purification kit (GE Healthcare).

pcDNA3.1+ plasmid and the amplified and purified human Strep-TagII-LAT2 cDNA were digested with *Eco*RI and *Xba*I restriction enzymes. FastDigest from Thermoscientific or New England Biolabs enzymes were used. The first ones digest in 5-15 minutes at 37°C, whereas the second ones need longer incubation times, and depending on the enzyme they were kept for 4 hours – overnight at 37°C. The amounts of DNA, enzymes and buffers used were strictly the ones recommended by the manufacturers.

To prevent self-annealing of linearized vector prior to the insertion of DNA fragments during ligation a phosphatase enzyme was used which dephosphorylate the 5' and 3' ends of DNA. To that aim, 1 µl of calf intestinal alkaline phosphatase (CIP) (Biolabs) was added to the pcDNA3.1+ digestion mixture for 1 hour at 37°C. Thereafter, the linearized and dephosphorylated pcDNA3.1+ plasmid was separated from the remaining circular DNA by gel electrophoresis and purified using the Gel Band Purification kit (GE Healthcare).

Ligation between the linearized vector and insert was performed using the T4 DNA ligase kit from Biolabs following their recommendations and incubating the sample overnight at 16^{o}C. The molar ratios used for each DNA were calculated using the NEBioCalculator (http://nebiocalculator.neb.com). The ligated product was then transformed using *E. coli* supercompetent cells by heat shock method.

XL1 Blue *E.coli* cells were thawed in ice, 80 µl of competent cells were mixed with the ligation mixture and kept on ice during 20 minutes. Heat shock was done at 42^{o}C for 90 seconds, and kept on ice for 2 minutes. Then, 900 µl of LB (1% tryptone, 0.5% yeast extract, 0.5% NaCl, pH 7.5) broth media was added and incubated for 1 hour at 37^{o}C shaking at 900 rpm. Cell culture was concentrated before plating to increase the probability of obtaining positive colonies: the 900 µl of cell culture were centrifuged for 3 minutes at 5,000 rpm, discarded 800 µl, resuspended the pellet in the remaining 100 µl cell culture and plated into LB plates containing ampicillin (50 µg/mL)**.** Plates were incubated overnight at 37^{o}C.

Colony PCR was used to screen the positive colonies containing the ligated gene of interest. In this case, instead of using purified DNA, it was used DNA directly from one bacterial colony which was diluted in 20 µl Milli-Q water. Then, a PCR was performed using 1 µl of the dilution as a DNA template, LAT2 (*Eco*RI) F and LAT2 (*Xba*I) R primers, and *Taq* DNA polymerase (Biotools). The PCR products were analyzed in a 1% agarose gel to check the positive colonies.

The pcDNA3.1+/Strep-TagII-LAT2 generated construct was sequenced using the BigDye® Terminator v3.1 Cycle Sequencing Kit following manufacturer's protocol. Samples were analyzed by the Genomic Unit of the UB Scientific and Technical Services (Parc Científic, Barcelona). LAT2 was sequenced with the universal primers T7 promoter and BGH reverse (see Annex I: Molecular Biology, 1.1. Sequencing oligonucleotides).

I.2. Generation of cysteine mutants for intersubunit crosslinking by site-directed mutagenesis

For crosslinking experiments, the following 4F2hc mutants were generated: S151C, S185C, S191C, S195C, S200C, S393C, S412C, S487C, G505C on a 4F2hc C330S or in a 4F2hc C109S/C330S background. The 4F2hc C330S and the 4F2hc C109S/C330S cDNA templates were already available in the laboratory since they were previously generated on His-4F2hc pcDNA4His-MaxC (Fort *et al.,* 2007). Using the pcDNA3.1+/Strep-TagII-LAT2 as a template (see 1.1.), the mutant C210S was firstly generated, and using this one as template the following different mutants were produced on a LAT2 C210S background: A235C, A315C, G392, S441C, C154S. Secondly, the LAT2 C154S/C210S mutant was created, on this background the mutants A235C, A315C, G392, S441C were generated.

The "QuikChange site-directed mutagenesis" kit (Stratagene) was used to introduce the previous site-directed point mutations in human 4F2hc and LAT2, although for the most difficult clones, the "QuickChange Lightning site-directed mutagenesis kit" (Stratagene) was used following strictly the commercial protocol. The proofreading PfuTurbo DNA polymerase, which replicates both plasmid chains with the incorporated mutation, was used. Two synthetic oligonucleotide primers, around 25 nucleotides long, containing the desired mutation in the middle of the sequence were designed (see Annex I: Molecular Biology, 1.2. Mutagenesis oligonucleotides). The DNA templates used were pcDNA4His-MaxC/His-4F2hc and pcDNA3.1+/Strep-TagII-LAT2 (see I.1). The PCR program used for that aim and the corresponding reaction mixture are described in Table 8.

A

Mix for QuickChange Site-directed	µL
dsDNA template (50 ng)	X
Reaction buffer (10x)	5
10 µM dNTP mix	1
Forward primer (125 ng)	1.25
Reverse primer (125 ng)	1.25
PfuTurbo DNA polymerase* (add last)	1
Milli-Q water, adjust until final volume	50

B

QuickChange PCR	Ta	Time	Cycles
Initial denaturation	95°C	1 min	1x
Denaturation	95 °C	30 seg	16x
Annealing	55 °C	1 min	
Elongation	68 °C	14 min	
Cooling	15 °C	∞	

Table 8. PCR conditions for site-directed mutagenesis. A) mix conditions used from QuickChange Site-directed mutagenesis kit and **B)** PCR conditions used to generate the different mutants.

To remove the parental methylated and hemimethylated DNA, 1 µl of Dpn I endonuclease was added to the PCR reaction mixture for 1 hour at 37°C, but to avoid not complete digestions it was kept 3 hours or overnight in some cases, thus, reducing the number of non-positive sequenced colonies. The nicked vector DNA containing the desired mutations is then transformed into XL1-Blue supercompetent cells (see I.1). Finally, mutants were selected and sequenced (see I.1). 4F2hc was sequenced using D1 forward and AS4 reverse primers (see Annex I: Molecular Biology, 1.1. Sequencing oligonucleotides).

I.3. Transitory co-transfection in HEK 293T

Human Embryonic Kidney 293T (HEK 293T) cell line is an adherent cell line that was originally derived from human embryonic kidney cells and grown in tissue culture. Nowadays represents one of the most used cell lines due to its high rate of transfection and fast growing. Cells were grown with Dulbecco's Modified Eagle Media (DMEM; Gibco ®) supplemented with 10% Fetal Bovine Serum (FBS; heat inactivated 30 min at 56 °C) and 1x penicillin/streptomycin (100 U/mL) (Gibco), at 37°C and in an humidified atmosphere containing 5% CO_2. DMEM is a basal medium consisting of amino acids, vitamins, glucose, salts, and a pH indicator, which contains no proteins or growth promoting agents.

Adherent cell lines grow until they cover the available surface area (100% confluence) or medium is depleted of nutrients. To maintain HEK 293T cell line, cells were sub-cultured before they reach 100% confluence (2-3 times a week) in order to avoid culture damage. To do this, cells were detached from the culture dish by trypsinization: removing medium carefully to not detach cells, washing with 10 mL of sterile Phosphate Buffered Saline (PBS), treating with 1 mL of trypsin (37°C) (Gibco) for 1-5 minutes until cells become into suspension, then the desired medium is added (depending on dilution desired) and finally cells are reseeded into new dishes. Mycoplasma tests were periodically done in the laboratory to avoid contamination in cell cultures (LookOut mycoplasma PCR detection kit). Cells were frozen gradually with 10% DMSO (Sigma Aldrich) until -80 °C or stored in liquid nitrogen for long term storage. HEK 293T cultures were manipulated in sterile conditions by using laminar flux hoods.

Cells were transiently transfected at 60-70% cell confluence using polyethylenimine (PEI) in 15 cm diameter plates with a mixture of DNA containing 15 µg of pcDNA4HisMax/ human 4F2hc mutants, and 15 µg of pcDNA3.1+ human/ Strep-TagII–tagged LAT2 (wild type or mutants) (see 1.1 and 1.2). The mixture was vortexed for 10 seconds and incubated for 15 minutes at room temperature, finally, it was added to the cell culture carefully not to detach cells. A transfection of a single construct was cotransfected with the same amount of the corresponding empty vector DNA. Cell medium was replaced 4 hours after transfection.

I.4. Protein purification from HEK 293T cells

Two days after transfection, cells were washed with 10 mL PBS (pH 6.5; this pH is set from the beginning of the protocol since it is the optimal for later crosslinking, see 1.5.), scrapped with 1 mL of lysate buffer [1% n-Dodecyl-β-D-maltopyranoside (DDM; Anatrace) in PBS (pH 6.5) supplemented with Complete EDTA free (Roche)] per 15 cm plate, and transferred into a 1.5 mL Eppendorf tube. Cell disruption was facilitated by homogenizing 15 times with a 25 gauge syringe and incubation on a rotating orbital shaker at 4°C. After 1 hour, the insoluble material was removed by centrifugation at 200,000*g* for 30 minutes at 4°C. A sample of the supernatant was used to quantify protein by the BCA

Protein Assay Kit (Pierce), and to analyze the protein expression by Western Blot (see I.6.). 75 µL of this sample were stored at −20°C after adding 25 µL of 4xLaemmli sample buffer (277.8 mM Tris-HCl, pH 6.8, 4.4% SDS, 44.4% (w/v) glycerol, 0.02% bromophenol blue).

Solubilized proteins were purified in batch by incubating in a rotor for 1 hour at 4 °C the supernatants (1 mL) with 150 µL of Ni-NTA-agarose beads (Qiagen), previously equilibrated with 1 mL of lysate buffer. Washing was done with 1 mL of lysate buffer containing 25 mM imidazole, and proteins were eluted by shaking incubation for 15 min at 4 °C with 250 µL of elution buffer (lysate buffer containing 250 mM imidazole).

I.5. 4F2hc-LAT2 intersubunit crosslinking

Bis(maleimido)ethane (BMOE; Pierce) and 1,8-bis(maelimido)diethylene glycol [$BM(PEG)_2$; Pierce] crosslinkers were selected due to these characteristics: they conjugate sulfhydryl groups (-SH), they are not cleavable by reducing agents, and they are the shorter spacer arm length crosslinkers available in Pierce with the previous characteristics, of 8 Å and 14.7 Å respectively (according to Pierce description). Actually, accurate measures of the lengths of these molecules (Green *et al.,* 2001) showed a range of distances between 6.27-10.52 Å and 3.51-14.26 Å for BMOE and $BM(PEG)_2$, respectively which make them even more convenient.

The main objective of the experiment was to determine specific pair residues in 4F2hc and LAT2 that were close enough to crosslink with BMOE or $BM(PEG)_2$, and that were located in positions that could triangulate with the endogenous disulfide bridge (see figure 40; see Chapter I: Results I.1 for detailed Cys residues selection for crosslinking). Thus, the region near the endogenous disulfide bridge was not explored since the disulphide bond is already a positive natural crosslinking position between both subunits. Crosslinking experiments were done in purified 4F2hc/LAT2 (through His-4F2hc) samples, and then the disulfide bridges that were not made through the crosslinker, including the endogenous disulfide bridge, were cleaved with reducing agents. Hence, a Western Blot visualizing the StrepII-LAT2 would show the positive crosslinked pair of mutants in the corresponding molecular weight for the heterodimer (see chapter 1: Results I.1).

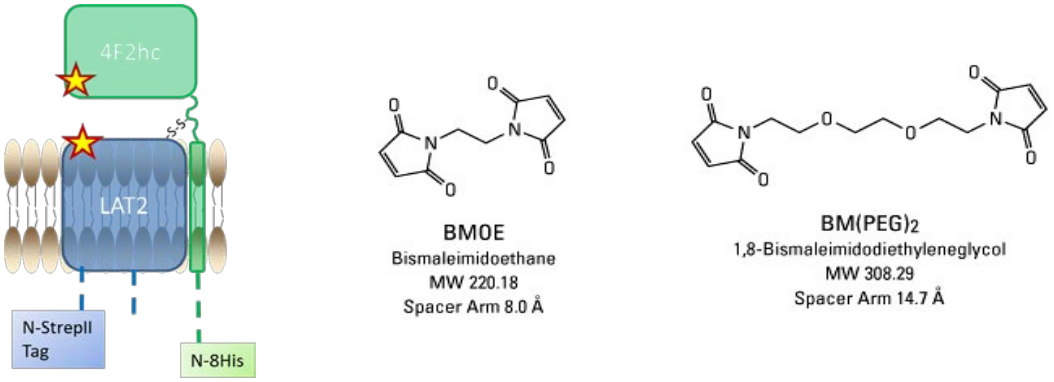


Figure 40. Schematic view of Cys location on 4F2hc/LAT2 heterodimer for crosslinking experiments (left) and crosslinkers used (right). On the left, schematic representation of His-4F2hc and Strep-LAT2 linked by the endogenous disulfide bridge (-S-S-), and the cysteine targets (yellow stars) candidates to crosslink. In the center and on the right, BMOE and $BM(PEG)_2$ (Pierce) molecules are represented respectively.

Once the heterodimers were purified with Ni-NTA, crosslinking protocol was followed. Firstly, BMOE or $BM(PEG)_2$ crosslinker 40Mm stocks were prepared by solubilizing in DMSO anhydrous each compound (2.2mg BMOE in 0.25mL DMSO, and 3.1mg $BM(PEG)_2$ in 0.25mL DMSO). Then, crosslinking buffers were prepared by diluting (1:1) these stocks to a final 20mM concentration by adding 0.25mL of lysate buffer (1% DDM in PBS at pH 6.5 supplemented with Complete EDTA free), generating the same stock concentration as the Pierce manufacturer suggests, but, diminishing 50% the amount of DMSO, which was observed to give better results.

Then, from the 250 µL of purified protein obtained (see I.4.), 2 µL were used to quantify protein by the BCA Protein Assay Kit (Pierce). Protein concentration of all the different conditions tested in a same experiment was adjusted to the lower one (around 2.5 µg/µL). Then, 100 µL of the purified protein was incubated with 30 µL of the crosslinking buffer [200 µM BMOE or $BM(PEG)_2$ crosslinking buffer] for 1 hours at 4°C. The crosslinking reaction was terminated by adding 100 mM DTT. Fourfold concentrated Laemmli sample buffer was added for Western blot analysis which was run directly on SDS-PAGE to avoid losing heterodimer fraction due to aggregation process by freezing and thawing the sample (see I.6.) (Hizukuri *et al.*, 2010).

I.6. Western blot analysis

Western blot was used to detect specifically His-4F2hc and StrepII-LAT2 coming from different kinds of samples: from HEK 293T total cell extracts (detergent-solubilized membrane protein) (see I.4), from purified and crosslinked protein expressed in HEK 293T cells (see I.5), from *P.pastoris* total membrane protein extracts (see I.9) and from detergent-solubilized *P.pastoris* membrane proteins (see I.10). When it was required, 100 mM dithiothreitol (DTT) was added with the aim to reduce the disulfide linkages.

The samples stored at -20°C with Laemmli sample buffer were preheated at 65°C, and 20 µg were loaded into a sodium dodecyl sulphate 8% polyacrylamide gel electrophoresis (SDS-PAGE). Then, proteins in the gel were transferred into an activated PVDF membrane (Immobilon-P, Millipore). Membranes were blocked with 5% skimmed powder milk or 5% BSA in PBS 0.1% Tween 20, for α-His or α-StrepIITag-antibodies, respectively, overnight at 4 °C. Then, membranes were incubated with the primary antibody (1:20,000 dilution) of anti–HisG-HRP (Invitrogen) or Strep MAB-Classic-HRP (IBA) for 1 hour at room temperature. Antibody was reused until signal started to lose (around 10 times). Then, three washes of 10 minutes with PBS 0.1% Tween 20 were done to remove unbound primary antibody,

and finally, to develop the membrane, it was incubated with enhanced chemiluminescent substrate (ECL, Millipore) for detection of horseradish peroxidase (HRP) activity.

I.7. Transport of L-Alanine in HEK 293T Cells

Transfected cells (see 1.3.) were scraped and transferred into 24-well dishes previously treated with poly-L-lysine (100 µL per well at 0.1 mg/mL for 30 min at 37 °C). One day later, L-alanine transport was measured. To this end, cells were washed twice with 2.5 mL of MGA buffer [10 mM Hepes (pH 7.4), 137 mM N-methyl-D-glucamine, 2.8 mM $CaCl_2 \cdot 2H_2O$, 1.2mM $MgSO_4 \cdot 7H_2O$, 5.4mMKCl] at 37 °C. The uptake solution (270 µL per well) contained 50 µM of cold L-alanine plus 0.5 µCi per well of L-[^{3}H]alanine (Hartman) in MGA buffer at 37 °C. After the indicated time, uptake was stopped by two washes with MGA buffer at 4 °C (2.5 mL per well). Cells were then lysed with 250 µL per well of lysis solution (100 mM NaOH, 0.1% SDS) by agitation for 30 min at room temperature. Then, 200 µL was used for radioactivity counting in a Packard Tri-Carb Liquid Scintillation Counter, and 25 µL duplicates were used to determine protein content by the BCA Protein Assay Kit (Pierce). Transport activity was calculated as pmol L-Ala/mg protein at 0, 15 and 30 seconds for control conditions. Since at 30 seconds the biggest differences were seen, transport activity of all the new mutants generated were measured at 30 seconds. All the transport conditions were normalized by the corresponding LAT2 transport activity (each heterodimer condition compared with its same LAT2 monomer transport). % transport activity was calculated comparing with the heterodimer control corresponding to each group.

I.8. 4F2hc-ED overexpression in *Escherichia coli* and purification

4F2hc-ED cloned in pTRC-His vector was overexpressed in *E.coli* BL21 (DE3) cells. *E.coli* cultures were grown at 37°C and 200 rpm until reach a 0.6-0.7 cell density (at 3.5 hours after inoculation). Overexpression induction was started by 1mM IPTG addition at 0.6-0.7 cell density. After 16 hours, cell culture was pelleted at 3,000g at 4°C for 20 minutes. Pellets were resuspended in lysis buffer (Tris 20 mM NaCl 150mM, pH 8.0) and frozen at -80°C (Fort *et al.,* 2007).

Cells were thawed on ice, added complete EDTA free protease inhibitor (Roche) and DNase. Cells were then lysed in a continuous flow Cell Disrupter (Constant Systems TS series) in two rounds of 22 Kpsi, and centrifuged the resulting cell lysate twice at 16,000*g* for 45 min (to remove the non-lysated cells). The soluble fraction was filtered. Purification was performed by using preequilibrated Ni-NTA resin, binding in batch, washing with same buffer with 20mM imidazol and eluting with 300mM imidazol in the same buffer. The final yield was around 2 mg of 4F2hc-ED/liter of initial culture. Buffer was then changed to remove imidazole by centrifugation using Amicon Ultra 15 10,000 Da and protein concentration was measured with BCA method (Fort et al., 2007).

I.9. Human LAT2 overexpressed in *Pichia pastoris* membrane obtantion

	Media preparation for heterologous overexpression in *Pichia pastoris*
YPD	**Yeast Extract Peptone Dextrose Medium:** 450mL were prepared with 10g peptone and 5g yeast extract and autoclaved. 50mL of autoclaved 20% dextrose were added previous use.
BMGY	**Buffered Glycerol-complex medium:** 10g of yeast extract, 20g peptone with 700 mL deionized water were autoclaved. The same day of use, it was complemented with: 100 mL 1M phosphate buffer (autoclaved), 100 mL 10% glycerol (autoclaved), 100 mL Yeast Nitrogen Base (134g YNB in 1 L deionized water filtered and light protected, 2 mL Biotin 500x (0.02%) (20 mg biotin in 100 mL deionized water and filtered).
BMMY	**Buffered Methanol-complex medium:** 10g of yeast extract, 20g peptone in 140 mL of deionized water were autoclaved, and, the same day of use, it was complemented with: 20 mL 1M phosphate buffer , 20 mL 10% methanol, 20 mL YNB and 400 µL biotin (0.02%).

Table 9. Media preparation to overexpress membrane proteins in *P.pastoris*.

Human LAT2 (with N-terminal StrepIItag) cloned in pPICZB vector and transformed into *P.pastoris* KM71H strain (Costa et al., 2013) was overexpressed following the next procedure: firstly, 5mL of YPD were inoculated from a glycerol of the best selected human LAT2 clone. It was grown at 30°C, 250 rpm, for 16 hours. Then, 20 µL from the first YPD culture were inoculated into selective media: 5 mL of YPD with Zeocin (100 µg/mL) and grown at 30°C, 250 rpm for 16 hours.

200 µL from the zeocin resistant starter (YPD) were inoculated in 1 L of BMGY, grown at 30°C, 250 rpm, until culture reaches an OD_{600} = 2–6 (approximately 16–18 hours). After pelleting at 1,500*g* and 4°C for 5 min, the cells were resuspended in 200 mL buffered complex methanol medium (BMMY) to start induction and grown for 72 hours at 30°C. Additional methanol was added after 24 and 48 h to a final concentration of 1% (v/v), in order to maintain inducing conditions. Cell were centrifuged at 2,500*g* for 5 minutes at room temperature and cell pellets were resuspended in 50 mM sodium phosphate, pH 7.4, 1 mM EDTA, 5% (w/v) glycerol and complete EDTA free protease inhibitors (Roche).

Cells were then lysed in a continuous flow Cell Disrupter (Constant Systems TS series) in three rounds of 25, 35 and 40 Kpsi. After centrifugation at 20,000*g* and 4°C for 15 min (to remove the non lysated cells), supernatant was recovered and ultracentrifuged at 250,000*g* and 4°C for 1 hour. The pellet was resuspended and homogenized (50 mL glass homogenizer) in 20 mM Tris–HCl, pH 7.4, 150 mM NaCl, 5% glycerol, and ultracentrifuged again. Finally, the membrane pellet was resuspended in 20 mM Tris–HCl, pH 7.4, 150 mM NaCl, 5% glycerol at a protein concentration between 20 and 25 mg/mL (the protein content was determined using the bicinchoninic acid protein (BCA) assay (Pierce)). Aliquots of 1 mL were frozen in liquid nitrogen and stored at -80°C until use.

I.10. 4F2hc-ED impact on solubilization/stability of LAT2 overexpressed in *P.pastoris*

To solve the question whether 4F2hc-ED interacts with LAT2, and whether it has a stabilizing role on LAT2, several experiments were designed to address this issue. In all of them, purified 4F2hc-ED was added in exceed to *P.pastoris* membranes overexpressing human LAT2 and mixed overnight. Next day sample containing 4F2hc-ED addition and sample without 4F2hc-ED were detergent-solubilized, and removed the aggregated protein by ultracentrifugation. Afterwards by western blot or Fluorescence Size Exclusion Chromatography (in the case of LAT2-GFP overexpressed (see 2.4)) the remaining soluble LAT2 was analyzed.

In all the experiments, 2 mg/mL of purified 4F2hc-ED in 20 mM Tris·HCl (pH 7.4), 150 mM NaCl (see 2.1) or just the same volume buffer where mixed with *P.pastoris* membranes overexpressing LAT2 (total membrane protein 20 mg/mL) on a rotating orbital shaker overnight at 4ºC. The final concentration of 4F2hc-ED was 1mg/mL. 2 mg/mL of Bovine Serum Albumin (BSA, Sigma) was added, instead of 4F2hc-ED, in the same buffer as a negative control condition, in order to check whether the possible 4F2hc-ED impact on LAT2 is exclusive.

In a first experiment, after 4F2hc-ED/*Pichia* membranes overnight incubation, three different decreasing concentrations of DDM were tried for LAT2 solubilization (1, 0.5 and 0.25%) to detect a more dramatic effect in case there was. Samples were solubilized for 1 h at 4°C and aggregates were removed by ultracentrifuged at 200,000*g* for 1 h. Supernatant samples (100 µL) were mixed with 4xLaemmly sample buffer and stored at -20°C. For protein stability analysis, the sample solubilized with 1% DDM was considered time 0. The rest of supernatant was kept 2, 24 and 72 hours previous a second ultracentrifugation. Again, supernatant samples (100 µL) were stored with 4xLaemmly sample buffer at -20°C. All samples were analyzed by western blot visualizing human StrepTagII-LAT2 by using the Strep MAB-Classic-HRP antibody (IBA) (see 1.6).

I.11. Bioinformatic tools

- Uniprot was used to find the amino acid sequence of the different HAT studied.
- Protein published 3D structures were searched in the Protein Data Bank (http://www. pdb. org),
- The PyMOL Molecular Graphics System (http://www. pymol. org) was used to visualize and generate the different figures of the 3D structures and models of HAT studied.
- ImageJ program to quantify the western blot bands by densitometry.

Chapter II

Identification of new HAT targets for 3D crystallization

II. Introduction

Chapter II: Introduction

The second part of the thesis was focused on finding a more stable eukaryotic light subunit than human LAT2, which could be suitable for 3D-crystallization studies. In the first Chapter, it was shown that purified human LAT2 alone with DDM is highly unstable and aggregates. On the other hand, a clear stabilizing effect of the heavy subunit 4F2hc on this light subunit was observed when co-expressed and purified as a heterodimer. Unfortunately, this effect was not sufficient to make the heterodimer 4F2hc/LAT2 suitable for 3D crystallization studies. Nevertheless, if we would succeed in finding a more stable light subunit, when trying the co-expression with its respective heavy subunit we would expect an increase of the heterodimer stability. In addition, the already 4F2hc-ED solved crystal structure (Fort *et al.,* 2007), would help us to solve the structure of the corresponding heterodimer.

Obtaining high-resolution membrane protein structures is still a difficult challenge that requires overcoming many difficulties: protein overexpression, detergent solubilization, and purification of stable and active membrane protein at high concentrations. Finally, even if pure and stable protein could be obtained, yet crystallization would not be certain. Nevertheless, over the last years novel strategies and improvement of crystallization technologies have raised the probability of success.

In our project we integrated several of the strategies commonly used to identify and optimize membrane protein targets for 3D crystallization. To mention some, it has been tried: different eukaryotic expression systems depending on the final goal: *Saccharomyces cerevisiae*, *Pichia pastoris* or Sf9 Insect cells; different detergents that have been successful in membrane protein crystallography; the addition of cholesterol and lipids; the elimination of the reactive cysteines to decrease the aggregation of the protein and the generation of truncation forms of the protein to reduce the flexibility and, hence, improve the chances of crystallizing it.

- **Strategy to find a more stable light subunit than human LAT2**

To identify a more stable eukaryotic light subunit than human LAT2, firstly different light subunits from different species were selected for the study. To speed up the process, it was adapted the pipeline developed by Dr. Drew from So Iwata laboratory (Imperial College London, Diamond Light Source) to achieve a fast screen of the expression levels and stability of eukaryotic membrane proteins. This high throughput screening is based on the overexpression in the baker yeast *S. cerevisiae* the protein candidates as GFP fusion proteins (Drew *et al.*, 2008).

The biggest advantage of this pipeline is the introduction of the GFP-fusion protein technology. This strategy allows: 1) to rapidly screen the protein overexpression levels by measuring the whole-cell fluorescence, which is specific and sensitive down to 5 ng; 2) thanks to the property of the GFP to remain folded in SDS, the integrity of the fusion protein can be detected after SDS-PAGE by in-gel

fluorescence; 3) to fast screen the detergent solubilization efficiency and 4) to monitor the monodispersity of recombinant proteins in different detergent-solubilized membranes by Fluorescence Size Exclusion Chromatography (FSEC) (Kawate and Gouaux, 2006) (Figure 41). In fact, the GFP technology could be adapted for other eukaryotic expression systems, as it has been done for *P. pastoris* (Mizutani *et al.,* 2011), and insect cells (Baconguis *et al.,* 2014).

The best advantage of the GFP-tag is the possibility to monitor at the first step the monodispersity of the sample in detergent-solubilized membranes by FSEC without the need to purify the protein. Symmetry and broadness of the peak obtained in FSEC indicates how homogeneous is the oligomeric state of the protein. This technology was developed in the laboratory of Dr. Gouaux (Howard Hughes Medical Institute (HHMI) and Oregon Health & Science University (OHSU)) as a pre-crystallization screening tool due to different advantages: it is time saving and low resource consuming, it requires little amount of sample, and it gives an initial and useful information of the quality of the membrane protein produced (Kawate and Gouaux, 2006).

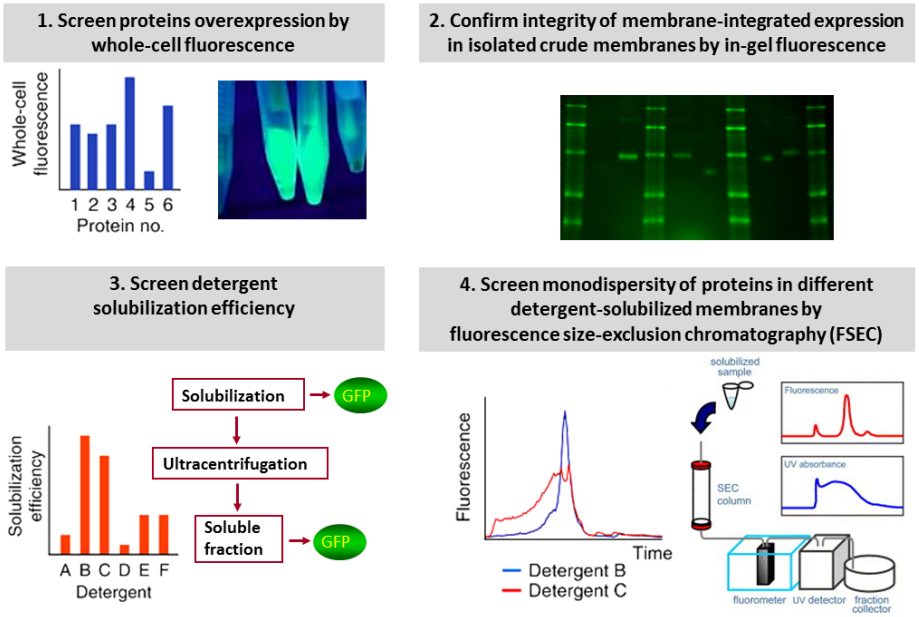


Figure 41. Advantages of the expression of GFP-tagged membrane proteins. Scheme representing the main advantages that allow to screen faster membrane proteins using the GFP-fusion protein technology (adapted from Kawate and Gouaux, 2006; Drew *et al.,* 2008).

In addition, *S. cerevisiae* proves to be the most suitable expression system to screen large amounts of eukaryotic membrane proteins (Drew *et al.,* 2008; Scharff-Poulsen and Pedersen, 2013; Parker and Newstead, 2014) because the cloning by homologous recombination and the expression screenings are less time consuming in comparison with other eukaryotic expression systems such as *P. pastoris*, insect cells or mammalian cells. Moreover, *Saccharomyces* have similar protein synthesis mechanisms, maturation machinery (such as disulfide bond formation or glycosylation) and membrane-trafficking pathways as mammalian cells. Furthermore, it has a highly regulated quality control system of membrane proteins in the endoplasmic reticulum. On the other hand, growing *Saccharomyces* is fairly

simple and cheap compared to other eukaryotic expression systems. Nevertheless, the differences in glycosylation patterns, in lipid membrane composition or lack of cholesterol may compromise in some cases the protein folding and function, or the expression levels. In these cases, it would be required to try other expression systems.

Interestingly, some groups have been successful in determining high-resolution structures of eukaryotic membrane proteins using this expression system. Some examples are: the human monoamine oxidase A at 2.2 Å (Se-Young Son *et al.,* 2008), the fungal (*Piriformospora indica*) high-affinity phosphate transporter at 2.9 Å (Pedersen *et al.,* 2013), the *Arabidopsis thaliana* H^+-ATPase at 3.6 Å (Pedersen *et al.,* 2007), the rabbit sarcoplasmic-endoplasmic reticulum Ca^{+2}-ATPase isoform 1a mutants D351A and P312A, obtained at 3.4 Å and 3.5 Å, respectively (Marchand *et al.,* 2008); the *S. cerevisiae* cytochrome bc(1) complex at 2.3 Å (Hunte *et al.*, 2000), the *S. cerevisiae* c10 ring of the ATP synthase at 2.0 Å (Pagadala *et al.,* 2011; Symersky *et al.*, 2012), the *Vigna radiate* H-translocating M-PPase at 2.35 Å (Lin *et al.,* 2012) or the *S. cerevisiae* vacuolar calcium/proton exchanger at 2.3 Å (Waight *et al.*, 2013).

- **Improvement of membrane protein environments by detergents, cholesterol and lipids**

Being successful in crystallization of membrane proteins requires the identification of detergents capable of maintaining the protein in a soluble and monodisperse state. Because of their hydrophobic nature, membrane proteins tend to form insoluble aggregates over time which reduces the probability of forming regular associations of proteins that are essential to form crystal lattices. Detergents provide convenient means for extracting membrane proteins but they poorly mimic the native membrane. Detergents should cover all the hydrophobic regions of the membrane protein once it is solubilized. This step seriously compromises the stability of this kind of proteins. At the moment, there is not any detergent that works for all the membrane proteins yet. For that reason, a careful selection of a variety of detergents is needed to find the more suitable one for each protein, since the solubilized proteins have to maintain their structural and functional properties and have to allow the generation of high quality crystals for X-ray diffraction.

Addition of cholesterol in the salt form cholesteryl hemisuccinate tris salt (CHS) during solubilization; furthermore, purification has been demonstrated to improve eukaryotic membrane proteins stability, especially in the case of the overexpressed proteins in *S. cerevisiae,* that lacks cholesterol in its membrane (Yao and Kobilka, 2005; Drew *et al.,* 2008). The use of cholesterol during purification and/or crystal screenings have led to solve crystal structures of eukaryotic proteins like the human Histamine H1 receptor complex with doxepin (Shimamura *et al.,* 2011).

Interestingly, some eukaryotic membrane protein structures have been solved with a cholesterol molecule inserted between their transmembrane domains. This is the case of the *Drosophila* Dopamine transporter (DAT) (Penmatsa *et al.,* 2013), or the crystal structure of the human β2-

adrenergic receptor–T4 lysozyme fusion protein (Cherezov *et al*., 2007). In the case of DAT, 0.1 mM CHS was added in the elution step, and in the second case, crystals were grown from a cholesterol-doped monoolein cubic phase. In addition, a palmitic acid was found in the structure of β2-adrenergic receptor (Cherezov *et al*., 2007). These are some evidences of the importance of cholesterol and lipids to maintain the native structure of membrane proteins, increase their stability and keep them functional.

The use of synthetic lipids to stabilize eukaryotic membrane proteins has become of common use in the recent years (lipids can be added during the solubilization, purification or in the crystallization conditions) since it has been demonstrated in many cases their stabilizing effect on the protein. In parallel, different new crystallographic techniques have been developed which require the use of lipids to form ordered lipidic structures containing the membrane protein. Some of these specific crystallographic techniques for membrane proteins are: HiLiDe (Gourdon *et al.,* 2011), Lipidic Cubic Phase (Caffrey *et al.,* 2009) and Bicelles (Faham *et al*., 2002; Agah and Faham, 2012). Some examples of structures that have been solved with addition of lipids are the *Drosophila* Dopamine transporter (*DAT*) (Penmatsa *et al.,* 2013), or the chicken acid-sensing ion channel 1-snake toxin complex (Baconquis *et al.,* 2014).

- **Choosing the best expression system:**

S. cerevisiae is a great expression system to fast screen a big number of membrane protein candidates and to identify putative good solubilization and stability conditions. The protein expression levels using this yeast are not usually very high. To overcome this problem, other expression systems are needed in order to increase the production of the identified target in *Saccharomyces*, but taking advantage of the optimized stability conditions using *Saccharomyce*s.

P. pastoris has been successful for more than 15 eukaryotic membrane protein structures. One of its major advantages is that it usually produces higher expression levels than other eukaryotic expression systems. Some examples are: the human AQP2 Aquaporin from kidney at 2.75 Å (Frick *et al.,* 2014), the human Leukotriene LTC4 Synthase at 2.75 Å (Niegowski *et al.,* 2014), the human Glioma pathogenesis-related protein 1 at 2.2 Å (Asojo *et al.,* 2011) or the rat voltage gated potassium channel subunit β2 at 2.9 Å (Tao *et al.,* 2010).

Insect cells result of special interest since it has been the most successful expression system used to solve the 43% of high resolution structures of eukaryotic membrane proteins, especially from human (Parker and Newstead, 2014). This fact could be probably due to its ability to produce stable proteins in high quantities. Some examples are: the human β2-adrenergic receptor–T4 lysozyme fusion protein at 2.4 Å (Cherezov *et al*., 2007), the human β2 adrenergic G-protein-coupled receptor at 3.4 Å (Rasmussen *et al.,* 2007), the human $A2_A$ adenosine receptor at 2.6 Å (Jaakola *et al.,* 2008), the human CXCR4 chemokine receptor at 2.5 Å (Wu *et al.,* 2010), the human Dopamine D3 receptor at 2.89 Å (Chien *et al.,* 2010), the human Sphingosine 1-phosphate receptor at 3.35 Å (Hanson *et al.,* 2012), the human

M2 muscarinic acetylcholine receptor at 3.0 Å (Haga *et al.,* 2012), the human κ-opioid receptor at 2.9 Å (Wu *et al.,* 2012), the human 5-HT_{2B} serotonin receptor at 2.7 Å (Wacker *et al.*, 2013), the human BK (SLO1) Channel Ca^{2+}-activation apparatus at 3.0 Å (Yuan *et al.,* 2010), the human microsomal prostaglandin E synthase 1 at 1.16 Å (Sjögren *et al.,* 2013) or the human GLUT1 N45T/E329Q mutant at 3.17 Å (Deng *et al.,* 2014).

Only a few structures have been obtained so far in mammalian cells due to its expensive costs. Some examples are: the *Drosophila melanogaster* Dopamine transporter at 3.0 Å (Penmatsa *et al.,* 2013), the human Rh C glycoprotein ammonia transporter at 2.10 Å expressed in HEK293S cells (Gruswitz *et al.,* 2010) or the human $GABA_AR$ receptor (β3 homopentamer) at 2.97 Å expressed in HEK293F cells (Miller *et al.,* 2014).

The variety of aforementioned methods and technologies made possible to find a good light subunit candidate for 3D crystallization.

II. Results

II.2. Protein expression screening in *Saccharomyces cerevisiae*

In order to find a more suitable HAT than human 4F2hc/LAT2 for 3D-crystallization trials our first aim was to find a more stable light subunit than human LAT2. Afterwards, the corresponding heavy subunit would be co-transformed and the stability of the heterodimer tested, if formed. We would expect to see an increase of the light subunit stability when adding the heavy subunit as it was observed in the case of human 4F2hc/LAT2 (see Chapter I).

To that aim, 24 eukaryotic light subunits belonging to different species, including human, where pre-selected to be cloned and expressed as GFP fusion proteins (table 10) in the baker's yeast *S. cerevisiae.* The corresponding cDNAs were cloned into pDDGFP-2 vector; then, transformed into *S. cerevisiae* cells (strain FGY217) and finally their expression was tested (see Chapter II: Materials & Methods II.1-4).

	asc1	LAT1	LAT2	y^+LAT1	y^+LAT2	xCT	Mnd	Jhl-21	$b^{0,+}$AT
Homo sapiens **(Human)**	X	X	X	X	X	X			X
Mus musculus **(Mouse)**	X	X	X	X	X	X			
Rattus norvegicus **(Rat)**									X
Bos taurus **(Cow)**				X					X
Xenopus tropicalis [1] **/** ***laevis*** [2] **(Toad)**		X[2]	X[2]	X[2]	X[2]				X[1]
Danio rerio **(Zebrafish)**				X					
Drosophila melanogaster **(Fly)**							X	X	

Table 10. Eukaryotic light subunits included for the expression screening in *Saccharomyces*. 24 eukaryotic light subunits from different species were selected for the expression screening (human, *Homo sapiens*; mouse, *Mus musculus*; rat, *Rattus norvegicus*; cow, *Bos taurus*; toad, *Xenopus laevis and Xenopus tropicalis*; Zebrafish, *Danio rerio* and fly, *Drosophila melanogaster*). $b^{0,+}$AT is the only light subunit of the screening that forms heterodimers with rBAT and not with 4F2hc in vivo. *Drosophila melanogaster* Mnd and Jhl-21 transporters are homologues of human system L (Reynolds *et al.*, 2009).

The expression levels of the different eukaryotic light subunits were established by measuring the whole-cell fluorescence in small scale cell cultures 22 h after protein expression and induction with 2% galactose (see Chapter II: Materials & Methods II.4). Only those light subunits with a fluorescence / absorbance ratio 5-fold or higher than the one of the yeast transformed with the empty vector (Figure 42) were selected as good targets.

In this first step, 14 eukaryotic light subunits were discarded for not reaching the minimum expression levels required for this project. These proteins were: asc1 (from human and mouse), LAT1

(from human and mouse), y^+LAT1 (from mouse and toad), y^+LAT2 (from human, mouse and toad), xCT (from mouse), Mnd (from fly) and $b^{0,+}$AT (from human, rat and cow).

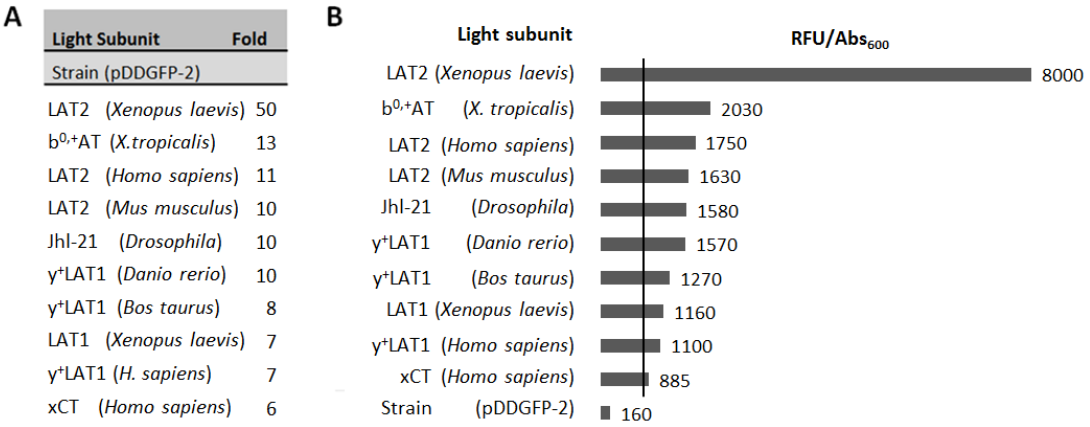


Figure 42. Selection of eukaryotic light subunits after expression screening in *S.cerevisiae*. A) List of the light subunits selected for overexpressing over the cut off (5 fold the negative control). Overexpression levels are measured by fluorescence (RFU) / final absorbance at 600 nm (Abs_{600} nm). **B**) Graphic representing the RFU/Abs_{600} corresponding to the selected light subunits. The line indicates the cut-off at 5 fold background signal.

After selecting the best candidates with the previous criteria, the amount of overexpressed protein related to the fluorescent measurements (Figure 43) was quantified. To that aim, cell culture expression of each selected protein was scaled up to 1 L. A stock of purified yEGFP from *Pichia pastoris* previously generated by Dr. Rosell and stored at -80 °C (Figure 43A) was used to prepare a yEGFP standard to determine the mg of expressed membrane protein fused to GFP for each candidate (Drew *et al.,* 2008) (see Chapter II: Materials & Methods, II.4). The mg of GFP-tagged light subunit produced per liter of cell culture ranged from 0.5 to almost 5 (Figure 42B).

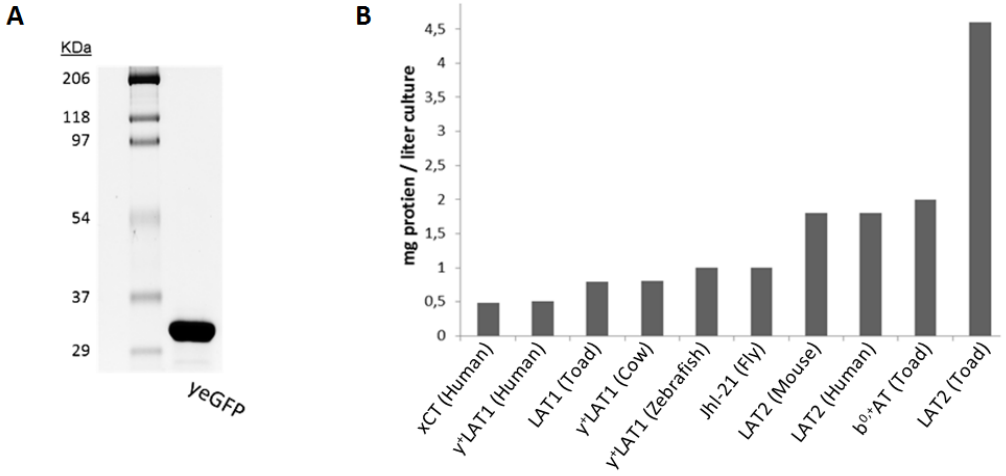


Figure 43. Determination of the amount of GFP-tagged light subunits produced in *Saccharomyces*. A) Purified yEGFP run in a SDS-PAGE, 10% polyacrylamide gel and stained with Comassie. **B**) Graphic indicating the mg of expressed GFP-tagged light subunits per liter of cell culture.

As a result of the quantification studies, LAT2 (*Xenopus laevis*) was the light subunit with the best expression levels (4.5 mg/L), followed by $b^{0,+}$AT (*Xenopus tropicalis*) (2 mg/L) and LAT2 (mouse and human) (1.8 mg/L) (Figure 43B). Interestingly, LAT2 from three different species has been well overexpressed in *S. cerevisiae* as it was human LAT2 in *P. pastoris* compared to the rest of light subunits (Costa *et al.*, 2013).

To evaluate the integrity of the GFP fusion proteins, SDS-PAGE in-gel fluorescence of crude-cell membrane extracts overexpressing the 10 selected light subunits was performed (Figure 44). The presence of free GFP may indicate proteolysis or problems during the translation process of the recombinant fusion-protein. Fortunately, all the different light subunits showed the expected molecular weight in the in-gel-fluorescence, indicating that there were no proteolysis or translation protein problems, probably due to the fact of GFP presence in the C-terminal and not in the N-terminal (Drew *et al.,* 2008). On the other hand, this result is important since it assures that the measured fluorescence in the whole cell culture to quantify the expression levels is directly measuring the overexpressed protein and not free GFP.

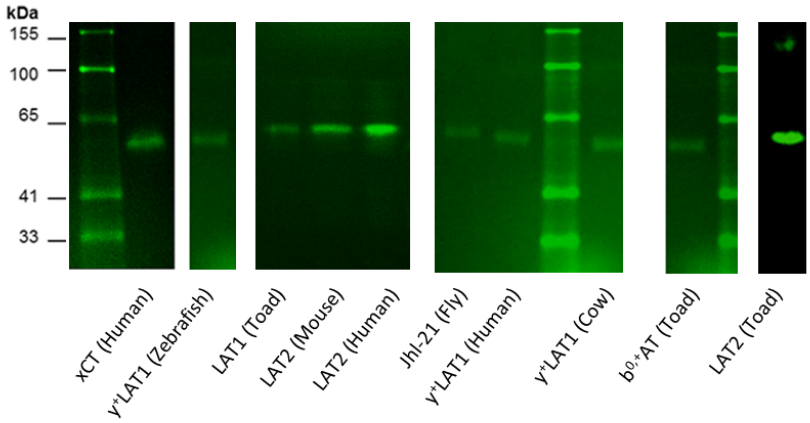


Figure 44. In-gel fluorescence of crude cell membrane extracts overexpressing GFP-tagged eukaryotic light subunits. The integrity of xCT (human), LAT2 (human), LAT2 (mouse), $b^{0,+}$AT(toad), Jhl-21 (fly), y^{+}LAT1 (zebrafish) and LAT2 (toad) was analyzed by in-gel fluorescence. All light subunits presented the expected molecular weight (~35 kDa + 29 kDa corresponding to the GFP fusion protein). In all the cases, only monomer were present, with the exception of LAT2 (toad) that presents dimer forms, and some aggregates.

II.3. Membrane protein-detergent solubilization screening

The ten light subunits showing the highest expression levels (II.2) were solubilized with 13 detergents of different characteristics and length that were selected for their relevance in membrane protein crystallography (Table 11). To that aim, 1 L cell culture of each light subunit were grown, and the resulting cell crude membrane extract was divided in 1 mL stocks to test the 13 different detergents from the same batch cell culture.

Non-ionic detergents	**Zwitterionic detergents**
6-cyclohexyl-1-hexyl-β-D-maltoside (Cymal-6)	N,N-dimethyldodecylamine N-oxide (DDAO)
7-cyclohexyl-1-heptyl-β-D-maltoside (Cymal-7)	Lauryl dimethylamine N-oxide (LDAO)
Nonyl maltoside (NM)	3-dodecylamido-N,N'-dimethylpropyl amine oxide (LAPAO)
Decyl maltoside (DM)	Foscholine-9 (Fos-9)
Dodecyl maltoside (DDM)	Foscholine-12 (Fos-12)
Octyl glucoside (OG)	Octaethylene glycol monododecyl ether ($C_{12}E_8$)
Nonyl glucoside (NG)	

Table 11. List of detergents selected for the eukaryotic light subunit solubilization screening. The detergents used to perform the solubilization efficiency screening for the selected light subunits belonged to two different families: non-ionic detergents (cymal-6, cymal-7, NM, DM, DDM, OG, NG) and zwitterionic detergents (DDAO, LDAO, LAPAO, Fos-9, Fos-12, $C_{12}E_8$).

In all the cases, crude cell membranes were solubilized to a final total membrane concentration of 3 mg/mL with 1% (w/v) detergent. The solubilization efficiency was calculated by using the GFP fluorescence values obtained from the detergent-solubilized membranes and the non-solubilized material (see Chapter II: Materials & Methods, II.6).

The cut-off value to select the more suitable detergent for each protein was determined at 40% solubilization yield. In this screening, 62 conditions fitted the cut-off from the total of 130 conditions tested. A protein was discarded, y^+LAT1 (from cow), since no detergent achieved to solubilize it over the cut-off (Table 12). In addition, two detergent conditions were discarded: octyl glucoside (OG) and foscholine-9 (Fos-9) because they were not able to solubilize any of the light subunits tested.

Thereafter, cholesterol effect in the detergent solubilization yield was tested on the 62 successful solubilization conditions. For all the conditions tested, 0.2% (w/v) cholesteryl hemisuccinate Tris salt (CHS) was added to the 1% (w/v) detergent. Cholesterol effect was tested since it was already described to have a stabilizing role on eukaryotic membrane proteins and it helps to increase the solubilization efficiency (Drew *et al.*, 2008). In our study, even though cholesterol effect on solubilization efficiency is modest, it produces an improvement in the solubilization of the protein in 49 out of the 62 conditions tested (Table 13).

Solubilization efficiency										
	LAT2 (Human)	LAT2 (Mouse)	LAT2 (Toad)	y^+LAT1 (Human)	y^+LAT1 (Zebrafish)	xCT (Human)	b^{0+}AT (Toad)	Jhl-21 (Fly)	LAT1 (Toad)	y^+LAT1 (Cow)
OG	<20%	<20%	<20%	<20%	<20%	<20%	<20%	<20%	<20%	<20%
NG	50%	<20%	<20%	<20%	<20%	<20%	<20%	<20%	<20%	<20%
NM	40%	<20%	<20%	<20%	<20%	<20%	40%	<20%	50%	<20%
DM	50%	40%	50%	<20%	<20%	<20%	40%	<20%	50%	<20%
DDM	50%	50%	50%	40%	40%	<20%	50%	40%	50%	<20%
Cymal-6	50%	50%	50%	<20%	40%	<20%	40%	<20%	40%	<20%
Cymal-7	<20%	50%	50%	<20%	<20%	<20%	50%	<20%	50%	<20%
Fos-9	<20%	<20%	<20%	<20%	<20%	<20%	<20%	<20%	<20%	<20%
Fos-12	<20%	50%	50%	50%	50%	50%	<20%	<20%	50%	<20%
DDAO	80%	40%	70%	40%	40%	<20%	40%	40%	50%	<20%
LDAO	<20%	50%	50%	50%	50%	<20%	50%	50%	50%	<20%
LAPAO	90%	80%	90%	60%	90%	40%	80%	60%	80%	<20%
$C_{12}E_8$	50%	50%	50%	<20%	<20%	<20%	50%	<20%	50%	<20%

Table 12. Screening of the most suitable detergents to solubilize each of the light subunit candidates by analyzing its solubilization efficiency. All the selected light subunits were detergent-solubilized with the non-ionic (OG, NG, NM, DM, DDM, Cymal-6 and Cymal-7) and zwitterionic detergents (Fos-9, Fos-12, DDAO, LDAO, LAPAO and $C_{12}E_8$) indicated in the table. Only conditions with 40% or higher solubilization efficiency were selected as good candidates (colored in grey gradient, being darker grey those conditions with higher solubilization efficiency). 68 detergent conditions were discarded in this screening. OG and Fos-9 detergents were not able to solubilize efficiently in any case.

Improvement (%) in the solubilization efficiency using cholesterol hemisuccinate Tris salt (0.2 % w/v)										
	LAT2 (Human)	LAT2 (Mouse)	LAT2 (Toad)	y^+LAT1 (Human)	y^+LAT1 (Zebrafish)	xCT (Human)	b^{0+}AT (Toad)	Jhl-21 (Fly)	LAT1 (Toad)	y^+LAT1 (Cow)
OG	d.c	d.c	d.c	d.c	d.c	d.c	d.c	d.c	d.c	d.c
NG	**11**	d.c	d.c	d.c	d.c	d.c	d.c	d.c	d.c	d.c
NM	**10**	d.c	d.c	d.c	d.c	d.c	**10**	d.c	**12**	d.c
DM	**15**	**10**	**10**	d.c	d.c	d.c	**17**	d.c	**16**	d.c
DDM	**N.D.**	**11**	**10**	**10**	**10**	d.c	**16**	**25**	**17**	d.c
Cymal-6	**10**	**N.D**	**10**	d.c	**N.D.**	d.c	**10**	d.c	**N.D**	d.c
Cymal-7	d.c	**13**	**13**	d.c	d.c	d.c	**12**	d.c	**10**	d.c
Fos-9	d.c	d.c	d.c	d.c	d.c	d.c	d.c	d.c	d.c	d.c
Fos-12	d.c	**16**	**18**	**15**	**14**	**64**	d.c	d.c	**16**	d.c
DDAO	**12**	**10**	**12**	**N.D**	**11**	d.c	**19**	**11**	**12**	d.c
LDAO	d.c	**12**	**15**	**16**	**17**	d.c	**14**	**15**	**14**	d.c
LAPAO	**N.D**	**N.D**	**10**	**N.D**	**N.D.**	**25**	**N.D.**	**N.D**	**N.D**	d.c
$C_{12}E_8$	**12**	**10**	**N.D**	d.c	d.c	d.c	**10**	d.c	**10**	d.c

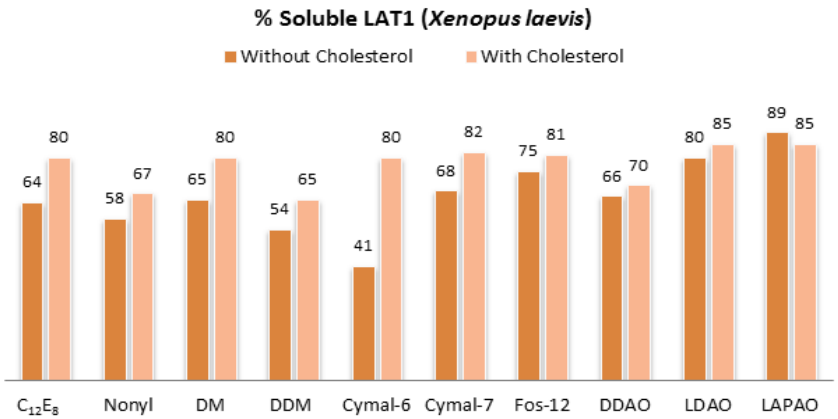
Table 13. Effect of cholesterol hemissucinate Tris salt (CHS) on detergent solubilization efficiency. Cholesterol effect was tested on the 62 previously selected conditions (Table 12). Conditions where the addition of CHS improved the solubilization yield are colored in grey (49 conditions from 62 did improve with cholesterol addition). No differences were observed (N.D.). Discarded conditions (d.c) in the previous screening were not tested with CHS.

II.4. Membrane protein stability screening by the Ultracentrifugation Dispersity Sedimentation assay

To get some information about the effect of the detergent conditions previously selected on the stability of our light subunit candidates, firstly we performed a stability study by the Ultracentrifugation Dispersity Sedimentation Assay (UDS; Gutmann *et al.*, 2007). UDS is faster and requires less use of detergent buffer than other methods which is a great advantage to screen a big number of conditions.

Then, UDS assay was tested in all detergent conditions that did solubilize the eukaryotic light subunits, with or without cholesterol, in a yield higher than 40%. In this method, the detergent solubilized-protein was kept overnight at 4 °C, ultracentrifuged at 80,000 rpm and the % of remaining solubilized protein compared to the initial one analyzed (see Chapter II: Materials & Methods, II.7). In the graphics from Figure 45 the average of 3 experiments per condition is represented. The cut-off value applied in this step was 80% of soluble protein after 16 h at 4 °C.

From the 124 detergent conditions (with and without CHS) tested using the UDS method, 59 conditions were discarded because they did not fulfilled the cut-off requirement. From the discarded conditions, 44 did not have cholesterol and 15 did have. Even though UDS technique made possible to identify detergent conditions able to decrease the kinetic aggregation of light subunits, it does not provide information about the quality of the remaining soluble protein. To evaluate the quality of the 8 remaining transporters, we studied the monodispersity of these proteins in the selected detergent conditions by Fluorescence Size-Exclusion Chromatography (FSEC) (Kawate and Gouaux, 2006).



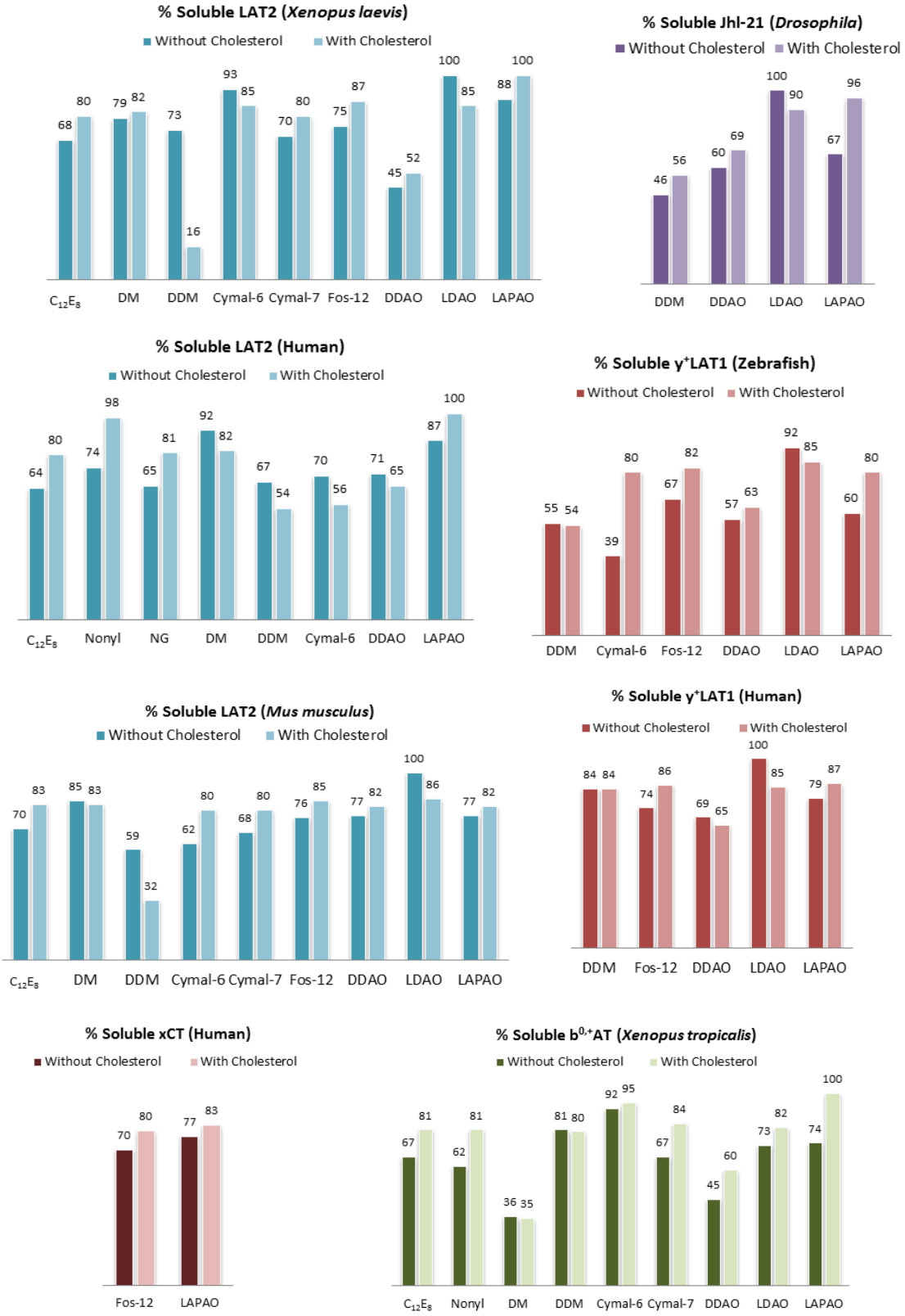


Figure 45. Analysis of the stability of the eukaryotic light subunits by the UDS assay. All the detergent conditions selected previously were tested with and without CHS addition. Graphic bars show eukaryotic light subunit % stability after UDS assay. The cut-off of this screening was 80% of remaining soluble protein after 16 h at 4°C.

II.5. Light subunit monodispersity study by Fluorescent Size-Exclusion Chromatography (FSEC)

Only those conditions with 80% or more of remaining soluble protein after UDS assay were tested for FSEC assay stability (see Chapter II: Materials & Methods, II.8) which allows to assess the monodispersity of the overexpressed and solubilized protein without the need of purifying it.

At the beginning of the study of the stability of these transporters analyzed by FSEC, all the conditions obtained in the previous step (Figure 45) were selected. However, after observing a clear improvement of those conditions which cholesterol was added compared to those without cholesterol (Figure 46), it was decided to continue the rest of FSEC screening only with detergent + cholesterol conditions. Stability improvement on eukaryotic membrane proteins after cholesterol addition has been described in other cases (Penmatsa *et al*., 2013; Drew *et al.*, 2008).

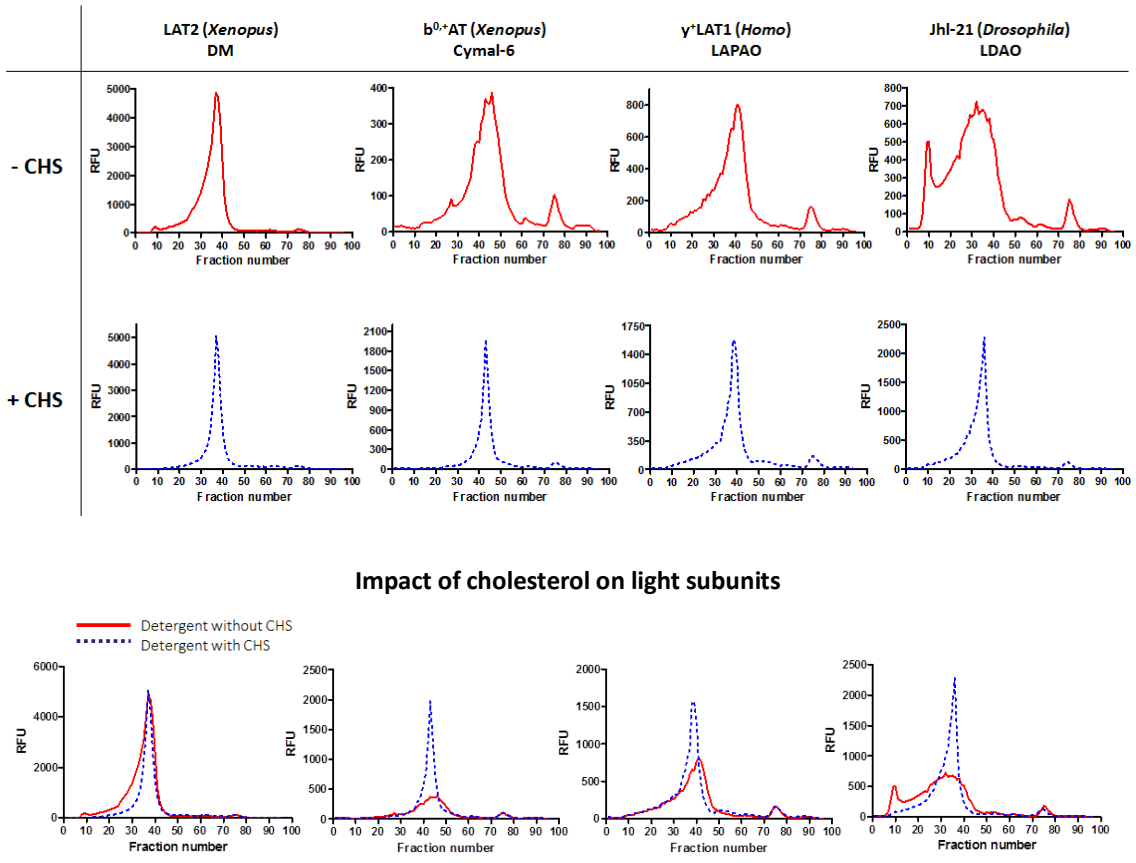
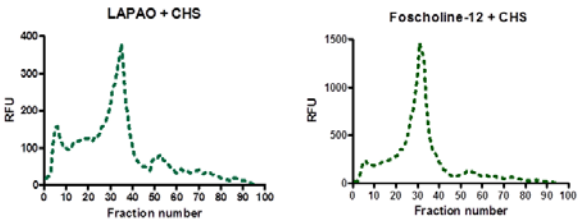


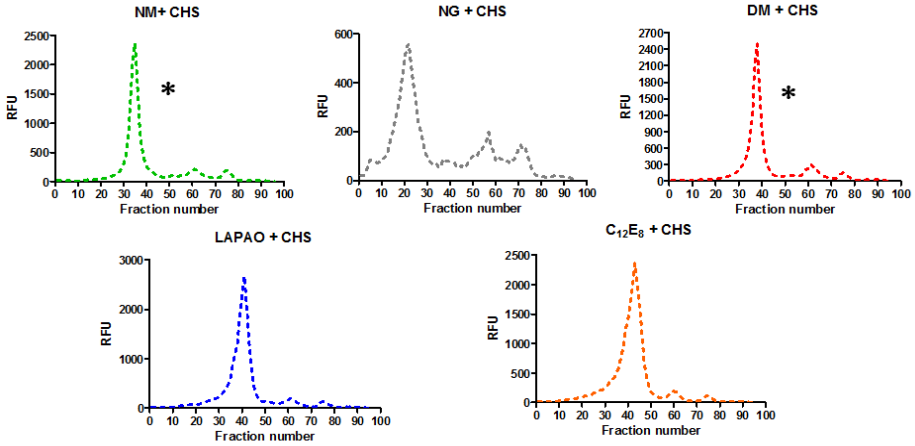
Figure 46. FSEC profiles showing the effect of cholesteryl hemisuccinate Tris salt (CHS) on the monodispersity of some eukaryotic light subunits of HATs. 4 different light subunits solubilized with 4 different detergents at 1% (w/v) were analyzed by FSEC without (in red line) or with 0.2% (w/v) CHS (blue dot line). Thereafter, the FSEC profiles were superposed to compare the impact of cholesterol on each protein. In all the conditions cholesterol addition improves the stability of the light subunit. RFU: Relative Fluorescence Units.

In Figure 47 the FSEC profiles corresponding to all the detergent/cholesterol conditions tested for the different proteins are presented. Conditions highlighted with an asterisk indicate that those conditions were considered good enough to pass to the next step of the project consisting on studying their behavior by size exclusion chromatography (SEC) after purification, tag removal and concentration.

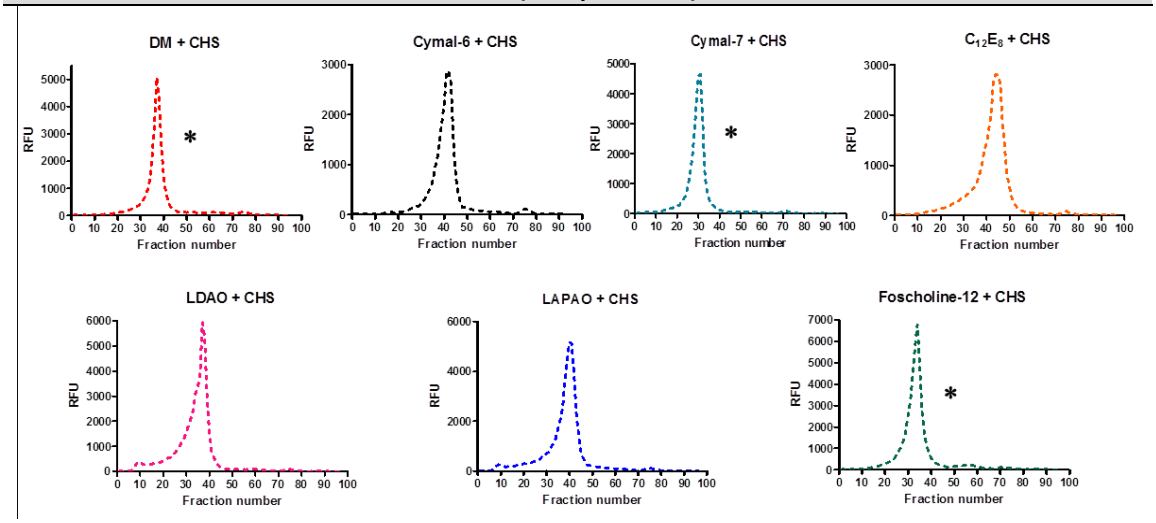
xCT (*Homo sapiens*)



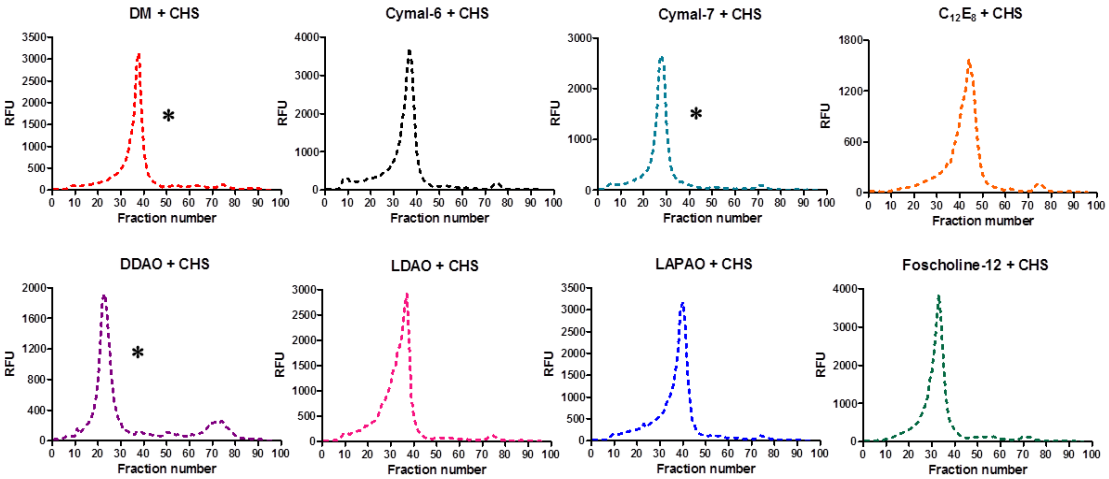
LAT2 (*Homo sapiens*)



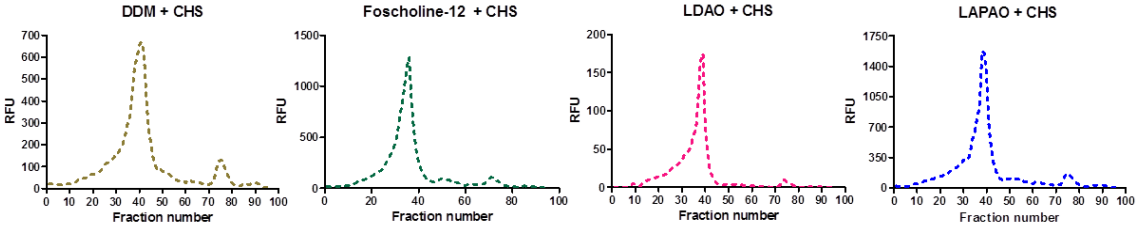
LAT2 (*Xenopus laevis*)



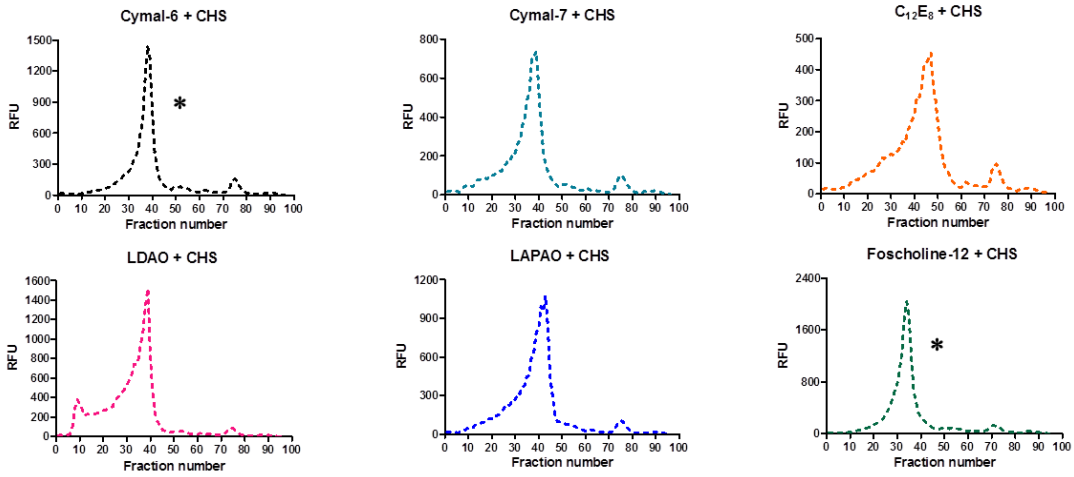
LAT2 (*Mus musculus*)



y$^+$LAT1 (*Homo sapiens*)



y$^+$LAT1 (Zebrafish)



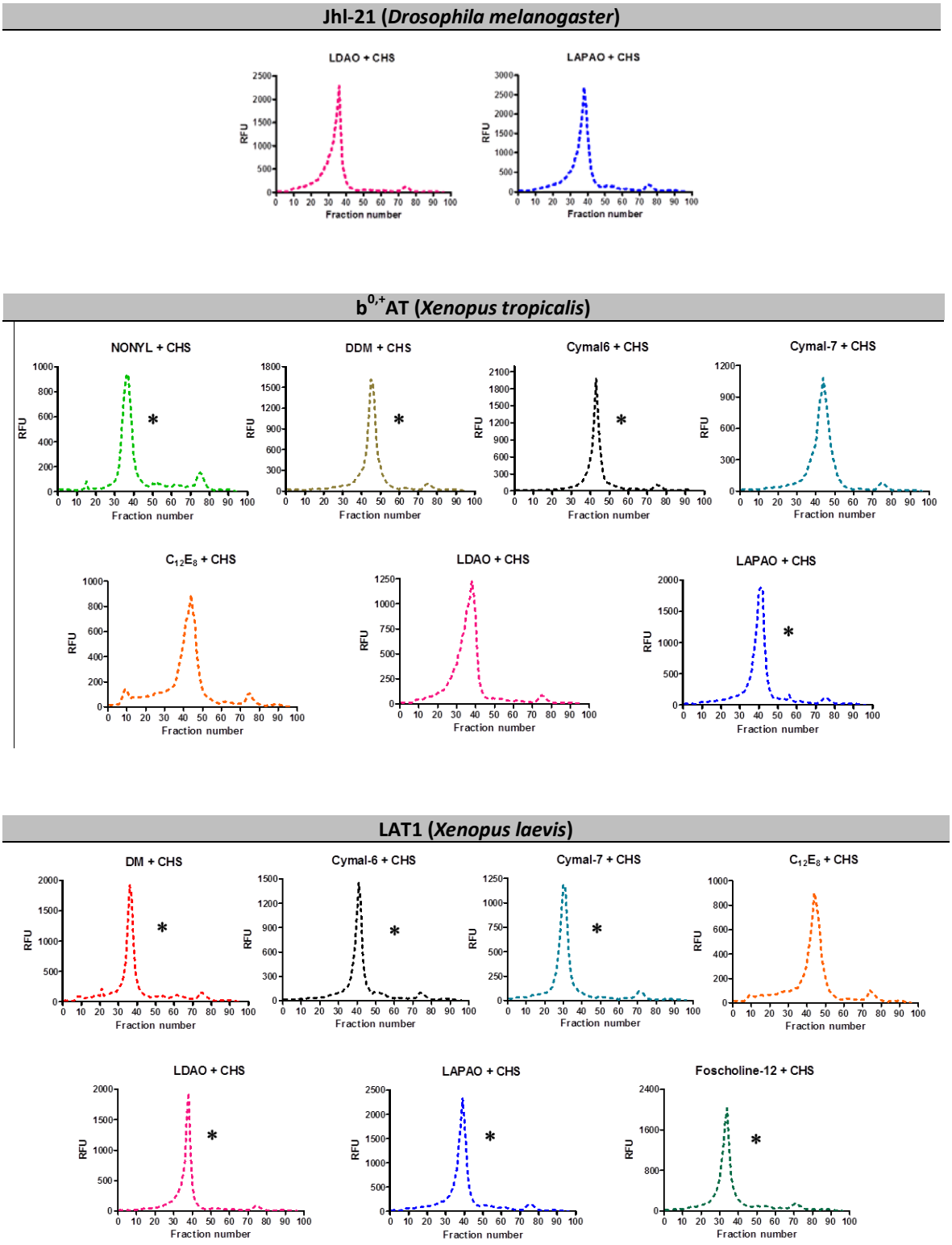


Figure 47. FSEC profiles of light subunits solubilized with the detergent/cholesterol conditions selected by UDS. All proteins were solubilized with 1%Detergent/0.2%Cholesterol at a final protein concentration of 3 mg/mL for 1 hour. Samples were ultracentrifugated at 250,000g for 1 hour and 500 µL were injected to Superose 6 previously equilibrated with 0.03% DDM. Detergent conditions marked with an asterisk were selected for the next step of the project.

Finally, 20 different detergent conditions and 6 light subunits were selected in this screening step, in which cholesterol was present in all of the conditions. These proteins were: LAT1 (from *Xenopus*), LAT2 (from *Xenopus*, human and mouse), y^+LAT1 (from Zebrafish) and $b^{0,+}$AT (from *Xenopus*). On the other hand, xCT (Human), Jhl-21(Drosophila) and y^+LAT1(Human) were discarded due to their non-monodisperse FSEC profiles. The effect of cholesterol stabilizing eukaryotic membrane proteins has been found in many proteins (Drew *et al.,* 2008).

II.6. Protein purification and size exclusion chromatography analysis

The 20 different conditions and the 6 different light subunits selected in the previous FSEC screening were overexpressed at large scale to screen their monodispersity by SEC after purification, tag removal and concentration for protein crystallization (see Chapter II: Materials & Methods, II.9). In all the cases it was used an equivalent start amount of total membrane protein ~250 mg, in order to follow the same protocol for the 20 conditions. The six light subunits were solubilized using 1% detergent and 0.2% CHS, and purified by Immobilized Metal ion Affinity Chromatography (IMAC) using nickel resin. Some steps were optimized to minimize protein loss. Firstly, 20 mM Tris, 150 mM NaCl buffer resulted better compared to others like PBS, but a further screening could be done. Protein binding was performed in batch, since it resulted in higher amount on attached protein than in gravity column. Washes were reduced to 2 without affecting protein purity, to avoid losing too much protein. Finally, 250 mM imidazole was enough to elute the protein reducing the time and protein manipulation when changing the buffer. Protein was eluted, concentrated and digested overnight with recombinant TEV protease at a 1:1 protein:protease ratio. TEV protease with a Hisx6 tag was produced and purified in the laboratory. Stocks at 2 mg/mL were stored at -80 °C (Figure 48) (see Chapter II: Materials & Methods, II.9.a).

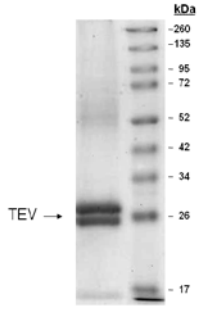


Figure 48. SDS-PAGE Comassie staining of recombinant His6-TEV protease produced in *E. coli* and purified by nickel affinity chromatography. A double band at 25-27 kDa is observed corresponding to the autoproteolytic activity of this TEV version.

Nevertheless, it was later replaced for a more efficient version of recombinant TEV produced by the Protein Expression IRB Core Facility, which cut 5-fold more the different light subunit-GFP fusion

proteins. It was noticed that the first recombinant TEV used was sensitive to imidazole; thus, in the previous digestions the buffer was thereafter replaced using a centricon of 100 kDa cut-off. In some occasions, low TEV digestion was observed presumably due to the presence of soluble protein aggregates. Afterwards, reverse IMAC was performed to eliminate the free GFP-His8 tag resulting from the protease digestion, as well as non-digested protein and the His6-TEV. During the purification the detergent, concentration was reduced to 3 times the critical micellelar concentration (CMC) of the corresponding detergent and the cholesterol amount was decreased as well to keep a 5:1 detergent/CHS ratio. The cleaved and purified protein was concentrated to ~ 4 mg/mL and then, injected into a Superdex 200 5/150 column (see Chapter II: Materials & Methods, II.9).

Finally, $b^{0,+}$AT, LAT1 and LAT2, all from *Xenopus* sp., showed promising SEC monodispersity profiles, being $b^{0,+}$AT the best profile when purified with DDM/CHS. The three best protein candidates with each corresponding best detergent condition are shown: $b^{0,+}$AT (DDM/CHS), LAT1 (Cymal-6/CHS) and LAT2 (Foscholine-12/CHS) in Figure 49.

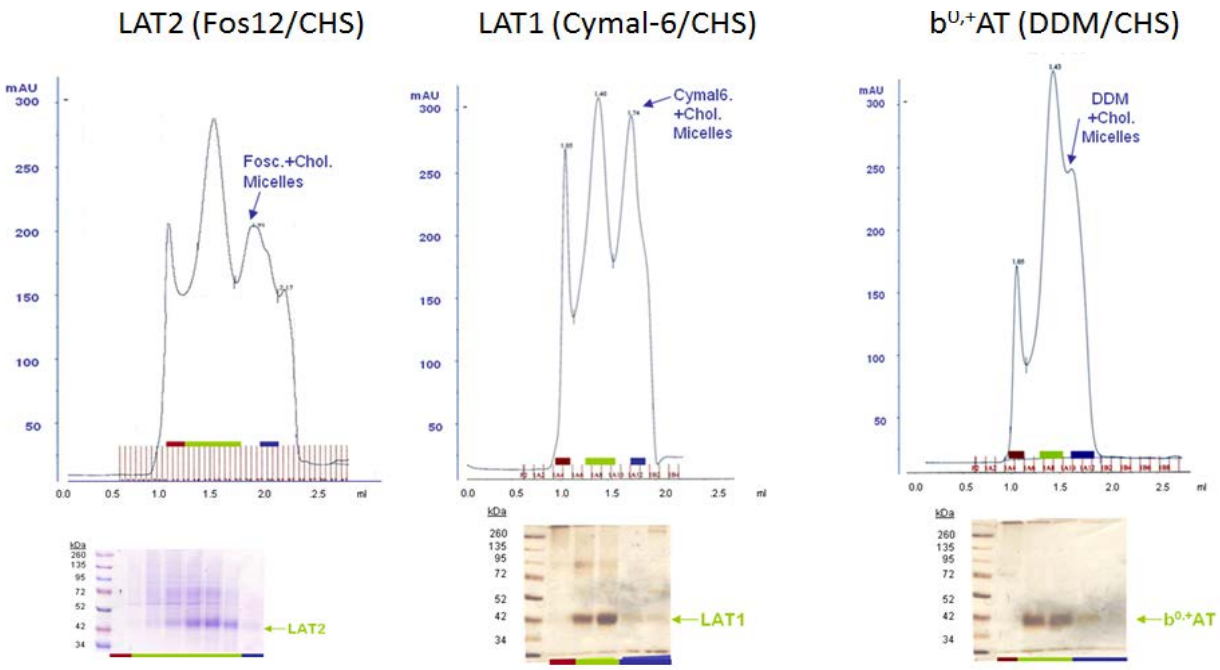


Figure 49. Size Exclusion Chromatography profiles and SDS-PAGE analysis of the best light subunits candidates after purification, tag removal and SEC monodispersity analysis using Superdex 200 5/150 column: LAT2 (*Xenopus laevis*) purified with foscholine-12 and CHS; LAT1 (*Xenopus laevis*) purified with cymal-6 and CHS; and $b^{0,+}$AT (*Xenopus tropicalis*) purified with DDM and CHS. (x axis: absorbance at 280nm (mAU); y axis: volume (mL)).

In all the cases, three different peaks were detected: the first one corresponding to the void (aggregated protein), the second one to the soluble transporter, and the last one to detergent/cholesterol micelles. It is worth mentioning that in those samples analyzed by tryptophan fluorescence size exclusion chromatography the third peak does not appear, indicating that it is not a

protein peak. The protein corresponding to the pure protein was quantified, and the yield was calculated per liter cell culture. In the case of $b^{0,+}$AT (DDM/CHS): 0.1 mg/L; LAT1 (cymal-6/CHS) 0.133mg/L and in the case of LAT2 (foscholine-12/CHS): 0.25 mg/L were obtained.

Although foscholine-12 was included in the screening since it is a common detergent used for this kind of tests and it was started to be used in crystallography screenings as well, we discarded this condition since Professor Gouaux, collaborating group (Vollum Institute, Portland), observed that this detergent inhibited the function of other transporters even though the peaks they observed were monodisperse during purification (personal communication, data not published). Other detergent conditions that could be taken into account for future detergent combination to improve the crystals resolution were: $b^{0,+}$AT with DM/CHS and LAPAO/CHS, and LAT1 with LAPAO/CHS (Figure 47C). The obtained protein yield were in this case: for $b^{0,+}$AT (DM/CHS): 0.07 mg/L, $b^{0,+}$AT (LAPAO/CHS): 0.133 mg/L, and LAT1 (LAPAO/CHS): 0.13 mg/L.

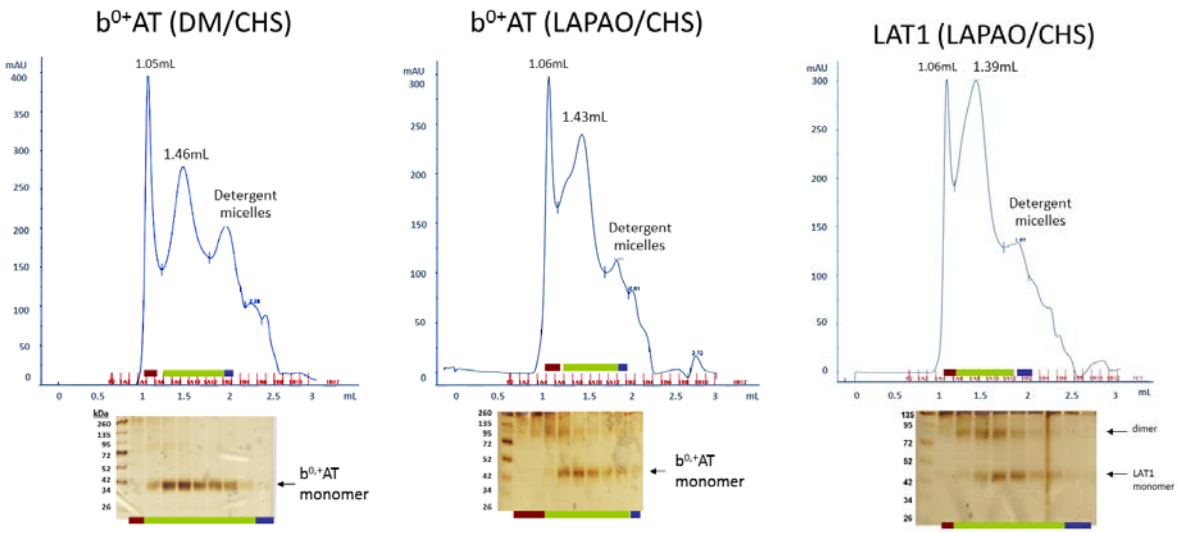


Figure 50. Other detergent conditions that could be taken into account. DM/CHS and LAPAO/CHS for $b^{0,+}$AT (*Xenopus tropicalis*), and LAPAO/CHS for LAT1 (*Xenopus laevis*). In all the cases in B) and C), 50 µL of purified and cleaved protein (~ 4 mg/mL) were injected into a Superdex 200 5/150 GL column equilibrated with the corresponding detergent and CHS. Proteins were detected by absorbance at 280 nm and peak collected and analyzed by SDS-PAGE (Coomassie and silver staining) (See Chapter II: Materials & Methods, II.9, and II.9.a).

As a result of the overall studies $b^{0,+}$AT (*Xenopus tropicalis*) was chosen as best candidate protein for crystallization studies after the screening of the different eukaryotic LATs. In addition, the detergent condition of DDM supplemented with CHS has been found, at this point of the study, the most promising condition for $b^{0,+}$AT. The first efforts to improve $b^{0,+}$AT monodispersity profile in the SEC were: firstly it was tried scaling up the purification in a Superdex 200 10/300 using the same conditions as in the previous assay. There is still more soluble transporter than aggregates, but there is still the third peak corresponding to detergent micelles and cholesterol (Figure 50).

To study the protein stability in less amount of detergent, it was tried to reduce from 3xCMC to 2xCMC during the purification process. Interestingly, the third peak disappeared (Figure 48 right), but the protein seems to aggregate more, since the void peak is higher than in 3xCMC.

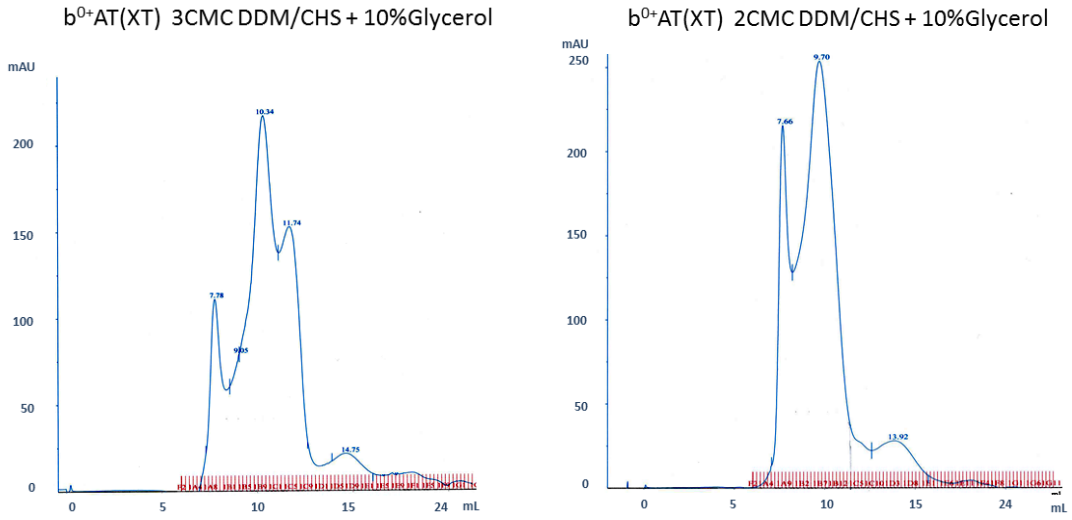


Figure 51. Best condition was scaled up and run into a Superdex 200 10/300 in the same purification conditions as it was done before (3xCMC) (on the left), or reducing the detergent amount during the purification to 2xCMC, (on the right) to check whether it is possible to reduce the detergent excess corresponding to the third peak in all the purifications using 3xCMC. Detergent reduction, required to avoid generating pseudocrystals in the cristalliaztion screenings produced by detergent excess, is achieved by reducing to 2xCMC. Membrane extract from 8 L *S.cervisiae* cell culture overexpressing $b^{0,+}$AT(*Xenopus tropicalis*) was solubilized with 1% DDM/0.2% CHS, then, divided to perform two purifications in parallel, one using 3xCMC DDM and adjusted the cholesterol: 0.027% DDM (w/v) and 0.0054% CHS (w/v); and the other purification detergent was reduced to 2xCMC of DDM and adjusted the cholesterol: 0.018% DDM / 0.0036% CHS. (x axis: absorbance at 280nm (mAU); y axis: volume (mL)).

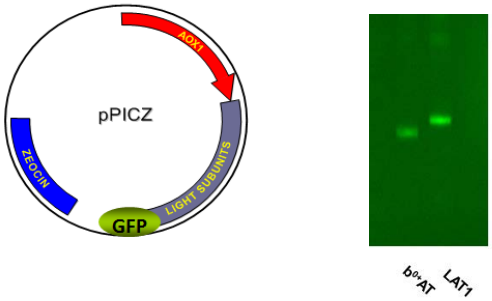
In addition, a small protocol change that gave notable difference was the buffer preparation, by advice of Dr. Rosell during his stay at Dr. Goaux laboratory. Sonication and filtration of the buffer previous to its use for protein solubilization and purification produced more homogeneous detergent micelles, improved protein solubility and higher monodispersity.

Even though the best conditions obtained analyzed by SEC are promising, they are still far from being used for crystallographic trials. One of the main reasons is that the protein expression should be higher in order to have from one batch of cell culture, sample enough so as to screen lots of crystallographic conditions, and be able to reproduce in big drop those that produce crystals. At the moment, to obtain 0.6 mg of pure $b^{0,+}$AT in DDM/CHS it is needed 6 L of cell culture, which is far from optimal. On the other hand, the SEC profile still shows a peak corresponding to protein aggregates that eluted in the void, indicating a tendency to aggregate, for that reason several strategies to improve stability were though and detailed in the next points.

II.7. Expressing the best light subunit candidates in *Pichia pastoris*

$b^{0,+}$AT (*Xenopus tropicalis*) and LAT1 (*Xenopus laevis*) were cloned by Dr. Rosell into pPICZ vector. In both cases a TEV protease site was introduced at the C terminus of the light subunits before yEGFP (Hisx8 tagged). The design of the fusion protein was the same as it was done for *S. cerevisiae*. Dr. Rosell and Dr. Costa transformed them into *Pichia pastoris* and selected the best overexpressing clones by whole-cell-fluorescence method as it was done for *S. cerevisiae* (see Chapter I: Materials & Methods, I.9, and Chapter II: Materials & Methods, II.4). Afterwards, they checked by In-gel-fluorescence the protein integrity (Figure 52A). Then, with Dr. Costa we overexpressed both proteins and purified them following the same protocol as it was done before to obtain the best candidate conditions (II.6, Figure 47, Materials & Methods, II.9). Minor modifications were adopted to perform the expression in *P. pastoris,* namely the solubilization detergent concentration which was increased to 2% detergent (w/v)/0.4%CHS (w/v), and an scale-up to run a Superdex 200 10/300 SEC column. The results indicated that even though the overexpression in *P. pastoris* was significantly higher (25,000 RFU and 12,000 RFU for $b^{0,+}$AT and LAT1 respectively, compared to the 2030 RFU and 1160 RFU obtained in *S. cerevisiae*), the quality of the protein was worse, since it aggregated more than the one expressed in *S. cerevisiae*.

A)



B)

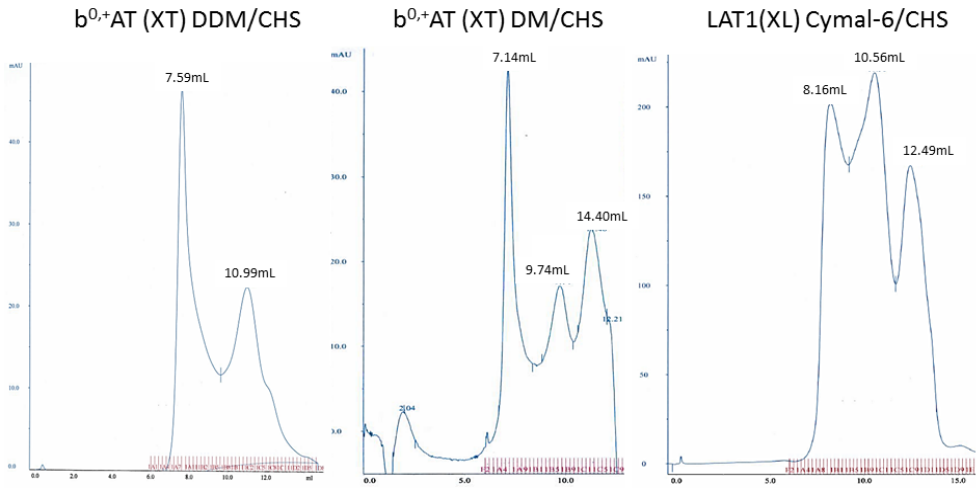


Figure 52. The 2 best light subunits selected in *S. cerevisiae* $b^{0,+}$AT (*Xenopus*) and LAT1 (*Xenopus*) were overexpressed in *P. pastoris* with GFP-His tag in C-terminal. Then, they were purified, removed their tag and SEC analyzed following the same procedure as it was done in *S. cerevisiae* using a Superdex 200 10/300 column (x axis: absorbance at 280nm (mAU); y axis: volume of the column (mL)). The purified protein elute between 9.74-10.99 mL (should be taken into account that the void volume value for each case) corresponds to a MW around 440 kDa (Ferritin in the calibration curve for the standard proteins on Tricorn Superdex 200 10/300).

II.8. Improvement of $b^{0,+}$AT stability for 3D crystallization by lipid addition

Since expressing the best light subunits in *P. pastoris* offered no advantage, *S. cerevisiae* was maintained as expression system to test the lipid addition effect on protein stability. Detergent-solubilized membrane proteins, even working with mild detergents, might not be stable enough and aggregate during concentration step. This problem might have the origin in the removal of key membrane lipids for this membrane protein which might let some hydrophobic areas of the protein not protected and then become aggregation start points. The addition of lipids in the solubilization step and during protein purification might mimic the protein membrane environment and reduce aggregation. Lipid requirement has been shown in many eukaryotic membrane proteins, for example, the voltage dependent K+ channel (Long *et al.,* 2007). Lipid screening allowed to check the stability for $b^{0,+}$AT expressed in *Saccharomyces cerevisiae* (Dr. Rosell, personal communication). In all cases the lipid was added during the solubilization and purification processes using the compounds and concentrations indicated in the Table 14.

Condition n°	LIPID	Solubilization	Purification
1	Asolectin	2 mg/mL	0.1 mg/mL
2	mC14PC	10 mM	0.5 mM
3	Brain Polar Lipid	0.05%	0.0025 %
4	POPC/POPG/POPE (3:1:1)	2 mM/0.7 mM/0.7 mM	0.1 mM/0.035 mM/0.035 mM
5	POPC/POPG/POPE (3:1:1)	1 mM/0.35 mM/0.35mM	0.05 mM/0.0175mM/0.0175 mM
6	POPC/POPE (2:1)	1 mM/0.5 mM	0.05 mM/0.025 mM
7	POPC/POPG (2:1)	1 mM/0.5 mM	0.0 5mM/0.025 mM
8	POPC/POPS (2:1)	1 mM/0.5 mM	0.05 mM/0.025 mM
9	POPC	2 mM	0.1 mM
10	POPE	2 mM	0.1 mM
11	POPS	1 mM	0.05 mM
12	POPG	1 mM	0.05 mM
13	DPPC	5 mM	0.25 mM
14	DPPC	2.5 mM	0.125 mM
15	DPPC	1 mM	0.05 mM
16	DMPC	5 mM	0.25 mM
17	DMPC	2.5 mM	0.125 mM
18	DMPC	1 mM	0.05 mM
19	DMPC/POPE (1:1)	1 mM/1 mM	0.05 mM/0.05 mM
20	DMPC/DPPC (2:1)	2.5 mM/1.25 mM	0.125 mM/0.0625 mM
21	DMPC/DPPC (2:1)	2 mM/1 mM	0.1 mM/0.05 mM
22	DMPC/POPC (1:1)	1 mM/1 mM	0.05 mM/0.05 mM
23	DMPC/DOPS (2:1)	2.5 mM/1.25 mM	0.125 mM/0.0625 mM
24	DPPC/DMPC (2:1)	2.5 mM/1.25 mM	0.125 mM/0.0625 mM
25	DPPC/DOPS (2:1)	2.5 mM/1.25 mM	0.125 mM/0.0625 mM

Table 14. Lipid conditions used to check $b^{0,+}$AT stability. $b^{0,+}$AT was solubilized from *Saccharomyces* membranes using 1%DDM/0.2%CHS and the corresponding lipid concentration (solubilization column). Solubilized protein was purified with 0.05%DDM/0.01%CHS and the indicated lipid concentration (purification column). POPC:1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine; POPG:1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoglycerol; POPE: 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine; POPS:1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine; DPPC: 1,2-dipalmitoyl-sn-glycero-3-phosphocholine; DMPC: 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine; DOPS: 1,2-dioleoyl-*sn*-glycero-3-phospho-L-serine.

To perform this screening, $b^{0,+}$AT was solubilized from *Saccharomyces* membranes and purified by nickel affinity chromatography. It was solubilized with DDM/CHS without lipid or together with the different lipid combinations (Table 14), then, kept at room temperature or at 4 °C for 16 h and analyzed by FSEC.

Interestingly, a clear improvement of the $b^{0,+}$AT stability checked by FSEC monodispersity was observed by addition of DMPC or DPPC in the solubilization and purification processes in comparison with the FSEC profiles corresponding to the purified transporter only in presence of DDM/CHS (Figure 53) when the purified samples were kept at 4 °C. The two best conditions (DMPC and DPPC) and two other conditions representatives of the non-working conditions (Asolectin and m14PC) are shown in Figure 53. Under more severe conditions (at room temperature) the effect was higher in the case of DPPC addition. This lipid was chosen to be included in the solubilization and purification of $b^{0,+}$AT.

Effect of lipid addition on $b^{0,+}$AT (*Xenopus*) stability

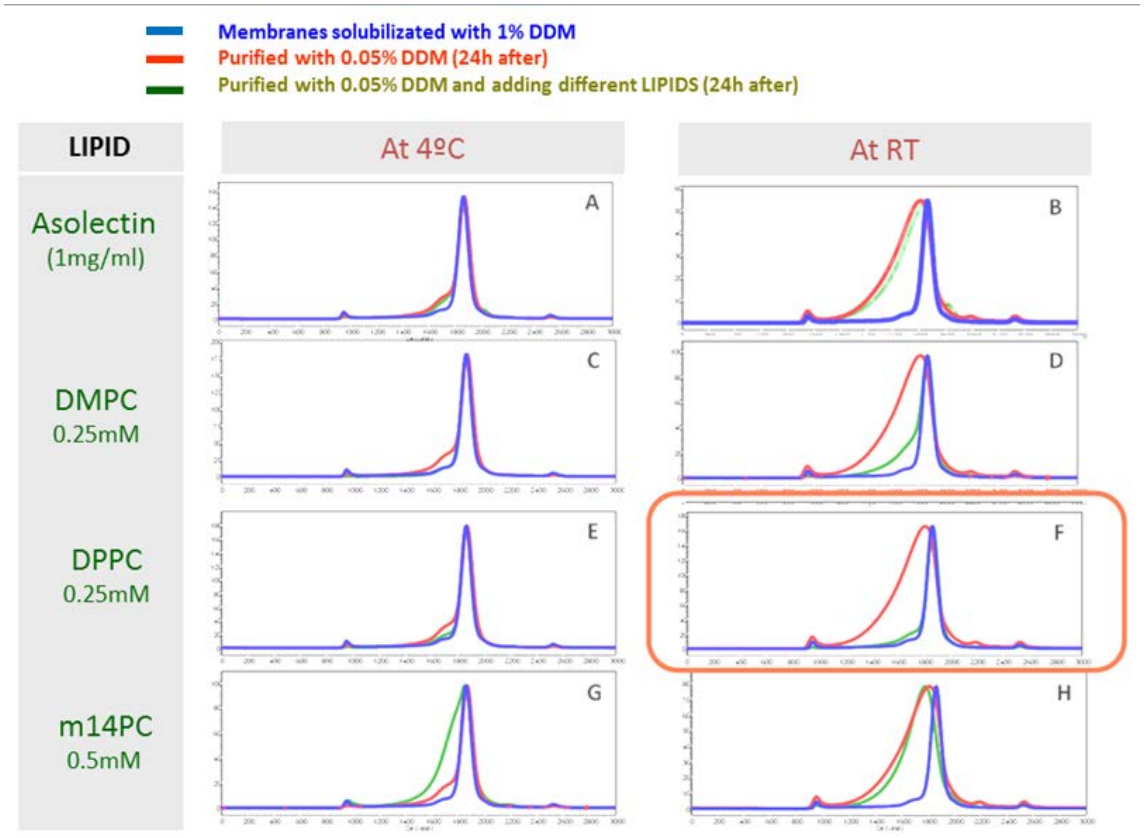


Figure 53. FSEC profiles showing the effect of different lipids on the monodispersity of purified $b^{0,+}$AT-GFP. $b^{0,+}$AT-GFP was solubilized with DDM/CHS (colored in blue), solubilized and purified with DDM/cholesterol (colored in red) and solubilized and purified with DDM/cholesterol and lipid (Asolectin (1 mg/mL), DMPC 0.25 Mm, DPPC 0.25 mM or m14PC 0.5 mM). Purified protein was kept at room temperature or at 4^0C for 16 h, and 100µL of each condition were injected into a Superose 6 column previously equilibrated with DDM 0.03% for FSEC analysis.

II.9. Expression and purification of $b^{0,+}$AT in insect cells

The project developed in *Saccharomyces* was very useful since it allowed us to identify $b^{0,+}$AT (from *Xenopus tropicalis*) as a putative candidate for 3D crystallography, and optimize its stability requirements (detergent, cholesterol and lipids). The problem using this expression system was that the final purification yield of $b^{0,+}$AT produced in *S. cerevisiae* was 0.1 mg of purified protein per liter of yeast culture. In order to improve the expression levels of this transporter, $b^{0,+}$AT was expressed in insect cell (Sf9 cell line) (see Chapter II: Materials & Methods, II.11-12). In addition, this expression system has been more successful in producing eukaryotic membrane proteins that ended with a solved crystal structure, than it has been *S. cerevisiae*. Actually, the 43% of solved membrane protein structures were obtained using baculovirus as expression system (Parker and Newstead, 2014).

$b^{0,+}$AT cDNA was cloned into the expression vector pFastBac, as a fusion protein with a GFP-His8 C-terminal tag and a thrombin protease site between the light subunit and the tag. DH10 *E. coli* strain was transformed with the generated construct and bacmid generated by transposition. Sf9 cells were transfected with the purified bacmid containing $b^{0,+}$AT-GFP-His8 and the first population of baculoviruses (P1) was obtained after 96 h, which was used to infect Sf9 cells and generate the P2 after 96 h post-infection. These baculoviruses were titrated by the end point dilution assay and used to infect 2 L of Sf9 cells culture (3.10^6 cells/mL) at a MOI of 2 (see Chapter II: Materials & Methods, II.12.a-13).

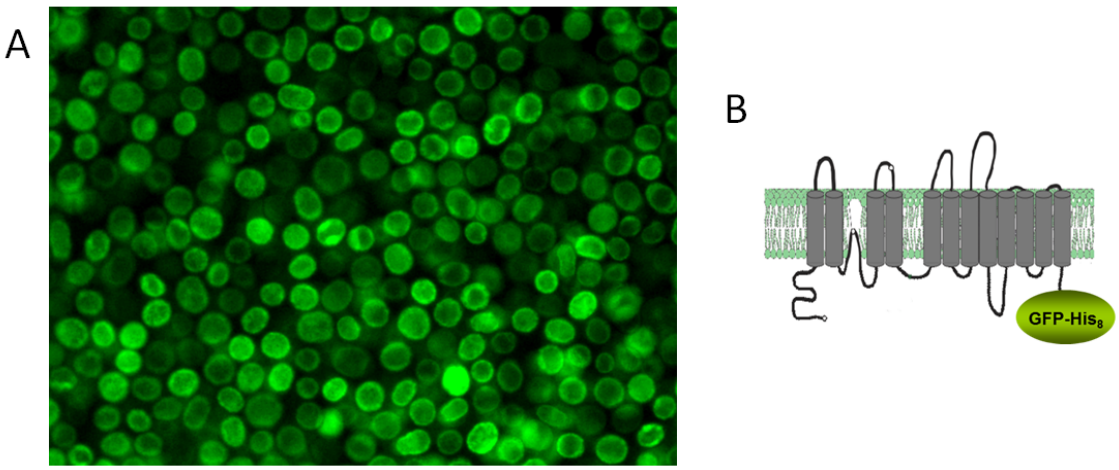


Figure 54. Expression of $b^{0,+}$AT-GFP in Sf9 insect cells. A) Checking of $b^{0,+}$AT-GFP production in Sf9 insect cells by fluorescence microscopy. **B**) Design of the recombinant $b^{0,+}$AT-GFP expressed in insect cells.

Sf9 cells were disrupted and membranes isolated 48 h post-infection. Protein was solubilized with the conditions that were optimized for *S. cerevisiae* (1% DDM/ 0.2% CHS/ 5 mM DPPC). Protein was purified by affinity chromatography (Ni^{2+} resin), digested with thrombin, concentrated until 4 mg/mL and analyzed its monodispersity by tryptophan fluorescence size exclusion chromatography (Trp-FSEC)

(see Figure 55). Promisingly, the stability conditions identified in *Saccharomyces* were reproduced on the protein overexpressed in insect cells. In addition, the final expression yield was improved from the one obtained in *S. cereviaae* which was 0.5 mg per liter of cell culture. Thus, we accomplished the objective to improve the final purification yield just with changing the expression system.

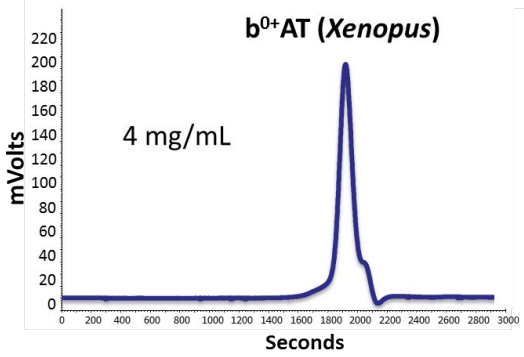


Figure 55. Tryptophan FSEC of $b^{0,+}$AT wild-type overexpressed and purified from Sf9 insect cells using DDM/CHS and DPPC selected conditions. A Superose 6 column was used in this case.

Trials to co-express the heavy subunit, rBAT (*X. tropicalis*), with $b^{0,+}$AT (*X. tropicalis*) were not successful, even trying several different expression systems (results by Dr. Rosell during his stage at Dr. Gouaux laboratory). For this reason we continued working on the improvement of $b^{0,+}$AT stability with the aim to crystallize it alone.

II.10. Improvement of $b^{0,+}$AT molecule by mutagenesis and generation of truncated versions

- Generation of the C150A $b^{0,+}$AT mutant (removing reactive cysteines)

The conserved cysteine involved in the disulphide bridge between the heavy and light subunit in HATs is facing the aqueous media, which makes it highly reactive, fact that can compromise the stability of the purified $b^{0,+}$AT. For this reason, the Cysteine 150 was mutated into Alanine, which was generated by site-directed mutagenesis (Figure 56) (see Chapter I: Materials & Methods, II.10.b).

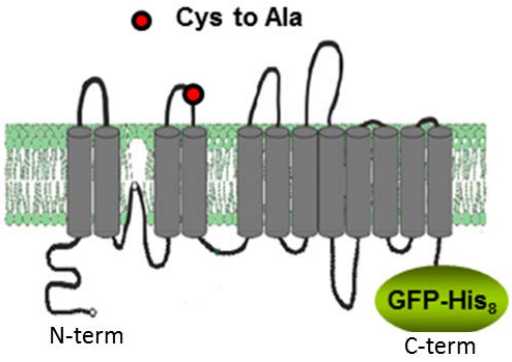


Figure 56. Scheme of $b^{0,+}$AT C150A design locating the reactive Cysteine 150 in the external loop II as a red circle. The GFP-Hisx8 tag is indicated at the C-terminal.

- Generation of truncated versions of C150A $b^{0,+}$AT mutant

In order to decrease the flexibility of the protein and improve the probability to get high quality crystals, different truncated forms of C150A $b^{0,+}$AT in N-terminus or in C-terminus or in both were generated taking into account the predicted topology of the protein (Figure 57). N-terminal Δ10 and Δ27, and C-terminal Δ28, Δ18, Δ13, Δ8 single truncation forms were generated on the C150A $b^{0,+}$AT. Initially, the single truncations were cloned into pFastBac vector and the respective viruses were generated (see Chapter II: Material & Methods, II.12.a-e).

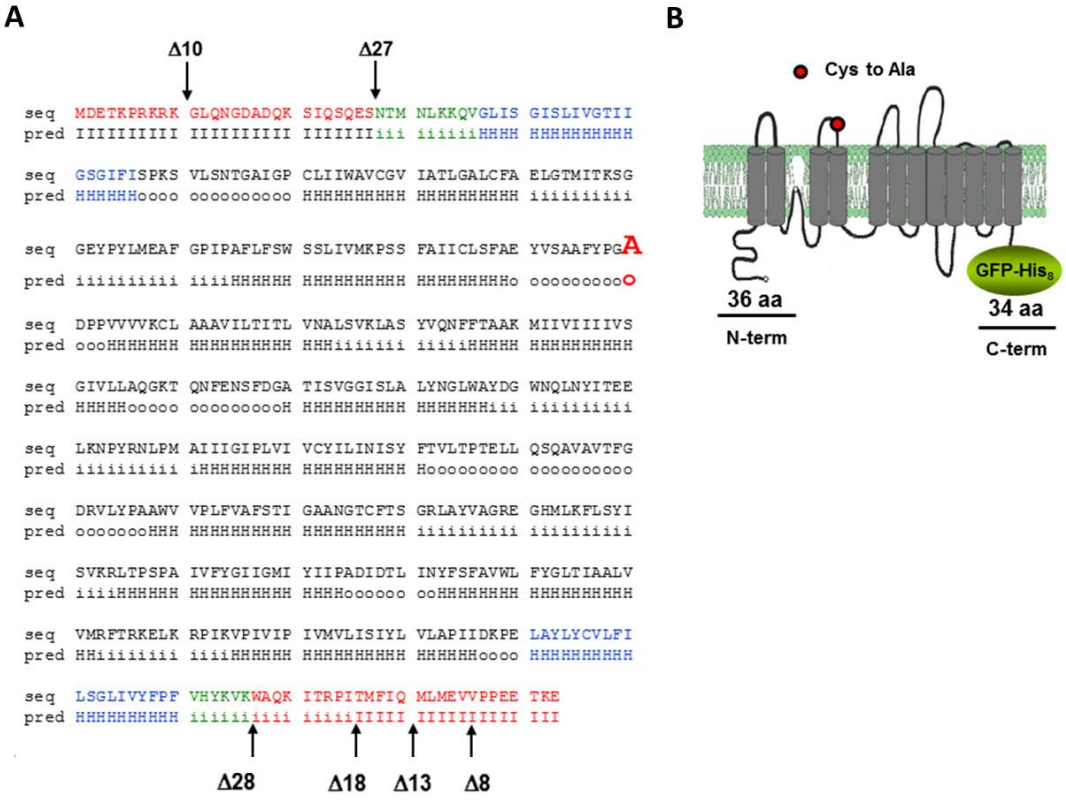


Figure 57. Amino acid sequence and predicted topology of $b^{0,+}$AT. A) The different truncations generated in N-terminal (Δ10 and Δ27) and in C-terminal (Δ28, Δ18, Δ13, Δ8) are indicated on the amino acid sequence. B) Scheme of the $b^{0,+}$AT topology indicating the N-terminal and C-terminal (previous GFP-His tag) amino acid lengths.

Then, 40 mL cell cultures of the different truncated versions cloned into Sf9 insect cells were grown and checked the protein expression and protein integrity by in gel fluorescence 48 h post-infection (Figure 58). Interestingly, in all the cases overexpression was observed with no sign of proteolysis.

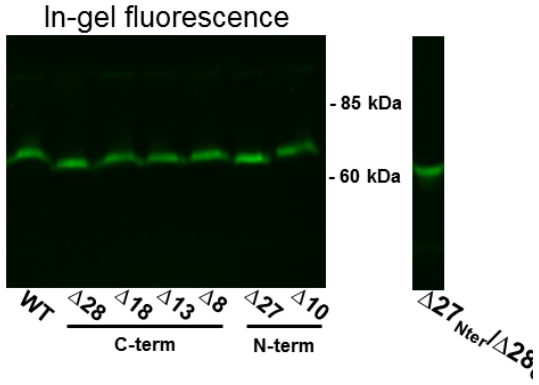


Figure 58. In-gel fluorescence of the different single truncation versions generated compared with the wild-type (WT) (on the left), and the double truncation Δ27Nterm/Δ28Cterm (on the right) generated on a C150A background. The molecular weight of the wild-type fusion protein corresponds to ~70 kDa (~40 kDa of the light subunit + 28 kDa of the GFP). Truncation forms run under the WT molecular weight as expected. In gel fluorescence showed good expression of all the forms compared with wild-type.

C150A $b^{0,+}$AT single truncated versions were solubilized with 1%DDM/0.2%CHS/5mM DPPC and their monodispersity analyzed by FSEC (Figure 59) (see Chapter II: Material & Methods, II.14-15). Since the longest single truncations Δ27Nterm and Δ28Cterm were possible to be overexpressed and they were monodisperse (Figure 56) it was decided to generate the double truncation form Δ27Nterm/ Δ28Cterm on C150A $b^{0,+}$AT. It was also cloned into pFastBac and generated its viruses. Thereafter, the expression and protein integrity was examined by in-gel fluorescence, observing that the expression levels comparable to the $b^{0,+}$AT wild-type and without observing proteolysis (Figure 55).

Furthermore, the monodispersity of the double mutant C150A $b^{0,+}$AT (Δ27Nterm/ Δ28Cterm) was analyzed by FSEC (Figure 59), but unfortunately the behavior observed was not as good as the single truncations, thus, it was discarded. Comparing the single truncated C150A $b^{0,+}$AT Δ28Cterm and C150A $b^{0,+}$AT Δ27Nterm monodisperisity profiles, the first one behaved better and it was the one selected for the first 3D crystallization trials.

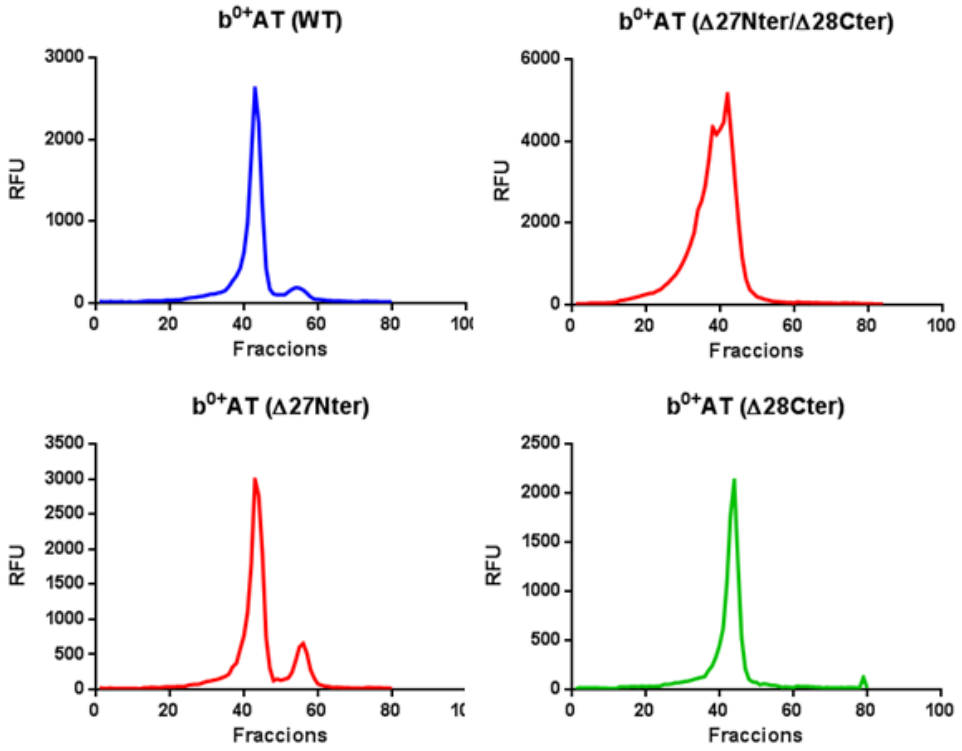


Figure 59. FSEC profiles of the truncated forms of C150A $b^{0,+}$AT compared with wild-type with the selected detergent-lipid conditions: DDM/CHS/DPPC. Δ28Cterm truncated form showed the best monodisperse profile, and thus, it was selected for 3D crystallization trials.

II.11. First 3D-crystallization trial of C150A $b^{0,+}$AT (Δ28 Cter)

The C-terminal truncated form Δ28 C150A $b^{0,+}$AT-GFP-Hisx8, previously selected, was overexpressed and purified for crystallization screening and 3D structure determination. Overexpression was done for 48 h post-infection at a MOI of 2 and a total of 3 liters of Sf9 cell culture were grown (see Chapter II: Materials & Methods, II.16.a). Isolated membranes with the overexpressed protein were solubilized using the previously optimized (see II.8 and table 14, condition 13) solubilization buffer (SB) [1% w/v DDM; 0.2% w/v CHS; 5mM DPPC in 20 mM Tris (pH 8.00), 150 mM NaCl].

The solubilized protein was purified by affinity chromatography with nickel resin (see Chapter II: Materials & Methods II.16.b). During purification, the detergent, cholesterol and lipid were reduced to 0.05% w/v DDM; 0.01% w/v CHS; 0.25mM DPPC in 20 mM Tris (pH 8.00), 150 mM NaCl and 25 mM or 250 mM imidazole were added in the washing or the elution buffers. The 9 mL of eluted protein were concentrated until 1.2 mL using a centrifugal filter of 100 nominal molecular weight limit (NMWL). Then, the protein was digested overnight at 4 °C with thrombin at 1:100 ratio (thrombin:fusion protein) to remove the GFP-Hisx8 tag. Then, it was concentrated to ~ 8 mg/mL and aggregates were eliminated by

high-speed centrifugation conditions. 500 µL of supernatant were injected into a Superdex 200 10/300 GL column (this last step was done twice because we had 1000 µL of sample). Purified, cleaved and concentrated truncated form C150A $b^{0,+}$AT Δ28 C-terminal was run into a SEC (Figure 60). The peak corresponding to the collected fractions 11-19 which correspond to 8.2 mL-9.8 mL was collected and dialyzed overnight at 4°C to remove the excess of detergent against Tris 50mM, NaCl 150 mM, DDM 0.05 %, CHS 0.01%, DPPC 0.25 mM with a 50 kDa cut-off.

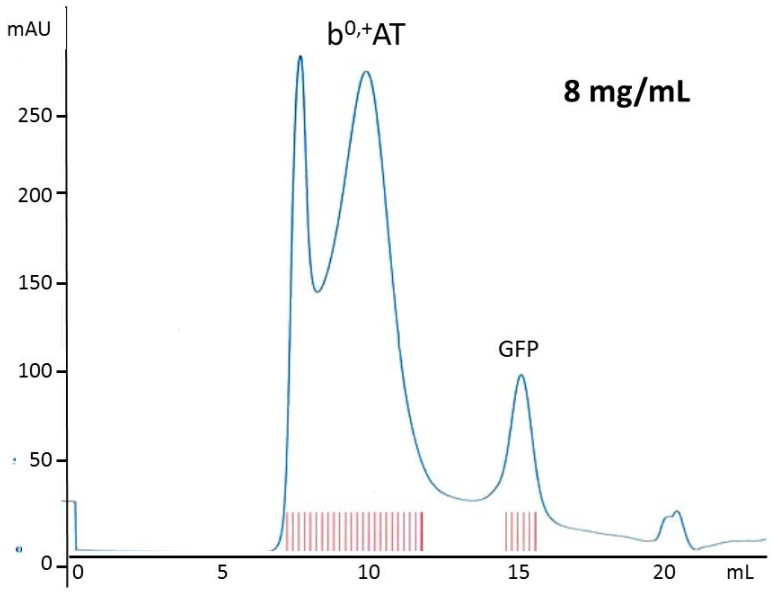


Figure 60. SEC profile of purified, cleaved and concentrated C150A $b^{0,+}$AT Δ28 C-terminal produced in Sf9 cells. 500 µL of purified protein and concentrated at 8mg/mL were injected into a Superose 6 column previously equilibrated with DDM 0.05%. The fractions between 11 and 19 which correspond to the peak and the volume 8.2mL-9.8mL were collected. Then, dialyzed and concentrated at 4.2 mg/mL and seeded for crystallization screening. The graphic represents the Absorbance at 280nm (y axis) and the volume of the column in mL (x axis).

Next day, 0.6 mg of protein at 1 mg/mL were concentrated until 4.2 mg/mL and first 3D-crystallization screening was performed (see Chapter II: Material & Methods II.16.b). There are different crystallization methodologies, but the most common one was firstly assayed, namely the vapor diffusion and sitting drop. This approach has been proven to be the most successful in crystallizing membrane proteins, even though it was originally designed for soluble proteins.

In the first trial the Memgold crystallization conditions were tested. The first crystallization trial ended with some conditions forming precipitates or pseudo-crystals as the showed in Figure 61. These pseudo-crystals were analyzed at Alba synchrotron but they did not diffract. Some possible reasons are: the excess of lipid and detergent, and the size of the detergent used, DDM, which is not the optimal to obtain high resolution structures.

Crystallization screening: MemGold plate, 20°C

Figure 61. Crystallization screening by hanging drop vapour diffusion at 20 °C using the MemGold pack screening. A drop ratio of 0.5 µL of protein (4.2 mg/mL concentrated) with 0.5 µL reservoir solution was used. Spherolites (detergent precipitates) were obtained in some conditions like the indicated on the left (0.1 M Ammonium Sulfate, 0.1 M Glycine pH 3.8, 28% (w/v) Triethylene Glycol), and others produced pseudo-crystals (0.1 M Calcium Chloride, 0.1 M TRIS pH 6.5 13% PEG (w/v) MME 2000) that did not diffract in X-Ray.

These are preliminary results, since not an exhaustive crystallization screening has been done. Usually, several commercial screening conditions are tried, like Membfac, Memplus, Memstart, Memsys and Memgold which are adapted to automatic sitting-drop seeding nL scale. Also, crystallization is tried at different temperatures, at 4 and 20 °C.

The purification process in the future should be optimized by reducing lipid and cholesterol concentrations previous to the next crystallization trials. Or, taking advantage of the lipid contained in the sample, other crystallographic techniques could be tried which are specific for membrane proteins: like bicelles (Faham *et al.,* 2002; Agah and Faham, 2012), HiLiDe (Gourdon *et al.,* 2011), or Lipidic Cubic Phase (Caffrey *et al.*, 2009).

In addition, replacement of DDM for shorter detergents in the last steps of the purification (during washes, elution or during SEC) can be tested since they would form smaller micelles around membrane proteins that would allow to solve in a higher resolution the protein structure.

To sum up, as a result of this screening and optimization process, a more stable light subunit than LAT2 for crystallization trials has been selected.

II. Discussion

II. Discussion

- **Selection of eukaryotic light subunits for crystallization studies**

The adopted GFP-based *S. cerevisiae* pipeline (Drew *et al.,* 2008) allowed to select three candidate light subunits for crystallization studies: $b^{0,+}$AT, LAT1 and LAT2. The three proteins belonging to *Xenopus* spp, are more stable than human LAT2 overexpressed in *P. pastoris* (Chapter I).

The three selected candidates were obtained by solubilization and purification with DDM/CHS in the case of $b^{0,+}$AT (*X. tropicalis*), cymal-6/CHS in the case of LAT1 (*X. laevis*), and foscholine-12/CHS for LAT2 (*X. laevis*). The foscholine-12 condition was discarded by recommendation of Dr. Gouaux (Vollum institute, Oregon) that was collaborating in the project, since in many cases the purified protein resulted not functional.

The quality of the proteins, analyzed by size exclusion chromatography (SEC), showed in the three cases a main peak corresponding to the purified protein, the void peak corresponding to the protein aggregates and a third peak corresponding to micelles of detergent and cholesterol (Figure 47). The protein with best SEC profile was $b^{0,+}$AT condition, and followed by LAT1. Even though $b^{0,+}$AT and LAT1 were selected as good targets, this thesis is mainly focused on the study of $b^{0,+}$AT.

- **Keys of the success in finding the good candidates**

The high-throughput method (Drew *et al.,* 2008) allowed fast screening of the overexpression levels, the quality of the recombinant protein and monitoring of the purification steps. *S. cerevisiae* cloning resulted less time-consuming than other eukaryotic expression systems like *P. pastoris* or Insect cells. Conversely, these two expression systems usually produce larger amounts of protein than *S. cerevisiae*, being the main reason to test them after having selected the candidate protein.

In addition, the GFP technology allowed quickly quantifying the proportion of soluble protein, measure the stability of the different detergent conditions in a fast way using the ultracentrifugation dispersity sedimentation (UDS) assay and FSEC to analyze the monodispersity of a solubilized-protein without the need to be purified (Kawate and Gouaux*,* 2006).

Nevertheless, once the optimal conditions of purification would have been set up for crystallization trials, the candidate protein could be cloned without GFP, and check whether the stability was the same. Seldom is the case where GFP impairs the function in a protein, but it could happen. On the other hand, some groups have been successful crystallizing their proteins with the Histidine tag, which would reduce the manipulation of the protein, hence reducing steps where the protein can aggregate.

- **Detergent conditions:**

DDM detergent condition selected for $b^{0,+}AT$ purification, is the most successful used in crystallization (Yamashita *et al.*, 2005; Kawate *et al.,* 2009; Lee *et al.*, 2014; Althoff *et al.,* 2014; Baconguis *et al.,* 2014). Nevertheless, due to its long carbonated tail, sometimes the X-ray diffracted crystals cannot be solved at a very high resolution; in those cases, changing the detergent during the purification process for shorter-length ones like OG or NG could rise the probability to increase the protein resolution (Figure 62) (an example is the protocol followed to screen replacement detergents for proteins like pirophosphatases (Kellosalo *et al.,* 2011); or the M2 muscarinic acetylcholine receptor (Haga *et al.,* 2012).

In the case of the LAT1 purification condition, Cymal-6 is a shorter detergent than DDM. Interestingly, it was used to determine the structure of the prokaryotic homologue of LATs, AdiC (Kowalczyk *et al.,* 2011) (see Introduction 3.b.2).

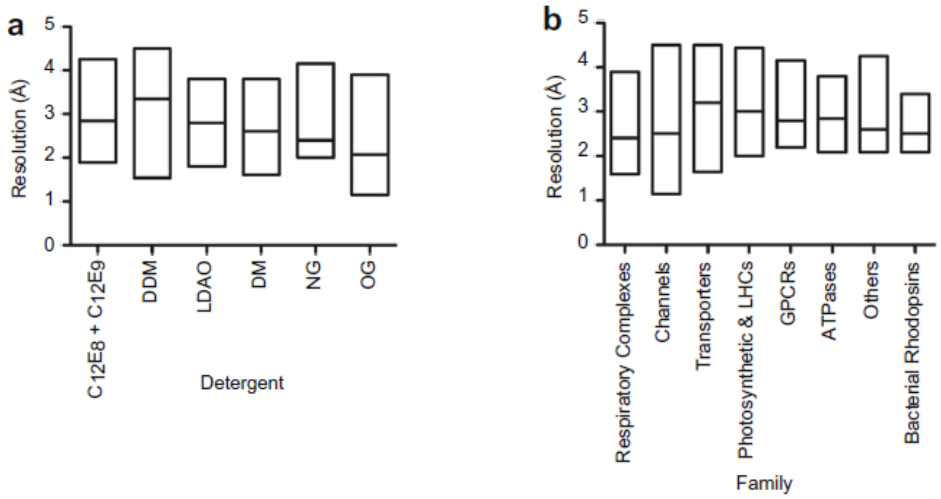


Figure 62. Graphics indicating the correlation between structural resolution and the type of crystallization detergent. The information to build the graphics was obtained from a database generated by collecting the crystallization information from all of the available unique alpha-helical MP structures in the PDB (in 2010). Only structures crystallized using a single defined detergent were included. This task was greatly facilitated by the 'Membrane Proteins of Known 3D Structure' web site from the Stephen White laboratory at UC Irvine (http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html). **A)** Square plots displaying the structural resolution for seven of the most successful crystallization detergents; the middle line in each plot represents the median. **B)** Deposited structural resolution for the different membrane protein family members (from Sonoda *et al.,* 2010).

In addition, new detergents like the MNG-2/3 could be tested, since recent structures solved used this detergent: the *Xenopus laevis* NMDA receptor (Lee *et al.,* 2014), or the β2 adrenaline receptor (Roenbaum *et al.,* 2011).

- **Cholesterol addition improves protein stability:**

Really interesting is the stabilizing effect of cholesterol addition (in the salt form of cholesteryl hemisuccinate Tris salt, CHS) on almost all light subunits tested. FSEC profile of the solubilized protein was significantly better when CHS was included in the solubilization step (Figure 45). This stabilizing effect was previously described for many eukaryotic membrane proteins (Drew *et al.,* 2008; Jones *et al.*, 2012) and it has successfully been used for purification and crystallization of GPCRs (Hanso *et al.,* 2008; Lee *et al.,* 2014). Cholesterol plays well established regulatory roles in many membrane proteins, either indirectly by modulating membrane properties, either by interacting directly by specific interactions with the membrane proteins (Burger *et al.,* 2000; Lee, 2004).

Indeed, it has been found a cholesterol binding site in some membrane proteins structures like the β2-adrenergic receptor (Hanson *et al.,* 2008) which showed that these receptors organize as monomers in an antiparallel association with two distinct cholesterol molecules bound per receptor. In this precise example, thermal stability analysis using isothermal denaturation confirmed that cholesterol analog significantly enhanced the stability of the receptor (Hanso *et al*., 2008).

But not only cholesterol is widely described to bind GPCR proteins, but also it has been found bound in the structure of eukaryotic transporters like the Drosophila dopamine transporter DAT (Penmatsa *et al.,* 2013). In fact, it was previously described that the conformation and function of this transporter can be affected by manipulating membrane cholesterol. When treating cells with methyl-β-cyclodextrin (mβCD), which depletes membrane cholesterol affecting raft and non-raft membrane domains, the dopamine uptake and efflux rate were reduced. They discarded that it was the raft localization because they performed an experiment to disrupt raft-localized DAT by using nystatin cholesterol chelating agent, and no effect was observed, whereas repletion with the analog desmosterol, which is a non-raft promoting sterol, restored the transport rates (Jones *et al.*, 2012). We expect to find specific interactions between cholesterol and HATs, since they have been associated with cholesterol as well, at least, in the lipid rafts (Schroeder *et al.,* 2012).

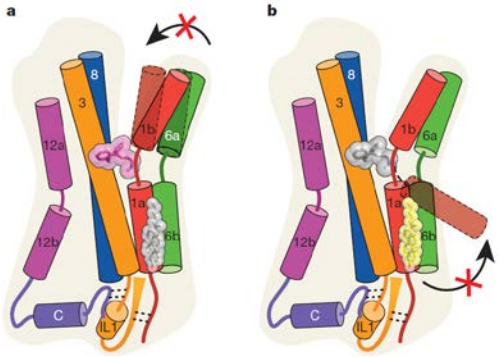


Figure 63. **Cholesterol bound in the solved structure of DAT transporter (Drosophila)** (from Penmatsa *et al.,* 2013).

In the case of human β2-adrenergic structure (Cherezov *et al*., 2007), in addition to cholesterol, it was also found palmitic acid bound to the protein, which may indicate that both are needed to

regulate the protein function or the interaction with other membrane proteins. The turkey β1-adrenergic G protein-coupled receptor structure was solved mutating the Cys residue responsible to bind palmitoil at the C-terminal, to reduce protein aggregation rate (Warne et al., 2010).

Many membrane proteins that interact with cholesterol have a cholesterol recognition / interaction amino acid consensus (CRAC) sequence like the GPCRs: β2-adrenergic receptor, serotoninA receptor or rhodopsin. The CRAC sequence consists on a sequence of: -L/V-(X)1–5-Y-(X)1–5-R/K-, in which (X)1–5 represents between one and five residues of any amino acid (Li *et al.,* 1998; Epand *et al.,* 2006; Gimpl *et al.,* 2010). The future study of these motifs in HAT transporters could be useful to design mutants that lose the interaction with cholesterol with the aim to achieve more stable proteins.

On the other hand, molecular-dynamics simulations predict that dimerization (calculated for the β2-adrenergic receptor) is modulated by membrane cholesterol. It has been found that cholesterol binds to transmembrane helix IV of this receptor blocking its involvement at the dimer interface. Dimer plasticity in many membrane proteins is relevant for its function (Prasanna *et al.,* 2014).

Since $b^{0,+}$AT (*Xenopus tropicalis*) was the best candidate, it was selected to improve its monodispersity profile. The scale-up into a larger SEC (Superdex 200 10/300) using the same conditions reproduced the three peaks observed before, and did not show a clear separation of the third peak, the one corresponding to micelles of detergent and cholesterol, and it also had a void peak corresponding to aggregated protein (Figure 48, on the left). For that reason, reduction of detergent/CHS from 3xCMC to 2xCMC during the purification process was tried and the third peak corresponding to detergent disappeared. But, the protein peak was displaced to the left, indicating a different oligomeric state of the protein (Figure 48, right).

- **Lipid addition to improve protein stability**

Solubilized-proteins tend to be less stable in detergent micelles than in the lipid bilayer. For this reason, it has become a more common practice during the last years the incorporation of synthetic lipids during purification and crystallization which can improve the quality of the membrane protein crystals allowing them to diffract at high resolution (Lemieux *et al.,* 2003; Zhang *et al.,* 2003; Yao *et al.,* 2005; Guan *et al.,* 2006; Hattori *et al.,* 2007; Sobolevsky *et al.,* 2009; Hibbs and Gouaux, 2011; Krishnamurthy and Gouaux, 2012). In our case, the addition of lipids to the solubilization and purification buffers has enormously improved the stability of the protein after purification -even at room temperature-, which is really advantageous for crystallization. After screening 25 different lipid conditions DPPC 0.25 mM was the best condition for $b^{0,+}$AT (results obtained by Dr. Rosell tested in his stay at Dr. Gouaux laboratory). In addition, the presence of lipids is a prerequisite for protein functionality, as demonstrated in some cases, for example in the cytochrome b6f complex. In this case the atomic structure solved at 2.5 Å allowed identifying seven lipid-binding sites, which might function

as cross-linking interactions to stabilize the super-complex with the PSI reaction center complex (Zhang *et al.,* 2003; Kurisu *et al.,* 2003; Phillips *et al.,* 2009; Hasan et al., 2014).

- **Changing the expression system to increase protein overexpression**

Changing the expression system into *P. pastoris* strain KM71H resulted in an increase in the protein expression of 12 and 10 fold for $b^{0,+}$AT (*X. tropicalis*) and of LAT1 (*X. laevis*) respectively, which represented a great advantage over *S. cerevisiae*, in addition to be a cheaper expression system. But unfortunately the FSEC profiles of solubilized and purified proteins using the same conditions as in *S. cerevisiae,* indicated that a bad quality of the protein in both cases (Figure 49). Probably the overexpression produced not properly folded protein.

Finally, insect cells Sf9 cell line, which was set up in Dr. Gouaux laboratory by Dr. Rosell and reproduced in our laboratory, was tried and proven to be better in expression yields than *S. cerevisiae*, 5-fold higher than *S. cerevisiae*; 0.5 mg of pure protein from 1 L were obtained in insect cells, whereas 0.1 mg from 1 L was obtained in *S. cerevisiae*. Furthermore, the protein produced in insect cells and purified showed a monodisperse peak without the third peak corresponding to excess detergent, which is another improvement (Figure 57). Finally, this expression system resulted of special interest since it has been successful for more than 43% of eukaryotic membrane proteins crystal structures, especially from human, being at the moment the best expression system (Parker and Newstead, 2014) (examples of structures solved with this expression system are mentioned in the introduction of this Chapter II). For all these reasons, this was the expression system chosen for crystallization studies. Other improvements that were introduced into the protocol when overexpressing in insect cells were: i) the protease TEV replacement for thrombin, which cleaved more efficiently, and ii) the protease removal in the SEC step avoiding the reverse IMAC step; thus, reducing protein manipulation steps.

- **Trying the heavy subunit co-expression to find an stability effect**

As it was mentioned in the introduction, the heavy subunit 4F2hc stabilizes LAT2. Hence, it was interesting to study the effect of the heavy subunit rBAT on $b^{0,+}$AT. Several trials to co-express both subunits in different expression systems failed (Insect cells, mammal cells and *P. pastoris* were tried) (results by Dr. Rosell). It was not possible to express rBAT, probably to the higher amount of glycosilations that make it more complex than 4F2hc. For that reason it was continued working with the light subunit alone.

- **Removing the reactive cysteines to reduce aggregation**

Once it was determined that the heterodimer was not forming, it was thought to remove the reactive cysteine in $b^{0,+}$AT responsible of the disulphide bridge in order to increase the protein stability

since it is thought that reactive cysteines tend to aggregate proteins. This is a strategy followed by many crystallographic groups to reduce the aggregation tendency of a membrane protein.

- **Reducing protein flexibility by generating truncation forms in N and C terminal**

In order to increase the crystal formation, one strategy that is followed in many groups in crystallography is the deletion of amino acids in the C and N terminal, in order to reduce the protein flexibility when these domains are really long. Some examples of structures solved using this strategy are the β2 adrenoceptor (Rasmussen *et al.,* 2011), the chicken acid-sensing ion channel 1-snake toxin complex (Baconguis *et al*., 2014) or the *C. elegans* GluClα ion channel (Althoff *et al.,* 2014) and the β1-adrenergic G protein-coupled receptor (Warne *et al.,* 2010). In the last case, also a truncated loop between TMD3 and TMD4 was included, which could be considered an idea for future crystallization studies. From all the truncated forms tested in C150A $b^{0,+}$AT *Xenopus tropicalis*: Δ27Nterm C150A $b^{0,+}$AT and Δ28Cterm C150A $b^{0,+}$AT were the best individual truncated forms (Figure 56). For that reason it was expected that the double truncated mutant (Δ27Nterm/Δ28Cterm C150A $b^{0,+}$AT) would be even more stable, but, unfortunately the protein FSEC profile was non-monodisperse (Figure 56), and this protein was discarded.

- **First crystallization trial**

Finally, Δ28Cterm C150A $b^{0,+}$AT produced in Sf9 insect cells, purified in presence of DDM, cholesterol and DPPC has been so far the best candidate condition for 3D crystallization studies after the different strategies followed to improve its stability and crystalizability.

The first crystallization trial was done in vapor diffusion sitting drop at 20 ^{o}C using the MemGold crystallization screening (from Molecular Dimensions), which is based on crystallization conditions reported for alpha helical membrane proteins up to 2007 from the Protein Data Bank. They cover a range of pH, precipitants (mainly PEGs) and salt additives. The two conditions that formed "crystals" were the 0.1M Ammonium Sulfate, 0.1M Glycine pH 3.8, 28% (w/v) Triethylene Glycol, and the 0.1M Calcium Chloride, 0.1M Tris pH 6.5, 13% PEG (w/v) MME 2000. Nevertheless, they did not diffract, probably for being pseudocrystals formed of excess detergent, lipid and cholesterol. Even though no good quality crystal was formed in the first crystallization trial, only 96 different conditions were tested.

Nowadays there are many commercial crystallization screenings available for membrane protein to be tested in vapor diffusion that could be tried in the future, like: Membfac (Hampton Research), Memstart, Memsys, MemBeta or Memgold (from Molecular Dimensions), BetaMem (Emerald Biosystems), JBScreen-Membrane (Jena Bioscience) or Mbclass screens (Quiagen), Memplus. Vapor diffusion is the method more widely used that has been successful in more membrane proteins, and it represents the 80% of structures of the PDB. Some examples of membrane protein structures

using this vapor diffusion method are: the *C.elegans* GluClα ion channel (Althoff *et al.,* 2014), the chicken acid-sensing ion channel 1-snake toxin complex (Baconguis *et al*., 2014), the AMPA receptor GluA2 (Dürr *et al.,* 2014), the β1-adrenergic G protein-coupled receptor (Warne *et al.,* 2010).

Since the excess of detergent can end up with spherulites formation during the crystallization trials, the removal of detergent previous crystal seeding is desired. Many strategies could be followed for that: 1) solubilize at the lower % detergent possible since the majority of excess detergent of the sample comes from the solubilization step, 2) dialyze the sample for longer periods of time (2 days), 3) concentrate the less possible to avoid concentrating the detergents as well, and run two SEC with the same sample instead of concentrating too much and run a single SEC.

- **Trying other crystallization methods**

In parallel to the introduction of lipids, different new crystallographic techniques have been developed which require the use of lipids to form ordered lipidic structures containing the membrane protein. Some of these specific crystallographic techniques for membrane proteins are: Lipidic Cubic Phase that form 3D crystals (Caffrey *et al.,* 2009) and Bicelles that form 2D crystals (Faham *et al*., 2002; Agah and Faham, 2012) (Figure 64). A specific variant developed in the group of Dr. Nissen (Centre for Membrane Pumps in Cells and Disease) which is an approach to achieve diffraction quality membrane protein crystals in high concentrations of lipid and detergent (HiLiDe) (Gourdon *et al.,* 2011). Since the selected light subunit best condition requires the addition of DPPC lipid to increase its stability, it might be a good idea to try any these methods.

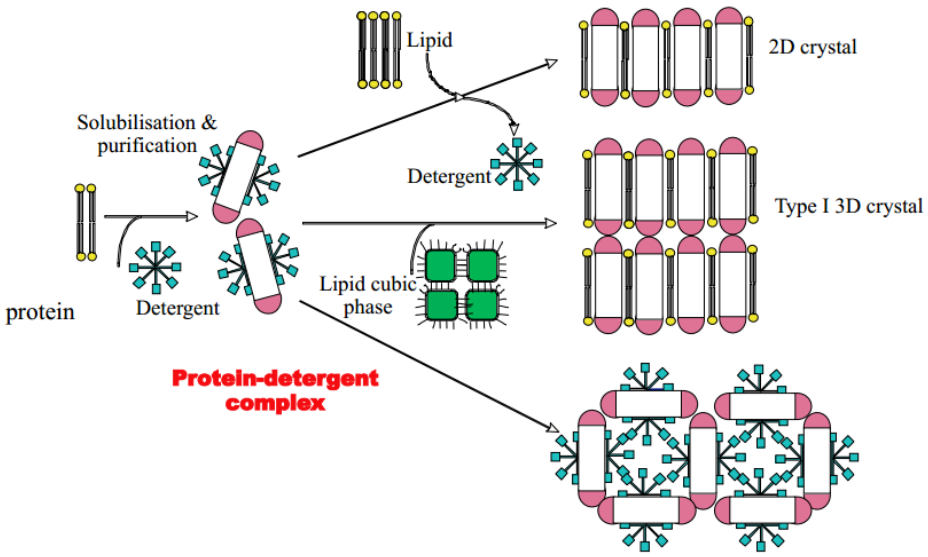


Figure 64. The solubilized and purified protein can be crystallized following several strategies.

- **Strategies to increase the crystal contacts**

One strategy to improve the chances of obtaining high-quality crystals is to increase crystal contacts by expanding the soluble domains. This can be carried out by several strategies. One is using epitope-specific antibodies that provide a large stable polar domain for crystal formation. Actually, a collaboration with Professor Gouaux is stablished in which they would build the Fab fragment for Δ28Cterm C150A $b^{0,+}$AT (*Xenopus tropicalis*) purified in DDM/CHS/DPPC. Many structures were obtained with this strategy like the β2 adrenaline receptor (Rosenbaum *et al.,* 2011), the *C. elegans* GluClα ion channel (Althoff *et al.,* 2014).

Other common strategy is to replace an external loop for a T4 lysozyme fusion protein which confers higher stability to the protein. It is also one common strategy, which has been useful for proteins like: membrane-bound pyrophosphatases (Kellosalo *et al.,* 2011), the human β2 adrenaline receptor (Cherezov *et al.,* 2007), the human histamine receptor (Shimamura *et al.,* 2011). This strategy could be followed if the ones before are finally not achieved.

In the case of HATs, instead of T4 lysozyme, the co-expression with the heavy subunit 4F2hc could achieve the same objective of increasing the contact surface to favor the packing. Even though 4F2hc/ $b^{0,+}$AT heterodimer is not found in the nature, it is known to be functional when co-expressed (Rajan *et al*., 1999; Rajan *et al*., 2000). The strategy of co-expressing *Xenopus tropicalis* 4F2hc/ $b^{0,+}$AT has already been started in the laboratory by Dr. Rosell.

- **Chosing a thermostable protein increases the chances of crystallization**

Thermostability of integral membrane proteins in detergent solution has proven to be an important parameter in the successful production of diffraction quality crystals for structure determination by X-ray crystallography. For that reason, screening different membrane proteins, testing different detergents, additives, substrates, inhibitors, antibodies, or screening mutants to find a more thermostable condition can be done with the FSEC thermostability assay for membrane protein precrystallization screening developed by Profesor Gouaux laboratory (Hattorie *et al.,* 2012). Finding a thermostable candidate has led to determine many membrane protein structures, like LeuT,ApcT, ASIC, P2X, GluA2, and GluCl (Hattori and Gouaux, 2012; Hibbs and Gouaux, 2011; Jasti *et al.,* 2007; Kawate *et al.,* 2009; Krishnamurthy and Gouaux, 2012; Shaffer *et al.,* 2009; Sobolevsky *et al.,* 2009; Yamashita *et al.,* 2005).

- **Immobilizing the transporter in a conformation by punctual mutations**

Mobility of the transporter could be blocked by finding mutants that could immobilize the transporter in a specific conformation. To that aim, residues to be mutated could be selected by multialignement with the prokaryotic homologues mutants that have been already crystallized in a precise conformation (described in the introduction 3.b.3); or by selecting point mutations found in

cystinuria patients where a loss of function was observed, and could be located in the substrate pathway. Recently, the human GLUT1 structure of the N45T and E329Q diffracted X-rays beyond 3.2 Å (Deng *et al.,* 2014). In this case they blocked the structure in the inward-facing conformation and mutated the residue responsible of the glycosylation.

- **Functionality of the protein**

Once the protein would be crystallized, the same conditions of overexpression and purification would be used to check that the crystallized protein is functional. To this aim, it is thought to use the scintillation proximity assay (SPA) which is already set up in the laboratory by Dr. Bartoccioni (Harder and Fotiadis, 2012). This assay can be applied to detergent-solubilized membrane transporters expressed with a histidine tag which would be immobilized to copper chelate SPA beads (fluoromicrospheres) containing the scintillant. Radiolabeled substrate will be tested on the protein-beads. Upon binding, the radioligand would emit radiation and induce photon emission from the beads that would be measured. Some advantages of this method are that it is possible to study specific transporters without interference of other endogenous transporters, that it is not require reconstitution into proteoliposomes, it does not need many steps, components are mixed in one well and signal is measured after incubation, which it makes it ideal tool for high-throughput screening; for example, to study new substrates, inhibitors, and for drug discovery. Alternatively, the purified protein could be reconstituted into proteoliposomes containing CHS and DPPC, as it was previously done for human 4F2hc/LAT2 (Rosell *et al.,* 2014). To that aim, *Xenopus tropicalis* $b^{0,+}$AT wild type and the Δ28Cterm C150A $b^{0,+}$AT version will be functionally tested to assess if the truncation and mutation had an effect on the transporter function.

To sum up, all the steps followed to find a suitable light subunit for crystallization studies have been successful. As a result, the Δ28Cterm C150A $b^{0,+}$AT (*Xenopus tropicalis*) expressed in Sf9 insect cells, solubilized and purified in DDM/CHS/DPPC is at the moment one good candidate to perform crystallization screenings.

II. Materials & Methods

II. Materials & Methods

II.1. Selected cDNAs of eukaryotic light subunits to be expressed in *S. cerevisiae*

For this project 24 light subunits from different species were selected (Table 15). The different cDNAs were obtained from different sources: either they were already available in the laboratory or either they were bought from the German Science Centre for Genome Research (RZPD: www.rzpd.de/). In the case of *Drosophila* light subunits, they were kindly provided by Dr. Goberdhan (University of Oxford).

Light Subunit	Specie	Uniprot code	% identity with human LAT	Obtained from
LAT1	*Homo sapiens*	Q01650	100	Laboratory
	Mus musculus	Q9Z127	86	Laboratory
	Xenopus laevis	Q32NQ3	83	Laboratory
JhI-21	*D. melanogaster*	Q9VKC2	51	Dr. Goberdhan
Minidiscs	*D. melanogaster*	Q9Y1A7	52	Dr. Goberdhan
LAT2	*Homo sapiens*	Q9UHI5	100	Laboratory
	Mus musculus	Q9QXW9	89	Laboratory
	Xenopus laevis	Q6P891	87	RZPD
y^+LAT1	*Homo sapiens*	Q9UM01	100	Laboratory
	Mus musculus	Q9Z1K8	90	Laboratory
	Xenopus tropicalis	Q6P4M7	83	RZPD
	Danio renio	Q2YDQ4	80	ZGC
	Bos taurus	A6QPX4	89	RZPD
y^+LAT2	*Homo sapiens*	Q92536	100	Laboratory
	Mus musculus	Q8BGK6	89	RZPD
	Xenopus laevis	Q28I80	83	RZPD
$b^{0,+}$AT	*Homo sapiens*	P82251	100	Laboratory
	Rattus norvegicus	P82252	83	RZPD
	Xenopus tropicalis	Q0P4R8	79	RZPD
	Bos taurus	Q3ZCL6	83	RZPD
asc1	*Homo sapiens*	Q9NS82	100	Laboratory
	Mus musculus	P63115	88	RZPD
xCT	*Homo sapiens*	Q9UPY5	100	Laboratory
	Mus musculus	Q9WTR6	89	RZPD

Table 15. List of the 24 eukaryotic light subunits selected for the project belonging to different species. Light subunit specie, Uniprot code and % amino acid identity corresponding to the human LAT are listed, as well as the source obtained from (laboratory, RZPD or Dr. Goberdhan laboratory).

II.2. Cloning eukaryotic light subunits cDNAs into GFP-8His-containing vector

The cDNAs encoding the different eukaryotic light subunits were inserted in the episomal pDDGFP-2 plasmid by homologous recombination in *S. cerevisiae*. This vector was generated by cloning the yeast enhanced green fluorescent protein (yEGFP) sequence (Cormack *et al.,* 1997) followed by an 8xHistidine tag and preceded by a Tobacco Etch Virus (TEV) protease site sequence into the 2 µ vector p424GAL1 (Mumberg *et al.,* 1995). Thus, the recombinant eukaryotic light subunits, expressed using this expression system, were produced as C-terminal GFP-tagged proteins. pDDGFP-2 contains a GAL1 promoter to regulate the galactose-inducible gene expression, a URA3 selection marker to select the yeast transformed colonies, and an ampicillin resistance to select the vector when transformed in *E. coli* for DNA amplification. This vector was kindly provided by Dr. Bernadette Byrne (Imperial College of London).

To perform the cloning of the eukaryotic light subunits into the pDDGFP-2 plasmid by homologous recombination, firstly, the genes of interest were amplified by high-fidelity PCR using the primers detailed in the Annex I (1.3. Primers to recombine into pDDGFP-2 plasmid). These primers contain: i) 5' overhangs (35 bp) complementing the upstream (coloured in red in Figure 65) and downstream (coloured in blue in Figure 65) recombination sequences to either side of the *Sma*I site in the pDDGFP-2 vector and ii) the corresponding forward and reverse sequence of the gene to be cloned.

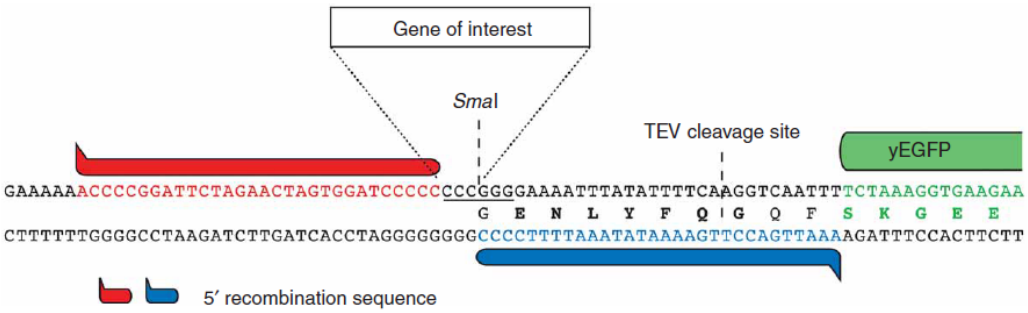


Figure 65. Cloning site used in the 2µ GFP-fusion vector pDDGFP-2. The yEGFP lacked the methionine start (from Drew *et al.,* 2008).

Secondly, pDDGFP-2 vector was linearized by *Sma*I digestion at 25 °C overnight. Then, the PCR products and the digested pDDGFP-2 plasmid were purified from 0.7% agarose gels (Quiagen kit) and quantified by using a NanoDrop spectrophotometer (Thermo Scientific). This purified material was used to transform the *S. cerevisiae* strain FGY217 (MATa, ura3-52, lys2D201, pep4δ) (Kota *et al.,* 2007), also provided by Dr. Bernadette Byrne (Figure 66). The main characteristic of this strain is that the gene encoding for the vacuolar proteinase A, pep4p, has been deleted, reducing possible proteolysis of the expressed protein.

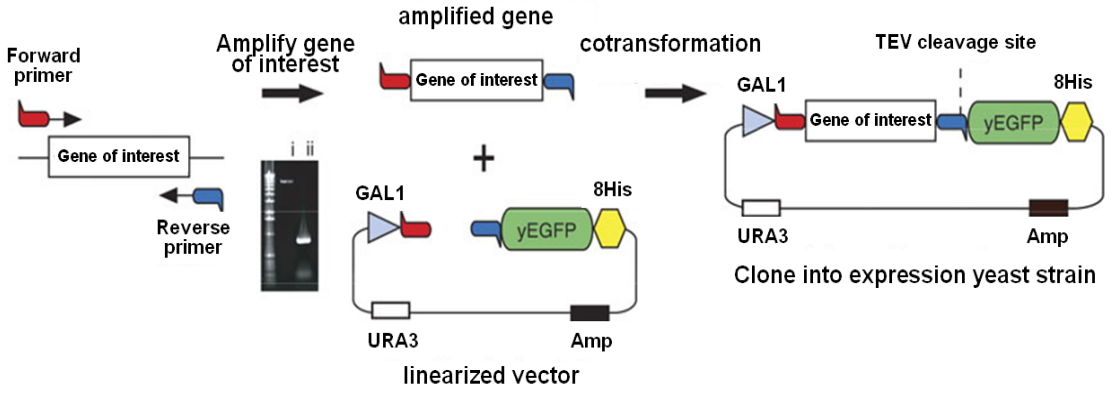


Figure 66. Cloning by homologous recombination into 2 µ GFP-fusion vector pDDGFP-2. i) *Sma*I linearized vector; ii) amplified gene with overhangs required for homologous recombination (from Drew *et al.,* 2008).

II.3. Preparation of *S. cerevisiae* competent cells and transformation

Firstly, 5 mL of Yeast Peptone Dextrose (YPD) medium were inoculated with the *S. cerevisiae* strain FGY217 into a 50 mL aerated falcon. This culture was grown overnight at 280 rpm and 30 °C in a shaker exclusive for yeast. The overnight culture was then diluted to an optical density of OD_{600} = 0.1 in a 250 mL shaker flask with 50 mL of YPD medium and grown in the same temperature until OD_{600} reached 0.5–0.6. At this cell density, cells were pelleted at 3,000*g* for 5 min at 4 °C, discarded supernatant and resuspended in 25 mL of sterile deionized H_2O (dH_2O hereafter) to wash the cell culture. Cells were pelleted again at 3,000*g* for 5 min at 4 °C, the supernatant discarded and the cell pellet resuspended in 1 mL of 100 mM litic acid (LiAc). Cell suspension was then transferred into a 1.5 mL eppendorf, cells centrifuged at 8,000*g* for 15 sec, supernatant discarded and cell pellet resuspended in 400 µL of 100 mM LiAc. 240 µL of 50% (wt/vol) PEG 3350 were prepared (on ice) in different eppendorf tubes for each transformation. 50 µL of cell suspension were added to each tube containing PEG 3350 and vortexed for 5 sec. 25 µl of 2 mg/mL single-stranded carrier DNA was added into each tube and vortexed for 5 sec. 50 µL of DNA mix were added (containing 3 µL of 25 ng/µL *Sma*I-digested vector, 5 µL of 150 ng/µL PCR product and 42 µL of sterile dH_2O) into each tube and vortexed for 5 sec. Cells were then incubated for 30 min at 30 °C, and heat shock was done for 25 min at 42 °C. After heat shock cells were washed (pelleted at 8,000*g* for 15 sec at room temperature, discarded supernatant and resuspended with 100 µL of sterile dH_2O). Cells were plated onto -URA selective plates and incubated for 2–3 days at 30 °C. Once the colonies reached the desired size, plates were stored at 4 °C previous overexpression screening.

II.4. Screening eukaryotic light subunit overexpression using whole-cell fluorescence

A colony of the transformant yeast was inoculated in 10 mL -URA medium supplemented with 2% glucose in an aerated 50 mL tube and grown overnight at 280 rpm and 30 °C. 10 µL of each overnight culture was spotted onto fresh -URA plates, dried at room temperature and incubated at 30 °C for 1–2 days. The overnight culture from the first step was diluted to a OD_{600}= 0.12 into two 50 mL aerated tubes, each containing 10 mL -URA medium with 0.1% glucose (high levels of glucose repress the GAL1 promoter) and incubated at 280 rpm and 30 °C. When OD_{600} was 0.6 (~7 hours) expression was induced by addition of galactose (2% final concentration). In order to avoid diluting culture medium, 20% galactose stock in -URA medium was prepared. Induction before OD_{600} = 0.6 causes a reduced biomass that lowers protein production, and later than OD_{600} = 0.6 it starts to be a greater protein degradation. Twenty-two hours after induction, cells were centrifuged at 3,000*g* for 5 min, removed the supernatant and resuspended the cell pellet in 200 µL of Yeast Suspension Buffer (YSB) (50 mM Tris–HCl (pH 7.6), 5 mM EDTA, 10% glycerol) supplemented with 1x complete protease inhibitor cocktail (Roche). Fluorescence measurements to check the eukaryotic light subunit expression were done by transferring 200 µL of the cell suspension to a black Nunc 96-well optical bottom plate and measuring GFP fluorescence at 512 nm by excitation at 488 nm in a microplate spectrofluorometer (FLx800, Biotek). Duplicates were done for each condition. In order to quantify the fluorescence, pure yEGFP at a concentration of 0.03 mg/mL (stocks of frozen purified protein were previously prepared in the laboratory by Dr. Rosell) was measured in the same plate and at the same final volume of 200 µL. The GFP concentration of the fusion proteins was calculated as follows:

$$\frac{\text{(Sample fluorescence - Background fluorescence)}}{\text{Pure yEGFP fluorescence}} \times \text{yEGFP stock (mg/mL)} = \text{GFP concentration (mg/mL)}$$

$$\frac{\text{Molecular Mass of Light Subunit (kDa)}}{\text{Molecular Mass of GFP (kDa)}} \times \text{Amount of GFP (mg/L)} = \text{Light Subunit overexpression (mg/L)}$$

To calculate the concentration in 200 µL, the previous value is divided by 40, corresponding to the 8,000 µL (cell culture)/200 µL (resuspension volume). It is considered 8mL instead of 10 mL because there is an effective 2 mL loss of volume by the transfer of only 200 µL of the resuspended cells to the 96 well plate, even though the final volume was around 250 µL (200 µL of buffer + cell pellet) (Drew *et al.*, 2008).

II.5. Confirmation of the integrity of GFP-tagged eukaryotic light subunits by in-gel fluorescence

In-gel fluorescence was done with crude cell membranes. For that, cells were broken using glass beads. The 200 µL of cell suspension (in last step of II.4) were transferred into a 1.5 mL eppendorf tube where glass beads (500 µm (Sigma)) were placed so that the final volume with the cell suspension was 500 µL, and an additional 500 µL of YSB was added. To break the cells a FastPrep was used (3 rounds for 10 sec each one at maximum intensity) at 4 °C. Unbroken cells were removed by centrifugation at 22,000*g* in a desktop centrifuge for 5 sec at 4 °C. Then, 500 µL of the supernatant was transferred into a new tube. And repeated the process with the mixture of unbroken cell pellet and glass beads where new 500 µL of YSB were added. The new supernatant was transferred to the 500 µL batch obtained from the first round of cell breakage. Crude membranes were pelleted by centrifuging the 1 mL supernatant using a desktop ultracentrifuge (120,000*g*) for 1 h at 4 °C. Crude membranes were resuspended in 50 µL of YSB, and 15 µL transferred into a tube containing 15 µL of Solubilization Buffer (SB; 50 mM Tris–HCl (pH 7.6), 5% glycerol, 5 mM EDTA (pH 8.0), 0.02% bromophenol blue, 2% SDS, 0.05 M DTT). 10 µL were loaded for SDS-PAGE (10% polyacrylamide) gels. Benchmark Fluorescent standard (Invitrogen) was included. Samples were not boiled to avoid GFP denaturation. SDS gel was rinsed with dH_2O and the fluorescent bands detected by exposing the gel to blue light. Bands were visualized using a GBOX gel reader (Syngene) and a Safe Imager (Invitrogen) of the Protein Expression IRB Core Facility.

II.6. Evaluation of the detergent solubilization efficiency on membranes containing overexpressed GFP-tagged eukaryotic light subunits

Those subunits showing the better overexpression levels were tested on detergents of different characteristics in order to determine the most efficient detergent solubilizing the overexpressed protein. To that aim, 1L of yeast culture for each protein was grown and the following detergents were tested for solubilization of light subunits:

- **Non-ionic detergents:** octyl glucoside (OG), nonyl glucoside (NG), nonyl maltoside (NM), decyl maltoside (DM), dodecyl maltoside (DDM), 6-cyclohexyl-1-hexyl-β-D-maltoside (Cymal-6) and 7-cyclohexyl-1-heptyl-β-D-maltoside (Cymal-7).
- **Zwitterionic detergents:** foscholine-9 (Fos-9), foscholine-12 (Fos-12), N,N-dimethyldodecylamine N-oxide (DDAO), lauryl dimethylamine N-oxide (LDAO), 3-dodecylamido-N,N'-dimethylpropyl amine oxide (LAPAO) and octaethylene glycol monododecyl ether ($C_{12}E_8$).

All detergents were prepared at a concentration of 10% (w/v) with or without cholesteryl hemisuccinate Tris salt (CHS) (Sigma Aldrich) 2% (w/v) in PBS. Since CHS is not easily soluble, it was added after detergent was completely soluble in the buffer.

- **Cell culture:** 10 mL of –URA, 2% glucose media in a 50 mL aerated falcon were inoculated with spotted yeast culture (see I.4). Next day, this cell culture was added to 150 mL of –URA, 2% glucose media in a 500 mL shaker flask, and incubated overnight in an orbital shaker at 280 rpm at 30°C. Next day, the 150 mL overnight culture was diluted to a final OD_{600} of 0.12 into 1 L of –URA, 0.1% glucose in a 2.5 L baffled shaker flask and incubated at 280 rpm and 30 °C until OD_{600} was 0.6. At this point, protein expression was induced by adding 100 mL of –URA media supplemented with 20% galactose to the 1 L cell culture. Cell culture was induced for 22 hours of incubation at 280 rpm and 30 °C. Then, cell culture was centrifuged at 4,000*g,* 4 °C for 10 min. The cell pellet was resuspended in 25 mL of Cell Resuspension Buffer (CRB; 50 mM Tris–HCl (pH 7.6), 1 mM EDTA, 0.6 M sorbitol) per liter of original cell culture.

- **Cell lysis:** Cells were broken using a heavy-duty cell disruptor (of the Protein Expression IRB Core Facility) for four passes at incremental pressures of 25, 30, 32 and 35 kpsi at 4–15 °C. The unbroken cells and debris were removed by centrifugation at 10,000*g* at 4 °C for 10 min and the supernatant containing membranes was collected.

- **Membrane extraction:** Membranes were collected by centrifuging the cleared supernatant at 150,000*g* at 4 °C for 120 min. Pellet was resuspended to a final volume of 6 ml per liter of the original cell culture with Membrane Resuspension Buffer (MRB; 20 mM Tris–HCl, (pH 7.5), 0.3 M sucrose, 0.1 mM $CaCl_2$). The total protein amount was calculated using the BCA protein assay kit following the manufacturer's instructions.

- **Detergent solubilization (with or without CHS):** The membrane suspension was adjusted to a final protein concentration of 3 mg/mL in PBS. To that aim, membrane suspension aliquots of 900 µL were transferred into 1.5 mL eppendorff tubes and 100 µL of each freshly prepared detergent with or without CHS (final 1%/0.2% or 1% respectively) were added to each of the 1.5 mL tubes containing 900 µL of membrane suspension (final 3 mg/mL). The mixtures were then incubated at 4 °C for 1 h with mild agitation. 100 µL were transferred into a Eppendorf tube and kept to measure later the fluorescence. The protein aggregates were removed by sedimentation at high g forces, 100,000*g* at 4 °C for 1 hour, assuming that they are orders of magnitude heavier than solubilized protein particles (the ultracentrifuge used in all the cases was Beckman-Coulter Optima Max, and the rotor TLA-55, Beckman Coulter). Then, 100 µL of the supernatant were transferred into a black Nunc 96-well optical bottom plate and the previous 100 µL previously kept were also transferred and fluorescence was measured as previously described (II.6). The detergent solubilization efficiency was calculated by comparing to the GFP fluorescence before and after the ultracentrifugation step.

II.7. Study of the GFP-tagged eukaryotic light subunits stability by the Ultracentrifugation Dispersity Sedimentation Assay (UDS)

The UDS is a high-throughput assay that allows to determinate rapidly the degree of protein aggregation in a detergent condition prior to crystallization, using small volumes of detergent solubilized membrane proteins that are removed by ultracentrifugation (Gutmann *et al.,* 2007).

Solubilized and ultracentrifuged protein as described in step II.6 was kept at 4 °C for 16 hours. Next day, 100 µL of sample were kept for later measurements, and the rest of the sample was ultracentrifuged at 100,000*g* at 4 °C for 1 hour (adapted from Gutmann *et al.,* 2007). Then, 100 µL of the supernatant were placed into a black Nunc 96-well optical bottom plate and the 100 µL previously kept were also transferred. Fluorescence was measured as described before (II.6). The tendency of these transporters to aggregate was evaluated by comparing to the GFP fluorescence measured before and after the ultracentrifugation, assuming that protein aggregates sediment at high *g*.

II.8. Analysis of the GFP-tagged eukaryotic light subunits monodispersity by Fluorescence Size Exclusion Chromatography (FSEC)

FSEC method was chosen because it allows screening the degree of monodispersity and the expression level of solubilized membrane proteins covalently fused to GFP without the need to purify the sample (Kawate and Gouaux, 2006).

To that aim, 0.5 mL of the solubilized and ultracentrifuged protein as described in step II.6 were injected into a Superose 6 10/300 column pre-equilibrated with 20 mM Tris–HCl (pH 7.5), 0.15 M NaCl and 0.03% DDM. The eluted sample was fractionated row-by-row in 0.2 mL into a 96-well plate after the first 6 mL (note that void volume usually appear at 7-8mL). GFP fluorescence at 512 nm by excitation at 488 nm emission was measured row-by-row. GFP fluorescence was plotted in each well against the fraction number. Finally, a graphic of the FSEC profile was generated using GraphPad Prism 5.

II.9. Purification of eukaryotic light subunits and monodispersity analysis by SEC

The 20 different conditions presenting monodisperse peaks in the FSEC assay were screened following this protocol. The purification followed several steps: GFP fusion protein solubilization, purification, cleavage of GFP-His tag, dialysis to remove detergent excess, removal of GFP-His tag by reverse IMAC, concentration of the flow through containing the purified protein, concentration of the purified protein to 4 mg/mL and finally SEC monodispersity analysis, which was also checked by SDS-PAGE silver staining.

Solubilization: 200 mg of total membrane protein (from 2L-6L cell culture depending on the expression level of each protein) (obtained as described in II.6) were solubilized at final total protein concentration of 5mg/mL in 20 mM Tris-Base, (pH 8), 150 mM NaCl, 1% (w/v) detergent / (w/v) 0.2% CHS, 10% glycerol for 1 hour at 4°C. Un-solubilized material was pelleted during 1h at 4°C at 100,000*g*.

GFP-fusion light subunit purification: Then, detergent-solubilized protein was purified by incubation of the supernatant (5 mM imidazole (Sigma) added) with 3 mL of Ni^{++}-NTA resin (Qiagen) for 2 hours. During the purification the detergent concentration was reduced to 3 times the respective CMC and the CHS amount was decreased as well to keep a 5:1 detergent/CHS ratio in 20 mM Tris-Base, (pH 8), 150 mM NaCl, 10% glycerol (Purification Buffer, PB:). After incubation, resin was washed twice with 30 mL of washing buffer (PB with 30mM imidazole (pH 8)). Elution was carried out incubating resin with 6 bed volumes of elution buffer (PB with 250 mM imidazole (pH 8)) for 30 min. All this process was carried out at 4°C. Then, the eluted protein was concentrated with a centricon (Millipore) of 100 kDa cut-off until 1 mg/mL.

GFP-His removal: TEV protease at a ratio of 1:5 was added and the sample was transferred into a dialysis tubing of 14 kDa molecular weight cut-off (Spectrumlabs) and dialyze overnight at 4°C against dialysis buffer (purification buffer without glycerol) to remove the excess of detergent produced during protein concentration and remove the glycerol. Note that TEV protease was initially produced in *E. coli* (see II.9.a), later it was provided by the Protein expression facility that produced in the same way, and eventually commercial TEV from Invitrogen was also used. **GFP-His and TEV-His removal:** Next day, free GFP-His and TEV-His were removed by reverse IMAC.

Concentration of the untagged light subunit: The flow through, containing the purified light subunit was collected and concentrated with a centrifugal filter device of 100 kDa cut-uff until a final concentration of 4mg/mL (concentration was measured in the Nanodrop and corrected the value by its extinction coefficient).

SEC: After ultracentrifugation (at 100,000*g*, 4 °C for 1 hour) to remove the aggregates, 500 µL of the supernatant of the sample were injected into a ÄKTA purifier in Superdex 200 10/300 GL column to analyze the monodispersity of the purified, digested and concentrated eukaryotic light subunits. The column was equilibrated with dialysis buffer. The sample was collected into a 96-well plate as in FSEC (see II.8) to analyze by SDS-PAGE the eluted sample.

SDS-PAGE and silver staining: those light subunits that did not aggregate all in the void and had a promising SEC profile were analyzed by SDS-PAGE and silver staining. A sample of the different fractions interesting to analyze were mixed with LSB and run a SDS-PAGE in a 8% polyacrylamide gel that was stained using the Silver staining Pierce kit which stains which can detect proteins easily at low-nanogram (1-5 ng).

II.9.a. TEV protease overexpression and purification

Part of the TEV used to digest the GFP-fusion light subunits was obtained as the following protocol details, and part were produced by the Protein Expression IRB Facility.

TEV overexpression: A glycerol stock stored at -80 °C of *E.coli* BL21(DE3)pLyss (resistant to chloramphenicol) clone containing TEV protease with Hisx6 N-terminal tag cloned into a kanamycin resistant vector was kindly provided by Dr. Fita laboratory. 5 mL of growing media (LB media supplemented with 1 % glucose, 50 mg/mL kanamycin, and 35 mg/mL of chloramphenicol) were grown overnight. Next day, this cell culture was used to inoculate 500 mL of growing media for 2 hours at 37 °C and 250 rpm, until OD_{600} reached 0.5. At this point, the incubator was cooled down to 25 °C and the culture grown until OD_{600} was 0.8 (around 30 min). Then, 0.5 mM (final concentration) of isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce TEV overexpression at 25 °C and 250 rpm overnight. Next day, cell culture was pelleted at 10,000 rpm for 10 min (using the rotor JA14) and the pellet was resuspended in 50 mL of lysis buffer (20 mM Tris (pH 8), 150 mM NaCl, 10 mM imidazole, 2 mM mercaptoethanol, and protease inhibitor EDTA free (Roche) (1 tablet/50 mL).

Cell culture lysis: cells were disrupted using the cell disruptor at the Protein Expression IRB Core Facility, for 1 cycle at 15 kpsi, and 2 cycles at 20 kpsi. Then, added 10 mL of lysis buffer for 1 cycle at 20 kpsi to collect the remaining cells. The non-broken cells were pelleted at 12,000 rpm (using the JA 25.50 rotor) for 30 min at 4 °C, and the supernatant was filtered for later purification.

TEV purification: purification was performed by nickel affinity chromatography. The filtered supernatant was mixed with pre-equilibrated Ni-NTA resin and let in stirring for 1 hour at 4 °C. Then, it was transferred into a column and 3 washes of 10x bed volumes with washing buffer (20mM Tris-HCl (pH 8), 1 M NaCl and 10 mM imidazole were performed. His-TEV was eluted with 1 bed volume of elution buffer (20mM Tris-HCl (pH 8), 150 mM NaCl and 500 mM imidazole).

Buffer exchange: using a centricon of 30 kDa cut-off (Millipore), the eluted sample was concentrated and replaced by 20 mM Tris-HCl (pH8), 150 mM NaCl, 2 mM DTT, 10 % glycerol, until 2 mg/mL. Concentration was measured with the BCA method. Aliquots of 200 µL at 1 mg/mL were stored at -80 °C. A 20 µL sample was used to check the purification purity by running a 8% polyacrylamide SDS-PAGE and the gel was stained by Comassie. In all the cases a double band was detected at ~28 kDa due to the autoproteolytic activity of TEV.

II.10. GFP-fusion light subunits overexpressed in *Pichia pastoris*

Xenopus tropicalis $b^{0,+}$AT, *Xenopus laevis* LAT1, selected as the two best candidate light subunits in the *S. cerevisiae* pipeline adapted from Drew *et al.,* 2008, were cloned as GFP-Hisx8 fusion proteins (at C-terminal) into pPICZB vector and transformed into *P. pastoris* KM71H strain by Dr. Rosell to check whether this expression system could produce higher amount of protein than *S. cerevisiae*. This was tried before to change into insect Sf9 cells, since it is a cheaper method, and it was already set up in the laboratory. The same fusion protein design as in *S. cerevisiae* cloning was followed (light subunit-TEV protease target-GFP-Hisx8).

The different *P. pastoris* glycerol stocks stored at -80 °C were used to overexpress the different light subunits following the protocol described before for the human LAT2 (see Chapter I: Materials & Methods I.9).

II.11. Baculovirus-Insect Cell Expression system

The use of insect cell culture for heterologous protein expression for protein crystallography has increased in the last years due to its ability to produce relatively large quantities of posttranslationally modified eukaryotic proteins in a relatively short period of time. The most extended technology is the baculovirus expression vector system since this family of virus are the most common infecting insect cells, capable of infecting more than 500 species of insects, and also are ideal vectors for a variety of mammalian cell lines (BacMan Technology) (an example is Penmatsa *et al.,* 2013). The most extended lepidopteran cells for baculovirus expression vector system are the Sf9 and Sf21 cell lines, which were isolated from ovarian tissue of the fall army worm *Spodoptera frugiperda*, and the High Five cell line, also designated BTI-Tn-5B1-4, that was stablished from the *Trichoplusia ni* embryonic tissue. In fact, Sf9 cells are a sub-clone of the Sf21 cells (the IPLBSF-Sf21 cell line) and were selected for their faster growth rate and higher cell densities than the Sf21 cells (Vaughn *et al.,* 1977; O'Reilly *et al.,* 1992).

This technology offers great advantages over bacterial, yeast or mammalian protein expression systems since the recombinant proteins they express show similar biological activity, antigenicity and immunogenicity to the authentic natural proteins. Moreover, they are not pathogenic for vertebrates or plants, and vectors do not depend on helper viruses. Sf9 cell line (Invitrogen, Novagen) was selected to overexpress *Xenopus* $b^{0,+}$AT, the best light subunit candidate from the screening performed in *S. cerevisiae*.

II.11.a. Growth and maintenance of the Sf9 cell line

Thawing and adaptation of Sf9 insect cells: -80 °C stored Sf9 insect cells were thawed quickly in 27 °C water bath, opened in sterile conditions under a laminar flow hood and carefully pipetted the content into a 15 mL centrifuge tube. 9 mL of pre-warmed 27ºC serum-free media Sf-900^{TM}III (Invitrogen) were added. This will be the media that will be used to manipulate these cells and in the expression of the correspondent proteins. Then, cells were resuspended 3 times and centrifuged at 400 g for 3 min. Then, medium was removed and 10 mL of fresh medium (pre-warmed at 27 °C) was added, and cells were resuspended and transferred into a 125 mL shake flask. Incubation was done at 27ºC at 120 rpm in a non-humidified and non-CO_2 regulated environment. After 48 hours, growth and viability was determined by Trypan Blue dye exclusion method (see Sf9 cell counting). Cell number over 10^6 cells/mL indicates viability over 85%.

Sf9 cell counting: Cell viability was analyzed by the trypan blue dye exclusion method. For that, 100 µL of cell culture was pipetted from the cell suspension culture into a 1.5 mL Eppendorf tube and mixed with 100 µL of trypan blue dye. Cells were resuspended and 10 µL were pipetted into a Newbauer chamber, counted in the several big squares and averaged them. The concentration of cells was calculated using: $Nx2x10^4$ cells/mL (the dilution factor 1:1 when applying the Trypan blue is here corrected).

Regular passage in cell suspension: To maintain Sf9 cell line at the maximum viability, cell culture was sub-cultured by diluting cells back to the optimal density every 2 to 3 days a week to maintain a log phase growth. Suspension cultures reached mid-log phase of growth at 2-4 × 10^6 cells/mL cell density. Then, they were diluted to 2.5× 10^5 (Friday) or 5× 10^5 cells/mL (Monday and Wednesday). Always a new flask was used, and maintained all the sterile conditions in all the process. The maximum volume that corresponds to each flask is 100 mL in 250 mL flask, 200 mL in 500 mL flask, 400 ml in 1L flask, and 0.8 L in 2 L flask. Media is always preheated at 27 °C. Cell culture is grown at 27 °C and 115 rpm in a non-humidified and non-CO_2 regulated environment. Maximum passage number for Sf9 cell for virus production is 25-30.

Freezing and storage of insect cells: For freezing insect cells, 300 mL of 2 x 10^6 cells/mL cell cultures were centrifuged at 1500 rpm for 2 min and the cell pellet re-suspended in 30 mL cryo-preservation media. Aliquots of the cell culture were slowly frozen in the blue-topped isopropanol container and cell culture vials were kept for a day at –80 °C before they were transferred to the liquid nitrogen tank for long-term storage.

II.12. Expression of *Xenopus* $b^{0,+}$AT in the Sf9 insect cell line

Xenopus tropicalis $b^{0,+}$AT wild type, the mutant C150A $b^{0,+}$AT and different truncated forms generated in N and C-terminal on the C150A $b^{0,+}$AT mutant were produced using this expression system. Sf9 cells were infected with baculoviruses containing the genes of interest using the Bac-to-Bac® Baculovirus Expression System (Invitrogen) which enables the efficient production of recombinant baculovirus for expression testing in insect cells.

The Bac-to-Bac® Expression System allows the generation of recombinant baculovirus by site-specific transposition in *E. coli* rather than homologous recombination in insect cells. The Bac-to-Bac® Vector Kit contains the pFastBac™ vector. This plasmid contains a strong polyhedrin promoter for protein expression and a large multiple cloning site for simplified cloning. The pFastBac™ constructs are used to transform DH10Bac™ *E. coli* competent cells, where takes place the recombination between the pFastBac™ expression cassette with the parent bacmid to form an expression bacmid. The bacmid is then purified and transfected into insect cells for production of recombinant baculovirus particles. The parent bacmid in DH10Bac™ *E. coli* contains a segment of the lacZα gene. The lacZα gene is disrupted upon transposition of the expression cassette into the bacmid allowing for blue/white selection of recombinants. This enables to identification of recombinant colonies easily (Figure 67).

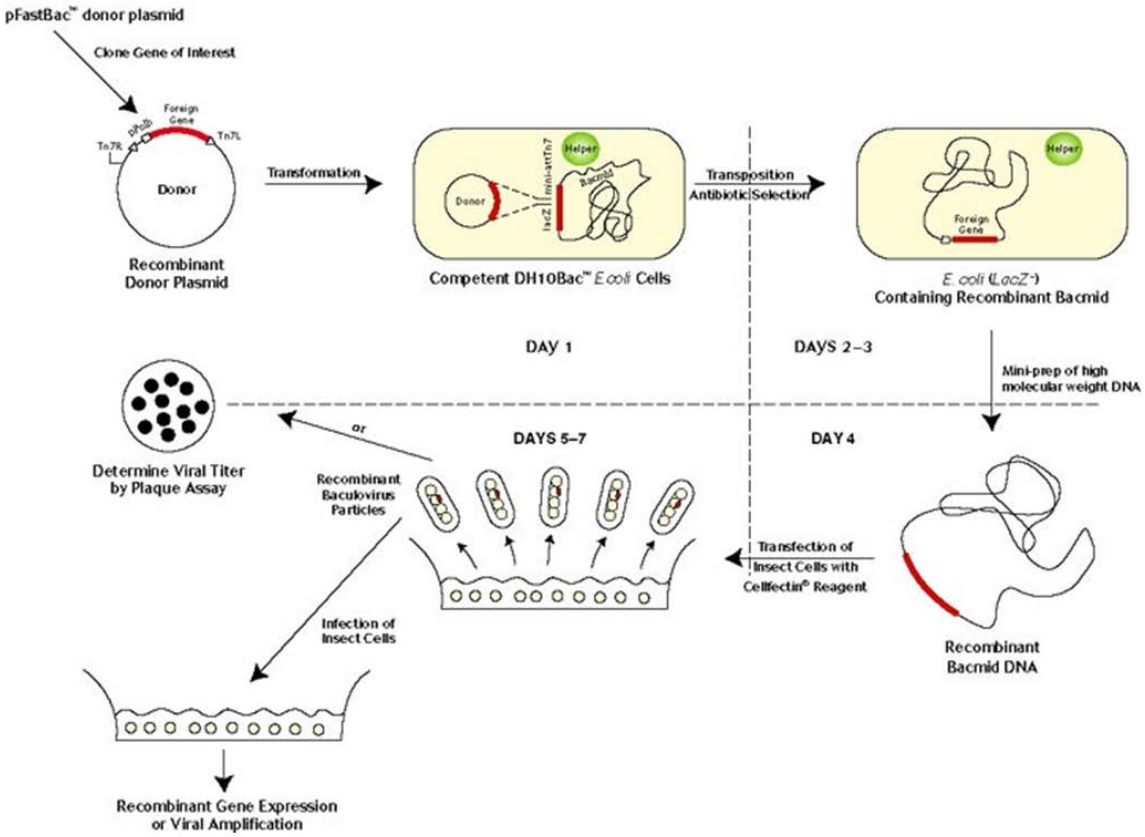


Figure 67. Scheme of the steps followed to generate the recombinant baculovirus to express the protein of interest using the Bac-to-Bac baculovirus expression system.

II.12.a. Cloning *Xenopus* $b^{0,+}$AT into pFastBac vector:

Xenopus $b^{0,+}$AT wild-type cDNA was cloned between EcoRI and XhoI sites into pFastBac™ vector (Invitrogen). This strategy facilitated the cloning of $b^{0,+}$AT wild-type cDNA in frame with a Thrombin site-GFP-His8 cDNA previously cloned into this plasmid. The design of this construct allowed the heterologous expression of *Xenopus* $b^{0,+}$AT wild-type as a C-terminal GFP-tagged protein. This cloning was performed following the protocol described before (see Chapter I: Materials & Methods I.1). In order to eliminate the reactive cysteine involved in the disulphide bridge with the corresponding heavy subunit (C150), the $b^{0,+}$AT C150A mutant was generated by site-directed mutagenesis (see Chapter I: Materials & Methods, I.2). The primers used for this purpose were (bolded nucleotides are the ones mutated): gca ttt tac cct ggc **gc**t gat ccc ccg gtg g (forward) and c cac cgg ggg atc a**gc** gcc agg gta aaa tgc (reverse). This mutant was used as a template to generate by PCR different truncated versions in the N- and C-terminus of $b^{0,+}$AT (Δ10, Δ27 and Δ8, Δ13, Δ18 and Δ28, respectively). The $b^{0,+}$AT C150A mutant together with its truncated versions were cloned also between EcoRI and XhoI sites into the pFastBac™ vector containing the Thrombin site-GFP-His8.

II.12.b. Transformation of DH10Bac *E. coli* cells with the pFastBac- $b^{0,+}$AT constructs:

Firstly, 50 ng of the pFastBac with the desired gene cloned were mixed with 50 μL of DH10Bac *E.coli* competent cells and incubate for 30 min on ice. Heat shock was performed for 45 sec at 42°C, and let them 2 min in ice. Then, 200 μL of super optimal broth with catabolite repression media (SOC) was added and incubated for 5 hours at 37°C (225 rpm). 100 μL of cells were plated on LB agar containing: 50 μg/mL kanamycin, 7 μg/mL gentamicin, 10 μg/mL tetracycline, 100 μg/mL Bluo-gal, and 40 μg/mL IPTG and incubated for 24-48 hours at 37°C (48 hours is preferable time because earlier it is hard to distinguish between white and blue colonies). Then, up 2-6 white colonies were peaked and grown overnight at 37°C in 6 mL of LB media containing 50 μg/mL kanamycin, 7 μg/mL gentamicin and 10 μg/mL tetracycline. From this 6 mL: 850 μL of cell suspension were mixed with 150 μL of autoclaved glycerol and stored at–80°C, and the rest of grown cells (~5 mL) were designated for bacmid isolation.

II.12.c. Isolation of the recombinant bacmid DNA from DH10Bac *E. coli* cells:

For bacmid isolation, DH10Bac *E.coli* cells from 5 mL cultures were spin down for 10 min at 4000 rpm. Supernatant was discarded and pellet dissolved in 200 μL of P1 buffer (Qiagen kits) or in Cell Resuspension Solution (Promega kits) and transferred into eppendorf tubes. Then, 200 μL of P2 reagent (Qiagen kits) or Cell Lysis Solution (Promega kits) was added and mixed by inverting eppendorf tubes few times. Then, 200 μL of N3 reagent (Qiagen kits) or Neutralization Solution (Promega kits) was added and mixed by inverting eppendorf tubes few times. Thereafter, a spin at maximum speed for 10 min was

done (using a table top centrifuge) and supernatant was collected in 2 mL eppendorf tubes. 1mL of isopropanol was added and tubes gently inverted and placed for 20 min at –20°C. Later on, tubes were spin for 15 min at RT and maximum speed in a 4ºC table top centrifuge. Supernatant was removed and pellet was washed with 1 mL of 70% ethanol by gently inverting eppendorf tube and spin for 15 min at room temperature (table top centrifuge). Then, supernatant was removed carefully and the pellet dried for 5 min. Finally, pellet was resuspended in 50 μL of autoclaved MilliQ water. Concentration of bacmid DNA was determined by using a NanoDrop spectrophotometer (Thermo Scientific) and diluted to 1 mg/mL.

II.12.d. Transfection of Sf9 cells with the recombinant bacmid DNA:

To produce the first population of virus (P1), Sf9 cells were seeded into 6-well plates at 0.9×10^6 cells/well in 2 mL of the serum-free media Sf-900™III (Invitrogen) and incubated at 27°C until they attach (around 30 min). While waiting, 8 μL of Cellfectin II (Invitrogen) were added to 100 μL of Sf-900™III media for each transfection condition. In parallel, 1 μL of 1mg/mL recombinant bacmid DNA was then added to 100 μL of Sf-900™III media. Then, the Cellfectin II containing tubes were mixed with the bacmid sample and incubated for 30 min at room temperature. DNA mix was added to each well and cells were incubated for 96 hours in 27 °C incubator (to prevent strong evaporation of media, plates were put into a recipient containing a slide of wet paper under them) in a non-humidified, non-CO2 regulated environment. Taking advantage of the generation of GFP fusion proteins, the expression was checked in a fluorescence microscope previous storage of P1. Then, supernatant containing P1 virus was collected by pipetting carefully (~2 mL from each well) and filtered into 2 mL eppendorf tube using 2 mL syringe and a filter of 0.2 μm. This stock of P1 virus was stored light protected at 4°C.

II.12.e. Infection of Sf9 cells with P1 virus (to produce P2 virus):

Sf9 cells were sub-cultured in 400 mL pre-warmed media in 1L flask at final concentration of 1.5×10^6 cells/mL. Then, infected with the P1 generated previously (see II.12.d) and incubated in an orbital shaker at 27 °C and 120 rpm during 96 hours in a non-humidified, non-CO_2 regulated environment. Then, cells were pelleted for 10-15 min at 4000-5000 rpm and collected supernatant containing P2 virus which was then filtered using disposable 0.2 μm filters (0.5 L Corning filter system was used) and stored at 4 °C light protected.

II.13. Baculovirus titration by the end point dilution assay:

A 10mL stock of Sf9 cells at a density of 0.75 x 10^6/mL in a sterile culture reservoir was needed. Then, 100 µL/well were seeded into a 96-well black plate (75,000 cells/well) and let them attach at 27 °C in a non-humidified and non-CO_2 regulated environment. Using a sterile deep-well 8-strip cluster, 10^{-1} to 10^{-8} stocks of viruses were made in Sf-900™III media by serial dilution (360 µL medium + 40 µL virus):

End point dilution preparation for baculovirus tritation		
Sf-900™III media	**P1 o P2 virus**	**Stock viruses**
360 µL	+ 40 µL	10^{-1}
360 µL	+ 40 µL 10^{-1} dilution	10^{-2}
360 µL	+ 40 µL 10^{-2} dilution	10^{-3}
360 µL	+ 40 µL 10^{-3} dilution	10^{-4}
360 µL	+ 40 µL 10^{-4} dilution	10^{-5}
360 µL	+ 40 µL 10^{-5} dilution	10^{-6}
360 µL	+ 40 µL 10^{-6} dilution	10^{-7}
360 µL	+ 40 µL 10^{-7} dilution	10^{-8}

Once cells were attached to the 96-well plate, medium was removed and replaced with 100 µL diluted virus. Infection was done in triplicates. 72 hours post-infection, green foci number of cells was counted in the dilution that gave <10 foci/well. To calculate the viral titer, the following equation was used:

(Average foci) x dilution factor x 10 = (pfu/ml)

And to calculate the volume required for subsequent infection, the following equation was used:

$$\frac{\text{(desired MOI) x (number of cells)}}{\text{Virus titer}} = \text{(mL)}$$

II.14. Expression of truncated versions of C150A $b^{0,+}$AT: overexpression and cell membrane extraction

40 mL Sf9 cultures at a cell density of 3-4×10^6 cells/mL were infected by adding P2 viruses, obtained as has been described previously, corresponding to the different truncated forms (Δ10, Δ27 in the N-terminus; Δ8, Δ13, Δ18 and Δ28 in the C-terminus) and to $b^{0,+}$AT wild-type as a control. In all cases cells were infected using a Multiplicity of Infection (MOI) of 2. Sf9 cultures were incubated for 48 hours in an orbital shaker at 27°C and 115 rpm. Cells were harvested by spinning for 15 min at 4000 rpm and pellets resuspended with 2 mL of cold resuspension buffer (20mM Tris (pH 8) 150mM NaCl with 2 tablets of protease inhibitor complete-EDTA free). The cell resuspension was lysated by sonication at 4 °C using 40% amplitude and 0.7 of intensity after adding DNAse to avoid viscosity. Sonication was performed for 15 sec 4 times. Then, non lysated cells were removed by centrifugation at 10,000 rpm for 20 min, and membranes were obtained by ultracentrifugation of the supernatant at 250,000*g* for 1 hour. Membranes were resuspended with 2 mL of resuspension buffer using a potter homogenizer. Membranes were flash-frozen with liquid N_2 and stored at -80 °C.

II.15. In-gel-fluorescence and FSEC to analyze protein expression and stability

The overexpressed protein in the membrane extracts (II.14) was analyzed by in-gel fluorescence (as in II.5) to check the protein expression levels and integrity. In addition, the quality of the C150A $b^{0,+}$AT truncated versions in comparison with $b^{0,+}$AT wild-type was analyzed by FSEC (similarly to II.8). 0.5mL of solubilization buffer (20mM Tris 20mM (pH 8) 150mM NaCl 2% DDM, 0.4% CHS) was added to an equal volume of the membrane stock. The mixture was incubated at 4 °C for 1 hour with mild agitation and the non-solubilized material removed by ultracentrifugation at 100,000*g* at 4 °C for 1 hour. 0.5 mL of the supernatant were injected onto a Superose 6 10/300 column equilibrated in 20 mM Tris–HCl (pH 7.5), 0.15 M NaCl and 0.03% DDM. The eluted sample was collected in fractions row by row in a 96-well plate and GFP fluorescence was measured. A graphic corresponding to the fluorescence of the GFP fusion protein was builded in GraphPad with the data obtained (as in II.8).

II.16. Large scale expression of $b^{0,+}$AT C150A (Δ28 C_{TER}) and purification

1,6L Sf9 culture (2x2L flasks) at a cell density of 3-4×10^6 cells/ml was infected with the $b^{0,+}$AT C150A (Δ28 C_{TER}) P2 virus at a final MOI of 2. Cells were incubated for 48 hours in an orbital shaker at 27°C and 115 rpm and harvested as has been described previously. The pellet was resuspended with 25 mL of cold resuspension buffer and cells lysed by sonication (see II.6). Membranes were purified following the previously described protocol and were resuspended with 20 mL of resuspension buffer using a Potter homogenizer. Membranes were flash-frozen with liquid N_2 and stored at -80 °C.

II.16.a. Buffer preparation for $b^{0,+}$AT C150A (Δ28 C_{TER}) purification

All the buffers used in the protein purification contained the resuspension buffer (RB) as a base (20mM Tris (pH 8) 150mM NaCl), on which protease inhibitor complete-EDTA free (Roche) was added previous addition of detergents. Thereafter, following this order it was added detergent, cholesterol and lipids. All buffers were sonicated at 40% amplitude and 0.7 intensity, for 1 minute at intervals of 10 sec, and filtered. Note: in all cases imidazole was added from a 1M (pH 8) stock to avoid changes on pH.

Solubilization buffer (2x DDM, CHS, lipid) (SB): first, fresh complete-EDTA free (1 tablet for each 50 mL of final volume) was added and dissolved. Then, DDM 2% w/v was added until completely dissolved. Then, Cholesteryl hemisuccinate Tris salt (CHS) 0.4% w/v (Sigma) and 10mM 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (DPPC) (Avanti Polar Lipids) were added. This buffer was sonicated as mentioned before to facilitate the solubility of CHS and lipids. Finally, imidazole was added to obtain a 10 mM final concentration. This buffer was added to an equal volume of the membrane stock volume to obtain a final 1x solubilization buffer with 1% DDM, 0.2% CHS, 5mM DPPC, 5mM imidazole.

- **Equilibration buffer:** 1x solubilization buffer was used.
- **Washing buffer (WB):** Resuspension buffer with 0.05% w/v DDM, 0.01% w/v CHS, 0.25 mM DPPC and 25 mM imidazole. It was obtained from diluting 40 times the solubilization buffer.
- **Elution buffer:** WB but with 250 mM imidazole.
- **Dialysis buffer (DB):** Resuspension buffer with 0.026% w/v DDM, 0.0052% w/v CHS, 0.25 mM DPPC and 5 mM $CaCl_2$.
- **Dialysis buffer after thrombin digestion:** DB without $CaCl_2$.

II.16.b. Solubilization, purification and preliminary 3D-crystallization screening of C150A $b^{0,+}$AT (Δ28 C_{TER})

Membranes from 1.6L cell culture were solubilized by adding the same volume of solubilization buffer (2x) to the membrane extract volume. Solubilization was performed by stirring for 2 hours and sample was ultracentrifuged for 1 hour at 100,000*g* at 4 °C to remove the aggregated material.

Supernatant was used to carry out the purification. 2.5 mL of Ni-NTA resin (5 mL of 50% slurry for 2.5 mg of GFP fusion to be purified) were equilibrated with five column volumes of equilibration buffer for 20 min in an orbital shaker at 4 °C. Detergent-solubilized membranes were added to the equilibrated nickel resin for binding. Magnetic stirrer was used to mix the solution at 4 °C for 2 hours. Then, it was transferred to a closed-flow column, let the resin precipitate, open the flow and collected the flow through. Before the column became dry, flow was stopped or directly added washing buffer. 100 µL flow through sample was collected to check binding efficiency by fluorescence quantification. Then, 20 column volumes of washing buffer were applied to the resin column without let it dry. After washes flow was stopped, 15 mL of elution buffer were applied, closed the column entrance and let for

15 min in orbital rotation at 4 °C. Then, opened the flow and collected the elution every 1 mL in different eppendorf 1.5 mL tubes. First and last tubes were discarded. The 13 mL resulted from elution were concentrated until 1.5 mL using a centricon (Amicon Ultra 15 mL 100kDa cut off). Protein amount was checked by nanodrop as previously described. A 1/10 dilution was done previous nanodrop measurement in order to avoid imidazole interference.

$CaCl_2$ was added to the sample at 5 mM final concentration, and thrombin from human plasma (Sigma) was added at 1:100 ratio (thrombin: fusion protein). Then, placed into a dialysis tubing of 14 kDa cut off membrane and dialyzed overnight at 4 °C against 100 times the sample volume with DB. As thrombin is ~37 kDa, and GFP tag is also under this MW, size exclusion chromatography will clearly separate them from the membrane protein which MW is higher ~54 kDa and together with detergent, cholesterol and lipids this MW increases much more.

A second dialysis was performed with new DB in order to diminish the excess of detergent, CHS, lipids, and remove the $CaCl_2$. After dialysis, sample was concentrated until 1 mL using centricon (Amicon Ultra 15 mL and 2 mL 100kDa cut off).

Äkta purifier (GE healthcare) for fast protein liquid chromatography (FPLC) and Superdex 200 (10/300) column (GE Healthcare Life Sciences) were used. 500 µL of purified sample were run at 4 °C with filtered new dialysis buffer. Peaks from 9.8 mL to 11 mL were collected and concentrated. Protein concentration was measured with nanodrop and sample was concentrated again with centricon (Amicon Ultra 2 mL 100 kDa cut off) until 4 mg/mL for crystallization screenings.

The crystallization screening was performed by hanging drop vapor diffusion at 20 °C using the MemGold pack screening containing 96 different conditions. A drop ratio of 0.5 µL of protein (4 mg/mL concentrated) with 0.5 µL reservoir solution was used. The obtained crystals were flash frozen in liquid N_2 directly and used for X-ray diffraction data collection in the Alba Synchrotron.

Conclusions

Conclusions

1. 4F2hc-ED is positioned tilted on LAT2 covering almost completely the external face of the transporter which is confirmed by three different lines of evidence. The low resolution of the 3D model of TEM negative-staining SPA at 21 Å of purified human 4F2hc/LAT2 overexpressed in *P. pastoris* suggested that the 4F2hc ectodomain was positioned on top of LAT2. Crosslinking experiments using two length arm bismaleimides allowed to identify specific residues that nearly completely (>80%) crosslinked 4F2hc and LAT2 co-expressed in HEK 293T and triangulate together with the endogenous intersubunit disulphide bridge the position of 4F2hc-ED on top of LAT2. In addition, the steric restrictions defined by the results of crosslinking fulfill the best docking energy model, as well as two additional and highly similar poses, obtained by docking 4F2hc-ED atomic structure with the human LAT2 homology model built based on the structure of the open-to-out AdiC.

2. The interaction between 4F2hc and the extracellular surface of LAT2 increases the stability of the detergent-solubilized light subunit. Specifically, the ectodomain of 4F2hc suffices to increase the stability of DDM-solubilized LAT2. The hydrophobic character of the interaction between 4F2hc-ED and LAT2 might be at the basis of the stabilization of this light subunit by 4F2hc.

3. The screening of 24 metazoan light subunits of HATs for protein expression, solubility and stability, based on GFP and *S. cerevisiae* pipeline, resulted in the selection of $b^{0,+}$AT (*X. tropicalis*) followed by LAT1 (*X. laevis*) as best candidate proteins for 3D crystallization studies. $b^{0,+}$AT and LAT1 were solubilized and purified in DDM/ Cholesteryl hemisuccinate Tris salt (CHS) and CYMAL-6/CHS, respectively.

4. The stability of the best candidate, $b^{0,+}$AT, was improved by addition of CHS and 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC) during solubilization and purification.

5. Changing the expression system to Sf9 insect cells increased the expression levels up to 5 fold, and improved the quality of the overexpressed protein when analyzed by size exclusion chromatography.

6. Furthermore, it was selected the most stable truncation form in N and C-terminal to reduce protein flexibility. In addition, it was mutated-off the reactive cysteine (C150A) to diminish protein aggregation, resulting finally $b^{0,+}$AT (*X. tropicalis*) C105A, Δ28 C-terminal, as the most suitable candidate for 3D crystallization studies.

Annexes

1. Oligonucleotides:

1.1. Sequencing oligonucleotides

Universal primers used to sequence inserts contained in commercial vectors (F: forward; R: reverse)

oligo name	5'-3' sequence
SP6	ATTTAGGTGACACTATAGAA
T7prom F	TAATACGACTCACTATAGGG
T7term R	GCTAGTTATTGCTCAGCGG
BGH R	TAGAAGGCACAGTCGAGG

Primers used to sequence human 4F2hc

oligo name	5'-3' sequence
D1	ATGAGCCAGACACCGAG
D2	TCTCGATTACCTGAGCTCTC
D3	TCGAATCCAACAAAGACTTGC
D4	TGAGTGACCAGCGGAGTAAG
AS1	CCTCGACGACTTCCACCGT
AS2	AACTCCAGAGCATCCTTCAC
AS3	CCCTGGGGACAAAAGTCGATG
AS4	GACTTTGACCTCGGAGTGCT

1.2. Mutagenesis oligonucleotides

The different oligonucleotides used to generate the cysteine mutants for crosslinking were the following (reverse primers were the complementary sequences to forward ones):

Human LAT2 mutagenesis oligonucleotides (forward primer, bolded nucleotides are the ones mutated)

oligo name	5'-3' sequence
C154S	CCGCTCTTCCCCACC**TCC**TTCCCCCCAGAGTCTGG
C210S	GGGATTGTACAGATA**TCC**AAAGGAGAGTACTTCTGGC
A235C	CCTGACATCGGCCTCGTC**TGC**CTGGCTTTCCTTCAGGGC
A315C	GCTCCTAGGAGTCATG**TGC**TGGATCATGCCCATTTCTG
G392C	CATCAACTATGTG**TGC**TTCATCAACTACC
S441C	CCTGCTGGTCTTC**TGC**CTGTGGTCAGAGCCGG

Human 4F2hc Mutagenesis oligonucleotides (forward primer, bolded nucleotides are the ones mutated)

oligo name	5'-3' sequence
S151C	CGTCTCGATTACCTGAGC**TGC**CTGAAGGTGAAGGGC
S185C	CCAATTTTGGC**TGC**AAGGAAGATTTTG
S191C	GGAAGATTTTGAC**TGC**CTCTTGCAATCG
S195C	GACAGTCTCTTGCAA**TGC**GCTAAAAAAAAGAGCATCCG
S200C	CGGCTAAAAAAAAG**TGC**ATCCGTGTCATTCTGG
S393C	CATCAACTATGTG**TGC**TTCATCAACTACC
S412C	CACCAAGGGCTTC**TGC**GAAGACAGGCTC
S487C	CTGCCAGCGCC**TGC**CTGCCAGCCAAG
G505C	GGCCGTGAGGAG**TGC**TCCCCTCTTGAGC

1.3. Primers to recombine into pDDGFP-2 *Saccharomyces cerevisiae* plasmid

Primers to recombine into pDDGFP-2 *Saccharomyces cerevisiae* plasmid
Human HATs

oligo	5'-3' sequence
ASC1NterSCHS	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGCCGGCCACACGCAGCAGCCG
ASC1CterSCHS	**AAATTGACCTTGAAAATATAAATTTTCCCC**TTGTGGCTTCGAGGGCTTGTCTGT
LAT1NterSCHS	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGCGGGTGCGGGCCCGAAGCGG
LAT1CterSCHS	**AAATTGACCTTGAAAATATAAATTTTCCCC**TGTCTCCTGGGGGACCACCTG
LAT2NterSCHS	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGAAGAAGGAGCCAGGCACCGA
LAT2CterSCHS	**AAATTGACCTTGAAAATATAAATTTTCCCC**GGGCTGGGGCTGCCCCGCCACGTC
LAT2NterSCHS(Trun)	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGCCCTGAAGAAAGAGATCGGATTG
LAT2CterSCHS(Trun)	**AAATTGACCTTGAAAATATAAATTTTCCCC**GGGCTTGTGTTGCCAGTAAACACC
y^+LAT1NterSCHS	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGTTGACAGCACTGAGTATGAA
y^+LAT1CterSCHS	**AAATTGACCTTGAAAATATAAATTTTCCCC**GTTAGACTTGGGATCCCGTTGCTT
y^+LAT2NterSCHS	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGAAGCCAGGGAGCCTGGG
y^+LAT2CterSCHS	**AAATTGACCTTGAAAATATAAATTTTCCCC**GTCAGTTTTCCTCTCATCCTTTTT
xCTNterSCHS	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGTCAGAAAGCCTGTTGTGTCC
xCTCterSCHS	**AAATTGACCTTGAAAATATAAATTTTCCCC**TAACTTATCTTCTTCTGGTACAAC
rBATNterSCHS	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGCTGAAGATAAAAGCAAGAGA
rBATCterSCHS	**AAATTGACCTTGAAAATATAAATTTTCCCC**ACACGAGGTATACAGTATGTTCAG
$b^{0,+}$ATNterSCHS	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGGGGATACTGGCCTGAGAAAG
$b^{0,+}$ATCterSCHS	**AAATTGACCTTGAAAATATAAATTTTCCCC**CTCAGGGTCTTCCTCCGGTGGGAC

Primers for homologous recombination into *Saccharomyces cerevisiae*
Mouse HATs

oligo	5'-3' sequence
ASC1NterSCMus	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGAGGCGGGACAGCGACATGGCA
ASC1CterSCMus	**AAATTGACCTTGAAAATATAAATTTTCCCC**TTGTGTCTTCAAGGGCTTGTCCGT
LAT1NterSCMus	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGCGGTCGCGGGCGCCAAGAGG
LAT1CterSCMus	**AAATTGACCTTGAAAATATAAATTTTCCCC**AGTCTCCTGAGGTACCACCTGCAT
LAT2NterSCMus	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGAAAAGGGAGCCCGCCAGCGA
LAT2CterSCMus	**AAATTGACCTTGAAAATATAAATTTTCCCC**GGGCTGCTCCTCCGAATCCGGGTC
y^+LAT2NterSCMus	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGAAGCACAGGAGCTTGGGAGC
y^+LAT2CterSCMus	**AAATTGACCTTGAAAATATAAATTTTCCCC**GTCAGTTTTCCTCTCAACATTTTT
xCTNterSCMus	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGTCAGAAAGCCAGTTGTGGCC
xCTCterSCMus	**AAATTGACCTTGAAAATATAAATTTTCCCC**TAATTCTTTAGAGTCTTCTGGTAC

Primers for homologous recombination into *Saccharomyces cerevisiae*
Rat HATs

oligo	5'-3' sequence
b^{0+}ATNterSCRAT	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGAGGAGACAAGCCCGAGGAGA
b^{0+}ATCterSCRAT	**AAATTGACCTTGAAAATATAAATTTTCCCCC**TCTGGGTCCTTCTCTGGTGGGAC

Primers for homologous recombination into *Saccharomyces cerevisiae*
Cow HATs (*Bos taurus*)

oligo	5'-3' sequence
y+LAT1NterSCBos	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGAAGCTGAAGAAGGAGATCTCGCTG
y+LAT1CterSCBos	**AAATTGACCTTGAAAATATAAATTTTCCCC**TTTAGACTTGGGATCCTGTTGTTT
b^{0+}ATNterSCBos	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGAAGAGACGAGCCTGAGGAAA
b^{0+}ATCterSCBos	**AAATTGACCTTGAAAATATAAATTTTCCCCC**TCTGGAGCTTCCTCGGGGGGGAC

Primers for homologous recombination into *Saccharomyces cerevisiae*
Toad HATs (*Xenopus tropicalis* and *Xenopus laevis*)

oligo	5'-3' sequence
LAT1NterSCXL	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGCCGCAGACAGCGTGAAGCGG
LAT1CterSCXL	**AAATTGACCTTGAAAATATAAATTTTCCCC**GGACTCCTGGGGGACAGCCTCCATGAC
LAT2NterSCXL	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGCAGAAGGTGCAAGACTGCGG
LAT2CterSCXL	**AAATTGACCTTGAAAATATAAATTTTCCCC**GGCATGTGCTCCACTTTCTTCCTC
y^{+}LAT1NterSCXT	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGGTACCTCTACCAAGTACACA
y^{+}LAT1CterSCXT	**AAATTGACCTTGAAAATATAAATTTTCCCC**ATCACTTTTTGAGAAGTCTTTTTC
b^{0+}ATNterSCXT	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGATGAGACAAAGCCCAGAAAA
b^{0+}ATCterSCXT	**AAATTGACCTTGAAAATATAAATTTTCCCC**TTCTTTTGTTTCTTCAGGGGGAAC

Primers for homologous recombination into *Saccharomyces cerevisiae*
Zebra Fish HATs (*Danio Renio*)

oligo	5'-3' sequence
y+LAT1NterSCDR	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGCGTCCTACAGCCCTACGCCT
y+LAT1CterSCDR	**AAATTGACCTTGAAAATATAAATTTTCCCCC**TCCGTTTTTTCCTCCTCCTTGTC

Primers for homologous recombination into *Saccharomyces cerevisiae*
Drosophila HATs

oligo	5'-3' sequence
MndNterSCDRO	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGCGCGAGTTCAGGCCAGCGAC
MndCterSCDRO	**AAATTGACCTTGAAAATATAAATTTTCCCC**GTCGAACTTCTCCTGGTTGGGCAT
jhl21NterSCDRO	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGACGGACCGATATGCCAATGGA
jhl21CterSCDRO	**AAATTGACCTTGAAAATATAAATTTTCCCC**GTTAATAGCCTCATTCGATTCGAT

Primers for homologous recombination into *Saccharomyces cerevisiae*
Drosophila HATs

oligo	5'-3' sequence
MndNterSCDRO	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGGCGCGAGTTCAGGCCAGCGAC
MndCterSCDRO	**AAATTGACCTTGAAAATATAAATTTTCCCC**GTCGAACTTCTCCTGGTTGGGCAT
jhl21NterSCDRO	**ACCCCGGATTCTAGAACTAGTGGATCCCCC**ATGACGGACCGATATGCCAATGGA
jhl21CterSCDRO	**AAATTGACCTTGAAAATATAAATTTTCCCC**GTTAATAGCCTCATTCGATTCGAT

ARTICLE: Costa *et al*., Protein Expression and Purification, 2013

Expression of human heteromeric amino acid transporters in the yeast *Pichia pastoris*

Protein Expression and Purification Vol. 87, pp.35-40, 2013

Meritxell Costa, Albert Rosell, Elena Álvarez-Marimon, Antonio Zorzano,
Dimitrios Fotiadis, Manuel Palacín.

Protein Expression and Purification 87 (2013) 35–40

Contents lists available at SciVerse ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Expression of human heteromeric amino acid transporters in the yeast *Pichia pastoris*

Meritxell Costa [a,b,c,1], Albert Rosell [a,b,1], Elena Álvarez-Marimon [a,b], Antonio Zorzano [a,b], Dimitrios Fotiadis [c,*,2], Manuel Palacín [a,b,*,2]

[a] Institute for Research in Biomedicine (IRB Barcelona), Department of Biochemistry and Molecular Biology, Faculty of Biology, University of Barcelona, E-08028 Barcelona, Spain
[b] Centro de Investigacion Biomedica en Red de Enfermedades Raras, E-08028 Barcelona, Spain
[c] Institute of Biochemistry and Molecular Medicine, and Swiss National Centre of Competence in Research (NCCR) TransCure, University of Bern, CH-3012 Bern, Switzerland

ARTICLE INFO

Article history:
Received 19 July 2012
and in revised form 8 October 2012
Available online 22 October 2012

Keywords:
4F2hc
Heteromeric amino acid transporter
LAT2
Membrane protein
Overexpression
Pichia pastoris

ABSTRACT

Human heteromeric amino acid transporters (HATs) play key roles in renal and intestinal re-absorption, cell redox balance and tumor growth. These transporters are composed of a heavy and a light subunit, which are connected by a disulphide bridge. Heavy subunits are the two type II membrane *N*-glycoproteins rBAT and 4F2hc, while L-type amino acid transporters (LATs) are the light and catalytic subunits of HATs. We tested the expression of human 4F2hc and rBAT as well as seven light subunits in the methylotrophic yeast *Pichia pastoris*. 4F2hc and the light subunit LAT2 showed the highest expression levels and yields after detergent solubilization. Co-transformation of both subunits in Pichia cells resulted in overexpression of the disulphide bridge-linked 4F2hc/LAT2 heterodimer. Two sequential affinity chromatography steps were applied to purify detergent-solubilized heterodimers yielding ~1 mg of HAT from 2 l of cell culture. Our results indicate that *P. pastoris* is a convenient system for the expression and purification of human 4F2hc/LAT2 for structural studies.

Introduction

Heteromeric amino acid transporters (HATs)[3] are present only in metazoans. They are the unique example known in all kingdoms of life of solute transporters composed of two subunits, a heavy (SLC3 family) and a light subunit (SLC7 family) linked by a conserved disulphide bridge [1–3]. The heavy subunit is essential for trafficking of the holotransporter to the membrane [4,5], whereas the light subunit catalyses the transport activity [6]. Functionally, HATs are amino acid antiporters (exchangers) with a 1:1 stoichiometry [7,8]. Two homologous human heavy subunits have been cloned and are called rBAT and 4F2hc. rBAT is related to the transport of neutral and cationic amino acids (system $b^{0,+}$) and 4F2hc is the heavy chain of the surface antigen 4F2hc (also named CD98 or fusion regulatory protein 1 (FRP1)) [2]. So far, eight human light subunits have been identified. Six of them are partners of 4F2hc (LAT1, LAT2, y^+LAT1, y^+LAT2, asc1, xCT), one heterodimerizes with rBAT ($b^{0,+}$AT) and one (AGT1) seems to interact with yet unknown heavy subunit(s) [1,3].

Most human cells, if not all, express HATs. It is therefore not surprising that these transporters have an impact on human health. Thus, the aminoacidurias cystinuria types A and B, and lysinuric protein intolerance are caused by mutations in rBAT, $b^{0,+}$AT and y^+LAT1, respectively [9,10]. The transporter 4F2hc/xCT (system x_c^-) controls cell redox balance [11], is required for Kaposi's sarcoma-associated herpesvirus infection [12,13] and is involved in cocaine relapse [14]. Moreover, 4F2hc/LAT1 is overexpressed in many human tumors, thus providing essential neutral amino acids for cell growth and stimulation of mTOR (mammalian target of rapamycin) [15]. In addition to the role of HATs in amino acid transport, 4F2hc mediates β1- and β3-integrin signaling [16,17]. Indeed, 4F2hc is necessary for the integrin-dependent rapid proliferation of B-lymphocytes during clonal growth and for growth of vascular smooth muscle cells after arterial injury [17,18].

Structural information on HATs is very scarce [3]. SLC3 members (heavy subunits) are type II membrane *N*-glycoproteins with a single transmembrane domain (TMD), an intracellular N-terminus and a large extracellular C-terminus (Fig. 1A). The extracellular

* Corresponding authors. Addresses: Institute of Biochemistry and Molecular Medicine, University of Bern, Bühlstrasse 28, CH-3012 Bern, Switzerland. (D. Fotiadis), Institute for Research in Biomedicine (IRB Barcelona), Park Científic de Barcelona, Josep Samitier 1-5, E-08028 Barcelona, Spain. (M. Palacín)

E-mail addresses: dimitrios.fotiadis@ibmm.unibe.ch (D. Fotiadis), manuel.palacin@irbbarcelona.org (M. Palacín).

[1] These authors contributed equally to this work.

[2] Share senior authorship and correspondence.

[3] *Abbreviations used:* HATs, heteromeric amino acid transporters; FRP1, fusion regulatory protein 1; mTOR, mammalian target of rapamycin; TMD, transmembrane domain; TIM, triose phosphate isomerase; BMGY, buffered complex glycerol medium; BMMY, buffered complex methanol medium; BCA, bicinchoninic acid protein assay; SEC, size exclusion chromatography.

1046-5928/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.pep.2012.10.003

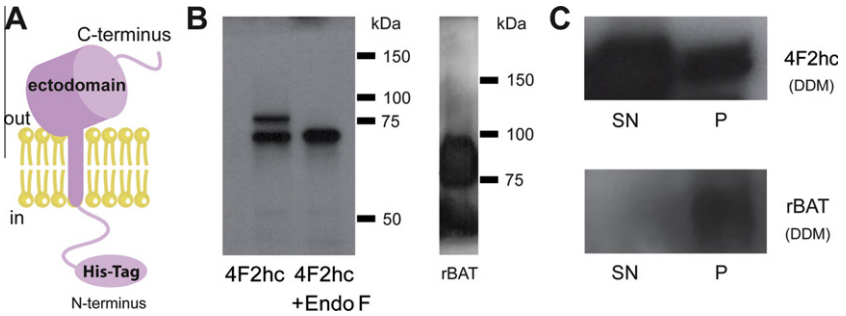


Fig. 1. Expression and solubilization of the heavy subunits of HATs: rBAT and 4F2hc. (A) Schematic representation of a heavy subunit. The locations of the termini and the engineered His-tag are indicated. (B) Left panel: Western blot analysis using anti-4F2hc antibody of untreated and Endo F treated membranes isolated from 4F2hc expressing Pichia cells. Right panel: Western blot analysis using anti-rBAT antibody of membranes isolated from rBAT expressing Pichia cells. (C) Solubilization with DDM of 4F2hc and rBAT from isolated membranes. An important fraction of 4F2hc is successfully solubilized by DDM and found in the supernatant (SN) after ultracentrifugation. This is not the case for rBAT, which aggregates and is exclusively found in the pellet (P).

domain of these proteins has amino acid sequence homology with bacterial α-amylases [19]. Moreover, X-ray diffraction of the extracellular domain of human 4F2hc revealed a three-dimensional structure similar to that of bacterial glucosidases. The structure of the 4F2hc ectodomain consists of a triose phosphate isomerase (TIM) barrel $(\beta\alpha)_8$ and eight antiparallel β-strands [20]. Despite this homology, 4F2hc lacks key catalytic residues and has no glucosidase activity [20]. The light subunits are highly hydrophobic and not glycosylated proteins. Cysteine-scanning mutagenesis studies with xCT support a twelve TMD topology with intracellular N- and C-termini (Fig. 2A) as a model for the light subunits of HATs [21]. Furthermore, the cysteine residue between TMD III and IV, which is involved in the inter-subunit disulphide bridge, is conserved. Structure–function studies [22,23] suggest that the structures of the light subunit of HATs should be similar to those of amino acid sequence-related (<20% amino acid sequence identity) prokaryotic transporters AdiC (L-arginine/agmatine antiporter) [24–27], ApcT (broad-specificity amino acid transporter) [28] and GadC (glutamate/GABA antiporter) [29]. These structures present the so-called '5 + 5 inverted repeat' fold, which was initially described in Na^+-coupled symporters [30–33]. This fold is characterized by five consecutive TMDs in the N- and C-terminal halves of the transporter, which are related by a pseudo-twofold symmetry.

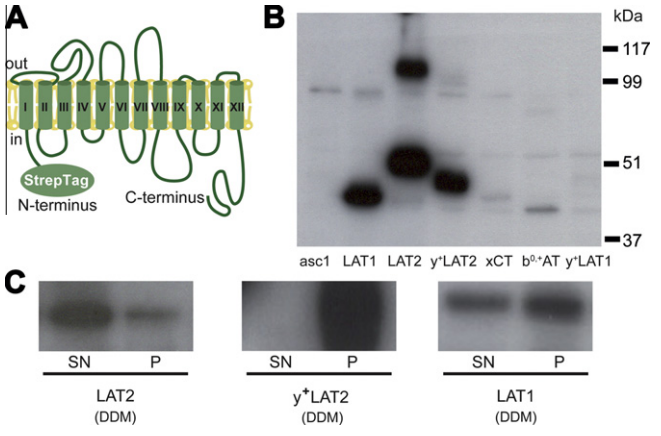


Fig. 2. Expression and solubilization of the light subunits of HATs: asc1, LAT1, LAT2, y^+LAT2, xCT, $b^{0,+}$AT and y^+LAT1. (A) Schematic representation of a light subunit. The locations of the termini and the engineered StrepTagII are indicated. TMDs are numbered. (B) Western blot analysis using anti-StrepTagII antibody of isolated membranes indicating the expression levels of the different light subunits of HATs. (C) Solubilization with DDM of the best expressing light subunits, i.e., LAT2, LAT1 and y^+LAT2. LAT2 was solubilized best and was mainly found in the supernatant (SN) after ultracentrifugation. LAT1 only solubilized partially and y^+LAT2 not at all, i.e., all the protein was found in the pellet (P).

The low amino acid sequence identity of AdiC, ApcT and GadC with the eukaryotic light subunits precludes the generation of a robust and reliable structural model of HATs based on these prokaryotic transporters. Moreover, in these prokaryotic structures there are no clues on the molecular interaction between the heavy and the light subunits of HATs.

In the present study, we report the first attempt to overexpress heterologously HATs for structure determination. Human HATs were expressed in the methylotrophic yeast *Pichia pastoris*, which can be used for high cell density fermentation and has been successfully used for the expression of eukaryotic membrane proteins [34–37]. Our results indicate that overexpression and purification of milligram amounts of recombinant human heterodimer 4F2hc/LAT2 is possible using the Pichia expression system. This opens the path for future structural studies and elucidation of the structure of human HATs.

Materials and methods

Microorganism strains and plasmid

Escherichia coli DH5 was used for cloning procedures. The *P. pastoris* expression kit, including vectors and host cells, was purchased from Invitrogen. Human HATs were expressed in the Pichia strain KM71H using the pPICZB vector.

Generation of expression constructs and P. pastoris transformation

The cDNAs of the full length human light subunit genes (asc1, LAT1, LAT2, y^+LAT1, y^+LAT2, xCT and $b^{0,+}$AT; accession numbers: Q9NS82, Q01650, Q9UHI5, Q9UM01, Q92536, Q9UPY5 and P82251) and human heavy subunit genes (4F2hc and rBAT; accession numbers: P08195 and Q07837) were inserted into the pPICZB vector between EcoRI and NotI sites. A Kozak consensus sequence (between brackets), a linker coding for Ala (italics), and a StrepTag-II sequence (underlined) or a 10× His-tag for the light or heavy subunits were incorporated before the codon corresponding to the second amino acid residue of each protein gene as follows: [ACC ATG *G*]-*CA*-TGGTCACATCCTCAATTTGAAAAA. Moreover, a DNA sequence encoding the Factor Xa cleavage site (ATCGAAGGTCGT) was added just after the tag sequences in all cases. To create the expression construct of 4F2hc/LAT2, the expression cassette containing the light subunit (AOX1 promoter plus cDNA for LAT2) was excised from the pPICZB–LAT2 construct using the BglII and BamHI sites. The cassette was then ligated into the pPICZB–4F2hc construct that was linearized with BglII. Before ligation, the PmeI restriction site in the AOX1 promoter of the heavy subunit was removed by site-directed mutagenesis to allow the

linearization of the 4F2hc/LAT2 construct by this enzyme. The constructs containing the heavy subunits (4F2hc or rBAT), the light subunits (asc1, LAT1, LAT2, y^+LAT1, y^+LAT2, xCT or $b^{0,+}$AT), or the 4F2hc/LAT2 heterodimer were used for transformation. *P. pastoris* cells were transformed by electroporation (BioRad Micropulser) using the constructs linearized with PmeI as described in the manufacturer's manual. Transformants were selected on YPD plates containing 100 μg/ml Zeocin (Invitrogen).

Expression screening of human HATs in P. pastoris

Zeocin resistant colonies were grown overnight in 15 ml of buffered complex glycerol medium (BMGY) until the cultures reached an OD_{600} of 2–6. After pelleting at 1500*g* and 4° C for 5 min, the cells were resuspended in buffered complex methanol medium (BMMY) using 1/5 of the original culture volume and grown for 72 h at 30° C. Additional methanol was added after 24 and 48 h to a final concentration of 1% (v/v), in order to maintain inducing conditions. Cell pellets were resuspended in 50 mM sodium phosphate, pH 7.4, 1 mM EDTA, 5% (w/v) glycerol and complete EDTA-free protease inhibitors (Roche). After addition of an equal volume of acid-washed glass beads (0.25–0.5 mm diameter), the cells were broken using the FastPrep system (MP Biomedicals) by three cycles of vortex for 10 s interspersed with 1 min incubations on ice. Glass beads, unbroken cells and other cell debris were removed by centrifugation at 1500*g* and 4° C for 10 min. The supernatant was collected and the pellet was resuspended in the original volume with fresh lysis buffer, subjected to an additional lysis cycle and recentrifuged. The combined supernatants were then ultracentrifuged at 250,000*g* and 4° C for 1 h. The crude membrane pellet was resuspended in 20 mM Tris–HCl, pH 7.4, 150 mM NaCl, 5% glycerol, and the protein content determined using the bicinchoninic acid protein (BCA) assay (Pierce).

For each protein and protein complex (i.e., heavy subunits, light subunits and 4F2hc/LAT2 heterodimer) isolated membranes from 20 clones were tested by Western blot analysis to identify the clone with the highest expression level and thus optimize expression. Light subunits, 4F2hc and rBAT were immunodetected using anti-StrepTagII (IBA GmbH), anti-4F2hc (H-300; Santa Cruz Biotechnology) and anti-rBAT [38] antibodies, respectively. The 4F2hc/LAT2 expression was analyzed using anti-StrepTagII and anti-4F2hc antibodies. No remarkable loss of 4F2hc/LAT2 expression, e.g. by loss of expression cassettes through recombination, was detected over repetitive rounds of cultivation and induction.

In vivo uptake of radiolabeled L-leucine into P. pastoris expressing 4F2hc/LAT2

The uptake assay was basically performed as described by Döring et al. [39] with small adaptations. Briefly, transformed and untransformed Pichia cells were harvested by centrifugation (2500*g*, 10 min, room temperature), washed twice with 10 mM Tris-HEPES, pH 7.4, 150 mM choline chloride, 1 mM $MgCl_2$, 1 mM $CaCl_2$ (transport buffer) and resuspended by adding 10 μl of the same buffer per OD_{600} unit. Uptake was initiated by mixing 10 μl of resuspended cells with 40 μl of transport buffer containing 0.5 μCi of [^{3}H]L-leucine (American Radiolabeled Chemicals). Uptake of L-leucine (10 μM final concentration) was terminated at different time points (see Fig. 3A) by addition of 3 ml of ice cold stop buffer (10 mM Tris-HEPES, pH 7.4, 150 mM choline chloride, 5 mM L-leucine) and filtered through membrane filters (Sartorius, 0.45 μm pore size). Filters were washed twice with 1–2 ml of ice cold stop buffer and dried. The trapped radioactivity was counted in a Packard Tri-Carb liquid scintillation counter. All experimental values were corrected by subtracting zero-time values.

Large-scale expression and purification of 4F2hc/LAT2

Large-scale cultures were performed in 2 l flasks using the same conditions as described for the expression screening, e.g. 10 flasks for a final cell culture volume of 2 l. Cell pellets were resuspended at a concentration of 0.2 g/ml in 50 mM sodium phosphate, pH 7.4, 1 mM EDTA, 5% glycerol and complete EDTA-free protease inhibitors (Roche). The cells were then lysed in a continuous flow Cell

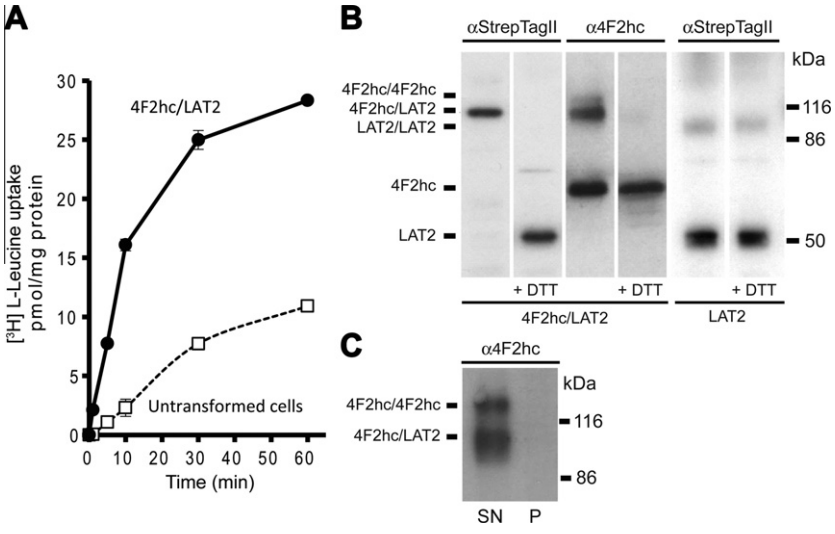


Fig. 3. Expression, function and solubilization of the human heterodimer 4F2hc/LAT2. (A) [^{3}H]L-leucine (10 μM) uptake into Pichia cells overexpressing 4F2hc/LAT2. Clear uptake of the radioligand is observed when compared with control (untransformed) cells. Bars represent mean ± SEM from triplicates of two independent experiments. (B) Western blot analysis of membranes isolated from Pichia cells expressing 4F2hc and LAT2. The anti-StrepTagII antibody indicated the presence of heterodimers only and no LAT2 monomers under non-reducing conditions. Under reducing conditions (+ DTT), the heterodimer disappeared yielding the LAT2 monomer band. The anti-4F2hc antibody indicated the presence of 4F2hc/4F2hc homodimer, 4F2hc/LAT2 heterodimer and 4F2hc monomer in the absence of DTT. Under reducing conditions (+ DTT), only 4F2hc monomers were visible. Western blot analysis using anti-StrepTagII antibody of isolated membranes from LAT2 expressing Pichia cells indicated the presence of LAT2 (major fraction; at ~50 kDa) and an artefactual LAT2 homodimer (minor fraction; at ~100 kDa), both under reducing and non-reducing conditions. Such dimerizations of light subunits in PAGE have been reported [41]. (C) The detergent DDM completely solubilized 4F2hc/LAT2 and 4F2hc/4F2hc from isolated membranes: all protein is found in the supernatant (SN) and none in the pellet (P) after ultracentrifugation.

Disrupter (Constant Systems TS series) in two rounds of 20,000 and 40,000 psi. After centrifugation at 20,000*g* and 4° C for 15 min, the supernatant was recovered and ultracentrifuged at 250,000*g* and 4° C for 1 h. The pellet was resuspended and homogenized (50 ml glass homogenizer) in 20 mM Tris–HCl, pH 7.4, 150 mM NaCl, 5% glycerol, and ultracentrifuged again. Finally, the membrane pellet was resuspended in 20 mM Tris–HCl, pH 7.4, 150 mM NaCl, 5% glycerol at a protein concentration between 18 and 29 mg/ml. Aliquots were frozen in liquid nitrogen and stored at −80 °C until use.

Frozen *P. pastoris* membranes containing overexpressed 4F2hc/LAT2 were thawed and solubilized for 1 h at 4 °C under gentle agitation in 2% *n*-dodecyl-β-D-maltopyranoside (DDM; Anatrace), 20 mM Tris–HCl, pH 8, 300 mM NaCl and 10% glycerol. The protein concentration during solubilization was 3 mg/ml. After ultracentrifugation for 1 h (100,000*g* at 4 °C), 4F2hc/LAT2 was purified by two sequential affinity chromatography steps. The first one was a metal affinity chromatography using cobalt beads (Talon, BD Science) and 20 mM Tris–HCl, pH 8, 300 mM NaCl, 10% glycerol, 20 mM imidazole, 0.25% DDM as washing buffer. 4F2hc/LAT2 was eluted with washing buffer supplemented with 200 mM imidazole. Next, the eluted heterodimer was purified by Strep-Tactin affinity chromatography (IBA) using 20 mM Tris–HCl, pH 8, 300 mM NaCl, 10% glycerol, 0.1% DDM as washing buffer. The protein was eluted with the same buffer supplemented with 3 mM desthiobiotin. The homogeneity of the purified 4F2hc/LAT2 was estimated by size-exclusion chromatography using a Superose 6 10/300 GL column (GE Healthcare). Protein concentration was determined by BCA assay.

Results and discussion

Expression screening of human HATs in P. pastoris

The cDNAs encoding the human heavy subunits (4F2hc and rBAT) were cloned into the pPICZB vector. After transformation of the constructs into the *P. pastoris* KM71H strain, expression tests were performed. His-tagged recombinant proteins were visualized by Western blot analysis from isolated membranes using anti-4F2hc and anti-rBAT antibodies, respectively (Fig. 1). Bands were observed at the expected size of monomers of 4F2hc (∼70 kDa) and rBAT (∼90 kDa) (Fig. 1B). Human 4F2hc is a glycoprotein with four putative *N*-glycosylation sites (N264, N280, N323 and N405). Its expression generated at least two bands corresponding to unglycosylated and glycosylated 4F2hc forms (Fig. 1B, left lane). The lowest band was most abundant representing the unglycosylated form of 4F2hc as demonstrated by endoglycosidase F treatment (Fig. 1B, right lane). The expression of rBAT rendered in addition to the monomer band (∼90 kDa) a diffuse, lower band (<75 kDa), most probably representing proteolytic fragments (Fig. 1B). When 4F2hc and rBAT were solubilized from Pichia membranes using DDM, only 4F2hc was extracted efficiently (Fig. 1C). The dramatic low solubilization yield for rBAT monomers suggests missfolding of this human protein when expressed in yeast. Recently, it has been reported that $b^{0,+}$AT, the light subunit associated to rBAT, facilitates the proper folding of rBAT in mammalian cells [40]. In order to determine if the co-expression of rBAT and $b^{0,+}$AT may increase the stability of the heavy subunit, we tried initially the expression of $b^{0,+}$AT alone as a Strep-tagged protein to check the capability of Pichia to express this transporter. $b^{0,+}$AT was not produced efficiently in this expression system (Fig. 2B) and therefore, the co-expression of both subunits (rBAT and $b^{0,+}$AT) was not considered.

The human light subunits associated to 4F2hc (asc1, LAT1, LAT2, y^+LAT1, y^+LAT2 and xCT) were expressed in *P. pastoris* as Strep-tagged proteins as well (Fig. 2B). Recombinant proteins were visualized by Western blotting using anti-StrepTagII antibody. Bands were detected at expected sizes, i.e., between 40 and 50 kDa (Fig. 2B). LAT1, LAT2 and y^+LAT2 were the light subunits with the best expression levels, although only LAT2 was efficiently solubilized with DDM (Fig. 2C). In Western blots, monomer and dimer bands of LAT2 were discerned at ∼50 and ∼100 kDa, respectively (Figs. 2B and 3B, rightmost two lanes). Such unspecific dimerizations have been reported for 4F2hc light subunits subjected to SDS–PAGE but not in native plasma membranes [41].

Expression of the human 4F2hc/LAT2 heterodimer in P. pastoris

Since the best light and heavy subunits, in terms of expression levels and detergent solubilization yields, were 4F2hc and LAT2, the 4F2hc/LAT2 heterodimer was expressed in *P. pastoris*. Uptake experiments with yeast cells expressing 4F2hc/LAT2 indicated transport of L-leucine (Fig. 3A), a characteristic substrate of this HAT [42]. To estimate the expression level of this heterodimer in isolated membranes by Western blotting, anti-StrepTagII and anti-4F2hc antibodies were used. In both cases, a band at ∼120 kDa appeared in the absence of DTT, but collapsed when samples were analyzed under reducing conditions (Fig. 3B). Thus, only LAT2 (∼50 kDa) and 4F2hc (∼70 kDa) were detected in the presence of DTT. We concluded that recombinant human 4F2hc/LAT2 heterodimer is functional and correctly expressed in *P. pastoris*, and contains the characteristic disulphide bridge connecting both subunits. A less conspicuous band corresponding to 4F2hc homodimers (∼140 kDa; above of the heterodimer band) was also detected, but only with anti-4F2hc antibodies under non-reducing conditions. In Western blots performed in the presence of DTT, this band collapsed to the height of 4F2hc monomers (Fig. 3B). Finally, 4F2hc monomers (∼70 kDa), but not LAT2 monomers or dimers, were detected under non-reducing conditions showing the excess expression of 4F2hc over LAT2 in isolated membranes (Fig. 3B). DDM efficiently solubilized 4F2hc/LAT2 heterodimer, 4F2hc homodimer (Fig. 3C) and 4F2hc monomers (data not shown). 4F2hc/LAT2 heterodimer was purified using DDM during the whole process by two sequential affinity chromatography steps: first by cobalt and then by Strep-Tactin affinity chromatography. The purified heterodimer was analyzed by SDS–PAGE, Western blotting and size exclusion chromatography (SEC) (Fig. 4). As seen in the silver-stain SDS–PAGE gel, the most abundant species corresponded to the 4F2hc/LAT2 heterodimer (Fig. 4A, lane 1). In addition, a low contribution of 4F2hc monomers and heterodimer aggregates/oligomers was found. The presence of the 4F2hc/LAT2 heterodimer, and 4F2hc and LAT2 monomers was also detected by Western blotting (Fig. 4A, lanes 2 and 3). Importantly, Western blot analysis confirmed the absence of 4F2hc dimers indicating the successful application of our approach of using two different tags on the two subunits to avoid co-purification of 4F2hc dimers. To illustrate the disruption of the heterodimer under reducing conditions, purified 4F2hc/LAT2 was incubated with DTT and subjected to SDS–PAGE. Two major bands corresponding to glycosylated 4F2hc and LAT2 are discerned, as well as a weak band of residual, non-disrupted 4F2hc/LAT2 heterodimer (Fig. 4A, lane 4). Interestingly, no aggregates/oligomers were found under reducing conditions (Fig. 4A, lane 4) compared to SDS–PAGE under non-reducing conditions (Fig. 4A, lane 1). This indicates that the observed higher molecular weight bands (Fig. 4A, lane 1) might be aggregates/oligomers formed during purification through oxidation of cysteine residues into disulphide bridges. The average yield was ∼1 mg of purified heterodimer per 2 l of yeast culture (see Table 1). SEC profiles of purified 4F2hc/LAT2 indicated a minor (at ∼13 ml) and a major peak (at ∼14 ml), both being relatively broad. A small amount of heavy protein aggregation was found in the void volume (Fig. 4B). The presence of 4F2hc and LAT2 monomers (as a

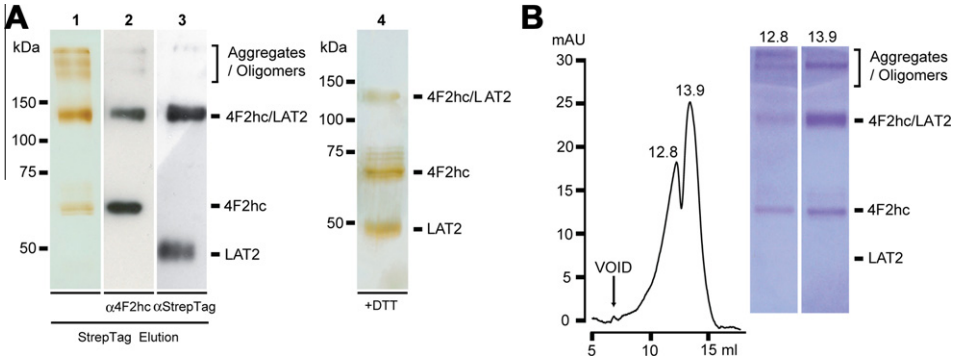


Fig. 4. SDS–PAGE, Western blot analysis and SEC of purified 4F2hc/LAT2. (A) Silver stained SDS–PAGE gel (lane 1), and anti-4F2hc (lane 2) and anti-StrepTagII (lane 3) Western blotting of purified 4F2hc/LAT2. The main silver stained fraction is composed of 4F2hc/LAT2 heterodimers. Minor silver stained fractions of 4F2hc monomers and aggregates/oligomers are also visible. Western blot analysis of purified 4F2hc/LAT2 (lane 2) using anti-4F2hc antibody yielded a similar pattern as with silver staining (lane 1). Western blotting using anti-StrepTagII antibody revealed two major bands corresponding to the heterodimer 4F2hc/LAT2 and the LAT2 monomer (lane 3). Lane 4 displays a silver stained SDS–PAGE gel of purified 4F2hc/LAT2 run under reducing conditions, i.e. in the presence of DTT (100 mM DTT in sample buffer). Bands for glycosylated 4F2hc and LAT2 are discerned, as well as a weak band of residual, non-disrupted 4F2hc/LAT2 heterodimer. (B) SEC of purified 4F2hc/LAT2. Two peaks were resolved at ~13 and ~14 ml corresponding to 4F2hc/LAT2 aggregates/oligomers and 4F2hc/LAT2 heterodimers – see also Coomassie blue stained SDS–PAGE gels of the corresponding peak fractions. If not specified, SDS–PAGE was always performed under non-reducing conditions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Purification table for the recombinant human 4F2hc/LAT2 heterodimer. Yield and composition of the purified proteins from isolated membranes (1200 mg total protein, i.e. from 2 l of cell culture) are indicated.

Purification step	Yield (mg)	Protein composition
Co^{2+} affinity chromatography	9.6–14.4	4F2hc/4F2hc 4F2hc LAT2 4F2hc/LAT2
Strep-Tactin affinity chromatography	0.8–0.96	4F2hc LAT2 4F2hc/LAT2

consequence of the reduction of the inter-subunit disulphide bridge during the process of membrane solubilization and protein purification), and of heterodimer aggregation in the SDS–PAGE gel (Fig. 4A, lane 2) explained the broad SEC profiles. SDS–PAGE confirmed the major contribution of 4F2hc/LAT2 heterodimer to fractions of the main peak at ~14 ml (Fig. 4B, right lane). In contrast, fractions of the less prominent peak at ~13 ml contained higher amounts of aggregates/oligomers than heterodimers (Fig. 4B, left lane). In summary, while additional efforts will be needed to improve the integrity of the inter-subunit disulphide bridge, we managed to isolate intact human 4F2hc/LAT2 heterodimer. The present study represents the first successful attempt to express and purify a HAT transporter at the milligram scale for future structural studies.

Acknowledgments

Financial support is gratefully acknowledged from the Spanish Ministry of Science and Innovation Grant SAF2009-12606-C02-01, the European Commission Frame Program 7 Grant 201924 EDICT and the Generalitat de Catalunya SGR2009-1355 (M.P.), and from the University of Bern, the Swiss National Science Foundation (Grants 31003A_125150 and 31003A_144168), the Bern University Research Foundation and the National Centre of Competence in Research (NCCR) TransCure (D.F.).

References

[1] F. Verrey, E.I. Closs, C.A. Wagner, M. Palacin, H. Endou, Y. Kanai, CATs and HATs: the SLC7 family of amino acid transporters, Pflugers Arch. 447 (2004) 532–542.
[2] M. Palacin, Y. Kanai, The ancillary proteins of HATs: SLC3 family of amino acid transporters, Pflugers Arch. 447 (2004) 490–494.
[3] D. Fotiadis, Y. Kanai, M. Palacín, The SLC3 and SLC7 families of amino acid transporters, Mol. Aspects Med.
[4] L. Mastroberardino, B. Spindler, R. Pfeiffer, P.J. Skelly, J. Loffing, C.B. Shoemaker, F. Verrey, Amino-acid transport by heterodimers of 4F2hc/CD98 and members of a permease family, Nature 395 (1998) 288–291.
[5] D. Torrents, R. Estevez, M. Pineda, E. Fernandez, J. Lloberas, Y.B. Shi, A. Zorzano, M. Palacin, Identification and characterization of a membrane protein (y+L amino acid transporter-1) that associates with 4F2hc to encode the amino acid transport activity y+L. A candidate gene for lysinuric protein intolerance, J. Biol. Chem. 273 (1998) 32437–32445.
[6] N. Reig, J. Chillaron, P. Bartoccioni, E. Fernandez, A. Bendahan, A. Zorzano, B. Kanner, M. Palacin, J. Bertran, The light subunit of system b(0,+) is fully functional in the absence of the heavy subunit, EMBO J. 21 (2002) 4906–4914.
[7] A.E. Busch, T. Herzer, S. Waldegger, F. Schmidt, M. Palacin, J. Biber, D. Markovich, H. Murer, F. Lang, Opposite directed currents induced by the transport of dibasic and neutral amino acids in Xenopus oocytes expressing the protein rBAT, J. Biol. Chem. 269 (1994) 25581–25586.
[8] J. Chillaron, R. Estevez, C. Mora, C.A. Wagner, H. Suessbrich, F. Lang, J.L. Gelpi, X. Testar, A.E. Busch, A. Zorzano, M. Palacin, Obligatory amino acid exchange via systems b0,+-like and y+L-like. A tertiary active transport mechanism for renal reabsorption of cystine and dibasic amino acids, J. Biol. Chem. 271 (1996) 17761–17770.
[9] M.J. Calonge, P. Gasparini, J. Chillaron, M. Chillon, M. Gallucci, F. Rousaud, L. Zelante, X. Testar, B. Dallapiccola, F. Di Silverio, P. Barceló, X. Estivill, A. Zorzano, V. Nunes, M. Palacín, Cystinuria caused by mutations in rBAT, a gene involved in the transport of cystine, Nat. Genet. 6 (1994) 420–425.
[10] L. Feliubadalo, M. Font, J. Purroy, F. Rousaud, X. Estivill, V. Nunes, E. Golomb, M. Centola, I. Aksentijevich, Y. Kreiss, B. Goldman, M. Pras, D.L. Kastner, E. Pras, P. Gasparini, L. Bisceglia, E. Beccia, M. Gallucci, L. de Sanctis, A. Ponzone, G.F. Rizzoni, L. Zelante, M.T. Bassi, A.L. George Jr., M. Manzoni, A. De Grandi, M. Riboni, J.K. Endsley, A. Ballabio, G. Borsani, N. Reig, E. Fernandez, R. Estevez, M. Pineda, D. Torrents, M. Camps, J. Lloberas, A. Zorzano, M. Palacin, Non-type I cystinuria caused by mutations in SLC7A9, encoding a subunit (bo,+AT) of rBAT, Nat. Genet. 23 (1999) 52–57.
[11] H. Sato, A. Shiiya, M. Kimata, K. Maebara, M. Tamba, Y. Sakakura, N. Makino, F. Sugiyama, K. Yagami, T. Moriguchi, S. Takahashi, S. Bannai, Redox imbalance in cystine/glutamate transporter-deficient mice, J. Biol. Chem. 280 (2005) 37423–37429.
[12] J.A. Kaleeba, E.A. Berger, Kaposi's sarcoma-associated herpesvirus fusion-entry receptor: cystine transporter xCT, Science 311 (2006) 1921–1924.
[13] M.V. Veettil, S. Sadagopan, N. Sharma-Walia, F.Z. Wang, H. Raghu, L. Varga, B. Chandran, Kaposi's sarcoma-associated herpesvirus forms a multimolecular complex of integrins (alphaVbeta5, alphaVbeta3, and alpha3beta1) and CD98–xCT during infection of human dermal microvascular endothelial cells, and CD98–xCT is essential for the postentry stage of infection, J. Virol. 82 (2008) 12126–12144.
[14] D.A. Baker, K. McFarland, R.W. Lake, H. Shen, X.C. Tang, S. Toda, P.W. Kalivas, Neuroadaptations in cystine–glutamate exchange underlie cocaine relapse, Nat. Neurosci. 6 (2003) 743–749.
[15] P. Nicklin, P. Bergman, B. Zhang, E. Triantafellow, H. Wang, B. Nyfeler, H. Yang, M. Hild, C. Kung, C. Wilson, V.E. Myer, J.P. MacKeigan, J.A. Porter, Y.K. Wang, L.C. Cantley, P.M. Finan, L.O. Murphy, Bidirectional transport of amino acids regulates mTOR and autophagy, Cell 136 (2009) 521–534.

[16] C.C. Feral, N. Nishiya, C.A. Fenczik, H. Stuhlmann, M. Slepak, M.H. Ginsberg, CD98hc (SLC3A2) mediates integrin signaling, Proc. Natl. Acad. Sci. USA 102 (2005) 355–360.
[17] J. Cantor, C.D. Browne, R. Ruppert, C.C. Feral, R. Fassler, R.C. Rickert, M.H. Ginsberg, CD98hc facilitates B cell proliferation and adaptive humoral immunity, Nat. Immunol. 10 (2009) 412–419.
[18] P. Fogelstrand, C.C. Feral, R. Zargham, M.H. Ginsberg, Dependence of proliferative vascular smooth muscle cells on CD98hc (4F2hc, SLC3A2), J. Exp. Med. 206 (2009) 2397–2406.
[19] M. Gabrisko, S. Janecek, Looking for the ancestry of the heavy-chain subunits of heteromeric amino acid transporters rBAT and 4F2hc within the GH13 alpha-amylase family, FEBS J. 276 (2009) 7265–7278.
[20] J. Fort, L.R. de la Ballina, H.E. Burghardt, C. Ferrer-Costa, J. Turnay, C. Ferrer-Orta, I. Uson, A. Zorzano, J. Fernandez-Recio, M. Orozco, M.A. Lizarbe, I. Fita, M. Palacin, The structure of human 4F2hc ectodomain provides a model for homodimerization and electrostatic interaction with plasma membrane, J. Biol. Chem. 282 (2007) 31444–31452.
[21] E. Gasol, M. Jimenez-Vidal, J. Chillaron, A. Zorzano, M. Palacin, Membrane topology of system xc- light subunit reveals a re-entrant loop with substrate-restricted accessibility, J. Biol. Chem. 279 (2004) 31228–31236.
[22] M. Jimenez-Vidal, E. Gasol, A. Zorzano, V. Nunes, M. Palacin, J. Chillaron, Thiol modification of cysteine 327 in the eighth transmembrane domain of the light subunit xCT of the heteromeric cystine/glutamate antiporter suggests close proximity to the substrate binding site/permeation pathway, J. Biol. Chem. 279 (2004) 11214–11221.
[23] P. Bartoccioni, C. Del Rio, M. Ratera, L. Kowalczyk, J.M. Baldwin, A. Zorzano, M. Quick, S.A. Baldwin, J.L. Vazquez-Ibar, M. Palacin, Role of transmembrane domain 8 in substrate selectivity and translocation of SteT, a member of the L-amino acid transporter (LAT) family, J. Biol. Chem. 285 (2010) 28764–28776.
[24] X. Gao, F. Lu, L. Zhou, S. Dang, L. Sun, X. Li, J. Wang, Y. Shi, Structure and mechanism of an amino acid antiporter, Science 324 (2009) 1565–1568.
[25] Y. Fang, H. Jayaram, T. Shane, L. Kolmakova-Partensky, F. Wu, C. Williams, Y. Xiong, C. Miller, Structure of a prokaryotic virtual proton pump at 3.2 Å resolution, Nature 460 (2009) 1040–1043.
[26] X. Gao, L. Zhou, X. Jiao, F. Lu, C. Yan, X. Zeng, J. Wang, Y. Shi, Mechanism of substrate recognition and transport by an amino acid antiporter, Nature 463 (2010) 828–832.
[27] L. Kowalczyk, M. Ratera, A. Paladino, P. Bartoccioni, E. Errasti-Murugarren, E. Valencia, G. Portella, S. Bial, A. Zorzano, I. Fita, M. Orozco, X. Carpena, J.L. Vazquez-Ibar, M. Palacin, Molecular basis of substrate-induced permeation by an amino acid antiporter, Proc. Natl. Acad. Sci. USA 108 (2011) 3935–3940.
[28] P.L. Shaffer, A. Goehring, A. Shankaranarayanan, E. Gouaux, Structure and mechanism of a Na^+-independent amino acid transporter, Science 325 (2009) 1010–1014.
[29] D. Ma, P. Lu, C. Yan, C. Fan, P. Yin, J. Wang, Y. Shi, Structure and mechanism of a glutamate-GABA antiporter, Nature 483 (2012) 632–636.
[30] A. Yamashita, S.K. Singh, T. Kawate, Y. Jin, E. Gouaux, Crystal structure of a bacterial homologue of Na^+/Cl^-dependent neurotransmitter transporters, Nature 437 (2005) 215–223.
[31] S. Faham, A. Watanabe, G.M. Besserer, D. Cascio, A. Specht, B.A. Hirayama, E.M. Wright, J. Abramson, The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na^+/sugar symport, Science 321 (2008) 810–814.
[32] S. Weyand, T. Shimamura, S. Yajima, S. Suzuki, O. Mirza, K. Krusong, E.P. Carpenter, N.G. Rutherford, J.M. Hadden, J. O'Reilly, P. Ma, M. Saidijam, S.G. Patching, R.J. Hope, H.T. Norbertczak, P.C. Roach, S. Iwata, P.J. Henderson, A.D. Cameron, Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter, Science 322 (2008) 709–713.
[33] S. Ressl, A.C. Terwisscha van Scheltinga, C. Vonrhein, V. Ott, C. Ziegler, Molecular basis of transport and regulation in the Na(+)/betaine symporter BetP, Nature 458 (2009) 47–52.
[34] M. Karlsson, D. Fotiadis, S. Sjovall, I. Johansson, K. Hedfalk, A. Engel, P. Kjellbom, Reconstitution of water channel function of an aquaporin overexpressed and purified from *Pichia pastoris*, FEBS Lett. 537 (2003) 68–72.
[35] S.B. Long, E.B. Campbell, R. Mackinnon, Crystal structure of a mammalian voltage-dependent Shaker family K^+ channel, Science 309 (2005) 897–903.
[36] J.D. Ho, R. Yeh, A. Sandstrom, I. Chorny, W.E. Harries, R.A. Robbins, L.J. Miercke, R.M. Stroud, Crystal structure of human aquaporin 4 at 1.8 Å and its mechanism of conductance, Proc. Natl. Acad. Sci. USA 106 (2009) 7437–7442.
[37] X. Tao, J.L. Avalos, J. Chen, R. MacKinnon, Crystal structure of the eukaryotic strong inward-rectifier K+ channel Kir2.2 at 3.1 Å resolution, Science 326 (2009) 1668–1674.
[38] E. Fernandez, M. Carrascal, F. Rousaud, J. Abian, A. Zorzano, M. Palacin, J. Chillaron, rBAT-b(0,+)AT heterodimer is the main apical reabsorption system for cystine in the kidney, Am. J. Physiol. Renal Physiol. 283 (2002) F540–F548.
[39] F. Döring, S. Theis, H. Daniel, Expression and functional characterization of the mammalian intestinal peptide transporter PepT1 in the methylotropic yeast *Pichia pastoris*, Biochem. Biophys. Res. Commun. 232 (1997) 656–662.
[40] M. Rius, J. Chillaron, Carrier subunit of plasma membrane transporter is required for oxidative folding of its helper subunit, J. Biol. Chem. 287 (2012) 18190–18200.
[41] E. Fernandez, M. Jimenez-Vidal, M. Calvo, A. Zorzano, F. Tebar, M. Palacin, J. Chillaron, The structural and functional units of heteromeric amino acid transporters. The heavy subunit rBAT dictates oligomerization of the heteromeric amino acid transporters, J. Biol. Chem. 281 (2006) 26552–26561.
[42] M. Pineda, E. Fernandez, D. Torrents, R. Estevez, C. Lopez, M. Camps, J. Lloberas, A. Zorzano, M. Palacin, Identification of a membrane protein, LAT-2, that co-expresses with 4F2 heavy chain, an L-type amino acid transport activity with broad specificity for small and large zwitterionic amino acids, J. Biol. Chem. 274 (1999) 19738–19744.

ARTICLE: Rosell *et al*., Proceedings of the National Academy of Sciences, 2014

Structural bases for the interaction and stabilization of the human amino acid transporter LAT2 with its ancillary protein 4F2hc

Proceedings of the National Academy of Sciences Vol. 111, pp.2966-2971, 2014

Albert Rosell, Marcel Meury, Elena Álvarez-Marimon, Meritxell Costa, Laura Pérez-Cano, Antonio Zorzano,, Juan Fernández-Recio, Manuel Palacín, and Dimitrios Fotiadis.

Structural bases for the interaction and stabilization of the human amino acid transporter LAT2 with its ancillary protein 4F2hc

Albert Rosell[a,b,1], Marcel Meury[c,d,1], Elena Álvarez-Marimon[a,b,1], Meritxell Costa[a,b,c,d,1], Laura Pérez-Cano[e], Antonio Zorzano[a,f,g], Juan Fernández-Recio[e,2], Manuel Palacín[a,b,f,2], and Dimitrios Fotiadis[c,d,2]

[a]Institute for Research in Biomedicine (IRB Barcelona), E-08028 Barcelona, Spain; [b]Centro de Investigación Biomédica en Red de Enfermedades Raras, E-08028 Barcelona, Spain; [c]Institute of Biochemistry and Molecular Medicine, and [d]Swiss National Centre of Competence in Research TransCure, University of Bern, CH-3012 Bern, Switzerland; [e]Joint Barcelona Supercomputing Center-IRB Program in Computational Biology, Department of Life Sciences, Barcelona Supercomputing Center, E-08034 Barcelona, Spain; [f]Department of Biochemistry and Molecular Biology, Faculty of Biology, University of Barcelona, E-08028 Barcelona, Spain; and [g]Centro de Investigación Biomédica en Red de Diabetes y Enfermedades Metabólicas Asociadas, E-08028 Barcelona, Spain

Edited* by Christopher Miller, Howard Hughes Medical Institute, Brandeis University, Waltham, MA, and approved January 15, 2014 (received for review December 24, 2013)

Heteromeric amino acid transporters (HATs) are the unique example, known in all kingdoms of life, of solute transporters composed of two subunits linked by a conserved disulfide bridge. In metazoans, the heavy subunit is responsible for the trafficking of the heterodimer to the plasma membrane, and the light subunit is the transporter. HATs are involved in human pathologies such as amino acidurias, tumor growth and invasion, viral infection and cocaine addiction. However structural information about interactions between the heavy and light subunits of HATs is scarce. In this work, transmission electron microscopy and single-particle analysis of purified human 4F2hc/L-type amino acid transporter 2 (LAT2) heterodimers overexpressed in the yeast *Pichia pastoris*, together with docking analysis and crosslinking experiments, reveal that the extracellular domain of 4F2hc interacts with LAT2, almost completely covering the extracellular face of the transporter. 4F2hc increases the stability of the light subunit LAT2 in detergent-solubilized *Pichia* membranes, allowing functional reconstitution of the heterodimer into proteoliposomes. Moreover, the extracellular domain of 4F2hc suffices to stabilize solubilized LAT2. The interaction of 4F2hc with LAT2 gives insights into the structural bases for light subunit recognition and the stabilizing role of the ancillary protein in HATs.

CD98hc | 4F2hc ectodomain

Heteromeric amino acid transporters (HATs) are composed of two subunits, a heavy (SLC3 family) and a light subunit [SLC7 or L-type amino acid transporter (LAT) family] linked by a conserved disulfide bridge (1). HATs are amino acid exchangers (1), and this transport activity resides in the light subunit (2). The heavy subunit (either 4F2hc or rBAT) is essential for trafficking of the holotransporter to the plasma membrane (3, 4). In mammals, six transporters heterodimerize with 4F2hc, and only one heterodimerizes with rBAT. The rBAT/$b^{0,+}$AT complex is a dimer of heterodimers in which the light subunit is required for proper rBAT folding and stability (5, 6). In contrast, 4F2hc-associated transporters are simple heterodimers (6), and possible stabilizing roles of the two subunits in the biogenesis of the heterodimer have not been described.

HATs have major impacts on human health and are involved directly in amino acidurias (cystinuria and lysinuric protein intolerance), tumor cell growth, glioma invasion, Kaposi's sarcoma-associated herpesvirus infection, and cocaine relapse (1). In addition to the role of HATs in amino acid transport, 4F2hc heterodimers mediate β1- and β3-integrin signaling (7).

Structural information about HATs is scarce (1). The heavy subunits are type II membrane *N*-glycoproteins with a single transmembrane domain (TMD), an intracellular N terminus, and a large extracellular C terminus with sequence homology with bacterial α-amylases. Indeed, the atomic structure of the extracellular domain (ED) of human 4F2hc (4F2hc-ED) is similar to that of bacterial glucosidases [i.e., a triose phosphate isomerase barrel, $(\beta\alpha)_8$, (subdomain A) and eight antiparallel β-strands (subdomain C)] but lacks glucosidase activity (8). The conserved cysteine residue participating in the intersubunit disulfide bridge is located between the single TMD and 4F2hc-ED. Physical and functional interaction of 4F2hc with integrins has been mapped to the TMD and cytosolic N-terminal domain (7, 9), whereas 4F2hc-ED is necessary for functional heterodimerization with the light subunit (9, 10). The light subunits are nonglycosylated proteins and have a 12-TMD topology with intracellular N and C termini (1). The conserved cysteine residue involved in the intersubunit disulfide bridge is located between TMD3 and TMD4 (11). Structure–function studies (12, 13) suggest that the structures of the light subunit of HATs should be similar to those of remote bacterial amino acid/polyamine/organocation (APC) transporters (~20% amino acid sequence identities), which present the LeuT-fold (14–17). In contrast, there is no structural information about the interaction between the two HAT subunits. In the present study, we show that information from transmission electron microscopy (TEM) and single-particle analysis (SPA) of human 4F2hc/L-type amino acid transporter 2 (LAT2) heterodimers is compatible with 4F2hc-ED interacting with the extracellular loops of LAT2. Docking analyses and crosslinking experiments indicated a location of 4F2hc-ED almost completely covering the external face of LAT2. Moreover, 4F2hc increases the stability of detergent-solubilized

Significance

Here we report the structural bases of the interaction between the catalytic light subunit and the heavy subunit of the heteromeric amino acid transporters. Our data show that the large ectodomain of the human heavy subunit 4F2hc plays a key role in the interaction and stability of the light subunit L-type amino acid transporter 2. This finding paves the way for structural and functional studies to elucidate the role of the heavy subunit in the regulation of these transporters; such studies will be highly relevant in human pathology.

Author contributions: A.R., M.M., E.Á.-M., M.C., L.P.-C., J.F.-R., M.P., and D.F. designed research; A.R., M.M., E.Á.-M., M.C., and L.P.-C. performed research; A.R., M.M., E.Á.-M., M.C., L.P.-C., A.Z., J.F.-R., M.P., and D.F. analyzed data; and J.F.-R., M.P., and D.F. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

[1]A.R., M.M., E.Á.-M., and M.C. contributed equally to this work.

[2]To whom correspondence may be addressed. E-mail: juanf@bsc.es, manuel.palacin@irbbarcelona.org, or dimitrios.fotiadis@ibmm.unibe.ch.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1323779111/-/DCSupplemental.

LAT2, allowing functional reconstitution of the heterodimer into proteoliposomes. The interaction of 4F2hc-ED with LAT2 provides insights regarding the structural bases for recognition of the light subunit and the stabilization of this complex and other 4F2hc associated transporters.

Results

TEM, SPA, and 3D Reconstruction of the Human 4F2hc/LAT2 Complex. Recombinant 4F2hc/LAT2 purified by Co^{2+}- and Strep-Tactin affinity chromatography (18) was analyzed by negative-stain TEM and SPA to unveil the supramolecular architecture of HATs (Fig. 1). On electron micrographs distinct particles corresponding to purified 4F2hc/LAT2 complexes are discerned as composed of two globular domains of different sizes (Fig. 1*A*, boxed particles and gallery). From 15,210 single-particle projections, class averages were calculated (Fig. S1*A*). Most class averages indicate a clear difference in size between the two domains. A few smaller particles reflecting 4F2hc or LAT2 from disrupted heterodimers were also found on electron micrographs (Fig. 1*A*, arrowheads). In addition, some small protein aggregates are seen on electron micrographs (Fig. 1*A*, asterisks). We determined the structure of 4F2hc/LAT2 by 3D reconstruction from negatively stained complexes (Fig. 1*B*). The resolution of the 3D map obtained from the final refinement run was 21 Å (Fig. S1*B*) and had a homogeneous angular distribution of projections (Fig. S1*C*). The calculated 3D map reflects a structure composed of a smaller and a larger density (Fig. 1*B*; see Fig. S1*D* for additional views). The smaller domain lies tilted (not flat) on the larger domain (Fig. 1*B*, black dotted line). Consequently, the complex has a cleft on one side (Fig. 1*B*, arrowhead) and a seal on the opposite side (Fig. 1*B*, white dotted curve). Importantly, the crystal structure of 4F2hc-ED fits nicely into the smaller density (Fig. 1*C*), making the assignment of 4F2hc and LAT2 in the 3D map possible. As determined by scanning TEM mass measurement, unstained n-dodecyl-β-D-maltoside (DDM)-purified membrane proteins prepared in a conventional way for TEM are associated with >55 kDa of copurified endogenous lipids and DDM molecules (19). Thus, the total mass of the LAT2/lipid/DDM ternary complex plus the single TMD and the cytoplasmic N-terminal domain of 4F2hc containing the His-tag and the protease cleavage site (>120 kDa) is significantly larger than that of the 4F2hc-ED (~50 kDa). This result, together with the fitting of the 4F2hc crystal structure into the smaller density, supports the assignment of LAT2 to the larger domain of the 3D map. In Fig. 1*C* the location of the N terminus in the 4F2hc-ED crystal structure is marked by an asterisk. This location is close to an additional density connecting the small and large domains that possibly arises from the N-terminal TMD of 4F2hc and extracellular loops of LAT2.

Crosslinking of 4F2hc-ED and LAT2. To confirm the interaction between 4F2hc-ED and LAT2, specific intersubunit crosslinking between cysteine residues were assessed in purified 4F2hc/LAT2 heterodimers (Fig. 2). To this end, homogenates from HEK293T cells transiently cotransfected with N-terminally His-tagged 4F2hc (His-4F2hc) and N-terminally Strep-TagII–tagged LAT2 (Strep-LAT2) were solubilized with DDM and purified by Ni^{2+}-affinity chromatography and then were crosslinked with the nonreducible cysteine crosslinkers bis(maleimido)ethane (BMOE) (with a spacer arm up to 10.5 Å) or 1,8-bis(maleimido)diethylene glycol [$BM(POE)_2$] (with a spacer arm up to 14.3 Å). Crosslinking experiments were performed in mammalian cells because mutants could be generated more conveniently in mammalian cells than in *Pichia* cells. Western blotting under reducing conditions and using an anti-Strep antibody revealed intersubunit crosslinking as DTT-resistant heterodimers. Human 4F2hc has two cysteine residues: Cys109, participating in the intersubunit disulfide bridge (located in the "neck" connecting the TMD and ectodomain) and Cys330, a partially hidden residue (located in the A-subdomain of the ectodomain). To avoid doubtful results, residue Cys330 was mutated to serine (C330S) in all mutants studied, and residue Cys109 was maintained to hold the disulfide intersubunit bridge. This strategy was validated by demonstrating heterodimerization of His-4F2hc C330S with Strep-LAT2 and induction of L-alanine transport in HEK293T cells (Fig. S2 *A–D*).

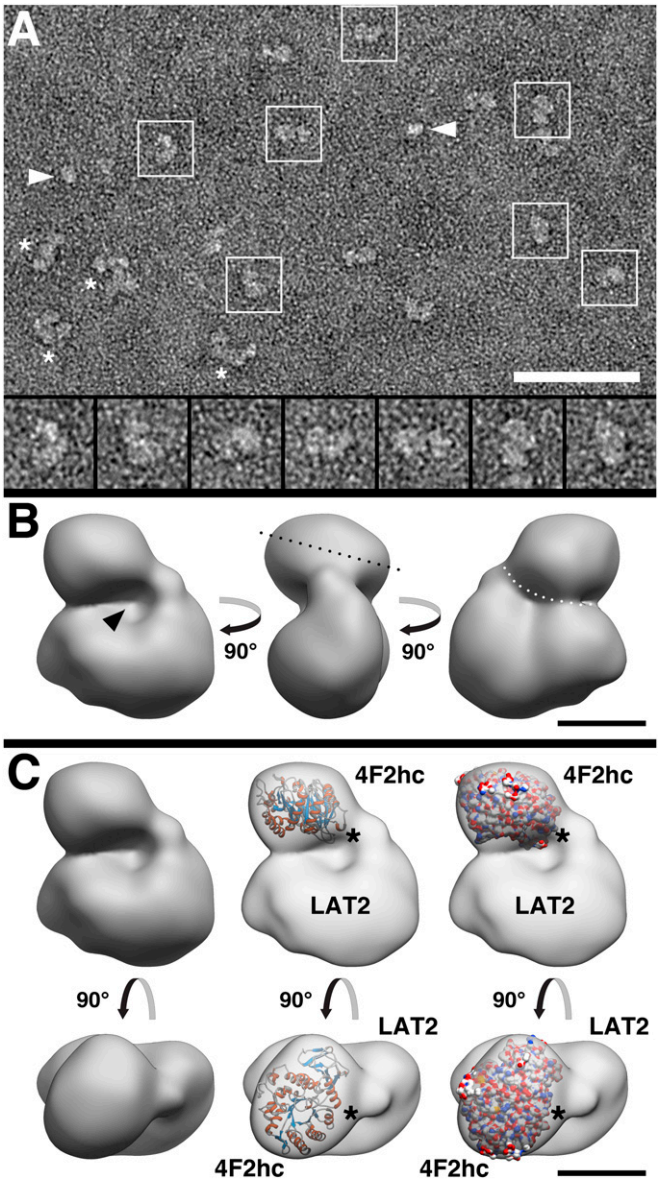


Fig. 1. TEM, SPA, and 3D reconstruction of human 4F2hc/LAT2. (*A*) Overview electron micrograph of purified, negatively stained 4F2hc/LAT2 heterodimers. The boxed 4F2hc/LAT2 complexes were magnified and are displayed in the gallery. Arrowheads mark 4F2hc or LAT2 monomers from disrupted heterodimers. Asterisks indicate small protein aggregates. (Scale bar: 50 nm.) The frame sizes of the magnified particles in the gallery are 21.8 nm. (*B*) 3D reconstruction of 4F2hc/LAT2 calculated from projections of negatively stained heterodimer particles. Different side views of the 3D model are shown. 4F2hc/LAT2 is composed of a large and a small density. The small density is located on top of the large density and is tilted, as indicated by the black dotted line. As indicated by an arrowhead, the 3D model features a distinct cavity. On the opposite side, both subunits are in close contact, as marked by the white, dotted curve. (Scale bar: 5 nm.) (*C*) Side (*Upper*) and top (*Lower*) views of the 4F2hc/LAT2 3D reconstruction without and with the fitted crystal structure of the 4F2hc-ED (Protein Data Bank ID code: 2DH2). The fitting assigns the small and large subunits to 4F2hc and LAT2, respectively. The structure of 4F2hc-ED is represented as a cartoon and surface model. Asterisks indicate the location of the N terminus in the 4F2hc-ED crystal structure. (Scale bar: 5 nm.)

BIOCHEMISTRY

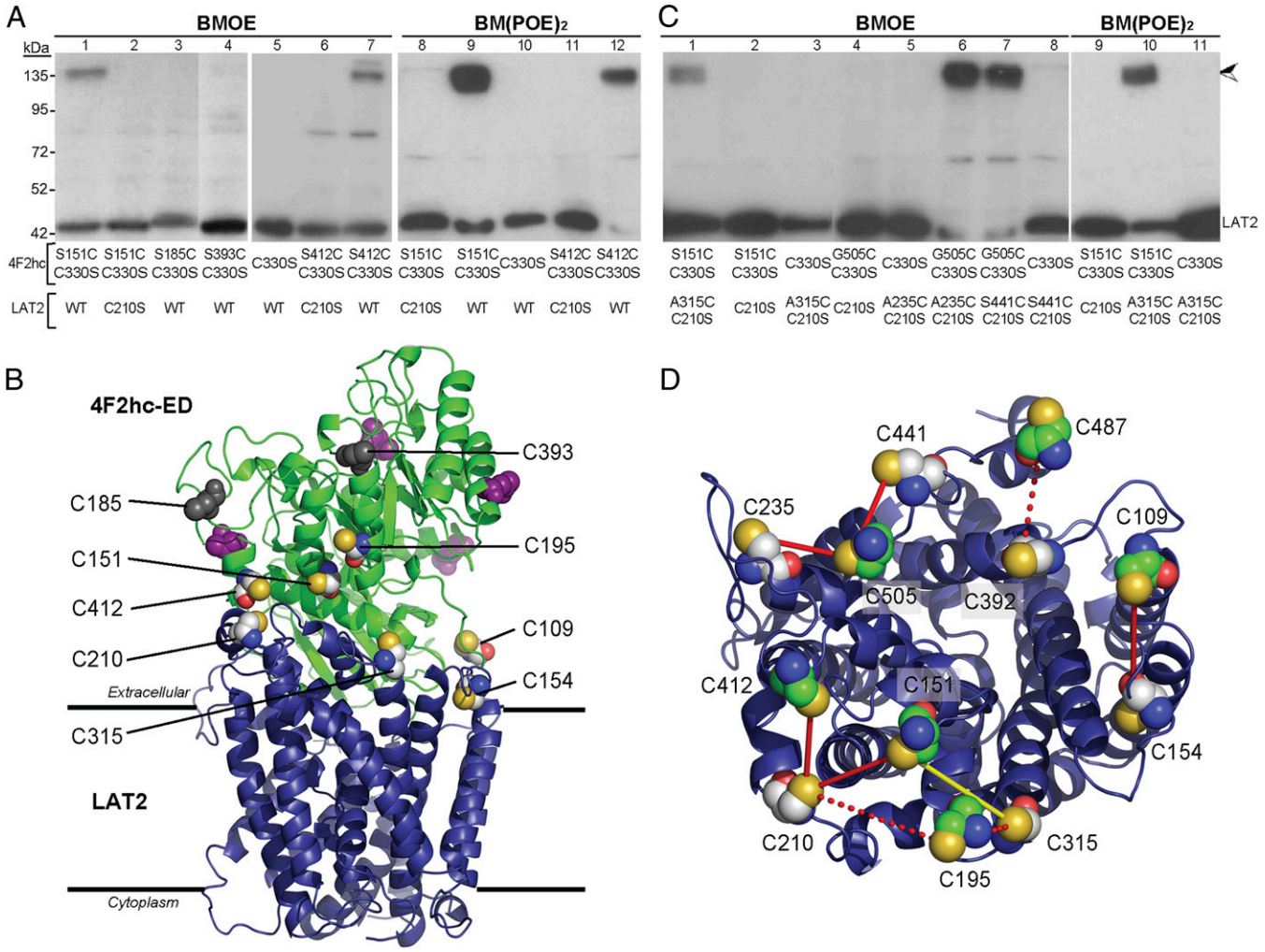


Fig. 2. Crosslinking of 4F2hc/LAT2 and docking model of the 4F2hc-ED and LAT2 complex. (*A* and *C*) Intersubunit crosslinking. Versions of His-4F2hc/Strep-LAT2 heterodimers expressed in HEK293T cells and purified by His-affinity chromatography were treated with the DTT-uncleavable crosslinkers BMOE or BM(POE)$_2$. Crosslinking was detected as DTT-resistant 4F2hc/LAT2 heterodimers (arrowhead) by Western blotting using αStrep antibodies under reducing conditions. The band of LAT2 monomer corresponds to LAT2 forming heterodimers with 4F2hc (i.e., copurified but not crosslinked). Single cysteine mutants were introduced in 4F2hc (C330S)/LAT2 wild-type (*A*) or LAT2 (C210S) (*C*). In all cases, crosslinking was abolished when 4F2hc (C330S) or LAT2 (C210S) was used, indicating the specificity of the crosslinked sites. (*B*) Lowest-energy model of the 4F2hc-ED–LAT2 complex. Highlighted cysteine residues (endogenous or added through mutation) are shown as sphere models (C atoms, gray; O atoms, red; N atoms, blue; S atoms, yellow). The putative *N*-glycosylation sites (N264, N280, N323, and N405 indicated in purple) and the non-crosslinked residues (C185 and C393 in gray) of 4F2hc are located in the most external face. (*D*) Summary of the crosslinked residues. The 4F2hc-ED–LAT2 model in *B* was rotated by 90°, i.e., viewed from 4F2hc-ED. Only the cysteine residues (endogenous or added through mutation) of the 4F2hc-ED are shown as sphere models with C atoms in green. In LAT2 (cartoon in blue) the cysteine residues (endogenous or added through mutation) are shown as sphere models with C atoms in gray. Crosslinked residues are connected by a line denoting the percentage of crosslinking (solid red line: 80–95%; solid yellow line: ∼60%; red dotted line: 5–10%). Similarly, the intersubunit disulfide bridge between C109 (4F2hc) and C154 (LAT2) is represented by a solid red line because it is present in >95% of the expressed LAT2 (see text for details).

According to our homology models of human LAT2 (Fig. S3 *A* and *B*), which are based on the atomic structure of the bacterial L-arginine/agmatine exchanger AdiC (17), seven of the nine endogenous cysteine residues in LAT2 are located in TMDs and internal loops (Fig. S3 *A* and *C*). The remaining two cysteine residues are in external loops: Cys154, participating in the intersubunit disulfide bridge, is located in the external loop TMD3–4, and Cys210 is located in the external loop TMD5–6 (Fig. S3 *A* and *C*). Nine serine residues spread over the surface of 4F2hc-ED potentially facing LAT2 (8) were selected and mutated individually to cysteine for crosslinking experiments (see locations in Fig. S2*E*). Five of the nine mutants were not considered for further studies because they produced proteolyzed proteins (S191C and S200C) (Fig. S2*F*, lanes 2 and 3) or lacked (S270C) (Fig. S2*F*, lane 4) or inefficiently formed (S497C and S506C) (Fig. S2 *G* and *H*, lanes 7 and 8) intersubunit disulfide bridges when cotransfected with Strep-LAT2. The other four 4F2hc mutants (S151C, S185C, S393C, and S412C), also containing the C330S mutation, heterodimerized with Strep-LAT2 (Fig. S2 *G* and *H*, lanes 1, 3, 5, and 6) and induced L-alanine transport similar to 4F2hc C330S (Fig. S2*I*). Of these four mutants, only S151C and S412C could be crosslinked with wild-type LAT2 using BMOE and BM(POE)$_2$ [Fig. 2*A*; BMOE: lanes 1 and 7; BM(POE)$_2$: lanes 9 and 12]. Interestingly, crosslinking with 4F2hc mutants S151C and S412C was totally abolished when LAT2 Cys210 was mutated to serine (C210S) [Fig. 2*A*; BMOE, lanes 2 and 6; BM(POE)$_2$, lanes 8 and 11]. The C210S (LAT2) mutant is fully functional (i.e., heterodimerized (Fig. S2 *G* and *H*, lanes 2 and 4) and induced L-alanine transport with 4F2hc (Fig. S2*I*). In contrast, mutant C330S (Fig. 2*A*, lanes 5 and 10), and mutants S185C (Fig. 2*A*, lane 3) and S393C (Fig. 2*A*, lane 4) incorporated in this background did not show crosslinking with

wild-type LAT2. These results demonstrated specific crosslinking between cysteine residues at positions 151 or 412 (which are located in the loop Aα1–Aβ2 and helix Aα8′′′, respectively) of subdomain A of 4F2hc-ED and Cys210, which is located in the extracellular loop TMD5–6 in LAT2.

Docking Analysis of 4F2hc-ED and LAT2. To characterize the interactions between 4F2hc-ED and LAT2, protein-docking analyses between the atomic structure of 4F2hc-ED (8) and homology models of human LAT2 (Fig. S3*B*) were performed with only two restrictions: (*i*) allowing the formation of the disulfide bridge between residues Cys109 (4F2hc-ED) and Cys154 (LAT2), and (*ii*) avoiding clashes between 4F2hc-ED residues and the plasma membrane (*SI Methods*). Of 3,145 possible docking poses, the model with the lowest energy is depicted in Fig. 2*B*. According to this model, 4F2hc-ED interacts with all the extracellular loops or ends of TMDs of LAT2. 4F2hc is *N*-glycosylated in *Pichia* (18) and mammalian cells (8), and in the model the four putative sites are located in the most external face of 4F2hc-ED, out of the contact interface with LAT2 (Fig. 2*B*). There is a large interaction surface between the two proteins, with the buried surface area of LAT2 upon 4F2hc binding being 1,735 Å^2 (Fig. S4*A*). As much as 60% of the LAT2 interaction surface is formed by hydrophobic residues, and 40% is formed by aromatic residues. This composition is consistent with the major contribution of desolvation to the binding energy as calculated by our scoring function. The most important contacting residues are shown in Table S1 and Fig. S4*B*.

Interestingly, the lowest-energy 4F2hc-ED–LAT2 model complies with the crosslinking analysis described above. Thus, crosslinked residues S151C (4F2hc)–Cys210 (LAT2) and S412C (4F2hc)–Cys210 (LAT2) were located at minimal interatomic distances of 13.4 Å and 11.5 Å, respectively (Table S2). These distances are in good agreement with the results obtained from crosslinking with BMOE and BM(POE)$_2$, in which the distances separating both reactive maleimide ends range from 6.3 to 10.5 Å and from 3.5 to 14.3 Å, respectively (20). In contrast, the 4F2hc residues that did not crosslink, S393C and S185C, are located at distances of 18.8 Å and 28.2 Å, respectively, from LAT2 Cys210, also corroborating the docking model (Table S2). Then, guided by this docking model, additional cysteine mutants in protein domains not studied previously were considered for crosslinking, i.e., S195C in helix 2 of subdomain A and S487C and G505C in subdomain C of 4F2hc-ED and A235C, A315C, G392C, and S441C in four external loops of LAT2 (Table S2). To ascribe the crosslinked residues unequivocally, 4F2hc mutants were generated in C330S background, and LAT2 mutants were generated in the C210S background, taking advantage of the fact that LAT2 C210S heterodimerized with the newly considered His-4F2hc mutants (Fig. S5 *A* and *B*, lanes 7, 12, and 17) in HEK293T cells. Moreover, the rest of the additional 4F2hc and LAT2 mutants heterodimerized and induced L-alanine transport in the transfected cells (Fig. S5). Six new paired positions in 4F2hc and LAT2 were crosslinked by BMOE and/or BM(POE)$_2$ (Fig. 2*C* and Fig. S6 *A* and *B*). Interestingly, the new pairing of G505C (4F2hc) with A235C and S441C (LAT2) (Fig. 2*C*, lanes 6 and 7) and the already mentioned pairing of S151C and S412C (4F2hc) with Cys210 (LAT2) (Fig. 2*A*) crosslinked almost completely (80–95%) (Table S2). Together with the intersubunit disulfide bridge between Cys109 (4F2hc) and Cys154 (LAT2), which links >95% of the expressed LAT2 mutants to 4F2hc (Fig. S2*G*, lanes 1 and 3 and Fig. S5*A*, lanes 15 and 16), these results triangulate the location of 4F2hc-ED over the external face of LAT2 (Fig. 2*D*). Indeed, our crosslinking studies showed that four of the six external loops of LAT2 (Cys210 and Ala235 in TMD5–6, Ala315 in TMD7–8, Gly392 in TMD9–10, and Ser441 in TMD11–12) (Fig. S3, asterisks) are in the vicinity of residues in subdomains A and C of the 4F2hc-ED in addition to the intersubunit disulfide bridge connecting Cys154 in TMD3–4 of LAT2 and Cys109 located five residues away from the C-terminal end of the 4F2hc TMD (Fig. 2*D* and Table S2). According to the lowest-energy 4F2hc-ED–LAT2 model, all the crosslinked positions in the present work are 8.1–17.5 Å apart (Table S2). In contrast, paired positions separated by >15 Å or >18 Å were not crosslinked by BMOE or BM(POE)$_2$, respectively (Fig. 2*A*, Fig. S6*A*, and Table S2). The lack of crosslinking in the indicated paired positions was not caused by cysteine residues being blocked by the formation of a direct disulfide bridge as the result of very close proximity (Fig. S6 *B* and *C*). Thus, the lowest-energy docking model is in full agreement with all crosslinking experiments. Moreover, if a slightly larger crosslinking distance is considered, only two very similar additional docking solutions comply with the five strongest (≥60% efficiency) crosslinking experiments (Fig. S4*C*). Like the 3D 4F2hc/LAT2 map from TEM and SPA, these docking models place 4F2hc-ED tilted on top of LAT2. Unfortunately, at the resolution of 21 Å, the large volume of the larger domain in the map (Fig. 1*C*), which contains LAT2, the TMD and the N-terminal domain of 4F2hc, lipid, and detergent, precludes a reliable fitting of the 4F2hc-ED–LAT2 docking model into the 3D 4F2hc/LAT2 map. In summary, TEM, SPA, docking, and crosslinking studies strongly indicate that 4F2hc-ED is positioned tilted on LAT2, almost completely covering the extracellular face of the transporter.

The Ectodomain of the Heavy Subunit 4F2hc Increases the Stability of LAT2. The interaction between 4F2hc-ED and the overall external face of LAT2 suggests that the heavy subunit may affect the stability of the light subunit. To investigate this possibility, we studied the behavior of purified 4F2hc/LAT2 and LAT2. Solubilization with DDM of isolated membranes from yeast cells coexpressing His-4F2hc and Strep-LAT2 and sequential purification by Co^{2+}- and Strep-Tactin affinity chromatography yielded 4F2hc/LAT2 heterodimers (18). In contrast, DDM solubilization of yeast membranes expressing Strep-LAT2 alone and subsequent Strep-Tactin affinity chromatography and size-exclusion chromatography yielded only aggregates of LAT2 (Fig. 3*A*). To study whether 4F2hc was able to conserve the functionality of LAT2, we tried to reconstitute purified 4F2hc/LAT2 heterodimers into liposomes, but our attempts failed. As an alternative approach to test the functionality of 4F2hc/LAT2 heterodimers and LAT2 monomers, isolated *Pichia* membranes were solubilized in DDM and reconstituted into proteoliposomes. Interestingly, only 4F2hc/LAT2 heterodimers, but not LAT2 monomers, could be reconstituted successfully into proteoliposomes as functional proteins (Fig. 3*B*), because LAT2 monomers aggregated before being inserting into liposomes (Fig. S7*A*). Surprisingly, expression of Strep-LAT2 alone, as well as coexpression of His-4F2hc and Strep-LAT2, resulted in a similar transport of 10 μM L-leucine over background in *P. pastoris* cells (Fig. S7*B*). These results suggest that LAT2 is properly folded and trafficked to the plasma membrane in yeast cells. However, the interactions with 4F2hc would be essential to maintain the correct folding of the light subunit in detergent. Kinetic analysis of L-leucine transport via His-4F2hc/Strep-LAT2 and Strep-LAT2 revealed similar K_ms: 178 ± 15 μM and 120 ± 7 μM, respectively (Fig. S7*C*). Thus, the interaction between 4F2hc and the extracellular surface of LAT2 increased the stability of the detergent-solubilized light subunit but did not alter dramatically the substrate apparent affinity of this transporter in vivo.

Next, we asked whether the ectodomain of 4F2hc is able to stabilize LAT2 without the generation of the conserved disulfide bridge between both HAT subunits. *Pichia* membranes expressing Strep-LAT2 were incubated with purified His-4F2hc-ED, and then LAT2 was solubilized with different concentrations of DDM (Fig. 4*A*). Interestingly, preincubation with 4F2hc-ED, but not with same concentration of BSA, increased the solubilization of LAT2 at the three concentrations of detergent tested (1%, 0.5%, and 0.25% DDM). These results suggested that 4F2hc-ED stabilizes LAT2 in solution. To test this possibility, LAT2 solubilized in 1% DDM in the absence or presence of 4F2hc-ED was ultracentrifuged at 2 h, 24 h, and 72 h after solubilization. Remarkably, the presence of 4F2hc-ED dramatically prevented the

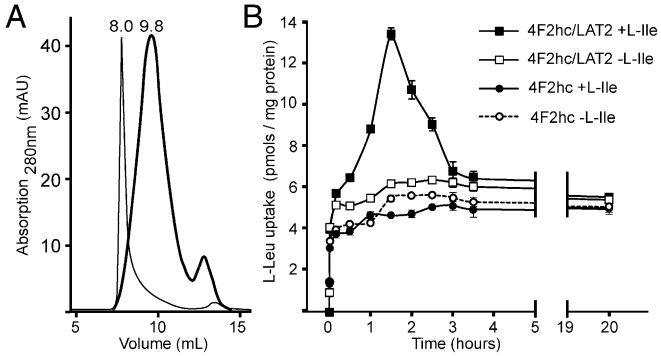


Fig. 3. (*A*) Size-exclusion chromatography profiles of purified 4F2hc/LAT2 (thick curve) and purified LAT2 (thin curve). 4F2hc/LAT2 elutes at 9.8 mL, whereas LAT2 is eluted in the void volume (8 mL). (*B*) Time-course of L-leucine uptake in 4F2hc/LAT2 and 4F2hc proteoliposomes. Transport of 10 μM L-[^{3}H]leucine into proteoliposomes was measured at 10, 30, 60, 90, 120, 150, 180, and 210 min and at 20 h. Proteoliposomes were loaded or not loaded with 4 mM cold L-isoleucine. Data are mean ± SEM of a representative experiment performed in triplicate. Transport of 10 μM L-[^{3}H]leucine into 4F2hc/LAT2 proteoliposomes preloaded with 4 mM L-isoleucine resulted in an overshoot. In contrast, L-leucine transport into 4F2hc/LAT2 proteoliposomes with no amino acids inside showed passive diffusion similar to that shown by 4F2hc proteoliposomes filled or not filled with L-isoleucine. This behavior is characteristic of coupled transporters such as the H^+/lactose cotransporter LacY (27) and exchangers such as the LAT transporter SteT, an L-serine/L-threonine exchanger from *Bacillus subtilis* (28).

precipitation of LAT2 (Fig. 4 *B* and *C*). In summary, these results indicate that the ectodomain of 4F2hc interacts with the external face of LAT2, thus increasing the stability of the detergent-solubilized transporter and allowing its functional reconstitution into proteoliposomes.

Discussion

This study reports on the supramolecular organization and the structural bases of the interactions of the heavy and light subunits of HATs. Specifically, three lines of evidence (TEM, SPA, and docking and crosslinking analyses) strongly support the notion that 4F2hc-ED almost completely covers the extracellular surface of the transporter subunit LAT2. Indeed, we identified specific residues that crosslinked 4F2hc and LAT2 nearly completely (>80%). This crosslinking, together with the endogenous intersubunit disulfide bridge, triangulates the position of the 4F2hc-ED over the external face of LAT2 (Fig. 2*D*). Importantly, the proposed docking model, as well as two additional and highly similar poses, fulfills the steric restrictions defined by the results from crosslinking (Fig. S4*C*).

Our docking analysis used the atomic structure of human 4F2hc-ED and homology models of human LAT2 based on the atomic structure of the amino acid transporter AdiC in an outward-facing conformation. Because of the low sequence conservation in the loops of AdiC and LAT2 (Fig. S3*A*), our docking models do not represent high-resolution structures of the 4F2hc-ED–LAT2 complex. Nevertheless, docking analysis revealed desolvation as the main contributor to the binding energy of 4F2hc-ED and LAT2. Indeed, the LAT2 residues that are predicted to be more relevant for the interactions are mainly hydrophobic, and they can be grouped into two clearly defined patches [see Normalized Interface Propensity (NIP) residues in Table S1 and Fig. S4 *A* and *B*]. One of the patches, which is totally hydrophobic, is composed of nine residues distributed along three external loops or TMD ends of LAT2 (TMD3–4, which harbors the Cys154 residue involved in the disulfide bridge, TMD9–10, and TMD11–12). The other is a mixed patch composed of four polar and five hydrophobic residues located in loops TMD1–2, TMD5–6, and TMD7–8. The interaction between 4F2hc-ED and LAT2 raises the question of whether the heavy subunit might have an impact on the LAT2 transport cycle. The structural paradigms of HAT light subunits, the bacterial APC transporters, present the LeuT-fold (14–17). Secondary transporters, upon the binding of an external substrate, transit to inward-facing conformations to release the substrate inside the cell. Besides transporter-specific changes, atomic structures of LeuT-fold transporters showed a commonality of features in this transition (21, 22): the "hash" domain (TMDs 3, 4, 8, and 9) and the extracellular ends of the "bundle" domain (TMDs 1, 2, 6, and 7) become closer, and the loop TMD7–8 occludes the external vestibule of the transporter. Our results indicate that these conformational changes are compatible with 4F2hc-ED directly contacting the external loops or ends of TMDs of LAT2. The highly conserved hydrophobic patch in LAT2 (Fig. S4*B* and Table S1), together with the intersubunit disulfide bridge, glue the hash domain and TMDs 11 and 12 to 4F2hc-ED. TMDs 11 and 12 are, according to the structural homolog AdiC (17), the most static TMDs. In contrast, the mixed patch is less conserved and involves the bundle domain and the occluding loop (TMD7–8). This architecture suggests that 4F2hc/LAT2, and probably other HATs, have evolved to bind the ectodomain of the heavy subunit firmly to the hash domain. In this scenario, energetically similar interactions of 4F2hc-ED with the bundle domain and loop TMD7–8 would be broken and replaced by the different conformations that the light subunit undergoes during the transport cycle.

According to amino acid transporters with LeuT-fold, the substrate-binding site in LAT2 is expected to be located at the bottom of the extracellular vestibule (1, 13) and therefore far away from the interacting 4F2hc-ED residues (Fig. S4 *D* and *E*). Moreover, our docking analysis located 4F2hc-ED in a tilted

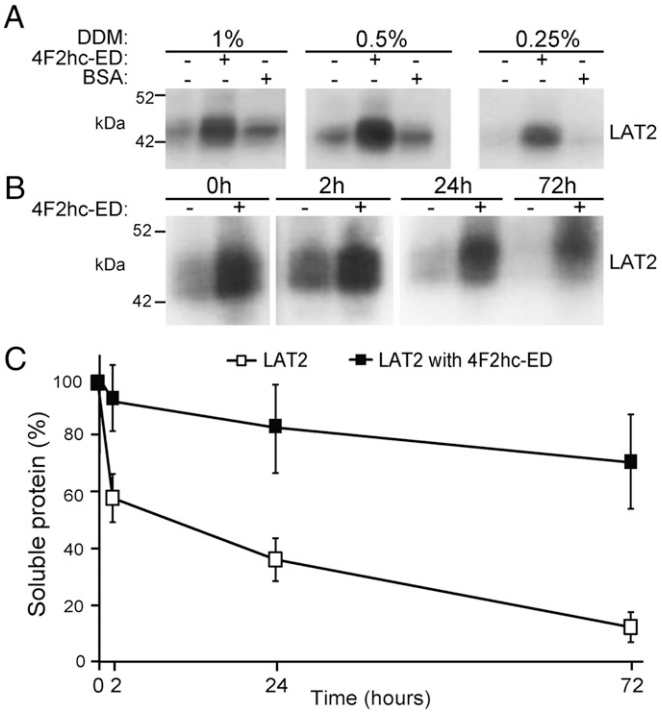


Fig. 4. 4F2hc-ED increases LAT2 solubility and stability. (*A*) Western blot analysis of DDM-solubilized LAT2 in the presence of 1 mg/mL of 4F2hc-ED or BSA or plain buffer, as indicated. (*B*) Western blot analysis of solubilized LAT2 (1% DDM) in the presence or absence of 1 mg/mL of 4F2hc-ED at different times. (*C*) Quantification of the soluble LAT2 at the indicated time. Data are the mean ± SEM of three independent experiments as described in *B*. LAT2 remaining in solution after ultracentrifugation was considered soluble LAT2. LAT2 was detected using αStrep antibody. The presence of 4F2hc-ED increased LAT2 stability at all analyzed times.

position, with subdomain C interacting with LAT2 TMD11–12 and leaving open a window in the opposite location flanked by LAT2 TMD7–8 for the entrance of the substrate to the binding site (Fig. S4*D*). Despite the different structures of the LAT2 monomer and the 4F2hc/LAT2 heterodimer, the measured K_m values for L-leucine transport in *Pichia* cells were comparable, suggesting that 4F2hc would have no significant impact on the substrate affinity of LAT2.

The hydrophobic character of the interaction between 4F2hc-ED and LAT2 might be the basis for the stabilization of this light subunit by 4F2hc. We show that 4F2hc allows functional reconstitution into proteoliposomes and purification of the transporter after solubilization with DDM. Moreover, we show here that 4F2hc-ED suffices to increase the stability of DDM-solubilized LAT2. The binding of 4F2hc-ED to the extracellular surface of LAT2 would bury a large hydrophobic patch that otherwise would be exposed to solvent, thus reducing aggregation and increasing stability. In addition, the compact and thermally stable structure (T_m ~59 °C) (23) of 4F2hc-ED may have additional effects in improving LAT2 stability.

HATs appeared in metazoans with an ancient heavy subunit that evolved to differentiated 4F2hc and rBAT in vertebrates (1). The fact that LAT2 monomers elicited amino acid transport in *Pichia* cells, which do not express 4F2hc or related proteins, indicates that 4F2hc is not required for proper folding of LAT2, as also has been demonstrated for the light subunit $b^{0,+}$AT in the absence of rBAT (2). In contrast, 4F2hc and rBAT are necessary for heterodimerization and functional expression of light subunits at the cell surface when expressed in *Xenopus* oocytes and mammalian cells (1). Specifically, 4F2hc-ED is necessary for these functions (9, 10). Desolvation of hydrophobic residues might be extended to the recognition of other light subunits by 4F2hc. The hydrophobic NIP residues in our docking analysis are highly conserved in 4F2hc-associated light subunits but are poorly conserved in cyanobacterial, yeast, and fungal LATs, which are the closest homologs to vertebrate LATs (up to 37% amino acid sequence identity) that do not interact with a heavy subunit (Table S1). Hydrophobic NIP residues also are largely conserved in $b^{0,+}$AT (Table S1). Thus, desolvation of hydrophobic residues also might contribute to the binding of the ectodomain of rBAT and its light subunit. Indeed, cotransfection of 4F2hc and $b^{0,+}$AT (i.e., the rBAT-associated light subunit) in mammalian cells yielded functional transporter at the cell surface (24). Interaction through a hydrophobic patch seems like a clever way to glue 4F2hc to six different light subunits (LAT1, LAT2, y^+LAT1, y^+LAT12, asc1, and xCT) that share as little as ~43% amino acid sequence identity (1). Our work revealed that 4F2hc-ED covers the external face of LAT2, increasing transporter stability. The idea that a common hydrophobic patch serves for subunit recognition by 4F2hc merits future experimental examination.

Methods

The indicated tagged versions of human 4F2hc and LAT2 were expressed and purified as reported previously (18). L-leucine transport by *Pichia* cells was performed as described (18). Transport activity of LAT2, 4F2hc, and 4F2hc/LAT2 was tested as described previously (13) by measuring their L-leucine/L-isoleucine exchange activity in reconstituted proteoliposomes. TEM, SPA, and 3D reconstruction were performed as described in *SI Methods*. Docking calculations of 4F2hc-ED–LAT2 interaction were performed based on the pyDock protocol (25), and the most relevant LAT2 residues for the interaction with 4F2hc were identified by using the NIP value (26). To validate the docking model, intersubunit crosslinking was performed in different versions of purified 4F2hc/LAT2 heterodimer expressed in HEK cells. The effect of 4F2hc-ED on the detergent solubilization and stability of human LAT2 was tested by incubating *Pichia* membranes expressing LAT2 (18) with purified 4F2hc-ED (8) before solubilization with DDM. Soluble LAT2 protein was detected and quantified by Western blotting. See *SI Methods* for details.

ACKNOWLEDGMENTS. This work was supported by Spanish Ministry of Science and Innovation Grants BIO2010-22324 (to J.F.-R.) and SAF2012-40080-C02-01, European Commission Frame Program 7 Grant 201924 (European Drug Initative on Channels and Transporters), Fundación Ramón Areces, and the Generalitat de Catalunya Grant SGR2009-1355 (to M.P.); by University of Bern, Swiss National Science Foundation Grants 31003A_125150 and 31003A_144168; the Bern University Research Foundation; the Novartis Foundation; the Marie Curie Actions International Fellowship Program; and National Center of Competence in Research TransCure (D.F.).

1. Fotiadis D, Kanai Y, Palacín M (2013) The SLC3 and SLC7 families of amino acid transporters. *Mol Aspects Med* 34(2-3):139–158.
2. Reig N, et al. (2002) The light subunit of system b(o,+) is fully functional in the absence of the heavy subunit. *EMBO J* 21(18):4906–4914.
3. Feliubadaló L, et al.; International Cystinuria Consortium (1999) Non-type I cystinuria caused by mutations in SLC7A9, encoding a subunit ($b^{o,+}$AT) of rBAT. *Nat Genet* 23(1):52–57.
4. Mastroberardino L, et al. (1998) Amino-acid transport by heterodimers of 4F2hc/CD98 and members of a permease family. *Nature* 395(6699):288–291.
5. Rius M, Chillarón J (2012) Carrier subunit of plasma membrane transporter is required for oxidative folding of its helper subunit. *J Biol Chem* 287(22):18190–18200.
6. Fernández E, et al. (2006) The structural and functional units of heteromeric amino acid transporters. The heavy subunit rBAT dictates oligomerization of the heteromeric amino acid transporters. *J Biol Chem* 281(36):26552–26561.
7. Feral CC, et al. (2005) CD98hc (SLC3A2) mediates integrin signaling. *Proc Natl Acad Sci USA* 102(2):355–360.
8. Fort J, et al. (2007) The structure of human 4F2hc ectodomain provides a model for homodimerization and electrostatic interaction with plasma membrane. *J Biol Chem* 282(43):31444–31452.
9. Fenczik CA, et al. (2001) Distinct domains of CD98hc regulate integrins and amino acid transport. *J Biol Chem* 276(12):8746–8752.
10. Bröer A, et al. (2001) Association of 4F2hc with light chains LAT1, LAT2 or y+LAT2 requires different domains. *Biochem J* 355(Pt 3):725–731.
11. Gasol E, Jiménez-Vidal M, Chillarón J, Zorzano A, Palacín M (2004) Membrane topology of system x_c^- light subunit reveals a re-entrant loop with substrate-restricted accessibility. *J Biol Chem* 279(30):31228–31236.
12. Jiménez-Vidal M, et al. (2004) Thiol modification of cysteine 327 in the eighth transmembrane domain of the light subunit xCT of the heteromeric cystine/glutamate antiporter suggests close proximity to the substrate binding site/permeation pathway. *J Biol Chem* 279(12):11214–11221.
13. Bartoccioni P, et al. (2010) Role of transmembrane domain 8 in substrate selectivity and translocation of SteT, a member of the L-amino acid transporter (LAT) family. *J Biol Chem* 285(37):28764–28776.
14. Shi Y (2013) Common folds and transport mechanisms of secondary active transporters. *Annu Rev Biophys* 42:51–72.
15. Shaffer PL, Goehring A, Shankaranarayanan A, Gouaux E (2009) Structure and mechanism of a Na+-independent amino acid transporter. *Science* 325(5943):1010–1014.
16. Fang Y, et al. (2009) Structure of a prokaryotic virtual proton pump at 3.2 Å resolution. *Nature* 460(7258):1040–1043.
17. Kowalczyk L, et al. (2011) Molecular basis of substrate-induced permeation by an amino acid antiporter. *Proc Natl Acad Sci USA* 108(10):3935–3940.
18. Costa M, et al. (2013) Expression of human heteromeric amino acid transporters in the yeast *Pichia pastoris*. *Protein Expr Purif* 87(1):35–40.
19. Jastrzebska B, et al. (2011) Rhodopsin-transducin heteropentamer: Three-dimensional structure and biochemical characterization. *J Struct Biol* 176(3):387–394.
20. Green NS, Reisler E, Houk KN (2001) Quantitative evaluation of the lengths of homobifunctional protein cross-linking reagents used as molecular rulers. *Protein Sci* 10(7):1293–1304.
21. Shimamura T, et al. (2010) Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science* 328(5977):470–473.
22. Krishnamurthy H, Gouaux E (2012) X-ray structures of LeuT in substrate-free outward-open and apo inward-open states. *Nature* 481(7382):469–474.
23. Turnay J, et al. (2011) Structural characterization and unfolding mechanism of human 4F2hc ectodomain. *Biochim Biophys Acta* 1814(5):536–544.
24. Rajan DP, et al. (2000) Differential influence of the 4F2 heavy chain and the protein related to b(0,+) amino acid transport on substrate affinity of the heteromeric b(0,+) amino acid transporter. *J Biol Chem* 275(19):14331–14335.
25. Cheng TM, Blundell TL, Fernandez-Recio J (2007) pyDock: Electrostatics and desolvation for effective scoring of rigid-body protein-protein docking. *Proteins* 68(2):503–515.
26. Fernández-Recio J, Totrov M, Abagyan R (2004) Identification of protein-protein interaction sites from docking energy landscapes. *J Mol Biol* 335(3):843–865.
27. Newman MJ, Foster DL, Wilson TH, Kaback HR (1981) Purification and reconstitution of functional lactose carrier from *Escherichia coli*. *J Biol Chem* 256(22):11804–11808.
28. Reig N, et al. (2007) Functional and structural characterization of the first prokaryotic member of the L-amino acid transporter (LAT) family: A model for APC transporters. *J Biol Chem* 282(18):13270–13281.

Supporting Information

Rosell et al. 10.1073/pnas.1323779111

SI Methods

Protein Expression and Purification and L-Leucine Transport in *Pichia pastoris* Cells. Tagged versions of human 4F2hc and L-type amino acid transporter 2 (LAT2) were expressed and purified as reported previously (1). The L-leucine transport assay using *Pichia* cells was performed as described previously (1).

Functional Reconstitution of LAT2, 4F2hc, and 4F2hc/LAT2 into Liposomes. Liposomes from soybean and pig brain lipids were prepared as described (2) in reconstitution buffer [5 mM Tris-SO_4 (pH 7.4), 120 mM KP_i, 0.5 mM EDTA, 1 mM $MgSO_4$, 1% glycerol] with or without 4 mM L-Ile. *Pichia* membranes containing expressed LAT2, 4F2hc, and 4F2hc/LAT2 were solubilized in 0.5% n-dodecyl-β-D-maltoside (DDM) for 30 min at 4 °C and ultracentrifuged (250,000 × *g*; 1 h; 4 °C). The supernatant was mixed with the preformed liposomes at a lipid/protein (wt/wt) ratio of 50 for 10 min on ice. Reconstitution was performed by adding 0.1 g wet Bio-Beads (Bio-Beads SM2 Adsorbent; Bio-Rad) to 500 μL of mixture followed by incubation for 3 h at 4 °C. Proteoliposomes were harvested by ultracentrifugation (250,000 × *g*; 1 h; 4 °C) and were resuspended in half of the original volume of reconstitution buffer. For transport assays, proteoliposomes (10 μL) were mixed with 180 μL of transport buffer [10 mM Tris-Hepes (pH 7.4), 150 mM choline chloride, 1 mM $MgCl_2$] containing 10 μM L-[^{3}H]Leu (0.5 μCi). Transport was stopped at the indicated times (Fig. 3*B*) by the addition of 850 μL ice-cold stop buffer [10 mM Tris-Hepes (pH 7.4), 150 mM choline chloride, 5 mM L-Leu] and filtered (Sartorius membrane filters, 0.45-μm pore size). Filters were washed three times with 2 mL of stop buffer and dried. Trapped radioactivity was counted in a Packard Tri-Carb Liquid Scintillation Counter. Experimental values were corrected by subtracting the signal at time 0.

Negative-Stain Transmission Electron Microscopy. Purified 4F2hc/LAT2 at ∼6 μg/mL was adsorbed for ∼5 s to Parlodion carbon-coated copper grids, which were rendered hydrophilic by glow discharge at low pressure in air. Grids were washed with three drops of double-distilled water and stained with two drops of 0.75% uranyl formate. Electron micrographs were recorded with a Philips CM12 transmission electron microscope operated at 80 kV and equipped with a Morada CCD camera (Soft Imaging System).

Image Processing and 3D Reconstruction. For image processing and 3D reconstruction, the software package EMAN2 (3) was used. Initially, 23,838 single particles of 4F2hc/LAT2 were picked with a box size of 128 pixels using the e2boxer program. Electron micrographs were corrected for the contrast transfer function of the microscope using the e2ctf program. Reference-free classification and averaging was done with the e2refine2d program, which yielded about 200 class averages. From these class averages the e2initialmodel algorithm built a starting 3D model. The preliminary model was refined against a final set of 15,210 projections by running the iterative refinement procedure e2refine. During the refinement process the angular spacing was decreased from 15° to 2.5°. The resolution of the 3D reconstruction was 21 Å, as estimated by calculating the Fourier shell correlation with the e2eotest program. The density threshold for volume rendering of the 4F2hc/LAT2 map was determined by fitting the surface representation of the 4F2hc ectodomain (4F2hc-ED) crystal structure into the small density of the 3D reconstruction and adjusting the threshold. Visualization of the 3D model and fitting of the 4F2hc-ED crystal structure was done in UCSF Chimera (4).

Docking Model of 4F2hc-ED with LAT2. This docking experiment had two major challenges: the lack of the 3D structure for one of the subunits, human LAT2, and the fact that this protein is a membrane protein. For this reason, we needed to develop ad hoc protocols based on our standard program, pyDock (5). To overcome the first problem, we used Modeler 8v1 (6) to model the structure of human LAT2, using as template the X-ray structure of the prokaryotic homolog AdiC from *Escherichia coli* [Protein Data Bank (PBD) ID code 3OB6, chain A; ∼20% amino acid sequence identity] (7), and an extensive multiple sequence alignment between AdiC and 250 close homologs including bacterial and simple eukaryotic and metazoan LATs (part of the alignment is shown in Fig. S3*A*). Given the structural uncertainty in some of the loops, we decided to select for docking the 20 models with the lowest digital optimized protein energy (DOPE) score as provided by Modeler. As shown in Fig. S3*B*, major differences between models were found in the transmembrane domain (TMD) loops TMD5-6 and TMD7–8, suggesting flexibility in these regions. We used each of these models and the structure of 4F2hc-ED (PDB ID code 2DH2, chain A) (8) to perform an independent docking search using FTDock 2.0 (9) and ZDock 2.1 (10), which generated 10,000 and 2,000 rigid-body docking poses, respectively, for each LAT2 model. This search resulted in a total of 240,000 different docking models. It is known that LAT2 Cys154 residue forms a conserved disulfide bond with the 4F2hc Cys109 residue (Fig. S6 *B* and *C*) (11). Therefore, we removed the docking solutions that could not satisfy this condition. For that purpose, and according to B factors of the crystal structures of 4F2hc-ED (PDB ID codes 2DH2 and 2DH3) (8), we assumed that the N terminus of the 4F2hc-ED (residues 109–116) was flexible. Therefore, we selected the docking solutions in which the minimal interatomic distance between 4F2hc Lys116 and LAT2 Cys154 residues was ≤20.9 Å; this value was defined from the pyDockTET linker length database (12), considering the flexible N terminus in 4F2hc as an interdomain linker. An important issue in modeling this interaction is that LAT2 is a transmembrane protein. Because the transmembrane regions were known, we also removed docking solutions that would clash badly with the membrane, i.e., in which >20% of residues in 4F2hc contained atoms within 3 Å distance of any atom of the membrane. The position of the membrane around 4F2hc was modeled based on a recent molecular dynamics simulation of the LAT2 homolog AdiC within the membrane (7). These two filtering steps reduced the initial 240,000 docking solutions to 3,145 true candidate docking models, which were scored and ranked according to their binding energy using pyDock (5). Finally the docking model with the lowest energy (the one ranked first) was minimized using Tinker (13). This model satisfies all crosslinking experiments with an error margin of <5 Å (Table S2).

Docking-Based Contributions of LAT2 Residues to Interactions with 4F2hc-ED. The docking model described above is a good representation of the overall binding mode between LAT2 and 4F2hc, but it is not intended to represent a high-resolution structure of the complex. Indeed, other, slightly different docking conformations are also possible (Fig. S4*C*). Furthermore, we should not disregard possible small side-chain and backbone rearrangements. As a consequence, the model is not sufficiently accurate for a detailed analysis of the contributions of every specific residue to the binding energy. Nevertheless, based on the analysis of the docking results (after filtering by disulfide-bridge restraints and membrane clashing), we can identify the LAT2

residues that are most relevant for the interaction with 4F2hc using the Normalized Interface Propensity (NIP) value, as previously described (14). The NIP value indicates how frequently a given residue is located at the interface in the 100 lowest-scoring docking orientations, after normalizing between 0 (random) and 1 (maximum). The residues with NIP >0.4 (usually indicating hot-spot residues for the interaction) are listed in Table S1.

Solubilization and Stability Assays of Human LAT2. N-terminally His-tagged human 4F2hc-ED was produced in *E. coli* and purified as previously described (8). Membranes of *P. pastoris* overexpressing N-terminally Strep-TagII–tagged human LAT2 were obtained as previously described and kept at −80 °C (1). Purified 4F2hc-ED or BSA at a final concentration of 1 mg/mL in solubilization buffer [20 mM Tris·HCl (pH 7.4), 150 mM NaCl] or solubilization buffer without protein was added to LAT2 membranes (total membrane protein concentration in solubilization buffer: 20 mg/mL) and incubated overnight at 4 °C on a rotating orbital shaker. Then, membranes were solubilized with the indicated concentration of DDM (Fig. 4*A*) for 1 h at 4 °C and ultracentrifuged at 200,000 × *g* for 1 h to remove nonsoluble proteins and aggregates. When analyzing protein stability, supernatants were ultracentrifuged again at the indicated times (Fig. 4 *B* and *C*). Supernatants in Laemmli sample buffer were loaded for SDS/PAGE (8% polyacrylamide) and analyzed by Western blotting.

Expression of 4F2hc and LAT2 in HEK293T Cells. Human His-4F2hc (N-terminally tagged) cloned in pcDNA4His-MaxC (8) was used. Human Strep-TagII-LAT2 (N-terminally tagged) cloned in pPICZ (1) was subcloned, adapting to the mammalian Kozak sequence, into EcoRI and XbaI sites in pcDNA3.1+. The QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent) was used to generate all mutants, which were confirmed by sequencing. HEK293T cells (11–30 passages) were grown in DMEM (Gibco 41966) supplemented with 10% FBS. Cells were kept at 37 °C in a humidified atmosphere containing 5% CO_2. Transient transfections were performed at 70% cell confluence with polyethylenimine in 15-cm-diameter plates with a mixture of DNA containing 2.5 μg of pEGFP (Clontech), 15 μg of pcDNA4HisMax human 4F2hc mutants, and 15 μg of pcDNA3.1+ human Strep-TagII–tagged LAT2 (wild-type or mutants). A transfection of a single construct was cotransfected with the same amount of the corresponding empty vector DNA. Cell medium was replaced 4 h after transfection.

Transport of L-Alanine in HEK293T Cells. Transfected cells were scraped and transferred into 24-well dishes previously treated with poly-L-lysine (100 μL per well at 0.1 mg/mL for 30 min at 37 °C). One day later, L-alanine transport was measured. To this end, cells were washed twice with 2.5 mL of MGA buffer [10 mM Hepes (pH 7.4), 137 mM *N*-methyl-D-glucamine, 2.8 mM $CaCl_2{\cdot}2H_2O$, 1.2 mM $MgSO_4{\cdot}7H_2O$, 5.4 mM KCl] at 37 °C. The uptake solution (270 μL per well) contained 50 μM of cold L-alanine plus 0.5 μCi per well of L-[^{3}H]alanine (Hartman) in MGA buffer at 37 °C. After the indicated time, uptake was stopped by two washes with MGA buffer at 4 °C (2.5 mL per well). Cells were then lysed with 250 μL per well of lysis solution (100 mM NaOH, 0.1% SDS) by agitation for 30 min at room temperature. Then 200 μL was used for radioactivity counting in a Packard Tri-Carb Liquid Scintillation Counter, and 25-μL duplicates were used to determine protein content using the Pierce BCA kit, Pierce.

Intersubunit Crosslinking. Two days after transfection, cells were washed with 10 mL PBS (pH 6.5), scrapped with 1 mL of lysate buffer [1% DDM in PBS (pH 6.5) with protease inhibitors (Complete EDTA free; Roche)] per 15-cm plate, and transferred into a 1.5-mL Eppendorf tube. Cell disruption was facilitated by homogenizing 15 times with a 25-gauge syringe and incubation on a rotating orbital shaker at 4 °C. After 1 h, the insoluble material was removed by centrifugation at 200,000 × *g* for 30 min at 4 °C. Part of the supernatant was frozen in Laemmli sample buffer and kept at −20 °C for analysis of expression by Western blot analysis; the rest was used for crosslinking experiments. To this end, supernatants (1 mL) were incubated for 1 h at 4 °C with 150 μL of Ni-NTA-agarose beads (Qiagen), previously equilibrated with 1 mL lysate buffer. Then beads were washed with 1 mL lysate buffer containing 25 mM imidazole, and proteins were eluted by shaking incubation for 15 min at 4 °C with 250 μL of elution buffer [PBS (pH 6.5) containing 250 mM imidazole]. The purified protein (100 μL) was incubated with 30 μL of the crosslinking buffer [200 μM bis(maleimido)ethane (BMOE) or 1,8-bis (maleimido)diethylene glycol [$BM(POE)_2$] (Pierce)] in lysate buffer containing 50% (vol/vol) DMSO] for 1 h at 4 °C. The crosslinking reaction was terminated by adding 100 mM DTT. Fourfold concentrated Laemmli sample buffer was added for Western blot analysis.

Western Blot Analysis. DDM-solubilized proteins (20 μg) from HEK293T cells or *Pichia* membranes were resolved in reducing (100 mM DTT) or nonreducing SDS/PAGE (8% polyacrylamide) and immunoblotted using αStrep (Strep MAB-Classic-HRP; IBA) and αHis (anti–HisG-HRP antibody; Invitrogen) antibodies.

1. Costa M, et al. (2013) Expression of human heteromeric amino acid transporters in the yeast *Pichia pastoris*. *Protein Expr Purif* 87(1):35–40.
2. Reig N, et al. (2002) The light subunit of system b(o,+) is fully functional in the absence of the heavy subunit. *EMBO J* 21(18):4906–4914.
3. Tang G, et al. (2007) EMAN2: An extensible image processing suite for electron microscopy. *J Struct Biol* 157(1):38–46.
4. Pettersen EF, et al. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612.
5. Cheng TM, Blundell TL, Fernandez-Recio J (2007) pyDock: Electrostatics and desolvation for effective scoring of rigid-body protein-protein docking. *Proteins* 68(2):503–515.
6. Eswar N, Eramian D, Webb B, Shen MY, Sali A (2008) Protein structure modeling with MODELLER. *Methods Mol Biol* 426:145–159.
7. Kowalczyk L, et al. (2011) Molecular basis of substrate-induced permeation by an amino acid antiporter. *Proc Natl Acad Sci USA* 108(10):3935–3940.
8. Fort J, et al. (2007) The structure of human 4F2hc ectodomain provides a model for homodimerization and electrostatic interaction with plasma membrane. *J Biol Chem* 282(43):31444–31452.
9. Gabb HA, Jackson RM, Sternberg MJ (1997) Modelling protein docking using shape complementarity, electrostatics and biochemical information. *J Mol Biol* 272(1):106–120.
10. Chen R, Weng Z (2003) A novel shape complementarity scoring function for protein-protein docking. *Proteins* 51(3):397–408.
11. Fotiadis D, Kanai Y, Palacín M (2013) The SLC3 and SLC7 families of amino acid transporters. *Mol Aspects Med* 34(2-3):139–158.
12. Cheng TM, Blundell TL, Fernandez-Recio J (2008) Structural assembly of two-domain proteins by rigid-body docking. *BMC Bioinformatics* 9:441.
13. Ponder W, Richards FM (1987) An efficient Newton-like method for molecular mechanics energy minimization of large molecules. *J Comput Chem* 8:1016–1024.
14. Fernández-Recio J, Totrov M, Abagyan R (2004) Identification of protein-protein interaction sites from docking energy landscapes. *J Mol Biol* 335(3):843–865.

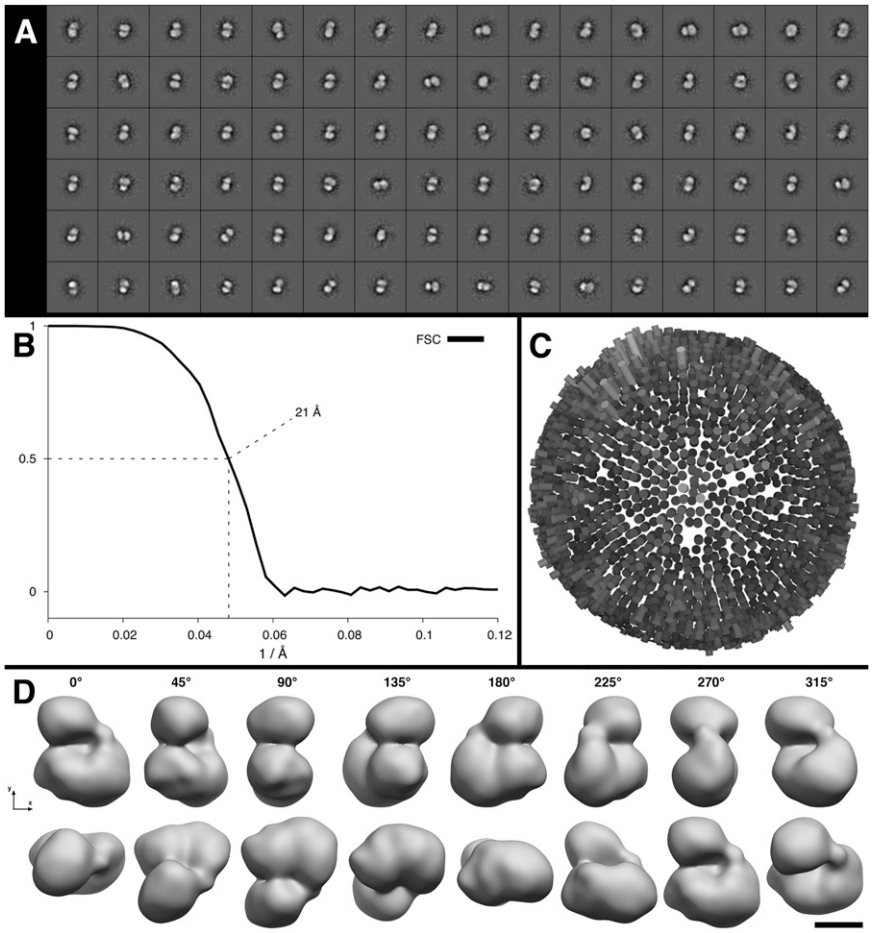


Fig. S1. Class averages and 3D reconstruction of human 4F2hc/LAT2. (*A*) Set of 4F2hc/LAT2 class averages. Two distinct densities of different sizes are discerned in most averages. Class averages were calculated from 15,210 single-particle projections. The frame sizes of the class averages are 39.6 nm. (*B*) The Fourier shell correlation (FSC) function indicates a resolution of 21 Å for the 3D map of 4F2hc/LAT2 using the 0.5 criterion. FSC was obtained by running the e2eotest program on the last refinement cycle. (*C*) The Euler angle distribution of the projections in the last refinement cycle indicates a homogenous sampling of the 4F2hc/LAT2 complex. (*D*) Different views of the 4F2hc/LAT2 3D reconstruction. The 3D map of 4F2hc/LAT2 is rotated in 45° increments along the *y* axis (*Upper*) and *x* axis (*Lower*). (Scale bar: 5 nm.)

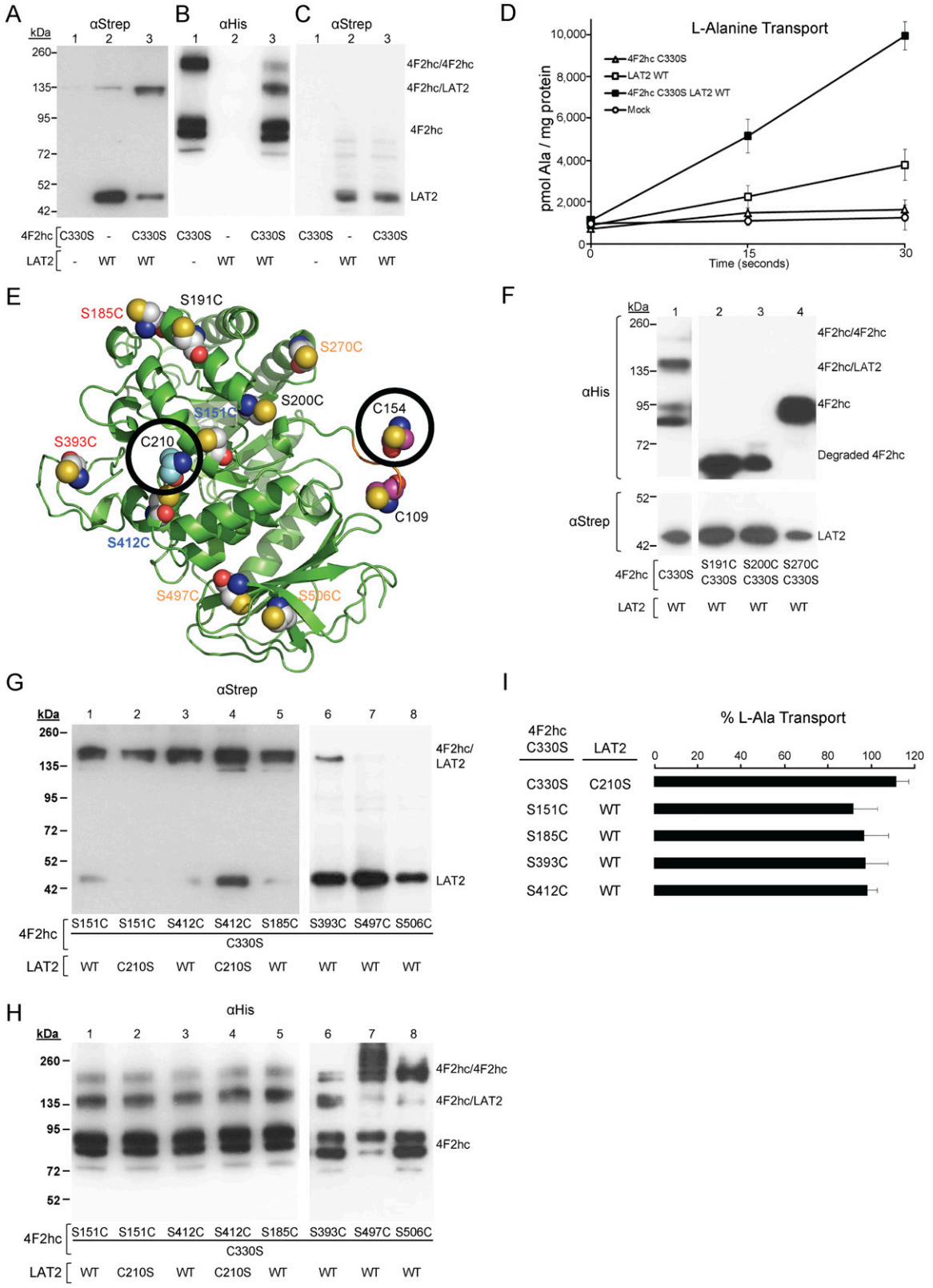


Fig. S2. Heterodimer formation and amino acid transport in HEK293T cells cotransfected with 4F2hc and LAT2 mutants generated before docking analyses. (*A*–*C*) Western blot analysis of 1% DDM-solubilized homogenates of cells transfected with the indicated versions of His-4F2hc and Strep-LAT2 using αStrep (*A* and *C*) or αHis (*B*) antibodies. Bands corresponding to heterodimers (4F2hc/LAT2), 4F2hc homodimers (4F2hc/4F2hc), and monomers (4F2hc or LAT2) are indicated under nonreducing conditions (*A* and *B*). In contrast, under reducing conditions only LAT2 monomers are visible (*C*). The low expression of 4F2hc/LAT2 heterodimers in cells transfected only with wild-type LAT2 (lane 2 in *A*) probably was caused by heterodimerization with endogenous 4F2hc. The minus sign (–) indicates transfection with the corresponding empty vector. 4F2hc monomers were revealed as doublets of bands (*B*) that correspond to core and

Legend continued on following page

mature *N*-glycosylation forms, as previously described (1). (*D*) Transport of L-alanine. Cells were analyzed for transport of 10 μM L-[^{3}H] alanine for the indicated times. Cotransfection with His-4F2hc (C330S) and wild-type Strep-LAT2 (closed squares) resulted in the induction of L-alanine transport over background [mock transfected or His-4F2hc (C330S) transfected cells]. Transfection with Strep-LAT2 alone induced L-alanine transport over background but less than cotransfection with both proteins ($P \leq 0.001$; Student *t*-test). Induction of L-alanine transport by LAT2 alone results from heterodimerization with endogenous 4F2hc and thus traffic to the plasma membrane (lane 2 in *A*) or from LAT2 monomers reaching the cell surface because of overexpression. Data (mean ± SEM) are from three experiments run in quadruplicate. (*E*) Positions of cysteine residues introduced in 4F2hc-ED before docking analysis. The two LAT2 extracellular cysteine residues (encircled) are also indicated according to the lowest-energy docking model (Fig. S3*B*), i.e., C210 with C atoms shown in cyan and C154 with C atoms shown in purple. LAT2 C154 forms the disulfide bridge with 4F2hc C109 (C atoms shown in purple). The N-terminal end of 4F2hc-ED [C109-S-E-L-P113 (one-letter-code amino acid sequence)] is colored orange to denote its high mobility. 4F2hc cysteine mutants are written in different colors: proteolytic 4F2hc production (black), inefficient formation of heterodimers with LAT2 (orange), lack of crosslinking with LAT2 C210 (red), and positive crosslink with LAT2 C210 (blue). C atoms for other cysteine residues are in gray, and in all cases O atoms are in red, N atoms are in blue, and S atoms are in yellow. (*F*) Lack of formation of 4F2hc/LAT2 heterodimers by some 4F2hc mutants in HEK293T cells. Western blot analysis of cells transfected with the indicated versions of His-4F2hc and Strep-LAT2 using αHis under nonreducing conditions (*Upper*) or αStrep under reducing conditions (*Lower*). Bands corresponding to heterodimers (4F2hc/LAT2), 4F2hc homodimers (4F2hc/4F2hc), and monomers (4F2hc or LAT2) are indicated. Degraded 4F2hc, bands corresponding to proteolytic fragments of the protein. WT, wild-type version of LAT2. (*G* and *H*) Expression of 4F2hc/LAT2 heterodimers by mutants of 4F2hc and LAT2 generated before the docking studies. Western blot analysis of 1% DDM-solubilized homogenates of HEK293T cells transfected with the indicated versions of His-4F2hc and Strep-LAT2 using αStrep (*G*) and αHis (*H*) antibodies under nonreducing conditions. All mutants in 4F2hc were generated in the C330S background. Bands corresponding to heterodimers (4F2hc/LAT2), 4F2hc homodimers (4F2hc/4F2hc), and monomers (4F2hc or LAT2) are indicated under nonreducing conditions (*G* and *H*). WT, wild-type version of LAT2. (*I*) Induction of L-alanine transport by the 4F2hc and LAT2 mutants generated before the docking analysis. HEK293T cells were transfected with wild-type or the C210S mutant of Strep-LAT2 alone or together with the indicated mutants of His-4F2hc. Cells were analyzed for transport of 10 μM L-[^{3}H]alanine for 15 s. Transport resulting from 4F2hc and LAT2 coexpression was calculated by subtraction of transport elicited by transfecting Strep-LAT2 alone. For 4F2hc C330S and wild-type LAT2 these transport values were 2463 ± 163 pmol L-alanine during 15 s (mean ± SEM from four experiments run in quadruplicate). For the mutants shown, transport is expressed as the percentage of the transport of 4F2hc C330S and wild-type LAT2. Data (mean ± SEM) are from two experiments run in quadruplicate. The coexpression of all of the indicated mutants induced L-alanine transport over the background caused by expression of the corresponding LAT2 version alone (i.e., without cotransfecting any version of 4F2hc). Thus, all LAT2 mutants studied were functional.

1. Fort J, et al. (2007) The structure of human 4F2hc ectodomain provides a model for homodimerization and electrostatic interaction with plasma membrane. *J Biol Chem* 282(43): 31444–31452.

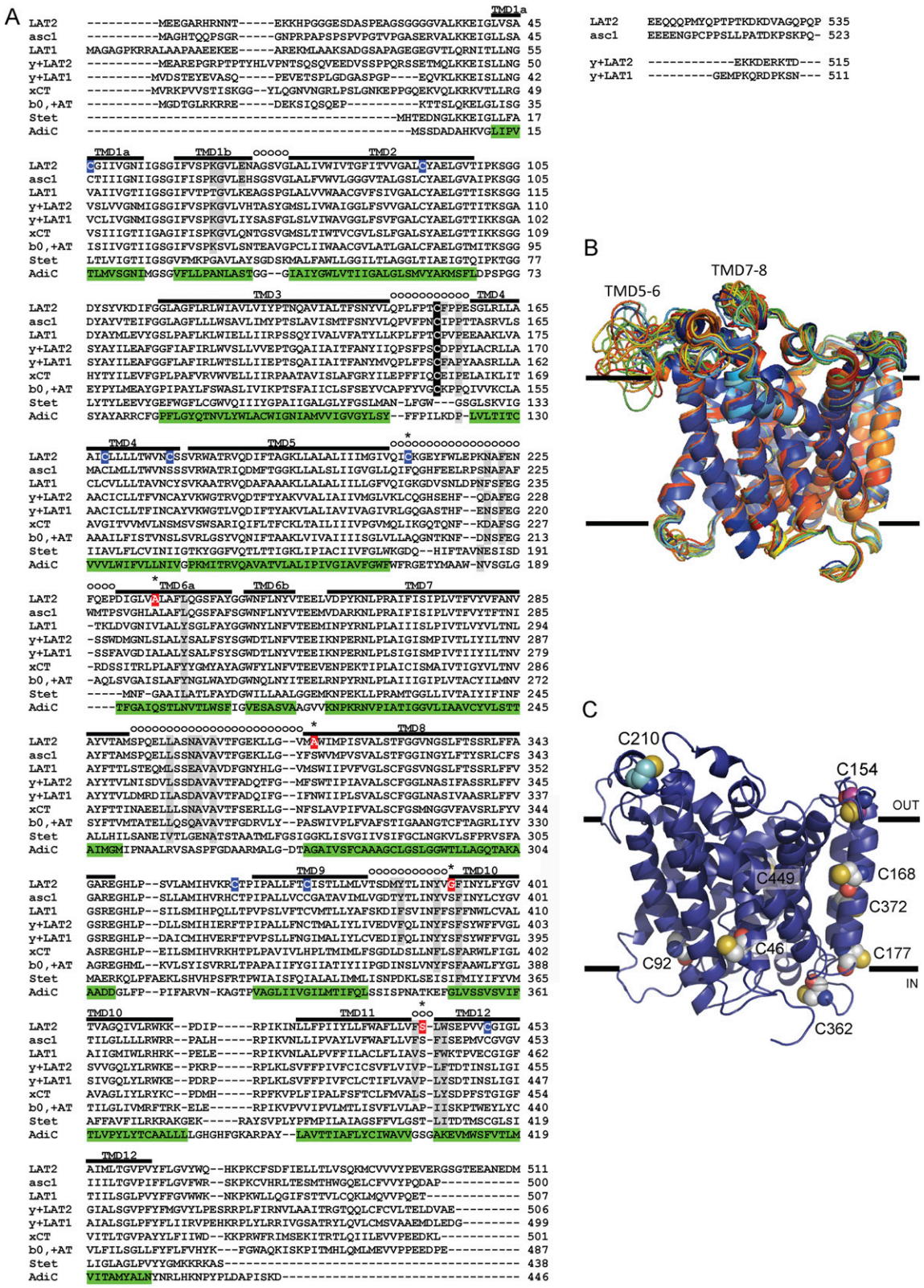


Fig. S3. Human LAT2 homology models. (*A*) Part of the extensive multiple amino acid sequence alignment between AdiC from *E. coli* and 250 close homologs (including bacterial, simple eukaryotic, and metazoan LATs) (1) used for homology modeling of human LAT2 is shown. The TMDs in human light subunits of HATs and the prokaryotic LAT SteT (serine/threonine exchanger) from *Bacillus subtilis* based on the crystal structure of AdiC are indicated over the alignment, and the TMD residues in AdiC are highlighted in green. Residues in external loops are shown as open circles. Conserved cysteine residues involved in the intersubunit disulfide bridge are boxed in black. Other endogenous cysteine residues in human LAT2 are boxed in blue. LAT2 residues mutated to cysteine are boxed in red. Human LAT2 residues crosslinked with 4F2hc-ED residues are indicated by an asterisk over the alignment. Conserved NIP residues are highlighted

Legend continued on following page

in gray. Small adjustments have been made by hand in loops TMD5–6, TMD7–8, TMD8–9, TMD10–11, and TMD11–12 of SteT and human LATs. (*B*) Superposition of the 20 homology models of human LAT2 built from AdiC monomer atomic structure (PDB ID code 3OB6) (2). Maximal divergence between the selected models occurs in the more flexible loops TMD5–6 and TMD7–8. (*C*) Cysteine residues in the best human LAT2 homology model. Lateral view with atoms of cysteine residues shown as spheres. The orientation of LAT2 in the membrane (marked by horizontal lines) was predicted using the OPM database (3). In *B* and *C* there are only two external cysteine residues: C154 (C atoms shown in purple), which is involved in the intersubunit disulfide bridge, and C210 (C atoms shown in cyan), which is located almost diametrically opposite.

1. Bartoccioni P, et al. (2010) Role of transmembrane domain 8 in substrate selectivity and translocation of SteT, a member of the L-amino acid transporter (LAT) family. *J Biol Chem* 285 (37):28764–28776.
2. Kowalczyk L, et al. (2011) Molecular basis of substrate-induced permeation by an amino acid antiporter. *Proc Natl Acad Sci USA* 108(10):3935–3940.
3. Lomize MA, Lomize AL, Pogozheva ID, Mosberg HI (2006) OPM: Orientations of proteins in membranes database. *Bioinformatics* 22(5):623–625.

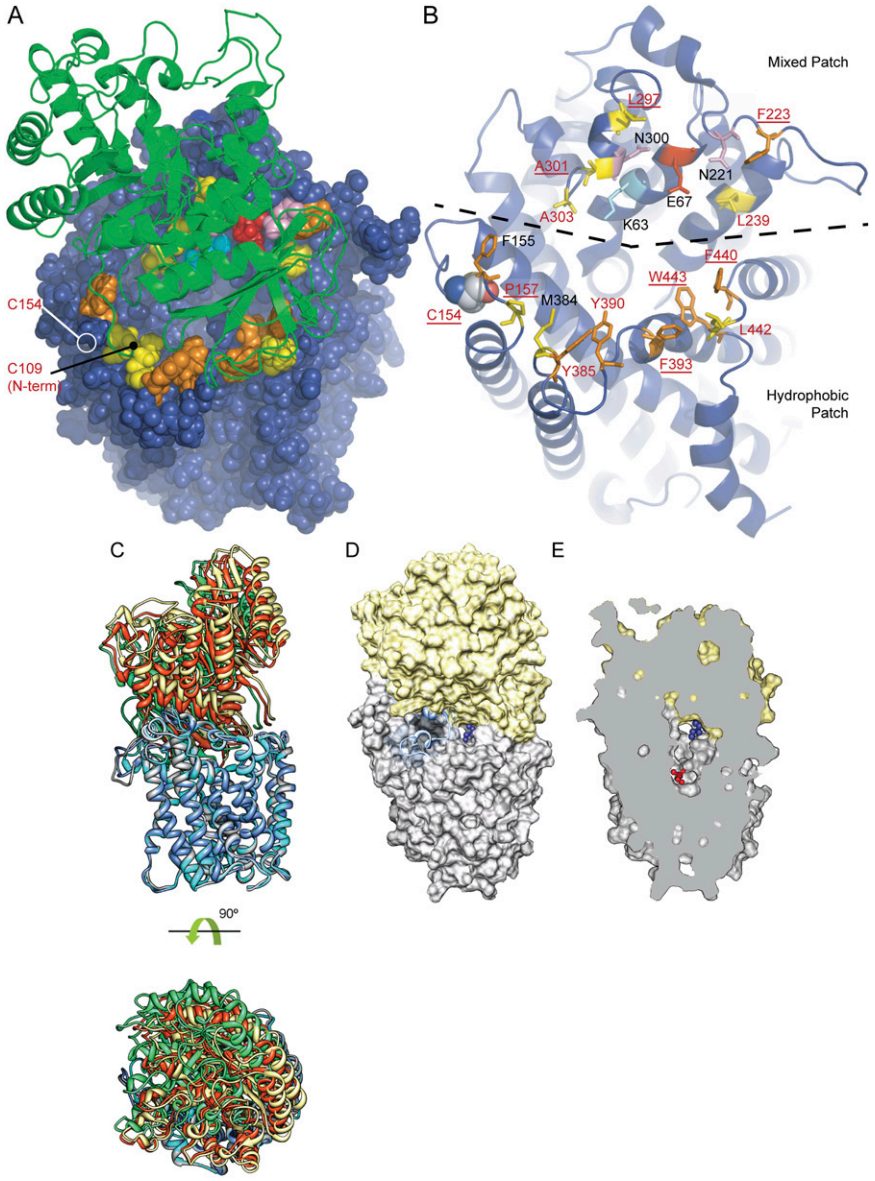


Fig. S4. 4F2hc-ED–LAT2 docking models showing LAT2 residues interacting with 4F2hc-ED and accessibility of substrate to binding site. (*A* and *B*) The most important residues of LAT2 for binding to 4F2hc-ED according to docking-derived NIP values. (*A*) Residues with NIP >0.4 are highlighted in the representation of LAT2 (blue, atoms shown as spheres) with the following color code: aromatic residues (orange), aliphatic residues (yellow), asparagine (light pink), glutamic acid (red), and lysine (cyan). The cartoon representation of 4F2hc-ED on top of LAT2 is shown in green. The hydrophobic patch of LAT2 interacts with subdomain C (antiparallel β8) of 4F2hc-ED. The cysteine residues involved in the disulfide bridge are indicated: LAT2 C154 is partially hidden, and one of its C atoms is indicated by a circle; 4F2hc C109 corresponds to the N-terminal end of the structure of 4F2hc-ED (indicated by black line). (*B*) Location of NIP-based interacting residues with cartoon representation of LAT2. Two defined patches (separated by the broken line) are evident: one hydrophobic and one mixed patch. The residues are represented by sticks color-coded as in *A*. C154 is highlighted with its atoms represented as spheres (C, gray; O, red; N, blue; and S, yellow). Totally or highly conserved residues are written in red, and nonconserved residues are written in black. Residues not conserved in cyanobacteria, yeast, and fungi homologs of LAT2 are underlined. Loop TMD5–6, containing residues N221 and F223, is larger in LAT2 than in AdiC and other bacterial LATs, precluding the generation of a robust homology model. A significant part of this loop is not at the interface with 4F2hc-ED, perhaps because of the poor modeling of this loop. (*C*) The three 4F2hc-ED–LAT2 docking models that satisfy the crosslinking experiments (Table S2) are displayed. The lowest-energy docking model (4F2hc-ED in red; LAT2 in gray) best satisfies the crosslinking experiments. Two additional docking models (ranked 3 and 15) satisfy the five strongest (≥60%) crosslinking experiments (considering an additional distance of 6 Å to account for small conformational changes). Docking models ranked 3 and 15 are shown with 4F2hc-ED/LAT2 in green/blue and yellow/cyan, respectively. A side view (*Upper*) and a top view (*Lower*) of the three best 4F2hc-ED–LAT2 docking models are shown. (*D* and *E*) The LAT2 external substrate vestibule is accessible in the lowest-energy 4F2hc-ED–LAT2 model. To visualize the accessibility of the LAT2 substrate vestibule from the exterior, we have docked L-isoleucine substrate with AutoDock and show here the most populated docking poses in blue and red. (*D*) The entrance of the substrate vestibule, where one of the most populated L-isoleucine positions is located, is accessible from the exterior. TMD7–8 loop (cartoon in cyan) flanks the entrance to this vestibule. (*E*) The interior of the substrate vestibule is visualized showing a second putative binding site of L-isoleucine (red) as determined by docking.

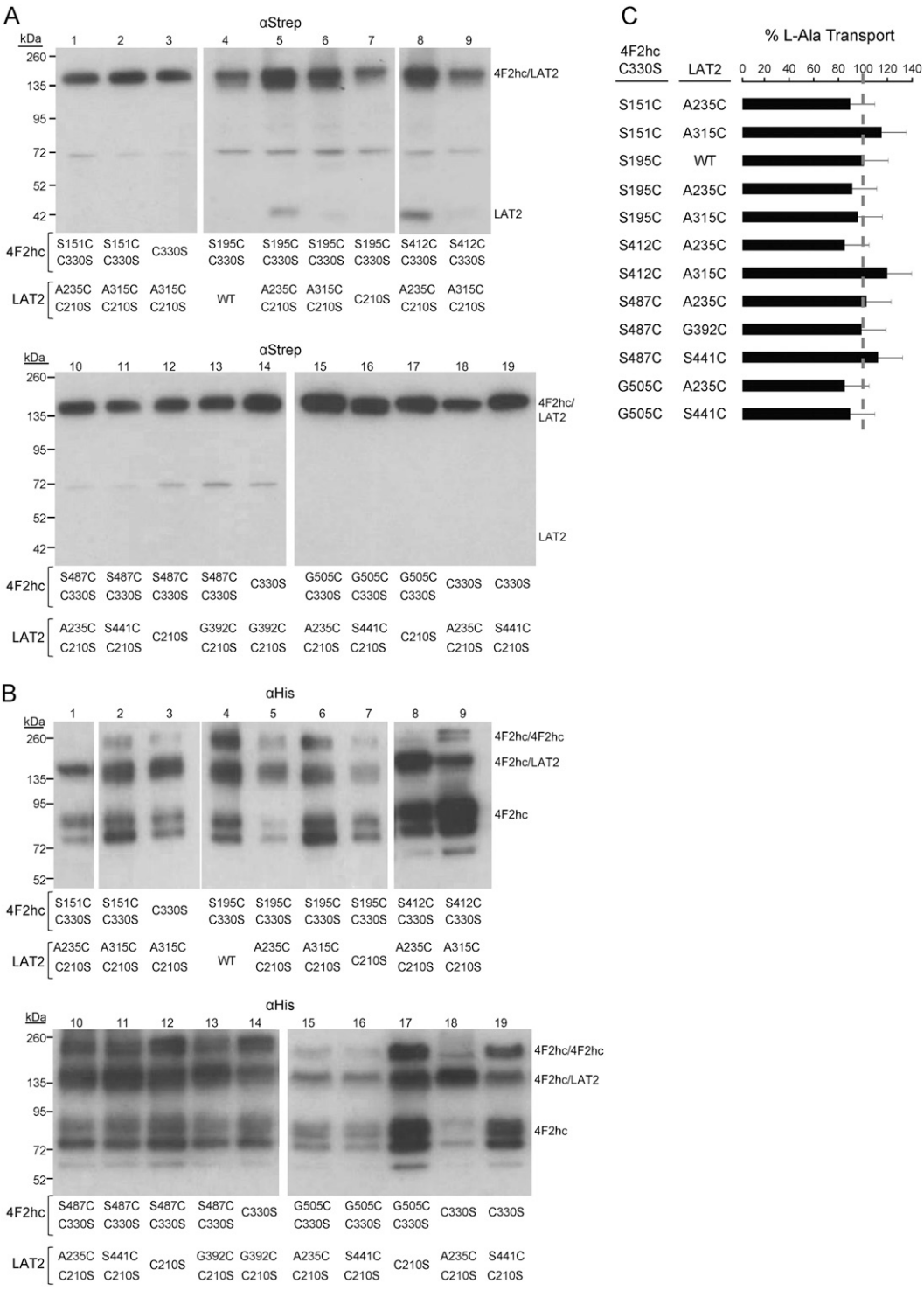


Fig. S5. Expression of 4F2hc/LAT2 heterodimers and induction of amino acid transport by mutants of 4F2hc and LAT2 generated after the docking studies. (*A* and *B*) Western blot analysis of 1% DDM-solubilized homogenates of HEK293T cells transfected with the indicated versions of His-4F2hc and Strep-LAT2 using αStrep (*A*) and αHis (*B*) antibodies under nonreducing conditions. Bands corresponding to heterodimers (4F2hc/LAT2), 4F2hc homodimers (4F2hc/4F2hc), and monomers (4F2hc or LAT2) are indicated. WT, wild-type version of LAT2. (*C*) Transport of L-alanine by the generated 4F2hc and LAT2 mutants. HEK293T cells were transfected with wild-type or the indicated mutants of Strep-LAT2 alone or together with the indicated mutants of His-4F2hc. One day later, cells were analyzed for transport of 10 μM L-[^{3}H]alanine for 15 s. Transport is expressed as the percentage of the transport elicited by the coexpression of 4F2hc C330S and wild-type LAT2 and has been calculated as indicated in the legend of Fig. S2*I*. Data are the mean ± SEM from one or two experiments run in quadruplicate. The coexpression of all of the indicated mutants induced L-alanine transport over the background caused by expression of the corresponding LAT2 version alone. Thus, all LAT2 mutants studied were functional.

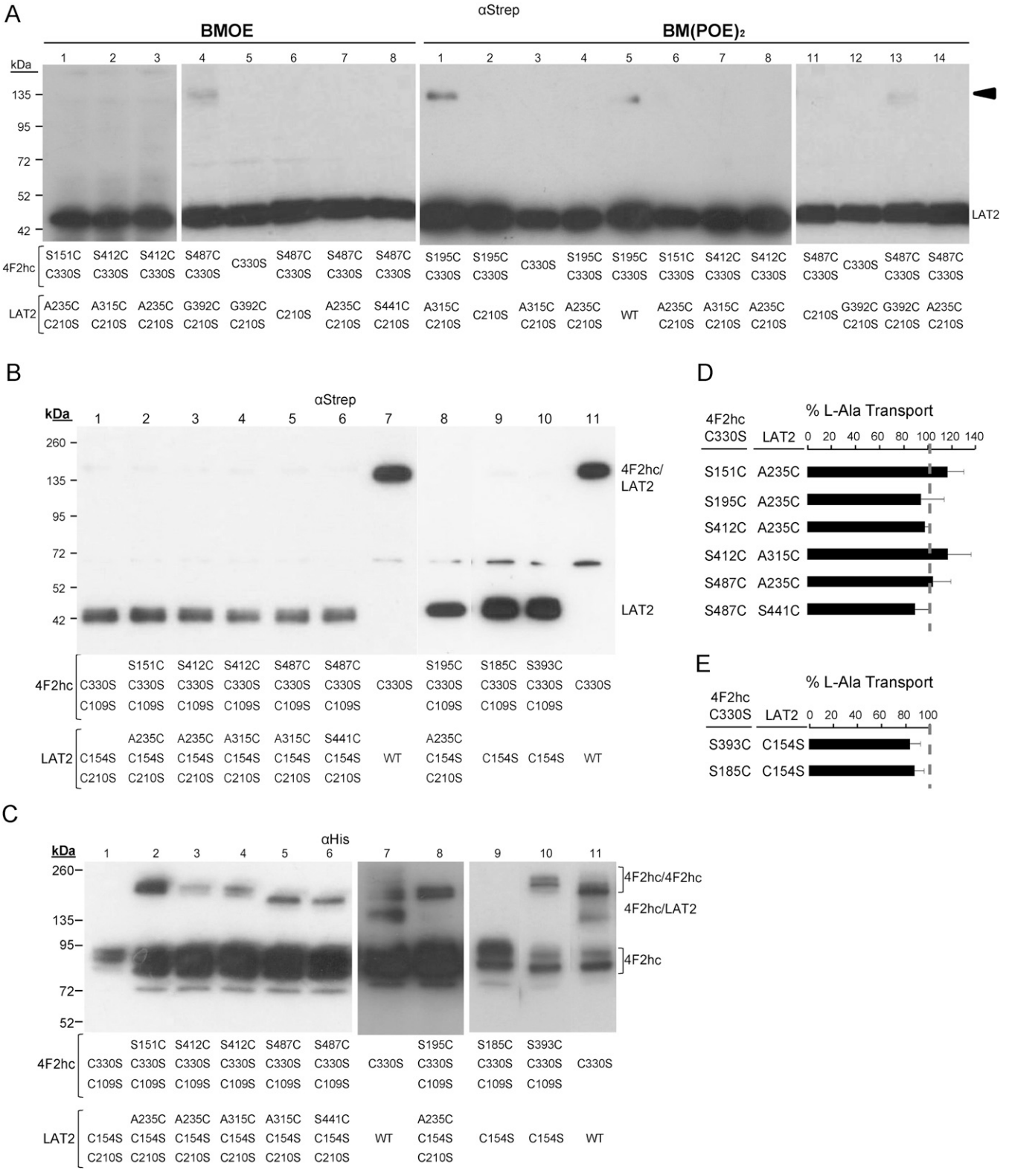


Fig. S6. Additional crosslinking analysis confirming the lowest-energy docking model of 4F2hc-ED and LAT2 complex. (*A*) Versions of His-4F2hc/Strep-LAT2 heterodimers from HEK293T cells purified by affinity chromatography were treated with the DTT-uncleavable crosslinkers BMOE or BM(POE)$_2$. Crosslinking was detected as DTT-resistant 4F2hc/LAT2 heterodimers (arrowhead) by Western blotting using αStrep antibodies under reducing conditions. The band of LAT2 monomer corresponds to LAT2 forming heterodimers with 4F2hc (i.e., copurified but not crosslinked). Single cysteine mutants were introduced in 4F2hc (C330S) and LAT2 (C210S). Weak but significant crosslinking in 4F2hc–LAT2 was detected between positions S487C–G392C (lane 4) by BMOE and between positions S195C–A315C (lane 1), S195C–Cys210 (wild-type LAT2) (lane 5), and S487C–G392C (lane 13) by BM(POE)$_2$. These crosslinking products were specific to these positions because crosslinking could be abolished by using the single mutants C330S (4F2hc) or C210S (LAT2). Crosslinking by BMOE and by BM(POE)$_2$ was

Legend continued on following page

negative for the following positions (4F2hc–LAT2): S151C–A235C, S412C–A235C, S412C–A315C, and S487C–A235C. Crosslinking by BMOE was negative for the paired positions S487C–S441C. (*B* and *C*) Cysteine pairs negative for crosslinking do not form direct intersubunit disulfide bridges. HEK293T cells were transfected with the indicated mutants of His-4F2hc and Strep-LAT2. Two days after transfection Western blot analyses using αStrep (*B*) or αHis (*C*) under nonreducing conditions were performed to detect the formation of intersubunit disulfide bridges. As indicated, analyzed mutants were introduced in 4F2hc (C330S/C109S) and LAT2 (C154S/C210S) or LAT2 (C154S) where the endogenous cysteine residues involved in the intersubunit disulfide bridge have been eliminated. In this way, 4F2hc/LAT2 heterodimers connected by a possible new disulfide bridge will be revealed under nonreducing conditions. With the exception of the positive control (cotransfection of 4F2hc C330S and wild-type LAT2 in lanes 7 and 11), none of the studied mutants showed the formation of 4F2hc/LAT2 heterodimers. Cysteine residues in several positions (S151C, S195C, S393C, S412C, and S487C), but not in position S185C, produced disulfide bridge-linked homodimers of 4F2hc (*C*). (*D* and *E*) Induction of L-alanine transport in transfected HEK293T cells. To validate our strategy, 1 d after transfection, the mutants studied in *B* and *C* were assayed for L-alanine transport for 15 s. Transport is presented as the percentage of the transport induced by cotransfection of 4F2hc (C109S/C330S) and LAT2 (C154S/C210S) (*D*), and 4F2hc (C109S/C330S) and LAT2 (C154S) (*E*), as described in the legend of Fig. S2*I*. Transport in these control conditions was 3,062 ± 167 and 3,017 ± 280 pmol L-alanine during 15 s (mean ± SEM from two experiments run in quadruplicate), respectively. Induction of L-alanine transport over the background elicited by LAT2 alone demonstrates the functional interaction of the mutants of 4F2hc and LAT2 that were studied.

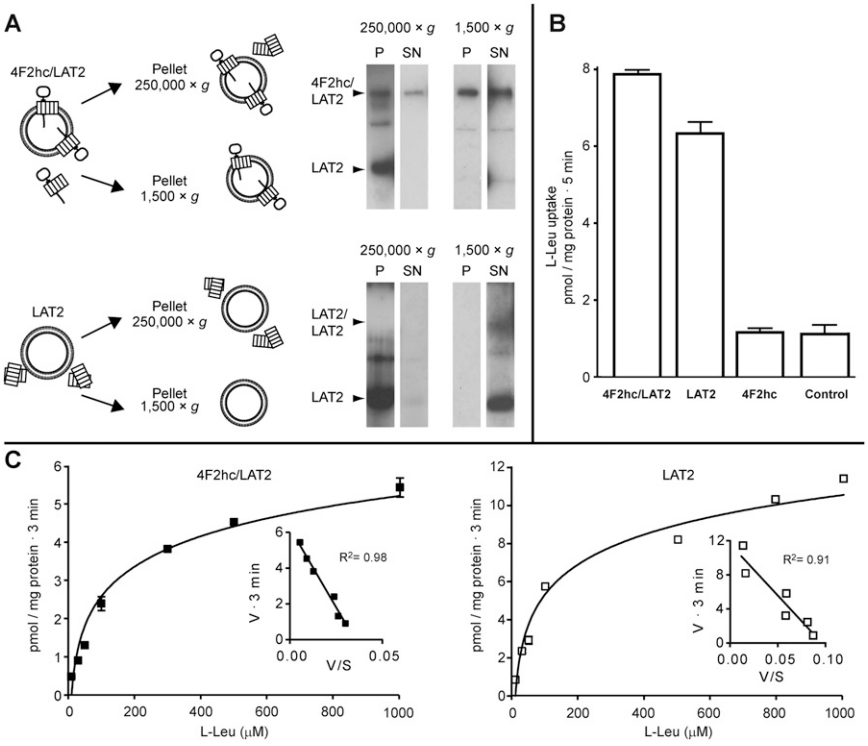


Fig. S7. Reconstitution into liposomes and functional characterization of 4F2hc/LAT2 and LAT2 in *Pichia*. (*A*) (*Left*) Reconstitution of heterodimers and LAT2 into liposomes. Membranes from *P. pastoris* coexpressing His-4F2hc and Strep-LAT2 or expressing Strep-LAT2 alone were solubilized with 1% DDM. Then detergent was removed to induce reconstitution of the membrane proteins into proteoliposomes. To distinguish between reconstituted and aggregated proteins, preparations were centrifuged (1,500 × *g*, 1 h, 4 °C) to precipitate proteoliposomes or were ultracentrifuged (250,000 × *g*, 1 h, 4 °C) to precipitate proteoliposomes and protein aggregates. (*Right*) Western blot analysis under nonreducing conditions with αStrep was used to detect Strep-LAT2 as monomers (LAT2) or heterodimers (4F2hc/LAT2). P, pellet; SN, supernatant. (*B*) In vivo uptake of L-leucine in different *Pichia* clones. Uptake of 10 μM L-[^{3}H]leucine over 5 min was determined in clones cotransformed with His-4F2hc and Strep-LAT2, transformed with Strep-LAT2, transformed with His-4F2hc, and in untransformed yeast cells (control). Data are the mean ± SEM of two independent experiments performed in triplicate. (*C*) Kinetic analysis of L-leucine transport in *P. pastoris* cells transformed with His-4F2hc and Strep-LAT2 (*Left*) or Strep-LAT2 alone (*Right*) yielded comparable K_ms of 178 ± 15 μM and 120 ± 7 μM, respectively. Transport was measured in linear conditions over 3 min. Data points were corrected for basal activity in untransformed yeast cells. The concentrations used were 10, 30, 50, 100, 300, 500, and 1,000 μM. (*Insets*) Eadie–Hofstee plots. Data are the mean ± SEM of representative experiments performed in triplicate. When not visible, error bars are smaller than symbol. R^2, correlation coefficient of the linear regression.

Table S1. Conservation of the NIP residues placed in the human 4F2hc–LAT2 interface

NIP			Conservation				
				Other 4F2hc LS			
Loop	Residue	Location	LAT2	Human	All	$b^{0,+}AT$	Cyanobacteria, yeast, and fungi
TMD1–2	K63	TMD1	Yes	No*	No*	Yes	No
	G64	TMD1	Yes	Yes	Yes	No	No
	E67	TMD1 (C-term)	Yes	No	No	No	No
TMD3–4	F155	Loop	Yes	No	No	No	No
	P157	Loop	Yes	Yes	Yes	Yes	No
TMD5–6	N221	Loop	Yes[†]	Polar	Polar	Polar	No
	F223	Loop	Yes	Yes	Yes	Yes	No
	L239	TMD6	Yes	L/Y	L/Y	L/Y	V/Y/I
TMD7–8	L297	Loop	Yes	Yes	L/I/V[‡]	Yes	No
	N300	Loop	Yes	Polar	Polar	Polar	Polar
	A301	Loop	Yes	Yes	Yes	Yes	No
	A303	Loop	Yes	Yes	Yes	Yes	Yes
TMD9–10	M384	Loop	Yes[§]	No	No	No	No
	Y385	Loop	Yes	Y/F[¶]	Y/F[¶]	NO	Y/F
	Y390	Loop	Yes	Y/F	Y/F	Y/F	Y/L/F
	F393	TMD11 (N-term)	Yes	Yes	Yes	Yes	No
TMD11–12	F440	Loop	Yes	F/V/L	F/V/L	No	No
	L442	TMD12	Yes	L/F	L/F/Y	L/Y/I	L/V/F/Y
	W443	TMD12	W/Y	W/Y[‖]	W/Y[‖]	I	No

The amino acid sequence of 72 proteins belonging to the LAT family (SLC7) and grouped in seven categories (LAT2, LAT1, y^+LAT1, y^+LAT2, asc1, xCT, $b^{0,+}AT$) (with eight vertebrate proteins per category) and cyanobacterial, yeast, and fungal LATs (with 16 proteins per category) were aligned by the ClustalW2 program. The conservation of the NIP residues was analyzed in vertebrate LAT2, in other vertebrate light subunits (LS) associated with 4F2hc, in vertebrate $b^{0,+}AT$, in cyanobacterial LATs, and in fungal LATs. The protein sequences were retrieved from the UniProt and GenBank databases. The following protein sequences were retrieved from the UniProt database: LAT1: *Homo sapiens* (Q01650), *Mus musculus* (Q9Z127), and *Xenopus tropicallis* (F6QFZ9); LAT2: *Homo sapiens* (Q9UHI5), *Mus musculus* (Q9QXW9), *Anolis carolinensis* (H9GCK9), *Xenopus tropicallis* (Q6P891), *Carassius carassius* (B8YMU3) and *Takifugu rubripes* (H2SUN0); y^+LAT1: *Anolis carolinensis* (H9G4U3), *Xenopus tropicallis* (F6S1W4), and *Latimeria chalumnae* (H3A009); y^+LAT2: *Homo sapiens* (Q92536), *Mus musculus* (Q8BGK6), *Gallus gallus* (F1ND99), *Xenopus tropicallis* (Q28I80), *Pelodiscus sinensis* (K7G1V4); asc1: *Homo sapiens* (Q9NS82), *Mus musculus* (P63115), *Xiphophorus maculatus* (M4AC17), and *Pelodiscus sinensis* (K7FII9); xCT: *Homo sapiens* (Q9UPY5), *Mus musculus* (Q9WTR6), and *Gallus gallus* (E1C734); $b^{0+}AT$: *Homo sapiens* (P82251), *Gallus gallus* (E1BT87), *Anolis carolinensis* (H9GCM6), *Xenopus laevis* (A9JS28), and *Danio rerio* (F1QGJ2). The following protein sequences were retrieved from the GenBank database: LAT1: *Ornithorhynchus anatinus* (XM_001508884), *Orcinus orca* (XM_004280070), *Gallus gallus* (XM_414194), and *Danio rerio* (NM_001128358); LAT2: *Ornithorhynchus anatinus* (XM_003429011), *Orcinus orca* (XM_004283215), and *Danio rerio* (NM_001271897.1); y^+LAT1: *Homo sapiens* (NP_001119577.1), *Rattus norvegicus* (AAF07216.1), *Ornithorhynchus anatinus* (XP_003430084.1), *Orcinus orca* (XP_004283244.1), and *Danio rerio* (NP_001032648.1); y^+LAT2: *Orcinus orca* (XM_004273192), *Anolis carolinensis* (XM_003222960), and *Danio rerio* (NM_001020474); asc1: *Orcinus orca* (XM_004284140), *Gallus gallus* (XM_414136), *Xenopus tropicallis* (XM_002935793), and *Danio rerio* (XM_690716); xCT: *Orcinus orca* (XM_004265135), *Anolis carolinensis* (XM_003221678), *Xenopus tropicallis* (XM_002932048), and *Danio rerio* (XM_002664447); $b^{0+}AT$: *Rattus norvegicus* (NM_053929), *Orcinus orca* (XM_004284134), and *Salmo salar* (NM_001173819); Cyanobacterial LATs: *Chroococcidiopsis sp.* (WP_015152622.1), *Cyanothece sp.* (WP_012630008), *Fischerella sp.* (WP_009454021.1), *Gloeobacter violaceus* (NP_924598), *Oscillatoria sp.* (WP_007357767), *Pleurocapsa minor* (WP_015143465), *Microcoleus sp.* (WP_015184929.1), and *Rivularia sp.* (WP_015120455.1); Yeast and fungal LATs: *Saccharomyces cerevisiae* (NM_001181184), *Arthroderma gypseum* (XM_003171846), *Schizophyllum commune* (XM_003036114), *Yarrowia lipolytica* (XM_505103), *Aspergillus niger* (XM_001396535), *Coprinopsis cinerea* (XM_001833155), *Nectria hematococca* (XM_003050065), and *Ajellomyces dermatitidis* (XM_002620161).

*K not present only in LAT1.

[†]H in *Anolis carolinensis* LAT2.

[‡]M in *Xenopus tropicallis* and H in *Mus musculus* y^+LAT2.

[§]I in *Anolis carolinensis* LAT2.

[¶]D in human xCT and H in *Xiphophorus maculatus* and *Danio rerio* asc1.

[‖]I in human asc1.

Table S2. Crosslinking analysis summary

Position	4F2hc residue	Location†	LAT2 residue	Location†	% crosslinking* BMOE (6.3–10.5 Å)‡	% crosslinking* $BM(POE)_2$ (3.5–14.3 Å)‡	Distance, Å§
Crosslinked	S151	Aα1-Aβ2	WT (C210)	TMD5–6	50	80	13.4
			A315	TMD7–8	25	60	14.9
	S195	Aα2	WT (C210)	TMD5–6	NA	5	17.5
			A315	TMD7–8	NA	10	15.5
	S412	Aα8‴	WT (C210)	TMD5–6	35	95	11.5
	S487	Cβ5	G392	TMD9–10	5	10	12.9
	G505	Cβ6	A235	TMD5–6	90	NA	8.4
			S441	TMD11–12	90	NA	8.1
Not crosslinked	S151	Aα1-Aβ2	A235	TMD5–6	No	No	27.0
	S185	Aα2	WT (C210)	TMD5–6	No	No	28.2
	S195		A235	TMD5–6	NA	No	37.9
	S393	Aα6-Aβ7	WT (C210)	TMD5–6	No	No	18.8
	S412	Aα8‴	A315	TMD7–8	No	No	27.1
			A235	TMD5–6	No	No	19.9
	S487	Cβ5	A235	TMD5–6	No	No	27.2
			S441	TMD11–12	No	NA	14.2

NA, not analyzed for crosslinking; No, residues not crosslinked.

*Percentage of LAT2 crosslinked was calculated as the amount of LAT2 crosslinked from the total LAT2 coimmunoprecipitated with His-4F2hc.

†The structural element where the residues are located is shown.

‡The range of crosslinking distance of BMOE and $BM(POE)_2$ is indicated in parentheses.

§Minimal interatomic distance between crosslinking residues in the lowest-energy docking model.

Summary in Spanish

1. Introducción

Debido a su relevancia fisiológica, las proteínas de membrana representan actualmente dos tercios de las dianas terapéuticas (Hediger *et al.,* 2013). No obstante, hasta el momento, sólo se han depositado 466 estructuras resueltas en la base de datos 'Protein Data Bank' (PDB, abril 2014), representando menos del 2% del total de estructuras resueltas por cristalización y, siendo mayormente proteínas de membrana de organismos procariotas. Este número tan bajo se debe principalmente a el carácter hidrofóbico de los segmentos integrales de las proteínas de membrana los cuáles se encuentran embebidos en bicapas de fosfolípido, dificultando así su sobreexpresión, purificación y cristalización (Arinaminpathy *et al.,* 2009; Bill *et al.,* 2011).

La importancia de los transportadores de aminoácidos recae básicamente en la importancia del papel que tienen los aminoácidos en las células, los cuales juegan papeles clave tales como la síntesis de proteínas, la regulación del metabolismo, crecimiento celular, regulación del volumen celular, producción de energía metabólica, transmisión neuronal, y moléculas de señalización celular. Los transportadores de aminoácidos son esenciales para la absorción de los aminoácidos en la nutrición, mediando la transferencia de aminoácidos entre órganos, entre células, y entre compartimentos celulares. Conjuntamente con otros transportadores, ayudan al mantenimiento de la concentración iónica celular, flujo de nutrientes, abasteciendo de aminoácidos a los tejidos del cuerpo (Christensen *et al.,* 1990).

Los transportadores de aminoácidos heteroméricos (HATs) sólo se encuentran en metazoos mediando el transporte de aminoácidos a través de la membrana celular y jugando papeles clave como reguladores de la síntesis de proteínas, crecimiento celular, regulando el metabolismo, el volumen celular, la producción de energía metabólica, la transmisión neuronal, y moléculas de señalización molecular.

Los HATs se caracterizan por ser el único ejemplo conocido en todos los reinos animales de transportadores de solutos constituidos por dos subunidades unidas distintas las cuales se encuentran unidas por un puente disulfuro (Palacín & Kanai, 2004). En mamíferos han sido identificadas dos subunidades pesadas (4F2hc o CD98, y rBAT) y 10 subunidades ligeras (Fernández *et al.,* 2005).

La subunidad pesada, perteneciente a la familia SLC3 de transportadores de solutos en mamíferos, es la responsable del tráfico del heterodímero a la membrana plasmática. Así como la subunidad ligera, pertenece a la familia de transportadores LAT, dentro de la familia SLC7, y es la unidad catalítica, la que confiere la especificidad de transporte de aminoácidos (Reig *et al*., 2002; Bröer & Palacín, 2011).

Estructuralmente, las subunidades pesadas son glicoproteínas de tipo II (N-terminal intracelular) con un solo segmento transmembrana, mientras que las cadenas ligeras no están glicosiladas y están formadas por 12 segmentos transmembrana.

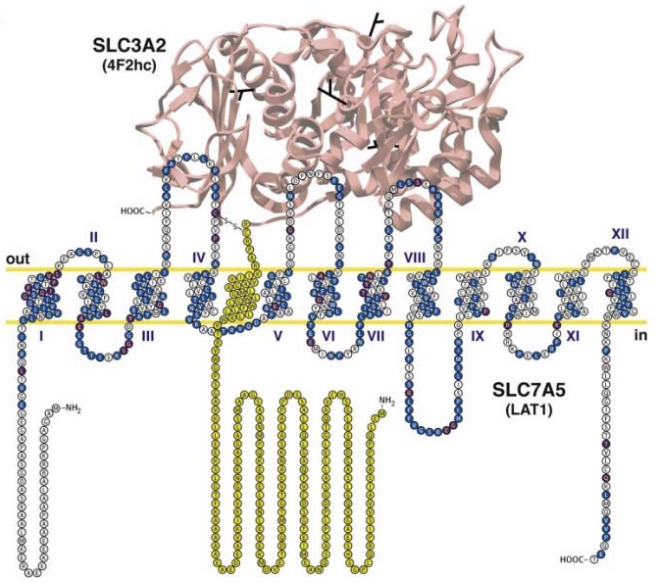


Figura 1. Modelo de 4F2hc/LAT1 (SLC3A2/SLC7A5) como miembro representativo de las proteínas HAT. Muestra la topología de LAT1 con 12 segmentos transmembrana, y la estructura atómica de 4F2hc-ED humano con 4 glicosilaciones se encuentra localizada en el esquema sobre el segmento transmembrana y N-terminal intracelular ilustrados de color amarillo. Adaptado de (Fotiadis, Kanai & Palacín, 2013).

Diversas patologías humanas destacan el papel fisiológico de los HAT. Dos transportadores HAT son responsables de aminoacidurias hereditarias. Así, mutaciones en cualquiera de los dos genes que codifican para las subunidades del sistema de transporte $b^{0,+}$ (rBAT y $b^{0,+}$AT) dan lugar a cistinuria (MIM 220100) (Calonge *et al.,* 1994; Feliubadaló *et al.,* 1999), mientras que mutaciones en y^+LAT1 (uno de los dos systemas y^+L asociados a 4F2hc) producen lisinuria con intolerancia a proteínas (LPI) (MIM 222700) (Torrents *et al.,* 1999; Borsani *et al.,* 1999).

Por otro lado, la subunidad ligera xCT asociada a 4F2hc media el intercambio de cistina (captación) y glutamato (eflujo) (sistema xC^-) (Sato *et al*., 1999; Bassi *et al.,* 2001), y es esencial para la infección del herpesvirus asociado al sarcoma de Kaposi (herpesvirus humano 8) (Kaleeba & Berger, 2006; Veettil *et al.,* 2008). El sistema xC^- también está involucrado en la recaída a la adicción a cocaína a través del control de los niveles basales de glutamato extrasináptico (Baker *et al.,* 2003).

Este sistema es además el principal regulador del balance redox en la membrana plasmática (Sato *et al.,* 2005). Finalmente, LAT1 (systema L asociado a 4F2hc) está sobreexpresado en diferentes tumores (Yamahuchi *et al.,* 2009) y transporta aminoácidos neutros esenciales para la señalización de mTOR (Nicklin *et al.,* 2009).

1.1. Subunidades pesadas (SLC3)

Existen dos subunidades distintas descritas: rBAT y 4F2hc. Hasta el momento, no se ha descrito que tengan una función en el transporte, ni en su modulación. Mientras que 4F2hc es una proteína expresada ampliamente en todos los tejidos, rBAT se localiza principalmente en las células epiteliales, tanto en la nefrona del riñón como en los enterocitos intestinales. Además, ambas subunidades tienen localización diferencial en células epiteliales, ya que rBAT se encuentra en la membrana apical, mientras que 4F2hc lo hace en la membrana basolateral.

4F2hc, además de llevar el transportador a membrana, juega múltiples funciones que se encuentran relacionadas con la regulación de la señalización de β-integrinas. Se ha demostrado que la interacción de la parte citoplasmática de 4F2hc induce la activación de la señalización de β1- y β3-integrinas (Feral *et al.,* 2005), las cuales regulan procesos de proliferación celular (Cantor *et al.,* 2009), adhesión, migración (Fogelstrand *et al.,* 2009) entre otros. Además, 4F2hc se encuentra sobreexpresado en multitud de tumores. Curiosamente, no se han descrito mutaciones en 4F2hc, ya que su deleción en ratones supone la muerte del feto, debido a que los embriones no consiguen implantarse (Tsumura *et al.,* 2003; Domínguez *et al.,* 2010).

rBAT Y 4F2hc comparten un 30% de identidad en secuencia y un 50% de similitud. Son proteínas muy glicosiladas: 4F2hc humano presenta 4 sitios de glicosilación y rBAT presenta 6.

Actualmente sólo se conoce la estructura atómica de la parte extracelular de 4F2hc (el ectodominio), la cual fue resuelta a 2.1Å en nuestro grupo (Fort *et al.*, 2007), la cual se vio era muy similar a las alfa-amilasas procariotas. 4F2hc-ED adolece de los residuos catalíticos clave pero mantiene la hendidura catalítica para la que no se conocen ligandos. Contienen el característico dominio A (formado por un barril TIM), el dominio B, presente solo en el caso de rBAT, y el domino C, formado por laminas β. La estructura de rBAT se desconoce, ya que no se ha conseguido su sobreexpresión debido a su inestabilidad, y hasta el momento solo existen modelos bioinformáticos de su estructura en base a su similitud con 4F2hc y las α-amilasas (Chillarón *et al.*, 2010) y se desconoce si posee actividad enzimática.

Se cree que podría existir alguna subunidad pesada sin identificar, ya que existen dos cadenas ligeras (asc2 y arpAT) de las que todavía no se conocen ninguna subunidad pesada interaccionando.

1.2. Subunidades ligeras (SLC7)

Las subunidades ligeras son intercambiadores de aminoácidos, y poseen distinta especificidad de sustrato. Las subunidades ligeras de los HAT pertenecen a la familia LAT, una de las integrantes de la gran superfamilia de transportadores APC (Aminoácidos, Poliaminas y Cationes orgánicos) (Jack *et al.,* 2000). Se han identificado 10 cadenas ligeras, de las cuales 8 conocemos su subunidad pesada: $b^{0,+}$AT y AGT1 se unen a rBAT, mientras que LAT1, LAT2, y^+LAT1, y^+LAT2, asc-1 y xCT son las cadenas ligeras de

4F2hc. Todas ellas son proteínas de ~50 kDa no glicosiladas y altamente hidrofóbicas, lo que resulta en que su movilidad en SDS-PAGE alrededor de 35-40 kDa.

Estudios bioquímicos indican una topología de 12 segmentos transmembrana, así como la localización intracelular de ambos extremos de la proteína (Gasol *et al.,* 2008).

La formación del puente disulfuro con la subunidad pesada no impide la formación del heterodímero, que igualmente consigue llegar a membrana, sugiriendo la existencia de otras interacciones entre ambas proteínas. No obstante, la reconstitución en liposomas muestra que la subunidad ligera es completamente funcional en ausencia de la pesada, en el caso de $b^{0,+}$AT (Reig *et al.*, 2002).

1.3. Información estructural de los LATs

Actualmente, no se dispone de ninguna estructura atómica de las subunidades ligeras. No obstante, sí se conoce la estructura de distintos homólogos procariotas:

LAT homologos	Estado Oligomérico	Co-crystallización	Resolución	Protein Data Bank (PDB)	Referencias
AdiC	Tetrámero	Fab Fragmento	3.2 Å	3NCY	Fang *et al.,* 2009
	Dímero	-	4 Å	3LRC	Gao *et al.,* 2009
		L-arginina	3 Å	3L1L	Gao *et al.,* 2010
		L-arginina	3 Å	3OB6	Kowalczyk *et al.,* 2011
ApcT	-	-	2.32 Å	3GIA	Shaffer *et al.,* 2009
		Fab Fragmento	2.48 Å	3GI9	
			2.59 Å	3GI8	
GadC	Dímero	-	3.1 Å	4DJK	Ma *et al.,* 2012

Tabla 1. Actuales estructuras atómicas dentro de la familia de transportadores procariotas APC (Hediger et al., 2013).

Así pues, las estructuras atómicas de dos transportadores procariotas de la superfamilia APC: AdiC (familia APA), ApcT (familia ABT) y GadC, representan los actuales paradigmas estructurales de los LATs. Estos transportadores muestran el plegamiento prototípico "5+5 repetición invertida", inicialmente descrito en el transportador de aminoácidos LeuT (Aquifex aeolicus) (Yamashita *et al.,* 2005), y posteriormente identificado también en otros transportadores procariotas no relacionados (~10% de identidad de secuencia de aminoácidos), (Faham *et al.,* 2008; Weyand *et al.,* 2008; Ressl *et al*., 2009). Todos ellos presentan un mismo plegamiento, descrito por primera vez para LeuT, que consiste en dos unidades de 5 segmentos transmembrana invertidos y que son compartidos por diversos tipos de transportadores. La existencia de estructuras para cada estado conformacional ha permitido la comprensión de su mecanismo de transporte de acceso alternante a sustrato (Fotiadis *et al.,* 2013).

La hidrofobicidad de las proteínas de membrana dificulta la obtención de cristales para estudios de cristalización. Particularmente, las subunidades ligeras son altamente inestables en solución aún en presencia de detergentes y la obtención de cristales a una resolución adecuada para resolver su estructura atómica es un reto.

Si bien las estructuras publicadas de los transportadores APC homólogos a las subunidades ligeras y del ectodominio de 4F2hc significan un claro avance en nuestro conocimiento de estos transportadores, su utilidad para modelar los transportadores HAT está limitada por tres claras deficiencias:

- Baja homología con las subunidades ligeras de los HAT humanos (<20%),
- falta de información de la interacción con el sustrato y
- no revelan las interacciones moleculares entre las subunidades ligeras y pesadas.

Por ello, en el laboratorio se iniciaron distintas estrategias para abordar el estudio estructural de los transportadores de aminoácidos. Por un lado, se ha estudiado el transportador de arginina/agmatina bateriano, homólogo procariota de las subunidades ligeras, AdiC (~18% homologo a los LATs), del cual se obtuvo la estructura en conformación abierta con sustrato unido (PDB 3BO6) (Kowalczyk *et al.*, 2011). Por otro lado, el estudio del transportador bacteriano más cercano a los LATs y que se encuentra bien caracterizado funcionalmente: SteT (intercambiador de L-serina/L-threonina) (comparte ~30% de identidad de secuencia de aminoácidos con los LATs). Desafortunadamente, los intentos de cristalización con la versión salvaje fueron fallidos.

Con este escenario, en el laboratorio se inició un proyecto para probar la sobreexpresión de los HATs humanos en un sistema de expresión eucariota, *Pichia pastoris*. Se clonaron las 7 subunidades ligeras humanas y 4F2hc y rBAT humanos en pPICZ con tags localizados ambos en N-terminal de 8 Histidinas y StrepII-tag para las subunidades pesadas y ligeras respectivamente. Se testó la expresión y solubilización en varios detergentes en todas las subunidades (resultados de Dr. Rosell y Dr. Costa) (Costa *et al.,* 2013).

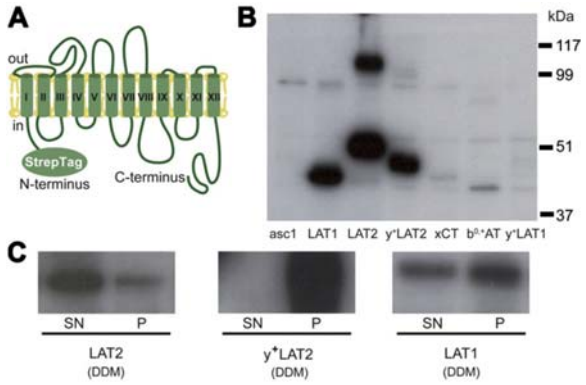


Figura 2. Resultados de expresión y solubilización de las subunidades ligeras de los HATs: asc1, LAT1, LAT2, y+LAT2, xCT, b0,+AT y y+LAT1. A) Representación esquemática del diseño de la cadena ligera con StrepIItag en N-terminal. **B)** Análisis por western blot usando anticuerpo anti-StrepTagII de la expresión de las diferentes subunidades ligeras a partir de membranas aisladas. **C)** Solubilización con DDM de las tres mejor subunidades ligeras que expresaron mejor: LAT2, y^+LAT2, LAT1, representando el western blot de la fracción solubilizada (SN: supernatant), y la precipitada después de ultracentrifugación a 55.000 rpm durante 1 hora (P: pellet) (Costa *et al*., 2013).

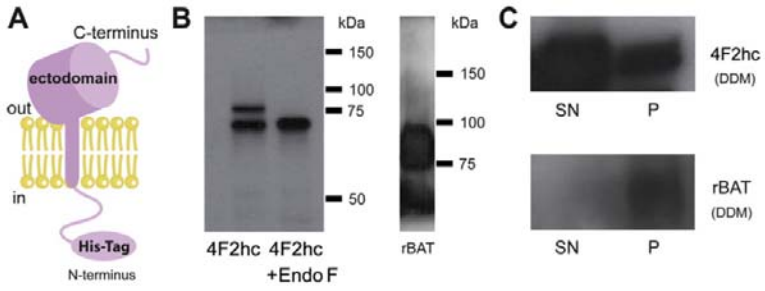


Figura 3. Resultados de expresión y solubilización de las subunidades pesadas de los HATs: 4F2hc y rBAT. A) Representación esquemática del diseño de la cadena pesada con tag de 8xHistidinas en N-terminal. **B)** Análisis por western blot usando anticuerpo anti-4F2hc(H300) de la expresión de las 4F2hc, 4F2hc digerido con EndoF para eliminar la parte glicasilada y rBAT a partir de membranas aisladas. C) Solubilización con DDM de 4F2hc (arriba) y rBAT (debajo) representando el western blot de la fracción solubilizada (SN: supernatant), y la precipitada después de ultracenctrifugación a 55.000 rpm durante 1 hora (P: pellet) (Costa *et al*., 2013).

De las cuales, 4F2hc y LAT2, presentaron los candidatos que mejor se expresaron y solubilizaron. Por ello, se probó su co-expresión que también resultó con buenos niveles de expresión y solubilización. A partir de un cultivo de 2 litros, se extrajeron membranas, solubilizaron en DDM y tras dos purificaciones secuenciales por cromatografía de afinidad de Ni^{++} y anti-strepTaqII se consiguió alrededor de 1mg de HAT. Se evaluó su perfil de estabilidad por cromatografía de exclusión molecular que presentó un pico formado mayormente por heterodímero y algunos monómeros debido a la rotura del puente disulfuro (figura 4A). Además, se expresó y purificó la subunidad ligera LAT2, pero sorprendentemente, sin la presencia de 4F2hc, la subunidad ligera agrega toda una vez purificada, eluyendo toda en el void (volumen 8mL) (figura 4A). Además se comprobó que 4F2hc/LAT2 fuese funcional (figura 4B). Ya que LAT2 humano no era estable suficientemente, no fue posible reconstituirlo en proteoliposomas, así pues, sólo se pudo reconstituir en proteoliposomas el heterodímero y 4F2hc, que no resultó estar implicado en transporte.

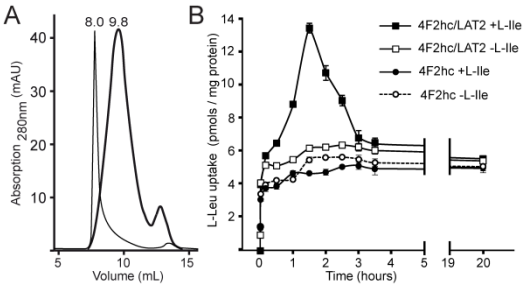


Figura 4. A) Perfiles de elución de LAT2 (línea estrecha) y el heterodímero 4F2hc/LAT2 (línea más gruesa) en cromatografía de exclusión molecular. 4F2hc/LAT2 eluye a 9.8mL, en un pico monodisperso, mientras que LAT2 eluye todo en el volumen void de la columna (8mL). **B)** Time-course de transporte de L-leucina en 4F2hc/LAT2 (por Dr. Costa, Rosell *et al.,* 2014).

A pesar de que la cantidad de heterodímero que llega al final de la purificación es una fracción importante, no representa suficiente cantidad para realizar estudios de cristalización en 3D, ya que se trata de una proteína no suficientemente estable.

No obstante, fue suficientemente estable para estudiar su estado oligomérico y estructura por microscopía electrónica de transmisión, haciendo tinción negativa y analizando computacionalmente una galería de más de 15,000 imágenes de las partículas solas o formando heterodímeros (en colaboración con Dr. Fotiadis i Dr. Meury).

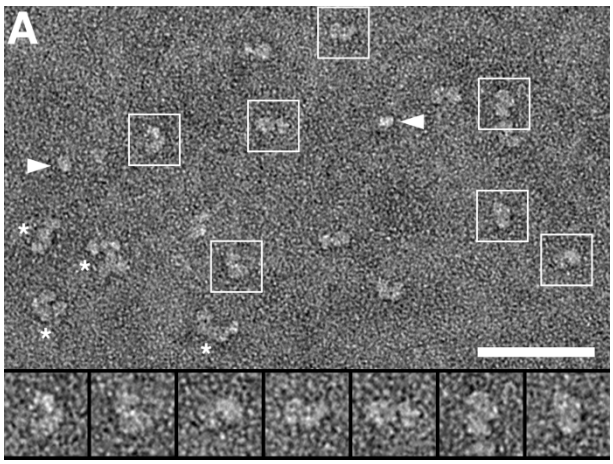


Figura 5. Micrografía 207lectrónica de heterodímeros de 4F2hc/LAT2 sobreexpresados en P.pastoris, purificados y teñidos negativamente para TEM. Recuadrados están los complejos que se seleccionaron par formar la galería de imágenes. Se descartaron los monómeros de 4F2hc o LAT2 provenientes de heterodímeros disgregados (flechas), así como los agregados (asteriscos) (en colaboración con Dr. Fotiadis i Dr. Meury) (Rosell *et a*l., 2014).

A partir de esta galería de imágenes se pudo reconstruir el primer modelo estructural de un HAT (4F2hc/LAT2) a baja resolución (21Å) (Rosell *et al.,* 2014).

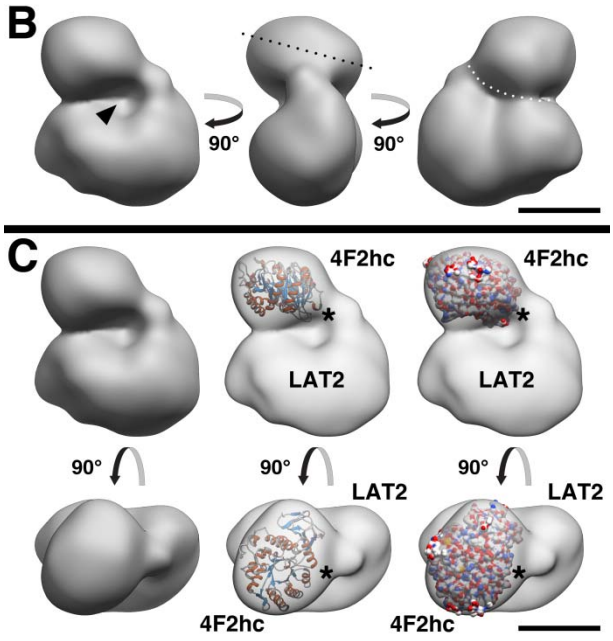


Figura 6. Reconstrucción en 3D del heterodímero 4F2hc/LAT2. La estructura de 4F2hc-ED se corresponde con la densidad más pequeña de la estructura. El asterisco muestra la posible localización del puente disulfuro y la densidad grande correspondería a la subunidad ligera rodeada de detergente y lípidos de membrana (en colaboración con el grupo de Prof. D. Fotiadis (University of Bern, Switzerland)) (Rosell *et al.,* 2014).

Este modelo 3D representa el primer modelo de un HAT humano, sugiriendo que 4F2hc-ED se localizaría en la densidad pequeña dónde cabe la estructura resuelta del ectodominio (Fort *et al.,* 2007) (Figura 6), y, la densidad grande correspondería a la subunidad ligera envuelta de detergente y lípidos.

Este modelo, una vez validado, constituiría la primera vez que se describe 4F2hc-ED posicionado sobre los loops extracelulares de LAT2. Mientras que anteriormente, se había creído estaría en contacto con la membrana plasmática (según modelización de la estructura de 4F2hc-ED Fort *et al.,* 2007).

Objetivos:

Debido a la baja resolución del modelo y la falta de más evidencias acerca de la composición de cada una de las subunidades, se presentó la necesidad de validar dicho modelo mediante técnicas complementarias. En este escenario, los dos objetivos de la tesis se han centrado en:

1. Validar la primera estructura atómica a baja resolución (21 Å) de un HAT (4F2hc/LAT2 humano) (capítulo I de la tesis) mediante estudios de crosslinking de cisteínas entre subunidades.

2. En segundo lugar, identificar mejores HAT, en términos de estabilidad, para estudios de cristalización 3D (capítulo II de la tesis).

Capítulo I:

La expresión del heterodímero 4F2hc/LAT2 humano en *Pichia* ha permitido obtener el primer modelo a baja resolución de un HAT (21 Å) por tinción negativa. En este capítulo se cuentan las estrategias seguidas para validar este modelo. Para ello se ha hecho análisis de residuos cercanos entre 4F2hc-ED y LAT2 parte extracelular mediante el uso de crosslinking de cisteínas. A la vez, se inició una colaboración con el grupo de Dr. Fernández-Recio y Dr. Pérez, para calcular la interacción de 4F2hc-ED sobre LAT2. Puesto que no se dispone de la estructura de LAT2 humano en primer lugar modelaron LAT2 humano a partir de la estructura atómica de AdiC (3OB6) y 250 homólogos cercanos. Debido a la gran variabilidad en los bucles extracelulares 3 y 4, se utilizaron 20 modelos de LAT2; se tuvieron en cuenta dos restricciones: el puente disulfuro y la membrana celular para calcular el acoplamiento de 4F2hc-ectodominio sobre los bucles extracelulares de LAT2.

I.1. Crosslinking de cisteínas entre 4F2hc y LAT2 humano:

Para realizar los estudios de crosslinking se cambió de sistema de expresión a la línea celular HEK 293T de mamífero, para poder clonar rápidamente distintos mutantes de 4F2hc y de LAT2hc, tuviese alta eficiencia de transfección y, permitiese testar la funcionalidad de los distintos mutantes de subunidades humanas en células humanas, en lugar de *Pichia pastoris*.

Para poder ingenierizar nuevas cisteínas sobre cada una de las subunidades, primeramente se analizaron las cisteínas endógenas. En el caso de 4F2hc sólo hay dos, la C109 formando el puente disulfuro con las subunidades ligeras, y la C330, más internalizada, que mutamos a serina para evitar cualquier tipo de especificidad en la reacción de crosslinking. En el entorno C330S se generaron 9 mutantes de cisteína localizados sobre la región más probable de estar interactuando con LAT2: S151C, S185C, S191C, S200C, S270C, S393C, S412C, S497C y S506C. En el caso de LAT2, al no tener ninguna estructura disponible, utilizamos el modelo de SteT y el alineamiento de secuencia entre SteT y LAT2 humano para localizar las distintas cisteínas de LAT2 sobre SteT (figura 7).

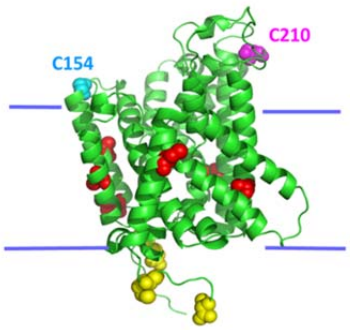


Figura 7. Cisteínas de LAT2 humano localizadas sobre el modelo de SteT: en amarillo se encuentran señaladas las cisteínas intracelulares; en rojo, las intermembrana; en azul la responsable del puente disulfuro y, en magenta, la cisteína C210, que utilizamos para testar crosslinking con los distintos mutantes de 4F2hc.

Como moléculas de crosslinking se eligieron: bismaloimidoetane (BMOE) por ser el crosslinker de brazo más corto (8 Å, según Pierce) y, $BM(PEG)_2$, de la misma familia de crosslinkers pero de brazo más largo (14.7 Å, según Pierce). Ambos se caracterizan por unir grupos sulfhidrilo y ser resistentes a la ruptura del enlace formado por agentes reductores como el DTT, con el fin de romper el puente disulfuro y poder seleccionar los pares de cisteínas crosslinkados como heterodímeros en un western blot.

El procedimiento seguido para crosslinkar fue: 1) sobreexpresión de los distintos heterodímeros en HEK293T durante 24 horas, 2) lisado celular y solubilización de las membranas con DDM 1% (condición igual que la usada para la obtención del modelo 3D), 3) purificación de His-4F2hc utilizando resina de Ni++, quedando una población de proteína: monómero de 4F2hc, homodímeros de 4F2hc-4F2hc, heterodímeros 4F2hc/LAT2 unidos por puente disulfuro, 4) reacción de crosslinking sobre el material purificado: 1 hora a 4ºC utilizando 200 µM BMOE o $BM(PEG)_2$ en 1%DDM en PBS a pH 6,5 y

suplementado con inhibidores de proteasas), 5) se paró la reacción con adición de DTT 100 mM y se añadió LSB 4x, calentó a 65 ºC durante 10 minutos 6) se corrió un gel al 8% de poliacrilamida SDS-PAGE, se transfirió a una membrana de nitrocelulosa y reveló con anticuerpo anti-StrepTag-HRP para visualizar LAT2. Los heterodímeros visualizados correspondían con la fracción crosslinkada.

Los distintos mutantes generados para hacer los experimentos de crosslinking fueron testados previo experimento de crosslinking mediante:

a) **Test de expresión:** por western blot de material lisado (anti-His(4F2hc) y anti-StrepIItag(LAT2)) y solubilizado con DDM1% (paso 1, 2 y 6 anterirores). Los mutantes S191C, S200C, S270C, S497C y S506C fueron descartados por no expresar suficiente heterodímero o presentar proteólisis.

b) **Test de funcionalidad** (transporte de L-Ala[H^3]): se testó el transporte y se comparó con el obtenido en los controles: 4F2hc C330S/LAT2 WT (WT: salvaje). Sólo se testó de aquellos mutantes que habían formado heterodímero: S151C, S185C, S393C, S412C. Todos ellos eran mutantes funcionales.

Resultando positivos los crosslinkings de las cisteínas 412 y 210 en 4F2hc con la cisteína endógena 210 en LAT2, representando la primera evidencia sobre el modelo 3D (Figura 8 A y C). En este momento justo acaban de calcular también el docking Dr.Fdz-Recio y Dr. Pérez, y se compararon las soluciones de crosslinking obtenidas en su modelo, coincidiendo las distancias en el modelo de mejor energía que tenían (docking representado en Figura 8B con las cisteínas señaladas). Así pues, tanto el docking como el modelo quedaron validados experimentalmente por experimentos de crosslinking (Rosell *et al.,* 2014).

Dado que el docking y los primeros resultados de crosslinking eran coincidentes, se utilizó el docking para dirigir el diseño de nuevos mutantes de cisteína sobre cada una de las subunidades para encontrar nuevos pares de residuos cercanos entre subunidades. En el caso de LAT2, todos los nuevos mutantes se generaron sobre LAT2 (C210S).

Finalmente, dirigiendo el diseño de nuevos mutantes sobre el docking calculado (Figura 8D, todas las posiciones de cisteína sobre LAT2 en blanco coloreado el carbono, o cysteínas de 4F2hc, en verde sobrevolando la estructura, se encuentran señaladas, se encontraron nuevas posiciones positivas de crosslinking señaladas en la estructura de la Figura 8D mediante líneas de color rojo, para aquellas posiciones que crosslinkaron fuertemente (80-95%), amarillo aquellas que croslinkaron alrededor de un 60% y en líneas punteadas aquellas que crosslinkaron 10-25% (Figura 8).

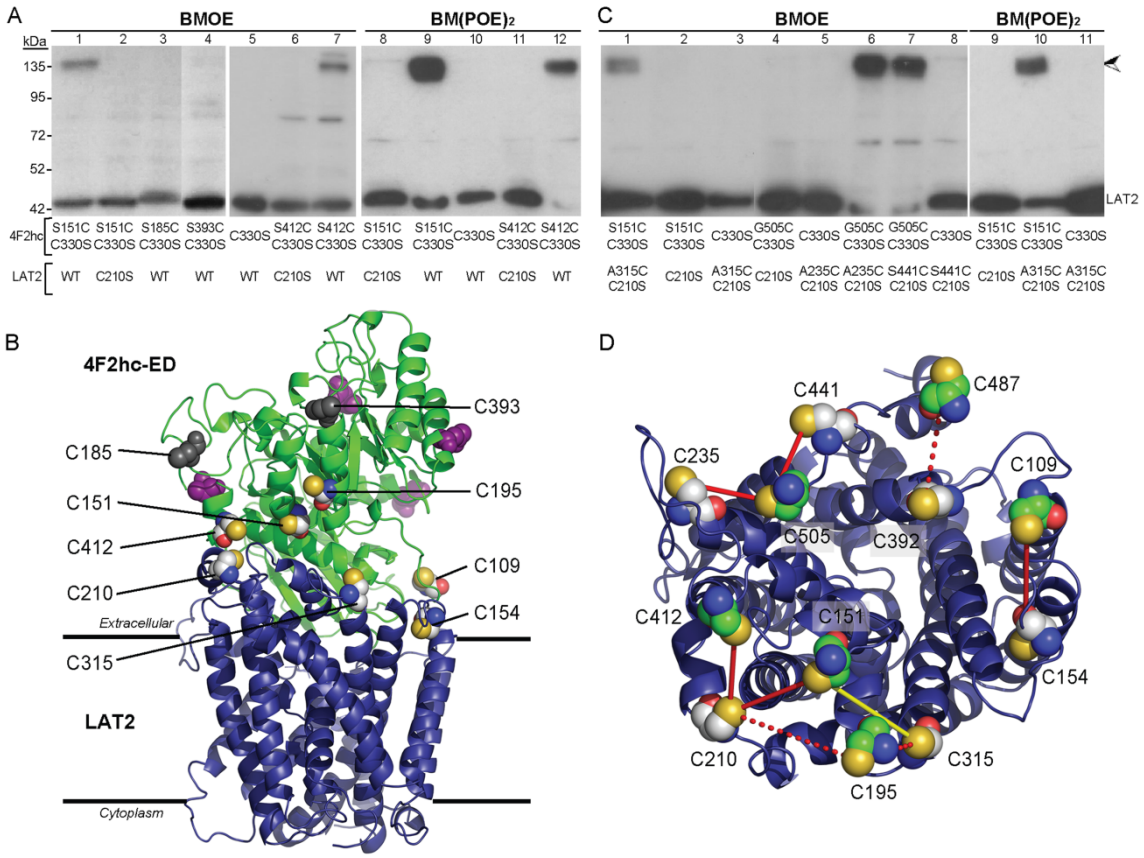


Figura 8. Crosslinking de 4F2hc/LAT2 y modelo de docking del ectodominio de 4F2hc y LAT2. A y C) Crosslinking entre subunidades. Versiones de heterodímeros His-4F2hc/StrepLAT2 expresados en células HEK 293T fueron purificadas por cromatografía de afinidad de histidinas y tratadas con los crosslinkers BMOE y $BM(PEG)_2$ resistentes a DTT. Crosslinking entre subunidades fue detectado como heterodímeros resistentes a DTT (flecha) por western blot utilizando anticuerpos anti-Strep bajo condiciones reductoras. La banda de monómero de LAT2 se corresponde con los heterodímeros formados con 4F2hc (co-purificados, pero no crosslinkados). Mutantes puntuales fueron introducidos en 4F2hc (C330S)/LAT2 salvaje (A) o en LAT2 (C210S) (C). No se observó crosslinking cuando se usó 4F2hc (C330S) o LAT2 (C210S), indicando la especificidad del crosslinking. B) El modelo de docking de menor energía con los residuos de cisteína marcados (endógenos o añadidos mediante mutación) se encuentran señalados por modelos de esferas (átomos C en gris, átomos de O, en rojo, átomos N, en azul; átomos de S, en amarillo). Los lugares putativos de N-glicosilación (N264, N280, N323, y N405 se encuentran indicados en lila) y los residuos no crosslinkados (C185 y C393 en gris) de 4F2hc se encuentran localizados en la parte más externa. D) Resumen de los residuos crosslinkados. El modelo 4F2hc-ED-LAT2 en B fue rotado 90° (mirado desde 4F2hc-ED). Sólo los residuos de cisteína (endógenos o añadidos por mutagénesis) de 4F2hc-ED se muestran como modelos de esfera con átomos C en verde. En LAT2 (en azul) los residuos de cisteína (endógenos o añadidos por mutagénesis) se muestran como modelos de esfera con átomos C en gris. Los residuos crosslinkados se encuentran conectados por una línea indicando el porcentaje de crosslinking (línea roja: 80-95%; línea amarilla: 60%; línea punteada: 5-10%). De manera similar, el puente disulfuro entre C109 (4F2hc) y C154 (LAT2) se representa por una línea roja por presentarse en >95% del LAT2 expresado.

I.2. Implicación de 4F2hc-ED en la estabilidad de la subunidad ligera:

Una vez se supo que 4F2hc/LAT2 podía ser purificado sin agregar (Figura 4A), mientras que LAT2 expresado solo agregaba todo sin la presencia de 4F2hc, se pensó en probar de añadir 4F2hc-ED a las membranas de *P.pastoris* sobreexpresando LAT2 para determinar si el ectodominio (sin el puente disulfuro, ni segmento transmembrana e intracelular) ejercía algún efecto sobre LAT2. Así pues, tras incubar toda la noche 4F2hcED purificado con membranas de *P.pastories* sobreexpresando LAT2, y al día siguiente se solubilizaba con DDM 1% o concentraciones decrecientes para aumentar el efecto, y se demostró que sólo el ectodominio es necesario para la interacción con LAT2. Sorprendentemente, se encontró que el ectodominio aumentaba la fracción soluble de LAT2hc además de su estabilidad cuando analizado con el tiempo, ver figura siguiente:

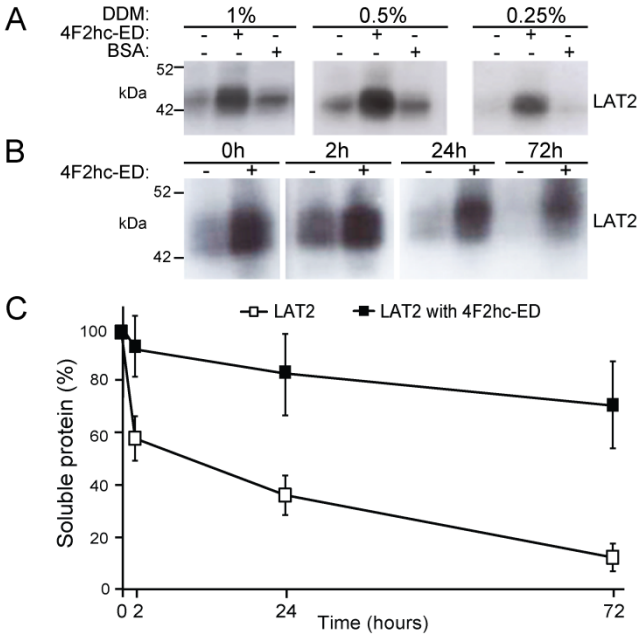


Figura 9. 4F2hc-ED aumenta la solubilidad y estabilidad de LAT2 humano.

Discusión I

La expresión del heterodímero 4F2hc/LAT2 humano en *Pichia* ha permitido obtener el primer modelo a baja resolución de un HAT (21 Å) por tinción negativa, análisis por docking y confirmación por entrecruzamiento de residuos de cisteína. Este modelo ha revelado por primera vez la posición relativa de 4F2hc-ED sobre la subunidad ligera, ofreciendo una explicación estructural de la estabilización del complejo. Asimismo, se ha comprobado que el papel estabilizador de 4F2hc reside en el ectodominio, el cual es capaz de estabilizar él solo a LAT2. Además, se calcularon los residuos de LAT2 con mayor probabilidad de interacción con 4F2hc-ED, observando dos patches distintos, uno hidrofóbico, el cual se está estudiando actualmente, y un patch mixto.

Conclusiones Objetivo I:

1. Tres líneas experimentales independientes evidencian que 4F2hc-ED se encuentra sobre la parte extracelular del transportador. Los resultados de los pares de residuos entre subunidades crosslinkados casi completamente (>80%) triangulan con el puente disulfuro endógeno la posición de 4F2hc-ED sobre LAT2, demostrando el modelo 3D de baja resolución, y, confirmando las predicciones del docking calculado de 4F2hc-ED sobre el modelo de LAT2 humano.

2. La interacción entre 4F2hc y la parte extracelular de LAT2 aumenta la estabilidad de la subunidad ligera solubilizada en detergente. Específicamente, el ectodominio de 4F2hc es suficiente para aumentar la estabilidad de LAT2 solubilizado en DDM. El carácter hidrofóbico de la interacción de 4F2hc-ED y LAT2 puede ser la base de la estabilización de esta subunidad ligera por 4F2hc.

Capítulo II:

El segundo objetivo se centró en la identificación del mejor candidato HAT para estudios de cristalización 3D. Debido a que 4F2hc/LAT2 era poco estable para estudios cristalografía, se inició una proyecto para en encontrar una nueva subunidad eucariota con mejores características. Para ello se seleccionaron 24 subunidades pequeñas de distintas especies, y se adaptó una estrategia desarrollada para seleccionar proteínas candidatas para cristalización (Drew *et al.,* 2008). Se expresaron todas con GFP en C-terminal tras insertarse en el plásmido pDDGFP-2 por recombinación homóloga. La introducción de GFP fusionada a las subunidades ligeras aporta muchas ventajas para monitorizar los niveles de expresión, integridad de la proteína, solubilidad y estabilidad de la proteína (Figura 10).

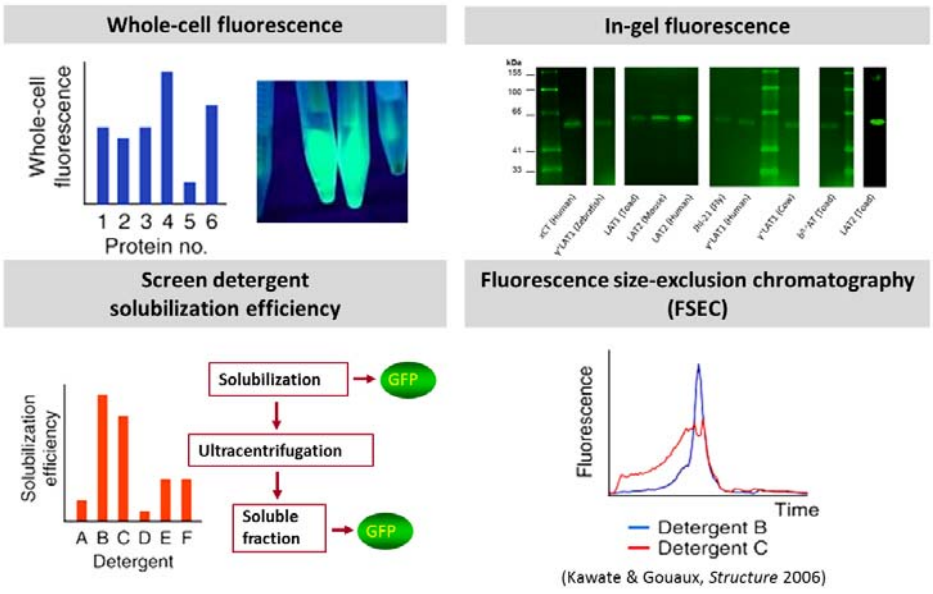


Figura 10. La selección de las subunidades ligeras se agilizó gracias a las ventajas de GFP unido a C-terminal.

	Asc-1	LAT1	LAT2	y^+LAT1	y^+LAT2	xCT	Mnd	Jhl21	$b^{0,+}$AT
Humano	X	X	X	X	X	X			X
Ratón	X	X	X	X	X	X			
Rata									X
Vaca				X					X
Sapo		X	X	X	X				X
Pez zebra				X					
Mosca							X	X	

Tabla 1. 24 subunidades ligeras de diferentes especies (desde mosca a humano) seleccionadas para ser clonadas en *S. cerevisiae* y testar su expresión.

Para estas 24 proteínas de fusión, se analizó la fluorescencia de GFP, se testó el nivel de expresión, integridad de la proteína y solubilidad con una batería de 8 detergentes (OG, NM, DM, DDM, CYMAL-6, LDAO, LAPAO y Foscolina-9) en presencia o no de colesterol (cholesteryl hemisuccinate tris salt: CHS). Por otro lado, se analizó la estabilidad de la proteína solubilizada pero no purificada mediante el "ultracentrifugation dispersity sedimentation assay" (Gutmann *et al.,* 2007). De esta manera se redujo el número de candidatos a 7 subunidades ligeras y 94 condiciones de detergente. Posteriormente, se examinó la estabilidad de las proteínas por el criterio de homogeneidad de la muestra en cromatografía de exclusión molecular, más sensible que el método anterior y se identificaron las mejores condiciones de estabilidad de los 7 candidatos, reduciendo el número de candidatos a 2 para posteriores estudios estructurales (Figura 12). El efecto del colesterol sobre la estabilidad de las subunidades ligeras fue clave. Observado en los perfiles de elución de las FSEC: cromatografías de exclusión molecular fluorescentes de la figura 11.

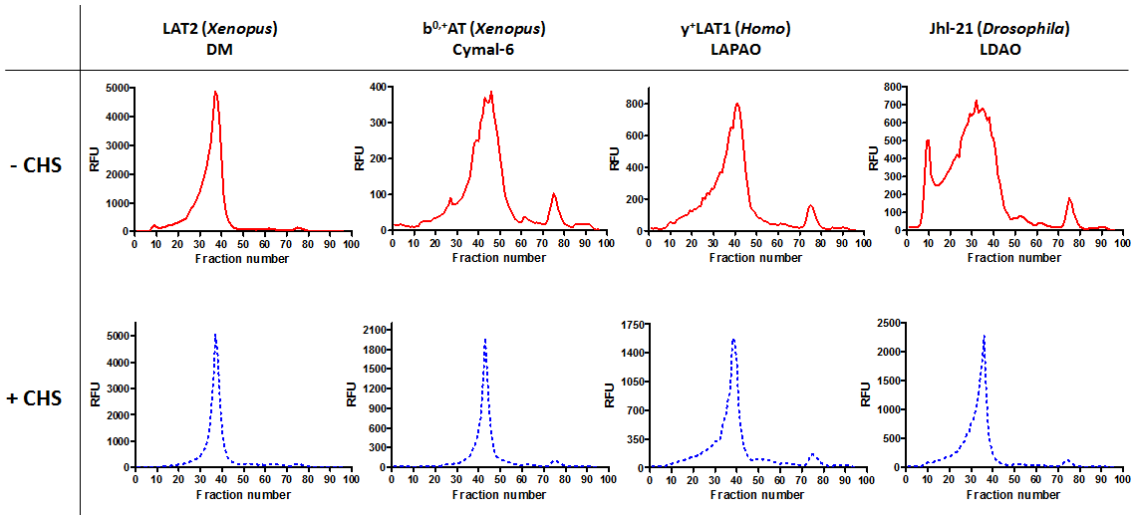


Figura 11. Efecto estabilizador del colesterol sobre todas las subunidades ligeras testadas. Mediante FSEC la monodispersión de distintas proteínas solubilizadas con distintos 1% (p/v) detergentes está representada (línea roja), en la primera línea; y solubilizadas con 1% (p/v) detergente y 0,2% (p/v) colesterol (línea azul punteada), segunda línea.

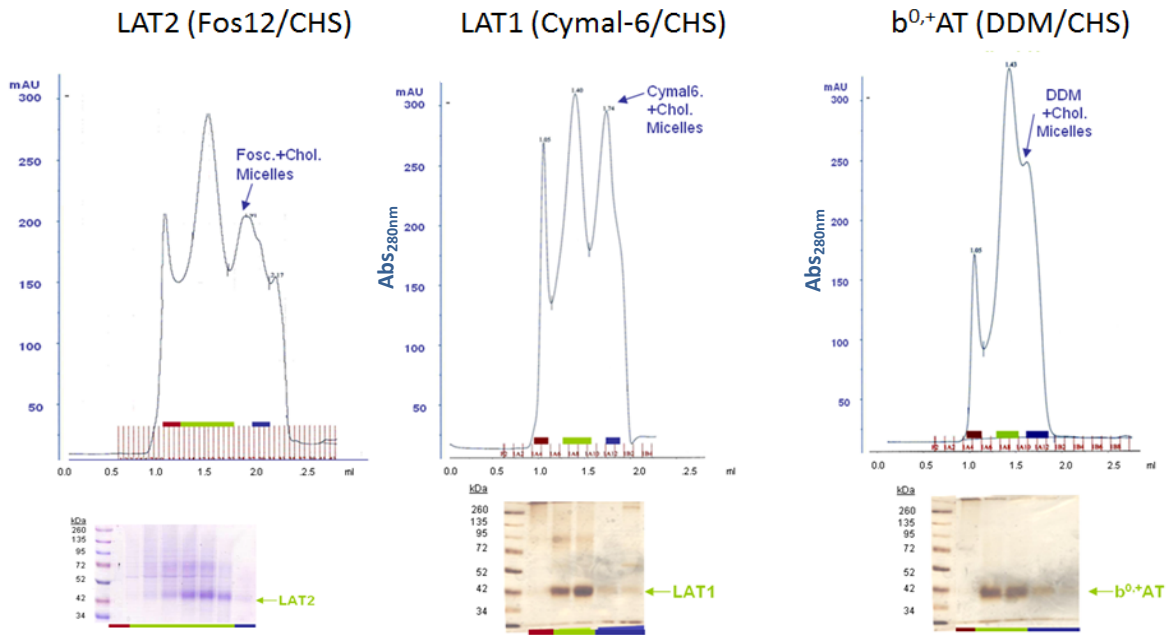


Figura 12. Perfiles de elución de las subunidades ligeras seleccionadas en el *screening* anterior, purificadas y concentradas a 4mg/mL. Inicialmente se seleccionaron LAT2 (*Xenopus laevis*), LAT1 (*Xenopus laevis*) y $b^{0,+}$AT (*Xenopus tropicalis*) para estudios de cristalización 3D después del *screening* descrito anteriormente.

$b^{0,+}$AT (*Xenopus tropicalis*) resultó ser la proteína con mejor perfil de elución después de ser purificada y concentrada a 4mg/mL, seguida de LAT1 (*Xenopus laevis*). LAT2 (*Xenopus laevis*) fue descartada por recomendación de Dr. Gouaux, con quien colaborábamos, debido a que la mayoría de proteínas purificadas en foscolina-12 no eran funcionales. A pesar de que LAT1 y $b^{0,+}$AT son ambas buenas candidatas para estudios de cristalización 3D, en esta tesis sólo se centró en el estudio de $b^{0,+}$AT, la cual se trabajó para ser optimizadas y aumentar las probabilidades de cristalización.

Mejora de la estabilidad: Las estrategias seguidas para la mejora de la estabilidad de $b^{0,+}$AT fueron:

1) Eliminar la cisteína reactiva posible causante de agregación, para ello la mutamos la cisteína 150, responsable del puente disulfuro, a alanina (C150A).

2) Adición de lípidos durante todo el proceso de purificación. El mejor candidato fue DPPC que incrementó enormemente la estabilidad de la proteína.

3) Aumentar la expresión de la proteína cambiando a otro sistema de expresión: Sf9 células de insecto.

4) Disminuir la flexibilidad de la proteína para aumentar la probabilidad de cristalización. Distintas truncaciones de la proteína en N y C-terminal fueron testadas, siendo la deleción de 28 residuos en C terminal la que tuvo un mejor perfil de elución (Figura13).

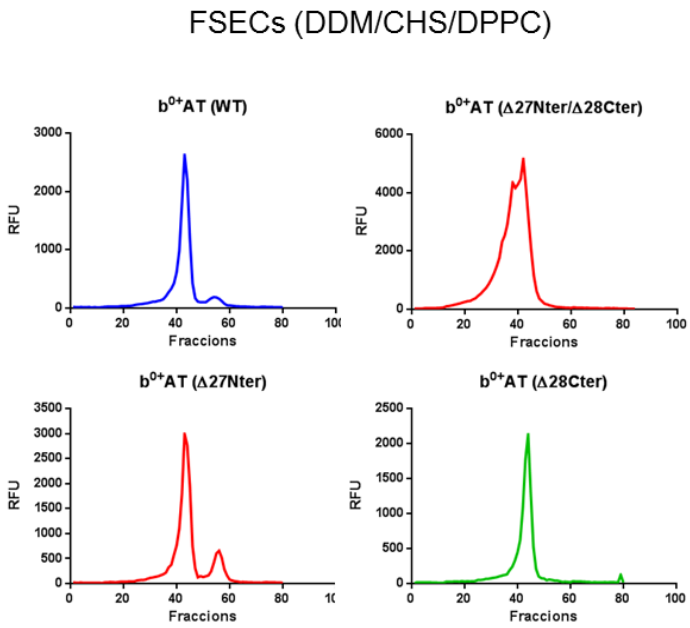


Figura 13. Perfiles FSEC de las formas truncadas de $b^{0,+}$AT C150A comparada con el WT (salvaje). La deleción 28 aminoácidos en C-terminal fue seleccionada para estudios de cristalización 3D.

La primera prueba de cristalización no resultó en ningún cristal de proteína que difractase, pero sólo se testó una placa de condiciones. Así pues, se trata de un buen candidato para iniciar estudios de cristalización.

Discusión II

Se pudo seleccionar $b^{0,+}$AT, LAT1 de *Xenopus* como mejores candidatos para estudios de estructura 3D. Se comprobó que el colesterol estabilizaba enormemente los transportadores eucariotas, como las estructuras resueltas recientemente: Transportador de Dopamina (Drosophila) (DAT) que contiene una molécula de colesterol en su estructura. Ello concuerda con el hecho de encontrar 4F2hc en los lípids rafts, zonas de la membrana ricas en colesterol. Por ello, se especula cual es el papel del colesterol sobre el transportador.

Además, la adición de lípidos DPPC ha aumentado enormemente la estabilidad de la proteína. La eliminación de la cisteína reactiva C150 junto con la deleción de las zonas móviles de la proteína hacen de $b^{0,+}$AT un candidato ideal para estudios de cristalización 3D. Los estudios preliminares no resultaron en ningún cristal de proteína pero sólo se testó una condición.

Conclusiones Objetivo II

1. $b^{0,+}$AT (*Xenopus tropicalis*) seguida de LAT1 (*Xenopus laevis*) fueron seleccionadas como las mejores candidatas para estudios de cristalización 3D tras testar la expresión, solubilidad y estabilidad de 24 subunidades ligeras de metazoos utilizando la tecnología de GFP y el sistema de expresión *S. cerevisiae*. $b^{0,+}$AT y LAT1 fueron solubilizadas y purificadas en DDM/Cholesteryl hemisuccinate Tris salt (CHS) y en cymal-6/CHS, respectivamente.

2. La estabilidad del mejor candidato, $b^{0,+}$AT, se mejoró mediante la adición de CHS y del lípido 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC) durante la solubilización y purificación de la proteína.

3. El cambio de sistema de expresión a células de insecto Sf9 incrementó los niveles de expresión a 5 veces y mejoró la cualidad de la proteína sobreexpresada cuando fue analizada por cromatografía de exclusión molecular.

4. Además, se seleccionó la forma truncada en N y C-terminal más estable con el fin de reducir la flexibilidad de la proteína. Además, se mutó la cisteína reactiva (C150A) para disminuir la agregación de la proteína, resultando finalmente $b^{0,+}$AT (C150A) (Δ28 C-terminal) (*X. tropicalis*) la mejor candidata para estudios de cristalización 3D.

Bibliography

1. Bibliography: Introduction

Adibi, S.A., & Gray, S.J. (1967). Intestinal absorption of essential amino acids in man. *Gastroenterology* **52**, 837-845.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). Cell junctions, cell adhesion, and the extracellular matrix. In *Molecular Biology of the Cell*. 4th edition. New York: Garland Science. Chap 19.

Arinaminpathy, Y., Khurana, E., Engelman, D. M., & Gerstein, M. B. (2009). Computational analysis of membrane proteins: the largest class of drug targets. *Drug Discov. Today* **14**, 1130-1135.

Armstrong, C. (2003). Voltage-gated K channels. *Sci. STKE.* **2003**, re10.

Baker, D., McFarland, K., Lake, R., Shen, H., Tang, X.C., Toda, S. & Kalivas, P.W. (2003). Neuroadaptations in cystine-glutamate exchange underlie cocaine relapse. *Nat. Neurosci.* **6**, 743–749.

Bannai, S. (1984). Transport of cystine and cysteine in mammalian cells. *Biochim. Biophys. Acta* **779**, 289-306.

Bannai, S., & Tateishi, N. (1986). Role of membrane transport in metabolism and function of glutathione in mammals. *J. Membrane Biol.* **89**, 1-8.

Barilli, A., Rotoli, B.M., Visigalli, R., Bussolati, O., Gazzola, G.C., Kadija, Z., & Dall'Asta, V. (2010). In lysinuric protein Intolerance system y^+L activity is defective in monocytes and in GM-CSF-differentiated macrophages. *Orphanet J Rare Dis*, **5**, 1-11.

Bartoccioni, P., Rius, M., Zorzano, A., Palacín, M. & Chillarón, J. (2008). Distinct classes of trafficking rBAT mutants cause the type I cystinuria phenotype. *Hum. Mol. Genet.* **17**, 1845–1854.

Bassi, M.T., Gasol, E., Manzoni, M., Pineda, M., Riboni, M., Martín, R., Zorzano, A., Borsani, G. & Palacín, M. (2001). Identification and characterization of human xCT that co-expresses, with 4F2 heavy chain, the amino acid transport activity system xc-. *Pflugers Arch.* **442**, 286–296.

Bauch, C. & Verrey, F. (2002). Apical heterodimeric cystine and cationic amino acid transporter expressed in MDCK cells. *Am. J. Physiol. Renal Physiol.* **283**, F181–F189.

Bauch, C., Forster, N., Loffing-Cueni, D., Summa, V. & Verrey, F. (2003). Functional cooperation of epithelial heteromeric amino acid transporters expressed in madin-darby canine kidney cells. *J. Biol. Chem.* **278**, 1316–1322.

Beitz, E. (2011). T(E)Xtopo: shaded membrane protein topology plots in LAT(E)X2epsilon. *Bioinformatics* **16**, 1050–1051.

Bertran, J., Magagnin, S., Werner, A., Markovich, D., Biber, J., Testar, X., Zorzano, A., Kühn, L.C., Palacin, M. & Murer, H. (1992a). Stimulation of system y(+)-like amino acid transport by the heavy chain of human 4F2 surface antigen in *Xenopus laevis* oocytes. *Proc. Natl. Acad. Sci.* **89**, 5606–5610.

Bertran, J., Werner, A., Moore, M.L., Stange, G., Markovich, D., Biber, J., Testar, X., Zorzano, A., Palacin, M. & Murer, H. (1992). Expression cloning of a cDNA from rabbit kidney cortex that induces a single transport system for cystine and dibasic and neutral amino acids. *Proc. Natl. Acad. Sci.* **89**, 5601–5605.

Bhatnagar, R. S., & Gordon, J. I. (1997). Understanding covalent modifications of proteins by lipids: where cell biology and biophysics mingle. *Trends Cell Biol.* **7**, 14-20.

Bill, R.M., Henderson, P.J.F., Iwata, S., Kunji, E.R.S., Michel, H., Neutze, R., Newstead, S., Poolman, B., Tate, C.G. & Vogel, H. (2011). Overcoming barriers to membrane protein structure determination. *Nat Biotech.* **29**, 335–340.

Bodoy, S., Martín, L., Zorzano, A., Palacín, M., Estévez, R. & Bertran, J. (2005). Identification of LAT4, a novel amino acid transporter with system L activity. *J. Biol. Chem.* **280**, 12002–12011.

Borsani, G., Bassi, M.T., Sperandeo, M.P., De Grandi, A., Buoninconti, A., Riboni, M., Manzoni, M., Incerti, B., Pepe, A., Andria, G., Ballabio, A. & Sebastio, G. (1999). SLC7A7, encoding a putative permease-related protein, is mutated in patients with lysinuric protein intolerance. *Nat. Genet.* **21**, 297–301.

Braun, D., Wirth, E.K., Wohlgemuth, F., Reix, N., Klein, M.O., Grüters, A., Köhrle, J. & Schweizer, U. (2011). Aminoaciduria, but normal thyroid hormone levels and signaling, in mice lacking the amino acid and thyroid hormone transporter Slc7a8. *Biochem. J.* **439**, 249–255.

Bridges, R.J., Natale, N.R. & Patel, S. a. (2012). System xc− cystine/glutamate antiporter: an update on molecular pharmacology and roles within the CNS. *Br. J. Pharmacol.* **165**, 20–34.

Bröer, S. & Palacín, M. (2011). The role of amino acid transporters in inherited and acquired diseases. *Biochem. J.* ***436***, 193–211.

Bröer, S. (2008). Amino acid transport across mammalian intestinal and renal epithelia. *Physiol. Rev.* **88**, 249-286.

Busch, A.E., Herzer, T., Waldegger, S., Schmidt, F., Palacin, M., Biber, J., Markovich, D., Murer, H. & Lang, F. (1994). Opposite directed currents induced by the transport of dibasic and neutral amino acids in Xenopus oocytes expressing the protein rBAT. *J. Biol. Chem.* **269**, 25581–25586.

Calonge, M.J., Gasparini, P., Chillarón, J., Chillón, M., Gallucci, M., Rousaud, F., Zelante, L., Testar, X., Dallapiccola, B. & Di Silverio, F. (1994). Cystinuria caused by mutations in rBAT, a gene involved in the transport of cystine. *Nat. Genet.* **6**, 420–425.

Cantor, J., Browne, C.D., Ruppert, R., Féral, C.C., Fässler, R., Rickert, R.C., Ginsberg, M.H. (2009). CD98hc facilitates B cell proliferation and adaptive humoral immunity. *Nat. Immunol.* **10**, 412–419.

Casals, F., Ferrer-Admetlla, A., Chillarón, J., Torrents, D., Palacín, M. & Bertranpetit, J. (2008). Is there selection for the pace of successive inactivation of the arpAT gene in primates? *J. Mol. Evol.* **67**, 23–28.

Chairoungdua, A., Kanai, Y., Matsuo, H., Inatomi, J., Kim, D.K. & Endou, H. (2001). Identification and characterization of a novel member of the heterodimeric amino acid transporter family presumed to be associated with an unknown heavy chain. *J. Biol. Chem.* **276**, 49390–49399.

Charania, M. A., Laroui, H., Liu, H., Viennois, E., Ayyadurai, S., Xiao, B., Ingersoll, S.A., Kalman, D., Merlin, D. (2013). Intestinal epithelial CD98 directly modulates the innate host response to enteric bacterial pathogens. *Infect Immun*. **81**, 923-934.

Chillarón, J., Roca, R., Valencia, A., Zorzano, A. & Palacín, M. (2001). Heteromeric amino acid transporters: biochemistry, genetics, and physiology. *Am. J. Physiol. Renal Physiol.* **281**, F995–F1018.

Chillarón, J., Font-Llitjós M., Fort J., Zorzano A., Goldfarb D.S., Nunes V., and Palacín, M. (2010). Pathophysiology and treatment of cystinuria. *Nat. Rev. Nephrol.* **6,** 424-434.

Chillarón, J., Estévez R., Mora C., Wagner C.A., Suessbrich H., Lang F., Gelpí J.L., Testar X., Busch AE, Zorzano A, Palacín M. (1996). Obligatory amino acid exchange via systems $b^{o,+}$-like and y^+L-like. A tertiary active transport mechanism for renal reabsorption of cystine and dibasic amino acids. *J. Biol. Chem.* **271**, 17761-17770.

Fernández, E., Carrascal M., Rousaud F., Abián J., Zorzano A., Palacín M., and Chillarón J.. (2002). rBAT-b(0,+)AT heterodimer is the main apical reabsorption system for cystine in the kidney. *Am. J. Physiol. Renal Physiol.* **283**, F540-F548.

Fernandez, E., Torrents D., Chillaron J., Martin del Rio R., Zorzano A., and Palacin M.. (2003). Basolateral LAT-2 has a major role in the transepithelial fux of L-cystine in the renal proximal tubule cell line OK. *J. Am. Soc. Nephrol.* **14**, 837-847.

Christensen, H. N. (1990). Role of amino acid transport and countertransport in nutrition and metabolism. *Physiol Rev*. **64**, 3151-3156.

Conrad, M., & Sato, H. (2012). The oxidative stress-inducible cystine/glutamate antiporter, system X_c^-: cystine supplier and beyond. *Amino Acids* **42**, 231-246.

Conte, M. R., & Matthews, S. (1998). Retroviral matrix proteins: a structural perspective. *Virology* **246**, 191-198.

Dall'Asta, V., Bussolati, O., Sala, R., Rotoli, B. M., Sebastio, G., Sperandeo, M. P., Andria G. & Gazzola, G. C. (2000). Arginine transport through system y^+L in cultured human fibroblasts: normal phenotype of cells from LPI subjects. *Am. J. Physiol. Cell Physiol*. **279**, C1829-C1837.

De Bundel, D., Schallier, A., Loyens, E., Fernando, R., Miyashita, H., Van Liefferinge, J., ... & Massie, A. (2011). Loss of system xc− does not induce oxidative stress but decreases extracellular glutamate in hippocampus and influences spatial working memory and limbic seizure susceptibility. *J. Neurosci.* **31**, 5792-5803.

Decoursey, T. E. (2003). Voltage-gated proton channels and other proton transfer pathways. *Physiol. Rev.* **83**, 475-579.

Del Amo, E.M., Urtti, A. & Yliperttula, M. (2008). Pharmacokinetic role of L-type amino acid transporters LAT1 and LAT2. *Eur. J. Pharm. Sci.* **35**, 161–174.

Dello Strologo, L., Pras, E., Pontesilli, C., Beccia, E., Ricci-Barbini, V., De Sanctis, L., Ponzone, A., Gallucci, M., Bisceglia, L., Zelante, L., et al. (2002). Comparison between SLC3A1 and SLC7A9 cystinuria patients and carriers: a need for a new classification. *J. Am. Soc. Nephrol.* **13**, 2547–2553.

Denoyer D, Kirby L, Waldeck K, Roselt P, Neels OC, Bourdier T, Shepherd R, Katsifis A, Hicks RJ. (2012) Preclinical characterization of 18F-D-FPHCys, a new amino acid-based PET tracer. *Eur. J. Nucl. Med. Mol. Imaging* **39**, 703-712.

Deves, R., & Boyd, C. A. R. (1998). Transporters for cationic amino acids in animal cells: discovery, structure, and function. *Physiol. Rev.* **78**, 487-545.

Deves, R., & Boyd, C. A. R. (2000). Surface antigen CD98 (4F2): not a single membrane protein, but a family of proteins with multiple functions. *J. Membrane Biol.* **173**, 165-177.

Domínguez, F., Simón, C., Quiñonero, A., Ramírez, M.Á., González-Muñoz, E., Burghardt, H., Cervero, A., Martínez, S., Pellicer, A., Palacín, M., Sánchez-Madrid, F. & Yáñez-Mó, M. (2010). Human endometrial CD98 is essential for blastocyst adhesion. *PLoS One* **5**, e13380.

Donnenberg, M. S. (2000). Pathogenic strategies of enteric bacteria. *Nature* **406**, 768-774.

Estévez, R., Camps, M., Rojas, A.M., Testar, X., Devés, R., Hediger, M.A., Zorzano, A., Palacín, M. (1998). The amino acid transport system y^+L/4F2hc is a heteromultimeric complex. *FASEB J.* **12**, 1319–1329.

Faham, S., Watanabe, A., Besserer, G. M., Cascio, D., Specht, A., Hirayama, B.A., Wright, E.M., Abramson, J. (2008). The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na+/sugar symport. *Science* **321**, 810-814.

Fan X., Ross D.D., Arakawa H., Ganapathy V., Tamai I., Nakanishi T. (2010). Impact of system L amino acid transporter 1 (LAT1) on proliferation of human ovarian cancer cells: a possible target for combination therapy with anti-proliferative aminopeptidase inhibitors. *Biochem Pharmacol*. **80**, 811-818.

Fang Y, Jayaram H, Shane T, Kolmakova-Partensky L, Wu F, Williams C, Xiong Y, Miller C. (2009). Structure of a prokaryotic virtual proton pump at 3.2 A resolution. *Nature* **460**, 1040-1043.

Feliubadaló, L., Font, M., Purroy, J., Rousaud, F., Estivill, X., Nunes, V., Golomb, E., Centola, M., Aksentijevich, I., Kreiss, Y., Goldman, B., Pras, M., Kastner, D.L., Pras, E., Gasparini, P., Bisceglia, L., Beccia, E., Gallucci, M., de Sanctis, L., Ponzone, A., Rizzoni, G.F., Zelante, L., Bassi, M.T., George, A.L., Manzoni, M., De Grandi, A., Riboni, M., Endsley, J.K., Ballabio, A., Borsani, G., Reig, N., Fernández, E., Estévez, R., Pineda, M., Torrents, D., Camps, M., Lloberas, J., Zorzano, A. & Palacín, M. (1999). Non-type I cystinuria caused by mutations in SLC7A9, encoding a subunit ($b^{o,+}$AT) of rBAT. *Nat. Genet*. **23**, 52–57.

Fenczik, C.A., Zent, R., Dellos, M., Calderwood, D.A., Satriano, J., Kelly, C. & Ginsberg, M.H. (2001). Distinct domains of CD98hc regulate integrins and amino acid transport. *J. Biol. Chem*. **276**, 8746–8752.

Feral, C.C., Nishiya, N., Fenczik, C. a., Stuhlmann, H., Slepak, M. & Ginsberg, M.H. (2005). CD98hc (SLC3A2) mediates integrin signaling. *Proc. Natl. Acad. Sci*. **102**, 355–360.

Fernández, E., Jiménez-Vidal, M., Calvo, M., Zorzano, A., Tebar, F., Palacín, M. & Chillarón, J. (2006). The structural and functional units of heteromeric amino acid transporters. The heavy subunit rBAT dictates oligomerization of the heteromeric amino acid transporters. *J. Biol. Chem*. **281**, 26552–26561.

Fernández E, Torrents D, Zorzano A, Palacín M, Chillarón J. (2005) Identification and functional characterization of a novel low affinity aromatic-preferring amino acid transporter (arpAT). One of the few proteins silenced during primate evolution. *J Biol Chem.* **280**, 19364-19372.

Fort, J., de la Ballina, L.R., Burghardt, H.E., Ferrer-costa, C., Turnay, J., Ferrer-orta, C., Usón, I., Zorzano, A., Fernández-Recio, J., Orozco, M., Lizarbe, M.A., Fita, I., Palacín, M., Ballina, L.R. De, Ferna, J., Palacı, M. & Hans, E. (2007). The structure of human 4F2hc ectodomain provides a model for homodimerization and electrostatic interaction with plasma membrane. *J. Biol. Chem.* **282**, 31444–31452.

Foster, J.W. (2004). *Escherichia coli* acid resistance: tales of an amateur acidophile. *Nat. Rev. Microbiol.* **2**, 898–907.

Fotiadis, D., Kanai, Y. & Palacín, M. (2013). The SLC3 and SLC7 families of amino acid transporters. *Mol. Aspects Med.* **34**, 139–158.

Fuchs, B. C., & Bode, B. P. (2005). Amino acid transporters ASCT2 and LAT1 in cancer: partners in crime? In *Semin Cancer Biol*. **15**, 254-266. Academic Press.

Fukasawa, Y., Segawa, H., Kim, J.Y., Chairoungdua, A., Kim, D.K., Matsuo, H., Cha, S.H., Endou, H. & Kanai, Y. (2000). Identification and characterization of a Na(+)-independent neutral amino acid transporter that associates with the 4F2 heavy chain and exhibits substrate selectivity for small neutral D- and L-amino acids. *J. Biol. Chem.* **275**, 9690–9698.

Furriols, M., Chillarón, J., Mora, C., Castelló, A., Bertran, J., Camps, M., Testar, X., Vilaró, S., Zorzano, A. & Palacín, M. (1993). rBAT, related to L-cysteine transport, is localized to the microvilli of proximal straight tubules, and its expression is regulated in kidney by development. *J. Biol. Chem.* **268,** 27060–27068.

Gabriško, M., & Janeček, Š. (2009). Looking for the ancestry of the heavy-chain subunits of heteromeric amino acid transporters rBAT and 4F2hc within the GH13 α-amylase family. *FEBS J.* **276**, 7265-7278.

Ganapathy, V., Thangaraju, M., & Prasad, P. D. (2009). Nutrient transporters in cancer: relevance to Warburg hypothesis and beyond. *Pharmacol. Ther.* **121**, 29-40.

Gao X., Lu F., Zhou L., Dang S., Sun L., Li X., Wang J., Shi Y., (2009). Structure and mechanism of an amino acid antiporter. *Science* **324**, 1565–1568.

Gao X., Zhou L., Jiao X., Lu F., Yan C., Zeng X., Wang J., Shi Y. (2010) Mechanism of substrate recognition and transport by an amino acid antiporter. *Nature* **463,** 828–832.

Gasol, E., Jiménez-Vidal, M., Chillarón, J., Zorzano, A. & Palacín, M. (2004). Membrane topology of system xc- light subunit reveals a re-entrant loop with substrate-restricted accessibility. *J. Biol. Chem*. **279**, 31228–31236.

Guo W, Zhao Y, Zhang Z, Tan N, Zhao F, Ge C, Liang L, Jia D, Chen T, Yao M, Li J, He X. (2011) Disruption of xCT inhibits cell growth via the ROS/autophagy pathway in hepatocellular carcinoma. *Cancer Lett*. **312**, 55-61.

Halestrap, A. P., & Meredith, D. (2004). The SLC16 gene family—from monocarboxylate transporters (MCTs) to aromatic amino acid transporters and beyond. *Pflügers Archiv*, **447**, 619-628.

Hashimoto, A., & Oka, T. (1997). Free D-aspartate and D-serine in the mammalian brain and periphery. *Progr. Neurobiol.* **52**, 325-353.

Hay, N., & Sonenberg, N. (2004). Upstream and downstream of mTOR. *Genes Dev.* **18**, 1926-1945.

Haynes, B.F., Hemler, M.E., Mann, D.L., Eisenbarth, G.S., Shelhamer, J., Mostowski, H.S., Thomas, C.A., Strominger, J.L. & Fauci, A.S. (1981). Characterization of a monoclonal antibody (4F2) that binds to human monocytes and to a subset of activated lymphocytes. *J. Immunol*. **126**, 1409–1414.

Hediger, M. A., Clémençon, B., Burrier, R. E., & Bruford, E. A. (2013). The ABCs of membrane transporters in health and disease (SLC series): introduction. *Mol. Aspects Med.* **34**, 95-107.

Hediger, M. A., Romero, M. F., Peng, J. B., Rolfs, A., Takanaga, H., & Bruford, E. A. (2004). The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteins. *Pflügers Archiv*, **447**, 465-468.

Helboe, L., Egebjerg, J., Møller, M., & Thomsen, C. (2003). Distribution and pharmacology of alanine–serine–cysteine transporter 1 (asc-1) in rodent brain. *Eur. J. Neurosci*. **18**, 2227-2238.

Henderson, N.C. (2004). CD98hc (SLC3A2) Interaction with β 1 integrins is required for transformation. *J. Biol. Chem*. **279**, 54731–54741.

Huang, S., & Ingber, D. E. (2005). Cell tension, matrix mechanics, and cancer development. *Cancer cell*, **8**, 175-176.

Imai, H., Kaira, K., Oriuchi, N., Shimizu, K., Tominaga, H., Yanagitani, N., Sunaga, N., Ishizuka, T., Nagamori, S., Promchan, K., Nakajima, T., Yamamoto, N., Mori, M. & Kanai, Y. (2010). Inhibition of L-type amino acid transporter 1 has antitumor activity in non-small cell lung cancer. *Anticancer Res.* **30**, 4819–4828.

Janecek, S., Svensson, B. & Henrissat, B. (1997). Domain evolution in the alpha-amylase family. *J. Mol. Evol.* **45**, 322–331.

Jeschke, G. (2013). A comparative study of structures and structural transitions of secondary transporters with the LeuT fold. *Eur Biophys J.* **42**,181-197.

Jiang, Y., Lee, A., Chen, J., Ruta, V., Cadene, M., Chait, B. T., & MacKinnon, R. (2003). X-ray structure of a voltage-dependent K+ channel. *Nature* **423**, 33-41.

Kageyama T, Nakamura M, Matsuo A, Yamasaki Y, Takakura Y, Hashida M, Kanai Y, Naito M, Tsuruo T, Minato N, Shimohama S. (2000). The 4F2hc/LAT1 complex transports L-DOPA across the blood-brain barrier. *Brain Res*, **879**, 115-121.

Kaira, K., Ishizuka, T., Yanagitani, N., Sunaga, N., Hisada, T., & Mori, M. (2007). Value of FDG positron emission tomography in monitoring the effects of therapy in progressive pulmonary sarcoidosis. *Clin. Nucl. Med.* **32**, 114-116.

Kaira, K., Ohde, Y., Endo, M., Nakagawa, K., Okumura, T., Takahashi, T., Murakami, H., Tsuya, A., Nakamura, Y., Naito, T., Kondo, H., Nakajima, T. & Yamamoto, N. (2011a). Expression of 4F2hc (CD98) in pulmonary neuroendocrine tumors. *Oncol. Rep*. **26**, 931–937.

Kaira, K., Oriuchi, N., Imai, H., Shimizu, K., Yanagitani, N., Sunaga, N., Hisada, T., Kawashima, O., Kamide, Y., Ishizuka, T., Kanai, Y., Nakajima, T. & Mori, M. (2009a). CD98 expression is associated with poor prognosis in resected non-small-cell lung cancer with lymph node metastases. *Ann Surg Oncol* **16**, 3473–3481.

Kaira, K., Oriuchi, N., Imai, H., Shimizu, K., Yanagitani, N., Sunaga, N., Hisada, T., Tanaka, S., Ishizuka, T., Kanai, Y., Endou, H., Nakajima, T. & Mori, M. (2008). l-type amino acid transporter 1 and CD98 expression in primary and metastatic sites of human neoplasms. *Cancer Sci*. **99**, 2380–2386.

Kaira K, Oriuchi N, Imai H, Shimizu K, Yanagitani N, Sunaga N, Hisada T, Ishizuka T, Kanai Y, Nakajima T, Mori M. (2009b). Prognostic significance of L-type amino acid transporter 1 (LAT1) and 4F2 heavy chain (CD98) expression in stage I pulmonary adenocarcinoma. *Lung Cancer* **6**, 120-126.

Kaira K, Oriuchi N, Imai H, Shimizu K, Yanagitani N, Sunaga N, Hisada T, Kawashima O, Kamide Y, Ishizuka T, Kanai Y, Nakajima T, Mori M. (2009c). CD98 expression is associated with poor prognosis in resected non-small-cell lung cancer with lymph node metastases. *Ann. Surg. Oncol*. **16**, 3473-3481.

Kaira, K., Oriuchi, N., Takahashi, T., Nakagawa, K., Ohde, Y., Okumura, T., Murakami, H., Shukuya, T., Kenmotsu, H., Naito, T., Kanai, Y., Endo, M., Kondo, H., Nakajima, T. & Yamamoto, N. (2011). L-type amino acid transporter 1 (LAT1) expression in malignant pleural mesothelioma. *Anticancer Res*. **31**, 4075–4082.

Kaira, K., Oriuchi, N., Takahashi, T., Nakagawa, K., Ohde, Y., Okumura, T., Murakami, H., Shukuya, T., Kenmotsu, H., Naito, T., Kanai, Y., Endo, M., Kondo, H., Nakajima, T. & Yamamoto, N. (2011c). LAT1 expression is closely associated with hypoxic markers and mTOR in resected non-small cell lung cancer. *Am. J. Transl. Res.* **3**, 468–478.

Kaira K, Sunose Y, Ohshima Y, Ishioka NS, Arakawa K, Ogawa T, Sunaga N, Shimizu K, Tominaga H, Oriuchi N, Itoh H, Nagamori S, Kanai Y, Yamaguchi A, Segawa A, Ide M, Mori M, Oyama T, Takeyoshi I. (2013). Clinical significance of L-type amino acid transporter 1 expression as a prognostic marker and potential of new targeting therapy in biliary tract cancer. *BMC Cancer* **3**:482.

Kaleeba, J. A., & Berger, E. A. (2006). Kaposi's sarcoma-associated herpesvirus fusion-entry receptor: cystine transporter xCT. *Science* **311**, 1921-1924.

Kanai, Y., Segawa, H., Miyamoto, K. i, Uchino, H., Takeda, E. & Endou, H. (1998). Expression cloning and characterization of a transporter for large neutral amino acids activated by the heavy chain of 4F2 antigen (CD98). *J. Biol. Chem*. **273**, 23629–23632.

Kanai, Y., Stelzner MG, Lee WS, Wells RG, Brown D, Hediger MA. (1992). Expression of mRNA (D2) encoding a protein involved in amino acid transport in S3 proximal tubule. *Am J Physiol.* **263**, F1087-F1092.

Kim JY, Kanai Y, Chairoungdua A, Cha SH, Matsuo H, Kim DK, Inatomi J, Sawa H, Ida Y, Endou H. (2001). Human cystine/glutamate transporter: cDNA cloning and upregulation by oxidative stress in glioma cells. *Biochim. Biophys. Acta.* **1512**, 335-344.

Kinne, A., Schülein, R. & Krause, G. (2011). Primary and secondary thyroid hormone transporters. *Thyroid Res.* **4**, 1–10.

Kobayashi K, Ohnishi A, Promsuk J, Shimizu S, Kanai Y, Shiokawa Y, Nagane M. (2008) Enhanced tumor growth elicited by L-type amino acid transporter 1 in human malignant glioma cells. *Neurosurgery* **62**, 493-503.

Kowalczyk, L., Ratera, M., Paladino, A., Bartoccioni, P., Errasti-Murugarren, E., Valencia, E., Portella, G., Bial, S., Zorzano, A., Fita, I., Orozco, M., Carpena, X., Vázquez-Ibar, J.L. & Palacín, M. (2011). Molecular basis of substrate-induced permeation by an amino acid antiporter. *Proc. Natl. Acad. Sci.* **108**, 3935–3940.

Krämer, S.D., Mu, L., Müller, A., Keller, C., Kuznetsova, O.F., Schweinsberg, C., Franck, D., Müller, C., Ross, T.L., Schibli, R. & Ametamey, S.M. (2012). 5-(2-18F-fluoroethoxy)-L-tryptophan as a substrate of system L transport for tumor imaging by PET. *J. Nucl. Med*. **53**, 434–442.

Kroemer, G., & Pouyssegur, J. (2008). Tumor cell metabolism: cancer's Achilles' heel. *Cancer cell* **13**, 472-482.

Krogh, A., Krogh, a, Larsson, B., von Heijne, G. & Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580.

Lemmon, M. A., & Schlessinger, J. (1994). Regulation of signal transduction and signal diversity by receptor oligomerization. *Trends Biochem. Sci.* **19**, 459-463.

Lo, M., Ling, V., Wang, Y.Z. & Gout, P.W. (2008a). The xc- cystine/glutamate antiporter: a mediator of pancreatic cancer growth with a role in drug resistance. *Br. J. Cancer* **99**, 464–72.

Lo, M., Wang, Y.-Z. & Gout, P.W. (2008b). The x c− cystine/glutamate antiporter: A potential target for therapy of cancer and other diseases. *J. Cell. Physiol.* **215**, 593–602.

Lüscher, B., Rousseaux, M., Lees, R., MacDonald, H.R. & Bron, C. (1985). Cell surface glycoproteins involved in the stimulation of interleukin 1-dependent interleukin 2 production by a subline of EL4 thymoma cells. II. Structure, biosynthesis, and maturation. *J. Immunol*. **135**, 3951–3957.

Ma, D., Lu, P., Yan, C., Fan, C., Yin, P., Wang, J. & Shi, Y. (2012). Structure and mechanism of a glutamate–GABA antiporter. *Nature* **483**, 632-636.

Makowske, M. & Christensen, H.N. (1982). Hepatic transport system interconverted by protonation from service for neutral to service for anionic amino acids. *J. Biol. Chem.* **257**, 14635–14638.

Mastroberardino, L., Spindler, B., Pfeiffer, R., Skelly, P.J., Loffing, J., Shoemaker, C.B. & Verrey, F. (1998). Amino-acid transport by heterodimers of 4F2hc/CD98 and members of a permease family. *Nature* **395**, 288–291.

Matsuo, H., Kanai, Y., Kim, J.Y., Chairoungdua, A., Kim, D.K., Inatomi, J., Shigeta, Y., Ishimine, H., Chaekuntode, S., Tachampa, K., Choi, H.W., Babu, E., Fukuda, J. & Endou, H. (2002). Identification of a novel Na+-independent acidic amino acid transporter with structural similarity to the member of a heterodimeric amino acid transporter family associated with unknown heavy chains. *J. Biol. Chem*. **277**, 21017–21026.

McLaughlin, S., & Aderem, A. (1995). The myristoyl-electrostatic switch: a modulator of reversible protein-membrane interactions. *Trends Biochem. Sci.* **20**, 272-276.

Meier, C., Ristic, Z., Klauser, S. & Verrey, F. (2002). Activation of system L heterodimeric amino acid exchangers by intracellular substrates. *EMBO J*. **21**, 580–589.

Milliner, D. S., & Murphy, M. E. (1993). Urolithiasis in pediatric patients. *Mayo Clinic Proceedings*. 68, 241-248

Mitchell, P. (1957). A general theory of membrane transport from studies of bacteria. *Nature* **180**, 134–136.

Mulgrew-Nesbitt, A., Diraviyam, K., Wang, J., Singh, S., Murray, P., Li, Z., Rogers, L. & Murray, D. (2006). The role of electrostatics in protein–membrane interactions. *Biochim. Biophys. Acta* **1761**, 812-826.

Mykkänen J1, Torrents D, Pineda M, Camps M, Yoldi ME, Horelli-Kuitunen N, Huoponen K, Heinonen M, Oksanen J, Simell O, Savontaus ML, Zorzano A, Palacín M, Aula P. (2000). Functional analysis of novel mutations in y(+)LAT-1 amino acid transporter gene causing lysinuric protein intolerance (LPI). *Hum. Mol. Genet*. **9**, 431-438.

Nabeyama, A., Kurita, A., Asano, K., Miyake, Y., Yasuda, T., Miura, I., Nishitai, G., Arakawa, S., Shimizu, S., Wakana, S., Yoshida, H. & Tanaka, M. (2010). xCT deficiency accelerates chemically induced tumorigenesis. *Proc. Natl. Acad. Sci.* **107**, 6436-6441.

Nakamura, E., Sato, M., Yang, H., Miyagawa, F., Harasaki, M., Tomita, K., Matsuoka, S., Noma, A., Iwai, K. & Minato, N. (1999). 4F2 (CD98) heavy chain is associated covalently with an amino acid transporter and controls intracellular trafficking and membrane topology of 4F2 heterodimer. *J. Biol. Chem*. **274**, 3009–3016.

Nakauchi, J., Matsuo, H., Kim, D.K., Goto, A., Chairoungdua, A., Cha, S.H., Inatomi, J., Shiokawa, Y., Yamaguchi, K., Saito, I., Endou, H. & Kanai, Y. (2000). Cloning and characterization of a human brain Na+-independent transporter for small neutral amino acids that transports D-serine with high affinity. *Neurosci. Lett*. **287**, 231–235.

Nicklin, P., Bergman, P., Zhang, B., Triantafellow, E., Wang, H., Nyfeler, B., Yang, H., Hild, M., Kung, C., Wilson, C., Myer, V. E., MacKeigan, J.P., Porter, J.A., Wang, Y. K., Cantley, L. C., Finan, P. M. & Murphy, L. O. (2009). Bidirectional transport of amino acids regulates mTOR and autophagy. *Cell* **136**, 521-534.

Oda, K., Hosoda, N., Endo, H., Saito, K., Tsujihara, K., Yamamura, M., Sakata, T., Anzai, N., Wempe, M.F., Kanai, Y. & Endou, H. (2010). L-type amino acid transporter 1 inhibitors inhibit tumor cell growth. *Cancer Sci.* **101**, 173–179.

Ohgimoto, S., Tabata, N., Suga, S., Tsurudome, M., Kawano, M., Nishio, M., Okamoto, K., Komada, H., Watanabe, N. & Ito, Y. (1996). Regulation of human immunodeficiency virus gp160-mediated cell fusion by antibodies against fusion regulatory protein 1. *J. Gen. Virol.* **77**, 2747–2756.

Ohkawa, M., Ohno, Y. Masuko, K., Takeuchi, A., Suda, K., Kubo, A., Kawahara, R., Okazaki, S., Tanaka, T., Saya, H., Seki, M., Enomoto, T., Yagi, H., Hashimoto, Y., & Masuko, T. (2011). Oncogenicity of L-type amino-acid transporter 1 (LAT1) revealed by targeted gene disruption in chicken DT40 cells: LAT1 is a promising molecular target for human cancer therapy. *Biochem. Biophys. Res. Commun*. **406**, 649–655.

Okamoto K., Tsurudome M., Ohgimoto S., Kawano M., Nishio M., Komada H., Ito M., Sakakura Y., Ito Y. (1997). An anti-fusion regulatory protein-1 monoclonal antibody suppresses human parainfluenza virus type 2-induced cell fusion. *J Gen Virol*. **78**, 83-89.

Palacín, M. Estévez, R., Bertran, J., & Zorzano, A. (1998). Molecular biology of mammalian plasma membrane amino acid transporters. *Physiol. Rev*. **78**, 969–1054.

Palacín, M., Fernández, E., Chillarón, J., & Zorzano, A. (2001). The amino acid transport system bo,+ and cystinuria. *Mol. Membr. Biol.* **18**, 21-26.

Palacín M, Kanai Y. (2004). The ancillary proteins of HATs: SLC3 family of amino acid transporters. *Pflugers Arch*. **447**, 490-494.

Palacín, M., Nunes, V., Font-llitjós, M., Jiménez-vidal, M., Fort, J., Gasol, E., Pineda, M., Feliubadaló, L., Chillarón, J. & Zorzano, A. (2005). The genetics of heteromeric amino acid transporters*. Physiology* **20**, 112–124.

Parmacek, M.S., Karpinski, B.A., Gottesdiener, K.M., Thompson, C.B. & Leiden, J.M. (1989). Structure, expression and regulation of the murine 4F2 heavy chain. *Nucleic Acids Res.* **17**, 1915–1931.

Patlak, C. S. (1957). Contributions to the theory of active transport: II. The gate type non-carrier mechanism and generalizations concerning tracer flow, efficiency, and measurement of energy expenditure. *B. Math. Biophys.* **19**, 209-235.

Perez, C., Koshy, C., Yildiz, O. & Ziegler, C. (2012). Alternating-access mechanism in conformationally asymmetric trimers of the betaine transporter BetP. *Nature* **490**, 126–130.

Pfeiffer, R., Rossier, G., Spindler, B., Meier, C., Kühn, L., & Verrey, F. (1999). Amino acid transport of y+ L-type by heterodimers of 4F2hc/CD98 and members of the glycoprotein-associated amino acid transporter family. *EMBO J.* **18**, 49-57.

Pfeiffer, R., Spindler, B., Loffing, J., Skelly, P.J., Shoemaker, C.B. & Verrey, F. (1998). Functional heterodimeric amino acid transporters lacking cysteine residues involved in disulfide bond. *FEBS Lett.* **439**, 157–162.

Pickel, V.M., Nirenberg, M.J., Chan, J., Mosckovitz, R., Udenfriend, S. & Tate, S.S. (1993). Ultrastructural localization of a neutral and basic amino acid transporter in rat kidney and intestine. *Proc. Natl. Acad. Sci.* **90**, 7779–7783.

Pineda, M., Fernández, E., Torrents, D., Estévez, R., López, C., Camps, M., Lloberas, J., Zorzano, A. & Palacín, M. (1999). Identification of a membrane protein, LAT-2, that Co-expresses with 4F2 heavy chain, an L-type amino acid transport activity with broad specificity for small and large zwitterionic amino acids*. J. Biol. Chem.* **274**, 19738–19744.

Pineda, M., Font, M., Bassi, M.T., Manzoni, M., Borsani, G., Marigo, V., Fernández, E., Río, R.M. del, Purroy, J., Zorzano, A., Nunes, V. & Palacín, M. (2004a). The amino acid transporter asc-1 is not involved in cystinuria. *Kidney Int*. **66**, 1453–1464.

Pineda, M., Wagner, C.A., Bröer, A., Stehberger, P.A., Kaltenbach, S., Gelpí, J.L., Martín Del Río, R., Zorzano, A., Palacín, M., Lang, F. & Bröer, S. (2004b). Cystinuria-specific rBAT(R365W) mutation reveals two translocation pathways in the amino acid transporter rBAT-b0,+AT. *Biochem. J.* **377**, 665–674.

Pujadas, G., & Palau, J. (2001). Evolution of α-amylases: architectural features and key residues in the stabilization of the (β/α) 8 scaffold. *Mol. Biol. Evol.* **18**, 38-54.

Quackenbush, E., Clabby, M., Gottesdiener, K.M., Barbosa, J., Jones, N.H., Strominger, J.L., Speck, S. & Leiden, J.M. (1987). Molecular cloning of complementary DNAs encoding the heavy chain of the human 4F2 cell-surface antigen: a type II membrane glycoprotein involved in normal and neoplastic cell growth. *Proc. Natl. Acad. Sci.* **84**, 6526–6530.

Quackenbush, E.J., Linsley, P. & Letarte, M. (1986). Mouse L cells express a molecular complex carrying the human epitopes recognized by monoclonal antibodies 44D7 and 44H7 after DNA-mediated gene transfer. *J. Immunol*. **137**, 234–239.

Reddy, P., Liu, L., Adhikari, D., Jagarlamudi, K., Rajareddy, S., Shen, Y., Du, C., Tang, W., Hämäläinen, T., Peng, S. L., Lan, Z. J., Cooney, A. J., Huhtaniemi, I., & Liu, K. (2008). Oocyte-specific deletion of Pten causes premature activation of the primordial follicle pool. *Science* **319**, 611-613.

Reig, N. et al., 2002. The light subunit of system b(o,+) is fully functional in the absence of the heavy subunit. *EMBO J*. 21(18), 4906–4914.

Ressl, S., van Scheltinga, A. C. T., Vonrhein, C., Ott, V., & Ziegler, C. (2009). Molecular basis of transport and regulation in the Na+/betaine symporter BetP. *Nature* **458**, 47-52.

Ritchie, J., & Taylor, P. (2001). Role of the System L permease LAT1 in amino acid and iodothyronine transport in placenta. *Biochem. J*. **356**, 719-725.

Rius, M. & Chillarón, J. (2012). Carrier subunit of plasma membrane transporter is required for oxidative folding of its helper subunit. *J. Biol. Chem*. **287**, 18190–18200

Rossier, G., Meier, C., Bauch, C., Summa, V., Sordat, B., Verrey, F. & Kühn, L.C. (1999). LAT2, a new basolateral 4F2hc/CD98-associated amino acid transporter of kidney and intestine. *J. Biol. Chem*. **274**, 34948–34954.

Russ, W. P., & Engelman, D. M. (2000). The GxxxG motif: a framework for transmembrane helix-helix association. *J. Mol. Biol.* **296**, 911-919.

Sanders, C. R., & Myers, J. K. (2004). Disease-related misassembly of membrane proteins. *Annu. Rev. Biophys. Biomol. Struct.*, **33**, 25-51.

Sato, H., Tamba, M., Ishii, T. & Bannai, S. (1999). Cloning and expression of a plasma membrane cystine/glutamate exchange transporter composed of two distinct proteins. *J. Biol. Chem*. **274**, 11455–11458.

Savaskan, E., Ehrhardt, R., Schulz, A., Walter, M., & Schächinger, H. (2008). Post-learning intranasal oxytocin modulates human memory for facial identity. *Psychoneuroendocrinology* **33**, 368-374.

Schlessinger, A., Yee, S. W., Sali, A. & Giacomini, K. M. (2013). SLC classification: an update. *Clin Pharmacol Ther*. **94**, 19–23.

Schulze, S., Köster, S., Geldmacher, U., Terwisscha van Scheltinga, A.C. & Kühlbrandt, W. (2010). Structural basis of Na(+)-independent and cooperative substrate/product antiport in CaiT. *Nature* **467**, 233–236.

Seib, T.M., Patel, S.A., Bridges, R.J. (2011). Regulation of the system x(C)- cystine/glutamate exchanger by intracellular glutathione levels in rat astrocyte primary cultures. *Glia*, **59**, 1387-1401.

Seiler, A., Schneider, M., Förster, H., Roth, S., Wirth, E.K., Culmsee, C., Plesnila, N., Kremmer, E., Rådmark, O., Wurst, W., Bornkamm, G.W., Schweizer, U. & Conrad, M. (2008). Glutathione peroxidase 4 senses and translates oxidative stress into 12/15-lipoxygenase dependent- and AIF-mediated cell death. *Cell Metab*. **8**, 237–248.

Senes, A., Engel, D. E., & DeGrado, W. F. (2004). Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Curr Opin Struct Biol.* **14**, 465-479.

Senes, A., Gerstein, M., & Engelman, D. M. (2000). Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with β-branched residues at neighboring positions. *J. Mol. Biol.* **296**, 921-936.

Sengupta, S., Peterson, T. R., & Sabatini, D. M. (2010). Regulation of the mTOR complex 1 pathway by nutrients, growth factors, and stress. *Mol. Cell* **40**, 310-322.

Shaffer, P.L., Goehring, A., Shankaranarayanan, A. & Gouaux, E. (2009). Structure and mechanism of a Na^+-independent amino acid transporter. *Science* **325**, 1010–1014.

Shigeta, Y., Kanai, Y., Chairoungdua, a, Ahmed, N., Sakamoto, S., Matsuo, H., Kim, D.K., Fujimura, M., Anzai, N., Mizoguchi, K., Ueda, T., Akakura, K., Ichikawa, T., Ito, H. & Endou, H. (2006). A novel missense mutation of SLC7A9 frequent in Japanese cystinuria cases affecting the C-terminus of the transporter. *Kidney Int*. **69**, 1198–1206.

Shimamura, T., Weyand, S., Beckstein, O., Rutherford, N. G., Hadden, J. M., Sharples, D., Sansom M.S., Iwata S., Henderson P.J., Cameron, A. D. (2010). Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science* **328**, 470-473.

Simell O. (2002) Lysinuric protein intolerance and other cationic aminoacidurias. In: Scriver CR, Beaudet AL, Sly WS, Valle D, Vogelstein B, eds. The Online Metabolic and Molecular Bases of Inherited Disease (OMMBID). New York, NY: McGraw-Hill. Chap 192. 4933-4956.

Suga, S., Tsurudome, M., Ito, M., Ohgimoto, S., Tabata, N., Nishio, M., Kawano, M., Komada, H., Sakurai, M. & Ito, Y. (1997). Human immunodeficiency virus type-1 envelope glycoprotein gp120 induces expression of fusion regulatory protein (FRP)-1/CD98 on CD4+ T cells: a possible regulatory mechanism of HIV-induced syncytium formation. *Med. Microbiol. Immunol*. **185**, 237–243.

Takeuchi S., Wada K., Toyooka T., Shinomiya N., Shimazaki H., Nakanishi K., Nagatani K., Otani N., Osada H., Uozumi Y., Matsuo H., Nawashiro H. (2013). Increased xCT expression correlates with tumor invasion and outcome in patients with glioblastomas. *Neurosurgery* **72**, 33-41.

Tang L., Bai L., Wang W.H., Jiang T. (2010). Crystal structure of the carnitine transporter and insights into the antiport mechanism. *Nat Struct Mol Biol*. **17**, 492-496.

Tate, S., Yan, N. & Udenfriend, S. (1992). Expression cloning of a Na(+)-independent neutral amino acid transporter from rat kidney. *Proc. Natl. Acad. Sci*. **89**, 1–5.

Toivonen, M., Tringham, M., Kurko, J., Terho, P., Simell, O., Heiskanen, K.M., Mykkanen, J. (2013). Interactions of y+LAT1 and 4F2hc in the y+l amino acid transporter complex: consequences of lysinuric protein intolerance-causing mutations. *Gen. Physiol. Biophys*. **32**, 479–488.

Torrents, D., Pineda, M., Ferna, E., Lloberas, J., Shi, Y., Zorzano, A. & Palacı, M. (1998). Identification and Characterization of a Membrane Protein (y+L Amino Acid Transporter-1) That Associates with 4F2hc to Encode the Amino Acid Transport Activity y+L. *J. Biol. Chem*. **273**, 32437–32445.

Tsumura, H., Suzuki, N., Saito, H., Kawano, M., Otake, S., Kozuka, Y., Komada, H., Tsurudome, M. & Ito, Y. (2003). The targeted disruption of the CD98 gene results in embryonic lethality. *Biochem. Biophys. Res. Commun*. **308**, 847–851.

Turnay, J., Fort, J., Olmo, N., Santiago-Gómez, A., Palacín, M., Lizarbe, M. a A. & Antonia, M. (2011). Structural characterization and unfolding mechanism of human 4F2hc ectodomain. *Biochim. Biophys. Acta* **1814**, 536–544.

Tusnady, G. E., & Simon, I. (2001). The HMMTOP transmembrane topology prediction server. *Bioinformatics* **17**, 849-850.

Vafa, O., Wade, M., Kern, S., Beeche, M., Pandita, T. K., Hampton, G. M., & Wahl, G. M. (2002). c-Myc can induce DNA damage, increase reactive oxygen species, and mitigate p53 function: a mechanism for oncogene-induced genetic instability. *Molecular cell* **9**, 1031-1044.

Verrey, F., Closs, E. I., Wagner, C. A., Palacin, M., Endou, H., & Kanai, Y. (2004). CATs and HATs: the SLC7 family of amino acid transporters. *Pflügers Archiv*. **447**, 532-542.

Verrey, F., Jack, D. L., Paulsen, I. T., Saier Jr, M. H., & Pfeiffer, R. (1999). New glycoprotein-associated amino acid transporters. *J. Memb. Biol.* **172**, 181-192.

Verrey, F., Meier, C., Rossier, G. & Kühn, L.C. (2000). Glycoprotein-associated amino acid exchangers: broadening the range of transport specificity. *Pflugers Arch*. **440**, 503–512.

Weidle, U.H., Scheuer, W., Eggle, D., Klostermann, S. & Stockinger, H. (2010). Cancer-related issues of CD147. *Cancer Genomics Proteomics* **7**, 157–169.

Wells, R. G., & Hediger, M. A. (1992). Cloning of a rat kidney cDNA that stimulates dibasic and neutral amino acid transport and has sequence similarity to glucosidases. *Proc. Natl. Acad. Sci*. **89**, 5596-5600.

Wells, R.G., Lee, W.S., Kanai, Y., Leiden, J.M. & Hediger, M.A. (1992). The 4F2 antigen heavy chain induces uptake of neutral and dibasic amino acids in Xenopus oocytes. *J. Biol. Chem.* **267**, 15285–15288.

Weyand, S., Shimamura, T., Yajima, S., Suzuki, S. I., Mirza, O., Krusong, K., Carpenter, E. P., Rutherford, N. G., Hadden, J. M., O'Reilly, J., Ma, P., Saidijam, M. Patching, S. G., Hope, R.J., Norbertczak, H. T., Roach, P. C., Iwata, S., Henderson, P. J. & Cameron, A. D. (2008). Structure and molecular mechanism of a nucleobase–cation–symport-1 family transporter. *Science* **322**, 709-713.

Wiriyasermkul, P., Nagamori, S., Tominaga, H., Oriuchi, N., Kaira, K., Nakao, H., Kitashoji, T., Ohgaki, R., Tanaka, H., Endou, H., Endo, K., Sakurai, H., & Kanai, Y. (2012). Transport of 3-fluoro-L-α-methyl-tyrosine by tumor-upregulated L-type amino acid transporter 1: a cause of the tumor uptake in PET. *J. Nucl. Med.* **53**, 1253-1261.

Wolf, D.A., Wang, S., Panzica, M.A., Bassily, N.H. & Thompson, N.L. (1996). Expression of a highly conserved oncofetal gene, TA1/E16, in human colon carcinoma and other primary cancers: homology to *Schistosoma mansoni* amino acid permease and *Caenorhabditis elegans* gene products. *Cancer Res*. **56**, 5012–5022.

Wullschleger, S., Loewith, R., & Hall, M.N. (2006). TOR signaling in growth and metabolism. *Cell* **124**, 471-484.

Xie X., Dumas T., Tang L., Brennan T., Reeder T., Thomas W., Klein R.D., Flores J., O'Hara B.F., Heller H.C., Franken P. (2005). Lack of the alanine-serine-cysteine transporter 1 causes tremors, seizures, and early postnatal death in mice. *Brain Res*. **1052**, 212-221.

Xu, D. & Hemler, M.E. (2005). Metabolic activation-related CD147-CD98 complex. *Mol. Cell. Proteomics* **4**, 1061–1071.

Yamashita, A., Singh, S.K., Kawate, T., Jin, Y. & Gouaux, E. (2005). Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters. *Nature* **437**, 215–223.

Yan, N., Mosckovitz, R., Udenfriend, S. & Tate, S.S. (1992). Distribution of mRNA of a Na(+)-independent neutral amino acid transporter cloned from rat kidney and its expression in mammalian tissues and Xenopus laevis oocytes. *Proc. Natl. Acad. Sci.* **89**, 9982–9985.

Yanagida, O., Kanai, Y., Chairoungdua, A., Kim, D.K., Segawa, H., Nii, T., Cha, S.H., Matsuo, H., Fukushima, J., Fukasawa, Y., Tani, Y., Taketani, Y., Uchino, H., Kim, J.Y., Inatomi, J., Okayasu, I., Miyamoto, K., Takeda, E., Goya, T. & Endou, H. (2001). Human L-type amino acid transporter 1 (LAT1):

characterization of function and expression in tumor cell lines. *Biochim. Biophys. Acta* **1514**, 291–302.

Yoon, J.H., Kim, I.J., Kim, H., Kim, H.-J., Jeong, M.J., Ahn, S.G., Kim, S.A., Lee, C.H., Choi, B.K., Kim, J.-K., Jung, K.Y., Lee, S., Kanai, Y., Endou, H. & Kim, D.K. (2005). Amino acid transport system L is differently expressed in human normal oral keratinocytes and human oral cancer cells. *Cancer Lett.* **222**, 237–245.

Yu, F.H. & Catterall, W. a. (2003). Overview of the voltage-gated sodium channel family. *Genome Biol.* **4**, 207.

Zent, R., Fenczik, C.A., Calderwood, D.A., Liu, S., Dellos, M. & Ginsberg, M.H. (2000). Class- and splice variant-specific association of CD98 with integrin beta cytoplasmic domains. *J. Biol. Chem*. **275**, 5059–5064.

2. Bibliography: Chapter I

Bartoccioni, P., Rio, C., Kowalczyk, L., Jocelyn, M., Zorzano, A., Quick, M., Baldwin, S. a, Vázquez-ibar, J.L., Palacín, M., Ratera, M., Baldwin, J.M., Va, L. & Del Rio, C. (2010). Role of Transmembrane Domain 8 in Substrate Selectivity and Translocation of SteT , a Member of the L -Amino Acid Transporter (LAT) Family. *J. Biol. Chem.* **285**, 28764–28776.

Bröer, a, Friedrich, B., Wagner, C. a, Fillon, S., Ganapathy, V., Lang, F. & Bröer, S. (2001). Association of 4F2hc with light chains LAT1, LAT2 or y+LAT2 requires different domains. *Biochem. J.* **355**, 725–731.

Costa, M., Rosell, A., Álvarez-Marimon, E., Zorzano, A., Fotiadis, D. & Palacín, M. (2013). Expression of human heteromeric amino acid transporters in the yeast Pichia pastoris. *Protein Expr. Purif.* **87**, 35–40.

Drew, D., Newstead, S., Sonoda, Y., Kim, H., von Heijne, G. & Iwata, S. (2008). GFP-based optimization scheme for the overexpression and purification of eukaryotic membrane proteins in Saccharomyces cerevisiae. *Nat. Protoc.* **3**, 784–798.

Estévez, R., Camps, M., Rojas, A.M., Testar, X., Devés, R., Hediger, M.A., Zorzano, A. & Palacín, M. (1998). The amino acid transport system y+L/4F2hc is a heteromultimeric complex. *FASEB J.* **12**, 1319–1329.

Eswar, N., Eramian D FAU - Webb, B., Webb B FAU - Shen, M.-Y., Shen MY FAU - Sali, A. & Sali, A. (2008). Protein structure modeling with MODELLER. *Methods Mol. Biol.* **426**, 145–159.

Fenczik, C. a, Zent, R., Dellos, M., Calderwood, D. a, Satriano, J., Kelly, C. & Ginsberg, M.H. (2001). Distinct domains of CD98hc regulate integrins and amino acid transport. *J. Biol. Chem.* **276**, 8746–8752.

Fort, J., de la Ballina, L.R., Burghardt, H.E., Ferrer-costa, C., Turnay, J., Ferrer-orta, C., Usón, I., Zorzano, A., Fernández-Recio, J., Orozco, M., Lizarbe, M.A., Fita, I., Palacín, M., Ballina, L.R. De, Ferna, J., Palacı, M. & Hans, E. (2007). The structure of human 4F2hc ectodomain provides a model for homodimerization and electrostatic interaction with plasma membrane. *J. Biol. Chem.* **282**, 31444–31452.

Fotiadis, D., Kanai, Y. & Palacín, M. (2013). The SLC3 and SLC7 families of amino acid transporters. *Mol. Aspects Med.* **34**, 139–158.

Green, N.S., Reisler, E. & Houk, K.N. (2001). Quantitative evaluation of the lengths of homobifunctional protein cross-linking reagents used as molecular rulers. *Protein Sci*. **10**, 1293-1304.

Hizukuri, Y., Kojima, S. & Homma, M. (2010). Disulphide cross-linking between the stator and the bearing components in the bacterial flagellar motor. *J. Biochem.* **148**, 309–318.

Kowalczyk, L., Ratera, M., Paladino, A., Bartoccioni, P., Errasti-Murugarren, E., Valencia, E., Portella, G., Bial, S., Zorzano, A., Fita, I., Orozco, M., Carpena, X., Vázquez-Ibar, J.L. & Palacín, M. (2011). Molecular basis of substrate-induced permeation by an amino acid antiporter. *Proc. Natl. Acad. Sci.* **108**, 3935–3940.

Krishnamurthy, H. & Gouaux, E. (2012). X-ray structures of LeuT in substrate-free outward-open and apo inward-open states. *Nature* **481**, 469–474.

Rajan, D.P. (2000). Differential Influence of the 4F2 Heavy Chain and the Protein Related to b0,+ Amino Acid Transport on Substrate Affinity of the Heteromeric b0,+ Amino Acid Transporter. *J. Biol. Chem.* **275**, 14331–14335.

Reig, N., Chillarón, J., Bartoccioni, P., Fernández, E., Bendahan, A., Zorzano, A., Kanner, B., Palacín, M. & Bertran, J. (2002). The light subunit of system b(o,+) is fully functional in the absence of the heavy subunit. *EMBO J.* **21**, 4906–4914.

Reig, N., del Rio, C., Casagrande, F., Ratera, M., Gelpí, J.L., Torrents, D., Henderson, P.J.F., Xie, H., Baldwin, S. a, Zorzano, A., Fotiadis, D. & Palacín, M. (2007). Functional and structural characterization of the first prokaryotic member of the L-amino acid transporter (LAT) family: a model for APC transporters. *J. Biol. Chem.* **282**, 13270–13281.

Rosell, A., Meury, M., Álvarez-Marimon, E., Costa, M., Pérez-Cano, L., Zorzano, A., Fernández-Recio, J., Palacín, M. & Fotiadis, D. (2014). Structural bases for the interaction and stabilization of the human amino acid transporter LAT2 with its ancillary protein 4F2hc. *Proc. Natl. Acad. Sci.* **111**, 2966–2971.

Shi, Y. (2013). Common folds and transport mechanisms of secondary active transporters. *Annu. Rev. Biophys.* **42**, 51–72.

Shimamura, T., Weyand, S., Beckstein, O., Rutherford, N.G., Hadden, J.M., Sharples, D., Sansom, M.S.P., Iwata, S., Henderson, P.J.F. & Cameron, A.D. (2010). Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science* **328**, 470–473.

Turnay, J., Fort, J., Olmo, N., Santiago-Gómez, A., Palacín, M., Lizarbe, M. a A. & Antonia, M. (2011). Structural characterization and unfolding mechanism of human 4F2hc ectodomain. *Biochim. Biophys. Acta* **1814**, 536–544.

3. Bibliography: Chapter II

Agah, S. & Faham, S. (2012). Crystallization of membrane proteins in bicelles. *Methods Mol. Biol.* **914**, 3–16.

Althoff, T., Hibbs, R.E., Banerjee, S. & Gouaux, E. (2014). X-ray structures of GluCl in apo states reveal a gating mechanism of Cys-loop receptors. *Nature* **512**, 333–337.

Asojo, O.A., Koski, R.A. & Bonafé, N. (2011). Structural studies of human glioma pathogenesis-related protein 1. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 847–855.

Baconguis, I., Bohlen, C.J., Goehring, A., Julius, D. & Gouaux, E. (2014). X-ray structure of acid-sensing ion channel 1-snake toxin complex reveals open state of a Na(+)-selective channel. *Cell* **156**, 717–729.

Burger, K., Gimpl, G. & Fahrenholz, F. (2000). Regulation of receptor function by cholesterol. *Cell. Mol. Life Sci.* **57**, 1577–1592.

Caffrey, M. & Cherezov, V. (2009). Crystallizing membrane proteins using lipidic mesophases. *Nat. Protoc.* **4**, 706–731.

Cherezov, V., Rosenbaum, D.M., Hanson, M. a, Rasmussen, S.G.F., Thian, F.S., Kobilka, T.S., Choi, H.-J., Kuhn, P., Weis, W.I., Kobilka, B.K. & Stevens, R.C. (2007). High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **318**, 1258–1265.

Chien, E.Y.T., Liu, W., Zhao, Q., Katritch, V., Han, G.W., Hanson, M. a, Shi, L., Newman, A.H., Javitch, J. a, Cherezov, V. & Stevens, R.C. (2010). Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* **330**, 1091–1095.

Cormack, B.P., Bertram, G., Egerton, M., Gow, N. a, Falkow, S. & Brown, a J. (1997). Yeast-enhanced green fluorescent protein (yEGFP): a reporter of gene expression in Candida albicans. *Microbiology* **143**, 303–311.

Costa, M., Rosell, A., Álvarez-Marimon, E., Zorzano, A., Fotiadis, D. & Palacín, M. (2013). Expression of human heteromeric amino acid transporters in the yeast Pichia pastoris. *Protein Expr. Purif.* **87**, 35–40.

Deng, D., Xu, C., Sun, P., Wu, J., Yan, C., Hu, M. & Yan, N. (2014). Crystal structure of the human glucose. *Nature* **510**, 121–125.

Drew, D., Newstead, S., Sonoda, Y., Kim, H., von Heijne, G. & Iwata, S. (2008). GFP-based optimization scheme for the overexpression and purification of eukaryotic membrane proteins in Saccharomyces cerevisiae. *Nat. Protoc.* **3**, 784–798.

Dürr, K.L., Chen, L., Stein, R.A., De Zorzi, R., Folea, I.M., Walz, T., Mchaourab, H.S. & Gouaux, E. (2014). Structure and dynamics of AMPA receptor GluA2 in resting, pre-open, and desensitized states. *Cell* **158**, 778–792.

Elbaz, Y., Danieli, T., Kanner, B.I. & Schuldiner, S. (2011). Expression of neurotransmitter transporters for structural and biochemical studies. *Protein Expr. Purif.* **73**, 152–160.

Epand, R.M. (2006). Cholesterol and the interaction of proteins with membrane domains. *Prog. Lipid Res.* **45**, 279–294.

Faham, S. & Bowie, J.U. (2002). Bicelle Crystallization : A new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *J. Mol. Biol.* **316**, 1–6.

Fort, J., de la Ballina, L.R., Burghardt, H.E., Ferrer-costa, C., Turnay, J., Ferrer-orta, C., Usón, I., Zorzano, A., Fernández-Recio, J., Orozco, M., Lizarbe, M.A., Fita, I., Palacín, M., Ballina, L.R. De, Ferna, J., Palacı, M. & Hans, E. (2007). The structure of human 4F2hc ectodomain provides a model for homodimerization and electrostatic interaction with plasma membrane. *J. Biol. Chem.* **282**, 31444–31452.

Frick, A., Kosinska, U., Mattia, F. De, Öberg, F., Hedfalk, K. & Neutze, R. (2014). X-ray structure of human aquaporin 2 and its implications for nephrogenic diabetes insipidus and trafficking. *Proc. Natl. Acad. Sci.* **111**, 6305–6310.

Gimpl, G. (2010). Cholesterol-protein interaction: methods and cholesterol reporter molecules. *Subcell. Biochem.* **51**, 1–45.

Gourdon, P., Andersen, J.L., Hein, K.L., Bublitz, M., Pedersen, B.P., Liu, X.-Y., Yatime, L., Nyblom, M., Nielsen, T.T., Olesen, C., Møller, J.V., Nissen, P. & Morth, J.P. (2011). HiLiDe—Systematic approach to membrane protein crystallization in lipid and detergent. *Cryst. Growth Des.* **11**, 2098–2106.

Gruswitz, F., Chaudhary, S., Ho, J.D., Schlessinger, A., Pezeshki, B., Ho, C.-M., Sali, A., Westhoff, C.M. & Stroud, R.M. (2010). Function of human Rh based on structure of RhCG at 2.1 A. *Proc. Natl. Acad. Sci.* **107**, 9638–9643.

Guan, L., Smirnova, I.N., Verner, G., Nagamori, S., Nagamoni, S. & Kaback, H.R. (2006). Manipulating phospholipids for crystallization of a membrane transport protein. *Proc. Natl. Acad. Sci.* **103**, 1723–1726.

Gutmann, D.A.P., Mizohata, E., Newstead, S., Ferrandon, S., Henderson, P.J.F., Veen, H.W.V.A.N. & Byrne, B. (2007). A high-throughput method for membrane protein solubility screening : The ultracentrifugation dispersity sedimentation assay. *Protein Sci.* **16**, 1422–1428.

Haga, K., Kruse, A.C., Asada, H., Yurugi-Kobayashi, T., Shiroishi, M., Zhang, C., Weis, W.I., Okada, T., Kobilka, B.K., Haga, T. & Kobayashi, T. (2012). Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature* **482**, 547–551.

Hanson, M. a, Cherezov, V., Griffith, M.T., Roth, C.B., Jaakola, V.-P., Chien, E.Y.T., Velasquez, J., Kuhn, P. & Stevens, R.C. (2008). A specific cholesterol binding site is established by the 2.8 A structure of the human beta2-adrenergic receptor. *Structure* **16**, 897–905.

Hanson, M.A., Roth, C.B., Jo, E., Griffith, M.T., Scott, F.L., Reinhart, G., Desale, H., Clemons, B., Cahalan, S.M., Schuerer, S.C., Sanna, M.G., Han, G.W., Kuhn, P., Rosen, H. & Stevens, R.C. (2012). Crystal structure of a lipid G protein-coupled receptor. *Science* **335**, 851–855.

Harder, D. & Fotiadis, D. (2012). Measuring substrate binding and affinity of purified membrane transport proteins using the scintillation proximity assay. *Nat. Protoc.* **7**, 1569–1578.

Hattori, M. & Gouaux, E. (2012). Molecular mechanism of ATP binding and ion channel activation in P2X receptors. *Nature* **485**, 207–212.

Hattori, M., Hibbs, R.E. & Gouaux, E. (2012). A fluorescence-detection size-exclusion chromatography-based thermostability assay for membrane protein precrystallization screening. *Structure* **20**, 1293–1299.

Hattori, M., Tanaka, Y., Fukai, S., Ishitani, R. & Nureki, O. (2007). Crystal structure of the MgtE Mg2+ transporter. *Nature* **448**, 1072–1075.

Hibbs, R.E. & Gouaux, E. (2011). Principles of activation and permeation in an anion-selective Cys-loop receptor. *Nature* **474**, 54–60.

Hunte, C., Koepke, J., Lange, C., Roßmanith, T. & Michel, H. (2000). Structure at 2 . 3 Å resolution of the cytochrome bc 1 complex from the yeast Saccharomyces cerevisiae co-crystallized with an antibody Fv fragment. *Structure* **8**, 669–684.

Jaakola, V.-P., Griffith, M.T., Hanson, M.A., Cherezov, V., Chien, E.Y.T., Lane, J.R., Ijzerman, A.P. & Stevens, R.C. (2008). The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **322**, 1211–1217.

Jasti, J., Furukawa, H., Gonzales, E.B. & Gouaux, E. (2007). Structure of acid-sensing ion channel 1 at 1.9 Å resolution and low pH. *Nature* **449**, 316–323.

Jidenko, M., Nielsen, R.C., Sørensen, T.L.-M., Møller, J. V, le Maire, M., Nissen, P. & Jaxel, C. (2005). Crystallization of a mammalian membrane protein overexpressed in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* **102**, 11687–11691.

Jones, K.T., Zhen, J. & Reith, M.E.A. (2012). Importance of cholesterol in dopamine transporter function. *J. Neurochem.* **123**, 700–715.

Kampmann, M. & Blobel, G. (2009). Three-dimensional structure and flexibility of a membrane-coating module of the nuclear pore complex. *Nat. Struct. Mol. Biol.* **16**, 782–788.

Kawate, T. & Gouaux, E. (2006). Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. *Structure* **14**, 673–681.

Kawate, T., Michel, J.C., Birdsong, W.T. & Gouaux, E. (2009). Crystal structure of the ATP-gated P2X(4) ion channel in the closed state. *Nature* **460**, 592–598.

Kellosalo, J., Kajander, T., Palmgren, M.G., Lopéz-marqués, R.L. & Goldman, A. (2011). Heterologous expression and purification of membrane-bound pyrophosphatases. *Protein Expr. Purif.* **79**, 25–34.

Koopmann, R., Cupelli, K., Redecke, L., Nass, K., Daniel, P., White, T.A., Stellato, F., Rehders, D., Liang, M., Aquila, A., Bajt, S., Barthelmess, M., Barty, A., Michael, J., Bostedt, C., Boutet, S., Bozek, J.D., Caleman, C., Davidsson, J., Doak, R.B., Ekeberg, T., Epp, S.W., Erk, B., Fleckenstein, H., Foucar, L., Graafsma, H., Hauser, G., Hirsemann, H., Holl, P., Hunter, M.S., Kirian, R.A., Lomb, L., Maia, F.R.N.C., Kimmel, N., Rudenko, A., Schlichting, I., Schulz, J., Marvin, M., Strüder, L., Timneanu, N., Ullrich, J., Wang, X., Weierstall, U., Williams, G.J., Wunderer, C.B., Spence, J.C.H., Stehle, T., Chapman, H.N., Betzel, C. & Duszenko, M. (2012). In vivo protein crystallization opens new routes in structural biology. *Nat. Meth.* **9**, 259–262.

Kowalczyk, L., Ratera, M., Paladino, A., Bartoccioni, P., Errasti-Murugarren, E., Valencia, E., Portella, G., Bial, S., Zorzano, A., Fita, I., Orozco, M., Carpena, X., Vázquez-Ibar, J.L. & Palacín, M. (2011). Molecular basis of substrate-induced permeation by an amino acid antiporter. *Proc. Natl. Acad. Sci.* **108**, 3935–3940.

Krishnamurthy, H. & Gouaux, E. (2012). X-ray structures of LeuT in substrate-free outward-open and apo inward-open states. *Nature* **481**, 469–474.

Kurisu, G., Zhang, H., Smith, J.L. & Cramer, W.A. (2003). Structure of the cytochrome b6f complex of oxygenic photosynthesis: tuning the cavity. *Science* **302**, 1009–1014.

Lee, A.G. (2004). How lipids affect the activities of integral membrane proteins. *Biochim. Biophys. Acta* **1666**, 62–87.

Lee, C.-H., Lü, W., Michel, J.C., Goehring, A., Du, J., Song, X. & Gouaux, E. (2014). NMDA receptor structures reveal subunit arrangement and pore architecture. *Nature* **511**, 191–197.

Lemieux, M.J., Song, J., Kim, M.J., Huang, Y., Villa, A., Auer, M., Li, X.-D. & Wang, D.-N. (2003). Three-dimensional crystallization of the Escherichia coli glycerol-3-phosphate transporter: a member of the major facilitator superfamily. *Protein Sci.* **12**, 2748–2756.

Li, H. & Papadopoulos, V. (1998). Peripheral-type benzodiazepine receptor function in cholesterol transport. Identification of a putative cholesterol recognition/interaction amino acid sequence and consensus pattern. *Endocrinology* **139**, 4991–4997.

Lin, S.-M., Tsai, J.-Y., Hsiao, C.-D., Huang, Y.-T., Chiu, C.-L., Liu, M.-H., Tung, J.-Y., Liu, T.-H., Pan, R.-L. & Sun, Y.-J. (2012). Crystal structure of a membrane-embedded H+-translocating pyrophosphatase. *Nature* **484**, 399–403.

Long, S.B., Tao, X., Campbell, E.B. & MacKinnon, R. (2007). Atomic structure of a voltage-dependent K+ channel in a lipid membrane-like environment. *Nature* **450**, 376–382.

Marchand, A., Winther, A.-M.L., Holm, P.J., Olesen, C., Montigny, C., Arnou, B., Champeil, P., Clausen, J.D., Vilsen, B., Andersen, J.P., Nissen, P., Jaxel, C., Møller, J.V. & le Maire, M. (2008). Crystal structure of D351A and P312A mutant forms of the mammalian sarcoplasmic reticulum Ca(2+) -ATPase reveals key events in phosphorylation and Ca(2+) release. *J. Biol. Chem.* **283**, 14867–14882.

Miller, P.S. & Aricescu, A.R. (2014). Crystal structure of a human GABAA receptor. *Nature* **512**, 270–275.

Mizutani, K., Yoshioka, S., Mizutani, Y., Iwata, S. & Mikami, B. (2011). High-throughput construction of expression system using yeast Pichia pastoris, and its application to membrane proteins. *Protein Expr. Purif.* **77**, 1–8.

Mumberg, D., Müller, R. & Funk, M. (1995). Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* **156**, 119–122.

Newstead, S., Kim, H., Heijne, G. Von, Iwata, S. & Drew, D. (2007). of eukaryotic membrane protein overexpression and purification in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci.* **104**, 13936–13941.

Niegowski, D., Kleinschmidt, T., Olsson, U., Ahmad, S., Haeggström, J.Z. & Rinaldo-matthis, A. (2014). Crystal Structures of Leukotriene C 4 Synthase in Complex with Product Analogs. *J. Biol. Chem.* **289**, 5199–5207.

Pagadala, V., Vistain, L., Symersky, J. & Mueller, D.M. (2011). Characterization of the mitochondrial ATP synthase from yeast Saccharomyces cerevisae. *J. Bioenerg. Biomembr.* **43**, 333–347.

Parker, J.L. & Newstead, S. (2014). Molecular basis of nitrate uptake by the plant nitrate transporter NRT1.1. *Nature* **507**, 68–72.

Pedersen, B.P., Buch-Pedersen, M.J., Morth, J.P., Palmgren, M.G. & Nissen, P. (2007). Crystal structure of the plasma membrane proton pump. *Nature* **450**, 1111–1114.

Pedersen, B.P., Kumar, H., Waight, A.B., Risenmay, A.J., Roe-zurz, Z., Chau, B.H., Schlessinger, A., Bonomi, M., Harries, W., Sali, A., Johri, A.K. & Stroud, R.M. (2013). transporter. *Nature* **496**, 533–536.

Pedersen, H.L., Willassen, N.P. & Leiros, I. (2009). The first structure of a cold-adapted superoxide dismutase (SOD): biochemical and structural characterization of iron SOD from Aliivibrio salmonicida. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* **65**, 84–92.

Penmatsa, A., Wang, K.H. & Gouaux, E. (2013). X-ray structure of dopamine transporter elucidates antidepressant mechanism. *Nature* **503**, 85–90.

Phillips, R., Ursell, T., Wiggins, P. & Sens, P. (2009). Emerging roles for lipids in shaping membrane-protein function. *Nature* **459**, 379–385.

Prasanna, X., Chattopadhyay, A. & Sengupta, D. (2014). Cholesterol modulates the dimer interface of the β_2-adrenergic receptor via cholesterol occupancy sites. *Biophys. J.* **106**, 1290–1300.

Rajan, D.P. (2000). Differential Influence of the 4F2 Heavy Chain and the Protein Related to b0,+ Amino Acid Transport on Substrate Affinity of the Heteromeric b0,+ Amino Acid Transporter. *J. Biol. Chem.* **275**, 14331–14335.

Rajan, D.P., Kekuda, R., Huang, W., Wang, H., Devoe, L.D., Leibach, F.H., Prasad, P.D. & Ganapathy, V. (1999). Cloning and expression of a b(0,+)-like amino acid transporter functioning as a heterodimer with 4F2hc instead of rBAT. A new candidate gene for cystinuria. *J. Biol. Chem.* **274**, 29005–29010.

Rasmussen, S.G.F., Choi, H.-J., Fung, J.J., Pardon, E., Casarosa, P., Chae, P.S., Devree, B.T., Rosenbaum, D.M., Thian, F.S., Kobilka, T.S., Schnapp, A., Konetzki, I., Sunahara, R.K., Gellman, S.H., Pautsch, A., Steyaert, J., Weis, W.I. & Kobilka, B.K. (2011). Structure of a nanobody-stabilized active state of the β(2) adrenoceptor. *Nature* **469**, 175–180.

Rasmussen, S.G.F., Choi, H.-J., Rosenbaum, D.M., Kobilka, T.S., Thian, F.S., Edwards, P.C., Burghammer, M., Ratnala, V.R.P., Sanishvili, R., Fischetti, R.F., Schertler, G.F.X., Weis, W.I. & Kobilka, B.K. (2007). Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature* **450**, 383–387.

Reynolds, B., Roversi, P., Laynes, R., Kazi, S., Boyd, C.A.R. & Goberdhan, D.C.I. (2009). Drosophila expresses a CD98 transporter with an evolutionarily conserved structure and amino acid-transport properties. *Biochem. J.* **420**, 363–372.

Rosell, A., Meury, M., Álvarez-marimon, E., Costa, M. & Pérez-cano, L. (2014). Structural bases for the interaction and stabilization of the human amino acid transporter LAT2 with its ancillary protein 4F2hc. *Proc. Natl. Acad. Sci.* **111**, 2966–2971.

Rosenbaum, D.M., Zhang, C., Lyons, J.A., Holl, R., Aragao, D., Arlow, D.H., Rasmussen, S.G.F., Choi, H.-J., Devree, B.T., Sunahara, R.K., Chae, P.S., Gellman, S.H., Dror, R.O., Shaw, D.E., Weis, W.I., Caffrey, M., Gmeiner, P. & Kobilka, B.K. (2011). Structure and function of an irreversible agonist-β(2) adrenoceptor complex. *Nature* **469**, 236–240.

Scharff-Poulsen, P. & Pedersen, P.A. (2013). Saccharomyces cerevisiae-based platform for rapid production and evaluation of eukaryotic nutrient transporters and transceptors for biochemical studies and crystallography. *PLoS One* **8**, e76851.

Schroeder, N., Chung, C.-S., Chen, C.-H., Liao, C.-L. & Chang, W. (2012). The lipid raft-associated protein CD98 is required for vaccinia virus endocytosis. *J. Virol.* **86**, 4868–4882.

Seddon, A.M., Curnow, P. & Booth, P.J. (2004). Membrane proteins, lipids and detergents: not just a soap opera. *Biochim. Biophys. Acta* **1666**, 105–117.

Shaffer, P.L., Goehring, A., Shankaranarayanan, A. & Gouaux, E. (2009). Structure and mechanism of a Na+-independent amino acid transporter. *Science* **325**, 1010–1014.

Shimamura, T., Shiroishi, M., Weyand, S., Tsujimoto, H., Winter, G., Katritch, V., Abagyan, R., Cherezov, V., Liu, W., Han, G.W., Kobayashi, T., Stevens, R.C. & Iwata, S. (2011). Structure of the human histamine H1 receptor complex with doxepin. *Nature* **475**, 65–70.

Sjögren, T., Nord, J., Ek, M., Johansson, P., Liu, G. & Geschwindner, S. (2013). Crystal structure of microsomal prostaglandin E 2 synthase provides insight into diversity in the MAPEG superfamily. *Proc. Natl. Acad. Sci.* **110**, 3806–3811.

Sobolevsky, A.I., Rosconi, M.P. & Gouaux, E. (2009). X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor. *Nature* **462**, 746–756.

Son, S., Ma, J., Kondou, Y., Yoshimura, M., Yamashita, E., Tsukihara, T., L, D.C., Natl, P. & Sci, A. (2008). Structure of human monoamine oxidase A at 2 . 2-Å resolution : The control of opening the entry for substrates / inhibitors. *Proc. Natl. Acad. Sci.* **105**, 5739-5944.

Symersky, J., Pagadala, V., Osowski, D., Krah, A., Meier, T., Faraldo-Gómez, J.D. & Mueller, D.M. (2012). Structure of the c(10) ring of the yeast mitochondrial ATP synthase in the open conformation. *Nat. Struct. Mol. Biol.* **19**, 485–91, S1.

Tao, X., Lee, A., Limapichat, W., Dougherty, D.A. & MacKinnon, R. (2010). A gating charge transfer center in voltage sensors. *Science* **328**, 67–73.

Vaughn, J.L., Goodwin, R.H., Tompkins, G.J. & McCawley, P. (1977). The establishment of two cell lines from the insect Spodoptera frugiperda (Lepidoptera; Noctuidae). *In Vitro* **13**, 213–217.

Wacker, D., Wang, C., Katritch, V., Han, G.W., Huang, X., Vardy, E., Mccorvy, J.D., Jiang, Y., Chu, M., Siu, F.Y., Liu, W., Xu, H.E., Cherezov, V., Roth, B.L. & Stevens, R.C. (2013). Structural Features for Functional. *Science* **340**, 615–619.

Waight, A.B., Pedersen, B.P., Schlessinger, A., Bonomi, M., Chau, B.H., Roe-zurz, Z., Risenmay, A.J., Sali, A. & Stroud, R.M. (2013). Structural basis for alternating access of a eukaryotic calcium/proton exchanger. *Nature* **498**, 107–110.

Wu, B., Chien, E.Y.T., Mol, C.D., Fenalti, G., Liu, W., Katritch, V., Abagyan, R., Brooun, A., Wells, P., Bi, F.C., Hamel, D.J., Kuhn, P., Handel, T.M., Cherezov, V. & Stevens, R.C. (2010). Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **330**, 1066–1071.

Wu, H., Wacker, D., Mileni, M., Katritch, V., Han, G.W., Vardy, E., Liu, W., Thompson, A. a, Huang, X.-P., Carroll, F.I., Mascarella, S.W., Westkaemper, R.B., Mosier, P.D., Roth, B.L., Cherezov, V. & Stevens, R.C. (2012). Structure of the human κ-opioid receptor in complex with JDTic. *Nature* **485**, 327–332.

Yamashita, A., Singh, S.K., Kawate, T., Jin, Y. & Gouaux, E. (2005). Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters. *Nature* **437**, 215–223.

Yao, Z. & Kobilka, B. (2005). Using synthetic lipids to stabilize purified beta2 adrenoceptor in detergent micelles. *Anal. Biochem.* **343**, 717–729.

Yuan, P., Leonetti, M.D., Pico, A.R., Hsiung, Y. & MacKinnon, R. (2010). Structure of the human BK channel Ca^{2+}-activation apparatus at 3.0 A resolution. *Science* **329**, 182–186.

Zhang, H., Kurisu, G., Smith, J.L. & Cramer, W.A. (2003). A defined protein-detergent-lipid complex for crystallization of integral membrane proteins: The cytochrome b6f complex of oxygenic photosynthesis. *Proc. Natl. Acad. Sci.* **100**, 5160–5163.