A random approach to stabilize a membrane transport protein for crystallization studies

Un enfoque aleatorio para estabilizar un transportador de membrana para estudios de cristalización

Arturo Rodríguez Banqueri

FACULTAT DE FARMÀCIA

DEPARTAMENT DE BIOQUIMICA I BIOLOGIA MOLECULAR

“A random approach to stabilize a membrane transport protein for crystallization studies”

“Un enfoque aleatorio para estabilizar un transportador de membrana para estudios de cristalización”

ARTURO RODRÍGUEZ BANQUERI

BARCELONA, 2013

FACULTAT DE BIOLOGIA

DEPARTAMENT DE BIOQUÍMICA I BIOLOGIA MOLECULAR

PROGRAMA DE DOCTORAT DE BIOTECNOLOGIA

“A random approach to stabilize a membrane transport protein for crystallization studies”

“Un enfoque aleatorio para estabilizar un transportador de membrana para estudios de cristalización”

Memòria presentada per Arturo Rodríguez Banqueri per optar al títol de doctor per la Universitat de Barcelona

DIRECTOR: JOSÉ LUIS, VÁZQUEZ-IBAR

TUTOR: MANUEL, PALACÍN PRIETO

ARTURO, RODRÍGUEZ BANQUERI

ARTURO RODRÍGUEZ BANQUERI

BARCELONA, 2013

A mis Padres

A Susana

A Gabriel

Agradecimientos

En primer lugar, quisiera agradecer al Dr. José Luis Vázquez-Ibar, por aceptarme para realizar mis estudios de doctorado bajo su dirección, por facilitarme su tiempo y medios para la realización de nuestra investigación y principalmente por su apoyo incondicional que en todo momento ha depositado en mí. Sin duda, eres una gran persona y me siento afortunado de haber realizado mi tesis bajo tu tutela. Agradezco también al Dr. Manuel Palacin, por su constante buena disposición hacia nosotros y por aceptarme en su grupo de investigación y laboratorio, gracias por contagiarnos siempre con tu espíritu científico y por tus constantes aportaciones dentro y fuera del laboratorio. También quisiera manifestar mi agradecimiento al Dr. Antonio Zorzano, por sus consejos, contribuciones y preguntas, no solo de mi trabajo sino de todos los seminarios que realizamos en el instituto.

A Lukasz ahora en Australia, se te echa de menos, gracias por hacer mi estancia en el laboratorio más fácil desde el primer día. A Ekaitz, que ha contribuido al desarrollo de mi tesis, enseñándome y aportando una mirada crítica que es tan necesaria para ver las cosas desde otro punto de vista, cuando más se necesita. A Paola, le agradezco sus enseñanzas para poder realizar los experimentos de transporte. A Meritxell C. por ser la mejor compañera de ordenador posible y a Elena por ser siempre amable con todos, seguro que te saldrá todo lo que te propongas! A Albert Rosell por sus consejos y su paciencia cuando discutimos sobre la TM. A Gonzalo por compartir su espacio de –80°C conmigo, cuando aún no tenía. A Susana Bodoy y a Joana, les agradezco su esfuerzo cotidiano por tener funcionando correctamente nuestro laboratorio. A Txell y su ayuda con las mutaciones puntuales. A Ana Obando y sus consejos de cristalografía. Y otra vez a Ekaitz, Gonzalo y a Elena siempre me acordaré de las birras y nuestros videos para las otras tesis. A los técnicos Jordi, Susana Bial y Juan Carlos (aunque ya no esté), quienes facilitan nuestras tareas y nos sacáis siempre de apuros. Por supuesto también a Natalia Mólner y Olga que hacéis mucho por nuestro laboratorio!!

A JP, Edu, Maribel, Shaska, Jana y recientemente Paula… Sin vosotros sin duda para mí el laboratorio no sería el mismo… Nuestras disputas Barça-Madrid… las reservas de energía… teorías sobre ciencia... el Marsella… el tiburón… Os echaré de menos. A mis compañeros de pasillo y vecinos Alba, Xevi, Mari Angels, Montse, David Sala, Yuliana, Maria Joao, Ana S., Laia, David Sebastian, Nacho, Manu, Lucia, Natalia Plana, Jessica y Víctor, por esos momentos tan divertidos dentro del laboratorio… el timer!!!!... las canciones del Fary… las conversaciones sobre política… las desapariciones misteriosas…

Espero no haberme olvidado a nadie, Muchas gracias a todos!

A mis amigos de la Universidad Rubén, Esteban, Albert, Carlos, Marc C., David, Marc B., Jordi. Simplemente sois muy grandes!!

A toda mi nueva gran familia chilena que he conocido durante estos años, en especial a la abuela Susana. A mis padres que siempre han sido un apoyo para mí, ahora más todavía desde que soy padre, muchas gracias por todo.

A Susana por todo, no solo porque ha sido un apoyo cuando la he necesitado durante la presente tesis, sino por todos estos años que hemos compartido juntos y que incluyen los momentos más felices de mi vida, te amo. Finalmente quisiera agradecer a mí hijo Gabriel, nunca hubiera imaginado lo mucho que ibas a cambiar mi vida y de qué manera, gracias por enseñarme lo que realmente importa.

INDEX

ABREVIATIONS

1 ABBREVIATIONS

2D	bi-dimensional
3D	three-dimensional
4f2ed	Heavy subunit 4F2 ectodominium
4F2hc	Heavy chain of 4F2
Ab	Antibody
Amp	ampiciline
ANTET	Anhydritetracycline
APA	The Basic Amino/Polyamine Antiporter Family
APC	The Amino acid-Polyamine-organoCation Superfamily
Asc-1	system asc amino acid transporter-1
Asc-2	system asc amino acid transporter-2
$B^{0+}AT$	System B0,+ amino Acid Transporter
bp	base pares
BSA	Bovine Serum Albumin
°C	Celsius grades
CAT	Cationic Amino Acid Transporter family
cDNA	Complementary DNA (from a RNA sequence synthetized from a retro-transcription)
CMC	Critic Micellar Concentration
Cpm	Counts per minute
Da	Dalton
DNA	Deoxyribonucleic Acid
DDM	n-dodecyl-β-D-Maltopyranoside
DM	n-decyl-β-D-Maltopyranoside
EDTA	Etilen Din amino tetra acetic acid
et al.	and others / and collaborators

FPLC	Fast Performance Liquid Chromatography
FSEC	Fluorescence Size Exclusion Chromatography
g	gram
GFP	Green Fluorescent Protein
h	hour
HAT	Heteromeric Amino Acid Transporter
IMAC	immobilized metal ion affinity chromatography
IMP	integral membrane protein
IPTG	isopropyl-β-tiogalactopiranoside
kb	kilobases
kDa	kilodalton
Kan	kanamicine
Kpsi	Kilo pounds per square inch
LDAO	Lauryldimethylamine-oxide
l	liter
LAT	L-type Amino acid Transporter
LB	Luria Bertani Broth
LPI	Lysinuric Protein Intolerance
LeuT	Leucine Na^+/Cl^--dependent transporter of *A. aeolicus*
M	molar (mol/l)
ES-MS	Electro spray - mass spectrometry
mg	miligram
min	minute
ml	milliliter
mM	milli molar
mm	millimeter
mRNA	Ribonucleic acid messenger
MWCO	Molecular Weight Cut-off

NG	n-Nonyl-β-D-Glucopyranoside
OD	Optical Density
OG	n-Octyl-β-D-Glucopyranoside
PBS	Phosphate Buffered Saline
PCR	Polimerase Chain Reaction
PDB	Protein Data Bank
psi	pounds per square inch
rBAT	related to b0,+-amino acid transporter
R.F.U	Relative Fluorescence Units
rpm	revolutions per minute
s	second
SDS	Sodium Dodecyl Sulfate
SDS PAGE	Sodium Dodecyl Sulfate Poliacrilamide Electroforesis gel
SEC	Size Exclusion Chromatography
Ser	Serine
SLC	Solute Carrier Family
sp.	Specie
Spc	Spectinomycin
SSS	Sodium solute symporter family
SteT	Serine / Threonine exchanger
Thr	Threonine
TM	Tramsmembrane
TMD	Transmembrane Domain
WT	wild type
x g	gravity acceleration
µmol	micro mol
µg	microgram
µl	micro liter

µM	micro molar
mM	mili molar

INTRODUCTION

2 INTRODUCTION

2.1 Integral membrane proteins

Membrane proteins are by definition, proteins that regardless their function; interact with the cell membrane or with the membrane of an intracellular organelle. Based on the type of association with the membrane, there are two main broad categories of membrane proteins: (i) peripheral membrane proteins that are associated (but not buried) with the membrane or with other membrane proteins; and (ii) integral membrane proteins (IMPs) that are partially or completely buried within the membrane. According the relationship with the lipid bilayer, IMPs can be classified as (a) monotopic or (b) polytopic depending whether they span across the membrane one or several times, respectively. IMPs can perform a large variety of important functions, being key players in maintaining cell homeostasis by transferring information between the extracellular and intracellular spaces of the cell or between intracellular compartments (White, 2009). For instance, membrane receptors are responsible for transducing signals across the cytoplasmic membrane that will trigger different physiological responses (Lefkowitz, 2007). Other IMPs like membrane transporters allow the selective passage of nutrients, metabolites or even drugs across the different membranes. Finally, ion channels are responsible for generating and propagating the action potential of excitatory cells like neurons or muscle cells (Kim, 2006).

It has been estimated that between 20 and 30% of the proteomes of most organisms are IMPs (Wallin and Von Heijne, 1998). Indeed, 30% of the human genome encodes integral membrane proteins (Wallin and Von Heijne, 1998; Fagerberg et al., 2010). The amounts and types of proteins in a membrane are highly variable, giving different functional properties to each membrane. For

instance, in the myelin membrane, which serves as conductor of electrical insulation for nerve cell axons, less than 25% of the membrane is protein (Alberts B, Molecular Biology of the Cell. 4th edition, 2002). On the other hand, in membranes involved in ATP production such the inner mitochondrial membrane, around 75% of the membrane is protein (Alberts B, Molecular Biology of the Cell. 4th edition, 2002). It is not surprised, therefore, that the malfunction of IMPs is direct cause of important pathologies (Sanders and Myers, 2004). In fact, almost 50% of current drugs target IMPs (Overington et al., 2006), playing as well an important role in drug discovery and development (Overington et al., 2006).

2.2 Membrane transport proteins

Membrane transport proteins (also known as carriers, permeases, translocators, translocases, or porters) are IMPs responsible for the selective transport across membranes of a wide variety of substrates that are required for the normal physiological activity of the cell (Kim, 2006). These proteins catalyze or mediate the movement of ions and molecules by physically binding and moving them across the membrane. Based on the energy dependence, membrane transport proteins can be classified into two types (Kaback, 1986). **Passive transporters** (or facilitators) transport their substrates from one area of high concentration to another of low concentration, which results in the equilibration of the concentration gradient of the substrate and no energy consumption during the process. **Active transporters** manage to transport substrates against their concentration gradient; a process thermodynamically unfavorable that requires the coupling to a form of energy. In this regard, active transporters can be classified as **primary transporters** if they use directly ATP as energy source or **secondary transporters** if they use the energy stored in electrochemical ion gradients. Coupling the downhill movement these ions

with the substrate is the mechanism by which secondary transporters utilizes the energy from electrochemical gradients (Kaback, 1987).

Membrane transport proteins can also be classified according their transport mechanism. **Uniporters** transport a single molecule at a time, while **symporters** and **antiporters** transport simultaneously two different molecules in the same or in opposite directions, respectively.

The physiological relevance of membrane transport proteins implies that genetic defects that affect their expression yield and/or functionality are the direct cause of severe pathologies (http://www.tcdb.org/disease_explore.php). In this line, membrane transport proteins have been explicitly identified as the primary molecular target in the action of several important drugs used for the treatment of hypertension, heart failure, gastro-intestinal disorders, atherosclerosis and various psychiatric disorders, including anxiety and depression (Overington et al., 2006). The mechanism by which the majority of these drugs perform their pharmacological function is by inhibiting their target. Mechanistically, this inhibition is accomplished by the direct competition of the drug for the substrate binding site or by an allosteric mechanism that results on the inhibition of the protein conformational changes needed during the transport cycle (Zhou et al., 2007, 2009). Finally, tissue distribution, organ-specific entry, and clearance of drugs (drug-resistance mechanisms) are often facilitated or hindered by the expression of membrane transporters (Kim, 2006).

2.3 Amino Acid Transporters

Amino acids are essential molecules for cell survival. They have important roles in protein synthesis, cell growth, cell size regulation, production of metabolic

energy and they are precursors for a variety of metabolites (Christensen, 1990). Chemically, amino acids are molecules containing an amine group, a carboxylic acid group, and a specific side-chain and they can be covalently modified after protein synthesis. The modified amino acids can be involved in a large variety of biological functions like protein regulation and function, intracellular signaling, genetic regulation or, simply, sensing (Tsiboli et al., 1997). Membrane transport proteins mediate the cellular intake and uptake of amino acids, passing through the hydrophobic domain of the cytoplasmatic membrane. These proteins are crucial for the absorption of amino acids required for nutrition mediating, as well, their intercellular, interorgan or inter compartmental transfer (Christensen, 1990). Dysfunctions of amino acid transporters are associated to metabolic disorders, (Bröer and Palacín, 2011). Indeed, mutations of renal and intestinal amino acid transporters affect whole-body homeostasis as result of amino acid malabsorption, causing, as well, renal problems as a consequence of impaired renal clearance (Bröer and Palacín, 2011). Importantly, amino acid uptake seems to be essential for cell tumor growth, thereby explaining their role in tumor progression (Ohkame et al., 2001; Kobayashi et al., 2008).

Based on sequence similarity, mammalian amino acid transporters are grouped in 11 solute carrier (SLC) families (Tables 1 and 2) (Bröer and Palacín, 2011). The lysosomal cystine transporter (Cystinosin) has not received an SLC number since it belongs to a family of proteins that appear to be involved in protein glycosylation (Kalatzis et al., 2001). In addition, there is a nomenclature based on functional criteria, such as substrate preference and Na^+-dependence, which categorizes amino acid transporters into systems (Bröer and Palacín, 2011) (Tables 1 and 2).

Table 1. Amino acid transporters, their properties and diseases. 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. AAT, amino acid transporter. **Adapted from (Bröer and Palacín, 2011)**

SLC	Acronym	Substrate(s)	Function	Disease/phenotype
SLC1A1	EAAT3	D,E,Cn	System X^-_{AG}	Dicarboxylic aminoaciduria, OCD
SLC1A2		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}	
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^+	
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

2.4 Heteromeric Amino Acid Transporters: propierties, functionlality and associated diseases of SLC3 and SLC7 families

The members of one of the most relevant families of mammalian amino acid transporters, the heteromeric amino acid transporters (HATs), are composed by two subunits connected in the plasma membrane through a disulfide bridge: the heavy subunit (SLC3 family) and the light subunit (SLC7 family) (Figure 1 and Table 2) (Bröer and Palacín, 2011). Functionally, HATs are obligatory amino acid exchangers (antiporters) with a 1:1 stoichiometry (Busch et al., 1994) (Chillarón et al., 1996). With very few exceptions, they are not coupled with electrochemical ion gradients of Na^+ or H^+. The heavy subunit is essential for trafficking of the holotransporter to the membrane (Mastroberardino et al., 1998; Torrents et al., 1998), whereas the light subunit catalyzes the transport function (Reig et al., 2002b).

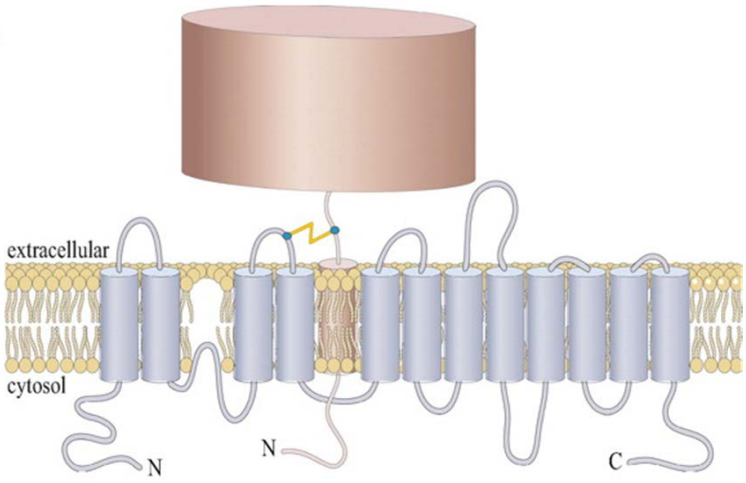


Figure 1. Structural organization of the Heteromeric Amino acid Transporter. Heavy subunit (brown) and light subunit (grey) are linked by disulfide bridge (yellow). The light subunit is represented in a 2D model based on the topological studies performed in xCT (Gasol et al., 2004). **Adapted and modified *from* (Fort et al., 2007)**

Two known heavy subunits have been identified: rBAT (related to $b^{0,+}$ amino acid transport; SLC3A1) and 4F2hc (4F2 cell-surface-antigen heavy chain, also named CD98; SLC3A2) (Table 2) (Bröer and Palacín, 2011). 4F2h is a multifunctional type II membrane glycoprotein involved, besides amino acid transport, in cell fusion, and β1 integrin-dependent adhesion (Fort et al., 2007). In humans, there are eight light subunits identified that compose the L-amino acid transporter (LAT) family; all of them belonging to the SLC7 family (Bröer and Palacín, 2011). Six members (LAT1, LAT2, y^+LAT1, y^+LAT2, asc1 and xCT) heterodimerize with 4F2hc, and two members ($b^{0,+}$AT and AGT1) with rBAT (Table 2) (Fernández et al., 2002) and (Nagamori S and Palacín M., unpublished results). The heavy subunit associated with asc2 is presently unknown (Table 2). The light subunits are not glycosilated and have the N- and C-terminal ends oriented towards the cytoplasm (Gasol et al., 2004). All light subunits have 12 TMDs and an apparent molecular weight on SDS-PAGE of 50 KDa (Gasol et al., 2004). As detailed in the next sections, congenital mutations of several HAT members are responsible of two main disorders: cystinuria and lysinuric protein intolerance (Palacín et al., 2001). Table 2 indicates the substrate specificity of each HAT member along with their associated pathologies. The closest family of LATs, the Cationic Amino Acid Transport (CAT) family (Table 1) (i.e. CAT1, CAT2, and CAT3), have 14 TMDs and present glycosylation (Wagner et al., 2001a). This family is involved in transport sytem y^+ transporting cationic amino acids with differential trans-stimulation by intracellular substrates (Verrey et al., 1999).

Table 2. Heteromeric Amino acid transporters, their properties and diseases. 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. **Adapted from (Bröer and Palacín, 2011)**

SLC	Acronym	Substrate(s)	Function	Disease/phenotype
SLC3A1	rBAT		Trafficking subunits Heavy chains of heteromeric AAT	Cystinuria
SLC3A2	4F2hc		Trafficking subunits Heavy chains of heteromeric AAT	Tumour growth
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	b0,+AT/rBAT	R,K,O,Cn	System b0,+	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/rBAT	D,E Asp,	Glu transporter	

2.4.1 rBAT/B0+AT and cystinuria

rBAT (SLC3A1) is the heavy chain of the renal cystine transport system $b^{0,+}$. It is expressed mainly in kidney and small intestine (Bertran et al., 1992; Tate et al., 1992; Wells and Hediger, 1992). $b^{0,+}$ AT (SLC7A9) was identified as the light subunit that co-expresses with rBAT and forms the $b^{0,+}$ amino acid transport system [(Feliubadaló et al., 1999), (Pfeiffer et al., 1999), (Chairoungdua et al., 1999)]. The rBAT/$b^{0,+}$AT heterodimer mediates the exchange of dibasic amino acids and cystine with neutral amino acids (except imino acids) (Bertran et al., 1992),(Chillarón et al., 1996). Thus, the exchange of dibasic by neutral amino acids is electrogenic (Busch et al., 1994).

Mutations in either of the genes that encodes the system $b^{0,+}$ (SLC3A1 and SLC7A9) cause **cystinuria** (OMIM 220100) (Palacin M, Goodyear P, Nunes V, Gasparini P. Cystinuria. Scriver C Baudet AL Sly WS Valle D eds The metabolic and molecular basis of inherited disease, Vol. III 8th ed. 2001:4909-4932 McGraw-Hill New York). This autosomal-recessive disorder is characterized by the hyperexcretion of dibasic amino acids and cystine in urine. In some cases, the low solubility of cystine leads to the formation of cystine calculi due the high concentration of this amino acid in the urinary tract (Chillarón et al., 2010).

The lack of genotype-phenotype correlation have let to a classification of cystinuria based on genetics (Dello Strologo et al., 2002). In this classification **Type A** cystinuria includes patients that have mutations in the SLC3A1 gene, **type B** includes patients with mutations in the SLC7A9 gene and, **type AB** comprises patients with one mutation in each gene.

.

2.4.2 4F2hc/LAT1 and tumor growth

The heterodimer 4f2hc/LAT1 (system L) is responsible for the Na^+-independent transport of branched and large neutral amino acids (Wagner et al., 2001). Indeed, LAT1 was the first subunit of sytem L found to interact with 4f2hc (Kanai et al., 1998). It has been shown that cell proliferation increases the expression of 4f2hc/LAT1, indicating the physiological importance of the uptake of neutral branched amino acids for cell growing and development. In this line, 4f2hc/LAT1 is highly expressed in nearly all tested tumor cell lines and in human tumors (Fuchs and Bode, 2005), suggesting the role of this transporter in angiogenesis, proliferation and tumor growth (Kaira et al., 2008). Recently, LAT1 has also been identified as a key transporter for mTOR (mammalian Target Of Rapamycin) regulation since it provides neutral branched amino acids to stimulate mTOR signaling (Nicklin et al., 2009).

2.4.3 4F2hc/y^+LAT1 and Lysinuric Protein intolerance

4F2hc/y^+LAT1 belongs to the system y^+L that mediates the electroneutral exchange of cytoplasmic cationic amino acids by external large neutral amino acids and Na^+ with a 1:1:1 stoichiometry (Torrents et al., 1998), (Pfeiffer et al., 1999) and (Chillaron et al., 2001). y^+LAT1 is mostly expressed in the basolateral membrane of the epithelial cells of the proximal tubule and small intestine (Bröer, 2008).

Mutations in the SLC7A7 gene (encoding y+LAT1) cause **lysinuric protein Intolerance** (LPI; OMIM 222700) (Palacín et al., 2005). LPI is an autosomal recessive disease mainly present in Finland (Palacín et al., 2005) and characterized by an urinary hyperexcretion of dibasic amino acids (arginine, lysine and ornithine) together with a poor intestinal absorption of these amino acids (Palacín et al., 2005). The low plasma levels of these amino acids are

thought to produce a functional deficiency of the urea cycle (Palacín et al., 2005). There are different clinical manifestations that include undernutrition, diarrhea, and vomits; however, the hyperammonemia resulting from the malfunction of the urea cycle can induce mental retardation or even coma (Palacín et al., 2004).

2.4.4 4F2hc/xCT and Kaposi's sarcoma-associated herpesvirus (KSHV)

4f2hc/xCT forms the x^-c transport system. This transport system exchanges the anionic form of cystine for glutamate with a 1:1 stoichiometry. The cysteine is rapidly reduced to cysteine and incorporated into glutathione and proteins. Indeed, transport of cystine and its intracellular reduction to cysteine are the rate-limiting steps in glutathione biosynthesis. xCT is expressed in most cell lines, in activated macrophages, and in the brain (Sato et al., 2002). Therefore, the key role of xCT in glutathione homeostasis suggests that this transporter contributes to the maintenance of the redox state, particularly in the central nervous system (Sato et al., 2005). xCT also serves as a fusion-entry receptor for the Kaposi's sarcoma-associated herpesvirus (KSHV), since overexpression of xCT increases severely the effectiveness of fusion of the KSHV in different cell lines (Kaleeba and Berger, 2006) (Qin et al., 2010).

2.5 Integral membrane proteins and structural studies

The three dimensional structure of IMPs not only provide valuable information about their structure-function relationships, but also it helps considerably during the discovery of new therapeutic agents against IMPs using the structure as scaffold for drug design (Kim, 2006). Unfortunately, due to the nature of these proteins, structural studies with IMPs continue to be a extremely challenging task.

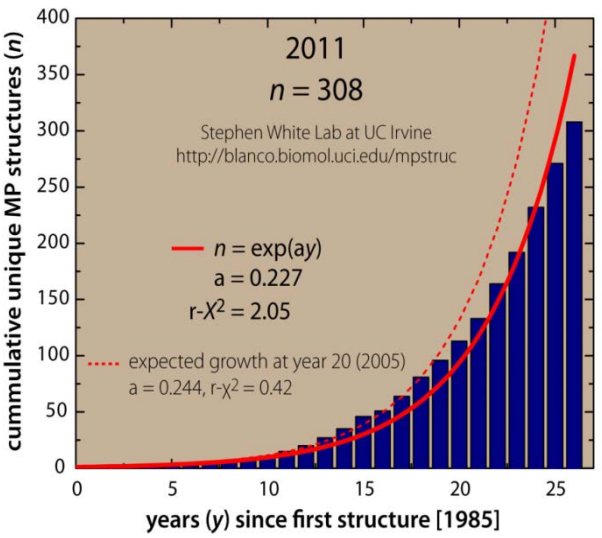


Figure 2. Representation of number of unique structures of membrane proteins solved since 1985 to 2011. (Red line) representing the tendency of the accumulation of membrane protein structures until 2011. (Red dotted line) represents the expected growth at year 20 (2005) since 1985. **Figure adapted from http://blanco.biomol.uci.edu/mpstruc**

Despite the fact that a reasonable number of macromolecules structures have been solved by electron crystallography or nuclear magnetic resonance (NMR), X-ray crystallography is the predominant technique for the determination of the structures of macromolecules at atomic resolution. Indeed, more than 70000 structures out of the 80000 deposited in the protein data bank (PDB) have been solved by X-ray crystallography **(http://www.rcsb.org/pdb/home/home.do).**

Unfortunately, high-resolution structural information of IMPs is still quite restricted relative to soluble protein. Specifically, among the 80000 protein structures deposited in the PDB, approximately 1000 are IMPs, representing near 400 different proteins **(http://blanco.biomol.uci.edu/mpstruc, http://www.rcsb.org/pdb/home/home.do)** (Figure 2). The initial bottleneck arises from the difficulty in obtaining the milligram amounts of recombinant

functional membrane protein necessary for crystallography studies using a heterologous expression system. The election of the best expression host is based on both yield and quality of the produced IMP and the final cost. IMPs have been successfully expressed in the bacteria *Escherichia Coli*, the most widely used host for protein overexpression (Grisshammer and Tate, 1995). Working with *E. coli* is quick, relative inexpensive and easy to use; facilitating, therefore, the multiple screening of different protein sequences and constructs. Alternatively, the bacteria *Lactoccoccus lactis* is also a good host for the heterologous expression of IMPs (Kunji et al., 2005). However, when working with eukaryotic IMPs the choices includes the yeasts *Pichia Pastoris* and *Saccharomyces Cerevisae*, different insect cell lines and, as final option, mammalian cell lines (Junge et al., 2008).

In general, overexpression of recombinant IMPs gives lower yields than those obtained for soluble proteins. In addition, the heterologous expression of IMPs frequently results in protein aggregation into inclusion bodies as a consequence of incorrect folding. There are different reasons that accounts for this. First, the codon usage by the ribosome is different between organisms, thus affecting the amount of expressed protein (Gvritishvili et al., 2010). Also, the mechanism by witch IMPs fold and insert into the membrane varies between species and, in addition, requires the presence of host-specific chaperons (Hendrick and Hartl, 1995). Indeed, the saturation of the Sec machinery in *E. coli* used for IMPs during biogenesis and insertion into the cytoplasmic membrane is the cause of incorrect folding and subsequent aggregation (Klepsch et al., 2011). Moreover, IMPs are embedded in a lipid media and since each organism has different membrane lipid composition, heterologous expression commonly results in an increase of the instability of the protein. This is particularly true for some GPCRs, since cholesterol plays essential roles on the organization, dynamics,

and function of this important class of IMPs (Zheng et al., 2012). Finally, eukaryotic IMPs often need post-translational modifications to maintain the correct folding and/or functionality of the protein that, logically, bacterial hosts cannot provide.

The majority of structural biology techniques applied to IMPs, including X-ray crystallography, need to extract and isolate these proteins from the membrane using detergents. This is a fairly delicate process since the detergent-solubilized protein tends to aggregate quite often due to its hydrophobic nature; complicating their manipulation for crystallization studies or even for any other structural or functional assay. Clearly, the choice of the detergent is a crucial part during the purification and crystallization of IMPs. The detergent n-dodecyl-β-D-maltopyranoside (DDM) is, perhaps, the detergent of choice to solubilize IMPs from their natural membrane environment since it is relatively cheap and, in the majority of the cases, preserves the integrity of the protein (Privé, 2007). However, DDM is not a good detergent for crystallization trials and, therefore, the protein needs to be subsequently exchanged to a variety of different detergents more suitable for crystallization (Iwata, 2003)

In addition, a common feature of IMPs is that they are notoriously resistant to crystallize due to the difficulty of forming well-ordered crystal lattices. The presence of the detergent in the IMP-detergent micelle complex reduces the probability to form the needed crystal contacts. Moreover, the dynamic nature of IMPs in the detergent-protein micelle complex (conformational dynamics) hampers considerably crystal formation and growing. In this line, when working with membrane proteins involves obtaining purified sample for either functional or crystallization studies, a time-consuming screening process is mandatory. The goal of this screening process is to identify optimal membrane

protein candidates (or versions of the selected IMP) with a reasonable expression yield in the chosen host and with acceptable stability after detergent solubilization.

2.6 Improving expression and stability of membrane proteins in solution for structural studies

There is no a single recipe to follow for optimizing an IMP before attempting structural studies. Multiples strategies have been proposed; however, it is almost impossible to predict whether a particular method will succeed or not. Nowadays, there is a high-throughput revolution in the structural biology field and new methods are emerging for effective expression, solubilization, purification and crystallization of membrane proteins ((Kawate and Gouaux, 2006), (Rasmussen et al., 2007) and (Simon Newstead, 2007). It is expected that these technical advances will lead to a rapid increase in the rate at which membrane protein structures are solved in the near future.

One of the first strategies largely and successfully used to improve IMPs structural knowledge is the extensive screening of a large number of homologues sequences of a chosen membrane protein target looking for optimal conditions of expression and detergent solubilization (Lewinson et al., 2008a). In other words, a so-called fishing expedition strategy aiming to find an optimal candidate that will be used as structural paradigm. In this regard, bacterial homologs have proven to be excellent structural and functional paradigms of mammalian membrane proteins (Singh et al., 2007), particularly those from thermophilic organisms. Alternatively, variants of a selected target (e.g., C-terminal and/or N-terminal modifications or single-point mutations) are routinely cloned in multiple expression vectors and tested for expression in order to identify a combination that will ultimately leads to and increasing

expression yield and/or stability. In either way, it is clear that a robust protocol to test the protein expression and stability of multiple samples in a fast and reliable manner is extremely beneficial in structural biology studies (Koth and Payandeh, 2009)

Using Antibody fragments is also one of the most popular strategies to increase the stability of IMPs. The 3D structure of the β_2-adrenoceptor (β_2AR), a GPCR, was determined by binding an F_{ab} antibody fragment to a cytoplasmic loop of the receptor (Rasmussen et al., 2007). This methodology was originally developed to crystallize mitochondrial membrane proteins (Hunte and Michel, 2002) and not only improves the stability of the IMP target but also helps on the formation of crystal contacts due to the presence of the F_{ab} fragments. Also, another crystal structure of β_2AR was solved by fusing the T4-lysozime to the receptor (Cherezov et al., 2007). Fusion of soluble proteins to extracellular or cytoplasmic domains of IMPS (either in loops or in the N- and C- terminal ends) was originally exploited for determining the topology of IMPs (Ehrmann et al., 1990). Similar to the F_{ab} fragments, the presence of the T4 lysozyme increases β_2AR stability and the formation of crystal contacts. It is worth mentioning that both strategies have raised some criticisms in the scientific community, since the presence of these proteins can, in some cases, induce the protein in a non-physiological conformation.

Mutagenesis is perhaps one of the most costless and effective strategy when trying to improve IMPs expression and stability. For instance, mutagenesis has improved the stability of some membrane proteins by deleting flexible parts of the protein without compromising severely the activity (Lemieux et al., 2003). Also, by deleting the C-terminal region of KtrA (the regulatory part of the bacterial potassium transporter KtrAB) a functional version of this protein was

constructed and subsequently crystallized revealing for the first time its octameric arrangement (Albright R.A, 2009). Moreover, and as commented before, engineering protein chimeras or fusion proteins that increases stability in solution have been largely used (Nishida M. et al, 2007; Cherezov V. et al, 2007).

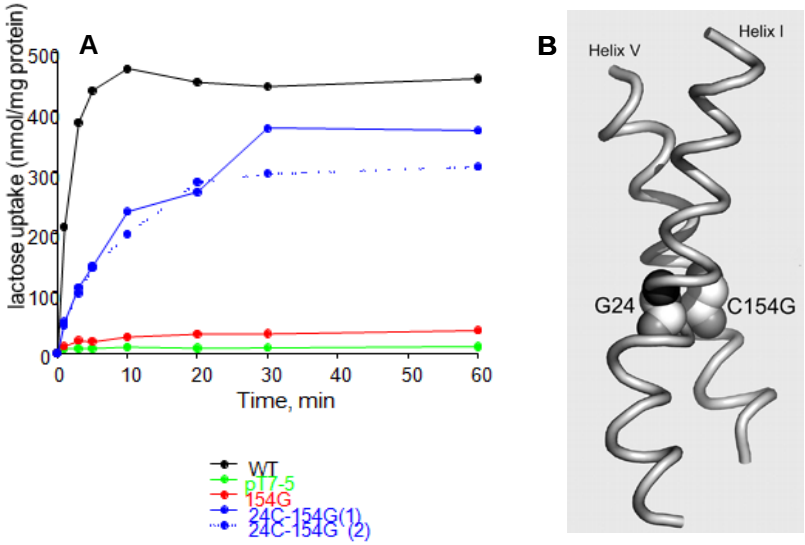


Figure 3. Effect of mutations of Gly 24 and Cys 154 in LacY transport activity in *E. coli* cells. (A) Time course of lactose accumulation of wild-type (black line), C154G-LacY (red line) and G24C/C154G-LacY (blue lines). Green line represents the control experiment (pT7-5 empty vector) with no expressed protein. (B) Spatial packing of TMD I and V in C154G-LacY (PDB 1PV7). TMD I and V are shown as spiral tubes. Gly residues at positions 24 and 154 are represented as spheres. **Adapted from (Ermolova et al., 2005)**

Many studies have pointed out that single side-chain substitutions in suitable positions of IMPs can significantly improve protein solubility and robustness being in some cases the key of getting well order diffracting crystals. Specifically, single point mutations in TMDs have demonstrated to increase the stability in solution of some membrane proteins after detergent solubilization and extraction from the membrane (Smirnova and Kaback, 2003); (Tate and Schertler, 2009). One of the most remarkable examples is the lactose permease

of *E. coli* (LacY), a β-galactosidase/proton symporter, member of the major facilitator superfamily (MFS) of membrane transporters. The key of obtaining the first X-ray structure of LacY was the use of a mutant (C154G-LacY) (Abramson et al., 2003). As seen in the X-ray structure and also proved by functional data, this mutation strongly favors one specific conformation of the protein by forming a Gly-Gly bridge between helices I and V (Ermolova et al., 2005) (Figure 3). As a result, the protein still binds the substrate but it is unable to translocate it due to the lack of conformational mobility. Furthermore, C154G-LacY is more thermostable than wild type with respect to ligand binding and aggregation (Smirnova and Kaback, 2003), a property that increases the probability of crystallization (Serrano-Vega et al., 2008).

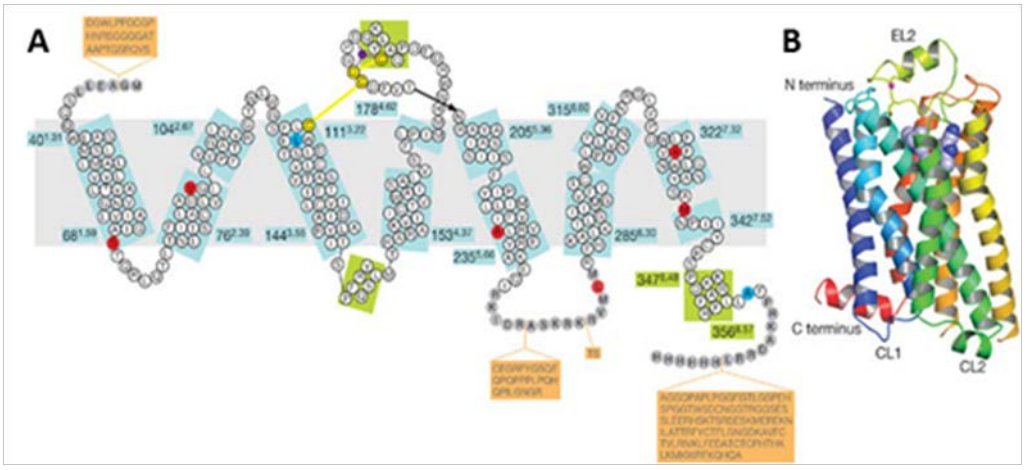


Figure 4. Schematic representations of the turkey β1AR structure. (A) Diagram of the turkey β1AR sequence in relation to secondary structure elements. Amino sequence in white circles indicates regions that are well ordered, but sequences in grey circles were not resolved in the structure. Representation of Beta-1-adrenergic receptor. Sequences on an orange background were deleted to make the β1AR construct for expression. Thermostabilising mutations are in red and two other mutations C116L (increases functional expression) (B) Ribbon representation of the β1AR structure in rainbow colouration (N-terminus blue, C-terminus red), with the Na+ ion in pink, the two disulphide bonds in yellow and cyanopindolol as a space-filling model. Extracellular loop 2 (EL2) and cytoplasmic loops 1 and 2 (CL1, CL2) are labelled. **Adapted *from* (Serrano-Vega et al., 2008)**

Interestingly, the improvement of the stability due to the C154G mutation was also extremely useful when performing luminescent experiments aiming to unravel the mechanistic role of a tryptophan residue sited the binding site of LacY during substrate binding and translocation (Vázquez-Ibar et al., 2003);(Vázquez-Ibar et al., 2004). In addition to a general effect of protein stabilization required for crystallography studies, a single mutation can, at the same time, stabilize a specific conformer of the protein during its catalytic cycle (Kowalczyk et al., 2011).

Finding single-point mutations that stabilize a selected IMP target is, quite often, a tedious work that requires a long-term systematic approach. For example, alanine scanning mutagenesis was used to build a thermostable mutant of βAR34–424, a truncated functional version of the β1-adrenergic receptor (Warne et al., 2003). A total of 318 mutations were made between positions 37 to 369, a region that encompasses all seven TMDs and the 23 amino acid residues at the C terminus. All the mutants were functionally expressed and assayed for thermostability by challenging their binding activity after heating the protein at increasing temperatures. The combinations of six of these mutations resulted in a mutant named βAR-m23 with a remarkable increase in thermostability compared to the native protein (Figure 4A) (Serrano-Vega et al., 2008). This multiple mutant was crystallized and, subsequently, the first crystal structure of β1-adrenergic receptor was solved (Warne et al., 2008) (Figure 4B). In addition, a further leucine scanning mutagenesis of βAR-m23 increased even more the thermostability of the receptor without affecting the function (Miller and Tate, 2011).

Taking into account the few high-resolution structures of IMPs, it is quite difficult to deduce and to rationalize (without a long-time consuming approach)

what amino acid combination will provide more stability of a target protein in detergent solution, while preserving its functional properties. Moreover, different studies from thermophilic organism have suggested that there is no a universal combination of factors that may be responsible for thermostability of IMPs or, even, soluble proteins (Razvi and Scholtz, 2006).

Random mutagenesis is a powerful tool that in combination with optimal screening methods allows the generation and selection of mutated versions of a chosen protein with new or improved properties (Labrou, 2010). Although extensively used in soluble proteins, very few examples can be found in membrane proteins. In the *E. coli* peptide transporter, YdgR, a member of the peptide transporter (PTR) family, 35 single point mutations resulted in a full or partial loss of transport activity of the transporter (Malle et al., 2011), a very useful strategy to study the structure-function relationships of IMPs. Recently, using error prone PCR random mutagenesis combined with fluorescence-activated cell sorting (FACs) and functional assays, Dodevski and Plückhun were able to find an evolved version of a GPCR with a 10-fold increase in functional expression and an improved stability in detergent solution (Dodevski and Plückhun et al, 2011).

2.7 Precrystallization studies and Green Fluorescent Protein

To apply a random mutagenesis strategy aiming to identify more stable mutants of a particular IMP is essential to combine it with fast and sensitive screening protocols. In this respect, the green fluorescent protein (GFP) fused to the intracellular C terminus of IMPs targets is a sensitive reporter that has enormously facilitated the precrystallization screening (Drew et al., 2006), Kawate and Gouaux, 2006), (Simon Newstead, 2007) and (Hammon et al., 2008). In this way, the GFP fluorescence is directly related to the quantity of

IMP. Therefore, the GFP fluorescence measured from either whole cell cultures or SDS-PAGE provides a direct measure of the amount of IMP (Drew et al., 2006); being, as well, extremely useful to follow the protein along the purification procedure. Most importantly, when GFP fluorescence is combined with size exclusion chromatography (fluorescence size exclusion chromatography, FSEC), it results in a powerful tool to characterize protein stability and quality under various conditions (e.g., homogeneity in different detergent solutions) using a very small amount of sample (Kawate and Gouaux, 2006).

GFP is a robust β-barrel protein containing 238 amino acids (Ormö et al., 1996) that in some cases can interfere with the expression and/or stability of the IMP target (Fucile et al., 2002). In this context, recent data reveals that the stability profile of a GFP-fused IMPs changes dramatically after removing the GFP (Hsieh et al., 2010). Indeed, similar observations were made previously with soluble proteins. Waldo and coworkers addressed this problem by successfully developing a GFP complementation assay to screen the solubility of globular proteins expressed in *E. coli* (Cabantous et al., 2005) and (Cabantous and Waldo, 2006). In this method, a 15-amino acid fragment of an engineered superfolded GFP (Pédelacq et al., 2005) (GFP_{11}) is expressed fused to the C terminus of the protein of interest (Cabantous et al., 2005). If the protein is stable and does not aggregate, the GFP_{11} fragment will complement with the remaining nonfluorescent 215-amino acid fragment of the GFP ($GFP_{1–10}$) independently expressed in the same cell. As a result, this complex emits GFP fluorescence and the method minimizes the effect of the GFP tag on the intrinsic properties of the protein under study.

.

2.8 Heteromeric Amino acid transports structural studies

The structural complexity of HATs is still far from being understood since only the human 4f2hc ectodomain has been crystallized (Figure 5) (Fort et al 2007). This structure was solved using two different crystals forms: monoclinic (PDB 2DH2) at 2.1 Å resolution and orthorhombic (PDB 2DH3) at 2.8 Å resolution (Fort et al., 2007) (Figure 5A). 4F2hc ectodomain structure is composed of a $(\beta\alpha)_8$ barrel and an antiparallel β_8 sandwich, a very similar folding than bacterial α-glycosidases (Figure 5A), although lacking key catalytic residues and consequently, catalytic activity (Fort et al., 2007).

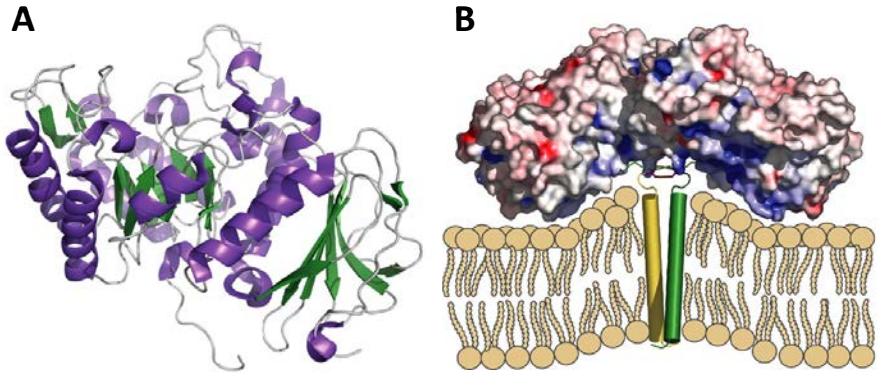


Figure 5. Structure of human 4f2hc ecotdomain. (A) Ribbon representation of the human 4f2hc ecotdomain. (B) Model of the of 4f2hc homodimer interacting with the citoplamatic membrane. **Figure adapted *from* (Fort et al., 2007)**

As observed in the crystal structure, and confirmed by in vivo cross-linking experiments, 4f2hc tend to form homodimers through a Cys 109 disulfide bridge (Fort et al 2007). Cys 109 is the same cysteine that uses 4f2hc to form disulfide bridges with the HATs light subunit. The exact physiological role of 4f2hc homodimerization is still under debate.

2.8.1 Structural studies of Ligth subunits of HATs (LATs)

Light subunits of HATs belong to the L-type amino acid transporter (LAT) family. A family included within the large amino acid, polyamines and organocation (APC) superfamily (Jack et al., 2000). LAT members have a 12 TMD topology (Figure 1) with the N- and C-terminal ends located inside the cell and with a TMD2-TMD3 intracellular loop accessible from the external medium (Gasol et al., 2004; Jimenez-Vidal et al., 2004) (Figure 1). The intervening cysteine residue that interacts with the heavy subunit in order to reach the plasma membrane is located in the extracellular loop between TMD3 and TMD4 and near to the TMD of the heavy subunit (Palacin et al., 2005) and (Fort et al., 2007).

To date, no 3D structures of a HAT light subunit have been determined. The closest structural models are several crystal structures of three prokaryotic members of the APC superfamily: AdiC, ApcT and GadC. All of them share less than 20 % amino acid identity with LAT members. AdiC, the arginine/agmatine exchanger from *E. Coli* has been crystallized in three conformational states: the inward-facing apo (PDB 3LRB and 3NCY) ((Gao et al., 2009)and (Fang et al., 2009)), the inward-facing substrate-bound (PDB 3OB6) (Kowalczyk et al., 2011) and the occluded substrate-bound conformation (PDB 3L1L) (Gao et al., 2010). The broad-specific amino acid transporter from *Methanocaldococcus jannaschii*, ApcT, was crystallized in the apo-occluded conformation (PDB 3GIA) (Shaffer et al., 2009). Finally, the crystal structure of GadC, the glutamate/GABA antiporter of *E. coli* reveals an apo-occluded conformation (PDB 4DJK) (Ma et al., 2012). All these transporters contain, like LATs, 12TMD; although they belong to a different family of APC transporters: the APA family (Wong et al., 2012). Notably, all these structures share the same structural fold: the so-called 5+5 inverted repeat fold; first described in LeuT, the Na^+-coupled leucine transporter of *Aquifex aeolicus* (Figure 6) (Yamashita et al., 2005), and a

prokaryotic model of mammalian serotonin transporters. This fold is characterize by a general structure of a core of 10 TMDs (numbered according to LeuT), where TMDs 1 to 5 are in a similar arrangement than helices 6 to 10 but in opposite orientation; that is, TMDs 1 to 5 are related with 6 to 10 by a two-fold symmetry (Figure 6) (Yamashita et al., 2005). It has been postulated (Forrest et al., 2008); and (Kowalczyk et al., 2011) based on molecular modeling and X-ray crystallography data, that this symmetrical arrangement is present in all conformational states of the transporters, dictating, therefore, the symmetrical relationship between conformers (e.g, the inward-facing state is symmetrical to the outward-facing state).

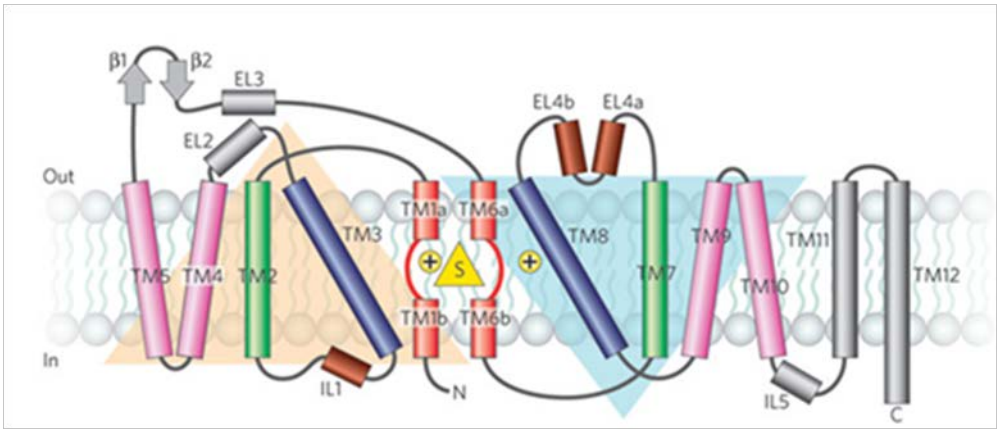


Figure 6. Architecture of the LeuT fold (5+5 inverted structural symmetry motif). Scheme of the topology of TM1-TM5 and TM6-TM10 is represented. **Figure Adapted from (Yamashita et al., 2005).**

Remarkable, different crystal structures have shown that the same fold is also shared by, at least, four other unrelated families of transporters with very little amino acid sequence identity. These families include the neurotransmitters sodium symporters (NSS) family where LeuT is included (Yamashita et al, 2005), the sodium/solute symporter (SSS) family (Faham et al., 2008), the betaine/choline/carnitine transporter (BCCT) family ((Ressl et al., 2009)), and

the nucleobase/cation symport-1 family (Weyand et al., 2008). Therefore, the increasing numbers of 3D structures of secondary transporters have revealed that distant families defined on the basis of sequence identity can be grouped into one single structural family.

2.8.2 Adic, the structural prokaryotic paradigm of LATs

AdiC, is an antiporter of *E. coli* that exchanges extracellular L-arginine (Arg^{+}) for intracellular agmatine (Agm^{2+}). This is a mechanism by which *E. coli* and other enteric bacteria achieve resistance to extreme acid environments (Ram Iyer, 2003). Amino acid decarboxylation in the cytoplasm is a proton-consuming reaction, thus preventing intracellular acidification in acid-rich environments. This transporter has a ~18% amino acid identity to eukaryotic LATs (Cassagrande et al, 2008). At present, the crystal structures of AdiC deposited in the PDB and the 2D structure represent the best structural models of LATs (Cassagrande et al., 2008), (Gao et al., 2009), (Fang et al., 2009), (Gao et al., 2010) and (Kowalczyk et al., 2011). These structures represent three conformations of AdiC transport cycle: the inward-facing apo (PDB 3LRB and 3NCY) (Gao et al., 2009) and (Fang et al., 2009), the inward-facing substrate-bound (PDB 3OB6) (Kowalczyk et al., 2011) and the occluded substrate-bound conformation (PDB 3L1L) (Gao et al., 2010) (Figure 9). Notably, mutagenesis has been crucial to obtain some of these structures.

A single point mutation in residue Asn 22 that improved the binding affinity for arginine by approximately six-fold compared to wild type while keeping intact the transport activity (Gao et al., 2010); permitted co-crystallize AdiC with arginine in an outward-facing occluded conformation (Gao et al., 2010) (Figure 9)

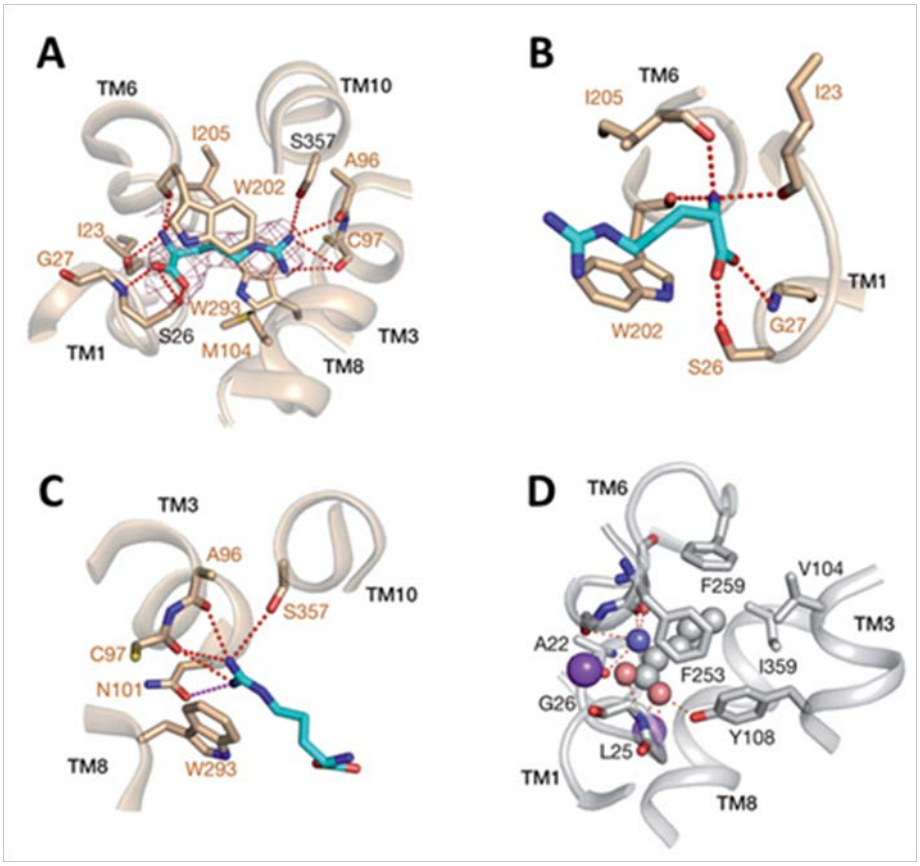


Figure 7. Substrate binding site of Adic and LeuT : A, Arg is bound at the centre of the transport path, recognized by amino acids from TM1, TM3, TM6, TM8 and TM10. The 2Fo - Fc electron density for Arg, coloured purple, is contoured at 1.0σ. B, Recognition of the α-carboxylate and α-amino groups of Arg. The α-amino group donates three hydrogen bonds, whereas the carboxylate group accepts two hydrogen bonds. C, Binding of the guanidinium group of Arg. The guanidinium group stacks against Trp 293, probably through cation–π interactions. The nitrogen atoms of the guanidinium group are located with approximate hydrogen-bond distance to four oxygen atoms in AdiC: the side chains of Ser 357 and Asn 101, and the carbonyl groups of Ala 96 and Cys 97. D, A close-up view of Leucine recognition by LeuT. Leu recognition by LeuT is similar to Arg recognition by AdiC. The structure of LeuT is displayed in a similar orientation as that of AdiC in a. Leu and the two Na+ ions (purple) are displayed as spheres. **Adapted from (Kowalczyk et al., 2011)**

The structure of N22A-AdiC revealed that the substrate arginine is located in the transport path, approximately 15 Å bellow the membrane surface from the periplasmic side and surrounded by five TMDs: 1, 3, 6, 8 and 10 (Figure 7). At one of end of the extended arginine molecule, the positively charged α-amino

group donates three hydrogen bonds to the carbonyl oxygen atoms of Ile 23 (TMD1), Trp 202 (TMD6) and Ile 205 (TMD6). The α-carboxylate group accepts two hydrogen bonds from the side chain of Ser 26 and the amide nitrogen of Gly 27; both residues located in the helix-breaking motif of TMD1 (Figure 7). The guanidinium group of the arginine stacks against Trp 293 (TMD8), probably through cation–π interactions (Figure 7). In addition, the nitrogen atoms of the guanidinium group are located at hydrogen-bond distances of Ser 357 (TMD10) , Asn 101 (TMD3), and the carbonyl groups of Ala 96 and Cys 97, both in TMD3 (Figure 7). The aliphatic portion of arginine interacts with the side chains of three hydrophobic amino acids, Met 104 in TMD3 and Trp 202 and Ile 205; both in TMD6 (Figure 7) (Gao et al., 2010). The binding site of AdiC is very similar to LeuT (Figure 7D). LeuT and Adic only share 10% amino acid identity; however, the orientation of the substrate in both transporters relative to the surrounding TMDs is similar (Figure 7). This similarity includes the coordination of the L-amino-carboxilate groups of the substrate with residues in the unwounded regions of TMD1 and TMD6 and the interaction of the lateral chain with residues in TMDs 3, 6 and 8 (Figure 7).

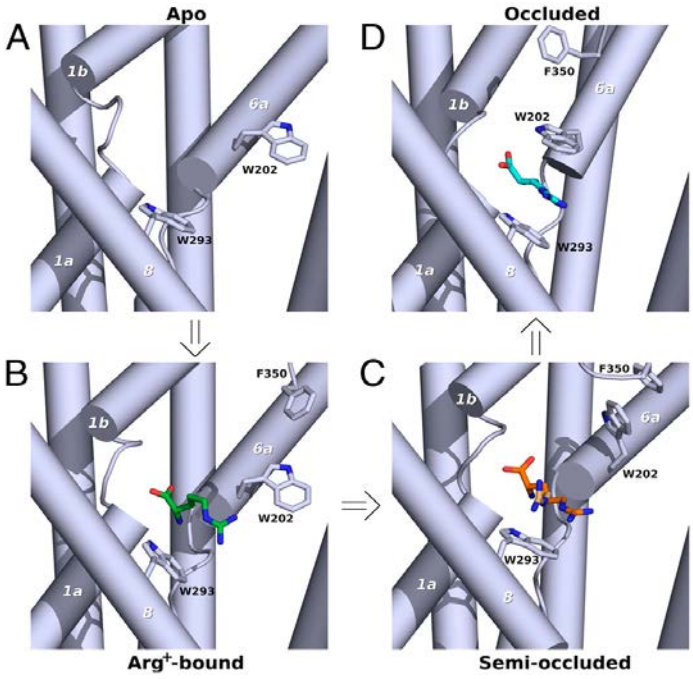


Figure 8. Proposed mechanism of Arg+ recognition and induced fit by AdiC. Periplasmic Arg+ is recognized by the apo conformation of AdiC (A; 3NCY) and binds with a similar orientation (B; current structure) as in the Arg+- occluded conformation (D; 3L1L). The proper Arg+ binding samples the semioccluded state (C; docked Arg+ in 3LRB) by stabilizing Trp202 (TM6a) and Phe350 (loop TMs 9–10) interaction. This semioccluded conformation evolves to the occluded state mainly by pivoting TM6a. Transition from the apo (A) to the semioccluded state (C) is defective in mutant N101A. TM segments are numbered in italics. **Figure adapted from (Kowalczyk et al., 2011)**

In 2010, our laboratory succeeded on crystallizing and solving the structure of the AdiC mutant N101A at 3.0 Å resolution (Kowalczyk et al., 2011) and PDB code (3OB6). By simply replacing Asn 101 by alanine, the protein lost its capacity to make a hydrogen bond with the substrate (Figure 7). N101A-AdiC exhibited a drastic reduction of the substrate translocation rate, although binding affinity was unaltered (Kowalczyk et al., 2011). As a result, the protein was stabilized in an outward-facing substrate-bound conformation that permitted to obtain well-diffracting crystals and to solve the structure. This new structural conformation of AdiC represented an intermediate state between

the previous outward-facing substrate free and outward-facing bound occluded conformations (Gao et al., 2009), (Fang et al., 2009) and (Gao et al., 2010).

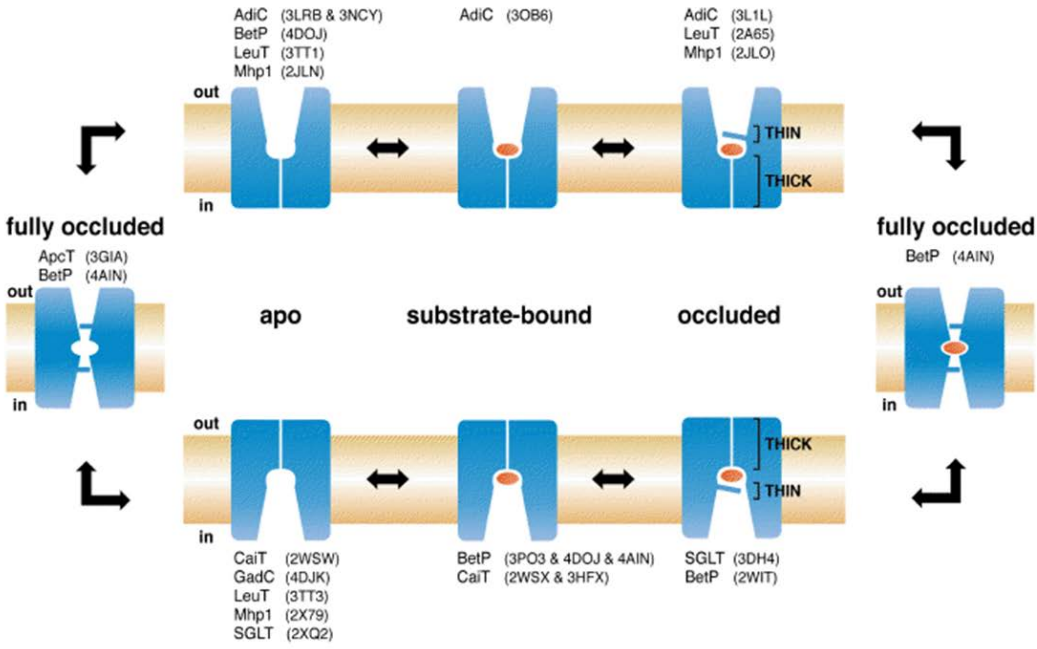


Figure 9. Symmetrical states along the alternative access mechanism of transporters with the 5 + 5 inverted repeat fold. Upon substrate (red ellipsoid) binding to the open-to-out apo state, the substrate-bound state (represented by N101A-AdiC structure) evolves to an occluded state, where two gates (thick and thin) prevent the diffusion of the substrate to either side of the membrane. Occlusion of the substrate by a thin gate is a common mechanism in the transport cycle of these transporters. The inward-facing states are symmetrically related to the outward-facing ones. Transition to the inward-facing states requires a transient fully occluded symmetrical intermediate. In ion-coupled symporters (LeuT, vSGLT, Mhp1, ApcT, and BetP) a free transition between the apo structures (outward- and inward-facing) is required to close the transport cycle. The apo-occluded structure of ApcT is close to this state. In antiporters (AdiC, GadC and CaiT), the return to the outward-facing states requires the binding and translocation of a new intracellular substrate that will move the transporter back through all the states but in the opposite direction. PDB access codes are indicated in parentheses. **Figure adapted (Kowalzyc et al, 2011)**

In addition, the N101A-AdiC structure informed about the role of the arginine guanidinium group in triggering the transition from the outward-facing to the outward-facing occluded state (Figure 8). N101A crystal structure together with functional and computational modeling revealed that the proper coordination or “productive pose” of this group stepping on Trp 293 (TMD 8) and interacting with Asn 101 (TMD 3) and Ser 357 (TMD 10), is of obligatory transit towards the occluded state (Figure 8B). Indeed, this work represented one of the few structural examples of substrate-induce fit of secondary transporters. Figure 9 schematizes this mechanism divided in four different steps and based on the available crystal structures of AdiC and other 5+5 inverted-repeat fold transporters together with molecular docking models using these structures (Kowalczyk et al., 2011).

As commented earlier, AdiC is the closest structural model of LAT transporters. However, in the absence of a high-resolution 3D structure of a mammalian LAT, it remains to be established whether LAT transporters, which are phylogenetically distant homologues of AdiC (an average of 18% amino acid sequence identity) (Figure 10), share the same substrate binding design and transport mechanism. New structures at atomic resolution of LAT prokaryotic homologs with closer amino acid identity must be released in order to unravel the architecture of the substrate binding site and the transport mechanism and, therefore, to better understand the molecular basis of the pathologies associated with congenital mutations of LATs.

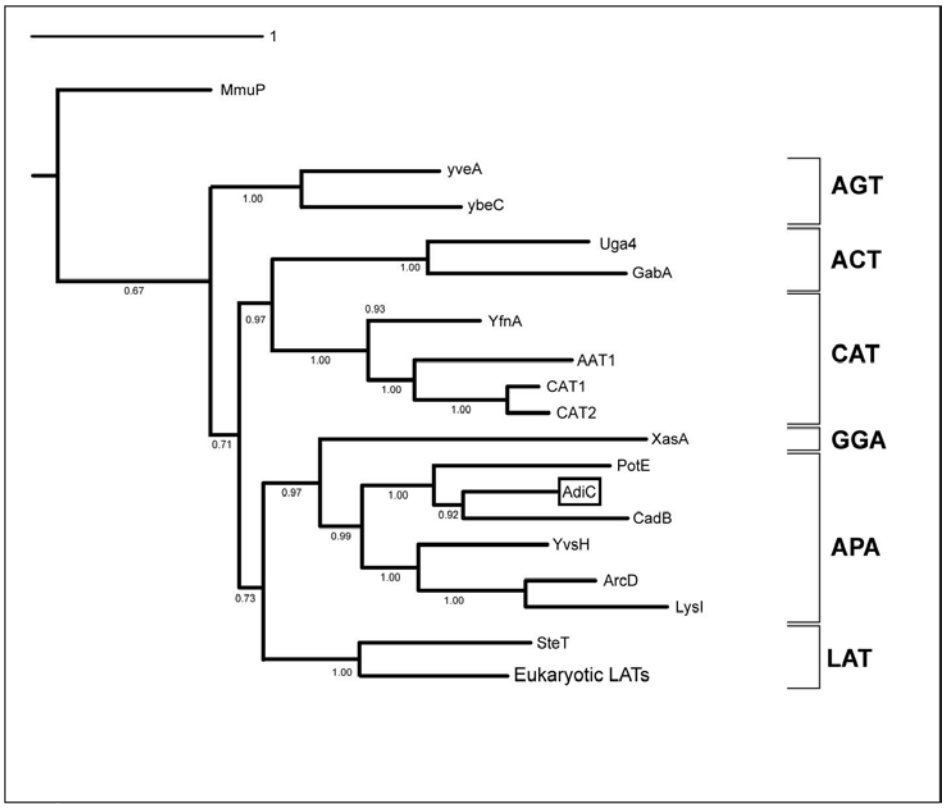


Figure 10. Phylogenetic relationship between AdiC and SteT and other members of the APC superfamily. The neighbor-joining tree illustrates the phylogenetic relationships of SteT and Adic with the closest alignable members of the APC superfamily that are either functionally characterized or putative amino acid transporters. Note that SteT is the closest prokaryotic homolog of LATs. **Figure adapted from (Casagrande et al., 2008)**

2.9 Serine/Threonine transporter (SteT), the first prokaryotic member of the L-Amino Acid Transporter (LAT) family

The L-serine/L-threonine antiporter of *Bacillus subtilis*, previously known as Ykba, was first cloned and functionally characterized in our laboratory, being the first prokaryotic member of the LAT family (Reig et al., 2007). SteT contains 438 amino acids and 12 predicted TMDs with a molecular mass of 48,879 KDa. Sequence alignments of SteT with LAT members show identities ranging from 26 to 30% amino acid identity (Reig et al., 2007) (Table 2). In addition, a

phylogenetic tree comparing SteT with the APC superfamily indicated that SteT clearly clusters with the members of the LAT family (Reig et al., 2007)

Table 3. The membrane topology of SteT is similar to eukaryotic LAT family and fits with the topology model of xCT, wich contains 12 transmebrane domains (Gasol et al., 2004).The most highly conserved regions between SteT and the members of the eukaryotic LAT family correspond to TM1, TM2 and the re-entrant loop between TM2 and TM3. The main differences are a shorter N and C termini in % amino acid identity between SteT and the LSHATs. In bold ≥25% amino acid identities. **Adapted from César Merino thesis**

	SteT	$b^{0,+}$AT	xCT	LAT-1	LAT-2	y^+LAT-1	y^+LAT-2	asc-1	AGT-1
$b^{0,+}$AT	**26**								
xCT	**29**	42							
LAT-1	**28**	43	45						
LAT-2	**30**	39	41	51					
y^+LAT-1	**28**	42	42	45	46				
y^+LAT-2	**29**	41	42	45	45	70			
asc-1	**30**	39	39	45	63	42	43		
AGT-1	20	31	30	29	23	30	31	28	
Mmasc-2	18	29	26	27	23	26	25	24	23

One of the main characteristics of SteT primary sequence is the absence of the highly conserved cysteine residue present in all mammalian LATs in the extracellular loop between TMD3 and TMD4 that is involved in the disulfide bridge with the heavy subunit. At present, it is still unknown the existence of homologs of HATs heavy subunits in prokaryotic organisms.

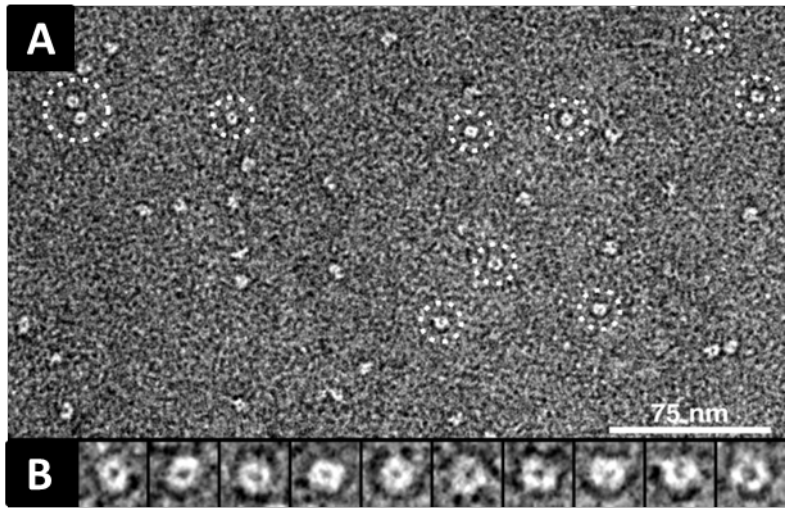


Figure 11. TEM of negatively stained SteT. The homogeneity of the DM-solubilized SteT proteins is reflected in the electron micrograph (A). The selected top view particles marked by broken circles were magnified and are displayed in the gallery (B). SteT proteins are elliptical (diameters ~6 ×~7 nm) and donut-shaped with a central depression. The frame size of the magnified particles in the gallery is 12 nm. **Adapted from (Reig et al., 2007)**

Interestingly, negative staining transmission electron microscopy (TEM) of detergent-purified SteT revealed that this protein is a monomer with elliptical shape according its dimensions (diameters 6x7 nm), presenting, as well, a central depression (Figure 11) (Reig et al., 2007).

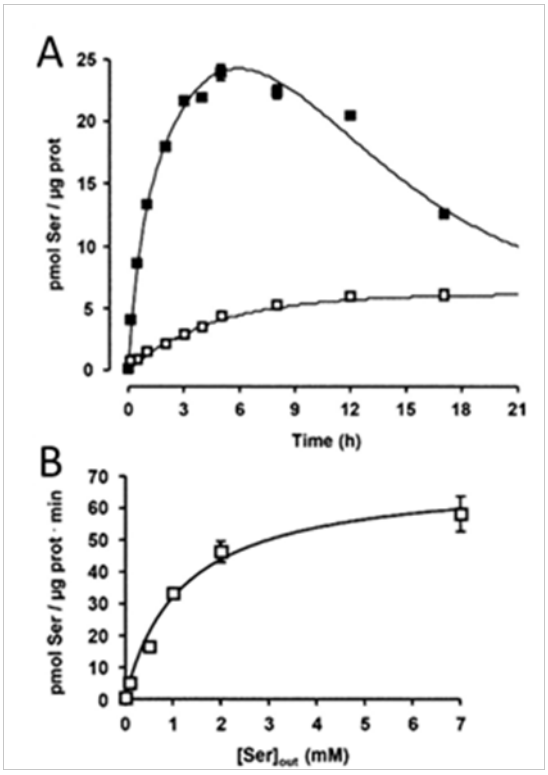


Figure 12. A) Time course of l-serine transport in SteT proteoliposomes. Influx (pmol/µg protein) of 10 µm radiolabeled l-serine into SteT proteoliposomes (SteT-PLs) containing 4 mml-serine (closed squares) or l-arginine (open squares). (white) Figure adapted (Reig et al., 2007) B) Kinetic analysis of SteT transport: transport of radiolabeled l-serine at different concentrations in the external (as indicated) and the internal (0.4, 1, 4, and 12 mM) medium was measured in SteT proteoliposomes (SteT-PLs) for 1 min under linear conditions. **Adapted from (Reig et al., 2007)**

A series of functional experiments with purified SteT reconstituted in proteoliposomes composed by *E. coli* phospholipids, revealed that SteT is a electroneutral obligatory exchanger (Figure 12) (Reig et al., 2007). It transport preferentially L-serine and L-threonine (L-Se/L-Thr exchanger), although it can also recognize other aromatic amino acids with less affinity in a (1:1) stoichiometry (Reig et al., 2007). Kinetic analysis of L-Ser/L-Ser exchange activity revealed an apparent half-saturation constant (K_M) of ~1.2 mM, a translocation rate (V_{MAX}) of 67 pmol·µg^{-1}·protein^{-1}·min^{-1}, and a turnover rate of 0.06 s^{-1} (Figure 12 B) (Reig et al., 2007).

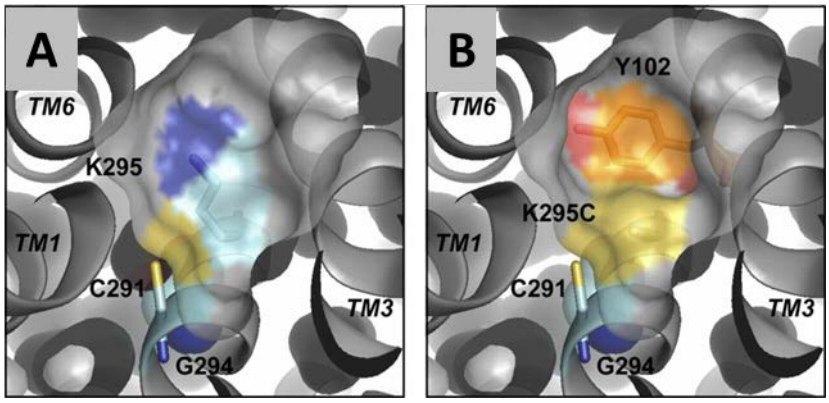


Figure 13. Putative substrate binding pocket of wild-type and K295C SteT. Upper view of the putative substrate binding pocket of wild-type SteT (A) and its K295C mutant (B), seen from the periplasmic space. This structural model is based on the open-to-out structure conformation of AdiC (Fang et al., 2009). The TM8 residues Cys-291 and Lys-295 are located at the surface of the bottom of the substrate binding pocket. In contrast, the TM8 residue Gly-294 (spheres) is not accessible to the solvent. Mutation K295C enlarges the substrate binding pocket in ~90 Å3 and residue Tyr-102, in TM3, became accessible at the bottom of the cavity. **Adapted from (Bartoccioni et al., 2010)**

Our understanding of the main residues of SteT involved in substrate affinity and specificity comes from a cysteine-scanning mutagenesis study of TMD 8 (Bartoccioni et al, 2010). Cysteine reactivity of single-cysteine mutants of TMD 8 introduced in a functional cysteine-less SteT mutant, combined with functional assays revealed the role of Cys 291 and Lys 295 in substrate recognition and specificity (Bartoccioni et al., 2010). In particular, Cys 291 modification (equivalent to Ser 284 in AdiC) with sulfhydryl agents inactivates transport activity in a substrate protectable manner suggesting that this residue is in a close proximity or interacting with the substrate. Moreover, substitution of Lys 295 (equivalent to Trp 293 in Adic) by Cys boards substrate specificity with the exception of imino acids (Bartoccioni et al, 2010). A structural model of SteT was built taking the outward-facing apo conformation structure of AdiC

(Gao et al., 2009) as template (Bartoccioni et al., 2010). The model locates Cys 291 at surface of the binding pocket whereas the side chain of Lys 295 is sitting at the bottom of the cavity (Figure 13).

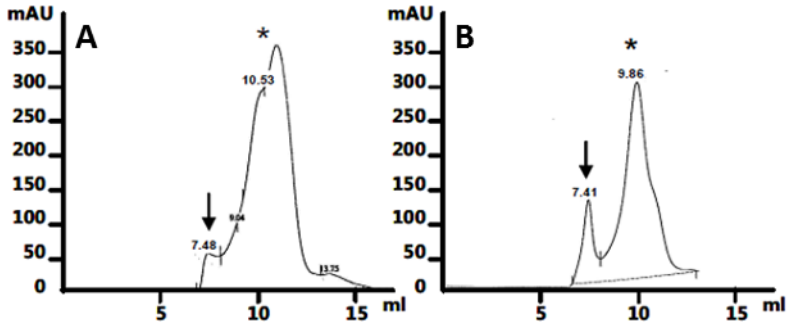


Figure 14. Size-exclusion chromatography of purified SteT. (a) SteT elution after injecting at 1 mg/mL to a Superdex 200 10/300 column. (b) SteT elution after injecting at 3 mg/ml in the same column. Asterisks point out the correct elution peak of the proteins. Arrows indicate the void volumes of the column in each plot. **(**Vázquez-Ibar JL and Palacín, M; unpublished results).

Further functional experiments showed the ability of the reducing agent, DTT, to activate the transport activity in proteoliposomes of the lousy-transporting SteT single cysteine mutants: S287C, G294C and S298C (Bartoccioni et al., 2010). In vivo functional analysis strongly suggested that these mutants were S-thiolated before detergent solubilization in the expressing host *E. coli*, indicating that this modification inhibits any further conformational change of the protein needed for substrate translocation. As all of them line in the same α-helical face in TMD 8, the authors concluded that TMD 8 undergoes conformational changes during the transport cycle of SteT (Bartoccioni et al., 2010).

SteT is an interesting target for crystallization studies, since it is the prokaryotic member of LATs with the closest amino acid identity (26 to 30% of amino acid

identity, Figure 10). It is reasonable to think that the atomic structure of SteT would be a better model for the eukaryotic LATs than the available crystal structures of AdiC (and average of 18% amino acid identity with LATs). Furthermore, the functional properties of SteT (broad substrate recognition, obligatory exchanger) resemble better the functional behavior of the majority of LATs. Unfortunately, SteT is very unstable after detergent solubilization from the membrane and subsequent purification. At 1 mg/ml concentration the DDM-purified protein stays in solution less than two days at 4°C and starts precipitating at concentrations higher that 3 mg/ml (Vázquez-Ibar JL and Palacín, M; unpublished results). Moreover, in size exclusion chromatography SteT has wide and multiple peaks in the elution profiles suggesting the presence of different aggregation states and, therefore, denoting instability (Figure 14) (Vázquez-Ibar JL and Palacín, M; unpublished results). All of this completely hampers any crystallography work. At this point, our group thought that high-throughput methods based on the combination of random mutagenesis and fast screening protocols will be the shortest path in order to identify more stable mutants of SteT suitable for crystallization work.

OBJECTIVES

3 OBJECTIVES

Membrane proteins are challenging targets for structural biologists. Finding optimal conditions to handle membrane proteins for crystallographic studies requires extensive and laborious screening of protein expression and/or stability in detergent. In fact, if the membrane protein is unstable and low expressed is almost impossible to initiate any crystallization work. Giving the successful use of mutagenesis for membrane proteins stabilization and crystallization, we reasoned that it was necessary to build and optimize an experimental protocol to obtain functional mutants of a membrane protein target with optimal characteristics to initiate crystallization studies. In this context, the combination of random mutagenesis with rapid and sensitive screening protocols of protein expression and stability seems to be the best approach for this goal.

3.1 Main objective

The main purpose of this thesis is to build up an experimental protocol with the objective to generate and characterize functional mutants of SteT with larger expression and improved stability after detergent solubilization and, therefore, with the required robustness for crystallization trials. This protocol is conceived to be a general methodology for any membrane protein of interest.

3.2 Specific objectives:

3.2.1 Construction of a library of random SteT mutants

The first goal will be to generate a random library of mutants of SteT that express and insert into the cytoplasmic membrane of the host, *E. coli*. For this

purpose, a combination of a random mutagenesis strategy using error prone PCR with a screening method based on a split version of GFP will be used.

3.2.2 Evaluation of the expression and stability of selected SteT mutants

The second goal will be to compare the relative expression and stability of the generated SteT random mutants with respect to wild-type SteT. The final aim will be to find out from the initial mutant library, optimal candidates for further studies of purification, functional studies and crystallization screening.

3.2.3 Analysis of purified SteT mutants

The third goal will be to select a final candidate for crystallization studies. For this purpose, the robustness of the previously selected mutants after purification and solubilization in different detergents commonly used for crystallization will be evaluated.

3.2.4 Functional studies of optimal mutants for crystallization

In this goal, transport assays of the selected mutants for crystallization will be performed in order to evaluate the functional impact of the mutations.

3.2.5 Crystallization screenings

Finally, the best mutant candidate will be sent to crystallization trials.

RESULTS AND DISCUSSION

4 RESULTS AND DISCUSSION

4.1 Construction of a library of SteT random mutants

In this first objective, the goal was to make a library of random mutants of SteT, where amino acid substitutions will be randomly distributed along the sequence of the protein. In order to reduce the effect of such mutations on SteT folding and function, we decided to include in the library mutants with only one or two amino acids substitutions. In this way, we can have a reasonable probability to preserve the integrity of SteT and, therefore, its function. In addition, in order to have more impact on the stability of the protein we decided to select only those mutants whose amino acid substitutions are found in the putative TMDs. Interactions between TMD are the major determinant in the assembly and stability of the native structure of IMPs (Thévenin and Lazarova, 2008)(Dalbey et al., 2011). Also, TMDs have different important roles in the expression and stability of IMPs (Langosch et al., 2007). Finally, to start with the mutagenesis, we decided to use a cDNA encoding SteT where all codons were optimized for *E. coli* expression. It is known that optimizing cDNA for a particular host can increase protein expression by simply codon usage or by increasing the stability of mRNA (Grunberg-Manago et al., 1999). This is a very well known strategy that has been successfully used for the heterologous expression of both soluble and membrane proteins in different organisms (Gvritishvili et al., 2010) (Wang Q. et al, 2012).

4.1.1 Random mutagenesis

SteT cDNA was cloned in vector pTETGFP11 (Cabantous and Waldo, 2006). Random mutagenesis of the cDNA encoding SteT was performed by error-prone PCR using a low fidelity DNA polymerase (Mutazyme II DNA® polymerase). PCR products obtained from the error-prone PCR were cloned into the $pTETGFP_{11}$ vector by a second PCR using a high fidelity polymerase, where the previous PCR products were the megaprimers and the template was $pTETGFP_{11}$-SteT (see materials and methods for details) (Figure 15).

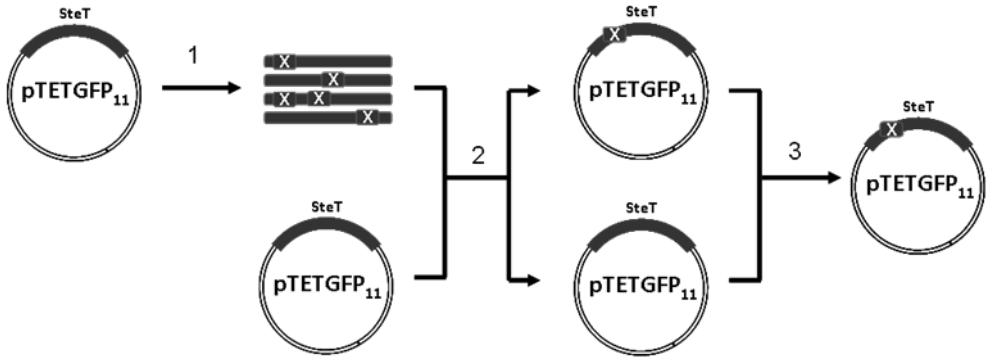


Figure 15. Scheme of the random mutagenesis: (1) Error prone PCR Genemorph EZClone, (2) EZClone Reaction to obtain the plasmids of interest with the random SteT mutant. (3) DpnI enzyme digestion

The resulting PCR was incubated with *Dpn*I in order to eliminate the parental non-mutated plasmid and transformed in *E. coli* BL21 (DE3) Star cells. Transformed cells were seeded in a LB-plate with the antibiotic spectinomycin and after overnight incubation at 37 °C, *E. coli* colonies appeared.

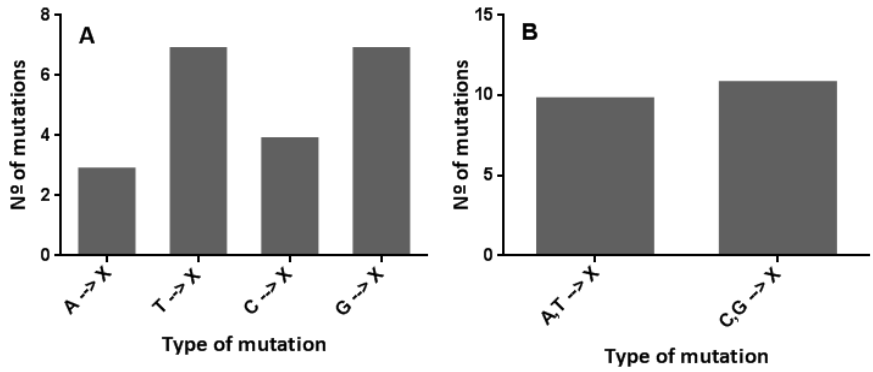


Figure 16. Mutagenesis analysis of the chosen error prone PCR of SteT cDNA. (a) Mutation rate of, adenine (A), thymine (T), cytosine (C) or guanine (G) by any nucleotide (X). (b) Mutation rate of A,T or C,G to other another nucleotide (X). Data comes from a total of 20 colonies obtained using the protocol of Figure 15.

In order to obtain the desired low mutation rate (one or two mutations per SteT sequence) we conducted calibration studies by varying the amount of DNA template, and the number of cycles during the error-prone PCR. 20 clones of each experiment were sequenced and analyzed, obtaining an optimal error prone PCR conditions when using 500 ng of DNA template and 28 cycles in the PCR. The mutation rate was 1.5 nucleotide changes per SteT cDNA sequence. Figure 16 shows the results of the mutational analysis using the previous conditions. Whereas the mutation rate of thymine and guanine were higher than the obtained with adenine and cytosine (Figure 16A), the total mutational rate of the complementary nucleotides (adenine/thymine or cytosine/guanine) were similar (Figure 16B), as expected from manufactured of the error prone PCR enzyme (Mutazyme II® DNA polymerase).

4.1.2 GFP split system as reporter screening

After calibrating the random mutagenesis reaction and optimizing the cloning strategy, a fast and sensitive screening protocol was necessary in order to

select the SteT mutants that readily express and insert into the cytoplasmic membrane of the expression host, *E. coli*. As introduced earlier, GFP fused to a IMP target is a sensitive reporter that has enormously facilitated precrystallization screening of these proteins (Drew et al., 2006), (Kawate and Gouaux, 2006), (Simon Newstead, 2007) and (Hammon et al., 2008). Unfortunately, in some cases GFP can interfere with the expression and/or the stability of the IMP target (Fucile et al., 2002). Moreover, recent data reveals that the stability profile of a GFP-fused IMP changes dramatically after removing the GFP by proteolysis, making very difficult to interpret the screening experiments (Hsieh et al., 2010). Taking into account all of this, a reporting system based on a split version of the GFP, originally developed as screening test for solubility of globular proteins (Cabantous and Waldo, 2006), was choose to identify those random mutants of SteT that retain the ability to express and properly fold in the cytoplasmic membrane of *E. coli*. This strategy not only allows checking the expression of the mutants in a fast way manner using the fluorescence of the GFP, but also minimizes the side effects of having a bulky and robust β-barrel protein, the GFP, fused on the C-terminal end of each mutant. We first conducted a series of experiments aiming to test the applicability and robustness of this method. In addition, other IMPs were also included in these preliminary experiments in order to evaluate the general use of the GFP split system in IMPs precrystallization tests.

4.1.2.1 Coexpression of SteT–GFP_{11} and $GFP_{1–10}$ in *E. coli* leads to GFP fluorescence

First, we studied whether the emission of fluorescence from GFP can be detected after the complementation of the non-fluorescent C-terminal end of GFP ($GFP_{1–10}$) with the remaining 15 amino acids of GFP (GFP_{11}) fused to the C-terminal end of a membrane protein target expressed in *E. coli* (Figure 17). We

challenged this GFP complementation assay by testing the expression of SteT with GFP_{11} fused at its C terminus in *E. coli* (Figure 17).

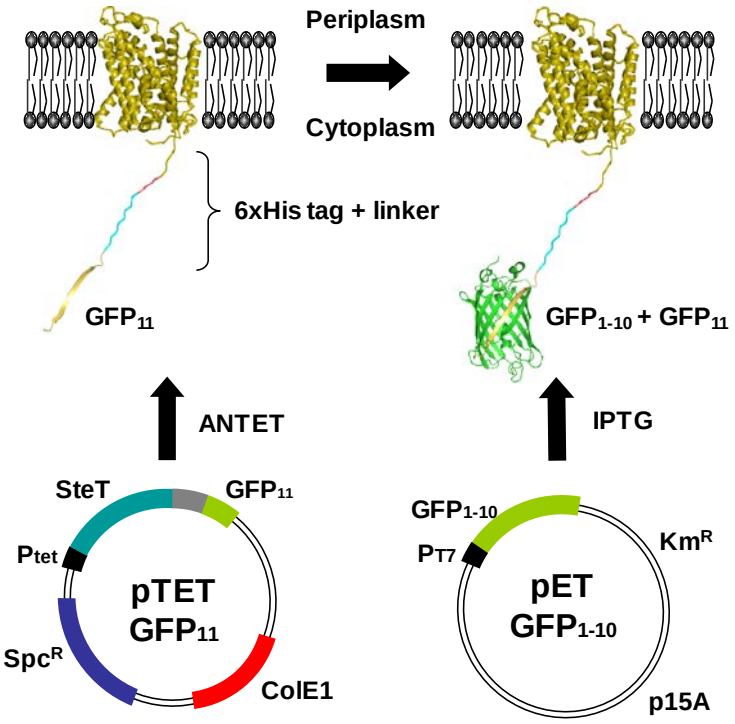


Figure 17. Schematic of the split GFP system adapted to membrane proteins. The split GFP system consists of two plasmids: pTET–GFP_{11} and pET–$GFP_{1–10}$ (Cabantous and Waldo, 2006). pTET carries the gene encoding the target membrane protein (SteT) fused to a small part (15 amino acids) of GFP (GFP_{11}) at its C terminus, and pET carries the gene encoding the rest of the GFP molecule ($GFP_{1–10}$, 215 amino acids). Plasmids are compatible, containing the ColE1 and the p15A origins of replication, respectively. They also encode two antibiotic resistance genes: spectinomycin (SpcR) in pTET and kanamycin (KmR) in pET. Protein expression is controlled by two promoters: Ptet (ANTET inducible) in pTET and PT7 (IPTG inducible) in pET. The expression of these genes can be induced simultaneously or sequentially, and complementation occurs when the GFP_{11}-fused membrane protein is expressed and inserted into the cytoplasmic membrane of *E. coli*.

The resulting fluorescent signal is expected to be proportional to the amount of protein expressed. In our assay, SteT–GFP_{11} and $GFP_{1–10}$ were encoded in two compatible expression vectors (modified versions of pTET and pET, respectively

(Figure 17) (Cabantous et al., 2005) and (Cabantous and Waldo, 2006). Importantly, the expression of the two proteins is regulated by two independent promoters; therefore, the expression of either SteT–GFP_{11} or $GFP_{1–10}$ is tightly controlled by simply adding or removing the appropriate inducing agent (ANTET for SteT–GFP_{11} and IPTG for $GFP_{1–10}$) (Figure 17). As described originally (Cabantous and Waldo, 2006), we also added a 10-amino-acid flexible linker (DGGSGGGSTS) between GFP_{11} and the C terminus of SteT to prevent steric restrictions that can hamper GFP_{11}–$GFP_{1–10}$ complementation.

The coexpression of SteT–GFP_{11} and $GFP_{1–10}$ in the same cell produced the typical spectrum of the GFP fluorescence after exciting the cells at 460 nm (Figure 18A). No fluorescence was detected when either SteT–GFP_{11} or $GFP_{1–10}$ was expressed independently (Figure 18A). Interestingly, a similar expression test can also be conducted with bacterial colonies, as shown in Figure 18B. By simply passing a nitrocellulose membrane with *E. coli* colonies cotransformed with the two expression vectors into separate agar plates containing the appropriate inducing agent (see Materials and methods for details), SteT expression can be monitored by observing the GFP fluorescence of the bacterial colony in a similar way as described previously for soluble proteins (Cabantous and Waldo, 2006). The feasibility of the split GFP assay for measuring the yield of SteT expression in *E. coli* was confirmed by Western blot analysis using the 6 × His tag epitope placed between the C terminus of SteT and the N terminus of GFP_{11} (Figure 19A). In these experiments, isolated cytoplasmic membranes of an *E. coli* culture sequentially expressing SteT–GFP_{11} followed by $GFP_{1–10}$ were subjected to SDS–PAGE and immunoblotted with HisProbe–HRP. These experiments showed that the nonfluorescent SteT–GFP_{11} band appeared only in the absence of $GFP_{1–10}$ (Figure 19A, lanes 1 and 2).

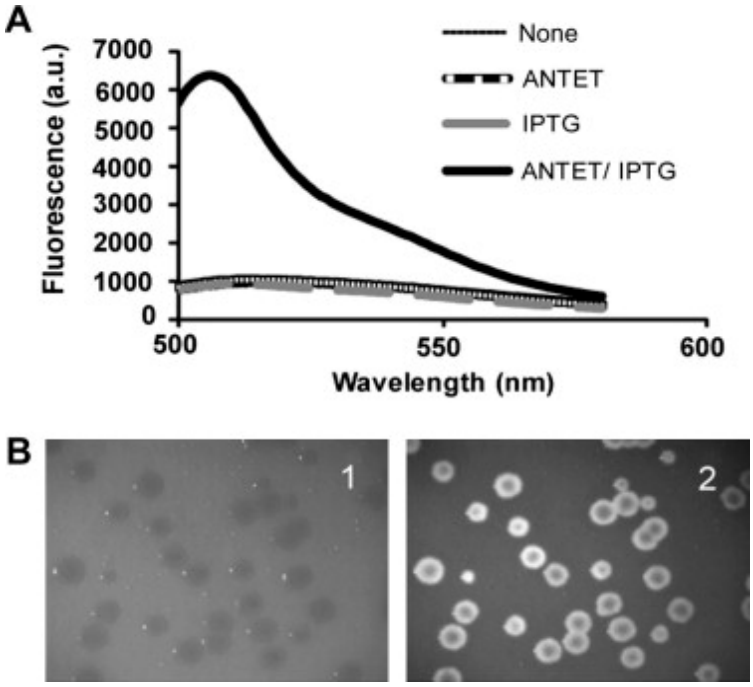


Figure 18. In vivo coexpression of SteT–GFP_{11} with $GFP_{1–10}$ leads to GFP fluorescence. (A) Fluorescent spectra measured in *E. coli* cultures harboring pTET and pET plasmids encoding SteT–GFP_{11} and $GFP_{1–10}$, respectively. Spectra were recorded using a 460-nm excitation wavelength in noninduced cells (none) or by inducing SteT–GFP_{11} (ANTET), $GFP_{1–10}$ (IPTG), or SteT–GFP_{11} and $GFP_{1–10}$ together (ANTET/IPTG) for 3 h at 30°C. (B) In vivo SteT–GFP_{11}–$GFP_{1–10}$ complementation can be detected in *E. coli* colonies. Colonies harboring pTET and pET plasmids encoding SteT–GFP_{11} and $GFP_{1–10}$, respectively, were grown in a nitrocellulose filter on top of an LB plate containing the appropriate antibiotics and inducing agents. After SteT–GFP_{11} and $GFP_{1–10}$ were expressed sequentially and complemented, GFP fluorescence from the colonies was observed under the fluorescence microscope without excitation (panel 1) or with UV excitation (panel 2).

Furthermore, a higher molecular weight band corresponding to SteT–GFP_{11}–$GFP_{1–10}$ appeared when $GFP_{1–10}$ was induced and increased in intensity over time (1 and 16 h after $GFP_{1–10}$ induction), whereas the intensity of the SteT–GFP_{11} band decreased (Figure 19A, lanes 3 and 4). Moreover, a fluorescent band at the expected molecular weight of SteT–GFP_{11}–$GFP_{1–10}$ appeared in an SDS–PAGE gel containing isolated *E. coli* membranes coexpressing SteT–GFP_{11} and $GFP_{1–10}$ (Figure 19B). This finding again confirms the presence of SteT–

GFP_{11}–$GFP_{1–10}$ in the cytoplasmic membrane. In addition, these results corroborate that complementation between SteT–GFP_{11} and $GFP_{1–10}$ occurs after SteT is fully translated (the GFP_{11} tag is at the C terminus) and inserted into the cytoplasmic membrane of the expression host. As described previously (Cabantous et al., 2005) the SteT–GFP_{11}–$GFP_{1–10}$ band increased with longer $GFP_{1–10}$ induction times (Figure 19A, lanes 3 and 4) as a result of a higher cytoplasmic concentration of $GFP_{1–10}$ and a longer time for $GFP_{1–10}$–GFP_{11} complementation.

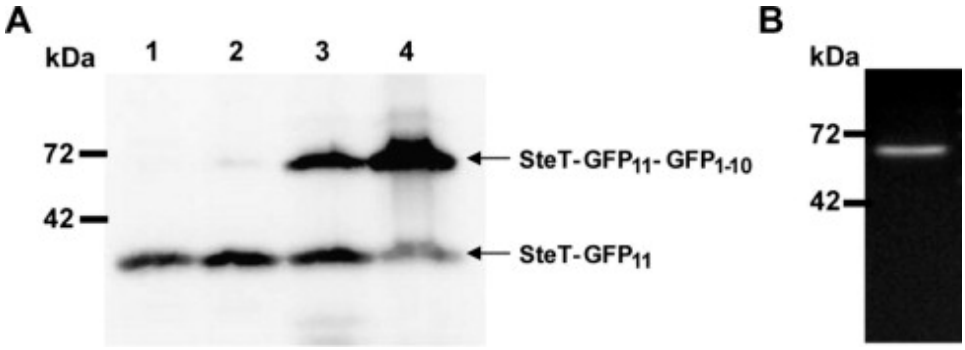


Figure 19. Complementation between SteT–GFP_{11} and $GFP_{1–10}$ occurs in the cytoplasmic membrane. (A) Anti-His-tag Western blot analysis of SteT–GFP_{11} and complemented SteT–GFP_{11}–$GFP_{1–10}$ expression in *E. coli* cytoplasmic membranes. *E. coli* membranes expressing 6 × His-tagged SteT–GFP_{11} before and after $GFP_{1–10}$ induction were solubilized with 1% DDM and subjected to SDS–PAGE before blotting. Lanes: SteT–GFP_{11} induction at 30°C for 1 h (lane 1) or 2 h (lane 2) and SteT–GFP_{11} induced for 3 h at 30°C followed by $GFP_{1–10}$ induction for 1 h (lane 3) or 16 h (lane 4). (B) In-gel fluorescence of an SDS–PAGE gel containing isolated *E. coli* membranes coexpressing SteT–GFP_{11} and $GFP_{1–10}$.

4.1.2.2 The split GFP system specifically measures SteT–GFP_{11} inserted into the cytoplasmic membrane

In some cases, the heterologous expression of membrane proteins in *E. coli* leads to the accumulation of these proteins as aggregates in inclusion bodies (Koth and Payandeh, 2009). Because GFP can be fluorescent in inclusion bodies (García-Fruitós et al., 2005), false positive errors can be generated from

misfolded or insoluble proteins located in these particles. Interestingly, GFP_{11}–$GFP_{1–10}$ does not complement inside inclusion bodies (Cabantous et al., 2005). Consequently, we reasoned that the sequential expression of each GFP fragment (GFP_{11}-fused membrane protein followed by $GFP_{1–10}$) could be a valuable expression assay to discard membrane proteins or SteT mutants that accumulate in inclusion bodies. To test this hypothesis, we induced the expression of SteT–GFP_{11} at two temperatures (30 and 37°C) for 2 and 16 h at each temperature (Figure 20). Subsequently, $GFP_{1–10}$ was induced for 1 h at 30°C in all of the conditions tested. The fluorescence signal measured in *E. coli* cultures after GFP_{11}–$GFP_{1–10}$ complementation indicates that the expression yield of SteT–GFP_{11} was substantially higher at 30°C than at 37°C (Figure 20A), as reported previously using a non-GFP-tagged version of SteT (Reig et al., 2007).

To ensure that the fluorescence signal came almost exclusively from the cytoplasmic membrane, we performed SDS–PAGE analysis of cytoplasmic membranes and inclusion bodies from the same *E. coli* cultures. The in-gel fluorescence of isolated membranes revealed the presence of the complemented SteT–GFP_{11}–$GFP_{1–10}$ at both temperatures (Figure 20B). Notably, the fluorescence intensity of each band was consistent with that measured previously in *E. coli* cultures (30°C > 37°C [Figures 20A and 20B]).

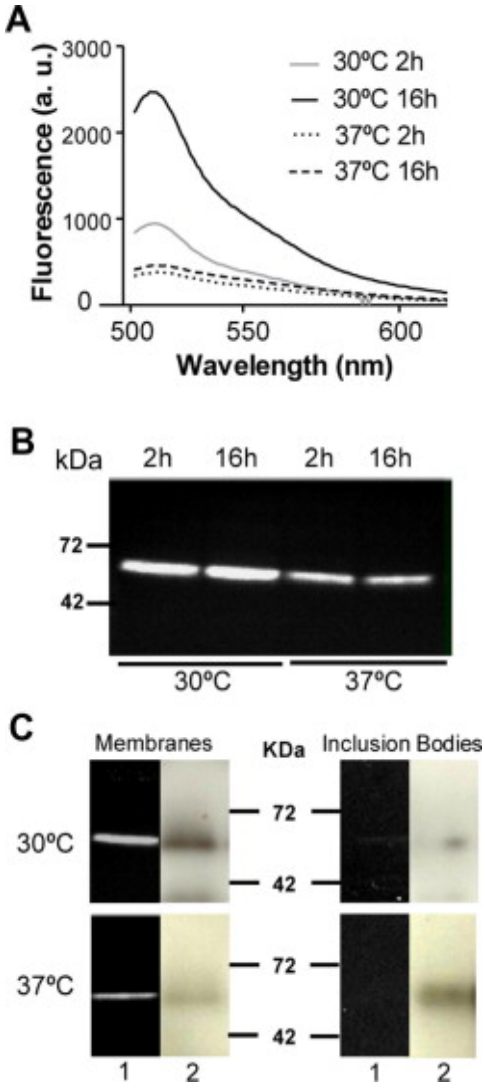


Figure 20. The split GFP system specifically measures the expression of SteT–GFP$_{11}$ in the cytoplasmic membrane. (A) GFP fluorescence spectra of complemented SteT–GFP$_{11}$–GFP1–10 measured with varying conditions of SteT–GFP$_{11}$ induction as indicated in the figure. After SteT–GFP$_{11}$ induction, GFP1–10 was induced for 3 h at 30°C in all experiments. (B) In-gel fluorescence of SDS–PAGE gels containing isolated *E. coli* cytoplasmic membranes expressing SteT–GFP$_{11}$ under a range of conditions as indicated followed by GFP1–10 induction for 3 h at 30°C in all experiments. (C) Analysis of SteT–GFP$_{11}$–GFP1–10 expression by GFP fluorescence emission (lanes 1) and anti-His-tag Western blot (lanes 2) of both the cytoplasmic membrane fraction and purified inclusion bodies of *E. coli* cultures expressing SteT–GFP$_{11}$ at either 30 or 37°C. After SteT–GFP$_{11}$ induction, GFP$_{1–10}$ was induced for 3 h at 30°C in all experiments.

Furthermore, anti-His-tag Western blot analysis of isolated cytoplasmic membranes confirms that the amount of SteT expressed at each temperature and inserted into the cytoplasmic membrane is also consistent with the fluorescence intensity measured in bacterial cultures and in-gel fluorescence (Figures 20A–C). Interestingly, the same Western blot analysis of the isolated inclusion bodies revealed that, at 37°C, SteT–GFP_{11} (complemented with $GFP_{1–10}$) accumulates in inclusion bodies to a larger extent than at 30°C, in contrast to the results obtained in the cytoplasmic membrane fraction (Figure 20C). Remarkably, only a nominal fluorescent signal was observed in the inclusion body fraction at the two temperatures (Figure 20C). These results confirm that even if the two GFP fragments interact in the cytoplasm before becoming confined to inclusion bodies, the fluorescence emission is almost completely quenched, so the protein fraction present in inclusion bodies does not contribute to the fluorescence signal. Therefore, when sequentially expressing SteT–GFP_{11} followed by $GFP_{1–10}$, the fluorescence signal not only reflects the expression yield of SteT but also specifically the amount of protein inserted into the cytoplasmic membrane (Figures 20A–C). This strategy is highly beneficial because it allowed us to quickly discard the protein fraction confined to inclusion bodies as a result of aggregation or misfolding, a common issue in the heterologous expression of membrane proteins (Koth and Payandeh, 2009), and indeed, it represents an extremely useful screening assay to test the expression yield of a mutant library (like the SteT random library), making sure that the amount of protein measured is folded and inserted in the membrane.

4.1.2.3 The split GFP can be used as general reporter of the expression yield of membrane proteins in *E. coli*

We next explored the robustness of this GFP complementation strategy by studying the expression and membrane insertion of four membrane proteins in *E. coli:* the lactose permease of *E. coli* (LacY), the small multidrug transporter of *E. coli* (EmrE), the small conductance mechanosensitive channel of *E. coli* (MscS), and the large conductance mechanosensitive channel of *E. coli* (MscL) (Figure 22). The selected proteins are well characterized structurally and differ in the number of TMDs and in their quaternary structures. LacY is a monomer composed of 12 TMDs (Abramson et al., 2003), EmrE is a homodimer with each monomer containing 4 TMDs (Chen et al., 2007), MscS is a homoheptamer containing 3 TMDs per monomer (Bass et al., 2002), and MscL forms a pentameric structure with 2 TMDs per monomer (Chang et al., 1998). Each protein was cloned in the pTET vector fused to GFP_{11} on its C terminus (Figure 17) and, as in the case of SteT; a linker comprising 10 amino acids was added between the C-terminal end of the membrane protein and GFP_{11}. Membrane protein expression was induced at 30°C for 3 h in all cases; subsequently $GFP_{1–10}$ expression was induced at 30°C for 1–16 h. For each tested protein, we measured the GFP fluorescence from the bacterial culture (Figure 21A) and from an SDS–PAGE gel of the isolated cytoplasmic membranes (Figure 21B). As found for SteT (Figure 21A), the fluorescence intensity of GFP increased with longer $GFP_{1–10}$ induction times (1–16 h) (Figure 21A). The fluorescent bands observed in the SDS–PAGE gels of cytoplasmic membranes expressing each target matched the molecular weight of each GFP-fused membrane protein, similar to the findings of the SteT studies (Figure 21B). These examples indicate that this split GFP system is a robust strategy to screen the expression of membrane protein candidates with distinct topologies. The only requirement

for the application of this approach is that the C-terminal end of the protein is oriented toward the cytoplasmic site.

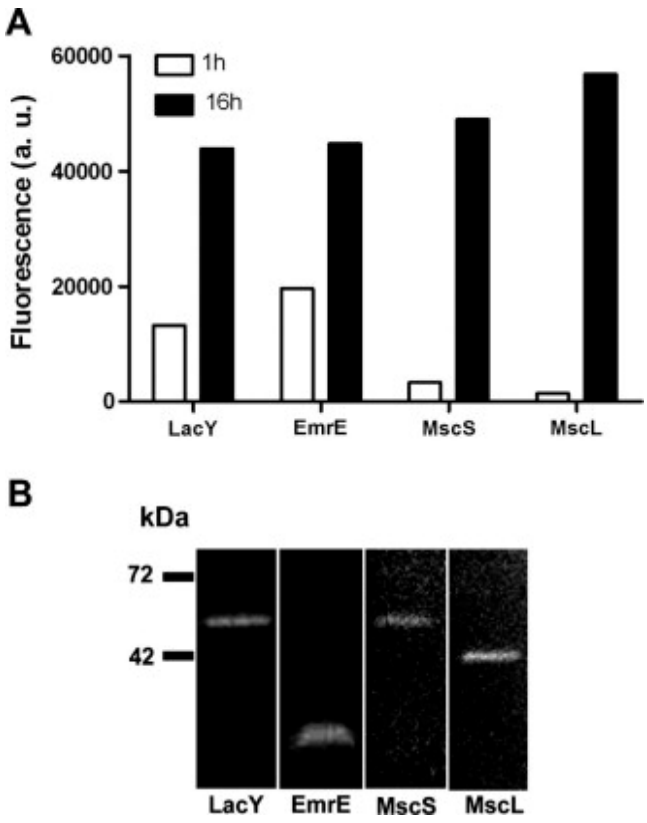


Figure 21. The split GFP system can measure expression yield of membrane proteins in *E. coli.* (A) GFP fluorescence measured in *E. coli* cultures expressing LacY–GFP_{11}, EmrE–GFP_{11}, MscS–GFP_{11}, or MscL–GFP_{11} plus $GFP_{1–10}$. Membrane proteins were induced at 30°C for 3 h followed by $GFP_{1–10}$ induction at 30°C for either 1 or 16 h as indicated. (B) In-gel fluorescence of SDS–PAGE gels containing isolated *E. coli* membranes sequentially expressing LacY–GFP_{11}, EmrE–GFP_{11}, MscL–GFP_{11}, or MscS–GFP_{11} (30°C for 3 h) followed by $GFP_{1–10}$ (16 h at 30°C).

4.1.3 Selection of the expressed SteT mutants

After calibrating the error prone PCR for a low mutation rate (Section 3.1.1), the library of SteT random mutants, cloned in the $pTETGFP_{11}$ vector, were transformed in *E. coli* BL21 (DE3) STAR cells harboring the $pETGFP_{1-10}$ vector. The transformation was plate in nitrocellulose membranes sitting on top of agar-LB plates where, after performing the "in vivo" GFP split assay in the bacterial colonies as described in section 3.1.2.2, clones with SteT mutants that express and insert in the membrane will be identified. The protocol is schematized in Figure 22 (see materials and methods for details).

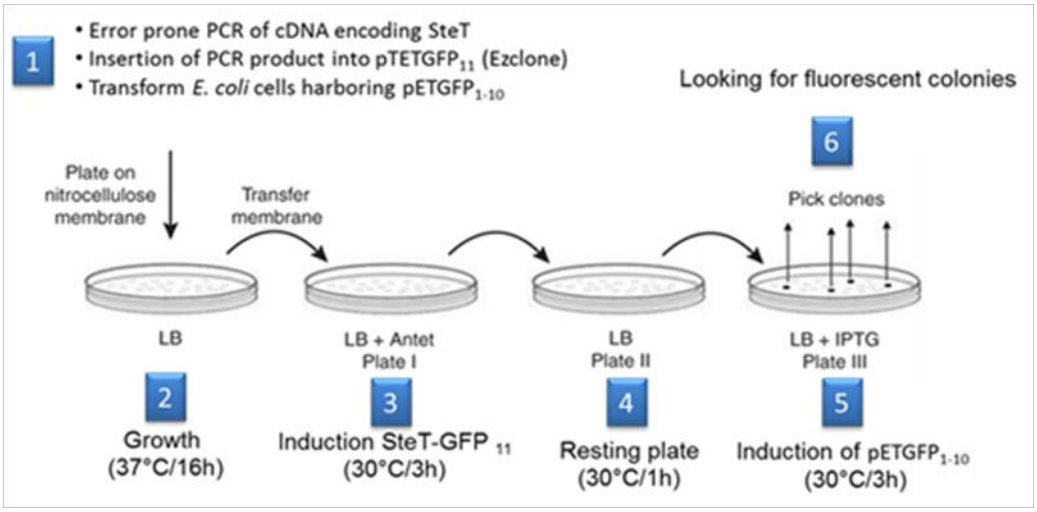


Figure 22. Scheme of the protocol used for SteT random mutants selection using the GFP split assay. (1) Creation by error prone PCR and cloning the SteT random mutant library. The resulting library was transformed in *E. coli* BL21 (DE3) cells harboring the $pETGFP_{1-10}$ vector and plated on nitrocellulose membranes. (2) Transformed *E. coli* colonies grew during 16 h at 30 °C. (3) SteT-GFP_{11} was induced after transferring the membrane to LB plates containing ANTET during 3 h at 30 °C. (4) Excess of ANTET was removed by transferring the membrane to LB plates with no inducing agent (resting plates) during 1 h at 30 °C. (5) GFP_{1-10} is induced by transferring the membranes to LB plates containing IPTG during 3 h at 30 °C. (6) Select the E. coli colonies that emit GFP fluorescence. Adapted from (Cabantous and Waldo, 2006)

The plates were incubated during 22 h at 37 °C and SteT-GFP_{11}, was induced by transferring the membrane to a LB plate with ANTET during 3 h at 30 °C. Next, the nitrocellulose membrane was transferred to another LB plate for 1 h with no inducers (resting plate) to eliminate the ANTET thus avoiding the co-expression of SteT-GFP_{11} with GFP_{1-10}. As expected, no fluorescent colonies were observed during this step, so there was no basal expression of GFP_{1-10} that after complementation with the GFP_{11} will emit fluorescence. To express the GFP_{1-10}, the nitrocellulose membrane was transferred to a LB plate containing IPTG during 3h at 30 °C. As expected, fluorescent colonies were observed under UV exposition (Figure 23).

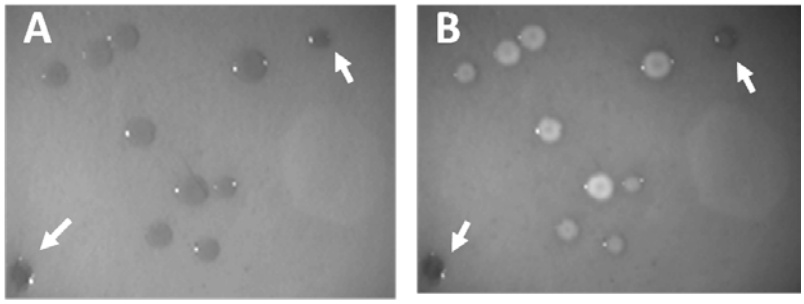


Figure 23. Selection of clones expressing SteT random mutants: *E. coli* colonies expressing sequentially SteT-GFP_{11} and GFP_{1-10} as seen in Figure 22, without excitation light (A) and after shading with UV light emission (B). The two pictures show exactly the same colonies. White arrows represent the clones that show no fluorescence and therefore no SteT mutant have expressed and/or inserted into de cytoplasmic membrane.

The clones that were selected were those that showed fluorescence under UV light using a Stereo Fluorescence Microscope (Figure 23). The plasmid DNA from the selected fluorescent colonies was extracted and sequenced using external primers to 5' and 3' of the cloning site. Overall, 533 fluorescent colonies were selected; and 395 out of these 533 colonies were fully sequenced. In addition, and as control, some clones showing no fluorescence were selected and sequenced, confirming the presence of a mutated version of

SteT. In several of these non-fluorescent clones, stop codons were found in the middle of the SteT sequence.

After sequencing, 149 mutants out of the 395 sent to the sequencing facility showed one or two amino acid substitutions. The other sequenced clones presented neutral mutations (nucleoside changed but no effective amino acid replacement), more than two amino acid substitutions or, simply no mutations (wild type SteT).

4.1.4 Localization of the mutations

After obtaining the 149 mutants with the desired amino acid substitution rate, the subsequently step was to select the ones that present mutations in the TMDs. For this purpose a SteT model (Bartochioni et al, 2010) was used (Figure 24). This model was built using, previous alignment (Figure 25), the X-ray structure of the close SteT homolog AdiC (PDB 3NCY), the L-arginine/agmatine exchanger of *Salmonella typhimurium* (~ 95 % identical to its *E. coli* ortholog) crystallized in an open outward-facing conformation (*Fang Y et al.*, 2009).

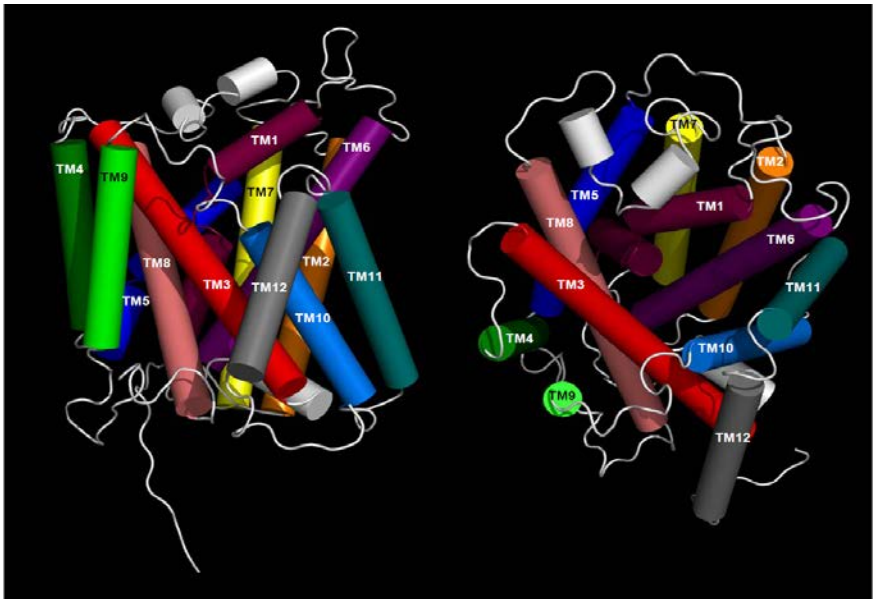


Figure 24. Representation of SteT 3D model. Side view (left) and periplasm view (right). Each TMD is indicated by its number and represented in a different color.

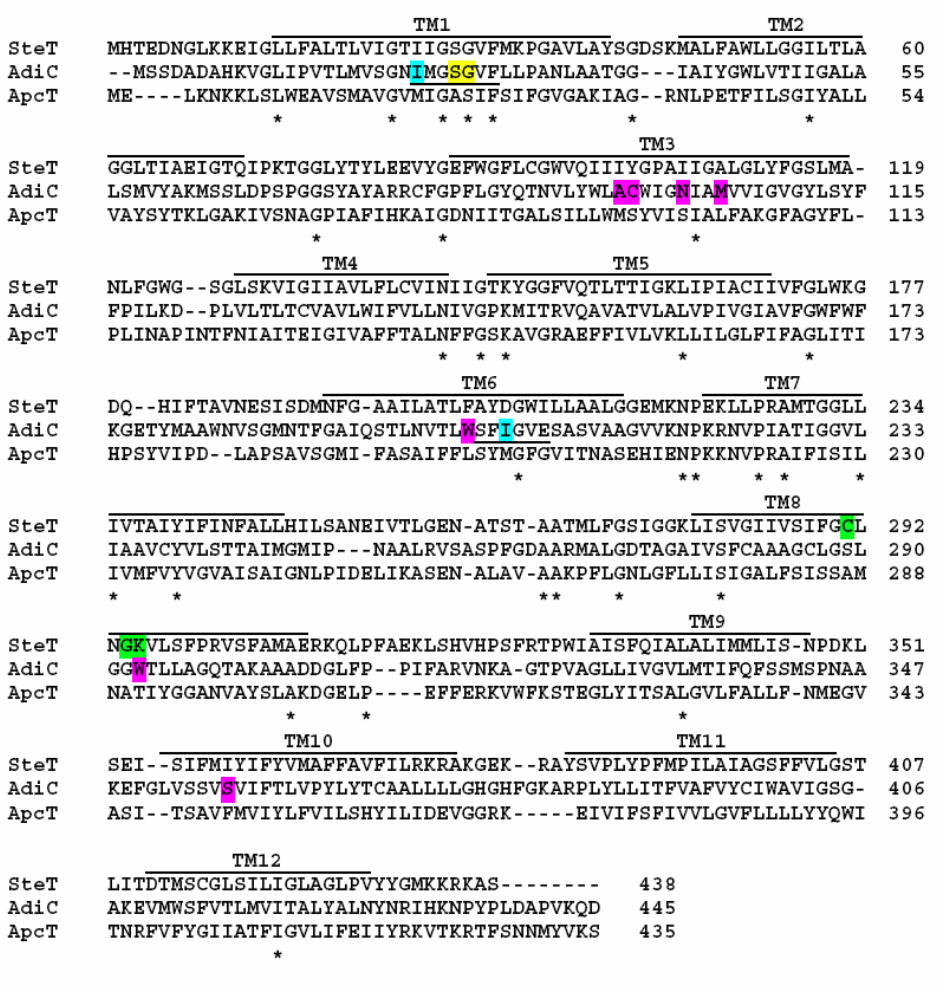


Figure 25. Multiple sequence alignment of the SteT, AdiC and ApcT proteins. This alignment was used to model SteT using the AdiC (*Salmonella typhimurium*) atomic structure (Fang Y. et al, 2009). TMDs are indicated by lines above the sequences. The unwound segments in TMD 1 and TMD 6 are underlined in the AdiC sequence. AdiC residues interacting with the α-amino group, the α-carboxylate group and the side chain of the L-arginine substrate are highlighted in blue, yellow and purple respectively. Functionally important residues in TMD 8 of SteT (Cys 291, Gly 294 and Lys 295) revealed in *Bartochioni et al*, 2010 are highlighted in green. *, identical residues in the three proteins. **Adapted from (*Bartoccioni et al*, 2010)**

Figure 25 shows the multiple sequence alignment of SteT, AdiC and ApcT. The positions of the TMD depicted in Figure 25 are the ones obtaining after building the SteT 3D model (Figure 24). Taking this model as reference, a total of 101 mutants from the previously 149 were selected, which were those that only

have amino acid substitutions in the predicted TMD. All these 101 mutants are listed in Table 4.

Table 4. SteT random mutants with one or two amino acids substitutions in TMD. The first 70 mutants of this list were further analyzed.

	MUTANT	TMD		MUTANT	TMD		MUTANT	TMD		MUTANT	TMD
1	G23R	1	**27**	G54S	2	**53**	L210Q-M229V	6-7	**79**	I409N	**12**
2	A398P	11	**28**	G152D-V370M	1-9	**54**	A369G	9	**80**	A136P	**4**
3	C168Y-L233M	5-7	**29**	L199M-L417M	6-12	**55**	C141W	4	**81**	A17G-G42V	**1**
4	G161N	5	**30**	T159I-S298T	5-8	**56**	F49Y-N347Y	2-9	**82**	F122S-Y205H	**6**
5	L53P	2	**31**	F89S-A105T	3-3	**57**	F139L	4	**83**	P328T	**9**
6	G339D	11	**32**	I235F	7	**58**	A197V-G232D	6-7	**84**	F304S	**8**
7	A267V	8	**33**	P34Q	1	**59**	F16L-I99V	1-3	**85**	E11D-A213T	**3-6**
8	G284V	8	**34**	F49Y	2	**60**	L63R	2	**86**	A369G	**10**
9	A109P	3	**35**	R374H	10	**61**	M32V-M342L	1-9	**87**	I56V-N293Y	**2-8**
10	C291S	8	**36**	G62C-F304S	2-8	**62**	A243P-A383S	7-11	**88**	A47P-K433P	**2**
11	G103S-L279P	3-8	**37**	P226Q	7	**63**	I107F-H249Y	3-7	**89**	I107T	**3**
12	A60E-C168R	2-5	**38**	A136E	4	**64**	I336N-M413S	9-12	**90**	L338V	**9**
13	F402S	11	**39**	V154E-G161S	5-5	**65**	L52Q-L174Q	2	**91**	Y150F-I170T	**5**
14	A39F	1	**40**	G69D	2	**66**	I164T	5	**92**	G91S	**3**
15	M392V	11	**41**	L14Q-T230A	1-7	**67**	F203S-R376P	6-10	**93**	K45E	**2**
16	A305T-T410S	8	**42**	L247V	7	**68**	G27A-T156S	1-5	**94**	G13A-L211P	**1-6**
17	I132F	4	**43**	F391Y	11	**69**	I134V-A377T	4-10	**95**	F16L	**1**
18	F31L	1	**44**	F31I	1	**70**	G35R-G55D	1-2	**96**	Y389A-L404M	**11**
19	W51R-L338Q	2-9	**45**	A196T	6	**71**	Y360C-A383V	10	**97**	L20Q	**1**
20	A424T	12	**46**	L199P-N254D	6	**72**	G399D	11	**98**	G87S-T201I	**6**
21	R374C	10	**47**	I285V	8	**73**	S190G-M392T	6-11	**99**	G151C	**5**
22	I235V	7	**48**	N193D	6	**74**	152V	5	**100**	A136T	**4**
23	G61D-L78V	2-3	**49**	S303T	8	**75**	F31C	1	**101**	G115S	**3**
24	G87D	3	**50**	E67K-E308K	2-8	**76**	A50S	2			
25	C49Y-A297V	2-8	**51**	G215D	6	**77**	G127C	4			
26	L247M	1-7	**52**	F371I-C415Y	10-12	**78**	F391Y-I409N	12			

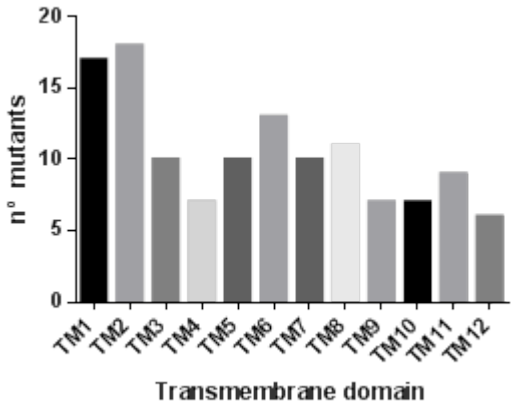


Figure 26. Number of mutations in each TMD of the 70 analyzed mutants.

For the final goal of this thesis, we analyzed in deep the expression and stability in detergent of only the first 70 mutants (Table 4). 41 of these mutants presented only one amino acid substitution and the other 29 have two amino acid substitutions in TMD. Figure 26 indicates the number of mutations present in each TMD of all 70 mutants selected for the screening. With the small exception of TMDs 1 and 2, the number of mutations in each TMD is fairly similar. Perhaps, this small bias in TMDs 1 and 2 indicates that these TM are more susceptible to be mutated. Figure 27 displays the 3D model of SteT indicating the positions of all the mutations found in the 70 selected mutants. As seen from the distribution of all the mutations, it is clear that the population of mutants chosen for expression and stability studies covers amino acid substitutions all over the protein and, therefore, it is a reasonable pool of randomly distributed SteT mutants.

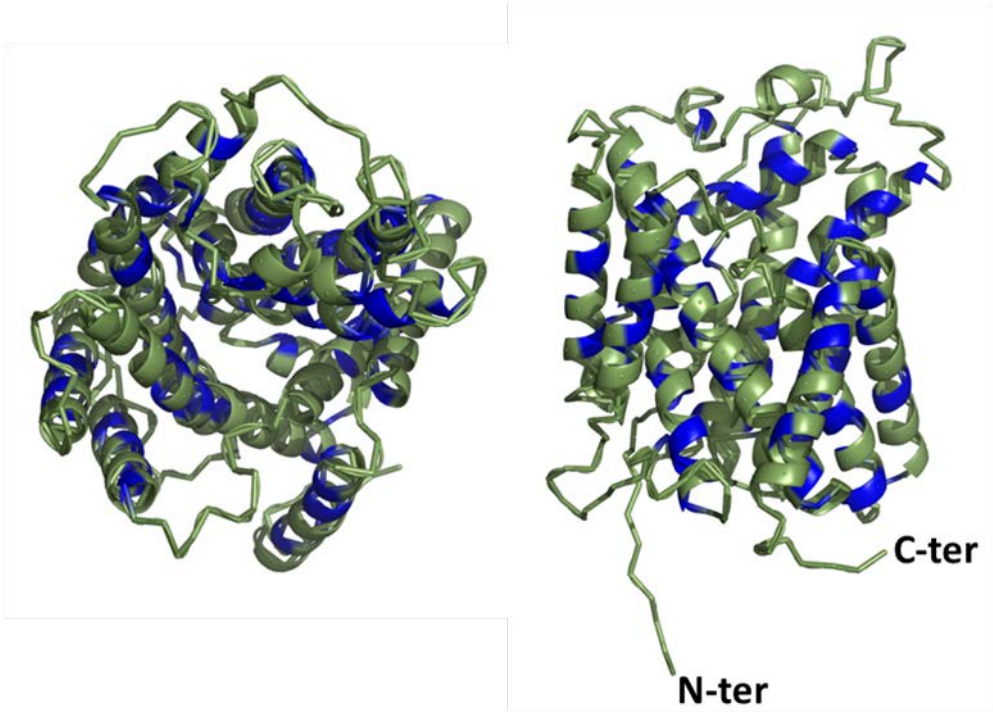


Figure 27. Localization of mutations in Transmembrane Domains of SteT. (left) SteT model from periplasmic perspective (right) SteT from side view of cytoplasmatic membrane

4.2 Evaluation of the expression and stability of selected SteT mutants

As mentioned in the introduction, SteT is very unstable in detergent solution and once purified, tends easily to aggregate as jugged by SEC experiments (Figure 14). Moreover is not stable more than two days at 1 mg/ml and cannot be concentrated more than 3 mg/ml. In order to compare whether the selected 70 mutants improved or not the expression and stability of SteT wild type, we combined the fluorescence of the GFP as reporter with SEC analysis.

4.2.1 Expression screening

Before comparing and evaluating the expression and stability of the 70 mutants selected, a preliminary assay was performed to find the optimal conditions to express SteT with the optimized codon for *E. coli* and the pTETGFP$_{11}$ vector. SteT expression was tested using different range of concentrations of ANTET (0,3 µg/ml, 0,5 µg/ml and 1 µg/ml) for 16 h at 30 °C. Protein expression yield was measured using whole cell fluorescence. We used 30 °C since it was the optimal temperature found previously to induce SteT expression (Reig et al., 2007) and (Bartoccioni et al, 2010). As seen in Figure 28A, no significant differences were found changing the inductor concentration at this temperature.

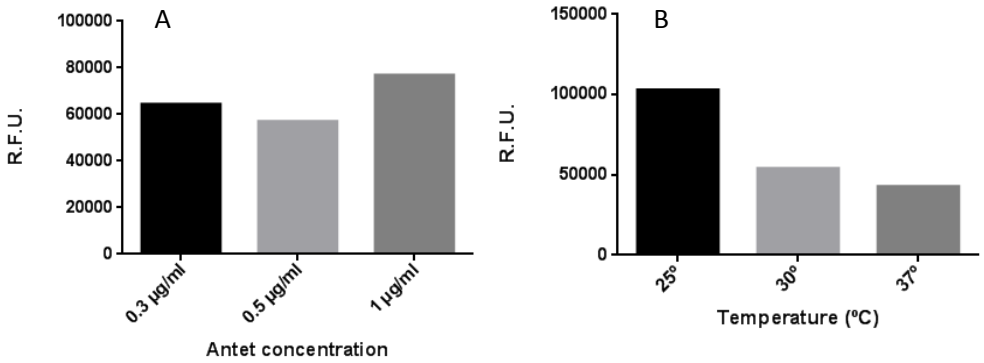


Figure 28. Optimizing expression conditions of SteT pTETGFP$_{11}$. (A) Expression yield of SteT-pTETGFP$_{11}$ in *E. coli* Bl21 (DE3) using different concentrations of ANTET (as indicated). The expression was carried out at 30 °C for 16 h. (B) Effect of temperature in the expression yield of SteT-pTETGFP$_{11}$ in *E. coli* Bl21 (DE3). The expression was carried out at different temperatures (as indicated) for 16 h using 0,3 µg/ml of ANTET. Expression was measured from whole-cell GFP fluorescence at OD$_{600}$ = 0,2. R.F.U. stands for relative fluorescence units.

In contrast, reasonable differences of SteT expression were found by changing the temperature of induction (Figure 28B). The expression yield at 25°C was over 2 fold larger than the expression at 30°C (Figure 28B). After finding the

optimal expression conditions (0,3 µg/ml of ANTET and 25°C), the expression yield of the selected SteT mutants was tested in small-volume cultures, using the whole-cell fluorescence from the GFP-fused tag.

Table 5 shows the expression yield of SteT mutants represented as percentage (%) of wild type expression. As seen in the table, 20 mutants (28 %) showed more expression than wild type. Notably, 7 of them (10%) improved wild type expression by a factor of 2; these mutants are: C168Y-L233M, G161N, C291S, N193D, S303T, L210Q-M229V and I134V-A377. Interestingly, mutant C291S was characterized previously in our laboratory and showed functional impairment (Bartoccioni et al, 2010). Since Cys 291 is one of the most conserved residues among LATs, it is reasonable to expect that any substitution will lead to inactivation. Remarkable, substitutions in SteT of conserved residues among APC members (se alignment in Figure 26) improved its expression yield (Table 5). These residues include Phe 31 (replaced by Leu or Ile), Leu 14 (replaced by Gln in the double mutant L14Q-T230A), Gly 27 (replaced by Ala in G27A-T156S), and Leu 338 (replaced by Gln in mutant W51R-L338Q) (Table 5). Since mutants, G27A-T156S, L14Q-230A and W51R-L338Q are double mutants; we cannot conclude that the mutations of these conserved regions are the only responsible for improving the expression as in the case of Phe 31.

Table 5. Expression yield of SteT random mutants. Expression levels of the SteT mutants expressed as % of wild type (WT). The TMD column indicates the SteT TMD number where the mutations are located.

Mutants	TMD	% of WT Expression	Mutants	TMD	% of WT Expression
G23R	1	74,34%	**G62C-F304S**	2-8	34,28%
A398P	11	110,70%	**P226Q**	7	29,85%
C168Y-L233M	5-7	230,09%	**A136E**	4	74,57%
G161N	5	204,39%	**V154E-G161S**	5-5	69,35%
L53P	2	137,81%	**G69D**	2	49,76%
A339D	11	167,52%	**L14Q-T230A**	1-7	143,21%
A267V	8	98,41%	**L247V**	7	73,39%
G284V	8	112,54%	**F391Y**	11	142,65%
A109P	3	110,60%	**F31I**	1	123,91%
C291S	8	209,31%	**A196T**	6	78,44%
G103S-L279P	3-8	74,73%	**L199P**	6	37,94%
A60E-C168R	2-5	98,09%	**I285V**	8	34,28%
F402S	11	97,74%	**N193D**	6	203,13%
A39F	1	163,59%	**S303T**	8	214,30%
M392V	11	29,93%	**E67K-E308K**	2-8	84,12%
A305T-T410S	8	23,80%	**G215D**	6	73,93%
I132F	4	63,77%	**F371I-C415Y**	10-12	63,96%
F31L	1	160,52%	**L210Q-M229V**	6-7	197,33%
W51R-L338Q	2-9	143,19%	**A369G**	9	76,43%
A424T	12	107,58%	**C141W**	4	81,92%
R374C	10	52,45%	**F49Y-N347Y**	2-9	95,24%
I235V	7	32,65%	**F139L**	4	87,58%
G61D-L78V	2-3	34,21%	**A197V-G232D**	6-7	46,50%
G87D	3	49,95%	**F16L-I99V**	1-3	50,24%
C49Y-A297V	2-8	26,65%	**L63R**	2	22,83%
L247M	1-7	21,90%	**M32V-M342L**	1-9	100,35%
G54S	2	93,65%	**A243P-A383S**	7-11	29,75%
G152D-V370M	1-9	46,57%	**I107F-H249Y**	3-7	93,26%
L199M-L417M	6-12	77,58%	**I336N-M413S**	9-12	97,74%
T159I-S298T	5-8	53,64%	**L52Q**	2	23,73%
F89S-A105T	3-3	32,83%	**G123D-I164T**	5	87,75%
I235F	7	33,45%	**F203S-R376P**	6-10	86,02%
P34Q	1	63,50%	**G27A-T156S**	1-5	133,04%
F49Y	2	50,13%	**I134V-A377T**	4-10	217,58%
R374H	10	135,42%	**G35R-G55D**	1-2	76,90%

4.2.2 Fluorescence Size Exclusion Chromatography

The stability of each mutant after solubilizing from the membrane with DDM was determined by fluorescence size exclusion chromatography (FSEC) (Kawate and Gouaux, 2006). SEC permits to evaluate accurately the degree of monosipersity of IMPs in different buffer conditions, by simply looking and comparing the elution profiles. Multiple and/or wide elution peaks (or polydispersity) indicates the protein's tendency to aggregate. Indeed, crystallographs use SEC as the mail tool to measure the quality of the protein right before initiate crystallographic screening. Also, the presence of GFP fused to the C-terminal end of each mutant enables to evaluate their stability in detergent solution by FESEC using only nanogram quantities of unpurified protein. DDM is one of the most popular detergents used for membrane solubilization and purification of IMPs because is a mild detergent and relatively cheap (Privé, 2007).

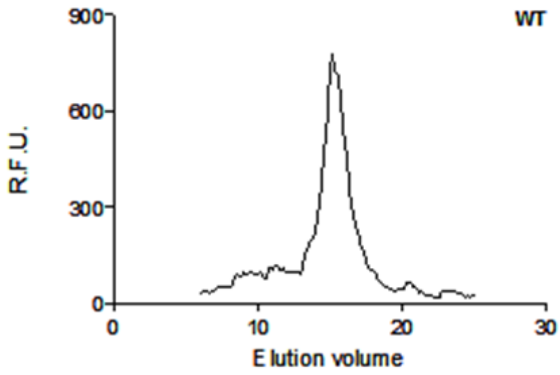


Figure 29. FSEC profile of SteT in 1 % of DDM. Solubilized *E. coli* membranes expressing SteT wild type were solubilized with 1 % DDM and injected into a Superose 6 10/30 column. Fractions of 200 µl were collected for fluorescence measurements and chromatogram construction. R.F.U. stands for relative fluorescence units.

.

For this assay, isolated cytoplasmic membranes of each of the 70 selected SteT random mutants (Table 5) were solubilized with DDM and injected into a size-exclusion chromatography column. SteT wild type eluted in an apparent molecular weight corresponding to the monomeric form of the protein plus the detergent-micelle (Figure 29). Subsequently, the 70 mutants were analyzed using the same experimental conditions (see Appendix I) and the resulting FSECs were compared to SteT wild type (Figures 30 and 31).

As seen in the examples of FSECs in Figure 30, those mutants that showed less expression than wild type (measured in whole cell, Table 5), displayed also a considerably less intensity in the FSECs profiles after solubilizing with 1 % DDM. In addition, the elution profile of these mutants was wider or similar to wild type. Notably, most of the mutants that expressed more than wild type (Table 5) presented also higher fluorescence intensities in the main elution peaks (Figure 32). This finding indicates that there is a correlation between the total amount of expressed protein measured by whole cell fluorescence using the split-GFP system (Table 5) and the total amount of solubilized protein with DDM from isolated membranes.

To evaluate and compare better the monodispersity of each mutant, we normalized all the FSECs and overlapped each one with the normalized SteT wild type FSEC (Figure 32, see material and methods for experimental details).

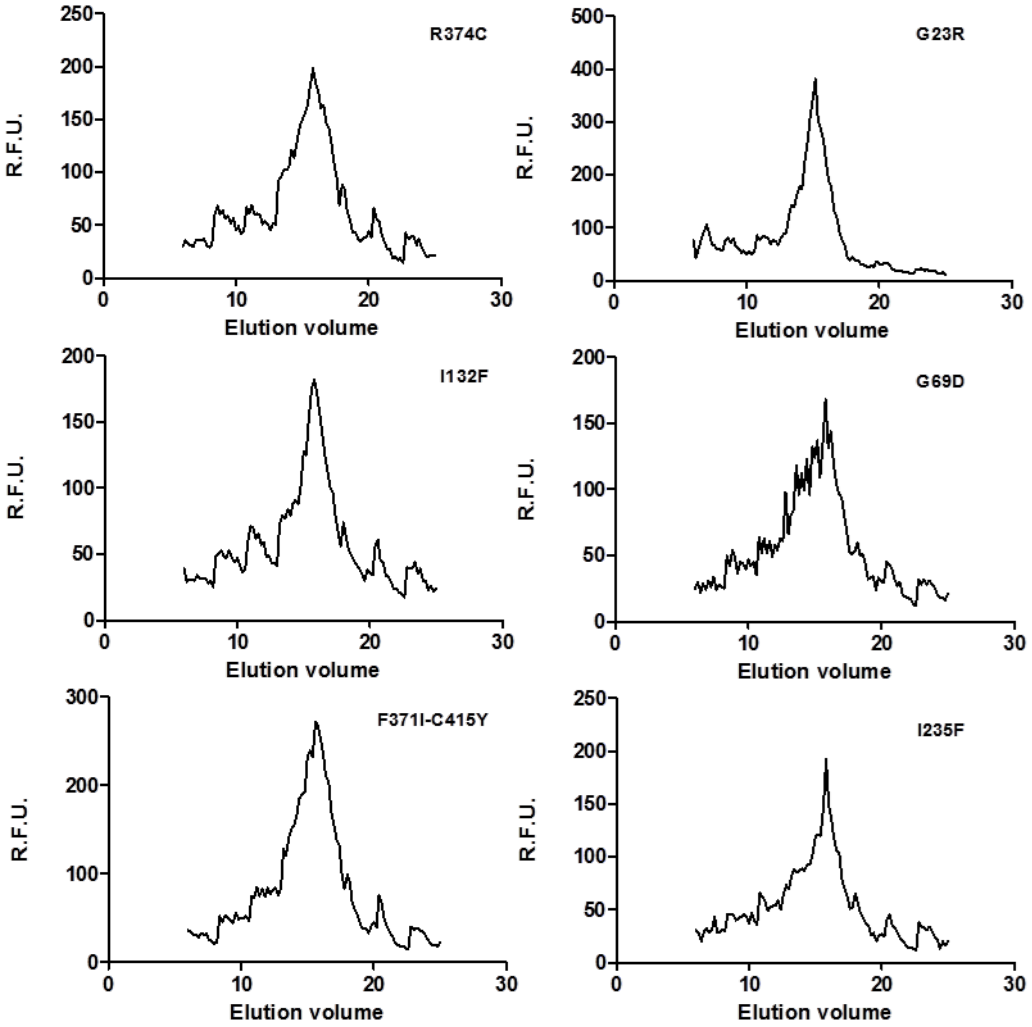


Figure 30. FSEC profiles of SteT random mutants expressing less than wild type. Solubilized *E. coli* membranes expressing each mutant (as indicated) were solubilized with 1 % DDM and injected into a Superose 6 10/30 column. Fractions of 200 µl were collected for fluorescence measurements and chromatogram construction. R.F.U. stands for relative fluorescence units.

As seen in the example of Figure 32 (right panels) a visual inspection of the overlapped chromatograms can, unambiguously, determine the mutants that are more monodisperse than wild type. In the example of Figure 32 (panel WT+A), the amino acid substitutions affect negatively the monodispersity (or stability) of DDM-solubilized SteT even though, the maximum intensity of the

elution peak was similar in both proteins (Figure 32, panel A). Conversely, other substitutions had a positive impact in the monodispersity of SteT (Figure 32, panel WT+B), in addition of improving the expression yield (Table 5).

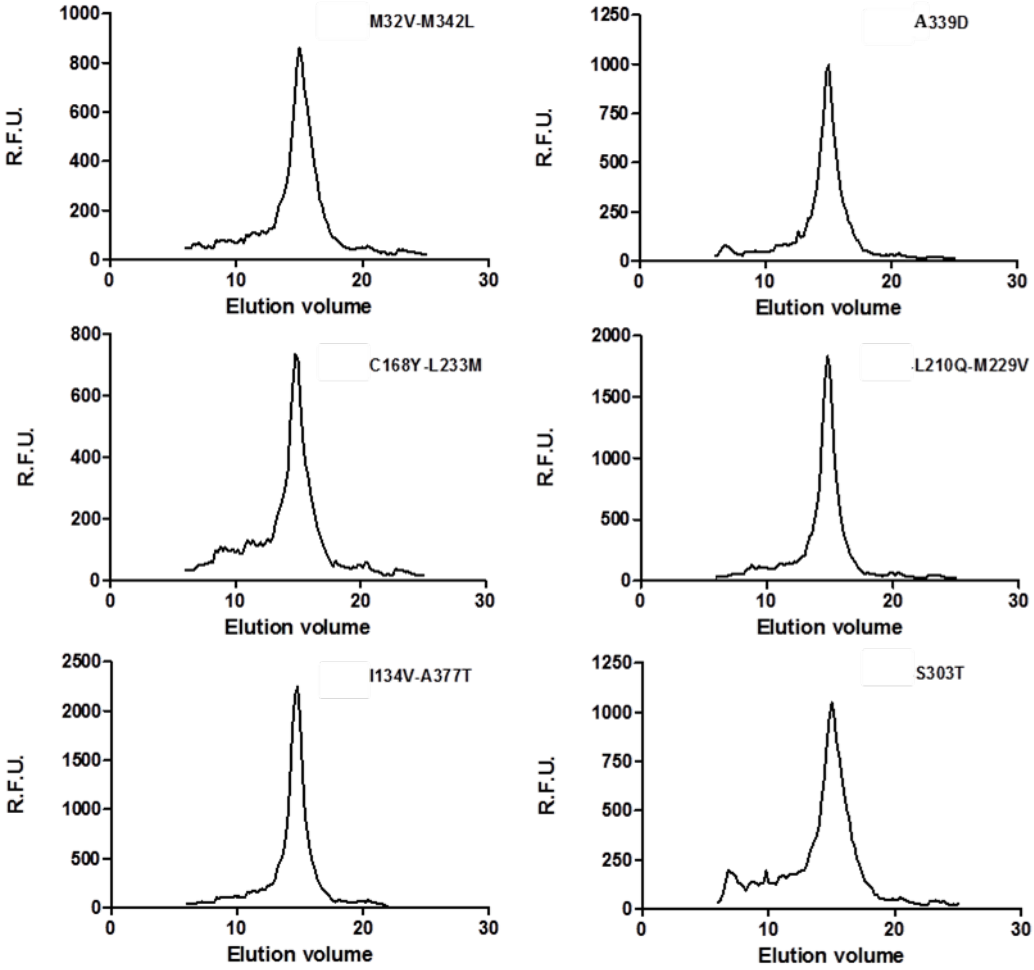


Figure 31. FSEC profiles of SteT random mutants expressing more than wild type Solubilized *E. coli* membranes expressing each SteT mutant (as indicated) were solubilized with 1 % DDM and injected into a Superose 6 10/30 column. Fractions of 200 µl were collected for fluorescence measurements and chromatogram construction. R.F.U. stands for relative fluorescence units.

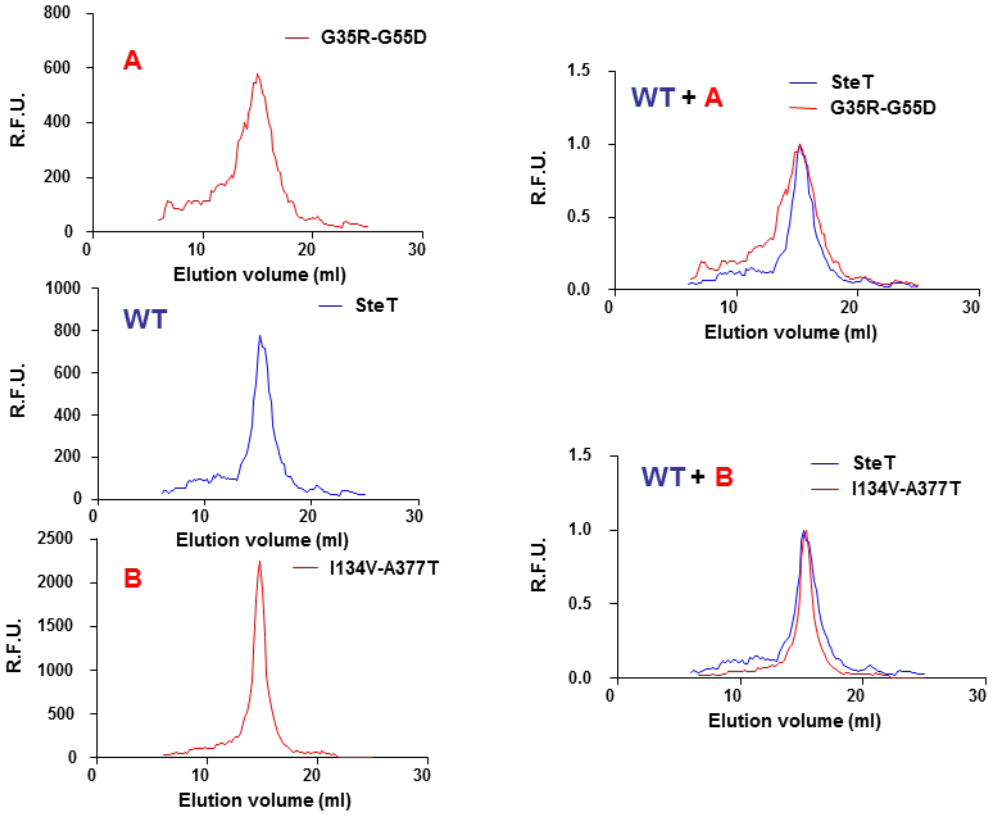


Figure 32. Normalization of FSEC profiles of mutants of SteT and its comparison with WT-SteT. (WT) FSEC profile of WT- SteT in 1% of DDM. (A) FSEC profile of mutant G53R-G55D-SteT in 1% of DDM. (B) FSEC profile of I134V-A377T-SteT in 1% of DDM. (WT+A) Normalization of mutant G35R-G55D-SteT and normalization of WT-SteT. (WT+B) Normalization of I134V-A377T-SteT and normalization of WT-SteT.

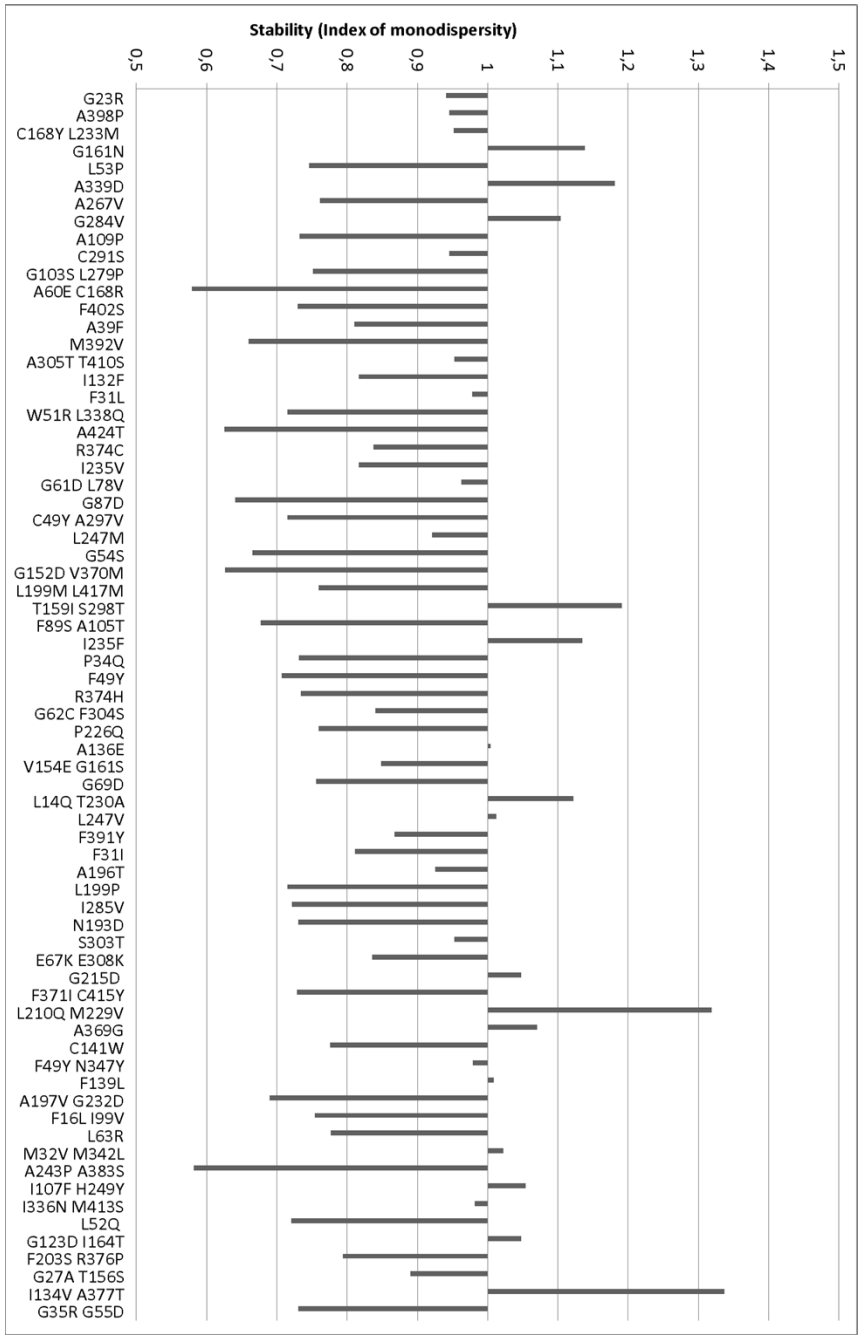


Figure 33. Calculated index of monodispersity of SteT random mutants selected for screening. This index is calculated by dividing the area of the normalized FSEC elution peak of wild type by the area of the normalized FSEC elution peak of each mutant (see materials and methods for full description).

At this point, we reasoned that it would be necessary to quantify in some extend the degree of monodispersity or stability with respect to wild type of each screened mutant in order to compare and evaluate better the impact of these mutations on SteT stability. Consequently, a numeric value of stability was assigned to each mutant (Figure 33). This numeric value, expressed as index of monodispersity, is calculated by dividing the area of the wild type normalized FSEC elution peak by the area of each mutant normalized FSEC elution peak (see materials and methods for full description). Therefore, values of this index of stability higher than 1 indicates better monodisperisity than wild type. On the contrary, values of the index of stability lower than 1, indicates worse monodispersity than wild type (Figure 33); thus, indicating more tendency to aggregate in this buffer conditions.

From the previous experiments, two outstanding mutants were selected: L210Q-M229V-SteT and I134V-A377T-SteT, both expressing about two-fold more than wild type (Table 5), and both having the largest index of stability (> 1,3) (Figure 33). Therefore, the stability of these two mutants was challenged in other detergents to further characterize their suitability as crystallization targets.

4.2.3 Analizing the I134V-A377T-SteT and L210Q-M229V-SteT in different detergents

Five different detergents (including DDM as control) were used to test both solubility and stability of the two selected SteT mutants from the previous screening: I134V-A377T and L210Q-M229V. These detergents were: n-decyl-β-D-maltopyranoside (DM), n-octyl-β-D-glucopyranoside (OG), 6-cyclohexyl-1-hexyl-β-D-maltoside (Cymal-6) and n-dodecyl-N,N-dimethylamine-N-Oxide (LDAO). The

solubilization efficiency of each detergent was calculated by measuring the GFP fluorescence before and after solubilizing the membranes expressing a given mutant and it was expressed as % (Table 6). Solubilization was carried out for 1 h at 4°C using 1 % of each detergent. As indicated in Table 6, the solubilization efficiency of all the detergents was remarkable high (≥ 80 %) in both mutants and wild type, but particularly high for mutant I134V-A377T in all the analyzed detergents.

Table 6. Detergents solubilization efficiency of SteT wild type and mutants. Solubilization efficiency was calculated by measuring the GFP fluorescence fused to each SteT version before and after solubilizing membranes expressing a given mutant with 1 % of the indicated detergent. Values were expressed as % of the initial GFP fluorescence before solubilization.

	Solubilization efficiency (%)				
	DDM	**DM**	**OG**	**Cymal-6**	**LDAO**
SteT-wt	89	84	82	80	81
L210Q-M229V-SteT	80	90	83	84	86
I134V-A377T-SteT	97	94	85	92	82

FSEC studies (Kawate and Gouaux, 2006) were then used to monitor the monodispersity of SteT wild type, I134V-A377T and L210Q-M229V in all the different detergents tested (Figures 34, 35 and 36).

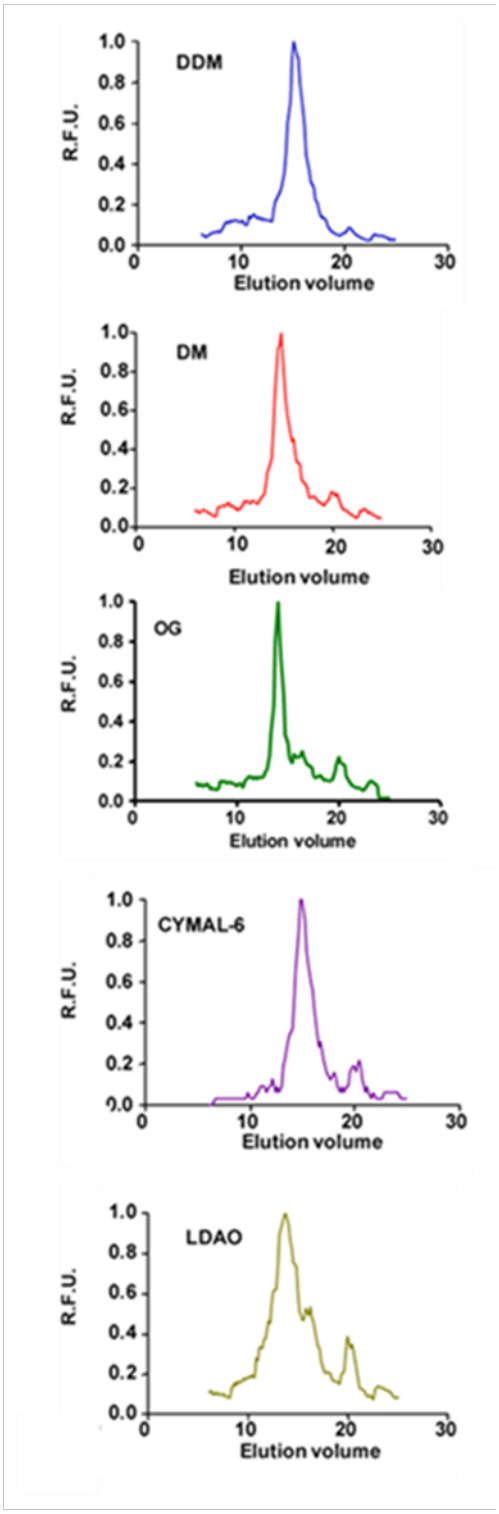


Figure 34. Normalized FSEC profile of SteT wild type solubilized in different detergents. Detergent-solubilized membranes expressing the protein were injected in a Superose 6 10/30 column and 200 µl fractions were collected for fluorescence detection and chromatogram construction. R.F.U. stands for relative fluorescence units.

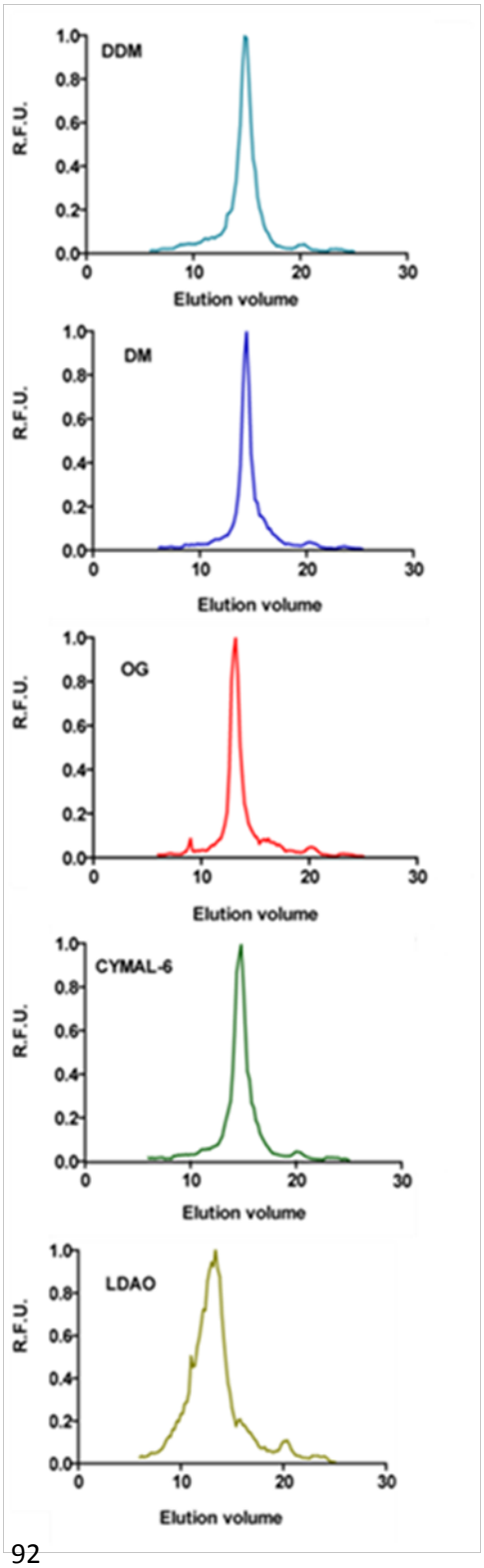


Figure 35. Normalized FSEC profile of L210Q-M229V-SteT solubilized in different detergents. Detergent-solubilized membranes expressing the protein were injected in a Superose 6 10/30 column and 200 µl fractions were collected for fluorescence detection and chromatogram construction. R.F.U. stands for relative fluorescence units.

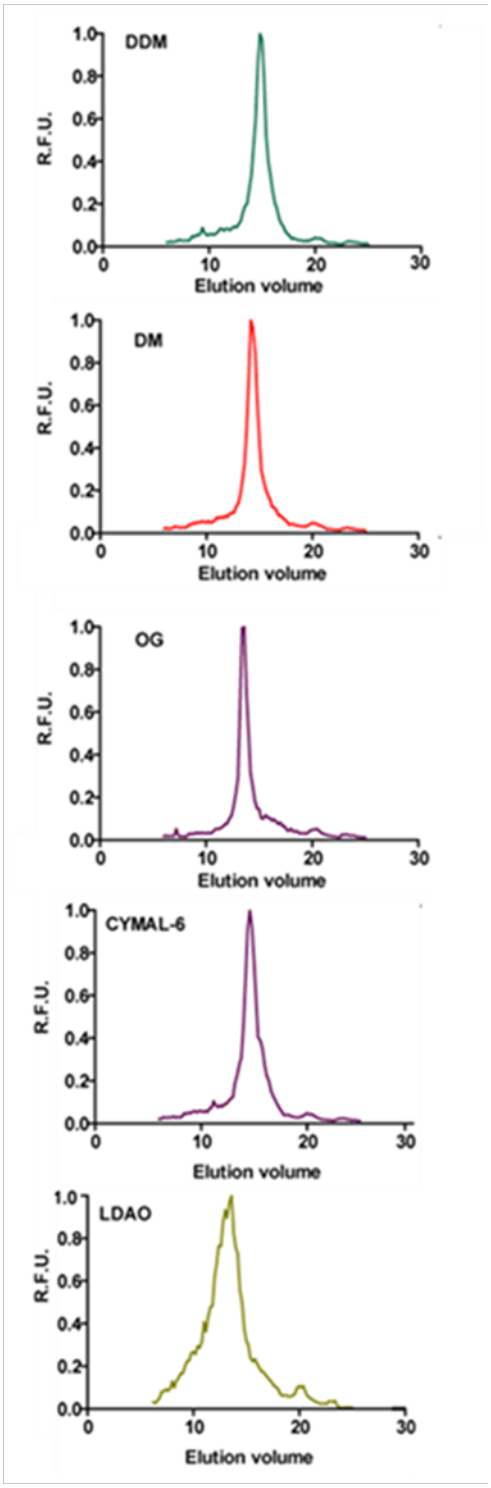


Figure 36. Normalized FSEC profile of I134V-A377T-SteT solubilized in different detergents. Detergent-solubilized membranes expressing the protein were injected in a Superose 6 10/30 column and 200 µl fractions were collected for fluorescence detection and chromatogram construction. R.F.U. stands for relative fluorescence units.

Notably, with the exception of LDAO, I134V-A377T and L210Q-M229V mutations substantially improve the stability of SteT in the different detergents tested as judged by the FSECs (Figures 34, 35 and 36). So we concluded that I134V-A377T-SteT and L210Q-M229V-SteT meet the criteria for large-scale purification, functional assays and crystallization trials.

4.3 Analysis of purified SteT mutant candidates

Purification of SteT requires larger volumes of *E. coli* cell growth in addition to a more robust expression vector. Therefore, we decided to clone I134V-A377T-SteT and L210Q-M229V-SteT into a modified pTTQ18 expression vector (Stark, 1987). This expression vector controlled by the TAC promoter (IPTG inducible) was modified in our laboratory (Errasti-Murugarren E. and Palacín M., unpublished results), and contains a superfolder GFP fused to the C-terminal end of the cloning site (Figure 37). The GFP is followed by a 10xHis tag for immobilized affinity chromatography (IMAC) purification (Figure 37). It also contains a PreScission protease site (3C in Figure 37) in order proteolitically remove both, the GFP and the 10xHis tags after IMAC purification.

Figure 37. pTTQ18 modified vector. P_{TAC}: TAC promoter, 3C: HRV 3C Presscision protease site, HIS(x10): tag of 10 histidine residues

As in the previous sections, we used whole-cell GFP fluorescence to test the best growing temperature and induction conditions to get the higher protein expression yield in this new construct (Drew et al, 2006). SteT wild type expression at, 25°C, 30°C or 37°C was compared with I134V-A377T-SteT and L210Q-M229V-SteT expression at the same temperatures (Table 7). As observed, 37°C and 0.1 mM of IPTG were the optimal conditions for SteT expression using this plasmid. Using these conditions, an expression yield of 1.4 mg/Liter culture for SteT-WT, 2.05 mg/Liter culture for L210Q-M229V-SteT and 5.16 mg/Liter of culture for I134V-A377T-SteT was obtained (table 7). It was previously reported that SteT wild type has an expression yield of 1 mg/Liter of culture at 30°C using the same expression vector (Reig N., 2007); therefore these new SteT constructs, clearly improve the amount of protein expressed.

Table 7. Screening of the expression of SteT variants in the modified pTTQ18 expression vector. Levels of expression are expressed as mg of protein per liter of culture (mg/L) calculated from the GFP fluorescence fused to the C-terminal end of SteT variants. In all conditions protein was induced with 0.1mM of IPTG

Temperature	25°C	30°C	37°C
WT	0,56 mg/L	0,6 mg/L	1,44 mg/L
L210Q-M229V	0,90 mg/L	0,87 mg/L	2,05 mg/L
I134V-A377T	3,5 mg/L	3,87 mg/L	5,16 mg/L

Further, for I134V-A377T-SteT, we studied protein expression versus different *E. coli* strains. Three *E. coli* strains were tested (Table 8): BL21(DE3), BL21(DE3)-STAR and BL21(DE3)C43. No appreciable expression differences between strains were found despite using different concentrations of IPTG (Table 8). Nevertheless, we selected BL21(DE3) as the most optimal expression host and 37°C and 1 mM IPTG as the best conditions for expressing SteT-wild type, L210Q-M229V-SteT and I134V-A37-SteT.

Table 8. The Screening of expression I134V-A377T-SteT in different *E. coli* strains. Levels of expression are expressed as mg of protein per liter of culture (mg/L), calculated from the GFP fluorescence fused to the C-terminal end of SteT variants. Protein was induced at different IPTG concentrations as indicated.

	Temperature			
	30°C		37°C	
[IPTG] (mM)	0,1	0,4	0,1	0,4
BL21(DE3)	1,1 mg/L	1,1 mg/L	4,2 mg/L	3,8 mg/L
STAR	0,7 mg/L	1,0 mg/L	3,7 mg/L	4,0 mg/L
C43	0,7 mg/L	0,8 mg/L	3,8 mg/L	3,3 mg/L

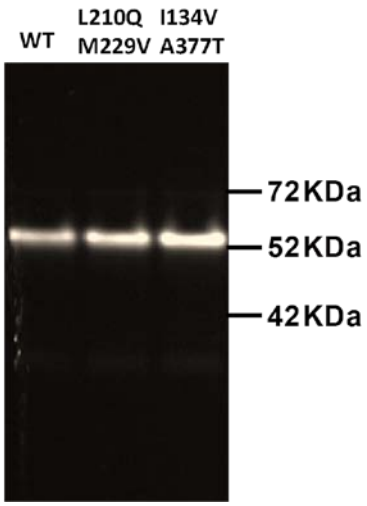


Figure 38. In gel fluorescence of SteT variants expressed in the pTTQ18-GFP vector. Isolated membranes of *E. coli* BL21(DE3) cells expressing each SteT variant were solubilized with 1 % DDM and subjected to SDS-PAGE. Images were taken using a gel transilluminator.

Moreover, the expressed proteins in this new expression vector are localized in the plasma membrane, as observed from the in-gel fluorescence experiments of isolated cytoplasmic membranes (Figure 38). We can observe that the amount of protein calculated from whole-cell fluorescence correlates very well with the amount of protein observed in isolated membranes. It is also worth of mentioning that although pTET plasmid produces less protein than the current vector; the expression ratios between all the SteT versions in both expression vectors were comparable. Finally, expression of soluble GFP was not observed.

4.3.1 Purification of I134V-A377T-SteT, L210Q-M229V-SteT and SteT wild type

Protein purification is mandatory prior any crystallization work. High purity and stability in detergent are necessary factors for any crystallization process. We therefore scaled up *E. coli* cultures expressing SteT wild type, L210Q-M229V-SteT or I134V-A377T-SteT and purified these proteins for subsequent evaluation of their stability after the purification protocol. Protein purification of wild type and the two SteT mutants, proteolytic GFP-His10 tag removal and

protein concentration were performed as described in *Materials and Methods* (section 7.6). All the purification steps were carefully optimized trying to minimize as much as we could the time length of each one in order to preserve the integrity of the protein. As discussed earlier, DDM is the detergent of choice for solubilizing IMPs and, therefore, the most convenient detergent for these preliminary experiments.

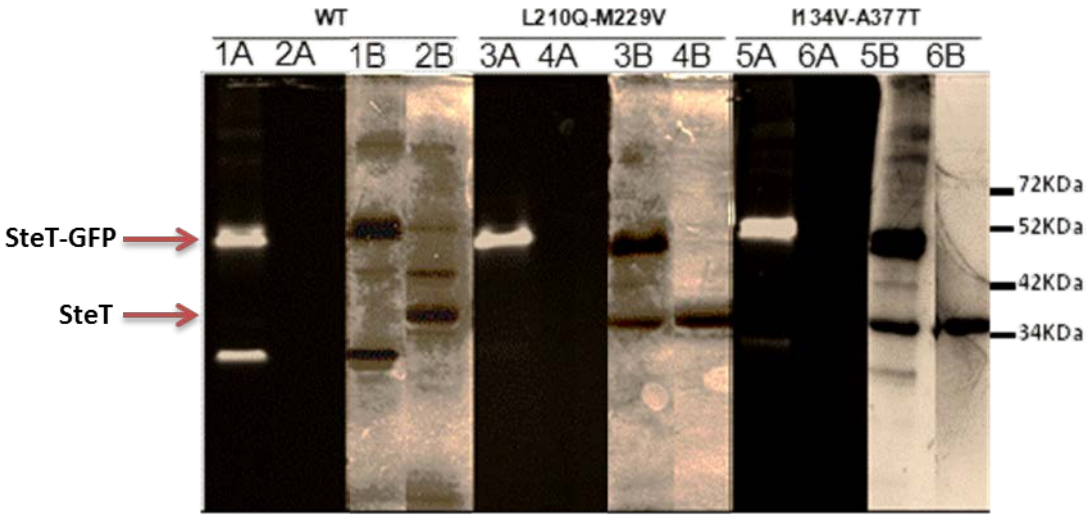


Figure 39. SDS-PAGE gels of purified SteT wild type (WT), L210Q-M229V-SteT and I134V-A377T-SteT (A): In gel fluorescence. (B): Coomassie–staining. (1,3,5): Elution fraction from IMAC purification of SteT wild type, L210Q-M229V-SteT and I134V-A377T-SteT, respectively. (2,4,6): Elution fraction from IMAC reverse purification of SteT wild type, L210Q.M229V-SteT and I134V-A377T-SteT, respectively. Arrows indicate the molecular weight in a SDS-PAGE gel of SteT-GFP and SteT purified, after the digestion with 3C.

Figure 39 compares the presence and purity of the three SteT variants (wild type, L210Q-M229V and I134V-A377T) during the first steps of purification: IMAC, PreScission digestion and IMAC reverse. Elution from the IMAC column revealed the presence of soluble GFP during the purification of wild type and I134V-A377T-SteT, a common characteristic of GFP-fusion proteins heterologous expressed. IMAC purification gave a fairly pure protein for both SteT mutants (Figure 39, lanes 4B and 6B). Due to the lower expression yield

and its poor stability in detergent, SteT wild type IMAC purification gave a semi-pure protein sample despite following exactly the same protocol used for the two mutants. As expected, PreScission digestion in all SteT variants eliminates both the soluble GFP and the 10xHis tag as observed in the coomassie-blue staining SDS-PAGE gels after IMAC reverse purification (Figure 39, lanes 2B, 4B and 6B) and in the in-gel fluorescence SDS-PAGE gels of the same samples (Figure 39, lanes 2A, 4A and 6A).

After IMAC and IMAC reverse purifications, proteins were concentrated to 2 mg/ml. Samples were then subjected to ultracentrifugation in order to remove any aggregated form and subsequently injected into a Superdex 200 5/150 GL SEC column equilibrated with SEC buffer containing 2xCMC of DDM. As before, SEC is the technique of choice to evaluate the monodispersity of each SteT version in similar conditions (buffer and protein concentration) used for crystallization.

All proteins eluted from the SEC column in the expected elution volume, corresponding to the monomeric size of the proteins plus the size of the DDM detergent micelle. SEC profiles of L210Q-M229V-SteT and I134V-A377T-SteT showed a clear monodisperse behavior with a single and symmetric SEC elution peak (Figure 40). Conversely, SteT wild type presented different high-molecular-weight aggregates as judged by the shoulder that appears in the SEC elution profile (Figure 40). The smaller intensity of the elution peak of SteT wild type compared with the two mutants (Figure 40), indicates that part of the SteT wild type protein was precipitated after ultracentrifugation and before injecting into the SEC column.

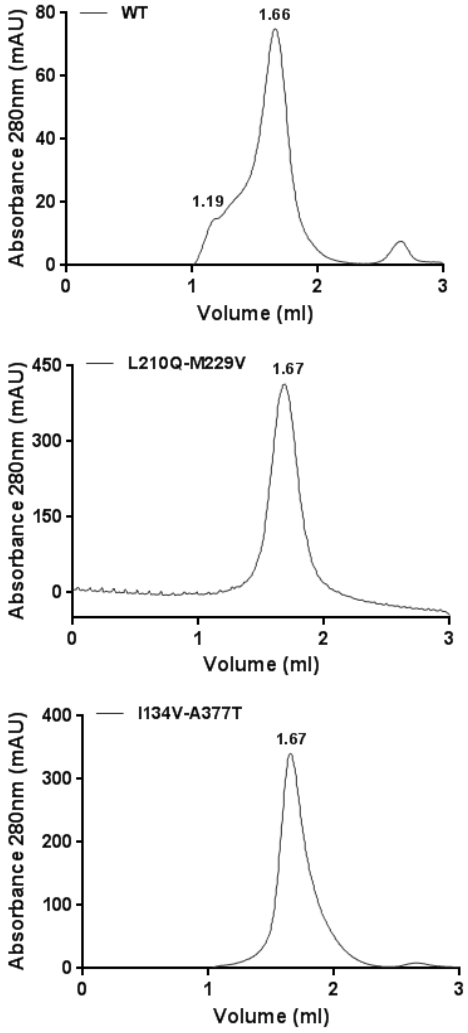


Figure 40. Size Exclusion Chromatography (SEC) profiles of SteT wild type (WT), L210Q-M229V-SteT and I134V-A377T-SteT. Purified samples from IMAC were concentrated to 2 mg/ml. 100µl of each sample was injected in Superdex 200 50/1 50 G column equilibrated with SEC buffer and 2xCMC DDM.

To better compare the stability (or monodispersity) of the three SteT variants in DDM, normalized SEC profiles of SteT wild type, I134V-A377T-SteT and L210Q-M229V-SteT were overlapped (Figure 41). The shape and symmetry of the mutants compared to wild type clearly points out the effect of these mutations in improving SteT monodispersity (or stability in detergent solution).

Furthermore, the main elution peaks of each SteT version; that is, the monodisperse fraction of each SEC, were collected and stored at ~ 1 mg/ml, 4°C for 1 week. This experiment was impossible to perform with wild type due to the low concentration of the remained protein before the SEC experiment as a consequence of aggregation. After 1 week at 4°C, the mutants samples were ultracentrifuged and protein concentration in the supernedat was measured and compared with the initial concentration before the incubation. Remarkable, almost 100% of the protein remained in solution after the incubation indicating, one more time, the gain in stability that mutants L210Q-M229V and I134V-A377T confer to SteT in this buffer condition that contains 2xCMC of DDM.

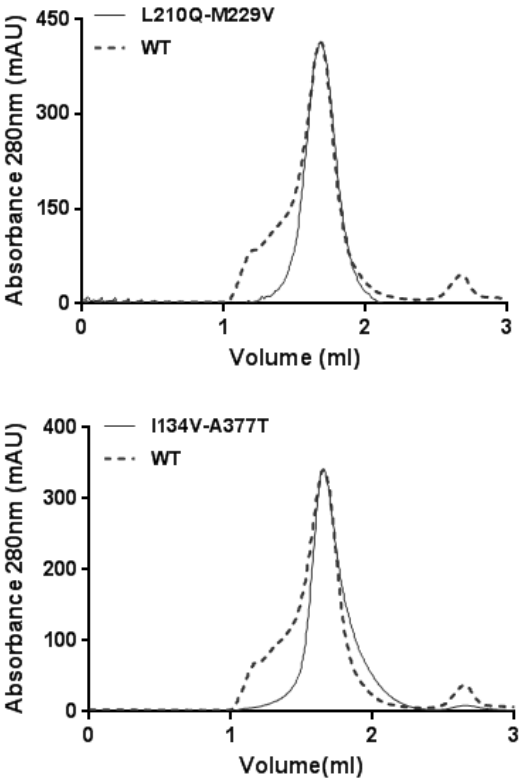


Figure 41. Overlapped SEC elution profiles of SteT wild type with either, L210Q-M229V-SteT and I134V-A377T-SteT. Normalized SEC chromatograms from Figure 40 were compared. The discontinuous lines correspond to wild type and the continues lines correspond to the indicated mutant.

4.3.2 Large scale purification of I134V-A377T-SteT

The previous experiments of expression and stability after purification in DDM indicate that I134V-A377T-SteT is probably de best candidate to initiate crystallographic studies. In this line, more stability tests were needed in similar experimental conditions as the ones required to initiate crystallization trials.

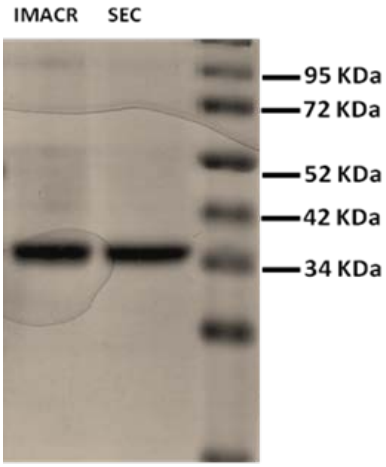


Figure 42. SDS-PAGE gel of purified I134V-A377T-SteT. Purity of I134V-A377T-SteT after the IMAC reverse purification step (IMACR lane) and after SEC purification (SEC line). SEC was performed in a Superdex 200 10/300 GL column equilibrated with SEC buffer containing 2xCMC of DDM and 10% glycerol.

Purification of I134V-A377T-SteT from large-scale volume cell culture was carried out. After the IMAC reverse (Figure 42), protein was concentrated to 2, 4 and 12 mg/ml. The three samples were ultracentrifuged and the supernatant was injected in a Superdex 200 10/300 GL SEC column equilibrated with SEC buffer containing 2xCMC of DDM and 10% of glycerol (Reig et al., 2007).

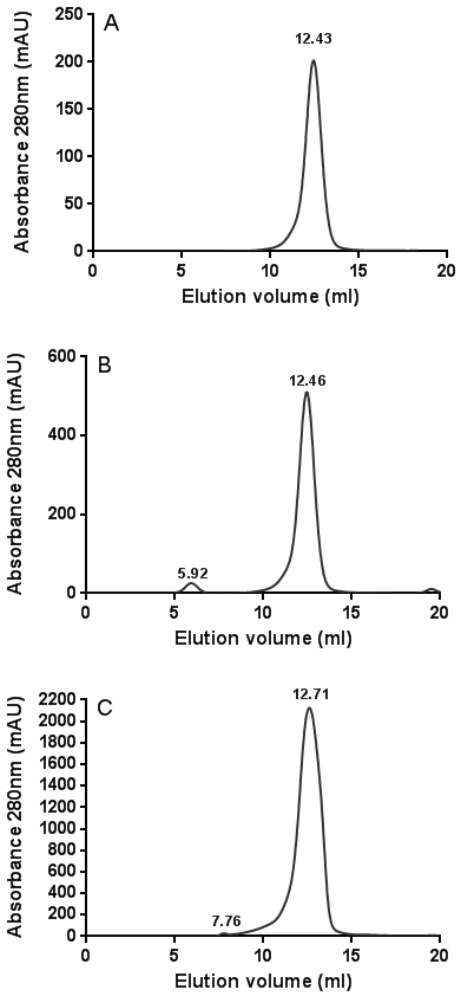


Figure 43. Size Exclusion Chromatography (SEC) profiles of I134V-A377T-SteT. Purified protein from IMAC was concentrated and injected into a Superdex 200 10/300 GL SEC column equilibrated with SEC buffer containing 2xCMC of DDM and 10% of glycerol at concentrations of 2 mg/ml (A), 4 mg/ml (B) and 12 mg/ml (C).

In all these three concentrations tested, I134V-A377T-SteT eluted as monomer (Figure 43). In addition, all SEC profiles of I134V-A377T-SteT display single and symmetric elution peaks indicating the absence of aggregation states (shoulders in the high-molecular-weight elution areas of the chromatogram), clearly indicating the great stability of the protein in these conditions even at a concentration of 12 mg/ml (Figure 43C). The main elution peaks of each SEC experiment of Figure 43 (or monodisperse protein fraction) was collected and joined. Thereafter, the joined fractions were concentrated to 2 mg/ml, 4 mg/ml

and 6 mg/ml and stored at 4°C during 1 week to evaluate more precisely the rate of aggregation in these experimental conditions. In these three concentrations tested, the purified protein was very stable within this period of time (1 week), since no variation of protein concentration was observed after ultracentrifugation.

SEC purified I134V-A377T-SteT was analyzed by electrospray mass spectrometry in order to ensure the correct molecular weight of the protein after digestion of the GFP tag by the PreScission protease (Figure 44). Mass spectrometry analysis of this sample revealed a well-resolved mass spectrum that, after deconvoluting, gave a main molecular weight of 48185 Da that accurately corresponds to I134V-A377T-SteT. The second main peak observed at 27092 Da (Figure 44) is exactly half of the I134V-A377T-SteT molecular mass (48185 Da) and it is a common outcome (and artifact) of the deconvoluting software. Overall, the mass spectrometry analysis of SEC purified I134V-A377T-SteT confirms (1) the correct digestion of the GFP fused SteT constructs by the Prescission protease in a single site and (2) a single I134V-A377T-SteT product of translation from the pTTQ18-GFP expression vector since no extra Met or N-formyl modifications are observed. The later are common consequences of the overexpression of membrane proteins in *E. coli* and can negatively affect protein crystallization due to the heterogeneity of the sample.

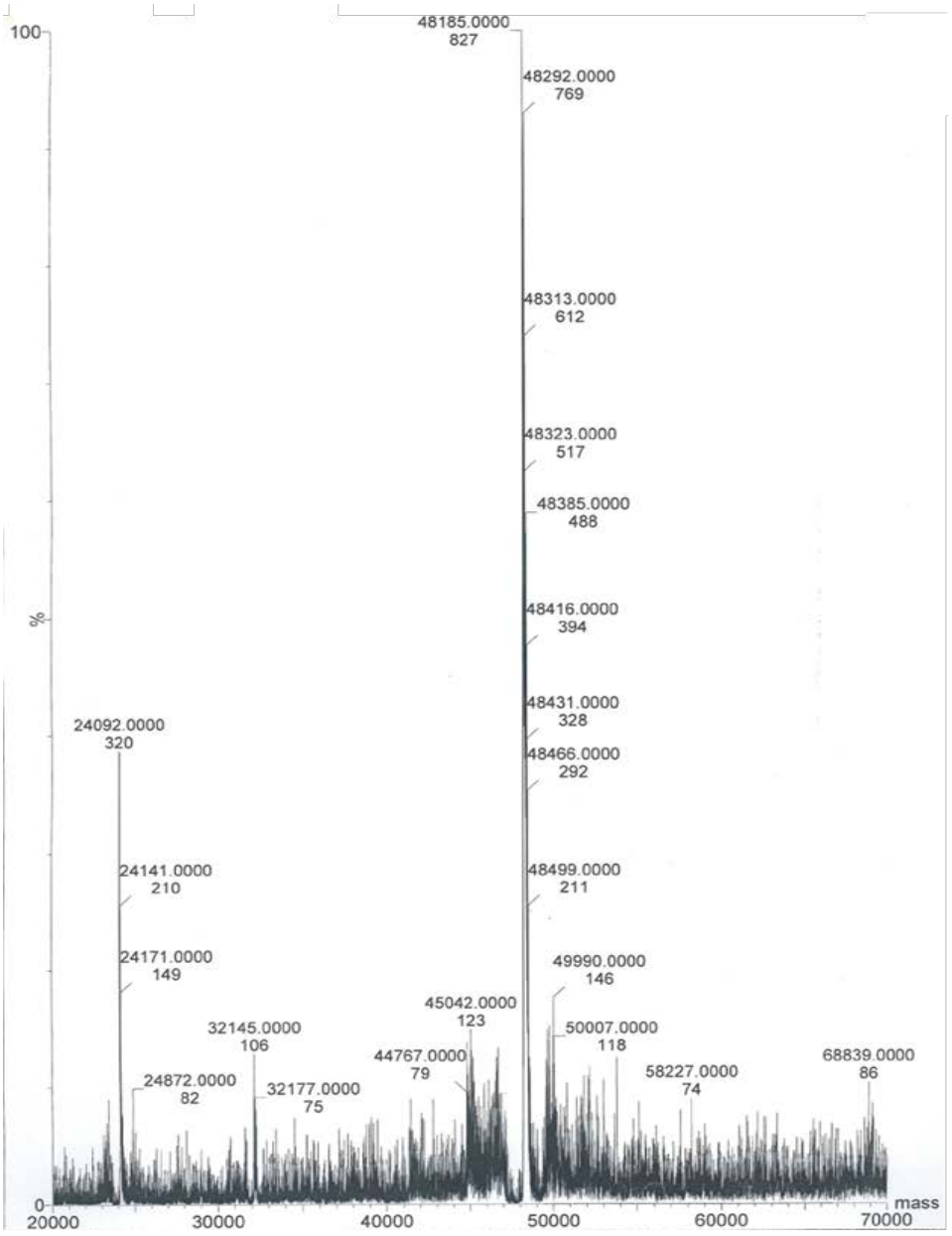


Figure 44. Electrospray mass spectrometry resolved spectrum of I134V-A377T-SteT. 50 μg of SEC purified protein was precipitated as described in Material and Methods. Sample was dissolved in 90 % formic acid and quickly injected in a HPLC SEC column in line with the mass spectrometer.

It is well accepted that by reducing the size of protein-detergent micelle complex, the probability of forming well-ordered crystals increases. In order to reduce the size of the micelle of the soluble protein-detergent complex we proceeded to solubilize and analyze the stability of I134V-A377T-SteT in DM. The purification procedure was similar as the one used before for DDM, but with some modifications. First, membranes expressing I134V-A377T-SteT were solubilized with 1 % of DM and no glycerol was added during all the purification process. It is known that glycerol increases the stability of purified proteins in detergent, however; if possible, it should be avoided in crystallization trials. Purified protein in this new detergent was concentrated up to 5 mg/ml and injected in a Superdex 200 5/150 GL SEC column equilibrated with SEC buffer and 2xCMC of DM. Similar to the DDM experiment, the SEC profile of I134V-A377T-SteT also showed a monodisperse behavior, as judged by the unique and symmetrical shape of the elution peak of the chromatogram (Figure 45A). The purity of the protein was also confirmed in a coomassie blue-stained SDS-PAGE gel (Figure 45B). The collected protein fraction from the eluted peak was concentrated to 1 mg/ml and dialyzed against SEC buffer containing 2xCMC of DM at 4°C during 2 days to evaluate protein stability. This dialysis step is sometimes necessary since protein concentrators are known to raise the concentration of the detergent in addition to the protein. This increase in detergent concentration can somehow improve protein stability; however, it affects negatively membrane protein crystallization. Notably, no significant variation (~8 %) of the initial protein concentration was observed after the dialyzed sample was ultracentrifuged, thus confirming that I134V-A377T –SteT was fairly stable in just 2xCMC of DM.

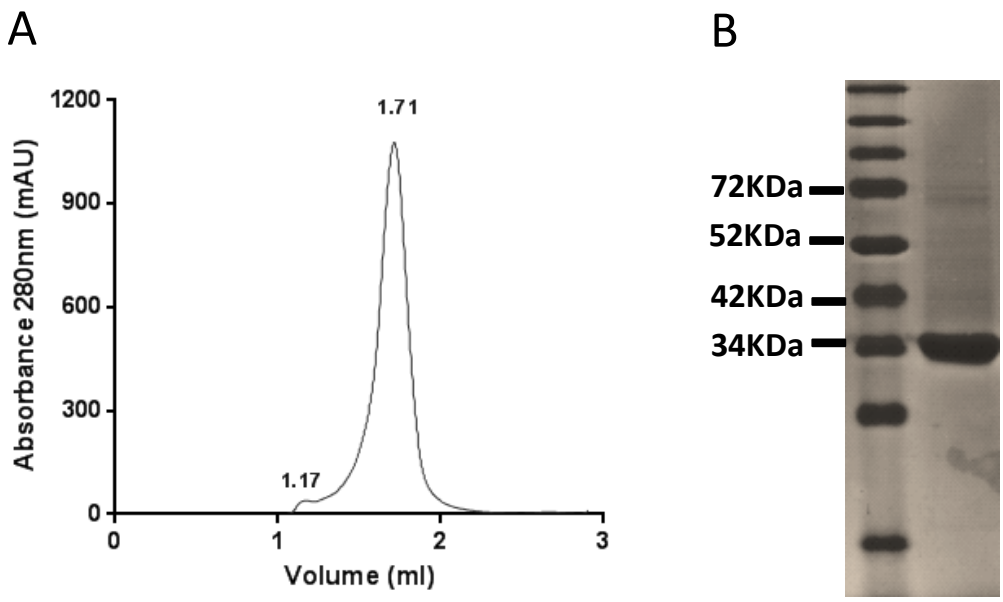


Figure 45. Solubilization, purification and stability of I134V-A377T-SteT in DM: (A) SEC profile of I134V-A377T-SteT. IMAC purified protein was concentrated at 5 mg/ml and injected in a Superdex 200 5/150 GL SEC column equilibrated with SEC buffer with 2xCMC of DM (B) SDS–PAGE gel of eluted I134V-A377T-SeT from (A). As in the previous examples, SteT (and its mutants) has an apparent molecular weight in SDS-PAGE gels of ~ 35 KDa.

4.3.3 Detergent screening of I134V-A377T-SteT

As commented in the previous section, success on getting crystals of a membrane protein lies on the possibility of getting stable protein in different detergents. Therefore, our next experiment with I134V-A377T-SteT was to evaluate its stability and monodispersity in three other detergents whose hydrophobic tail and polar head group are sequentially reduced allowing, therefore, more protein-protein interactions needed for crystallization.

In these experiments, we first solubilized membranes expressing I134V-A377T-SteT in DM. Protein was purified as usual and, after IMAC reverse elution, the

purified protein was divided in four tubes and concentrated in parallel up to ~ 4 mg/ml. During this concentration process the protein was subjected to buffer exchange containing 2xCMC either, DM, NG, OG or Cymal-6. Possible protein aggregates were eliminated by ultracentrifugation and the supernatant was injected into a Superdex 200 5/150 SEC column equilibrated with SEC buffer and 2xCMC of each detergent. DM was chosen as the solubilizing detergent because it has a higher CMC compared to DDM. In this way, detergent exchange is more effective during concentration.

SEC profiles of I134V-A377T-SteT in the selected detergents showed a monodisperse behavior in all cases (Figure 46). The elution volume (equivalent to the retention time) was almost identical in all cases. It is worth mentioning that the reduced intensity of the experiment with NG (Figure 46) was likely due to a protein lost during the ultracentrifugation process, prior to the injection to the SEC column, as a consequence of protein aggregation during detergent exchange. The SDS-PAGE gel of the elution peaks of each detergent, confirmed the identity of I134V-A377T-SteT (Figure 47).

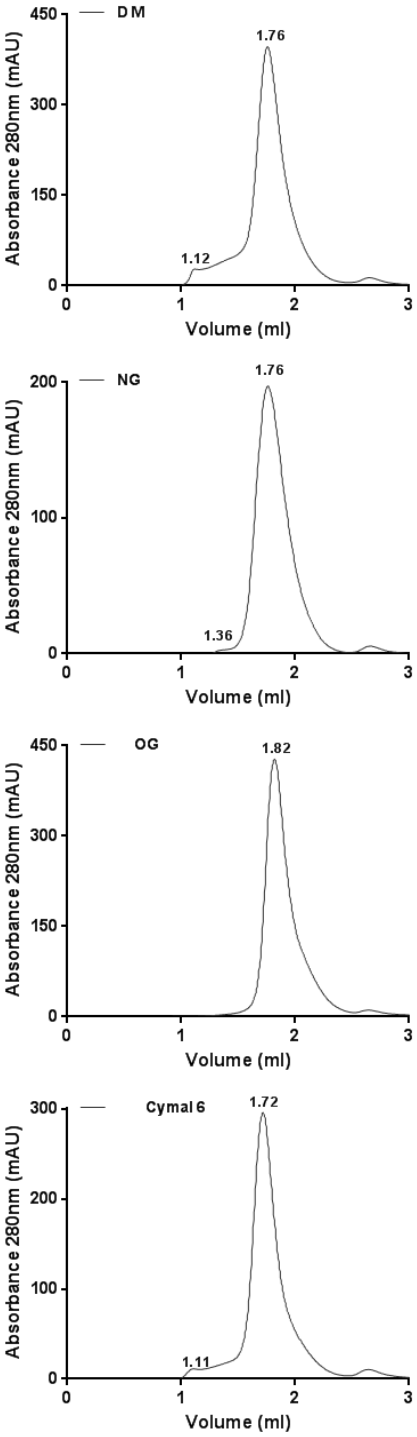


Figure 46. Stability of I134V-A377T-SteT in different detergents. IMAC purified protein in DM was concentrated while exchanging the detergent to, DM, NG, OG or Cymal-6. Samples exchanged to the indicated detergent were injected at ~4 mg/ml to a Superdex 200 5/150 SEC column equilibrated with SEC buffer containing 2xCMC of the given detergent.

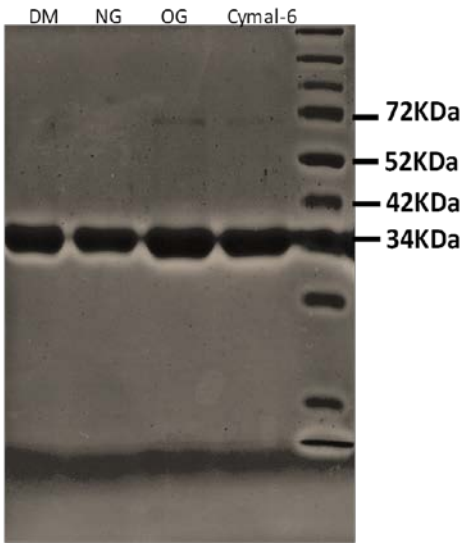


Figure 47. SDS-PAGE gel of the SEC I134V-A377T-SteT fraction from the detergent screening experiment. The gel was stained with coomassie blue. Each line corresponds to the experiment done with the indicated detergent (top).

Eluted I134V-A377T-SteT from the four SEC experiments (Figure 47) were collected, concentrated at ~1 mg/ml and stored up to three days at 4°C. In all four detergents the purified protein was stable after one day at 4°C with no variation of the protein concentration due to aggregation and subsequent precipitation. After 3 days at 4°C samples, the totality of I134V-A377T-SteT remained in solution in DM and Cymal-6 (Table 9). Some protein lost was observed in the detergents NG and OG due to protein aggregation, although most of the protein remained in solution (86 % and 73 % for NG and OG, respectively) (Table 9).

Table 9. Solubility of I134V-A377T-SteT in different detergents. Peaks containing monodisperse-behaving I134V-A377T-SteT in the indicated detergents (Figure 46) were collected and stored at ~1 mg/ml and 4°C. Protein concentration was measured before and after 3 days after ultracentrifugation.

DETERGENT	DM	NG	OG	Cymal-6
Solubility after 3 days	100 %	86 %	73 %	100 %

The remarkable stability of the I134V-A377T-SteT in detergents with proved success in membrane protein crystallography (particularly OG and Cymal-6) clearly indicates the optimal behavior of this mutant to initiate crystallographic screenings.

4.3.4 Detergent screening of L210Q-M229V-SteT

We also challenged the stability of purified L210Q-M229V-SteT, the second SteT mutant candidate obtained from the initial screening, in the same detergents used for I134V-A377T-SteT.

Similar to I134V-A377T-SteT, L210Q-M299V-SteT was solubilized with 1 % DM and after IMAC purification, the protein was concentrated up to ~4 mg/ml. As in the previous section, during this concentration process the protein was subjected to a buffer exchange containing 2xCMC of DM, NG, OG or Cymal-6. After eliminating the aggregates by ultracentrifugation, the supernatant was injected into a Superdex 200 5/150 column equilibrated with SEC buffer and 2xCMC of each detergent. In the case of the L210Q-M299V-SteT sample exchanged by NG, it was impossible to concentrate the protein as a consequence of protein aggregation, indicating the instability of L210Q-M299V-SteT in NG. As in the case of the previous SteT mutant and as judged by the SEC profiles of L210Q-M299V-SteT, the protein showed a monodisperse behavior in DM, OG and Cymal-6, clearly indicating its stability and robustness (Figure 48). L210Q-M299V-SteT volume of elution in the different detergents was also similar as the previous mutant (Figures 46 and 47) with subtle differences between detergents. The SDS-PAGE gel of the elution peaks of each detergent, confirmed the identity and purity of SteT (Figure 49). A higher molecular band in Cymal-6 is observed, very likely due to the typical dimerization artifacts that SDS-PAGE gels cause in membrane proteins (Figure 49).

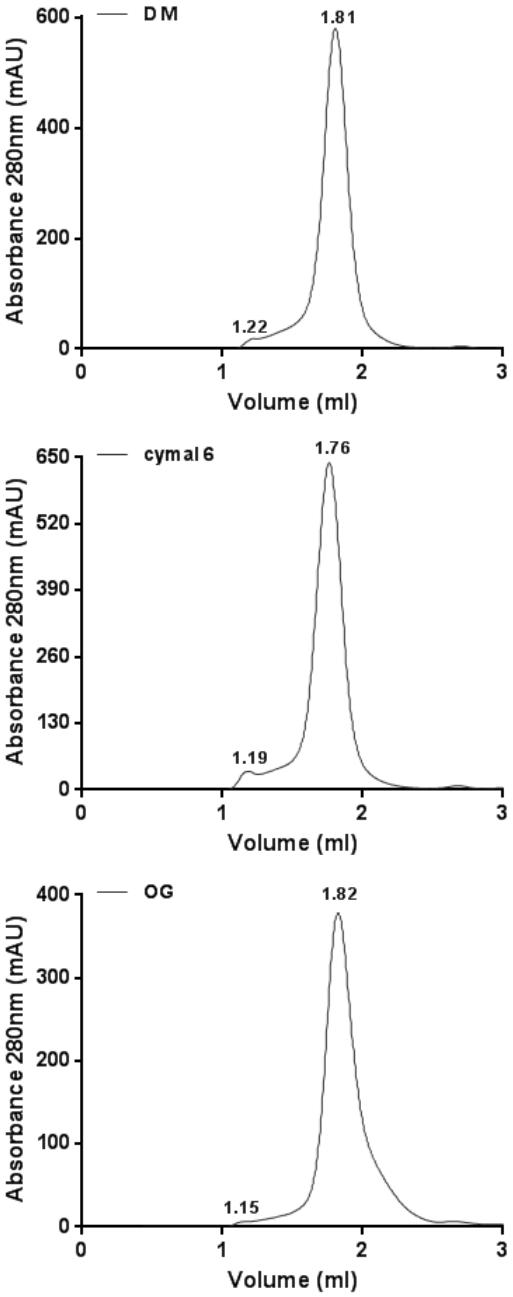


Figure 48. Size Exclusion Chromatography (SEC) profile of L210Q-M229V-SteT at 4 mg/ml. Experiments were done in a Superdex 200 5/150 GL column equilibrated with 20mM Tris pH 8.0, 150mM NaCl and 2xCMC of DM, OG and Cymal-6.

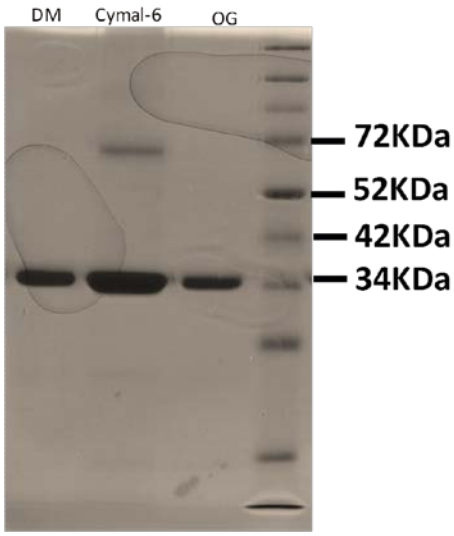


Figure 49. SDS-PAGE gel of the SEC L210Q-M299V-SteT fraction from the detergent screening experiment. The gel was stained with coomassie blue. Each line corresponds to the experiment done with the indicated detergent (top).

The monodisperse-eluted protein from the SEC experiments (Figure 48) was collected to further analyze its stability in each detergent. Following the same procedure used for the previous mutant, the collected protein was concentrated up to 1 mg/ml and kept at 4°C. No substantial variation of protein concentration after ultracentrifugation was observed after one day of incubation. After 3 days at 4°C only L210Q-M299V-SteT in OG aggregates since 32 % of protein was lost with this detergent (Table 10).

Table 10. Solubility of L210Q-M299V-SteT in different detergents. Peaks containing monodisperse-behaving I134V-A377T-SteT in the indicated detergents (Figure 48) were collected and stored at ~1 mg/ml and 4°C. Protein concentration was measured before and after 3 days after ultracentrifugation.

DETERGENT	DM	OG	Cymal-6
Solubility after 3 days	100%	68%	100%

4.4 Functional studies of L210Q-M229V-SteT and I134V-A377T-SteT: transports assays in proteoliposomes

The previous experiments have clearly demonstrated that the mutants found in our pipeline, L210Q-229V-SteT and I134V-A377T-SteT, have sufficient stability in detergent-micelle complexes and reasonable expression yield to send them for crystallization screenings. The criteria to only include mutants with no more than 2 amino acid substitutions in the mutant library obeys the objective to minimize the impact of such mutations over SteT functionality. In this regard, we wanted to characterize how the double mutations L210Q-229V or I134V-A377T alter the amino acid exchange activity of SteT. As commented in the Introduction, transport experiments using radiolabeled amino acids are the fastest way to study the functional properties of SteT, being the experiments performed in proteoliposomes the most indicated for membrane transporters previously purified. For these experiments, SteT wild-type, L210Q-229V-SteT and I134V-A377T-SteT with the GFP fused to the C-terminus were overexpressed in *E. coli* and purified by IMAC. Purified SteT variants were reconstituted into proteoliposomes by mixing them with *E. coli* lipids at a protein/lipid ratio of 1:100 (see Materials and Methods for details).

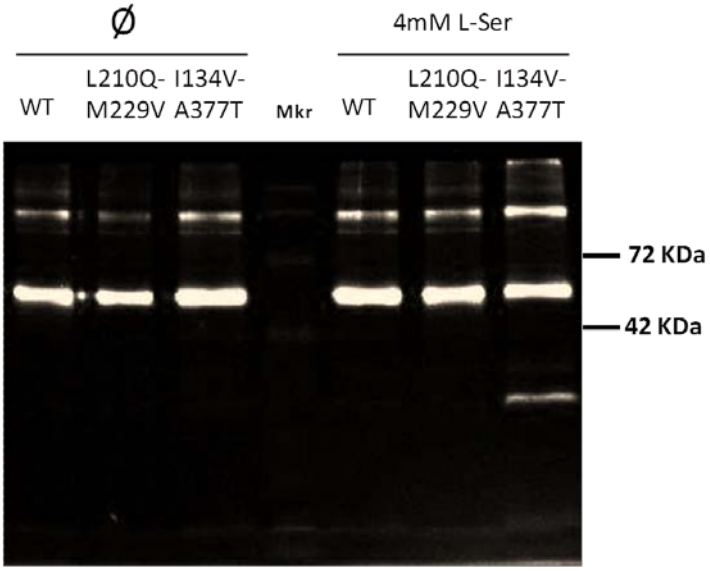


Figure 50. In gel fluorescence of SteT variants reconstituted in proteoliposomes. Proteoliposomes containing SteT wild type (WT) or mutants L210Q-229V-SteT and I134V-A377T-SteT were solubilized with 1 % SDS and subjected to SDS-PAGE electrophoresis. (Ø) Indicates proteoliposomes containing no L-Ser and (4mM Ser) indicates proteoliposomes loaded with 4 mM of L-Ser. Mkr is the protein marker.

Figure 50 displays the SDS-PAGE in-gel fluorescence of SteT-wild type and the two mutants after being incorporated in lipid vesicles. As judged by the intensity of the fluorescence bands observed at the apparent molecular weight of SteT-GFP (between 42 and 72 KDa, see Figure 50), the efficiency of protein reconstitution into proteoliposomas was similar in all three SteT variants.

Since SteT is an obligatory exchanger (as the mammalian LATs), uptake of radioactive L-Ser (SteT main substrate) only will take place if the interior of the proteoliposome contains another SteT substrate (in our experiment, a saturated concentration of 4 mM of non-radioactive L-Ser).

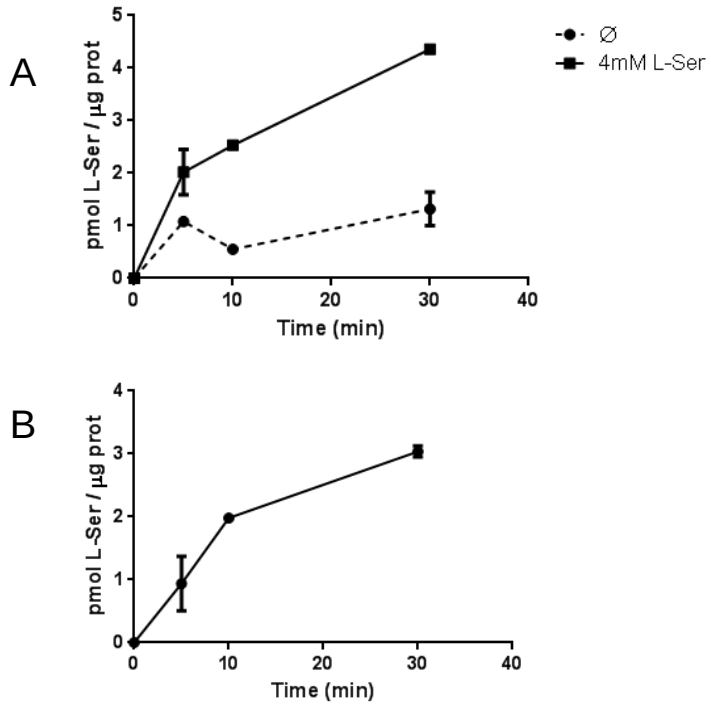


Figure 51. Transport activity of purified SteT wild type reconstituted into proteoliposomes. (A) Time-dependent uptake of 10 µM of radioactive L-Ser in proteoliposomes loaded with 4 mM of cold L-Ser (continuous line) or no loaded (ϕ and dashed lines). (B) Net transport activity of SteT wild type calculated from the data of panel A.

Proteoliposomes loaded with 4 mM of L-Ser were incubated with 10 µM or radioactive L-Ser and the amount of accumulated substrate was measured at different time intervals (Figure 51). Since L-Ser can diffuse through the lipid vesicles, the specific L-Ser uptake by SteT (net uptake) was calculated after subtracting the data obtained in the L-Ser loaded proteoliposomes from the one obtained from empty proteoliposomes. In our experiments of uptake we found that the activity of SteT wild type fused to the GFP was very similar as the reported previously in the absence of GFP (Reig et al., 2007) (Figure 51), indicating that GFP fused to the C-terminal end of SteT does not affect its

transport activity when comparing assays under the same experimental conditions (Reig et al., 2007). Notably, the double mutants studied L210Q-M229V-SteT and I134V-A377T-SteT transport L-Ser with similar kinetics (rate and maximum of accumulated L-Ser) as wild type (Figure 52 and 53).

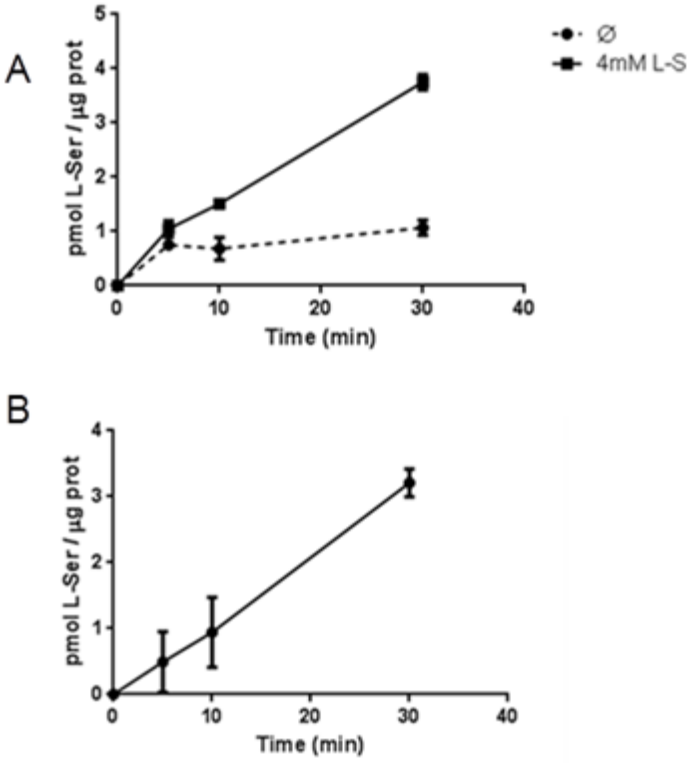


Figure 52. Transport activity of purified L210Q-M229V-SteT reconstituted into proteoliposomes. (A) Time-dependent uptake of 10 µM of radioactive L-Ser in proteoliposomes loaded with 4 mM of cold L-Ser (continuous line) or no loaded (ϕ and dashed lines). (B) Net transport activity of L210Q-M229V-SteT calculated using the data of panel A..

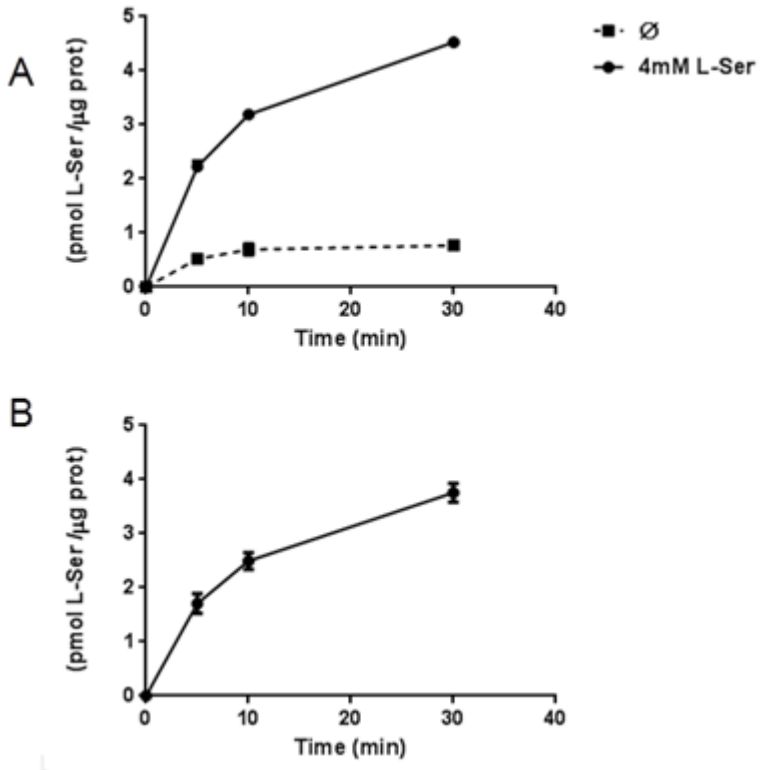


Figure 53. Transport activity of purified I134V-A377T-SteT reconstituted into proteoliposomes. (A) Time-dependent uptake of 10 µM of radioactive L-Ser in proteoliposomes loaded with 4 mM of cold L-Ser (continuous line) or no loaded (ϕ and dashed lines). (B) Net transport activity of I134V-A377T-SteT calculated using the data of panel A.

From these functional experiments, we could conclude that both set of double mutations I134V-A377T and L210Q-M229V do not alter substantially the transport activity of SteT. In addition, all the mutated positions are sitting away from the proposed substrate binding site of SteT (Bartoccioni et al, 2010) as well as from the binding site of its crystallized homolog, Adic (Kowalczyk et al., 2011). It is clear that our transport experiments (Figures 52 and 53) do not explain all the functional characteristics of these SteT variants; however we can conclude that these mutations, if any, would affect mostly the maximum velocity of substrate translocation rather to the ability of the protein to

recognize the substrate, L-Ser. In this sense, these functional experiments prove that I134V-A377T-SteT and L210Q-M229V-SteT are still excellent structural paradigms of the mammalian LATs.

4.5 Crystallization screenings

The main objective of this thesis has been to validate our protocol as a useful and general approach to build stability in a membrane protein target to initiate crystallographic studies. The data from the previous sections have pointed I134V-A377T-SteT and L210Q-M229V-SteT as two optimal variants of SteT obtained from random mutagenesis that have completely changed both the expression yield and the stability of SteT allowing for the first time to perform crystallization assays of this LAT homolog. Logically, the next step was to undergo crystallization trials.

As revealed earlier, I134V-A377T-SteT was the mutant that showed the highest expression yield using the pTTQ18-GFP expression vector (Table 7); consequently I134V-A377T-SteT was the target chosen for crystallization. The data presented in this section contains three different crystallization experiments of I134V-A377T-SteT. The difference between them is the detergent used for solubilizing (and stabilizing) the protein since, as commented, the nature of the protein-detergent micelle complex is determinant for getting protein crystals.

Each experiment consisted on testing a total number of 288 crystallizing conditions spread in three 96-well plates. These conditions were obtained from commercial sources and are suited for IMP crystallization (Membfac, Memplus, Memstart, Memsys and Memgold). Each crystallization condition was tested at two temperatures (20°C and 4°C) and, in some cases, at two different protein

concentrations. Sitting drops were seeded using robots by mixing 15 μl of protein and the same volume of crystallization condition or precipitant.

4.5.1 OG screening

We have found several detergents with optimal properties after solubilizing I134V-A377T-SteT for crystallization trials (Figure 46). In order to increase the probability of forming crystals, our first choice was the detergent that has the smaller micelle size: OG. In addition, OG was used to crystallize ApcT and GadC (Shaffer et al., 2009) and (Ma et al., 2012), the APC homologs of SteT. As in the previous experiments in section 4.3.3, membranes expressing I134V-A377T-SteT were solubilized with DM and purified by IMAC. DM was changed by OG during protein concentration and the concentrated protein in OG was injected in the SEC column for further purification and completely detergent exchange (Figure 57). The final buffer conditions of the purified protein were 2xCMC of OG and 10% of glycerol. The glycerol was added to increment the stability of the protein, since it is believed that glycerol generates a more native environment surrounding the protein-detergent micelle by reducing the water concentration and increasing the hydrophobicity (Iwata, 2003 and Byrne and Jormakka, 2006). As we expected from previous experiments, SEC elution profile of purified I134V-A377T-SteT in OG showed a monodisperse behavior and eluted as monomer (Figure 54).

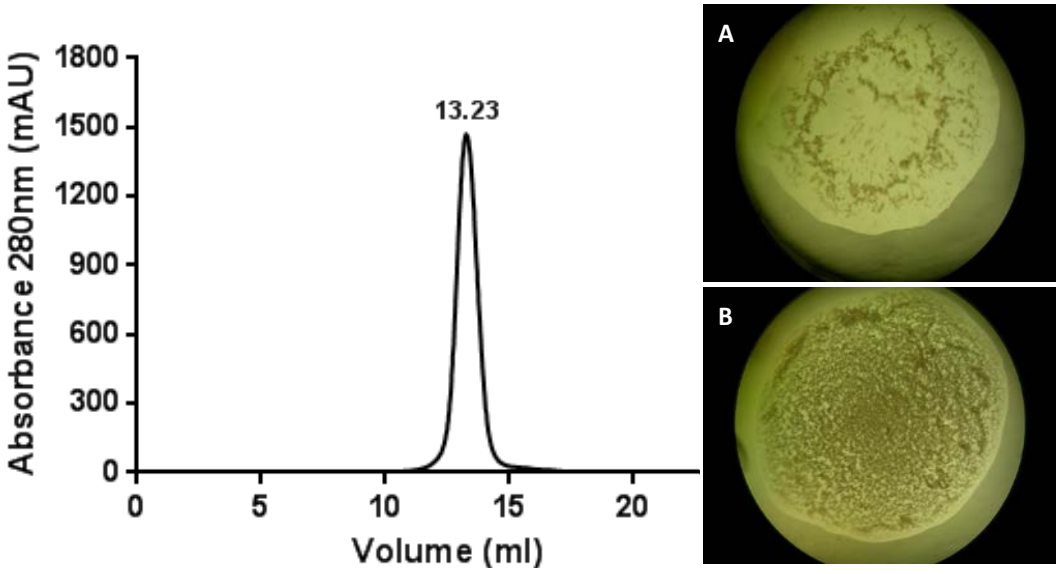


Figure 54. Crystalization screening of I134V-A377T-SteT (left) Size exclusion chromatography of I134V-A377T with 2xCMC of OG. (right) (A) memstar well with I134V-A377T concentrated to 2 mg/ml (B) memstar well with I134V-A377T concentrated to 8 mg/ml.

After seeding the protein mixed with the different precipitant conditions, I134V-A377T-SteT aggregates appeared shortly (~ 1 day) in most of the crystallization conditions. The presence of aggregates increased over the days. Amorphous precipitation was observed in the majority of the conditions (Figure 54, right panels A and B). Interestingly, some precipitates were observed with the so-called “skins” (small and very thin layer that covers the drop), indicating denatured protein.

The fact that none of the 576 crystallization conditions induced I134V-A377T-SteT crystallization was obviously indicating that, although the protein was not enough stable in OG (Figure 54, right panels A and B), the protein aggregated faster than the time required for nucleation. In this line, it would be interesting to repeat these assays with OG, varying protein concentration, glycerol content

(sometimes glycerol interferes crystallization) and temperature aiming to control better the rate of aggregation and crystal nuclei formation.

4.5.2 DM screening

Giving that the screening done in DM exchanged to OG resulted on aggregates; a second screening with I134V-A377T-SteT solubilized DM and, therefore, without exchanging detergent, was performed. We expected to increase protein stability in the crystallization tests. IMAC purified I134V-A377T-SteT was concentrated to ~ 8 mg/ml and further purified by SEC in a Superdex 200 10/300 column equilibrated with SEC buffer and 2xCMC DM. As we expected, SEC profile of I134V-A377T-SteT purified in DM showed a monodisperse behavior (Figure 55).

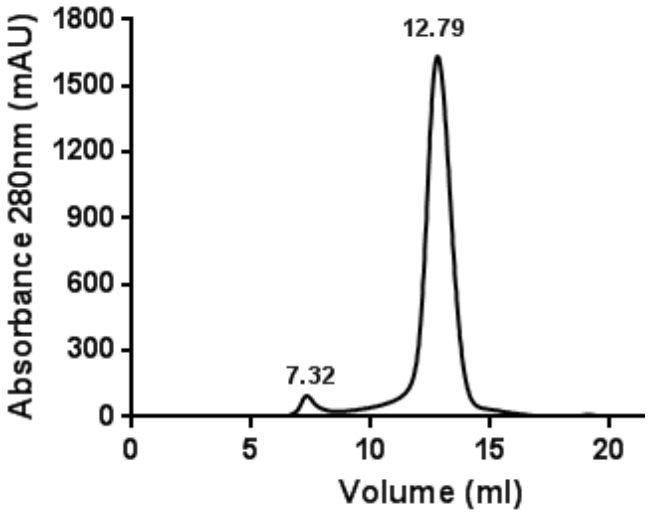


Figure 55. SEC profile of I134V-A377T-SteT solubilized and purified in DM. SEC elution profile of I134V-A377T-SteT. IMAC purified protein was concentrated up to ~8 mg/ml and injected in a Superdex 200 10/300 SEC column equilibrated with SEC buffer and 2xCMC of DM.

SEC purified Superdex 200 10/300 column was concentrated to 2 and 8 mg/ml and dialyzed for 24 h. Again, in this step the objective was to remove as much as possible the excess of detergent that accumulates during protein concentration. After dialysis, possible aggregates were removed by ultracentrifugation and protein was seeded. In this detergent, spherulites of I134V-A377T-SteT started to appear after 1 day of seeding at the two temperatures used for screening (20°C and 4°C) and its number and size increased over the next weeks (Figure 56).

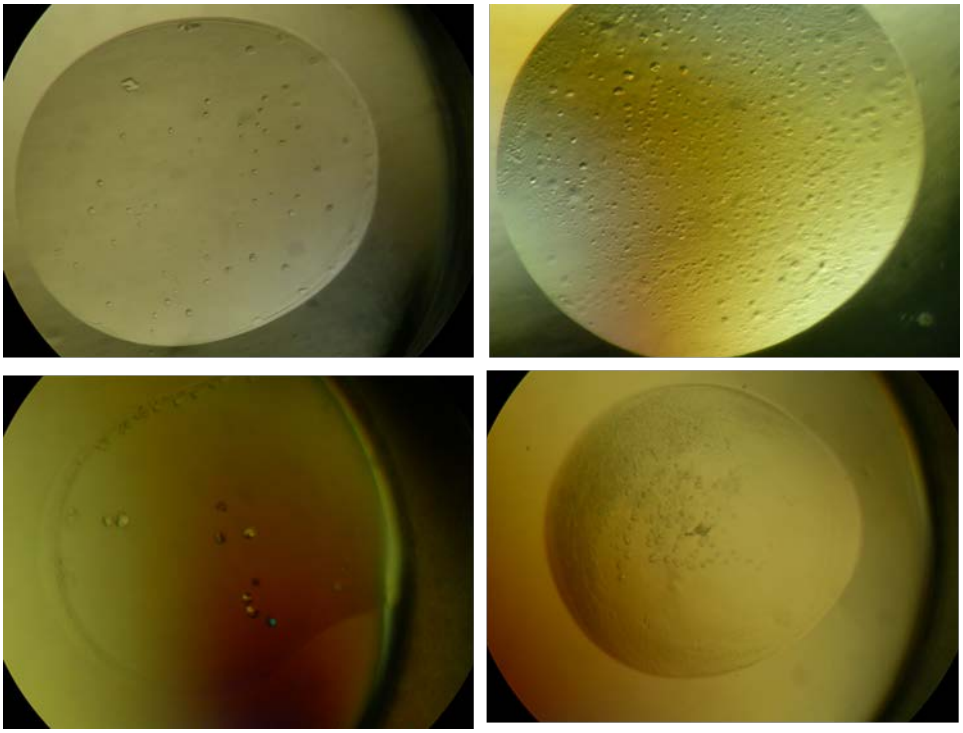


Figure 56. Crystalizzation screening using 2xCMC of DM of I134V-A377T-SteT (left): Protein concentrated to 2 mg/ml. (right): Protein concentrated to 8 mg/ml. Photos of wells in diferent PEGs mixture (350 400 550 600) and pH around 8.0.

However no crystal formation was found, the spherulites formation was very interesting. These structures can often be precursors of crystal formation and in

Interestingly, in the conditions where the majority of the spherulites appeared, the precipitant contained low molecular weight PEGs (PEG 350, PEG 400, PEG 550, PEG 600 and mixtures) and pH around 8.0; similar to the conditions where AdiC, ApcT and GadC were crystallized (Kowalczyk et al., 2011), (Shaffer et al., 2009) and (Ma et al., 2012) (Figure 56)

4.5.3 Cymal-6 screening

In the next crystallization tests of I134V-A377T-SteT, the protein was solubilized in Cymal-6. This detergent produces a small protein-detergent micelle size than DM while maintaining I134V-A377T-SteT stability in solution (Figure 46). In addition, the excess of Cymal-6 during protein concentration is easier to eliminate than DM excess and, therefore, we expected to switch the equilibrium of I134V-A377T-SteT crystallization from spherulites (Figure 56) to crystals. Cymal-6 was used at 2xCMC concentration during all the purification steps of I134V-A377T-SteT, including protein solubilization from *E. coli* membranes where 1 % of Cymal-6 was used. SEC profile of I134V-A377T-SteT in 2xCMC of Cymal-6 was also showing a monodisperse protein (Figure 57). After SEC purification, I134V-A377T-SteT was concentrated to 2 and 8 mg/ml and dialyzed for 24 h to remove the excess of detergent. After dialysis, aggregates were removed by ultracentrifugation and protein concentration was adjusted to 2 mg/ml and 4 mg/ml for crystallization tests.

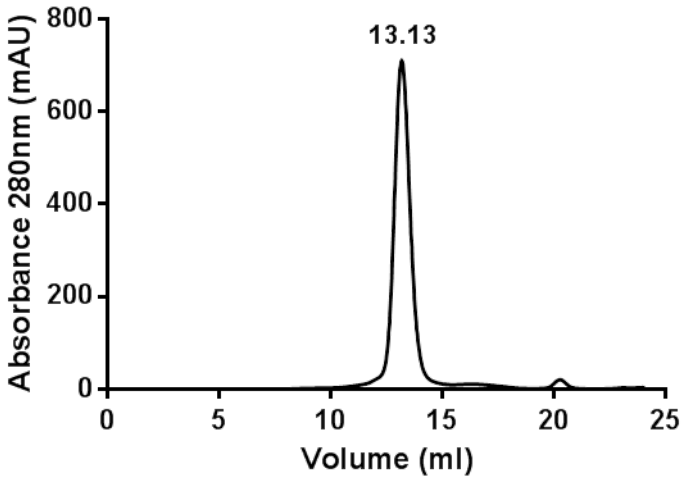


Figure 57. SEC elution profile of I134V-A377T-SteT solubilized and purified in Cymal-6. IMAC purified protein was concentrated up to ~5 mg/ml and injected in a Superdex 200 10/300 SEC column equilibrated with SEC buffer and 2xCMC of Cymal-6.

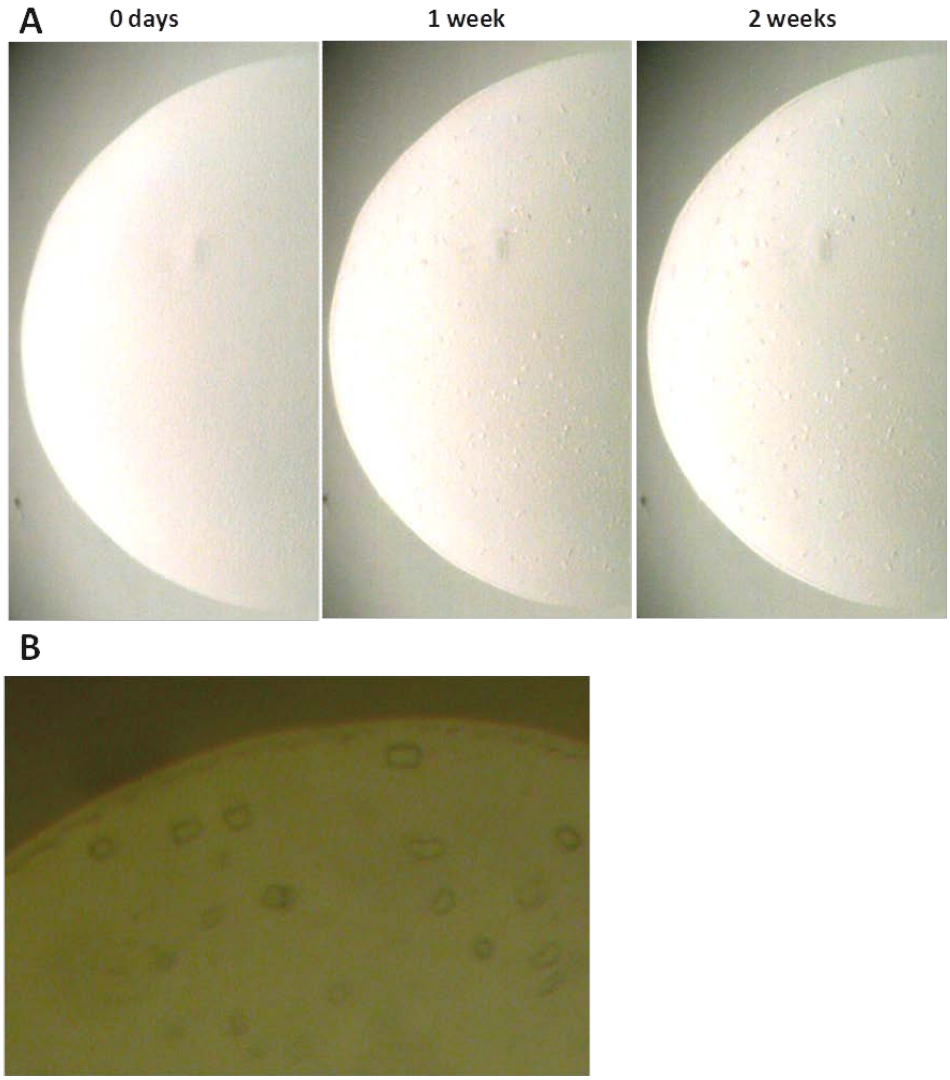


Figure 58. Crystallization screening of I134V-A377T-SteT solubilized in Cymal-6. (A) Automatic pictures of crystallization drops from a 96-well plate taken by the Crystal Farm plate-housing incubator. Pictures of the same well were taken at the indicated times after protein seeding. Precipitant conditions: 25 % (w/v) of PEG 1500, 4.3 % (w/v) PEG 4000, 0,1 M sodium acetate pH 4.6. (B) Magnified image of the 2-week picture from panel (A) taken by a Leica lupe.

After 1 week of protein seeding, small crystals started to appear at 20°C when the precipitant contained 25 % (w/v) of PEG 1500, 4.3 % (w/v) PEG 4000, 0,1 M sodium acetate pH 4.6 (Figure 58). The size of these crystals was not bigger

than 5-10 micrometers (Figure 58). In addition, spherulites and crystal-like forms were also observed in similar precipitant conditions. It is worth mentioning that most of well-diffracting crystals of APC transporters appeared at basic pH (Kowalczyk et al., 2011), (Shaffer et al., 2009) and (Ma et al., 2012) , contrarily of these crystals hits of I134V-A377T-SteT. In addition, some spherulites were observed in similar precipitant conditions as the one used to obtain crystals of AdiC, ApcT or GadC.

These exciting initial crystal hits of a SteT mutant are, perhaps, an excellent proof of concept of the experimental approach developed in this thesis. Based on the found crystallizing conditions, further experiments of crystal growth of I134V-A377T-SteT in large-volume drops will be carried out shortly in order to obtain optimal SteT crystals regarding size and quality to diffract X-rays for structure determination.

GLOBAL DISCUSSION

5 GLOBAL DISCUSSION

Unraveling the molecular architecture and the transport mechanism of membrane transport proteins is crucial to understand the majority of physiological processes and for current drug development. In this regard, x-ray crystallography is, now at days, the most powerful technique to study the structural elements that determine the function and regulation of this class of proteins. Consequently, over the last years, several consortiums have been created with the main goal of joint efforts for IMPs structural elucidation (e.g., the New York consortium on Membrane Protein structure, http://nycomps.org, or the European Drug Initiative on Channels and Transporters, http://www.edict-project.eu), where membrane transporters represent one of the main targets.

Despite their biological importance, IMPs are extremely difficult targets for structural biologists. They are notoriously resistant to crystallize due to their hydrophobic nature; however, the bottleneck when working with IMPs is the difficulty of producing enough quantity and quality of recombinant versions of these proteins. Structural studies of IMPs often require a previous and extensive search for candidates that must fulfill two main properties: reasonable expression yield in the chosen expression host and good stability after detergent solubilization and purification. In this sense, the development of experimental strategies to facilitate and optimize this search has been, undoubtedly, the key for speeding up the 3D structural resolution of IMPs. Perhaps, the most successful one (and widely used as well) is the combination of the “funnel approach” (Lewinson et al., 2008b), consisting on testing in parallel many prokaryotic sequence homologs of a highly relevant but untreatable IMP target, with fast and efficient screening protocols of

expression and thermostability; preferentially in a high throughput manner (Mancia and Love, 2010); (Vergis et al., 2010) and (Eshaghi, 2009). This confirms the idea that new screening approaches are extremely beneficial in the field of structural biology of membrane proteins.

Our laboratory identified SteT as a highly potential structural model of eukaryotic LATs due to the closest amino acid identity (~30 %) to its human counterparts together with its functional properties (broad substrate recognition, obligatory exchanger) (Figure 12) (Reig et al., 2007) and (Bartoccioni et al., 2010). Unfortunately, SteT showed quite instability in detergent after membrane extraction and purification, precipitating at concentrations near 3 mg/ml and, consequently, making impossible to crystallize it (Vázquez-Ibar JL and Palacín, M; unpublished results) (Figure 14). We, therefore, decided to build stability in SteT using mutagenesis, taking into account that this tool has been the key for stabilizing and crystallizing a few MTPs (Abramson et al., 2003) and (Kowalczyk et al., 2011)and, most importantly, for obtaining the first atomic structures of GPCR (Miller and Tate, 2011). Indeed our final goal was to establish a general protocol that eventually could be applied to any membrane protein of interest (see specific objectives).

As the first step in the protocol, a library of SteT random mutants was created with amino acid substitutions within TMDs. Punctual mutations in these regions have the biggest impact on both the expression and stability of IMPs (Kang et al., 2013). Since, it is highly difficult to predict what amino acid sequence or what amino acid replacement is going to improve the expression of a particular IMP expressed heterologously in *E. coli*, random mutagenesis was chosen as the optimal alternative. Indeed, when working with orphan or poorly biochemically characterized proteins, random mutagenesis has proved to be an

excellent strategy to find functionally relevant residues in membrane transporters (Malle et al., 2011) and (Zhao et al., 2011) or in GPCR (Li et al., 2005). Most importantly, systematic cycles of random mutagenesis has been also employed to improve expression and stability of GPCRs (Sarkar et al., 2008).

An initial library of ~ 400 random mutants of SteT that express and insert in the cytoplasmic membrane of *E. coli* was easily created in less than a week. Our methodology of combining a previously optimized error-prone PCR reaction and the split-GFP assay adapted and validated for us for membrane proteins (Rodríguez-Banqueri et al., 2012), allowed us to quickly generate this library with very little cost. As screening tool for MTP expression and "in vivo" stability in *E. coli*, the split-GFP assay appears to be extremely useful not only by minimizing any side effect caused by the full-length GFP fused to the target (in our case, a particular random SteT mutant), but also by providing a more precise information in a very fast manner regarding protein folding and membrane insertion (Rodríguez-Banqueri et al., 2012).

150 out of the initial 400 mutants (38 %) contained the desired rate of amino acid substitutions: 1 or 2. This restriction in the number of amino acid substitutions on each SteT random mutant was imposed to preserve the protein from possible interfering effects of such mutations on protein folding and/or function. Among these 150 mutants, 101 (67 %) contained the amino acid substitutions in the predicted TMDs (Table 4), indicating that our mutagenesis reaction protocol was able to efficiently distribute random mutations along the protein sequence, since the amino acid located in SteT TMDs represent 65 % among the total number of amino acid of the protein.

A representative pool of 70 random SteT mutants (Table 5) covering a homogeneous amino acid replacements along the 12 TMDs of SteT (Figures 26 and 27) was finally selected for further analysis of protein expression and stability after DDM solubilization. Although different methods of MTPs stability (or thermostability, since this property is related with the probability of crystallization) have been proposed (Alexandrov et al., 2008), SEC is still the most sensitive technique for this purpose, and highly convenient to test many samples like our random mutants library without the necessity of purifying by simply using the GFP as reporter (FSEC) (Kawate and Gouaux, 2006). Good behavior of FSEC profiles is a direct proof of protein stability, the rate-limiting step of obtaining crystals for structure elucidation. Moreover, for detergent-solubilized secondary transporters, it is very difficult to set up a medium to high throughput substrate binding assay to evaluate the resistance of the protein from thermal denaturation, as commonly used to evaluate the stability of GPCR (Tate, 2012). Normally, binding affinities expressed as dissociation constants of most secondary transporters like SteT in detergent micelles range from high micromolar to millimolar while GPCRs binding affinities in the same conditions are in the range of nanomolar. Therefore, FSEC was the optimal technique to evaluate the degree of monodispersity (or stability) of each mutant after DDM solubilization and compared with SteT wild type. We assigned a numeric value (we called index of monodispersity) related to the sharpness of the SEC elution peak of each mutant (see Materials and Methods for details) to objectively evaluate and compare the effect of every mutation on the stability of SteT wild type after DDM solubilization (Figure 59).

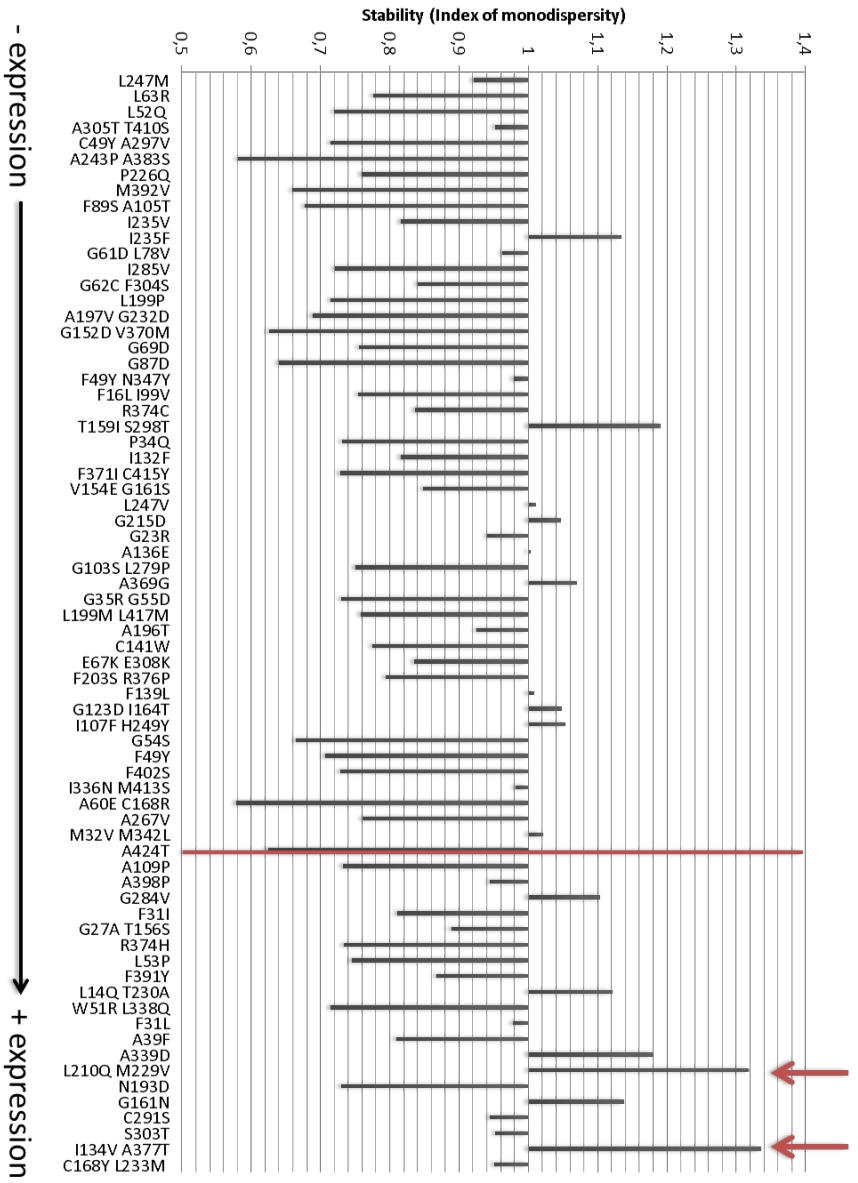


Figure 59. Calculated index of monodispersity of SteT random mutants selected for screening. This index is calculated by dividing the area of the normalized FSEC elution peak of wild type by the area of the normalized FSEC elution peak of each mutant (see materials and methods for full description). The mutants are ordered regarding their expression level from the top to the bottom. Mutants down the red line express more than wild type. Mutants that are more stable than wild type have a Index of monodispersity above 1. Selected mutants for large-scale purification are indicated by red arrows.

Putting together the expression yield and index of monodispersity of the studied SteT random mutants (Figure 59) not only allowed us to discriminate those mutants with ideal robustness for next step screening tests (I134V-A377T-SteT and L210Q-M229V-SteT, see Figure 59) but also, the results provided us relevant information regarding the relationships between heterologous expression yield of MTPs in *E. coli* and stability after detergent solubilization. As found by other laboratories regarding mutations in TMDs (Martinez Molina et al., 2008), most of the mutations had a negative effect in both SteT expression yield and stability in DDM. A large number of amino acid replacements (71 % of mutants) decreased wild type expression. Also, 60 % of mutants were less stable in detergent than wild type. Notably, among the mutants that showed more expression than wild type, 6 of them (30 %) presented more stability than SteT, particularly, I134V-A377T-SteT and L210Q-M229V-SteT (Figure 59). Conversely, only 2 mutants (4 %) out of the 50 expressing less than wild type were more stable in DDM than wild type, indicating some level of relationship between amount of expressed protein and stability in DDM. In summary, the screening of expression yield and stability in DDM pointed I134V-A377T-SteT and L210Q-M229V-SteT as the optimal candidates for large-scale purification (Figure 59). Furthermore, both mutants were also stable in detergents more suitable for crystallization like DM, OG or Cymal-6 (Figures 35 and 36), a criteria also used by other laboratories when searching for crystallization candidates (Sonoda et al., 2011).

Although we cannot state that there is a strong relationship between the level of expression and detergent stability (Figure 59), our results in SteT suggest that structural elements (induced by side-chain substitutions) that improve membrane protein production in *E. coli* due to one or different factors (e.g., increasing mRNA stability or the rate of translation and membrane insertion or

facilitating the hydrophobic matching between the protein and the phospholipids of the membrane) can also stabilize the protein after detergent solubilization. Unraveling what is the precise molecular mechanism that explains changes in expression and/or stability of SteT as a consequence of a particular side chain substitution will require a large number of experimental data and, logically, it is out of the scope of this thesis. Nevertheless, observing the positions of the mutated positions in I134V-A377T-SteT and L210Q-M229V-SteT (Figure 60) we can assure that these mutated positions are not affecting substantially the tertiary structure of SteT nor its functional properties since the putative substrate binding site is far away from those positions (Figure 13) (Bartoccioni et al., 2010).

In addition, by analyzing the transport activity of I134V-A377T-SteT and L210Q-M229V-SteT in proteoliposomes, we observed that the translocation activity of SteT was almost unaltered after these amino acid replacements (Figures 51 and 52). Therefore, the observed stabilizing effect of these mutations cannot be attributed to a conformational stabilization of the transporter as found before for LacY (Abramson et al., 2003) or AdiC (Kowalczyk et al., 2011). Clearly, more precise functional characterization will be needed to figure out in what extend I134V-A377T or L210Q-M229V replacements alter the functional properties of SteT.

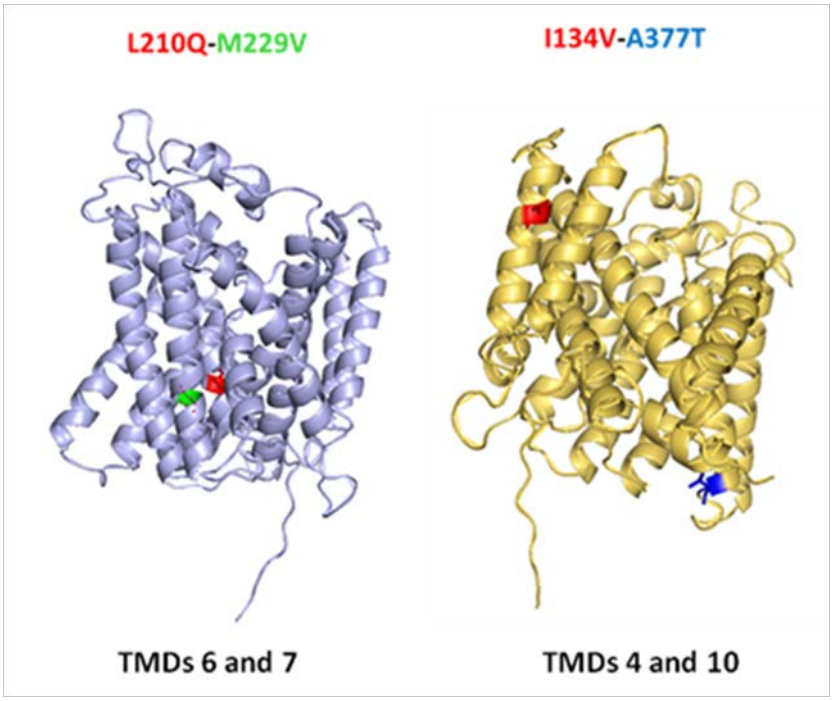


Figure 60. Localization of positions L210/M229 and I134/A377 in the SteT 3D model. TMD refers to transmembrane domain. The position of the mutations in the model are colored in red, green or blue.

In the next step, we performed large-scale purification of I134V-A377T-SteT and L210Q-M229V-SteT followed by stability tests using SEC after protein concentration and after removal the GFP by proteolysis. All these experiments reinforced the candidature of these mutants as ideal targets to initiate crystallization trials (Figures 40, 41 and 43). Different reasons made us to further analyze the stability of I134V-A377T-SteT and L210Q-M229V-SteT after purification. First, it is necessary to stress out the protein using similar experimental conditions used for crystallization studies; for example evaluate the stability in detergent at high protein concentration (Figure 43). Second, FSECs studies not always correlate with SECs studies after purification. One

reason of the later is that the loss of lipids bound to the protein during the affinity purification can alter the properties of the protein in the detergent-micelle complex. In addition, as discussed earlier, the presence GFP can, sometimes, alter the solubility and/or the stability of the fused protein. In this sense, we observed absolutely no effect regarding protein stability in any of the two SteT mutants after GFP removal (Figure 40). Perhaps, the initial screening of SteT random mutants using the split-GFP assay was determinant to discard any possible side effect of full-length GFP in SteT.

Finally, stability tests using detergents with smaller polar head groups and shorter aliphatic chains (DM, NG, OG and Cymal-6) validated the suitability of I134V-A377T-SteT and L210Q-M229V-SteT to initiate crystallization studies (Figures 46 and 48). With the only exception of L210Q-M229V-SteT in NG, both mutants are stable in these detergents at concentrations and buffer conditions (specially, 2xCMC of a given detergent) commonly used for crystallographic studies.

In conclusion, this thesis has validated and optimized an experimental pipeline strategy to build the necessary robustness in SteT for crystallization studies. However, this method can be applied to any membrane protein heterologous expressed in *E. coli*. Our results suggest that random mutagenesis combined with quick and sensitive screening tests is a valuable strategy to find mutated versions of a membrane transport protein with a notable improvement of expression yield and stability after detergent solubilization, ideal properties of an IMP target for structural studies. This method has allowed us to find two mutants of SteT, I134V-A377T-SteT and L210Q-M229V-SteT that are currently undergoing crystallization screening for structure resolution. The initial crystal

hits of I134V-A377T-SteT (Figure 58) are preliminary results that reinforce the general applicability of this work.

CONCLUSIONS

6 CONCLUSIONS

I. Using SteT as a proof of principle, we have validated a general methodology to build stability in a membrane transport protein for crystallization studies.

II. Random mutagenesis combined with quick and sensitive screening assays appears to be an optimal strategy to find and characterize mutations in a membrane transport protein that considerably improves its expression yield and stability after detergent solubilization.

III. Using the split-GFP as reporter of membrane proteins it is possible to obtain a more precise and fast information regarding protein folding and membrane insertion of these proteins minimizing any possible side effect caused by the presence of full-length GFP.

IV. Single or double amino acid replacements in TMDs of SteT have a profound effect on both protein expression yield and protein stability after detergent solubilization.

V. Although the most common effect after amino acid replacements in SteT TMDs is a decrease in both expression and stability, a few mutants show a substantial improvement of both properties.

VI. Side-chain substitutions that improve SteT expression in *E. coli* are also more likely to stabilize the protein after detergent solubilization.

VII. The double amino acid substitutions in SteT, I134V-A377T or L210Q-M229V provides a notably enhancement of protein expression and stability after detergent solubilization and purification without altering the functionality of the transporter.

VIII. The screening experiments performed with non-purified I134V-A377T-SteT and L210Q-M229V-SteT correlates very well with the experiments with purified samples, validating our strategy of combining the split-GFP assay with FSEC to screen protein expression and stability after detergent solubilization.

IX. I134V-A377T-SteT is an excellent structural paradigm to study the structure and mechanism of the eukaryotic LATs.

MATERIALS AND METHODS

7 MATERIALS AND METHODS

7.1 Molecular biology protocols

DNA manipulations were carried out according to Sambrook (Sambrook 2001).

7.1.1 DNA extraction

The clones were grown in 200ml culture to extract the DNA with maxiprep kit (Qiagen) 5 ml of culture for miniprep kit (Qiagen), both using manufacter instructions. Both kits are based in the cellular membrane lysis by a salt treatment and the DNA adsorpsion in a silica resin in presence of high salt concentration buffer. The elution was done a low salt concentration buffer or deionized water. DNA plasmid concentration was measure by nanodrop.

7.1.2 Random Mutagenesis

Random mutagenesis of the cDNA encoding SteT purchased from GenScript (Piscataway, NJ, USA) was performed using GeneMorph II EZClone Domain Mutagenesis Kit.

The first reaction consisted in an error prone PCR with Mutazyme II DNA polymerase. This reaction was performed according the kit instructions, using the next procedure setting in the thermo cycler after using 500ng of template target, in order to obtain a low mutation rate:

Table 11. Error prone PCR conditions

Segment	Temperature	Time	N° of cycles
1	95°C	2 mins	1 cycle
2	95°C	30 s	28 cycles
	55°C	30 s	
	70°C	2 mins 30 s	
3	72°C	10 mins	1 cycle

This reaction was purified using GE healthcare gel purification kit, after the purification this pcr reaction was used as a megaprimer of the EZclone reaction, this reaction was done following the kit instructions, using pTETGFP$_{11}$ with cDNA encoding SteT purchased from GenScript (Piscataway, NJ, USA) . Finally, this reaction was digested by 1µl of dpnI during 3 h. After the digestion, the reaction containing the mutants in pTET were transformed in *E. coli* BL21 (DE3) harboring pET encoding GFP$_{1-10}$, to start the fluorescence screening using the GFP split system in vivo with bacterial colonies.

7.1.3 Sequencing

Sequencing was used in order to check if a mutation has been incorporated successfully in the DNA. In the present work the kit "ABI PRISM Dye terminator Cycle Sequencing Ready Reaction" was fully used. The kit consists in a PCR reaction with fluorescent dinucleotides. These fluorescence dinucleotides are placed in the DNA during the amplification. The screening of the reaction of the present thesis work were carried out by "Serveis Cientifico-Tècnics de la UB " in Parc Cientific de Barcelona (PCB).

7.2 *E. coli* strains, transformation protocols and cloning procedures

E. coli XL1BLUE strain has been used for extraction of plasmidic DNA and for cloning and sequencing. *E. coli* BL21 (DE3), BL21 (DE3) Star and C-43 (DE3) have been used for expression of the membrane proteins. *E. coli* BL21 Star (DE 3) cells have a genotype that promotes high mRNA stability and protein yield; and *E. coli* C-43 (DE3) cells are effective expressing toxic and membrane proteins

7.2.1 Competent cells BL21(DE3) Star + pETGFP$_{1-10}$

materials and reagents:

- LB media
- $CaCl_2$ 0.1M
- 0.1 M $CaCl_2$ + 15% glycerol

procedure:

pETGFP$_{1-10}$ plasmid was transformed in BL21(DE3) Star competent cells. One colony was inoculated into 2 ml LB medium incubated at 37°C overnight with agitation. 1-ml overnight cell culture was diluted into 100 ml LB medium (in a 500 ml flask). This media was incubated at 37°C until OD_{600} = ~0.25-0.3 with agitation. Then, the culture was chilled on ice for 15 min. the cells were centrifuged for 10 min at 3300 x g at 4°C. The medium was discarded and the cell pellet resuspended in 30-40 ml cold 0.1M $CaCl_2$. The cells were kept on ice for 30 min. Then, the cells were centrifuged for 10 min at 3300 x g at 4°C. The supernatant was removed, and the cell pellet was resuspended in 6 ml of buffer 0.1 M CaCl2 solution plus 15% glycerol. 0.4 - 0.5 ml of the cell suspension was dispensed into sterile 1.5 ml micro-centrifuge tubes. The tubes were frozen on liquid nitrogen and then transferred them to -70 °C freezer.

7.2.2 Transformation in XL1BLUE, BL21 (DE3), BL21 (DE3) Star and C-43 (DE3) competent cells

Competent *E. coli* cells were taken from -80°C freezer and kept at 4°C. Then 80 µl of competent cells with 100 ng of DNA were kept on ice during 20 min. The competent cells were incubated with the DNA at 42°C during 90 s. And then, the sample was incubated on ice for 2 min. After this step, 900 µl of LB was added and competent cells with LB media were incubated for 1h at 37°C. Then, a centrifugation was performed at 15.000 rpm during 10 s, and the supernatant was discarded. Finally the pellet was resuspended in 100 µl of LB and was

spread on LB agar plate with the corresponding antibiotic. The plate was incubated overnight at 37°C.

7.2.3 Cloning membrane proteins on GFP-split system

pTET-GFP_{11} and pET-GFP_{1-10} vectors encoding GFP_{11} (GFP residues 1 to 16) and GFP_{1-10} (GFP residues 16 to 230) (Cabantous, 2006), respectively were generously provided by Dr. Geoffrey S. Waldo (Los Alamos National Laboratory, USA). The genes encoding SteT-$(His)_6$, SteT-G294V-$(His)_6$, EmrE-$(His)_6$, LacY, MscL and MscS and the mutated versions of some of them were cloned into the *NcoI* and *BamHI* sites of the pTET-GFP_{11} vector to generate the corresponding C-terminal GFP_{11} fusion of each protein. In SteT and EmrE constructs, a 6XHis tag is present between the C-terminal of the membrane protein and the N-terminal of the GFP_{11} was cloned in *NdeI* and *BamHI* sites. All constructs were verified by DNA sequencing.

7.2.4 Cloning modified SteT with codon usage from *E. coli* proteins on GFP-split system

pTET-GFP_{11} and pET-GFP_{1-10} vectors encoding GFP_{11} (GFP residues 1 to 16) and GFP_{1-10} (GFP residues 16 to 230) (Cabantous, 2006), respectively were generously provided by Dr. Geoffrey S. Waldo (Los Alamos National Laboratory, USA). The genes encoding SteT purchased from GenScript (Piscataway, NJ, USA), were cloned into the NcoI and BamHI sites of the pTET-GFP_{11} vector to generate the corresponding C-terminal GFP_{11} fusion of each protein. All constructs were verified by DNA sequencing. This vector was used for generate the random mutant library

7.2.5 Cloning membrane proteins on PTTQ18-His(x10)-GFP

The genes encoding SteT purchased from GenScript (Piscataway, NJ, USA), L210Q-M229V and SteT-I134V-A377T were cloned into the *EcoRI* and *PstI* sites of the PTTQ18-$(His)_{10}$-GFP vector to generate the corresponding protein fusion with GFP of each protein. Prescission protease target Leu-Glu-Val-Leu-Phe-Gln-Gly-Pro was introduced with three Pro residues between the Prescision target and the C-terminus of the membrane protein using PCR. All constructs were verified by DNA sequencing.

7.3 Measurement of protein concentration and detection by "In-gel fluorescent"

7.3.1 "In-gel fluorescence" on SDS-PAGE

Membranes and purified protein were visualized in a sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) gel (Laemmli, 1970). Isolated cytoplasmic membranes were solubilized with 1% n-dodecyl-β-d-maltopyranoside (DDM, Affymetrix, Santa, Clara, CA, USA) and subjected to SDS–PAGE. "In-gel" GFP fluorescence from the distinct membrane protein–GFP_{11}–$GFP_{1–10}$ gel bands was visualized using a GBOX gel reader (Syngene, Cambridge, UK) and a Safe Imager (Invitrogen, Carlsbad, CA, USA).

7.3.2 Measurement of membrane and protein concentration

Membranes and solubilized protein from membranes were measured by *BCA Protein Assay Reagent* (Pierce), using kit instructions. Purified protein was measured by nanodrop and using protparam (http://expasy.org) and with *BCA Protein Assay Reagent* kit (Pierce), using kit instructions.

7.4 GFP split system methods

7.4.1 In vivo fluorescence screening assay in bacterial cultures.

materials and reagents:

- BL21(DE3) co-transformed with pETGFP$_{1-10}$ and pTET with membrane protein fused to GFP$_{11}$
- LB media
- spectinomycin 35 μg/ml
- kanamycin 75 μg/ml
- Anhydrous tetracycline, (ANTET , ACROS organics) 0.3 μg/ml
- Isopropyl β-D-thiogalactoside (IPTG, Sigma) 0.4 mM
- PBS (Phosphate buffer solution)
- QuantaMasterTM spectrofluorimeter (Photon Technology International, Inc., Lawrenceville, New Jersey)

procedure:

E. coli BL21(DE3) cells freshly co-transformed with pET-GFP$_{1-10}$ and pTET encoding different membrane protein targets with GFP$_{11}$ fused to their C-terminus were cultured in LB media containing spectinomycin (35 μg/ml) and kanamycin (75 μg/ml). After the culture reached a cell density of OD_{600} = 0.6, the membrane protein was induced by adding 0.3 μg/ml of anhydrous tetracycline (ANTET, ACROS organics) for a given time at 30 or 37°C. After the first induction, ANTET was washed out by pelleting the cells followed by re-suspension in a pre-warmed ANTET-free LB media. Thereafter, GFP$_{1-10}$ was induced by adding 0.4 mM of isopropyl β-D-thiogalactoside (IPTG, Sigma) for 1 or 3 h at 30°C. For fluorescence measurements, cells were washed twice with PBS and resuspended in the same buffer, adjusting the cell density to OD_{600} = 0.2. Fluorescence intensity and spectra were recorded in a QuantaMasterTM

spectrofluorimeter (Photon Technology International, Inc., Lawrenceville, New Jersey) between 500 and 600 nm using an excitation wavelength of 460 nm.

7.4.2 Preparation of isolated cytoplasmic membranes and inclusion bodies.

materials and reagents:

- PBS (Phosphate buffer solution) + 45% of sucrose
- lysozyme 0.5 mg/ml
- EDTA 1 mM
- Triton X-100 1 %

procedure:

An *E. coli* cell pellet expressing a GFP_{11}-fused membrane protein and GFP_{1-10} was washed once with PBS and subjected to osmotic shock with 45 % sucrose followed by incubation with lysozyme (0.5 mg/ml) and 1 mM EDTA. Subsequently sample was briefly sonicated and subjected to centrifugation (13.000 x g, 5 min, 4°C). The resulting supernatant was ultracentrifuged (200.000 x g, 10 min, 4°C) and the pellet containing the cytoplasmic membranes was kept at -20°C until use. The pellet from the first centrifugation containing mostly inclusion bodies was washed twice with 1 % Triton X-100 in order to remove unbroken cells (Rodríguez-Carmona et al., 2010).

7.4.3 "In-gel" fluorescence and western blot visualization of GFP_{11}–$GFP_{1–10}$ fused to membrane proteins

materials and reagents:

- PBS (Phosphate buffer solution) + 45% of sucrose
- lysozyme 0.5 mg/ml

- EDTA 1 mM
- Triton X-100 1 %

procedure:

GFP fluorescence achieved after the complementation of GFP_{11}-fused membrane proteins with $GFP_{1–10}$ was visualized in a sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) gel. Briefly, isolated cytoplasmic membranes or purified inclusion bodies of *E. coli* cells co-expressing the different GFP_{11}-fused membrane proteins and $GFP_{1–10}$ were solubilized with 1% of DDM (Affymetrix, Santa, Clara, CA, USA) and subjected to SDS–PAGE. “In-gel” GFP fluorescence from the distinct membrane protein–GFP_{11}–$GFP_{1–10}$ gel bands was visualized using a GBOX gel reader (Syngene, Cambridge, UK) and a Safe Imager (Invitrogen, Carlsbad, CA, USA). Western blot analyses were performed using the HisProbe–HRP (horseradish peroxidase) kit (Thermo Scientific, Rockford, IL, USA).

7.4.4 In vivo visualization expression in bacterial colonies.

materials and reagents:

- *E. coli* BL21(DE3) co-transformed with pETGFP$_{1-10}$ and pTET with membrane protein fused to GFP_{11}
- spectinomycin 35 µg/ml
- kanamycin 75 µg/ml
- ANTET 0.3 µg/ml
- IPTG 0.4 mM
- PBS (Phosphate buffer solution)
- LB agar plates with spectinomycin 35 µg/ml and kanamycin 75 µg/ml
- (add ANTET 0.3 µg/ml or IPTG 0.5 mM, if required)
- Nitrocellulose filter membrane (Amersham Hybond-N, GE Healthcare)

procedure:

E. coli BL21(DE3) cells freshly co-transformed with pET-GFP_{1-10} and pTET encoding different membrane protein targets with GFP_{11} fused to their C-terminus were grown overnight at 37°C on nitrocellulose filter paper (Amersham Hybond-N, GE Healthcare) lying on top of a LB-agar plate containing spectinomycin and kanamycin, as previously described (Cabantous and Waldo, 2006). Expression of SteT-GFP_{11} was initiated by placing the filter paper in a new plate containing 0.3 µg/ml of ANTET for 3 h at 30°C. After the incubation, the filter paper was moved to a new LB-agar plate containing no inducing agent. Finally, GFP_{1-10} was induced by transferring the filter paper into a new LB-agar plate containing 0.4 mM of IPTG and incubating it for 3 h at 30°C. Green colonies indicating the expression of protein complemented with GFP_{1-10} were visible under either UV or blue light using a Stereo Fluorescence Microscope (Leica).

7.5 Library of mutants

In order to generate the mutants and the posterior screening the next pipeline has been followed:

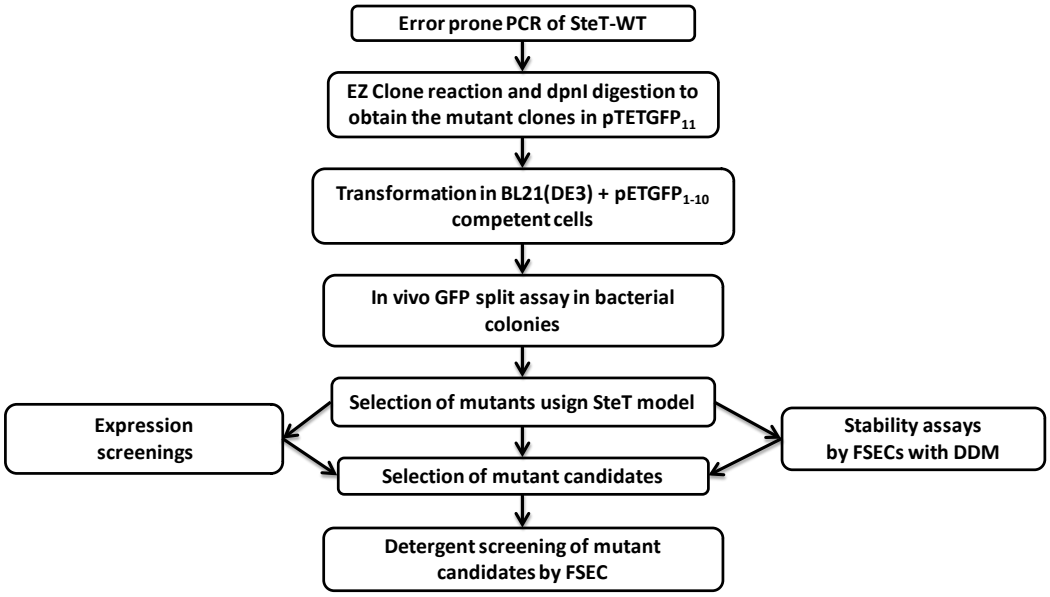


Figure 61. Pipeline for the generation of the SteT mutant random library and screenings for selection of the mutant candidates

7.5.1 Generation of Mutant library and selection of expressed mutants

materials and reagents:

- BL21(DE3) co-transformed with pETGFP$_{1-10}$ and pTET with membrane protein fused to GFP$_{11}$
- spectinomycin 35 µg/ml
- kanamycin 75 µg/ml
- ANTET 0.3 µg/ml
- IPTG 0.4 mM

- PBS (Phosphate buffer solution)
- LB agar plates with spectinomycin and kanamycin
- Nitrocellulose filter membrane (Amersham Hybond-N, GE Healthcare)

procedure:

Random mutagenesis of the cDNA encoding SteT purchased from GenScript (Piscataway, NJ, USA) was performed as described in section 7.1.2 of materials and methods. After the digestion with *dpnI*, the reaction containing the mutants in pTET was transformed in *E. coli* BL21(DE3) harboring pET encoding GFP_{1-10}. The transformation was grown overnight at 37°C in a nitrocellulose filter membrane lying on top of a LB-agar plate containing spectinomycin and kanamycin, as previously described (Cabantous, 2006). Expression of the mutants was initiated by placing the filter paper in a new plate containing 0.3 µg/ml of ANTET for 3 h at 30°C. After the incubation, the filter paper was moved to a new LB-agar plate containing no inducing agent. Finally, GFP_{1-10} was induced by transferring the filter paper into a new LB-agar plate containing 0.4 mM of IPTG and incubating it for 3 h at 30°C. Green colonies indicating the expression of mutants complemented with GFP_{1-10} were visible under either UV or blue light using a Stereo Fluorescence Microscope (Leica). After the selection of the fluorescence clones, the sequences were analyzed for the localization of the mutations on transmembrane domains. For this purpose we use a SteT model (Bartochioni et al, 2010) based on the x-ray structure of the close SteT homolog: AdiC, the L-Arginine-agmatine exchanger of *E. coli* (Fang Y et al., 2009).

7.5.2 Expression screening of mutants using the GFP split system

materials and reagents:

- *E. coli* BL21(DE3) co-transformed with pETGFP$_{1-10}$ and pTET with membrane protein fused to GFP$_{11}$
- LB media
- spectinomycin 35 µg/ml
- kanamycin 75 µg/ml
- Anhydrous tetracycline, (ANTET , ACROS organics)
- Isopropyl β-D-thiogalactoside (IPTG, Sigma) 0.4 mM
- PBS (Phosphate buffer solution)
- QuantaMasterTM spectrofluorimeter (Photon Technology International, Inc., Lawrenceville, New Jersey)

procedure:

To start the expression screening of the mutants, the temperature of incubation and the concentration for ANTET inductor were optimized for SteT-WT. In the case of ANTET, 20 ml of *E. coli* BL21 (DE3) harboring pTET-SteT-GFP$_{11}$ and pET encoding GFP$_{1-10}$, were grown and induced overnight with either 0.3 µg/ml or 0,5 µg/ml or 1 µg/ml of ANTET and 0.4 mM of IPTG at 30°C. For temperature screening 20 ml of *E. coli* BL21 (DE3) harboring pTET-SteT-GFP$_{11}$ and pET encoding GFP$_{1-10}$, were grown and induced overnight with 0.3 µg/ml of ANTET and 0.4 mM of IPTG at 20°C, 25°C, 30°C and 37°C. For the fluorescence measurements, cells were washed twice with PBS and resuspended in the same buffer, adjusting the cell density to OD$_{600}$ = 0.2. Fluorescence intensity and spectra were recorded in a QuantaMasterTM spectrofluorimeter (Photon Technology International, Inc., Lawrenceville, New Jersey) between 500 and 600 nm using an excitation wavelength of 460 nm.

For expression screening of mutants, 20 ml of *E. coli* BL21(DE3) harboring pTET-mutant or pTET-SteT and pET encoding GFP$_{1-10}$ were grown and induced with 0.4mM IPTG overnight at 25°C. Also, *E. coli* BL21(DE3) harboring or pTET-SteT

and pET encoding GFP_{1-10} were grown without and not induced overnight at 25°C. Cells were washed twice with PBS and resuspended in 1ml of the same buffer. 200 µl of each suspension was put in a microplate 96 well plate, and fluorescence was measured. OD_{600} of each suspension was measured.

Expression of each mutant was estimated as % of SteT-wt, using the next formulas, being R.F.U. (Relative Fluorescence Units):

WT expression was calculated as:

WT expression = (R.F.U. SteT-wt-GFP/O.D.$_{600}$ SteTwt-GFP.) – (R.F.U. SteT-wt /O.D.$_{600}$ SteTwt)

The formula to calculate the expression of each mutant:

Mutant expression =

[(R.F.U. SteT-mutant-GFP/O.D.$_{600}$ mutant.) – (R.F.U. SteT-wt /O.D.$_{600}$ SteTwt.)] / WT expression

7.5.3 Analysis of stability of mutants by Fluorescence Size Exclusion Chromatography (FSEC) in DDM

materials and reagents:

FSEC buffer: 20 mM Tris-Base, 150 mM NaCl and 0.05% of DDM

procedure:

For FSEC analysis, isolation of bacterial plasma membranes was performed as previously described (Newby et al., 2009). 10 mg/ml of plasma membranes were solubilized in 1% of DDM and after high-speed spin (55,000 rpm 1 h) 500 µl of supernatant were injected into Superose 6 column equilibrated with 0.05% of DDM. 200 µl fractions were collected and fluorescence associated to

each fraction was analyzed in microplate fluorimeter. Data were represented using GraphPad software.

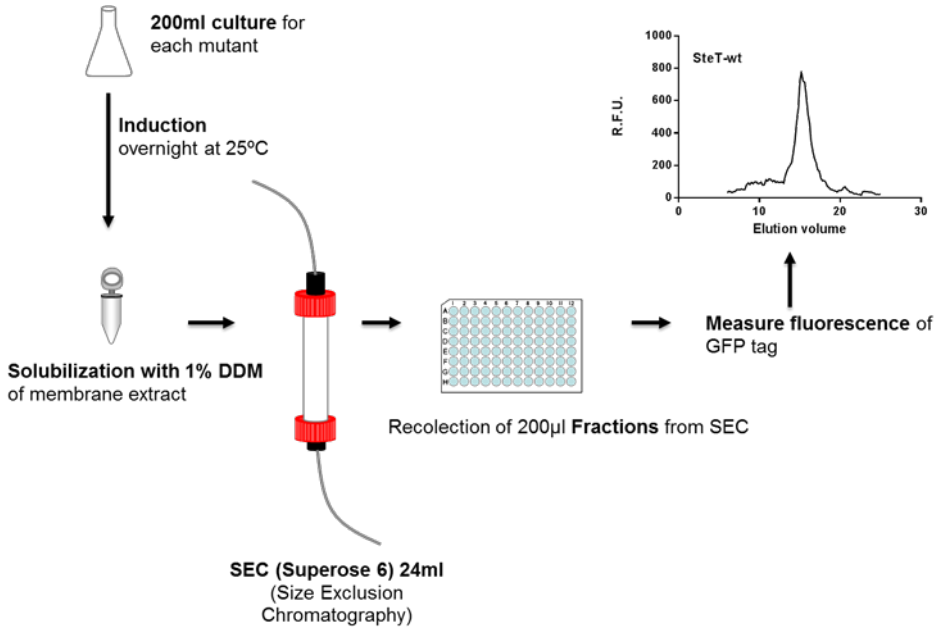


Figure 62. Scheme of FSEC used for screen the monodispersity the library of the mutants.

7.5.4 Normalization of the FSEC mutant area and % of Stability

After obtaining the FSEC profiles of mutants and WT, normalization of the profile has been performed by dividing all fluorescent values points of the FSEC by the maximum fluorescence value of the elution peak. Data was represented using Graphpad software (Figure 61).

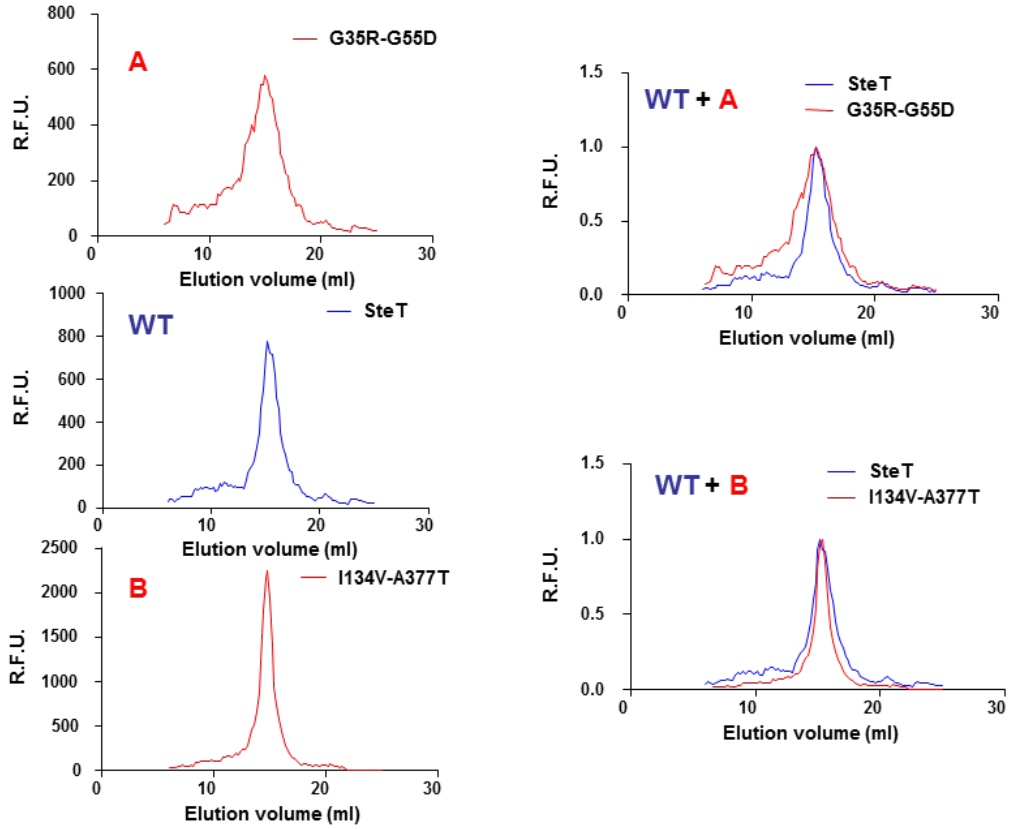


Figure 63. Normalization of FSEC profiles of mutants of SteT and its comparison with WT-SteT. (WT) FSEC profile of WT- SteT in 1% of DDM. (A) FSEC profile of mutant G53R-G55D-SteT in 1% of DDM. (B) FSEC profile of I134V-A377T-SteT in 1% of DDM. (WT+A) Normalization of mutant G35R-G55D-SteT and normalization of WT-SteT. (WT+B) Normalization of I134V-A377T-SteT and normalization of WT-SteT..

The % stability for each mutant was calculated with the next formulas:

The % of stability of WT was calculated after measure the area of WT FSEC between 13 ml and 17 ml (because the maximum of the elution peak for WT and for the mutants is expected at 15ml) with the next formula:

$$\%\ \textbf{of stabilty of SteT} - \textbf{WT} = \int_{13ml}^{17ml} \boldsymbol{Area\ Normalized\ WT\ FSEC} \Big/ \int_{13ml}^{17ml} \boldsymbol{Area\ Normalized\ WT\ FSEC}$$

In order to select the mutants with more stability than WT, the % of stability of each mutant was calculated after measure the area of each mutant FSEC between 13 ml and 17 ml:

$$\%\ \textbf{of stabilty of SteT mutant} = \int_{13ml}^{17ml} \boldsymbol{Area\ Normalized\ WT\ FSEC} \Big/ \int_{13ml}^{17ml} \boldsymbol{Area\ Normalized\ mutant\ FSEC}$$

With this approach, higher values of % of stability indicate better monodispersity than WT.

7.5.5 Detergent Screening using FSECs

materials and reagents:

- FSEC buffer: 20 mM Tris-Base (pH 8), 150 mM NaCl and 0.05% of DDM
- Solubilization buffer contains 20 mM Tris-Base, 150 mM NaCl, 1 mM Pefabloc and detergent 1%.

procedure:

Isolation of bacterial plasma membranes was performed as previously described (Newby et al., 2009). Solubilization was performed during 1 h at 4°C, in solubilization buffer. Solubilization efficiency was measured comparing the amount of target protein present in the solubilized membrane fraction before the high-speed (55,000 rpm) spin with the amount of protein present in the

supernatant after the spin. Unsolubilized material will be pelleted during the spin (Newby et al., 2009). 10 mg/ml of plasma membranes were used for solubilization in agitation for 1 h at 4°C. Solubilitation efficiencies using 1% detergent were also compared. For FSEC analysis, 10 mg/ml of plasma membranes were solubilized in 1% of particular detergents and after high-speed spin (55,000 rpm 1 h) 500 µl of supernatant were injected into Superose 6 column equilibrated with 0.05% of DDM. 200 µl fractions were collected and fluorescence associated to each fraction was analyzed in microplate fluorimeter. Data were represented using GraphPad software.

7.6 Purification and analysis of the SteT candidates L210Q-M299V and I134V-A377T

For the expression, purification, functional assays, crystallization and stability screenings the next workflow has been followed:

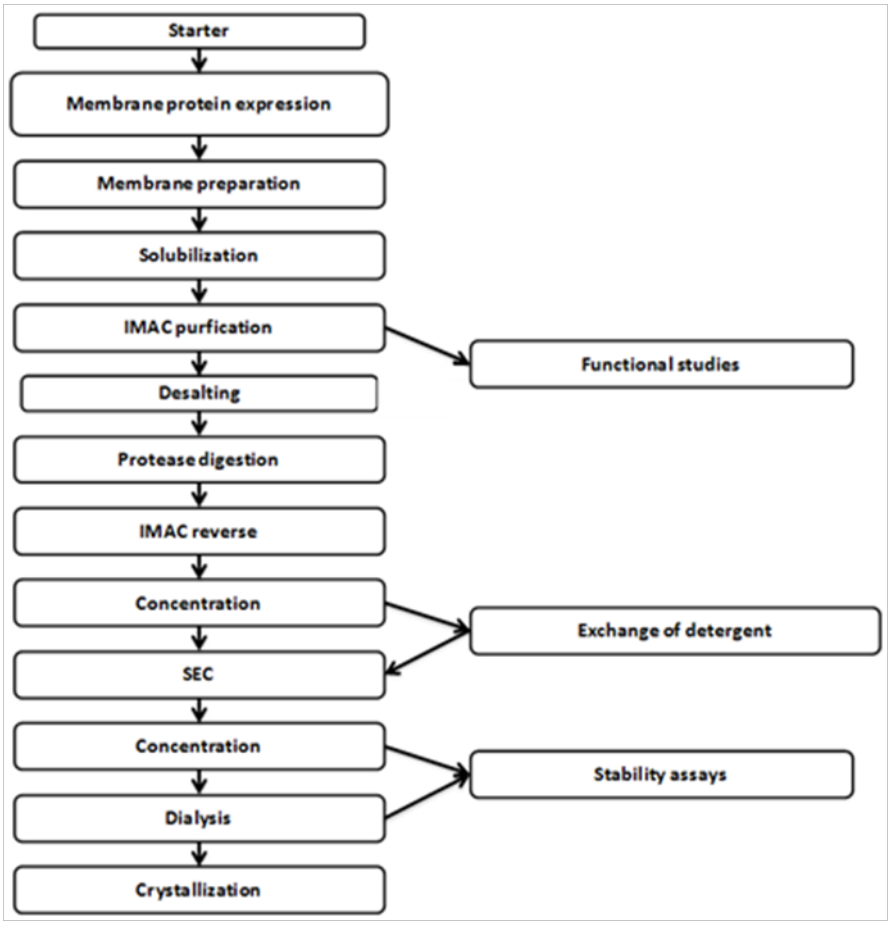


Figure 64. Workflow for purification of the mutant candidates for functional assays, detergent screening, stability assays and crystallization trials.

7.6.1 Protein expression analysis in pTTQ18-GFP vector

materials and reagents:

- PBS buffer
- LB media
- ampiciline 100 µg/ml
- IPTG

procedure:

PTTQ18-protein-GFP-His10 construction was transformed into *E. coli* BL21(DE3), *E. coli* BL21(DE3) Star and C43(DE3) strains. To determine expression levels in the different growing conditions we estimated membrane protein overexpression from whole cell using fluorescence. 3 different growing temperatures were assayed for each construction (25°C, 30°C and 37°C) during an overnight growth with LB media. These conditions were tested using 0.1 or 0.4 mM IPTG as inductor in Luria-Broth (LB) medium with Ampiciline 100 µg/ml. 15 ml cultures were grown until reaching an optical density of 0.5 and then induced with the indicated IPTG concentrations. Then, after incubation, bacterial cells were harvested by centrifugation (5000 x g, 15 min) and cell pellets were resuspended in 500 µl of PBS 1X (Drew et al., 2006) and immediately frozen in liquid nitrogen and conserved at -80°C until fluorescence measurement. GFP-associated fluorescence was measured in a microplate fluorimeter and transformation of fluorescent units to mg/L was done as reported (Drew et al., 2006).

7.6.2 Protein expression and Plasma membrane isolation

materials and reagents:

- Cell wash buffer: Mix 20 mM Tris base (pH 8) and 150 mM NaCl. The buffers were prepared at room temperature and therefore the pH of Tris is 7.4. When the experiment is performed at 4°C, the pH is 8. The buffer is filtered with a 0.2-µm sterile-filter and stored at 4°C.
- Cell lysis buffer: 20 mM Tris base (pH 8), 250–500 mM NaCl and one complete protease inhibitor tablet EDTA-free (Roche). The buffers were prepared at room temperature and therefore the pH of Tris is 7.4 at this time. When the experiment is performed at 4°C, the pH is 8. The buffer is filtered with a 0.2-µm sterile-filter and stored at 4°C.
- Membrane resuspension buffer: 20 mM Tris base (pH 8), 150 mM NaCl. The buffers were prepared at room temperature and therefore the pH of Tris is 7.4 at this time. When the experiment is performed at 4°C, the pH is 8. The buffer is filtered with a 0.2-µm sterile-filter and stored at 4°C.

procedure:

Isolation of bacterial plasma membranes was performed as follows. Erlenmeyers of 800 ml LB with ampiciline 100 µg/ml, were inoculated with 16 ml of an overnight bacterial growth (200 ml of LB bacterial growth in 1L Erlenmeyer with ampiciline (100 µg/ml) at 37°C and 220 rpm). Once an OD_{600}=0.5 was reached, induction with 0.1 mM IPTG was carried out and incubation was performed for 22 h at 37°C and 220 rpm. After this incubation period, bacterial growth was centrifuged at 5000 x g for 15 min at 4°C and resulting pellets were mixed and washed with 500 ml of ice cold washing buffer and centrifuged again at 5000 x g for 15 min at 4°C. Resulting pellet was resuspended in 200 ml of lysis buffer. Resuspended pellet was homogenized in

glass-teflon homogenizer and DNAse was added to the bacterial suspension. Afterwards, bacterial cells were broken by passing 3 times at 22 Kpsi though a cell disruptor at 4°C and the resulting suspension was centrifuged at 14,000 rpm for 1h at 4°C. Supernatant was then centrifuged at 55,000 rpm for 2h at 4°C to pellet bacterial membranes.

7.6.3 Solubilization

materials:

- Solubilization buffer: 20 mM Tris-Base (pH 8), 150 mM NaCl and 1% of detergent.

procedure:

Once plasma membranes obtained, solubilization was carried out. Unsolubilized material will be pelleted during the 55.000 rpm spin during 1h at 4°C (Newby et al., 2009). In all the cases, 3 mg/ml of plasma membranes were used for solubilization in agitation for 1 h at 4°C.

7.6.4 Protein purification: Immobilized Metal Affinity Chromatography (IMAC)

materials:

- Washing buffer: 20 mM Tris-Base (pH 8), 150 mM NaCl, 2xCMC detergent, Imidazole and with or without 10% of glycerol
- Elution buffer: 20 mM Tris-Base (pH 8), 150 mM NaCl, 2xCMC detergent and 350 mM imidazole and with or without 10% glycerol

procedure:

After solubilization step, supernatant containing solubilized plasma membrane protein has to be purified from the rest of the solubilized membrane proteins. First IMAC step consisted of an incubation of the supernatant (10 mM

imidazole added) with NiNTA (Qiagen) resin for 2 h. In this particular case, protein resulting from the solubilization of membranes from 6.4 L was incubated with 5 ml of resin (10 ml of total resin including ethanol). After incubation, 5 ml of resin were washed with 50 ml of washing buffer with 10mM of Imidazole, 50 ml of additional washing buffer with 20 mM Imidazole and 50 ml of additional washing buffer with 40 mM Imidazole. Elution was carried out incubating resin with Elution buffer for 30 min. All this process was carried out at 4°C.

7.6.5 Desalting

materials and reagents:

SEC buffer: 20 mM Tris-Base (pH 8), 150 mM NaCl and 2xCMC detergent.

procedure:

Eluted protein was desalted in a PD-10 desalting column following the manufacturer´s protocol. Desalting column was equilibrated in using the SEC buffer.

7.6.6 Protease digestion

materials and reagents:

- Protease buffer: 20 mM Tris-Base, 150 mM NaCl, 2xCMC detergent, 0.5 mM EDTA and 1 mM DTT.

procedure:

Resulting desalted protein was digested with HVR 3C protease incubating with agitation 1:20 (protease:protein) in protease buffer, during 22h at 4°C.

7.6.7 IMAC reverse

After protease digestion, His10-tagged GFP as well as His10-tagged non-digested protein has to be removed from protein solution. To do this, a reverse

IMAC was carried out incubating supernatant with NiNTA (Qiagen) resin for 2 h. All this process was carried out at 4°C.

7.6.8 Concentration of the protein and Size Exclusion Chromatography (SEC)

materials and reagents:

- SEC buffer: 20 mM Tris-Base, 150 mM NaCl, 10% glycerol when OG was used as detergent and 2xCMC detergent
- Disalysis buffer: 20 mM Tris-Base, 150 mM NaCl and 2xCMC detergent

procedure:

After IMAC reverse, purified digested protein was eluted with the flowthrough and concentrated in a 100 MWCO concentrator device (Centricon) VIVASPIN. In the case to study the effect of different detergents on protein behaviour, DM detergent was changed during concentration and in Size Exclusion Chromatography (SEC) by the SEC buffer. Before injecting, aggregates were eliminated by high-spin centrifugation (55,000 rpm, 30 min) and the protein was injected into a Superdex 200 column. Superdex 200 5/150 column and Superdex 200 10/300 GL were used for protein purification and detergent screening.

7.6.9 Dialysis

materials:

- Disalysis buffer: 20 mM Tris-Base (pH 8), 150 mM NaCl and 2xCMC detergent

procedure:

Dialysis was perform incubating purified protein with agitation during 24 h 1:100 v/v (protein/dialysis buffer) using dialysis membrane of 50 MWCO. After

this period, protein was recovered from dialysis membranes and centrifuged at 55,000 rpm for 30 min at 4°C to eliminate aggregates.

7.6.10 Stability detergent assays

Purified protein from SEC was concentrated to desired concentration and the protein concentration of supernatant was measured after centrifugation at 55.000 rpm at 4°C during 30 min. Incubation at 4°C from 1 day to 1 week of the supernatant was performed depending on the experiment. After the incubation at 4°C, protein concentration was measured again, after centrifugation at 55.000 rpm at 4°C during 30 min. The difference of protein concentration, after and before the incubation at 4°C was calculated.

7.7 Transport experiments

7.7.1 Reconstitution of SteT and L210Q-M299V and I134V-A377T into Proteoliposomes

materials:

- *E. coli* polar lipid (Avanti Polar Lipids) 50 mg/ml
- Disalysis buffer: 120 mM KPi, pH 7.4, 0.5 mM EDTA, 1 mM MgSO4, 5 mM TrisSO4, 1% glycerol, and a 4 mM concentration of L-Ser (unless otherwise indicated)
- 1.25% β-d-octyl glucoside (Roche Applied Science)

procedure:

E. coli polar lipid extract solubilized in chloroform (50 mg/ml) was dried under a stream of nitrogen to remove the solvent and to obtain a thin layer of dry lipids in a glass tube. The dried lipids were resuspended in dialysis buffer to yield a final lipid concentration of 40 mg/ml. After four 30-s sonication and vortexing cycles, the liposomes were extruded in a LiposoFast-Pneumatic Actuator

(Avestin) through a 400-nm polycarbonate filter (Avestin) to obtain unilamellar vesicles of homogeneous size. Liposomes were mixed with purified protein at a 1:100 protein/lipid ratio (w/w). To destabilize the liposomes, 1.25% β-d-octyl glucoside was added and incubated on ice with occasional agitation for 5 min. DDM and β-d-octyl glucoside were removed by dialysis for 40 h at 4°C against 100 volumes of dialysis buffer. Finally, proteoliposomes were ultracentrifuged (100,000 × g, 1 h at 4°C), and the pellet was resuspended in one-third of the initial volume of dialysis buffer without amino acids.

7.7.2 Transport Measurements

materials and reagents:

- Transport buffer: 150 mM choline chloride, 10 mM Tris-HEPES, pH 7.4, 1 mM $MgCl_2$, 1 mM CaCl2, 0.5 µCi of radiolabeled l-amino acid, and unlabeled amino acid to the desired final concentration
- Disalysis buffer: 120 mm KPi, pH 7.4, 0.5 mm EDTA, 1 mm $MgSO_4$, 5 mm $TrisSO_4$, 1% glycerol, and a 4 mM concentration of L-Ser (unless otherwise indicated)
- Stop buffer: 150 mM choline chloride, 10 mM Tris-HEPES (pH 7.4), and 5 mM L-serine
- L-[3H]serine. (American Radiolabeled Chemicals)

procedure:

Influx measurements in proteoliposomes were made as described (Reig et al., 2002) with minor changes. Cold proteoliposomes (10 µl) in dialysis buffer were mixed with 180 µl of transport buffer and incubated at room temperature for different periods of time. Reactions were stopped by the addition of 850 µl of ice-cold stop buffer and filtration through membrane filters (Sartorius; 0.45-µm pore size). Filters were then washed three times with 2 ml of stop buffer and dried, and the trapped radioactivity was counted. All experimental values were

corrected by subtracting zero time values obtained by adding the stop solution before the proteoliposomes into the transport buffer. Protein concentration in the proteoliposomes was determined using the Amido Black protein assay (Schaffner and Weissmann, 1973), and the transport was expressed as pmoles/µg of protein per unit of time and reported as the mean ± S.E.

7.8 Crystallization screenings in microplates

materials and reagents:

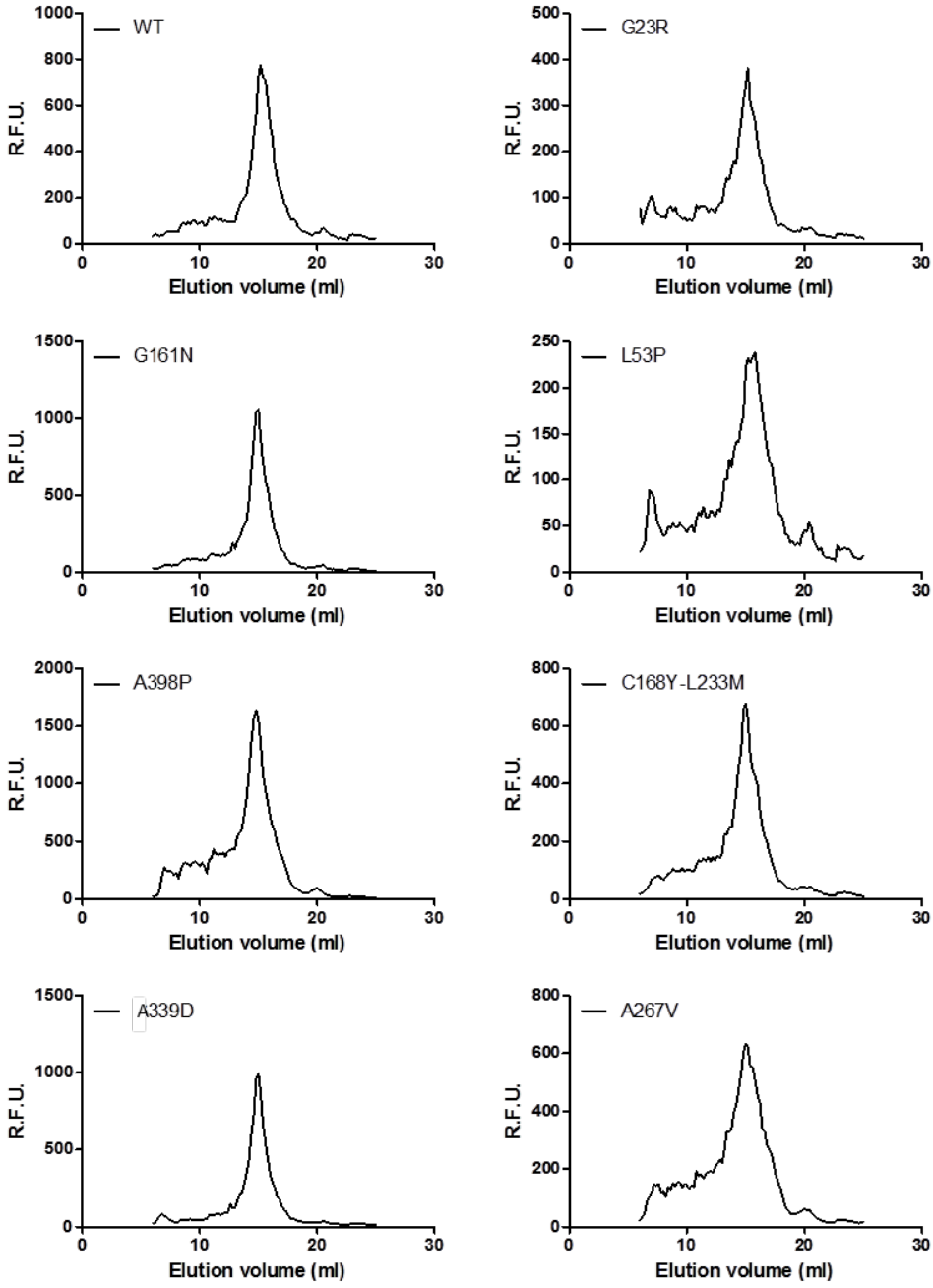
- SEC buffer: 20 mM Tris-Base (pH 8), 150 mM NaCl, 2xCMC detergent. 10% glycerol was added when OG was used as detergent.
- Disalysis buffer: 20 mM Tris-Base (pH 8), 150 mM NaCl and 2xCMC detergent

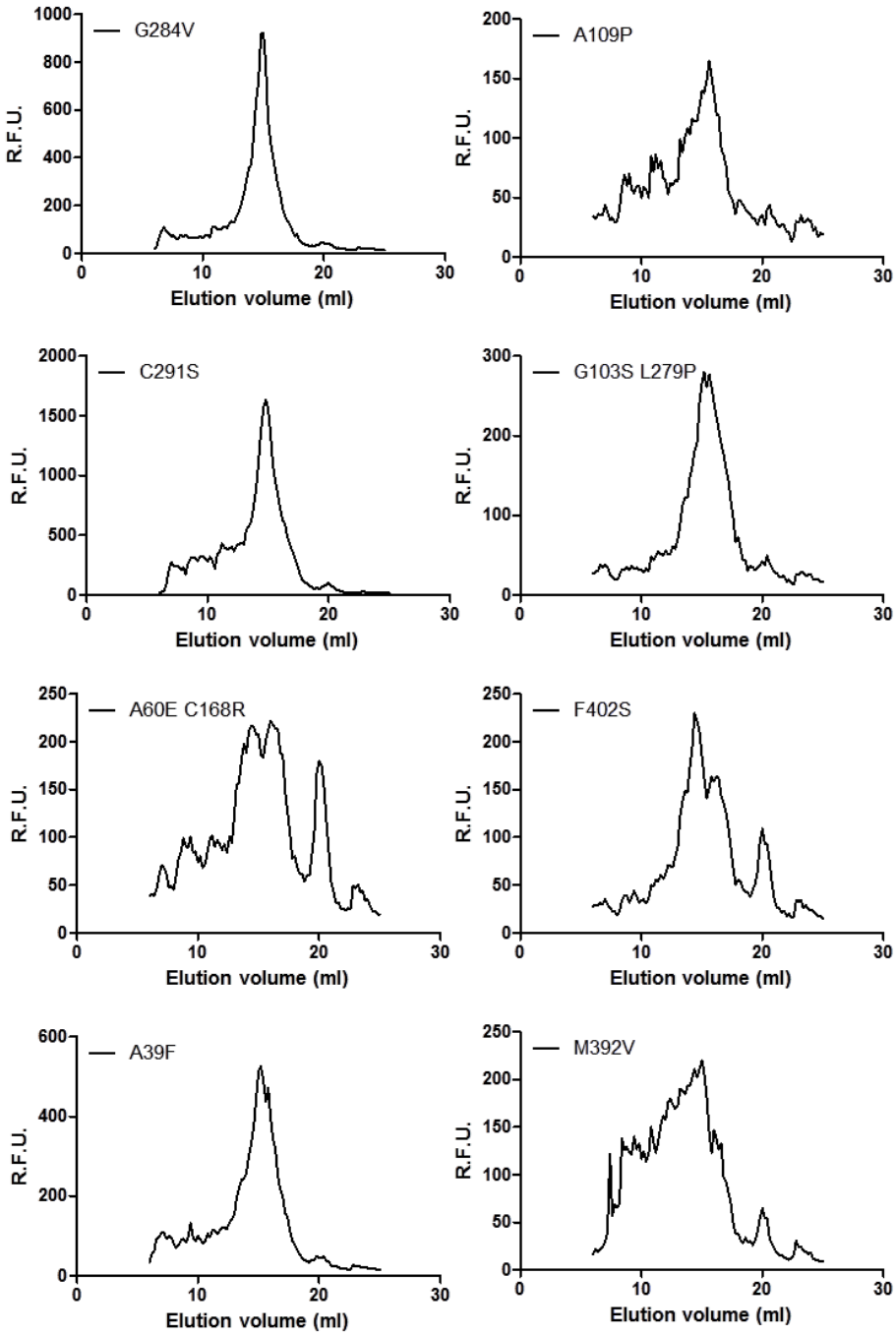
procedures:

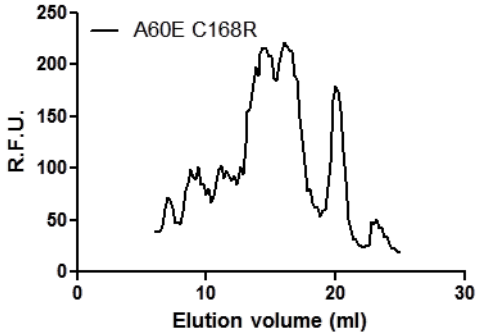
Superdex 200 10/300 GL was used for protein purification for crystallization assays. Superdex 200 buffer consists of 20 mM Tris-Base, 150 mM NaCl, 10% glycerol and 2xCMC detergent. For crystallization assays, after SEC analysis, peaks were collected, concentrated until ~4 or ~10 mg/ml and dialyzed 24 h in the case of DM and Cymal-6. After this period, protein was recovered from dialysis membranes and centrifuged at 55,000 rpm for 30 min at 4°C to eliminate aggregates. Initial crystallization conditions were obtained with Membfac, Memplus, Memstart, Memsys and Memgold crystallization screening conditions (Hampton Research) using sitting-drop in 96 well microplate at 20°C or 4°C. 15 µl of protein were used for each well of the microplates. Photos were taken using Leica lupe microscope. In the case of the screening with Cymal-6 programmed photos of the microplates wells were taken by crystall farm.

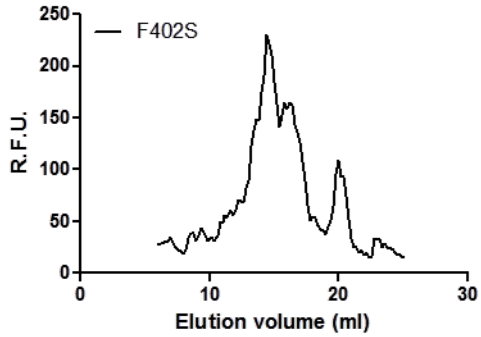
APPENDIX I

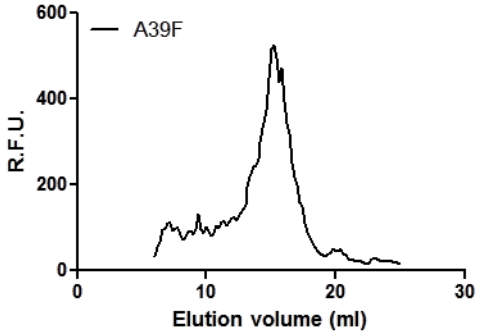
APPENDIX I FSEC profiles of the 70 random mutants of STeT analyzed

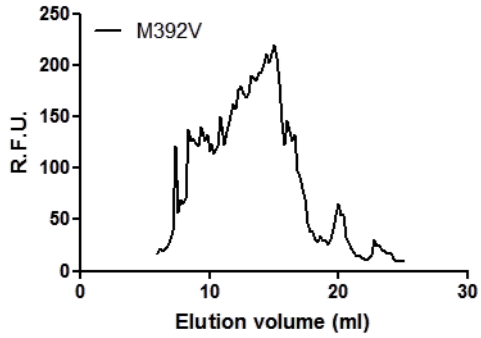


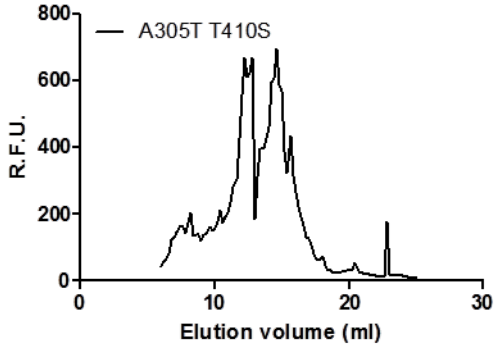

G284V
R.F.U.
Elution volume (ml)
A109P
R.F.U.
Elution volume (ml)
C291S
R.F.U.
Elution volume (ml)
G103S L279P
R.F.U.
Elution volume (ml)
A60E C168R
R.F.U.
Elution volume (ml)
F402S
R.F.U.
Elution volume (ml)
A39F
R.F.U.
Elution volume (ml)
M392V
R.F.U.
Elution volume (ml)

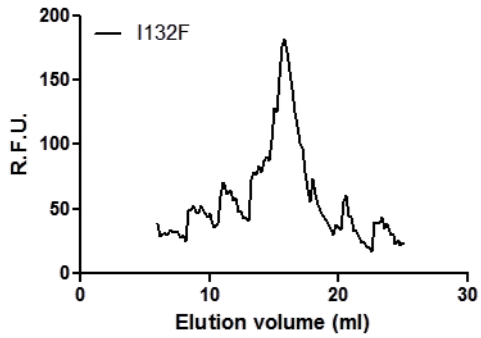

A60E C168R
R.F.U.
Elution volume (ml)

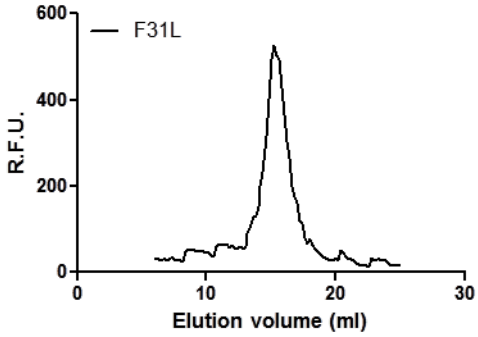

F402S
R.F.U.
Elution volume (ml)

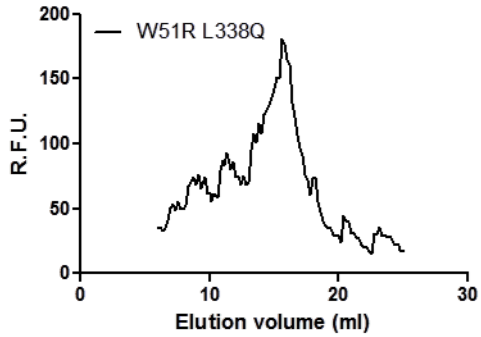

A39F
R.F.U.
Elution volume (ml)

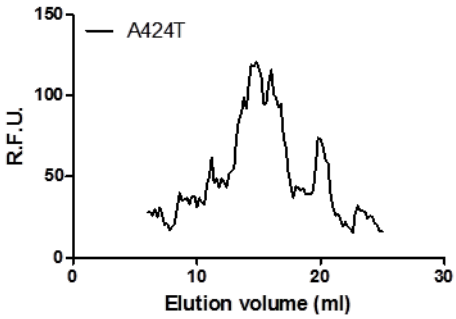

M392V
R.F.U.
Elution volume (ml)

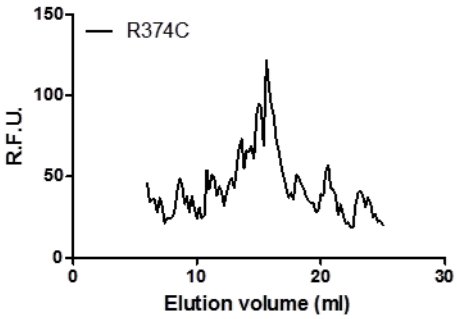

A305T T410S
R.F.U.
Elution volume (ml)

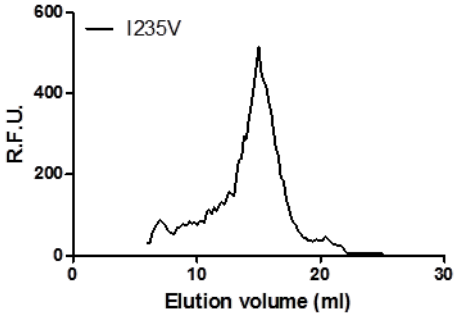

I132F
R.F.U.
Elution volume (ml)

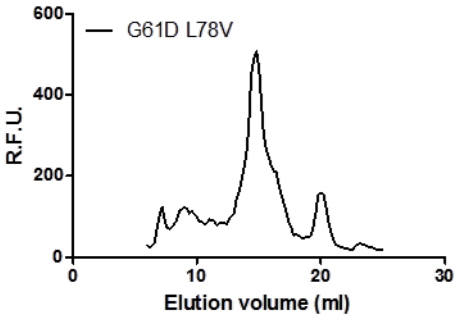

F31L
R.F.U.
Elution volume (ml)

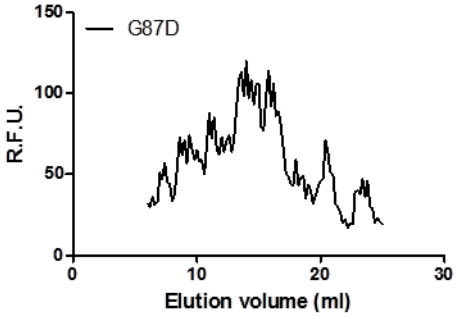

W51R L338Q
R.F.U.
Elution volume (ml)

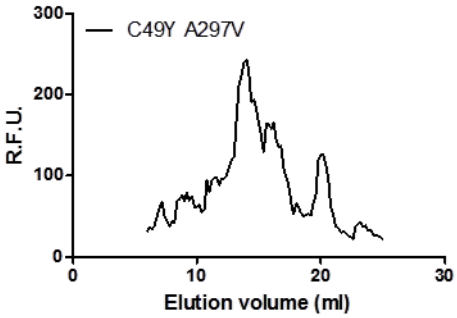

A424T
R.F.U.
150
100
50
0
0
10
20
30
Elution volume (ml)

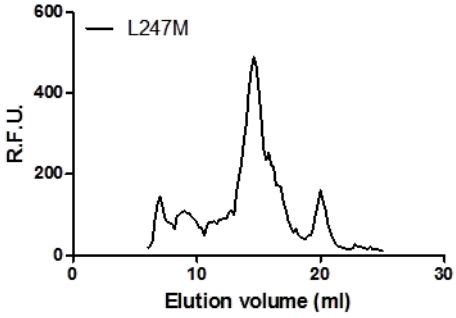

R374C
R.F.U.
150
100
50
0
0
10
20
30
Elution volume (ml)

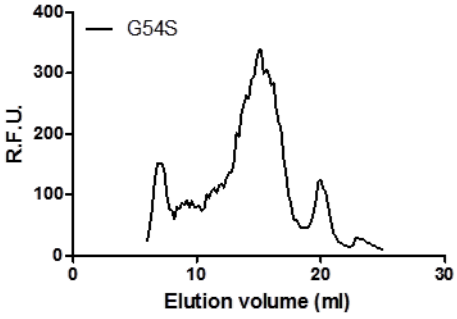

I235V
R.F.U.
600
400
200
0
0
10
20
30
Elution volume (ml)

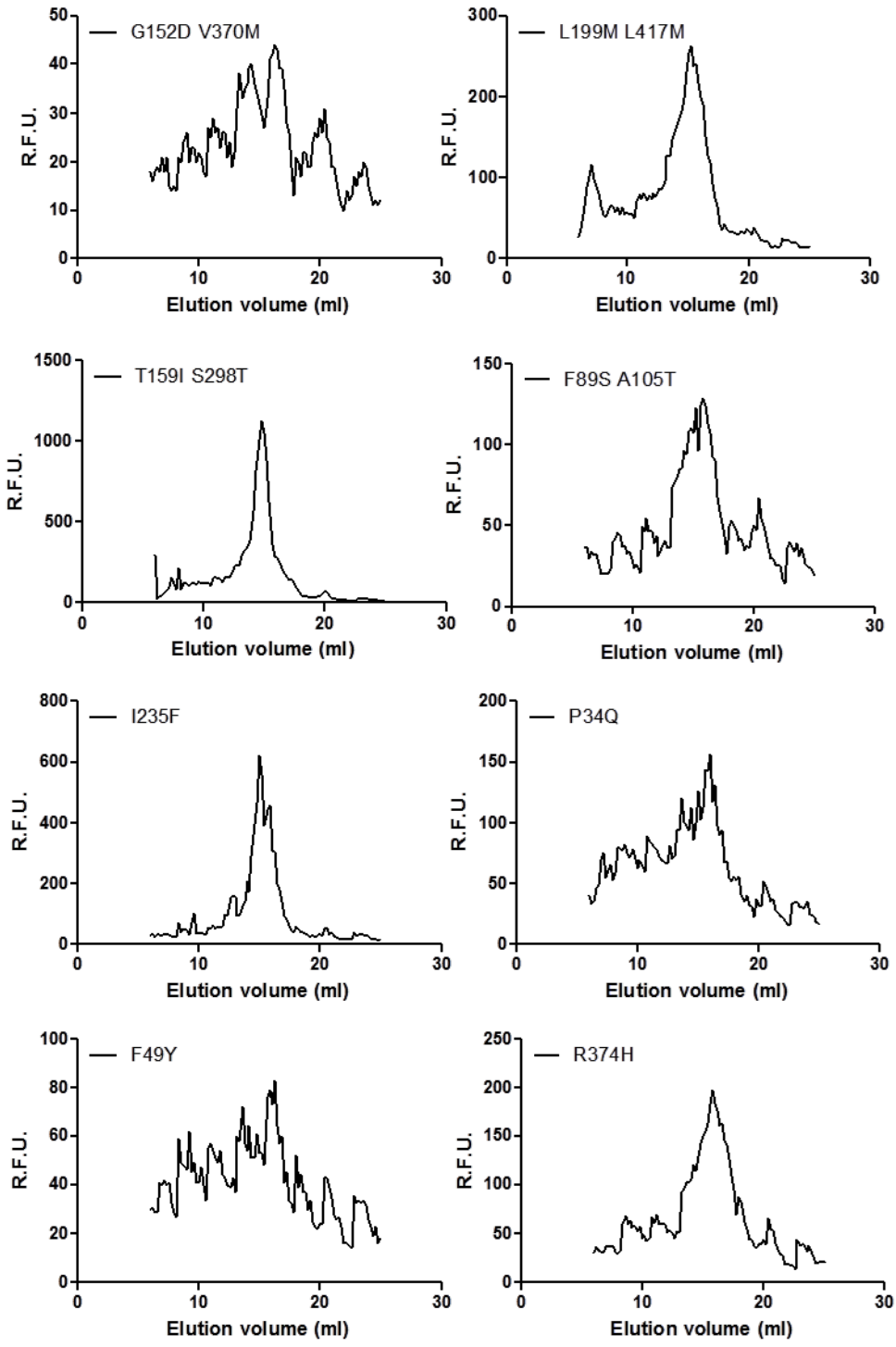

G61D L78V
R.F.U.
600
400
200
0
0
10
20
30
Elution volume (ml)

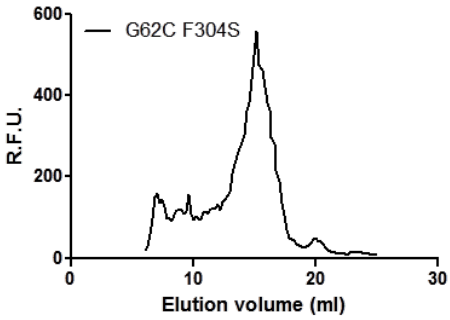

G87D
R.F.U.
150
100
50
0
0
10
20
30
Elution volume (ml)

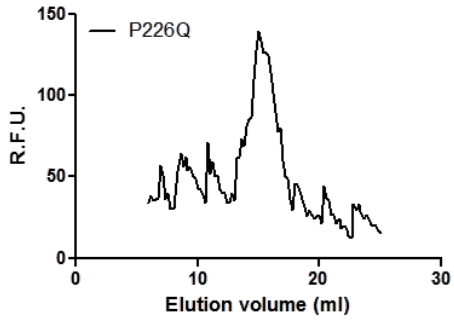

C49Y A297V
R.F.U.
300
200
100
0
0
10
20
30
Elution volume (ml)

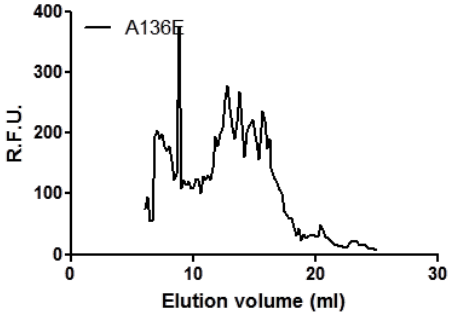

L247M
R.F.U.
600
400
200
0
0
10
20
30
Elution volume (ml)

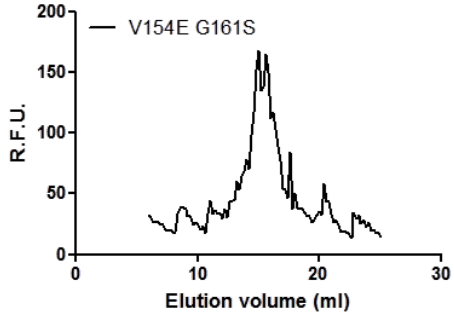

G54S
R.F.U.
400
300
200
100
0
0
10
20
30
Elution volume (ml)

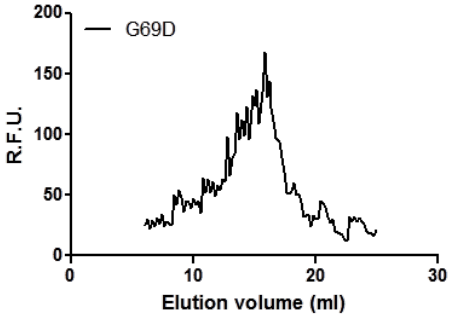

G152D V370M
R.F.U.
Elution volume (ml)
L199M L417M
R.F.U.
Elution volume (ml)
T159I S298T
R.F.U.
Elution volume (ml)
F89S A105T
R.F.U.
Elution volume (ml)
I235F
R.F.U.
Elution volume (ml)
P34Q
R.F.U.
Elution volume (ml)
F49Y
R.F.U.
Elution volume (ml)
R374H
R.F.U.
Elution volume (ml)

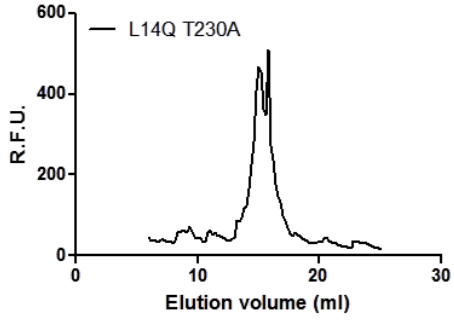

G62C F304S
R.F.U.
Elution volume (ml)

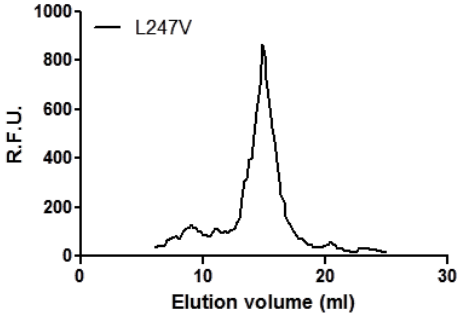

P226Q
R.F.U.
Elution volume (ml)

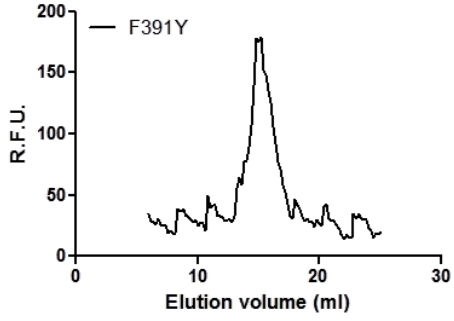

A136E
R.F.U.
Elution volume (ml)

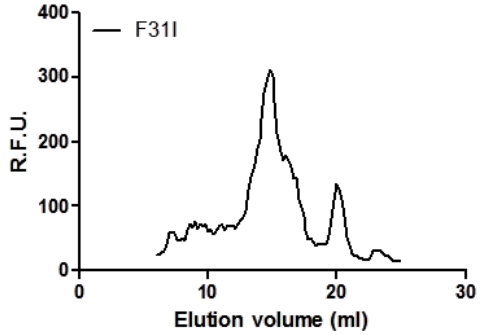

V154E G161S
R.F.U.
Elution volume (ml)

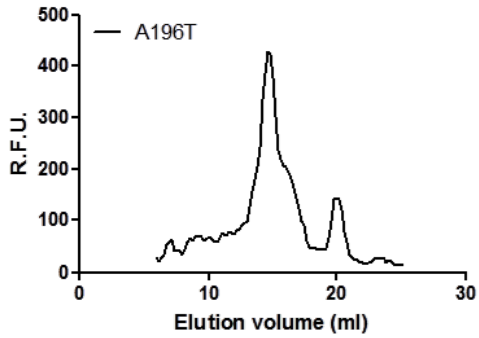

G69D
R.F.U.
Elution volume (ml)

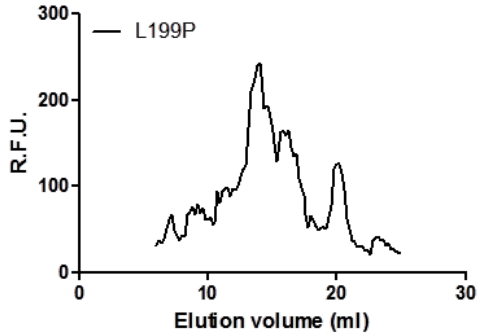

L14Q T230A
R.F.U.
Elution volume (ml)

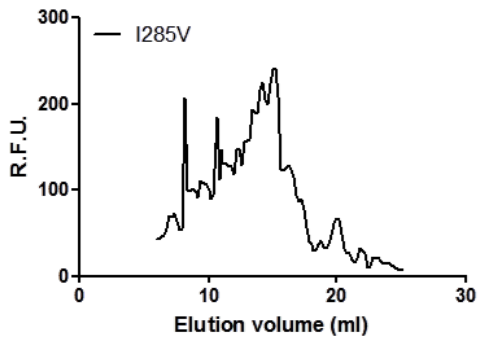

L247V
R.F.U.
Elution volume (ml)

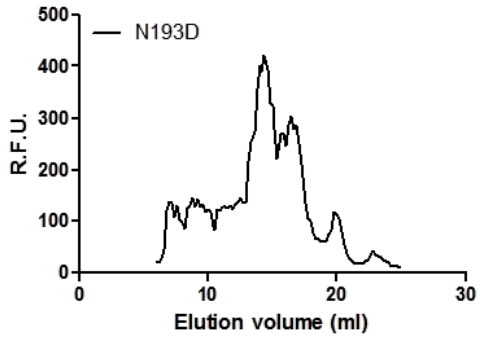

F391Y
R.F.U.
Elution volume (ml)

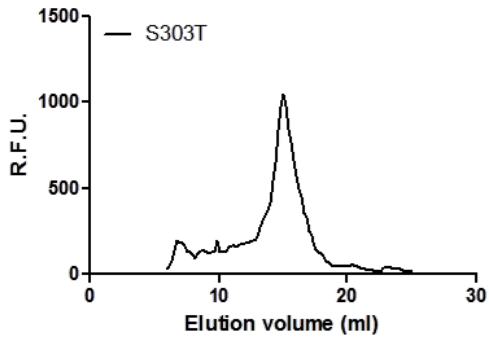

F31I
R.F.U.
Elution volume (ml)

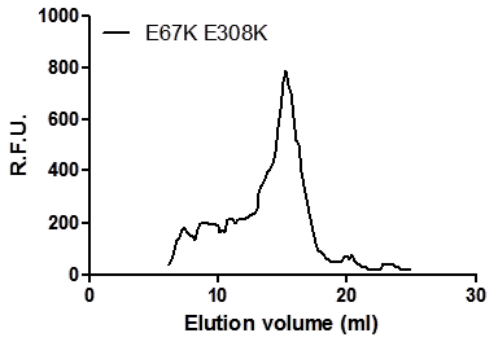

A196T
R.F.U.
Elution volume (ml)

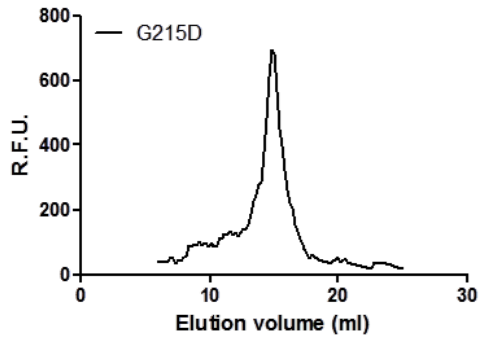

L199P
R.F.U.
Elution volume (ml)

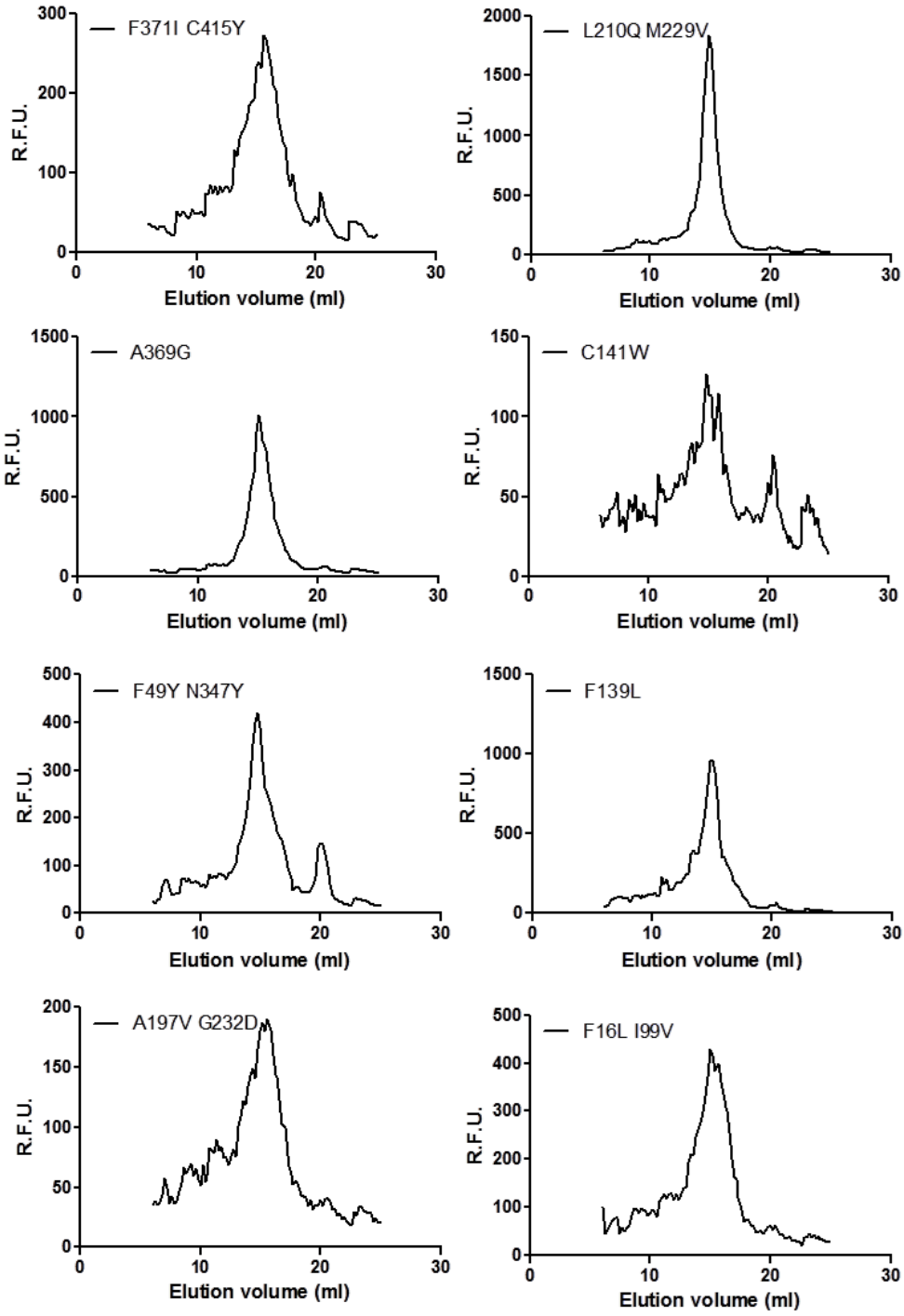

I285V
R.F.U.
Elution volume (ml)

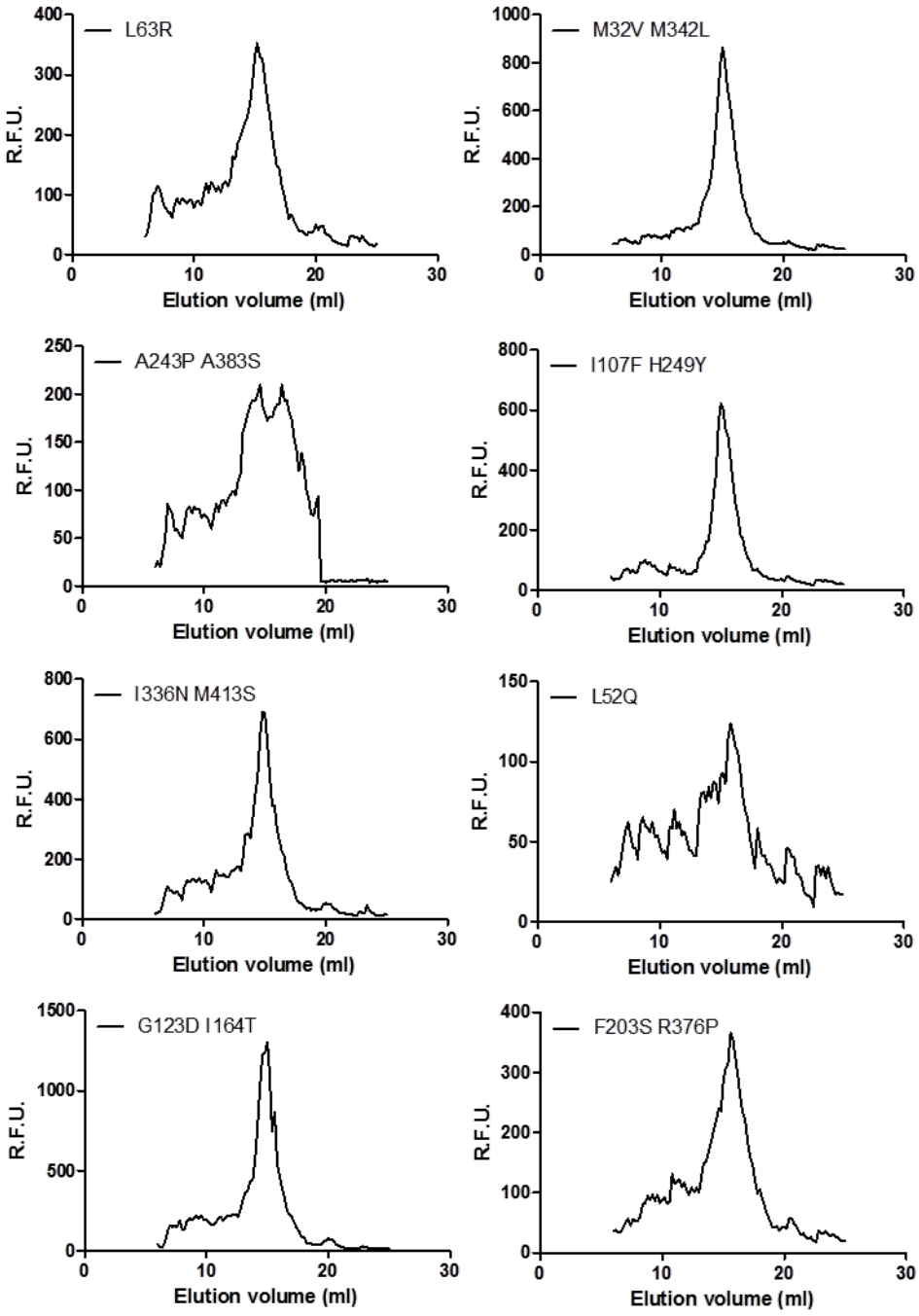

N193D
R.F.U.
Elution volume (ml)

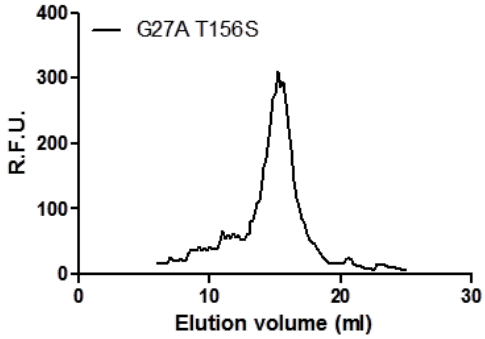

S303T
R.F.U.
Elution volume (ml)

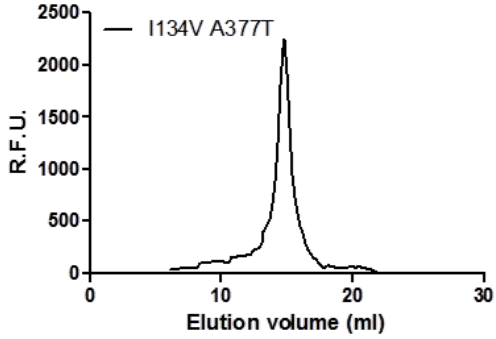

E67K E308K
R.F.U.
Elution volume (ml)

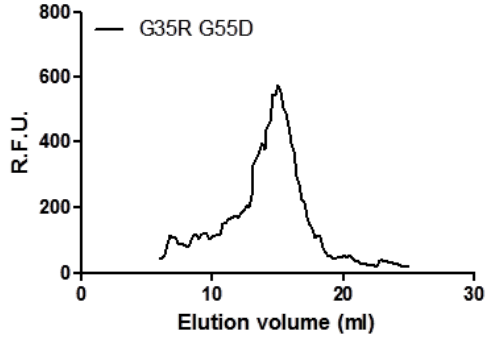

G215D
R.F.U.
Elution volume (ml)


F371I C415Y
R.F.U.
Elution volume (ml)
L210Q M229V
R.F.U.
Elution volume (ml)
A369G
R.F.U.
Elution volume (ml)
C141W
R.F.U.
Elution volume (ml)
F49Y N347Y
R.F.U.
Elution volume (ml)
F139L
R.F.U.
Elution volume (ml)
A197V G232D
R.F.U.
Elution volume (ml)
F16L I99V
R.F.U.
Elution volume (ml)


L63R
R.F.U.
Elution volume (ml)
M32V M342L
R.F.U.
Elution volume (ml)
A243P A383S
R.F.U.
Elution volume (ml)
I107F H249Y
R.F.U.
Elution volume (ml)
I336N M413S
R.F.U.
Elution volume (ml)
L52Q
R.F.U.
Elution volume (ml)
G123D I164T
R.F.U.
Elution volume (ml)
F203S R376P
R.F.U.
Elution volume (ml)


G27A T156S
R.F.U.
400
300
200
100
0
0
10
20
30
Elution volume (ml)


I134V A377T
R.F.U.
2500
2000
1500
1000
500
0
0
10
20
30
Elution volume (ml)


G35R G55D
R.F.U.
800
600
400
200
0
0
10
20
30
Elution volume (ml)

APENDIX II

ARTICLE: Rodríguez-Banqueri et al., Analytical Biochemistry, 2012

Assessment of membrane protein expression and stability using a split green fluorescent protein reporter

Analytical Biochemistry Vol. 423, pp.7-14, April 2012

Arturo Rodríguez-Banqueri, Lukasz Kowalczyk, Manuel Palacín, José Luis Vázquez-Ibar

Analytical Biochemistry 423 (2012) 7–14

Contents lists available at SciVerse ScienceDirect

Analytical Biochemistry

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

Assessment of membrane protein expression and stability using a split green fluorescent protein reporter

Arturo Rodríguez-Banqueri [a,1], Lukasz Kowalczyk [a,1], Manuel Palacín [a,b,c], José Luis Vázquez-Ibar [a,d,*]

[a] Institute for Research in Biomedicine (IRB Barcelona), 08028 Barcelona, Spain
[b] Center for Biomedical Research on Rare Diseases (CIBERER), 08025 Barcelona, Spain
[c] Department of Biochemistry and Molecular Biology, Faculty of Biology, Universitat de Barcelona, 08028 Barcelona, Spain
[d] Catalan Institute for Advanced Studies and Research (ICREA), 08010 Barcelona, Spain

ARTICLE INFO

Article history:
Received 10 October 2011
Received in revised form 28 December 2011
Accepted 29 December 2011
Available online 8 January 2012

Keywords:
Membrane protein
Protein crystallography
GFP
SteT

ABSTRACT

Membrane proteins are challenging targets for structural biologists. Finding optimal candidates for such studies requires extensive and laborious screening of protein expression and/or stability in detergent. The use of green fluorescent protein (GFP) as a reporter has enormously facilitated these studies; however, its 238 residues can potentially alter the intrinsic properties of the target (e.g., expression or stability). With the aim of minimizing undesired effects of full-length GFP, here we describe the utility of a split GFP reporter during precrystallization studies of membrane proteins. GFP fluorescence appeared by complementation of the first 15 residues of GFP (GFP_{11}) (fused to the C terminus of a membrane protein target) with the remaining nonfluorescent GFP (GFP_{1-10}). The signal obtained after sequential expression of SteT (L-serine/L-threonine exchanger of *Bacillus subtilis*) fused to GFP_{11} followed by GFP_{1-10} specifically measured the protein fraction inserted into the *Escherichia coli* cytoplasmic membrane, thereby discarding protein aggregates confined as inclusion bodies. Furthermore, in vitro complementation of purified SteT–GFP_{11} with purified GFP_{1-10} was exploited to rapidly assess the stability of wild-type and G294V mutant versions of SteT–GFP_{11} following detergent solubilization and purification. This method can be applied in a medium- to high-throughput manner with multiple samples.

Approximately 30% of the human genome encodes membrane proteins [1]. These perform critical functions maintaining cell homeostasis by transferring information between the extracellular and intracellular sides of the cytoplasmic membrane or the cytoplasm and lumen of intracellular compartments. Genetic defects that affect the expression and/or functionality of many membrane proteins are the direct cause of severe pathologies. In addition, tissue distribution, organ-specific entry, and clearance of drugs are often facilitated or hindered by the expression of membrane transporters [2–4].

Unfortunately, despite the important role of many membrane proteins in human pathophysiology, high-resolution structural information is still lacking relative to soluble protein counterparts. Specifically, although there are more than 70,000 protein structures deposited in the Protein Data Bank, only 862 of these represent membrane proteins (302 unique proteins). Although the hydrophobic nature and poor stability of membrane proteins in solution present formidable obstacles for structural studies, the initial bottleneck arises from the difficulty in obtaining the milligram amounts of recombinant functional membrane protein necessary for crystallography or nuclear magnetic resonance studies using a heterologous expression system (e.g., the bacterium *Escherichia coli*, the most widely used host for protein overexpression) [5]. Normally, overexpression of membrane proteins results in lower yields than those obtained for soluble proteins. In addition, the heterologous expression of membrane proteins frequently results in protein aggregation into inclusion bodies as a consequence of incorrect folding. In *E. coli*, this problem has been attributed to saturation of the Sec machinery used for membrane proteins during biogenesis and insertion into the cytoplasmic membrane [6]. Consequently, for the purpose of protein purification for either functional or crystallization studies, a time-consuming screening process is required to identify optimal candidates with a reasonable expression yield and acceptable stability after detergent solubilization.

A generalized strategy in the membrane protein structure–function field consists of working in parallel with various homologs of a selected membrane protein target [7]. In this regard, bacterial homologs have proven to be excellent structural and functional paradigms of mammalian membrane proteins [8], particularly those from thermophilic organisms. In addition, variants of a select

* Corresponding author at: Institute for Research in Biomedicine (IRB Barcelona), 08028 Barcelona, Spain.
E-mail address: jl.vazquez.ibar@gmail.com (J.L. Vázquez-Ibar).
[1] These authors contributed equally to this work.

0003-2697/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.ab.2011.12.044

target (e.g., C-terminal and/or N-terminal modifications or single-point mutations) are routinely expressed in multiple expression vectors to identify a combination increasing expression yield and/or stability. Therefore, a robust protocol to test the protein expression and stability of multiple samples in a fast and reliable manner is extremely beneficial in structural studies [9].

Green fluorescent protein (GFP)[2] fused to the intracellular C terminus of membrane protein targets is a sensitive reporter that has enormously facilitated precrystallization screening [10–13]. GFP fluorescence is directly related to protein expression, and when GFP fluorescence is combined with size exclusion chromatography (fluorescence size exclusion chromatography, FSEC), it results in a powerful tool to characterize protein stability under various conditions (e.g., homogeneity in different detergent solutions) using a very small amount of sample [11]. Despite the popularity of this tool, GFP is a robust β-barrel protein containing 238 amino acids [14] that in some cases can interfere with the expression and/or stability of the membrane protein target [15]. In this context, recent data reveal that the solution behavior of a GFP-fused membrane protein changes dramatically after removing the GFP [16]. Indeed, similar observations were made previously with soluble proteins. Waldo and coworkers addressed this problem by successfully developing a GFP complementation assay to screen the solubility of globular proteins expressed in *E. coli* [17,18]. In this method, a 15-amino-acid fragment of an engineered "superfolder" GFP [19] (GFP_{11}) is expressed following the C terminus of the protein of interest [17]. If the protein is stable and does not aggregate, the GFP_{11} fragment will complement with the remaining nonfluorescent 215-amino-acid fragment of GFP (GFP_{1-10}) independently expressed in the same cell. As a result, this complex emits GFP fluorescence and the method minimizes the effect of the GFP tag on the intrinsic properties of the protein under study.

Here we adapted and explored the benefits of this GFP complementation assay for precrystallization screening of membrane proteins. Similar to the method described by Waldo and coworkers, our approach consists of sequentially coexpressing a membrane protein target (e.g., the L-serine/L-threonine exchanger from *Bacillus subtilis,* SteT) [20] fused on its C terminus with GFP_{11} (SteT–GFP_{11}) followed by the GFP_{1-10} fragment (Fig. 1) in *E. coli.* Importantly, GFP fluorescence appeared only when SteT–GFP_{11} was expressed and inserted into the cytoplasmic membrane. In contrast, no fluorescence was detected when SteT–GFP11 aggregated into inclusion bodies. We demonstrate the general applicability of this approach by testing four other unrelated membrane proteins with distinct topologies and quaternary structures. GFP fluorescence can be observed after in vitro GFP_{11}–GFP_{1-10} complementation even in the presence of detergents in the buffer. We exploited this property to build a rapid assay to assess the stability of membrane proteins after detergent solubilization and purification. Our results indicate that this GFP complementation strategy is a fast, sensitive, and reliable tool that facilitates the structural study of membrane proteins.

Materials and methods

Cloning membrane proteins fused to GFP_{11}

pTET and pET vectors encoding GFP_{11} (GFP residues 1–16) and GFP_{1-10} (GFP residues 16–230) [18], respectively, were generously provided by Geoffrey S. Waldo (Los Alamos National Laboratory, Los Alamos, NM, USA). The complementary DNAs (cDNAs) encoding SteT–$(His)_6$, SteT–G294V–$(His)_6$, EmrE–$(His)_6$, LacY, MscL, and MscS and mutated versions were cloned into the *Nco*I and *Bam*HI sites of the pTET–GFP_{11} vector to generate the corresponding C-terminal GFP_{11} fusion of each protein. In the SteT and EmrE constructs, a hexahistidine tag is present between the C terminus of the membrane protein and the N terminus of GFP_{11}. All constructs were verified by DNA sequencing.

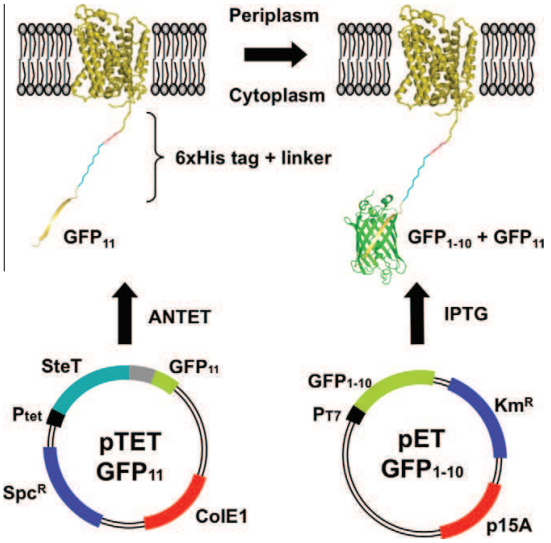


Fig.1. Schematic of the split GFP system adapted to membrane proteins. The split GFP system consists of two plasmids: pTET–GFP_{11} and pET–GFP_{1-10} [17]. pTET carries the gene encoding the target membrane protein (SteT) fused to a small part (15 amino acids) of GFP (GFP_{11}) at its C terminus, and pET carries the gene encoding the rest of the GFP molecule (GFP_{1-10}, 215 amino acids). Plasmids are compatible, containing the ColE1 and the p15A origins of replication, respectively. They also encode two antibiotic resistance genes: spectinomycin (Spc^R) in pTET and kanamycin (Km^R) in pET. Protein expression is controlled by two promoters: P_{tet} (ANTET inducible) in pTET and P_{T7} (IPTG inducible) in pET. The expression of these genes can be induced simultaneously or sequentially, and complementation occurs when the GFP_{11}-fused membrane protein is expressed and inserted into the cytoplasmic membrane of *E. coli*.

In vivo fluorescence screening assay in bacterial cultures

E. coli BL21(DE3) cells freshly cotransformed with pET–GFP_{1-10} and pTET encoding different membrane protein targets with GFP_{11} fused to their C terminus were cultured in Luria–Bertani (LB) medium containing spectinomycin (35 μg/ml) and kanamycin (75 μg/ml). After the culture reached a cell density of $A_{600} \sim 0.6$, the membrane protein was induced by adding 0.3 mg/ml anhydrous tetracycline (ANTET, Acros Organics, Morris Plains, NJ, USA) for a given time at 30 or 37 °C. After the first induction, ANTET was washed out by pelleting the cells followed by resuspension in a prewarmed ANTET-free LB medium. Thereafter, GFP_{1-10} was induced by adding 0.5 mM isopropyl β-D-thiogalactoside (IPTG, Sigma, St. Louis, MO, USA) for 1 or 3 h at 30 °C. GFP_{11}-fused membrane proteins and GFP_{1-10} were also simultaneously expressed by coinduction with IPTG and ANTET at 30 °C for the indicated time. For fluorescence measurements, cells were washed twice with phosphate-buffered saline (PBS) and resuspended in the same buffer, adjusting the cell density to $OD_{600} = 0.2$. Fluorescence intensity and spectra were recorded in a QuantaMaster spectrofluorimeter

[2] *Abbreviations used:* GFP, green fluorescent protein; FSEC, fluorescence size exclusion chromatography; SteT, L-serine/L-threonine exchanger from *Bacillus subtilis*; cDNA, complementary DNA; LB, Luria–Bertani; ANTET, anhydrous tetracycline; IPTG, isopropyl β-D-thiogalactoside; PBS, phosphate-buffered saline; EDTA, ethylenediaminetetraacetic acid; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; DDM, *n*-dodecyl-β-D-maltopyranoside; HRP, horseradish peroxidase; UV, ultraviolet; TM, transmembrane domain; SEC, size exclusion chromatography.

(Photon Technology International, Lawrenceville, NJ, USA) between 500 and 600 nm using an excitation wavelength of 460 nm.

Preparation of isolated cytoplasmic membranes and inclusion bodies

An *E. coli* cell pellet expressing a GFP_{11}-fused membrane protein and GFP_{1-10} was washed once with PBS and subjected to osmotic shock with 45% sucrose followed by incubation with lysozyme (0.5 mg/ml) and 1 mM ethylenediaminetetraacetic acid (EDTA). Subsequently, sample was briefly sonicated and subjected to centrifugation (13,000*g*, 5 min, 4 °C). The resulting supernatant was ultracentrifuged (200,000*g*, 10 min, 4 °C), and the pellet containing the cytoplasmic membranes was kept at −20 °C until use. The pellet from the first centrifugation containing mostly inclusion bodies was washed twice with 1% Triton X-100 to remove unbroken cells [21].

"In-gel" fluorescence visualization of GFP_{11}–GFP_{1-10} fused to membrane proteins

GFP fluorescence achieved after the complementation of GFP_{11}-fused membrane proteins with GFP_{1-10} was visualized in a sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) gel. Briefly, isolated cytoplasmic membranes or purified inclusion bodies of *E. coli* cells coexpressing the different GFP_{11}-fused membrane proteins and GFP_{1-10} were solubilized with 1% *n*-dodecyl-β-D-maltopyranoside (DDM, Affymetrix, Santa, Clara, CA, USA) and subjected to SDS–PAGE. "In-gel" GFP fluorescence from the distinct membrane protein–GFP_{11}–GFP_{1-10} gel bands was visualized using a GBOX gel reader (Syngene, Cambridge, UK) and a Safe Imager (Invitrogen, Carlsbad, CA, USA). Western blot analyses were performed using the HisProbe–HRP (horseradish peroxidase) kit (Thermo Scientific, Rockford, IL, USA).

In vivo visualization of SteT–GFP_{11} expression in bacterial colonies

E. coli BL21(DE3) colonies cotransformed with the plasmids pTET encoding SteT–GFP_{11} and pET encoding GFP_{1-10} were grown overnight at 37 °C on nitrocellulose filter paper (Amersham Hybond-N, GE Healthcare, Little Chalfont, UK) lying on top of an LB agar plate containing spectinomycin and kanamycin, as described previously [18]. Expression of SteT–GFP_{11} was initiated by placing the filter paper in a new plate containing 0.3 mg/ml ANTET for 3 h at 30 °C. After the incubation, the filter paper was moved to a new LB agar plate containing no inducing agent. Finally, GFP_{1-10} was induced by transferring the filter paper into a new LB agar plate containing 0.5 mM IPTG and incubating it for 3 h at 30 °C. Green colonies indicating the expression of SteT–GFP_{11} complemented with GFP_{1-10} were visible under either ultraviolet (UV) or blue light using a Stereo Fluorescence Microscope (Leica).

Expression and purification of SteT–G294V–GFP_{11}

E. coli BL21(DE3) cells harboring pTET–SteT–G294V–$(His)_6$–GFP_{11} were grown in LB medium. After cell density reached $OD_{600} \sim 0.6$, SteT–G294V–$(His)_6$–GFP_{11} expression was induced by adding 0.3 mg/ml ANTET and decreasing the temperature to 30 °C. After 3 h, cells were harvested, resuspended in 20 mM Tris–HCl (pH 8.0) and 0.5 mM EDTA, and disrupted using an EmulsiFlex-C3 homogenizer (Avestin Europe, Mannheim, Germany). Cell debris was removed by centrifugation (10,000*g*, 30 min, 4 °C), and the supernatant was subjected to ultracentrifugation (100,000*g*, 1 h, 4 °C). Membrane pellet was resuspended in 20 mM Tris–HCl (pH 7.6) and 150 mM NaCl, flash-frozen in liquid nitrogen, and stored at −80 °C. For purification, membranes were thawed and solubilized for 30 min at 4 °C in 2% DDM followed by ultracentrifugation (100,000*g*, 30 min, 4 °C). The soluble fraction was incubated with TALON His-tag resin beads (Clontech–Takara Bio Europe, Saint-Germain-en-Laye, France) for 3 h at 4 °C, washed first with 20 mM Tris–HCl (pH 7.6), 200 mM NaCl, 0.02% DDM, and 10 mM imidazole and washed second with the previous buffer plus 20 mM imidazole. Protein was eluted using the same buffer containing 200 mM imidazole. The protein was concentrated using Vivaspin 20 100-kDa MWCO (molecular weight cutoff) concentrators (Sartorius–Stedim, Aubagne, France) and subjected to size exclusion chromatography using a Superdex 200 10/300 GL column (GE Healthcare) equilibrated with 20 mM Tris–HCl (pH 7.6), 150 mM NaCl, and 0.02% DDM.

Purification and refolding of GFP_{1-10} from inclusion bodies

GFP_{1-10} purification from inclusion bodies was carried out following the protocol described by Cabantous and Waldo with some modifications [18]. Briefly, 800 ml of *E. coli* BL21(DE3) cell culture harboring pET–GFP_{1-10} plasmid was induced with 0.5 mM IPTG for 5 h at 37 °C. The cell pellet was resuspended in 24 ml of TNG buffer (50 mM Tris [pH 7.4], 0.1 M NaCl, and 10% glycerol), disrupted, and centrifuged at 16,000*g* for 20 min. The pellet was resuspended in TNG buffer containing 2% Triton X-100 and centrifuged at 16,000*g* for 20 min. The resulting pellet was resuspended in TNG buffer containing 1 M NaCl and centrifuged again for 20 min at 16,000*g*. The final pellet was resuspended with TNG buffer, split in 1-ml aliquots, and centrifuged for 10 min at 16,000*g*. The resulting pellets were purified inclusion bodies and were stored at −80 °C until use. For GFP_{1-10} refolding and purification, inclusion bodies were resuspended with 1 ml of 9 M urea and 5 mM dithiothreitol (DTT) at 37 °C. After a centrifugation step at 16,000*g* for 1 min, the supernatant was diluted by adding 25 ml of TNG buffer. The solution was passed through a 0.2-mm filter, and protein purity was evaluated by Coomassie blue staining SDS–PAGE analysis. Finally, pure protein was quantified using UV absorbance.

In vitro GFP complementation assay with purified SteT–G294V–GFP_{11} and GFP_{1-10}

A range of concentrations of purified GFP_{1-10} (1–100 μM) were added to a quartz fluorescence cuvette containing 1 μM purified SteT–G294V–GFP_{11} in 20 mM Tris–HCl (pH 7.6), 150 mM NaCl, and 0.02% DDM. GFP_{11}–GFP_{1-10} complementation was measured by scanning the GFP fluorescence between 500 and 600 nm with an excitation wavelength of 460 nm as indicated previously.

Stability assay using in vitro GFP complementation assay

E. coli BL21(DE3) cells (15 ml) harboring either SteT–GFP_{11} or SteT–G294V–GFP_{11} were grown at 37 °C. After cell density reached $OD_{600} \sim 0.6$, SteT expression was induced by adding 0.3 mg/ml ANTET and decreasing the temperature to 30 °C. After 3 h, cells were harvested, resuspended in 20 mM Tris–HCl (pH 8.0) and 0.5 mM EDTA, and solubilized using a lysis buffer containing 2% DDM, 30 μg/ml DNase, 0.5 mM Pefabloc (Roche Diagnostics, West Sussex, UK), and lysozyme (0.5 mg/ml) for 2 h at 4 °C. The suspension was ultracentrifuged at 200,000*g* for 10 min and 4 °C. The supernatant was incubated for 1 h at 4 °C in TALON His-tag resin beads, and further purification was performed using empty spin columns following the large-scale purification protocol of SteT–G294V–GFP_{11}. Concentration of purified protein was measured using UV absorbance and characterized by Coomassie blue staining SDS–PAGE. After this step, 20 μl of purified GFP_{11}-fused membrane protein (∼1 μM) was mixed with 180 μl of GFP_{1-10}, and the intensity of GFP fluorescence was measured as described earlier (final concentration 1:50, mol/mol, SteT–GFP_{11}/GFP_{1-10}). After 20 h of

incubation at 4 °C, purified SteT–GFP_{11} was ultracentrifuged (200,000*g*, 10 min, 4 °C) and 20 µl of this supernatant was also mixed with 180 µl of GFP_{1-10} and subjected to fluorescence measurement. The percentage of remaining fluorescence (% *F*) after overnight incubation at 4 °C indicative of the relative stability of each protein was calculated using the following equation:

$$\%F = (F_{4^\circ C}/F_0) * 100, \quad (1)$$

where F_0 and $F_{4\ ^\circ C}$ are the GFP fluorescence intensities measured before and after overnight incubation at 4 °C, respectively.

Results and discussion

Coexpression of SteT–GFP_{11} and GFP_{1-10} in E. coli leads to GFP fluorescence

As the first objective of this work, we studied whether the emission of fluorescence from GFP can be detected after the complementation of the nonfluorescent C-terminal end of GFP (GFP_{1-10}) with the remaining 15 amino acids of GFP (GFP_{11}) fused to the C-terminal end of a membrane protein target expressed in *E. coli* (Fig. 1). We challenged this GFP complementation assay by testing the expression of SteT with GFP_{11} fused at its C terminus in *E. coli* (Fig. 1). The resulting fluorescent signal is expected to be proportional to the amount of protein expressed. SteT is a polytopic membrane transport protein composed of 12 transmembrane domains (TMs) with monomeric assembly in the membrane [20]. SteT can be expressed in *E. coli*, solubilized with DDM, purified, and functionally reconstituted into proteoliposomes composed of *E. coli* phospholipids [20,22]. In our assay, SteT–GFP_{11} and GFP_{1-10} were encoded in two compatible expression vectors (modified versions of pTET and pET, respectively [Fig. 1]) [17,18]. Importantly, the expression of the two proteins is regulated by two independent promoters; therefore, the expression of either SteT–GFP_{11} or GFP_{1-10} is tightly controlled by simply adding or removing the appropriate inducing agent (ANTET for SteT–GFP_{11} and IPTG for GFP_{1-10}) (Fig. 1). As described originally [18], we also added a 10-amino-acid flexible linker (DGGSGGGSTS) between GFP_{11} and the C terminus of SteT to prevent steric restrictions that can hamper GFP_{11}–GFP_{1-10} complementation. The coexpression of SteT–GFP_{11} and GFP_{1-10} in the same cell produced the typical spectrum of the GFP fluorescence after exciting the cells at 460 nm (Fig. 2A). No fluorescence was detected when either SteT–GFP_{11} or GFP_{1-10} was expressed independently (Fig. 2A). Interestingly, a similar expression test can also be conducted with bacterial colonies, as shown in Fig. 2B. By simply passing a nitrocellulose membrane with *E. coli* colonies cotransformed with the two expression vectors into separate agar plates containing the appropriate inducing agent (see Materials and methods for details), SteT expression can be monitored by observing the GFP fluorescence of the bacterial colony in a similar way as described previously for soluble proteins [18]. The feasibility of the split GFP assay for measuring the yield of SteT expression in *E. coli* was confirmed by Western blot analysis using the 6 × His tag epitope placed between the C terminus of SteT and the N terminus of GFP_{11} (Fig. 3A). In these experiments, isolated cytoplasmic membranes of an *E. coli* culture sequentially expressing SteT–GFP_{11} followed by GFP_{1-10} were subjected to SDS–PAGE and immunoblotted with HisProbe–HRP. These experiments showed that the nonfluorescent SteT–GFP_{11} band appeared only in the absence of GFP_{1-10} (Fig. 3A, lanes 1 and 2). Furthermore, a higher molecular weight band corresponding to SteT–GFP_{11}–GFP_{1-10} appeared when GFP_{1-10} was induced and increased in intensity over time (1 and 16 h after GFP_{1-10} induction), whereas the intensity of the SteT–GFP_{11} band decreased (Fig. 3A, lanes 3 and 4). Moreover, a fluorescent band at the expected molecular weight of SteT–GFP_{11}–GFP_{1-10} appeared in an SDS–PAGE gel containing isolated *E. coli* membranes coexpressing SteT–GFP_{11} and GFP_{1-10} (Fig. 3B). This finding again confirms the presence of SteT–GFP_{11}–GFP_{1-10} in the cytoplasmic membrane. In addition, these results corroborate that complementation between SteT–GFP_{11} and GFP_{1-10} occurs after SteT is fully translated (the GFP_{11} tag is at the C terminus) and inserted into the cytoplasmic membrane of the expression host. As described previously [17], the SteT–GFP_{11}–GFP_{1-10} band increased with longer GFP_{1-10} induction times (Fig. 3A, lanes 3 and 4) as a result of a higher cytoplasmic concentration of GFP_{1-10} and a longer time for GFP_{1-10}–GFP_{11} complementation.

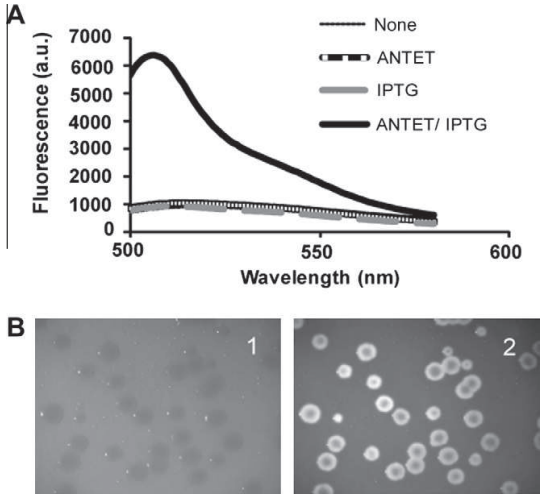


Fig.2. In vivo coexpression of SteT–GFP_{11} with GFP_{1-10} leads to GFP fluorescence. (A) Fluorescent spectra measured in *E. coli* cultures harboring pTET and pET plasmids encoding SteT–GFP_{11} and GFP_{1-10}, respectively. Spectra were recorded using a 460-nm excitation wavelength in noninduced cells (none) or by inducing SteT–GFP_{11} (ANTET), GFP_{1-10} (IPTG), or SteT–GFP_{11} and GFP_{1-10} together (ANTET/IPTG) for 3 h at 30 °C. (B) In vivo SteT–GFP_{11}–GFP_{1-10} complementation can be detected in *E. coli* colonies. Colonies harboring pTET and pET plasmids encoding SteT–GFP_{11} and GFP_{1-10}, respectively, were grown in a nitrocellulose filter on top of an LB plate containing the appropriate antibiotics and inducing agents. After SteT–GFP_{11} and GFP_{1-10} were expressed sequentially and complemented, GFP fluorescence from the colonies was observed under the fluorescence microscope without excitation (panel 1) or with UV excitation (panel 2).

The split GFP system specifically measures SteT–GFP_{11} inserted into the cytoplasmic membrane

In some cases, the heterologous expression of membrane proteins in *E. coli* leads to the accumulation of these proteins as aggregates in inclusion bodies [9]. Because GFP can be fluorescent in inclusion bodies [23], false positive errors can be generated from misfolded or insoluble proteins located in these particles. Interestingly, GFP_{11}–GFP_{1-10} does not complement inside inclusion bodies [17]. Consequently, we reasoned that the sequential expression of each GFP fragment (GFP_{11}-fused membrane protein followed by GFP_{1-10}) could be a valuable expression assay to discard membrane proteins that accumulate in inclusion bodies. To test this hypothesis, we induced the expression of SteT–GFP_{11} at two temperatures (30 and 37 °C) for 2 and 16 h at each temperature (Fig. 4). Subsequently, GFP_{1-10} was induced for 1 h at 30 °C in all of the conditions tested. The fluorescence signal measured in *E. coli* cultures after GFP_{11}–GFP_{1-10} complementation indicates that the expression yield of SteT–GFP_{11} was substantially higher at 30 °C than at

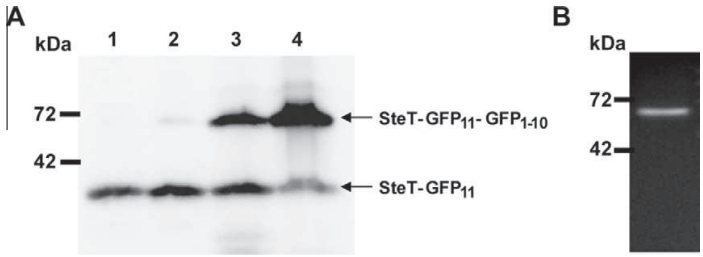


Fig.3. Complementation between SteT–GFP_{11} and GFP_{1-10} occurs in the cytoplasmic membrane. (A) Anti-His-tag Western blot analysis of SteT–GFP_{11} and complemented SteT–GFP_{11}–GFP_{1-10} expression in *E. coli* cytoplasmic membranes. *E. coli* membranes expressing 6 × His-tagged SteT–GFP_{11} before and after GFP_{1-10} induction were solubilized with 1% DDM and subjected to SDS–PAGE before blotting. Lanes: SteT–GFP_{11} induction at 30 °C for 1 h (lane 1) or 2 h (lane 2) and SteT–GFP_{11} induced for 3 h at 30 °C followed by GFP_{1-10} induction for 1 h (lane 3) or 16 h (lane 4). (B) In-gel fluorescence of an SDS–PAGE gel containing isolated *E. coli* membranes coexpressing SteT–GFP_{11} and GFP_{1-10}.

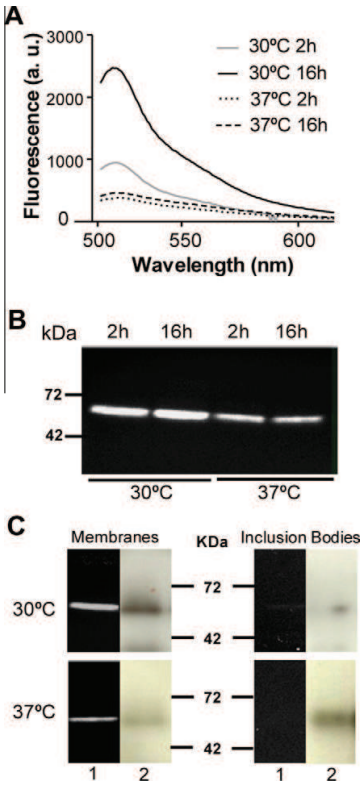


Fig.4. The split GFP system specifically measures the expression of SteT–GFP_{11} in the cytoplasmic membrane. (A) GFP fluorescence spectra of complemented SteT–GFP_{11}–GFP_{1-10} measured with varying conditions of SteT–GFP_{11} induction as indicated in the figure. After SteT–GFP_{11} induction, GFP_{1-10} was induced for 3 h at 30 °C in all experiments. (B) In-gel fluorescence of SDS–PAGE gels containing isolated *E. coli* cytoplasmic membranes expressing SteT–GFP_{11} under a range of conditions as indicated followed by GFP_{1-10} induction for 3 h at 30 °C in all experiments. (C) Analysis of SteT–GFP_{11}–GFP_{1-10} expression by GFP fluorescence emission (lanes 1) and anti-His-tag Western blot (lanes 2) of both the cytoplasmic membrane fraction and purified inclusion bodies of *E. coli* cultures expressing SteT–GFP_{11} at either 30 or 37 °C. After SteT–GFP_{11} induction, GFP_{1-10} was induced for 3 h at 30 °C in all experiments.

37 °C (Fig. 4A), as reported previously using a non-GFP-tagged version of SteT [20]. To ensure that the fluorescence signal came almost exclusively from the cytoplasmic membrane, we performed SDS–PAGE analysis of cytoplasmic membranes and inclusion bodies from the same *E. coli* cultures. The in-gel fluorescence of isolated membranes revealed the presence of the complemented SteT–GFP_{11}–GFP_{1-10} at both temperatures (Fig. 4B). Notably, the fluorescence intensity of each band was consistent with that measured previously in *E. coli* cultures (30 °C > 37 °C [Figs. 4A and 4B]). Furthermore, anti-His-tag Western blot analysis of isolated cytoplasmic membranes confirms that the amount of SteT expressed at each temperature and inserted into the cytoplasmic membrane is also consistent with the fluorescence intensity measured in bacterial cultures and in-gel fluorescence (Figs. 4A–C). Interestingly, the same Western blot analysis of the isolated inclusion bodies revealed that, at 37 °C, SteT–GFP_{11} (complemented with GFP_{1-10}) accumulates in inclusion bodies to a larger extent than at 30 °C, in contrast to the results obtained in the cytoplasmic membrane fraction (Fig. 4C). Remarkably, only a nominal fluorescent signal was observed in the inclusion body fraction at the two temperatures (Fig. 4C). These results confirm that even if the two GFP fragments interact in the cytoplasm before becoming confined to inclusion bodies, the fluorescence emission is almost completely quenched, so the protein fraction present in inclusion bodies does not contribute to the fluorescence signal. Therefore, when sequentially expressing SteT–GFP_{11} followed by GFP_{1-10}, the fluorescence signal not only reflects the expression yield of SteT but also specifically the amount of protein inserted into the cytoplasmic membrane (Figs. 4A–C). This strategy is highly beneficial because it allowed us to quickly discard the protein fraction confined to inclusion bodies as a result of aggregation or misfolding, a common issue in the heterologous expression of membrane proteins [9].

The split GFP can be used as general reporter of the expression yield of membrane proteins in E. coli

We next explored the robustness of this GFP complementation strategy by studying the expression and membrane insertion of four membrane proteins in *E. coli:* the lactose permease of *E. coli* (LacY), the small multidrug transporter of *E. coli* (EmrE), the small conductance mechanosensitive channel of *E. coli* (MscS), and the large conductance mechanosensitive channel of *E. coli* (MscL) (Fig. 5). The selected proteins are well characterized structurally and differ in the number of TMs and in their quaternary structures. LacY is a monomer composed of 12 TMs [24], EmrE is a homodimer with each monomer containing 4 TMs [25], MscS is a homoheptamer containing 3 TMs per monomer [26], and MscL forms a pentameric structure with 2 TMs per monomer [27]. Each protein was cloned in the pTET vector fused to GFP_{11} on its C terminus (Fig. 1) and, as in the case of SteT, a linker comprising 10 amino acids was added between the C-terminal end of the membrane protein and

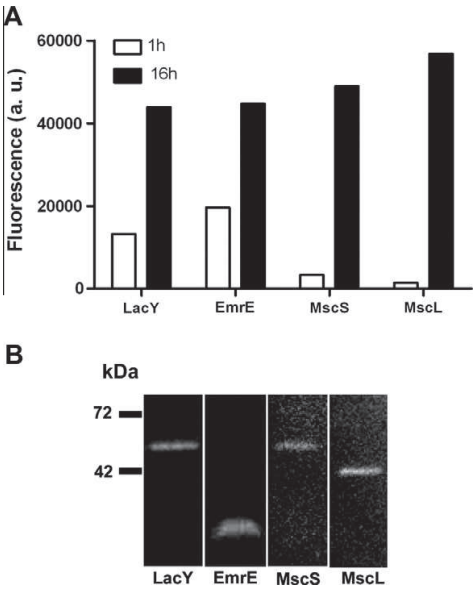


Fig.5. The split GFP system can measure expression yield of membrane proteins in *E. coli*. (A) GFP fluorescence measured in *E. coli* cultures expressing LacY–GFP_{11}, EmrE–GFP_{11}, MscS–GFP_{11}, or MscL–GFP_{11} plus GFP_{1-10}. Membrane proteins were induced at 30 °C for 3 h followed by GFP_{1-10} induction at 30 °C for either 1 or 16 h as indicated. (B) In-gel fluorescence of SDS–PAGE gels containing isolated *E. coli* membranes sequentially expressing LacY–GFP_{11}, EmrE–GFP_{11}, MscL–GFP_{11}, or MscS–GFP_{11} (30 °C for 3 h) followed by GFP_{1-10} (16 h at 30 °C).

GFP_{11}. Membrane protein expression was induced at 30 °C for 3 h in all cases; subsequently GFP_{1-10} expression was induced at 30 °C for 1–16 h. For each tested protein, we measured the GFP fluorescence from the bacterial culture (Fig. 5A) and from an SDS–PAGE gel of the isolated cytoplasmic membranes (Fig. 5B). As found for SteT (Fig. 3A), the fluorescence intensity of GFP increased with longer GFP_{1-10} induction times (1–16 h) (Fig. 5A). The fluorescent bands observed in the SDS–PAGE gels of cytoplasmic membranes expressing each target matched the molecular weight of each GFP-fused membrane protein, similar to the findings of the SteT studies (Fig. 5B). These examples indicate that this split GFP system is a robust strategy to screen the expression of membrane protein candidates with distinct topologies. The only requirement for the application of this approach is that the C-terminal end of the protein is oriented toward the cytoplasmic site.

The stability of detergent-purified membrane proteins can be evaluated using the split GFP system

Membrane protein stability after detergent solubilization and purification is often a key factor in achieving the high-quality protein crystals required for atomic resolution structural determination [28–30]. To achieve this, it is necessary to conduct a large screening process to characterize the stability of a target protein solubilized in a variety of detergents and/or the stability of homologs of a selected protein (sometimes including mutants) in one particular detergent. Consequently, we explored how the GFP complementation assay performed in vitro after protein solubilization and purification can speed up such a screening regimen. First, we examined whether GFP can be complemented in vitro after mixing purified SteT–G294V–GFP_{11} and purified GFP_{1-10}. We found that the G294V mutation in SteT substantially improves protein stability after detergent solubilization and purification. The size exclusion chromatography (SEC) profile of wild-type SteT solubilized with DDM and purified by affinity chromatography [22] indicated the presence of aggregates at concentrations of >1 mg/ml (see Supplementary Fig. 1A in supplementary material). Under these conditions, wild-type SteT began to precipitate after the protein concentration reached $\geqslant$3 mg/ml. In contrast, 10 mg/ml injected SteT–G294V eluted predominantly as a monomer in the same SEC experiment (Supplementary Fig. 1B). Therefore, we considered SteT–G294V to be a better candidate than wild type for these preliminary in vitro complementation experiments. SteT–G294V–GFP_{11} was extracted from *E. coli* membranes with DDM, and protein was purified by affinity and gel filtration chromatography, as described previously [22]. Purified SteT–G294V–GFP_{11} (1 μM) was incubated with a range of concentrations of purified GFP_{1-10} functionally refolded from inclusion bodies [18] (see Materials and methods for details). After a short incubation (15 min), the typical GFP fluorescence spectrum appeared (Fig. 6A), thereby demonstrating that GFP_{11} fused to SteT successfully complements GFP_{1-10} in a buffer containing detergent at the concentration typically used in membrane protein crystallization (2 × CMC of DDM). Moreover, in our hands, a molar ratio of 50:1 (GFP_{1-10}/SteT–G294V–GFP_{11}) resulted in an optimal fluorescence signal (Fig. 6B). We further exploited this assay to build a rapid protocol to test the stability of SteT–GFP_{11} after DDM solubilization and purification by affinity chromatography. For this purpose, we used the two versions of SteT (wild-type and mutant G294V) whose stabilities in detergent-purified solution are distinct, as discussed earlier (Supplementary Fig. 1). Consequently, affinity mini-purifications of each

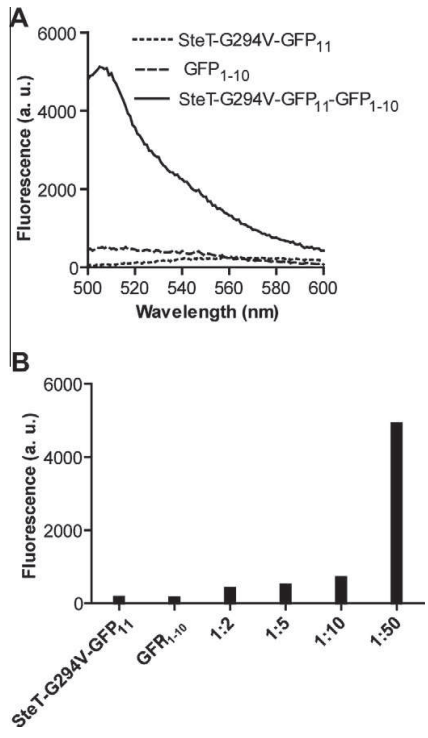


Fig.6. In vitro complementation of purified SteT–G294V–GFP_{11} with GFP_{1-10} leads to GFP fluorescence. (A) Fluorescence spectra of purified SteT–G294V–GFP_{11}, purified GFP_{1-10}, and mixed SteT–G294V–GFP_{11} and GFP_{1-10} at a 1:50 molar ratio. (B) Fluorescence intensity of GFP resulting from the incubation of 1 μM purified SteT–G294V–GFP_{11} with GFP_{1-10} at different molar ratios as indicated.

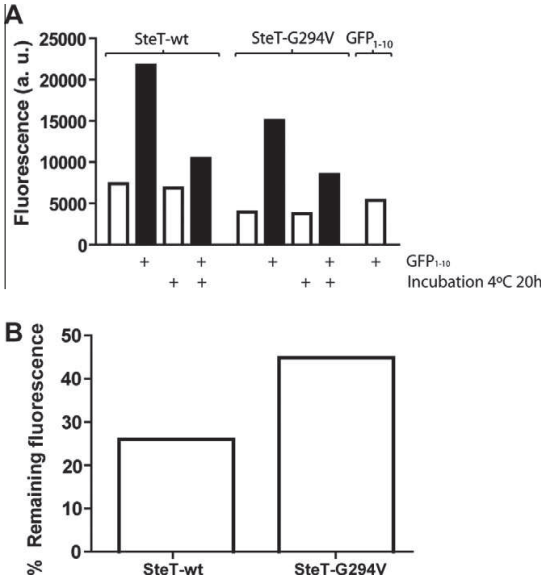


Fig.7. Stability of purified SteT–GFP_{11} and SteT–G294V–GFP_{11} measured using in vitro GFP complementation. (A) GFP fluorescence of purified SteT–GFP_{11} and SteT–G294V–GFP_{11} before (white bars) and after (black bars) complementation with purified $GFP_{1–10}$. As indicated at the bottom of the graph, the complementation was performed immediately after SteT–GFP_{11} or SteT–G294V–GFP_{11} purification (no label) or after an overnight incubation of both SteT versions at 4 °C (+). (B) Quantification of the stability of purified SteT–GFP_{11} and SteT–G294V–GFP_{11} measured as the percentage of fluorescence remaining after overnight incubation at 4 °C (% *F*) from the results obtained in panel A. % *F* was calculated using Eq. (1).

SteT–GFP_{11} version were incubated with 50 μM purified $GFP_{1–10}$ immediately after purification and an overnight incubation at 4 °C. Protein aggregates were removed after ultracentrifugation in each case. GFP fluorescence was measured after each complementation assay (Fig. 7A) and expressed for each SteT version as the percentage of fluorescence signal (% *F*) remaining in solution after the overnight incubation and subsequent ultracentrifugation (Eq. (1) and Fig. 7B). Consequently, the higher the percentage, the greater the stability of the protein in the buffer conditions used. As expected, SteT–G294V–GFP_{11} proved to be more stable in solution than SteT–GFP_{11} because 45% of the mutant remained in solution after the overnight incubation compared with 23% "survival" of wild-type SteT (Fig. 7B). These results are consistent with our previous finding using non-GFP-tagged versions of wild type and G294V (Supplementary Fig. 1); thus, they reinforce the use of this strategy to measure the stability of membrane proteins in detergent-solubilized micelles using the in vitro split GFP system.

Conclusions

Structural studies of membrane proteins often require an extensive search for optimal targets with two main characteristics: a reasonable yield in the chosen expression host and good stability after detergent solubilization and purification. The use of full-length GFP as a reporter has allowed expedition and simplification of this search; however, in some cases, the fused GFP interferes with the expression and stability of the fused membrane protein partner. Here we sought to set up a new approach minimizing the undesired effects of the GFP tag while preserving the considerable benefits (rapid and sensitive) of using luminescent GFP as a reporter for structural studies of membrane proteins. In the split GFP system, the tag contains only 15 amino acids, minimizing its interfering effect, allowing more precise information regarding protein folding and membrane insertion of a particular target to be obtained as well as better knowledge of its stability in detergent-solubilized micelles. Furthermore, given that the split GFP method has the ability to discriminate between protein aggregates in inclusion bodies and protein inserted into the cytoplasmic membrane, false positives during the heterologous expression of membrane proteins can be avoided.

Finally, our results indicate that the split GFP system described here is a rapid and sensitive tool that improves the use of GFP as a fluorescent fusion reporter of membrane proteins while simultaneously overcoming some of the limitations associated with its use. We performed this study using *E. coli* as an expression host; however, we believe that this method could easily be implemented in a medium- to high-throughput manner in other hosts such as yeast, insect cells, and mammalian cell lines.

Acknowledgments

We are very grateful to Geoffrey S. Waldo and members of his laboratory for sending us the split GFP expression plasmids. We thank Ronald H. Kaback, Shimon Schuldiner, and Oded Lewinson for providing the LacY, EmrE, MscS, and MscL cDNA. We are indebted to Adam Weinglass and Ekaitz Errasti Murugarren for critical reading of the manuscript, to Antonio Zorzano and members of his group for their advice, and to Susana Bial for technical support. We also thank T. Yates for editing the English. This work was supported in part by the Spanish Ministry of Science and Innovation, Grants BFU2008-04637 (J.L.V-I.) and SAF2009-12606-C02-01 (M.P.), and by the EC FP7 Grant 201924 (EDICT) (M.P. and L.K.).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ab.2011.12.044.

References

[1] E. Wallin, G. von Heijne, Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms, Protein Sci. 7 (1998) 1029–1038.
[2] K.M. Giacomini, S.M. Huang, D.J. Tweedie, L.Z. Benet, K.L. Brouwer, X. Chu, A. Dahlin, R. Evers, V. Fischer, K.M. Hillgren, K.A. Hoffmaster, T. Ishikawa, D. Keppler, R.B. Kim, C.A. Lee, M. Niemi, J.W. Polli, Y. Sugiyama, P.W. Swaan, J.A. Ware, S.H. Wright, S.W. Yee, M.J. Zamek-Gliszczynski, L. Zhang, Membrane transporters in drug development, Nat. Rev. Drug Discov. 9 (2010) 215–236.
[3] E. Errasti-Murugarren, M. Pastor-Anglada, Drug transporter pharmacogenetics in nucleoside-based therapies, Pharmacogenomics 11 (2010) 809–841.
[4] R.B. Kim, Transporters and drug discovery: why, when, and how, Mol. Pharm. 3 (2006) 26–32.
[5] R. Grisshammer, C.G. Tate, Overexpression of integral membrane proteins for structural studies, Q. Rev. Biophys. 28 (1995) 315–422.
[6] M.M. Klepsch, J.O. Persson, J.W. de Gier, Consequences of the overexpression of a eukaryotic membrane protein, the human KDEL receptor, in *Escherichia coli*, J. Mol. Biol. 407 (2011) 532–542.
[7] O. Lewinson, A.T. Lee, D.C. Rees, The funnel approach to the precrystallization production of membrane proteins, J. Mol. Biol. 377 (2008) 62–73.
[8] S.K. Singh, A. Yamashita, E. Gouaux, Antidepressant binding site in a bacterial homologue of neurotransmitter transporters, Nature 448 (2007) 952–956.
[9] C.M. Koth, J. Payandeh, Strategies for the cloning and expression of membrane proteins, Adv. Protein Chem. Struct. Biol. 76 (2009) 43–86.
[10] D. Drew, M. Lerch, E. Kunji, D.J. Slotboom, J.W. de Gier, Optimization of membrane protein overexpression and purification using GFP fusions, Nat. Methods 3 (2006) 303–313.
[11] T. Kawate, E. Gouaux, Fluorescence detection size-exclusion chromatography for precrystallization screening of integral membrane proteins, Structure 14 (2006) 673–681.
[12] S. Newstead, H. Kim, G. von Heijne, S. Iwata, D. Drew, High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in *Saccharomyces cerevisiae*, Proc. Natl. Acad. Sci. USA 104 (2007) 13936–13941.

[13] J. Hammon, D.V. Palanivelu, J. Chen, C. Patel, D.L. Minor Jr., A green fluorescent protein screen for identification of well-expressed membrane proteins from a cohort of extremophilic organisms, Protein Sci. 18 (2009) 121–133.
[14] M. Ormo, A.B. Cubitt, K. Kallio, L.A. Gross, R.Y. Tsien, S.J. Remington, Crystal structure of the *Aequorea victoria* green fluorescent protein, Science 273 (1996) 1392–1395.
[15] S. Fucile, E. Palma, A. Martinez-Torres, R. Miledi, F. Eusebi, The single-channel properties of human acetylcholine α7 receptors are altered by fusing α7 to the green fluorescent protein, Proc. Natl. Acad. Sci. USA 99 (2002) 3956–3961.
[16] J.M. Hsieh, G.M. Besserer, M.G. Madej, H.Q. Bui, S. Kwon, J. Abramson, Bridging the gap: a GFP-based strategy for overexpression and purification of membrane proteins with intra- and extracellular C-termini, Protein Sci. 19 (2010) 868–880.
[17] S. Cabantous, T.C. Terwilliger, G.S. Waldo, Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein, Nat. Biotechnol. 23 (2005) 102–107.
[18] S. Cabantous, G.S. Waldo, In vivo and in vitro protein solubility assays using split GFP, Nat. Methods 3 (2006) 845–854.
[19] J.D. Pedelacq, S. Cabantous, T. Tran, T.C. Terwilliger, G.S. Waldo, Engineering and characterization of a superfolder green fluorescent protein, Nat. Biotechnol. 24 (2006) 79–88.
[20] N. Reig, C. del Rio, F. Casagrande, M. Ratera, J.L. Gelpi, D. Torrents, P.J. Henderson, H. Xie, S.A. Baldwin, A. Zorzano, D. Fotiadis, M. Palacin, 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 (2007) 13270–13281.
[21] E. Rodriguez-Carmona, O. Cano-Garrido, J. Seras-Franzoso, A. Villaverde, E. Garcia-Fruitos, Isolation of cell-free bacterial inclusion bodies, Microb. Cell Fact. 9 (2010) 71.
[22] P. Bartoccioni, C. Del Rio, M. Ratera, L. Kowalczyk, J.M. Baldwin, A. Zorzano, M. Quick, S.A. Baldwin, J.L. Vazquez-Ibar, M. Palacin, Role of transmembrane domain 8 in substrate selectivity and translocation of SteT, a member of the l-amino acid transporter (LAT) family, J. Biol. Chem. 285 (2010) 28764–28776.
[23] E. Garcia-Fruitos, N. Gonzalez-Montalban, M. Morell, A. Vera, R.M. Ferraz, A. Aris, S. Ventura, A. Villaverde, Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins, Microb. Cell Fact. 4 (2005) 27.
[24] J. Abramson, I. Smirnova, V. Kasho, G. Verner, H.R. Kaback, S. Iwata, Structure and mechanism of the lactose permease of *Escherichia coli*, Science 301 (2003) 610–615.
[25] Y.J. Chen, O. Pornillos, S. Lieu, C. Ma, A.P. Chen, G. Chang, X-ray structure of EmrE supports dual topology model, Proc. Natl. Acad. Sci. USA 104 (2007) 18999–19004.
[26] R.B. Bass, P. Strop, M. Barclay, D.C. Rees, Crystal structure of *Escherichia coli* MscS, a voltage-modulated and mechanosensitive channel, Science 298 (2002) 1582–1587.
[27] G. Chang, R.H. Spencer, A.T. Lee, M.T. Barclay, D.C. Rees, Structure of the MscL homolog from *Mycobacterium tuberculosis:* a gated mechanosensitive ion channel, Science 282 (1998) 2220–2226.
[28] M.J. Serrano-Vega, F. Magnani, Y. Shibata, C.G. Tate, Conformational thermostabilization of the β1-adrenergic receptor in a detergent-resistant form, Proc. Natl. Acad. Sci. USA 105 (2008) 877–882.
[29] S. Eshaghi, High-throughput expression and detergent screening of integral membrane proteins, Methods Mol. Biol. 498 (2009) 265–271.
[30] Y. Sonoda, S. Newstead, N.J. Hu, Y. Alguel, E. Nji, K. Beis, S. Yashiro, C. Lee, J. Leung, A.D. Cameron, B. Byrne, S. Iwata, D. Drew, Benchmarking membrane protein detergent stability for improving throughput of high-resolution X-ray structures, Structure 19 (2011) 17–25.

RESUMEN

8 RESUMEN

8.1 Introducción

Aproximadamente el 30% del genoma humano codifica para proteínas de membrana integrales (Wallin et al, 1998). Defectos genéticos que afectan tanto su expresión como su funcionalidad de muchas proteínas de membrana son la causa directa de importantes patologías. De hecho, más del 50% de las proteínas de membrana hoy en día son dianas de fármacos (Lundstrom et al, 2006). Las proteínas de integrales de membrana hacen funciones cruciales como mantener la homeostasis celular transfiriendo información entre cada lado de la membrana citoplasmática o entre compartimentos intracelulares. La distribución entre tejidos o entrada específica en los órganos es generalmente facilitada o mediad por la expresión de transportes que permiten la entrada o transferencia de estos fármacos (Kim et al., 2006).

La cristalografía por rayos X es hoy en día una de las técnicas más potentes para estudiar estas proteínas a nivel atómico. Las estructuras 3D generadas no solo proveen información sobre su función, pero también ayudan a descubrir nuevos agentes terapéuticos utilizando su estructura como base del diseño del fármaco, el cual modulará la funcionalidad de la proteína (Blundell et al, 2002). Desafortunadamente, obtener cristales de alta calidad para la difracción con rayos X es una labor altamente difícil, debido a la naturaleza hidrofóbica de estas proteínas y su flexibilidad conformacional en solución (Wiener MC et al, 2004). Aunque, el primer problema empieza por la dificultad de obtener cantidades de porteína recombínate y funcional de membrana que se necesitan para estudios de cristalografía o de MNR usando sistemas de expresión heterólogos como es el caso de la bacteria *E. coli*, el organismo más ampliamente usado para la sobreexpresión de proteínas (Grisshammer and

Tate, 1995). Normalmente, las proteínas de membrane comparadas con las solubles, exiben bajo rendimiento de expresión. Además, la expression heteróloga de las proteínas de membrane resulta frecuentemente en aggregados de proteína en cuerpos de inclusión como consecuencia de un plegamineto incorrecto. En *E. coli*, este problema se ha atribuido a una saturación de la maquinaria Sec (Klepsch et al., 2011) utilizada durante el proceso de biogénesis y de la inserción de todas las proteínas de membrana en la membrana citoplasmática. Consecuentemente, es necesario un proceso incial de proceso de búsqueda incial de candidatos óptimos para la cristalización y estudios funcionales con un rendimiento de expresión y estabilidad en detergente después de la solubilización razonables.

Actualmente, hay un revolución de métodos y procesos empleados en el campo de la biología structural de alto rendimiento, tanto para la expresión, solubilización, purificación y cristalización de proteínas de membrana (Kawate and Gouaux, 2006), (Rasmussen et al., 2007) and (Simon Newstead, 2007). Se espera por tanto que estos avances técnicos lleven a incrementar el número de estructuras resueltas de proteínas de membrana en un futuro cercano.

Una de las estrategias inciales, altamente y eficientemente usadas es utilizar homólogos bacterianos, estos han probado ser un excelente paradigma estructural y funcional de las proteínas de membrana eucariotas (Singh et al., 2007). Por otro lado, la mutagénesis es de una de las técnicas más efectivas y menos costosas cuando se trata de mejorar tanto la expresión como la estabilidad de estas proteínas. Normalmente mutaciones puntuales en segmentos transmembrana han demostrado incrementar la estabilidad en solución de algunas de estas proteínas después de su solubilización en y extracción con detergente desde la membrana (Smirnova and Kaback HR, 2003)

and (Tate CG and Schertler GF, 2009). Además, una mutación simple puede estabilizar una conformación específica durante su ciclo catalítico (Abramson et al, 2003) and (Kowalczyk et al., 2010). Métodos de mutagénesis sistemáticos, también se han empleado para encontrar mutantes que tengan mayor estabilidad y expresión, como el uso del la *alanine scanning mutagénesis*, usado para la termoestabilización del β1-adrenergic receptor un receptro acoplado a proteína G (GPCRs) (Warne et al., 2003). A pesar de esto, encontrar dichos mutantes no es una tarea trial y de hecho es muy difícil y a veces imposible detectar que mutaciones van a ser efectivas para la estabilización de la proteína.

Los Transportadores de aminoácidos L (LAT) corresponden a una de las 13 famílias de la superfamília de transportadores APC (Jack et al., 2000). La familia LAT tiene tanto miembros procariotas como eucariotas. Los miembros eucariotas son las subunidades ligeras de los transportadores de aminoácidos heteroméricos (HAT) (Palacin et al., 2005). Los HAT son intercambiadores compuestos por dos subunidades, un proteína de membrana politópica (subunidad ligera; familia SLC 7) y una glicoproteína N-glicosilada tipo II, unida por un puente disulfuro a la ligera (sub unidad pesada, familia SLC3) (Palacin et al., 2005). La subunidad ligera es el componente catalítico del transporte (Reig et al., 2002), mientras que la subunidad pesada parace ser esencial unicamente en la transferencia de la subunidad pesada hacia la membrana plasmática (Bartoccioni et al., 2008). Diversas enfermedades intervienen a causa del papel de los HATs (Broër and Palacin, 2011).

SteT (serine/threonine antiporter) de *Bacillus subtillis* es el primer miembro procariota caracterizado de la familia LAT y tiene una identidad amino acídica aproximadamente de un 30% con las subunidades ligeras HAT en humanos

(Reig et al., 2007). SteT es una proteína de membrane politópica compuesta por 12 segmentos transmembrana (TM) con una inserción monomérica en la membrana (Reig et al., 2007). SteT puede ser expresada en *E. coli*, solubilizada con detergente DDM (Dodecyl-maltopiranoside), purificada y reconstituida en proteoliposomas compuestos de fosfolípidos de *E. coli* (Reig et al., 2007) (Bartoccioni et al., 2010), a pesar de esto SteT es bastante inestable en solución y empieza a precipitar a partir de los 2-3 mg/ml, haciendo imposible cualquier trabajo de cristalización.

El trabajo con proteínas de membrana sigue siendo un desafío para los biólogos estructurales. Encontrar tanto las condiciones óptimas para estos estudios es una tarea laboriosa tanto de cribaje de expresión y estabilidad de la proteína. De hecho como ocurre con SteT, cuando una proteína es inestable es casi imposible trabjar con estas proteínas para hacer cristalización. Similar al trabajo de cristalización, métodos de alto-rendimiento basados en la combinación de mutagénesis al azar con protocolos que permitan una detección y selección de alto cribaje, serán la vía más rápida para identificar mutantes más estables de una proteína de membrana particular. Por eso el objetivo principal de ésta tesis es desarrollar un protocolo de medio-alto rendimiento para generar y caracterizar mutantes funcionales de SteT con más estabilidad una vez solubilizados en detergente y por lo tanto con más probabilidad de cristalizar, sin necesidad de purificarlos previamente. Este protocolo está concebido para ser una metodología general para cualquier proteína de interés.

8.2 Resultados y discusión

8.2.1 Libreria de mutantes

Hemos puesto a punto una metodología para generar y seleccionar de manera rápida mutantes de la proteína SteT con una o dos substituciones aminoacídicas que se expresan y se insertan en la membrana citoplasmática de *E. coli*. El ensayo utilizado consite en usar una GFP divida en 2 fragmentos. Este ensayo nos ha facilitado enormemente este trabajo (Rodriguez-Banqueri et al, 2012). Este método consiste en 2 vectores uno que expresa SteT unida a un fragmento de 16 aminoácidos de la GFP (GFP_{11}) y un segundo vector que expresa el resto de la GFP (GFP_{1-10}). Si la proteína de membrana está correctamente expresada y plegada en la membrana citoplasmática, el fragmento GFP_{1-10} se complementa con el GFP_{11} obteniendo fluorescencia. En el caso de que la proteína agrega y va a parar a cuerpos de inclusión no hay fluorescencia (Figura 1). Utilizamos la técnica *error prone PCR* para generar los mutantes al azar y una vez clonados en el vector $pTETGFP_{11}$, transformamos los clones en células BL21(DE3) conteniendo el segundo vector de este sistema ($pETGFP_{1-10}$) en una membrana de nitrocelulosa sobre una placa de Petri, después de hacer el ensayo in vivo (mirar materiales y métodos) induciendo secuencialmente en las placas de Petri cada una de las proteínas. Observamos las placas en una lupa binocular con exposición a UV, para detectar las colonias fluorescentes (Figura 2).

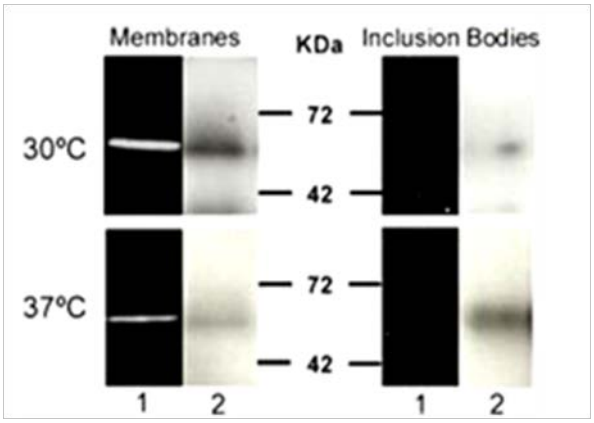


Figura 1. "Split GFP system" mide específicamente la expression de SteT–GFP$_{11}$ en la membrane citoplasmática. Analisis de la expression SteT–GFP$_{11}$–GFP$_{1–10}$ por emission de fluorescencía de la GFP (línea 1) y un anti-His-tag Western blot (línea 2) tanto de las fracciones correspondientes a la membrane citoplasmática como a los cuerpos de inclusion purificados de *E. coli* expresando SteT-GFP$_{11}$ tanto a 30°C como a 37°C. Después de la inducción de SteT-GFP$_{11}$, GFP$_{1-10}$ fue inducida durante 3h a 30°C en todos los experimentos.

Seleccionamos unas 533 colonias fluorescentes y de éstas secuenciamos unas 395. De las colonias secuenciadas, encontramos 149 que contenían de una a dos mutaciones en segmentos transmembrana. Para analizar la localización de estos mutantes usamos un modelo de SteT (Bartoccioni et al., 2010) basado en la estructura 3D de Adic (Fang et al., 2009). De los 149 mutantes, localizamos 101 que contenían al menos una mutación en estos dominios transmembrana. De los 101 mutantes, 70 fueron analizados tanto para expresión como para ver su estabilidad sin necesidad de purificarlos.

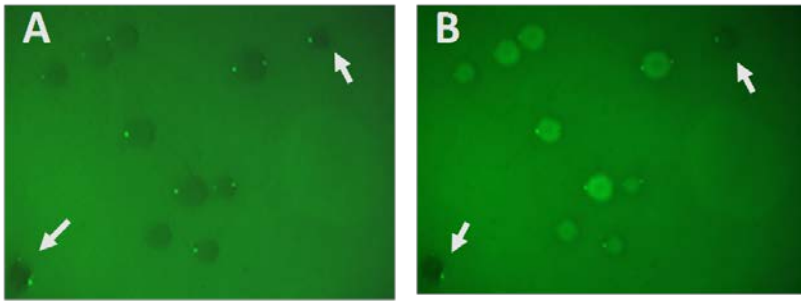


Figura 2. Selección de los mutantes en placas de petri: (A) mutantes sin exposición a luz UV (right) mutantes con exposición a luz UV y un filtro para la GFP. Las flechas blancas indican mutantes que no expresan fluorescencia.

8.2.2 Analisis de los mutantes

Se analizaron los 70 mutantes buscando aquellos que mejoraran subtancialmente las propiedades de SteT WT tanto a nivel de expression como a nivel de estabilidad, utilizando únicamente la fluorescencia de la GFP y sin la necesidad de purificar previamente las proteínas. Para analizar la expresión utilizamos pequeños cultivos de 20 ml y observamos la fluorescencia de los mutantes y WT. Después del crecimiento e inducción de estos y corrigiendo por la densidad óptica los cultivos, obtuvimos los valores relativos de expresión respecto a WT.

En la figura 3 se muestran los mutantes ordenados de menor a mayor nivel de expresión, donde mutantes que expresaban más que WT, son aquellos con valores superiores al 100%. Para el análisis de la estabilidad usamos FSECs (cromatografía de exclusión por tamaño fluorescente), sin necesidad de purificar la muestra y únicamente solubilizando las membranas con DDM (dodecyl-maltopiranosido), comparamos el grado de homogeneidad de la muestra con WT. Una vez normalizados los valores de cada perfil de elución de FSEC, para no tener en cuenta la intensidad de los mismos, comparamos la

anchura de los picos de elución y calculamos un valor numérico dividiendo el área de elución del pico de WT por el área del pico de elución de cada mutante (Figura 4). Los valores obtenidos por encima del 1 nos indican un grado mayor de monodispersidad que WT y por tanto mayor estabilidad de la proteína en solución. Así que concluimos que tanto el mutante I134V-A377T y L210Q-M229V (Figura 5) eran candidatos para hacer estudios de purificación, funcionales y ensayos de pre-cristalización. Antes de realizar los ensayos de purificación, se decidió realizar un screnning de detergentes por FSEC para estos dos mutantes con diferentes detergentes, más usados para la cristalización de proteínas de membrana.

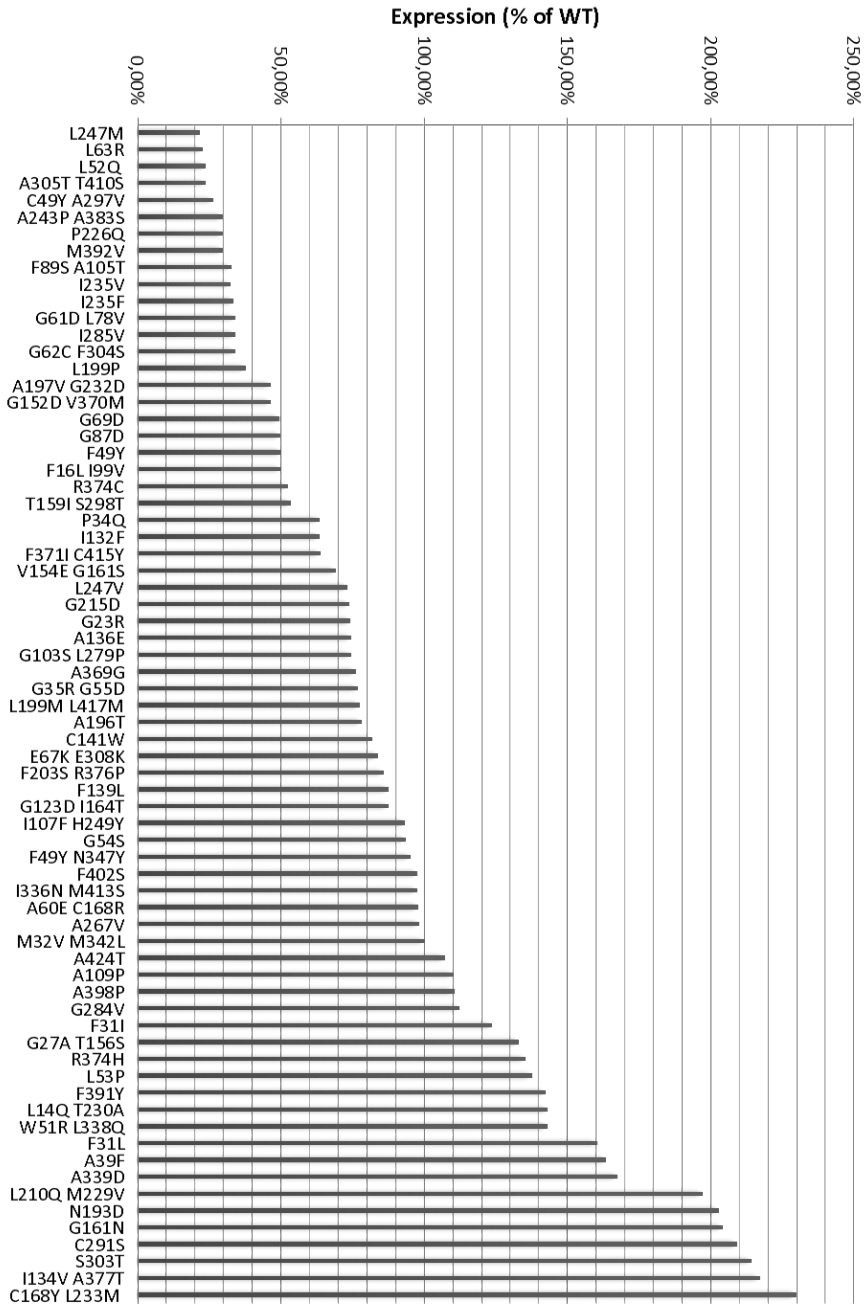


Figura 3. % de expresión de los 70 mutantes de SteT respecto a WT. Valores de expresión calculados a partir de la fluorescencia en cultivo de los mutantes comparados con WT (ambos corregidos por la densidad óptica de los cultivos y de la cepa sin expresar).

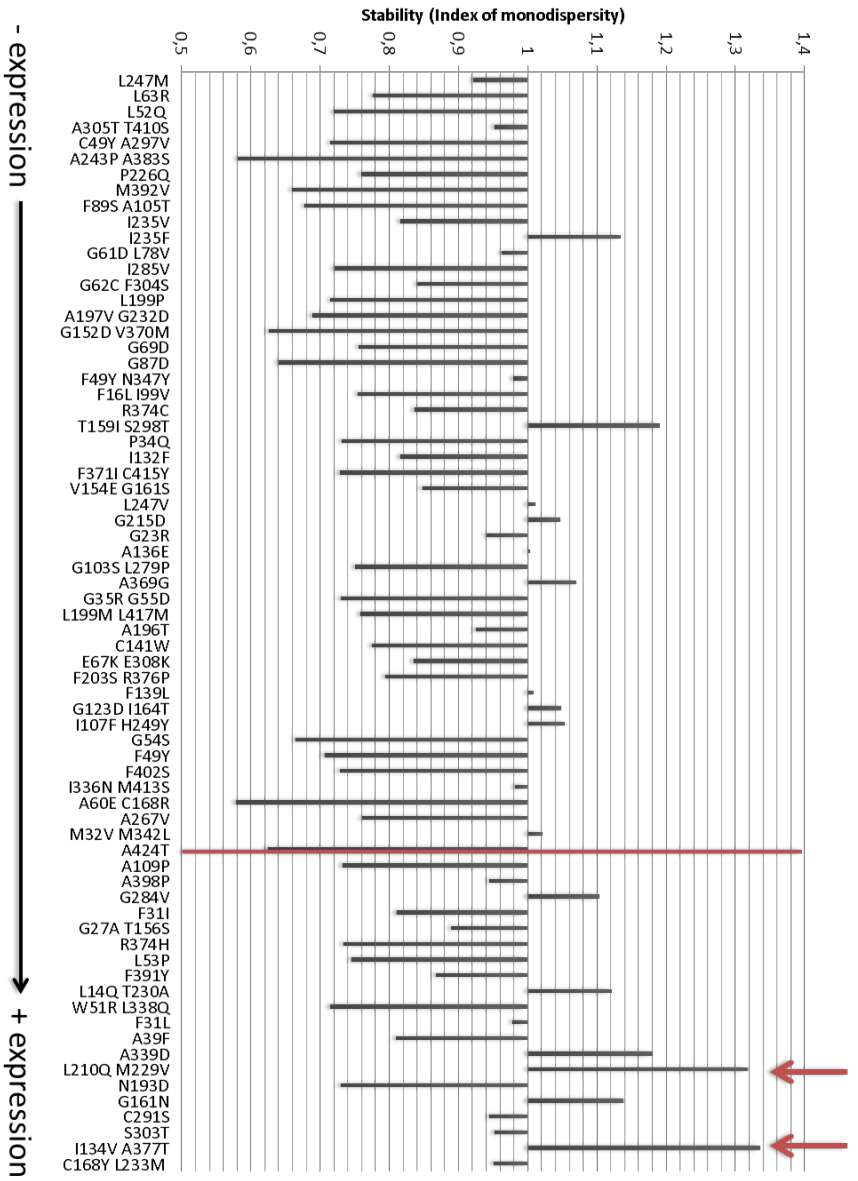


Figura 4. Análisis de la estabilidad de los 70 mutantes de SteT seleccionados. Normalización de las áreas correspondientes a los picos de elución de las FSEC de cada mutante de SteT y WT fue realizado, para calcular el índice de monodispersidad, el área del pico de elución de WT normalizado fue divido con el área del pico de elución de cada mutante normalizada. Los mutantes están ordenados de menor a mayor expresión. Los mutantes representados debajo de la línea roja son aquellos que expresan más que WT y los más estables son aquellos que tienen valores superiores a 1. Los mutantes L210Q-M229V y I134V-A377T que fueron seleccionados están representados por 2 flechas rojas.

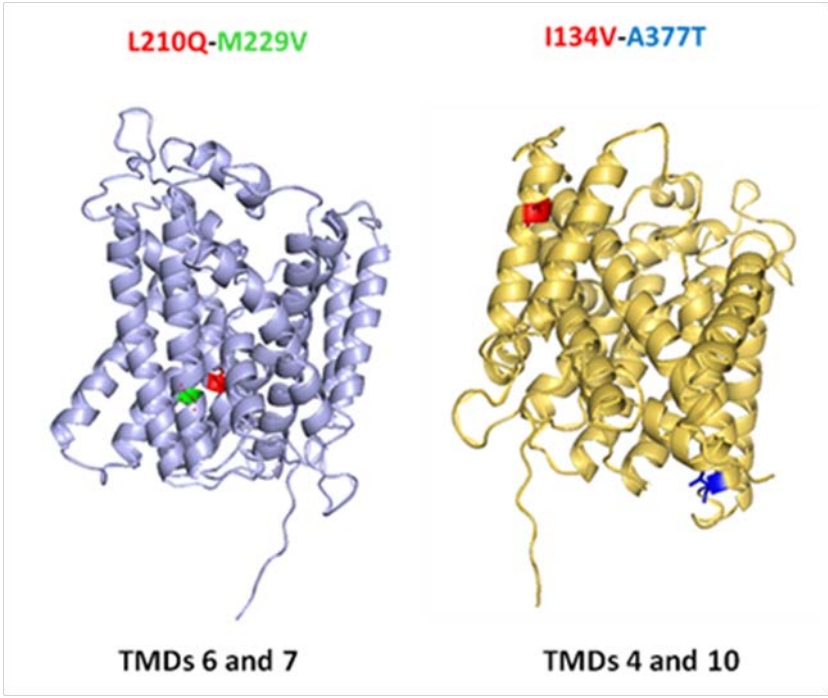


Figura 5. Localización de las mutaciones de los mutantes L210Q-M229V y I134V-A377T utilizando el modelo de SteT (Bartoccioni et al, 2010). En color rojo, verde y azul se representan las localizaciones de las mutaciones en ambos modelos. TMD: Dominio Transmembrana

8.2.3 Purificación y análisis de los mutantes I134V-A377T y L210Q-M229V

La purificación de los mutantes de SteT requería de cultivos más grandes de *E. coli* además de un vector de expresión más robusto. Así que decidimos clonar los dos mutantes en un vector de sobreexpresión único pTTQ18 (Stark, 1987). Este vector fue modificado en nuestro laboratorio por el doctor Ekaitz Errasti y contiene la versión súper-plegada de la GFP fusionada al extremo C-terminal del sito de clonación así como cola de 10 aminoácidos de His. Este vector también contiene una diana del enzima HRV 3C para eliminar tanto la GFP como la cola de histidinas.

Tabla 1. Evaluación de la expresión de WT, L210Q-M229V y I134V-A377T utilizando el vector pTTQ18-GFP. La tabla contiene los niveles de expression de los mutantes en mg/L. Todas las condiciones están expresadas usando 0.1 mM of IPTG en un crecimiento de 16 h.

Temperature	25°C	30°C	37°C
WT	0,56 mg/L	0,6 mg/L	1,44 mg/L
L210Q-M229V	0,90 mg/L	0,87 mg/L	2,05 mg/L
I134V-A377T	3,5 mg/L	3,87 mg/L	5,16 mg/L

El primer ensayo fue el análisis de la expresión de estos dos mutantes en este nuevo vector. En la tabla 1 podemos ver los niveles de expresión de los 2 mutantes y WT, siendo I134V-A377T el mutante con mayor expresión. Después se purificó la proteína para ver el comportamiento de está una vez purificada y concentrada, utilizando DDM (Figura 6). En ambos casos los 2 mutantes se mostraron monodispersos y más estables que WT. También se midió la estabilidad a una concentración de 1 mg/ml a 4°C; y tanto L210Q-M229V y

I134V-A377T, no mostraron pérdida de proteína en solución. A diferencia de WT que no es estable una vez purificada a 1 mg/ml más de 2 días.

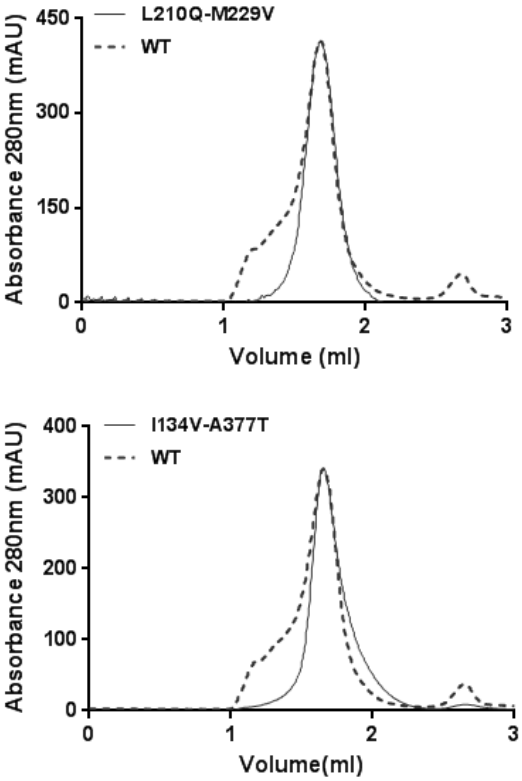


Figura 6. Perfiles de elución de SEC normalizados y superpuestos. SteT-WT (línea discontinua) y SteT-L210Q-M229V y SteT-I134V-A377T (línea continua). 100 µl por muestra fueron inyectados en una superdex 200 50/150 GL.

Decidimos escoger el mutante I134V-A377T para continuar haciendo ensayos de pre-cristalización, observando el comportamiento de este mutante en concentraciones mayores en DDM y analizando su estabilidad, no sólo en éste detergente sino también en otros más adecuados para la cristalización ya que forman micelas más pequeñas favoreciendo los contactos proteína-proteína Estos detergentes son el Decyl-maltopiranosido (DM), el Nonyl-glucosido (NG), el Octyl-glucosido (OG) y el Cymal-6. En todos los casos el mutante tenía un perfil de elución monodisperso, además la proteína I134V-A377T en DDM se mantuvo estable a 6 mg/ml durante una semana a 4°C, sin que la proteína

precipitara. También fue concentrada a 1 mg/ml para el análisis de su estabilidad tanto en DM, NG, OG y Cymal-6 a 4°C durante 3 días. En todos los casos se mantuvo estable, excepto para la proteína solubilizada en NG y OG donde un 86% y un 73% permanecieron estables respectivamente, después de 3 días a 4°C. Este ensayo se realizó también para L210Q-M229V, y se obtuvieron resultados similares, excepto en NG donde no se pudo llegar a concentrar la proteína, seguramente debido a la precipitación de la misma.

8.2.4 Estudios funcionales

Después de los ensayos de purificación, quedaba aún sin resolver si los mutantes seguían siendo funcionales después de la mutagénesis. Para determinar si los mutantes eran funcionales: WT-SteT, L210Q-229V-SteT y I134V-A377T-SteT con GFP en el extremo C-terminal se sobreexpresaron en *E. coli*, se purificaron mediante IMAC y se reconstituyeron en proteoliposomas para experimentos de transporte. Los experimentos de transporte revelaron que ambos mutantes son funcionales (Figura 7). Aún son necesarios estudios adicionales de los mutantes I134V-A377T y L210Q-M229V para caracterizarlos. Aunque en el caso del mutante I134V-A377T, se puede asumir que, basándose en el modelo de SteT (Bartoccioni et al, 2010), que las mutaciones de este mutante están situadas muy lejos del sitio de unión al sustrato (Figura 5).

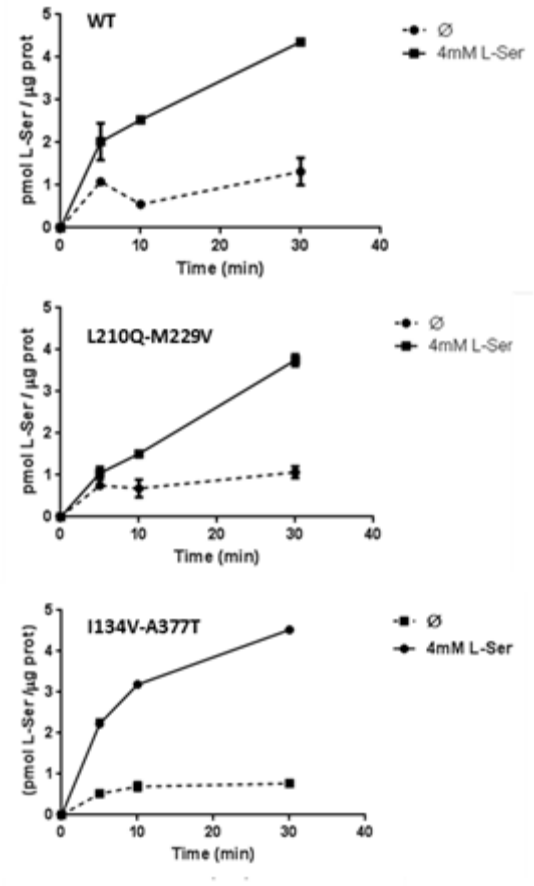


Figura 7. Transporte en proteoliposomas de SteT-WT, SteT-L210Q-M229V y SteT-I134V-A377T usando 10µM de L-Ser radioactiva. (Ø) Proteoliposomas sin 4mM de L-Ser en su interior (4 mM L-Ser) Proteoliposomas que contienen 4 mM de L-Ser en su interior.

8.2.5 Ensayos de crystalización

Después de que se confirmó que ambos mutantes eran funcionales, se iniciaron los ensayos de cristalización. Decidimos comenzar con I134V-A377T, porque aunque ambos mutantes tienen una estabilidad similar, el mutante I134V-A377T expresa 3 veces más que L210Q-M229V. Se realizaron pruebas de diálisis previas a los ensayos de cristalización. Después de purificar con 2xCMC de DM el mutante por cromatografía de afinidad, se intercambió por 2xCMC de OG durante la concentración y la SEC, buscando disminuir el tamaño de micela para aumentar los contactos proteína-proteína. Además se añadió un 10% de glicerol para aumentar la estabilidad del mutante. Después de la purificación, la proteína se concentró a 2 mg/ml y 8 mg/ml y se sembraron las 2

concentraciones en microplacas con pocillos que contenían diferentes condiciones de cristalización adecuadas para proteínas de membrana (Membfac, Memplus, Memstart, Memsys y Memgold). Dos temperaturas diferentes se utilizaron para estas placas (20°C y 4°C). No se formaron cristales en este cribaje, únicamente se observó proteína precipitada y en algunos casos desnaturalizada. En un segundo ensayo utilizando únicamente DM, para aumentar la estabilidad del complejo proteína-detergente, se observaron la formación de esferolitos. Esto sugiere que la no formación de estos probablemente es debida a la inestabilidad de los mutantes en el cambio de OG ya es un detergente de cadena más corta. Además la formación de esferolitos fue en condiciones cercanas a las condiciones de cristalización de AdiC, GadC and ApcT. En éste ensayo, tampoco se observaron cristales. Finalmente, la selección usando Cymal-6 permitió obtener cristales en microplaca, en condiciones con PEG 1500 25,0% w / v, PEG 4000 4,286% w / v de acetato de sodio 0,1 M y pH 4,6 (Figura 8). Curiosamente, también aparecieron esferolitos en condiciones similares a las encontradas en el ensayo en microplaca en DM.

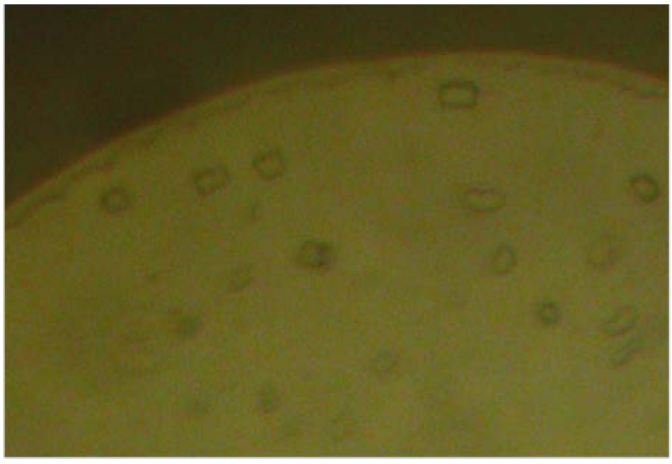


Figura 8. Cristales de I134V-A377T a 5 mg/ml purificada con 2xCMC de Cymal-6, 20mM tris pH 8 y 150mM NaCl. Condiciones del pocillo: PEG 1500 25,0% w / v, PEG 4000 4,286% w / v de acetato de sodio 0,1 M y pH 4,6.

Nuestros resultados sugieren que la mutagénesis al azar puede ser una herramienta muy útil y rápida para encontrar mutantes de una proteína de membrana dado que incrementan la estabilización cuando se combina con un sistema como el de la GFP divida como método decribaje, evitando falsos positivo durante la expresión heteróloga. El uso de la GFP permitió el cribaje de una manera rápida de la biblioteca de mutantes en búsqueda de aquellos mutantes que incrementan la expresión y la estabilidad respecto a los WT, no siendo necesario purificar los mutantes previamente. Este método nos permitió encontrar I134V-A377T y L210Q-M229V, que no sólo se expresan más que WT, pero también son capaces de concentrarse en altas concentraciones, necesarios para los ensayos de cristalización en diferentes detergentes, en contraste de WT. Además, este método nos permitió obtener cristales del mutante I134V-A377T en un cribado de microplacas. Estos cristales se han de reproducir en placa grande para y para determinar la reproducibilidad y si es así su difracción.

BIBLIOGRAPHY

9 BIBLIOGRAPHY

Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H.R., and Iwata, S. (2003). Structure and mechanism of the lactose permease of Escherichia coli. Science *301*, 610–615.

Alexandrov, A.I., Mileni, M., Chien, E.Y.T., Hanson, M.A., and Stevens, R.C. (2008). Microscale fluorescent thermal stability assay for membrane proteins. Structure *16*, 351–359.

Bartoccioni, P., Del Rio, C., Ratera, M., Kowalczyk, L., Baldwin, J.M., Zorzano, A., Quick, M., Baldwin, S.A., Vázquez-Ibar, J.L., and Palacín, M. (2010). Role of transmembrane domain 8 in substrate selectivity and translocation of SteT, a member of the L-amino acid transporter (LAT) family. J. Biol. Chem. *285*, 28764–28776.

Bass, R.B., Strop, P., Barclay, M., and Rees, D.C. (2002). Crystal structure of Escherichia coli MscS, a voltage-modulated and mechanosensitive channel. Science *298*, 1582–1587.

Bertran, J., Werner, A., Moore, M.L., Stange, G., Markovich, D., Biber, J., Testar, X., Zorzano, A., Palacin, M., and Murer, H. (1992). Expression cloning of a cDNA from rabbit kidney cortex that induces a single transport system for cystine and dibasic and neutral amino acids. Proc Natl Acad Sci U S A *89*, 5601–5605.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. *72*, 248–254.

Bröer, S. (2008). Amino Acid Transport Across Mammalian Intestinal and Renal Epithelia. Physiol Rev *88*, 249–286.

Bröer, S., and Palacín, M. (2011). The role of amino acid transporters in inherited and acquired diseases. Biochem. J. *436*, 193–211.

Busch, A.E., Herzer, T., Waldegger, S., Schmidt, F., Palacin, M., Biber, J., Markovich, D., Murer, H., and Lang, F. (1994). Opposite directed currents induced by the transport of dibasic and neutral amino acids in Xenopus oocytes expressing the protein rBAT. J. Biol. Chem. *269*, 25581–25586.

Cabantous, S., Terwilliger, T.C., and Waldo, G.S. (2005). Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. Nat. Biotechnol. *23*, 102–107.

Cabantous, S., and Waldo, G.S. (2006). In vivo and in vitro protein solubility assays using split GFP. Nat Meth *3*, 845–854.

Casagrande, F., Ratera, M., Schenk, A.D., Chami, M., Valencia, E., Lopez, J.M., Torrents, D., Engel, A., Palacin, M., and Fotiadis, D. (2008). Projection Structure of a Member of the Amino Acid/Polyamine/Organocation Transporter Superfamily. J Biol Chem *283*, 33240–33248.

Chairoungdua, A., Segawa, H., Kim, J.Y., Miyamoto, K., Haga, H., Fukui, Y., Mizoguchi, K., Ito, H., Takeda, E., Endou, H., et al. (1999). Identification of an amino acid transporter associated with the cystinuria-related type II membrane glycoprotein. J. Biol. Chem. *274*, 28845–28848.

Chang, G., Spencer, R.H., Lee, A.T., Barclay, M.T., and Rees, D.C. (1998). Structure of the MscL homolog from Mycobacterium tuberculosis: a gated mechanosensitive ion channel. Science *282*, 2220–2226.

Chen, Y.-J., Pornillos, O., Lieu, S., Ma, C., Chen, A.P., and Chang, G. (2007). X-ray structure of EmrE supports dual topology model. Proc. Natl. Acad. Sci. U.S.A. *104*, 18999–19004.

Cherezov, V., Rosenbaum, D.M., Hanson, M.A., Rasmussen, S.G.F., Thian, F.S., Kobilka, T.S., Choi, H.-J., Kuhn, P., Weis, W.I., Kobilka, B.K., et al. (2007). High-Resolution Crystal Structure of an Engineered Human β2-Adrenergic G Protein–Coupled Receptor. Science *318*, 1258–1265.

Chillarón, J., Estévez, R., Mora, C., Wagner, C.A., Suessbrich, H., Lang, F., Gelpí, J.L., Testar, X., Busch, A.E., Zorzano, A., et al. (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*, 17761–17770.

Chillarón, J., Font-Llitjós, M., Fort, J., Zorzano, A., Goldfarb, D.S., Nunes, V., and Palacín, M. (2010). Pathophysiology and treatment of cystinuria. Nat Rev Nephrol *6*, 424–434.

Chillarón, J., Roca, R., Valencia, A., Zorzano, A., and Palacín, M. (2001). Heteromeric amino acid transporters: biochemistry, genetics, and physiology. Am. J. Physiol. Renal Physiol. *281*, F995–1018.

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

Dalbey, R.E., Wang, P., and Kuhn, A. (2011). Assembly of Bacterial Inner Membrane Proteins. Annual Review of Biochemistry *80*, 161–187.

Dello Strologo, L., Pras, E., Pontesilli, C., Beccia, E., Ricci-Barbini, V., De Sanctis, L., Ponzone, A., Gallucci, M., Bisceglia, L., Zelante, L., et al. (2002). Comparison between SLC3A1 and SLC7A9 cystinuria patients and carriers: a need for a new classification. J. Am. Soc. Nephrol. *13*, 2547–2553.

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

Drew, D., Lerch, M., Kunji, E., Slotboom, D.-J., and De Gier, J.-W. (2006). Optimization of membrane protein overexpression and purification using GFP fusions. Nat. Methods *3*, 303–313.

Ehrmann, M., Boyd, D., and Beckwith, J. (1990). Genetic analysis of membrane protein topology by a sandwich gene fusion approach. Proc. Natl. Acad. Sci. U.S.A. *87*, 7574–7578.

Ermolova, N.V., Smirnova, I.N., Kasho, V.N., and Kaback, H.R. (2005). Interhelical packing modulates conformational flexibility in the lactose permease of Escherichia coli. Biochemistry *44*, 7669–7677.

Eshaghi, S. (2009). High-throughput expression and detergent screening of integral membrane proteins. Methods Mol. Biol. *498*, 265–271.

Fagerberg, L., Jonasson, K., Von Heijne, G., Uhlén, M., and Berglund, L. (2010). Prediction of the human membrane proteome. Proteomics *10*, 1141–1149.

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

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

Feliubadaló, L., Font, M., Purroy, J., Rousaud, F., Estivill, X., Nunes, V., Golomb, E., Centola, M., Aksentijevich, I., Kreiss, Y., et al. (1999). Non-type I cystinuria caused by mutations in SLC7A9, encoding a subunit (bo,+AT) of rBAT. Nat. Genet. *23*, 52–57.

Fernández, E., Carrascal, M., Rousaud, F., Abián, J., Zorzano, A., Palacín, M., and Chillarón, J. (2002). rBAT-b0,+AT heterodimer is the main apical reabsorption system for cystine in the kidney. Am J Physiol Renal Physiol *283*, F540–F548.

Forrest, L.R., Zhang, Y.-W., Jacobs, M.T., Gesmonde, J., Xie, L., Honig, B.H., and Rudnick, G. (2008). Mechanism for alternating access in neurotransmitter transporters. Proc. Natl. Acad. Sci. U.S.A. *105*, 10338–10343.

Fort, J., De la Ballina, L.R., Burghardt, H.E., Ferrer-Costa, C., Turnay, J., Ferrer-Orta, C., Usón, I., Zorzano, A., Fernández-Recio, J., Orozco, M., et al. (2007). The structure of human 4F2hc ectodomain provides a model for homodimerization and electrostatic interaction with plasma membrane. J. Biol. Chem. *282*, 31444–31452.

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

Fucile, S., Palma, E., Martínez-Torres, A., Miledi, R., and Eusebi, F. (2002). The single-channel properties of human acetylcholine α7 receptors are altered by fusing α7 to the green fluorescent protein. PNAS *99*, 3956–3961.

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

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

García-Fruitós, E., González-Montalbán, N., Morell, M., Vera, A., Ferraz, R., Arís, A., Ventura, S., and Villaverde, A. (2005). Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. Microbial Cell Factories *4*, 27.

Gasol, E., Jiménez-Vidal, M., Chillarón, J., Zorzano, A., and Palacín, M. (2004). Membrane topology of system xc- light subunit reveals a re-entrant loop with substrate-restricted accessibility. J. Biol. Chem. *279*, 31228–31236.

Grisshammer, R., and Tate, C.G. (1995). Overexpression of integral membrane proteins for structural studies. Q. Rev. Biophys. *28*, 315–422.

Gvritishvili, A.G., Leung, K.W., and Tombran-Tink, J. (2010). Codon Preference Optimization Increases Heterologous PEDF Expression. PLoS ONE *5*, e15056.

Hammon, J., Palanivelu, D.V., Chen, J., Patel, C., and Minor, D.L. (2008). A green fluorescent protein screen for identification of well-expressed membrane proteins from a cohort of extremophilic organisms. Protein Science NA–NA.

Hendrick, J.P., and Hartl, F.U. (1995). The role of molecular chaperones in protein folding. FASEB J. *9*, 1559–1569.

Hsieh, J.M., Besserer, G.M., Madej, M.G., Bui, H.-Q., Kwon, S., and Abramson, J. (2010). Bridging the gap: a GFP-based strategy for overexpression and purification of membrane proteins with intra and extracellular C-termini. Protein Sci. *19*, 868–880.

Hunte, C., and Michel, H. (2002). Crystallisation of membrane proteins mediated by antibody fragments. Current Opinion in Structural Biology *12*, 503–508.

Iwata, S. (2003). Methods and Results in Crystallization of Membrane Proteins (Internat'l University Line).

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

Junge, F., Schneider, B., Reckel, S., Schwarz, D., Dötsch, V., and Bernhard, F. (2008). Large-scale production of functional membrane proteins. Cellular and Molecular Life Sciences *65*, 1729–1755.

Kaback, H.R. (1986). Active transport in Escherichia coli: passage to permease. Annu Rev Biophys Biophys Chem *15*, 279–319.

Kaback, H.R. (1987). Molecular biology of active transport: from membrane to molecule to mechanism. Harvey Lect. *83*, 77–105.

Kaira, K., Oriuchi, N., Imai, H., Shimizu, K., Yanagitani, N., Sunaga, N., Hisada, T., Tanaka, S., Ishizuka, T., Kanai, Y., et al. (2008). l-type amino acid transporter 1

and CD98 expression in primary and metastatic sites of human neoplasms. Cancer Sci. *99*, 2380–2386.

Kalatzis, V., Cherqui, S., Antignac, C., and Gasnier, B. (2001). Cystinosin, the protein defective in cystinosis, is a H+-driven lysosomal cystine transporter. EMBO J *20*, 5940–5949.

Kaleeba, J.A.R., and Berger, E.A. (2006). Broad target cell selectivity of Kaposi's sarcoma-associated herpesvirus glycoprotein-mediated cell fusion and virion entry. Virology *354*, 7–14.

Kanai, Y., Segawa, H., Miyamoto, K. i, Uchino, H., Takeda, E., and Endou, H. (1998). Expression cloning and characterization of a transporter for large neutral amino acids activated by the heavy chain of 4F2 antigen (CD98). J. Biol. Chem. *273*, 23629–23632.

Kang, H.J., Lee, C., and Drew, D. (2013). Breaking the barriers in membrane protein crystallography. Int. J. Biochem. Cell Biol. *45*, 636–644.

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

Kim, R.B. (2006). Transporters and drug discovery: why, when, and how. Mol. Pharm. *3*, 26–32.

Klepsch, M.M., Persson, J.O., and De Gier, J.-W.L. (2011). Consequences of the overexpression of a eukaryotic membrane protein, the human KDEL receptor, in Escherichia coli. J. Mol. Biol. *407*, 532–542.

Kobayashi, K., Ohnishi, A., Promsuk, J., Shimizu, S., Kanai, Y., Shiokawa, Y., and Nagane, M. (2008). Enhanced tumor growth elicited by L-type amino acid transporter 1 in human malignant glioma cells. Neurosurgery *62*, 493–503; discussion 503–504.

Koth, C.M.M., and Payandeh, J. (2009). Strategies for the cloning and expression of membrane proteins. Adv Protein Chem Struct Biol *76*, 43–86.

Kowalczyk, L., Ratera, M., Paladino, A., Bartoccioni, P., Errasti-Murugarren, E., Valencia, E., Portella, G., Bial, S., Zorzano, A., Fita, I., et al. (2011). Molecular basis of substrate-induced permeation by an amino acid antiporter. Proc. Natl. Acad. Sci. U.S.A. *108*, 3935–3940.

Kunji, E.R.S., Chan, K.W., Slotboom, D.J., Floyd, S., O'Connor, R., and Monné, M. (2005). Eukaryotic membrane protein overproduction in Lactococcus lactis. Curr. Opin. Biotechnol. *16*, 546–551.

Labrou, N.E. (2010). Random mutagenesis methods for in vitro directed enzyme evolution. Curr. Protein Pept. Sci. *11*, 91–100.

Langosch, D., Hofmann, M., and Ungermann, C. (2007). The role of transmembrane domains in membrane fusion. Cell. Mol. Life Sci. *64*, 850–864.

Lefkowitz, R.J. (2007). Seven transmembrane receptors: something old, something new. Acta Physiologica *190*, 9–19.

Lemieux, M.J., Song, J., Kim, M.J., Huang, Y., Villa, A., Auer, M., Li, X.-D., and Wang, D.-N. (2003). Three-dimensional crystallization of the Escherichia coli glycerol-3-phosphate transporter: a member of the major facilitator superfamily. Protein Sci. *12*, 2748–2756.

Lewinson, O., Lee, A.T., and Rees, D.C. (2008a). The funnel approach to the precrystallization production of membrane proteins. J. Mol. Biol. *377*, 62–73.

Lewinson, O., Lee, A.T., and Rees, D.C. (2008b). The funnel approach to the precrystallization production of membrane proteins. J. Mol. Biol. *377*, 62–73.

Li, B., Nowak, N.M., Kim, S.-K., Jacobson, K.A., Bagheri, A., Schmidt, C., and Wess, J. (2005). Random mutagenesis of the M3 muscarinic acetylcholine receptor expressed in yeast: identification of second-site mutations that restore function to a coupling-deficient mutant M3 receptor. J. Biol. Chem. *280*, 5664–5675.

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

Malle, E., Zhou, H., Neuhold, J., Spitzenberger, B., Klepsch, F., Pollak, T., Bergner, O., Ecker, G.F., and Stolt-Bergner, P.C. (2011). Random mutagenesis of the prokaryotic peptide transporter YdgR identifies potential periplasmic gating residues. J. Biol. Chem. *286*, 23121–23131.

Mancia, F., and Love, J. (2010). High-throughput expression and purification of membrane proteins. J. Struct. Biol. *172*, 85–93.

Martinez Molina, D., Cornvik, T., Eshaghi, S., Haeggström, J.Z., Nordlund, P., and Sabet, M.I. (2008). Engineering membrane protein overproduction in Escherichia coli. Protein Sci. *17*, 673–680.

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

Miller, J.L., and Tate, C.G. (2011). Engineering an ultra-thermostable β(1)-adrenoceptor. J. Mol. Biol. *413*, 628–638.

Nicklin, P., Bergman, P., Zhang, B., Triantafellow, E., Wang, H., Nyfeler, B., Yang, H., Hild, M., Kung, C., Wilson, C., et al. (2009). Bidirectional Transport of Amino Acids Regulates mTOR and Autophagy. Cell *136*, 521–534.

Ohkame, H., Masuda, H., Ishii, Y., and Kanai, Y. (2001). Expression of L-type amino acid transporter 1 (LAT1) and 4F2 heavy chain (4F2hc) in liver tumor lesions of rat models. Journal of Surgical Oncology *78*, 265–272.

Ormö, M., Cubitt, A.B., Kallio, K., Gross, L.A., Tsien, R.Y., and Remington, S.J. (1996). Crystal structure of the Aequorea victoria green fluorescent protein. Science *273*, 1392–1395.

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

Palacín, M., Bertran, J., Chillarón, J., Estévez, R., and Zorzano, A. (2004). Lysinuric protein intolerance: mechanisms of pathophysiology. Mol. Genet. Metab. *81 Suppl 1*, S27–37.

Palacín, M., Borsani, G., and Sebastio, G. (2001). The molecular bases of cystinuria and lysinuric protein intolerance. Curr. Opin. Genet. Dev. *11*, 328–335.

Palacín, M., Nunes, V., Font-Llitjós, M., Jiménez-Vidal, M., Fort, J., Gasol, E., Pineda, M., Feliubadaló, L., Chillarón, J., and Zorzano, A. (2005). The genetics of heteromeric amino acid transporters. Physiology (Bethesda) *20*, 112–124.

Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T.C., and Waldo, G.S. (2005). Engineering and characterization of a superfolder green fluorescent protein. Nature Biotechnology *24*, 79–88.

Pfeiffer, R., Loffing, J., Rossier, G., Bauch, C., Meier, C., Eggermann, T., Loffing-Cueni, D., Kühn, L.C., and Verrey, F. (1999). Luminal heterodimeric amino acid transporter defective in cystinuria. Mol. Biol. Cell *10*, 4135–4147.

Privé, G.G. (2007). Detergents for the stabilization and crystallization of membrane proteins. Methods *41*, 388–397.

Qin, Z., Freitas, E., Sullivan, R., Mohan, S., Bacelieri, R., Branch, D., Romano, M., Kearney, P., Oates, J., Plaisance, K., et al. (2010). Upregulation of xCT by KSHV-encoded microRNAs facilitates KSHV dissemination and persistence in an environment of oxidative stress. PLoS Pathog. *6*, e1000742.

Rasmussen, S.G.F., Choi, H.-J., Rosenbaum, D.M., Kobilka, T.S., Thian, F.S., Edwards, P.C., Burghammer, M., Ratnala, V.R.P., Sanishvili, R., Fischetti, R.F., et al. (2007). Crystal structure of the human β2 adrenergic G-protein-coupled receptor. Nature *450*, 383–387.

Razvi, A., and Scholtz, J.M. (2006). Lessons in stability from thermophilic proteins. Protein Sci *15*, 1569–1578.

Reig, N., Chillarón, J., Bartoccioni, P., Fernández, E., Bendahan, A., Zorzano, A., Kanner, B., Palacín, M., and Bertran, J. (2002a). The light subunit of system b(o,+) is fully functional in the absence of the heavy subunit. EMBO J. *21*, 4906–4914.

Reig, N., Chillarón, J., Bartoccioni, P., Fernández, E., Bendahan, A., Zorzano, A., Kanner, B., Palacín, M., and Bertran, J. (2002b). The light subunit of system bo,+ is fully functional in the absence of the heavy subunit. EMBO J *21*, 4906–4914.

Reig, N., Del Rio, C., Casagrande, F., Ratera, M., Gelpí, J.L., Torrents, D., Henderson, P.J.F., Xie, H., Baldwin, S.A., Zorzano, A., et al. (2007). Functional and structural characterization of the first prokaryotic member of the L-amino acid transporter (LAT) family: a model for APC transporters. J. Biol. Chem. *282*, 13270–13281.

Ressl, S., Terwisscha van Scheltinga, A.C., Vonrhein, C., Ott, V., and Ziegler, C. (2009). Molecular basis of transport and regulation in the Na(+)/betaine symporter BetP. Nature *458*, 47–52.

Rodríguez-Banqueri, A., Kowalczyk, L., Palacín, M., and Vázquez-Ibar, J.L. (2012). Assessment of membrane protein expression and stability using a split green fluorescent protein reporter. Anal. Biochem. *423*, 7–14.

Rodríguez-Carmona, E., Cano-Garrido, O., Seras-Franzoso, J., Villaverde, A., and García-Fruitós, E. (2010). Isolation of cell-free bacterial inclusion bodies. Microbial Cell Factories *9*, 71.

Sanders, C.R., and Myers, J.K. (2004). Disease-Related Misassembly of Membrane Proteins. Annual Review of Biophysics and Biomolecular Structure *33*, 25–51.

Sarkar, C.A., Dodevski, I., Kenig, M., Dudli, S., Mohr, A., Hermans, E., and Plückthun, A. (2008). Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. Proc. Natl. Acad. Sci. U.S.A. *105*, 14808–14813.

Sato, H., Shiiya, A., Kimata, M., Maebara, K., Tamba, M., Sakakura, Y., Makino, N., Sugiyama, F., Yagami, K., Moriguchi, T., et al. (2005). Redox imbalance in cystine/glutamate transporter-deficient mice. J. Biol. Chem. *280*, 37423–37429.

Sato, H., Tamba, M., Okuno, S., Sato, K., Keino-Masu, K., Masu, M., and Bannai, S. (2002). Distribution of cystine/glutamate exchange transporter, system x(c)-, in the mouse brain. J. Neurosci. *22*, 8028–8033.

Schaffner, W., and Weissmann, C. (1973). A rapid, sensitive, and specific method for the determination of protein in dilute solution. Anal. Biochem. *56*, 502–514.

Serrano-Vega, M.J., Magnani, F., Shibata, Y., and Tate, C.G. (2008). Conformational thermostabilization of the β1-adrenergic receptor in a detergent-resistant form. PNAS *105*, 877–882.

Shaffer, P.L., Goehring, A., Shankaranarayanan, A., and Gouaux, E. (2009). Structure and mechanism of a Na+-independent amino acid transporter. Science *325*, 1010–1014.

Simon Newstead, H.K. (2007). High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in Saccharomyces cerevisiae. Proceedings of the National Academy of Sciences of the United States of America *104*, 13936–13941.

Singh, S.K., Yamashita, A., and Gouaux, E. (2007). Antidepressant binding site in a bacterial homologue of neurotransmitter transporters. Nature *448*, 952–956.

Smirnova, I.N., and Kaback, H.R. (2003). A mutation in the lactose permease of Escherichia coli that decreases conformational flexibility and increases protein stability. Biochemistry *42*, 3025–3031.

Sonoda, Y., Newstead, S., Hu, N.-J., Alguel, Y., Nji, E., Beis, K., Yashiro, S., Lee, C., Leung, J., Cameron, A.D., et al. (2011). Benchmarking Membrane Protein Detergent Stability for Improving Throughput of High-Resolution X-ray Structures. Structure *19*, 17–25.

Stark, M.J.R. (1987). Multicopy expression vectors carrying the lac represser gene for regulated high-level expression of genes in Escherichia coli. Gene *51*, 255–267.

Tate, C.G. (2012). A crystal clear solution for determining G-protein-coupled receptor structures. Trends in Biochemical Sciences *37*, 343–352.

Tate, C.G., and Schertler, G.F.X. (2009). Engineering G protein-coupled receptors to facilitate their structure determination. Curr. Opin. Struct. Biol. *19*, 386–395.

Tate, S.S., Yan, N., and Udenfriend, S. (1992). Expression cloning of a Na(+)-independent neutral amino acid transporter from rat kidney. Proc Natl Acad Sci U S A *89*, 1–5.

Thévenin, D., and Lazarova, T. (2008). Stable interactions between the transmembrane domains of the adenosine A2A receptor. Protein Sci *17*, 1188–1199.

Torrents, D., Estévez, R., Pineda, M., Fernández, E., Lloberas, J., Shi, Y.B., Zorzano, A., and Palacín, M. (1998). Identification and characterization of a membrane protein (y+L amino acid transporter-1) that associates with 4F2hc to encode the amino acid transport activity y+L. A candidate gene for lysinuric protein intolerance. J. Biol. Chem. *273*, 32437–32445.

Tsiboli, P., Konstantinidis, G., Skendros, Y., Katsani, A., and Choli-Papadopoulou, D.T. (1997). Identification of post-translational modified amino acids. Amino Acids *13*, 13–23.

Vázquez-Ibar, J.L., Guan, L., Svrakic, M., and Kaback, H.R. (2003). Exploiting luminescence spectroscopy to elucidate the interaction between sugar and a tryptophan residue in the lactose permease of Escherichia coli. Proc. Natl. Acad. Sci. U.S.A. *100*, 12706–12711.

Vázquez-Ibar, J.L., Guan, L., Weinglass, A.B., Verner, G., Gordillo, R., and Kaback, H.R. (2004). Sugar recognition by the lactose permease of Escherichia coli. J. Biol. Chem. *279*, 49214–49221.

Vergis, J.M., Purdy, M.D., and Wiener, M.C. (2010). A high-throughput differential filtration assay to screen and select detergents for membrane proteins. Anal. Biochem. *407*, 1–11.

Verrey, F., Jack, D.L., Paulsen, I.T., Saier, M.H., Jr, and Pfeiffer, R. (1999). New glycoprotein-associated amino acid transporters. J. Membr. Biol. *172*, 181–192.

Verrey, F., Meier, C., Rossier, G., and Kühn, L.C. (2000). Glycoprotein-associated amino acid exchangers: broadening the range of transport specificity. Pflugers Arch. *440*, 503–512.

Wagner, C.A., Lang, F., and Bröer, S. (2001a). Function and structure of heterodimeric amino acid transporters. Am J Physiol Cell Physiol *281*, C1077–C1093.

Wagner, C.A., Lang, F., and Bröer, S. (2001b). Function and structure of heterodimeric amino acid transporters. Am J Physiol Cell Physiol *281*, C1077–C1093.

Wallin, E., and Von Heijne, G. (1998). Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. Protein Sci. *7*, 1029–1038.

Warne, T., Chirnside, J., and Schertler, G.F.X. (2003). Expression and purification of truncated, non-glycosylated turkey beta-adrenergic receptors for crystallization. Biochim. Biophys. Acta *1610*, 133–140.

Warne, T., Serrano-Vega, M.J., Baker, J.G., Moukhametzianov, R., Edwards, P.C., Henderson, R., Leslie, A.G.W., Tate, C.G., and Schertler, G.F.X. (2008). Structure of a beta1-adrenergic G-protein-coupled receptor. Nature *454*, 486–491.

Wells, R.G., and Hediger, M.A. (1992). Cloning of a rat kidney cDNA that stimulates dibasic and neutral amino acid transport and has sequence similarity to glucosidases. Proc. Natl. Acad. Sci. U.S.A. *89*, 5596–5600.

Weyand, S., Shimamura, T., Yajima, S., Suzuki, S., Mirza, O., Krusong, K., Carpenter, E.P., Rutherford, N.G., Hadden, J.M., O'Reilly, J., et al. (2008). Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. Science *322*, 709–713.

White, S.H. (2009). Biophysical dissection of membrane proteins. Nature *459*, 344–346.

Wong, F.H., Chen, J.S., Reddy, V., Day, J.L., Shlykov, M.A., Wakabayashi, S.T., and Saier, M.H., Jr (2012). The amino acid-polyamine-organocation superfamily. J. Mol. Microbiol. Biotechnol. *22*, 105–113.

Yamashita, A., Singh, S.K., Kawate, T., Jin, Y., and Gouaux, E. (2005). Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters. Nature *437*, 215–223.

Zhao, R., Shin, D.S., Diop-Bove, N., Ovits, C.G., and Goldman, I.D. (2011). Random mutagenesis of the proton-coupled folate transporter (SLC46A1), clustering of mutations, and the bases for associated losses of function. J. Biol. Chem. *286*, 24150–24158.

Zheng, H., Pearsall, E.A., Hurst, D.P., Zhang, Y., Chu, J., Zhou, Y., Reggio, P.H., Loh, H.H., and Law, P.-Y. (2012). Palmitoylation and membrane cholesterol stabilize μ-opioid receptor homodimerization and G protein coupling. BMC Cell Biol. *13*, 6.

Zhou, Z., Zhen, J., Karpowich, N.K., Goetz, R.M., Law, C.J., Reith, M.E.A., and Wang, D.-N. (2007). LeuT-desipramine structure reveals how antidepressants block neurotransmitter reuptake. Science *317*, 1390–1393.

Zhou, Z., Zhen, J., Karpowich, N.K., Law, C.J., Reith, M.E.A., and Wang, D.-N. (2009). Antidepressant specificity of serotonin transporter suggested by three LeuT-SSRI structures. Nat. Struct. Mol. Biol. *16*, 652–657.