4F2hc/LAT1 heterodimer: a promising candidate for solving the first atomic structure of a Heteromeric Amino acid Transporter

Heterodímero 4F2hc/LAT1: un candidato prometedor para resolver la primera estructura atómica de un transportador heteromérico de aminoácidos

Ana Zuleima Obando Martínez

Esta tesis ha sido realizada en el Instituto de investigación Biomédica de Barcelona (IRB Barcelona) dentro del programa de Doctorado de Biotecnología de la Universidad de Barcelona (UB)
Facultad de Biología, Departamento de Bioquímica y Biología Molecular (UB)

4F2hc/LAT1 heterodimer: a promising candidate for solving the first atomic structure of a Heteromeric Amino acid Transporter

Heterodímero 4F2hc/LAT1: un candidato prometedor para resolver la primera estructura atómica de un transportador heteromérico de aminoácidos

Memoria para optar al grado de Doctor por la Universidad de Barcelona
presentada por:

Ana Zuleima Obando Martínez

El Director,	El Codirector,	La Interesada,
Manuel Palacín Prieto	Albert Rosell Febres	Ana Zuleima Obando Martínez

Barcelona, 2016

A mis padres, a mi hermana, y a Mike...

Gracias!

Acknowledgments

Es difícil para mí resumir en pocas palabras el gran agradecimiento que tengo hacia todas aquellas personas que, de modo directo o indirecto, han contribuido en el desarrollo de esta tesis.

Empezaré agradeciendo a mis directores de tesis:

A Manuel Palacín, por su gran apoyo desde el principio para poder venir a hacer mi tesis en su laboratorio, aquí en Barcelona, y seguir preocupándose por mí una vez llegué, no solo en lo referente a la tesis, si no también a mi bienestar en general, lo cual nunca olvidaré. Ya que sí, ha sido bastante difícil estar lejos de mi familia y seres queridos durante todo este tiempo, y haber recibido su apoyo y preocupación hacia mí, fue muy importante. Además de lo anterior, Manuel Palacín ha sido para mí un gran profesor y un ejemplo de científico; curioso, apasionado, objetivo, exhaustivo, y un gran etc., además de una gran persona. Gracias Manuel por dirigir esta tesis, por sus preguntas que nos motivan a pensar más allá, por las discusiones científicas, y por la ilusión que nos infunde para seguir investigando

De igual forma, quiero agradecer a mi otro director, Albert Rosell, sin el cual esta tesis no habría sido posible, ya que en el momento necesario, me permitió trabajar con él en este proyecto, que forma parte de una ardua, pero muy interesante investigación que lleva años desarrollando con rigor y perseverancia. Gracias por tu apoyo, enseñanzas y guía. Has sido para mí, un excelente profesor y un científico integral, pero por lo que más te estaré agradecida siempre, es por tu enorme preocupación hacia mí en lo referente a la tesis, pero sobre todo en lo referente al plano personal, ya que como he dicho antes, no ha sido nada fácil, y tú me mostraste "la luz al final del camino". No quería ser tan informal, pero tratándose de la persona que "me rescato" del "malvivir" de los primeros años de mis estudios, hacia un proyecto bastante interesante y prometedor, y hacia un muchísimo mejor vivir, no puedo mas que agradecerte enormemente en las palabras que me salen... quería decir del corazón, pero esto puede ser demasiado dramático y sentimental en España y otras partes del mundo. Gracias por ser un gran ser humano, por los buenos y graciosos momentos que pasamos durante el desarrollo de esta tesis, por el trabajo constante en equipo, y por todo lo demás que ya he mencionado antes.

También quiero agradecer especialmente a Lukasz Kowalczyk, porque fue él quien me inicio en el mundo de la estructura de proteínas de membrana y cristalografía, y me enseño muchísimas cosas durante el tiempo que coincidimos en el laboratorio. La parte de AdiC, la pude desarrollar gracias a sus enseñanzas. Lukasz también fue un ejemplo para mí de rigor científico, y de cómo deberían abordarse este tipo de estudios. Además, fue un apoyo y una persona que se preocupó por mí y mi aprendizaje mientras trabajamos juntos.

Quiero agradecer además a otras personas que también colaboraron directamente en el desarrollo de esta tesis:

A la obra social "La Caixa", por la beca que me permitió desarrollar mis estudios. A los miembros de mi comité de asesoramiento académico, Antonio Zorzano, Ignasi Fita y Marçal Pastor, los cuales año tras año fueron evaluando el desarrollo de mis estudios doctorales, en pos de mejorarlos y para ayudarme en mi formación doctoral. Quiero agradecer a Antonio Zorzano, el cual por su cercanía con el laboratorio de Manuel Palacín, no solo contribuyó como parte del comité de asesoramiento, si no que también lo hizo durante nuestros seminarios de

grupo y conversaciones por el laboratorio, aportando ideas y preguntas interesantes. A Ignasi Fita quiero agradecerle además por su gran apoyo y enseñanzas en lo relacionado a cristalografía y estructura de proteínas, ya que además él, y su grupo, contribuyeron con el desarrollo de los experimentos y análisis de resultados. Además de lo anterior, valoro muchísimo la calidez humana con la cual Ignasi mostraba interés por mis estudios, cuando nos encontrábamos por los pasillos y me preguntaba acerca de mis proyectos.

Quiero agradecer a todos los miembros del grupo de estructura de Ignasi por su colaboración, y a aquellos a los cuales tuve la oportunidad de conocer un poquito más, tanto a los que están ahora como los que ya se han ido. A Luca, a Maria a David. A Xavier Carpena por ayudar en la difracción de rayos X de algunos de los cristales de AdiC, y en el análisis de datos, además de por sus enseñanzas de cristalografía. A Rosa por su gran ayuda cuando necesitaba algún reactivo, material, o tenía alguna pregunta, además de por ser una persona muy agradable.

Quiero agradecer al grupo de Modesto Orozco, especialmente a Guillem Portella, antiguo miembro, por su gran colaboración en el proyecto de AdiC y la fabricación de uno de los modelos de AdiC en "inward-facing conformation", que se muestra en el apéndice de esta tesis. También por su amabilidad y simpatía. En la misma línea, quiero agradecer especialmente a Rafael Nuñez (CIB-CSIC, Madrid) por su colaboración en la toma y análisis de imágenes de las tinciones negativas de 4F2hc/LAT1, y su reconstrucción 3D, así como por explicarme algunos de los pormenores de estos estudios y responder a todas mis preguntas. Todo esto siempre con gran amabilidad y simpatía. Especiales gracias a Oscar Llorca (CIB-CSIC, Madrid) por colaborar con nosotros en los estudios para intentar resolver la estructura de 4F2hc/LAT1 por cryo-EM.

Agradezco a todos los miembros del servicio de expresión de proteínas del IRB, y de igual forma, a la Plataforma Automatizada de Cristalografía (PAC) del Parque Científico de Barcelona, al servicio de purificación de proteínas, y al servicio de microscopía, por los servicios prestados para poder realizar esta tesis. En especial, quiero agradecer a un ex miembro del PAC, Sonia Rodríguez, y a Joan Pous por su trabajo y ayuda, también por las conversaciones durante mis idas a la plataforma. A Queralt también quiero agradecerle por los mismos motivos, pero en el servicio de expresión de proteínas del IRB. A Lydia y a Xavi, del servicio de microscopía, por su colaboración y enseñanzas en el microscopio electrónico.

Quiero agradecer muy especialmente a Manuel Elkin Patarroyo, a Manuel Alfonso Patarroyo, y a Hernando Curtidor, por permitirme trabajar junto a ellos en la Fundación Instituto de Inmunología de Colombia (FIDIC), la cual fue mi primera experiencia real en investigación, que me motivo enormemente a seguir por este camino. Gracias por sus enseñanzas, y por ser una referencia en la investigación científica. Gracias además por apoyarme en la realización de mis estudios con sus recomendaciones, sin ellas esto no habría sido posible.

Respecto a las personas que mas he tenido cerca durante estos años, para no dejar sin mencionar a alguien, quisiera agradecer a todas las personas del grupo de Manuel y Antonio, con las que he compartido buenos momentos, risas, estrés y demás cosas del día a día, así como a aquellas que me han ayudado en el desarrollo de experimentos. A Paola Bartoccioni por enseñarme a hacer reconstituciones, transportes y SPA. Por tus enseñanzas, por tu ayuda y un interminable etc., pero sobre todo por tu buen rollo italiano colombian-like ja ja ja, majísima, como dirían aquí, divertida, buena persona. De igual forma quiero agradecer a Susana Bial, que me enseño mucho durante los primeros años y me ayudo en varios experimentos, también por

ser buena persona y divertida, como su amiga de arriba. A Jordi Seco quiero agradecerle por su excelente trabajo que nos facilita la vida, por las risas que nos causa, cuando está de buen humor ji ji (ya sé que es difícil lidiar con tanta gente indisciplinada y testaruda), también por sus silbidos y sus cantares que a mí me hacían gracia, y subían el ánimo. Quiero agradecer a Olga por su ayuda con todos los temas burocráticos y demás, y ante todo por su gracia y alegría. Gracias a Joana por su ayuda con el AKTA cuando la necesité, y por enseñarme mas cosas referentes a éste.

Tranquilos que no he terminado, vamos a hablar también de mis queridísimos "catalans". En primer lugar Ma. Angels, Alba y Aida, así en grupo también evito el peligro de confundir una con la otra, aunque en mi mente ya estoy confundida ;). Han dejado, pero que muy en alto, el nombre de Catalunya, junto con Albert, Xavi, Edu, Elena, Arturo, Meri y Jordi Coste, ya creo que no es verdad lo que otra gente, no yo, dicen por ahí sobre lo malos que son los catalanes. Yo los veo a ustedes y digo, ¡que amorosos y buenas personas que son!. Así que gracias niñas por los buenos momentos, por ser tan majas, por reírse con mis tonterías, por lo "liantas", por los chistes de Ma Angels, boníssims! (pero el ver como ella se ríe y lo demás no), y por sus copitas de vermut. A las AlbAida, porque era chévere que hubiera alguien de mi misma edad en el lab, estaba harta de tanta gente vieja. Quiero agradecer a Edu por lo gracioso que es, por su personalidad que no permitía que nos aburriéramos, y sobre todo por los bailes tipo Michael Jackson. Tú continúa que te veo un futuro prometedor por ese lado. A Arturo por sus conversaciones profundas y altamente argumentativas, acerca de todo y acerca de nada, relacionadas o no con el lab, gracias a ello he aprendido datos interesantes y curiosos. También me has mostrado como un cuerpo puede doblarse sobre sí mismo más allá de sus capacidades. A Xavi Duran, por querer salvar el mundo y a los animales, está muy bien, necesitamos mas personas como tú, pero menos intensos ;) , ji ji. A Elena, por su simpatía cuando llegue al laboratorio, por los momentos de distracción en nuestro pasillo, y charlas sobre los experimentos y demás. Te agradezco enormemente la ayuda que me diste cuando estaba escribiendo esta tesis, el pymol, los consejos, entre otros. A Meri Costa, que aunque hace mucho ya que se fue del lab, en el tiempo que coincidimos me encantó su personalidad; tranquila, generosa, amable y graciosa. Gracias también por, junto con Elena, llevar a cabo los estudios precedentes a esta tesis, que han permitido que esta sea posible.

Ahora desplazándonos hacia otras partes de España, pero no menos importantes, tenemos a la antítesis de los catalanes, digo, no..., verdad que habíamos quedado que eran "chéveres", me he equivocado; corrijo, en igual línea que los catalanes, majísimos, tenemos a los madrileños Montse, Paula y Fernando, al pompeu fabradino Nacho, y a la burgalesa Sara, a los cuales les quiero agradecer por los buenos momentos, por ser divertidos y por lo siguiente. A Montse por ser tan buena gente, pero buena de las de verdad, te agradezco tu preocupación hacia a mí y tu ayuda en todo momento. Me encanta que eres una buena persona, sincera, y estas siempre dispuesta a ayudar a los demás. Además, tú y tu Xavi, me han salvado en más de una ocasión de los problemas informáticos que se me presentaron durante la escritura de la tesis, así que muchas gracias a los dos por esto. A Pau por lo maja, por las conversaciones e indignaciones compartidas ;), por los "cinco minutos mas", por su sencillez, y por su buen rollo hacia los latinoamericanos je je, se nota que si sabes apreciar la calidad humana ;) . A Nacho por acompañarme en mi primer cumple en Barcelona, por irse de fiesta conmigo, por presentarme personas súper agradables como Ricardo, al cual recuerdo con muchísimo cariño, por ayudarme a conseguir piso y preocuparse por mí en general intentando ayudar. Y tranquilo Nacho que sí sé que eres de Ponferrada. A Sari por ser tan agradable y hacernos la vida mas

alegre, por entretener a (...) y darle alegrías, y por ayudarme cuando lo he necesitado. A Fernando por las risas, las conversaciones, y las cervecitas compartidas junto a los demás.

Por otro lado quiero darle las gracias especialmente a Juan Pablo, un ejemplo de científico y de persona. Gracias por tu forma de ser, por las cervecitas y las hamburguesas compartidas después del lab, y las charlas de vuelta a casa. Por las conversaciones sobre "los grandes misterios de este y otros mundos", por llamarlo de alguna forma, y por los momentos divertidos. También quiero agradecer a mi querida Reicy, cuasi compatriota. Eso que te digo es muy lindo, pues no a cualquiera le doy el honor de llamarse colombiano, bueno cuasi. Y pues es que sí, mas cercana, querida, agradable y chistosa no puedes ser, gracias por eso, que me hace sentir como si nos conociéramos desde hace mucho, y por los ánimos que me diste para terminar la tesis.

Quiero agradecer también a Maribel, a Saska y a Susana Barros, por los buenos momentos, las conversaciones y las risas, con Saska sobretodo en esas horas en las que solo andan por ahí los guardias de seguridad. También te agradezco mucho que te hayas preocupado por mí y que me hayas dicho lo del curso de cristalografía, que estuvo bastante bien. A ti Maribel te agradezco por tu apoyo y compañía durante los primeros años, por haber confiado en mí, y por haber sido de las primeras personas en las que confié en el laboratorio.

También he tenido la oportunidad de compartir buenos momentos con la gente nueva del lab, aunque desafortunadamente, el tiempo de conocerlos ha sido muy corto. Sin embargo, quiero decirles a los nuevos estudiantes de Albert, Bas Olde y Elena Vazquez, que se les nota por encima lo majos que son, ya hemos tenido la oportunidad de reír y trabajar un poco juntos, así que les agradezco por trabajar en la continuación de los estudios de esta tesis, y quiero decirles que están en buenas manos con Albert, seguro que aprenderán y se divertirán mucho. Él también necesita cuidados especiales ;) , así que a cuidarlo mucho. A Fernanda, minha querida amiga, quiero decirle que me habría encantado conocerla mas, y que espero que podamos hacerlo, pues me pareces una persona súper agradable, tranquila, buena gente y divertida, y seguro que si te conozco mejor descubriré muchísimas mas cualidades. Han sido muy divertidas nuestras conversaciones, y me encanta practicar mi portugués contigo.

Quiero agradecer a Willi Largo, que fue un gran amigo durante esta dura etapa. Fue una alegría haberme encontrado contigo en Barcelona, fuiste el responsable de que me empezara a divertir aquí sin parar. También quiero agradecer a mis amigas en Colombia que me subían el ánimo cuando hablábamos por chat, o cuando finalmente nos podíamos encontrar. No necesito nombrarlas, ustedes saben quiénes son. Siempre que nos veíamos era como si nunca me hubiese ido. Nunca he dejado de extrañarlas.

Por supuesto también quiero agradecerles a los amores más grandes de mi vida:

A mi familia, para la cual no tengo palabras de agradecimiento suficientes. A mi mamá y a mi papá les doy gracias enormes, pues no solo son responsables de que haya podido realizar esta tesis, apoyando siempre las decisiones que consideran mejores para mi hermana y para mí, y haciendo hasta lo imposible para que podamos llevarlas a cabo, sino porque son responsables de todo lo que soy, y creo que no he salido tan mal, je je. Siempre me han dado todo lo necesario, sobretodo amor incondicional, y sé que son las personas que estarán para, y por mí, siempre que lo necesite. Papis, son las mejores personas que conozco, siempre interesados por el bienestar de los demás, generosos, honestos, cariñosos, trabajadores y fuertes. Gracias por darnos un excelente ejemplo y motivarnos a superarnos. Por eso he terminado aquí haciendo

esta tesis. Quiero agradecer a mi hermana, que si hablamos de buenos ejemplos, me ha dado el mejor, a pesar de que se supone que debería ser al revés. Sandris, le agradezco enormemente por ser un ejemplo de responsabilidad y esfuerzo, y sobre todo por siempre preocuparse por mí y ayudarme cuando lo he necesitado, además de por las risas y buenos momentos que hemos compartido desde pequeñas.

A ti Mike, te agradezco por haberme cambiado la vida a muchísimo mejor, contigo me he sentido como en casa, como no me sentía desde hace mucho tiempo; desde que llegue aquí, me has dado tranquilidad y alegría. Gracias por apoyarme y ayudarme durante todo este tiempo, por esperarme hasta muy tarde a que saliera del laboratorio y recogerme, por escuchar mis seminarios aunque no tuvieran nada que ver con lo que haces, por escuchar mis problemas, consolarme e intentar resolverlos.

Para terminar quiero agradecer muy especialmente:

A todos mis familiares, tíos, tías, primos y primas de todos los órdenes, y a mis abuelos, gracias por ser tan especiales, forman parte de lo que soy, y han hecho que no los deje de extrañar.

A la familia de Mike por recibirme y tratarme como a alguien de la familia, por preocuparse por nosotros, por cuidarnos, y por lo bien que nos la pasamos cuando nos encontramos.

A Dios, por haberme permitido estar en donde debía estar, por ayudarme siempre, por haberme inspirado, y por haberme dado fe y la fortaleza necesaria para seguir a pesar de las dificultades.

The main body of the present thesis is focused on studies to find a highly stable metazoan Heteromeric Amino acid Transporter (HAT) suitable for structural biology, and the optimal conditions for its production at enough quality for such studies. Nevertheless, other structural studies were performed during the first years of my thesis. In them, crystallographic studies of the arginine/agmatine exchanger AdiC were performed to obtain the structure of this transporter in inward-facing conformations. AdiC is the closest prokaryote homologue of the catalytic light subunits of HATs with atomic structure solved. However, AdiC structure has been solved in outward-facing conformations, leaving unknown the complete mechanism of transport cycle, like for instance, substrate translocation after the extracellular substrate induced fitting, and the recognition of substrates from the intracellular side. For this reason, we tried to solve the structure of inward-facing conformations of AdiC. Although we were able to obtain crystals of AdiC mutants, designed to stabilize inward-facing conformations, that diffracted until 3.5 Å resolution, diffraction quality was not adequate to solve the structure of these mutants. Moreover, the cell dimensions of these crystals led to suggest that these mutants could be crystalizing in outward-facing conformation. Thus, for this reason, and due to the risk of not obtain new structural data after a probable long time of crystals optimization, in agreement with Dr. Manuel Palacín and Dr. Albert Rosell, we decided to pursue other goals that could offer additional structural knowledge about HATs, which are developed in my thesis. However, as a consequence of some promising results in our studies with AdiC, and the deep analysis generated during those years about how to proceed to stabilize AdiC inward-facing conformations to solve these structures, we wanted to summarize and discuss these data in the Appendix section entitled "Towards the molecular mechanism of the Arg/Agm exchanger AdiC".

Contents

List of Figures

List of tables

Abbreviations

4F2hc 4F2 heavy chain

4F2hc-ED 4F2hc-Ectodomain

APA amino acid/polyamine antiporter

APC transporters amino acid, polyamine and organocation transporters

asc sytem asc amino acid transport system

ATP adenosine-5'-triphosphate

BCA bicinchoninic acid protein assay

bp base pairs

BSA bovine serum albumin

cAMP 3'-5'-cyclic adenosine monophosphate

CATs Cationic Amino acid Transporters

cDNA DNA copy

CHS cholesteryl hemisuccinate Tris salt

CMC critical micellar concentration

CoIP coimmunoprecipitation

CRAC cholesterol recognition / interaction amino acid consensus

Cryo-EM cryo-electron microscopy

Cymal-6 6-cyclohexyl-1-hexyl-β-D-maltoside

DAT Drosophila Dopamine transporter

DDM n-Dodecyl-β-D-maltoside

dH2O deionized H2O

DM n-Decyl-b-D-maltoside

DMEM Dulbecco's modified Eagle's medium

DMPC 1,2-dimyristoyl-sn-glycero-3-phosphocholine

DMPE 2-dimyristoyl-sn-glycero-3-phosphoethanolamine

DMSO dimethyl sulfoxide

DPPC 1,2-dipalmitoyl-sn-glycero-3-phosphocholine

DNA deoxyribonucleic acid

dNTP deoxyribonucleotide

DTT dithiothreitol

EDTA ethylenediaminetetraacetic acid

ER endoplasmic reticulum

FBS fetal bovine serum

FPLC fast protein liquid chromatography

FSEC fluorescent size exclusion chromatography

g gravitational-force

GABA glutamate/c-aminobutyric acid

GFP green fluorescent protein

GSH reduced tripeptide glutathione

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

HMMTOP Hidden Markov Model for Topology Prediction

HRP horseradish peroxidase

HUGO Human Genome Organization

Ig immunoglobulin

IMAC immobilized metal ion affinity chromatography

IMPs Integral Membrane Proteins

IPTG isopropyl β-D-1-thiogalactopyranoside

kb kilobase

kDa kiloDalton

Ki inhibiton constant

Km Michaelis–Menten constant

KSHV Kaposi's sarcoma herpesvirus

LAT L- amino acid transporter

LB medium Luria Broth medium

LDAO Lauryldimethyl amine oxide

LPI lysinuric protein intolerance

LSB Laemmli sample buffer

mAb monoclonal antibody

MOI Multiplicity of infection

mRNA messenger RNA

OD optical density

OMIM online Mendelian inheritance in man

ORF Open Reading Frame

P1 and P2 Population 1 and Population 2 of viruses

PBS phosphate buffer saline

PCR polymerase chain reaction

PDB protein data bank

PEG polyethylene glycol

PEI polyethylenimine

PET Positron Emission Tomography

POPC 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine

POPE 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine

POPG 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)

PS projection structure

rBAT related to basic amino acid transporter

RFU: Relative Fluorescence Units

RNA ribonucleic acid

RT room temperature

SDS sodium dodecyl sulfate

SDS-PAGE SDS-polyacrylamide gel electrophoresis

SEC size exclusion chromatography

SLC Solute carrier

TCDB the transporter classification database

TEM transmission electron microscopy

TIM Triose phosphate isomerase

TMD transmembrane domain

Tris tris(hydroxymethyl)aminomethane

UDS Ultracentrifugation Dispersity Sedimentation Assay

WPRE motif Woodchuck Hepatitis Virus Post-transcriptional Regulatory Element

WT wild type

xCT system xC- amino acid transporter

y+LAT system y+L amino acid transporter

YPD Yeast extract peptone dextrose media

YSB Yeast suspension buffer

1. Introduction

1.1 Transport across cellular membranes

In multicellular organisms, the plasma membrane that envelops the cytoplasm with all the relevant organelles and molecules for life, provides structure and protection to the cell, and is involved in the maintenance of the cellular homeostasis. Moreover, is the gate for the uptake and efflux of compounds essential for nutrition, cellular growing, metabolism and signal transduction inside the cell, and from cell to cell in all the organism. Membrane organelles are also involved in the same function than the plasma membrane does, but for regulating the internal conditions in the cells, the communication between cellular compartments, and the transport of essential compounds to the place in which they are needed.

Passive diffusion is the simplest mechanism by which molecules move across the cellular membranes. In this process, molecules able to dissolve in the phospholipid bilayer, diffuse across the membrane for being dissolved at the other side of the membrane in the aqueous solution, and thus, equilibrating the inside and outside concentrations of molecules (Cooper 2000). In this way, small and relatively hydrophobic molecules can diffuse at substantial rates across the phospholipid bilayers. For instance, hydrophobic molecules, like benzene, gases, like O_2 and CO_2, and small polar uncharged molecules, like ethanol, crosses by passive diffusion the cellular membranes (Figure 1), moving from a compartment with a high concentration to other with a lower concentration of the molecule (Cooper 2000). On the other hand, other physiologically relevant larger uncharged polar or charged compounds are unable to cross cellular membranes by passive diffusion, and then, transport proteins are required for their selective pass through the phospholipid bilayers (Figure 1). In some instance, water also needs transport proteins to ensure sufficient flux across biological membranes (Figure 2).

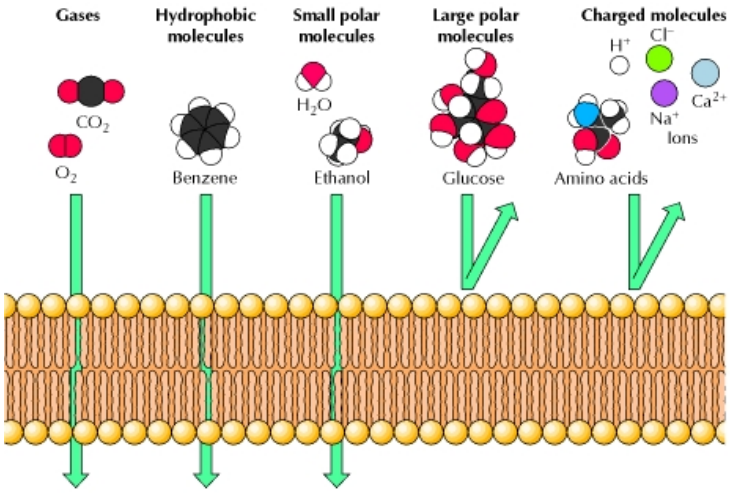

Figure 1. Transport across the cellular membranes by passive diffusion. Gases, hydrophobic molecules, and small polar uncharged molecules can diffuse through cellular membranes, but larger polar molecules and charged molecules cannot. (Cooper 2000).

In higher eukaryotes such as mammals, sugars, amino acids, nucleotides, inorganic ions and drugs, are transported across the cellular membranes in processes regulated by integral membrane proteins (IMPs), from the solute carrier (SLC) transporters and non-SLC transporters series (Figure 2). Among them, transport proteins are most widely considered as either channels or transporters. Channels drive the movement of solutes down their

electrochemical gradients, similar to transporters do, but channels and transporters differentiate between them by their transport mechanism. First, in channels, relatively little energy is required for the interaction between the channel and the transported ion, and second, the major conformational changes are present in the domains that act as "gates", determining an open or close- state, and are induced by extrinsic factors, like changes in membrane potential or the binding of small regulatory molecules. (Dubyak 2004). In contrast, in transporters, the conformational changes required for substrate translocation, are induced by the specific interaction of the substrate with defined residues of the transporter, and generally involve the movement of domains extended along the protein. Thus, the rate of transport through transporters is generally many times slower than that of transport through channels (Dubyak 2004).

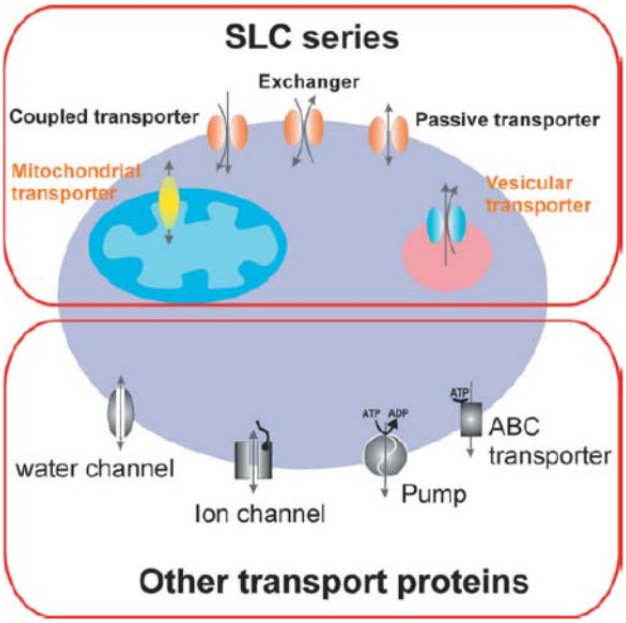

Figure 2. Illustrated representation of solute carrier (SLC) and non-SLC transporters series. Both types of transporters are expressed in the plasma membrane, and in intracellular compartments membranes. (Hediger *et al.* 2004).

Transporters are classified in passive or facilitated transporters, and active transporters (Hediger *et al.* 2004). Passive transporters facilitate crossing through the membranes of molecules, like glucose or amino acids, down their electrochemical gradients. On the other hand, active transporters use mechanisms that require energy to generate ion/solute gradients across membranes. At the same time, they are divided in primary- or secondary-active transporters, depending on the direct or indirect use of the cellular energy contained in the ATP molecule for transport. Primary-active transporters are dependent on the binding to the ATP, or on the ATP hydrolysis for their function, like members of the ATP-binding cassette (ABC) transporter family and ion pumps. Thus, mammalian ABC transporters bind or hydrolyze ATP for the transport of substances such as lipids, carbohydrates, ions, xenobiotics and drugs out of cells or into cellular compartments (Borst and Elferink 2002). Additionally, to pump ions, like Na^+, K^+, H^+, Ca^{2+} or Cu^{2+}, out of cells or into organelles, ion pumps hydrolyze ATP, and create and sustain electrochemical ion gradients (Hediger *et al.* 2004). On the other hand, secondary-active transporters couple ions from the electrochemical gradient generated by ion pumps to enable transport of nutrients across cellular membranes.

1.2 Solute Carrier (SLC) Transporters

The SLC series includes families of passive transporters, ion coupled transporters and exchangers (Figure 2). These transporters are involved in the transport of nutrients relevant for the cell and organism survival, like vitamins, trace minerals, sugars and amino acids, and also ions for maintaining cellular homeostasis and membrane potential. As a consequence, in humans, mutations in these genes or impaired function of SLC transporters are involved in several diseases. The last revision about human SLC transporters (Hediger *et al.* 2013), indicates that 9 new gene families (a total of 100 new genes) have been identified between 2004 and 2013, being recognize nowadays 52 families of these transporters (Table 1). Generally, members in one family encode proteins with at least 20% amino acid sequence identity between them. In general, the genes names are the root symbol, SLC, followed by the family number and the letter that indicates the subfamily, thus, the last number refers to the specific transporter.

The HGNC Solute Carrier Family Series	Total 2004	Total 2013
SLC1: The high affinity glutamate and neutral amino acid transporter family	7	7
SLC2: The facilitative GLUT transporter family	14	14
SLC3: The heavy subunits of the heteromeric amino acid transporters	2	2
SLC4: The bicarbonate transporter family	10	10
SLC5: The sodium glucose cotransporter family	8	12
SLC6: The sodium- and chloride-dependent neurotransmitter transporter family	16	21
SLC7: The cationic amino acid transporter/glycoprotein-associated amino-acid transporter family	14	14
SLC8: The Na^+/Ca^{2+} exchanger family	3	3
SLC9: The Na^+/ H^+ exchanger family	8	13
SLC10: The sodium bile salt cotransport family	6	7
SLC11: The proton coupled metal ion transporter family	2	2
SLC12: The electroneutral cation-Cl cotransporter family	9	9
SLC13: The human Na^+-sulfate/carboxylate cotransporter family	5	5
SLC14: The urea transporter family	2	2
SLC15: The proton oligopeptide cotransporter family	4	5
SLC16: The monocarboxylate transporter family	14	14
SLC17: The vesicular glutamate transporter family	8	9
SLC18: The vesicular amine transporter family	3	4
SLC19: The folate/thiamine transporter family	3	3
SLC20: The type III Na^+-phosphate cotransporter family	2	2
SLC21/SLCO: The organic anion transporting family	11	12
SLC22: The organic cation/anion/zwitterion transporter family	18	23
SLC23: The Na^+-dependent ascorbic acid transporter family	4	4
SLC24: The $Na^+/(Ca^{2+}–K^+)$ exchanger family	5	6
SLC25: The mitochondrial carrier family	27	53
SLC26: The multifunctional anion exchanger family	10	11
SLC27: The fatty acid transport protein family	6	6
SLC28: The Na^+-coupled nucleoside transport family	3	3
SLC29: The facilitative nucleoside transporter family	4	4
SLC30: The zinc efflux family	9	10
SLC31: The copper transporter family	2	2
SLC32: The vesicular inhibitory amino acid transporter family	1	1
SLC33: The Acetyl-CoA transporter family	1	1
SLC34: The type II Na^+-phosphate cotransporter family	3	3
SLC35: The nucleoside-sugar transporter family	17	30
SLC36: The proton-coupled amino acid transporter family	4	4
SLC37: The sugar-phosphate/phosphate exchanger family	4	4
SLC38: The System A & N, sodium-coupled neutral amino acid transporter family	6	11
SLC39: The metal ion transporter family	14	14
SLC40: The basolateral iron transporter family	1	1
SLC41: The MgtE-like magnesium transporter family	3	3
SLC42: The Rh ammonium transporter family (pending)	3	3
SLC43: Na^+-independent, system-L like amino acid transporter family	2	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	298	395

Table 1. Identified families of human SLC transporters. Human SLC transporters classified by the Human Genome Organization (HUGO) gene nomenclature committee. (Hediger *et al.* 2013).

1.3 Amino acid/Polyamine/Organocation (APC) Superfamily of Transporters

At the same time, SLC transporters families have been grouped into superfamilies together with prokaryotic transporters, based on their evolutionary relationship and substrate specificity. One of those superfamilies is the amino acid/polyamine/organocation superfamily (APC), which includes solute:cation symporters and solute:solute antiporters (Isnard *et al.* 1996; Kashiwagi *et al.* 1997; Deves and Boyd 1998). Members of the APC superfamily together with other transporters have been classified in families and subfamilies based on the homology amongst proteins that shares common ancestors (Jack *et al.* 2000; Wong *et al.* 2012). By the moment, eleven families have been identified (Table 2), as recorded in the Transporter Classification Database (TCDB; www.tcdb.org). Almost half of them are amino acid transporters, such as the amino acid/polyamine/organocation (APC) family, the amino acid/auxin permease family (AAAP), the alanine or glycine:cation symporter family (AGCS) and the hydroxy/aromatic amino acid permease (HAAAP) family (Wong *et al.* 2012). Different human amino acid transporters, members of SLC families, belong to the APC superfamily, such as, SLC7: the cationic amino acid transporter/glycoprotein-associated amino-acid transporter family (light subunits of heteromeric amino acid transporters, explained later); SLC12: the electroneutral cation-Cl cotransporter family; SLC32: the vesicular inhibitory amino acid transporter family; SLC36: the proton coupled amino acid transporter family; or the SLC38: the System A & N sodium-coupled neutral amino acid transporter family.

2.A.3	APC family
2.A.15	BCCT family
2.A.18	AAAP family
2.A.21	SSS family
2.A.22	NSS family
2.A.25	AGCS family
2.A.30	CCC family
2.A.39	NCS1 family
2.A.40	NCS2 family
2.A.42	HAAAP family
2.A.53	SulP family

Description, protein members and references of these families can be found in TCDB (www.TCDB.org).

Table 2. Families belonging to the APC superfamily of transporters. First column indicates the Transporter Classification number (TC), and the second column the family name. **APC:** amino acid/polyamine/organocation. **BCCT:** betaine/carnitine/choline. **AAAP:** amino acid/auxin permease. **SSS:** solute:sodium symporter. **NSS:** neurotransmitter:sodium symporter. **AGCS:** alanine or glycine:cation symporter. **CCC:** cationchloride cotransporter. **NCS1:** nucleobase:cation symporter-1. **NCS2:** nucleobase:cation symporter-2. **HAAAP:** hydroxy/aromatic amino acid permease. **SulP:** sulfate permease. (Wong *et al.* 2012).

1.4 Amino acid Transport

Amino acids are the principal components of proteins, which structure cells and organisms and are involved in all the relevant processes that take place for sustention of life, such as cellular division and differentiation, protein synthesis, metabolism, signal transduction, neural transmission and cellular growth. Additionally, amino acids are also relevant not as constituent of all proteins in the organism, but they are also directly involved in some processes acting as signaling molecules, energy fuels or neurotransmitters, such as glutamate, glycine or aspartate

(Broer and Palacin 2011). In humans, some amino acids can be synthetized by the organism, but other, known as essential amino acids, cannot, and they must be obtained from the diet. Thus, during digestion, proteins from food are degraded into peptides, that are absorbed in the small intestine and hydrolyzed into amino acids, which are delivered in the blood stream to reach all tissues. When blood is filtered by the kidney, amino acids, together with other relevant metabolic compounds, are reabsorbed by the organism.

During the whole process, from the feeding until absorption in the intestine and kidney reabsorption, and from here to the final destination in the organism, amino acids cross cellular membranes helped by amino acid transporters. They mediate the transfer of amino acids between organs, between cells, and between cellular compartments (Muth and Schuldiner 2000). Amino acid transport can be carried out simultaneously with ions in a co- or counter-transport in varied stoichiometries (Jack *et al.* 2000). Eleven families of SLC transporters (Table 1) are involved in the amino acid transport. Some transporters in these families share amino acid specificity , in fact, unrelated evolutionarily families, such as amino acid transporter families SLC1 and SLC7 have chemically similar substrates (Schlessinger *et al.* 2013). However, amino acid transporters are differentially expressed in organs, and some are more abundant in determined tissues than in other. Thus, when a specific transporter is mutated or is not expressed, its function is impaired in specific tissues or in the whole body, which lead to development of symptoms and inherited and acquired diseases, as a consequence of the variety of physiological processes in which amino acid transport is required (Broer and Palacin 2011). For instance, normal neuronal excitability can be negatively affected in brain or, regular homoeostasis can be hindered in the whole body, causing malabsorption of amino acids and renal complications. On the other hand, overexpression of some of these transporters have been involved in tumor progression (Broer and Palacin 2011). In table 3, the families of amino acid transporters are shown with their main characteristics, as well as the diseases they have been related to. However, despite most of the amino acid transporters have been identified and characterized, some remain orphan, as in the case of SLC25 and SLC38 families (Broer and Palacin 2011).

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 3. Human amino acid transporters families. Substrates are given in one-letter code. Cit, citrulline; Cn, cystine; O, ornithine. The 'Function' column includes references to amino acid transport systems. These systems have acronyms indicating the substrate specificity of the transporter. 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− indicates transporters for anionic amino acids (as in X− AG and x− c). The subscript AG indicates that the transporter accepts aspartate and glutamate, and the subscript c indicates that the transporter also accepts cystine. Y+ or y+ refer to transporters for cationic amino acids (an Na+-dependent cationic amino acid transporter has not been unambiguously defined and as a result Y+ is not used), B or b refers to amino acid transporters of broad specificity with superscript 0 indicating a transporter accepting neutral amino acids and superscript + indicating a transporter for cationic amino acids. T stands for a transporter for aromatic amino acids, and system N indicates selectivity for amino acids with nitrogen atoms in the side chain. In the remaining cases, the preferred substrate is indicated by the one-letter code for amino acids. For example, system L refers to a leucine-preferring transporter and system ASC to a transporter preferring alanine, serine and cysteine. Proline and hydroxyproline are referred to as imino acids. Owing to historic idiosyncrasies, the nomenclature for plasma-membrane amino acid transport systems is not completely consistent, but is widely used in the field. AAT, amino acid transporter. Substrates, function system and disease involvement are shown. (Broer and Palacin 2011).

1.5 Heteromeric Amino acid Transporters (HATs)

HATs are the unique example known in all kingdoms of life of solute transporters liking two different protein chains by a conserved disulfide bridge. These heterodimers consist of a heavy subunit or heavy chain (hc), from SLC3 family (4F2hc (SLC3A2) or rBAT (SLC3A1)), and a light subunit of the L-Amino acid Transporters (LATs) from SLC7 family (Figure 3). Heavy subunits traffic the holotransporter to the membrane, and LATs are on charge of the amino acid transport activity (Estevez *et al.* 1998; Pfeiffer *et al.* 1998; Deves and Boyd 2000; Palacin and Kanai 2004).

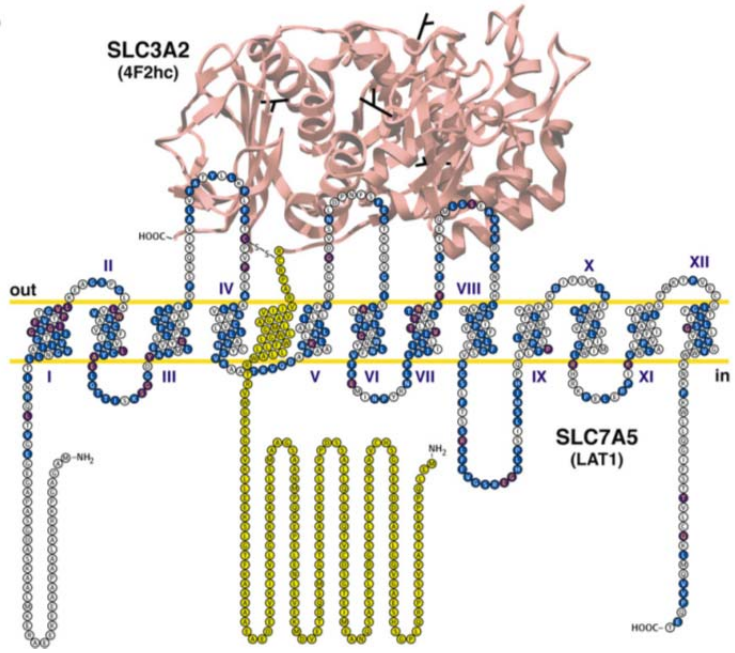

Figure 3. Representation of the topology of HATs. Model of human 4F2hc/LAT1, as a representative member of HATs, is shown. Light subunits of HATs (LATs) are associated with the heavy subunit, 4F2hc or rBAT, through a conserved disulfide bridge. Heavy subunits are type II membrane *N*-glycoproteins with an intracellular N-terminus and one transmembrane domain (TMD) (showed in yellow), and a bulky ectodomainh. The four glycosylation sites of 4F2hc are indicated in black in the crystal structure of 4F2hc ectodomain, shown as a pink cartoon (Fort *et al.* 2007). LATs are not glycosylated proteins, and consist of 12 TMDs (indicated in roman numerals) and intracellular N- and C- terminal ends. (Fotiadis *et al.* 2013).

The two characterized heavy subunits (rBAT and 4F2hc) are type II membrane *N*-glycoproteins with a single transmembrane domain (TMD), an intracellular N-terminus and an extracellular ectodomain (ED) (Figure 3). LATs are polytopic not glycosylated membrane proteins, belonging to the APC superfamily of transporters, which metazoan members are predicted to exhibit 12 α-helical TMDs, and cytosolic N- and C-terminal ends. This prediction has been supported by topology experiments of prokaryotic homologues from the APC superfamily (Ellis *et al.* 1995; Cosgriff and Pittard 1997; Hu and King 1998), and by the crystal structure of some of them, such as the Arg/Agm exchanger AdiC (Fang *et al.* 2009, Gao *et al.* 2009, Gao *et al.* 2010, Kowalczyk *et al.* 2011), the glutamate-GABA antiporter GadC (Ma *et al.* 2012) and the broad-specificity amino acid transporter ApcT (Shaffer *et al.* 2009). In fact, studies of one LAT, xCT, have determined that this transporter exhibit 12 transmembrane segments (Gasol *et al.* 2004), compatible with the topology of AdiC, GadC and ApcT.

Disulfide bridge, connecting both subunits in HATs, is established between highly conserved extracellular cysteines located in the putative extracellular loop 2 of light subunits, and a few residues away from the TMD in the heavy subunit (Figure 3). Interestingly, when these cysteines are mutated, the interaction between heavy and light subunits is not disrupted, suggesting that other non-covalent interactions are established between both proteins (Estevez *et al.* 1998; Pfeiffer *et al.* 1998).

Ten light subunits have been identified by the moment, six of them are associated to 4F2hc (LAT1, LAT2, y+LAT1, y+LAT2, asc1, and xCT) and two, AGT-1 and $b^{0,+}$AT, are associated with rBAT (Fotiadis *et al.* 2013; Nagamori *et al.* 2016). The remaining two light subunits are asc-2 and ArpAT, for which associated heavy subunit has not been identified yet (Table 4). Heterodimer formation between heavy and light subunits determines the normal functioning of each HAT in the cellular membranes. Thus, loss of function of heavy subunit does not allow the membrane localization of the transporter, hampering the amino acid transport across such membranes, as well as it does mutations in light subunits that leads to transport malfunction or depletion of LATs expression. Due to LATs are the catalytic subunits (Reig *et al.* 2002), the substrate selectivity is determined by these proteins, then being involved in different systems of transport, and as a consequence, being also related with different diseases because of its dysfunction or overexpression in some tissues (Table 4).

1.5.1 rBAT and its light subunits

rBAT was the first heavy subunit characterized, which expression in Xenopus oocytes induced a Na^+-independent uptake of dibasic and neutral amino acids (Bertran *et al.* 1992; Tate *et al.* 1992; Wells and Hediger 1992). Two light subunits have been found to associate with rBAT, $b^{0,+}$AT and recently AGT-1 (Nagamori *et al.* 2016).

1.5.1.1 rBAT/$b^{0,+}$AT ($b^{0,+}$ transport system)

It is expressed in the apical side of the plasma membranes of epithelial cells, mainly in the small intestine and in the renal proximal tubule. This transporter mediates 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). Direction of the substrate exchange is driven by the high intracellular concentration of neutral amino acids, the low intracellular concentration of cystine, and the electric potential across the plasma membrane. Co-expression of $b^{0,+}$AT and rBAT in cells is required for system $b^{0,+}$ activity. rBAT and $b^{0,+}$AT are interdependent for trafficking from the endoplasmic reticulum (ER) to the plasma membrane (Bartoccioni *et al.* 2008). Assembly with $b^{0,+}$AT prevents ER-associated degradation (ERAD) of unassembled rBAT and is required for the oxidative folding of the rBAT ectodomain (Rius and Chillaron 2012). Recently, it has been reported that the *N*-glycan N575 of human rBAT is indispensable for early traffic, and the C-terminal loop of this heavy subunit relevant for stability and/or early traffic of rBAT–$b^{0,+}$AT (Rius *et al.* 2016). Some evidence suggests that, in native tissues, rBAT/$b^{0,+}$AT dimerizes forming a functional tetramer of heterodimers, although the rBAT/$b^{0,+}$AT heterodimer is the functional unit (Fernandez *et al.* 2006). In this complex, two $b^{0,+}$AT molecules homodimerize within the heterotetramer, as it has been demonstrated by fluorescence resonance energy transfer studies and crosslinking assays (Fernandez *et al.* 2006). The rBAT-$b^{0,+}$AT oligomerization process might be dictated by the heavy subunit rBAT and not by $b^{0,+}$AT. This possibility was suggested by

experiments performed using the functional chimera formed by rBAT and the 4F2hc-associated light subunit xCT, where this transporter retained the capability to form heterotetramers (Fernandez *et al.* 2006).

Heavy subunit	Associated light subunit	Gene	Human Chromosome	Transport system	Predominant Substrates/function	Disease linkage
4F2hc		SLC3A2	11q13		Traffic of the holotransporter to the membrane.	Overexpressed in tumors
	LAT1	SLC7A5	16q24.3	L	Transport of neutral L-AA essential for cellular growing through mTOR signaling. Transport of T3, T4, L-DOPA, BCH	Overexpressed in tumors
	LAT2	SLC7A8	14q11.2	L	Neutral L-AA, T3, T4, BCH	
	y+LAT1	SLC7A7	14q11.2	y+L	Na+ independent:cationic AA Na+/large neutral L-AA	Lysinuric protein intolerance (LPI)
	y+LAT2	SLC7A6	16q22.1	y+L	Na+ independent: cationic AA Na+/large neutral amino acid	
	asc1	SLC7A10	19q13.1	Asc	Small neutral AA D-Ser transport in CNS	
	xCT	SLC7A11	4q28.3	xc-	Cystine (anionic form) L-glutamate Control cell redox balance	Overexpressed in tumors. Kaposi Sarcoma Herpes Virus receptor. Cocaine relapse.
rBAT		SLC3A1			Traffic of the holotransporter to the membrane	Type A Cystinuria
	$b^{0,+}$AT	SLC7A9	19q13.1	$b^{0,+}$	Neutral/dibasic AA	Type B Cystinuria
	AGT-1	SLCA13	8q21.3		L-Aspartate and L-glutamate	
Unknown	**asc-2**	Slc7a12		Asc		
	ArpAT	Slc7a15	Silenced in primates		Aromatic-preferring AA	

Table 4. Function and disease involvement of heavy and light subunits of HATs. Summary of the characteristics of HATs as indicated in each column. AA: amino acids. Adapted from the doctoral thesis (Álvarez-Marimon. 2014).

rBAT/$b^{0,+}$AT in cystinuria

The main apical reabsorption system for cystine in kidney is the transport activity of rBAT/$b^{0,+}$AT (Fernandez *et al.* 2002). On the extracellular side, rBAT/$b^{0,+}$AT has a high affinity for cationic amino acids and cystine (Km ~100 μM), while on the intracellular side the affinity for neutral amino acids is lower, in the milimolar range (Reig *et al.* 2002). Mutations in rBAT or $b^{0,+}$AT cause cystinuria (Online Mendelian Inheritance in Man (OMIM) database, entry 220100), which is the most common primary inherited aminoaciduria (Calonge *et al.* 1994; Feliubadalo *et al.* 1999). Cystinuria produces hyperexcretion of dibasic amino acids and cystine in urine, due to the failure in the reabsorption of filtered cystine and dibasic amino acids in the proximal tubule. One of the complications caused by cystinuria is the formation of renal calculi, known as urolithiasis, due to the high concentration of the hyperexcreted amino acid and the poor solubility of cystine in the urinary tract. This, can cause obstruction, infection and ultimately chronic kidney disease (Chillaron *et al.* 2010). Cystinuria is responsible for 1–2% of all cases of renal calculi formation in adults and 6–8% in pediatric patients (Milliner and Murphy 1993).

Three types of cystinuria have been described according to the protein affected; type A cystinuria, caused by mutations in rBAT gene (SLC3A1), of which 133 have identified so far, type B cystinuria, caused by mutations in b^{0+}AT gene (SLC7A9), of which 95 have been reported by now, and type AB cystinuria, in which both genes are mutated (Chillaron *et al.* 2010). In the case of rBAT mutants, they show problems in the trafficking of the heterodimer to the plasmatic membrane (Bartoccioni *et al.* 2008), being distinguishable two different mechanisms underlying the trafficking defects. One makes the assembly of rBAT with $b^{0,+}$AT really slow, which results in very small amount of functional and mature glycosylated heterodimer, and other, could allow efficient assembly with $b^{0,+}$AT and disulfide-link formation between both subunits, but due to the remaining rBAT core-glycosylation, oligomerization is defective, and heterodimer is probably degraded via the proteasome. On the other hand, the studied $b^{0,+}$AT mutations causing cystinuria, could case no expression of the protein, defects in the trafficking to the plasma membrane, and defects in the transport function (Reig *et al.* 2002; Shigeta *et al.* 2006).

1.5.1.2 rBAT/AGT-1

AGT-1 (Aspartate/Glutamate transporter 1) is in reality a sodium-independent exchanger of anionic amino acids, with high affinities for L-aspartate, L-glutamate and L-cystine (Nagamori *et al.* 2016). It was cloned from a cDNA library of mouse kidney and characterized in Xenopus oocytes (Matsuo *et al.* 2002). AGT-1 shares 35-37 % amino acid identity with the other LATs, conserving the cysteine residue responsible for the disulphide bond between the heavy and light subunits, although the highest identity is exhibited with asc2 (48 %). Its localization is mainly in kidney, at the apical membrane of epithelial of the straight part of the proximal tubule (Nagamori *et al.* 2016). AGT-1 heterodimerizes with rBAT, as revealed by co-inmunoprecipitation studies and analyzing renal brush border membranes of a mutated mouse model of rBAT. Indeed, in the rBAT mutant mouse D140G (i.e., missense mutation aspartate 140 to glycine), expression in the apical membrane of the renal epithelial cells of both b^{0+}AT and AGT-1 is completely abolished (Nagamori *et al.* 2016). Under reducing conditions, AGT-1 appears as a band of 40 kDa, and under absence of such reducing conditions, is detected in a

band of 250 kDa, suggesting the formation of heterotetramers similar to rBAT/b^{0+}AT. Because rBAT/AGT-1 transport cysteine, this transporter is also a candidate transporter to be defective in cystinuria.

1.5.2 4F2hc and its light subunits

4F2hc was discovered as a surface antigen of activated lymphocytes, and identified as 4F2 antigen heavy chain (hc), or CD98, based on the systematic CD designation for antigens (Haynes *et al.* 1981; Quackenbush *et al.* 1987). Latter, 4F2hc was also named 'Fusion Regulatory Protein' (FRP-1) because of its function in cell fusion events that lead to multinucleated giant cells such as osteoclasts (Ohgimoto *et al.* 1996; Suga *et al.* 1997). In mammals, 4F2hc is almost ubiquitously expressed in cellular membranes through the whole body (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). Additionally to the trafficking of light subunits to the plasma membrane, 4F2hc seems to be directly involved in cell migration, fusion and proliferation by mediating β1- and β3-integrin signaling, although how this mediation occurs is by now unknown (Feral *et al.* 2005; Cantor *et al.* 2009; Fogelstrand *et al.* 2009). It has been suggested that the dual function of 4F2hc plays a key role in integrating integrin signaling and amino acid transport, but these relationship between both processes via 4F2hc has not been experimentally addressed (Broer and Palacin 2011). 4F2hc-associated light subunits carry out a variety of transport functions as described below.

1.5.2.1 4F2hc/LAT1

This HAT is a sodium-independent obligatory exchanger with a 1:1 stoichiometry. It transports mainly large and aromatic neutral amino acids, including leucine, isoleucine and tyrosine (Mastroberardino *et al.* 1998; Meier *et al.* 2002). It has been observed that the affinity of LAT1 by intracellular substrates is lower than for extracellular substrates, suggesting that the exchange velocity of 4F2hc/LAT1 is regulated by the intracellular amino acid concentration (Meier *et al.* 2002). Additionally, L-leucine and L-isoleucine are relatively better efflux than influx substrates, which would indicate that LAT1 uses intracellular amino acids accumulated by other transporters (Meier *et al.* 2002). In human tissues, mRNA of SLC7A5 (LAT1) has been found expressed the most in placenta, brain, spleen, testes, ovary, colon and bone marrow (Kanai *et al.* 1998; Segawa *et al.* 1999). Thyroid hormone is transported through this transporter (Kinne *et al.* 2011), thus LAT1 supplies thyroid hormones and amino acids to the developing fetus in the placental membrane (Ritchie and Taylor 2001). Additionally, it participates in the transport of L-3,4-dihydroxyphenylalanine (L-DOPA) across the blood–brain barrier (Kageyama *et al.* 2000), and in the transport of L-leucine across the inner blood-retinal barrier (Tomi *et al.* 2005).

4F2hc/LAT1 and cancer

Like in normal cells, tumor cells require an ample supply of oxygen and nutrients to proliferate, and glucose and amino acids are one of those relevant nutrients (Shchors and Evan 2007). 4F2hc/LAT1 has been found overexpressed in several types of primary human cancers and metastasis, including lung, colon, breast, prostate, head, neck and 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). This has been suggested like an adaptation strategy of tumor cells for obtaining essential amino acids for tumor metabolism, growth and survival. There has been found positive correlation between the overexpression of 4F2hc/LAT1 with the aggressiveness of tumors, cell proliferation, angiogenesis, and metastases (Fuchs and Bode 2005; Kaira *et al.* 2008; Kaira *et al.* 2009; Kaira *et al.* 2011).

Often, co-ordinated overexpression of neutral amino acid transporter 2 (ASCT2; SLC1A5) with 4F2hc/LAT1 is found in tumor cells (Fuchs and Bode 2005). 4F2hc/LAT1 exchanges large neutral amino acids with high affinity (Km ≤ 50 μM), and glutamine and asparagine with lower affinity (Km ~2 mM)(Yanagida *et al.* 2001). In contrast, ASCT2 has high affinity for small neutral amino acids, asparagine and glutamine (Km ~20 μM), and lower affinity for some large neutral amino acids, including some of the 4F2hc/LAT1 substrates (Broer *et al.* 1999; Broer *et al.* 2000). As a consequence, ASCT2 mediates net uptake of glutamine to drive cell growth and survival (Fuchs *et al.* 2007), and it has been proposed that in tumor cells LAT1 uses intracellular ASCT2 substrates to adjust the essential amino acid concentrations for metabolic demands, and mTOR (mammalian target of rapamycin) signaling (Fuchs and Bode 2005). mTOR is a serine/threonine protein kinase that controls essential cellular processes, such as cell growth, proliferation and survival (Sengupta *et al.* 2010). Transport of neutral branched amino acids, and specifically, leucine uptake via LAT1 in exchange for intracellular glutamine accumulated by ASCT2, activates mTOR signaling pathway, as demonstrated in cultured cells (Nicklin *et al.* 2009).

Due to the relevant role of LAT1 in tumor growth and metastasis, different anticancer therapies have been developed targeting this protein. Some of these drugs could be transported and internalized, and other only binds and inhibit the transport (del Amo *et al.* 2008). For instance, BCH (2-aminobicyclo-(2,2,1)-heptane-2-carboxylic acid), T3, and melphalan, have been reported to inhibit amino acid uptake in tumor cells (Mastroberardino *et al.* 1998; Kim *et al.* 2002; Uchino *et al.* 2002), inhibiting tumor growth in cultured cells (Kim *et al.* 2008). However, their growth inhibitory effects on transplanted tumors in nude mice have not been reported (Oda *et al.* 2010). Interestingly, a new molecule, KYT-0353, was found to inhibit [^{14}C]-leucine uptake and cellular growth in human colon cancer-derived HT-29 cells, in cells from mouse renal proximal tubule expressing human 4F2hc/LAT1, and also inhibited tumor growth in mice transplanted with HT-29 cells (Oda *et al.* 2010). The IC_{50} found to inhibit [^{14}C]-leucine uptake in HT-29 cells via 4F2hc/LAT1 was 0.06 μM, suggesting high affinity of KYT-0353 for human LAT1. These results point to this molecule as a promising anticancer drug targeting LAT1.

To image and evaluate the stage of tumors in vivo, tracers for positron emission tomography (PET), transported inside tumor cells by LAT1, have been investigated. Radiolabeled amino acids that are transport substrates for LAT1 have been mainly used for this purpose: tyrosine (L-[$^{3\text{-}18}$F]-α-methyl tyrosine), which is the most specific for LAT1 due to its α-methyl moiety (Kaira *et al.* 2007; Kaira *et al.* 2009; Wiriyasermkul *et al.* 2012), phenylalanine (L-p-(2-[^{18}F]fluoroethyl)-phenylalanine) (Wang *et al.* 2011), tryptophan (5-(2-18F-fluoroethoxy)-Ltryptophan) (Kramer *et al.* 2012) and methionine (S-(3-[^{18}F]fluoropropyl)-D-homocysteine) (Denoyer *et al.* 2012). However, limited success has been obtained for tumor imaging outside the brain (Ikotun *et al.* 2013). The main issue causing limited sensitivity by using these tracers, is the nature of amino acid exchanger of LAT1, which avoids concentrating substrates inside tumor cells directly. Then, other strategies, like the use of a zirconium-89 labeled monoclonal

antibody, that targets the extracellular domain of LAT1, showing specificity and high affinity for LAT1 in vitro and in vivo, are being developed (Ikotun *et al.* 2013).

1.5.2.2 4F2hc/LAT2

It is a sodium-independent obligatory exchanger, as 4F2hc/LAT1, but with high-affinity for a broad variety of neutral amino acids, including small ones (alanine, glycine, cysteine and serine), scarcely transported by LAT1 (Pineda *et al.* 1999; Rossier *et al.* 1999; Meier *et al.* 2002). 4F2hc/LAT2 is more ubiquitously expressed than LAT1, and is found in many tissues: prostate, testes, ovaries, kidney, placenta, brain, liver, spleen, skeletal muscle, heart, small intestine or lung (Pineda *et al.* 1999; Park *et al.* 2005). Moreover, 4F2hc/LAT2 localizes in the basolateral side in renal and intestinal epithelial cells (Rossier *et al.* 1999). Null knockout mice of Slc7a8 (LAT2) have not shown renal reabsorption defects of neutral amino acids, and cystine in particular (Braun *et al.* 2011). However, when it is heterologously expressed, efficiently exports L-cysteine, as also determined in cell culture models of renal epithelia (Bauch *et al.* 2003; Fernandez *et al.* 2003). In fact, Slc7a8 knockout mice have normal development and growth, which suggest a functional compensation by other amino acid transporters.

1.5.2.3 4F2hc/y+LAT1

This transporter mediates the efflux of a dibasic amino acid, in exchange for an extracellular neutral amino acid plus Na^+, with a stoichiometry of 1:1:1 (Broer and Palacin 2011). It is mainly expressed in kidney, lung, epithelial cells and circulating monocytes and macrophages, but also has been found in the small intestine, placenta and spleen (Torrents *et al.* 1998; Pfeiffer *et al.* 1999; Broer *et al.* 2000; Barilli *et al.* 2010). y+LAT1 is basolaterally located in the epithelial cells of the small intestine and in the renal proximal tubule (Torrents *et al.* 1998). The apparent affinity of 4F2hc/y^+LAT1 for substrates is lower at the intracellular binding site than at the extracellular (Kanai *et al.* 2000). It seems that in cells with active uptake of cationic amino acids, e.g., cells expressing rBAT/$b^{0,+}$AT or cationic amino acid transporters (CATs), 4F2hc/y+LAT1 mediates the export of cationic amino acids (Verrey *et al.* 2004).

4F2hc/y+LAT1 in Lysinuric Protein Intolerance (LPI)

Impairment of the y+LAT1 function, due to mutations, causes Lysinuric Protein Intolerance (LPI; OMIM database, entry 222700) (Borsani *et al.* 1999; Torrents *et al.* 1999). Polypeptide products of mutations shifting the ORF of SLC7A7 (y+LAT1 gene) have been found retained in the cytoplasm (Mykkanen *et al.* 2000; Toivonen *et al.* 2013), and some SLC7A7 point mutations have been identified to produce inactive y+LAT1 (Mykkanen *et al.* 2000). On the other hand, no mutations in 4F2hc (SLC3A2) have been identified to be involved in this pathology. 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 (Oyanagi and Nagao 1998; Sebastio *et al.* 2011). As a consequence, intestinal malabsorption of cationic amino acids causes urine hyperexcretion of these amino acids (Palacin *et al.* 2005), and thus, metabolic imbalances, characterized by low plasma levels of dibasic amino acids and dysfunction of the urea cycle. This fact leads to hyperammonaemia, orotic aciduria and protein aversion. Clinical manifestations of LPI appear after weaning with varied symptoms, like vomiting, diarrhea, failure in development, hepatosplenomegaly, bone-marrow and lung

abnormalities, osteoporosis, mental retardation, episodes of coma, altered immune response and chronic renal disease (Broer and Palacin 2011).

1.5.2.4 4F2hc/y+LAT2

Although 4F2hc/y+LAT2 shows transport characteristics very similar to those of 4F2hc/y+LAT1 (Torrents *et al.* 1998; Pfeiffer *et al.* 1999; Kanai *et al.* 2000), it has a narrower substrate specificity than the latter transporter. y+LAT2 preferentially mediates the efflux of L-arginine in exchange of L-glutamine plus Na^+ (Bröer et al., 2000). It has been determined that arginine has the highest affinity for the intracellular binding site, being the main physiological function of this transporter, the arginine efflux in tissues and cell types, in which there is less expression of y^+LAT1. In fact, distribution of y+LAT2 is ampler than the distribution of y+LAT1 in the body, being expressed in brain, heart, testis, kidney, small intestine and parotid gland (Bröer et al., 2000), and it has been observed compensation of the system y^+L transport in fibroblasts of some LPI patients by 4F2hc/y+LAT2 activity (Dall'Asta *et al.* 2000). One peculiarity of y+LAT2, in comparison with other systems transporting glutamine, is that it recognizes neutral amino acids in the presence of Na^+, and cationic amino acids in the absence of Na^+ (Dye *et al.* 2004).

1.5.2.5 4F2hc/asc1

This protein transports short chain neutral amino acids like glycine, L-alanine, L-serine, L-threonine, L-cysteine, α-aminoisobutyric acid and β-alanine, in a Na^+-independent manner (Fukasawa *et al.* 2000; Nakauchi *et al.* 2000). Despite its preferential function as an exchanger, among all light subunits, asc1, can function as a facilitate diffuser, and is able to transport D-isomers of small neutral amino acids, particularly D-serine, with high apparent affinity (Km ~50 μM) (Fukasawa *et al.* 2000; Nakauchi *et al.* 2000). It is expressed in lung, skeletal muscle, heart, kidney, small intestine and placenta, although is particularly relevant its distribution all over the central nervous system. Here acts as a neuronal transporter of glutamate co-agonists, like D-serine and glycine, that activate the N-methyl-D-aspartate (NMDA) receptors (Helboe *et al.* 2003). 4F2hc/asc1 controls the synaptic concentration of such compounds probably avoiding over activation of NMDA receptors (Hashimoto and Oka 1997). In fact, asc1 (Slc7a10) knockout mice model, shows tremors, ataxia, seizures and early postnatal death (Xie *et al.* 2005), probably due to elevated extracellular D-serine concentrations that over-activate NMDA receptors causing hyper-excitability. Contrary to this view, a more recent study using this knockout model, demonstrates that 4F2hc/asc1 controls glycine levels in the brain and is required for glycinergic inhibitory transmission (Safory *et al.* 2015).

1.5.2.6 4F2hc/xCT

This transporter exchange, in an electroneutral and Na+-independent manner, extracellular anionic cystine and intracellular glutamate, with a 1:1 stoichiometry (Makowske and Christensen 1982; Bannai 1984; Sato *et al.* 1999; Bassi *et al.* 2001). xCT is expressed in neuronal cells, macrophages, fibroblast, pancreas, hepatocytes, kidney and cell culture lines (Sato *et al.* 1999; Bassi *et al.* 2001). At physiological conditions, there are high intracellular levels of glutamate and low levels of cysteine, which drives the direction of the exchange of both amino acids by xCT. This is due to that intracellular cystine is rapidly reduced to cysteine, and incorporated into glutathion and protein biosynthesis pathways (Bannai 1984; Bannai and Tateishi 1986).

Additionally, cysteine is the rate-limiting substrate for reduced tripeptide glutathione (GSH) synthesis, which is required for proliferation, redox cycling and antioxidative defense (Reddy *et al.* 2008; Seiler *et al.* 2008). Thus, 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; Kim *et al.* 2001; Huang and Ingber 2005; Lo *et al.* 2008; Seib *et al.* 2011).

4F2hc/xCT impact in health

Due to the excessive glucose metabolism in cancer cells via anaerobic glycolysis, the demand for nutrients in those cells is increased (Ganapathy *et al.* 2009). Thus, to maintain increased the energy demand, cancer cells express transporters for lactate and amino acids, which increase the reactive oxygen species (ROS) and redox cycle regulation (Vafa *et al.* 2002; Kroemer and Pouyssegur 2008). In this way, overexpression of 4F2hc/xCT may protect tumor cells against oxidative stress by increasing GSH synthesis (Takeuchi *et al.* 2013). In the recent years, 4F2hc/xCT has been related to a new form of cell death named ferroptosis, which is caused by genetic ablation of chemical inhibition of this transporter (Dixon *et al.* 2014). Interestingly, xCT is downregulated by p53 inducing ferroptosis, as a novel tumor suppressor mechanism (Jiang *et al.* 2015). Finally, it has been suggested that overexpression of 4F2hc/xCT in tumor cells, could protect them against radiotherapy and chemotherapy by avoiding the oxidative stress caused by the two anti-cancer therapies (Takeuchi *et al.* 2013).

On the other hand, xCT function impairment has led to defective protection against oxidative stress in activated macrophages, in cellular and mouse models (Nabeyama *et al.* 2010), and additionally, it provokes the activation of the ROS/autophagy pathway in hepatocellular carcinoma (Guo *et al.* 2011). Furthermore, xc- malfunctioning can cause excitotoxicity, which together with oxidative stress are characteristics of Parkinson's disease. Additionally, in brain, xCT seems to be more related to the homeostasis of glutamate, and not so much with contributing to the balance of glutathione levels for protecting cells from oxidative stress as it was suggested before (De Bundel *et al.* 2011). It is known that system xc- can affect excitability, having behavioral consequences, due to the transport of glutamate. It has been found to be relevant for optimal spatial working memory, and its inactivation decreases susceptibility to limbic epileptic seizures (De Bundel *et al.* 2011). xCT overexpression in gliomas has been found related to neurodegeneration, and brain edema in the periphery of glioma tumors, by the excess of glutamate. These effects were alleviated by silencing xc- system (Savaskan *et al.* 2008). Conversely, chronic cocaine consumption has been associated with down regulation of system xc-, which results in reduced extracellular levels of glutamate, and diminished stimulation of extrasynaptic group II metabotropic glutamate receptors in the accumbens nucleus and prefrontal cortex. This fact causes relapse in cocaine addiction (Baker *et al.* 2003). Additionally, xCT has been found as the fusion-entry receptor for the Kaposi's sarcoma-associated herpesvirus (KSHV; human herpesvirus 8), since overexpression of xCT, significantly increases the effectiveness of fusion of the KSHV to different cell lines (Kaleeba and Berger 2006).

1.6 Relevance of the determination of the atomic structure of HATs

As it has been described by now, HATs are involved in several important physiological processes, from cellular growth and metabolism to synaptic transmission. Moreover, these transporters are related with metabolic disorders and inherited and acquired diseases, as a consequence of their malfunctioning or overexpression. Due to the tight function-structure relationship, structure of HATs can reveal the molecular mechanisms of the amino acid transport in normal conditions, and help to understand the molecular basis of some of these disorders. In this way, could be better illustrated how the transporter is affected, and which stages of the transport are impaired, and probably, therapeutic drugs for some of these diseases could be developed or improved in the future.

Nowadays, it is well known the valuable use of atomic structure of biomolecules for drug designing, being very helpful to find agonist or inhibitors with therapeutic effect. In the same line, the structure of these therapeutic targets can be very useful to increase the selectivity of a drug, and/or to raise the drug bioavailability without decreasing its selectivity and affinity. Unfortunately, due to the amphipathic nature of HATs, structural data of them at atomic level is very scarce, being restricted to the human 4F2hc ectodomain (Fort *et al.* 2007). The structure determination of LAT1, a potential anti-cancer molecule, or of some of those closely related light subunits (LATs of the SLC7 family share around 60% identity between each other), could help to increase the selectivity of such inhibitor by LAT1 in the tumor cells. Additionally, could be useful, for instance, to develop better tracers for imaging tumors via LAT1 by PET, as those developed by the moment lacks of sensitivity, as previously discussed (see Introduction section 1.5.2.1). In addition, structure determination of any heterodimer, or of any of its light subunits, could be valuable to understand the main principles of the amino acid transport by these type of transporters. For instance, structural information about the putative involvement of the heavy subunit (rBAT and/or 4F2hc) in the catalytic properties of light subunits, and about the mechanism of amino acid transport mediated by integrin signaling via 4F2hc could be obtained.

1.7 Structural information of HATs

1.7.1 Crystal structure of human 4F2hc ectodomain (4F2hc-ED)

Atomic structure of the 4F2hc-ED, at maximum resolution of 2.1 Å, was solved in our lab almost ten years ago (Fort *et al.* 2007). 4F2hc-ED exhibits structural homology with bacterial glucosidases (α-amylases), having in common domains A and C, but lacking of the B domain, which has the glucosidase catalytic residues (Figure 4) (Chillaron *et al.* 2001; Fort *et al.* 2007). The domain A is a TIM barrel (or (βα)8 barrel), and the domain C, which is always present in amylases with some variability in sequence and length, is a C-terminal β-sheet domain, which consists in 8 antiparallel β sheets (Pujadas and Palau 2001). Domain B is a long loop region inserted between the third β strand and the third α-helix of domain A, that is present in rBAT-ED (Fort *et al.* 2007). Thus, 4F2hc lacks glucosidase activity, since no enzymatic activity was detected for purified 4F2hc-ED when tested D-glucose, D-galactose, or D-manose as substrates (Fort *et al.* 2007). On the other hand, glucosidase-like function has not been yet investigated for rBAT.

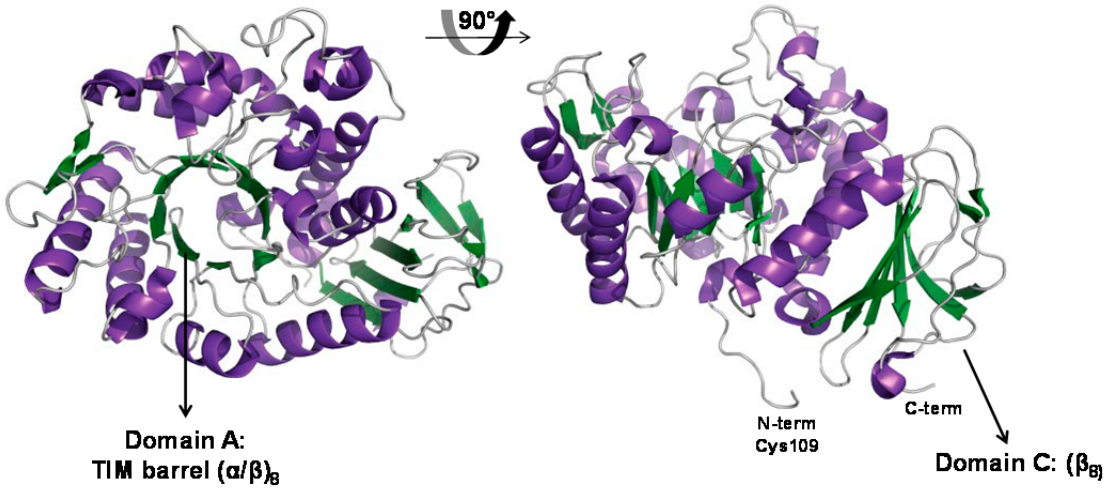

Figure 4. Human 4F2hc-ED structure. From left to right: top view and lateral view of the structure of 4F2hc-ED. This protein has in common with α-amylases the A (TIM barrel (β/α)8) and C (8 antiparallel β-sheets) domains. α-helices are colored in purple and β-sheets in green. Adapted figure (Fort *et al.* 2007; Álvarez-Marimon. 2014).

Interestingly, it was observed a macromolecular dipole of 4F2hc-ED with two charged faces (Figure 5). On top and in the lateral sides, residues negatively charged were predominant, and on the bottom face, and at N-terminal of the ectodomain, predominated residues positively charged (Fort *et al.* 2007). These two differentiated patches may dictate the interactions of 4F2hc with other moleules. A model of 4F2hc-ED orientated on the membrane, showed an electrostatic interaction between 4F2hc-ED and the phospholipid polar heads of the plasma membrane (Fort *et al.* 2007). These kind of interactions are typical in several peripheral and integral membrane proteins (McLaughlin and Aderem 1995; Bhatnagar and Gordon 1997; Conte and Matthews 1998). This orientation correlated with the positioning of the cysteine involved in the disulphide bridge with light subunits, close to the N-terminal transmembrane segment of 4F2hc.

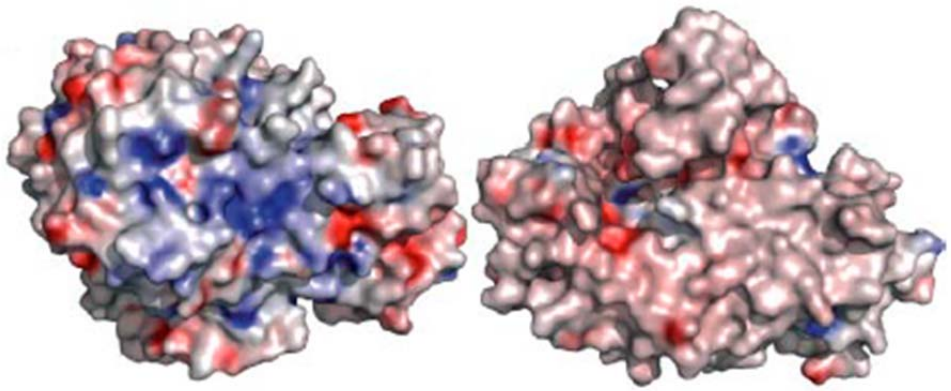

Figure 5. Electrostatic surface of 4F2hc-ED. Electrostatic surface of monomer in the monoclinic structure (2DH2) has a dipolar distribution of charges with a positive patch in the N-terminal region (left) and negative surface residues in the other sites of the molecule (right). (Fort *et al.* 2007).

Human 4F2-ED was crystallized as a monomer (PDB 2DH2) and as a homodimer (PDB 2DH3) (Fort *et al.* 2007). When human 4F2hc has been overexpressed in different cell types, homodimerization has been evidenced by the appearance of a band of around 190 kDa in SDS-PAGE, which is eliminated upon DTT treatment. This indicates the presence of a disulfide bridge between both 4F2hc molecules, formed by the same cysteine residue (Cys109)

responsible for the disulfide bond with light subunits (Fort *et al.* 2007). When are overexpressed, these 4F2hc homodimers are formed in the endoplasmic reticulum, and arrived to the plasma membrane, as demonstrated by endo H susceptibility and biotin labeling assays, respectively (Fort *et al.* 2007). However, 4F2hc-ED does not have Cys109. This correlates with the ability of 4F2hc C109S mutant to form homodimers, even in the absence of disulfide bond, suggesting an important interaction area between 4F2hc monomers (Estevez *et al.* 1998; Fort *et al.* 2007). Cys109 would be relevant for the stabilization of the 4F2hc-4F2hc interaction but not for the homodimerization process. Homodimers has not been found in normal tissues or tumors (Turnay *et al.* 2011), suggesting a preference of 4F2hc to form heterodimer with its light subunits, than homodimer with itself, in physiological expression levels.

It has been found in 4F2hc, but not in rBAT, the Engelman's motif, GxxxG, which is involved in the homodimerization of transmembrane α-helices (Russ and Engelman 2000; Senes *et al.* 2000; Senes *et al.* 2004). These two glycines in 4F2hc are highly conserved in all vertebrates, but this motif is not found in the 4F2hc-associated light subunits. Interestingly, it is found in integrins (Berger *et al.* 2010), suggesting a possible involvement in the 4F2hc-integrin interaction (see Introduction section 1.5.2). Other cysteine in 4F2hc, Cys330, highly conserved in mammals, is buried into the TIM barrel (Figure 6). Despite the inaccessibility of this cysteine, it has been reported the participation of this residue in processes of cell fusion (Okamoto *et al.* 1997). Thus, it is probable that 4F2hc undergoes conformational movements that renders Cys330 accessible and able to interact with other proteins like β1-integrins. In fact, studies to visualize the mobile areas of 4F2hc-ED by the determination of the B factor, showed that movements of Aα5 and Aα6 could make accessible Cys330 (Figure 6) (Rodríguez-de-la-Ballina. 2011).

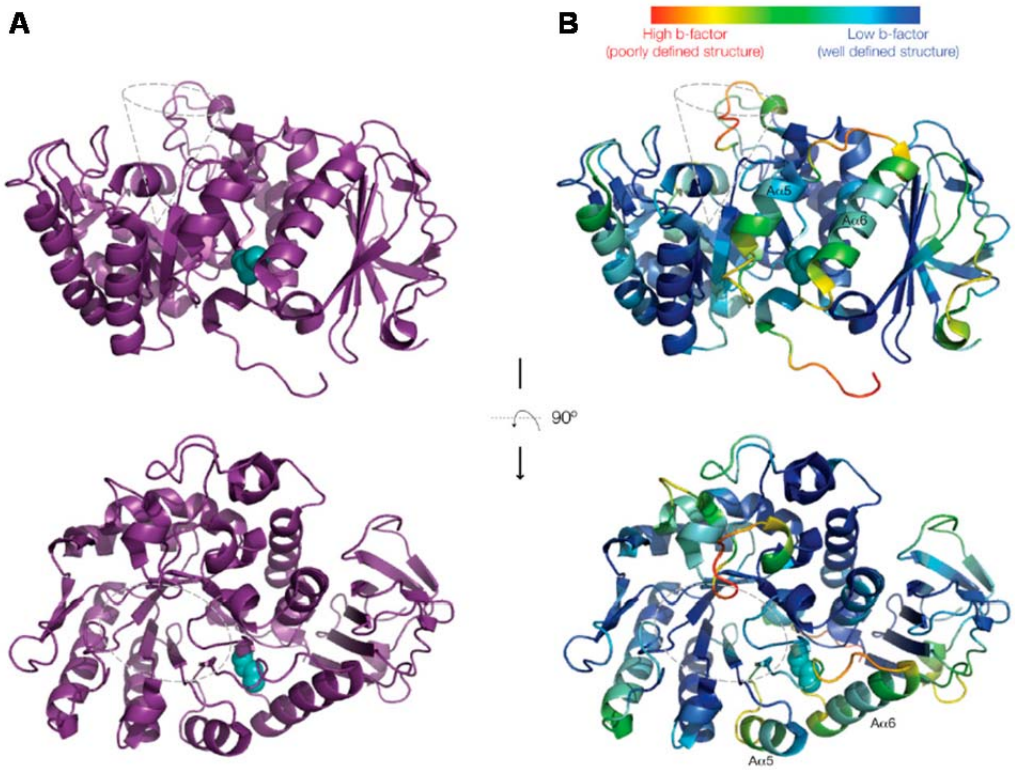

Figure 6. Analysis of the 4F2hc-ED mobile areas by B factor data. Cys330 colored in cyan is highlighted on the 4F2hc-ED structure (A). Regions of the 4F2hc-ED are colored based on the B factor (B). Regions of the protein that are well defined are colored in blue (low B factor) and those that are not so well defined are colored in red

(high B factor). Helices α5 and α6 showed a higher B factor than the rest of the helices from domain A, indicating that they are may be more mobile. A color code bar that represent the B factor values is shown in the upper part of panel B. Temperature factor or B factor describes the attenuation of X-ray scattering caused by thermal motion. Thus, a low B factor means that the position of the atom has been determined with accuracy, whereas a high B factor indicates a poorly defined area. Lateral and top views of the protein are represented in the upper and bottom parts of the figure, respectively. (Rodríguez-de-la-Ballina. 2011).

1.7.2 Structure of prokaryotic homologues of LATs

The closest structural paradigms of light subunits are structures of prokaryotic membrane transporters of the APC superfamily (Table 5), since nowadays none structure of a metazoan LAT has been solved. These structures correspond to prokaryotic homologues that share ≤ 20% amino acid sequence identity with metazoan LATs, although have in common the presence of 12 TMDs. Additionally, these prokaryotic transporters exhibit a structural fold, called "5+5 inverted repeat fold" or LeuT fold (Yamashita *et al.* 2005), that it is also expected to appear in metazoan light subunits. In the 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 5. Crystal structures from prokaryotic APC superfamily members homologues to LATs (SLC7). This table lists the available structures in the protein data bank (PDB) of prokaryotic homologues of LAT family. The oligomeric state, resolution (Å), if they were co-crystallized, and the PDB accession number are shown. (Hediger *et al.* 2013).

1.7.2.1 The LeuT or *5 + 5* inverted repeat fold

This fold was discovered for the first time in the bacterial leucine transporter LeuT (Na^+-dependent and Cl^--dependent amino acid transporter) (Yamashita *et al.* 2005). The LeuT fold is characterized by a pseudosymmetry, where the 1 to 5 TMDs (first repeat), are related to the 6 to 10 TMDs (second 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 7) (Yamashita et al., 2005). Distant families of secondary transporters exhibit this fold, like prokaryotic APC superfamily members: AdiC (arginine/agmatine antiporter) (Fang *et al.* 2009; Gao *et al.* 2009; Gao *et al.* 2010; Kowalczyk *et al.* 2011); ApcT (H^+-dependent amino acid transporter) (Shaffer *et al.* 2009) and GadC (glutamate/ GABA antiporter) (Ma *et al.* 2012); and transporters of other families: Mhp1 (benzyl hydantoin transporter) (Weyand *et al.* 2008), CaiT (carnitine/γ-butyrobetaine antiporter) (Tang *et al.* 2010), BetP (betaine transporter) (Ressl *et al.* 2009; Perez *et al.* 2012) and vSGLT (Na^+-coupled glucose transporter) (Faham *et al.* 2008). Interestingly, these transporters share low amino acid sequence identity among them, and with LeuT (<10%). Thus, the presence of this fold in these unrelated sequence protein may indicate a common evolution process, drive by gene duplication and fusion events. The result is a fold

that allows these proteins to undergo conformational changes among symmetry-related states to carry out the transport activity (Khafizov *et al.* 2010).

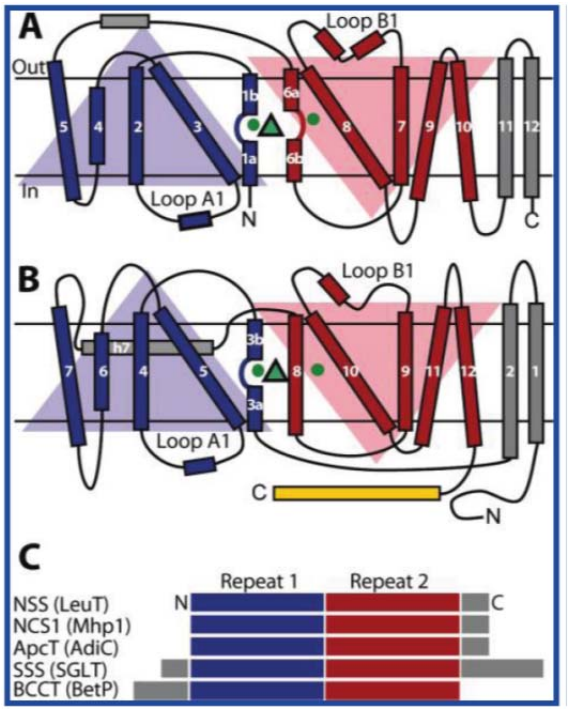

Figure 7. Schematic representation of the LeuT fold. (**A**) LeuT, (**B**) BetP, and (**C**) comparison of the known structures. 5 TMDs helix inverted-topology repeat are represented in blue and red. Peripheral TMD helices and the amphipathic helix h7 of BetP are represented in gray. The first loop of each repeat, loops A1 and B1, are shown. In BetP, the long cytoplasmic C-terminal helix is represented in yellow. Green triangles and circles represent substrates and ions, respectively. (Khafizov *et al.* 2010).

1.7.2.2 Structure of the closest prokaryotic homologue of LATs, the Arginine/Agmatine exchanger AdiC

AdiC is a member of the basic Amino acid/Polyamine Antiporter (APA) subfamily, and constitutes the first structure solved of an amino acid antiporter (Fang *et al.* 2009; Gao *et al.* 2009; Gao *et al.* 2010; Kowalczyk *et al.* 2011). Together with the LAT subfamily, they form part of the APC superfamily of transporters (Casagrande *et al.* 2008). AdiC shares around 18% identity with LATs, being the closest structural paradigm for LATs. This protein is an Arginine (Arg)/Agmatine (Agm) exchanger used by enteric bacteria, as *E. coli*, during extreme acid-resistance (pH 1.5-4). AdiC is expressed as a dimer in the plasma membrane. Arg is imported from the extracellular space to the cytoplasm, where it is decarboxylated by a specific decarboxylase (AdiA), which catalyzes the replacement of the α-carboxyl group of Arg for a proton in Agm, which is then expelled out. For this reason, it has been mentioned that AdiC acts as a "virtual proton pump". Arg/Agm exchange is carried out with 1:1 stoichiometry. Some *E. coli* strains are pathogenic, causing several types of urine and gastrointestinal infections that can be lethal, like the hemolytic-uremic syndrome (HUS) (Griffin and Tauxe 1991). Establishment of *E. coli* in this type of illness is possible due to the resistance of enteric bacteria to the extreme acidic conditions in the stomach, and to the mild acidic conditions in the intestine and in the urine tract. This is favored by the amino acid antiporter AdiC, together with the other three amino acid–dependent acid-resistance systems (glutamic acid–, lysine- and ornithine-dependent acid resistance systems).

According to the alternating access model for secondary transporters (Figure 8) (Jardetzky 1966), AdiC, and LeuT fold transporters, are expected to undergo conformational changes between outward- and inward-facing conformations (open-to-out and open-to-in, respectively) for carrying out the transport. In the case of AdiC, to catalyze the exchange of Arg/Agm (Figure 8) (Kowalczyk *et al.* 2011). Upon initial interaction with the substrate (Arg or Agm), there is an induced-fitting that improves substrate binding and occludes the substrate (thin gate), blocking free diffusion to either side of the membrane (Kowalczyk *et al.* 2011). Occlusion of the substrate has also been reported for other ion coupled symporters sharing the same fold, like LeuT, vSGLT, Mhp1 and BetP (Yamashita *et al.* 2005; Faham *et al.* 2008; Weyand *et al.* 2008; Ressl *et al.* 2009) (Figure 8). In exchangers like AdiC, the transit between outward- and inward-facing conformations occurs only from substrate-bound occluded states (Figure 8). Moreover, ion-coupled symporters also transit between these states without substrate bound (apo state) (Figure 8).

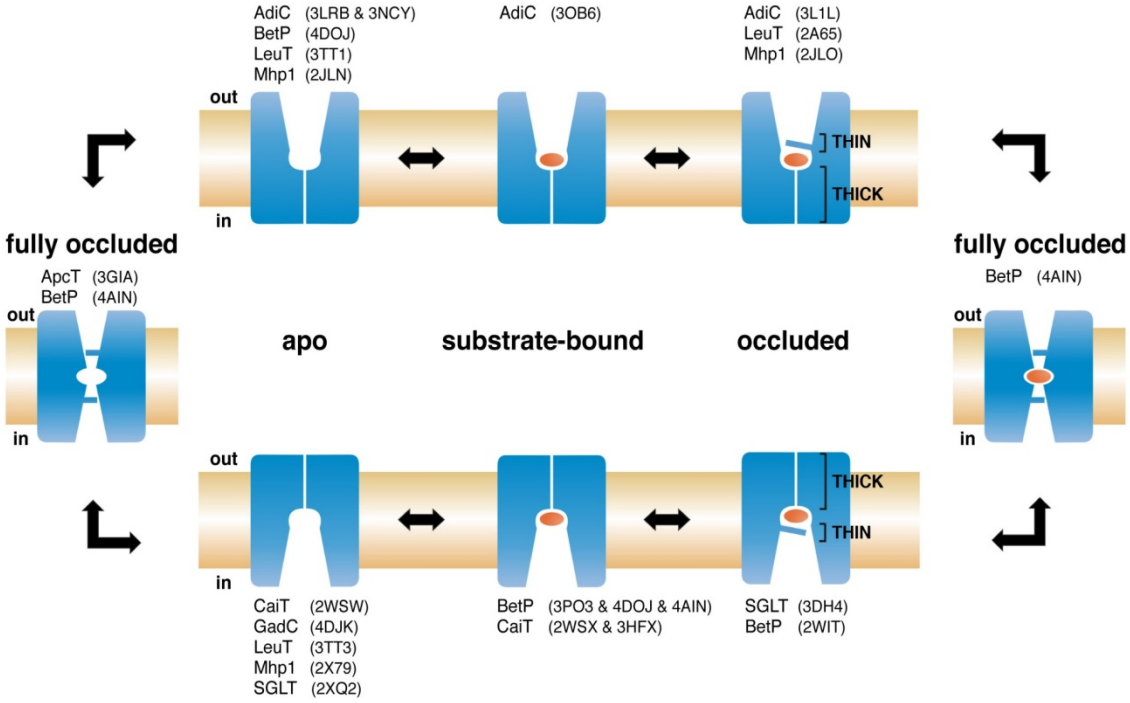

Figure 8. Symmetrical states along the alternative access mechanism of transporters with the LeuT fold. Upon substrate (red ellipsoid) binding to the open-to-out (outward-facing) apo state, the substrate-bound state (AdiC (3OB6)) evolves to an occluded state, where two gates (thick and thin) prevent the diffusion of the substrate to either side of the membrane (Krishnamurthy *et al.* 2009). 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 (Yamashita *et al.* 2005), vSGLT (Faham *et al.* 2008), Mhp1 (Weyand *et al.* 2008; Shimamura *et al.* 2010), BetP (Ressl *et al.* 2009; Perez *et al.* 2012), GadC (Ma *et al.* 2012), and AdiC (Fang *et al.* 2009; Gao *et al.* 2009; Gao *et al.* 2010; Kowalczyk *et al.* 2011). 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 (Shaffer *et al.* 2009) 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. Protein Data Bank (PDB) access codes are indicated in parentheses. (Kowalczyk *et al.* 2011).

AdiC structure has been solved in three different states, but all facing outwards (Figure 8): outward-facing without substrate (PDB: 3LRB and 3NCY) (Fang *et al.* 2009; Gao *et al.* 2009),

outward-facing with substrate bound (mutant N101A, PDB: 3OB6) (Kowalczyk *et al.* 2011), and outward-facing with substrate occluded (mutant N22A; PDB: 3L1L) (Gao *et al.* 2010). One of them (mutant N101A, PDB: 3OB6) was solved in our lab at 3.0Å resolution (Figure 9). This structure has brought insights about the substrate-induced fitting mechanism of AdiC, however the complete mechanism of transport (e.g., translocation) remains unknown, because crystal structures of the inward-facing states are missing. In the Appendix section, a description of the substrate-induced fitting mechanism of AdiC, and the studies conducted by us, for trying to crystalize AdiC in inward-facing conformation for determining its atomic structure, is provided.

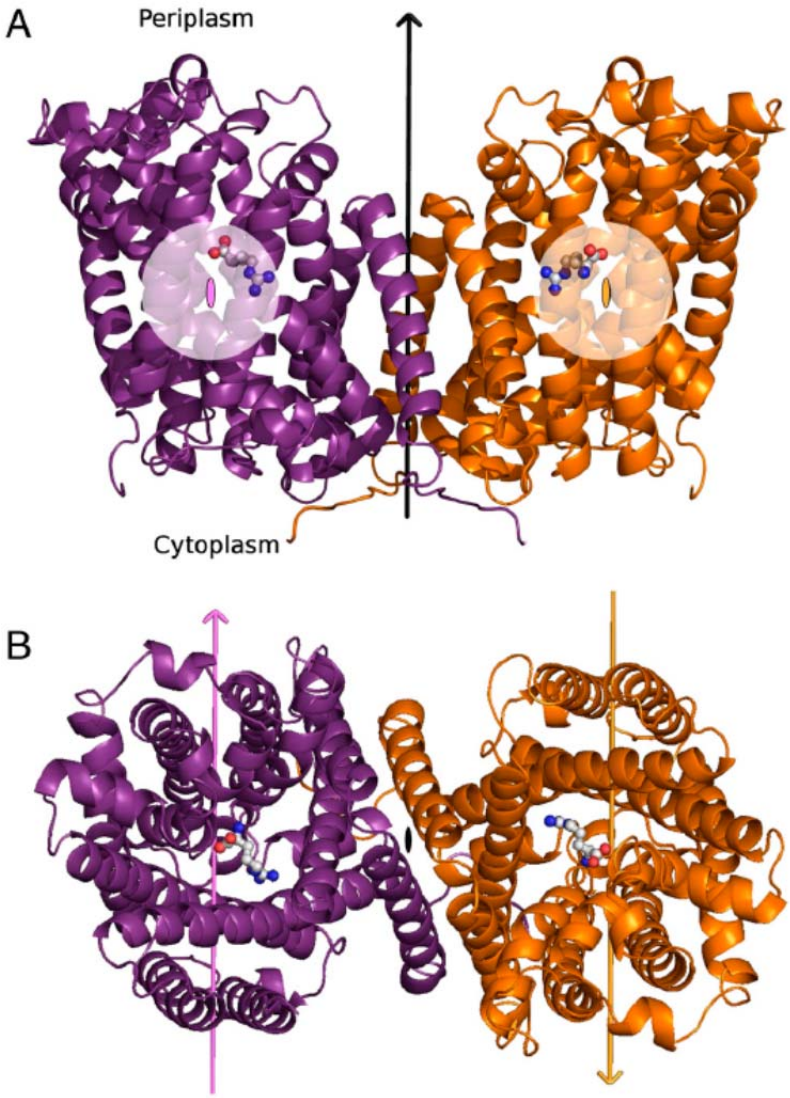

Figure 9. Structure of AdiC-N101A bound to Arg^+. (**A**) Lateral and (**B**) periplasmic views of the AdiC-N101A-Arg^+ complex homodimer (in purple, protomer 1 and in orange, protomer 2). The bound substrate is depicted with a ball and-stick model. The two-fold subunit axes (ovals or arrows) are colored as their corresponding protomers. The dimer axis (arrow or ovals) is indicated in black. (Kowalczyk *et al.* 2011).

1.7.2.3 The closest characterized bacterial homologue of metazoan LATs: SteT (L-serine/L-threonine antiporter)

SteT was the first prokaryotic member of the SLC7 family cloned and functionally characterized in our laboratory (Reig *et al.* 2007). SteT is an electroneutral obligatory exchanger, with preferential transport of L-serine and L-threonine in a 1:1 stoichiometry, although it can also transport with less affinity aromatic amino acids (Reig *et al.* 2007). This transporter share around 30% amino acid identity with metazoan light subunit, although lacks of the conserved cysteine between TMD3 and TMD4 responsible for the disulfide bridge formation.

Single particle analysis by negative-staining electron microscopy (TEM) of detergent-purified SteT, showed that this protein is a monomer with an elliptical shape. The images showed a central cavity as well (Figure 10) (Reig *et al.* 2007). Functional studies involving TMD8, revealed that Lys 295 is implicated in the substrate recognition and specificity (Bartoccioni et al., 2010). Due to the troubles to crystalize and/or to obtain well-diffracting crystals of SteT, atomic structure of SteT has not been solved to date. For this reason, a structural model of SteT based on the structure of AdiC was generated. In such model, it can be observed that substitution of the residue Lys295 for a smaller residue, leads to a broader space in the substrate binding site that results in wider selectivity of substrates (Bartoccioni et al., 2010). In the last years, two SteT mutants, I134V-A377T and L210Q-229V, generated by random mutagenesis and selected using the GFP-split system (Cabantous and Waldo 2006), have been found in our lab as good candidates for crystallization studies (Rodríguez-Banqueri 2013).

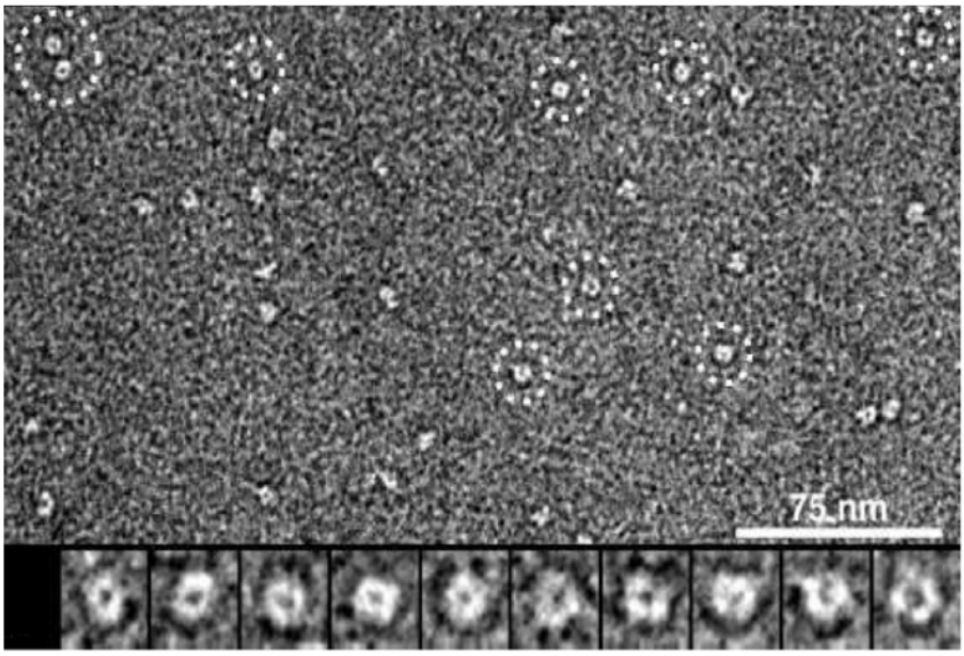

Figure 10. Single particles of purified SteT visualized by negative staining and TEM. Detergent-solubilized and purified SteT showed a homogeneous population of monomers, with a diameter around 7 nm. This size corresponds to the protein, together with the additional mass arising from the detergent belt and endogenous lipid bound to the protein. (Reig *et al.* 2007).

1.7.3 First structural model of a HAT, human 4F2hc/LAT2, and the pursuit for a more stable HAT

Human 4F2hc-LAT2 heterodimer could be expressed in *Pichia pastoris* and purified in DDM (n-Dodecyl-β-D-maltoside). This fact, allowed to solve a structural model of the heterodimer at low resolution, 21Å, from negatively stained complexes by TEM (Figure 11) (Rosell *et al.* 2014). Two densities of different size were observed (Figure 11A and 11B), fitting properly the structure of the 4F2hc-ED inside the small body (Figure 11C). In this way, it was observed that 4F2hc and LAT2 interacts almost completely along the interface of both proteins; 4F2hc covering the extracellular face of the transporter (Figure 11B and 11C). Docking analysis and crosslinking assays also corroborated such findings (Rosell *et al.* 2014). This structure showed some details about the interaction between the heavy chain and the light subunit, but the low stability of this heterodimer did not allow going further in structural studies to obtain atomic resolution. Docking analysis and crosslinking assays also corroborated such findings (Rosell *et al.* 2014).

Additionally, in that work, it was observed that human 4F2hc highly increases the stability of a poorly stable human LAT2 in solution (Rosell *et al.* 2014). LAT2 purified from Pichia membranes in DDM, produced a single peak of aggregated protein in the void volume by Size Exclusion Chromatography (SEC), meanwhile co-expression of 4F2hc-LAT2 in the same conditions generated a protein peak in an elution volume corresponding with soluble protein (Figure 12A left). On the other hand, due to the low stability of LAT2 and to the stabilization effect of 4F2hc on LAT2 in solution, only 4F2hc/LAT2 heterodimer, but not LAT2 alone, could be functionally reconstituted in liposomes (Figure 12A right). It was also evidenced that the ectodomain (ED) of 4F2hc was sufficient to the stabilization effect of 4F2hc on LAT2 in solution (Figure 12B and 12C).

This leads to the rationality that, it would be possible to obtain a more stable heterodimer by finding first a more stable light subunit than human LAT2. This hypothesis could be also supported by the nature of both subunits forming the heterodimer. 4F2hc has only one TMD, meanwhile metazoan light subunits are predicted to have twelve (Gasol *et al.* 2004), as also it has been shown in the structure of the prokaryotic homologue of LATs AdiC (Fang *et al.* 2009; Gao *et al.* 2009; Gao *et al.* 2010; Kowalczyk *et al.* 2011). This suggests that the stabilizaty of the whole heterodimer in solution would be more compromised by the light subunit than by the heavy chain. For this reason, in our lab, Dr. Elena Álvarez developed a project in order to get a metazoan light subunit with good expression levels and highly stable (Álvarez-Marimon. 2014). Twenty four metazoan light subunits, from seven different species, were cloned and expressed as C-terminal GFP fusion proteins in the yeast *Saccharomyces cerevisiae*. The protocol described by Drew and co-workers (Drew *et al.* 2008) was followed. After expression screening, stability studies by Ultracentrifugation Dispersity Sedimentation assay (UDS) (Gutmann *et al.* 2007) and Fluorescence Size Exclusion Chromatography (FSEC) (Kawate and Gouaux 2006), three non-human light subunits from a vertebrate source (LAT1, LAT2 and b^{0+}AT) were identified as the most stable. Thus, the corresponding vertebrate heterodimers (4F2hc/LAT1, 4F2hc/LAT2 and rBAT/b^{0+}AT) became putative good candidates for structural studies. This work is unpublished, and we would like to keep confidential the source organism of these proteins. Therefore, hereafter these proteins will be referred by their names indicated as vertebrate proteins, and unless otherwise specified, we are referring to these non-human proteins from a vertebrate source.

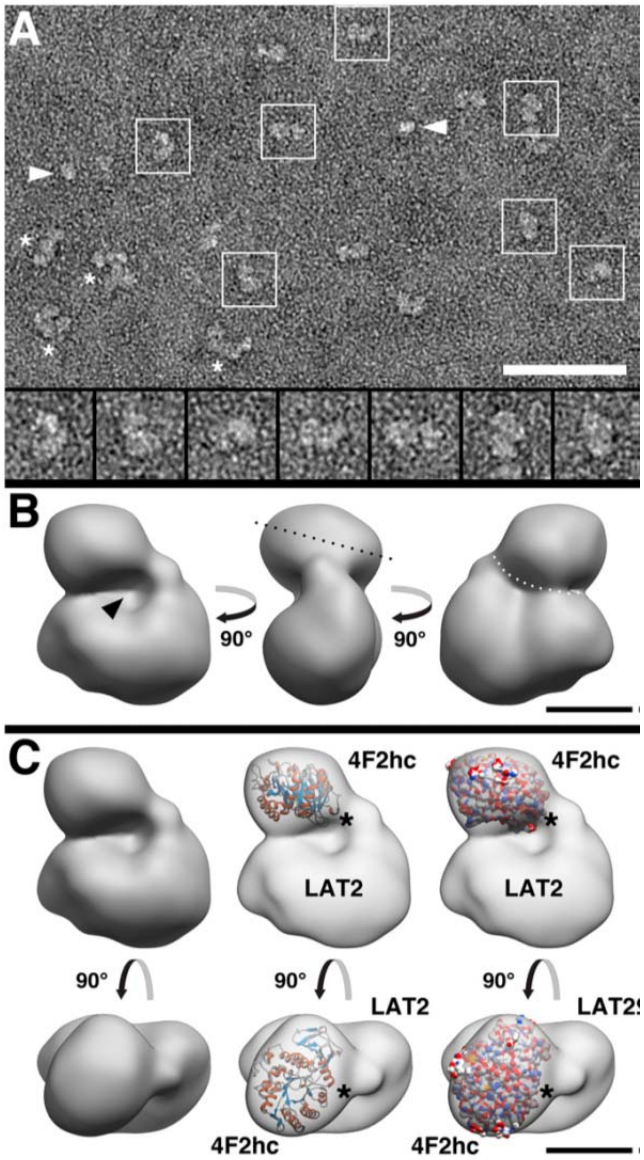

Figure 11. TEM and 3D reconstruction of human 4F2hc/LAT2. (**A**) Overview electron micrograph of purified and negatively stained 4F2hc/LAT2 heterodimers. The boxed 4F2hc/LAT2 complexes were magnified and displayed in the gallery. Arrowheads mark 4F2hc or LAT2 monomers from disrupted heterodimers. Asterisks 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). (Rosell *et al.* 2014).

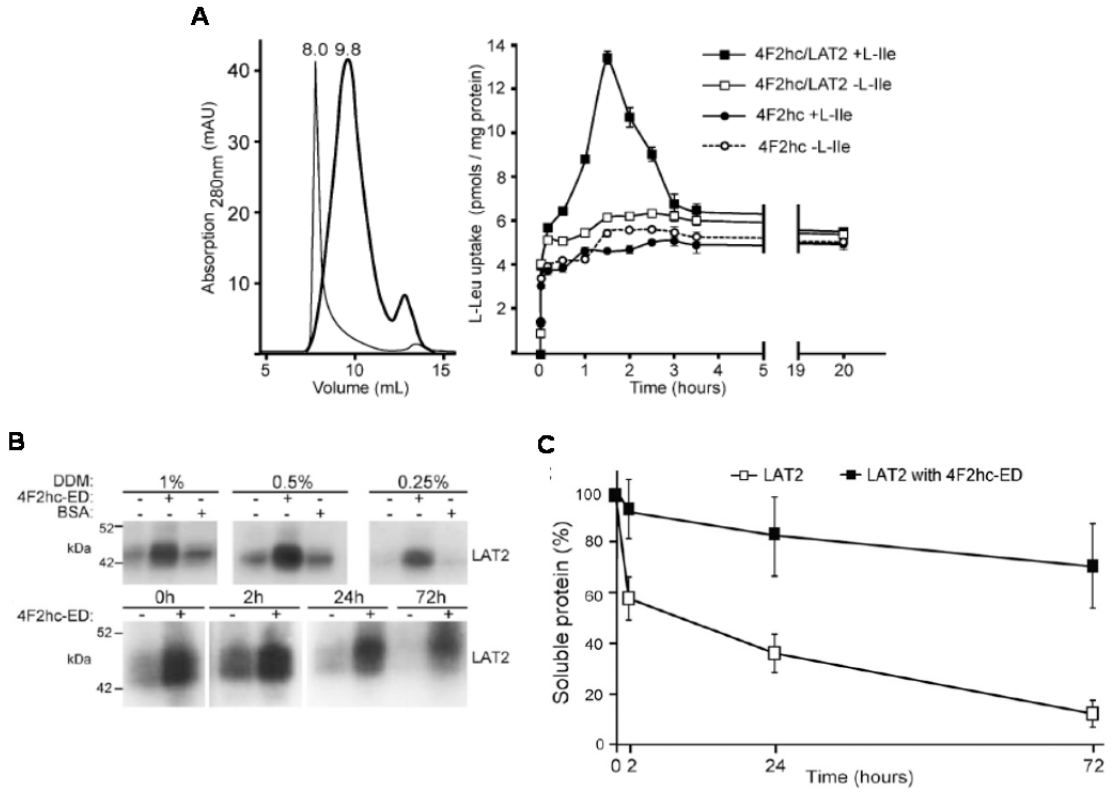

Figure 12. Human LAT2 is stable in solution only in the presence of human 4F2hc. Size-exclusion chromatography profiles of purified 4F2hc/LAT2 (thick curve) and purified LAT2 (thin curve) (**A**). 4F2hc/LAT2 elutes at 9.8 mL, whereas LAT2 is eluted in the void volume (8 mL). Time-course of L-leucine uptake in 4F2hc/LAT2 and 4F2hc proteoliposomes (**A**). To the right, 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 (Newman *et al.* 1981) and exchangers such as the LAT transporter SteT (Reig *et al.* 2007). (**B**) In top, western blot analysis of DDM-solubilized LAT2 in the presence of 1 mg/mL of 4F2hc-ED or BSA or plain buffer, as indicated. In bottom, 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. The presence of 4F2hc-ED increased LAT2 stability at all analyzed times. Adapted figure (Rosell *et al.* 2014).

2. Objectives

Objectives

Owing to the relevance of determining the atomic structure of HATs, and also to the obstacles for studying this type of metazoan membrane proteins, the main objectives of this thesis were the following:

I. To identify an heterodimer with high expression levels and stable enough to be undertaken in structural studies:

1. To analyze the expression levels and the stability, in different expression systems, of three vertebrate heterodimers formed by light subunits previously identified as the most stable among 24 light subunits investigated: rBAT/$b^{0,+}$AT, 4F2hc/LAT1 and 4F2hc/LAT2.

2. To modify the heterodimer selected in the previous objective (vertebrate 4F2hc/LAT1) to increase its homogeneity and to reduce its flexibility for 3D crystallography.

3. To determine the functionality of vertebrate 4F2hc/LAT1 heterodimer. This would indicate its correct folding in the expression system and in solution.

II. To optimize the conditions for the study of vertebrate 4F2hc/LAT1 by cryo-EM:

1. To determine the stability of 4F2hc/LAT1 in different detergents and lipids.

2. To study the stability of 4F2hc/LAT1 in amphipol A8-35. This polymer has been successfully used for solving atomic structures of metazoan membrane proteins by cryo-EM.

3. To analyze the heterogeneity of 4F2hc/LAT1 solubilized in amphipol A8-35 by negative staining.

4. To perform the 3D reconstruction of 4F2hc/LAT1 solubilized in amphipol A8-35, from negative stained samples, and compare its overall shape with the reported 3D reconstructions of detergent-solubilized human 4F2hc/LAT2.

3. Results and Discussion

Results and Discussion

It has been determined that around 30% of the human proteome is composed of integral membrane proteins (IMPs) (Overington *et al.* 2006; Murray *et al.* 2012). They are involved in a wide variety of physiological processes, as result of their particular function and expression in different cellular membranes. Thus, they are involved in metabolism, signal transduction, energy conversion and utilization, cellular growth and survival. As a consequence, dysfunction of these proteins causes different types of disorders and diseases (Sanders and Myers 2004), thus constituting IMPs, two thirds of all therapeutic targets (Hediger *et al.* 2013). However, membrane proteins represent less than the 2% of all protein structures available in the PDB, with a major representation of prokaryotic membrane proteins (PDB, February 2016), and with minority of structures solved at high resolution. These statistics reflect the challenges for obtaining membrane proteins structures.

For structural studies of relatively high molecular weight proteins or macromolecular complexes, the two predominant techniques widely used are 3D crystallography and cryo-EM. Both techniques require highly pure, monodisperse and stable protein in solution. However, much less protein is needed in the cryo-EM samples in comparison with 3D crystallography, where a high amount of protein is required to screen several conditions of crystallization, and find any in which well-ordered and well-diffracting crystals can grow. Both techniques have advantages and limitations that will be discussed through this Results and Discussion section.

Production of high quality protein, with the characteristics described above, is more difficult when the protein, or some of the proteins forming a macromolecular complex, are metazoan integral membrane proteins. At first instance, metazoan membrane proteins are difficult to obtain at high yield by using heterologous expression systems, by which finding the appropriate eukaryotic or prokaryotic expression system is imperative. It is usually considered that prokaryotic expression systems, like *Escherichia coli*, yeast (*Pichia pastoris*) and insect cells, produce the highest amounts of protein (He *et al.* 2014; Rosano and Ceccarelli 2014). However, the final yield of a protein heterologously expressed depends on the individual protein, by which, for obtaining the highest yield, several expression systems are usually tested (Demain and Vaishnav 2009; Bernaudat *et al.* 2011). Bacteria and yeast share the possibility of reaching high cellular densities in relatively short periods of time, improving protein final yields and saving time (He *et al.* 2014; Rosano and Ceccarelli 2014). In addition, culture materials are more inexpensive in comparison with insect or mammalian cells. In some instance, the most widely used prokaryotic microorganism for overexpressing proteins, *E. coli*, has been used to produce eukaryotic integral membrane proteins for structural studies (Ferguson *et al.* 2007; Nishida *et al.* 2007; Newby *et al.* 2008; Lu *et al.* 2014). However, nowadays, these structures represent less than the 8 % of the structures from eukaryotic membrane proteins deposited in the PDB, corresponding the majority to proteins of lower eukaryotes (He *et al.* 2014). This evidences the *E. coli* deficiencies in the production of metazoan membrane proteins.

Prokaryotes lacks of some of the requirements for appropriate eukaryote protein folding and protein insertion in the membrane, due to the more complex mechanisms of transcription, translation and secretion present in eukaryotes. Thus, *E. coli* cannot carry out some post-translational modifications neither express relevant chaperon proteins, and may lacks of some essential lipids for metazoan membrane protein stability and function (Sahdev *et al.* 2008). Also, it has been recently suggested that post-translational modifications, specifically glycosylation, can

be relevant for protecting proteins from the air–water interface, that can denature them, during sample preparation in thin vitrified layers for cryo-EM (Bai *et al.* 2015). Thus, eukaryotic expression systems are more commonly used than prokaryotic ones for producing well folded, mature and active metazoan membrane proteins for structural and functional determination.

Other consideration to be taken into account in the production of metazoan membrane proteins, is that membrane proteins are embedded in a highly hydrophobic lipid bilayer, interacting their hydrophobic regions with the hydrophobic inner part of the membrane. Then, when a membrane protein is extracted from its natural environment, may aggregate due to unspecific hydrophobic interactions between its hydrophobic regions, hampering protein maintenance in aqueous solution. For this reason, detergents and lipids, also a cholesterol derivative, are used in the purification of membrane proteins to mimic the interaction of the protein with the membrane, avoiding protein denaturation and aggregation. As a result, hydrophobic residues from the transmembrane domains (TMDs) interact with hydrophobic chains of detergents and lipids in which such proteins are purified (le Maire *et al.* 2000). Nevertheless, detergents can destabilize some membrane proteins or lead to their inactivation upon time. This, together with the low or no expression of some membrane proteins in heterologous expression systems, hinder the structural determination of great number of membrane proteins. A close evidence of the complexity of structural determination of metazoan membrane proteins, are the studies with human 4F2hc/LAT2 (see Introduction section 1.7.3). A low resolution model was obtained for this heterodimer, however, its low stability did not allow going further in structural studies to obtain atomic resolution.

In the present work, formation of the vertebrate heterodimers conformed by the three previously selected vertebrate light subunits (see Introduction section 1.7.3), 4F2hc/LAT1, 4F2hc/LAT2 and rBAT/$b^{0,+}$AT, was tried, and their stability in solution was studied to find which of these HATs would be the most promising candidate for structural studies. For avoiding repetition, hereafter, unless otherwise specified, we are referring to these non-human proteins from a vertebrate source, object of our studies. Concerning to the detergent condition for the solubilization and purification of these complexes, Dr. Elena Álvarez found that the most widely used detergent for maintaining membrane proteins well folded in solution, n-dodecyl-β-D-maltoside (DDM), plus cholesteryl hemisuccinate (CHS), was in general effective in maintaining light subunits soluble (Álvarez-Marimon. 2014). For this reason, the basic solubilization and purification condition used as starting point from which compare other conditions was DDM/CHS. Different concentrations were used depending on the tested expression system, but always in the same DDM/CHS ratio (5:1, w/w). This DDM/CHS ratio was used in other studies in which CHS was added to increase the stability of metazoan membrane proteins for their structural determination (Drew *et al.* 2008; Zhang *et al.* 2012; Penmatsa *et al.* 2015).

As already mentioned, selection of an appropriate expression system is capital to obtain enough well folded and stable membrane protein for structural studies, and in some occasions, like the current, testing several types of microorganisms and cell lines may be important to find the one that accomplish with the desired requirements. Due to the complexity of HATs; metazoan transporters composed of one-spanning membrane protein (heavy subunit) disulfide bond linked with a 12-spanning membrane protein (light subunit), it is feasible thinking that a complex machinery of translation, membrane insertion and translocation, may be required to the proper folding and assemble of these heterodimers. Additionally, membrane proteins integrity and activity are also conditioned by the interaction with specific lipids (van Meer *et al.* 2008; Phillips *et al.* 2009), by which, the lipid composition of the cellular membrane in which the protein is

expressed can be of big relevance. For these reasons, we focus attention in eukaryotic expression systems to ensure as much as possible that the protein biogenesis and the lipidic composition were the most similar to the native one.

In the initial studies vertebrate, rBAT, 4F2hc, LAT1, LAT2 and $b^{0,+}$AT, as well as vertebrate heterodimers formed between them, were GFP tagged. GFP fusion makes possible rapid and feasible pre-structural studies of membrane proteins, since there is no requirement of extensive protein purification owing to the high sensitivity of the GFP fluorescence detection. In this way, protein expression levels can be rapidly screened by whole cell fluorescence (Drew *et al.* 2008), and protein stability easily monitored by Ultracentrifugation Dispersity Sedimentation assay (UDS) (Gutmann *et al.* 2007) and Fluorescent Size Exclusion Chromatography (FSEC) (Kawate and Gouaux 2006). This strategy was previously used in our lab to identify the most stable light subunits (see Introduction section 1.7.3) (Álvarez-Marimon. 2014).

3.1 Expression of vertebrate 4F2hc/LAT1, 4F2hc/LAT2 and rBAT/$b^{0,+}$AT in human cells

As a starting point, human cells, tsA201 cell line, seemed to be a good choice for testing the expression and stability of 4F2hc/LAT1, 4F2hc/LAT2 and rBAT/$b^{0,+}$AT vertebrate heterodimers. Mammalian cells are used for the heterologous expression of metazoan membrane proteins due to their higher eukaryote properties, enabling the proper post-translational modifications and the correct protein folding, and due to the membrane lipid composition, which could enhance the stability and activity of these proteins (Khan 2013). Mammalian cells have been used for obtaining atomic structures of metazoan membrane proteins by 3D crystallography, like human rhodopsin (Standfuss *et al.* 2007), human RhCG (Gruswitz *et al.* 2010), Drosophila dopamine transporter (Penmatsa *et al.* 2013; Penmatsa *et al.* 2015) and chicken acid-sensing ion channel (Baconguis and Gouaux 2012), and by cryo-EM, like rat TRPV1 channel (Liao *et al.* 2013), and human γ-secretase (Lu *et al.* 2014; Bai *et al.* 2015).

The human cell line tsA201 could be appropriate for the production of vertebrate HATs for structural studies, taking into account the high identity shared between LAT1, LAT2 and $b^{0,+}$AT with the corresponding human homologues (around 80 %), and between rBAT and 4F2hc with the human heavy subunits (around 50 %). High identity between human and vertebrate LATs suggests that the amino acid transport function carried out by these light subunits is conserved through vertebrates, and that a human expression system for producing vertebrate LAT1, LAT2 and $b^{0,+}$AT, could generate similar protein folding and activity to that of human homologues.

In order to check the ability of these cells to express the different components of the heterodimer, the initial trials were made expressing separately the heavy and the light subunits. GFP was fused to the C-terminus (Figure 13B and 13D) or to the N-terminus (Figure 13A and 13C) of both proteins (see Materials and Methods section 5.1.1). We placed GFP in both termini of each subunit to look for the optimal position in which GFP does not affect protein expression or integrity, and later, the formation of the heterodimer.

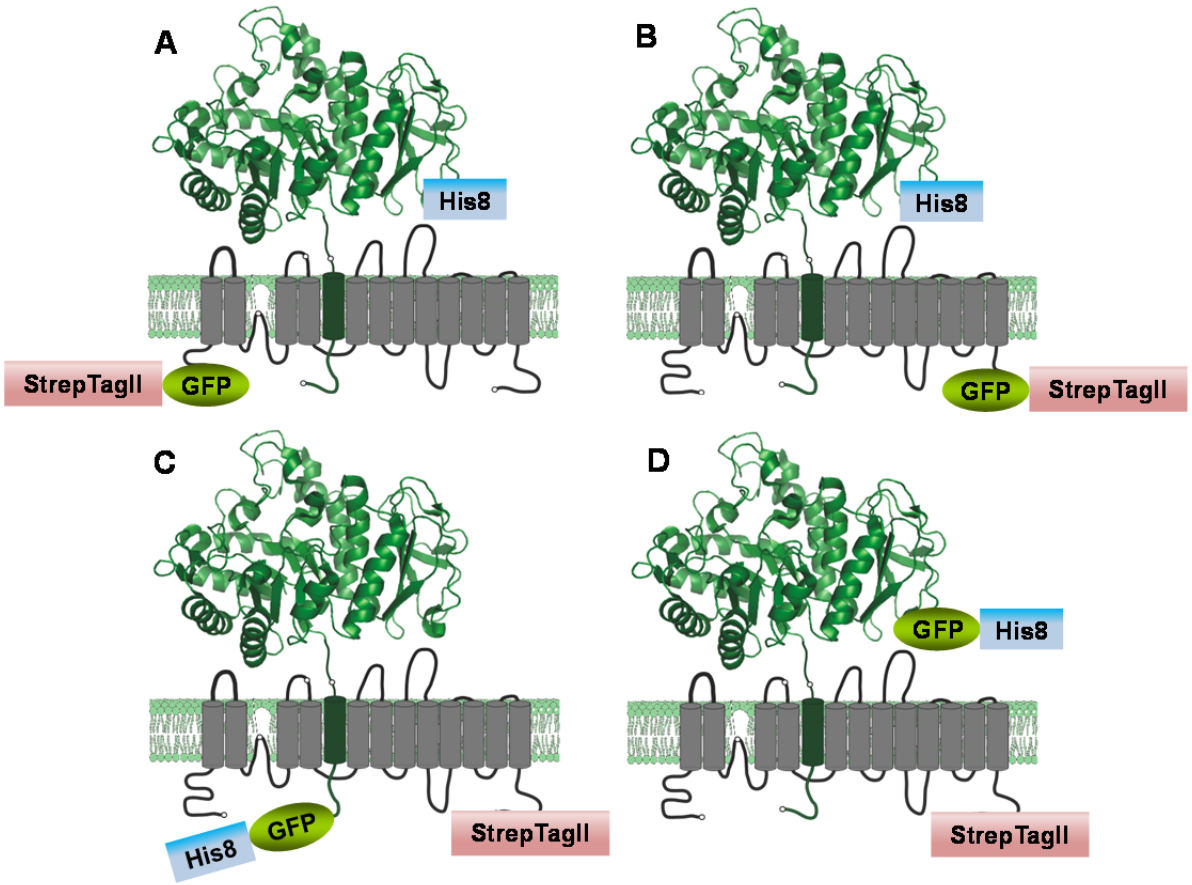

Figure 13. Representation of the GFP-fusion heterodimers intended to be expressed in tsA201 cells. GFP-Strep-tag II was fused to the N-terminus (**A**) or C-terminus (**B**) of light subunits (LAT1, LAT2 and $b^{0,+}$AT), and GFP-His8 was placed in the N-terminus (**C**) or C-terminus (**D**) of heavy subunits (rBAT and 4F2hc). The GFP-fused light subunits would be intended to be co-expressed with their corresponding heavy subunit with a His8 at the C-terminus (**A, B**), and GFP-fused heavy subunits would be intended to be co-expressed with their corresponding light subunits with a Strep-tag II at the C-terminus (**C, D**), as indicated in the figure. Heavy and light subunits are colored green and grey, respectively.

rBAT was expressed with the GFP at the N- or C-terminal at small-scale, and analyzed by FSEC just after membrane protein solubilization in 1 % of detergent plus 0.2 % CHS (see Materials and Methods sections 5.1.2 and 5.1.3). The detergents tested were DDM, DM and C_{12}MNG (Figure 14). DDM and DM are the two detergents more extensively used to extract membrane proteins and to maintain them stable in solution. Despite of having long alkyl chains and therefore form large micelles that can obstruct crystallization, around half of the eukaryotic membrane proteins crystals have been obtained in DDM or DM (He *et al.* 2014). Moreover, close to 82% of the crystallized membrane proteins have been extracted using one of these two detergents.

rBAT with GFP fused to the C-terminus (rBAT-GFP) showed complete proteolysis, being observable by FSEC a single protein peak in the elution volume of GFP (results not shown). On the other hand, expression of rBAT with the GFP at the N-terminus (GFP-rBAT), produced two main peaks in all the detergents tested (Figure 14A-C). In all cases, the protein peak of smallest molecular weight corresponds to free GFP, as it has been demonstrated by the superposition of the FSEC profiles corresponding to GFP-rBAT solubilized in DDM/CHS and

purified GFP (Figure 14D). This result indicates that part of the GFP-rBAT produced in this cell line is proteolyzed. Additionally, it is evident, by the trace of the first main peak, that the non-proteolyzed GFP-rBAT shows different aggregation stages in all the detergent conditions (Figure 14A-C), suggesting low stability of this protein.

The poor stability of vertebrate rBAT could be indicating the necessity of other factors that stabilizes the protein. In fact, it has been found that human rBAT without its light subunit, $b^{0,+}AT$, is rapidly degraded in the endoplasmic reticulum in mammalian cells, requiring assembly with $b^{0,+}AT$ before entering in contact with the calnexin chaperon system for its proper folding and heterodimer maturation (Bauch and Verrey 2002; Bartoccioni *et al.* 2008; Rius and Chillaron 2012). It is possible that this vertebrate rBAT also needs to assemble with its corresponding $b^{0,+}AT$ for folding and maturation. This could be the reason why it has been impossible to obtain vertebrate rBAT (GFP-rBAT) without significant aggregation and proteolysis in human cells (Figure 14).

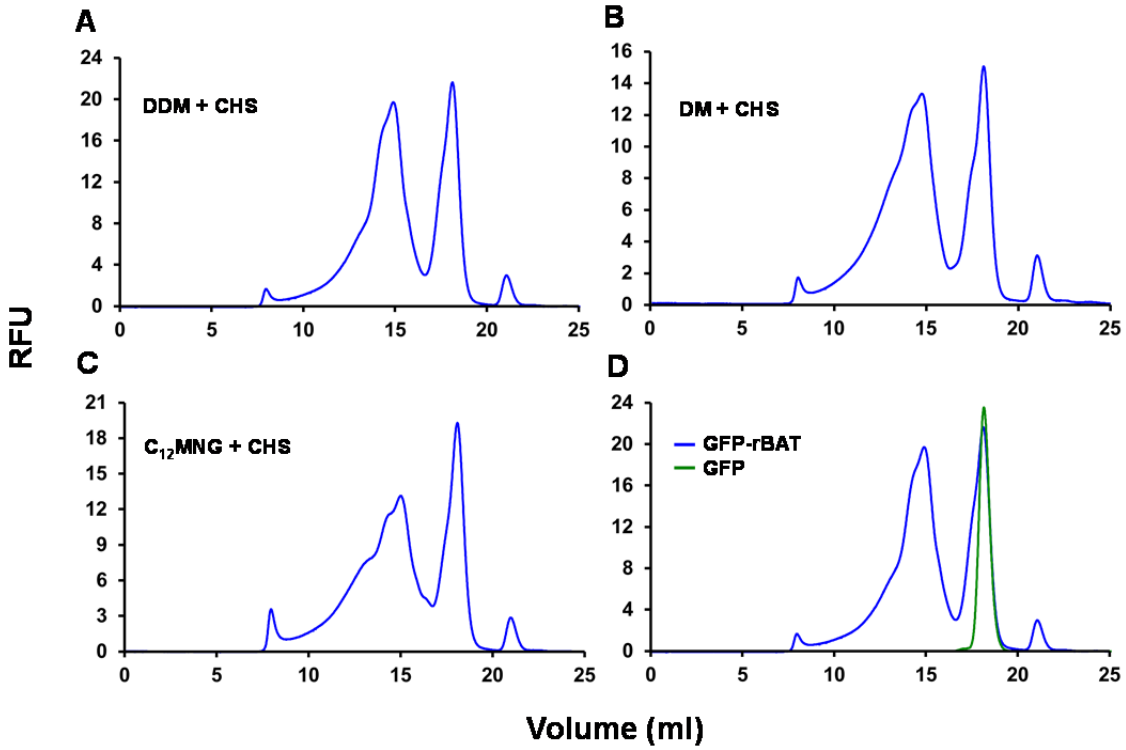

Figure 14. FSEC profiles of GFP-rBAT solubilized from tsA201 cells membranes in different detergents. rBAT with His8-GFP placed in the N-terminus (GFP-rBAT) was produced in tsA201 cells. This protein was solubilized in 1 % of DDM, DM or C_{12}MNG (**A**, **B** and **C**, respectively), adding 0.2 % CHS in all cases. 100µl of the soluble fraction were run in a Superose 6, 10/300 GL column. The superposition of the FSEC profile corresponding to the DDM + CHS condition with the FSEC trace of purified GFP (**D**), indicates that the higher peak present in all the solubilization conditions corresponds to free GFP and close molecular weight proteolysis products containing GFP. 100µl of purified GFP (0.013 mg/ml) were loaded in a Superose 6, 10/300 GL column. DDM: n-Dodecyl-β-D-Maltoside. DM: n-Decyl-β-D-Maltoside. C_{12}MNG: Lauryl maltose-neopentyl glycol. RFU: Relative Fluorescence Units.

As tested by FSEC just after solubilization in DDM/CHS (see Materials and Methods sections 5.1.2 and 5.1.3), 4F2hc cloned with the GFP at the N-terminus (GFP-4F2hc) produced only proteolyzed protein with a single FSEC peak in the elution volume of GFP (data not shown). However, for 4F2hc fused to GFP at the C-terminus (4F2hc-GFP), although showing similar

proteolysis level to that observed in GFP-rBAT (Figure 15), the peak corresponding with the non-proteolyzed 4F2hc-GFP was much narrower (Figure 15B), indicating a higher stability of 4F2hc-GFP in compare with GFP-rBAT.

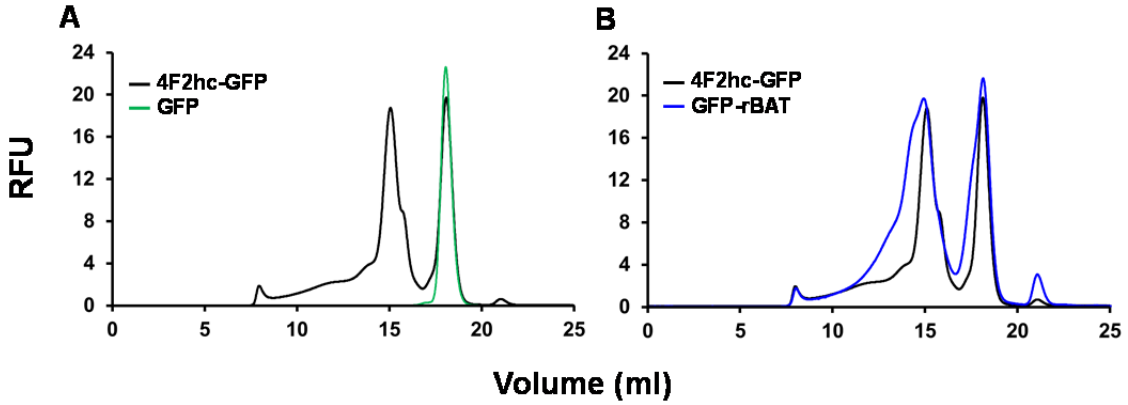

Figure 15. FSEC traces of solubilized 4F2hc-GFP expressed in tsA201 cells. 4F2hc with GFP-His8 placed at the C-terminus (4F2hc-GFP) was produced in tsA201 cells. The FSEC profile of 4F2hc-GFP solubilized with 1% DDM/0.2% CHS (**A**) was significantly better in comparison with the obtained for GFP-rBAT in the same detergent condition (B), although free GFP is also detected (**A, B**). 100µl of the soluble fraction of the tsA201 cells membranes, and 100µl of purified GFP (0.013 mg/ml), were run in a Superose 6, 10/300 GL column. RFU: Relative Fluorescence Units.

The expression of GFP-tagged versions, at the N- or C- terminal, of vertebrate LAT1, LAT2 and $b^{0,+}$AT was attempted in tsA201 cells. Cellular membranes were solubilized in DDM/CHS (1 %/0.2 %) and analyzed by FSEC (see Materials and Methods sections 5.1.2 and 5.1.3). No full protein was detected, but peaks corresponding to the GFP and proteolysis products containing GFP (results not shown). For this reason, in order to generate heterodimers, non-GFP-tagged versions of these light subunits were co-expressed with their corresponding GFP-tagged heavy subunits. Thus, the expression of the heterodimers shown in Figure 13A and 13B was discarded, and co-expression of GFP-rBAT with $b^{0,+}$AT-Strep-tag II (Figure 13C), and of 4F2hc-GFP with LAT1-Strep-tag II or LAT2-Strep-tag II (Figure 13D) was tested (see Materials and Methods sections 5.1.2 and 5.1.3).

Co-expression of GFP-rBAT with $b^{0,+}$AT-Strep-tag II in tsA201 cells did not generate vertebrate heterodimer. When tsA201 cells co-expressing both subunits were solubilized with DDM/CHS and analyzed by FSEC, the resulting chromatogram was very similar to that of GFP-rBAT when is expressed alone (data not shown). Since the condition DDM/CHS was the one in which GFP-rBAT showed the best FSEC behavior (Figure 14), although not good, DM/CHS and C_{12}MNG/CHS were not tested. Thus, it seems that GFP-rBAT is not able to interact with $b^{0,+}$AT-Strep-tag II when expressed in tsA201 cells.

On the other hand, the same experiment but co-expressing 4F2hc-GFP with LAT1-Strep-tag II or LAT2-Strep-tag II (Figure 16A and 16B, respectively) produced a shift towards the left in the elution volume of the fluorescent peak, in comparison with the peak corresponding to 4F2hc-GFP alone. This result indicates formation of heterodimer in both cases. However, a second peak of protein, eluting at the same volume than free 4F2hc-GFP, indicates that some 4F2hc-GFP is not forming heterodimer with LAT1-Strep-tag II or LAT2-Strep-tag II (Figure 16). Furthermore, a peak of proteolyzed protein eluting at the same volume than free GFP is still observed (Figure 16).

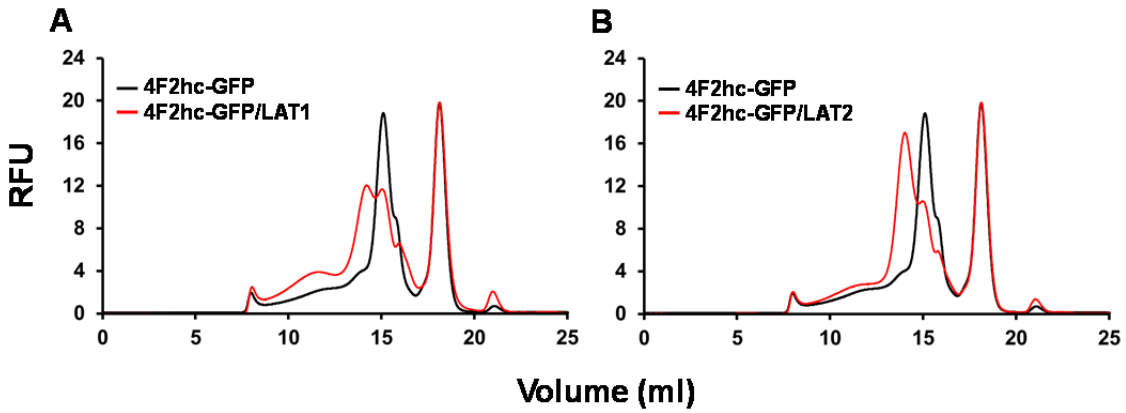

Figure 16. 4F2hc-GFP/LAT1 and 4F2hc-GFP/LAT2 formation in tsA201 cells. Co-expression of 4F2hc-GFP with LAT1-Strep-tag II (**A**) or with LAT2-Strep-tag II (**B**) generated a shift towards the left in the fluorescent peak of 4F2hc-GFP, indicating the formation of 4F2hc-GFP/LAT1 and 4F2hc-GFP/LAT2 heterodimers, respectively. Free 4F2hc-GFP and GFP are also observed in both FSEC traces (**A, B**). tsA201 cells membranes were solubilized in DDM/CHS, and 100µl of the soluble fraction were run in a Superose 6, 10/300 GL column for each co-expression. RFU: Relative Fluorescence Units.

These results suggest that, although tsA201 cells are able to generate 4F2hc-GFP/LAT1 and 4F2hc-GFP/LAT2 heterodimers, the expression of these proteins with a GFP tag, as an expression reporter and monitoring marker of the protein stability, could be inappropriate. The main reason is the high level of proteolysis affecting both heterodimers, illustrated by the large amount of free GFP generated respect to the amount of heterodimer produced (Figure 16). For this reason, other expression systems broadly used for producing proteins for structural studies were subsequently tested in the present study.

3.2 Expression of 4F2hc/LAT1, 4F2hc/LAT2 and rBAT/$b^{0,+}$AT vertebrate heterodimers in insect cells

Nowadays, baculovirus-insect cell expression system is being considered the main method for obtaining metazoan integral membrane proteins for structural studies (Contreras-Gomez *et al.* 2014; He *et al.* 2014). In fact, protein expression using baculovirus-insect cells has led to the major number of structures of metazoan integral membrane proteins in comparison with the other most used expression systems (He *et al.* 2014). These structures include receptors, ion channels, and with less but significant representation, solute transporters (Feng *et al.* 2010; Levin *et al.* 2012; Shintre *et al.* 2013; Deng *et al.* 2014; Sun *et al.* 2014). For full references of representatives protein structures obtained in insect cells see the review of He and co-workers (He *et al.* 2014). It is worth to mention that more than a half of these proteins are from human, indicating the suitability of this system for producing human-like active proteins.

Additionally, the methodology including, generation of the recombinant bacmid for insect cells transfection, production of baculovirus for infecting the cells and expressing the protein, and insect cell culture, are relatively quick (Ciccarone *et al.* 1998; Contreras-Gomez *et al.* 2014). There are two most widely used insect cell lines for expression of integral membrane proteins, Sf9 from *Spodoptera frugiperda* and Hi5 from *Trichoplusia ni*. However, more than 80% of the structures of integral membrane proteins were generated by using Sf9 line (He *et al.* 2014), which was also the one used in the present work to express 4F2hc/LAT1, 4F2hc/LAT2 and rBAT/$b^{0,+}$AT.

Similarly to what was done in tsA201 human cells, the different components of the heterodimer were expressed individually in order to check their expression levels and stability. The heavy (rBAT and 4F2hc) and the light ($b^{0,+}$AT, LAT1 and LAT2) subunits were cloned into pFastBac plasmid vector, where GFP was previously cloned (see Materials and Methods section 5.2.1). In all cases, these subunits were cloned to be expressed fused to GFP at the C-terminus (Figure 13B and 13D) or at the N-terminus (Figure 13A and 13C). Finally, formation of heterodimer between the best GFP-tagged subunit with their corresponding partner, without GFP, was also tested, with the aim to obtain a heterodimer appropriate to be undertaken in structural studies.

Recombinant baculovirus for each construct where generated by the Bac-to-Bac system (see Materials and Methods section 5.2.2). Initially, the P2 baculovirus was used to test at small-scale the expression levels, integrity and stability of GFP-tagged subunit separately (see Materials and Methods section 5.2.3). As it was observed in tsA201 human cells (Figures 14 and 15), only the expression of rBAT with GFP fused to the N-terminal (GFP-rBAT), and 4F2hc with GFP at the C-terminal (4F2hc-GFP), produced not fully degraded protein in insect cells (Figure 17A and 17B, respectively). After Sf9 cells membranes solubilization with 1 % DDM and 0.2 % CHS, soluble fraction was subjected to FSEC. It can be observed in the chromatograms that both proteins behave better in solution when they are expressed in insect cells than in mammal cells (compare Figure 17 with Figure 15B). However, 4F2hc-GFP continue to be more stable in DDM/CHS than GFP-rBAT, as observed by the presence of a major narrow peak of 4F2hc-GFP, less aggregation species, and lower proteolysis peaks in comparison with the GFP-rBAT FSEC chromatogram (Figure 17).

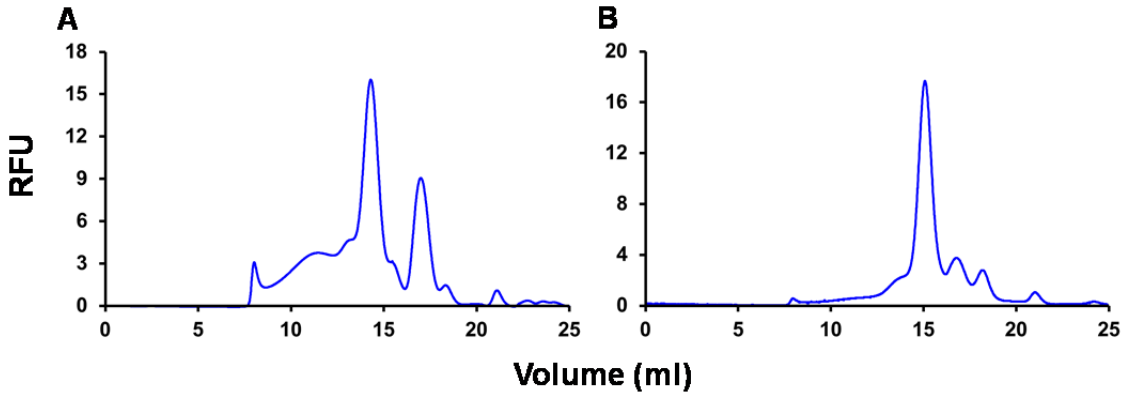

Figure 17. FSEC profiles of vertebrate rBAT and 4F2hc heavy subunits solubilized from Sf9 insect cells.
rBAT with His8-GFP at the N-terminus (**A**) and 4F2hc with GFP-His8 at the C-terminus (**B**) were expressed in Sf9 cells. Cellular membranes were solubilized in 1 % DDM and 0.2 % CHS. 100 µl of the soluble fraction were run in a Superose 6, 10/300 GL column in both cases. FSEC profiles of both heavy subunits were better than the ones obtained when both proteins were expressed in tsA201 cells. Similarly to the mammalian cells, in insect cells, 4F2hc FSEC profile showed more monodispersity and reduced proteolysis than the corresponding to rBAT. RFU: Relative Fluorescence Units.

Regarding the vertebrate light subunits, they were also expressed with GFP for being able to compare rapidly their expression levels, integrity and stability. Expression of LAT1, LAT2 and $b^{0,+}$AT with GFP fused at the N- or at C-terminus was tested. Sf9 cells membranes were solubilized in 1 % DDM/0.2 % CHS, and the soluble fraction was subjected to FSEC. For all light subunits, when the GFP was placed at the N-terminus, only some peaks corresponding to proteolytic fragments of low molecular weight were detected (data not shown). However, when the GFP was fused to the C-terminal of these light subunits (LAT1-GFP, LAT2-GFP and $b^{0,+}$AT-

GFP), FSEC showed peaks corresponding to the entire protein without proteolysis peaks (Figure 18).

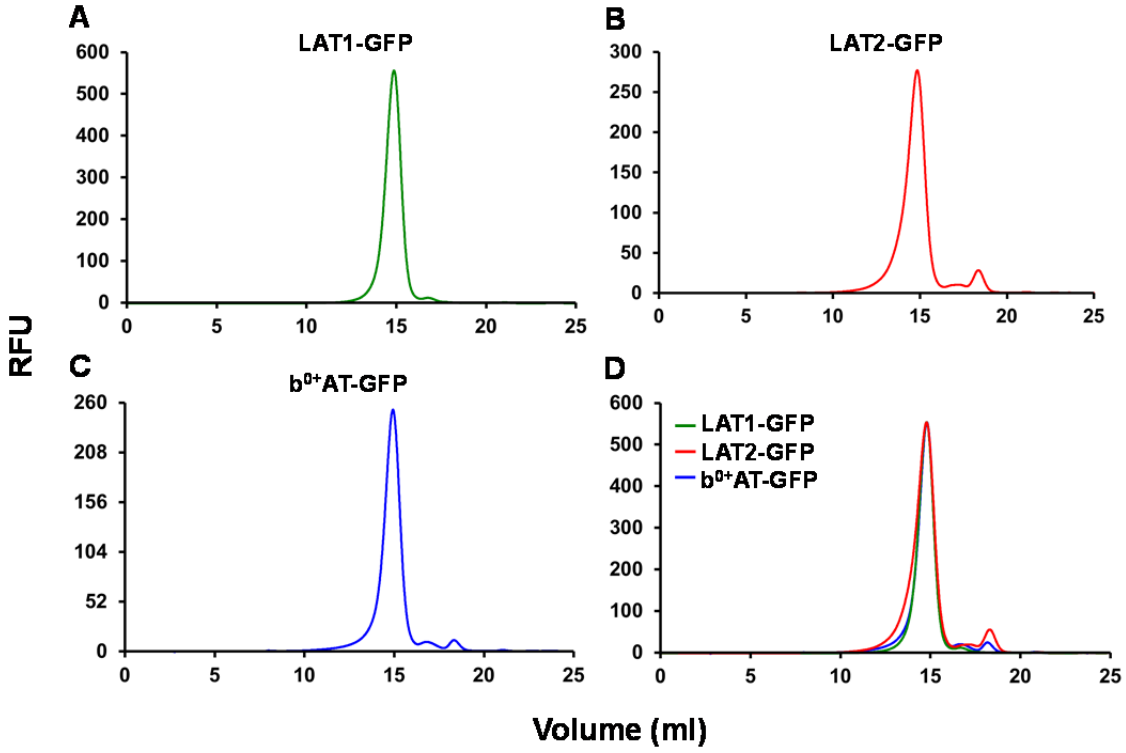

Figure 18. FSEC profiles of vertebrate LAT1, LAT2 and $b^{0,+}$AT solubilized from Sf9 insect cells. The three light subunits were expressed with GFP-Strep-tag II fused to the C-terminus and solubilized in 1 % DDM/0.2 % CHS. 100µl of the soluble fraction were run in a Superose 6, 10/300 GL column in all cases. LAT1 (**A**), LAT2 (**B**) and $b^{0,+}$AT (**C**) showed good behavior by FSEC, however the peak that shows more monodispersity was that of LAT1 (**D**). RFU: Relative Fluorescence Units.

It can be observed in Figure 18, that all the light subunits analyzed with GFP at the C-terminal showed a major single narrow and symmetric peak, indicating good stability of these proteins solubilized in DDM/CHS. Deeper analysis of the three peaks revealed that the best ones were from LAT1-GFP and $b^{0,+}$AT-GFP (Figure 18D). They were more symmetric and narrower than LAT2-GFP, indicating less heterogeneity in the former samples and possible less tendency to aggregation. However, as stabilization effect of human 4F2hc on human LAT2 has been previously described (see Introduction section 1.7.3), it is also possible that a similar stabilization effect of vertebrate 4F2hc on vertebrate LAT2 takes place. For this reason, vertebrate LAT2 was not discarded at this point, since it remains mainly soluble when expressed without its corresponding 4F2hc, in contrast to human LAT2 that completely aggregates without its 4F2hc partner (see Figure 12A in Introduction section 1.7.3). Thus, combination of a more stable LAT2 (vertebrate) plus the stabilizing effect of its 4F2hc partner (vertebrate), could produce a good candidate for structural studies, by which this heterodimer must be also tested with the other two (4F2hc/LAT1 and rBAT/$b^{0,+}$AT). Despite that LAT1-GFP and $b^{0,+}$AT-GFP peaks are quite similar, the peak of $b^{0,+}$AT-GFP is slightly wider, which could be indicating more polydispersity than LAT1-GFP. However, the difference is so small that could be due to an intrinsic effect of this experiment in particular, being possible slightly variations in the results between LAT1 and $b^{0,+}$AT in other tests.

The next step in this study was trying to express properly the full vertebrate heterodimers using the GFP-tagged subunits. Like in this system, GFP-rBAT and $b^{0,+}$AT-GFP were expressed (Figure 17A and 18C), heterodimer formation between rBAT and $b^{0,+}$AT was tested by using two different combinations: rBAT/$b^{0,+}$AT-GFP and GFP-rBAT/$b^{0,+}$AT (Figure 19A and 19B, respectively).

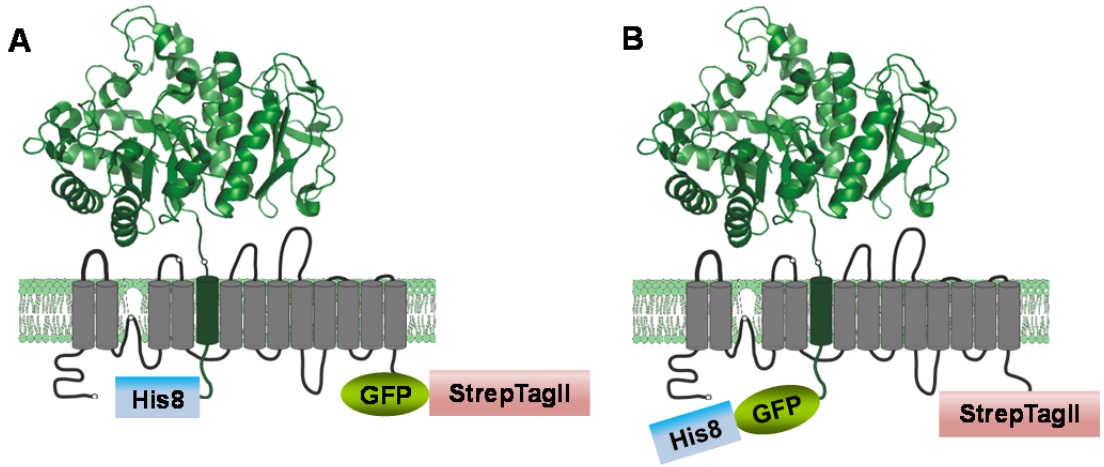

Figure 19. GFP tagged versions of rBAT/$b^{0,+}$AT heterodimer for expression in Sf9 insect cells. rBAT with a His8 at the N-terminal was co-expressed with $b^{0,+}$AT with a GFP-Strep-tag II at the C-terminus (**A**). rBAT fused to His8-GFP at the N-terminus was co-expressed with $b^{0,+}$AT with a Strep-tag II at the C-terminus (**B**).

Sf9 cells intended for the expression of the two versions of rBAT/ $b^{0,+}$AT were solubilized in 1 % DDM/0.2 % CHS, and the soluble fractions were subjected to FSEC (see Materials and Methods section 5.2.3). Similar to what happened in tsA201 human cells (see section 3.1), no generation of rBAT/$b^{0,+}$AT heterodimer was observed using both combinations (data not shown). It is possible that GFP fusion to the N-terminal of rBAT, or to C-terminal of $b^{0,+}$AT, may affect the interaction with the corresponding subunit, by altering at some extent a relevant region for the interaction between both subunits. Since a GFP-tagged rBAT/$b^{0,+}$AT heterodimer was not formed in insect nor in human cells, this complex was discarded as a candidate for structural studies.

Formation of 4F2hc/LAT1 and 4F2hc/LAT2 was investigated by using the GFP-tagged versions of the proteins that were successfully expressed in Sf9 cells: 4F2hc-GFP, LAT1-GFP and LAT2-GFP, as indicated in Figure 20. In this way, the generation of these heterodimers was tried by these combinations: 4F2hc/LAT1-GFP and 4F2hc/LAT2-GFP (Figure 20A), and 4F2hc-GFP/LAT1 and 4F2hc-GFP/LAT2 (Figure 20B).

Sf9 cells were solubilized in 1 % DDM/0.2 % CHS, and the soluble fraction was subjected to FSEC (see Materials and Methods section 5.2.3). For 4F2hc/LAT1-GFP and 4F2hc/LAT2-GFP, the resulting chromatograms were identical to those of LAT1-GFP and LAT2-GFP when are expressed alone (Figure 18A and 18B). Thus, positioning of the GFP in the C-terminal of these light subunits could be hindering the interaction between heavy and light subunit like in the case of rBAT/$b^{0,+}$AT-GFP.

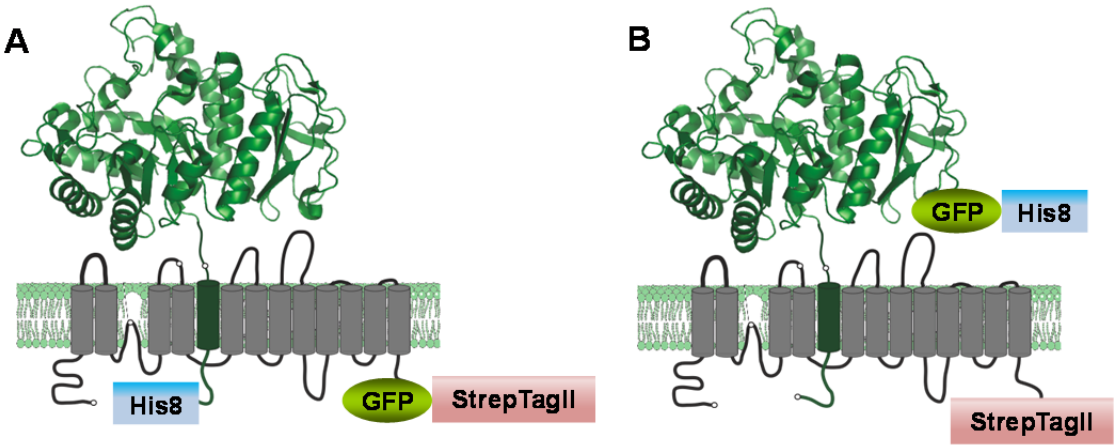

Figure 20. GFP tagged versions of 4F2hc/LAT1 and 4F2hc/LAT2 for expression in Sf9 insect cells. 4F2hc with a His8 at the N-terminal was co-expressed with LAT1 or LAT2 with a GFP-Strep-tag II at the C-terminus (**A**). 4F2hc fused to GFP-His8 at the C-terminus was co-expressed with LAT1 or LAT2 with a Strep-tag II at the C-terminus (**B**).

On the other hand, co-expression in Sf9 cells of 4F2hc-GFP with LAT1 or with LAT2, generated the appearance of a new peak toward the left in the FSEC profiles (Figure 21A and 21B, respectively), in comparison with the chromatogram corresponding to 4F2hc-GFP. The shift towards the left in the elution volume of the fluorescent peak suggests the formation of both heterodimers in Sf9 insect cells, as previously observed also in tsA201 human cells (Figure 16). However, less amount of proteolytic fragments is observable, suggesting that Sf9 insect cells are able to overexpress better these GFP-tagged heterodimers than tsA201 mammalian cells.

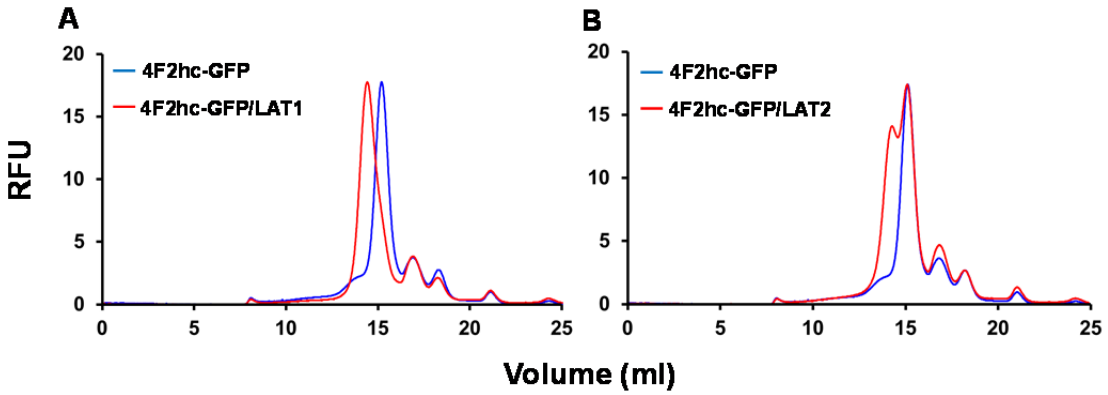

Figure 21. 4F2hc-GFP/LAT1 and 4F2hc-GFP/LAT2 formation in Sf9 cells. 4F2hc fused to GFP-His8 at the C-terminal (4F2hc-GFP) was co-expressed with LAT1 (**A**) and LAT2 (**B**) with a Strep-tag II at the C-terminus, in Sf9 insect cells. Cellular membranes were solubilized in 1 % DDM/0.2 % CHS, and 100 µl of the soluble fraction were run in a Superose 6, 10/300 GL column for each co-expression. Superposition of the FSEC profile corresponding to solubilized 4F2hc-GFP with the FSEC traces obtained co-expressing 4F2hc-GFP with LAT1 (**A**), or with LAT2 (**B**), showed the presence of a peak with an elution volume corresponding to a specie of higher molecular weight, suggesting that 4F2hc-GFP/LAT1 and 4F2hc-GFP/LAT2 heterodimers have been formed. RFU: Relative Fluorescence Units.

To confirm that the peak toward the left that appear in the FSEC chromatograms after co-expressing 4F2hc-GFP/LAT1 or 4F2hc-GFP/LAT2, corresponds to formed heterodimers, both putative complexes were purified by strep-tactin affinity chromatography from solubilized membranes of Sf9 cells co-expressing both subunits, taking advantage of the Strep-tag II added to the C-terminal of the light subunits (Figure 20B) (see Materials and Methods section 5.2.4).

Then, the purified sample was analyzed by FSEC, for the detection of the putative heterodimers by the GFP placed on the C-terminal of the heavy subunit 4F2hc (Figure 20B). In this way, FSEC traces comparable to those obtained without purification, after the DDM/CHS solubilization (Figure 21), would corroborate the formation of both heterodimers in Sf9 cells.

For doing this, membranes from insect cells co-infected with the corresponding P2 baculovirus to produce 4F2hc-GFP/LAT1 and 4F2hc-GFP/LAT2, were solubilized in 1 % DDM/0.2 % CHS. A part of the soluble fraction was subjected to FSEC (Figures 22A and 23A, respectively), and the remaining part was strep-tactin affinity purified, eluting the protein with D-desthiobiotin in 0.05 % DDM/0.01 % CHS. Then, the eluted samples were also analyzed by FSEC (Figures 22C and 23C, for 4F2hc-GFP/LAT1 and 4F2hc-GFP/LAT2, respectively).

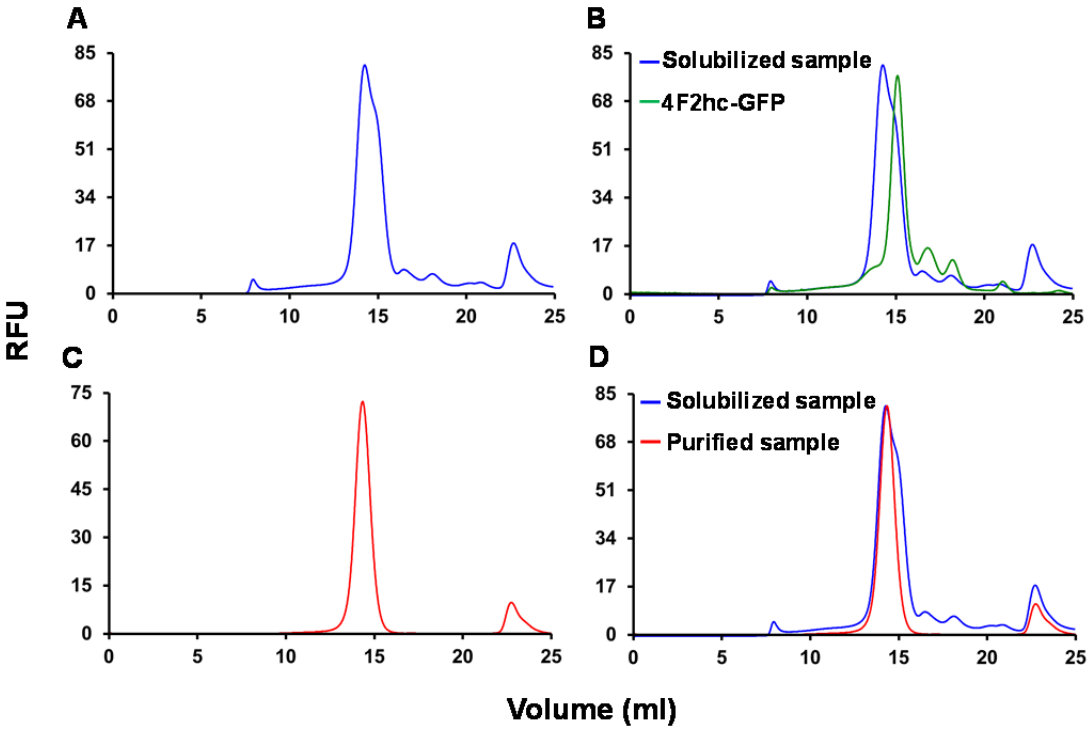

Figure 22. Validation of 4F2hc-GFP/LAT1 generation by strep-tactin affinity chromatography and FSEC. Sf9 cells co-expressing 4F2hc-GFP and LAT1 with a Strep-tag II at the C-terminus, were solubilized in 1 % DDM/0.2 % CHS. 100µl of the soluble fraction were run in a Superose 6, 10/300 GL column and analyzed by FSEC (**A**). The main peak would correspond with the putative heterodimer and the secondary one with free 4F2hc-GFP (**B**). The remaining part of the sample was subjected to strep-tactin affinity chromatography, and 100µl of the eluted protein, in presence of 0.05 % DDM/0.01 % CHS, were analyzed by FSEC (**C**). Comparison of the FSEC profiles of solubilized and strep-tactin purified samples (**D**) indicates that both main peaks correspond to the same molecular specie, corroborating the 4F2hc-GFP/LAT1 formation in Sf9 cells. RFU: Relative Fluorescence Units.

Superposition of the FSEC chromatograms of the solubilized and purified samples, 4F2hc-GFP/LAT1 (Figure 22D) and 4F2hc-GFP/LAT2 (Figure 23D), showed that the major peak obtained by FSEC after DDM/CHS solubilization of Sf9 membranes co-expressing 4F2hc-GFP with LAT1, or with LAT2, corresponds with the respective heterodimer. In addition, superposition of the chromatograms of DDM/CHS-solubilized heterodimers with the FSEC profile of DDM/CHS-solubilized 4F2hc-GFP, showed that the "right shoulder" observed in the FSEC profiles of the solubilized heterodimers corresponds to free 4F2hc-GFP (Figures 22B and 23B).

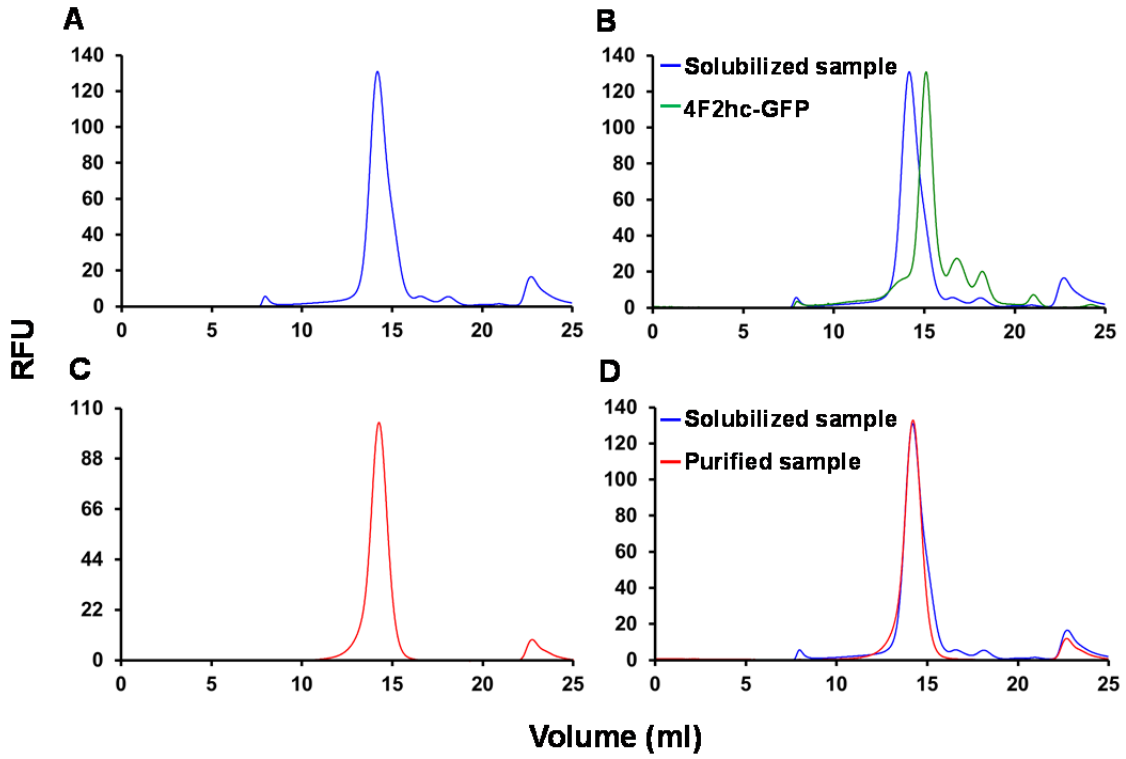

Figure 23. Validation of 4F2hc-GFP/LAT2 generation by strep-tactin affinity chromatography and FSEC. Sf9 cells co-expressing 4F2hc-GFP and LAT2 with a Strep-tag II at the C-terminus, were solubilized in 1 % DDM/0.2 % CHS. 100µl of the soluble fraction were run in a Superose 6, 10/300 GL column and analyzed by FSEC (**A**). The main peak would correspond with the putative heterodimer and the right shoulder with free 4F2hc-GFP (**B**). The remaining part of the sample was subjected to strep-tactin affinity chromatography, and 100µl of the eluted protein, in presence of 0.05 % DDM/0.01 % CHS, were analyzed by FSEC (**C**). Comparison of the FSEC profiles of solubilized and strep-tactin purified samples (**D**) indicates that both main peaks correspond to the same molecular specie, corroborating the 4F2hc-GFP/LAT2 formation. RFU: Relative Fluorescence Units.

Superposition of FSEC profiles of strep-tactin affinity purified 4F2hc-GFP/LAT1 and 4F2hc-GFP/LAT2 (Figure 24), showed narrow and practically symmetrical peaks, suggesting protein monodispersity. Additionally, both heterodimers are stable in solution after one round of affinity purification and reduction of the DDM/CHS concentration, from 1 %/0.2 %, used in the solubilization, to 0.05 %/0.01 %, used during purification. However, the peak of 4F2hc-GFP/LAT1 is slightly better than that from 4F2hc-GFP/LAT2. The latter is a little wider toward the left at the base, indicating the presence of aggregates, similar to the previously observed in the FSECs of solubilized LAT1-GFP and LAT2-GFP (Figure 18D), in which LAT1-GFP was significantly better than LAT2-GFP. These two results suggest that LAT1 could be a more stable protein than LAT2. Nevertheless, the peak of 4F2hc-GFP/LAT2 seemed to be a little more symmetrical and narrower than the peak of LAT2-GFP (compare traces in Figures 18B and 23C), which could be pointing to the stabilization effect of vertebrate 4F2hc on vertebrate LAT2, similar to what was observed in their human homologues (see Introduction section 1.7.3).

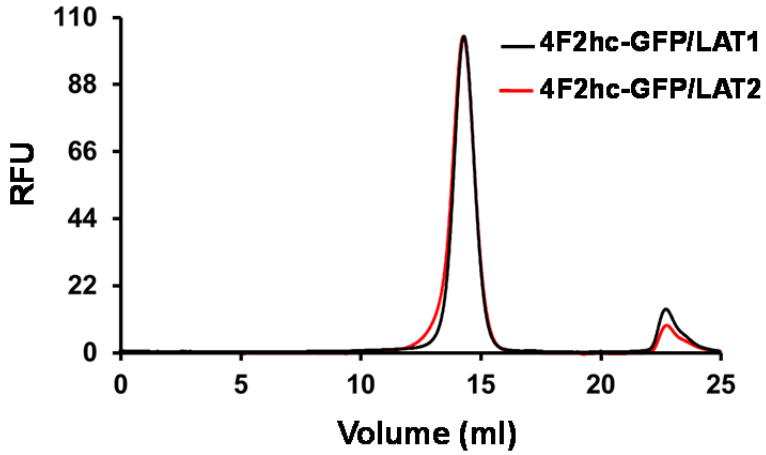

Figure 24. Comparison of FSEC profiles of purified 4F2hc-GFP/LAT1 and 4F2hc-GFP/LAT2. Sf9 insect cells membranes expressing both heterodimers (independently), with a Strep-tag II at the C-terminal of the light subunit, were solubilized in 1 % DDM/0.2 % CHS. Strep-tactin affinity chromatography was performed and proteins were eluted in presence of 0.05 % DDM/0.01 % CHS. 100 µl of the eluted protein were run in a Superose 6, 10/300 GL column and analyzed by FSEC for each heterodimer. The peak of 4F2hc-GFP/LAT1 indicates slightly less heterogeneity than that of 4F2hc-GFP/LAT2, suggesting that the first one is a little more stable than the second one at low concentrations. RFU: Relative Fluorescence Units.

3.2.1 4F2hc-GFP/LAT1 is more thermostable than 4F2hc-GFP/LAT2

To analyze more in deep which of both 4F2hc heterodimers, 4F2hc/LAT1 or 4F2hc/LAT2, was the most stable, a thermostability assay was performed using their GFP-tagged forms, 4F2hc-GFP/LAT1 and 4F2hc-GFP/LAT2 (see Materials and Methods section 5.2.5). Sf9 membranes were solubilized in 1 % DDM/0.2 % CHS, and both heterodimers were purified by strep-tactin affinity chromatography. Both eluted samples were left at the same Relative Fluorescence Units (RFU), and divided in aliquots of equal volume to ensure that all samples were at the same conditions during the experiment, except for the temperature challenge, and also to enable peak comparison. Each aliquot was subjected to a different temperature, as indicated in Figure 25, for 10 min, and after ultracentrifugation the soluble fraction was subjected to FSEC.

The normalized FSEC traces indicate the loss of monodispersity of 4F2hc-GFP/LAT1 and 4F2hc-GFP/LAT2 as they are subjected to a higher temperature (Figure 25A and 25B, respectively). However, the negative effect of the temperature increasing in the stability of 4F2hc-GFP/LAT1 is less profound than it is in 4F2hc-GFP/LAT2 (Figure 25A and 25B, respectively). This is more evident when the FSEC profiles at 4°C and 70°C, of each heterodimer, are superposed (Figure 25C and 25D, respectively), and when the FSEC traces of both complexes heated at 70°C, are compared (Figure 25E). FSECs profiles of 4F2hc-GFP/LAT1 are very similar from the minor (4°C) up to the major (70°C) temperature, with a little widening of the protein peak as temperature increases. On the other hand, FSECs chromatograms of 4F2hc-GFP/LAT2 get wider more evidently as temperature increases, indicating a highest tendency to aggregate in comparison with 4F2hc-GFP/LAT1. These results suggest that 4F2hc-GFP/LAT1 is more stable than 4F2hc-GFP/LAT2, which could be directly related to the stability of the light subunits. When vertebrate LAT1-GFP and LAT2-GFP were expressed alone, chromatographic profile of LAT2 was worse than that of LAT1 (Figure 18D). The melting temperature (Tm) for 4F2hc-GFP/LAT1, i.e., the temperature at which half of the

protein remains soluble, was of around 65°C. This value is high for a metazoan membrane protein indicating high stability.

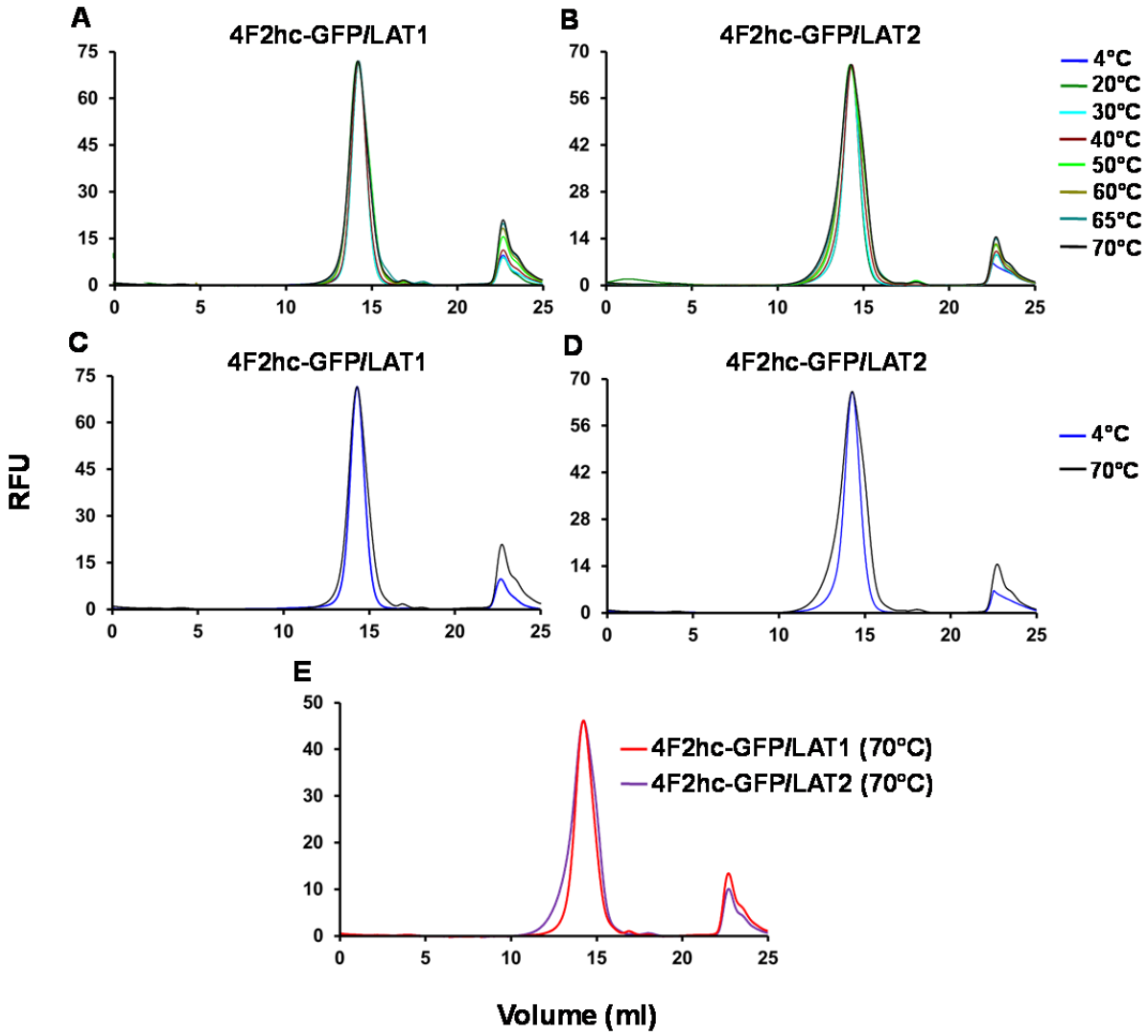

Figure 25. 4F2hc-GFP/LAT1 is more thermostable than 4F2hc-GFP/LAT2. Sf9 membranes expressing both heterodimers were solubilized in DDM/CHS. Strep-tactin affinity chromatography was performed taking advantage of the Strep-tag II placed at the C-terminus of both light subunits. The eluted heterodimers were left at the same concentration, and fractioned aliquots subjected to the temperatures indicated at the top right part of the figure. After ultracentrifugation, 100 µl of the soluble fraction were run in a Superose 6, 10/300 GL column and analyzed by FSEC. Normalized FSEC profiles of heated 4F2hc-GFP/LAT1 (**A**) and 4F2hc-GFP/LAT2 (**B**) indicated that 4F2hc-GFP/LAT1 monodispersity didn´t get significantly worst through the heating, but, in contrast, 4F2hc-GFP/LAT2 peak quality decreased significantly during heating. This fact can be also observed when normalized FSEC traces of 4F2hc-GFP/LAT1 (**C**) and 4F2hc-GFP/LAT2 (**D**), both non-heated or heated at 70°C, were compared, or when normalized FSEC profiles of both heterodimers heated at 70°C were superposed (**E**). RFU: Relative Fluorescence Units.

At this time, all results pointed out to vertebrate 4F2hc/LAT1 as the most stable heterodimer to start working with for structural studies. For this reason, from now and on, this study was focused on the strategies for obtaining at enough yield, pure and stable vertebrate 4F2hc/LAT1, as well as versions of this heterodimer suitable for structural studies by 3D crystallography.

3.2.2 Effect of different detergents on the stability of LAT1

We tried to find other appropriate detergent conditions to purify 4F2hc/LAT1, in which heterodimer show a behavior comparable or better than that observed when DDM/CHS is used. Finding other detergents in which a protein could be stable can be vital. The length and chemical characteristics of the alkyl chains, like the presence of aromatic rings, as well as the size and charge of the polar chain, influence the stability in solution of membrane proteins and their activity (Carpenter *et al.* 2008). Thus, a detergent in which the protein is more stable, monodisperse or active, could be found by screening the stability of a membrane protein in different detergents. Only the light subunit (vertebrate LAT1-GFP) was used in this screening, for making more stringent the process by avoiding the possible stabilizing effect of 4F2hc on LAT1, and taking into account that LAT1 has more transmembrane segments (twelve) than 4F2hc (one), being more compromised the stability of the heterodimer by the light subunit than by the heavy.

We compared four detergent + CHS conditions with the basal DDM + CHS condition. These detergents were DM, 6-cyclohexyl-1-Hexyl-β-D-Maltoside (Cymal-6), lauryldimethyl amine oxide (LDAO) and lauryl maltose-neopentyl glycol (C_{12}MNG). All of them have been broadly used in the crystallization of membrane proteins by virtue of their different biochemical properties, useful in the solubilization and stabilization of membrane proteins, as well as for crystals formation. Cymal-6 has been one of the smaller chain maltoside detergents used in 3D crystallography of membrane proteins (Long *et al.* 2007; Kowalczyk *et al.* 2011), due to its smaller micelles that can be more adequate for crystal formation and thus, for diffraction improvement (Carpenter *et al.* 2008). LDAO is known by its small micelle, that forms smaller belts around the transmembrane regions of the protein, in comparison with big micelle detergents like DDM and DM, which potentially allows more crystal contacts between the exposed polar surfaces of the membrane protein (Prive 2007). LDAO has been widely used in 3D crystallography for structural determination (Zhou *et al.* 2001; Shultis *et al.* 2006; Prive 2007), and furthermore, Cymal-6 and LDAO have been also used in the 3D crystallization of AdiC (Kowalczyk *et al.* 2011), the structural paradigm of LATs. C_{12}MNG is one of a series of maltose-neopentyl glycol detergents that has shown to increase the stabilization of membrane proteins in comparison with conventional detergents, as observed for several membrane protein systems, which structures has been solved by using MNG detergents (Chae *et al.* 2010)(Rosenbaum *et al.* 2011; Haga *et al.* 2012; Manglik *et al.* 2012; White *et al.* 2012; Suzuki *et al.* 2014).

To test the effect of these detergents on the stability of our transporter, LAT1-GFP was solubilized from Sf9 cells in 1 % of DDM, DM, C_{12}MNG, Cymal-6 or LDAO, in all cases supplemented with 0.2 % CHS, and the soluble fraction was subjected to FSEC (see Materials and Methods section 5.2.3) (Figure 26). The addition of CHS in all conditions was performed because it was determined the requirement of CHS for stabilizing this type of transporters, independently of the kind of detergent used (Álvarez-Marimon. 2014). Chromatograms showed that LAT1-GFP behaves similar in all the detergents tested, but LDAO (Figure 26D and 26F). This could be because LDAO, despite of favoring crystal contacts and improving diffraction for some membrane protein crystals, it has been also reported as a destabilizing detergent, being estimated that only 20 % of membrane proteins are resistant to its denaturing effects (Michel 2001).

Although, LAT1-GFP in C_{12}MNG + CHS behave similar than in the other detergents (DDM, DM and Cymal-6), the former peak showed a very slight widening at the base towards the left (Figure 26B and 26F), possibly indicating a major tendency to aggregation of LAT1-GFP in C_{12}MNG than in the other three detergents. However, due to the minimal difference, it is possible that this FSEC result does not mean a real difference in the stability of LAT1 using C_{12}MNG, or the other detergents that showed good behavior (DDM, DM and Cymal-6). These results suggest that other detergents with shorter hydrocarbon chain than DDM, like DM, even with half of the length of the DDM alkyl chain, like Cymal-6, or with additional physicochemical properties, like C_{12}MNG, could be used in combination with CHS for stabilizing LAT1 in solution, and probable 4F2hc/LAT1. Moreover, this result indicates the possible suitability of other detergents than DDM for the purification of 4F2hc/LAT1 in future experiments, in which DDM properties could negatively affect structure determination of this heterodimer, e.g., inability of 4F2hc/LAT1 to form crystals or poor diffraction of such crystals.

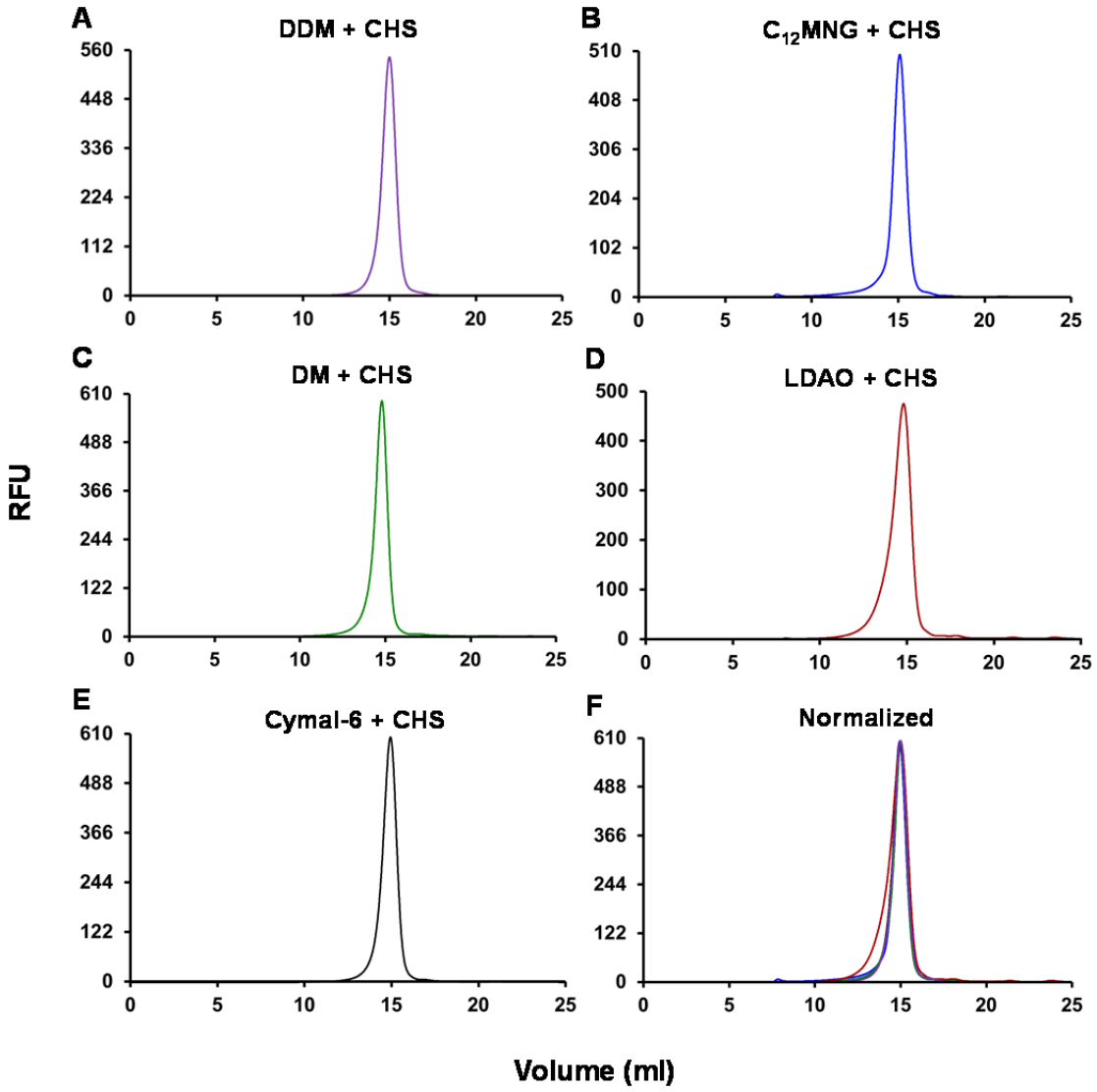

Figure 26. FSEC profiles of vertebrate LAT1 solubilized from Sf9 cells using several detergents. LAT1 was expressed in Sf9 insect cells with GFP-Strep-tag II fused to the C-terminus (LAT1-GFP), and cellular membranes were solubilized in 1 % (w/v) of the indicated detergent (DDM, C_{12}MNG, DM, LDAO or Cymal-6) plus 0.2 % CHS. 100 µl of the soluble fraction were run in a Superose 6, 10/300 GL column for each sample.

Chromatographic profiles of LAT1-GFP in all the detergents tested, except for LDAO (**D**), were very similar (see panel F for a comparison). Normalized peaks from **A** to **E**, in the same color code, are shown in panel **F**. LDAO: lauryldimethyl amine oxide. Cymal-6: 6-cyclohexyl-1-Hexyl-β-D-Maltoside. C_{12}MNG: lauryl maltose-neopentyl glycol. RFU: Relative Fluorescence Units.

3.2.3 4F2hc-GFP/LAT1 remains stable after double affinity chromatography

To test the yield and quality of purified 4F2hc-GFP/LAT1 produced in Sf9 cells, this heterodimer was expressed at large-scale and purified by two sequential affinity chromatographies (see Materials and Methods section 5.2.6). Insect cell membranes were isolated and solubilized in 1 % DDM/0.2 % CHS. For full purification of the heterodimer, strep-tactin affinity chromatography was performed as first step, and the eluted protein was subjected to immobilized metal ion affinity chromatography (IMAC). In Figure 27, FSEC profiles of samples recovered during the whole purification process are shown. "DDM+CHS solubilization" corresponds to the soluble fraction obtained after membrane solubilization in DDM/CHS (Figure 27A). Superposition of the chromatograms from solubilization, strep-tactin elution and Ni-NTA (IMAC) elution (Figure 27D), confirms the presence of heterodimer from the Sf9 membranes until the final purification step. Moreover, 4F2hc-GFP-LAT1 peaks showed monodispersity through the full purification process, suggesting a relevant stability of this heterodimer. Curiously, in contrast to the observed in small-scale cultures, a second peak of protein corresponding to free GFP-His8 appeared after solubilization, indicating that 4F2hc-GFP/LAT1 was affected by some proteolytic process during its expression at large-scale.

Western blot analysis using α-Strep-tag II antibody (see Materials and Methods section 5.2.6.1), detecting the heterodimer by the Strep-tag II placed at the C-terminus of LAT1, confirmed the presence of 4F2hc-GFP/LAT1 through the whole purification process (Figure 27E). Additionally, coomassie staining of purified 4F2hc/LAT1 showed integrity and purity of this heterodimer in the final sample. Addition of DTT to the former sample, caused the disassembly of heterodimer in 4F2hc-GFP and LAT1 subunits, by the reduction of the disulfide bond, confirming again the identity of the purified protein (Figure 27E, lane 4, and Figure 27F). Two bands are observed after DTT addition due to differences in the glycosylation of 4F2hc (Figure 27F).

The FSEC profile of purified and concentrated 4F2hc-GFP/LAT1 (0.1 mg/ml) showed a narrow and symmetric peak (Figure 27C), indicating stability and monodispersity of the protein. However, the purification yield was of around 0.075 mg per L of culture, a very little amount of protein to apply X-ray crystallography. Because the protein was at low concentration in the sample, it is unknown how stable 4F2hc-GFP/LAT1 is going to be after concentration at higher magnitudes, which is generally required for reaching desired protein densities for crystal growing. Thus, it is possible that the protein losses stability during concentration and, as a result, significant amount of protein would be lost, reducing more the final yield. For these reasons, it was decided testing the production of 4F2hc-GFP/LAT1 in other expression system, the methylotrophic yeast *Pichia pastoris* (see section 3.4).

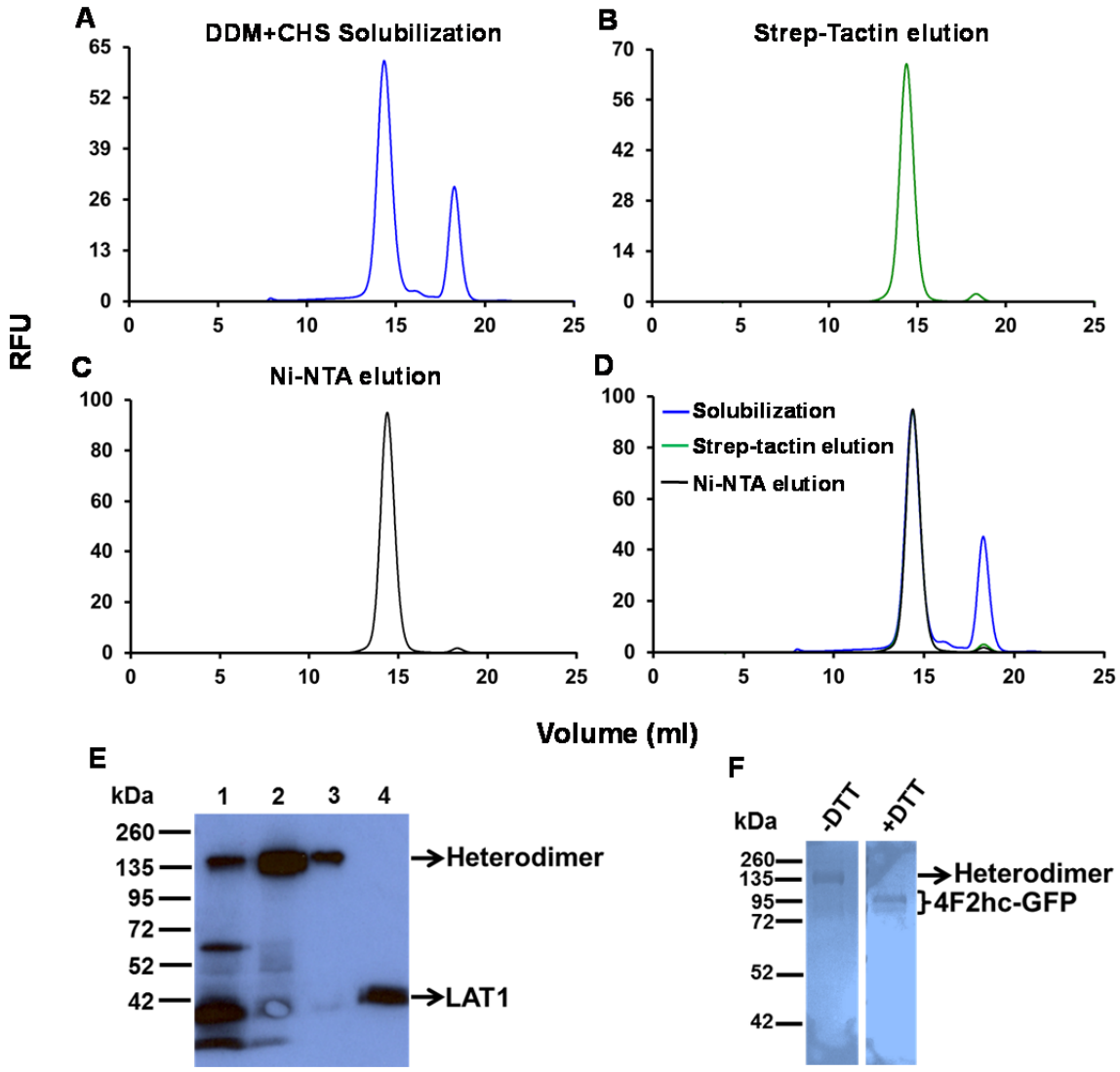

Figure 27. Purification of 4F2hc-GFP/LAT1 by sequential double-affinity chromatography. After heterodimer expression in Sf9 cells, 4F2hc-GFP/LAT1 solubilization was performed with 1 % DDM/ 0.2 % CHS from isolated cell membranes. As a first step, the heterodimer was purified by strep-tactin affinity chromatography taking advantage of the Strep-tag II present in the C-terminus of LAT1, and finally, purified by Immobilized Metal Ion Affinity Chromatography (IMAC) using the His8 tag placed in the C-terminus of 4F2hc-GFP. During the two chromatographies the concentration of DDM/CHS was reduced until 0.05 %/0.01 %. 100 µl of the solubilized sample and of both elutions (from strep-tactin and Ni-NTA resins) were run in a Superose 6, 10/300 GL column and analyzed by FSEC (**A**, **B** and **C**, respectively). The comparison of the FSEC profiles corresponding to the solubilized sample and both elutions shows the integrity and stability of the heterodimer during the whole purification process (**D**). 5 µl of the solubilized sample (lane 1), 5 µl of the strep-tactin elution (lane 2), and 5 µl from the Ni-NTA elution, without or with DTT (lanes 3 and 4, respectively), were analyzed by western blot using a α-Strep-tag II antibody (**E**). 39 µl from the Ni-NTA elution were run in a SDS-PAGE and the purity of the sample analyzed by coomassie blue staining (**F**).

3.3 Generating a less flexible and non-glycosylated vertebrate 4F2hc/LAT1

Sometimes, proteins have long N- or C- terminal ends, as well as long loops connecting their secondary structural elements (e.g. β-strands and α-helixes). Occasionally, these loops are related to the activity of the protein, for instance, in phosphatases some loops have been found to be implicated in their binding to inhibitors, as an activity regulatory mechanism during signaling pathways to maintain and increase the protein phosphorylation (Connor *et al.* 1999; Liu *et al.* 2004). However, generally, loops and N- or C-terminal ends in integral membrane proteins are not directly involved in the protein function, but in the correct insertion and translocation to the plasma membrane (van Geest and Lolkema 2000; Aseervatham *et al.* 2015). Nevertheless, these long unstructured regions can obstruct protein structural studies. At first instance, flexible parts of protein can hinder crystal formation or result in disordered crystal packing, leading to low resolution models. As a consequence, removal of long loops, N- and/or C-terminals, for structural studies, has produced functional protein cores that have allowed obtaining well-ordered crystals, from which their atomic structures have been solved (Hattori and Gouaux 2012; Penmatsa *et al.* 2013; Althoff *et al.* 2014). In fact, there is a study in which authors described how modifications performed to a glutamate receptor, among them, deletion of several residues in the C-terminus and in a linker domain, significantly increase structure resolution (Sobolevsky *et al.* 2009).

For this reason, truncated versions of vertebrate 4F2hc/LAT1 were investigated, with the aim of finding a less flexible heterodimer. Taking into account that LAT1-GFP, 4F2hc-GFP and 4F2hc-GFP/LAT1 were produced in Sf9 cells at enough amounts to compare expression levels of different versions by western blot, the first approximation to the truncated versions of LAT1 were made using insect cells, meanwhile preparing constructs to the first trials of expression in *Pichia pastoris*.

LAT1 topology was predicted by using the Hidden Markov Model for Topology Prediction (HMMTOP) server (Tusnady and Simon 1998). This light subunit was predicted to have a long cytoplasmic N-terminal (51 residues), and a shorter intracellular C-terminal (29 residues). Based on this topology prediction, and to find how many residues could be trimmed without losing expression, different truncated versions at the N- or C-terminal, or both, were cloned with a Strep-tag II placed in the C-terminal, and expressed in Sf9 cells (see Materials and Methods section 5.2.7). Western blot analysis, using α-Strep-tag II antibody, of Sf9 isolated membranes expressing the different truncated versions of LAT1, showed that LAT1 with up to 40 amino acid deletions at the N-terminus (Δ10Nter, Δ20Nter, Δ30Nter and Δ40Nter) could be successfully expressed (Figure 28A). However, not even 10 amino acids could be removed from the C-terminal without considerably affecting the expression of LAT1 in comparison with wild type LAT1 (wt) (Figure 28A). Significantly reduced expression of the double truncated versions (at the C- and N-terminal) was observed as well, probably due to the truncation of the C-terminal end (Figure 28A).

For knowing if LAT1(Δ40Nter) was able to form heterodimer with 4F2hc-GFP, LAT1(Δ40Nter) was co-expressed with 4F2hc-GFP in Sf9 cells as represented in Figure 28B. Green fluorescence of Sf9 cells, as observed by fluorescence microscope (Nikon TB200),

indicates the expression of 4F2hc-GFP (Figure 28C). In order to corroborate 4F2hc-GFP/LAT1(Δ40Nter) formation, isolated membranes from the cells previously checked at the microscope, were analyzed by western blot, using α-Strep-tag II and α-His tag antibodies (Figure 28D and 28E, respectively). The appearance of a band in the expected size using both antibodies, and its disappearance by the addition of DTT, demonstrated the generation of 4F2hc-GFP/LAT1(Δ40Nter) heterodimer in Sf9 cells (Figure 28D and 28E).

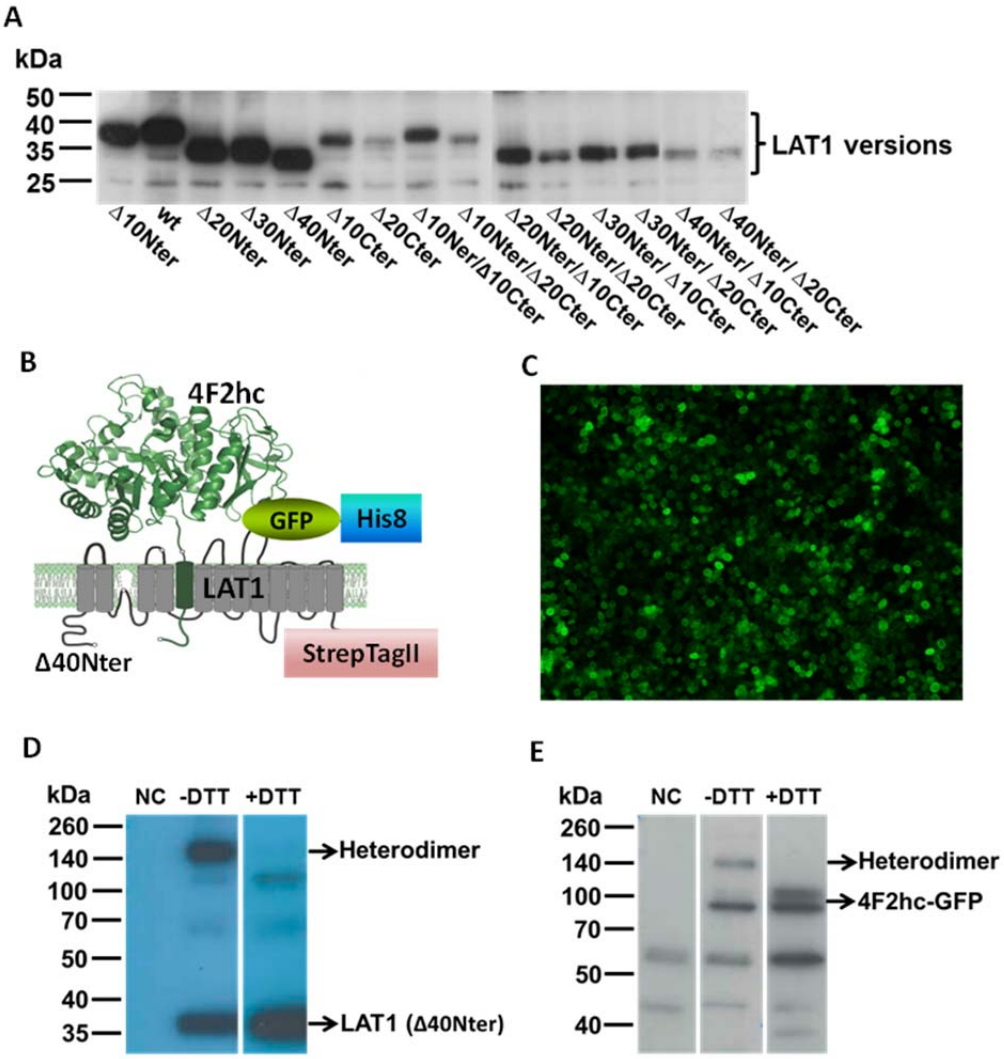

Figure 28. Sf9 cells express 4F2hc-GFP/LAT1(Δ40Nter) heterodimer. Baculoviruses carrying cDNA of wild-type LAT1 (wt), or of the different LAT1 truncations at the N- or C- terminal, or at both, were generated and used to infect Sf9 cells. In all cases, a Strep-tag II was placed in the C-terminus. Isolated Sf9 membranes were subjected to western blot using α-Strep-tag II antibody (**A**). 5 μg of total protein was loaded in each lane to screen the expression levels of the different LAT1 truncated versions. After the Δ symbol, the number of removed residues and the position of the deletion, N-terminal (Nter) or C-terminal (Cter), are indicated. Baculoviruses expressing 4F2hc-GFP and LAT1(Δ40Nter) were used to co-infect Sf9 cells (MOI: 2/2) to test if LAT1(Δ40Nter) formed heterodimer with 4F2hc-GFP. A schematic representation of the recombinant heterodimer expressed is shown (**B**). Sf9 cells co-infected with 4F2hc-GFP and LAT1(Δ40Nter) viruses viewed at the fluorescence microscope (**C**). Sf9 membranes expressing 4F2hc-GFP and LAT1(Δ40Nter) were analyzed 48 h post-infection by western blot using α-Strep-tag II and α-His tag antibodies (**D** and **E**, respectively). In both cases heterodimer formation was detected, and its dissociation was observed when the disulfide bond established between 4F2hc-GFP and LAT1(Δ40Nter) is

reduced with DTT. Free LAT1(Δ40Nter) (**D**) and free 4F2hc-GFP (**E**) are also observed. Membranes of uninfected Sf9 cells were used as negative control. NC: Negative control.

Despite the low yield obtained for 4F2hc-GFP/LAT1 in insect cells, small-scale screening of truncated versions of LAT1 has allowed identifying terminal regions relevant for its correct expression, and which region can be trimmed without affecting the heterodimer formation. Vertebrate 4F2hc has been predicted, by using the HMMTOP server (Tusnady and Simon 1998), to have a long cytoplasmic N-terminal (78 amino acids). Therefore, this subunit also could provide flexibility to 4F2hc/LAT1 heterodimer hindering structural studies with the entire heterodimer. For this reason, a version of 4F2hc/LAT1(Δ40Nter) with a deletion in the N-terminal of 4F2hc could be of great interest for being studied. For doing this, expression in Pichia was latter used (see section 3.4.1), due to the high yield of heterologous protein that is generally obtained by using this yeast (Macauley-Patrick *et al.* 2005).

Additionally, elimination or reduction of glycosylation in glycoproteins is other strategy used in crystallography for increasing protein homogeneity in the sample, and for avoiding interference of these sugars in the crystal contacts. The possibility of obtaining less glycosylated version of 4F2hc/LAT1 was also explored in the present study, by using Pichia as expression system (see section 3.4.1). Production of considerably less glycosylated glycoproteins in insect cells is quite inefficient, which in combination with the low yield of 4F2hc/LAT1 produced in them, could significantly reduce the final yield of protein, or even completely abolish the expression of 4F2hc. Additionally, non-glycosylated human 4F2hc has been successfully expressed in Pichia, being able to form heterodimer (Costa i Torres 2012). Therefore, it is probable that vertebrate 4F2hc behave similarly when expressed in Pichia, since both homologues share 47% amino acid identity among them.

3.4 Expression of vertebrate 4F2hc/LAT1 in *Pichia pastoris*

To investigate if 4F2hc/LAT1 could be produced stable and at higher levels than in Sf9 insect cells, expression and formation of this heterodimer was investigated in *Pichia pastoris*, owing to the advantages and properties of this expression system in the production of heterologous membrane proteins as described below.

Yeast is one of the most used expression systems for producing metazoan integral membrane proteins, being *Pichia pastoris* the most used, as it is considered the best expression system between yeast species (Cereghino and Cregg 2000). In fact, most of the crystallographic structures of metazoan integral membrane proteins produced in yeast, have been obtained by expressing proteins in *Pichia pastoris* (He *et al.* 2014). Pichia is relatively easy to manipulate genetically, it has the capability to produce proteins with eukaryotic post-translational modifications, and can reach higher cellular densities than *Saccharomyces cereviseae*, which improve protein yield (Macauley-Patrick *et al.* 2005). Additionally, reagents to prepare Pichia growth media are less expensive than those required by other expression systems, like mammalian and insect cells.

Some of the membrane proteins expressed in *P. pastoris*, which structure have been solved, were of human origin, like potassium and water channels, or the histamine receptor (Horsefield *et al.* 2008; Ho *et al.* 2009; Shimamura *et al.* 2011; Brohawn *et al.* 2012; Miller and Long 2012). This suggests the suitability of this system to produce in good condition proteins of higher eukaryotes. In fact, human 4F2hc/LAT2 heterodimer could be obtained in *Pichia pastoris*

allowing functional and structural studies (Costa *et al.* 2013; Meury *et al.* 2014; Rosell *et al.* 2014). Thus, it was feasible thinking that Pichia could correctly express, fold and assemble vertebrate 4F2hc/LAT1, which shares high identity with human homologue (78 % for LAT1 and 47 % for 4F2hc).

3.4.1 Expression of non-glycosylated and less flexible versions of 4F2hc/LAT1

Experiments in Sf9 insect cells showed that LAT1, with a deletion of 40 residues at the N-terminus, could be produced at expression levels comparable to the wild-type LAT1 (Figure 28A). Moreover, this deleted version was able to form heterodimer with 4F2hc-GFP (Figure 28D and 28E). Thus, for obtaining a less flexible and homogenous version of 4F2hc/LAT1, that potentially could form well-ordered crystal lattices diffracting at high resolution, we tried to obtain the less glycosylated version of 4F2hc as possible, and investigated its ability to form heterodimer with LAT1(Δ40Nter). Then, if successful, production of this heterodimer with a deletion of the long N-terminal region of 4F2hc would be tested to determine its stability in solution and its final yield.

3.4.1.1 Expression of less-glycosylated vertebrate 4F2hc–GFP versions

Glycosylation in proteins is variable in the type and number of sugars bound to the glycosylation target, as well as, the occupancy of such site is also variable between different molecules in the same protein sample. Such heterogeneity, and the flexibility of these long branched structures, can hinder crystal contacts impeding crystallographic studies. For this reason, attempts to reduce this heterogeneity are worth. One way to eliminate glycosylation is to treat proteins with glycosidases, as it has been reported in the case of the complex between gastric intrinsic factor (IF), cobalamin (Cbl, vitamin B12) and the cubilin IF-Cbl binding-region. The structure of this complex was solved at 3.3 Å resolution from protein produced in Pichia and treated with Endo H (Andersen *et al.* 2010). On the other hand, site directed mutagenesis has been usually used for eliminating glycosylation sites from glycoproteins, in order to improve the resolution of the crystal diffraction (Sobolevsky *et al.* 2009; Andersen *et al.* 2010).

The whole sequence of vertebrate 4F2hc was subjected to *N*-glycosylation prediction using the NetNGlyc 1.0 server tool (Table 6). Asparagines 228, 254, 264, 268, and 294 were predicted as putative targets for glycosylation. N228 (+++), N268 (++) and N294 (++) have the higher probability of being glycosylated, and N254 and N264 the lower probability (+). Sequence alignment between human 4F2hc and vertebrate 4F2hc, indicated that these asparagine residues would be exposed in the vertebrate 4F2hc ectodomain (ED), as observed by homology on the structure of human 4F2hc-ED (Fort *et al.* 2007), supporting the glycosylation prediction for these asparagines.

Some glycosylation can be relevant for protein folding and activity, by which, elimination of those glycosylation sites can compromise protein expression and stability. Thus, studies to find which of the vertebrate 4F2hc glycosylations could be removed from the protein without affecting its expression and stability were required. For doing this, expression screening of different less-glycosylated 4F2hc mutants was performed. These mutants were, a version of vertebrate 4F2hc in which the five asparagines predicted to be glycosylated (N228, N254, N264, N268 and N294), were mutated by glutamine, being denominated non-glycosylated 4F2hc (NG-

4F2hc), and mono-glycosylated versions of 4F2hc (Q228N, Q254N, Q264N, Q268N and Q294N), in which only the indicated asparagine has been recovered by site directed mutagenesis using NG-4F2hc as a template (see Materials and Methods section 5.3.1.1). NG-4F2hc, mono-glycosylated 4F2hc versions and wild-type 4F2hc, were cloned into pPICZ B expression vector with GFP-His8 fused to the C-terminal (see Materials and Methods section 5.3.1.1). Protein expression screening was performed by directly measuring the GFP-fluorescence of the cultures providing from each positive transformant clone (see Materials and Methods section 5.3.3), to determine the expressions levels.

SeqName	Position[1]	Potential[2]	Jury Agreement[3]	N-Glyc Result[4]
4F2hc	228 NVTE	0.7556	(9/9)	+++
4F2hc	254 NASS	0.5240	(5/9)	+
4F2hc	264 NMTA	0.6289	(6/9)	+
4F2hc	268 NYSK	0.6959	(9/9)	++
4F2hc	294 NETV	0.6896	(9/9)	++
4F2hc	313 NFSS	0.3740	(9/9)	--
4F2hc	402 NNSP	0.0966	(9/9)	---
4F2hc	414 NVTF	0.4880	(4/9)	-

Table 6. Predicted *N*-glycosylation sites in vertebrate 4F2hc. [1]Indicates the position in the protein of asparagines (N) in the N-X-S/T sequons that were analyzed looking for glycosylation sites. X is any amino acid except proline. [2]It is the averaged output of nine neural networks, thus any potential crossing the threshold of 0.5, represents a predicted glycosylation site. [3]Indicates how many of the nine neural networks support the prediction. [4]Glycosylation prediction. +++ indicates the higher glycosylation probability and + the lower. - - - indicates the higher non-glycosylation probability and - the lower.

All mono-glycosylated mutants of 4F2hc, but Q268N, were successfully expressed in Pichia (Figure 29C-G), showing expression levels similar to those obtained for wild type 4F2hc (Figure 29A). However, non-glycosylated 4F2hc expressed almost half of the wild type (Figure 29B). This can be observed more clearly in Figure 30A, where a comparison among the measured fluorescence of the most expressing clones, from each 4F2hc version, is shown. These results, suggest that the presence of at least one of these four glycosylation sites (N228, N254, N264 or N294) in vertebrate 4F2hc, could be relevant for 4F2hc expression in *Pichia pastoris*. However, these glycosylation sites do not seem to be indispensable for protein expression, since 4F2hc is still produced without the five targets of glycosylation (Figures 29B and 30A). This could be expectable, because as it has been said previously, a non-glycosylated human 4F2hc has been also expressed in Pichia (Costa i Torres 2012).

Membranes from *P. pastoris* clones that most express the indicated 4F2hc version (Figure 30A) were analyzed by α-His western-blot and by in-gel fluorescence, confirming the expression of the 4F2hc-GFP less glycosylated mutants (Figure 30B). Western blot band for wild type 4F2hc is wider than the other of 4F2hc less glycosylated versions (Figure 30B), although Q254N, Q264N and Q294N bands are wider than those from non-glycosylated 4F2hc and Q228N. This result suggests that heterogeneity in the sample has been reduced by elimination of the five glycosylation sites, or at least of four, in the case of the mono-glycosylated versions. Thus, the more homogenous samples regarding glycosylation seem to be non-glycosylated 4F2hc (NG-

4F2hc) and 4F2hc Q228N. The same can be also slightly observed by in-gel fluorescence (Figure 30B), since for the other mono-glycosylated mutants (Q254N, Q264N, Q294N) two bands seem to be observable. As a result, NG-4F2hc or 4F2hc Q228N could be suited for the co-expression with LAT1(Δ40Nter), however, we decided to prioritize non-glycosylated 4F2hc, investigating whether NG-4F2hc could be able to form heterodimer and if this heterodimer could be produced at enough yield for structural studies.

3.4.1.2 Expression of the non-glycosylated NG-4F2hc(Δ36Nter)/LAT1 heterodimer

As non-glycosylated 4F2hc (NG-4F2hc) could be expressed in Pichia*,* and previously a vertebrate LAT1 version with 40 amino acid deleted at N-terminus was expressed at comparable levels than the wild-type, being able to form heterodimer, 4F2hc-GFP/LAT1(Δ40Nter) in Sf9 cells (Figure 28), we tried to generate the NG-4F2hc/LAT1(Δ40Nter) heterodimer. In this way, a less heterogenic and flexible version of 4F2hc/LAT1 could be available for possible crystallographic studies.

Thus, pPICZ B vector carrying the cassettes for the co-expression of NG-4F2hc and LAT1(Δ40Nter) was used to transform Pichia (see Materials and Methods sections 5.3.2 and 5.3.3). Protein expression screening at small-scale was performed in 16 zeocin resistant transformants clones from two independent transformations, but 48 h after protein induction none of the clones expressed LAT1(Δ40Nter), but did 4F2hc (data not shown). The full pPICZ B-LAT1(Δ40Nter) construct was sequenced, however, LAT1(Δ40Nter) was in correct frame for its expression, and no mutations in the protein and Kozak sequences, nor in the promoter region were found. The cause of the lack of expression of this LAT1 version in Pichia remains unknown. Then, further studies expressing LAT1(Δ40Nter) in this yeast could not be performed.

On the other hand, vertebrate 4F2hc has been predicted to have a long cytoplasmic N-terminal (78 amino acids), longer than LAT1, which may provide extensive flexibility to this region of the protein, hindering structural studies with the entire heterodimer. For this reason, NG-4F2hc version with a deletion of 36 amino acids at the N-terminus (Δ36Nter) was tested for heterodimer formation with wild type LAT1. Only 36 residues were trimmed because these residues were not conserved in 4F2hc between several vertebrate species, which would suggest little or no involvement in protein insertion and translocation in the plasma membrane. If NG-4F2hc(Δ36Nter)/LAT1 would be stable in solution and expressed at enough yield, it would have considerable potential for 3D crystallography.

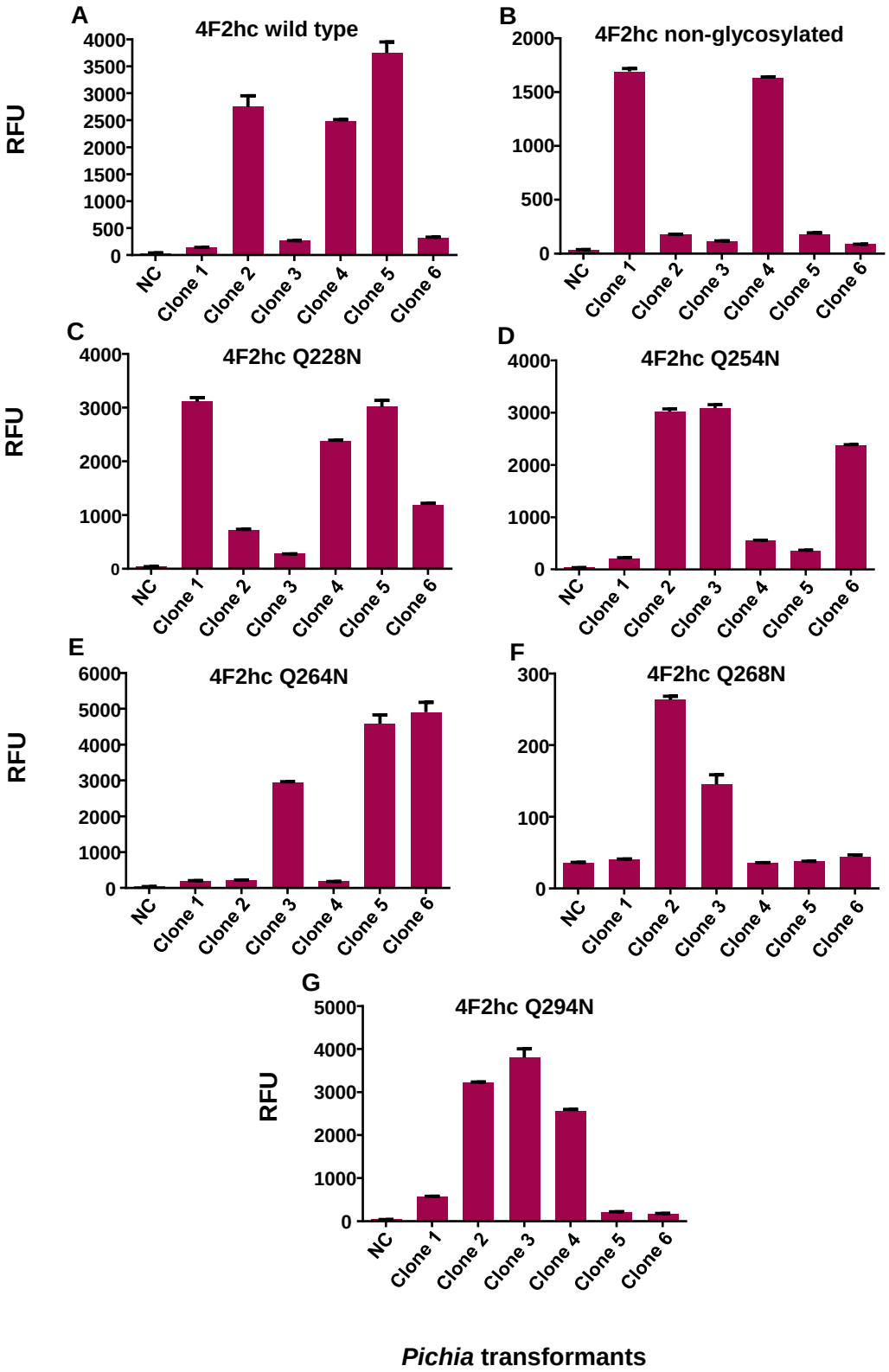

Figure 29. Expression screening of less-glycosylated versions of vertebrate 4F2hc in *Pichia pastoris*. Pichia strain KM71H was transformed with pPICZ B vector carrying the cDNA of wild type 4F2hc, 4F2hc with the five predicted targets of glycosylation eliminated (NG-4F2hc), and five 4F2hc mutants with only one of the five glycosylation targets (Q228N, Q254N, Q264N, Q268N and Q294N). All versions were expressed fused to GFP-

His8 at the C-terminus. Six zeocin resistant transformants of each construct were used to do the expression screening at small scale. Whole cell fluorescence measured after 48 h of protein expression induction for each analyzed clone is shown (**A-G**). As a negative control, Pichia strain KM71H transformed with empty pPICZ B was used (NC). Q: glutamine, N: asparagine. RFU: Relative Fluorescence Units.

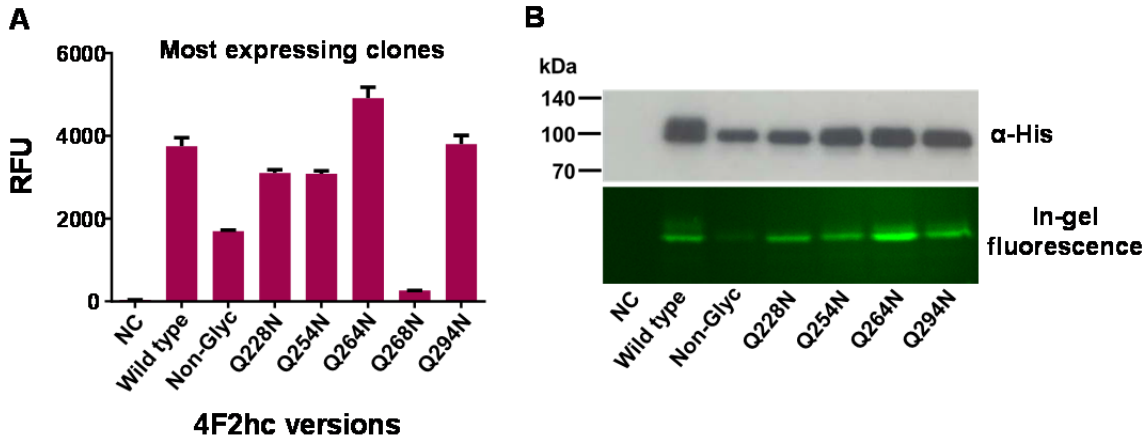

Figure 30. Analysis of the best expressing clones for less-glycosylated 4F2hc versions. Comparison of the whole cell fluorescence of the best clones expressing the different 4F2hc versions (**A**). Western blot detection, with α-His tag antibody, and in-gel fluorescence, of membranes from the clones showed in panel A (**B**). Non-Glyc: 4F2hc with the five predicted targets of glycosylation mutated from asparagine (N) to glutamine (Q). As a negative control, Pichia strain KM71H transformed with empty pPICZ B vector was used. NC: negative control. RFU: Relative Fluorescence Units.

pPICZ B vector carrying the cassettes for the co-expression of NG-4F2hc(Δ36Nter) and wild-type LAT1 (see Materials and Methods sections 5.3.1.2 and 5.3.2), was used to transform Pichia. Protein expression screening at small-scale was performed in 5 zeocin resistant transformants clones. The five clones expressed NG-4F2hc(Δ36Nter) and LAT1, and they formed heterodimer among both subunits, as checked by western blot using α-Strep-tag II and α-His tag antibodies (not shown). The clone with major levels of formed NG-4F2hc(Δ36Nter)/LAT1(wild type) was used for large-scale expression (see Materials and Methods section 5.3.4).

Isolated membranes from a 6L culture of Pichia expressing NG-4F2hc(Δ36Nter)/LAT1 (see Materials and Methods section 5.3.4), were solubilized in 2 % DDM/ 0.4 % CHS. The complex was double affinity purified, in 0.05 % DDM/0.01 % CHS, by IMAC followed by strep-tactin chromatography (see Materials and Methods section 5.3.5.2). Eluted protein was concentrated to 0.5 mg/ml, and the soluble fraction was subjected to size exclusion chromatography (SEC). Chromatogram showed a major peak of soluble protein, getting wider at the base toward the left, and thus, indicating also the presence of partial aggregation (Figure 31). A second minor peak is also observed. This second peak could be micelles of DDM/CHS plus some lipids coming from *Pichia pastoris* membranes, that absorb at 280 nm, since it is known that some lipids absorb at such wavelength (Cunningham. 1994). In addition, in previous studies where different HATs were purified from Pichia or Saccharomyces membranes, this peak was generally obtained (Costa i Torres 2012; Álvarez-Marimon. 2014). Western blot (α-His and α-Strep-tag II) of fractions from the most symmetric part of the peak corroborates the formation of the heterodimer (NG-4F2hc(Δ36Nter)/LAT1), although showed the presence of lesser amounts of aggregates and free subunits, compare to the amount of soluble heterodimer (Figure 31).

After the second affinity chromatography, the yield obtained for NG-4F2hc(Δ36Nter)/LAT1 was around 0.16 mg per liter of culture. This yield was twice higher than that obtained in Sf9 cells for the wild type 4F2hc-GFP/LAT1 (0.075 mg from 1L culture). However, in each step of the NG-4F2hc(Δ36Nter)/LAT1 purification process, and upon protein concentration and ultracentrifugation, important amounts of heterodimer were lost, suggesting a not optimal stability of this 4F2hc/LAT1 version. This low stability would be related with essential failures in the 4F2hc protein folding due to the lack of glycosylation and/or because of the removal of 36 residues at the N-terminus. This, at the same time, could difficult the proper interaction of 4F2hc with LAT1, reducing the final yield of heterodimer and decreasing the potential stabilization effect of 4F2hc on LAT1 and thus, of the whole heterodimer.

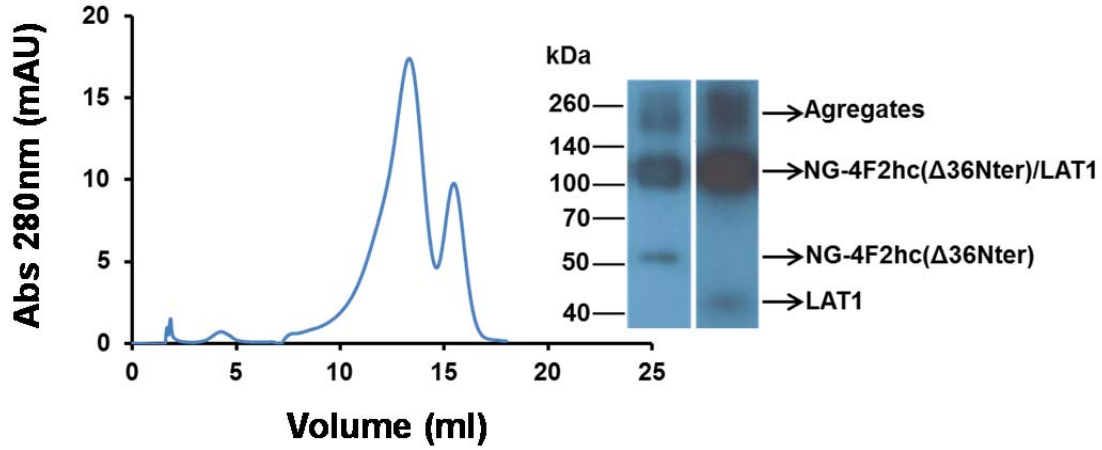

Figure 31. SEC profile of purified NG-4F2hc(Δ36Nter)/LAT1 from *Pichia pastoris*. Non-glycosylated 4F2hc with a deletion of 36 amino acids at the N-terminus, NG-4F2hc(Δ36Nter), was successfully co-expressed with LAT1 in *Pichia pastoris,* forming heterodimer. NG-4F2hc(Δ36Nter)/LAT1 was purified by a sequential double-affinity chromatography (IMAC followed by strep-tacting affinity chromatography) in 0.05% DDM/0.01% CHS. The eluted heterodimer was concentrated to 0.5 mg/ml and subjected to SEC using a Superose 6, 10/300 GL column. The most symmetric part of the protein peak was collected and subjected to α-His and α-Strep-tag II (from left to right) western blot, which is shown at the right side of the chromatogram.

The yield and stability of NG-4F2hc(Δ36Nter)/LAT1 expressed in Pichia, and in the actual DDM/CHS condition, is not enough to undertake crystallographic studies, since big amounts of protein are required to reach the concentrations needed for crystal growing. Additionally, the presence of some aggregation in the sample at a relatively low concentration of protein (0.5 mg/ml), in compare with the concentrations used for 3D crystallization, as well as the loss of protein during the purification process, suggest low stability of this 4F2hc/LAT1 version. This low stability could avoid crystal formation or hinder well-ordered crystallization, even at the desired concentrations of protein.

There is a high probability of having to modify our wild type 4F2hc/LAT1 heterodimer to obtain well-diffracting crystals, as this has been probed for great number of membrane proteins (Sobolevsky *et al.* 2009; Andersen *et al.* 2010; Hattori and Gouaux 2012; Penmatsa *et al.* 2013; Althoff *et al.* 2014). For this reason, we tried to do it by using the same approaches used in such works, but the protein show low stability. Thus, other similar modifications would have to be tested looking to produce enough amount of stable and monodisperse protein for crystallization trials. However, other no crystal based technique offer some advantages related to these issues, among other, by which we decided to focus our efforts toward solving the structure of

4F2hc/LAT1 by using cryo-EM. In this technique, protein modifications, like de-glycosylation or deletions at the N- or C-terminus, are not highly relevant to solve 3D structures of proteins, and the amount and concentration of purified protein used for structural determination, is generally several fold lower than those used in crystallography (Bai *et al.* 2015; Du *et al.* 2015). Thus, wild type version of 4F2hc/LAT1 could be undertaken to cryo-EM studies if it is correctly expressed and more stable than purified NG-4F2hc(Δ36Nter)/LAT1. Additionally, the recent advances in this area make this technique a good choice for solving the structure of 4F2hc/LAT1, as it will be explained in the next section.

3.4.2 Cryo-EM as a promising technique for solving the structure of vertebrate 4F2hc/LAT1

In the last years, the great development in the cryo-EM technology, like more powerful electron microscopes and direct electrons detectors (Binshtein and Ohi 2015), has made possible the structure determination, at near-atomic resolution, of integral membrane proteins of smaller molecular weight than it was possible in the past. This is the case of human γ-secretase (170 kDa) (Lu *et al.* 2014; Bai *et al.* 2015), the α1 glycine receptor (260 kDa) (Du *et al.* 2015) and the TRPV1 (300 kDa) (Liao *et al.* 2013) and TRPV2 (Zubcevic *et al.* 2016) channels. Thus, cryo-EM may be a very useful and powerful tool to solve the structure of the ~120 kDa vertebrate 4F2hc/LAT1.

Additionally, preparation of protein sample for vitrification in cryo-EM studies, can offer additional advantages when working with troubled metazoan membrane proteins, since lower amount and concentration of protein is required in comparison with 3D crystallography. In this way, low protein purification yields are less limiting. The concentration of protein required for cryo-EM imaging usually range between 0.3 to 1 mg/ml in 3 to 5 µl per cryo-EM grid. These numbers are dramatically lower respect to the 3-50 mg/ml usually required, in large sample volumes, to test several crystallization conditions, optimize the crystals quality, and reproduce well-diffracting crystals in order to obtain complete data sets at high resolution.

Despite the less restrictions in the protein amount required for cryo-EM, obtaining high resolution data from a vitrified protein sample is almost as challenging as it is by 3D crystallography. Finding the appropriate buffer and vitrification conditions for correct protein visualization at the electron microscope, and for obtaining particle images with enough contrast for structural determination is a laborious task. Furthermore, protein sample must accomplish the same criteria of quality than for 3D crystallography. It must be highly pure and monodisperse. This monodispersity is an indication of correct protein folding and stability. At this point, requirement of less protein concentration can be relevant for 4F2hc/LAT1, as avoiding having to extensively concentrate the protein can help in the maintenance of protein stability.

Thus, for studying 4F2hc/LAT1 expressed in Pichia as a possible candidate for cryo-EM, the same GFP-tagged heterodimer correctly expressed in Sf9 insect cells (4F2hc-GFP/LAT1) was used to screen rapidly the expression and stability of 4F2hc/LAT1 in different purification conditions, owing to the advantages of the GFP fusion for pre-structural studies of membrane proteins (Kawate and Gouaux 2006; Drew *et al.* 2008). Additionally, 4F2hc/LAT1 without GFP was also generated for testing the best conditions found with the GFP version, since non-GFP tagged 4F2hc/LAT1 will be the version used for cryo-EM studies. Although GFP removal could

be done by digestion with thrombin, because a target for this enzyme had been added between 4F2hc and GFP, incomplete proteolysis reaction would produce undesired sample heterogeneity.

3.4.3 Small-scale expression screenings of 4F2hc/LAT1 and 4F2hc-GFP/LAT1

Expression screening of 4F2hc-GFP/LAT1

Expression screening in Pichia cells transformed with the pPICZ B construct generated to co-express 4F2hc-GFP (GFP-His8 fused at the C-terminus), and LAT1 (Strep-tag II at the C-terminus) (see Materials and Methods section 5.3.1.1), was performed as described in Section 5.3.3 of Material and Methods. Only three clones, of the ten selected in YPD + Zeocin medium, were zeocin resistant. Results showed that 4F2hc-GFP/LAT1 heterodimer is formed in *Pichia pastoris* (Figure 32A and 32C), as also it is confirmed by the disruption of the disulfide bond that links both subunits upon DTT addition (Figure 32B and 32D). Western blot using α-His tag antibody also showed 4F2hc dimers (Figure 32C) that disappear upon DTT addition (Figure 32D). Clones 2 and 3 expressed more heterodimer than clone 1 (Figure 32A and 32C), although there are significant amount of free 4F2hc-GFP and LAT1 that is not forming heterodimer. Clone 2 (lane 2) was selected for large-scale expression because it seems to express the highest amount of 4F2hc-GFP/LAT1.

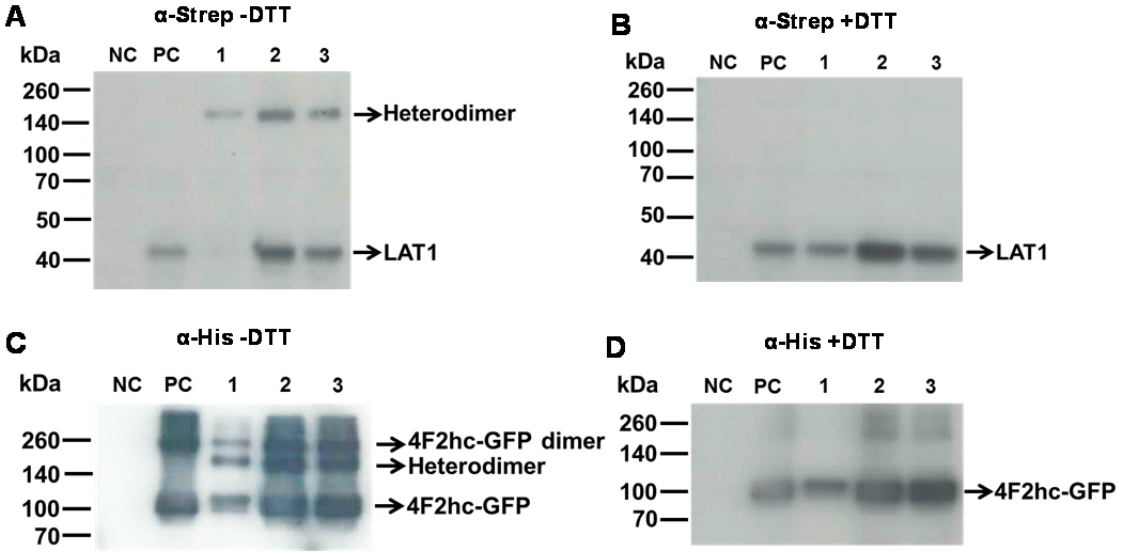

Figure 32. Expression screening of 4F2hc-GFP/LAT1 in *Pichia pastoris*. Western blot analysis of the expression levels of 4F2hc-GFP/LAT1 in isolated membranes from zeocin resistant *P. pastoris* clones, transformed with pPICZ B vector carrying the cDNAs of 4F2hc fused to GFP-His8 at the C-terminus (4F2hc-GFP), and LAT1 with a C-terminal Strep-tag II (LAT1). α-Strep-tag II (**A** and **B**) and α-His (**C** and **D**) antibodies were used, and samples untreated and treated with DTT are showed as indicated above of each panel. The anti-Strep-tag II and anti-His tag antibodies indicated the presence of heterodimers in absence of DTT (**A**, **C**). Under reducing conditions 4F2hc-GFP/LAT1 partially or totally disappeared increasing the LAT1 or 4F2hc-GFP monomer bands (**B**, **D**). The α-His tag antibody showed the presence of 4F2hc-GFP homodimer in the absence of DTT (**C**). Under reducing conditions practically only 4F2hc-GFP monomers were visible (**D**). Membranes from Pichia cells expressing LAT1 (**A**, **B**) or 4F2hc-GFP (**C**, **D**) were included as positive controls (PC). Membranes from *P. pastoris* transformed with empty pPICZ B vector were used as negative control (NC). 5μg of total protein were loaded in all cases.

Expression screening of 4F2hc/LAT1

Expression screening in Pichia cells transformed with the pPICZ B construct designed to co-express 4F2hc (His8 at the N-terminus), and LAT1 (Strep-tag II at the C-terminus) (see Materials and Methods section 5.3.1.2), was performed as described in section 5.3.3 of Material and Methods. Six out of eight zeocin resistant clones expressed 4F2hc/LAT1 heterodimer, as better observed using the antibody against Strep-tag II (Figure 33A). In the western blot performed using the antibody against His tag, a band of bigger molecular weight than the heterodimer is observed (Figure 33B). This band corresponds to the reported dimerization of 4F2hc (Fort *et al.* 2007), as it can be evidenced in the clone 5, in which high expression of 4F2hc (Figure 33B, lane 5) but not of LAT1 (Figure 33A, lane 5), led to the formation of great amount of 4F2hc dimers but not of 4F2hc/LAT1 heterodimers. In an attempt to obtain a clone producing the major amount as possible of 4F2hc and LAT1 forming heterodimer, clone 1 was selected for large-scale expression, because it showed almost no free 4F2hc or LAT1 subunits (Figure 33A and 33B, lane 1).

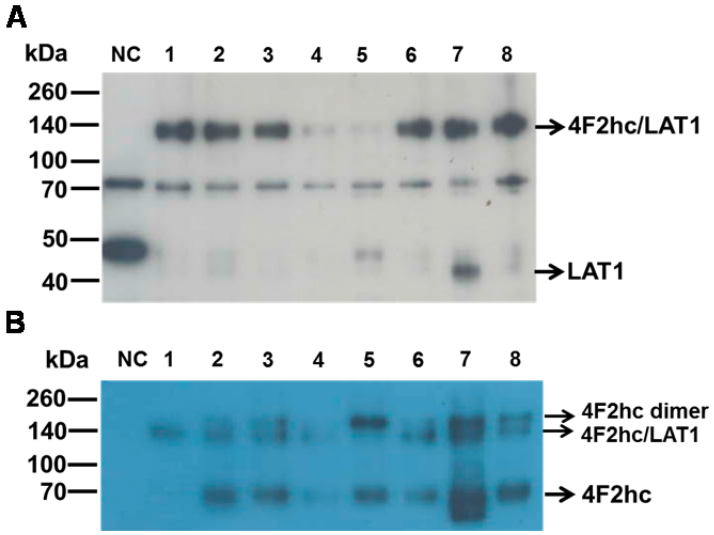

Figure 33. Expression screening of 4F2hc/LAT1 in *Pichia pastoris*. Western blot analysis of the expression levels of 4F2hc/LAT1, in isolated membranes from eight zeocin resistant clones, transformed with pPICZ B vector carrying the cDNAs of 4F2hc fused to His8 at the C-terminus, and LAT1 with a C-terminal Strep-tag II. α-Strep-tag II (**A**) and α-His tag (**B**) antibodies were used. Membranes of *P. pastoris* transformed with empty pPICZ B vector were used as a negative control. NC: negative control. 5 μg of total protein were loaded in all cases.

3.4.4 Analysis of the effect of lipids on the stability of 4F2hc-GFP/LAT1

Detergent solubilized membrane proteins, even in mild detergents, might not be stable enough and aggregate, at different extents, during purification or concentration steps. This could be due to the removal, during solubilization by detergents, of membrane lipids that are relevant for the stability and activity of these proteins (van Meer *et al.* 2008; Phillips *et al.* 2009). Thus, some hydrophobic regions of a detergent solubilized membrane protein could be deprived of relevant interactions with membrane lipids, being exposed to the water solution, and then establishing unspecific interactions with hydrophobic regions of other molecules, in this way acting as aggregation sources. The addition of lipids in the solubilization step and during protein purification, might fulfill for these lipid-protein interactions, reducing aggregation. Lipid

requirement has been shown in many eukaryotic membrane proteins, for example, the voltage dependent K+ channel (Long *et al.* 2007).

In order to find the best lipidic condition for the stabilization of 4F2hc/LAT1 expressed in Pichia, addition to the basal DDM/CHS condition of several lipids, individually or combined, was used during the solubilization and purification of 4F2hc-GFP/LAT1. These lipidic conditions have been used broadly in structural biology of membrane proteins. Two approach were followed to test these lipidic conditions. First, Ultracentrifugation Dispersity Sedimentation assay (UDS) (Gutmann *et al.* 2007) was used to select some of the best behaving lipidic conditions, and then, FSEC of 4F2hc-GFP/LAT1 in those conditions was carried out to investigate the stability and monodispersity of this heterodimer in such conditions.

As has been described through this work, tagging membrane proteins with GFP, allows time-saving screening of stability conditions. Moreover, the high sensitivity of GFP fluorescence emission is useful for testing several conditions at a time using very low protein amounts. For these reasons, 4F2hc-GFP/LAT1 version was used to develop this objective.

3.4.4.1 Study of the lipid addition on the 4F2hc-GFP/LAT1 stability by UDS

Isolated membranes from a 6 L culture expressing 4F2hc-GFP/LAT1 (see Materials and Methods section 5.3.4), were divided in 7 equal samples and each one was solubilized with a different lipidic condition. One of the samples was solubilized with 2 % DDM/0.4 % CHS, which is the basal condition (BC), and the other 6 were solubilized in the basal condition plus one of the following lipids: 5 mM DPPC, 5 mM DMPC, 2 mM POPC, 1mM POPG, 2 mM POPE or 2 mM POPC/ 0.7 mM POPG/ 0.7 mM POPE. Soluble fractions were subjected to strep-tactin affinity chromatography (see Materials and Metdhods section 5.3.5.1), and protein was eluted in buffer with the same lipidic composition but less concentrated (0.05 % DDM/ 0.01 % CHS (BC), BC + 0.25 mM DPPC, BC + 0.25 mM DMPC, BC + 0.1 mM POPC, BC + 0.05 mM POPG, BC + 0.1 mM POPE and BC + 0.1 mM POPC/ 0.035 mM POPG/ 0.035 mM POPE. The obtained heterodimers were subjected to UDS. To perform this analysis, purified 4F2hc-GFP/LAT1 in each of the lipidic conditions, was ultracentrifuged, and part of the soluble fraction was stored for western blot analysis (Figure 34, Day 1). The remaining part of the samples were left for 16 h at room temperature, and at the next day, soluble fractions obtained after ultracentrifugation were collected for western blot analysis (Figure 34, Day 2).

By UDS, major stability in solution of membrane proteins is evidenced by a major proportion of protein in the soluble fraction after ultracentrifugation. Proteins are more susceptible to degradation and denaturation as much time elapses after protein extraction from the membrane, and this process is speed up as temperature is higher. Determination of the remaining soluble part of 4F2hc-GFP/LAT1 in each of the lipidic conditions, was performed after 16h at room temperature, in order to subject the protein to a non-appropriate condition for its stability, and in this way, being able to see more evidently a possible effect of lipids in the protein quality.

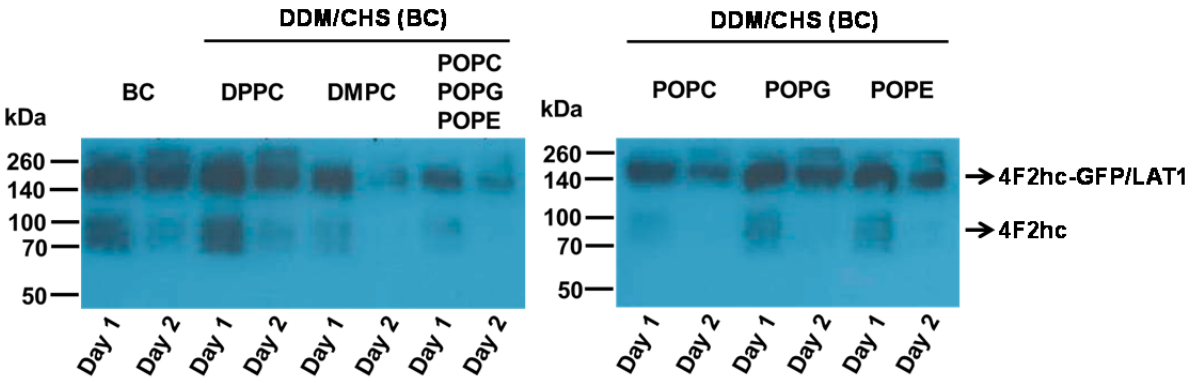

Figure 34. Analysis by UDS of the lipidic conditions used to stabilize 4F2hc-GFP/LAT1. 4F2hc-GFP/LAT1 was solubilized from isolated Pichia membranes in 7 different conditions: **1**) 2 % DDM/0.4 % CHS (basal condition (BC)), **2**) BC + 5 mM DPPC, **3**) BC + 5 mM DMPC, **4**) BC + 2 mM POPC/ 0.7 mM POPG/ 0.7 mM POPE, **5**) BC + 2 mM POPC, **6**) BC + 1 mM POPG and **7**) BC + 2 mM POPE. The heterodimer was purified by strep-tacting affinity chromatography and eluted in **1**) 0.05 % DDM/ 0.01 % CHS (basal condition (BC)), **2**) BC + 0.25 mM DPPC, **3**) BC + 0.25 mM DMPC, **4**) BC + 0.1 mM POPC/ 0.035 mM POPG/0.035 mM POPE, **5**) BC + 0.05 mM POPG, **6**) BC + 0.1 mM POPC and **7**) BC + 0.1 mM POPE. Part of the eluted heterodimer (Day 1) was kept for 16 h at room temperature, ultracentrifuged, and the soluble fraction recovered (Day 2). The remaining soluble 4F2hc-GFP/LAT1 in both times (Day 1 and Day 2), in the different lipidic conditions tested, was detected by western blot using the antibody against the His-tag placed in the GFP fused to 4F2hc. 10 µl of all samples were loaded. DPPC: 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine; DMPC: 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine; POPC: 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine; POPG: 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol); POPE: 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine.

Purified heterodimer in DDM/CHS and in DDM/CHS/DPPC behaves very similarly, with the major amounts of solubilized protein obtained right after purification (Figure 34, Day 1). This difference in the amount of purified heterodimer was particularly marked in comparison with DDM/CHS/DMPC and DDM/CHS/POPC/POPG/POPE conditions. This result indicates that these lipids are reducing the protein purification yield. It is possible that physicochemical properties of the DDM/CHS/lipid micelles for DMPC, and POPC/POPG/POPE, were not as efficient as those of DDM/CHS and DDM/CHS/DPPC in the solubilization of 4F2hc-GFP/LAT1, or that they could be interfering in some way with the binding of 4F2hc-GFP/LAT1 to the strep-tactin resin. The latter explanation could make sense if it is considered that Strep-tag II is present at the C-terminus of LAT1, the subunit in the heterodimer to be significantly more embedded into the DDM/CHS/lipid micelle, due to its twelve predicted transmembrane domains. Curiously, despite all the lipidic conditions contain the same concentration of DDM/CHS, it seems that less protein remains in solution 16 h after purification (Day 2) in the DMPC, POPC, POPG, POPE and POPC/POPG/POPE conditions (Figure 34), as if heterodimer was less stable in these lipids.

Regarding the proportion of protein that remains soluble after 16h at room temperature (Day 2), in compare to day 1, none of the lipids showed an improvement in protein stabilization in compare to DDM/CHS basal condition (Figure 34). DMPC, POPC/POPG/POPE and POPC seemed to be the worst conditions, as less protein is maintained soluble at day 2 in these lipidic conditions. DPPC condition also behaves very similar to DDM/CHS, regarding to the maintenance in solution of 4F2hc-GFP/LAT1 after 16h at room temperature. Both conditions maintained more protein soluble at day 2 than the other lipids. POPG and POPE were also effective in maintaining protein in solution, as observed when comparing day 1 with day 2, although slightly less than DDM/CHS and DPPC. The lipidic conditions in which more protein

remained soluble (DPPC, POPG and POPE), as identified by UDS, were able to avoid large aggregation of 4F2hc-GFP/LAT1. However, the real potential of these conditions to increase 4F2hc-GFP/LAT1 stability, in comparison with DDM/CHS, would be determined by investigating the monodispersity of the complex. For doing this, FSEC analysis of 4F2hc-GFP/LAT1 in the best lipidic conditions was also done.

3.4.4.2 Study of the lipid addition on the 4F2hc-GFP/LAT1 stability by FSEC

Based on the results obtained by UDS (Figure 34), 4F2hc-GFP/LAT1 was solubilized and purified in the best three lipidic conditions together with the basal condition (DDM/CHS). Isolated membranes from a 6 L culture, were divided in 4 equal samples and each one was solubilized with 2 % DDM + 0.4 % CHS, basal condition (BC), and the other 3 were solubilized in the basal condition plus one of these lipids: 5 mM DPPC, 1 mM POPG or 2 mM POPE. Soluble fractions were subjected to strep-tactin affinity chromatography, and protein was eluted in buffer with the same lipidic composition but less concentrated (0.05 % DDM/ 0.01 % CHS (basal condition; BC), BC + 0.25 mM DPPC, BC + 0.05 mM POPG or BC + 0.1 mM POPE).

To obtain a more pure 4F2hc-GFP/LAT1, strep-tactin affinity purification was followed by ion metal affinity chromatography (IMAC) (see Materials and Methods section 5.3.5.2). Testing protein stability after two-round affinity purification is relevant, because this protocol will be finally used to purify heterodimer for structural studies. In this way, FSECs of doubled affinity purified 4F2hc-GFP/LAT1 could bring a more realistic view about the stability of the heterodimer in each lipidic condition. Purified 4F2hc-GFP/LAT1 in all lipidic conditions was concentrated by 6-fold, and GFP fluorescence was checked as a way to rapidly quantify 4F2hc-GFP/LAT1 in each sample. For leaving all samples at the same concentration, samples with the highest fluorescence were diluted to reach comparable Relative Fluorescence Units (RFU) values. In this way, all samples were at the same conditions during the experiment except for the differential lipid addition. Purified 4F2hc-GFP/LAT1 was left for 16 h at room temperature, ultracentrifuged and subjected to FSEC (Figure 35).

It can be observed a little distortion, in all conditions, in the FSEC traces shown in Figure 35. This was due to a failure in the AKTA protein collector, which sometimes misplaced some drops in the black 96-well plates used for measuring fluorescence. However, the quality of FSECs is good enough to make some conclusions. In first place, 4F2hc-GFP/LAT1 produced in *Pichia pastoris,* and purified by two consecutive affinity chromatographies is stable in DDM/CHS, as well as it is in DDM/CHS/lipids (DPPC, POPE and POPG) (Figure 35A-D). In all the conditions, aggregated protein is not observed, suggesting that 4F2hc-GFP/LAT1 is sufficiently stable for undergoing double affinity purification. The heterodimer peaks would have been quite symmetrical, if protein collector had been working properly, as it can be inferred from the FSEC profiles (Figure 35).

On the other hand, none of the lipids seems to increase the stability of 4F2hc-GFP/LAT1, or to increase monodispersity in the sample, as FSECs chromatograms among all the conditions tested are very similar in shape and width of the protein peak (Figure 35E and 35F). Because the buffer of the protein sample for cryo-EM or crystallography must be kept as simple as possible, with only the required additives for protein stabilization, and in fact, the simplest condition (DDM/CHS), is sufficient for 4F2hc-GFP/LAT1stabilization, as suggested by UDS and FSEC experiments, this condition continued to be used for the next experiments.

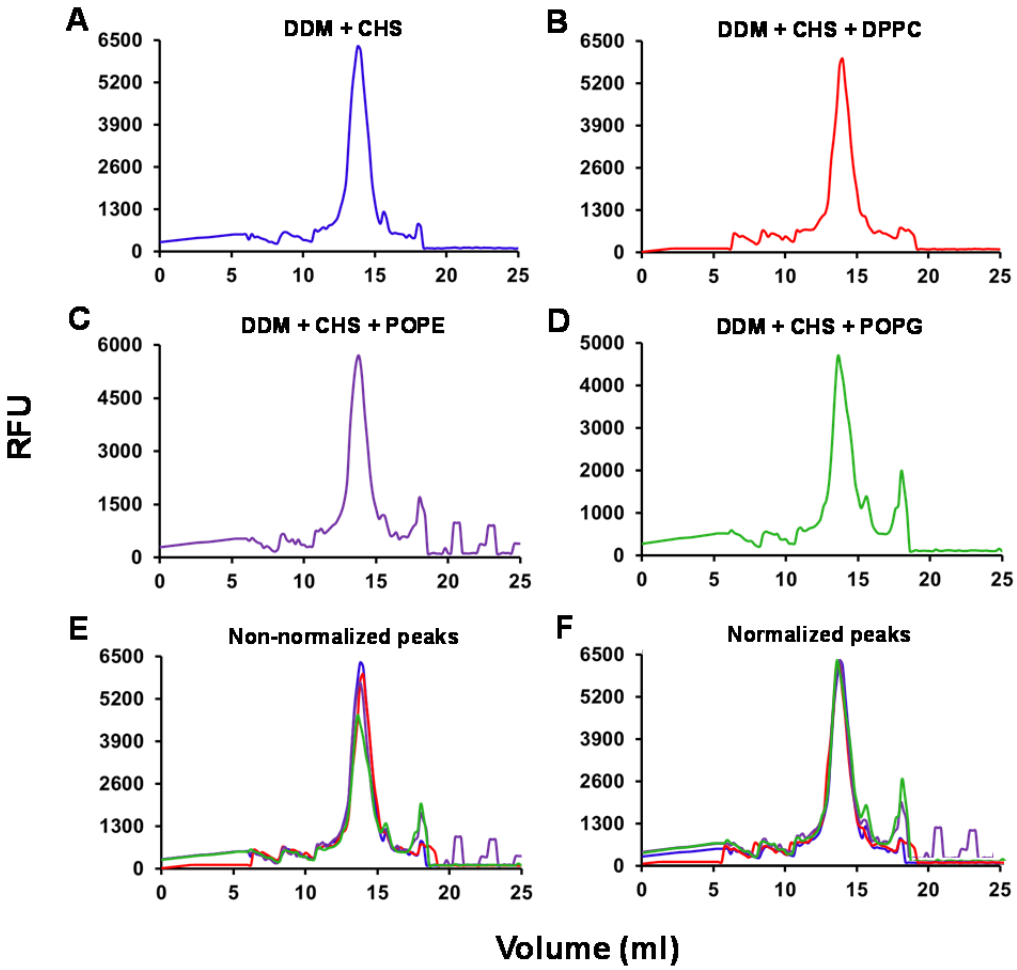

Figure 35. FSEC profiles of purified 4F2hc-GFP/LAT1 in different lipidic conditions. 4F2hc-GFP/LAT1 was solubilized from isolated Pichia membranes in 4 different conditions: **1)** 2 % DDM / 0.4 % CHS (basal condition (BC)), **2)** BC + 5 mM DPPC, **3)** BC + 1 mM POPG and **4)** BC + 2 mM POPE. The heterodimer was purified by a sequential double-affinity chromatography, first by strep-tactin affinity chromatography, and finally, by IMAC. Elution of 4F2hc-GFP/LAT1 from IMAC was performed in: **A)** 0.05 % DDM / 0.01 % CHS (basal condition (BC), **B)** BC + 0.25 mM DPPC, **C)** BC + 0.05 mM POPG and **D)** BC + 0.1 mM POPE, respectively. Purified heterodimer was left for 16 h at room temperature, and 100 µl of each sample analyzed by FSEC in a Superose 6, 10/300 GL column. Non-normalized (**E**) and normalized (**F**) peaks from A to D are shown merged in the same color code for comparison. DPPC: 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine; POPE: 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine. POPG: 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol). RFU: Relative Fluorescence Units.

3.4.5 Stability in DDM/CHS of vertebrate 4F2hc/LAT1 expressed in Pichia

As no improvement in the 4F2hc-GFP/LAT1 stability was observed in any of the tested lipids, and in all of these conditions including the basal (DDM/CHS), heterodimer seemed to be stable and monodisperse, the behavior of purified non-GFP tagged 4F2hc/LAT1 was explored using DDM/CHS. Isolated membranes from a 6 L culture expressing 4F2hc/LAT1 (see Materials and Methods section 5.3.4), were solubilized in 2 % DDM / 0.4 % CHS. The complex was doubled affinity purified in 0.05 % DDM / 0.01 % CHS by IMAC, followed of strep-tactin chromatography (see Materials and Methods section 5.3.5.2). Then, the 4F2hc/LAT1 thus obtained was quantified. The yield of the heterodimer purification was generally from 1 mg to 3

mg from 6 L of culture (0.16-0.5 mg from 1 L culture). This means that the yield of wild type heterodimer produced in *Pichia pastoris* was 2-6-fold higher than the yield obtained in Sf9 cells (0.075mg from 1 L culture). Taking into account that the heterodimer expressed in insect cells was the GFP tagged version (4F2hc-GFP/LAT1), having to be removed the GFP yet, we concluded that the change of expression system was effective to increase the production of heterodimer.

To test the quality of purified 4F2hc/LAT1 upon concentration and upon time, this heterodimer was concentrated to 1 mg/ml, and the soluble fraction obtained after ultracentrifugation was analyzed by SEC (Figure 36A). The remaining part of the sample was left for 16 h at 4°C, and at the next day, it was concentrated at 2 mg/ml and the soluble fraction obtained after ultracentrifugation subjected to SEC (Figure 36A). The protein quality is mainly good, since a single peak of soluble protein elutes from the column and big aggregates are not observed in the void volume (Figure 36A). Additionally, the peak of 4F2hc/LAT1 after one day and twice as concentrated as at day 1, showed the same behavior in SEC than at 1 mg/ml the day before (Figure 36A). This indicates that 4F2hc/LAT1 is quite stable in solution when purified in DDM/CHS. Western blot analysis, by α-His-tag and α-Strep-tag II antibodies, together with coomassie staining, of the top fraction of the 2 mg/ml peak, showed that 4F2hc/LAT1 is highly pure in the final purification stage, and that heterodimer is not even partially disrupted into its heavy and light subunits (Figure 36B), supporting stability of the non-GFP tagged heterodimer produced in Pichia.

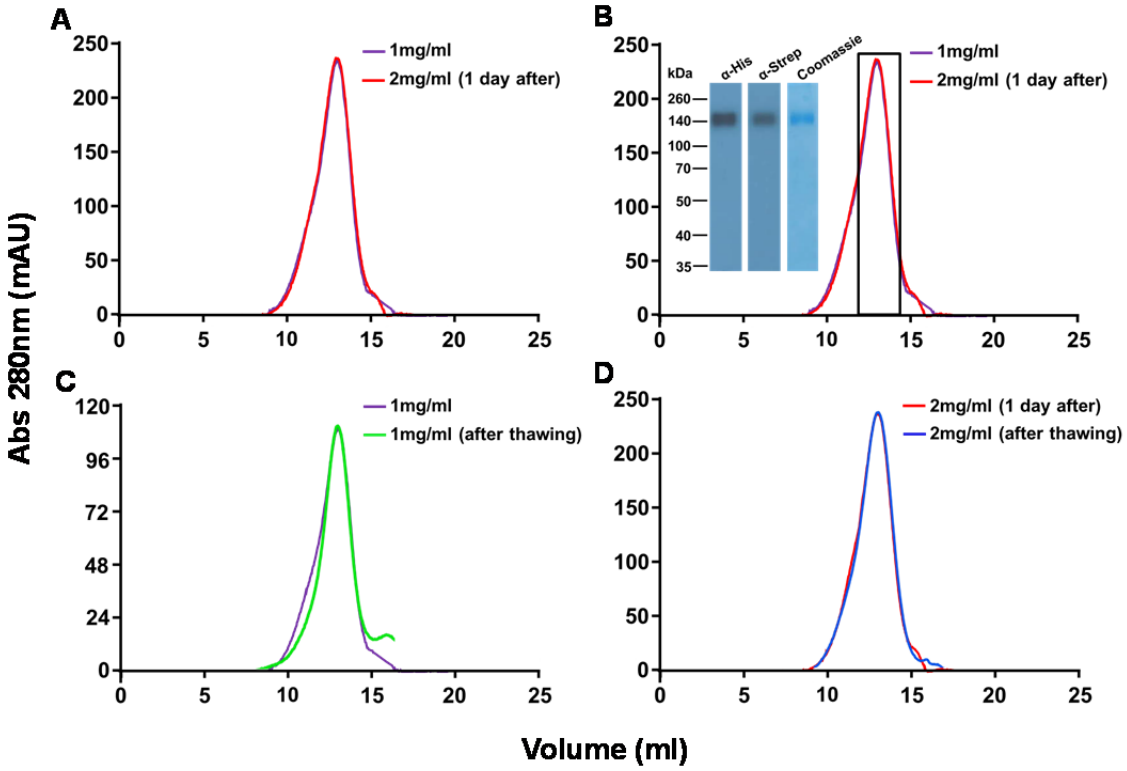

Figure 36. SEC profiles of 4F2hc/LAT1 purified from *Pichia pastoris*. The heterodimer was purified by a sequential double-affinity chromatography, first by Immobilized Metal Ion Affinity Chromatography (IMAC), and finally, by strep-tactin affinity chromatography. The purification process was performed in presence of 0.05 % DDM/ 0.01 % CHS. A part of the eluted heterodimer from strep-tactin resin was concentrated to 1 mg/ml and subjected to SEC (**A**), and the remaining sample left for 16 h at 4°C, concentrated to 2 mg/ml and analyzed by SEC

(**A**). Normalized SEC profiles of both samples indicate that the protein remains as stable as the first day even after being twice concentrated (**A**). Fractions from both SECs, enclosed in a black rectangle (**B**), were pooled and frozen at -80°C. After thawing, these samples were concentrated until 1 mg/ml and 2 mg/ml and analyzed by SEC. Normalized SEC profiles comparing unfrozen and thawed samples at 1 mg/ml and 2 mg/ml are shown in Panels **C** and **D**, respectively. The integrity and purity of 4F2hc/LAT1 were analyzed by western blot, using antibodies against His tag and Strep-tag II, and by coomassie brilliant blue staining (**B**).

Being able to reach concentrations of 1 and 2 mg/ml is relevant, since concentrations used for cryo-EM studies are usually bellow these numbers. For instance, 0.3 mg/ml of TRPVI channel (Liao *et al.* 2013) and 0.6 mg/ml of γ-secretase (Bai *et al.* 2015), were used in the structural determination of these proteins by cryo-EM. However, despite the peak of protein is mainly good (Figure 36), as it is not too broad and not big aggregation species are observed, it is not completely symmetrical, showing a slight widening towards the left at the base. This indicates some degree of 4F2hc/LAT1 polydispersity in the sample, which could be due to heterodimer tendency to moderately aggregate.

This polydispersity is neither appropriated for crystallography nor cryo-EM studies, as it has been discussed before, by which actions for obtaining a monodisperse or less heterogeneous sample must be addressed. Thus, the fractions containing the bigger species of protein from the 1 mg/ml and 2 mg/ml peaks (outside of the black rectangle in Figure 36B) were discarded, and protein coming from the more symmetrical part of the peaks (inside the black rectangle in Figure 36B), was frozen in liquid nitrogen for their posterior analysis by SEC. This would allow to obtain a more monodisperse sample if the protein remains stable after a freeze-thaw cycle.

After thawing, protein coming from each peak was concentrated to its original concentration (1 mg/ml and 2 mg/ml), and soluble fraction subjected to SEC. 4F2hc/LAT1 remained stable in solution after freezing, as observed by the presence of a single major peak of soluble protein (Figure 36C and 36D). The SEC profiles were very similar to those obtained before freezing. Additionally, in the case of the 1 mg/ml peak, it can be observed that the protein in the fractions collected is quite more monodisperse, as the widening of the original peak has been significantly reduced (Figure 36C). This is relevant because it is probable that 4F2hc/LAT1 in the cryo-EM sample is required at 1 mg/ml or less concentrated. In the case of the 2 mg/ml peak, it seems that the polydispersity of protein is equal to the original sample, however, this could be also because this protein has been subjected to more rigorous conditions than the 1 mg/ml sample. In first place, 2 mg/ml protein stays one day more at 4°C before SEC, and second, it was concentrated twice more in two occasions than the 1 mg/ml sample.

3.4.5.1 Vertebrate 4F2hc/LAT1 is more stable than human 4F2hc/LAT2

As already mentioned, first structural studies that were conducted in our lab with a metazoan HAT, led to determine a low resolution model of human 4F2hc/LAT2 (21Å) from negatively stained complexes (Rosell *et al.* 2014). Other studies for solving its atomic structure could not be possible due to its low stability, probably due to the low stability of human LAT2 (see Figure 12 in Introduction section 1.7.3) (Rosell *et al.* 2014). Therefore, a screening to find a more stable light subunit was performed, in which vertebrate LAT1 and LAT2 were two of the best candidates (Álvarez-Marimon. 2014). In the present work, vertebrate LAT1, has been found to form the most stable heterodimer (4F2hc/LAT1), in comparison with vertebrate 4F2hc/LAT2 (Figure 25). Thus, to confirm that all these studies have led to find a more stable heterodimer (vertebrate 4F2hc/LAT1) than human 4F2hc/LAT2, the human heterodimer was expressed in

Pichia pastoris and purified under the same conditions than 4F2hc/LAT1, to compare both stabilities. The pPICZ B construct generated by Dr. Rosell to co-express 4F2hc-His10 and LAT2-Strep-tag II (Rosell et al., 2014) was used for this purpose.

Isolated membranes from a 6 L culture expressing human 4F2hc/LAT2 (see Materials and Methods section 5.3.4), were solubilized in 2 % DDM / 0.4 % CHS. The soluble fraction was purified by IMAC followed by strep-tactin affinity chromatography in 0.05 % DDM/0.01 % CHS (see Materials and Methods section 5.3.5.2). Then, protein was concentrated to 1.5 mg/ml and subjected to SEC. The chromatogram of human 4F2hc/LAT2 showed a peak quite wide and asymmetrical (Figure 37A), indicating great polydispersity, much more than that of vertebrate 4F2hc/LAT1 (Figure 37B). Additionally, it can be observed by coomassie staining of the top fraction of the peak, that there is significant amount of free 4F2hc and LAT2, suggesting disruption of human 4F2hc/LAT2 heterodimer (Figure 37A). This significant heterodimer dissociation into its heavy and light subunits does not occur in vertebrate 4F2hc/LAT1 (Figure 36B). These results are a clear evidence that, as expected, vertebrate 4F2hc/LAT1 is much more stable in solution than human 4F2hc/LAT2, suggesting the possibility of going further in structural studies with the new metazoan heterodimer.

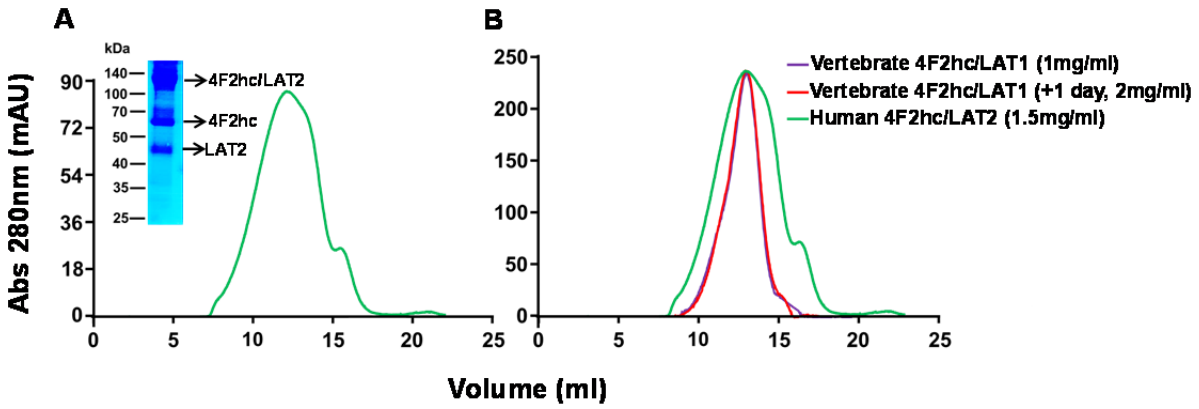

Figure 37. Comparison of the stability of vertebrate 4F2hc/LAT1 and human 4F2hc/LAT2. Human 4F2hc/LAT2 was purified by a sequential double-affinity chromatography, following the same strategy, and in the same DDM/CHS concentrations than those in which vertebrate 4F2hc/LAT1 was purified. The eluted heterodimer from strep-tactin resin was concentrated to 1.5 mg/ml and subjected to SEC using a Superose 6, 10/300 GL column (**A**). The integrity and purity of purified human 4F2hc/LAT2 were analyzed by coomassie brilliant blue staining, in which besides heterodimer, a lot of 4F2hc and LAT2 monomers can be observed (**A**). Normalized SEC profiles of purified human 4F2hc/LAT2 and vertebrate 4F2hc/LAT1 indicate that, in comparable protein concentrations, 4F2hc/LAT1 is more stable (**B**).

3.4.6 Functional validation of vertebrate 4F2hc/LAT1 expressed in *Pichia pastoris*

Certainly, membrane protein activity is directly related with its proper folding in the native plasma membrane. When a membrane protein is heterologously expressed, the way in which it is translated, translocated and assembled, together with the membrane lipid composition, determine the protein folding and function. Additionally, to the harmful effects that expression in a foreign organism could cause, protein extraction and purification from the membrane could cause denaturation and/or degradation, at any extent, with the consequent protein inactivation, which hinders or impedes structural studies. In this sense, it is relevant to determine the activity

of heterologously expressed and purified membrane proteins, as it is an indication of a well-structured and matured protein.

Ideally, reconstitution assays of a purified membrane protein in liposomes would lead to the determination of the protein function in a membrane-like environment. However, these studies are usually complicated, requiring to test several lipids to obtain the proper lipidic composition for the studied protein. For this reason, alternative strategies in which protein reconstitution is not needed have been designed. A good approximation to study the functionality of a protein is by determining the binding of the corresponding substrates. A binding signal indicates native folding of the binding site, which can be reflecting correct folding of the whole purified protein, suggesting its suitability for structural studies.

One of these binding assays is the Scintillation Proximity Assay (SPA) (Udenfriend *et al.* 1985), in which scintillation beads are functionalized to bind protein tags. Thus, tagged protein is incubated with the corresponding beads and radioactive ligand. In this way, when the radio-ligand emitting β-rays binds to the protein, that is in turn bound to the scintillation bead, this β-ray emission reaches the bead by proximity, causing a light signal that can be determined in a scintillation counter (Figure 38).

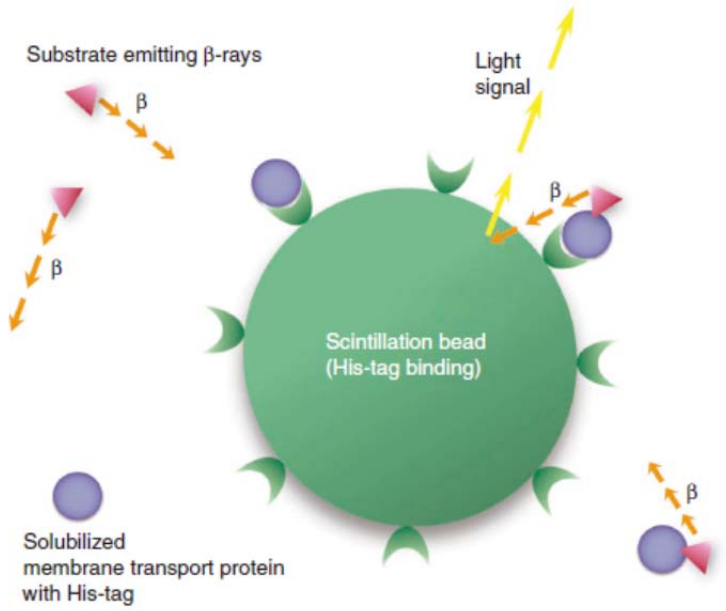

Figure 38. Representation of substrate binding determination by SPA. Scintillation beads immobilize acceptor proteins on their surfaces by tags (e.g., His tag). This beads are stimulated to emit light when the radioligand interacts with the bound protein. This luminescence signal is measured in a scintillation counter. (Harder and Fotiadis 2012).

Therefore, due to the practicability of the SPA set up in a 96 well plate and its high throughput; as nor washing nor filtration steps are required to separate bound of unbound ligand, we decided to test the binding properties of 4F2hc/LAT1 by using this technique. Additionally, SPA has been also used to determine the binding properties of amino acid transporters, like the prokaryotic homologue of LATs AdiC (Harder and Fotiadis 2012) and human 4F2hc/LAT2 (Meury *et al.* 2014), as well as of other type of transporters (Quick and Javitch 2007), receptors, and other membrane proteins since 30 years ago (Nelson 1987). The SPA assays were performed in our laboratory under the advice of Dr. Paola Bartoccioni.

3.4.6.1 First approximation to the binding properties of purified vertebrate 4F2hc/LAT1

SPA is also broadly used to determine the inhibition constants of non-radiolabeled substrates or protein inhibitors. Often, ligands and inhibitors are useful for structural studies to maintain the conformational homogeneity in a protein sample, since protein heterogeneity hinders structure determination by 3D crystallography and cryo-EM, as discussed through this work. This conformational heterogeneity, sometimes, can be due to an intrinsic flexibility of the protein related with its function. Thus, different conformational stages of a protein, besides the most energetically favorable, can be present in a sample. For this reason, finding a molecule able to block in a single conformation the protein could be very useful.

In a previous study, KYT-0353 was found to inhibit [^{14}C]-leucine uptake and cellular growth in human colon cancer-derived HT-29 cells, and in cells from mouse renal proximal tubule expressing human 4F2hc/LAT1, and also was found to inhibit tumor growth in mice transplanted with HT-29 cells (Oda *et al.* 2010). The IC_{50} found to inhibit [^{14}C]-leucine uptake in HT-29 cells via 4F2hc/LAT1 inhibition was of 0.06 μM, suggesting high affinity of KYT-0353 for human LAT1. Like human LAT1 is 78 % identical to vertebrate LAT1, we wanted to investigate the potential use of KYT-0353 as vertebrate 4F2hc/LAT1 inhibitor for structural studies.

For doing this, isolated membranes from a 6 L culture expressing vertebrate 4F2hc/LAT1 (see Materials and Methods section 5.3.4), were solubilized in 2 % DDM/0.4 % CHS. The soluble fraction was purified by IMAC followed by strep-tactin affinity chromatography in 0.05 % DDM/0.01 % CHS (see Materials and Methods section 5.3.5.2). Then, protein was concentrated to 1 mg/ml, and soluble fraction subjected to SPA experiments (see Materials and Methods section 5.3.6), and SEC for analyzing the quality of the protein in the SPA sample (Figure 39). A major peak of highly pure and stable soluble protein was obtained (Figure 39A), indicating a good quality of 4F2hc/LAT1 sample to carry out the binding assay.

SPA was performed by mixing during different times, yttrium silicate (YSi) streptavidin coated beads, 10 μg of purified 4F2hc/LAT1 and 0.5 μCi [^{3}H]-isoleucine as a substrate (see Materials and Methods section 5.3.6). The same mixture plus 6 μM KYT-0353 (human LAT1 inhibitor) was used to determine preliminarily the potential binding of this inhibitor to vertebrate LAT1 in the heterodimer. To determine the specific binding signal of the L-[^{3}H]-isoleucine in the experiment, the same initial mixture, without inhibitor, and plus 5 mM L-leucine or 10 mM D-desthiobiotin was used (Figure 39B). In this experiments 4F2hc/LAT1 was bound to the YSi streptaviding coated beads by the C-terminal Strep-tag II of LAT1.

The addition of the non-radiolabeled substrate (5 mM Leu) to the SPA mixture, induced a modest specific binding signal, i.e., the difference among the CPM of 0.5 μCi [^{3}H]-Ile and the 5 mM cold Leu in Figure 39B. However, this 5 mM Leu binding signal would correspond, in this experiment, with the maximum reduction of the SPA signal caused by the binding of a cold substrate to 4F2hc/LAT1, as also determined by the addition of 10 mM D-desthiobiotin, which would cause the detachment of 4F2hc/LAT1 from the streptavidin coated SPA beads (Figure 39B). Moreover, the addition of 6 μM human LAT1 inhibitor KYT-0353 showed binding comparable to 5 mM L-leucine, and thus similar SPA signal to the latter condition, and to the 10 mM D-desthiobiotin control. These results suggest that purified vertebrate 4F2hc/LAT1 could be functional, and also that KYT-0353 may work as a vertebrate LAT1 inhibitor, and consequently, could be used in structural studies of 4F2hc/LAT1.

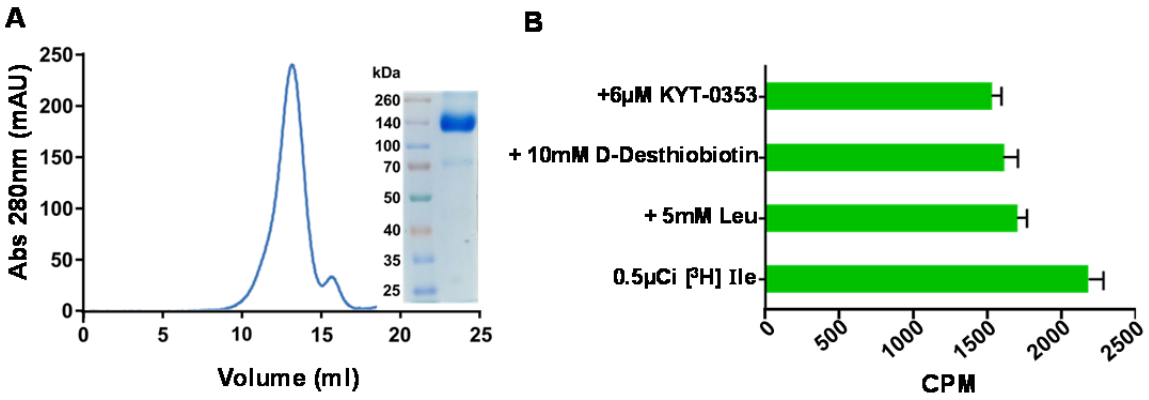

Figure 39. First approximation to the binding properties of 4F2hc/LAT1 by SPA. Heterodimer was purified by a sequential double-affinity chromatography in presence of 0.05 % DDM / 0.01 % CHS, and concentrated to 1 mg/ml. The quality and purity of the sample was analyzed by SEC using a Superose 6, 10/300 GL column, and coomassie brilliant blue staining, respectively (**A**). SEC profile and coomassie staining indicated that purified and concentrated 4F2hc/LAT1 used for SPA experiments was not aggregated and was highly pure. SPA was performed incubating for 1, 3 or 16 h, streptavidin-YSi beads (220 µg per well), 10 µg of purified 4F2hc/LAT1 and 0.5 µCi L-[^{3}H]-radiolabeled isoleucine, as a substrate (**B**). A saturating concentration of non-radiolabeled L-leucine (+ 5 mM Leu) was also incubated with the mix described before to determine the maximum binding of could substrate, and thus also the maximum reduction of the SPA signal (CPM). Determination of the binding of human LAT1 inhibitor, KYT-0353, was performed by using the same conditions but in absence of could substrate. D-desthiobiotin at 10 mM final concentration was also incubated with the same mix (beads/protein/L-[^{3}H]-Ile) as a control. Data points indicate the means ± s.e.m. from triplicate determinations. Results shown in panel B corresponds to 3 h of incubation, but are representatives for the other two times. CPM: counts per minute.

The high background signal observed in this experiment has been problematic to detect a significant binding signal. This noise signal could be caused by non-specific binding of the radio-ligand to the beads, but not to the protein, or by non-proximity effects, in which the concentrations of the radiolabeled ligand or the SPA beads provokes a signal, even though the radio-ligand is not linked to the bead (Auld *et al.* 2004). On the other hand, it is also probable that 4F2hc position relative to LAT1, could compromise the binding of substrate and inhibitor to LAT1 binding site. To check this last possibility, SPA experiments for determination of the binding properties of LAT1 without 4F2hc were conducted.

3.4.6.2 Binding properties of purified vertebrate LAT1

Isolated membranes from a 6 L Pichia culture expressing vertebrate LAT1 (see Materials and Methods section 5.3.4), were solubilized in 2 % DDM/0.4 % CHS. LAT1 was purified from the soluble fraction by strep-tactin affinity chromatography in 0.05 % DDM/0.01 % CHS (see Materials and Methods section 5.3.5.1). Then, protein was concentrated to 0.5 mg/ml, and soluble fraction was subjected to SPA (see Materials and Methods section 5.3.6), and SEC for the analysis of the protein quality in the SPA sample (Figure 40A). SEC chromatogram showed a major peak in the expected elution volume for LAT1, with a "left shoulder" corresponding to aggregated protein eluting in the column void volume, and a second minor non-protein peak, as showed by coomassie staining (Figure 40A). This non-protein peak could be micelles of DDM/CHS plus some lipids from *Pichia pastoris* membranes, that absorb at 280 nm, since it is known that some lipids absorb at such wavelength (Cunningham. 1994). In addition, in previous studies where different HATs were purified from Pichia or Saccharomyces membranes, this peak was generally observed (Álvarez-Marimon. 2014).

LAT1 SEC profile indicates that, although a part of LAT1 was aggregated, the major amount of it remains soluble, thus, it was possible that these part of the protein was well folded and able to bind substrate. On the other hand, LAT1 aggregation supports thinking that vertebrate 4F2hc also increases the stability of vertebrate LAT1, as it has been described for human 4F2hc/LAT2 (Rosell *et al.* 2014), since vertebrate 4F2hc/LAT1 heterodimer does not show significant aggregation as vertebrate LAT1 does (Figures 39A and 40A, respectively). Moreover, this result demonstrates again, that vertebrate LAT1 is more stable than human LAT2, which fully aggregated at comparable concentrations (Rosell *et al.* 2014).

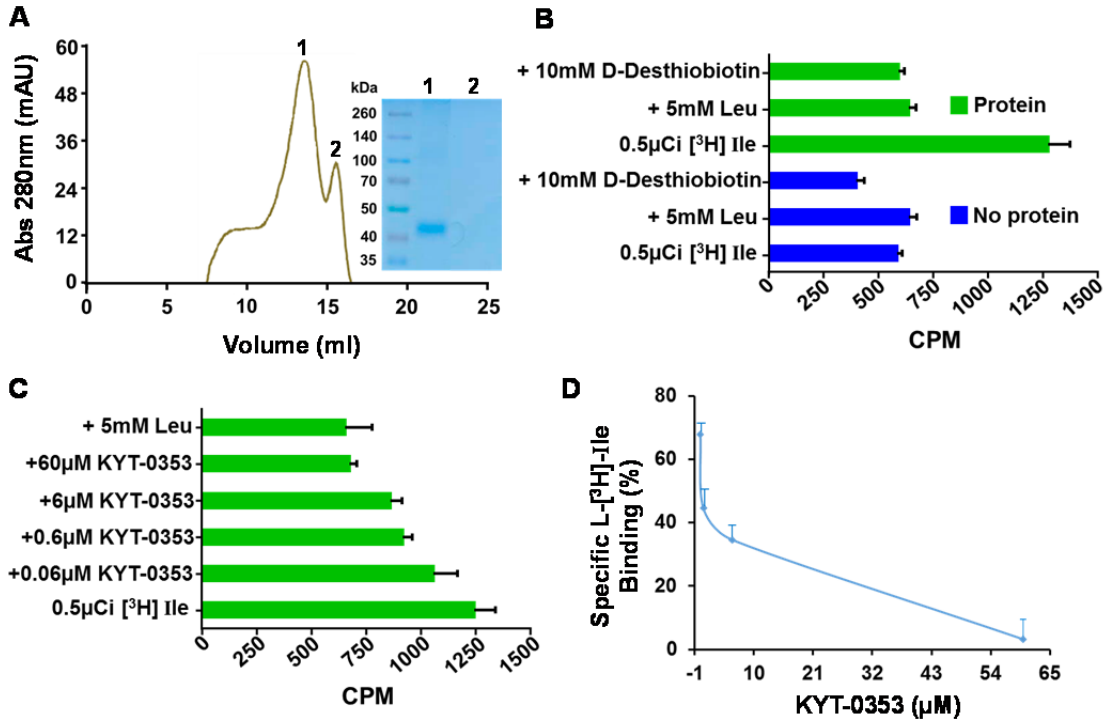

Figure 40. Substrate and KYT-0353 inhibitor binding assays by SPA in purified LAT1. LAT1 was purified by strep-tactin affinity chromatography in presence of 0.05 % DDM/0.01 % CHS, and concentrated to 0.5 mg/ml. The quality and purity of the sample were analyzed by SEC using a Superose 6, 10/300 GL column, and coomassie brilliant blue staining, respectively (**A**). SEC profile and coomassie staining indicated that purified and concentrated LAT1 used for SPA experiments was highly pure, although with some aggregates. Peak 1 correspond with LAT1 and peak 2 contains no protein. SPA was performed incubating for 1, 3 or 16 h streptavidin-YSi beads (220 µg per well), 5 µg of purified LAT1 and 0.5 µCi L-[^{3}H]-radiolabeled isoleucine, as a substrate (**B**). Determination of the unspecific binding of L-[^{3}H]-isoleucine to the YSi beads was performed including no-protein controls (**B**). A saturating concentration of non-radiolabeled L-leucine (+ 5 mM Leu) was also incubated with the mix described before to determine the maximum binding of could substrate, and thus also the maximum reduction of the SPA signal (CPM). Different concentrations of human LAT1 inhibitor KYT-0353 were also incubated with the mix, but in absence of cold substrate (**C**). From data in **C**, percentage of specific binding of L-[^{3}H]-isoleucine versus KYT-0353 inhibitor concentration was plotted. Data points indicate the means ± s.e.m. from triplicate determinations. D-desthiobiotin at 10 mM final concentration was also incubated with the same mix (beads/protein/L-[^{3}H]-Ile) as a control. Results shown in panels **B-D** correspond to 3 h of incubation, but are representatives for the other two times.

SPA was performed by incubating for different times, yttrium silicate (YSi) streptavidin coated beads, 5 µg of purified LAT1 and 0.5 µCi [^{3}H]-isoleucine, in all the conditions. Increasing amounts of KYT-0353 (0.06 µM to 60 µM) were used to characterize the inhibition of vertebrate LAT1 by KYT-0353. In the control conditions, 5 mM L-leucine or 10 mM D-desthiobiotin

were used to determine the specificity of the L-[^{3}H]-isoleucine binding signal. Significant specific binding of radiolabeled isoleucine to LAT1 was detected by SPA (Figure 40B and 40C). This was higher than the specific binding of radiolabeled isoleucine previously determined for the heterodimer (Figure 39B). This result suggest that LAT1 expressed in Pichia and purified in DDM/CHS is functional. On the other hand, SPA signal linked to the binding of L-[^{3}H]-isoleucine to LAT1 was displaced by increasing concentrations of KYT-0353 inhibitor (Figure 40C and 40D), suggesting specific binding of KYT-0353 to this vertebrate LAT1. These results further support the inhibition of LAT1 by KYT-0353, as the highest concentration of KYT-0353 showed similar reduction in the SPA signal to that obtained in the 5mM cold L-leucine or 10 mM D-desthiobiotin control conditions (Figure 40B and 40C). However, as these results are very preliminary, further SPA optimization must be performed, as well as, determination of the type of inhibition and the Ki of KYT-0353 for vertebrate LAT1, by using more inhibitor and substrate concentrations.

Curiously, despite chromatographic profile of 4F2hc/LAT1 is quite better than that of LAT1, not being detectable significant aggregation in the heterodimer like it is in LAT1, higher specific binding of substrate and inhibitor was determined for LAT1 than for 4F2hc/LAT1. As a consequence, it is probable that position of 4F2hc relative to LAT1, could be interfering in the substrate or inhibitor interaction with LAT1 in the SPA experiment. Thus, 4F2hc/LAT1 reconstitution in liposomes must be further tested for studying the functionality of this heterodimer. In fact, it has been recently described a new protocol for the reconstitution in liposomes of other metazoan HAT; rBAT/AGT-1 complex (Nagamori *et al.* 2016). Due to the efficiency in the reconstitution of a HAT, such protocol could be successfully used in the reconstitution of vertebrate 4F2hc/LAT1 and must be tested.

3.4.7 Vertebrate LAT1 shows size heterogeneity as determined by negative staining

To analyze sample dispersity, negative staining of LAT1 purified in DDM/CHS was performed. We started with this subunit, because it could be the most susceptible to heterogeneity, among both proteins, due to its major content of TMDs (twelve) that can cause protein aggregation, and can be surrounded by different amounts of DDM/CHS and endogenous membrane lipids. Moreover, the analysis of the heterogeneity linked only to one subunit, could be easier to study as a first step. For doing this, the top fraction of a LAT1 SEC peak was directly subjected to negative staining with uranyl formate as described in Materials and Methods section 5.4.1. Most of the sample showed particles of LAT1 corresponding to putative monomeric forms of the protein, and few aggregates (Figure 41). However, it is observed heterogeneity in the size of the LAT1 particles, ranging from 8 to 23 nm in diameter. If the same heterogeneity is present in 4F2hc/LAT1, particle analysis for structural determination by cryo-EM could be difficult, which also may lead to low resolution structural models.

In fact, very preliminary studies by negative staining performed with purified vertebrate 4F2hc/LAT1 in DDM/CHS, also showed polydispersity in the size of the particle corresponding to LAT1 (data not shown). This heterogeneity could be related with a variable size of the DDM/CHS micelles in the buffer used for the solubilization and purification steps, affecting the apparent size of the protein surrounded by these micelles. In addition, these micelles could provide inefficient stripping of lipids from 4F2hc/LAT1 during solubilization, generating 4F2hc/LAT1 molecules differentially delipidated, which could be responsible for the observed

heterogeneity. For these reasons, in the next sections, 3.4.8 and 3.4.9, possible strategies for reducing or eliminating this polydispersity are discussed and their results are shown.

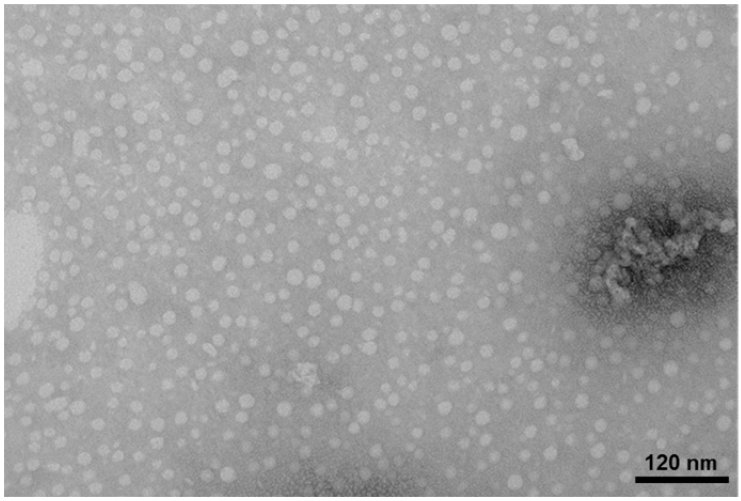

Figure 41. Analysis by negative staining of purified LAT1. Vertebrate LAT1 was purified by strep-tactin affinity chromatography, concentrated to 0.5 mg/ml, and subjected to SEC using a Superose 6, 10/300 GL column. Purified LAT1 from the SEC peak, at 30 μg/ml, was subjected to negative staining using uranyl formate. A representative electron micrograph of purified negatively-stained LAT1 is shown. Although there is some aggregation, most of the protein is observed as single particles, however, heterogeneity in the particle size is observed, ranging from 8 to 23 nm in diameter.

3.4.8 CHS concentration reduction compromises vertebrate 4F2hc/LAT1 stability

Detergent solubilization of membrane proteins, generally results in protein–detergent complexes (PDCs) (le Maire *et al.* 2000). However, depending on the membrane protein, the detergent and the protocol used during the solubilization and purification, PDCs are going to contain different amounts and types of lipids interacting with transmembrane domains of the protein. Some structures of membrane proteins have revealed essential protein-lipid interactions for the protein function, and crystallization (Ferguson *et al.* 2000; Jones 2007). Similarly, in recent years, it has been proposed a role of cholesterol in the control of transport in neurotransmitter sodium symporters, since it was found in the crystal structure of Drosophila DAT between two TMDs, determining the movement of one of them during the transport (Penmatsa *et al.* 2013). For these reasons, it is important to preserve relevant interactions between membrane proteins and lipids during the solubilization and purification processes. However, insufficient membrane solubilization or insufficient membrane protein delipidation can provide excess of lipids in the protein, which can increase heterogeneity in the sample (Garavito *et al.* 1996), and hinder cryo-EM studies. For this reason, an equilibrium between the preservation of lipidic interactions relevant for protein stabilization and activity, and elimination of those that are not essential, must be achieved by testing different detergent conditions during the solubilization and purification processes.

In our case, during this whole work, vertebrate 4F2hc/LAT1 has been solubilized with DDM in the presence of CHS, because in previous studies it was found an increment in the stabilization of metazoan HATs by the presence of CHS (Álvarez-Marimon. 2014). Moreover, CHS has demonstrated its suitability in the stabilization of membrane proteins for structural studies, since

it has been used in the crystallization of several proteins, mainly G-protein coupled receptors, leading to the determination of their high resolution structures (He *et al.* 2014). Nevertheless, due to the heterogeneity found by negative staining in LAT1 subunit (Figure 41), it is feasible thinking that the emulsification of the CHS in the DDM-containing buffer that is used for the solubilization of Pichia membranes, could provide different DDM/CHS micelles sizes and inefficient stripping of lipids from 4F2hc/LAT1, which could be responsible for the different LAT1 particle sizes. In this way, we thought that it could be possible to increase CHS solubilization yield in DDM by the reduction of CHS concentration. This, could cause two possible related effects: First, an increased homogeneity of the DDM/CHS micelles sizes, and second, more homogeneous 4F2hc/LAT1 delipidation. In this way, both effects could increase LAT1 monodispersity. Thus, if 4F2hc/LAT1 could be obtained stable enough in less concentration of CHS, dispersity of LAT1 size could be investigated by negative staining.

4F2hc/LAT1 was expressed in Pichia, and membrane fraction was divided in three equal samples. Each of them was solubilized in 2 % DDM but in the presence of a different concentration of CHS (0.4 %, which is the original condition, 0.2 % and 0.1 %). Soluble fraction from each condition was subjected to double-affinity purification (see Materials and Methods section 5.3.5.2) in 0.05 % DDM and the corresponding concentration of CHS (0.01 % that is the original condition, 0.005 % and 0.0025 %, respectively). Eluted protein, in each condition, was concentrated to 0.5 mg/ml and subjected to SEC analysis. Chromatograms showed that 4F2hc/LAT1 losses stability as the concentration of CHS decreases; peaks get wider and more asymmetric as the concentration of CHS is reduced (Figure 42).

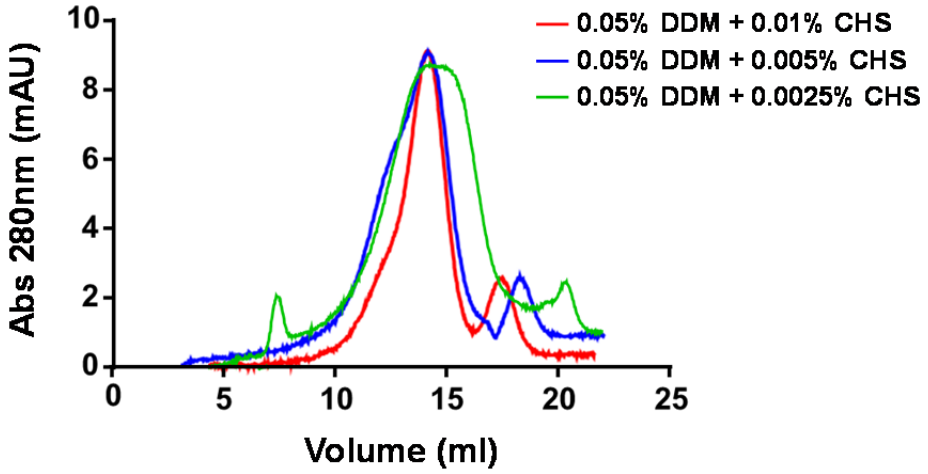

Figure 42. Impact of the CHS concentration in the stability of 4F2hc/LAT1. 4F2hc/LAT1 was solubilized from isolated Pichia membranes in 3 different conditions: **1)** 2 % DDM/0.4 % CHS, **2)** 2 % DDM/0.2 % CHS and **3)** 2 % DDM/0.1 % CHS. Heterodimer was purified by a sequential double-affinity chromatography in presence of 0.05 % DDM in all cases but with a different CHS concentration, respectively for each condition: **1)** 0.01 % CHS, which is the standard condition used in the present work, **2)** 0.005 % CHS, or **3)**0.0025 % CHS. The eluted 4F2hc/LAT1 from strep-tactin resin was concentrated to 0.5 mg/ml and subjected to SEC using a Superose 6, 10/300 GL column. SEC profiles of 4F2hc/LAT1 clearly get worse as CHS concentration decreases.

This result corroborates the relevance of CHS in the stabilization of vertebrate 4F2hc/LAT1, and indicates that CHS concentrations less than a half of the generally used here (basal condition), cannot be used to try to reduce the heterogeneity in the sample, as it is required for the protein stability. Remains to be known if other concentrations of CHS, between the basal condition (0.01 %) and the half tested here (0.005 %), could stabilize the protein sufficiently.

3.4.9 Vertebrate 4F2hc/LAT1 is stable solubilized in amphipol A8-35

The advances in electron microscopes and direct electron detectors for cyo-EM studies, have led to the determination of structures belonging to membrane proteins of less than 200 kDa. However, there has been one additional factor that has enable obtaining these proteins more stable and better defined for cryo-EM studies, than usually detergents do, amphipol A8-35. This molecule has been successfully used to solve the structures of the γ-secretase, and TRPV1 and TRPV2 channels (Liao *et al.* 2013; Bai *et al.* 2015; Zubcevic *et al.* 2016).

Amphipols are amphipathic polymers composed of many hydrophilic and many hydrophobic groups, that are able to self-assemble into small and globular well-defined particles formed by a few molecules (Popot *et al.* 2011). Each molecule is composed of a short backbone chain and an appropriate number and size of hydrophobic chains, that makes that the equilibrium of particle size is reached by few molecules (Popot *et al.* 2011). Amphipols are designed to avoid the formation of large assembles like long hydrophilic loops, due to the high content of hydrophobic chains, and for this reason, these chains are relatively short and are quiet separated between each other and interspersed with highly polar groups (Popot *et al.* 2011).

The presence of large amount of hydrophobic groups per amphipol molecule, causes that single molecules of amphipol can be more tightly bound to the hydrophobic parts of a membrane protein, than a single molecule of most of the detergents (Popot *et al.* 2011). Thus, amphipols have higher affinity for membrane proteins than detergents generally do. As a result, some membrane proteins with bound amphipol have been more stably maintained in solution than when they have been bound to detergents, even after sample dilution (Popot *et al.* 2011; Kleinschmidt and Popot 2014). However, this effect could depend on the particular protein, amphipol and detergent evaluated. In contrast, amphipols are not efficient in solubilizing membranes, and for this reason, membrane proteins are solubilized in detergents and then these are exchanged by amphipols (Popot *et al.* 2011).

The following characteristics of amphipol A8–35, that could make it useful for cryo-EM studies of 4F2hc/LAT1, were summarized by Popot and co-workers (Popot *et al.* 2011). A8–35 (Figure 43) is highly soluble in water and at above pH 7, because most of their carboxylates are ionized. For this reason, it must be used at these pHs, which was in concordance with the working pH in the 4F2hc/LAT1 purification. A8-35 has an average molecular weight of 9 or 10 kDa, depending on the average length of the polyacrylic acid used for its synthesis. This amphipol forms particles, composed of four self-assembled molecules, of around 40 kDa, which makes possible working at very dilute conditions (Giusti *et al.* 2012). A8-35 particles are monodisperse, and bound to membrane proteins may restrict large movements of transmembrane domains, locking proteins in a single conformation, which would increase protein sample homogeneity significantly, and thus, being favorable for cryo-EM studies (Althoff *et al.* 2011; Popot *et al.* 2011). Therefore, A8-35 could be useful to decrease the 4F2hc/LAT1 heterogeneity, observed in the size of LAT1 particles by negative staining (Figure 41), since also the heterogeneity probably related to the different size of DDM/CHS micelles, could be eliminated by solubilizing the protein in amphipol. Additionally, A8-35 has proven to better separate particles into the holes of cryo-EM grids, and to facilitate sample preparation in comparison with detergents samples (Althoff *et al.* 2011; Cabra and Samso 2015).

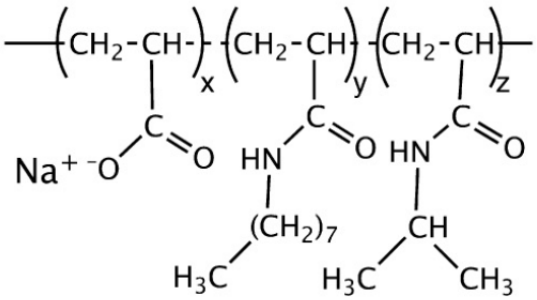

Figure 43. Chemical structure of polyacrylate-based amphipol A8-35. As most of amphipols, A8–35 is composed of hydrophobic monomers with carbon alkyl groups of 8 to 12 carbons, interspersed by one or a few hydrophilic monomers. A8–35 is carboxylated, amidated and anionic. The indices x, y, z denote the proportion of the various units: x: 35 %, y: 25 %, z: 40 %. (Kleinschmidt and Popot 2014).

Thus, we decided, to test the stability in solution of 4F2hc/LAT1 in amphipol A8-35, to investigate its potential for being used in cryo-EM studies with our transporter. The heterodimer, expressed in Pichia, was solubilized and double affinity purified in the basal condition used through this work (2 % DDM/0. 4% CHS, and 0.05 % DDM/0.01 % CHS, respectively) (see Materials and Methods section 5.3.5.2). After elution from the strep-tactin affinity chromatography, 4F2hc/LAT1 was concentrated to 0.1 mg/ml, and incubated with A8-35 in a ratio 1:3 (protein/amphipol, w/w), for exchanging the detergent bound to the protein by the amphipol A8-35. Finally, free DDM was removed by addition of Bio-Beads SM-2 (see Materials and Methods section 5.3.7).

At next day, SEC of the soluble fraction, at 1 mg/ml, was performed in buffer with neither detergent nor lipid. The SEC profile of 4F2hc/LAT1, in which exchange of DDM for amphipol A8-35 was done, showed a major peak of stable and soluble protein (Figure 44). The peak of 4F2hc/LAT1 in amphipol A8-35 is narrower and more symmetric than that in DDM/CHS (Figure 44A), which suggests that A8-35 is as, or more effective in stabilizing 4F2hc/LAT1 than DDM, and could had improved protein monodispersity. However, this has to be proven by negative staining experiments. No reference to CHS depletion is made, because it is probable that the interaction between the CHS and the protein remains intact when the DDM is exchanged by amphipol. It has been described that amphipols respect lipid-protein interactions (Althoff *et al.* 2011; Popot *et al.* 2011), and in fact, sometimes lipids required for the stabilization and activity of a protein, are added before the exchange of detergent by amphipol (Popot *et al.* 2011; Liao *et al.* 2013). Therefore, it is highly probable that CHS remains bound to 4F2hc/LAT1 upon A8-35 exchange.

A shift to the right in the elution volume of the amphipol-solubilized heterodimer (Figure 44B), indicates an apparent reduction in the size of the 4F2hc/LAT1. This could correlate with the exchange of the ~72 kDa micelles of DDM by the ~40 kDa particles of amphipol. This result supports that the stability of amphipol-solubilized 4F2hc/LAT1 observed by SEC, would be due to the binding of the amphipol to the protein and the stripping of DDM. This experiment was performed several times showing reproducibility. In Figure 44 a representative SEC profile of 4F2hc/LAT1 solubilized in amphipol A8-35 is shown.

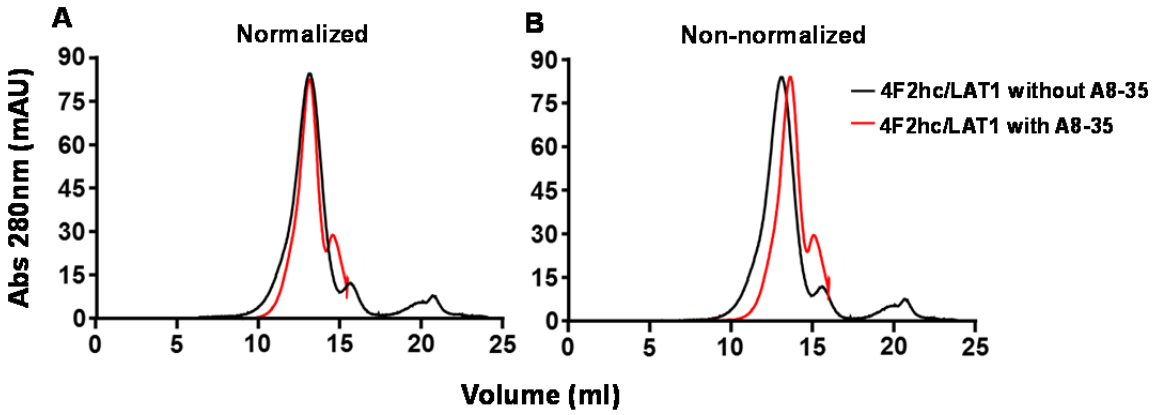

Figure 44. 4F2hc/LAT1 is stable solubilized in amphipol A8-35. Heterodimer was purified by a sequential double-affinity chromatography in presence of 0.05 % DDM/0.01 % CHS. The eluted 4F2hc/LAT1 was incubated with amphipol A8-35 during 4h, and the detergent was removed from the sample by stirring for 9h in presence of Bio-Beads SM-2. After concentration to 1 mg/ml, sample was analyzed by SEC using a Superose 6, 10/300 GL column (**A**, **B**). The practically monodisperse SEC peak corresponding to 4F2hc/LAT1 with amphipol A8-35 bound, shows that, in this condition, this heterodimer is at least, as stable as the complex purified in DDM/CHS (**A**). The difference in the 4F2hc/LAT1 elution volume in presence or absence of A8-35, indicates a reduction in the apparent size probably due to the DDM/A8-35 exchange (**B**).

3.4.9.1 Stability of vertebrate 4F2hc/LAT1 in amphipol A8-35 over time

Like 4F2hc/LAT1 was obtained in a new environment, A8-35/CHS, it was important to analyze the stability of the protein in such condition, since also this would be the final condition of the sample for cryo-EM. Thus, double affinity purified 4F2hc/LAT1 was subjected to DDM exchange by A8-35 as described previously (see Materials and Methods section 5.3.7), and concentrated to 1 mg/ml or 0.8 mg/ml (Figure 45A and 45B, respectively). After SEC, fractions from the peak enclosed in the blue rectangle (Figure 45A), were concentrated to 1 mg/ml and left for 16 h at 4°C, and fractions from the peak enclosed in the green rectangle (Figure 45B), were left without concentrate for 16 h at 4°C.

Eighteen hours after the SEC, protein concentrated the day before (1 mg/ml) was subjected to SEC (Figure 45C), meanwhile protein that has not been concentrated the day before, was concentrated to 0.7 mg/ml and analyzed by SEC (Figure 45D). Both experiments revealed that once the protein is in amphipol A8-35, and there is not presence of DDM/CHS micelles in the sample, 4F2hc/LAT1 starts to aggregate upon time in a non-concentration dependent manner. This is evidenced by the widening to the left in the base of the peak in both samples, when protein was concentrated the first day (Figure 45C), and when it was concentrated the next day (Figure 45D). This indicates that cryo-EM experiments must be conducted as soon as the DDM/amphipol exchange has been performed. However, it is not known if amphipol-solubilized protein could be maintained stable in solution if preserved at -80°C, in which case protein could be kept freeze until cryo-EM grids preparation. The decreased stability showed over time by A8-35-solubilized 4F2hc/LAT1 (Figure 45), in comparison with the stability maintained by DDM/CHS-solubilized 4F2hc/LAT1 over time (Figure 36), suggests that amphipol A8-35 has not been as successful as expected in the stabilization of this heterodimer. Nevertheless, it is worthy to mention that the heterodimer after the DDM/amphipol exchange and SEC (day 1), is in a buffer with no detergent, no CHS, and no amphipol, and additionally, at day 2, it has stayed one night more in solution, due to the DDM/amphipol exchange, than

heterodimer purified in DDM/CHS (1 day after in Figure 36). Thus, the fact that great part of the amphipol-solubilized protein remains soluble at the second day after the exchange and in this buffer (Figure 45), suggests a positive effect of amphipol A8-35 in the stabilization of 4F2hc/LAT1. The great part of the protein that seems to be stable at second day is observed by a major peak of protein, which would be symmetric unless by the deformation caused by the elution of protein partially aggregated (Figure 45C and 45D).

It is possible that the DDM/A8-35 exchange was not complete, and that some DDM/CHS that could remain bound to some protein molecules after the exchange, was stripped from the protein when it is diluted in buffer without DDM, CHS or amphipol during SEC, causing the partial aggregation of the protein. Additionally, it has been reported that amphipols respect relevant protein-lipid interactions, but in cryo-EM studies of γ-secretase (Lu *et al.* 2014; Bai *et al.* 2015) and TRPV1 channel (Liao *et al.* 2013) solubilized in amphipol, other lipids but not CHS were used, therefore the efficiency of the detergent/amphipol exchange in presence of CHS is not widely reported, neither it is whether CHS would remain bound to the protein after such exchange or if is lost during the process. CHS depletion from the protein after the DDM/amphipol exchange, probably during SEC, may explain the aggregated protein observed in the SEC profiles (Figure 45C and 45D), and it would be in concordance with the essential role of CHS for the stabilization of 4F2hc/LAT1, as it was demonstrated here (Figure 42).

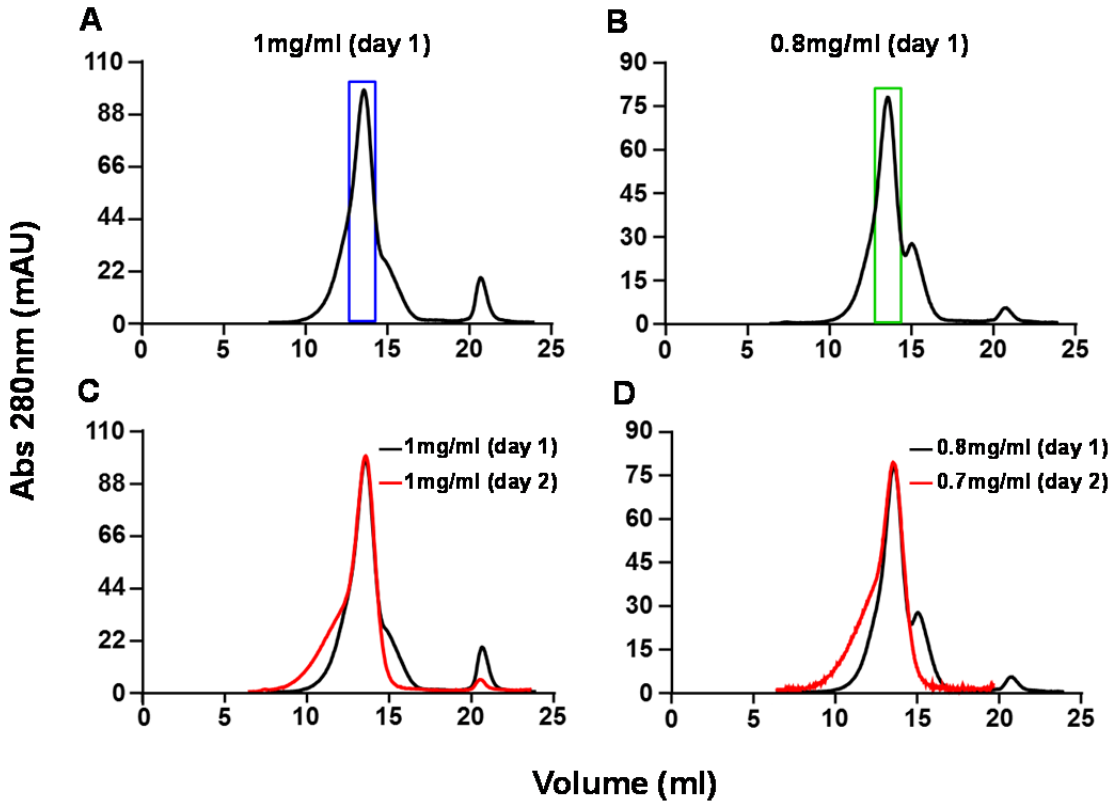

Figure 45. Stability of 4F2hc/LAT1 solubilized in amphipol A8-35 over time. Double-affinity purified 4F2hc/LAT1 in presence of 0.05 % DDM / 0.01 % CHS was subjected to detergent exchange into A8-35. Then, protein was concentrated to 1 mg/ml or 0.8 mg/ml and analyzed by SEC using a Superose 6, 10/300 GL column (**A** and **B,** respectively). Pooled fractions, enclosed in a blue or green rectangle, from **A** and **B,** were concentrated to 1 mg/ml or no concentrated, respectively, and, in both cases, were left for 16 h at 4°C. At the next day, 4F2hc/LAT1 concentrated at 1 mg/ml (from **A**) and at 0.7 mg/ml (from **B**) were subjected to SEC using a Superose 6, 10/300 GL column (**C** and **D**, respectively). Comparison of the normalized SEC profiles from heterodimer obtained at days 1

and 2, shows a widening to the left respect to the peaks from the day before, indicating that the protein in A8-35 is aggregating upon time (**C**, **D**).

In conclusion, the apparent good stability of 4F2hc/LAT1 in amphipol A8-35 at day 1, and the fact that most of this protein seems to be stable over time (Figure 44 and Figure 45), suggest quite potential of amphipol A8-35 for the stabilization of 4F2hc/LAT1. Thus, due to the advantages of these amphipathic polymers for cryo-EM studies, other experiments to optimize the exchange of DDM/CHS by amphipol must be performed, for trying to increase 4F2hc/LAT1 stability in amphipol, if this is the cause of the loss of stability. For doing this, a higher concentration of A8-35 could be used, as it will be discussed later in section 3.5. Additionally, the time of incubation with amphipol could be also incremented, and the DDM depletion from the sample by using BioBeads could be also optimized by adding more beads, or by incrementing the time of incubation. In any case, it could be advisable not to extend the incubation periods, as this could be harmful for the protein.

3.4.9.2 Analysis of amphipol A8-35-solubilized 4F2hc/LAT1 by negative staining

In order to analyze 4F2hc/LAT1 dispersity in amphipol, and to evaluate its suitability for cryo-EM studies, negative staining was performed. Hereafter negative staining, particle analysis and 3D-reconstructions, were performed in collaboration with Professor Oscar Llorca (CIB-CSIC, Madrid) under the support and guidance of Dr. Rafael Nuñez-Ramírez.

Double affinity purified 4F2hc/LAT1 was subjected to DDM exchange by A8-35, and concentrated to 1 mg/ml (see Material and Methods section 5.3.7 for detailed information). After SEC, fractions from the 4F2hc/LAT1 peak (Figure 46A) were pooled and left at 4°C. At next day, 20 hours after SEC, protein without a further SEC step was concentrated to 0.6 mg/ml and subjected to negative staining as described in Materials and Methods section 5.4.1 .

Three protein concentrations (60 μg/ml, 12 μg/ml and 6 μg/ml) were used to find the optimal condition for single particle analyses by negative staining, and to estimate the potential protein concentration that could be used for further vitrification in cryo-EM. Generally, a concentration around 10-fold than the optimal for negative staining studies is suited for cryo-EM (personal communication, Dr. Rafael Nuñes-Ramirez). The optimal concentration of 4F2hc/LAT1 for particle analysis by negative staining was 6 μg/ml, since at this concentration enough images of discrete particles could be analyzed. Thus, it seems that the protein is being obtained in the range of concentration needed for cryo-EM experiments, and it is probable that protein does not have to be concentrated after SEC, which could be worthy for maintaining protein stability.

By using the Xmipp software (de la Rosa-Trevin *et al.* 2013), automatic particle selection and classification were done, and reference-free 2D class averages were determined (Figure 46B), as described in Materials and Methods section 5.4.2. Most of the particles seemed to be composed of two bodies, which would correspond with the heterodimer formed by 4F2hc and LAT1. However, in some images the view of two associated subunits is less notable, probably because the orientation of 4F2hc/LAT1 particles, in such images, impedes to see one of the subunits. Since no free 4F2hc or LAT1 were observed by western blot and coomassie staining (Figure 46B), it is improbable that this apparent single particles could be one of these subunits that are not forming heterodimer. Similar particles formed by two bodies of different size has been previously observed by negative staining of human 4F2hc/LAT2 (Meury *et al.* 2014; Rosell *et al.* 2014). However, human 4F2hc/LAT2 purified in presence of DDM had significant tendency to dissociate in its two subunits, whereby free subunits were also observed by negative staining

(Rosell *et al.* 2014). In this work, we have also showed, significant disruption of human 4F2hc/LAT2 purified under the same conditions than vertebrate 4F2hc/LAT1 (DDM/CHS) (Figure 37). This result supports that 4F2hc/LAT1 is a more stable complex, being more suitable for structural studies by cryo-EM, since less disruption of heterodimer would facilitate particle classification. In the 3D-reconstruction of human 4F2hc/LAT2, the atomic structure of human 4F2hc-ED fitted well inside the 3D map of the smaller density (Figure 50B and 50C), despite of being a bigger protein (heavy subunit) than LAT2 (light subunit). The bigger size of the density corresponding to LAT2 in relation with the expected was then attributed to the DDM, CHS and lipids bound to the 12 TMDs of LAT2 and to the single N-terminal TMD of 4F2hc (Meury *et al.* 2014; Rosell *et al.* 2014), which would be similar for vertebrate 4F2hc/LAT1, as we will show later here (Figure 50A).

It is notable the heterogeneity observed in the size of the bigger body in 4F2hc/LAT1 (Figure 46B), similar to it was previously observed in the micrographs of LAT1 (Figure 41), although in the latter case, further analysis like particle classification and determination of 2D class averages were not made. This indicates that the polydispersity in the sample remains present even after the exchange of DDM by amphipol, suggesting that this exchange is not complete, or that the amount of CHS and/or lipids that remain bound to the protein is variable between the 4F2hc/LAT1 particles, as also discussed before. Therefore, further studies to increase monodispersity of the 4F2hc/LAT1 complexes in the sample for cryo-EM must be conducted.

For being able to analyze 4F2hc/LAT1 particles, a more homogenous set of particles was obtained by discarding those excessively big, as estimated by the previously observed for human 4F2hc/LAT2 (Meury *et al.* 2014; Rosell *et al.* 2014). After classification of this population, 2D class averages showed more clearly the two bodies, and less heterogeneity between them regarding the LAT1 particle (Figure 46C). However, it is probable that some of this heterogeneity is not caused by differences in the size of LAT1 particle, attributed to the differential incorporation of amphipol, lipids and CHS, but because particles are being observed from different angles. These results strongly suggest the presence of a subpopulation of non-aggregated and monodisperse 4F2hc/LAT1 in the sample, which should be enriched for future cryo-EM studies.

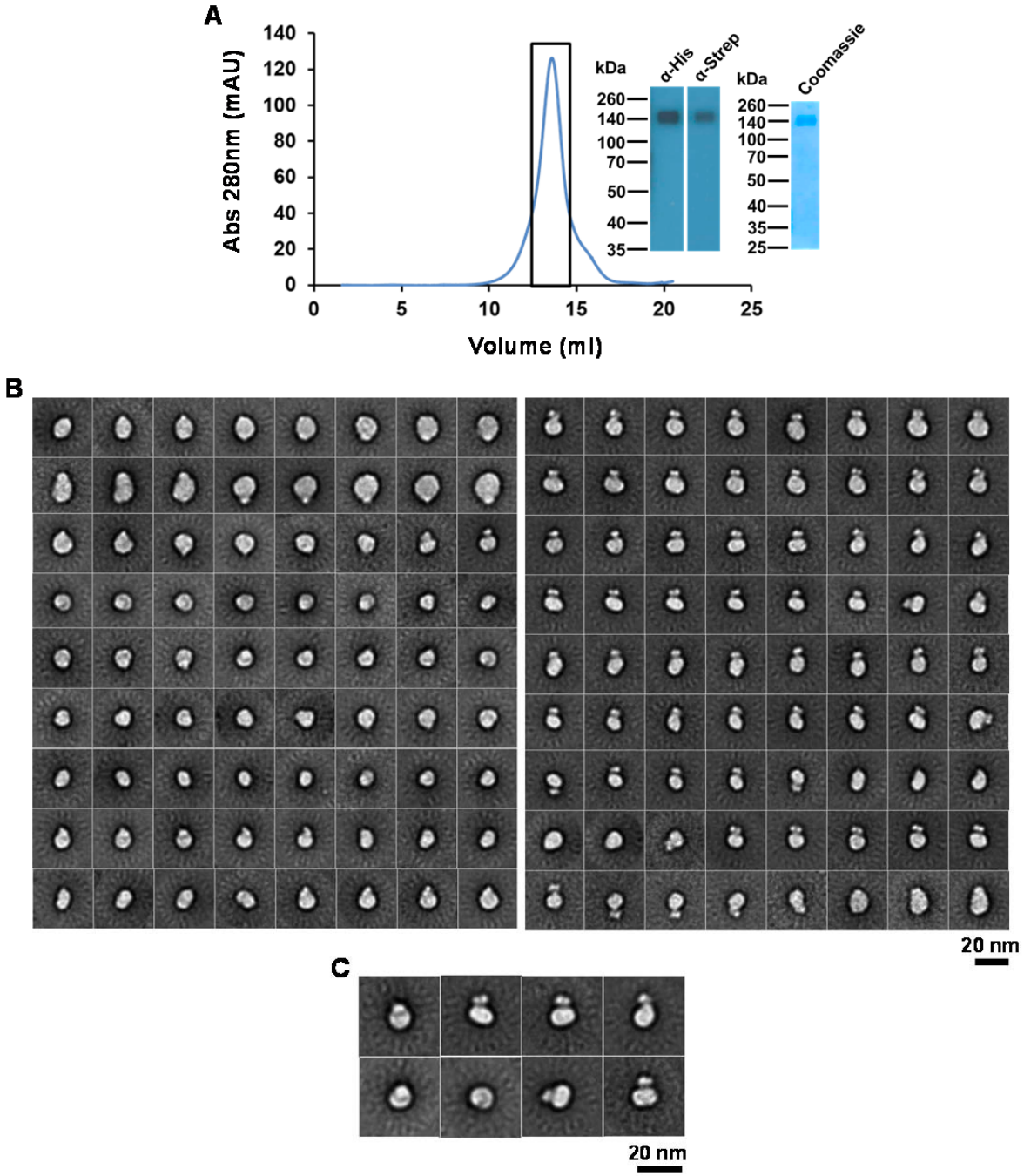

Figure 46. Negative staining of A8-35-solubilized 4F2hc/LAT1. Double affinity purified 4F2hc/LAT1 in presence of 0.05 % DDM/0.01 % CHS was subjected to detergent exchange into amphipol A8-35. Sample was concentrated to 1 mg/ml and subjected to SEC using a Superose 6, 10/300 GL column (**A**). Fractions of 4F2hc/LAT1 peak, enclosed in a black rectangle, were pooled and left for 16 h at 4°C. At the next day, sample was concentrated to 0.6 mg/ml and subjected to negative staining with 2 % uranyl acetate at 3 different concentrations (60 µg/ml, 12 µg/ml and 6 µg/ml). Images for the 6 µg/ml sample analysis are shown, as this concentration was found to be the optimal for these studies. Galleries of 2D class averages of 4F2hc/LAT1 particles from the whole sample (**B**). A selection of 2D class averages that showed more structural details after reclassification are magnified and displayed in Panel **C.** In general, two bodies of different size are observed.

3.4.9.3 Effect of CHS depletion in the stability of 4F2hc/LAT1 solubilized in amphipol A8-35

Size heterogeneity in the LAT1 particles was evident in the single particle analysis by negative staining of A8-35-solubilized 4F2hc/LAT1 (Figure 46B). Amphipols are supposed to reduce protein heterogeneity by replacing detergent, in our case DDM. However, lipids remain bound to the transmembrane domains of membrane proteins after the detergent-amphipol exchange. In our case, CHS is probably retained in 4F2hc/LAT1, but the percentage of bound molecules may be variable, one possibility, which would explain the heterogeneity of the particle corresponding to LAT1. Additionally, it is possible that the emulsion of CHS in DDM buffers used for solubilization and purification of the protein, could hinders homogeneous protein delipidation, producing 4F2hc/LAT1 molecules with different incorporation of lipids and, as a consequence, with different sizes in the LAT1 particle.

We evaluated the possibility to obtain a more homogeneous sample by eliminating the possible excess of CHS and/or lipids from the protein during the final steps of purification, and after exchange of DDM by amphipol. As it was observed previously, that using less concentration of CHS during the solubilization and purification of 4F2hc/LAT1 highly decreases its stability (Figure 42), this time, CHS was maintained during solubilization and in the initial steps of purification, and then was eliminated from the purification buffers. In this way, it could be possible that 4F2hc/LAT1 would retain bound lipids and CHS, that would be essential for its stability, and loses the excess of those not essential or weakly bound, making more homogeneous the sample.

For doing this, 4F2hc/LAT1 was solubilized in 2 % DDM/0.4 % CHS, and affinity purified by IMAC in DDM/CHS (0.05 %/0.01 %), as usual. Eluted sample was divided in two equal parts, and they were independently subjected to strep-tactin affinity chromatography. After protein binding, for one of them, column washings and protein elution were made using only 0.05 % DDM, whereas the other sample was purified in 0.05 % DDM and 0.01 % CHS. Then, in both samples exchange of DDM for A8-35 was performed. SEC analysis for both samples, concentrated at 1 mg/ml, was carried out to compare the stability of 4F2hc/LAT1 in amphipol A8-35 after CHS depletion, with that in which CHS is present until the end of purification. Superposition of SEC profiles of both samples showed that protein in which CHS has been eliminated during the final steps of the purification process, is less stable, showing much more polydispersity and tendency to aggregation than the sample that was maintained in presence of CHS (Figure 47). This result strongly supports that, by the moment, and unless other condition is found, CHS is essential for the stability of vertebrate 4F2hc/LAT1, even when DDM is exchanged by amphipol A8-35. Then, other strategies must be further tested for increasing protein monodispersity and stability for cryo-EM studies.

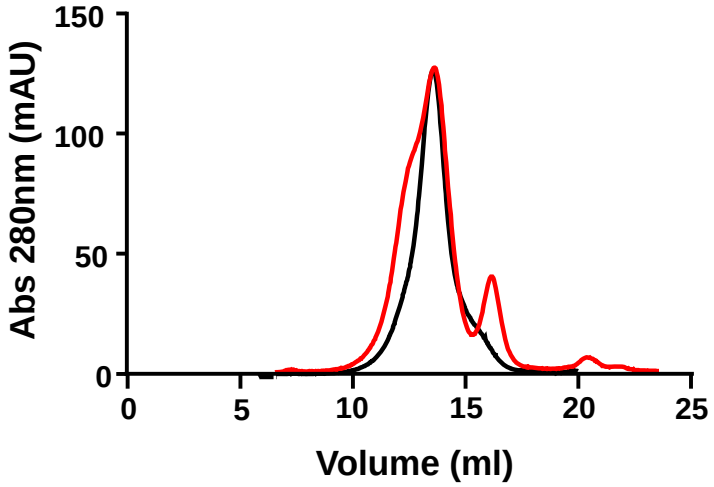

Figure 47. 4F2hc/LAT1 solubilized in amphipol A8-35 needs CHS to remain stable. Heterodimer was purified by IMAC in 0.05 % DDM/0.01 % CHS. After protein binding with strep-tactin resin, column washings were performed with buffer containing 0.05 % DDM but not CHS. Elution was performed in the same buffer and the exchange of DDM by amphipol A8-35 was carried out. Sample was concentrated to 1 mg/ml and analyzed by SEC using a Superose 6, 10/300 GL column (red line). The comparison with the SEC profile of the protein in amphipol when it has been purified without eliminating the CHS during the purification process (black line), indicates that 4F2hc/LAT1 stability decreases when CHS is reduced even though the protein was going to be solubilized in amphipol. Normalized peaks are shown.

3.4.10 3D reconstruction of amphipol-solubilized 4F2hc/LAT1

Taking into account that the particle heterogeneity in 4F2hc/LAT1, solubilized in A8-35, was observed 20 hours after SEC (Figure 46B), and that in previous experiments, it had been shown that similar sample (18 hours after SEC), was polydisperse and showed tendency to aggregate (Figure 45). It was probable that part of the heterogeneity observed by microscopy, was due to the deterioration or loss of stability of 4F2hc/LAT1 upon time. For this reason, A8-35-solubilized heterodimer, obtained as usual (see Materials and Methods section 5.3.7), was concentrated to 1 mg/ml, and just after SEC, central fractions of the peak were collected and immediately subjected to negative staining at three different protein concentrations (80 µg/ml, 15 µg/ml and 8 µg/ml).

The best concentration of protein found for analyzing single particles was 8 µg/ml, close to the concentration used in the previous experiment (6 µg/ml). Particles were classified and reference-free 2D class averages (Figure 48B) were determined by using the Xmipp software (de la Rosa-Trevin *et al.* 2013), as described in Materials and Methods section 5.4.3. Similar to the previously observed with protein that stayed 20 hours at 4°C after its purification by SEC (Figure 46), the protein subjected to negative staining immediately after its SEC purification, showed heterogeneity in the size of the big particle (LAT1). This indicates that the polydispersity observed in the sample is not due to the protein destabilization upon time, but it is maybe caused by differential content of amphipol, CHS and/or lipids, as previously discussed. However, in this occasion, less heterogeneity between particles in each class was found (Figure 48) in comparison with the previous experiment (Figure 46), although this difference was subtle. Achieving homogenous binding of amphipol, CHS and lipids to the protein, would increase the monodispersity in the sample. This would improve chances to determine 4F2hc/LAT1 structure

by cryo-EM, since particle images could be more properly classified in structural homogenous subsets and aligned for structural modeling.

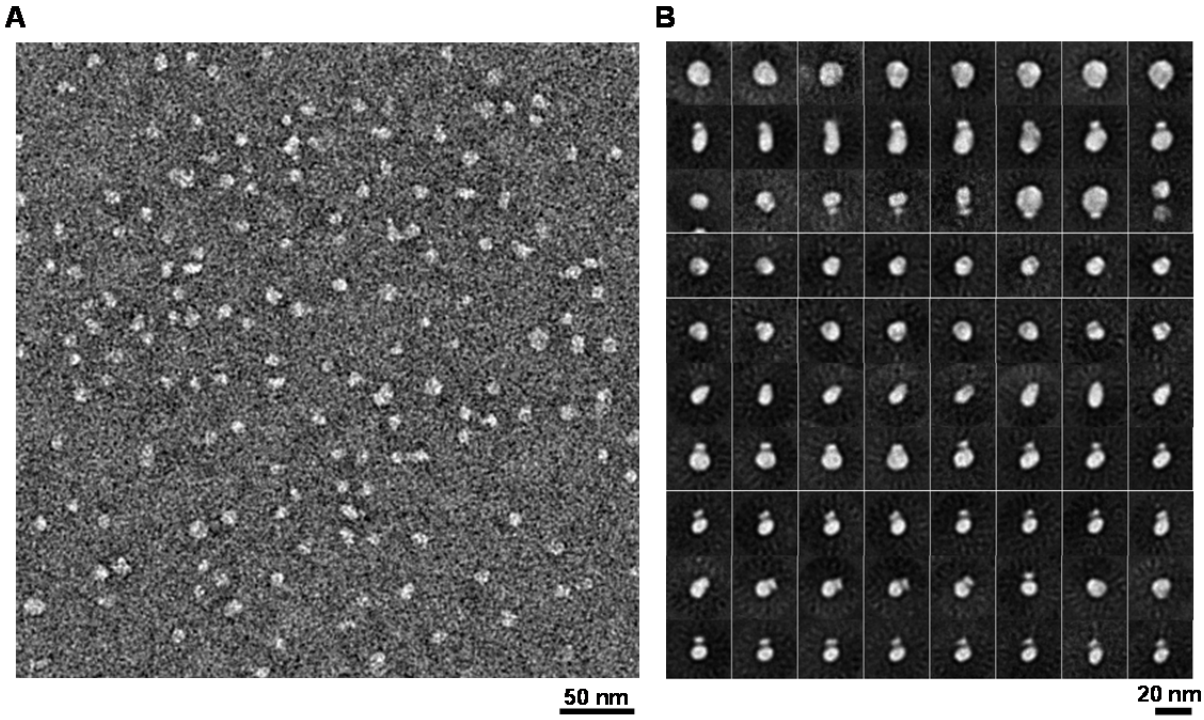

Figure 48. Negative staining of amphipol A8-35-solubilized 4F2hc/LAT1 after decreasing the time of processing. Micrograph of 4F2hc/LAT1 solubilized in amphipol and negatively stained with 2 % uranyl acetate the same day of purification by SEC (**A**). No big aggregates are observed. 2D class averages of particles in the sample (**B**). Heterogeneity in the big particle is also observed.

A homogeneous group of particles was used to calculate an initial 3D model, that was used as a 3D reference to classify particles into 3 groups (Figure 49A) with RELION (Scheres 2012), as described in Materials and Methods section 5.4.3. The third group, or class 3, was the most abundant in the sample and the better structurally defined, and for this reason was further analyzed. 2D class averages in this group showed more similar particle sizes, and better definition (Figure 49B) than the 2D averages obtained before the 3D classification (Figure 48B). Thus, for obtaining some structural information, 3D refinement was done for this group of particles (Figure 49C).

Similar to it was found for human 4F2hc/LAT2 (Figure 50B and 50C) (Meury *et al.* 2014; Rosell *et al.* 2014), in the present vertebrate 4F2hc/LAT1 3D map, two densities of different size are observed (Figure 50A). The good fitting of the structure of human 4F2hc-ED inside the small density (Figure 50A) corroborates that this density would correspond to vertebrate 4F2hc, and the bigger density to LAT1, as previously described for human 4F2hc/LAT2 (Figure 50B and 50C) (Meury *et al.* 2014; Rosell *et al.* 2014). Thus, here we also observed 4F2hc on top of LAT1, correlating this with previous findings, in which molecular docking between human 4F2hc-ED and LAT2 model, fabricated based on the atomic structure of the amino acid transporter AdiC, and cross-linking experiments, supported this position of heavy subunit in relation to light subunit, probably establishing 4F2hc-ED interactions with some of the LAT2 loops (Rosell *et al.* 2014). However, in contrast to the observed in human 4F2hc/LAT2, a space between 4F2hc and LAT1 densities is observed in the present 3D map, as well as the presence

of two densities connecting both subunits. One of the densities linking 4F2hc and LAT1 (Figure 50A) would be indicating the position of the N-terminal TMD of 4F2hc, as previously described for 4F2hc/LAT2 (Figure 50B and 50C) (Meury *et al.* 2014; Rosell *et al.* 2014). Thus, also this density could be showing the approximate position of the disulphide bond among both subunits, since the cysteine of 4F2hc forming the disulphide bond with LAT1 is a few residues away from the 4F2hc TMD, towards the extracellular space.

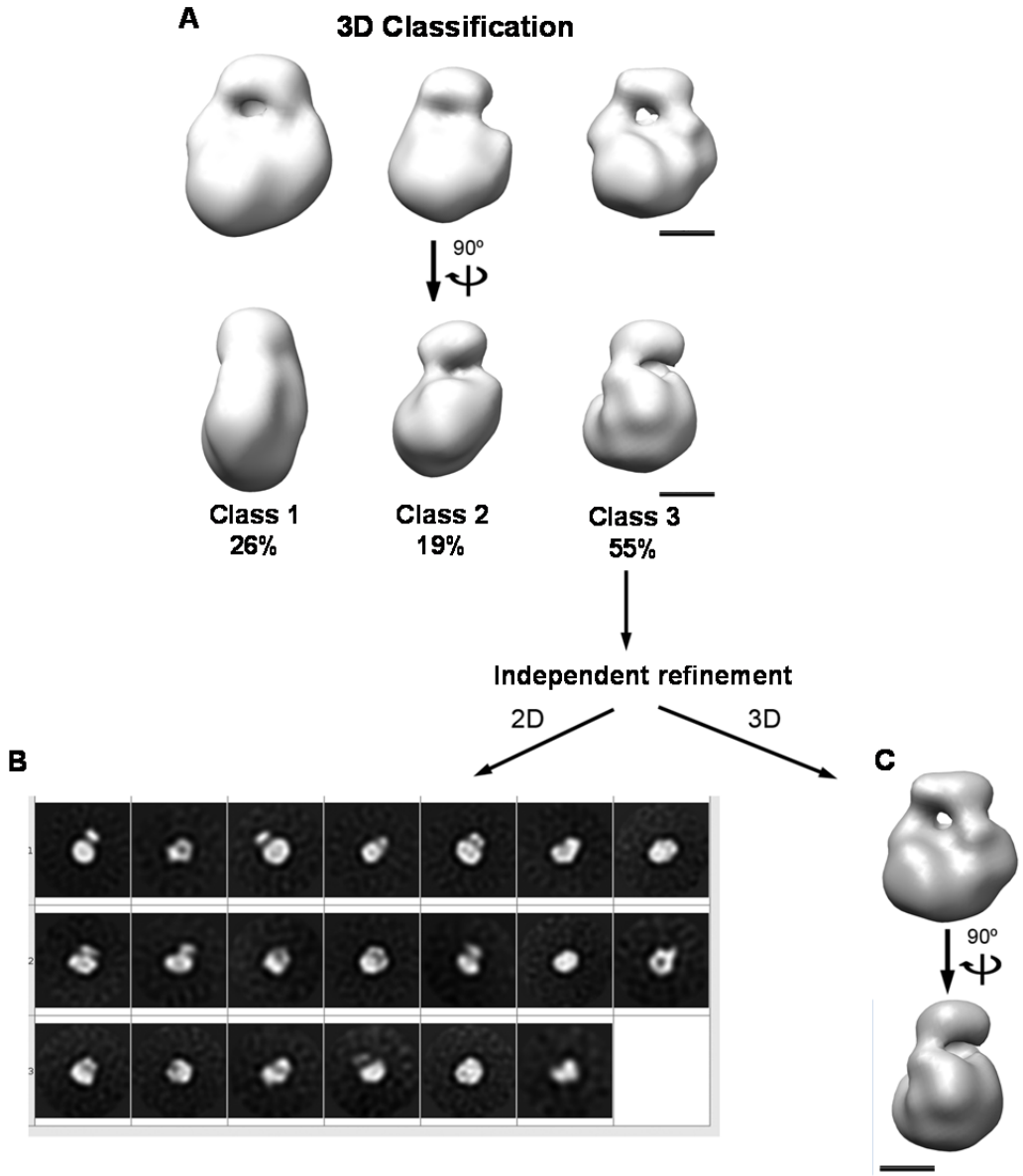

Figure 49. 3D reconstruction of 4F2hc/LAT1 solubilized in amphipol A8-35. An initial 3D reconstruction of the 4F2hc/LAT1 particles solubilized in amphipol A8-35 was obtained from the 2D class averages. A homogenous group of particles were classified in three groups (classes). The percentage of particles in the sample, in each class, is indicated (**A**). The most abundant particles (Class 3) seemed better structurally defined than the other two, and because of this, it was refined in 2D and in 3D. 2D class averages of particles from the class 3 (**B**). 3D map of 4F2hc/LAT1 particles from the class 3 (**C**). Scale bars represent 5 nm.

In fact, the best fitting of human 4F2hc-ED inside the small density of the 3D map of vertebrate 4F2hc/LAT1 (Figure 50A), shows the N-terminal of the 4F2hc-ED very close to the density that

we found, by comparison with the 3D maps of human 4F2hc/LAT2, as being the density close to the N-terminal of 4F2hc-ED, and probably to the disulphide bridge (indicated with an asterisk in Figure 50). The other density connecting both subunits in 4F2hc/LAT1, might correspond with a point of strong interaction between 4F2hc and LAT1, that was not similarly observed in the 3D map of 4F2hc/LAT2 (compare Figure 50A with 50B and 50C). In fact, in the first reported 3D reconstruction of 4F2hc/LAT2 (Rosell *et al.* 2014), docking and cross-linking experiments indicated other potential interactions among both subunits in the opposite site of the 4F2hc-ED N-terminal, and thus of the putative disulphide bond site. Indeed in such study, 4F2hc density resides almost completely above the LAT2 density in the 3D map (Figure 50B) (Rosell *et al.* 2014). Thus, it is possible that this interaction between the heavy and the light subunit, in the opposite site of the potential disulphide bond site, was also present in 4F2hc/LAT2 model. However, as the space between 4F2hc and LAT1, observed here (Figure 50A), is not observed in the 4F2hc/LAT2 reconstruction (Figure 50B and 50C), such interaction density was not observed in the same way as for 4F2hc/LAT1.

On the other hand, a defined cavity is observed in the center of the LAT1 density, towards the interface with 4F2hc (black arrowhead in Figure 50A), which could be related with the extracellular substrate binding site, that is slightly observed in the previous 4F2hc/LAT2 3D maps (Figure 50B and 50C) (Meury *et al.* 2014; Rosell *et al.* 2014). Interestingly, the less abundant and less defined class found in the 4F2hc/LAT1 sample (Class 2 in Figure 49A), showed 3D reconstruction more similar to those of 4F2hc-LAT2 (Figure 50B and 50C) than the most abundant and better defined class 3 (Figure 49C and 50A). The differences between vertebrate 4F2hc/LAT1 and human 4F2hc/LAT2 maps, may be due to the current model shows better definition than the previous 4F2hc-LAT2 maps, since showed more defined and protuberant shapes, despite of being apparently more voluminous than those of the detergent- or detergent/CHS-solubilized human 4F2hc/LAT2 (Figure 50). The more voluminous map could be reflecting the differential content of endogenous lipid, CHS, detergent and/or amphipol that stay bound to the protein particles after the DDM exchange for amphipol A8-35. As this is the first time that the map of a HAT is obtained in an amphipol, we can speculate that the more voluminous form of 4F2hc/LAT1 map could be also due to the amphipol, which would improve definition, but may increase the apparent size of 4F2hc/LAT1 particles, mainly of LAT1, in negative staining, as it is observed here (Figure 50).

On the other hand, 4F2hc seems to be a little disinclined and turned in the 4F2hc/LAT1 model, in comparison with the position found for human 4F2hc in the previous reconstructions, as determined by the 4F2hc-ED structure fitting in all the models (compare the position of the 4F2hc small densities and of the 4F2hc-ED inside them in Figure 50A-C). Thus, it is also probable that vertebrate 4F2hc/LAT1 is adopting a conformation a little different than human 4F2hc/LAT2, causing some of the main differences, described above, among these two heterodimers maps, together or not with the effect of the amphipol presence instead of detergent. For instance, this subtle turned of 4F2hc-ED in relation to LAT1, could have uncover a bit more the extracellular substrate binding site in LAT1, that together with the amphipol addition could have slightly increase the definition on this area (black arrowheads Figure 50). However, due to the preliminary of this reconstruction at low resolution, from a sample with some level of polydispersity, this model could not be quite accurate and thus no asseverations can be made.

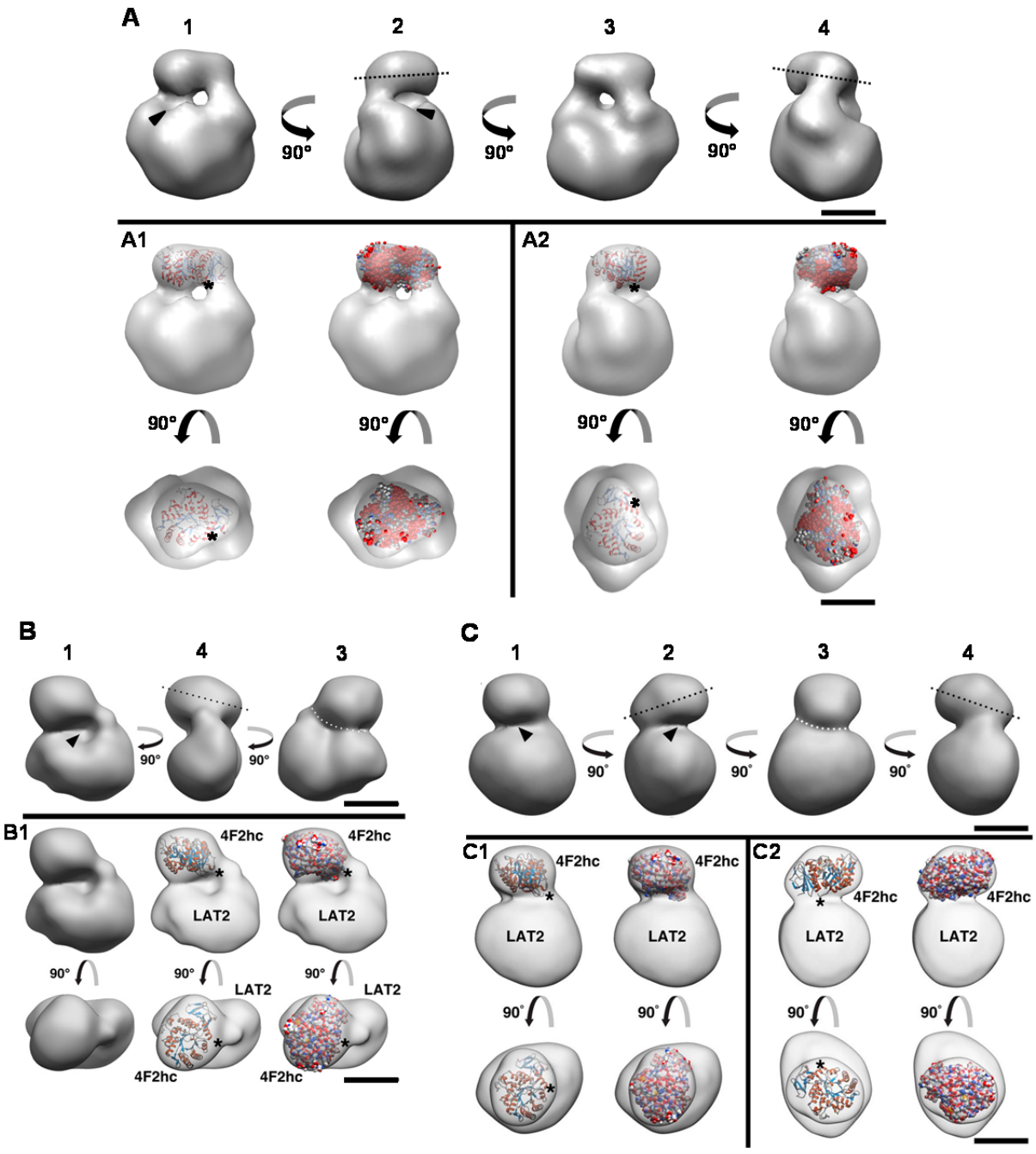

Figure 50. Comparison of the 3D maps of vertebrate 4F2hc/LAT1 and human 4F2hc/LAT2. Vertebrate 4F2hc/LAT1 3D reconstruction (**A**) was made by single particle analysis TEM from protein solubilized in amphipol A8-35. 3D reconstruction of human 4F2hc/LAT2 was previously obtained by single particle analysis TEM from protein purified in DDM (**B**) (Rosell *et al.* 2014) or DDM/CHS/LMNG (**C**) (Meury *et al.* 2014). For enable comparisons, 3D reconstruction of vertebrate 4F2hc/LAT1 (**A**) was orientated and depicted equal to 4F2hc/LAT2 (**B, C**). Four different orientations of 4F2hc/LAT1 are shown, that correspond with the orientations of 4F2hc/LAT2, as indicated by the numbers above each orientation. The structure of human 4F2hc-ED fits well inside the small density of the 4F2hc/LAT1 map, as showed in orientation 1 (**A1**) and orientation 2 (**A2**). Similar to it was shown for 4F2hc/LAT2, in DDM and orientation 1 (**B1**), and in DDM/CHS/LMNG and orientation 1 (**C1**) and orientation 2 (**C2**). Thus, the small density correspond to heavy subunit (4F2hc) and the big to light subunits (human LAT2 in **B** and **C,** and vertebrate LAT1 in **A**). N-terminal of 4F2hc-ED structure is indicated by an asterisk (*), showing comparable orientation of 4F2hc in all maps (**A**, **B** and C), and close proximity of the N-

terminal of 4F2hc-ED to the potential site of the N-terminal TMD of 4F2hc in all maps. In the previous 4F2hc/LAT2 maps, 4F2hc density seemed tilted, as indicated by the black dotted line (**B** and **C**), which lead to a site sealed among both subunits (indicated by the white dotted curve), and at the opposite site to a distinct cavity (black arrowheads **B** and **C**). In the present 4Fhc/LAT1 map, the tilt is slightly observed (black dotted line in **A**), and the density of 4F2hc/LAT1 is not sealed completely among both subunits, leaving a space in the interface of both proteins. Taking into account this slightly different position of the 4F2hc densities (orientation 2 and 4 in **A, B** and **C**) and the fitting of 4F2hc-ED in them (compare top views in **A1** and **A2** with those in **B1**, **C1** and **C2**), an apparent little disinclination and rotation of the vertebrate 4F2hc density in 4F2hc/LAT1 (**A**) is observed in compare with human 4F2hc/LAT2 (**B** and **C**). All this could have revealed a better defined cavity whih could be related with the extracellular substrate binding site (indicated by black arrowheads in **A**). Scale bars represent 5 nm and all maps are at the same scale.

3.5 Perspectives

Concerning to the polydispersity observed in the LAT1 particle, optimization of the exchange of DDM by amphipol must be tried. The ratio amphipol/protein used until now (3:1) will be increased to enhance the binding of amphipol to the TMDs of the protein, displacing at the same time the DDM-protein interactions. In fact, in a recent study of TRPV2 channel (Zubcevic *et al.* 2016), a higher concentration of amphipol A8-35 (protein/amphipol; 1:10; w/w) is used for solving the high resolution structure of this channel. Interestingly, Zubcevic and co-workers used similar purification conditions to those used by us for purifying 4F2hc/LAT1, since they also add CHS during the purification of TRPV2 channel, besides DDM and lipids, before exchanging the detergent for amphipol A8-35 (Zubcevic *et al.* 2016). Thus, it could be possible that the presence of CHS requires at least 3-fold higher amount of amphipol for an efficient protein solubilization in amphipol. This could be supported by the fact that in the structures of γ-secretase (Lu *et al.* 2014; Bai *et al.* 2015) and TRPV1 channel (Liao *et al.* 2013) in amphipol A8-35, the protein/amphipol ratio used is the same used by us (1:3; w/w), and before amphipol solubilization, protein is purified in detergent (CHAPSO and digitonin, and DDM, respectively) plus lipids, but in absence of CHS. This would correlate with our hypothesis about that the protein solubilization in amphipol could not be complete by the presence of CHS, which could be also bound in different amounts among 4F2hc/LAT1 molecules by the incomplete DDM/amphipol exchange, and besides, could avoid homogeneous delipidation of the protein. In this sense, by increasing the concentration of amphipol in further experiments, we would expect to improve the homogeneity of the preparation by the enrichment of 4F2hc/LAT1 molecules solubilized in amphipol A8-35.

On the other hand, it is known that a more homogeneous sample of a membrane protein can be obtained by delipidating as much as possible the protein, ideally until only lipids require for the stability and function remain bound to the protein. Reaching the equilibrium among enough delipidation of the protein and the preservation of relevant lipids is challenging and requires empirical optimization. It has been described that the amount of endogenous phospholipid that remains bound to a membrane protein after purification, strongly varies with the detergent type and concentration (Ilgu *et al.* 2014). Thus, a more homogeneous sample of 4F2hc/LAT1 could be obtained by using other detergents. Here we showed good stability of LAT1 in DM, Cymal-6 or C_{12}MNG in presence of CHS (Figure 26), suggesting that 4F2hc/LAT1 could be also stable in such detergents. Therefore, they should be tested for purifying 4F2hc/LAT1 before detergent/amphipol exchange, looking to decrease heterogeneity. It is also probable that other detergent than DDM could be more efficiently exchanged by amphipol A8-35, increasing also protein homogeneity.

Moreover, as mentioned through this thesis sometimes lipids are added during the purification of membrane proteins for increasing stability. In these cases, it is probable that during solubilization, the protein was stripped from endogenous lipids relevant for their stability, thus needing the addition of lipids to be stable. Here, we tested the stability of 4F2hc/LAT1 in DDM/CHS and several lipidic conditions, however, none of the lipidic conditions increased the stability of 4F2hc/LAT1, in compare with DDM/CHS. Nevertheless, some of these lipids, or other, could be tested in combination with other detergent/CHS conditions (e.g., different concentration of DDM and/or CHS, or different detergent), looking to increase protein stability, if it is reduced by the new purification condition used, but mainly to increase amphipol/detergent exchange and sample momodispersity. For instance, a condition in which the detergent causes homogeneous delipidation, and CHS is required at lower concentrations, because of the stabilization effect of added lipids, could solve some of the posible issues related with the presence of CHS. These detergent/CHS/lipids conditions would be similar to the purification condition of TRPV2 channel used for solving its structure in amphipol A8-35 by cryo-EM (Zubcevic *et al.* 2016).

Additionally, an alternative strategy could be the production of 4F2hc/LAT1 in expression systems having endogenous cholesterol, like insect or mammalian cells. In this way, it is possible that the cell could provide the protein with the essential cholesterol, without being required the addition of CHS. It is worth to remember that insect (Sf9) and mammalian cells were able to produce 4F2hc-GFP/LAT1 heterodimer. In addition, the peak of 4F2hc-GFP/LAT1 obtained in Sf9 cells showed monodispersity in FSEC, and despite of being expressed at low yield, it is probable that this production is good enough for cryo-EM studies. However, non-GFP tagged 4F2hc/LAT1 must be generated in both expression systems, and its stability evaluated under the required concentrations for cryo-EM. In this new scenario, the amphipol/protein ratio would have to be optimized also.

Other way that could be useful for improving 4F2hc/LAT1 monodispersity, is the use of Styrene-Maleic Acid co-polymers (SMAs) (Dörr et al. 2016). SMA particles are amphipathic, and capable of inserting into biological membranes, due to alternated disposition of hydrophobic styrene and hydrophilic maleic acid in such particles. These molecules are able to solubilize lipid bilayers in the form of nanodiscs without the addition of detergents. Thus, by using this approach, the hydrophobic regions of recombinant membrane proteins solubilized in SMA, will be surrounded by its associated phospholipids, allowing the purification of membrane proteins in a native-like environment (Figure 51). SMA co-polymers have been demonstrated to be efficient in the solubilization of recombinant membrane proteins expressed in bacteria, yeast, insect cells and human cells (Long *et al.* 2013; Dorr *et al.* 2014; Gulati *et al.* 2014; Paulin *et al.* 2014; Jamshad *et al.* 2015; Prabudiansyah *et al.* 2015; Skaar *et al.* 2015).

Similar to amphipols, the absence of detergents in SMA-solubilized protein samples, could increase the quality of cryo-EM data by avoiding the lower contrast between particle and ice that the detergent causes. In fact, low-resolution 3D structures of SMA-solubilized proteins, like AcrB transporter or the ABC transporter Pgp, have already been solved by cryo-EM (Gulati *et al.* 2014; Postis *et al.* 2015). The application of this strategy to our project would imply the use of the same condition during the full process, 4F2hc/LAT1 would be solubilized and purified in presence of SMA, without the delipidating and destabilizing effects of detergents and without adding CHS. Thus, the heterogeneity probably related with the different amount of CHS and lipids bound to the protein, and with the DDM/amphipol exchange, would be eliminated. The

first attempt of purifying 4F2hc/LAT1 in SMA is being carried out in collaboration with Dr. Vincent Postis (University of Leeds, UK), with protein expressed in *Pichia pastoris*, and we are waiting for the first results about the 4F2hc/LAT1 stability in the presence of these molecules. However, if the absence of cholesterol in yeast membranes compromise the stability of SMA-solubilized 4F2hc/LAT1, this approach would be applied in protein produced in insect and mammalian cells. The use of SMA would help to keep relevant interactions between the endogenous cholesterol, present in these expression systems, and the purified protein. This fact could be determinant to maintain stable 4F2hc/LAT1 from the solubilization step to the cryo-EM sample.

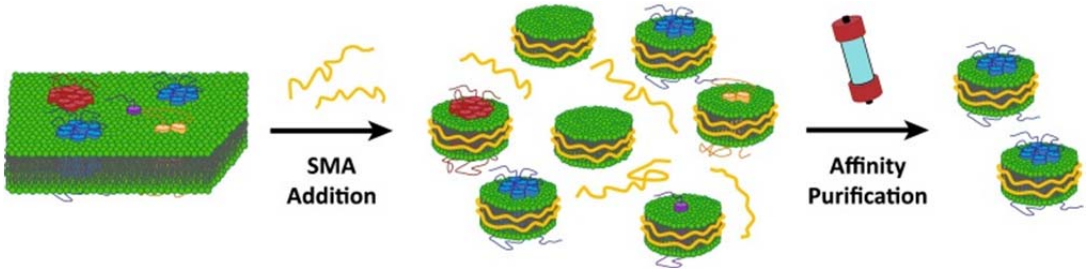

Figure 51. Extraction of membrane proteins in a native lipid environment by SMA. SMA additions leads to the formation of native nanodiscs containing different MPs or only lipid material. Subsequent affinity purification allows for the isolation of native nanodiscs with the protein of interest. (Dorr *et al.* 2016).

The previous described strategies to improve the homogeneity of 4F2hc/LAT1 samples, for facilitating the classification and analysis of heterodimer particles by cryo-EM, will be conducted. However, of not being possible, software and algorithms that allows reliable particle classification on heterogeneous samples, like the maximum likelihood based 3D classification algorithm implemented in RELION (Scheres 2012), will be used. In fact, in the TRPV1 channel structural determination, RELION was used to discard particle heterogeneity, being solved the structure at 3.4 Å resolution from particles representing only 20 % of the total particle selected (Liao *et al.* 2013; Liao *et al.* 2014). Making an approximation by the observed in our preliminary 3D reconstruction, promising results could be expected in the future with 4F2hc/LAT1, as the better structurally defined class was also the most populated in the sample (55 %; Figure 49).

Taken together, the results showed through the present study suggest high potential of vertebrate 4F2hc/LAT1 for solving the first atomic structure of a Heteromeric Amino acid Transporter (HAT) by cryo-EM. 4F2hc/LAT1 shares several characteristics with other metazoan membrane proteins, whose structures have been solved at atomic resolution by cryo-EM (Liao *et al.* 2013; Bai *et al.* 2015; Du *et al.* 2015). Similar to those proteins, 4F2hc/LAT1 is composed of more than one subunit, and it shows a comparable size, 120 kDa for the heterodimer, and from 170 kDa to 300 kDa for γ-secretase (Lu *et al.* 2014; Bai *et al.* 2015), glycine receptor (Du *et al.* 2015), TRPV1(Liao *et al.* 2013) and TRPV2 (Zubcevic *et al.* 2016) channels. Importantly, 4F2hc/LAT1 is stable, at protein concentrations compatible with cryo-EM, in amphipol A8-35, the polymer used to solve the structure of all the proteins listed before except the glycine receptor.

On the other hand, γ-secretase, with a predicted size of 170 kDa, could not be well-visualized in hydrated-freeze samples by electron microscopes a few years ago. However, taking advantage of the technical improvements in microscopes and detectors, last year its structure was solved at 3.4 Å resolution. Vertebrate 4F2hc/LAT1, despite of being smaller than γ-secretase, may be a good target to apply these new advances to speed up the process to obtain its structure at high

resolution in the next years. In fact, 20 years ago, Richard Henderson indicated that, under ideal sample and imaging conditions, it is possible to achieve 3D reconstructions of molecules of around 100 kDa at 3 Å resolution (Henderson 1995). Thus, some of these ideal conditions are being developed, like powerful electron microscopes and direct detectors.

The structure of human 4F2hc/LAT2 at 20 Å resolution and related studies, revealed that 4F2hc-ED may interact with several extracellular loops of LAT2 (Rosell *et al.*, 2014). The precise characterization of these interactions by a high-resolution structure of a vertebrate HAT, together with functional studies and single-molecule FRET, would allow to know the impact of the heavy subunit on light subunit conformational changes associated to substrate translocation during transport cycle. Thus, determination of the atomic structure of 4F2hc/LAT1, will be very valuable to the better understanding of the molecular mechanism of HATs amino acid transport through cellular membranes, as well as, to increase the knowledge about the molecular bases of inherited aminoacidurias and tumor growth.

4. Conclusions

Conclusions

1. Among the heterodimers formed by vertebrate LAT1, LAT2 or $b^{0,+}$AT, light subunits previously determined as the most stable, 4F2hc/LAT1 and 4F2hc/LAT2, but not rBAT/$b^{0,+}$AT, could be expressed in mammalian and insect cells. Vertebrate 4F2hc/LAT1, expressed in Sf9 cells, was more thermostable than 4F2hc/LAT2, being a better candidate for structural studies.

2. The methylotrophic yeast *Pichia pastoris* improved 2-6 fold the expression yield of vertebrate 4F2hc/LAT1 in compare with insect cells. Pichia was selected to produce 4F2hc/LAT1 for structural studies.

3. Our hypothesis about that finding a quite stable light subunit of HAT, together with the described stabilizing effect of 4F2hc, would improve the stability of the 4F2hc-containing heterodimer, was corroborated. Vertebrate LAT1 presented higher stability than human LAT2 expressed in Pichia, and vertebrate 4F2hc increased the stability of vertebrate LAT1. Then, as expected, purified vertebrate 4F2hc/LAT1 was more stable than the previously studied human 4F2hc/LAT2 expressed in Pichia, being a better candidate for structural studies.

4. A version of vertebrate 4F2hc with a 36 residues delition at its N-terminal, and with mutations to glutamine in all its asparagines predicted to be glycosylated, was expressed in Pichia and formed heterodimer with wild type vertebrate LAT1 (NG-4F2hc(Δ36Nter)/LAT1). Although, this heterodimer version could be suitable for 3D crystalography, because would have less flexibility and homogeneous de-glycosilation, the low stability and yield of this heterodimer produced in Pichia, hinder crystallographic studies with this vertebrate 4F2hc/LAT1version.

5. None of the lipid conditions tested improved the stability of vertebrate 4F2hc/LAT1 solubilized in n-Dodecyl β-D-maltoside (DDM) plus cholesteryl hemisuccinate (CHS), as determined by FSEC.

6. Scintillation proximity assays showed a robust substrate binding (e.g., L-isoleucine) to vertebrate LAT1. Interestingly, this binding was blocked by KYT-0353, a human LAT1 inhibitor. This result opens the possibility to use this molecule to generate 4F2hc/LAT1-Inhibitor complexes, diminishing the conformational heterogeneity of the sample and the intramolecular flexibility of the heterodimer.

7. Vertebrate 4F2hc/LAT1 solubilized in amphipol A8-35 is stable. This polymer has been used successfully to solve the structure of γ-secretase, TRPV1 and TRPV2 channels by high-resolution cryo-EM. This fact would facilitate the resolution of 4F2hc/LAT1 structure by this technique.

8. 3D reconstruction from negative stained 4F2hc/LAT1, solubilized in amphipol A8-35, showed a bilobular particle comparable with the reported low-resolution models of human 4F2hc/LAT2. The fact that 4F2hc/LAT1 preparations contain only

heterodimers and no single subunits, in contrast to human 4F2hc/LAT2 samples, facilitated classification of the heterodimer particles in subclasses. This made vertebrate 4F2hc/LAT1 a suitable candidate for cryo-EM studies, although the size heterogeneity corresponding to the light subunit density has to be improved.

5. Materials and Methods

5.1 Expression screenings in the tsA201 human cell line of GFP-fused rBAT/$b^{0,+}$AT, 4F2hc/LAT1 and 4F2hc/LAT2 vertebrate heterodimers

5.1.1 Generation of the expression constructs

The cDNAs of the rBAT and 4F2hc heavy subunits, and of the $b^{0,+}$AT, LAT1 and LAT2 light subunits were cloned into pEG BacMam expression vector (Figure 52) by Dr. Albert Rosell in our laboratory.

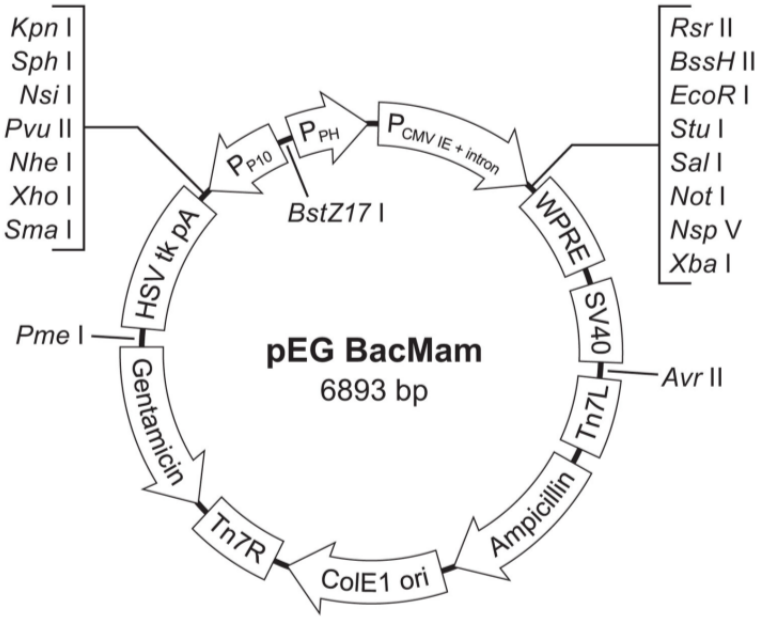

Figure 52. Map of the pEG BacMam expression vector. The cDNAs of interest were cloned into the multiple cloning site behind the CMV promoter using unique restriction sites. Elements that are important for high level expression in mammalian cells are shown, including those that are important for transcription initiation (CMV promoter), transcription termination (SV40 poly A late signal) and mRNA processing (intron and WPRE motif). (Goehring *et al.* 2014).

Heavy and light subunits were cloned fused to GFP at the N-terminus or at the C-terminus (Figure 53A and 53B), to test the impact of the GFP position in the expression and integrity of each subunit, and later in the formation of the heterodimer. GFP was tagged with a poly-histidine tag (His8) when was fused to a heavy subunit, or with a Strep-tag II (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) when was added to a light subunit. In all cases, the sequence Leu-Glu-Val-Leu-Phe-Gln-Gly-Pro, target for the human rhinovirus 3C protease, was introduced between GFP and the corresponding subunit (Figure 53A and 53B). In order to check the generation of the heterodimers, these GFP-fusion proteins would be intended to co-expression with the corresponding subunit tagged only with 3C-His8 (heavy subunits) or with 3C-Strep-tag II (light subunits) at the C-terminal (Figure 53C).

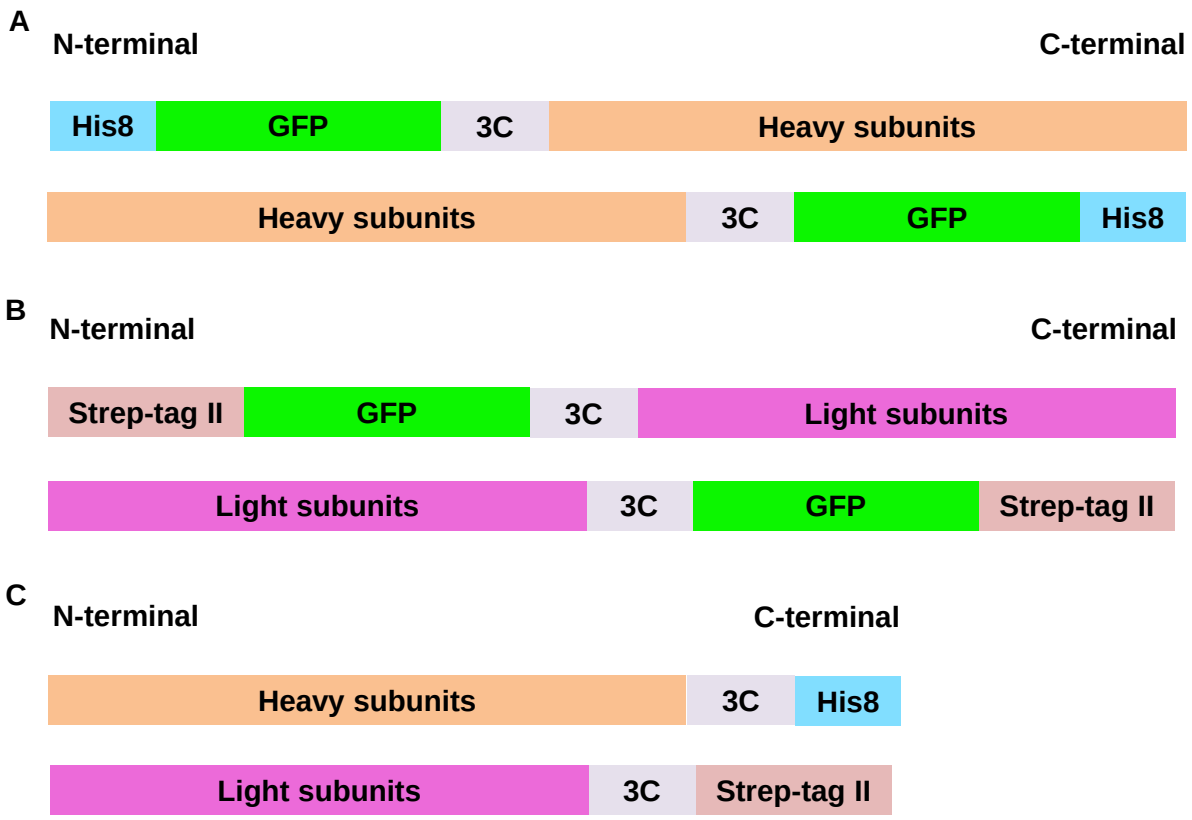

Figure 53. Representation of the constructs expressed in tsA201 cells. GFP was fused to the N- and C- terminals of each protein with a His8 (heavy subunits) and a Strep-tag II (light subunits) as indicated in **A** and **B**, respectively. (**C**) Heavy and light subunits intended for the co-expression with their corresponding GFP-tagged partners were cloned with a His8 tag (heavy subunits) or with a Strep-tag II tag (light subuinits) at the C-terminus. The position of the target sequence for human rhinovirus 3C protease between the protein and the GFP, or protein and the tags is indicated (3C).

5.1.2 Protein expression by small-scale transfection and co-transfection

In a six-well dish, $1x10^6$ tsA201 cells per well were seeded in 2ml of DMEM medium (4.5 g/l glucose, L-glutamine, and sodium pyruvate; Gibco) supplemented with fetus bovine serum (10 % v/v). After 24 hours (h) of incubation at 37°C and 5 % CO_2, transfection was carried out as described: Per well, 4µl of Lipofectamine 2000 (Life Technologies) were mixed with 50 µl of Opti-MEM I reduced serum medium (Life Technologies) (MixA), and incubated by 5 minutes (min) at room temperature. Separately, 1.5 µg of one single DNA construct, or 0.75 µg *plus* 0.75 µg, when co-transfecting, were mixed with 50 µl of Opti-MEM I (MixB). MixA and MixB were mixed and incubated for 20 min at room temperature. During this incubation, and taking care not to unattach nor perturb the cells, media was eliminated and 2ml of fresh DMEM was added. Then, transfection or co-transfection mixture was dropwise added. After 24 h of incubation at 37°C and 5 % CO_2, the DMEM medium was replaced with fresh DMEM supplemented with 5 mM sodium butyrate to boost protein expression. 48 h post-transfection, DMEM was aspirated off, adherent cells gently rinse with 2 ml PBS and collected with 1 ml PBS into 1.5 ml centrifuge tubes.

5.1.3 Protein solubilization and Fluorescence Size Exclusion Chromatography (FSEC) analysis

The cell suspensions obtained from small-scale transfections and co-transfections, were centrifuged at 1,500xg for 5 min at 4°C. After supernatant removal, cells were resuspended in 200 µl of solubilization buffer containing, 50 mM Tris pH 8.0, 150 mM NaCl, 1 mM EDTA, a protease inhibitor cocktail (1 mM PMSF, 200 µM aprotinin, 2 µg/ml leupeptine and 2 µM pepstatin A), 1 % (w/v) of the desired detergent, and 0.2 % (w/v) of Cholesteryl Hemisuccinate tris salt (CHS). For sustaining CHS solubilization in the buffer, the detergent was first dissolved, and after CHS addition, the buffer was sonicated at 4°C during 6 cycles of 15 seconds (sec) using 80 % of amplitude and 0.7 of intensity (Sonicator KIKA Labortechnik U2005 Control™).

Protein solubilization was performed by rotation during 1 h at 4°C. The solubilized sample was centrifuged at 70,000xg in a TL100 ultracentrifuge for 40 min at 4°C, and the supernatant collected. 100 µl of sample were injected into an ÄKTA purifier (GE healthcare life sciences) for FSEC analysis (Kawate and Gouaux 2006) in a Superose 6, 10/300 GL column (GE healthcare life sciences). FSEC buffer containing, 50 mM Tris pH 8.0, 150 mM NaCl and 0.05% n-Dodecyl-β-D-Maltoside (DDM) was prepared just before use and filtered by a 0.2 µm pore filter.

5.2 GFP-based expression screening of rBAT/b^{0+}AT, 4F2hc/LAT1 and 4F2hc/LAT2 vertebrate heterodimers in Sf9 cells

5.2.1 Generation of the expression constructs

The cDNAs of rBAT and 4F2hc heavy subunits, and $b^{0,+}$AT, LAT1 and LAT2 lights subunits, were cloned into pFastBac1 vectors (Figure 54), where previously Dr. Rosell has cloned GFP-His8 or GFP-Strep-tag II. These pFastBac1-GFP vectors were named as $pNGFP_{HIS}$ and $pNGPP_{STREPTAG}$, to produce N-terminal GFP-fusion proteins, or $pCGFP_{HIS}$ and $pCGFP_{STREPTAG}$, to express C-terminal GFP-fusion proteins. In this way, similarly to the strategy it was done in tsA201 cells, all subunits were cloned fused to GFP at the N-terminus and at the C-terminus (Figure 55) to test the impact of GFP position in their expression, integrity and stability, and later in the formation of heterodimer. GFP was tagged with a His8 when it was fused to a heavy subunit, or with a Strep-tag II when it was placed to a light subunit (Figure 55A and 55B, respectively). In whichever position of the GFP at N- or C-terminal ends, a target sequence for thrombin protease (Leu-Val-Pro-Arg-Gly-Ser) was added between the protein and the GFP (Figure 55). In order to check the generation of the heterodimers, the GFP-fusion proteins that were successfully expressed, were then co-expressed with the corresponding subunit tagged only with Trombin-His8 (heavy subunits), or with Thrombin-Strep-tag II (light subunits) at the C-terminal (Figure 55C). These last constructs were generated cloning the corresponding subunits directly into pFastBac1 vector (Figure 54).

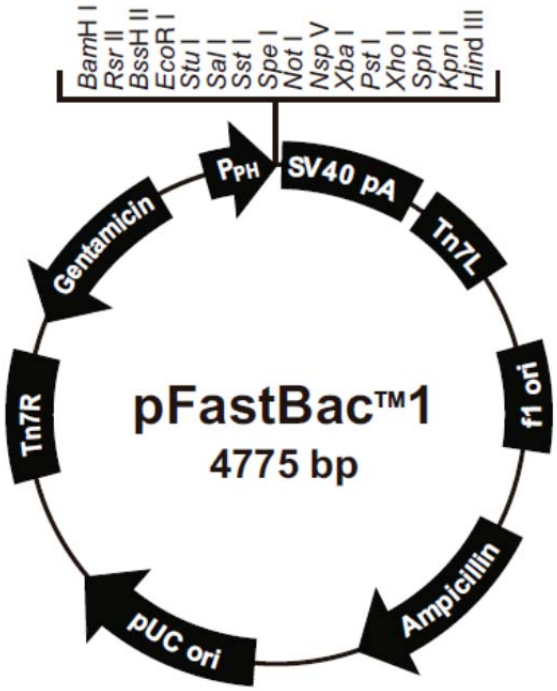

Figure 54. Map of the pFastBac1 expression vector. The main elements are indicated: origin: bases 2-457; ampicillin resistance gene: bases 589-1449; pUC origin: bases 1594-2267; Tn7R: bases 2511-2735; gentamicine resistance gene (complementary strand): bases 2802-3335; polyhedrin promoter (P_{PH}): bases 3904-4032; multiple cloning site: bases 4037-4142; SV40 polyadenilation signal: bases 4160-4400 and Tn7L: 4429-4594. Figure and information from Bac-to-Bac® Baculovirus Expression System user guide (www.lifetechnologies.com). Additional information about this vector and its sequence is available in the former link.

Cloning of the HATs into pNGFP$_{HIS}$, pNGPP$_{STREPTAG}$, pCGFP$_{HIS}$, pCGFP$_{STREPTAG}$ and pFastBac1 vectors.

The cDNAs encoding rBAT and 4F2hc were cloned into pNGFP$_{HIS}$ vector into *Xho*I site and between *Bam*HI and *Xho*I sites, respectively. These same cDNAS, were cloned into pCGFP$_{HIS}$ plasmid into *Xho*I site (rBAT) and between *Eco*RI and *Not*I sites in the case of 4F2hc. Regarding the cDNAS encoding the light subunits, LAT1 was cloned into *Hin*dIII site into both vectors (pNGFP$_{STREPTAG}$ and pCGFP$_{STREPTAG}$), LAT2 into *Hin*dIII site and between *Eco*RI and *Hin*dIII sites into pNGFP$_{STREPTAG}$ and pCGFP$_{STREPTAG}$ plasmids, respectively, and finally, $b^{0,+}$AT was cloned between *Bam*HI and *Hin*dIII sites (pNGFP$_{STREPTAG}$) and between *Eco*RI and *Xho*I sites (pCGFP$_{STREPTAG}$). In all cases, the cDNAs were cloned in frame with the GFP cDNAs cloned previously. Regarding the generation of the constructs without GFP, rBAT and 4F2hc cDNAS were cloned into pFastBac1 vector into *Xho*I site and between *Eco*RI and *Not*I sites, respectively. Finally, $b^{0,+}$AT, LAT1 and LAT2 were cloned between *Eco*RI and *Xho*I sites, into *Hin*dIII site and between *Eco*RI and *Hin*dIII sites, respectively.

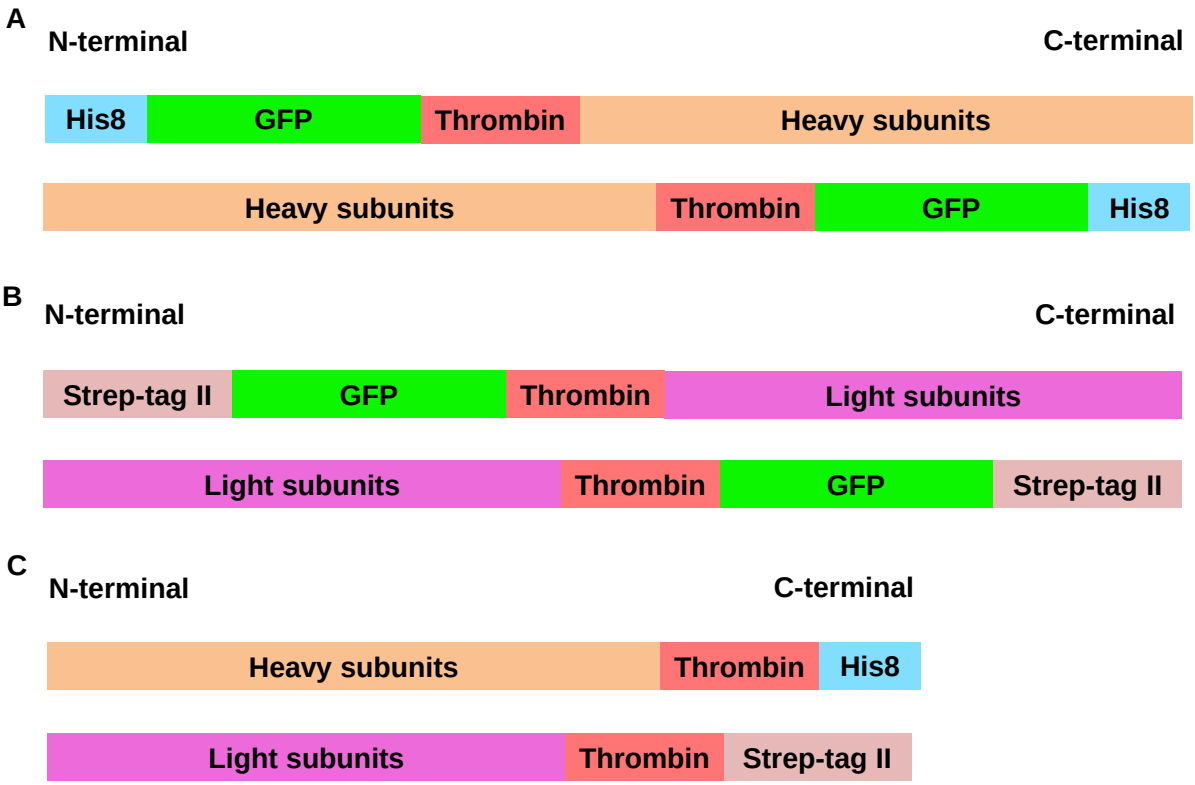

Figure 55. Representation of the constructs expressed in Sf9 insect cells. GFP was fused to the N- and C- terminals of each protein with a His8 (heavy subunits) and Strep-tag II (light subunits) as indicated in **A** and **B**, respectively. GFP tagged versions of heavy and light subunits that could be expressed, were co-expressed with their corresponding partner subunits cloned with a His8 tag (heavy subunits) or a Strep-tag II (light subunits) at the C-terminus (**C**). The position of the target sequence for thrombin protease between the protein and the GFP, or the protein and the tag, is indicated.

HATs cDNAs were amplified by PCR (Polymerase Chain Reaction) using the proofreading enzyme Expand High Fidelity Taq polymerase (Roche), and following the conditions showed in Table 7.

Reaction mix insert amplification	(µl)
Template plasmid (20 ng)	X
Reaction buffer with MgCl2 (10X)	5
dNTP mix Forward primer (10 µM)	1
Reverse primer (10 µM)	2
Expand High Fidelity enzyme	2
Milli-Q water (adjust to final volume)	X
Final volume	50

PCR insert amplification	T°	Time (min)	Cycles
Initial denaturation	95°C	5	1x
Denaturation	95°C	1	
Anneling	54°C	1	30x
Elongation	72°C	2	
Final Elongation	72°C	7	1x
Cooling	4°C	∞	

Table 7. High fidelity PCR reaction standard conditions

The amplified cDNAs were purified from agarose gels (1 % w/v) (Seakem LE). These gels were prepared with TAE buffer (40 mM Tris base, 20 mM acetic acid, 1 mM EDTA) and samples run at 80-100 V for 1 h. SyberSafe (Invitrogen) at a 1/10,000 dilution was used as intercalating DNA agent to visualize the DNA, and 1 kb ladder from Fermentas was used to reference our samples. Gels were visualized with the GeneGenius Gel Imaging system (Syngene). The PCR product was purified from agarose gel using the Gel Band Purification kit (GE Healthcare).

Amplified and purified cDNAS were ligated into pGEM-T Easy vector (Promega) according to the manufacturer's protocol. pGEM-T/HATs ligations were transformed into XL1 Blue *E.coli* cells by the heat shock method: Competent cells were thawed on ice, 80 µl of cells were mixed with the ligation mixture and kept on ice during 20 min. Heat shock was done at 42°C for 90 sec, and kept on ice for 2 min. Then, 900 µl of LB medium (1 % tryptone, 0.5 % yeast extract, 0.5 % NaCl, pH 7.5) was added and incubated for 1 h at 37°C shaking at 900 rpm. The 900 µl of cell culture were centrifuged for 3 min at 5,000 rpm, discarded 800 µl, and resuspended the pellet in the remaining 100 µl of medium, and finally, plated into LB medium plates containing ampicillin (50 µg/ml).

pGEMt/HATs constructs were digested with the corresponding restriction enzymes (New England Biolabs), and the released DNA fragments were purified from agarose gels by the Gel Band Purification kit (GE Healthcare). $pNGFP_{HIS}$, $pNGPP_{STREPTAG}$, $pCGFP_{HIS}$, $pCGFP_{STREPTAG}$ and pFastBac1 plasmids were processed with the corresponding endonucleases, and, in order to prevent self-annealing of linearized vector, were treated for 1 h at 37°C with calf intestinal alkaline phosphatase (CIP) (New England Biolabs). Thereafter, the linearized and dephosphorylated plasmids were purified from agarose gels by the Gel Band Purification kit (GE Healthcare). Ligation between the linearized vectors and the corresponding HATs inserts was performed for 16 h at 37°C using the T4 DNA ligase (New England Biolabs), and following the manufacturer's recommendations. The products from the different ligation reactions were transformed into XL1 Blue *E.coli* cells by the heat shock method. After the screening of positive clones, the protein constructs generated into $pNGFP_{HIS}$, $pNGPP_{STREPTAG}$, $pCGFP_{HIS}$, $pCGFP_{STREPTAG}$ and pFastBac1 were 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).

5.2.2 Production of recombinant baculovirus by the Bac-to-Bac expression system

5.2.2.1 Bacmid generation in DH10Bac *E. coli* strain

Transformation of DH10Bac *E. coli* cells

For each cDNA cloned into $pNGFP_{HIS}$, $pNGPP_{STREPTAG}$, $pCGFP_{HIS}$, $pCGFP_{STREPTAG}$ and pFastBacI vectors, 20 ng of plasmid were added to 30 µl of DH10Bac cells, and incubated on ice for 30 min. Cells were subjected to heat shock for 45 sec at 42°C, and immediately, 200 µl of S.O.C medium (2 % tryptone, 0.5 % yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 20 mM glucose) were added. Vials were incubated at 37°C during 4 h at 225 rpm. After this time, 100 µl of cells were plated on LB agar plates containing 50 µg/ml kanamycin, 7 µg/ml gentamicin, 10 µg/ml tetracycline, 100 µg/ml Bluo-gal and 40 µg/ml IPTG. Plates were incubated at 37°C during 48 h, and then, 3-6 white colonies were picked up and inoculated into

6 ml of LB medium containing 50μg/ml kanamycin, 7μg/ml gentamicin and 10μg/ml tetracycline. Cell cultures were incubated for 16 h at 37°C and 225 rpm. For storage at -80°C, 850 μl of cell suspension were mixed with 150 μl of sterilized glycerol, and put into –80°C freezer, the remaining part of the culture was used to isolate the recombinant bacmids.

Isolation of the recombinant bacmids

The DH10Bac cell suspensions were centrifuged at 1,500xg for 10 min at room temperature. The cell pellet was resuspended in 200 μl of P1 solution (Qiagen kit) and transferred to a 1.5 ml centrifuge tube. 200 μl of P2 solution (Qiagen kit) were added and mixed carefully by inverting the vial 6 times, then, after 5 min, 300 μl of neutralization buffer (N3; Qiagen kit) were added and mixed by inversion 5 times. The mixture was centrifuged at 1,500xg for 10 min at room temperature, and the supernatant was transferred into 2 ml centrifuge tubes. 1ml of isopropanol was added to the vial and mixed by inverting the tube carefully, and then, the tube was placed into –20°C freezer during 20 min. Sample was spin at 1,500xg for 15 min at room temperature, and the supernatant discarded. 1 ml of 70 % ethanol was used for washing the pellet by gently inverting the tube and then centrifuging at 1,500xg for 15 min at room temperature. After ethanol removal, the pellet was dried by leaving the tube open by 15 min, and finally, resuspended in 50 μl of autoclaved MilliQ water. Concentration of bacmid DNA was determined by using a NanoDrop spectrophotometer (Thermo Scientific). Bacmids were diluted to 1 mg/ml for being used, and then stored at 4°C until being needed again.

5.2.2.2 Transfection of Sf9 cells with recombinant bacmids to produce P1 virus

For generating population 1 (P1) of viruses, six-well plate were used, corresponding each well to one transformation with one recombinant bacmid. 9×10^5 cells per well were seeded in 2 ml of 27°C preheated serum-free Sf-900™III media (Life Technologies). The cells were incubated at 27°C until they attach (about 30 min). Meanwhile, 8 μl of Cellfectin II (Life Technologies) were added to 100 μl of Sf-900™III media in a centrifuge tube for each transfection. In a different centrifuge tube, 1 μg of bacmid DNA was added to 100 μl of Sf-900™III media. Both mixtures were mixed together and incubated for 30 min at room temperature.

Taking care of not to disturb or unattach the cells, media was removed, and 2 ml of fresh 27°C preheated Sf-900™III were added per well. Mixture containing bacmid DNA was added to the well. The six-well plates were incubated at 27°C for 72 h in a non-humidified and non-CO_2 regulated atmosphere, making sure to have water inside the incubator to avoid strong media evaporation. After this time, media containing P1 virus was collected and filtered into a sterile 2 ml centrifuge tube by using small 0.2 μm filters. When P1 viruses were not immediately used, they were stored light protected at 4°C until use. The transfection efficiency of bacmids providing from $pNGFP_{HIS}$, $pNGPP_{STREPTAG}$, $pCGFP_{HIS}$ and $pCGFP_{STREPTAG}$ constructs was established by checking the transfected cells in a fluorescence microscope. Those P1 viruses that produced green-fluorescent cells were tittered and used for generating P2 of viruses. The transfection efficiency of bacmids providing from pFastBacI constructs was not possible to be determined by this method. Therefore, in this last case, the P1 viruses were always directly tittered.

5.2.2.3 Infection of Sf9 cells with P1 virus to produce P2 virus

Based on the desired volume of P2 virus, P1 viruses were added to a Multiplicity Of Infection (MOI) of 0.01 to Sf9 suspension cell cultures at a cell density of 1.5×10^6 cells/ml. The Sf9 cells infected with the P1 virus, were incubated at 27°C and 120 rpm for 96 h in a non-humidified and non-CO_2 regulated atmosphere orbital shaker. Cells were centrifuged for 15 min at 8,000xg and 4°C. The supernatants containing P2 viruses were collected and filtered using sterile 0.2 µm filters. P2 viruses were tittered and stored light protected at 4°C.

P2 virus production for small-scale expression in attached Sf9 insect cells

6×10^6 Sf9 cells were seeded in 10 cm plates in 10 ml of 27°C preheated Sf-900™III media. The cells were incubated at 27°C until they attach (about 30 min). P1 viruses were added to a MOI of 0.01. Incubation was performed at 27°C in a non-humidified and non CO_2 regulated incubator during 96 h. After this time, supernatant containing P2 viruses was filtered using sterile 0.2 µm filters, and P2 virus stored light protected at 4°C. This strategy was used to test the expression of truncated versions of LAT1 (see section 5.2.7.2).

5.2.2.4 Determination of the virus titers by using the end point dilution assay

Under sterile conditions, 100 µl of Sf9 cells (to titer baculovirus from $pNGFP_{HIS}$, $pNGPP_{STREPTAG}$, $pCGFP_{HIS}$ and $pCGFP_{STREPTAG}$ constructs) or 100 µl of Sf9 Easy Titer cells (to titer baculovirus from pFastBac1 constructs) (Hopkins and Esposito 2009), both at $0.75x10^6$ cells/ml, were transferred to 96-well black plates (75,000 cells/well) for cellular attachment. Plates were left at 27°C in a non-humidified and non-CO_2 regulated incubator for 15 min. Serial dilutions of P1 viruses in Sf-900™III media were carried out by using a sterile deep-well 8-strip clusters. Dilutions of viruses from 10^{-1} to 10^{-8} were obtained by mixing 360 µl of medium with 40 µl of P1 virus (10^{-1}), and then, taking 40 µl of the former dilution into new 360 µl of medium. Once the cells were attached to the 96-well black plate, media was removed and 100 µl of each virus dilution were added in each well with a multi-channel pipet. Each virus dilution was tested by triplicate. At 72 h post-infection, the number of green foci were counted in the dilution that gave <10 foci/well. To calculate the viral titer this equation was used:

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

5.2.3 Small-scale expression and FSEC analysis of vertebrate HATs

For each protein to be expressed, 10 ml of Sf9 cells, at a cell density of $2.0x10^6$ cells/ml, were infected with the corresponding P2 virus at a MOI of 2. The infected Sf9 cultures were incubated at 27°C for 48 h in a non-humidified and non-CO_2 regulated orbital shaker at 120 rpm. 48 h post-infection, 1 ml of culture was centrifuged at 1,500xg for 5 min at 4°C. The cell pellet was resuspended in 2 ml of PBS, and cells centrifuged again at 1,500xg for 5 min at 4°C. After supernatant removal, cells were resuspended in 200 µl of solubilization buffer containing, 50 mM Tris pH 8.0, 150 mM NaCl, 1 mM EDTA, a protease inhibitor cocktail (1 mM PMSF, 200 µM aprotinin, 2 µg/ml leupeptine and 2 µM pepstatin A), 1 % (w/v) of the desired detergent and 0.2 % (w/v) of Cholesteryl Hemisuccinate tris salt (CHS). CHS was solubilized as has been

described previously (see section 5.1.3). Protein solubilization was performed by rotation during 1 h at 4°C. The solubilized sample was centrifuged at 70,000xg in a TL100 ultracentrifuge for 40 min at 4°C, and the supernatant collected. 100 µl of sample were injected into an ÄKTA purifier (GE healthcare life sciences) for FSEC analysis in a Superose 6, 10/300 GL column (GE healthcare life sciences). FSEC buffer containing, 50mM Tris pH 8.0, 150mM NaCl and 0.05% DDM was prepared just before use and filtered by a 0.2 µm pore filter.

5.2.4 Small-scale purification of 4F2hc-GFP/LAT1 and 4F2hc-GFP/LAT2 by strep-tactin affinity chromatography

Sf9 cell cultures of 50 ml at a density of 2.0×10^6 cells/ml were co-infected with P2 viruses of 4F2hc-GFP-His8 and vLAT-Strep-tag II, or of 4F2hc-GFP-His8 and LAT2-Strep-tag II, at a final MOI of 2:2. The infected cultures were incubated for 48 h at 27°C in an orbital shaker at 120 rpm. Cells were centrifuged for 15 min at 8,000xg, and resuspended in 100 ml of PBS. Cell suspension was centrifuged again for 15 min at 8,000xg, and resuspended in 8 ml of solubilization buffer containing, 50 mM Tris pH 8.0, 150 mM NaCl, 1 mM EDTA, a protease inhibitor cocktail (1 mM PMSF, 200 µM aprotinin, 2 µg/ml leupeptine and 2 µM pepstatin A), 1 % (w/v) of DDM and 0.2 % (w/v) of CHS. This mixture was rotated during 1 h at 4°C and ultracentrifuged at 70,000xg during 40 min. The supernatant was used to purify both heterodimers by affinity chromatography using Strep-Tactin Superflow resin (IBA). 0.5 ml of resin were incubated by rotation in batch with the corresponding supernatant during 1 h at 4°C. Samples and resin were transferred to an empty column, and after beads precipitation, flow through was discarded and resin washed with 10 ml of washing buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.05 % DDM, 0.01 % CHS). Heterodimers were eluted with 1.5 ml of washing buffer containing 2.5 mM D-desthiobiotin (IBA).

5.2.5 Thermostability assay applied to 4F2hc-GFP/LAT1 and 4F2hc-GFP/LAT2

4F2hc-GFP/LAT1 and 4F2hc-GFP/LAT2 were expressed in Sf9 cell cultures of 50 ml and purified by strep-tag affinity chromatography as has been described previously in Section 5.2.4. 100µl of purified sample were injected into an ÄKTA purifier (GE healthcare life sciences) for FSEC analysis in a Superose 6, 10/300 GL column (GE healthcare life sciences), in order to check that both heterodimers samples were at the same concentration before performing the thermostability assay. Taking as a reference the height of each FSEC peak, the protein concentration correction was performed. Base on the method described by Hattori and co-workers (Hattori *et al.* 2012), heterodimer samples were divided in 150µl aliquots, and each aliquot was heated to a different temperature (20°C, 30°C, 40°C, 50°C, 60°C and 70°C) during 10 min by using a gradient thermocycler (TGradient96, Biometra). After this time, samples were ultracentrifuged at 70,000xg during 15 min. 100 µl of supernatant were analyzed by FSEC into a Superose 6, 10/300 GL column (GE healthcare life sciences).

5.2.6 Large-scale purification of 4F2hc-GFP/LAT1 by double affinity chromatography

2L flasks with 0.8 L of Sf9 cell culture per flask, at a density of $3\text{x}10^6$ cells/ml, were co-infected with P2 viruses of 4F2hc-GFP and LAT1 at a final MOI of 2:2. The infected cultures were

incubated for 48 h at 27°C in an orbital shaker at 120 rpm. After this time, cells were harvested by centrifugation at 8,000xg for 20 min, resuspended in 400 ml of PBS, and centrifuged again by the same previous conditions. Cells were resuspended in 50ml of resuspension buffer containing, 50 mM Tris pH 8.0, 150 mM NaCl, and a protease inhibitor cocktail (1 mM PMSF, 200 µM aprotinin, 2 µg/ml leupeptine and 2 µM pepstatin A). Cellular lysis was performed in 6 cycles of 15 sec of sonication, 60 % of amplitude and 0.6 of intensity, using a Sonicator KIKA Labortechnik U2005 Control™. Samples were centrifuged at 8,000xg, and the supernatant ultracentrifuged at 70,000xg for 90 min. The pelleted membranes were homogenized in 20ml of membrane resuspension buffer containing, 50 mM Tris pH 8.0 and 150 mM NaCl. Then, 20ml of the same buffer but supplemented with 2 % (w/v) of DDM and 0.4 % (w/v) of CHS were added. Membrane protein solubilization was performed during 2 h at 4°C by stirring, and then, the sample was ultracentrifuged at 70,000xg for 1h (500 µl of supernatant was reserved for western blot and FSEC analysis).

In order to perform two sequential affinity chromatographies, first, the supernatant was incubated by rotation with 2 ml of Strep-Tactin superflow resin (IBA) during 90 min at 4°C. Then, sample and resin were transferred into an empty column, and left until resin precipitation, discarding the flow through. The resin was washed with 40 ml of washing buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.05 % DDM and 0.01 % CHS). Protein elution was made in 6 ml of buffer containing the same components of the washing buffer plus 2.5 mM D-desthiobiotin (IBA) (500 µl of eluted sample were taken for western blot and FSEC analysis). The remaining part of the eluted sample was incubated by stirring with 1 ml of Ni-NTA resin (QIAGEN) for 1 h at 4°C. After this time, sample and resin were placed into an empty column and left until resin precipitation. Flow through was discarded, and column was washed with 20 ml of washing buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.05 % DDM, 0.01 % CHS and 20 mM Imidazole). Protein elution was performed with 3 ml of elution buffer, containing the same components of the former washing buffer but with 250 mM Imidazole. The sample was concentrated until 0.1 mg/ml, and analyzed by FSEC injecting 100 µl of sample into an ÄKTA purifier (GE healthcare life sciences) using a Superose 6 10/300 GL column (GE healthcare life sciences). The purity and integrity of the concentrated sample was analyzed by western blot and coomassie staining, loading 5µl and 39µl of sample, respectively (see section 5.2.6.1).

5.2.6.1 Detection of HATs by western blot and coomassie staining

Western blot

Always 5 µg of protein was charged per well. The samples were mixed with loading sample buffer (LSB: 50 mM Tris/HCl pH 6.8, 10 mM DTT, 2 % (w/v) SDS, 10 % glycerol and 0.1 % (w/v) bromophenol blue), and heated during 5 minutes at 60°C in order to improve protein mobility and band detection. Only when it was required, DTT was added in a final concentration of 100 mM with the aim to reduce the disulfide linkages. The molecular weight marker was provided by Bio-Rad, the range as general moved from 25 kDa to 260 kDa, with variations depending on the batch.

The proteins separated by SDS-PAGE, were transferred during 90 min at 250 mA to a nitrocellulose membrane previously activated with methanol, using the mini-protean 3 kit (Bio-Rad). The detection of 4F2hc and LAT1 was performed following different western blot protocols: 4F2hc detection was made following the protocol from the fabricant for α-His tag

clone His.H8 (05-949, Millipore), and LAT1 was detected following the protocol from the fabricant for StrepMAB Classic HRP antibody (21509-001, IBA).

Coomassie brilliant blue staining

Proteins run in the SDS-PAGE gel, were stained by incubating the gel for 1 h with the staining buffer (45 % methanol, 10 % acetic acid, 45 % distilled water and 0.1 % w/v Brilliant Blue R, Sigma). The gel was washed with a cleaning solution (7.5 % acetic acid, 7.5 % isopropanol and distilled water), until the background in the gel was as transparent as possible.

5.2.7 Expression screening of truncated versions of vertebrate LAT1 in Sf9 cells

With the aim to find a 4F2hc/LAT1 heterodimer with low intra molecular flexibility for structural studies, a screening of the expression of truncated versions of LAT1 at the N-terminus (Δ10Nter, Δ20Nter, Δ30Nter and Δ40Nter), at the C-terminus (Δ10Cter and Δ20Cter) and at both (Δ10Nter/Δ10Cter, Δ20Nter/Δ10Cter, Δ30Nter/Δ10Cter, Δ40Nter/Δ10Cter; Δ10Nter/Δ20Cter, Δ20Nter/Δ20Cter, Δ30Nter/Δ20Cter and Δ40Nter/Δ20Cter) were conducted.

5.2.7.1 Cloning of the truncated versions of LAT1 into pFastBac1 vector

The cDNAs corresponding to the different truncated versions of LAT1, tagged with Strep-tag II in the C-terminus, were cloned into pFastBac1 vector in *Hin*dIII site. cDNAs were amplified by PCR, cloned into pGEM-T Easy vector, subcloned into pFastBac1plasmid, and the corresponding ligation products transformed into XL1 Blue *E.coli* cells by the heat shock method (see section 5.2.1 for detailed information about the cloning protocol). The pFastBac1 generated constructs were sequenced using the BigDye® Terminator v3.1 Cycle Sequencing Kit following manufacturer's protocol. These constructs were used to produce the corresponding P1 and P2 baculovirus (see section 5.2.2).

5.2.7.2 Small-scale expression of truncated versions of LAT1

6×10^6 Sf9 cells were seeded in 10cm plates in 10 ml of 27°C preheated Sf-900™III media. The cells were incubated at 27°C until they attach (about 30 min). P2 viruses were added to a MOI of 2 (see section 5.2.2.3). Expression was performed at 27°C in a non-humidified and non CO_2 regulated incubator during 48 h. After this time, cells were resuspended in 10 ml of PBS, centrifuged for 15 min at 8,000xg, and resuspended in 10 ml of PBS containing a protease inhibitor cocktail (1mM PMSF, 200 µM aprotinin, 2 µg/ml leupeptine and 2 µM pepstatin A). Cellular lysis was performed in 6 cycles of 15 sec of sonication, 60 % of amplitude and 0.6 of intensity, using a Sonicator KIKA Labortechnik U2005 Control™. Samples were centrifuged at 8,000xg, and the supernatant ultracentrifuged at 70,000xg for 40 min. The pelleted membranes were resuspended in 200 µl of membrane resuspension buffer containing 50 mM Tris pH 8.0 and 150 mM NaCl. The expression levels of truncated versions of LAT1 were checked by western blot using a Strep-tag II antibody (see section 5.2.6.1).

5.3 Expression of vertebrate 4F2hc/LAT1 in *Pichia pastoris*

P. pastoris is one of the yeast species capable of metabolizing methanol. The methanol metabolic pathway involves different enzymatic processes, being the first step the oxidation of methanol to formaldehyde. This reaction is catalyzed by the enzyme alcohol oxidase (AOX) in the peroxisome, generating hydrogen peroxide, a toxic compound expulsed away from the rest of the cell. The AOX gene and its promoter were isolated in 1980 by the Salk Institute Biotechnology/Industrial Associate Inc., who developed different vectors and strains, converting *Pichia pastoris* in a strong expression system under the control of this AOX promoter. Later, two genes in *P. pastoris* that code for AOX (*AOX1* and *AOX2*) were found, although the *AOX1* gene is responsible for the vast majority of alcohol oxidase activity in the cell. Expression of the *AOX1* gene is tightly regulated, and induced essentially by methanol to high levels (Ellis *et al.* 1985; Cregg *et al.* 1989). Is important to remark that, *Pichia pastoris,* is useful with both, heterologous expression of intracellular proteins or secreted proteins. In this last case, the recombinant protein requires the presence of a signal sequence to target it to the secretory pathway (for further information visit www. pichia.com).

Three types of host strains derived from NRRL-Y 11430 (Northern Regional Research Laboratories, Peoria, Ill.) are available. These kinds of strains vary with regard to their ability to utilize methanol, based on deletions in one or both AOX genes. To allow the selection, sometimes these strains have mutations in one or more auxotrophic genes, like *His4*, requiring a supplementation with histidine, in this case, to grow on poor media. The most common used strain is GS115 (His4) (also called GS115H); this strain is wild-type for both genes, *AOX1* and *AOX2*, and grows at the normal methanol rate. This kind of strains are called Mut+ phenotype (methanol utilization plus), consuming a large amount of methanol (from 5 % to 30 % of total media). The strains with deletions in one or both AOX genes are usually better producers of foreign proteins. In the present study, the strain KM71H has been used, a strain with a large deletion in the gene *AOX1*. This strain is a Mut^s phenotype (or methanol utilization slow), due to the methanol metabolic pathway is regulated only by the weak gene *AOX2*. The third, and less known host, MC100-3, is deleted for both AOX genes, and is totally unable to grow on methanol.

5.3.1 Design and cloning of the different vertebrate 4F2hc/LAT1 versions expressed in Pichia

The expression vector used to produce the different versions of our transporter in *Pichia pastoris* was pPICZ B (Figure 56) (Life Technologies). This vector allows the methanol inducible expression of the gene of interest, via the AOX1 promoter, in any Pichia strain. In addition, pPICZ B carries a resistance to Zeocin (Life Technologies), which can be used for selection in Pichia and *Escherichia coli*.

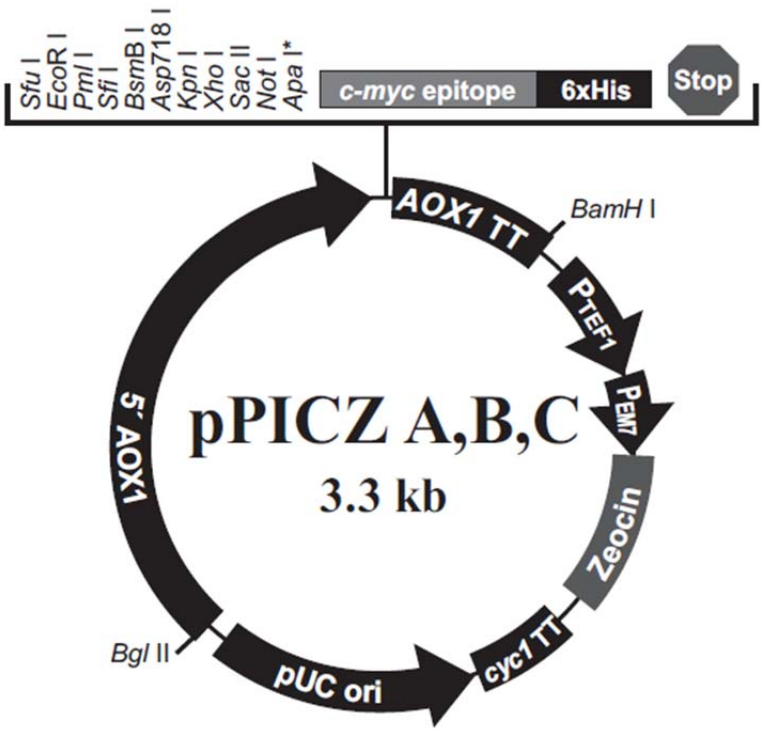

Figure 56. Map of pPICZ A, B, and C. Summary of the features of the pPICZ A, B, and C vectors. 5´AOX1: AOX promoter that allows methanol-inducible high-level expression in Pichia, and targets plasmid integration to the *AOX1* locus. Multiple cloning site with 10 unique restriction sites. AOX1 Transcription Termination (TT). TEF1 promoter: transcription elongation factor 1 gene promoter from *Saccharomyces cerevisiae* that drives expression of the *Sh ble* gene in Pichia, conferring zeocin resistance (GenBank Acc. no. D12478, D01130). EM7 (synthetic prokaryotic promoter): constitutive promoter that drives expression of the *Sh ble* gene in *E. coli*, conferring zeocin resistance. *Sh ble* gene (*Streptoalloteichus hindustanus ble* gene): Zeocin resistance gene for selection in *E. coli.* CYC1 transcription termination region: 3´ end of the *Saccharomyces cerevisiae* CYC1 gene that allows efficient 3´ mRNA processing of *the Sh ble* gene for increased stability (GenBank Acc. no. M34014). pUC origin: allows replication and maintenance of the plasmid in *E. coli*. *Sac* I, *Pme* I, *Bst*X I: Unique restriction sites that permit linearization of the vectors at the *AOX1* locus for efficient integration into the Pichia genome. *The restriction site in pPICZ B between *Not*I and the myc epitope is *Xba*I in pPICZ B. The vector sequences are available in www.invitrogen.com. Figure and information from EasySelect ™ Pichia Expression Kit manual, 2010.

5.3.1.1 GFP-tagged versions: less-glycosylated versions of 4F2hc-GFP and 4F2hc-GFP/LAT1

4F2hc cDNA was amplified by PCR, cloned into pGEM-T Easy vector, and subcloned into a pPICZ B plasmid, between *Eco*RI and *Not*I sites, where previously a Thrombin-GFP-His8 tag was cloned between *Not*I and *Xba*I sites. pPICZB-4F2hc-GFP was used as a template to generate, by PCR site-directed mutagenesis, the less-glycosylated versions of 4F2hc. These mutants were obtained using the QuickChange Site-Directed Mutagenesis Kit (Stratagene), following manufacturer's protocol.

To generate the construct to co-express 4F2hc-GFP and LAT1, first, the cDNA corresponding to the light subunit, tagged with Strep-tag II in the C-terminus and with a Thrombin site between LAT1 and Strep-tag II, was amplified by PCR, cloned into pGEM-T Easy vector, and subcloned into pPICZ B plasmid into *Not*I site. Then, the expression cassette from pPICZB-LAT1 was excised (*Bam*HI/*Bgl*II) and cloned into pPICZB-4F2hc-GFP, linearized previously with *Bam*HI

(Figure 57). The resulting construct, pPICZB-4F2hc-GFP/LAT1, showed both cDNAs (light and heavy subunits), each of them with its own promotor AOX1.

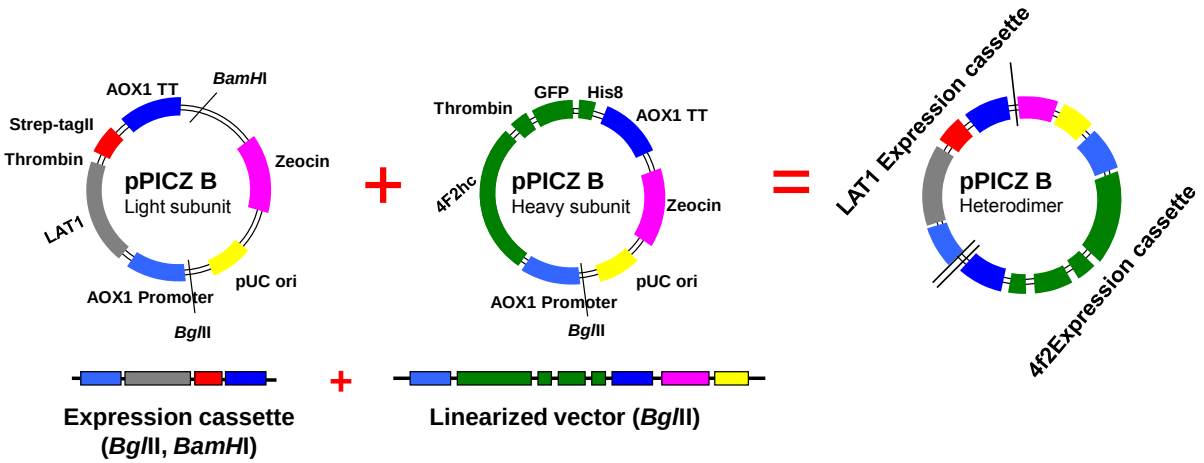

Figure 57. Generation of the pPICZ B vector for the expression of 4F2hc-GFP/LAT1 in Pichia. Both subunits were independently cloned into pPICZ B vector. Then, LAT1 carrying vector was digested with *Bgl*II and *Bam*HI enzymes, and 4F2hc-GFP carrying vector was linearized with *Bgl*II enzyme. Expression cassette from the former digestion (LAT1 carrying) was cloned inside the linearized vector (carrying 4F2hc-GFP) to generate a pPICZ B vector with the expression cassettes for the expression of both proteins in Pichia. The same methodology was used to express non-GFP tagged versions of the heterodimer, but cloning 4F2hc without GFP.

The corresponding ligation products were transformed into XL1 Blue *E.coli* cells by the heat shock method, and plated into LB medium plates containing zeocin (25 μg/mL) (see section 5.2.1 for more detailed information about the cloning protocol). The pPCZ B generated constructs were sequenced using the BigDye® Terminator v3.1 Cycle Sequencing Kit following manufacturer's protocol.

5.3.1.2 NG-4F2hc(Δ36Nter)/LAT1 and 4F2hc/LAT1

NG-4F2hc(Δ36Nter) was generated by PCR using pPICZB-4F2hc-GFP as a template. The resulting amplified cDNA was cloned into pGEM-T Easy vector, and subcloned into a pPICZ B plasmid, between *Eco*RI and *Not*I sites. The same cloning strategy was followed for 4F2hc. In both cases, a His8 tag was placed in N-terminus and a thrombin site was introduced between 4F2hc and His8 tag. To generate the construct to produce NG-4F2hc(Δ36Nter)/LAT1 and 4F2hc/LAT1, the expression cassette from pPICZB-LAT1 was cloned into pPICZB-NG-4F2hc(Δ36Nter) and pPICZB-4F2hc constructs (Figure 57) (see section 5.3.1.1). The corresponding ligation products were transformed into XL1 Blue *E.coli* cells by the heat shock method, and plated into LB medium plates containing zeocin (25 μg/mL). These constructs were sequenced using the BigDye® Terminator v3.1 Cycle Sequencing Kit following manufacturer's protocol.

5.3.2 Transformation of Pichia cells with the pPICZB-HAT constructs by electroporation

To promote construct integration into Pichia genome, the linearization of the pPICZ vector is recommended. This process is done within the 5′ AOX1 promotor using unique restriction sites placed in this region. The cleavage site used in this work was the corresponding to *Pme*I

endonuclease (New England Biolabs). 10 μg from the different constructs generated (See Section 5.3.1), were linearized according to manufacturer's protocol, concentrated until 10 μl, and stored at –20°C. The efficiency of each digestion was checked, using small aliquots, by agarose gel electrophoresis.

Generation of electro-competent *Pichia pastoris* cells

1. Grow a 5 mL culture of Pichia in Yeast Extract Peptone Dextrose (YPD) medium in a 30°C shaking incubator for 16 h at 230 rpm. Preparation protocol of YPD medium is shown in Supplemental Materials and Methods section 5.5.2.

2. Dilute the previous culture to 0.15–0.20 (Abs600nm) in a volume of 50 mL YPD. The flask have to be large enough to provide good aeration.

3. Grow yeast to an absorbance of 0.8–1.0 (Abs600nm) in a 30°C shaking incubator at 230 rpm. It should take 4-5 h.

4. Centrifuge the culture at 500×g for 5 min at room temperature and pour off the supernatant.

5. Resuspend the pellet in 9 mL of ice-cold BEDS (10 mM bicine-NaOH, pH 8.3, 3 % (v/v) ethylene glycol, 5 % (v/v) DMSO, and 1 M sorbitol), supplemented with 200 mM DTT.

6. Incubate the cell suspension for 5 min at 100 rpm in the 30°C shaking incubator.

7. Centrifuge the culture again at 500×g for 5 min at room temperature, and resuspend the cells in 1 mL of BEDS solution without DTT.

8. The competent cells are now ready for transformation. It is possible in that point, to freeze the cells slowly in small aliquots at -80°C. Do not store the cells more than 6 months.

Transformation of *Pichia pastoris* by electroporation

1. Mix approximately 10 μg of linearized pPICZ constructs with 80 μl of electro-competent cells in an electroporation cuvette (0.2 cm gap sterile electroporation cuvette, Bio-Rad). Incubate for 2 min on ice.

2. Electroporate the samples using a Gene Pulser® II electroporator (Bio-Rad Laboratories), following the next parameters: charging voltage, 1,500 V; resistance, 200 Ω and capacitance, 25 μF.

3. Immediately after electroporation, add 0.5 mL 1 M sorbitol and 0.5 mL YPD, and resuspend the cells. Incubate the cell suspension in a 30°C shaker for 3h at 100 rpm.

4. Plate on YPD media containing zeocin (100 μg/mL). Note: Increasing the Zeocin concentration from 100 μg/ml to 1000 μg/ml is useful in the selection of multicopy integrants.

5.3.3 Small-scale expression screenings in Pichia of 4F2hc, LAT1 and 4F2hc/LAT1 versions

A small-scale initial screening is required to identify recombinant *Pichia pastoris* clones able to express, at good levels, the desired protein. 6-10 colonies, depending on the construct, providing from the YPD-zeocin plates from the Pichia transformation (See Section 5.3.2) were used to inoculate 5 ml of YPD medium with zeocin (100 µg/ml). These cultures were grown in a shaking incubator at 30°C and 230 rpm, for 16 h. The clones growing in these conditions, were considered positive zeocin resistant transformants, and they were used to do the expression screening. To perform the small-scale expression screening the next protocol was performed with all positive transformants:

1. Inoculate 15 ml of Buffered Glycerol-complex medium (BMGY) with 10 µl from the YPD-zeocin cultures, in a 50 ml conical polypropylene tube. Preparation protocol of BMGY medium is shown in Supplemental Materials and Methods section 5.5.2.

2. Grow at 30°C in a shaking incubator at 230 rpm until culture reaches an Abs600nm close to 2–6, approximately 16–18 h.

3. Centrifuge the cells at 1,500xg for 5 min at room temperature.

4. Decant the supernatant, and resuspend the cell pellet in 3 ml of Buffered Methanol-complex medium (BMMY) to start the induction of protein expression. Cover the tubes with 2 layers of sterile gauze to allow a proper aeration. The protein expression was performed at 30°C and 230 rpm for 48 h. Methanol was added again after the first 24 h. The concentration of methanol used to induce protein nexpression was 1 % (v:v). Preparation protocol of BMMY medium is shown in Supplemental Materials and Methods section 5.5.2.

5. Centrifuge cells at 1,500xg for 5 min at room temperature.

6. Decant the supernatant and store the cell pellets at –80°C until ready to process.

In order to disrupt the *Pichia pastoris* cells expressing our proteins at small-scale, the pellet was resuspended in 300 µl of breaking buffer (50 mM sodium phosphate, pH 7.4, 1 mM EDTA, 5% glycerol), supplemented with protease inhibitors (Complete, EDTA-free; Roche). Then, the same volume of glass beads (212-300 µm, Sigma-Aldrich) was added. The samples were subjected to 3 cycles of 10 sec of lysis using a FastPrep disruptor (FP120A Instrument (120W), QBiogene). After this first lysis step, samples were centrifuged at 600xg during 10 min at 4°C, and the supernatant removed and kept it in ice. 300 µl of new breaking buffer were added again to the sample, and the lysis was repeated under the same previous conditions. To obtain the total membrane fraction, both supernatants were pooled and ultracentrifuged at 250,000xg and 4°C for 1 h. The membranes pellets were resuspended in 200 µl of Tris 20mM, 150 mM NaCl, 5 % glycerol, and stored at -80°C. The amount of total membrane protein was evaluated by the BCA Assay (Pierce). The protein expression levels were analyzed by western blot using antibodies against Strep-tag II or His8, detecting LAT1or 4F2hc, respectively (see section 5.2.6.1). Additionally, the expression levels of the less-glycosylated versions of 4F2hc-GFP, were analyzed measuring the whole cell GFP fluorescence and by in-gel fluorescence. To measure the fluorescence, the cell pellets providing from the small-scale cultures were resuspended in 800 µl of breaking buffer, and the fluorescence detected using a fluorimeter, with excitation and

emission wavelengths of 488 nm and 512 nm, respectively). The in-gel fluorescence analysis was performed using isolated membranes obtained as has been described previously. These samples were run in a SDS-PAGE, and GFP-tagged proteins visualized exposing the gels to blue light. The samples have not to be boiled.

5.3.4 Expression at large-scale of LAT1 and 4F2hc/LAT1 versions in Pichia

Independently of the overexpressed protein, the protocol was always the same:

1. Inoculate 70 ml of BMGY medium with a single colony providing from a good clone in terms of protein expression levels.

2. Grow at 30°C in a shaking incubator at 230 rpm for 16 h.

3. Inoculate 1L of BMGY medium, placed in a 2L flask, with 10 ml of the previous BMGY culture. 6L of BMGY medium were used.

4 Grow at 30°C in a shaking incubator at 230 rpm until culture reaches an Abs600nm close to 2–6, approximately 16–18 h.

5. Centrifuge the cells using sterile centrifuge tubes at 3,500xg for 15 min at room temperature.

6. Decant the supernatant, and resuspend the cell pellet in 200 ml of BMMY, for each flask, to start the induction of protein expression. Cover the flasks with 2 layers of sterile gauze to allow a proper aeration. The protein expression was performed at 30°C and 230 rpm for 48 h. Methanol was added again after the first 24 h. The concentration of methanol used was 1 % (v:v).

7. Centrifuge cells at 3,500xg for 15 min at room temperature.

8. Decant the supernatant and store the cell pellets at –80°C until ready to process.

The cell pellets were resuspended in a final volume of 200 ml of breaking buffer, supplemented with protease inhibitors (Complete, EDTA-free; Roche). This cell suspension was subjected to lysis, using a cell disruptor (of the Protein Expression IRB Core Facility), by 2 initial cycles at 30 KPSI, and 2 final cycles at 40 KPSI. Once the cellular lysis finished, the sample was centrifuged for 15 min at 4ºC and 20,000xg. The supernatant was ultracentrifuged at 250,000xg and 4°C for 1 h. The resulting membrane pellet was resuspended in 60 ml of Tris 50mM/150 mM NaCl/5 % glycerol, and stored at -80°C. The amount of total membrane protein was evaluated by the BCA Assay (Pierce).

5.3.5 Large-scale purification of LAT1 and 4F2hc/LAT1 versions from Pichia

5.3.5.1 Purification of LAT1 by strep-tactin affinity chromatography

Isolated membranes from a 6L culture of Pichia expressing LAT1 (see section 5.3.4), were solubilized in 2 % DDM/ 0.4 % CHS by stirring for 1 h at 4°C. During this step, the total protein concentration was kept around 5 mg/ml. After that, the sample was ultracentrifuged at 250,000xg and 4°C for 1 h, and the supernatant subjected to strep-tactin affinity chromatography using 10 ml of strep-tactin superflow resin (IBA). Before binding, the resin was equilibrated with 300 ml of equilibration buffer (Tris 50 mM, pH 8 / 150 mM NaCl). The first binding was performed by stirring for 1 h at 4°C, after this step, the mixture was transferred to an empty

column, where a second binding was done. 300 ml of washing buffer (Tris 50 mM, pH 8 / 150 mM NaCl / 0.05 % DDM / 0.01 % CHS) were used to clean the resin. The elution was carried out adding 20 ml of elution buffer (Tris 50mM, pH 8 / 150 mM NaCl / 0.05 % DDM / 0.01 % CHS / 2.5 mM D-desthiobiotin). The purified protein was concentrated until the desired concentration by means of a concentrator with a cut off of 100 KDa (Millipore). The monodispersity of the purified protein was analyzed by Size Exclusion Chromatography (SEC) in an Äkta purifier (GE healthcare) using a Superose 6 10/300 GL column (GE healthcare). This same protocol was followed to purify the 4F2hc-GFP/LAT1 used to perform the UDS assays (see Results and Discussion section 3.4.4.1).

5.3.5.2 Purification of 4F2hc/LAT1 versions by double-affinity chromatography

The protocol described below was used to purify all versions of 4F2hc/LAT1. Depending on the version, could be slight variations that are described in the corresponding section of Results and Discussion.

Two tags were present, His8 tag in the heavy subunit (4F2hc), and Strep-tag II in the light subunit (LAT1). This fact allows the possibility of a double-affinity chromatography to obtain a purer heterodimer. Isolated membranes from a 6L culture of Pichia expressing 4F2hc/LAT1 (see section 5.3.4), were solubilized in 2 % DDM / 0.4 % CHS by stirring for 1 h at 4°C. During this step, the total protein concentration was kept around 5 mg/ml. After that, the sample was ultracentrifuged at 250,000xg and 4°C for 1 h, and the supernatant subjected to an Immobilized Metal Ion Affinity Chromatography (IMAC) using 15 ml of Ni-NTA resin (Qiagen). Before binding, the resin was equilibrated with 450 ml of equilibration buffer (Tris 50mM, pH 8 / 150 mM NaCl). The first binding was performed by stirring for 1 h at 4°C, after this step, the mixture was transferred to an empty column, where a second binding was done. 450 ml of washing buffer (Tris 50mM, pH 8 / 150 mM NaCl / 0.05 % DDM / 0.01 % CHS / 20 mM Imidazole) were used to clean the resin. The elution was carried out adding 60 ml of elution buffer (Tris 50mM pH 8 / 150 mM NaCl / 0.05 % DDM / 0.01 % CHS / 250 mM Imidazole). The eluted fraction was subjected to strep-tactin affinity chromatography following exactly the same protocol used to purify LAT1 (see section 5.3.5.1). The monodispersity of the purified protein was analyzed by Size Exclusion Chromatography (SEC) in an Äkta purifier (GE healthcare) using a Superose 6 10/300 GL column (GE healthcare).

5.3.6 Scintillation Proximity Assay (SPA)

The scintillation proximity assay was performed based on the protocol described for the prokaryotic LAT homologue AdiC (Harder and Fotiadis 2012), with slight variations. After LAT1 or 4F2hc/LAT1 purification (see sections 5.3.5.1 and 5.3.5.2), reactive were added like described in Table 8, from left to right. Each condition (Table 8, column 1) was performed per triplicate in 96 well plates for scintillation counting (TopCount: white) at a final volume of 100 µl. The buffer (1x) indicated in the second column (Table 8) is the same buffer used in the purification of LAT1 or 4F2hc/LAT1, 50 mM Tris pH 8, 150 mM NaCl, 0.05 % DDM and 0.01 % CHS. For control conditions with D-desthiobiotin, this buffer additionally contained 10 mM D-desthiobiotin. Cold L-leucine, L-[^{3}H]-Isoleucine, and Ysi beads, were prepared in the buffer described before, with or without D-desthiobiotin depending on the sample, prior SPA experiment.

When human LAT1 inhibitor, KYT-0353 (Oda *et al.* 2010), was tested, four concentrations of it were investigated for LAT1 (60 µM, 6 µM, 0.6 µM, 0.06 µM) and one for 4F2hc/LAT1 (6 µM). KYT-0353 compound was kindly provided by Dr. Michael F. Wempe, associate research professor of the School of Pharmacy and Pharmaceutical Sciences of the University of Colorado. This inhibitor was prepared at concentrations of 600 µM, 60 µM, 6 µM and 0.6 µM by serial dilution in buffer (50 mM Tris pH 8, 150 mM NaCl, 0.05 % DDM and 0.01 % CHS), from a stock solution in DMSO anhydrous at 10mM. Then, 10 µl of these solutions were added to the SPA mix for reaching the corresponding final concentration.

Condition	Buffer (1x) without or with D-desthiobiotin (10 mM)	Cold L-leucine (50 mM) or KYT-0353 (different concentrations)	L-[^{3}H]-isoleucine (0.5 µCi in 10ml buffer)	Transporter (0.5 mg/ml)	YSi beads in 1x buffer
No protein	No D-desthiobiotin	0	10 µl	0	220µg
No protein + cold substrate	No D-desthiobiotin	10 µl L-leucine	10 µl	0	220µg
No protein + KYT-0353	No D-desthiobiotin	10 µl KYT-0353	10 µl	0	220µg
No protein + D-desthiobiotin	With D-desthiobiotin	0	10 µl	0	220µg
Protein	No D-desthiobiotin	0	10 µl	5-10 µg	220µg
Protein + cold substrate	No D-desthiobiotin	10 µl L-leucine	10 µl	5-10 µg	220µg
Protein + KYT-0353	No D-desthiobiotin	10 µl KYT-0353	10 µl	5-10 µg	220µg
Protein + D-desthiobiotin	With D-desthiobiotin	0	10 µl	5-10 µg	220µg

Table 8. Description of the experimental conditions included in the SPA assays.

After all the conditions were settled, plates were sealed with a transparent sticky film and incubated light protected by agitation at 4°C during 1 h. Plates were measured in a scintillation counter, and after, they were incubated again under the same conditions by 2 h more, then counting was repeated, and plates were incubated again for 16 h more.

5.3.7 Exchange of DDM detergent for amphipol A8-35

The strategy used here was based on the protocols followed in the structure determination of TRPV1 channel (Liao *et al.* 2014) and γ-secretase (Bai *et al.* 2015). In our case, after 4F2hc/LAT1 purification by the sequential affinity chromatographies (IMAC followed by strep-tactin chromatography), protein was concentrated to 0.1 mg/ml, and amphipol A8-35 was added in a ratio 1:3 (protein/amphipol, w/w). Amphipol stock has been previously prepared in MilliQ water at 20mg/ml by stirring ON at 4°C. The protein/amphipol mixture was incubated by gentle stirring at 4°C for 4 h. DDM was adsorbed by incubation with Bio-Beads SM-2 Adsorbent (Bio-Rad Laboratories) at 4°C during 9 h by stirring (15mg of Bio-Beads/ml of sample). To check the stability of 4F2hc/LAT1, protein was concentrated around 1 mg/ml, by using centricon with a cut-off of 100 KDa (Millipore), and SEC was performed by using buffer, 50 mM Tris pH 8, 150 mM NaCl, without DDM and without CHS.

5.4 Single particle analysis and 3D reconstruction of vertebrate 4F2hc/LAT1

5.4.1 Negative staining

Negative staining is a technique widely used in electron microscopy to visualize with major contrast biomolecules. Macromolecular complexes or proteins can be observed with increased definition due to the high contrast provided by the staining, which allows their analysis and the generation of their 3D models at low resolution. Stains usually used are heavy metal salts, generally derived from molybdenum, uranium, or tungsten, they surrounds the organic samples but do not covered them, and for this reason is defined as a negative staining. In this way, at the microscope, electrons are more deflected by the staining than by the biomolecules causing phase contrast when those highly deflected electrons are filtered out by the objective aperture located below the sample. For negative staining, first the protein has to be deposited in grids in which the staining is performed.

Protein from the central fractions of the peaks of DDM/CHS-solubilized LAT1 and of amphipol-solubilized 4F2hc/LAT1, obtained by SEC, were subjected to negative staining. Dilutions of 4F2hc/LAT1 (80μg/ml, 60μg/ml, 15μg/ml 12μg/ml 8μg/ml and 6μg/ml) were prepared in the SEC buffer, containing 150 mM NaCl and 50 mM Tris pH 8.0. LAT1 (30 μg/ml, top fraction of the peak) and 4F2hc/LAT1 were adsorbed to glow-discharged Formvar (only LAT1) or carbon film grids (200 and 400 MESH) by incubating the grid above a 5 μl protein drop. Adsorption time was of 1 min for LAT1 and of 15 sec for the heterodimer. The surfaces of grids with LAT1 were successively incubated above 3 drops (10 μl) of Milli-Q water (1 min in total) and stained with uranyl formate by incubation during 1 min. The excess of stain was blotted carefully touching with filter paper the borders of the grids and the opposite surface, where there was no protein. In the case of 4F2hc/LAT1, the blotting of the excess of solutions was performed after the contact of the grid with each drop, in the way described before. In this occasion, the time of incubation with Milli-Q water was reduced to simple contact with two drops of 50 μl, then, after blotting the staining was performed with 10 μl of 2% uranyl acetate and the excess of stained was blotted as already described.

5.4.2 Single particle analysis TEM

Sample of 4F2hc/LAT1 was solubilized in amphipol A8-35 as previously described. After the SEC, protein from the middle fractions of the peak was left during 20 hours at 4°C. Soluble fraction was obtained by ultracentrifugation and subjected to negative staining as described before.

Single particle analysis and 3D reconstruction of amphipol-solubilized 4F2hc/LAT1 was performed by Dr. Rafael Núñez-Ramírez from the Electron Microscopy service of the CIB-CSIC. Briefly, by using the Xmipp software (de la Rosa-Trevin *et al.* 2013), automatic particle selection (32725 particles) and classification were made. Particle alignment in each class, led to determine reference-free 2D class averages. In this point, heterogeneity between particles was found, mainly in the LAT1 density, as discussed in Results and Discussion section 3.4.9.2. Thus, for being able to analyze 4F2hc/LAT1 particles, a more homogenous set of particles was obtained by discarding those excessively big, as estimated by the previously observed for human 4F2hc/LAT2 (Meury *et al.* 2014; Rosell *et al.* 2014). A total of 14339 particles were obtained in

this way, and 2D class averages were determined with Xmipp software (de la Rosa-Trevin *et al.* 2013).

5.4.3 3D reconstruction

Negative staining was performed, as previously described, with amphipol-solubilized 4F2hc/LAT1 eluted from the central fractions of the peak, and immediately after its final purification by SEC. A total of 51312 particles were selected and classified by using the Xmipp software (de la Rosa-Trevin et al. 2013). Particle alignment in each class led to the reference-free 2D class averages, then, due to the heterogeneity in particle size was also present this time (mainly in LAT1), a homogeneous group of 35530 particles was used to calculate an initial 3D model using Xmipp software (de la Rosa-Trevin et al. 2013). This model was used as reference to classify particles into 3 groups with RELION (Scheres 2012). Group 3 of particles was the most abundant in the sample and the better structurally defined, as discussed in Results and Discussion section 3.4.10. For this reason, 2D class averages and the 3D model of group 3 of particles were refined by using RELION (Scheres 2012). For volume rendering of the 4F2hc/LAT1 map, the density threshold was determined by fitting the surface representation of the crystal structure of 4F2hc ectodomain (PDB: 2DH2) into the small density.

5.5 Supplemental Materials and Methods

5.5.1 Growth and maintenance of the Sf9 insect cell line

Sf9 insect cells recovery after freezing

Under sterile conditions, using a laminar flow hood, 9ml of serum-free media Sf-900™III (Life Technologies) pre-warmed at 27ºC were added into a sterile 15 ml falcon, meanwhile frozen Sf9 cells were thawed by 1-2 min at 27°C in a water bath. Just before complete thawing, Sf9 vial was decontaminated by spraying 70% ethanol, and cells carefully resuspended into the pre-warmed Sf-900™III media. Cell suspension was centrifuged at 400xg for 3 min at room temperature, the medium was removed under sterile conditions, and cells were resuspended in other 9 ml of fresh pre-warmed Sf-900™III medium. This procedure was performed twice more. After the last medium removal, 10 ml of fresh pre-warmed Sf-900™III were added and the cell suspension transferred into a sterile 125 ml shake flask. Sf9 cells were incubated at 27°C and at 120 rpm under a non-humidified and non-CO_2 regulated atmosphere. Growth and viability was determined 48 hours later by trypan blue dye exclusion assay.

Cell counting and viability

For cell counting and viability determination, 100 µl of cell culture were transferred into a 1.5 ml tube and mixed with 100 µl of trypan blue. After mixture homogenization, 10 µl of sample were placed into a Newbauer chamber. Total number of cells, and number of non-dye (viable) cells were determined in the four big squares, and then average of both counts was determined. The number of cells is given by the formula N x 2 x 10^4 cells/ml, in which N is the average of cells (total or viable). The percentage of cellular viability was calculated from the quotient between the average of viable cells and the average of total cells.

Maintenance of Sf9 suspension cultures

Sf9 insects cells were maintain viable in suspension by regular passage from cultures at mid-log phase of growth ($2x10^6$ to $4x10^6$ cells/ml) to a culture density around $0.5x10^6$ cells/ml. The cell passage was performed using serum-free media Sf-900™III pre-warmed at 27°C. The final volume of the cell suspension should be around 40% of the whole shake flask volume for optimal culture aeration. The entire process was always performed under sterile conditions and using new sterile shake flasks. Suspension cell cultures were grown at 27°C and 120 rpm in a non-humidified and non-CO_2 regulated atmosphere. The maximum passage number for Sf9 cells to produce protein is 20-25, and to generate baculovirus is 25-30.

Cryo-preservation and storage of Sf9 insect cells

For storage of frozen Sf9 cells, 300 mL of suspension cell cultures at $2x10^6$ cells/ml were centrifuged at 400xg for 2 min at room temperature. Sf900™III media was removed, cells resuspended in 30 ml of cryo-preservation media (46.25 % fresh Sf-900™III, 46.25 % conditioned Sf-900™III and 7.5 % DMSO), and aliquoted in cryo-tubes vials. Aliquots were kept for 16 h at –80 °C, and the next day they were transferred into a liquid nitrogen storage tank.

5.5.2 Media and buffers preparation for protein expression in *Pichia pastoris*

YPD (Yeast Extract Peptone Dextrose Medium)

For 1L: 20 g of peptone (CONDA) and 10 g of yeast extract (CONDA) were dissolved in 900 ml of milliQ water. This solution was sterilized for 30 min at 121ºC. Just before use, when medium is colder than 55ºC and under sterile conditions, 100 ml of 10X dextrose, previously prepared and sterilized, were added and mixed. Medium can be stored at 4ºC until use.

YPD + Zeocin

YPD medium was prepared as above described, and in sterile conditions, when medium is under 55ºC for the addition of 100 ml of 10X dextrose, zeocin antibiotic (Life Technologies) was also added at a final concentration of 100 µg/ml. Medium containing Zeocin has a shelf life of one to two weeks light protected.

YPD agar plates

For 1L: 20 g of peptone (CONDA), 10 g of yeast extract (CONDA) and 20 g of agar, were dissolved in 900 ml of milliQ water and sterilized by 30 min at 121ºC. Then, when medium is under 55ºC and in sterile conditions, 100 ml of 10X dextrose, previously prepared and sterilized, were added and mixed. YPD agar was added in plates, and after solidification, they were stored at 4ºC until used.

YPD agar + Zeocin plates

YPD agar was prepared as above described, and in sterile conditions, when medium is under 55ºC for the addition of 100 ml of 10X dextrose, zeocin antibiotic (Life Technologies) was also added at a final concentration of 100 µg /ml. YPD agar (+Zeocin) was added in plates, and after

solidification, they were stored light protected at 4ºC until used. YPD agar (+Zeocin) plates have a shelf life of one to two weeks.

BMGY (buffered complex glycerol medium)

For 1L: 20 g of peptone (CONDA) and 10 g of yeast extract (CONDA) were dissolved in 700 ml of milliQ water, and sterilized by 30 min at 121ºC. Just before use, media was supplemented with 100 ml of 10X YNB, 100 ml of potassium phosphate pH 6.0, 100 ml of 10 % glycerol and 2 ml of 500X biotin. All stock solutions for media supplementation were previously prepared and sterilized as described later.

BMMY (buffered complex methanol medium)

For 1L: 20 g of peptone (CONDA) and 10 g of yeast extract (CONDA) were dissolved in 700 ml of milliQ water, and sterilized by 30 min at 121ºC. Just before use, media was supplemented with 100 ml of 10X YNB, 100 ml of potassium phosphate pH 6.0, 2 ml of 500X biotin and 100 ml of 10 % methanol.

10X (20 %, w:w) Dextrose

1L stock: 200 g of D-glucose were dissolved in 1L of milliQ water, and sterilized for 30 min at 121ºC. Stock was stored at 4ºC until use.

10X (13.4 %, w:w) YNB (Yeast Nitrogen Base)

1L stock: 134 g of yeast nitrogen base (YNB), with ammonium sulfate and without amino acids, were dissolved in 1L of milliQ water and filter sterilized by using 0.2 μm pore filters. Stock was stored at 4ºC until use. The shelf life of this solution is approximately one year.

500X (0.02 %, w:w) Biotin

100 ml stock: 20 mg of biotin were dissolved in 100 ml of water and filter sterilized by using 0.2 μm pore filters. Stock was stored at 4ºC until use. The shelf life of this solution is 1 year approximately.

1M potassium phosphate buffer, pH 6.0

1L stock: 132 ml of 1M K_2HPO_4 were mixed with 868 ml of 1M KH_2PO_4, and if needed, pH was adjusted to 6.0 ± 0.1 with KOH. The buffer was sterilized at 121ºC for 30 min, and stored at room temperature.

10X (10%, v:v) Glycerol

1L stock: 100 ml of 100 % glycerol were mixed with 900 ml of water and sterilized for 30 min at 121ºC. Stock was stored at 4ºC until use.

10X (10%, v:v) Methanol

300 ml stock: 30 ml of 99 % methanol were mixed with 270 ml of milliQ water and filter sterilized. Stock was stored at 4ºC until use.

6. Summary in Spanish

6.1 Introducción

Los transportadores de aminoácidos son de suma relevancia debido a la gran cantidad de procesos celulares en los cuales los aminoácidos son esenciales. Entre ellos se encuentran la síntesis de proteínas, la regulación del metabolismo, el crecimiento celular, la producción de energía metabólica, la transmisión neuronal y la señalización celular. Estas proteínas son esenciales para la absorción de los aminoácidos, y su transferencia entre órganos, células, y compartimentos celulares. Adicionalmente, con otros transportadores, también ayudan al mantenimiento de la concentración iónica celular y al flujo de nutrientes (Christensen 1990).

Los transportadores heteroméricos de aminoácidos (HATs por sus siglas en inglés), se caracterizan por ser el único ejemplo conocido, en todos los reinos de la vida, de transportadores de solutos formados por dos subunidades, una pesada y una ligera, unidas por un puente disulfuro (Figura 1) (Broer and Palacin 2011). Dos subunidades pesadas (4F2hc y rBAT) y 10 subunidades ligeras, han sido identificadas en mamífero. Seis subunidades ligeras se asocian con 4F2hc (LAT1, LAT2, y+LAT1, y+LAT2, asc1 y xCT) y dos, AGT-1 y b^{0+}AT, se asocian con rBAT (Fotiadis *et al.* 2013). La subunidad pesada, perteneciente a la familia SLC3 de transportadores de solutos, es la responsable del tráfico del heterodímero a la membrana plasmática. La subunidad ligera, pertenece a la familia de transportadores LATs, dentro de la familia SLC7, y es la subunidad a cargo del transporte de aminoácidos (Reig *et al.* 2002; Broer and Palacin 2011). Las subunidades pesadas son *N*-glicoproteínas de membrana de tipo II, con un solo dominio transmembrana (TMD), un N-terminal intracelular y un ectodominio (ED) extracelular (Figura 1). Por otro lado, los LATs son proteínas de membrana no glicosiladas con 12 TMDs, y extremos N- y C-terminales citosólicos (Figura1).

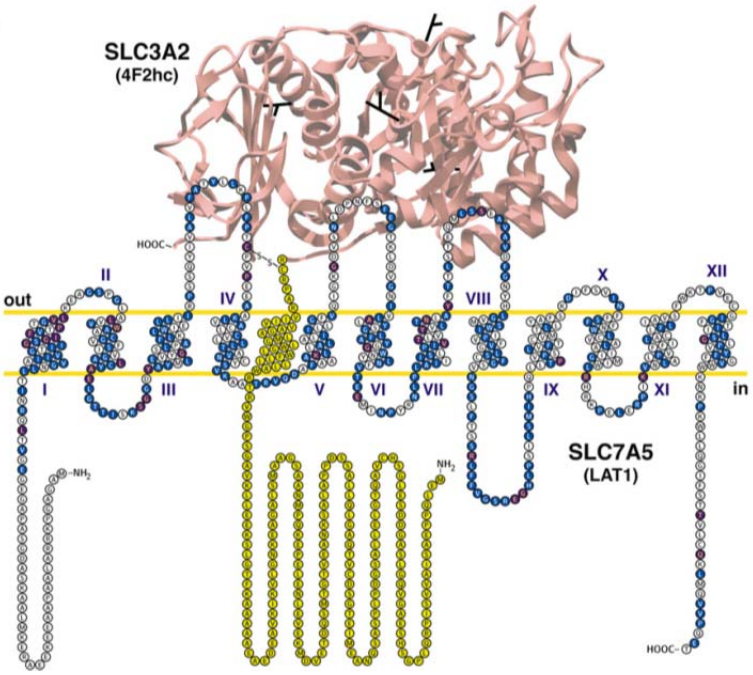

Figura 1. Representación de la topología de los HATs. Se muestra el modelo de 4F2hc/LAT1 como miembro representativo de los HATs. Subunidades ligeras de los HATs (LATs) se asocian con la subunidad 4F2hc o rBAT por puente disulfuro. Las subunidades pesadas son *N*-glicoproteínas de membrana de tipo II. Los 4 sitios de glicosilación en 4F2hc están indicados en negro en la estructura atómica del ectodominio de 4F2hc, mostrado como "cartoon" rosa (Fort *et al.* 2007). Los LATs no son glicosilados y tienen de 12 TMDs (indicados en números romanos). (Fotiadis *et al.* 2013).

Debido a que los HATs están relacionados desde crecimiento celular, tumoral, y metabolismo, hasta transmisión sináptica, están diferencialmente expresados en las membranas celulares de diferentes tejidos. Como consecuencia de esto, la sobreexpresión de algunos de ellos, o por el contrario, la ausencia de su función está relacionada con diferentes patologías. Por ejemplo, 4F2hc/LAT1 y 4F2hc/xCT han sido encontrados sobrexpresados en diferentes tumores y carcinomas, proveyendo nutrientes para el crecimiento y supervivencia tumoral y metástasis (Wolf *et al.* 1996; Sato *et al.* 1999; Kim *et al.* 2001; Yanagida *et al.* 2001; Fuchs and Bode 2005, Huang and Ingber 2005; Kaira *et al.* 2008; Kobayashi *et al.* 2008; Lo *et al.* 2008; Seib *et al.* 2011; Takeuchi *et al.* 2013). Por otro lado, mutaciones en la subunidad ligera y+LAT1 son causantes de lisinuria con intolerancia a proteínas (Oyanagi and Nagao 1998; Mykkanen *et al.* 2000; Sebastio *et al.* 2011). De forma similar, mutaciones en la subunidad pesada rBAT y en la ligera $b^{0,+}$AT causan cistinuria (Chillaron *et al.* 2010).

Es así que, debido a la gran relevancia de los HATs para la salud humana, es de gran importancia obtener la estructura de estos transportadores, ya que hay una estrecha relación entre la estructura y la función de las proteínas. En esta forma se podría elucidar el mecanismo molecular del transporte de aminoácidos en condiciones normales, y ayudar a entender las bases moleculares de algunas de estas enfermedades. Esto podría ayudar a desarrollar o mejorar drogas terapéuticas para estas condiciones (Kim et al., 2008). Desafortunadamente, debido a la gran hidrofobicidad de las proteínas de membrana, y a su baja producción en sistemas de expresión heterólogos, la única información estructural de la que disponemos para los HATS es: la estructura atómica del ectodominio de 4F2hc humano (Fort *et al.* 2007), la estructura de un homólogo procariota de los LATs, el intercambiador de Arginina/Agmatina AdiC (Fang *et al.* 2009; Gao *et al.* 2009; Gao *et al.* 2010; Kowalczyk *et al.* 2011), con un 19 % de identidad de secuencia, y un modelo a baja resolución (21Å) del heterodímero 4F2hc/LAT2 humano (Rosell *et al.* 2014). Aunque este modelo reveló nueva información acerca de la interacción entre ambas subunidades, la baja estabilidad de 4F2hc/LAT2 no permitió realizar otros estudios para obtener resolución atómica.

La baja estabilidad de este heterodímero pareció estar asociada con la baja estabilidad de la subunidad ligera LAT2, ya que cuando esta subunidad era purificada agregaba completamente (Figura 2A, izquierda) (Rosell *et al.* 2014). Este hecho, impidió la reconstitución funcional de LAT2 humano en liposomas, mientras que su co-expresión con 4F2hc producía heterodímeros funcionales que podían ser reconstituidos y mostraban transporte de L-Isoleucina (Figura 2A, derecha). Interesantemente, se encontró que 4F2hc incrementaba en gran medida la estabilidad de LAT2 en solución, incluso sólo el ectodominio (ED) de 4F2hc producía este efecto (Figura 2B y 2C).

Esto sugirió que, posiblemente obteniendo una subunidad ligera más estable que LAT2 humano, se podría obtener un heterodímero más estable. En nuestro laboratorio se realizaron estudios para encontrar una subunidad ligera más estable (Álvarez-Marimon. 2014). 24 subunidades ligeras de metazoos, de siete especies diferentes, fueron expresadas en *Saccharomyces cerevisiae*, y sus niveles de expresión y estabilidad analizados. Entre todos los LATs estudiados, LAT1, LAT2 y $b^{0,+}$AT fueron encontrados como los más estables. Debido a que este trabajo no está publicado preferimos mantener la confidencialidad de la especie, entendiéndose que a menos que se indique lo contrario, las proteínas a las que nos referimos de ahora en adelante son las nombradas anteriormente y sus respectivas subunidades pesadas de la misma especie. Este resultado sugirió que los heterodímeros formados por estas subunidades

ligeras (4F2hc/LAT1, 4F2hc/LAT2 y rBAT/$b^{0,+}$AT), podrían ser buenos candidatos para estudios estructurales.

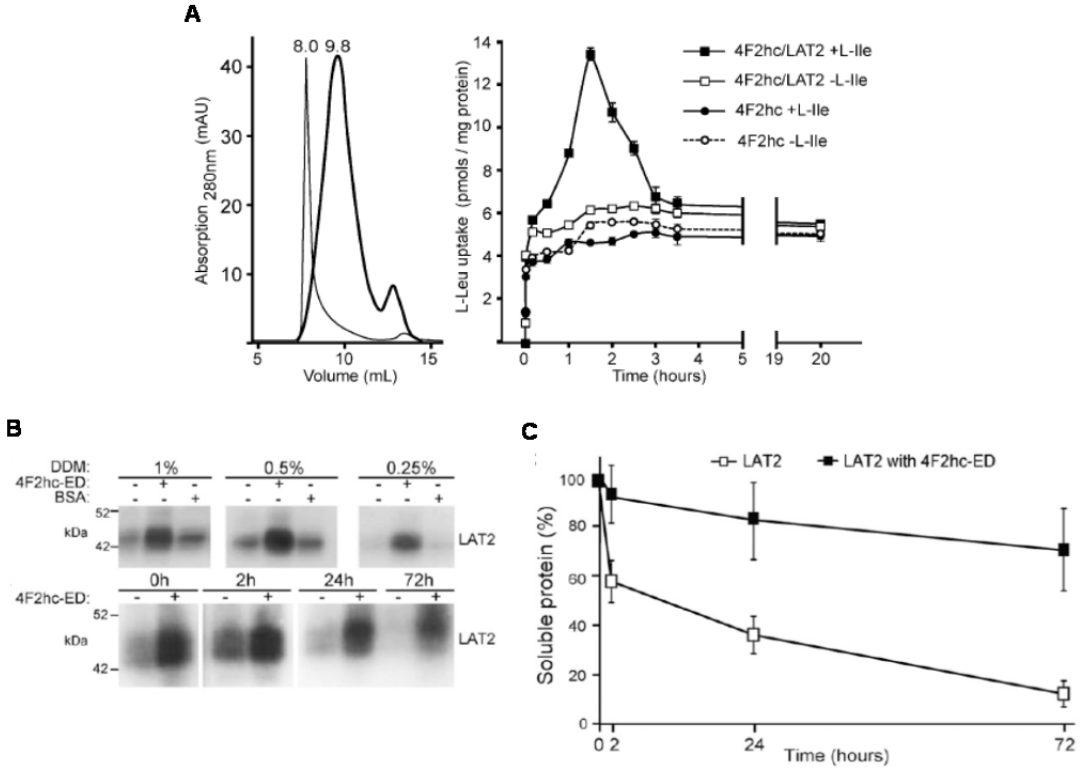

Figura 2. LAT2 humano es estable en solución solo en presencia de 4F2hc. (**A**) Panel de la izquierda: un pico de LAT2 en el "void volume" indicando su completa agregación (8.0 ml), pero un pico de proteína soluble aparece cuando LAT2 es co-expresado con 4F2hc (9.8 ml). (**A**) panel de la derecha: solamente 4F2hc/LAT2 pudo ser reconstituido en liposomas mostrando transporte de L-[^{3}H]-Isoleucina, a diferencia de LAT2 expresado solo. (**B**) El ectodominio (ED) de 4F2hc permitió la solubilización de LAT2 en diferentes concentraciones de detergente DDM (arriba) y al paso del tiempo (**C**). Figura adaptada (Rosell *et al.* 2014).

6.2 Objetivos

I. Identificar un heterodímero con altos niveles de expresión y lo suficientemente estable para abordar estudios estructurales:

1. Analizar los niveles de expresión y la estabilidad en diferentes sistemas de expresión, de tres heterodímeros de vertebrado formados por subunidades ligeras previamente identificadas como las mas estables entre 24 subunidades ligeras estudiadas: rBAT/$b^{0,+}$AT, 4F2hc/LAT1 y 4F2hc/LAT2.

2. Modificar el heterodímero seleccionado en el objetivo anterior (4F2hc/LAT1) para incrementar su homogeneidad y disminuir su flexibilidad para cristalografía 3D.

3. Determinar la funcionalidad del heterodímero 4F2hc/LAT1 de vertebrado, ya que de ser activo esto indicaría su correcto plegamiento en el sistema de expresión y en solución.

II. Optimizar las condiciones para el estudio de 4F2hc/LAT1 mediante *cryo-EM:*

1. Determinar la estabilidad de 4F2hc/LAT1 en diferentes detergentes y lípidos.

2. Estudiar la estabilidad de 4F2hc/LAT1 en amphipol A8-35. Este polímero ha sido usado con éxito para resolver, a nivel atómico, estructuras de proteínas de membrana de metazoos mediante *cryo-EM.*

3. Analizar la heterogeneidad de 4F2hc/LAT1 solubilizado en amphipol A8-35 mediante tinción negativa.

4. Intentar la reconstrucción 3D de 4F2hc/LAT1 solubilizado en amphipol A8-35, a partir de tinción negativa, y comparar su estructura global con la obtenida a partir de las reconstrucciones 3D del 4F2hc/LAT2 humano solubilizado en detergente.

6.3 Resultados y Discusión

En el presente estudio, inicialmente, se investigó la formación de los heterodímeros 4F2hc/LAT1, 4F2hc/LAT2 y rBAT/$b^{0,+}$AT de un vertebrado, así como su estabilidad, en diferentes sistemas de expresión.

6.3.1 Expresión de 4F2hc/LAT1, 4F2hc/LAT2 y rBAT/$b^{0,+}$AT en células humanas y de insecto.

Las células de mamífero y de insecto, se han usado ampliamente para obtener la estructura de proteínas de membrana de metazoos por cristalografía 3D (Standfuss *et al.* 2007; Feng *et al.* 2010; Gruswitz *et al.* 2010; Baconguis and Gouaux 2012; Levin *et al.* 2012; Shintre *et al.* 2013; Deng *et al.* 2014; Sun *et al.* 2014; Penmatsa *et al.* 2015), y por cryo-EM (Liao *et al.* 2013; Bai *et al.* 2015). Inicialmente, las subunidades ligeras y pesadas fueron expresadas independientemente con la GFP fusionada al C- o al N- terminal en la línea celular humana tsA210. El objetivo fue determinar la habilidad de estas células para expresar los diferentes componentes del heterodímero, y poder evaluar su integridad y estabilidad por *Fluorescence Size Exclusion Chromatography* (FSEC) (Kawate and Gouaux 2006).

Respecto a las subunidades pesadas, rBAT solo pudo ser expresada con la GFP al N-terminal (GFP-rBAT), y 4F2hc con la GFP al C-terminal (4F2hc-GFP). La FSEC de ambas proteínas mostró proteólisis. Sin embargo, el pico correspondiente a 4F2hc-GFP fue mejor que para GFP-rBAT. Por otro lado, la expresión de LAT1, LAT2 y b0,+AT, fusionadas con la GFP al N- o al C-terminal, resultó en picos de proteína proteolizada. Por esta razón, para investigar la formación de heterodímero, LAT1, LAT2 y b0,+AT fueron co-expresadas en células tsA201, sin GFP, con su correspondiente subunidad pesada marcada con GFP: GFP-rBAT/b^{0+}AT, 4F2hc-GFP/LAT1 y 4F2hc-GFP/LAT2. GFP-rBAT/b^{0+}AT no se detectó por FSEC. Por otro lado, la expresión de 4F2hc-GFP/LAT1 y 4F2hc-GFP/LAT2 produjo un cambio en el volumen de elución, respecto al pico de 4F2hc-GFP (Figura 3A y 3B), indicando la generación de ambos heterodímeros. No obstante, se continúa observando una elevada proteólisis. Esto sugiere que estas células, tsA201, no son apropiadas para expresar estos heterodímeros usando como marcador para monitorizar su estabilidad la GFP.

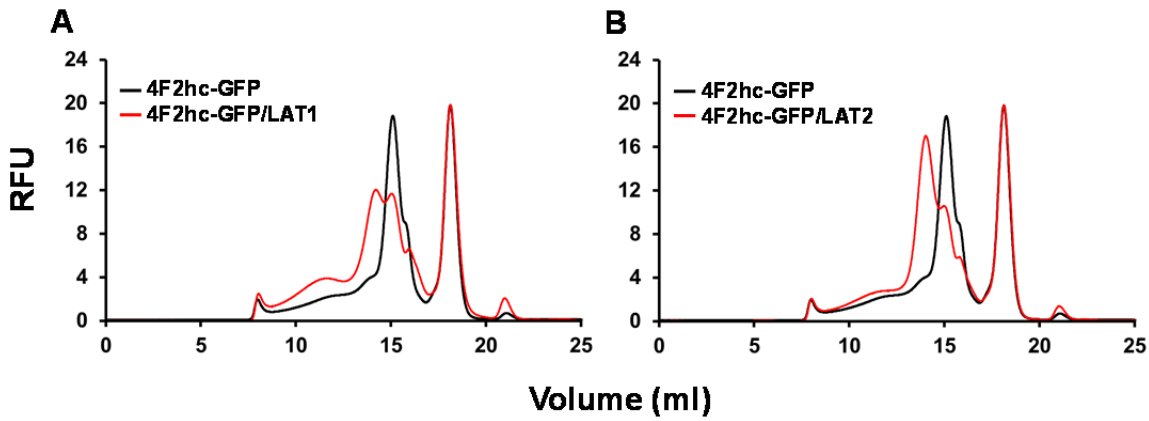

Figura 3. Formación de los heterodímeros 4F2hc-GFP/LAT1 y 4F2hc-GFP/LAT2 en células tsA201. FSECs correspondientes a la co-expresión de 4F2hc-GFP con LAT1 (**A**) o con LAT2 (**B**). Se genera un cambio en la elución del pico de 4F2hc-GFP por la formación de heterodímero con la respectiva subunidad ligera. RFU: Unidades relativas de Fluorescencia.

Por esta razón, se probó la expresión de estos heterodímeros en células de insecto (línea celular Sf9). La expresión de las subunidades pesadas (rBAT y 4F2hc), y ligeras (b0,+AT, LAT1 y LAT2), también se probó independientemente y con la GFP al N- y al C- terminal. A continuación, se testó la formación de heterodímero entre las subunidades que se expresaron mejor con GFP, y sus correspondientes parejas sin GFP. Al igual que en las células tsA201, solamente la expresión de rBAT con la GFP al N-terminal (GFP-rBAT) y 4F2hc con la GFP al C-terminal (4F2hc-GFP), funcionó. Ambas proteínas mostraron mejor comportamiento en FSEC al ser expresadas en Sf9, respecto a tsA201. De igual forma, 4F2hc-GFP continua siendo más estable en solución que GFP-rBAT.

Respecto a las subunidades ligeras, LAT1, LAT2 y $b^{0,+}$AT, únicamente pudieron ser obtenidas con la GFP al C-terminal (LAT1-GFP, LAT2-GFP y $b^{0,+}$AT-GFP). Todas mostraron alta estabilidad en solución al ser analizadas por FSEC. Sin embargo, los picos de LAT1-GFP y $^{b0,+}$AT-GFP fueron más estrechos y simétricos, indicando mayor monodispersidad respecto a LAT2-GFP.

Para intentar obtener el heterodímero rBAT/$b^{0,+}$AT, se usaron las versiones GFP-rBAT y $b^{0,+}$AT-GFP, co-expresándose cada una con su respectiva subunidad sin GFP: rBAT/$b^{0,+}$AT-GFP y GFP-rBAT/$b^{0,+}$AT. Sin embargo, el heterodímero no se formó en ninguna de las dos combinaciones. Por otro lado, la co-expresión en células Sf9 de 4F2hc-GFP con LAT1 y con LAT2, resultó en la generación de ambos heterodímeros (Figura 4). Sin embargo, en células de insecto, se observó una menor cantidad de fragmentos proteolíticos en comparación con las células tsA201.

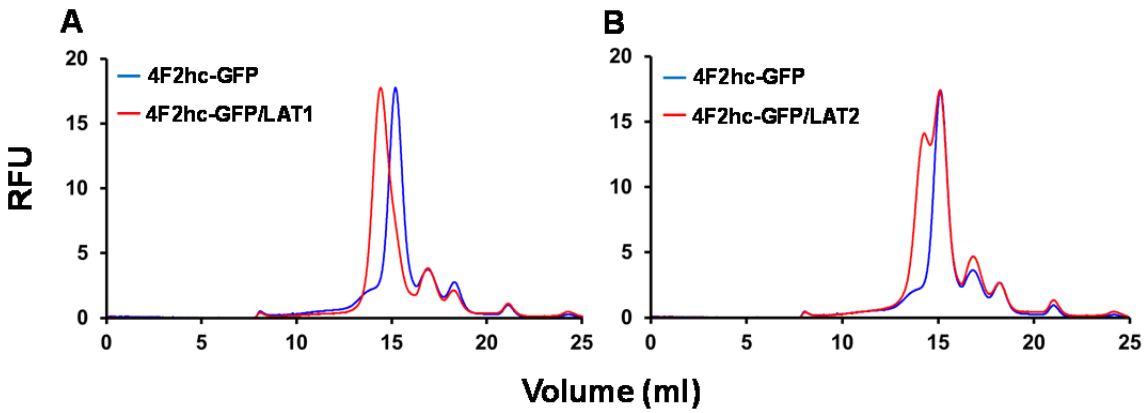

Figura 4. Formación de los heterodímeros 4F2hc-GFP/LAT1 y 4F2hc-GFP/LAT2 en células de insecto. FSEC correspondientes a la co-expresión de 4F2hc-GFP con LAT1 (**A**) o con LAT2 (**B**). Se genera un cambio en la elución del pico de 4F2hc-GFP por la formación de heterodímero con la respectiva subunidad ligera. RFU: Unidades relativas de Fluorescencia.

6.3.1.1 4F2hc/LAT1 es más estable que 4F2hc/LAT2

Para determinar cuál de los dos heterodímeros, 4F2hc/LAT1 o 4F2hc/LAT2, era el más estable, se realizó un ensayo de termoestabilidad utilizando sus versiones marcadas con GFP (4F2hc-GFP/LAT1 y 4F2hc-GFP/LAT2). Ambos heterodímeros purificados fueron sometidos a diferentes temperaturas durante 10 minutos, y la fracción soluble obtenida después de ultracentrifugar sometida a FSEC, observándose que 4F2hc-GFP/LAT1 era más termoestable que 4F2hc-GFP/LAT2 (Figura 5). Se determinó la *melting temperature* de 4F2hc-GFP/LAT1, siendo esta de 65°C, la cual es alta para una proteína de membrana, indicando una gran estabilidad de este heterodímero.

Debido a que diferentes resultados han señalado a 4F2hc/LAT1 como el heterodímero más estable, decidimos seleccionar este complejo para intentar resolver su estructura.

6.3.1.2 4F2hc-GFP/LAT1 permanece estable después de una doble purificación por cromatografía de afinidad

Para investigar el rendimiento y calidad finales de 4F2hc-GFP/LAT1 expresado en células de insecto, el heterodímero se purificó, primero, mediante cromatografía de afinidad a strep-tactin, y el eluído resultante sometido a una cromatografía de afinidad por iones metálicos inmovilizados (IMAC por sus siglas en inglés). El heterodímero purificado mostró una alta estabilidad durante todo el proceso, mostrando los picos de FSEC una elevada monodispersidad de la muestra (Figura 6 A-D). La pureza e integridad del heterodímero purificado también se confirmó por western blot y tinción por coomassie (Figura 6E y 6F).

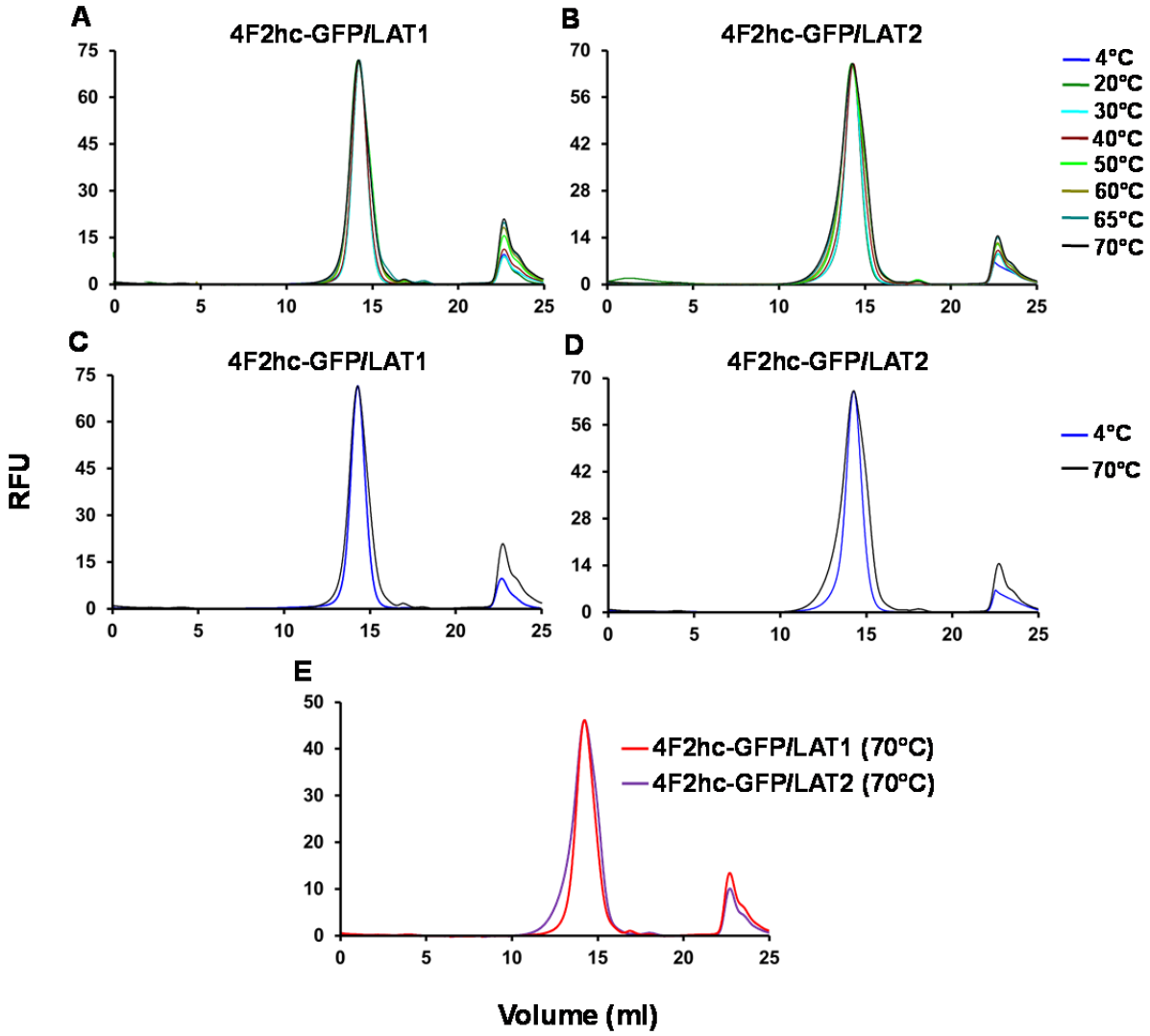

Figura 5. 4F2hc-GFP/LAT1 es más termoestable que 4F2hc-GFP/LAT2. Membranas de células Sf9 expresando ambos heterodímeros fueron solubilizadas en 1 % DDM / 0.2 % CHS, y la fracción soluble purificada por *strep-tactin affinity chromatography* en presencia de 0.05 % DDM /0.01 % CHS. Alícuotas de ambas proteínas, a la misma concentración, fueron sometidas independientemente a las temperaturas mostradas en la figura durante 10 min y luego analizadas por FSEC. Los picos de 4F2hc/LAT1 son bastante similares en todas las temperaturas, sugiriendo mayor estabilidad en solución que 4F2hc/LAT2, el cual mostró mayor deformación del pico a medida que la temperatura era mayor. RFU: Unidades relativas de Fluorescencia.

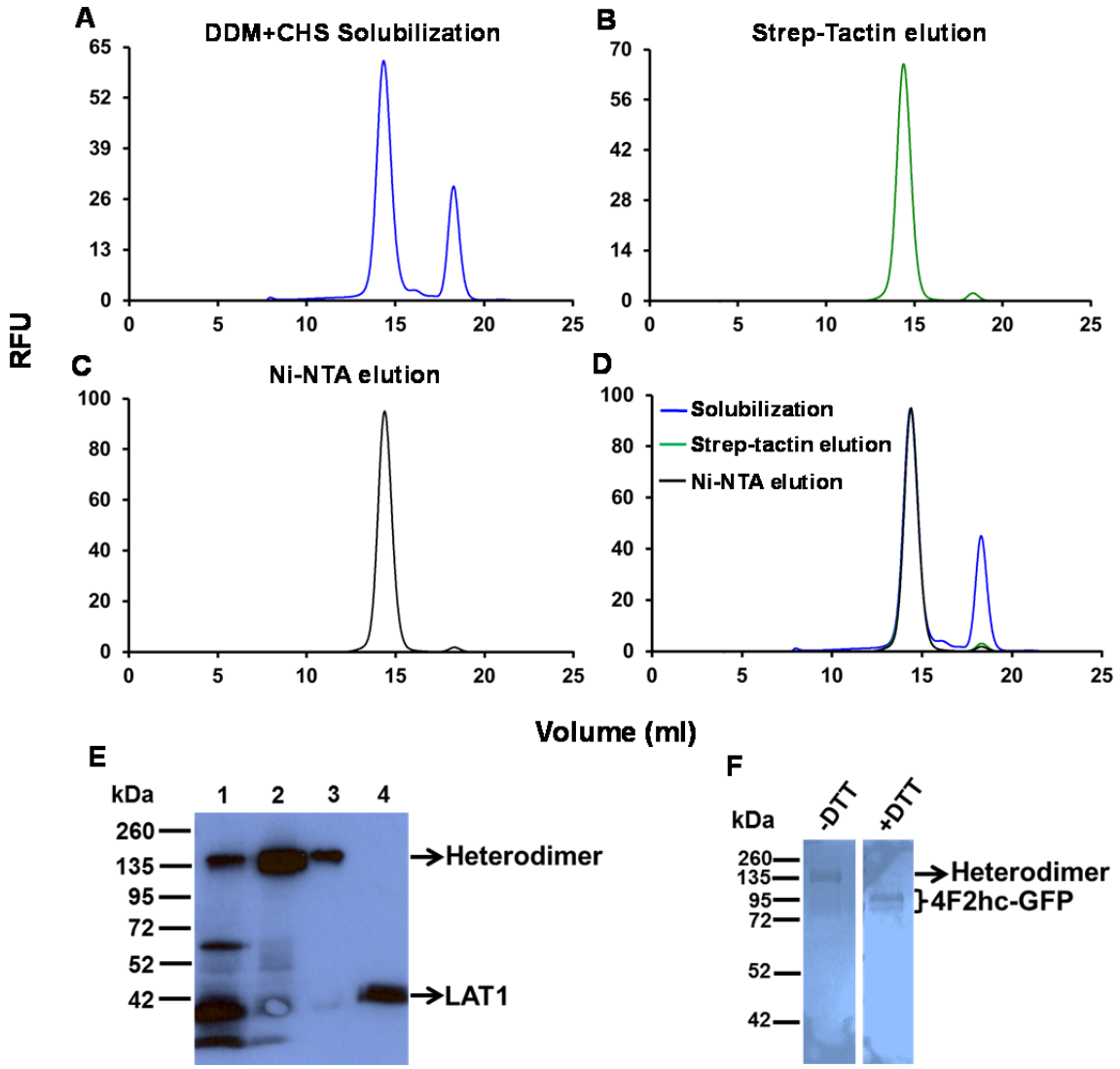

Figura 6. Purificación de 4F2hc-GFP/LAT1 mediante dos cromatografías de afinidad secuenciales. (**A**) Fracción soluble de las membranas solubilizadas en DDM/CHS. (**B**) Elución de la cromatografía de strep-tactin. (**C**) Elución de la cromatografía de afinidad a iones metálicos inmovilizados (IMAC, usando resina NI-NTA). (**D**) Se confirma la formación del heterodímero y se observa su alta estabilidad después de ser purificada. (**E**) Western blot α-Strep-tag II de los tres pasos de purificación mostrados de A a C (líneas 1 a 3), y de la muestra 3 después de adición de DTT. (**F**) Coomassie de la muestra del panel C (3) con y sin adición de DTT. RFU: Unidades relativas de Fluorescencia.

6.3.2 Expresión de 4F2hc/LAT1 en *Pichia pastoris*

Pichia pastoris es uno de los sistemas de expresión más usados, para producir proteínas integrales de membrana eucariotas para fines estructurales (Cereghino and Cregg 2000; He *et al.* 2014) (Horsefield *et al.* 2008; Ho *et al.* 2009; Shimamura *et al.* 2011; Brohawn *et al.* 2012; Miller and Long 2012). Adicionalmente, 4F2hc/LAT2 humano pudo ser expresado en *Pichia pastoris*, permitiendo estudios estructurales y funcionales (Costa *et al.* 2013; Meury *et al.* 2014; Rosell *et al.* 2014), por lo cual es factible que produzca también en forma activa 4F2hc/LAT1.

6.3.2.1 Estabilidad y actividad de 4F2hc/LAT1 expresado en Pichia

4F2hc/LAT1 sin GFP se expresó en Pichia, y se purificó por IMAC y por cromatografía de afinidad a strep-tactin. Los perfiles de SEC mostraron que el heterodímero es bastante estable después de ser concentrado incluso a 2 mg/ml, y también después de ser congelado a -80°C (Figura 7). La muestra de heterodímero purificado mostró una alta pureza e integridad (Figura 7B). Adicionalmente, la obtención de proteína después de la purificación fue de 0,16 a 0,5 mg por L de cultivo, es decir, de 2 a 6 veces más alta que en células de insecto.

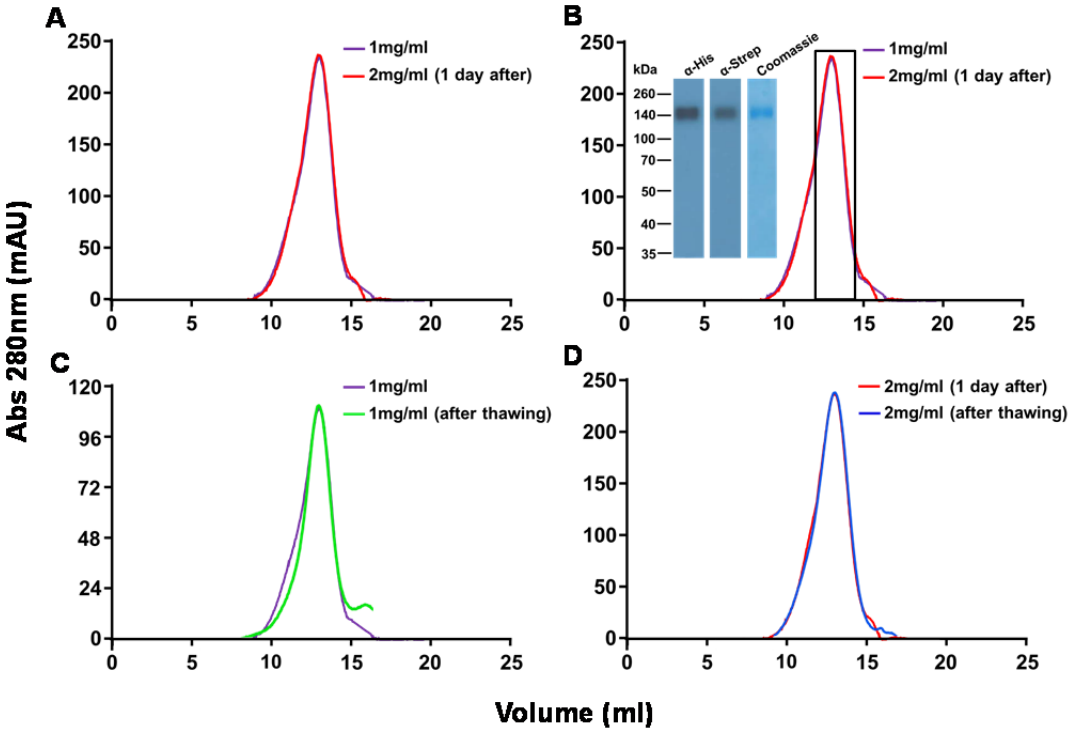

Figura 7. Perfiles de SEC de 4F2hc/LAT1 expresado en *Pichia pastoris*. (**A** y **B**) El heterodímero, después de las dos cromatografías de afinidad para su purificación, muestra buena estabilidad y alta pureza (en DDM/CHS), a dos diferentes concentraciones y al paso del tiempo, aunque se observa una moderada tendencia a la agregación. (**C** y **D**) 4F2hc/LAT1 es estable en DDM/CHS, incluso después de ser congelado. Esto es mostrado por la SEC de las fracciones del pico en B (mostradas dentro del rectángulo), que fueron congeladas con nitrógeno líquido y concentradas a la concentración original después de su descongelación.

Por otro lado, expresando y purificando 4F2hc/LAT2 humano, de igual forma que 4F2hc/LAT1, se demostró que el nuevo heterodímero es más estable (Figura 8). 4F2hc/LAT2 humano fue el heterodímero con el cual se obtuvo el modelo a baja resolución (21Å) (Rosell *et al.* 2014), citado en la introducción, pero que no fue suficientemente estable como para proseguir con estudios estructurales. Este resultado es relevante, ya que valida la selección de 4F2hc/LAT1 como heterodímero más estable.

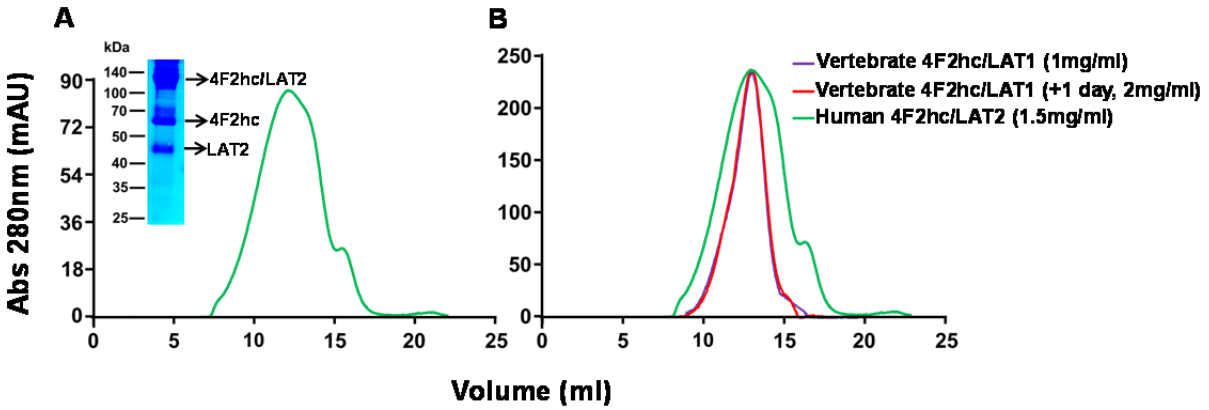

Figura 8. Comparación de la estabilidad de 4F2hc/LAT1 y 4F2hc/LAT2 humano. Los dos heterodímeros fueron expresados y purificados en las mismas condiciones. (**A**) se observa un pico que indica gran inestabilidad de 4F2hc/LAT2 humano, se observa también la disrupción del heterodímero en sus subunidades. (**B**) Se observa una significativa mayor estabilidad de 4F2hc/LAT1 frente a 4F2hc/LAT2 humano.

Adicionalmente, por "Scintillation Proximity Assay" (SPA), LAT1 mostró unión a sustrato y al inhibidor de LAT1 humano KYT-0353 (Oda *et al.* 2010) (Figura 9). También fue detectada unión al sustrato en el caso de 4F2hc/LAT1, aunque más modesta. Ambos resultados, estabilidad y función, sugieren que este heterodímero es un buen candidato para estudios estructurales.

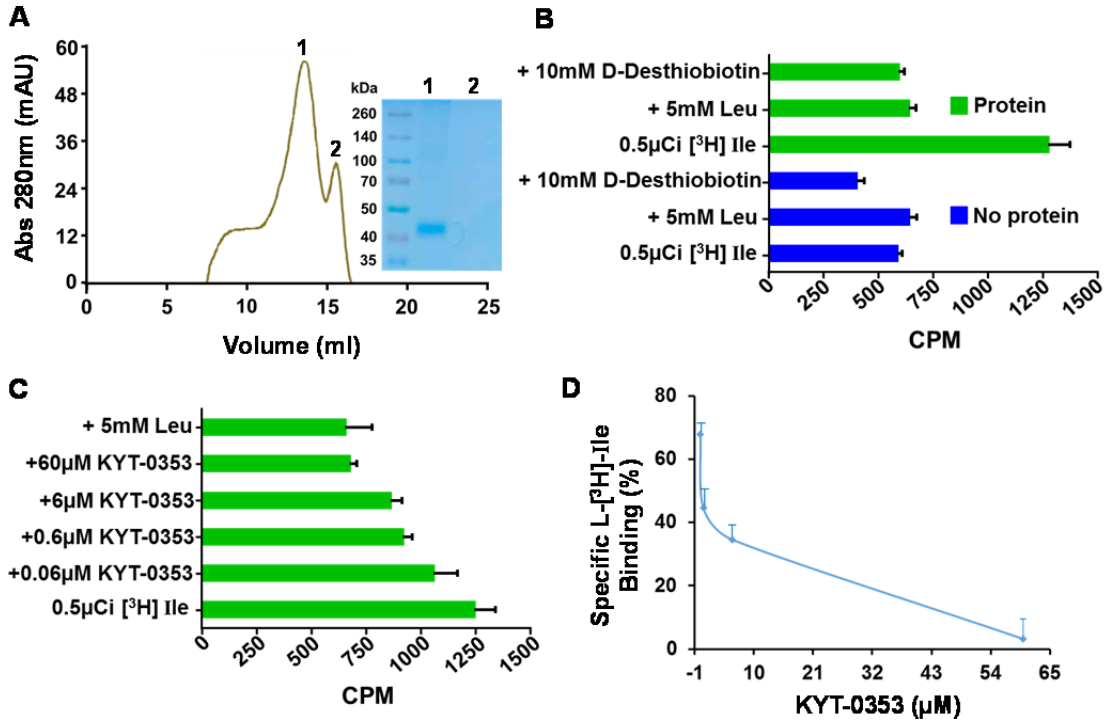

Figura 9. LAT1 purificado une sustrato e inhibidor de LAT1 humano. Las propiedades de unión de LAT1 fueron estudiadas por "Scintillation Proximity Assay" (SPA). LAT1 fue purificado de membranas de Pichia en DDM/CHS e incubado con L-[^{3}H]-Isoleucina, cuya unión a LAT1 fue desplazada por leucina fría e inhibidor KYT-0353, en una forma dependiente de la concentración, mostrando especificidad. Esto se evidencia por la disminución de los conteos por minuto (CPM) en presencia de sustrato frío (leucina) o inhibidor (KYT-0353).

6.3.3 4F2hc/LAT1 podría ser un buen candidato para resolver su estructura por cryo- EM

Debido a los recientes avances en los microscopios electrónicos y en los detectores directos de electrones, ha sido posible, en los últimos años, resolver la estructura atómica de proteínas de membrana eucariotas de menor peso molecular de lo que era posible anteriormente (Liao *et al.* 2013), (Bai *et al.* 2015) (Du *et al.* 2015). La buena estabilidad de 4F2hc/LAT1, el rango de concentración en el que puede ser obtenido de manera estable, así como los avances tecnológicos antes citados, convierten a este heterodímero en un buen candidato para resolver la primera estructura atómica de un HAT por *cryo-EM*. En los trabajos citados anteriormente, donde lograron la resolución a nivel atómico de proteínas de membrana de bajo peso molecular, utilizaron como agente solubilizador de sus proteínas el amphipol A8-35.

Este polímero anfipático se caracteriza por mantener más estables las proteínas de membrana, así como por evitar el bajo contraste entre la proteína y el medio en las muestras de cryo-EM. Por estas razones, decidimos probar la estabilidad de nuestro heterodímero solubilizado en este polímero. Para hacer esto, el heterodímero fue purificado en presencia de DDM/CHS, y finalmente, intercambiado el detergente por el amphipol A8-35. La calidad de 4F2hc/LAT1 fue analizada por SEC, mostrándose una buena estabilidad del heterodímero solubilizado en esta molécula (Figura 10).

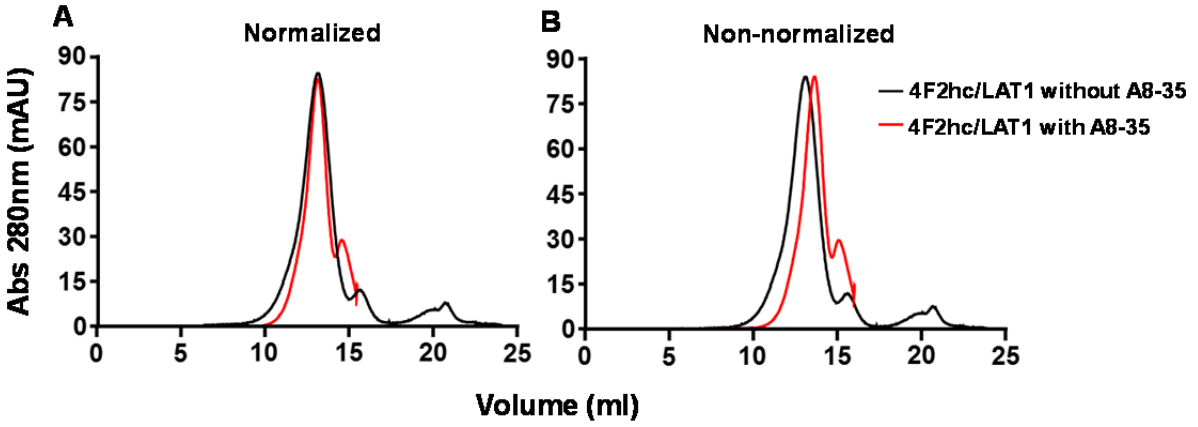

Figura 10. 4F2hc/LAT1 es estable después de su solubilización en amphipol A8-35. Heterodímero purificado en 0.05 % DDM / 0.01 % CHS, fue sometido a intercambio de DDM por amphipol A8-35. El pico de la proteína solubilizada en amphipol mostró buena estabilidad y monodispersidad, aparentemente más que cuando se encontraba solubilizado en DDM/CHS (**A**). Se observa un cambio en la elución del heterodímero, lo que podría indicar el intercambio de las micelas de DDM por las partículas de amphipol (de menor peso molecular que el DDM).

Single particle analysis de muestras del heterodímero solubilizado en A8-35, mostró partículas formadas por dos cuerpos (Figura 11), similares a las descritas anteriormente para 4F2hc/LAT2 humano (Meury *et al.* 2014; Rosell *et al.* 2014). Sin embargo, en nuestra muestra, se observó heterogeneidad en el tamaño de la partícula grande, que corresponde con LAT1. Esto debe ser mejorado para poder resolver la estructura de este heterodímero por cryo-EM, ya que dificultaría la selección y clasificación de las partículas en subgrupos con similares características estructurales, y su alineamiento para la reconstrucción 3D a alta resolución. Es probable que esta heterogeneidad se deba a que el intercambio DDM/amphipol no sea total y/o a que diferentes cantidades de CHS y/o lípidos endógenos de membrana permanezcan unidos en

forma diferencial a 4F2hc/LAT1 después del intercambio. Lo anterior, podría deberse a la presencia del CHS en la muestra, que podría evitar la delipidación homogénea de la proteína y/o ser parcialmente desalojado de las TMDs de la proteína causando heterogeneidad entre las partículas de 4F2hc/LAT1. De hecho, en estudios resientes en los cuales se resolvió la estructura del canal TRPV2 solubilizado en amphipol y purificado en presencia de CHS (Zubcevic *et al.* 2016), la concentración de amphipol es 3 veces mayor a la usada por nosotros. En los otros trabajos con proteínas de membrana de tamaño cercano a nuestro 4F2hc/LAT2, y con proteína solubilizada en amphipol, no se usó CHS, y la ratio proteína/amphipol fue la misma usada por nosotros. Esto sugiere que la presencia de CHS podría afectar el intercambio del detergente por el amphipol, requiriéndose una mayor ratio amphipol/proteína para una solubilización en amphipol más eficiente. De tal forma que en futuros experimentos el intercambio DDM/CHS debe ser optimizado para intentar reducir la heterogeneidad, así mismo como otras condiciones (e.g., otros lípidos y/o detergentes para la purificación y solubilización) podrían ser probadas para obtener una unión mas homogénea de lípidos y CHS a la proteína, lo que también podría mejorar la heterogeneidad y el intercambio detergente/amphipol.

A pesar de la heterogeneidad encontrada, un grupo homogéneo de partículas fue usado para construir un modelo 3D inicial, que se usó como referencia para clasificar estas partículas en tres clases y realizar una reconstrucción 3D de cada una de ellas (Figura 12A). La clase de partículas 3, fue la más abundante en la muestra y la que mejor se definía estructuralmente, por lo cual se refinó, obteniéndose un nuevo mapa 3D (Figura 12C). En comparación con los modelos anteriores (4F2hc/LAT2 humano) (Meury *et al.* 2014; Rosell *et al.* 2014), el presente modelo parece mostrar algo más de definición (Figura 13). Esto indicaría el potencial del amphipol A8-35 en la obtención de imágenes estructurales mejor definidas de 4F2hc/LAT1 por cryo-EM. Adicionalmente, el presente modelo también parece ser más voluminoso que los anteriores modelos de 4F2hc/LAT2 humano, lo cual pude ser debido a la unión del amphipol a las TMDs en lugar del detergente (Figura 13A), y/o estar reflejando un contenido diferente de, DDM, CHS, lípidos endógenos y amphipol entre las partículas de 4F2hc/LAT1, como ha sido discutido en el párrafo anterior. Los principales objetivos que nos marcamos a partir de ahora son: bajar los niveles de heterogeneidad de la muestra para incrementar las opciones de obtener una estructura a baja resolución, y demostrar de manera más robusta, mediante reconstitución en proteoliposomas, la funcionalidad de 4F2hc/LAT1.

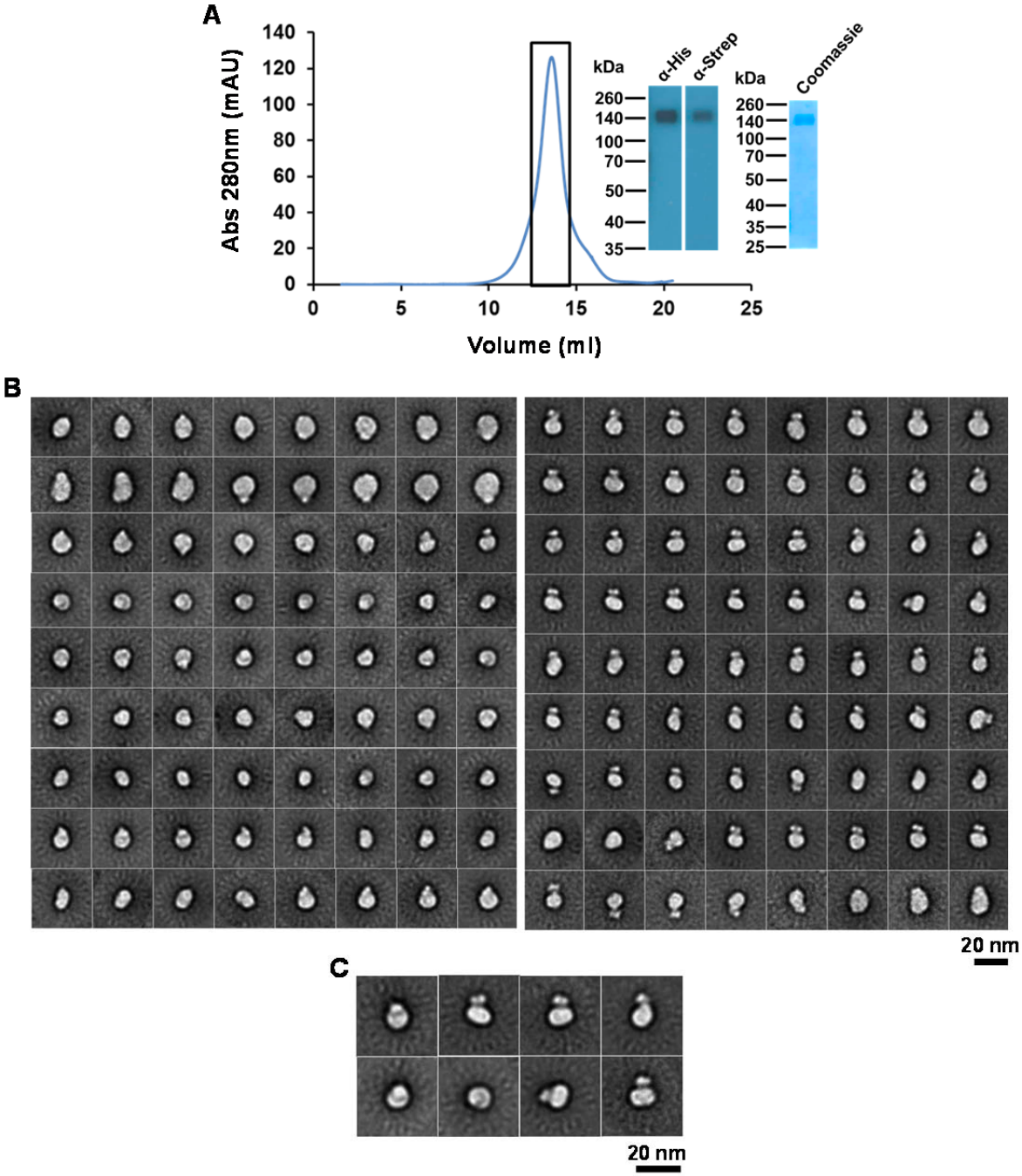

Figure 11. Tinción negativa de 4F2hc/LAT1 de vertebrado solubilizado en amphipol A8-35. Las fracciones del pico mostradas dentro del rectángulo en A fueron tomadas para hacer la tinción negativa. Estas fracciones mostraron gran pureza e integridad del heterodímero (**A**). Promedios de clase 2D libres de referencia de las partículas en las micrografías de la tinción negativa de 4F2hc/LAT1 (**B**), mostrando partículas formadas por dos densidades correspondientes a 4F2hc (densidad más pequeña) y a LAT1 (densidad más grande). Se observa heterogeneidad principalmente en la densidad grande (LAT1). Una selección de promedios de clase 2D que mostraron mas detalles estructurales después de la reclasificación se muestra en mayor tamaño en el panel **C**. Aparentemente, no se observan subunidades libres que no forman heterodímero.

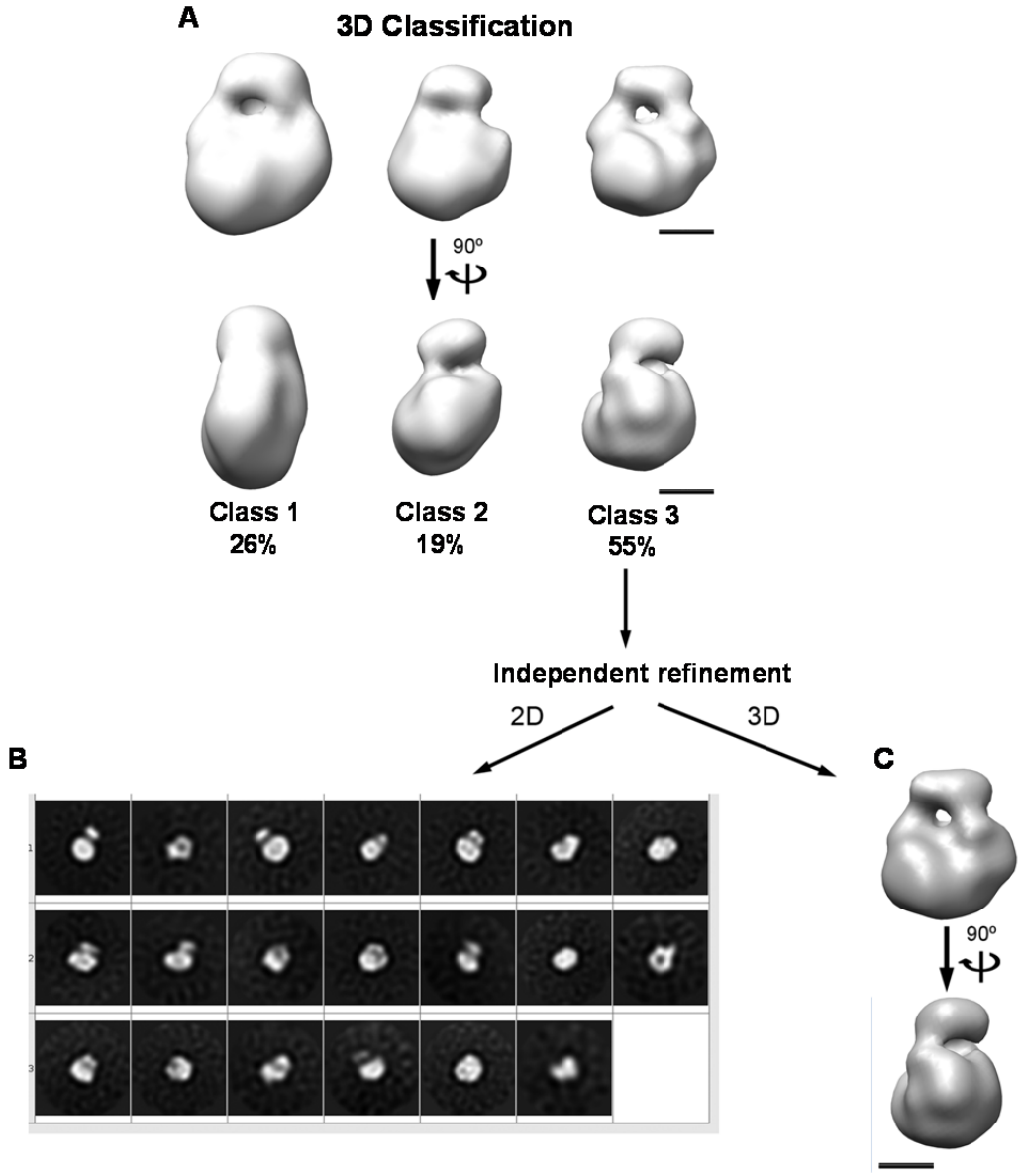

Figure 12. Reconstrucción 3D de 4F2hc/LAT1 en amphipol A8-35. Una reconstrucción 3D inicial se obtuvo a partir de los promedios de clase 2D de las partículas de 4F2hc/LAT1 solubilizadas en amphipol A8 -35. Usando como referencia este modelo, un grupo homogéneo de partículas se clasifico en tres grupos (classes) **(A)**. Se indica el porcentaje de partículas en cada clase en la muestra. La case mas abundante y la que parecía estar mejor definida estructuralmente (class 3) se refinó en 2D (B) y en 3D (C).

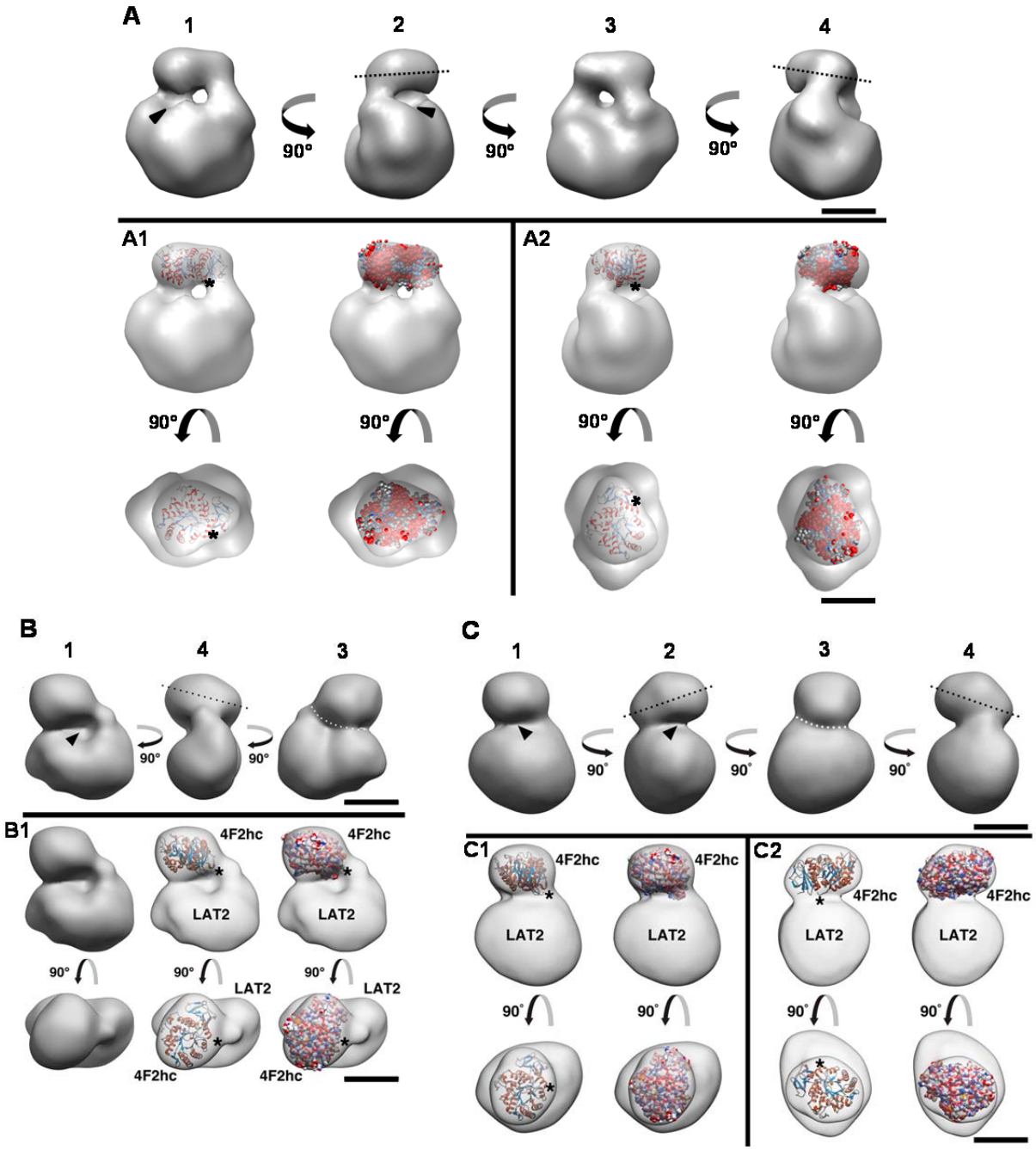

Figura 13. Comparación de los mapas 3D correspondientes a 4F2hc/LAT1 y 4F2hc/LAT2 humano. La reconstrucción 3D de 4F2hc/LAT1 solubilizado en amphipol A8-35 (**A**), se obtuvo a partir del análisis de partículas discretas visualizadas por TEM. Las reconstrucciones 3D de 4F2hc/LAT2 humano corresponden a proteína purificada en DDM (**B**) (Rosell et al., 2014) o en DDM/CHS/LMNG (**C**) (Meury et al., 2014). Para facilitar su comparación, la reconstrucción 3D de 4F2hc/LAT1 (**A**) se muestra en las mismas orientaciones que 4F2hc/LAT2 (**B**, **C**), orientaciones 1, 2, 3 o 4. La estructura de 4F2hc-ED humano se ajustó bien a la densidad pequeña del mapa 3D de 4F2hc/LAT1 (**A1**, **A2**). Este resultado indica que esta densidad corresponde a la subunidad pesada del heterodímero, y la densidad grande a LAT1, como también ha sido descrito para 4F2hc/LAT2 (**B1**, **C1**). La localización del extremo N-terminal de 4F2hc-ED está indicada con un (*), mostrando una orientación comparable de 4F2hc en todos los mapas (**A**, **B**, **C**), y una localización próxima a la teórica posición del TMD de 4F2hc. En los mapas previos de 4F2hc/LAT2, la densidad de 4F2hc está inclinada, indicado por una línea negra discontinua (**B**, **C**). En cambio esta inclinación es mucho menor en la reconstrucción 3D de

4F2hc/LAT1 (A). En esta última además parece haber más definición que en los anteriores modelos de 4F2hc/LAT2. Por ejemplo, se observa una cavidad en el centro de LAT1 en la interface con 4F2hc (puntas de flecha en A), que no se observa de igual forma en los otros modelos, y que podría estar relacionada con el sitio extracelular de unión a sustrato. Sin embargo, debido a lo preliminar de esta reconstrucción y a su baja resolución no se pueden hacer aceveraciones. Todos los mapas están a una misma escala, representando la barra 5 nm.

7. Bibliography

Bibliography

Althoff, T., R. E. Hibbs, S. Banerjee andE. Gouaux (2014). X-ray structures of GluCl in apo states reveal a gating mechanism of Cys-loop receptors. *Nature,* 512(7514): 333-337.

Althoff, T., D. J. Mills, J. L. Popot andW. Kuhlbrandt (2011). Arrangement of electron transport chain components in bovine mitochondrial supercomplex I1III2IV1. *EMBO J,* 30(22): 4652-4664.

Álvarez-Marimon., M. E. (2014). 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. Doctoral thesis, Universitat de Barcelona.

Andersen, C. B., M. Madsen, T. Storm, S. K. Moestrup andG. R. Andersen (2010). Structural basis for receptor recognition of vitamin-B(12)-intrinsic factor complexes. *Nature,* 464(7287): 445-448.

Aseervatham, J., L. Tran, K. Machaca andO. Boudker (2015). The Role of Flexible Loops in Folding, Trafficking and Activity of Equilibrative Nucleoside Transporters. *PLoS One,* 10(9): e0136779.

Auld, D. S., M. W. Farmen, S. D. Kahl, A. Kriauciunas, K. L. McKnight, C. Montrose andJ. R. Weidner (2004). Receptor Binding Assays for HTS and Drug Discovery. Assay Guidance Manual. G. S. Sittampalam, N. P. Coussens, H. Nelson*et al.* Bethesda (MD).

Baconguis, I. andE. Gouaux (2012). Structural plasticity and dynamic selectivity of acid-sensing ion channel-spider toxin complexes. *Nature,* 489(7416): 400-405.

Bai, X. C., C. Yan, G. Yang, P. Lu, D. Ma, L. Sun, R. Zhou, S. H. Scheres andY. Shi (2015). An atomic structure of human gamma-secretase. *Nature,* 525(7568): 212-217.

Baker, D. A., K. McFarland, R. W. Lake, H. Shen, X. C. Tang, S. Toda andP. W. Kalivas (2003). Neuroadaptations in cystine-glutamate exchange underlie cocaine relapse. *Nat Neurosci,* 6(7): 743-749.

Bannai, S. (1984). Transport of cystine and cysteine in mammalian cells. *Biochim Biophys Acta,* 779(3): 289-306.

Bannai, S. andN. Tateishi (1986). Role of membrane transport in metabolism and function of glutathione in mammals. *J Membr Biol,* 89(1): 1-8.

Barilli, A., B. M. Rotoli, R. Visigalli, O. Bussolati, G. C. Gazzola, Z. Kadija, G. Rodi, F. Mariani, M. L. Ruzza, M. Luisetti andV. Dall'Asta (2010). In Lysinuric Protein Intolerance system y+L activity is defective in monocytes and in GM-CSF-differentiated macrophages. *Orphanet J Rare Dis,* 5: 32.

Bartoccioni, P., C. Del Rio, M. Ratera, L. Kowalczyk, J. M. Baldwin, A. Zorzano, M. Quick, S. A. Baldwin, J. L. Vazquez-Ibar andM. Palacin (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.

Bartoccioni, P., M. Rius, A. Zorzano, M. Palacin andJ. Chillaron (2008). Distinct classes of trafficking rBAT mutants cause the type I cystinuria phenotype. *Hum Mol Genet,* 17(12): 1845-1854.

Bassi, M. T., E. Gasol, M. Manzoni, M. Pineda, M. Riboni, R. Martin, A. Zorzano, G. Borsani andM. Palacin (2001). Identification and characterisation of human xCT that co-expresses, with 4F2 heavy chain, the amino acid transport activity system xc. *Pflugers Arch,* 442(2): 286-296.

Bauch, C., N. Forster, D. Loffing-Cueni, V. Summa andF. Verrey (2003). Functional cooperation of epithelial heteromeric amino acid transporters expressed in madin-darby canine kidney cells. *J Biol Chem,* 278(2): 1316-1322.

Bauch, C. andF. Verrey (2002). Apical heterodimeric cystine and cationic amino acid transporter expressed in MDCK cells. *Am J Physiol Renal Physiol,* 283(1): F181-189.

Berger, B. W., D. W. Kulp, L. M. Span, J. L. DeGrado, P. C. Billings, A. Senes, J. S. Bennett andW. F. DeGrado (2010). Consensus motif for integrin transmembrane helix association. *Proc Natl Acad Sci U S A,* 107(2): 703-708.

Bernaudat, F., A. Frelet-Barrand, N. Pochon, S. Dementin, P. Hivin, S. Boutigny, J. B. Rioux, D. Salvi, D. Seigneurin-Berny, P. Richaud, J. Joyard, D. Pignol, M. Sabaty, T. Desnos, E. Pebay-Peyroula, E. Darrouzet, T. Vernet andN. Rolland (2011). Heterologous expression of membrane proteins: choosing the appropriate host. *PLoS One,* 6(12): e29191.

Bertran, J., A. Werner, M. L. Moore, G. Stange, D. Markovich, J. Biber, X. Testar, A. Zorzano, M. Palacin andH. Murer (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 U S A,* 89(12): 5601-5605.

Bhatnagar, R. S. andJ. I. Gordon (1997). Understanding covalent modifications of proteins by lipids: where cell biology and biophysics mingle. *Trends Cell Biol,* 7(1): 14-20.

Binshtein, E. andM. D. Ohi (2015). Cryo-electron microscopy and the amazing race to atomic resolution. *Biochemistry,* 54(20): 3133-3141.

Borsani, G., M. T. Bassi, M. P. Sperandeo, A. De Grandi, A. Buoninconti, M. Riboni, M. Manzoni, B. Incerti, A. Pepe, G. Andria, A. Ballabio andG. Sebastio (1999). SLC7A7, encoding a putative permease-related protein, is mutated in patients with lysinuric protein intolerance. *Nat Genet,* 21(3): 297-301.

Borst, P. andR. O. Elferink (2002). Mammalian ABC transporters in health and disease. *Annu Rev Biochem,* 71: 537-592.

Braun, D., E. K. Wirth, F. Wohlgemuth, N. Reix, M. O. Klein, A. Gruters, J. Kohrle andU. Schweizer (2011). Aminoaciduria, but normal thyroid hormone levels and signalling, in mice lacking the amino acid and thyroid hormone transporter Slc7a8. *Biochem J,* 439(2): 249-255.

Broer, A., N. Brookes, V. Ganapathy, K. S. Dimmer, C. A. Wagner, F. Lang andS. Broer (1999). The astroglial ASCT2 amino acid transporter as a mediator of glutamine efflux. *J Neurochem,* 73(5): 2184-2194.

Broer, A., C. Wagner, F. Lang andS. Broer (2000). Neutral amino acid transporter ASCT2 displays substrate-induced Na+ exchange and a substrate-gated anion conductance. *Biochem J,* 346 Pt 3: 705-710.

Broer, S. andM. Palacin (2011). The role of amino acid transporters in inherited and acquired diseases. *Biochem J,* 436(2): 193-211.

Brohawn, S. G., J. del Marmol andR. MacKinnon (2012). Crystal structure of the human K2P TRAAK, a lipid- and mechano-sensitive K+ ion channel. *Science,* 335(6067): 436-441.

Busch, A. E., T. Herzer, S. Waldegger, F. Schmidt, M. Palacin, J. Biber, D. Markovich, H. Murer andF. Lang (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(41): 25581-25586.

Cabantous, S. andG. S. Waldo (2006). In vivo and in vitro protein solubility assays using split GFP. *Nat Methods,* 3(10): 845-854.

Cabra, V. andM. Samso (2015). Do's and don'ts of cryo-electron microscopy: a primer on sample preparation and high quality data collection for macromolecular 3D reconstruction. *J Vis Exp,*(95): 52311.

Calonge, M. J., P. Gasparini, J. Chillaron, M. Chillon, M. Gallucci, F. Rousaud, L. Zelante, X. Testar, B. Dallapiccola, F. Di Silverio andet al. (1994). Cystinuria caused by mutations in rBAT, a gene involved in the transport of cystine. *Nat Genet,* 6(4): 420-425.

Cantor, J., C. D. Browne, R. Ruppert, C. C. Feral, R. Fassler, R. C. Rickert andM. H. Ginsberg (2009). CD98hc facilitates B cell proliferation and adaptive humoral immunity. *Nat Immunol,* 10(4): 412-419.

Carpenter, E. P., K. Beis, A. D. Cameron andS. Iwata (2008). Overcoming the challenges of membrane protein crystallography. *Curr Opin Struct Biol,* 18(5): 581-586.

Casagrande, F., M. Ratera, A. D. Schenk, M. Chami, E. Valencia, J. M. Lopez, D. Torrents, A. Engel, M. Palacin andD. Fotiadis (2008). Projection structure of a member of the amino acid/polyamine/organocation transporter superfamily. *J Biol Chem,* 283(48): 33240-33248.

Cereghino, J. L. andJ. M. Cregg (2000). Heterologous protein expression in the methylotrophic yeast Pichia pastoris. *FEMS Microbiol Rev,* 24(1): 45-66.

Ciccarone, V. C., D. A. Polayes andV. A. Luckow (1998). Generation of Recombinant Baculovirus DNA in E.coli Using a Baculovirus Shuttle Vector. *Methods Mol Med,* 13: 213-235.

Connor, J. H., T. Kleeman, S. Barik, R. E. Honkanen andS. Shenolikar (1999). Importance of the beta12-beta13 loop in protein phosphatase-1 catalytic subunit for inhibition by toxins and mammalian protein inhibitors. *J Biol Chem,* 274(32): 22366-22372.

Conte, M. R. andS. Matthews (1998). Retroviral matrix proteins: a structural perspective. *Virology,* 246(2): 191-198.

Contreras-Gomez, A., A. Sanchez-Miron, F. Garcia-Camacho, E. Molina-Grima andY. Chisti (2014). Protein production using the baculovirus-insect cell expression system. *Biotechnol Prog,* 30(1): 1-18.

Cooper, G. M. (2000). The Cell: A Molecular Approach 2nd edition, Sunderland (MA): Sinauer Associates.

Cosgriff, A. J. andA. J. Pittard (1997). A topological model for the general aromatic amino acid permease, AroP, of Escherichia coli. *J Bacteriol,* 179(10): 3317-3323.

Costa i Torres, M. (2012). The first 3D structural model of an eukaryotic heteromeric amino acid transporter. Doctoral thesis, Universitat de Barcelona.

Costa, M., A. Rosell, E. Alvarez-Marimon, A. Zorzano, D. Fotiadis andM. Palacin (2013). Expression of human heteromeric amino acid transporters in the yeast Pichia pastoris. *Protein Expr Purif,* 87(1): 35-40.

Cregg, J. M., K. R. Madden, K. J. Barringer, G. P. Thill andC. A. Stillman (1989). Functional characterization of the two alcohol oxidase genes from the yeast Pichia pastoris. *Mol Cell Biol,* 9(3): 1316-1323.

Cunningham., F. M. (1994). Lipid Mediators. Academic Press. Harcourt Brace & Company.

Chae, P. S., S. G. Rasmussen, R. R. Rana, K. Gotfryd, R. Chandra, M. A. Goren, A. C. Kruse, S. Nurva, C. J. Loland, Y. Pierre, D. Drew, J. L. Popot, D. Picot, B. G. Fox, L. Guan, U. Gether, B. Byrne, B. Kobilka andS. H. Gellman (2010). Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nat Methods,* 7(12): 1003-1008.

Chillaron, J., R. Estevez, C. Mora, C. A. Wagner, H. Suessbrich, F. Lang, J. L. Gelpi, X. Testar, A. E. Busch, A. Zorzano andM. Palacin (1996). Obligatory amino acid exchange via systems bo,+-like and y+L-like. A tertiary active transport mechanism for renal reabsorption of cystine and dibasic amino acids. *J Biol Chem,* 271(30): 17761-17770.

Chillaron, J., M. Font-Llitjos, J. Fort, A. Zorzano, D. S. Goldfarb, V. Nunes andM. Palacin (2010). Pathophysiology and treatment of cystinuria. *Nat Rev Nephrol,* 6(7): 424-434.

Chillaron, J., R. Roca, A. Valencia, A. Zorzano andM. Palacin (2001). Heteromeric amino acid transporters: biochemistry, genetics, and physiology. *Am J Physiol Renal Physiol,* 281(6): F995-1018.

Christensen, H. N. (1990). Role of amino acid transport and countertransport in nutrition and metabolism. *Physiol Rev,* 70(1): 43-77.

Dall'Asta, V., O. Bussolati, R. Sala, B. M. Rotoli, G. Sebastio, M. P. Sperandeo, G. Andria andG. C. Gazzola (2000). Arginine transport through system y(+)L in cultured human fibroblasts: normal phenotype of cells from LPI subjects. *Am J Physiol Cell Physiol,* 279(6): C1829-1837.

De Bundel, D., A. Schallier, E. Loyens, R. Fernando, H. Miyashita, J. Van Liefferinge, K. Vermoesen, S. Bannai, H. Sato, Y. Michotte, I. Smolders andA. Massie (2011). Loss of system x(c)- does not induce oxidative stress but decreases extracellular glutamate in hippocampus and influences spatial working memory and limbic seizure susceptibility. *J Neurosci,* 31(15): 5792-5803.

de la Rosa-Trevin, J. M., J. Oton, R. Marabini, A. Zaldivar, J. Vargas, J. M. Carazo andC. O. Sorzano (2013). Xmipp 3.0: an improved software suite for image processing in electron microscopy. *J Struct Biol,* 184(2): 321-328.

del Amo, E. M., A. Urtti andM. Yliperttula (2008). Pharmacokinetic role of L-type amino acid transporters LAT1 and LAT2. *Eur J Pharm Sci,* 35(3): 161-174.

Demain, A. L. andP. Vaishnav (2009). Production of recombinant proteins by microbes and higher organisms. *Biotechnol Adv,* 27(3): 297-306.

Deng, D., C. Xu, P. Sun, J. Wu, C. Yan, M. Hu andN. Yan (2014). Crystal structure of the human glucose transporter GLUT1. *Nature,* 510(7503): 121-125.

Denoyer, D., L. Kirby, K. Waldeck, P. Roselt, O. C. Neels, T. Bourdier, R. Shepherd, A. Katsifis andR. J. Hicks (2012). Preclinical characterization of 18F-D-FPHCys, a new amino acid-based PET tracer. *Eur J Nucl Med Mol Imaging,* 39(4): 703-712.

Deves, R. andC. A. Boyd (1998). Transporters for cationic amino acids in animal cells: discovery, structure, and function. *Physiol Rev,* 78(2): 487-545.

Deves, R. andC. A. Boyd (2000). Surface antigen CD98(4F2): not a single membrane protein, but a family of proteins with multiple functions. *J Membr Biol,* 173(3): 165-177.

Dixon, S. J., D. N. Patel, M. Welsch, R. Skouta, E. D. Lee, M. Hayano, A. G. Thomas, C. E. Gleason, N. P. Tatonetti, B. S. Slusher andB. R. Stockwell (2014). Pharmacological inhibition of cystine-glutamate exchange induces endoplasmic reticulum stress and ferroptosis. *Elife,* 3: e02523.

Dorr, J. M., M. C. Koorengevel, M. Schafer, A. V. Prokofyev, S. Scheidelaar, E. A. van der Cruijsen, T. R. Dafforn, M. Baldus andJ. A. Killian (2014). Detergent-free isolation, characterization, and functional reconstitution of a tetrameric K+ channel: the power of native nanodiscs. *Proc Natl Acad Sci U S A,* 111(52): 18607-18612.

Dorr, J. M., S. Scheidelaar, M. C. Koorengevel, J. J. Dominguez, M. Schafer, C. A. van Walree andJ. A. Killian (2016). The styrene-maleic acid copolymer: a versatile tool in membrane research. *Eur Biophys J,* 45(1): 3-21.

Drew, D., S. Newstead, Y. Sonoda, H. Kim, G. von Heijne andS. Iwata (2008). GFP-based optimization scheme for the overexpression and purification of eukaryotic membrane proteins in Saccharomyces cerevisiae. *Nat Protoc,* 3(5): 784-798.

Du, J., W. Lu, S. Wu, Y. Cheng andE. Gouaux (2015). Glycine receptor mechanism elucidated by electron cryo-microscopy. *Nature,* 526(7572): 224-229.

Dubyak, G. R. (2004). Ion homeostasis, channels, and transporters: an update on cellular mechanisms. *Adv Physiol Educ,* 28(1-4): 143-154.

Dye, J. F., S. Vause, T. Johnston, P. Clark, J. A. Firth, S. W. D'Souza, C. P. Sibley andJ. D. Glazier (2004). Characterization of cationic amino acid transporters and expression of endothelial nitric oxide synthase in human placental microvascular endothelial cells. *FASEB J,* 18(1): 125-127.

Ellis, J., A. Carlin, C. Steffes, J. Wu, J. Liu andB. P. Rosen (1995). Topological analysis of the lysine-specific permease of Escherichia coli. *Microbiology,* 141 (Pt 8): 1927-1935.

Ellis, S. B., P. F. Brust, P. J. Koutz, A. F. Waters, M. M. Harpold andT. R. Gingeras (1985). Isolation of alcohol oxidase and two other methanol regulatable genes from the yeast Pichia pastoris. *Mol Cell Biol,* 5(5): 1111-1121.

Estevez, R., M. Camps, A. M. Rojas, X. Testar, R. Deves, M. A. Hediger, A. Zorzano andM. Palacin (1998). The amino acid transport system y+L/4F2hc is a heteromultimeric complex. *FASEB J,* 12(13): 1319-1329.

Faham, S., A. Watanabe, G. M. Besserer, D. Cascio, A. Specht, B. A. Hirayama, E. M. Wright andJ. Abramson (2008). The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na+/sugar symport. *Science,* 321(5890): 810-814.

Fang, Y., H. Jayaram, T. Shane, L. Kolmakova-Partensky, F. Wu, C. Williams, Y. Xiong andC. Miller (2009). Structure of a prokaryotic virtual proton pump at 3.2 A resolution. *Nature,* 460(7258): 1040-1043.

Feliubadalo, L., 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 andM. Palacin (1999). Non-type I cystinuria caused by mutations in SLC7A9, encoding a subunit (bo,+AT) of rBAT. *Nat Genet,* 23(1): 52-57.

Fenczik, C. A., R. Zent, M. Dellos, D. A. Calderwood, J. Satriano, C. Kelly andM. H. Ginsberg (2001). Distinct domains of CD98hc regulate integrins and amino acid transport. *J Biol Chem,* 276(12): 8746-8752.

Feng, L., E. B. Campbell, Y. Hsiung andR. MacKinnon (2010). Structure of a eukaryotic CLC transporter defines an intermediate state in the transport cycle. *Science,* 330(6004): 635-641.

Feral, C. C., N. Nishiya, C. A. Fenczik, H. Stuhlmann, M. Slepak andM. H. Ginsberg (2005). CD98hc (SLC3A2) mediates integrin signaling. *Proc Natl Acad Sci U S A,* 102(2): 355-360.

Ferguson, A. D., B. M. McKeever, S. Xu, D. Wisniewski, D. K. Miller, T. T. Yamin, R. H. Spencer, L. Chu, F. Ujjainwalla, B. R. Cunningham, J. F. Evans andJ. W. Becker (2007). Crystal structure of inhibitor-bound human 5-lipoxygenase-activating protein. *Science,* 317(5837): 510-512.

Ferguson, A. D., W. Welte, E. Hofmann, B. Lindner, O. Holst, J. W. Coulton andK. Diederichs (2000). A conserved structural motif for lipopolysaccharide recognition by procaryotic and eucaryotic proteins. *Structure,* 8(6): 585-592.

Fernandez, E., M. Carrascal, F. Rousaud, J. Abian, A. Zorzano, M. Palacin andJ. Chillaron (2002). rBAT-b(0,+)AT heterodimer is the main apical reabsorption system for cystine in the kidney. *Am J Physiol Renal Physiol,* 283(3): F540-548.

Fernandez, E., M. Jimenez-Vidal, M. Calvo, A. Zorzano, F. Tebar, M. Palacin andJ. Chillaron (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.

Fernandez, E., D. Torrents, J. Chillaron, R. Martin Del Rio, A. Zorzano andM. Palacin (2003). Basolateral LAT-2 has a major role in the transepithelial flux of L-cystine in the renal proximal tubule cell line OK. *J Am Soc Nephrol,* 14(4): 837-847.

Fogelstrand, P., C. C. Feral, R. Zargham andM. H. Ginsberg (2009). Dependence of proliferative vascular smooth muscle cells on CD98hc (4F2hc, SLC3A2). *J Exp Med,* 206(11): 2397-2406.

Fort, J., 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 andM. Palacin (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.

Fotiadis, D., Y. Kanai andM. Palacin (2013). The SLC3 and SLC7 families of amino acid transporters. *Mol Aspects Med,* 34(2-3): 139-158.

Fuchs, B. C. andB. P. Bode (2005). Amino acid transporters ASCT2 and LAT1 in cancer: partners in crime? *Semin Cancer Biol,* 15(4): 254-266.

Fuchs, B. C., R. E. Finger, M. C. Onan andB. P. Bode (2007). ASCT2 silencing regulates mammalian target-of-rapamycin growth and survival signaling in human hepatoma cells. *Am J Physiol Cell Physiol,* 293(1): C55-63.

Fukasawa, Y., H. Segawa, J. Y. Kim, A. Chairoungdua, D. K. Kim, H. Matsuo, S. H. Cha, H. Endou andY. Kanai (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(13): 9690-9698.

Ganapathy, V., M. Thangaraju andP. D. Prasad (2009). Nutrient transporters in cancer: relevance to Warburg hypothesis and beyond. *Pharmacol Ther,* 121(1): 29-40.

Gao, X., F. Lu, L. Zhou, S. Dang, L. Sun, X. Li, J. Wang andY. Shi (2009). Structure and mechanism of an amino acid antiporter. *Science,* 324(5934): 1565-1568.

Gao, X., L. Zhou, X. Jiao, F. Lu, C. Yan, X. Zeng, J. Wang andY. Shi (2010). Mechanism of substrate recognition and transport by an amino acid antiporter. *Nature,* 463(7282): 828-832.

Garavito, R. M., D. Picot andP. J. Loll (1996). Strategies for crystallizing membrane proteins. *J Bioenerg Biomembr,* 28(1): 13-27.

Gasol, E., M. Jimenez-Vidal, J. Chillaron, A. Zorzano andM. Palacin (2004). Membrane topology of system xc- light subunit reveals a re-entrant loop with substrate-restricted accessibility. *J Biol Chem,* 279(30): 31228-31236.

Giusti, F., J. L. Popot andC. Tribet (2012). Well-defined critical association concentration and rapid adsorption at the air/water interface of a short amphiphilic polymer, amphipol A8-35: a study by Forster resonance energy transfer and dynamic surface tension measurements. *Langmuir,* 28(28): 10372-10380.

Griffin, P. M. andR. V. Tauxe (1991). The epidemiology of infections caused by Escherichia coli O157:H7, other enterohemorrhagic E. coli, and the associated hemolytic uremic syndrome. *Epidemiol Rev,* 13: 60-98.

Gruswitz, F., S. Chaudhary, J. D. Ho, A. Schlessinger, B. Pezeshki, C. M. Ho, A. Sali, C. M. Westhoff andR. M. Stroud (2010). Function of human Rh based on structure of RhCG at 2.1 A. *Proc Natl Acad Sci U S A,* 107(21): 9638-9643.

Gulati, S., M. Jamshad, T. J. Knowles, K. A. Morrison, R. Downing, N. Cant, R. Collins, J. B. Koenderink, R. C. Ford, M. Overduin, I. D. Kerr, T. R. Dafforn andA. J. Rothnie (2014). Detergent-free purification of ABC (ATP-binding-cassette) transporters. *Biochem J,* 461(2): 269-278.

Guo, W., Y. Zhao, Z. Zhang, N. Tan, F. Zhao, C. Ge, L. Liang, D. Jia, T. Chen, M. Yao, J. Li andX. He (2011). Disruption of xCT inhibits cell growth via the ROS/autophagy pathway in hepatocellular carcinoma. *Cancer Lett,* 312(1): 55-61.

Gutmann, D. A., E. Mizohata, S. Newstead, S. Ferrandon, V. Postis, X. Xia, P. J. Henderson, H. W. van Veen andB. Byrne (2007). A high-throughput method for membrane protein solubility screening: the ultracentrifugation dispersity sedimentation assay. *Protein Sci,* 16(7): 1422-1428.

Haga, K., A. C. Kruse, H. Asada, T. Yurugi-Kobayashi, M. Shiroishi, C. Zhang, W. I. Weis, T. Okada, B. K. Kobilka, T. Haga andT. Kobayashi (2012). Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature,* 482(7386): 547-551.

Harder, D. andD. Fotiadis (2012). Measuring substrate binding and affinity of purified membrane transport proteins using the scintillation proximity assay. *Nat Protoc,* 7(9): 1569-1578.

Hashimoto, A. andT. Oka (1997). Free D-aspartate and D-serine in the mammalian brain and periphery. *Prog Neurobiol,* 52(4): 325-353.

Hattori, M. andE. Gouaux (2012). Molecular mechanism of ATP binding and ion channel activation in P2X receptors. *Nature,* 485(7397): 207-212.

Hattori, M., R. E. Hibbs andE. Gouaux (2012). A fluorescence-detection size-exclusion chromatography-based thermostability assay for membrane protein precrystallization screening. *Structure,* 20(8): 1293-1299.

Haynes, B. F., M. E. Hemler, D. L. Mann, G. S. Eisenbarth, J. Shelhamer, H. S. Mostowski, C. A. Thomas, J. L. Strominger andA. S. Fauci (1981). Characterization of a monoclonal antibody (4F2) that binds to human monocytes and to a subset of activated lymphocytes. *J Immunol,* 126(4): 1409-1414.

He, Y., K. Wang andN. Yan (2014). The recombinant expression systems for structure determination of eukaryotic membrane proteins. *Protein Cell,* 5(9): 658-672.

Hediger, M. A., B. Clemencon, R. E. Burrier andE. A. Bruford (2013). The ABCs of membrane transporters in health and disease (SLC series): introduction. *Mol Aspects Med,* 34(2-3): 95-107.

Hediger, M. A., M. F. Romero, J. B. Peng, A. Rolfs, H. Takanaga andE. A. Bruford (2004). The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteinsIntroduction. *Pflugers Arch,* 447(5): 465-468.

Helboe, L., J. Egebjerg, M. Moller andC. Thomsen (2003). Distribution and pharmacology of alanine-serine-cysteine transporter 1 (asc-1) in rodent brain. *Eur J Neurosci,* 18(8): 2227-2238.

Henderson, R. (1995). The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q Rev Biophys,* 28(2): 171-193.

Ho, J. D., R. Yeh, A. Sandstrom, I. Chorny, W. E. Harries, R. A. Robbins, L. J. Miercke andR. M. Stroud (2009). Crystal structure of human aquaporin 4 at 1.8 A and its mechanism of conductance. *Proc Natl Acad Sci U S A,* 106(18): 7437-7442.

Hopkins, R. andD. Esposito (2009). A rapid method for titrating baculovirus stocks using the Sf-9 Easy Titer cell line. *Biotechniques,* 47(3): 785-788.

Horsefield, R., K. Norden, M. Fellert, A. Backmark, S. Tornroth-Horsefield, A. C. Terwisscha van Scheltinga, J. Kvassman, P. Kjellbom, U. Johanson andR. Neutze (2008). High-resolution x-ray structure of human aquaporin 5. *Proc Natl Acad Sci U S A,* 105(36): 13327-13332.

Hu, L. A. andS. C. King (1998). Membrane topology of the Escherichia coli gamma-aminobutyrate transporter: implications on the topography and mechanism of prokaryotic and eukaryotic transporters from the APC superfamily. *Biochem J,* 336 (Pt 1): 69-76.

Huang, S. andD. E. Ingber (2005). Cell tension, matrix mechanics, and cancer development. *Cancer Cell,* 8(3): 175-176.

Ikotun, O. F., B. V. Marquez, C. Huang, K. Masuko, M. Daiji, T. Masuko, J. McConathy andS. E. Lapi (2013). Imaging the L-type amino acid transporter-1 (LAT1) with Zr-89 immunoPET. *PLoS One,* 8(10): e77476.

Ilgu, H., J. M. Jeckelmann, M. S. Gachet, R. Boggavarapu, Z. Ucurum, J. Gertsch andD. Fotiadis (2014). Variation of the detergent-binding capacity and phospholipid content of membrane proteins when purified in different detergents. *Biophys J,* 106(8): 1660-1670.

Isnard, A. D., D. Thomas andY. Surdin-Kerjan (1996). The study of methionine uptake in Saccharomyces cerevisiae reveals a new family of amino acid permeases. *J Mol Biol,* 262(4): 473-484.

Jack, D. L., I. T. Paulsen andM. H. Saier (2000). The amino acid/polyamine/organocation (APC) superfamily of transporters specific for amino acids, polyamines and organocations. *Microbiology,* 146 (Pt 8): 1797-1814.

Jamshad, M., J. Charlton, Y. P. Lin, S. J. Routledge, Z. Bawa, T. J. Knowles, M. Overduin, N. Dekker, T. R. Dafforn, R. M. Bill, D. R. Poyner andM. Wheatley (2015). G-protein coupled receptor solubilization and purification for biophysical analysis and functional studies, in the total absence of detergent. *Biosci Rep,* 35(2).

Jardetzky, O. (1966). Simple allosteric model for membrane pumps. *Nature,* 211(5052): 969-970.

Jiang, L., N. Kon, T. Li, S. J. Wang, T. Su, H. Hibshoosh, R. Baer andW. Gu (2015). Ferroptosis as a p53-mediated activity during tumour suppression. *Nature,* 520(7545): 57-62.

Jones, M. R. (2007). Lipids in photosynthetic reaction centres: structural roles and functional holes. *Prog Lipid Res,* 46(1): 56-87.

Kageyama, T., M. Nakamura, A. Matsuo, Y. Yamasaki, Y. Takakura, M. Hashida, Y. Kanai, M. Naito, T. Tsuruo, N. Minato andS. Shimohama (2000). The 4F2hc/LAT1 complex transports L-DOPA across the blood-brain barrier. *Brain Res,* 879(1-2): 115-121.

Kaira, K., T. Ishizuka, N. Yanagitani, N. Sunaga, T. Hisada andM. Mori (2007). Value of FDG positron emission tomography in monitoring the effects of therapy in progressive pulmonary sarcoidosis. *Clin Nucl Med,* 32(2): 114-116.

Kaira, K., N. Oriuchi, H. Imai, K. Shimizu, N. Yanagitani, N. Sunaga, T. Hisada, O. Kawashima, Y. Kamide, T. Ishizuka, Y. Kanai, T. Nakajima andM. Mori (2009). CD98 expression is associated with poor prognosis in resected non-small-cell lung cancer with lymph node metastases. *Ann Surg Oncol,* 16(12): 3473-3481.

Kaira, K., N. Oriuchi, H. Imai, K. Shimizu, N. Yanagitani, N. Sunaga, T. Hisada, S. Tanaka, T. Ishizuka, Y. Kanai, H. Endou, T. Nakajima andM. Mori (2008). l-type amino acid transporter 1 and CD98 expression in primary and metastatic sites of human neoplasms. *Cancer Sci,* 99(12): 2380-2386.

Kaira, K., N. Oriuchi, T. Takahashi, K. Nakagawa, Y. Ohde, T. Okumura, H. Murakami, T. Shukuya, H. Kenmotsu, T. Naito, Y. Kanai, M. Endo, H. Kondo, T. Nakajima andN. Yamamoto (2011). L-type amino acid transporter 1 (LAT1) expression in malignant pleural mesothelioma. *Anticancer Res,* 31(12): 4075-4082.

Kaleeba, J. A. andE. A. Berger (2006). Kaposi's sarcoma-associated herpesvirus fusion-entry receptor: cystine transporter xCT. *Science,* 311(5769): 1921-1924.

Kanai, Y., Y. Fukasawa, S. H. Cha, H. Segawa, A. Chairoungdua, D. K. Kim, H. Matsuo, J. Y. Kim, K. Miyamoto, E. Takeda andH. Endou (2000). Transport properties of a system y+L neutral and

basic amino acid transporter. Insights into the mechanisms of substrate recognition. *J Biol Chem,* 275(27): 20787-20793.

Kanai, Y., H. Segawa, K. Miyamoto, H. Uchino, E. Takeda andH. Endou (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(37): 23629-23632.

Kashiwagi, K., S. Shibuya, H. Tomitori, A. Kuraishi andK. Igarashi (1997). Excretion and uptake of putrescine by the PotE protein in Escherichia coli. *J Biol Chem,* 272(10): 6318-6323.

Kawate, T. andE. Gouaux (2006). Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. *Structure,* 14(4): 673-681.

Khafizov, K., R. Staritzbichler, M. Stamm andL. R. Forrest (2010). A study of the evolution of inverted-topology repeats from LeuT-fold transporters using AlignMe. *Biochemistry,* 49(50): 10702-10713.

Khan, K. H. (2013). Gene expression in Mammalian cells and its applications. *Adv Pharm Bull,* 3(2): 257-263.

Kim, C. S., S. H. Cho, H. S. Chun, S. Y. Lee, H. Endou, Y. Kanai andK. Kim do (2008). BCH, an inhibitor of system L amino acid transporters, induces apoptosis in cancer cells. *Biol Pharm Bull,* 31(6): 1096-1100.

Kim, D. K., Y. Kanai, H. W. Choi, S. Tangtrongsup, A. Chairoungdua, E. Babu, K. Tachampa, N. Anzai, Y. Iribe andH. Endou (2002). Characterization of the system L amino acid transporter in T24 human bladder carcinoma cells. *Biochim Biophys Acta,* 1565(1): 112-121.

Kim, J. Y., Y. Kanai, A. Chairoungdua, S. H. Cha, H. Matsuo, D. K. Kim, J. Inatomi, H. Sawa, Y. Ida andH. Endou (2001). Human cystine/glutamate transporter: cDNA cloning and upregulation by oxidative stress in glioma cells. *Biochim Biophys Acta,* 1512(2): 335-344.

Kinne, A., R. Schulein andG. Krause (2011). Primary and secondary thyroid hormone transporters. *Thyroid Res,* 4 Suppl 1: S7.

Kleinschmidt, J. H. andJ. L. Popot (2014). Folding and stability of integral membrane proteins in amphipols. *Arch Biochem Biophys,* 564: 327-343.

Kobayashi, K., A. Ohnishi, J. Promsuk, S. Shimizu, Y. Kanai, Y. Shiokawa andM. Nagane (2008). Enhanced tumor growth elicited by L-type amino acid transporter 1 in human malignant glioma cells. *Neurosurgery,* 62(2): 493-503; discussion 503-494.

Kowalczyk, L., 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 andM. Palacin (2011). Molecular basis of substrate-induced permeation by an amino acid antiporter. *Proc Natl Acad Sci U S A,* 108(10): 3935-3940.

Kramer, S. D., L. Mu, A. Muller, C. Keller, O. F. Kuznetsova, C. Schweinsberg, D. Franck, C. Muller, T. L. Ross, R. Schibli andS. M. Ametamey (2012). 5-(2-18F-fluoroethoxy)-L-tryptophan as a substrate of system L transport for tumor imaging by PET. *J Nucl Med,* 53(3): 434-442.

Krishnamurthy, H., C. L. Piscitelli andE. Gouaux (2009). Unlocking the molecular secrets of sodium-coupled transporters. *Nature,* 459(7245): 347-355.

Kroemer, G. andJ. Pouyssegur (2008). Tumor cell metabolism: cancer's Achilles' heel. *Cancer Cell,* 13(6): 472-482.

le Maire, M., P. Champeil andJ. V. Moller (2000). Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim Biophys Acta,* 1508(1-2): 86-111.

Levin, E. J., Y. Cao, G. Enkavi, M. Quick, Y. Pan, E. Tajkhorshid andM. Zhou (2012). Structure and permeation mechanism of a mammalian urea transporter. *Proc Natl Acad Sci U S A,* 109(28): 11194-11199.

Liao, M., E. Cao, D. Julius andY. Cheng (2013). Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature,* 504(7478): 107-112.

Liao, M., E. Cao, D. Julius andY. Cheng (2014). Single particle electron cryo-microscopy of a mammalian ion channel. *Curr Opin Struct Biol,* 27: 1-7.

Liu, P., C. Huang, H. L. Wang, K. Zhou, F. X. Xiao andW. Qun (2004). The importance of Loop 7 for the activity of calcineurin. *FEBS Lett,* 577(1-2): 205-208.

Lo, M., V. Ling, Y. Z. Wang andP. W. Gout (2008). The xc- cystine/glutamate antiporter: a mediator of pancreatic cancer growth with a role in drug resistance. *Br J Cancer,* 99(3): 464-472.

Long, A. R., C. C. O'Brien, K. Malhotra, C. T. Schwall, A. D. Albert, A. Watts andN. N. Alder (2013). A detergent-free strategy for the reconstitution of active enzyme complexes from native biological membranes into nanoscale discs. *BMC Biotechnol,* 13: 41.

Long, S. B., X. Tao, E. B. Campbell andR. MacKinnon (2007). Atomic structure of a voltage-dependent K+ channel in a lipid membrane-like environment. *Nature,* 450(7168): 376-382.

Lu, P., X. C. Bai, D. Ma, T. Xie, C. Yan, L. Sun, G. Yang, Y. Zhao, R. Zhou, S. H. Scheres andY. Shi (2014). Three-dimensional structure of human gamma-secretase. *Nature,* 512(7513): 166-170.

Lu, P., D. Ma, C. Yan, X. Gong, M. Du andY. Shi (2014). Structure and mechanism of a eukaryotic transmembrane ascorbate-dependent oxidoreductase. *Proc Natl Acad Sci U S A,* 111(5): 1813-1818.

Ma, D., P. Lu, C. Yan, C. Fan, P. Yin, J. Wang andY. Shi (2012). Structure and mechanism of a glutamate-GABA antiporter. *Nature,* 483(7391): 632-636.

Macauley-Patrick, S., M. L. Fazenda, B. McNeil andL. M. Harvey (2005). Heterologous protein production using the Pichia pastoris expression system. *Yeast,* 22(4): 249-270.

Makowske, M. andH. N. Christensen (1982). Hepatic transport system interconverted by protonation from service for neutral to service for anionic amino acids. *J Biol Chem,* 257(24): 14635-14638.

Manglik, A., A. C. Kruse, T. S. Kobilka, F. S. Thian, J. M. Mathiesen, R. K. Sunahara, L. Pardo, W. I. Weis, B. K. Kobilka andS. Granier (2012). Crystal structure of the micro-opioid receptor bound to a morphinan antagonist. *Nature,* 485(7398): 321-326.

Mastroberardino, L., B. Spindler, R. Pfeiffer, P. J. Skelly, J. Loffing, C. B. Shoemaker andF. Verrey (1998). Amino-acid transport by heterodimers of 4F2hc/CD98 and members of a permease family. *Nature,* 395(6699): 288-291.

Matsuo, H., Y. Kanai, J. Y. Kim, A. Chairoungdua, D. K. Kim, J. Inatomi, Y. Shigeta, H. Ishimine, S. Chaekuntode, K. Tachampa, H. W. Choi, E. Babu, J. Fukuda andH. Endou (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(23): 21017-21026.

McLaughlin, S. andA. Aderem (1995). The myristoyl-electrostatic switch: a modulator of reversible protein-membrane interactions. *Trends Biochem Sci,* 20(7): 272-276.

Meier, C., Z. Ristic, S. Klauser andF. Verrey (2002). Activation of system L heterodimeric amino acid exchangers by intracellular substrates. *EMBO J,* 21(4): 580-589.

Meury, M., M. Costa, D. Harder, M. Stauffer, J. M. Jeckelmann, B. Bruhlmann, A. Rosell, H. Ilgu, K. Kovar, M. Palacin andD. Fotiadis (2014). Detergent-induced stabilization and improved 3D map of the human heteromeric amino acid transporter 4F2hc-LAT2. *PLoS One,* 9(10): e109882.

Michel, H. (2001). Crystallization of membrane proteins. International Tables for Crystallograpy F. Kluwer Academic Publishers.

Miller, A. N. andS. B. Long (2012). Crystal structure of the human two-pore domain potassium channel K2P1. *Science,* 335(6067): 432-436.

Milliner, D. S. andM. E. Murphy (1993). Urolithiasis in pediatric patients. *Mayo Clin Proc,* 68(3): 241-248.

Murray, C. W., M. L. Verdonk andD. C. Rees (2012). Experiences in fragment-based drug discovery. *Trends Pharmacol Sci,* 33(5): 224-232.

Muth, T. R. andS. Schuldiner (2000). A membrane-embedded glutamate is required for ligand binding to the multidrug transporter EmrE. *EMBO J,* 19(2): 234-240.

Mykkanen, J., D. Torrents, M. Pineda, M. Camps, M. E. Yoldi, N. Horelli-Kuitunen, K. Huoponen, M. Heinonen, J. Oksanen, O. Simell, M. L. Savontaus, A. Zorzano, M. Palacin andP. Aula (2000). Functional analysis of novel mutations in y(+)LAT-1 amino acid transporter gene causing lysinuric protein intolerance (LPI). *Hum Mol Genet,* 9(3): 431-438.

Nabeyama, A., A. Kurita, K. Asano, Y. Miyake, T. Yasuda, I. Miura, G. Nishitai, S. Arakawa, S. Shimizu, S. Wakana, H. Yoshida andM. Tanaka (2010). xCT deficiency accelerates chemically induced tumorigenesis. *Proc Natl Acad Sci U S A,* 107(14): 6436-6441.

Nagamori, S., P. Wiriyasermkul, M. E. Guarch, H. Okuyama, S. Nakagomi, K. Tadagaki, Y. Nishinaka, S. Bodoy, K. Takafuji, S. Okuda, J. Kurokawa, R. Ohgaki, V. Nunes, M. Palacin andY. Kanai (2016). Novel cystine transporter in renal proximal tubule identified as a missing partner of cystinuria-related plasma membrane protein rBAT/SLC3A1. *Proc Natl Acad Sci U S A,* 113(3): 775-780.

Nakamura, E., M. Sato, H. Yang, F. Miyagawa, M. Harasaki, K. Tomita, S. Matsuoka, A. Noma, K. Iwai andN. Minato (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(5): 3009-3016.

Nakauchi, J., H. Matsuo, D. K. Kim, A. Goto, A. Chairoungdua, S. H. Cha, J. Inatomi, Y. Shiokawa, K. Yamaguchi, I. Saito, H. Endou andY. Kanai (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(3): 231-235.

Nelson, N. (1987). A novel method for the detection of receptors and membrane proteins by scintillation proximity radioassay. *Anal Biochem,* 165(2): 287-293.

Newby, Z. E., J. O'Connell, 3rd, Y. Robles-Colmenares, S. Khademi, L. J. Miercke andR. M. Stroud (2008). Crystal structure of the aquaglyceroporin PfAQP from the malarial parasite Plasmodium falciparum. *Nat Struct Mol Biol,* 15(6): 619-625.

Newman, M. J., D. L. Foster, T. H. Wilson andH. R. Kaback (1981). Purification and reconstitution of functional lactose carrier from Escherichia coli. *J Biol Chem,* 256(22): 11804-11808.

Nguyen, H. T. andD. Merlin (2012). Homeostatic and innate immune responses: role of the transmembrane glycoprotein CD98. *Cell Mol Life Sci,* 69(18): 3015-3026.

Nicklin, P., 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 andL. O. Murphy (2009). Bidirectional transport of amino acids regulates mTOR and autophagy. *Cell,* 136(3): 521-534.

Nishida, M., M. Cadene, B. T. Chait andR. MacKinnon (2007). Crystal structure of a Kir3.1-prokaryotic Kir channel chimera. *EMBO J,* 26(17): 4005-4015.

Oda, K., N. Hosoda, H. Endo, K. Saito, K. Tsujihara, M. Yamamura, T. Sakata, N. Anzai, M. F. Wempe, Y. Kanai andH. Endou (2010). L-type amino acid transporter 1 inhibitors inhibit tumor cell growth. *Cancer Sci,* 101(1): 173-179.

Ohgimoto, S., N. Tabata, S. Suga, M. Tsurudome, M. Kawano, M. Nishio, K. Okamoto, H. Komada, N. Watanabe andY. Ito (1996). Regulation of human immunodeficiency virus gp160-mediated cell fusion by antibodies against fusion regulatory protein 1. *J Gen Virol,* 77 (Pt 11): 2747-2756.

Okamoto, K., M. Tsurudome, S. Ohgimoto, M. Kawano, M. Nishio, H. Komada, M. Ito, Y. Sakakura andY. Ito (1997). An anti-fusion regulatory protein-1 monoclonal antibody suppresses human parainfluenza virus type 2-induced cell fusion. *J Gen Virol,* 78 (Pt 1): 83-89.

Overington, J., B. Al-Lazikani andA. Hopkins (2006). How many drug targets are there? *Nat Rev Drug Discov,* 5(12): 993–996.

Oyanagi, K. andM. Nagao (1998). [Lysinuric protein intolerance and other cationic aminoacidurias]. *Ryoikibetsu Shokogun Shirizu,*(19 Pt 2): 562-564.

Palacin, M. andY. Kanai (2004). The ancillary proteins of HATs: SLC3 family of amino acid transporters. *Pflugers Arch,* 447(5): 490-494.

Palacin, M., V. Nunes, M. Font-Llitjos, M. Jimenez-Vidal, J. Fort, E. Gasol, M. Pineda, L. Feliubadalo, J. Chillaron andA. Zorzano (2005). The genetics of heteromeric amino acid transporters. *Physiology (Bethesda),* 20: 112-124.

Park, S. Y., J. K. Kim, I. J. Kim, B. K. Choi, K. Y. Jung, S. Lee, K. J. Park, A. Chairoungdua, Y. Kanai, H. Endou andD. K. Kim (2005). Reabsorption of neutral amino acids mediated by amino acid

transporter LAT2 and TAT1 in the basolateral membrane of proximal tubule. *Arch Pharm Res,* 28(4): 421-432.

Parmacek, M. S., B. A. Karpinski, K. M. Gottesdiener, C. B. Thompson andJ. M. Leiden (1989). Structure, expression and regulation of the murine 4F2 heavy chain. *Nucleic Acids Res,* 17(5): 1915-1931.

Paulin, S., M. Jamshad, T. R. Dafforn, J. Garcia-Lara, S. J. Foster, N. F. Galley, D. I. Roper, H. Rosado andP. W. Taylor (2014). Surfactant-free purification of membrane protein complexes from bacteria: application to the staphylococcal penicillin-binding protein complex PBP2/PBP2a. *Nanotechnology,* 25(28): 285101.

Penmatsa, A., K. H. Wang andE. Gouaux (2013). X-ray structure of dopamine transporter elucidates antidepressant mechanism. *Nature,* 503(7474): 85-90.

Penmatsa, A., K. H. Wang andE. Gouaux (2015). X-ray structures of Drosophila dopamine transporter in complex with nisoxetine and reboxetine. *Nat Struct Mol Biol,* 22(6): 506-508.

Perez, C., C. Koshy, O. Yildiz andC. Ziegler (2012). Alternating-access mechanism in conformationally asymmetric trimers of the betaine transporter BetP. *Nature,* 490(7418): 126-130.

Pfeiffer, R., G. Rossier, B. Spindler, C. Meier, L. Kuhn andF. Verrey (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(1): 49-57.

Pfeiffer, R., B. Spindler, J. Loffing, P. J. Skelly, C. B. Shoemaker andF. Verrey (1998). Functional heterodimeric amino acid transporters lacking cysteine residues involved in disulfide bond. *FEBS Lett,* 439(1-2): 157-162.

Phillips, R., T. Ursell, P. Wiggins andP. Sens (2009). Emerging roles for lipids in shaping membrane-protein function. *Nature,* 459(7245): 379-385.

Pineda, M., E. Fernandez, D. Torrents, R. Estevez, C. Lopez, M. Camps, J. Lloberas, A. Zorzano andM. Palacin (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(28): 19738-19744.

Popot, J. L., T. Althoff, D. Bagnard, J. L. Baneres, P. Bazzacco, E. Billon-Denis, L. J. Catoire, P. Champeil, D. Charvolin, M. J. Cocco, G. Cremel, T. Dahmane, L. M. de la Maza, C. Ebel, F. Gabel, F. Giusti, Y. Gohon, E. Goormaghtigh, E. Guittet, J. H. Kleinschmidt, W. Kuhlbrandt, C. Le Bon, K. L. Martinez, M. Picard, B. Pucci, J. N. Sachs, C. Tribet, C. van Heijenoort, F. Wien, F. Zito andM. Zoonens (2011). Amphipols from A to Z. *Annu Rev Biophys,* 40: 379-408.

Postis, V., S. Rawson, J. K. Mitchell, S. C. Lee, R. A. Parslow, T. R. Dafforn, S. A. Baldwin andS. P. Muench (2015). The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy. *Biochim Biophys Acta,* 1848(2): 496-501.

Prabudiansyah, I., I. Kusters, A. Caforio andA. J. Driessen (2015). Characterization of the annular lipid shell of the Sec translocon. *Biochim Biophys Acta,* 1848(10 Pt A): 2050-2056.

Prive, G. G. (2007). Detergents for the stabilization and crystallization of membrane proteins. *Methods,* 41(4): 388-397.

Pujadas, G. andJ. Palau (2001). Evolution of alpha-amylases: architectural features and key residues in the stabilization of the (beta/alpha)(8) scaffold. *Mol Biol Evol,* 18(1): 38-54.

Quackenbush, E., M. Clabby, K. M. Gottesdiener, J. Barbosa, N. H. Jones, J. L. Strominger, S. Speck andJ. M. Leiden (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 U S A,* 84(18): 6526-6530.

Quackenbush, E. J., P. Linsley andM. Letarte (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(1): 234-239.

Quick, M. andJ. A. Javitch (2007). Monitoring the function of membrane transport proteins in detergent-solubilized form. *Proc Natl Acad Sci U S A,* 104(9): 3603-3608.

Reddy, P., L. Liu, D. Adhikari, K. Jagarlamudi, S. Rajareddy, Y. Shen, C. Du, W. Tang, T. Hamalainen, S. L. Peng, Z. J. Lan, A. J. Cooney, I. Huhtaniemi andK. Liu (2008). Oocyte-specific deletion of Pten causes premature activation of the primordial follicle pool. *Science,* 319(5863): 611-613.

Reig, N., J. Chillaron, P. Bartoccioni, E. Fernandez, A. Bendahan, A. Zorzano, B. Kanner, M. Palacin andJ. Bertran (2002). The light subunit of system b(o,+) is fully functional in the absence of the heavy subunit. *EMBO J,* 21(18): 4906-4914.

Reig, N., C. del Rio, F. Casagrande, M. Ratera, J. L. Gelpi, D. Torrents, P. J. Henderson, H. Xie, S. A. Baldwin, A. Zorzano, D. Fotiadis andM. Palacin (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.

Ressl, S., A. C. Terwisscha van Scheltinga, C. Vonrhein, V. Ott andC. Ziegler (2009). Molecular basis of transport and regulation in the Na(+)/betaine symporter BetP. *Nature,* 458(7234): 47-52.

Ritchie, J. W. andP. M. Taylor (2001). Role of the System L permease LAT1 in amino acid and iodothyronine transport in placenta. *Biochem J,* 356(Pt 3): 719-725.

Rius, M. andJ. Chillaron (2012). Carrier subunit of plasma membrane transporter is required for oxidative folding of its helper subunit. *J Biol Chem,* 287(22): 18190-18200.

Rius, M., L. Sala andJ. Chillaron (2016). The role of N-glycans and the C-terminal loop of the subunit rBAT in the biogenesis of the cystinuria-associated transporter. *Biochem J,* 473(3): 233-244.

Rodríguez-Banqueri, A. (2013). "A random approach to stabilize a membrane transport protein for crystallization studies". Doctoral thesis, Universitat de Barcelona.

Rodríguez-de-la-Ballina., L. (2011). Characterization of the multifunctional protein 4F2hc. Doctoral thesis, Universitat de Barcelona.

Rosano, G. L. andE. A. Ceccarelli (2014). Recombinant protein expression in Escherichia coli: advances and challenges. *Front Microbiol,* 5: 172.

Rosell, A., M. Meury, E. Alvarez-Marimon, M. Costa, L. Perez-Cano, A. Zorzano, J. Fernandez-Recio, M. Palacin andD. Fotiadis (2014). Structural bases for the interaction and stabilization of the human amino acid transporter LAT2 with its ancillary protein 4F2hc. *Proc Natl Acad Sci U S A,* 111(8): 2966-2971.

Rosenbaum, D. M., C. Zhang, J. A. Lyons, R. Holl, D. Aragao, D. H. Arlow, S. G. Rasmussen, H. J. Choi, B. T. Devree, R. K. Sunahara, P. S. Chae, S. H. Gellman, R. O. Dror, D. E. Shaw, W. I. Weis, M. Caffrey, P. Gmeiner andB. K. Kobilka (2011). Structure and function of an irreversible agonist-beta(2) adrenoceptor complex. *Nature,* 469(7329): 236-240.

Rossier, G., C. Meier, C. Bauch, V. Summa, B. Sordat, F. Verrey andL. C. Kuhn (1999). LAT2, a new basolateral 4F2hc/CD98-associated amino acid transporter of kidney and intestine. *J Biol Chem,* 274(49): 34948-34954.

Russ, W. P. andD. M. Engelman (2000). The GxxxG motif: a framework for transmembrane helix-helix association. *J Mol Biol,* 296(3): 911-919.

Safory, H., S. Neame, Y. Shulman, S. Zubedat, I. Radzishevsky, D. Rosenberg, H. Sason, S. Engelender, A. Avital, S. Hulsmann, J. Schiller andH. Wolosker (2015). The alanine-serine-cysteine-1 (Asc-1) transporter controls glycine levels in the brain and is required for glycinergic inhibitory transmission. *EMBO Rep,* 16(5): 590-598.

Sahdev, S., S. K. Khattar andK. S. Saini (2008). Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Mol Cell Biochem,* 307(1-2): 249-264.

Sanders, C. R. andJ. K. Myers (2004). Disease-related misassembly of membrane proteins. *Annu Rev Biophys Biomol Struct,* 33: 25-51.

Sato, H., M. Tamba, T. Ishii andS. Bannai (1999). Cloning and expression of a plasma membrane cystine/glutamate exchange transporter composed of two distinct proteins. *J Biol Chem,* 274(17): 11455-11458.

Savaskan, E., R. Ehrhardt, A. Schulz, M. Walter andH. Schachinger (2008). Post-learning intranasal oxytocin modulates human memory for facial identity. *Psychoneuroendocrinology,* 33(3): 368-374.

Scheres, S. H. (2012). A Bayesian view on cryo-EM structure determination. *J Mol Biol,* 415(2): 406-418.

Scheres, S. H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol,* 180(3): 519-530.

Schlessinger, A., S. W. Yee, A. Sali andK. M. Giacomini (2013). SLC classification: an update. *Clin Pharmacol Ther,* 94(1): 19-23.

Sebastio, G., M. P. Sperandeo andG. Andria (2011). Lysinuric protein intolerance: reviewing concepts on a multisystem disease. *Am J Med Genet C Semin Med Genet,* 157C(1): 54-62.

Segawa, H., Y. Fukasawa, K. Miyamoto, E. Takeda, H. Endou andY. Kanai (1999). Identification and functional characterization of a Na+-independent neutral amino acid transporter with broad substrate selectivity. *J Biol Chem,* 274(28): 19745-19751.

Seib, T. M., S. A. Patel andR. J. Bridges (2011). Regulation of the system x(C)- cystine/glutamate exchanger by intracellular glutathione levels in rat astrocyte primary cultures. *Glia,* 59(10): 1387-1401.

Seiler, A., M. Schneider, H. Forster, S. Roth, E. K. Wirth, C. Culmsee, N. Plesnila, E. Kremmer, O. Radmark, W. Wurst, G. W. Bornkamm, U. Schweizer andM. Conrad (2008). Glutathione peroxidase 4 senses and translates oxidative stress into 12/15-lipoxygenase dependent- and AIF-mediated cell death. *Cell Metab,* 8(3): 237-248.

Senes, A., D. E. Engel andW. F. DeGrado (2004). Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Curr Opin Struct Biol,* 14(4): 465-479.

Senes, A., M. Gerstein andD. M. Engelman (2000). Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *J Mol Biol,* 296(3): 921-936.

Sengupta, S., T. R. Peterson andD. M. Sabatini (2010). Regulation of the mTOR complex 1 pathway by nutrients, growth factors, and stress. *Mol Cell,* 40(2): 310-322.

Shaffer, P. L., A. Goehring, A. Shankaranarayanan andE. Gouaux (2009). Structure and mechanism of a Na+-independent amino acid transporter. *Science,* 325(5943): 1010-1014.

Shchors, K. andG. Evan (2007). Tumor angiogenesis: cause or consequence of cancer? *Cancer Res,* 67(15): 7059-7061.

Shigeta, Y., Y. Kanai, A. Chairoungdua, N. Ahmed, S. Sakamoto, H. Matsuo, D. K. Kim, M. Fujimura, N. Anzai, K. Mizoguchi, T. Ueda, K. Akakura, T. Ichikawa, H. Ito andH. Endou (2006). A novel missense mutation of SLC7A9 frequent in Japanese cystinuria cases affecting the C-terminus of the transporter. *Kidney Int,* 69(7): 1198-1206.

Shimamura, T., M. Shiroishi, S. Weyand, H. Tsujimoto, G. Winter, V. Katritch, R. Abagyan, V. Cherezov, W. Liu, G. W. Han, T. Kobayashi, R. C. Stevens andS. Iwata (2011). Structure of the human histamine H1 receptor complex with doxepin. *Nature,* 475(7354): 65-70.

Shimamura, T., S. Weyand, O. Beckstein, N. G. Rutherford, J. M. Hadden, D. Sharples, M. S. Sansom, S. Iwata, P. J. Henderson andA. D. Cameron (2010). Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science,* 328(5977): 470-473.

Shintre, C. A., A. C. Pike, Q. Li, J. I. Kim, A. J. Barr, S. Goubin, L. Shrestha, J. Yang, G. Berridge, J. Ross, P. J. Stansfeld, M. S. Sansom, A. M. Edwards, C. Bountra, B. D. Marsden, F. von Delft, A. N. Bullock, O. Gileadi, N. A. Burgess-Brown andE. P. Carpenter (2013). Structures of ABCB10, a human ATP-binding cassette transporter in apo- and nucleotide-bound states. *Proc Natl Acad Sci U S A,* 110(24): 9710-9715.

Shultis, D. D., M. D. Purdy, C. N. Banchs andM. C. Wiener (2006). Outer membrane active transport: structure of the BtuB:TonB complex. *Science,* 312(5778): 1396-1399.

Skaar, K., H. J. Korza, M. Tarry, P. Sekyrova andM. Hogbom (2015). Expression and Subcellular Distribution of GFP-Tagged Human Tetraspanin Proteins in Saccharomyces cerevisiae. *PLoS One,* 10(7): e0134041.

Sobolevsky, A. I., M. P. Rosconi andE. Gouaux (2009). X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor. *Nature,* 462(7274): 745-756.

Standfuss, J., G. Xie, P. C. Edwards, M. Burghammer, D. D. Oprian andG. F. Schertler (2007). Crystal structure of a thermally stable rhodopsin mutant. *J Mol Biol,* 372(5): 1179-1188.

Suga, S., M. Tsurudome, M. Ito, S. Ohgimoto, N. Tabata, M. Nishio, M. Kawano, H. Komada, M. Sakurai andY. Ito (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(4): 237-243.

Sun, J., J. R. Bankston, J. Payandeh, T. R. Hinds, W. N. Zagotta andN. Zheng (2014). Crystal structure of the plant dual-affinity nitrate transporter NRT1.1. *Nature,* 507(7490): 73-77.

Suzuki, H., T. Nishizawa, K. Tani, Y. Yamazaki, A. Tamura, R. Ishitani, N. Dohmae, S. Tsukita, O. Nureki andY. Fujiyoshi (2014). Crystal structure of a claudin provides insight into the architecture of tight junctions. *Science,* 344(6181): 304-307.

Swainsbury, D. J., S. Scheidelaar, R. van Grondelle, J. A. Killian andM. R. Jones (2014). Bacterial reaction centers purified with styrene maleic acid copolymer retain native membrane functional properties and display enhanced stability. *Angew Chem Int Ed Engl,* 53(44): 11803-11807.

Takeuchi, S., K. Wada, T. Toyooka, N. Shinomiya, H. Shimazaki, K. Nakanishi, K. Nagatani, N. Otani, H. Osada, Y. Uozumi, H. Matsuo andH. Nawashiro (2013). Increased xCT expression correlates with tumor invasion and outcome in patients with glioblastomas. *Neurosurgery,* 72(1): 33-41; discussion 41.

Tang, L., L. Bai, W. H. Wang andT. Jiang (2010). Crystal structure of the carnitine transporter and insights into the antiport mechanism. *Nat Struct Mol Biol,* 17(4): 492-496.

Tate, S. S., N. Yan andS. Udenfriend (1992). Expression cloning of a Na(+)-independent neutral amino acid transporter from rat kidney. *Proc Natl Acad Sci U S A,* 89(1): 1-5.

Toivonen, M., M. Tringham, J. Kurko, P. Terho, O. Simell, K. M. Heiskanen andJ. Mykkanen (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(4): 479-488.

Tomi, M., M. Mori, M. Tachikawa, K. Katayama, T. Terasaki andK. Hosoya (2005). L-type amino acid transporter 1-mediated L-leucine transport at the inner blood-retinal barrier. *Invest Ophthalmol Vis Sci,* 46(7): 2522-2530.

Torrents, D., R. Estevez, M. Pineda, E. Fernandez, J. Lloberas, Y. B. Shi, A. Zorzano andM. Palacin (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. A candidate gene for lysinuric protein intolerance. *J Biol Chem,* 273(49): 32437-32445.

Torrents, D., J. Mykkanen, M. Pineda, L. Feliubadalo, R. Estevez, R. de Cid, P. Sanjurjo, A. Zorzano, V. Nunes, K. Huoponen, A. Reinikainen, O. Simell, M. L. Savontaus, P. Aula andM. Palacin (1999). Identification of SLC7A7, encoding y+LAT-1, as the lysinuric protein intolerance gene. *Nat Genet,* 21(3): 293-296.

Turnay, J., J. Fort, N. Olmo, A. Santiago-Gomez, M. Palacin andM. A. Lizarbe (2011). Structural characterization and unfolding mechanism of human 4F2hc ectodomain. *Biochim Biophys Acta,* 1814(5): 536-544.

Tusnady, G. E. andI. Simon (1998). Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J Mol Biol,* 283(2): 489-506.

Uchino, H., Y. Kanai, D. K. Kim, M. F. Wempe, A. Chairoungdua, E. Morimoto, M. W. Anders andH. Endou (2002). Transport of amino acid-related compounds mediated by L-type amino acid transporter 1 (LAT1): insights into the mechanisms of substrate recognition. *Mol Pharmacol,* 61(4): 729-737.

Udenfriend, S., L. D. Gerber, L. Brink andS. Spector (1985). Scintillation proximity radioimmunoassay utilizing 125I-labeled ligands. *Proc Natl Acad Sci U S A,* 82(24): 8672-8676.

Vafa, O., M. Wade, S. Kern, M. Beeche, T. K. Pandita, G. M. Hampton andG. M. Wahl (2002). c-Myc can induce DNA damage, increase reactive oxygen species, and mitigate p53 function: a mechanism for oncogene-induced genetic instability. *Mol Cell,* 9(5): 1031-1044.

van Geest, M. andJ. S. Lolkema (2000). Membrane topology and insertion of membrane proteins: search for topogenic signals. *Microbiol Mol Biol Rev,* 64(1): 13-33.

van Meer, G., D. R. Voelker andG. W. Feigenson (2008). Membrane lipids: where they are and how they behave. *Nat Rev Mol Cell Biol,* 9(2): 112-124.

Verrey, F., E. I. Closs, C. A. Wagner, M. Palacin, H. Endou andY. Kanai (2004). CATs and HATs: the SLC7 family of amino acid transporters. *Pflugers Arch,* 447(5): 532-542.

Wang, L., W. Qu, B. P. Lieberman, K. Plossl andH. F. Kung (2011). Synthesis, uptake mechanism characterization and biological evaluation of (18)F labeled fluoroalkyl phenylalanine analogs as potential PET imaging agents. *Nucl Med Biol,* 38(1): 53-62.

Wells, R. G. andM. A. Hediger (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 U S A,* 89(12): 5596-5600.

Weyand, S., 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 andA. D. Cameron (2008). Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science,* 322(5902): 709-713.

White, J. F., N. Noinaj, Y. Shibata, J. Love, B. Kloss, F. Xu, J. Gvozdenovic-Jeremic, P. Shah, J. Shiloach, C. G. Tate andR. Grisshammer (2012). Structure of the agonist-bound neurotensin receptor. *Nature,* 490(7421): 508-513.

Wiriyasermkul, P., S. Nagamori, H. Tominaga, N. Oriuchi, K. Kaira, H. Nakao, T. Kitashoji, R. Ohgaki, H. Tanaka, H. Endou, K. Endo, H. Sakurai andY. Kanai (2012). Transport of 3-fluoro-L-alpha-methyl-tyrosine by tumor-upregulated L-type amino acid transporter 1: a cause of the tumor uptake in PET. *J Nucl Med,* 53(8): 1253-1261.

Wolf, D. A., S. Wang, M. A. Panzica, N. H. Bassily andN. L. Thompson (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(21): 5012-5022.

Wong, F. H., J. S. Chen, V. Reddy, J. L. Day, M. A. Shlykov, S. T. Wakabayashi andM. H. Saier, Jr. (2012). The amino acid-polyamine-organocation superfamily. *J Mol Microbiol Biotechnol,* 22(2): 105-113.

Xie, X., T. Dumas, L. Tang, T. Brennan, T. Reeder, W. Thomas, R. D. Klein, J. Flores, B. F. O'Hara, H. C. Heller andP. Franken (2005). Lack of the alanine-serine-cysteine transporter 1 causes tremors, seizures, and early postnatal death in mice. *Brain Res,* 1052(2): 212-221.

Yamashita, A., S. K. Singh, T. Kawate, Y. Jin andE. Gouaux (2005). Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters. *Nature,* 437(7056): 215-223.

Yanagida, O., Y. Kanai, A. Chairoungdua, D. K. Kim, H. Segawa, T. Nii, S. H. Cha, H. Matsuo, J. Fukushima, Y. Fukasawa, Y. Tani, Y. Taketani, H. Uchino, J. Y. Kim, J. Inatomi, I. Okayasu, K. Miyamoto, E. Takeda, T. Goya andH. Endou (2001). Human L-type amino acid transporter 1 (LAT1): characterization of function and expression in tumor cell lines. *Biochim Biophys Acta,* 1514(2): 291-302.

Zhang, C., Y. Srinivasan, D. H. Arlow, J. J. Fung, D. Palmer, Y. Zheng, H. F. Green, A. Pandey, R. O. Dror, D. E. Shaw, W. I. Weis, S. R. Coughlin andB. K. Kobilka (2012). High-resolution crystal structure of human protease-activated receptor 1. *Nature,* 492(7429): 387-392.

Zhou, M., J. H. Morais-Cabral, S. Mann andR. MacKinnon (2001). Potassium channel receptor site for the inactivation gate and quaternary amine inhibitors. *Nature,* 411(6838): 657-661.

Zubcevic, L., M. A. Herzik, Jr., B. C. Chung, Z. Liu, G. C. Lander andS. Y. Lee (2016). Cryo-electron microscopy structure of the TRPV2 ion channel. *Nat Struct Mol Biol,* 23(2): 180-186.

Appendix

Towards the molecular mechanism of the Arg/Agm exchanger AdiC

Introduction

The present structural paradigm of the light subunits of HATs (LATs), is the arginine/agmatine (Arg/Agm) exchanger (AdiC) from *E. coli* (Fang *et al.* 2009; Gao *et al.* 2009; Gao *et al.* 2010; Kowalczyk *et al.* 2011) (PDB codes, 3NCY; 3LRB; 3L1L; 3OB6, respectively). AdiC belongs to the Amino acid/Polyamine Antiporter APA subfamily, which together with the LAT subfamily, are included in the APC superfamily of transporters (Casagrande *et al.* 2008). Although AdiC shows only about 18 % of sequence identity with LATs, structure-function studies suggests that AdiC and LATs share a similar protein fold (Gasol *et al.* 2004; Bartoccioni *et al.* 2010).

AdiC is expressed as a dimer in the plasma membrane, and takes part in the acid resistance system 3 (AR3) used by the enteric bacteria, like *E. coli*, to survive to extreme acid conditions in the host´s stomach (pH 1 to 3), and to weak acids in the intestine (pH 4.5 to 7 with high concentrations of volatile fatty acids) (Lin *et al.* 1996). This is one out of four amino acid-dependent acid resistance systems used by these bacteria, to deal with such an acidic environments (arginine- glutamic acid–, lysine- and ornithine-dependent acid resistance systems). These mechanisms allow pathogenic *E. coli* strains to establish several types of gastrointestinal and urine infections that can be lethal, like the hemolytic-uremic syndrome (HUS) (Griffin and Tauxe 1991). When Arg enters to the cell, the decarboxylase AdiA catalyzes the conversion of Arg^{1+} to Agm^{2+} inside bacteria, which is then expelled out (Gong *et al.* 2003; Iyer *et al.* 2003). In this process, one proton is consumed increasing the internal pH from 3.6 to 4.7 in the stomach.

AdiC, and some secondary prokaryotic ion-coupled symporters from the amino acid/polyamine/organocation (APC) family, which share low amino acid sequence identity among them (<10 %), exhibit the fold denominated as 5+5 inverted repeat or LeuT fold (extensively described in the Introduction section 1.7.2.1). It has been proposed that this fold allows these transporters to undergo conformational changes among symmetry-related states to carry out the transport (Yamashita *et al.* 2005; Khafizov *et al.* 2010). AdiC is expected to undergo conformational changes between outward- and inward-facing conformations (open-to-out and open-to-in, respectively), to catalyze the exchange of Arg/Agm (Figure 1) (Kowalczyk *et al.* 2011). Upon initial interaction with the substrate (Arg or Agm), there is an induced-fitting that improves substrate binding and occludes the substrate (thin gate, Figure 1), blocking free diffusion to either side of the membrane (Kowalczyk *et al.* 2011). AdiC structure has been solved in three different outward facing conformations (Figure 1): outward-facing without substrate (PDB: 3LRB and 3NCY) (Fang *et al.* 2009; Gao *et al.* 2009), outward-facing with substrate bound (mutant N101A, PDB: 3OB6) (Kowalczyk *et al.* 2011), and outward-facing with substrate occluded (mutant N22A; PDB: 3L1L) (Gao *et al.* 2010). One of them (mutant N101A, PDB: 3OB6) was solved in our lab at 3.0Å resolution (Figure 2).

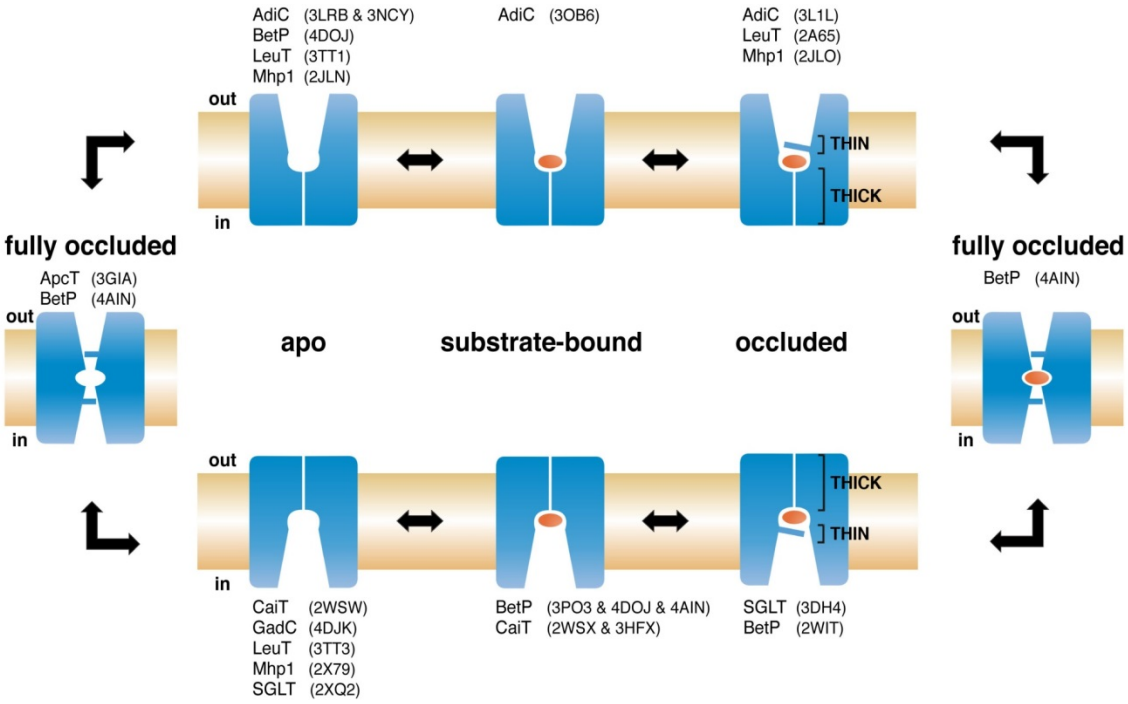

Figure 1. Symmetrical states along the alternative access mechanism of transporters with the LeuT fold. Upon substrate (red ellipsoid) binding to the open-to-out (outward-facing) apo state, the substrate-bound state (AdiC (3OB6)) evolves to an occluded state, where two gates (thick and thin) prevent the diffusion of the substrate to either side of the membrane (Krishnamurthy *et al.* 2009).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 (Yamashita *et al.* 2005), vSGLT (Faham *et al.* 2008), Mhp1 (Weyand *et al.* 2008; Shimamura *et al.* 2010), BetP (Ressl *et al.* 2009; Perez *et al.* 2012), GadC (Ma *et al.* 2012; Ma *et al.* 2013), and AdiC (Fang *et al.* 2009; Gao *et al.* 2009; Gao *et al.* 2010; Kowalczyk *et al.* 2011). 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 (Shaffer *et al.* 2009) 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. Protein Data Bank (PDB) access codes are indicated in parentheses. (Kowalczyk *et al.* 2011).

This structure has brought insights about the substrate-induced fitting mechanism of AdiC. TMD1, TMD3, TMD6, TMD8 and TMD10 are disposed to the center of the transporter surrounding the substrate, Arg (Gao *et al.* 2009; Gao *et al.* 2010). Upon Arg binding, TMD2, TMD6 and TMD10 undergoes conformational changes. TMD1 and TMD6 have a helix unwounded segment dividing the helices in two sub TMD segments denominated as, TMD1a and TMD1b, and TMD6a and TMD6b, respectively. The initial recognition of Arg by AdiC from outside the cell, implies the interaction of the α-amino carboxyl moiety with residues in the unwound region of TMD1 and TMD6 (Figure 3C-E) (Kowalczyk *et al.* 2011). Π-cation interaction of the guanidinium group of Arg with residue Trp293 (TMD8), and hydrogen bonding with residue Asn101 (TMD3), are necessary to induce the occlusion of the substrate (Figure 3F-E) (Kowalczyk *et al.* 2011). Arg occlusion occurs by 40° tilting of TMD6a, which positions residue Trp202 over the hydrophobic part of the substrate, and by 10° pivoting of TMD10, which further stabilizes Trp293-Arginine interaction by a hydrogen bound between the guanidinium group of Arg and Ser357 (TMD10) (Gao *et al.* 2010) (Figure 3). Despite this knowledge about the initial stages of the transport, the complete mechanism of transport (e.g.,

translocation) remains unknown, because crystal structures of the inward-facing states are missing.

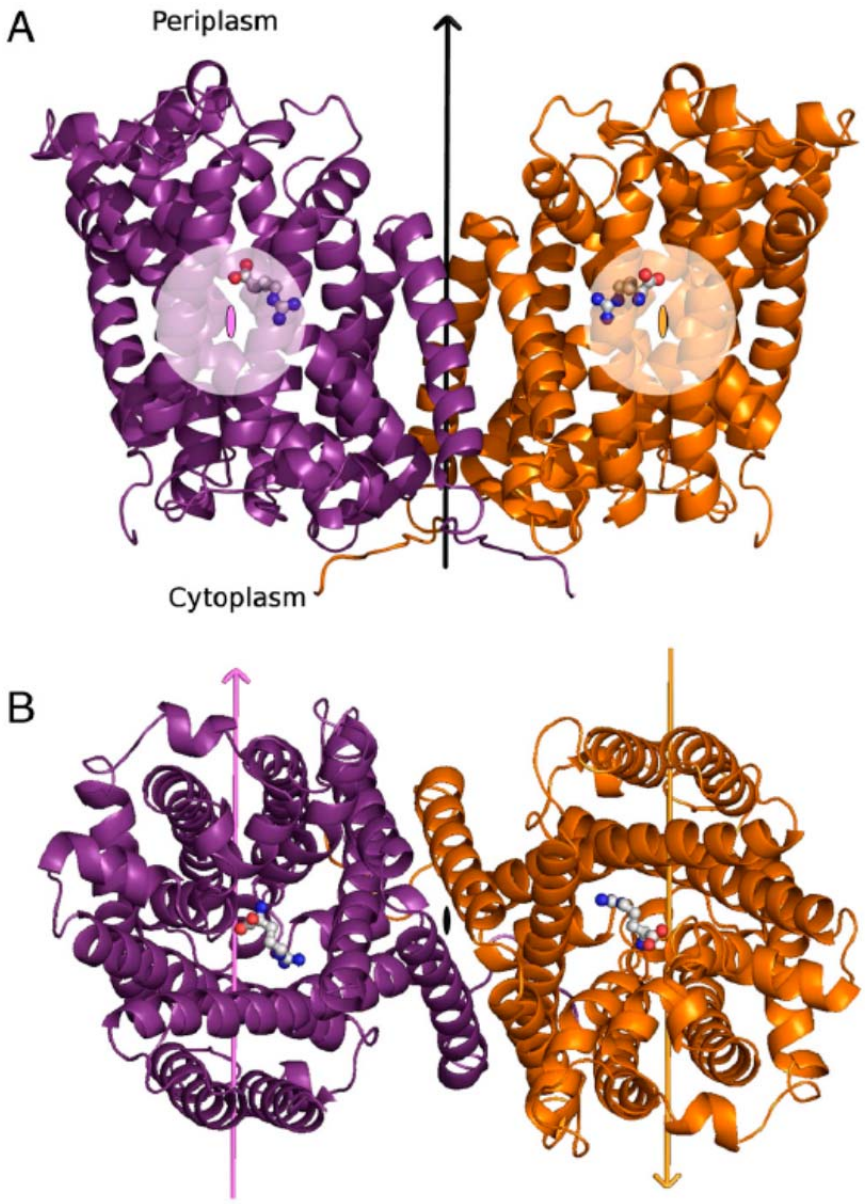

Figure 2. Structure of AdiC-N101A bound to Arg^+ (PDB: 3OB6). (**A**) Lateral and (**B**) periplasmic views of the AdiC-N101A-Arg^+ complex homodimer (in purple, protomer 1 and in orange, protomer 2). The bound substrate is depicted with a ball and-stick model. The two-fold subunit axes (ovals or arrows) are colored as their corresponding protomers. The dimer axis (arrow or ovals) is indicated in black. Adapted figure (Kowalczyk *et al.* 2011).

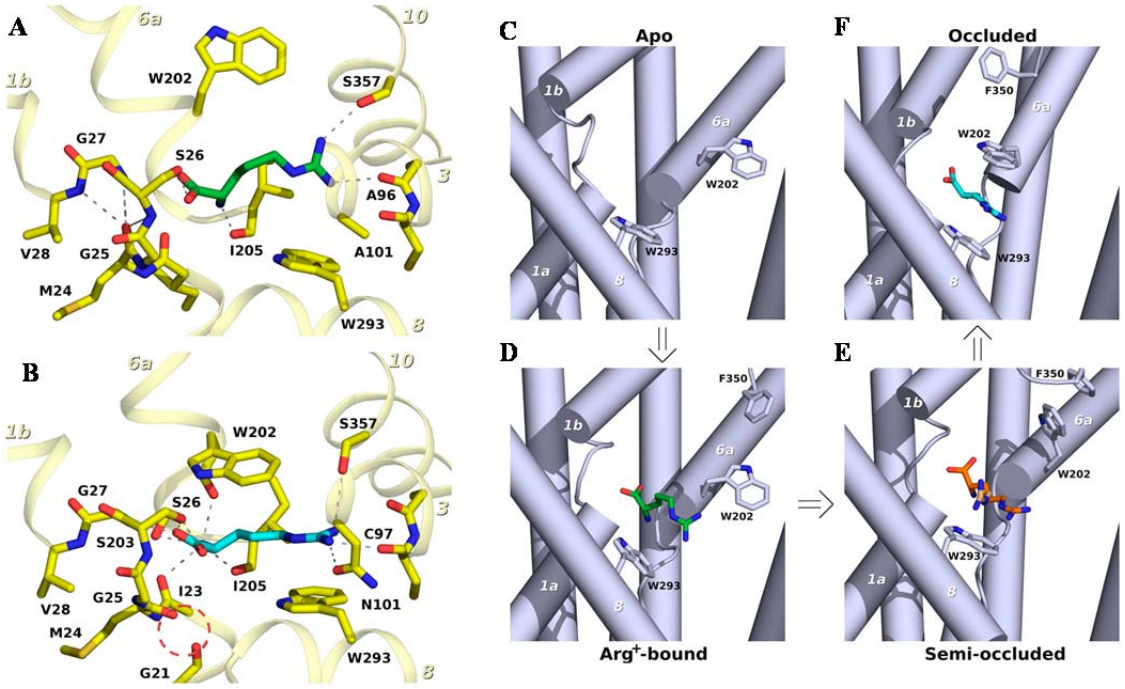

Figure 3. Transition from the apo to the semi-occluded state in AdiC. (**A**) Open-to-out conformation, where Trp202 and Trp293 are separated, a characteristic of this state. Other bulky residues have been included for density quality assessment. Arg^+ coordination in the AdiC-N101A structure is shown. The Arg^+ α-amino group is next to the negatively charged end of TMD6a helix dipole, and at hydrogen-bond distance of Ile205 carbonyl group (unwound segment of TMD6). The α-carboxy group lies in the vicinity of the positively charged end of TMD1b helix dipole and participates in a hydrogen bond with Ser26 side chain (unwound segment of TMD1). The guanidinium group is 4.6 Å away from Trp293 (TMD8) and its nitrogen atoms are at a hydrogen-bond distance of Ala96 carbonyl group (TMD3) and Ser357 side chain (TMD10). (**B**) Same view as in (A) of the Arg^+-bound occluded conformation of AdiC (3L1L (Gao et al. 2010)). The outward-facing Arg-occluded conformation (3L1L) differ in the position of the Arg^+ guanidinium group as well as in the unwound segment of TMD1 (Ile23 to Ser26), where the carbonyl group of Met24 is now further stabilized by hydrogen bonds with residues Ser26, Gly27 and Val28 (**B**), thereby preventing the collision with Gly21 (dashed circle). Periplasmic Arg^+ is recognized by the apo conformation of AdiC (**C**; 3NCY), and binds with a similar orientation (**D**; 3OB6) as in the Arg^+-occluded conformation (**E**; 3L1L). The proper Arg^+ binding, samples the semi-occluded state (**F**; docked Arg^+ in 3LRB) by stabilizing Trp202 (TMD6a) and Phe350 (loop TMDs 9–10) interaction. This semi-occluded conformation evolves to the occluded state mainly by pivoting TMD6a. Transition from the apo (**C**) to the semi-occluded state (**E**) is defective in mutant N101A. TMD segments are numbered in italics. Adapted figure (Kowalczyk *et al.* 2011).

In this regard, it is fundamental to solve inward-facing structures of AdiC, since furthermore the molecular mechanism of transport has not been fully revealed yet for any exchanger. By default, in the absence of substrate, AdiC crystallizes in the outward facing conformation (Fang *et al.* 2009; Gao *et al.* 2009). However, the combination of site-directed mutagenesis and substrate has allowed the resolution of crystal structures of the transporter bound to a substrate at different conformations. Thus, mutant N22A (Gao *et al.* 2010) rendered the outward facing arginine bound, and occluded structure, whereas mutant N101A rendered the open to out arginine bound structure (Kowalczyk *et al.* 2011). All these data validate the use of mutants as a strategy to get different conformational states from a particular transporter. In this regard, the use of mutated versions of transporters has been also used by Eric Gouaux and coworkers to solve the structure of LeuT in the inward facing conformation (Krishnamurthy and Gouaux 2012). In similar way, Christine Ziegler and collaborators obtained the structure of the inward-facing

conformation of other prokaryotic LeuT fold transporter, the Na^+/betaine transporter BetP, mutating one residue that change the specificity for substrate (Perez *et al.* 2011).

The pseudo-two fold symmetry of LeuT fold transporters can be used to model the symmetrical conformations (i.e., outward- and inward-facing related conformations) (Forrest *et al.* 2008; Kowalczyk *et al.* 2011). In the AdiC inward-facing arginine occluded model (open-to-in), fabricated from the outward-facing (open-to-out) Arg-occluded state (Gao *et al.* 2010), a new substrate cavity close to the cytoplasm is generated (internal cavity, open-to-in in Figure 4) (Kowalczyk *et al.* 2011). This is in agreement with a possible interaction of the substrate with Asn22 (TMD1) in this internal cavity. This fact would explain why Gao and collaborators crystallized AdiC, in the Arg-bound outward-facing occluded state, using mutant N22A. Thus, disruption of a putative interaction between Arg and Asn22 might have facilitated the outward-facing Arg-bound occluded state (Gao *et al.* 2010). Modeling of the inward-facing conformations of AdiC, would allow to analyze possible strategies to stabilize and crystalize this transporter in inward-facing conformation, for instance by point mutation of the internal and external binding sites.

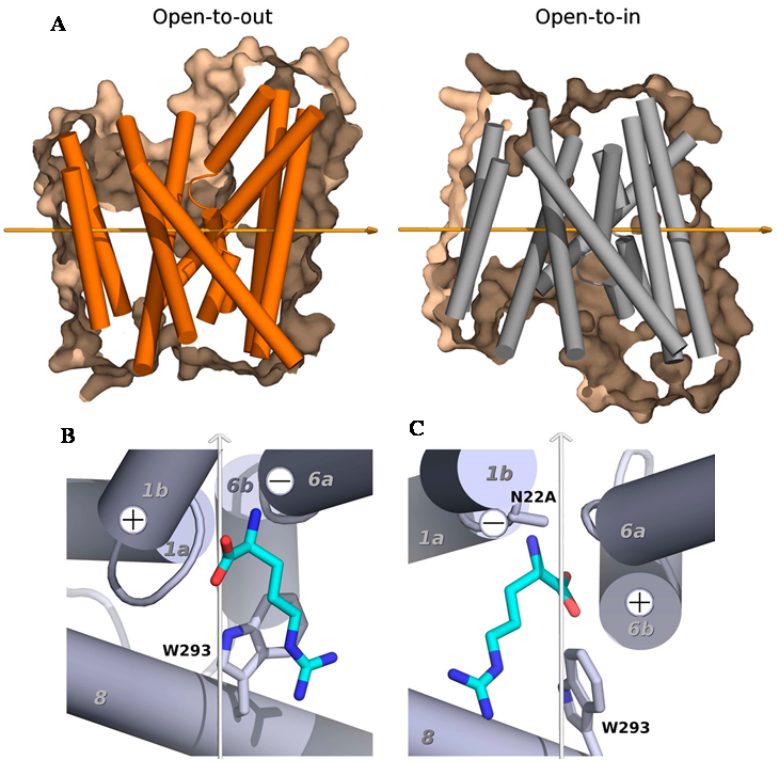

Figure 4. Transition from the open-to-out to the open-to-in conformation in AdiC. The model of the first 10 TMD segments of the open-to-in conformation of AdiC (gray cylinders), was built by swapping the conformations of the two *5* + *5* inverted repeats of the open-to-out structure (orange cylinders), by a nearly pure rotation along the subunit axis of the two repeats (light orange arrow). Both the bundle (TMDs 1, 2, 6, and 7) and the hash (TMDs 3, 4, 8, and 9) domains undergo major conformational changes during this transition. TMDs *5* and 10 undergo minor changes. (**B** and **C**) Proposed mechanism of Arg^+ translocation by AdiC. An Arg^+ internal binding site model (**C**) can be generated after applying the *5* + *5* inverted repeat symmetry to the external binding site of the occluded state (**B**; N22A mutant, 3L1L). Rotation of Arg^+ through the subunit axis (arrow) results in its turning upside down to this new internal cavity (**C**). Trp293 (TMD8) moves along with the guanidinium group of Arg+. The α-amino group of Arg+ in the internal site sits next to position 22. Positive and negative signs indicate the dipole charge of the corresponding TMD ends. TMD segments are numbered in italics. Adapted figure (Kowalczyk et al. 2011).

Objectives

1. To solve the structure of AdiC in inward-facing conformations:

    1.1. To generate and crystallize AdiC mutants designed to stabilize the transporter in inward-facing conformations.
    1.2. To crystallize wild-type AdiC by the High Lipid-Detergent (HiLiDe) method as an alternative strategy for the crystallization of the inward-facing conformation.
    1.3. To analyze, designed and test other alternative strategies to stabilize inward-facing conformations of AdiC, to be valid for solving the open-to-in atomic structure by crystallography.

Results and Discussion

Looking for mutants that stabilize the inward-facing conformation of AdiC

AdiC N22D mutant

The AdiC inward-facing arginine occluded model (open-to-in), was generated by Kowalczyk and co-workers by rotating the TMDs of the outward-facing (open-to-out) Arg-occluded state structure (Gao *et al.* 2010) along the subunit axis. Then, every TMD from the *5 + 5* inverted repeat (TMDs 1–10) were replaced by the corresponding pseudosymmetric TMD (e.g., TMD3 in place of TMD8 rotated), keeping the position of the subunit axis invariable (Kowalczyk *et al.* 2011). This model was validated when the same methodology applied to Mhp1 and LeuT open-to-out conformations, leads to open-to-in models resembling the inward-facing structures of both transporters, very approximately in the case of the Mhp1 (Forrest *et al.* 2008; Kowalczyk *et al.* 2011). Docking studies of Arg^+ in the AdiC inward-facing model, showed substrate in an internal binding cavity, newly generated in this open-to-in model, towards the cytoplasmic side (Figure 4 and Figure *5*B). In this internal cavity, Arg^+ is positioned close to Asn22 (TMD1), suggesting an interaction of the substrate with this residue, which would explain why when Gao and collaborators mutate Asn22 for Ala, crystallized AdiC in the Arg-bound outward-facing occluded state (Figure *5*A). Asn22 is one of the polar and charged conserved amino acids located in the central transport pathway (Gao *et al.* 2009). Thus, disruption of a putative interaction between Arg^+ and Asn22 might have facilitated the outward-facing Arg-bound occluded state. For this reason, looking for stabilizing AdiC in inward-facing conformation, we mutated Asn22 for Asp to strength the interaction of the substrate in this internal binding cavity.

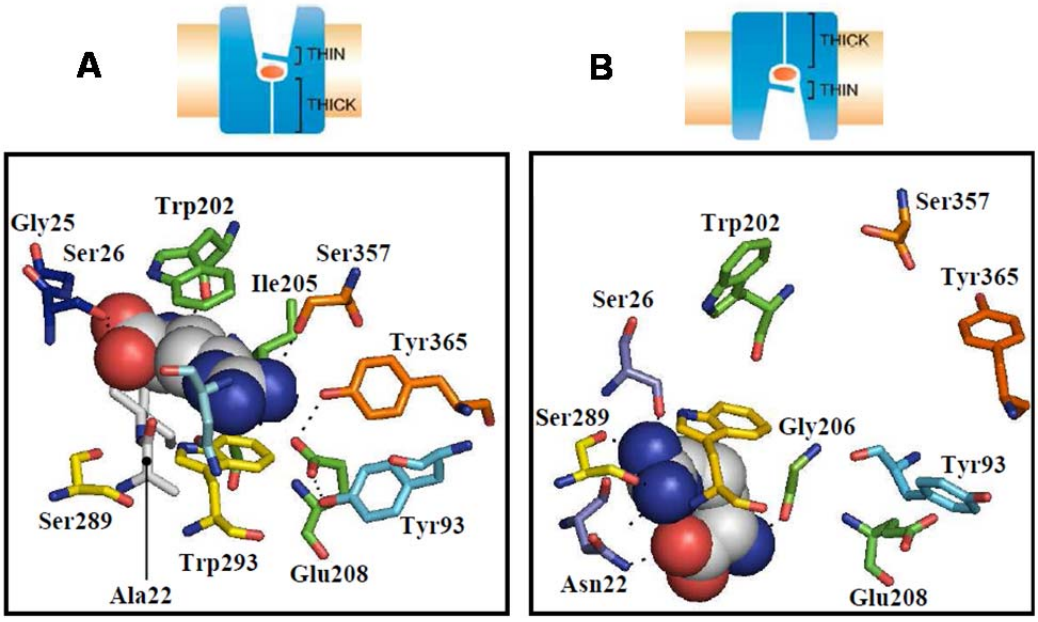

Figure 5. Position of Arg^+ in the binding site of two opposite symmetry related AdiC stages. A view of the external binding site of AdiC in the outward-facing Arg^+ bound occluded state obtained by mutating Asn 22 by Ala (A). A view of the docking of Arg^+ in the internal binding site of AdiC in the inward-facing model built from the outward-facing Arg^+ bound occluded state represented in A (B). The pass of Arg^+ from the external binding site to the internal binding site in the transition from the outward- to the inward-facing conformation during the substrate translocation is observed. Substrate leaves behind Trp293 (TM8). Changes in the position of some residues in the transport pathway are observed during this translocation. Arg^+ is represented in spheres and residues of AdiC in sticks.

AdiC N22D mutant was expressed in *E. coli*, solubilized in Cymal-6 detergent, and purified by IMAC and SEC in the same detergent (see Materials and Methods section in this Appendix). The most symmetric part of the SEC peak (Figure 6A) was collected, concentrated to 3 mg/ml, and subjected to crystallization screenings in small drop by sitting-drop vapor diffusion at 20 °C (see Materials and Methods section in this Appendix). Crystal hits were found at basic (7.5-9.5) and acidic pHs (4.3-7.5), at 24-34 % of PEG 400 (Figure 6B), and one at 24 % PEG 350 MME (Methoxypolyethylene glycol 350). Crystals were reproduced in big drop by sitting-drop vapor diffusion at 20 °C at pHs 8.0 to 9.5, and at the same percentages of PEG 400 (24-34 %) (Figure 6C). Crystals grown in acidic pHs could not be reproduced in big drop. Crystals of N22D mutant diffracted between 3.8Å to 20Å, but showed high mosaicity. The best diffracting crystals grew when mixing equal volumes of protein and reservoir solution containing 0.1M Glycine pH 8.5 or 9.5, and 28 % to 32 % of PEG 400. Data sets obtained from crystals with the best diffractions allows determining the unit cell dimensions of the crystals (Table 1), which were very similar to those of the N101A mutant that was crystallized in an outward-facing conformation (Table 1) (Kowalczyk *et al.* 2011).

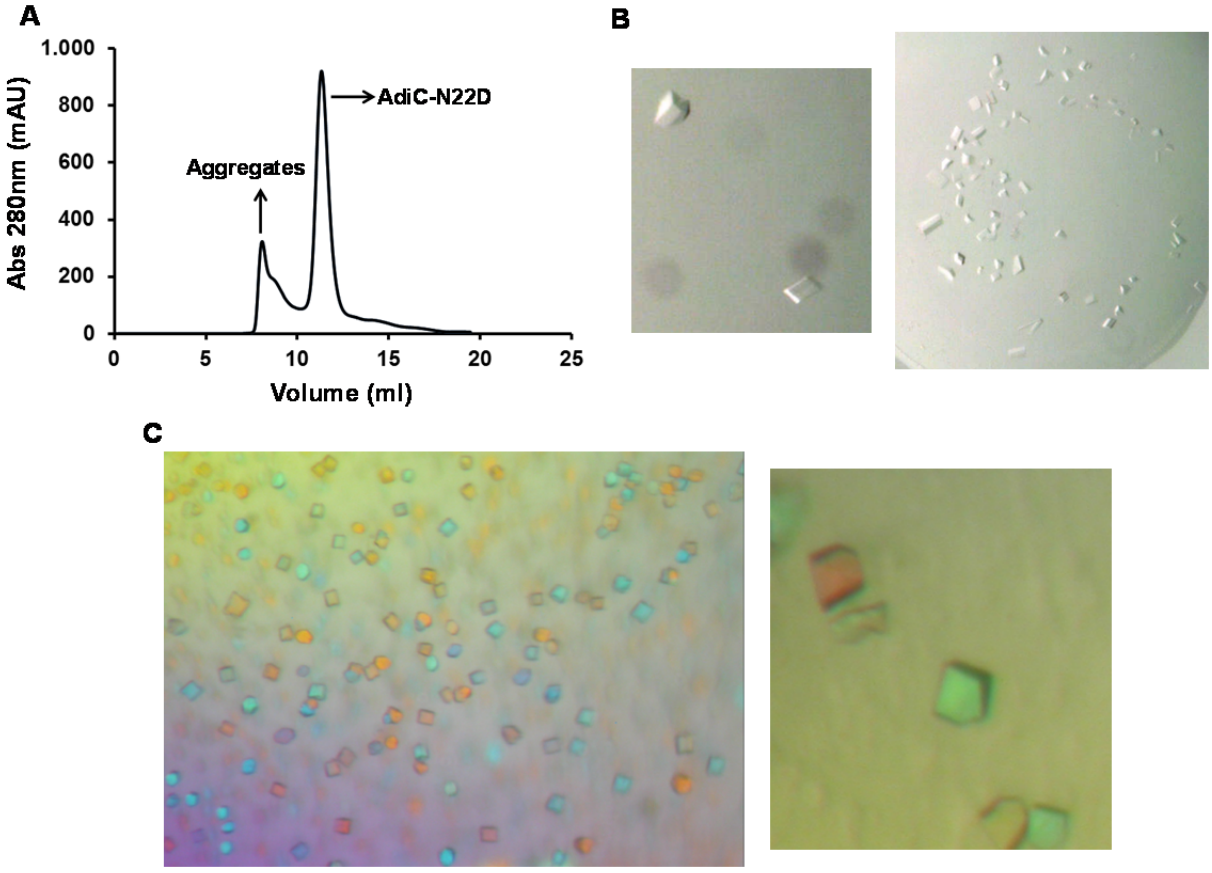

Figure 6. 3D crystallization of AdiC N22D mutant. Representative SEC chromatogram of N22D purified in 2xCMC Cymal-6 detergent (**A**). Some of the initial crystal hits obtained in basic pHs in small drop by sitting-drop vapor diffusion technique (**B**). Some of the crystals reproduced in big drop that showed the best diffractions (3.8 Å) (**C**).

AdiC S26A mutant

On the other hand, other of the polar and charged conserved amino acids disposed to the central transport cavity is Ser26, located in the helix-breaking GSG motif of TMD1 (Gao *et al.* 2009). In the Arg bound, with and without occlusion of the substrate, (mutants N22A and N101A, respectively), the α-carboxylate group accepts a hydrogen bond from the side chain of Ser26 (Figure 7) (Gao *et al.* 2010; Kowalczyk *et al.* 2011). In fact, mutation of Ser26 for Lys leads to abolishment of the transport function of AdiC (Gao *et al.* 2009). Additionally, Ser26, Gly27 and Val28 were found to stabilized the carbonyl group of Met24 by hydrogen bonds in the external binding cavity in the inward-facing Arg^+ bound state, previous to the occlusion of the substrate (Figure 7) (Kowalczyk *et al.* 2011). Thus, due to the putative stabilization of the substrate in the external binding cavity by Ser26, and then, its probable involvement in the stabilization of outward-facing conformation, we mutated this residue for Ala. In this way, it was probable favoring substrate fitting in the internal binding cavity, thus stabilizing inward-facing conformations.

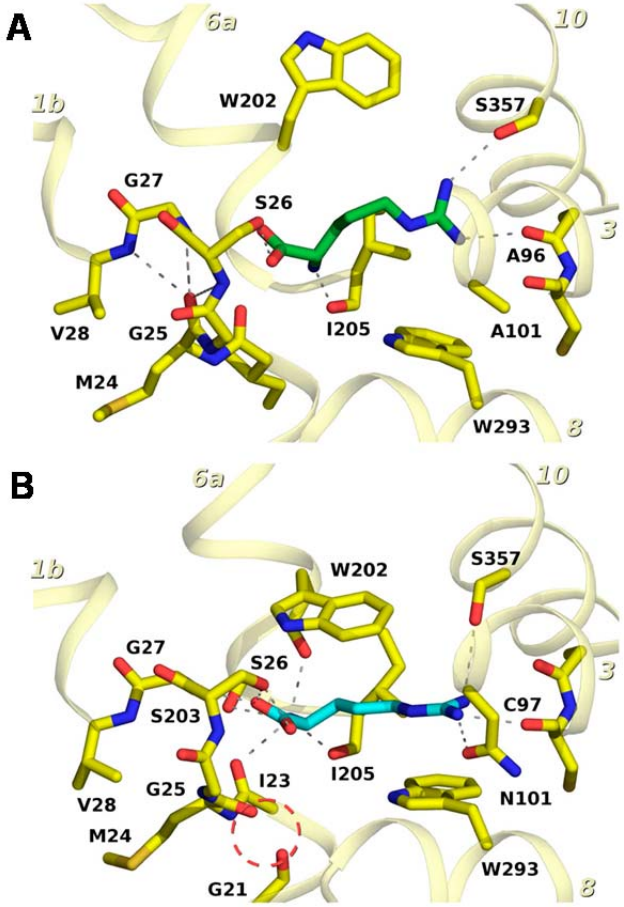

Figure 7. Arg^+ in AdiC open-to-out conformations. Arg^+ coordination in the AdiC-N101A structure (**A**). The Arg^+ α-amino group is next to the negatively charged end of TMD6a helix dipole, and at hydrogen-bond distance of Ile205 carbonyl group (unwound segment of TMD6). The α-carboxy group lies in the vicinity of the positively charged end of TMD1b helix dipole and participates in a hydrogen bond with Ser26 side chain (unwound segment of TMD1). The guanidinium group is ~4.6 Å away from Trp293 (TMD8) and its nitrogen atoms are at hydrogen-bond distance of Ala96 carbonyl group (TMD3) and Ser357 side chain (TMD10). (**B**) Same view as in (A) of the Arg^+-bound occluded conformation of AdiC (3L1L) (Gao et al. 2010). (A and B) differ in the position of the Arg^+ guanidinium group as well as in the unwound segment of TMD1 (Ile23 to Ser26), where the carbonyl group of Met24 is now further stabilized by hydrogen bonds with residues Ser26, Gly27 and Val28 (B), thereby preventing the collision with Gly21 shown in B (dashed circle). (Kowalczyk *et al.* 2011).

AdiC S26A mutant was expressed in *E. coli*, solubilized in Cymal-6 detergent, and purified by IMAC and SEC in the same detergent (see Materials and Methods section in this Appendix). The most symmetric part of the peak was collected (Figure 8A), concentrated to 3 mg/ml, and subjected to crystallization screenings in small drop by sitting-drop vapor diffusion at 20 °C (see Materials and Methods section in this Appendix). Less crystal hits were found for S26A than for N22D mutant. These were at 20 %-30 % PEG400, and at pHs of 4.0-6. In general terms these crystals were not well shaped nor stable. They were deteriorating quickly after forth day without reaching an appropriated size and shape. For these reasons, looking for improve crystal quality, a second screening in small drop was performed with protein from the SEC peak concentrated to 7 mg/ml. Crystals appeared at pHs 5.0-5.5 (Figure 8B), and seemed more defined and longer than those in the previous screening. However, when trying to reproduce those crystals in big drop, crystals start to deteriorate quickly and disappeared, indicating high instability. Further analysis of the screenings in small drop, showed that some crystals start to grow after 2 weeks,

and remains longer until one or two months at basic pHs (7.5-9.4), and 22-30 % PEG 400 or in 22 % and 26 % PEG 350 MME. These crystals at basic pHs were reproduced in big drop, but they were irregular and had grooves and breaks (Figure 8C), furthermore, it seemed that some of the S26A crystals were still unstable after 4 or 5 days of setting up crystallization drops. This means, that when crystals have not reached the maximum size, they start to deteriorate. Thus, other protein:reservoir ratio was tested, (1:2), trying to make the crystal to grow more slowly and maybe to form better and more stable crystals. Using this new condition, the crystals start to grow after 3 days, and they were better than the previous after 5 days of setting up the crystal drops, reaching its final size without disintegrate. Crystals of S26A mutant diffracted between 3.5Å to 22Å and with lower mosaicity than the crystals of N22D mutant. The unit cell dimensions of S26A crystals were determined for crystals with the best diffractions, and they were also similar to those of the N101A mutant (Kowalczyk *et al.* 2011) (Table 1).

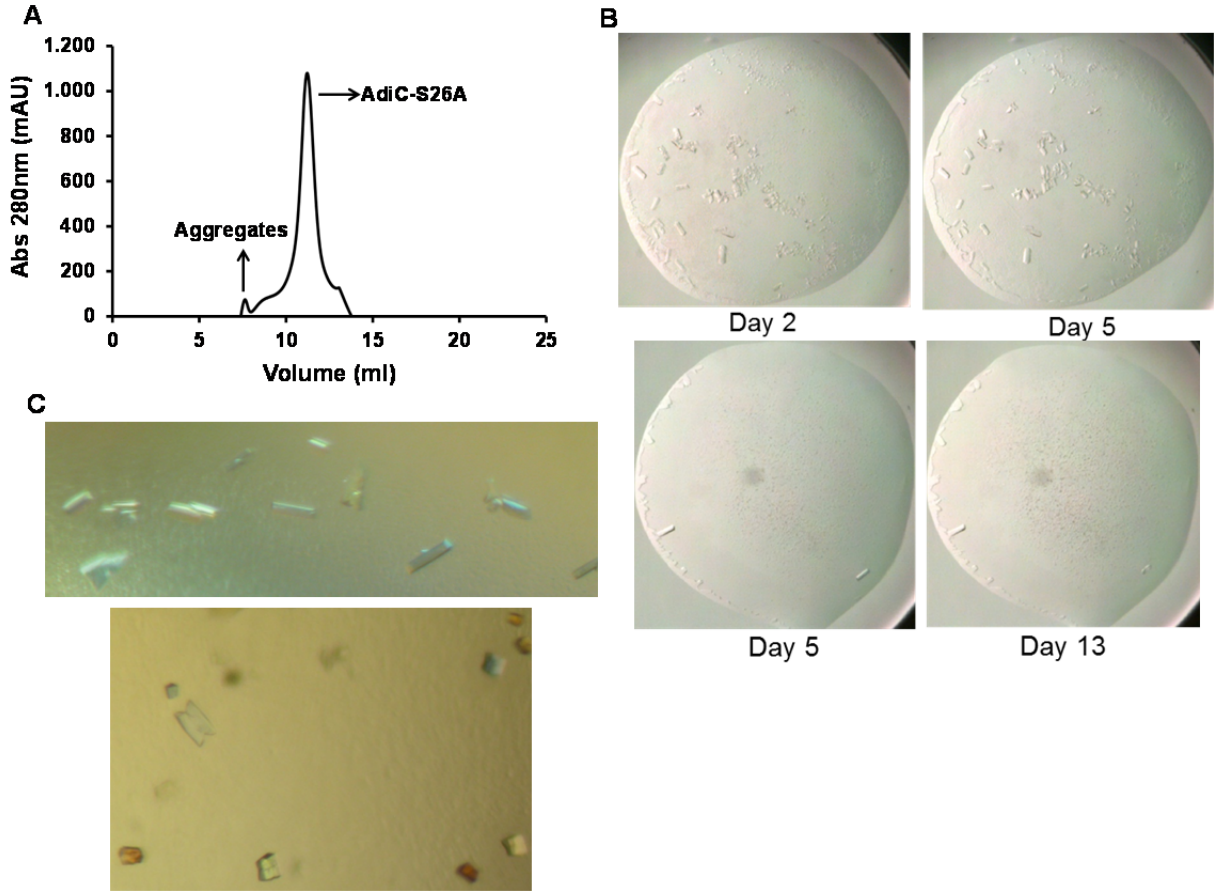

Figure 8. 3D crystallization of AdiC S26A mutant. Representative SEC chromatogram of S26A purified in 2xCMC Cymal-6 detergent (**A**). The protein showed the same SEC profile at 3 or 7 mg/ml. Some of the initial crystal hits obtained in the second screening in acidic pHs in small drop by sitting-drop vapor diffusion technique (**B**). Some of the crystals of basic conditions reproduced in big drop (**C**). Some of the crystals in the upper panel of C diffracted at 3.5 Å.

AdiC N22D-S26A mutant

We also tried to obtain crystals from the double N22D-S26A mutant, to combine the destabilization of the interaction of the substrate in the external binding cavity, and improving the interaction of AdiC in the internal binding cavity, looking to displace the equilibrium

towards the inward-facing stages. The double mutant was expressed, purified in Cymal-6 detergent, and subjected to crystal screenings in small drop, in the same way than N22D and S26A single mutants (see Materials and Methods section in this Appendix). Initial crystal hits of the double mutant N22D-S26A were obtained in acidic (pH 5.0-5.5) and basic (pH 9 and pH 10) conditions, and in 28-30 % of PEG 400 (Figure 9B). Crystals were reproduced in big drop, in acidic and basic conditions (Figure 9C and 9D, respectively), like needles and thin bars, being more predominant these shapes at basic pHs. In general terms, the crystals of the double mutant in the acidic conditions seemed to be more stable than those in acidic conditions of the S26A mutant. They also tend to be much more voluminous than the crystals of the basic conditions (Figure 9). All crystals of the N22D-S26A double mutant diffracted very poorly, by which not even the unit cell dimensions of these crystals could be obtained.

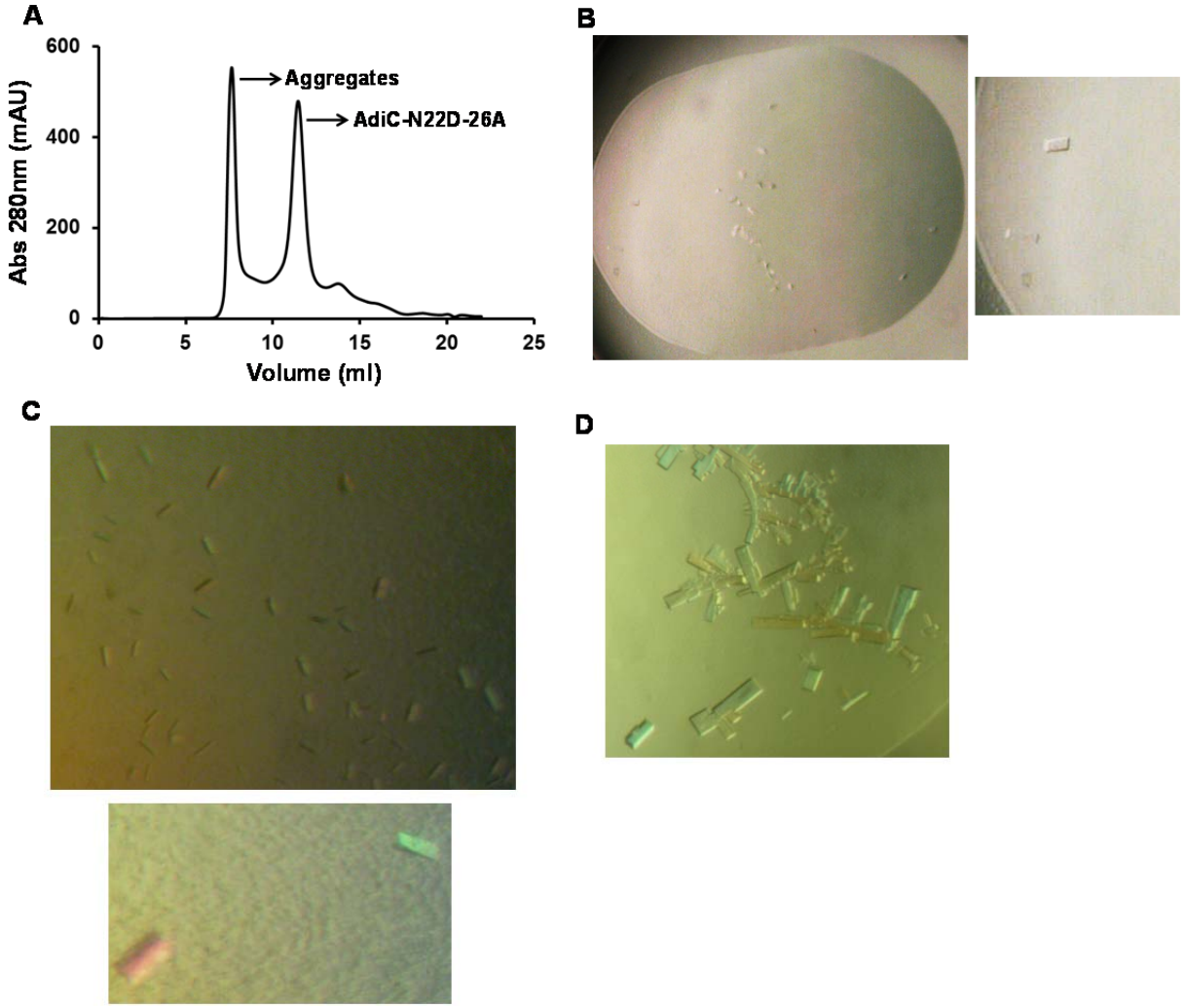

Figure 9. 3D crystallization of AdiC N22D-S26A double mutant. Representative SEC chromatogram of N22D-S26A purified in 2xCMC Cymal-6 detergent (**A**). Protein from the peak concentrated at 3mg/ml was used for seeding the crystallization drops. Some of the initial crystal hits obtained in small drop by sitting-drop vapor diffusion technique (**B**). Some of the crystals were reproduced in big drop in acidic (**C**) and in basic conditions (**D**).

Interestingly, N22D-S26A double mutant showed differences in the crystallization in compare to the N22D and S26A mutants that could be suggesting some kind of differences in the physicochemical properties, among the double and the single mutants. Such differences could

be attributed to a different conformation adopted by N22D-S26A. Crystals of double mutant grew more in acidic conditions than those of single mutants, and in general less crystal hits were obtained for the double mutant than for the single mutants. All the structures of AdiC outward-facing conformation have been obtained from crystals grown in basic conditions, and the physiological pH for AdiC function is around 2.5, therefore it is possible that the stabilization of inward-facing conformation could be obtained at acidic pHs, by which a mutant crystallizing in acidic conditions and different to the wild type or to previously studied mutants, that crystalize in outward-facing, could be interesting for being studied. On the other hand, in other studies with AdiC, crystals also grew in acidic conditions but showed the worst diffractions, compare with the crystals of basic conditions (Gao *et al.* 2009). Therefore, it is also posible that the growing of crystlas of AdiC mutants in acidic or basic conditions is not related with the stabilization of one conformation over the other, and that wild type AdiC and several mutants could crystalize in acidic conditions that produce poor diffractions. This suggests that acidic pHs might destabilice the protein or cause that this exists in different conformations.

Unit cell dimensions of N22D and S26A mutants indicate an outward-facing conformation

The similarity between the unit cell dimensions of the crystals of N22D and S26A mutants with the N101A mutant, crystalized in outward-facing Arg^+ bound conformation (Kowalczyk *et al.* 2011), could be suggesting that these two mutants are also crystalizing in outward facing conformation. This rationality could be supported by the differences in the unit cell dimensions, found between crystals of Mhp1 transporter in outward- (Weyand *et al.* 2008) and inward-facing conformations (Shimamura *et al.* 2010) (Table 1). For Mhp1, the major changes in the transition from outward-facing substrate occluded state to the inward-facing, are a rigid body movement of the hash domain (TMDs 3, 4, 8 and 9) relative to the bundle domain (TMDs 1, 2, 6 and 7) (Shimamura *et al.* 2010). However, the inward-facing model of AdiC obtained from the outward-facing substrate occluded state, proposed that in the transition from the outward- to the inward-facing states of this antiporter, there is a pivoting of both, the bundle and the hash domains (Kowalczyk *et al.* 2011). Then, it is probable than a transporter undergoing more profound changes during substrate translocation, between outward- and inward- facing conformations, will also form crystals with different crystal contacts in each of these different conformations, and thus also with different unit cell dimensions.

Additionally, crystals of these mutants diffracting at higher resolution (3.5-3.8 Å), diffracted with high mosaicity (N22A), and datasets were not good or complete enough to solve the structure of both mutants to corroborate their conformations. Several attempts to improve crystal quality were made, like screening different detergents as additives in the crystallization drop, similar to what Kowalczyk and co-workers used for obtaining crystals of N101A mutant, which diffractions allowed to solve the structure of this mutant at 3.0 Å resolution. Addition of a second detergent, in the crystallization drop, is useful to cover hydrophobic regions of membrane proteins that have not been completely shielded by the detergent used during the solubilization and purification steps. In this way, increased stability of the protein during crystals growing by elimination of hydrophobic unspecific interaction between protein molecules can be achieved, avoiding the formation of crystals with irregular forms or ruptures. In fact, cubic big crystals of N22D were obtained adding to the drop 1xCMC of LDAO (Figure 10), which is one out of two detergents used as additives to obtain the high quality diffracting crystals of N101A. Unfortunately, these crystals diffracted only between 14-26 Å. Similar results were obtained with S26A, whose crystals grown in drops with an addition of 1xCMC of Lauryl Maltose Neopentyl

Glycol or 1xCMC of DM (see Materials and Methods section in this Appendix), diffracted at 8Å or at 10-22 Å, respectively.

Protein	Unit cell dimensions a, b, c (Å)
AdiC-N101A	86.1, 77.2, 104.4
AdiC-N22D	85.5, 80, 105.5
AdiC-S26A	86.9, 77.6, 106.5
Outward-facing Mhp1	86.7, 108.0, 109.8
Inward-facing Mhp1	173.9, 173.9, 74.5

Table 1. Comparison of the unit cell dimensions of AdiC mutants crystals. Cell dimensions are very similar in all AdiC mutants (N101A, N22D and S26A). However, evident differences are observable between the crystals of the outward- and the inward-facing conformation of Mhp1, (Weyand *et al.* 2008) and (Shimamura *et al.* 2010), respectively.

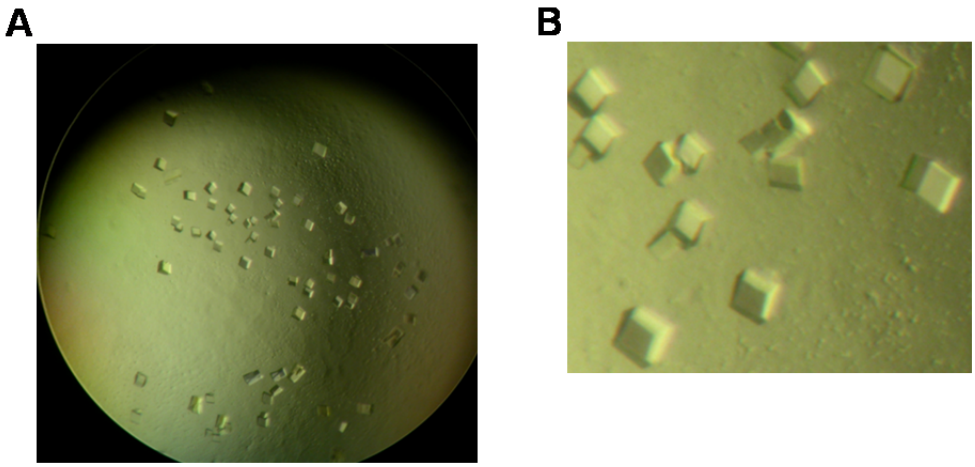

Figure 10. Crystals of N22D mutant purified in Cymal-6 with the addition of LDAO in the drop. Drop was composed of 1 µl of protein (3 mg/ml), 0.2 µl LDAO (10xCMC), and 0.8 µl reservoir solution (28 % PEG 400, Glycine pH 9.5). A close up view from the crystals in A is shown in B.

Therefore, because of the high probability that mutants N22D and S26A crystallized in outward-facing conformation like, wild-type AdiC, N22A (Gao *et al.* 2010) and N101A (Kowalczyk *et al.* 2011) mutants, and due to the datasets were not good or complete enough to solve the structure of both mutants, other strategies looking to stabilize AdiC in inward-facing conformation were investigated.

3D Crystallization of AdiC by the High Lipid-Detergent (HiLiDe) method

The use of lipids with detergents to solubilize and purify membrane proteins for crystallization is widely used, as lipids can increase the stability of the protein in solution and its crystallization, as well as, to improve crystal quality and diffraction (Hunte and Richers 2008). The residence of membrane proteins in solution in an environment, as much as similar to the membrane

environment, is responsible for such advantages of using lipids. Due to the membrane proteins are stabilize in a different way by the interactions with the lipids than when they are interacting only with the detergent, it is probable that membrane proteins were stabilized in a different conformation when lipids are present than when they are not. In fact, conformational changes for many membrane proteins, that can be essential for activity or regulation, comprise transmembrane regions (Palsdottir and Hunte 2004). Thus, the structural integrity of membrane proteins is affected by the lateral pressure imposes by the lipid bilayer (Palsdottir and Hunte 2004). This lateral pressure can be diminished by the destabilization caused by detergent solubilization, increasing conformational freedom of the protein (Palsdottir and Hunte 2004).

In fact, Gouaux and coworkers used, in addition to specific mutants and conformational monoclonal antibodies, a five-fold molar excess of 1,2-dimyristoyl-sn-glycero-3-phosphoethanolamine (DMPE) in order to obtain the inward-facing conformation of LeuT (Krishnamurthy and Gouaux 2012). Thus, the presence of lipids may also facilitate crystallization of AdiC in inward-facing conformation. Indeed, projection structure (PS) of lipid-embedded AdiC W293L mutant, showed differences in the position of TMDs proposed to pivot during the transition from the outward- to the inward-facing conformation in LeuT fold transporters, when compared to the PS of the outward-facing conformations of detergent-solubilized AdiC (Jeckelmann *et al.* 2011). In such study, 2D crystals of AdiC W293L mutant, obtained in DDM and *E. coli* polar lipids, were used to get the projection structure (PS) of this mutant at 6.5 Å resolution by Cryo-TEM (Casagrande *et al.* 2008). Projection structures (PSs) of outward-facing substrate-free (PDB code 3LRB) and Arg^+-bound occluded (PDB code 3L1L) AdiC, were calculated from the PDB-coordinates by using the projection structure visualization-method (PSV-method) validated in the same study (Jeckelmann *et al.* 2011). W293L mutant showed a strong density near TMDs 1, 2 and 6 (area 3 in Figure 11A), that was split into two weaker densities in the PS of the outward-facing substrate-free AdiC (area 3 in Figure 11B). These TMDs has been shown to undergo conformational changes during substrate occlusion (Gao *et al.* 2010), and, in fact, the PS of the outward-facing Arg^+-bound occluded AdiC , showed similar density in this area to the W293L PS, although in the substrate occluded state it is less protuberant (area 3 in Figure 11C). Interestingly, it seems to have changes mainly in the position of the TMDs 4 and 9 among the calculated PSs of AdiC outward-facing states (substrate-free and with substrate occluded), and the W293L PS (area 4 in Figure 11). These differences between detergent-solubilized (i.e., AdiC without substrate and with substrate bound occluded), and lipid-solubilized AdiC (W293L mutant), suggested that the latter could be adopting an inward-facing conformation, probably because the lateral pressure exerted on the protein by the lipid bilayer. For this reason, crystallization of wild-type AdiC in lipid was also addressed in the present work to stabilize AdiC in inward-facing conformation. Lipids can help also to improve crystallization of AdiC purified in Cymal-6 or to optimize the crystals.

A few years ago, Nissen and coworkers developed a new strategy for 3D crystallization of membrane proteins in the presence of high concentration of lipid and detergent (HiLiDe). This approach was designed to obtain crystals with good diffraction quality (Gourdon *et al.* 2011). The presence of lipids, produces crystals formed by continuous bilayers of the lipids with the membrane protein embedded by the TMDs. In comparison with other crystallization techniques using lipids, like lipidic cubic-, sponge- and bicelle-phase methods, HiLiDe is more practical since it does not require specific lipid compositions nor phase control, and it is not necessary to deal with the high viscosity of the sample or with the temperature dependence in the lipid bilayer states (Gourdon *et al.* 2011). Additionally, HiLiDe is completely compatible with crystallization vapor-diffusion technique, not as the other lipid techniques. Based on the

protocol described by Gourdon and co-workers (Gourdon *et al.* 2011), we performed HiLiDe, solubilizing and purifying wild-type AdiC in Cymal-6 detergent, and incubating the purified protein with *E. coli* polar lipids in four different protein/detergent/lipid ratios (see Materials and Methods section in this Appendix for detailed information).

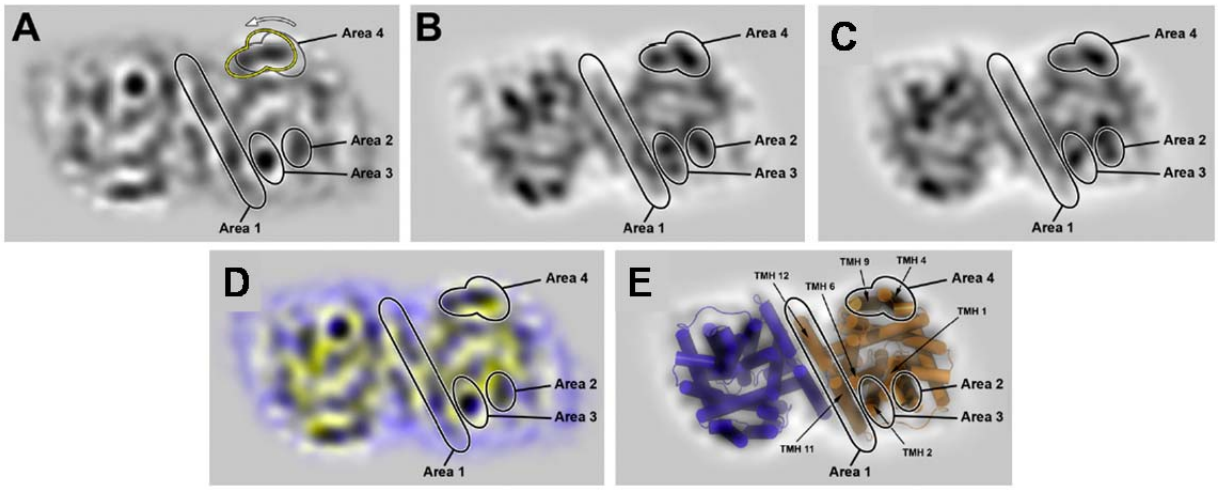

Figure 11. Comparison of the projection structures (PSs) between AdiC-W293L and apo and holo states of wild-type AdiC. (**A**) PS of AdiC-W293L, at 6.5 Å resolution, determined by cryo-TEM of 2D crystals (Casagrande et al, 2008). (**B**) Calculated PS of apo-AdiC at 6.0 Å resolution (PDB code 3LRB) (Gao et al., 2009). (**C**) Calculated PS of holo-AdiC at 6.0 Å resolution (PDB code 3L1L) (Gao et al., 2010). (**D**) Blue-yellow superposition of the AdiC-W293L cryo-TEM (blue) and calculated holo-AdiC (yellow) PSs. (**E**) Superposition of the calculated holo-AdiC PS with the corresponding X-ray structure shown as cartoon representation. Monomers are colored in blue and brown. Specific areas in the PS, and TMDs of AdiC, discussed in the Results and Discussion section of this Appendix, are contoured and labeled (Areas 1–4). The location of Area 4 differs in the calculated PS (black contour line) from that in the PS from cryo-TEM (yellow contour line). In all panels (A)–(D), only one of the two AdiC monomers is contoured and labeled. The panel frame sizes are 125 x 75 Å. Adapted figure and caption from (Jeckelmann *et al.* 2011).

Only spherulites were obtained in some of the conditions of the initial screening in small drop (results not shown). Due to the exclusive spherulites formation, that could be indicating excess of detergent and/or lipids respect to the protein concentration, and/or that the particular phase behavior of the different lipids negatively affected crystallization, other experiment with less detergent concentration (see Materials and Methods section in this Appenddix), and with a single lipid was performed. In fact, HiLiDe was standardized for using a single lipid and not a mix of lipids (Gourdon *et al.* 2011). The lipid 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) was used, and four different protein/detergent/lipid ratios were tested in an initial crystallization screening in small drop (see Materials and Methods section in this Appendix for detailed information). This time, small crystals like thin bars or irregular bars with split ends or ruptures grew in some conditions (Figure 12B-F). Some of these crystals were diffracted, showing low resolution but seemed to be protein crystals that can be improved.

Due to the low probability to obtain the structure of AdiC in inward-facing conformation, or new structural data by solving the structure of the AdiC mutants (N22D and S26A), and also due to the risk of not reaching some interesting result after a probable long time of improving the HiLiDe crystals of wild type AdiC, we decided in agreement with Dr. Manuel Palacín to pursue other project. This project is developed through the main part of this doctoral thesis. However, as a consequence of the studies and results showed in this appendix, and to the

analysis of them, ideas and guides about how to proceed with the studies to obtain the inward-facing conformation of AdiC were developed, and are shown in the next section of this Appendix (Perspectives).

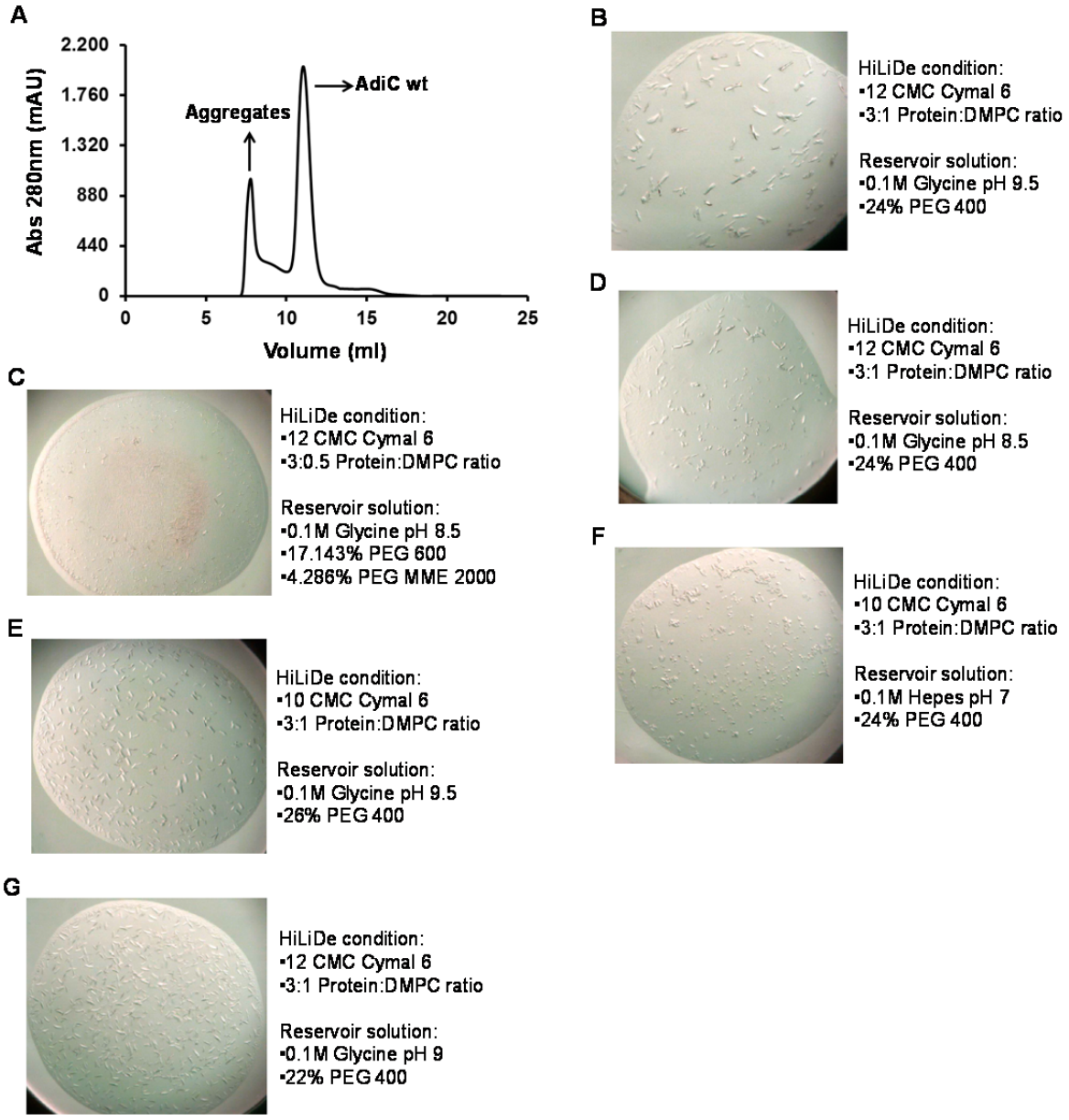

Figure 12. Initial crystal hits obtained for wild-type AdiC using HiLiDe. Representative SEC chromatogram of wild-type AdiC purified in 2xCMC Cymal-6 detergent (**A**). Pictures of the drops in which AdiC crystals grew (**B-G**). Commercial screenings in small drop were tested for the crystallization of AdiC relipidated with DMPC by HiLiDe. Sitting-drop vapor diffusion technique was used in all cases. Drops consisted of the same volume of reservoir solution and relipidated protein. Description of the HiLiDe and crystallization conditions is indicated for each crystal hit. DMPC: 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine.

Perspectives

3D crystallization of wild-type AdiC by HiLiDe to obtain the inward-facing conformation

To optimize the quality of the crystals obtained by HiLiDe, a screening of 96 detergent conditions (Hampton Research), as additives in the drop, could be performed, as proposed by Gourdon and coworkers (Gourdon, et al. 2011). As already mentioned, addition of a second detergent can help to enhance the formation of well-diffracting crystals without grooves or breakages. On the other hand, a screening in big drop changing the vapor diffusion strength, by varying the PEG concentration, and the ionic strength, by varying the concentration of the buffer, could be performed to scale up the crystals, looking also to improve them (Figure 13). For being able to determine whether, in fact, HiLiDe technique using DMPC is useful to solve the crystal structure of AdiC in a conformation different to those previously solved, probably inward-facing conformation, the characteristics and diffractions of further optimized HiLiDe AdiC crystals should be compared with wild-type AdiC crystals obtained in detergent and without lipids.

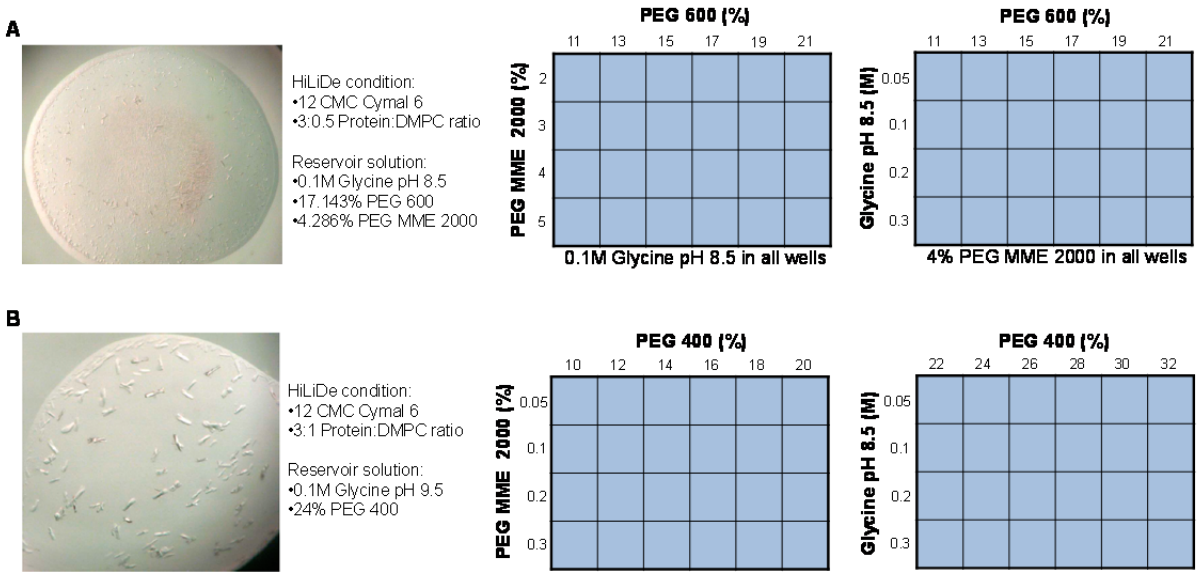

Figure 13. Proposed scale-up and optimization of AdiC HiLiDe crystals in big drop. A possible experiment to optimize, in 24 well-plate, by sitting drop (big drop) vapor diffusion technique, AdiC HiLiDe crystals obtained in small drops in the conditions showed in A and B. Description of the HiLiDe and crystallization conditions for the crystals showed (A and B) is indicated at the right of the crystals images. AdiC would be obtained in the same HiLiDe condition (A or B) and, in big drop, reservoir solution would be screened using related conditions as indicated in the 24-well plates represented in blue for each conditions (i.e, the two plates at the right of the indicated conditions, are the proposed screenings for optimization of the crystals in A and B, respectively). Usually (even whiteout applying HiLiDE) the crystallization drops consists of 1µl of protein and 1 µl of reservoir solution, but other ratios are also used.

It could be worthy to try the crystallization of the W293L mutant by HiLiDe to obtain AdiC inward-facing structure, or new mechanistic information of this transporter, since in this mutant, were observed the differences in the position of TMDs probably involved in the transition to the inward-facing state (Figure 11). Trp293 takes part of the thin gate closing the transport pathway towards the internal binding cavity sided to the cytoplasm. In the external cavity, in outward-

facing conformation, after Arg^+ binding, proper coordination of Arg^+-guanidinium group with residues Asn101 and Trp293 is required to the transition of AdiC to occlude the substrate in outward-facing states. The guanidinium group interacts with the side chain of Trp293 (TMD8), through a π-cation interaction that is essential for transport activity (Kowalczyk *et al.* 2011). Thus, W293L mutant is not able to bind Arg, by which was initially designed to diminish the number of possible conformations of AdiC for crystallization (Casagrande *et al.* 2008). Therefore, besides the proposed effect of the lipid lateral pressure in the stabilization of AdiC in inward-facing state, it is probable that the replacement of Trp293 by Leu opens the pathway towards the internal binding cavity, helping to stabilize inward-facing conformation. This also probably due to the abolished interaction of the substrate with the outward-facing (external) binding cavity.

Chemical cross-linking to lock AdiC in inward-facing conformation

Other way to obtain specific conformations of transporters, is the use of chemical cross-linking to lock the transporter in a single conformation. In the sodium/aspartate symporter from *Pyrococcus horikoshii* (Glt_{Ph}), a bacterial homologue of glutamate transporters from the SLC1 family, positioning of two cysteines in specific regions in a cys-less mutant lead to the stabilization of the inward-facing conformation. In this case, the experiments were performed by oxidative cross-linking in the presence of copper 1,10-phenantroline (CuPhen), or by bi-functional thiol-specific cross-linking with divalent mercury (Hg^{2+}) (Reyes *et al.* 2009).

In such study, cysteines were placed in specific positions of the TMD2 and in the extracellular hairpin 2. These positions were found 25 Å apart from each other in the outward-facing structure of Glt_{Ph}, but formed spontaneous disulfide bond when cysteines were placed in the corresponding residues in a glutamate transporter homologue, the excitatory amino acid transporter 1 (EAAT 1) (Ryan *et al.* 2004). Thus, in outward-facing conformation, the residues replaced by cysteines are at far distance between them as to form disulfide bridge, but seemed to be close in some conformation adopted by these transporters (Reyes *et al.* 2009). Therefore, with the aim of crystallizing AdiC inward-facing conformation, a similar cross-linking strategy could be implemented in a cys-less mutant. For doing this, comparison among AdiC outward-facing structures and inward-facing models was performed, looking for pairs of residues at far distances in outward-facing states, getting significant closer in inward-facing conformations.

A few years ago, the structure of the glutamate/c-aminobutyric acid (GABA) antiporter (GadC) from *E. coli* was solved in inward-facing conformation at 3.1 Å resolution (Ma *et al.* 2012). Same as AdiC, GadC belongs to the APC family of transporters and is required in the acid resistance systems (ARs) of enteric bacteria. Thus, similar to the Arg/Agm exchange by AdiC, GadC expels intracellular protons through the exchange of extracellular L-glutamate by its intracellular decarboxylated product, GABA. In the GadC inward-facing structure, there is a folded C-terminal plug that is not present in AdiC, blocking the internal binding cavity from the cytoplasm (Figure 14A). Additionally, an extensive network of H-bonds surrounds three residues of the loop 7 (L7), Ile265, Asn266, and Ser268, positioning L7 in the surface of three loops, L3, L9, and L11, in the periplasmic side (Figure 14B and 14C). The location of the c-plug and L7 seemed to be relevant features in the stabilization of the GadC inward-facing conformation. In fact, restriction of the movement of loop L7 by disulfide bond with TMD10, in residues in close proximity in inward-facing conformation (Figure 14D), leads to abolish substrate transport (Ma *et al.* 2012). This result indicated that L7 has to move from its inward-

facing position, during the substrate transport, suggesting that the position of L7 in this structure was essential to the stabilization of inward-facing state (Ma *et al.* 2012).

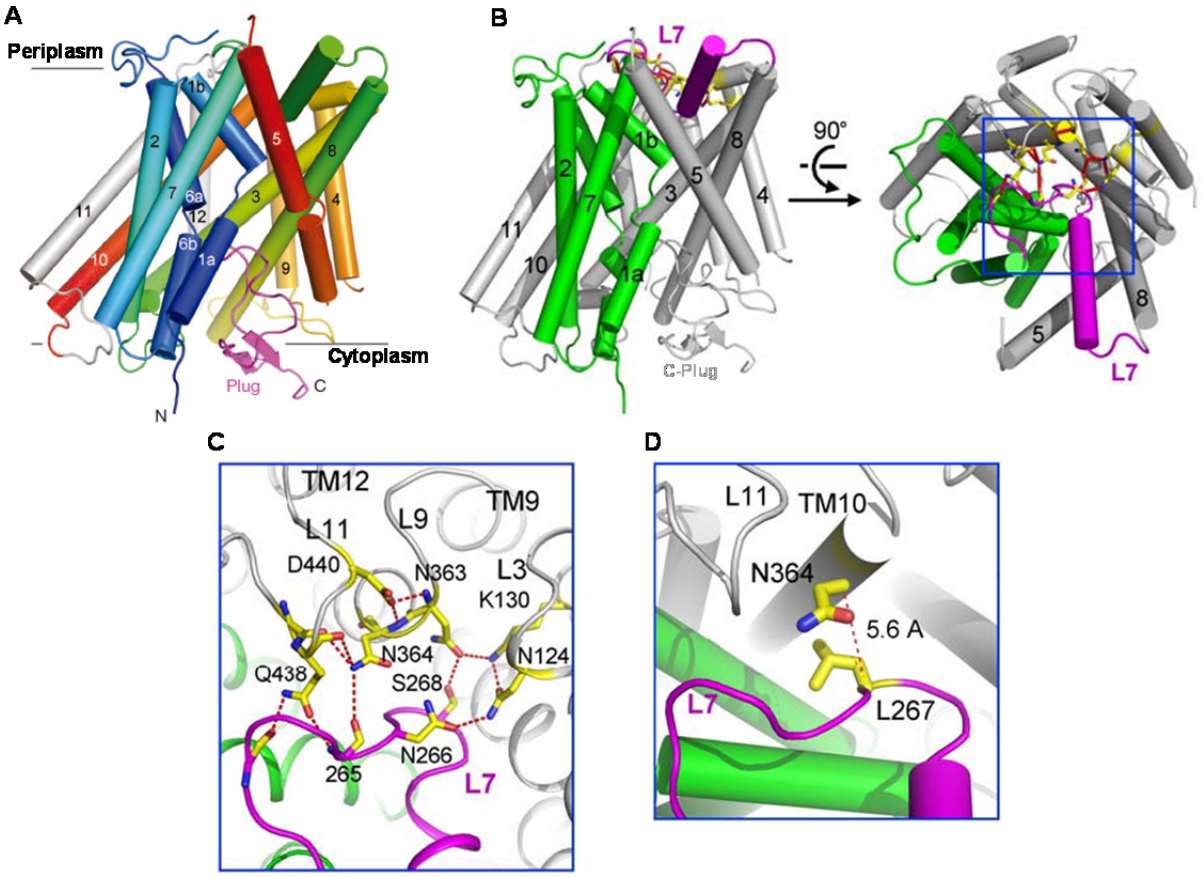

Figure 14. Overall structure of GadC and interactions of loop 7 stabilizing inward-facing conformation. (**A**) Overall structure of the wild-type GadC. TMD1–TMD10 are rainbow-colored, with TMD1 in blue and TMD10 in red. TMD11 and TMD12 are shown in grey. The C-terminal fragment (C-plug) is colored magenta. The C-plug of GadC blocks an otherwise inward-open conformation. (**B**) Two perpendicular views show the position of the L7 loop (magenta). Residues from L7 interact with amino acids from the L3, L9, and L11 loops. **(C)** A close-up view of the hydrogen bond networks surrounding the L7 loop (magenta). There is an extensive network of H-bonds surrounding three residues of the L7 loop: Ile265, Asn266, and Ser268. These H-bonds appear to glue the L7 loop with three surface loops: L3, L9, and L11. In particular, Asn266 accepts a H-bond from Asn124 on L3, whereas Ser268 donates a H-bond to Asn363 on L9. Asn124 and Asn363 each accept a H-bond from Lys130. (**D**) Restricting the movement of loop L7 relative to TM10 led to abrogated substrate transport. This panel shows the relative positions of Leu267 on the L7 loop and Asn364 on TM10. Figure and caption adapted from (Ma *et al.* 2012).

In the AdiC inward-facing model (Kowalczyk *et al.* 2011), generated from the Arg^+ bound occluded conformation (Gao *et al.* 2010), it could be observed a disposition of L7 (loop from TMD7 to TMD8) in the periplasmic side, similar to GadC, showing L7 close to TMD3 (L3 in the Arg^+ bound occluded structure) (Figure 15B). Serine 260 from L7 is at disulfide bond distance (5.04 Å) of Y114 from TMD3 (Figure 15B). These residues are 15.76 Å apart in the outward-facing occluded structure (Gao *et al.* 2010) (Figure 15A), the proposed state to precede the transition to open-to-in conformation of AdiC during substrate transport. Thus, a crosslinking between these two positions could help in the stabilization of inward-facing conformation. Additionally, Ala253 from L7 is also close to Ala181 from L5 (loop from TMD5

to TMD6), although at a farther distance (7.06 Å) than the before mentioned pair of residues (Figure 15B). However, as this is a model based on structural data from other conformation of the transporter, some deviation can be present, by which, cysteine cross-linking should be also tested in residues that seemed to significant move, getting closer during the outward- to inward-transition, despite not being at disulfide bond distance. For the same reason, testing the cross-linking of the equivalent residues in AdiC, described to form hydrogen bonds stabilizing L7 with L3, L9, and L11 in GadC (Figure 14C), should be also tested. Additionally, chemical cross-linkers exhibiting larger arms to cross-link residues at longer distances than 6 Å could be also used to ligate more distant residues, but that seemed to approach in open-to-in conformation. In fact, in AdiC inward-facing model, L9 is much father from L7 than in GadC, however, L9 is observed in AdiC as a long loop, by which, it can be flexible and movable, probably approaching to L7 in the native protein during transport.

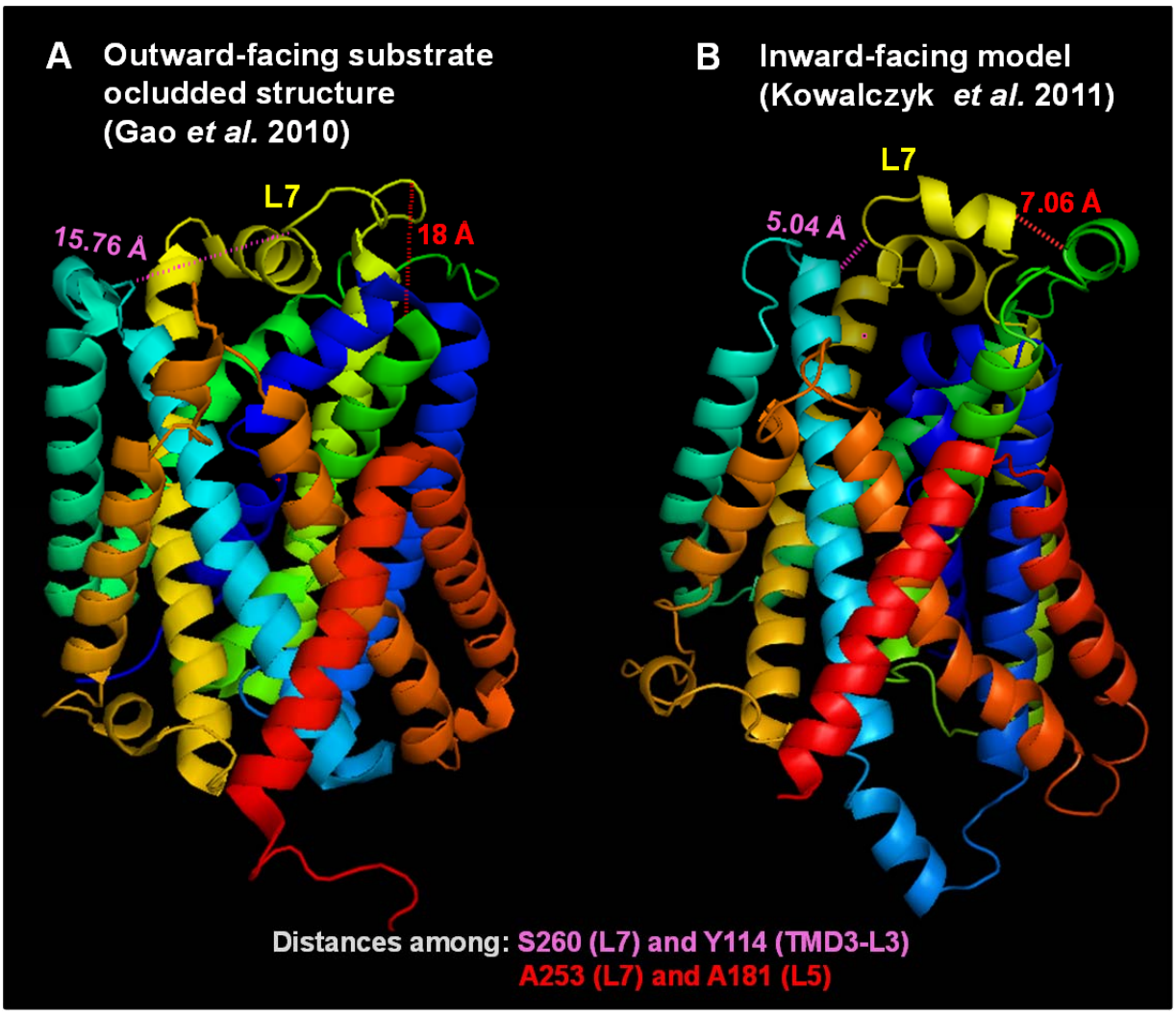

Figure 15. Position of potential pairs of residues to cross-link AdiC inward-facing conformation based on inward-facing model fabricated from the Arg+ occluded structure. Loop 7 (L7) is important in GadC to stabilize inward-facing conformation through its interaction with residues from loop 3, 9 and 11. In the AdiC inward-facing model (Kowalczyk *et al.* 2011) (B), fabricated from the outward-facing substrate bound occluded structure (A), similar to GadC, L7 is displaced towards the center of the transport pathway, in the periplasmic side, getting closer to the L3 (pale blue), but also to L5 (dark green), instead of L9 and L11. This is observed by the reduction of the distance among Ser260 from L7 and Tyr114 from TMD3-L3 (magenta dashed line and label), passing from 15.76

Å, in the outward-facing Arg^+ occluded state (A), to 5.04 Å in the inward-facing model (B). Additionally, the distance among Ala253 from L7 and Ala181 from L5 (red dashed line and label), is reduced from 18 Å (A) to 7.06 Å (B) in the proposed transition from outward- to inward-facing conformation. Therefore, such residues or neighbor residues could be good candidates for adding cysteines in a cys-less mutant for cross-linking experiments. TMD1–TMD12 are rainbow-colored, with TMD1 in dark blue and TMD12 in red. L7 is indicated in yellow. The color code for the distances showed, and the amino taken for performing such measurements between them, are indicated in the bottom of the figure.

In contrast with the predicted for AdiC, for GadC it was proposed that the most pronounced conformational change, in the transition from open-to-out to open-to-in conformations, is a rigid-body rotation of about 35° of bundle domain (TMD1, TMD2, TMD6 and TMD7). This observation was done when the GadC inward-facing structure was compared with the AdiC outward-facing structure. Thus, due to the similarities between both proteins, and for being able to analyze more precisely the possible positions to perform a cysteine cross-linking of the inward-facing state, with the structural data offered by GadC, a new AdiC inward-facing model based in the inward-facing structure of GadC was designed in collaboration with Dr. Guillem Portella Carbó (Cambridge University) (Figure 16B). This model, hereafter referred as AdiC like GadC model, was compared with outward-facing Arg^+ bound occluded AdiC structure (Gao *et al.* 2010), looking for pairs of residues significantly approaching between them in the inward-facing model (AdiC like GadC), compare to the outward-facing structure. In table 2 are listed residues moving more than 7 Å between both conformations, and being at a final distance of less than 9 Å in the AdiC like GadC model. In addition, in figure 16 are depicted the distances between residues moving more than 8 Å and being at a final distance of less than 6. As it could be observed in the table 2 and figure 16, all possible cross-linking positions are in close proximity between them.

In the lab, we already have the AdiC cys-less mutant from *Salmonella* kindly donated by professor Christopher Miller from the Department of Biochemistry, Howard Hughes Medical Institute, Brandeis University. *Salmonella* and *E. coli* AdiC share high identity between them (95 %), making the previous analysis performed for *E. coli* applicable to *Salmonella* homologue. The selected pairs of residues that are not at a final distance of less than 6 Å, in the inward-facing state, should be also tested for cysteine cross-linking, just in case the model present some deviation from the native inward-facing structure. A promising AdiC cross-linked version, to obtain inward-facing of AdiC, would be that showing spontaneous formation of disulfide bond and/or upon cross-linker treatment.

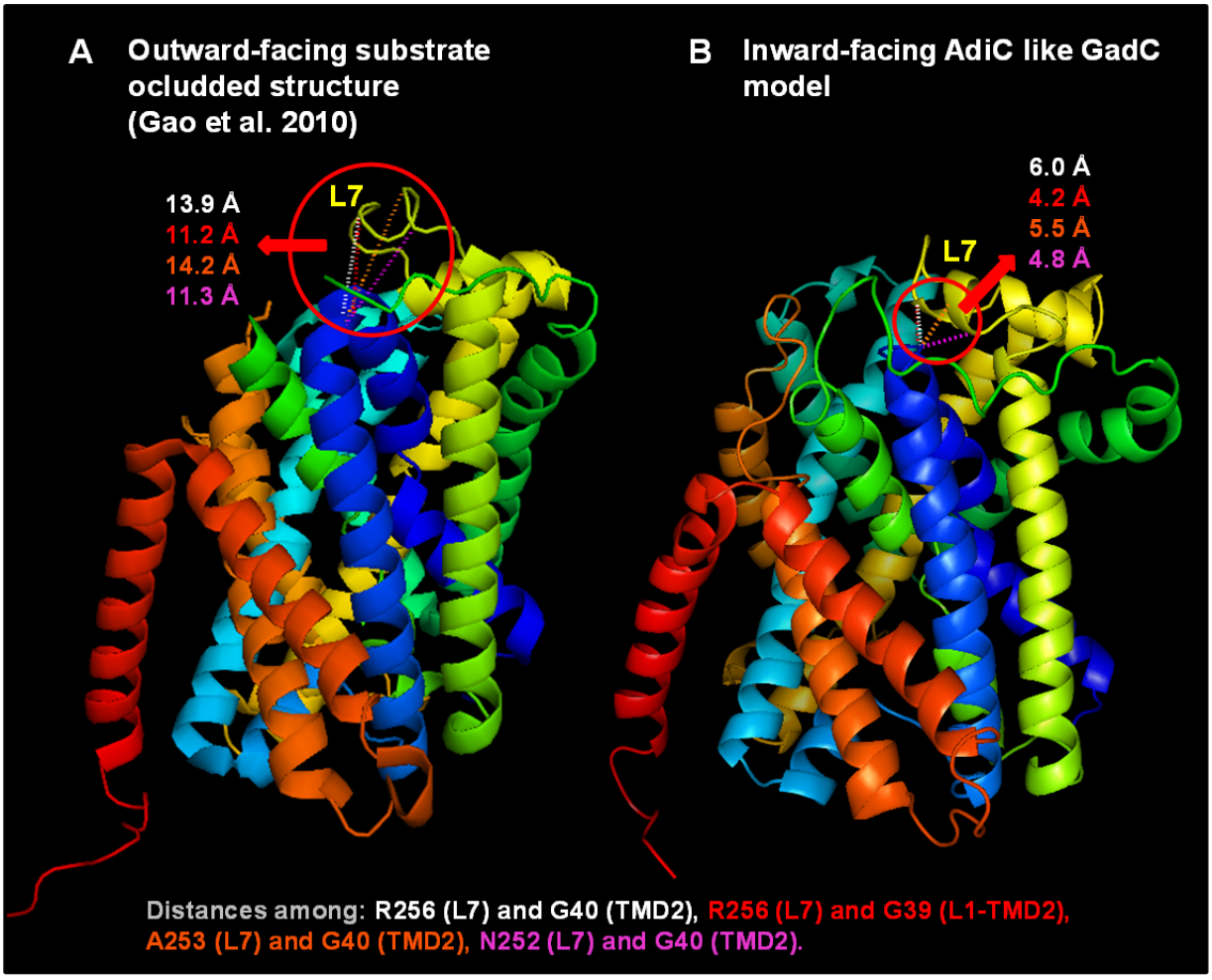

Figure 16. Position of potential pairs of residues to cross-link AdiC inward-facing conformation based on the AdiC like GadC inward-facing model. Loop 7 (L7) is important in GadC to stabilize inward-facing conformation through its interaction with residues from loop 3, 9 and 11. In the AdiC like GadC model (B), in contrast to GadC, L7 gets closer to the loop 1 and mainly to the TMD2 that follows (blue, from which distances encircled in red were measured). This is observed by the reduction of the distance among some residues in the proposed transition from the outward-facing (A) to the inward-facing state (B) (red circle), as indicated in the figure. Distances among, Arg256 from L7 and Gly40 from TMD2 (white dashed line), Arg256 from L7 and Gly39 from L1-TMD2 (red dashed line), Ala253 from L7 and Gly40 from TMD2 (orange dashed line), and Asn252 from L7 and Gly40 from TMD2 (magenta dashed line) are showed in each model. These distances are indicated by the red arrow in the respective line color code representing each of the measurements. Such residues or neighbor residues could be good candidates for adding cysteines in a cys-less mutant for cross-linking experiments. TMD1–TMD12 are rainbow-colored, with TMD1 in dark blue and TMD12 in red. L7 is indicated in yellow. The color code for the distances showed, and the amino taken for performing such measurements between them, are indicated in the bottom of the figure.

Residue 1	Residue 2	Distance in outward- facing Arg^+ bound occluded structure	Distance in inward-facing AdiC like GadC model
Ser37 (End TMD1)	Lys348 (L9)	21.3 Å	8.3 Å
Thr38 (TMD1-L1)		21.7 Å	8.9 Å
Gly40 (TMD2)	Asn252 (L7)	11.3 Å	4.8 Å
	Ala253 (L7)	14.2 Å	5.5 Å
	Ala254 (L7)	14.9 Å	7.8 Å
	Arg256 (L7)	13.9 Å	6.0 Å
Gly 39 (L1-TMD2)		11.2 Å	4.2 Å

Table 2. Pairs of residues getting closer in one of the proposed transitions to the inward-facing conformation. When looking for pairs of residues moving more than 7 Å, getting closer between them, and being at a final distance of less than 9 Å, in the transition from the outward- to the inward-facing conformation, the pairs of residues (residue 1 and residue 2) showed in the table were found. The models compared among them for doing this were the outward-facing Arg^+ occluded state (Gao *et al.* 2010) and the inward-facing conformation model, built based on the structure of inward-facing GadC (Ma *et al.* 2012) (AdiC like GadC model). The distance of the pairs of residues in each of the before mentioned models are indicated in column 3 and 4, as indicated. The residues that are at the greater final distances in the AdiC like GadC model (e.g., Ser37 and Lys348), exhibit the major displacement from one conformation to the other (more than 10 Å), as can be calculated from the differences of the distances in each conformation. In parenthesis is showed the location of the residues in TMDs or in loops (L) of AdiC.

Searching for molecules that block AdiC in inward-facing conformation

Another way to stabilize inward-facing conformation of AdiC, could be finding a molecule that binds to the inward-facing binding cavity and not to the outward-facing cavity, opening the transporter to the cytoplasm. This could be done in two ways, the first, by directing a methanethiosulfonate (MTS) reagent to the internal binding cavity of a version of AdiC with only one cysteine placed in a strategic place in the internal binding cavity. Similar to the strategy used by Dr. Miller and collaborators with AdiC (S26C), for avoiding the substrate transport by AdiC dimers orientated outside-out (impermeable MTS reagent), or inside-out (permeable MTS) reagent (Tsai *et al.* 2012). Moreover, an inward-facing structure of a single cysteine mutant of lactose permease of *E. coli* (LacY), very similar to the wild-type inward-facing structure, was crystallized and solved with MTS-galactopyranosides bound to the internal binding cavity (Chaptal *et al.* 2011). This, suggests the possibility of obtaining feasible structural data from a single cysteine mutant with a MTS reagent modified substrate.

The second approach, is to find molecules that bind to the inward-facing binding cavity with much higher affinity than to the outward-facing cavity. For doing this, docking experiments of compounds from the ZINC ("ZINC is not commercial") database (Irwin and Shoichet 2005; Irwin *et al.* 2012) in the outward-facing substrate occluded AdiC structure, and in the AdiC like GadC inward-facing model was done. In this way, compounds binding to AdiC like GadC, but not to outward-facing substrate occluded AdiC, were found, and their ZINC ID numbers are listed in table 3. They are ordered from the best to the worst as found in the docking experiments. The identity, characteristics and purchasability of those compounds can be obtained in the ZINC web page (zinc.docking.org) by introducing the ID number in the search

bar. In this way, it is possible to investigate which of them are appropriate for structural studies, addressed to solve the inward-facing structure of AdiC.

ZINC identification number	
10225162	12795499
12806065	12796341
12809209	13426236
12913440	10978343
12804935	11738722
12806047	12821217
13583970	12806074
12806870	12858581
12813545	05279033
12805546	12746855
18131926	05279036
22621974	12769852
12795497	17963843
17587657	64701303
12804940	64701306
14941049	64701309

Table 3. ZINC identification number for compounds found by docking analysis to bind to the hypothetical AdiC inward-facing conformation (AdiC like GadC model), but not to the outward-facing substrate occluded AdiC structure. They are ordered from the best to the worst behaving (from column one to column two). The compounds, their molecular structure and characteristics, as well as, their purchasability can be found in zinc.docking.org introducing the ID number in the search bar.

Other mutagenic approaches for the stabilization of AdiC in inward-facing conformations

Transporters mutated has been used to obtain other conformations of these proteins, different from those obtained with the wild type transporter or previously obtained by other means, like in the case of AdiC (Gao *et al.* 2010; Kowalczyk *et al.* 2011), LeuT (Krishnamurthy and Gouaux 2012) and BetP (Perez *et al.* 2011). Particularly, the AdiC substrate bound occluded (Gao *et al.* 2010) and the inward-facing substrate bound BetP (Perez *et al.* 2011) structures were obtained by introducing a single point mutation that changes the affinity or specificity of the transporter for substrates, respectively. Thus, it would be advisable to analyze in deep the transport and binding kinetics of N22D, S26A and the double mutant, to find if there are differences with the wild type transporter or previously studied mutants that lead to suggest if these mutants could offer new structural information, or conversely, to find evidences that support the similarity of these mutants with wild type AdiC and mutants that crystallize in outward-facing state. This would help to decide whether to continue or not with the improvement of N22D, S26A and the double mutant crystals for solving their structures. On the other hand, in the studies of AdiC and BetP, mentioned before (Gao *et al.* 2010; Perez *et al.* 2011), mutants stabilizing such conformations were found by screening the transport and binding of several mutants in conserved residues potentially important for the substrate recognition and translocation. For AdiC, such residues were those in the putative transport pathway that when mutated for specific residues could increase the affinity of AdiC for Arg (Gao *et al.* 2010), since the wild type AdiC crystalizes without substrate (Fang *et al.* 2009; Gao *et al.* 2009). Regarding BetP, an alanine scanning was performed in the glycine rich motif to investigate the role of such motif in the substrate specificity and ion coupling, finding that one of the mutations changes the specificity of BetP for betaine, towards choline (Perez *et al.* 2011). Indeed, the inward-facing conformation of this BetP mutant was solved with choline bound. Therefore, a similar strategy to find potential

mutants stabilizing AdiC inward-facing conformations should be implemented, like looking for mutants increasing the affinity of AdiC for agmatine, or modifying its binding properties for substrates.

On the other hand, if the sole use of mutants does not stabilize inward-facing conformation for solving the structure, a combination of mutagenesis together with other strategies could be investigated. These strategies could be chemical cross-linking, as described before, or the use of monoclonal antibodies that would further stabilize AdiC in inward-facing conformation. Eric Gouaux and coworkers reported a strategy to crystallize the inward-facing conformation of LeuT (Krishnamurthy and Gouaux 2012) that could be applied for AdiC. Thus, monoclonal antibodies raised against AdiC would be screened for positive interaction with the mutated versions of AdiC, and for negative interaction with wild type AdiC. For evidencing the interaction of the antibodies with AdiC mutants, Fluorescence Size Exclusion Chromatography (FSEC) could be used to detect the displacement of the chromatogram of the GFP-fusion mutants, caused by the antibody binding.

Conclusions

1. Although crystals of N22D and S26A mutants diffracted at a maximum 3.5-3.8 Å not all crystals of the same kind (i.e., same purification and crystallization condition) diffracted the same, and if further experiments with these mutants will be conducted the quality and reproducibility of the crystals must be improved. This could be aimed by using another detergent to solubilize and crystalize these AdiC mutants, or by using additives like lipids and/or detergents. Although, by the moment none of the detergent additives were effective to improve crystal quality and diffraction.

2. Unit cell dimensions of N22D and S26A crystals suggest that these mutants are crystalizing in outward-facing conformation. For this reasons studies with these mutants were suspended. However, if new evidences about the transport and binding kinetics of these mutants points to significant differences with wild type AdiC and the mutants that crystalize in outward-facing conformation could be interesting to solve the structure of these mutants. Additionally, because the unit cell dimension could be very similar between crystals of AdiC in outward-, and inward-facing conformation if the changes are not too pronounced among both conformations, such that the crystal contacts change.

3. Depending on the results of the above mentioned points, the improvement of the crystals of N22D-S26A double mutant could be attempted. Also, because good diffractions of these crystals were not obtained, then no useful information about the cell dimensions of these crystals is known. Thus, the potential of this double mutant to help in the stabilization of inward-facing conformation was not determined. This mutant could exhibit binding and transport properties different from the single mutants, that could favor whichever conformation over other.

4. Crystals of wild type AdiC obtained by HiLiDe should be improved, as proposed in this work, to solve the structure and to determine if the presence of lipid is stabilizing conformations of AdiC, different from those previously reported. Eventually, other lipids and detergents could be tested to improve these crystals.

5. Crystallization of W293L mutant, in the prescence or not of lipids (only detergent, HiLiDe or other lipidic tecniques), should be tested, since movemente of TMDs proposed to move in the transition from the outward- to the inward- facing conformation were observed in 2D crystals of this mutant solubilized in DDM and lipids.

6. It is possible that other strategies, in combination or not with mutated version of the protein, could be the key for solving the inward-facing conformation of AdiC, as performed with other transporters, like LeuT, Glt_{Ph} and the permease LacY. For them, combination of mutagenesis with the use of monoclonal antibodies (LeuT), or cross-linking (Glt_{Ph}), or a substrate linked to a MTS reagent, was determinant to obtain the inward-facing conformation of these proteins. To stabilize, in similar way, AdiC in inward-facing conformation, pairs of residues or close positions for a potential cross-linking were found here, as well as a list of molecules that could bind to the internal binding cavity of AdiC and then stabilizing the inward-facing conformation.

Materials and Methods

Wild type and AdiC mutants' expression

AdiC previously cloned into pTrcA vector (Invitrogen), with a His6 tag at the N-terminus followed by a target for Enterokinase enzyme, was used to introduce punctual mutations using the QuikChange site-directed mutagenesis kit (Stratagene) and the following pairs of primers for each mutant.

CGGGGAATATTATGGGGGCAGGTGTTTTTCTG	*AdiC S26A Forward*
CAGAAAAACACCTGCCCCCATAATATTCCCCG	*AdiC S26A Reverse*
CTGATGGTGTCGGGGGACATTATGGGGTCAGG	*AdiC N22D Forward*
CCTGACCCCATAATGTCCCCCGACACCATCAG	*AdiC N22D Reverse*
GATGGTGTCGGGGGACATTATGGGGGCAGGTGTTTTTCTG	*AdiC N22D-S26A Forward*
CAGAAAAACACCTGCCCCCATAATGTCCCCCGACACCATC	*AdiC N22D-S26A Reverse*

Mutations were confirmed by sequencing. *E. coli* BL21 strain (DE3) was transformed with pTrcAHis-His6-AdiC, mutants or wild type, and grown in LB media. Expression was induced at 0.6 optical density (A600) with 0.5 mM isopropyl-ß-D-thiogalactopyranoside (IPTG) at 37 °C. Bacteria were harvested by centrifugation (30 min, 4000 x g at 4ºC) after five hours, as previously realized with other AdiC mutants (Kowalczyk *et al.* 2011). Cells were resuspended in 20 mM Tris-HCl, pH 8.0, 0.5 mM EDTA buffer and disrupted using a cell disruptor (30.000 psi, Constant Systems L_{TD}). Cell debris were pelleted by centrifugation (1 h at 14.000 rpm, 4 °C) and the supernatant was subjected to ultracentrifugation (1-2 h at 60.000 rpm 4 °C). Membrane pellet was resuspended in 50 mM NaPi buffer pH 7.6, frozen in liquid nitrogen, and stored at −80 °C.

Purification of wild type and AdiC mutants

Membranes were solubilized chilled on ice in 1.5% of Cymal-6 (Affymetrix) for 3 h at 4 °C by gentle stirring. Then, samples were ultracentrifuged for 1 h at 100000 × g at 4 °C,. The supernatant was incubated in batch with TALON metal affinity resin (Clontech-Takara Bio Europe, Saint-Germain-en-Laye, France) for 3 h at 4 °C. Column was washed with 4 column volumes of 50 mM NaPi buffer pH 7.6, 20 mM imidazole, and two-fold the critical micelle concentration (2xCMC) of detergent (CMC Cymal-6=0.028%). After washings resins with bound protein was equilibrated in 50 mM Tris, pH 8.0, 1 mM $CaCl_2$ and 0.056% Cymal-6. The protein was eluted by cleaving the His6 tag with Enterokinase (Invitrogen) over night at 20°C. Protein was concentrated and subjected to Size Exclusion Chromatography (SEC) in 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.056% Cymal-6, using a Superdex 200 10/300 GL column (GE Healthcare, United Kingdom). Purified protein was concentrated to 5 (all mutants) or 10 mg/ml (S26A) and subjected to ON dialysis against the buffer mentioned before. Protein preparations were ultracentrifuged at 55000 rpm for 30 min at 4 °C, and soluble fraction was used at 3 mg/ml (all mutants) or 7 mg/ml (S26A) for drop seeding.

Crystallization of wild type and AdiC mutants

Initial screenings were always performed in small drop with Cartesian liquid handling robot, and using commercial screenings: MemSys, MemGold (Molecular Dimensions, Newmarket, England), MembFac (Hampton Research, Aliso Vijeo, USA), and pi-PEG. Additionally, optimization screenings (previously used with other AdiC mutants) were also tested except for the HiLiDe samples. Small drops mixing equal volumes of the protein and reservoir solution were seeded in 96 well plates by sitting-drop vapor diffusion technique, in the presence and absence of the substrate (arginine or agmatine), that was added to the protein just after drop seeding. Plates were incubated at 20°C. Screenings in 24 well plates were designed using the conditions in which the crystals grew as central condition, from which make variations of the concentration of the precipitant and/or the pH of the buffer, trying to improve the quality and size of the crystals. Sitting-drop and hanging-drop were seeded at 4°C and 20°C, by mixing equal or different volumes of protein and reservoir solution containing; 0.1 M Tris-HCl, pH 8.0–9.5 or 0.1 M glycine pH 8.5–9.5 and 24–30% PEG400 or 24-30% PEG350 monomethyl ether (MME). Sometimes arginine or agmatine were added to the protein at 2mM final concentration. Best diffracting crystals grew in both, presence and absence of substrate (arginine), although some crystals also grew in presence of agmatine. Crystals were cryoprotected by soaking the crystal in reservoir solution containing 30% PEG400, or the precipitant in which crystals grew, immediately they were flash-frozen in liquid nitrogen and diffracted at the Swiss Light Source (Paul Scharer Insitute, Villigen, Switzerland).

Screening of detergents additives in AdiC-mutants crystallization.

Screenings of 18 detergents as additives in the crystallization drop were made in sitting-drop vapor diffusion at 20 °C, by mixing 1 µl of protein (3mg/ml or 7mg/ml), 0.2 µl of the additive detergent (1 CMC) and 0.8 µl of reservoir solution in which crystals of the mutant have been previously obtained (Glycine pH 9.0 or 9.5 or Sodium Citrate pH 5.0 or 5.2 and 24-30% PEG400).

HiLiDe

Wild type AdiC was expressed and purified in Cymal-6 as described above, and the most symmetric part of the peak (Figure 12A) was collected and concentrated to 10mg/ml for HiLiDe. For the first experiment, *E. coli* polar lipids previously dissolved in chloroform (40 mg/ml), to avoid oxidation, were added to a glass tubes (recommended, 5-10 cm long and no more than 1 cm in diameter) under N_2 gas flow to form a film in the bottom of the tube with the required amount of lipid. Two protein/lipid ratios (w/w) were tested, 3:0.5 and 3:1, each in two different Cymal-6 concentrations, 14-fold and 18-fold the Cymal-6 CMC (0.028%). Cymal-6 dissolved in water at 40mg/ml was added to the protein to the concentration required. After complete evaporation of the chloroform, protein mixed with the excess of detergent, was added to the lipid film under N_2 flow, being careful not to splash the sample. A small magnet was introduced into the tube and the lid was tightly placed to close the tube. Parafilm was employed to fix the cap to ensure a closure as hermetical as possible. Samples were left by ~16 h in gentle stirring at 4°C, and then they were ultracentrifuged at 55000 for 30 min. Soluble fraction was subjected to crystallization screnning in small drop, as described above (Crystallization of wild type and AdiC mutants).

AdiC like GadC inward-facing model

The model was generated thanks to the collaboration of Dr. Guillem Portella Carbó (Cambridge University). AdiC (PDB: 3L1L) TMDs were separated in two groups: 1, 2 6 and 7 in one group, and 3, 4, 5, 8, 9 and 10 in other, since these groups were suggested as the two distinguishable domains (gate and core domains, respectively) of GadC and AdiC, when comparing the structure of inward-facing GadC with the outward-facing AdiC (Ma *et al.* 2012). By using Chimera software a structural alignment was performed to reorientate the TMDs of AdiC like those of GadC, conserving the relative position of the amino acids in each TMD. By using molecular dynamics, AdiC like GadC model was built from the TMD coordinates obtained in the previous step. Thus, AdiC TMDs and loops adopted the geometry of GadC.

Bibliography

Bartoccioni, P., C. Del Rio, M. Ratera, L. Kowalczyk, J. M. Baldwin, A. Zorzano, M. Quick, S. A. Baldwin, J. L. Vazquez-Ibar andM. Palacin (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.

Casagrande, F., M. Ratera, A. D. Schenk, M. Chami, E. Valencia, J. M. Lopez, D. Torrents, A. Engel, M. Palacin andD. Fotiadis (2008). Projection structure of a member of the amino acid/polyamine/organocation transporter superfamily. *J Biol Chem,* 283(48): 33240-33248.

Chaptal, V., S. Kwon, M. R. Sawaya, L. Guan, H. R. Kaback andJ. Abramson (2011). Crystal structure of lactose permease in complex with an affinity inactivator yields unique insight into sugar recognition. *Proc Natl Acad Sci U S A,* 108(23): 9361-9366.

Faham, S., A. Watanabe, G. M. Besserer, D. Cascio, A. Specht, B. A. Hirayama, E. M. Wright andJ. Abramson (2008). The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na+/sugar symport. *Science,* 321(5890): 810-814.

Fang, Y., H. Jayaram, T. Shane, L. Kolmakova-Partensky, F. Wu, C. Williams, Y. Xiong andC. Miller (2009). Structure of a prokaryotic virtual proton pump at 3.2 A resolution. *Nature,* 460(7258): 1040-1043.

Forrest, L. R., Y. W. Zhang, M. T. Jacobs, J. Gesmonde, L. Xie, B. H. Honig andG. Rudnick (2008). Mechanism for alternating access in neurotransmitter transporters. *Proc Natl Acad Sci U S A,* 105(30): 10338-10343.

Gao, X., F. Lu, L. Zhou, S. Dang, L. Sun, X. Li, J. Wang andY. Shi (2009). Structure and mechanism of an amino acid antiporter. *Science,* 324(5934): 1565-1568.

Gao, X., L. Zhou, X. Jiao, F. Lu, C. Yan, X. Zeng, J. Wang andY. Shi (2010). Mechanism of substrate recognition and transport by an amino acid antiporter. *Nature,* 463(7282): 828-832.

Gasol, E., M. Jimenez-Vidal, J. Chillaron, A. Zorzano andM. Palacin (2004). Membrane topology of system xc- light subunit reveals a re-entrant loop with substrate-restricted accessibility. *J Biol Chem,* 279(30): 31228-31236.

Gong, S., H. Richard andJ. W. Foster (2003). YjdE (AdiC) is the arginine:agmatine antiporter essential for arginine-dependent acid resistance in Escherichia coli. *J Bacteriol,* 185(15): 4402-4409.

Gourdon, P., J. L. Andersen, K. L. Hein, M. Bublitz, B. P. Pedersen, X.-Y. Liu, L. Yatime, M. Nyblom, T. T. Nielsen, C. Olesen, J. V. Møller, P. Nissen andJ. P. Morth (2011). HiLiDe—Systematic Approach to Membrane Protein Crystallization in Lipid and Detergent. *Crystal Growth and Design,* 11(6): 2098-2106.

Griffin, P. M. andR. V. Tauxe (1991). The epidemiology of infections caused by Escherichia coli O157:H7, other enterohemorrhagic E. coli, and the associated hemolytic uremic syndrome. *Epidemiol Rev,* 13: 60-98.

Hunte, C. andS. Richers (2008). Lipids and membrane protein structures. *Curr Opin Struct Biol,* 18(4): 406-411.

Irwin, J. J. andB. K. Shoichet (2005). ZINC--a free database of commercially available compounds for virtual screening. *J Chem Inf Model,* 45(1): 177-182.

Irwin, J. J., T. Sterling, M. M. Mysinger, E. S. Bolstad andR. G. Coleman (2012). ZINC: a free tool to discover chemistry for biology. *J Chem Inf Model,* 52(7): 1757-1768.

Iyer, R., C. Williams andC. Miller (2003). Arginine-agmatine antiporter in extreme acid resistance in Escherichia coli. *J Bacteriol,* 185(22): 6556-6561.

Jeckelmann, J. M., M. Palacin andD. Fotiadis (2011). A tool for the qualitative comparison of membrane-embedded and detergent-solubilized membrane protein structures in projection. *J Struct Biol,* 173(2): 375-381.

Khafizov, K., R. Staritzbichler, M. Stamm andL. R. Forrest (2010). A study of the evolution of inverted-topology repeats from LeuT-fold transporters using AlignMe. *Biochemistry,* 49(50): 10702-10713.

Kowalczyk, L., 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 andM. Palacin (2011). Molecular basis of substrate-induced permeation by an amino acid antiporter. *Proc Natl Acad Sci U S A,* 108(10): 3935-3940.

Krishnamurthy, H. andE. Gouaux (2012). X-ray structures of LeuT in substrate-free outward-open and apo inward-open states. *Nature.*

Krishnamurthy, H., C. L. Piscitelli andE. Gouaux (2009). Unlocking the molecular secrets of sodium-coupled transporters. *Nature,* 459(7245): 347-355.

Lin, J., M. P. Smith, K. C. Chapin, H. S. Baik, G. N. Bennett andJ. W. Foster (1996). Mechanisms of acid resistance in enterohemorrhagic Escherichia coli. *Appl Environ Microbiol,* 62(9): 3094-3100.

Ma, D., P. Lu andY. Shi (2013). Substrate selectivity of the acid-activated glutamate/gamma-aminobutyric acid (GABA) antiporter GadC from Escherichia coli. *J Biol Chem,* 288(21): 15148-15153.

Ma, D., P. Lu, C. Yan, C. Fan, P. Yin, J. Wang andY. Shi (2012). Structure and mechanism of a glutamate-GABA antiporter. *Nature,* 483(7391): 632-636.

Palsdottir, H. andC. Hunte (2004). Lipids in membrane protein structures. *Biochim Biophys Acta,* 1666(1-2): 2-18.

Perez, C., C. Koshy, S. Ressl, S. Nicklisch, R. Kramer andC. Ziegler (2011). Substrate specificity and ion coupling in the Na+/betaine symporter BetP. *EMBO J,* 30(7): 1221-1229.

Perez, C., C. Koshy, O. Yildiz andC. Ziegler (2012). Alternating-access mechanism in conformationally asymmetric trimers of the betaine transporter BetP. *Nature,* 490(7418): 126-130.

Ressl, S., A. C. Terwisscha van Scheltinga, C. Vonrhein, V. Ott andC. Ziegler (2009). Molecular basis of transport and regulation in the Na(+)/betaine symporter BetP. *Nature,* 458(7234): 47-52.

Reyes, N., C. Ginter andO. Boudker (2009). Transport mechanism of a bacterial homologue of glutamate transporters. *Nature,* 462(7275): 880-885.

Ryan, R. M., A. D. Mitrovic andR. J. Vandenberg (2004). The chloride permeation pathway of a glutamate transporter and its proximity to the glutamate translocation pathway. *J Biol Chem,* 279(20): 20742-20751.

Shaffer, P. L., A. Goehring, A. Shankaranarayanan andE. Gouaux (2009). Structure and mechanism of a Na+-independent amino acid transporter. *Science,* 325(5943): 1010-1014.

Shimamura, T., S. Weyand, O. Beckstein, N. G. Rutherford, J. M. Hadden, D. Sharples, M. S. Sansom, S. Iwata, P. J. Henderson andA. D. Cameron (2010). Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science,* 328(5977): 470-473.

Tsai, M. F., Y. Fang andC. Miller (2012). Sided functions of an arginine-agmatine antiporter oriented in liposomes. *Biochemistry,* 51(8): 1577-1585.

Weyand, S., 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 andA. D. Cameron (2008). Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science,* 322(5902): 709-713.

Yamashita, A., S. K. Singh, T. Kawate, Y. Jin andE. Gouaux (2005). Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters. *Nature,* 437(7056): 215-223.