UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE CHIMICHE

SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE MOLECOLARI

INDIRIZZO: SCIENZE CHIMICHE

CICLO XXV

Structural analysis of SulP/SLC26 anion transporters

Direttore della Scuola: Ch.mo Prof. Antonino Polimeno

Coordinatore d'indirizzo: Ch.mo Prof. Antonino Polimeno

Supervisore: Ch.mo Prof. Roberto Battistutta

Dottoranda: Greta Bonetto

Credo di poter affermare che nella ricerca scientifica
né il grado di intelligenza né la capacità di eseguire
e portare a termine il compito intrapreso
siano fattori essenziali per la riuscita e per la soddisfazione personale.
Nell'uno e nell'altro contano maggiormente la totale dedizione
e il chiudere gli occhi davanti alle difficoltà:
in tal modo possiamo affrontare i problemi che altri,
più critici e più acuti, non affronterebbero.

Rita Levi Montalcini

Table of contents

Abbreviations

3D	Three-Dimensional
A8-35	Amphipol A8-35
Aa	Amino acid
AAAP	Amino Acid/Auxin Permease
ABC	ATP-Binding Cassette
ACP	Acyl Carrier Protein
AFM	Atomic Force Microscopy
AGCS	Alanine or Glycine:Cation Symporter
ALA (A)	Alanine
APC	Amino Acid-Polyamine-Organocation
ARG (R)	Arginine
ASA	Anti-Sigma factor Antagonist
ASN (N)	Asparagine
ASP (D)	Aspartic acid
ATP	Adenosine TriPhosphate
BCCT	Betaine/Carnitine/Coline Transporter
BODIPY	Boron-DIPYrromethene
bp	base pair
cAMP	cyclic Adenosine MonoPhosphate
CCC	Cation-Chloride Cotransporter
CCD	Congenital Chloride Diarrhoea
CCM	CO_2 Concentrating Mechanism
CD	Circular Dichroism
CECF	Continuous-Exchange Cell-Free
CF	Cell-Free
CFTR	Cystic Fibrosis Transmembrane conductance Regulator
CHO	Chinese Hamster Ovary
CMC	Critical Micelle Concentration
CPM	N-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]-maleimide
CSP	Chemical Shift Perturbation
CTP	Cytidine-5'-Triphosphate
CYS (C)	Cysteine

Abbreviations

Da	Dalton
D-CF	Detergent based mode Cell-Free
DDM	n-dodecyl β-D-maltoside
DHPC	l,2-diheptanoyl-sn-glycero-3-phosphocholine
DLS	Dynamic Light Scattering
DM	n-Decyl-β-D-Maltopyranoside
DM NG	Decyl Maltose Neopentyl Glycol
DPC	n-dodecylphosphocholine
DNA	DeoxyriboNucleic Acid
DRA	Downregulated in Adenomas
DSF	Differential Scanning Fluorimetry
DTDST	Diastrophic Dysplasia Sulphate Transporter
DTT	DiThioThreitol
EDTA	EthyleneDiamineTetraacetic Acid
FM	Feeding Mixture
FOS-12	n-dodecylphosphocholine
FOS-16	n-hexadecylphosphocholine
FPLC	Fast Protein Liquid Chromatography
FRET	Fluorescence Resonance Energy Transfer
FT	Flow Through
GAP	GTPase Activating Protein
GDP	Guanosine DiPhosphate
GFP	Green Fluorescent Protein
GLN (Q)	Glutamine
GLU (E)	Glutamic acid
GLY (G)	Glycine
GTP	Guanosine TriPhosphate
HAAAP	Hydroxy/Aromatic Amino Acid Permease
HEK	Human Embryonic Kidney
HIS (H)	Histidine
HSQC	Heteronuclear Single Quantum Coherence
ID	IDentity
IHC	Inner Hair Cell

ILE (I)	Isoleucine
IMAC	Immobilized-Metal Affinity Chromatography
IMP	IntraMembrane Particle
IPTG	IsoPropyl β-D-1-ThioGalactopyranoside
IVS	InterVening Sequence
LB	Luria Bertani
L-CF	Lipid based mode Cell-Free
LDAO	n-Dodecyl-N,N-Dimethylamine-N-Oxide
LEU (L)	Leucine
LMPC	1-myristoyl-2-hydroxy-sn-glycero-3-phosphocholine
LMPG	lyso-myristoylphosphatidylglycerol
LPPG	1-palmitoyl-2-hydroxy-sn-glycero-3-[phospho-RAC-(1-glycerol)]
LOV	Light-Oxygen-Voltage
LYS (K)	Lysine
MAP1S	Microtubule Associated Protein 1S
MES	4-MorpholineEthaneSulfonic acid
MET (M)	Methionine
MR	Molecular Replacement
MSP	Membrane Scaffold Protein
MW	Molecular Weight
MWCO	Molecular Weight Cut-Off
NA	Not Available
NADPH	Nicotinamide Adenine Dinucleotide PHosphate
NAT/NCS2	Nucleobase-Ascorbate Transporter/Nucleobase:Cation Symporter-2
NCS1	Nucleobase:Cation Symporter-1
ND	NanoDisc
NG	n-Nonyl-β-D-Glucopyranoside
NLC	NonLinear Capacitance
NMR	Nuclear Magnetic Resonance
NSS	Neurotransmitter:Sodium Symporter
NTA	NitriloTriacetic Acid
NTP	Nucleoside TriPhosphate
OAS	O-AcetylSerine
OASTL	O-AcetylSerine (Thiol)Lyase

OD	Optical Density
OG	Octyl-β-D-Glucopyranoside
OHC	Outer Hair Cell
ON	Over Night
OMIM	Online Mendelian Inheritance in Man
PAT-1	Putative Anion Transporter-1
PBS	Phosphate Buffered Saline
P-CF	Precipitate forming mode Cell-Free
PCR	Polymerase Chain Reaction
Pd	Polydispersion
PDB	Protein Data Bank
PDZ	Post-synaptic density 95/Discs large/Zona occludens 1
PEG	PolyEthylene Glycol
PHE (F)	Phenylalanine
PL	Pellet
PoPMuSiC	Prediction of Protein Mutations Stability Changes
PRO (P)	Proline
PS	Pendred Syndrome
PKA	Protein Kinase A
RF	Restriction-Free
RM	Reaction Mixture
RNA	RiboNucleic Acid
RT	Room Temperature
RuBisCO	Ribulose-1,5-Bisphosphate Carboxylase Oxygenase
SAD	Single-wavelength Anomalous Dispersion
SANS	Small Angle Neutron Scattering
Sat-1	Sulphate Anion Transporter-1
SDS-PAGE	Sodium Dodecyl Sulphate - PolyAcrylamide Gel Electrophoresis
SEC	Size Exclusion Chromatography
SER (S)	Serine
SN	Supernatant
SNHL	SensoriNeutral Hearing Loss
SSM	Solid Supported Membrane
SSS	Solute:Sodium Symporter

STAS	Sulfate Transporter and Anti-Sigma factor antagonist
SulP/SLC26	Sulfate Permease/Solute Carrier 26
SUMO	Small Ubiquitin-like MOdifier
Tat1	Testis Anion Transporter 1
TC	Transporter Classification
TCDB	Transporter Classification Database
TEM	Transmission Electron Microscopy
THR (T)	Threonine
TLS	Translation/Libration/Screw parameterization
Tm	Temperature of melting
TMs	TransMembrane segments
TRP (W)	Tryptophan
TYR (Y)	Tyrosine
UM	n-Undecyl-β-D-Maltopyranoside
UTP	Uridine-5'-TriPhosphate
VAL (V)	Valine
VAPA	Vesicle-Associated membrane protein-associated Protein A
wt	wild type

Abstract

The subject of this thesis is a family of anion transporters known as SulP/SLC26 (Sulfate Permease/Solute Carrier 26) family, a large and ubiquitous family of membrane proteins capable of transporting a wide variety of monovalent and divalent anions, whose members were found in eubacteria, plants, fungi, and mammals. The clinical relevance of the SulP/SLC26 gene family has been highlighted with the identification of pathogenic mutations related to hereditary genetic human diseases with diverse symptoms that arise as a result of the different substrate specificities and tissue localizations of the different transporters, such as dystrophic dysplasia (SLC26A2), congenital chloride diarrhoea (SLC26A3) and Pendred syndrome (SLC26A4). The SulP/SLC26 family belongs to the APC (Amino Acid-Polyamine-Organocation) superfamily, one of the largest superfamily of secondary carriers. While some members of other families of the APC superfamily have been structurally characterized, very little is known about the molecular organization of the SulP/SLC26 proteins and no high-resolution three-dimensional structure of full-length sequences is available. The SulP/SLC26 anion transporters share a common structural organization: a highly conserved transmembrane domain and a less conserved cytoplasmic C-terminal portion mainly composed of a STAS domain. The name STAS (Sulfate Transporter and Anti-Sigma factor antagonist) is due to a remote but statistically significant sequence similarity with bacterial ASA (Anti-Sigma factor Antagonist) proteins (Aravind and Koonin, 2000). The bacterial ASA proteins are functionally and structurally well characterized in their 3D structure both by NMR spectroscopy and X-ray crystallography. Unlike these proteins, the STAS domains present in anion transporters are poorly characterized in terms of both their function and structure. Despite the fact that their precise role is unclear, the STAS domains play a fundamental role in the function/regulation of SulP/SLC26 anion transporters. In particular, it has been proposed that the STAS domain, like ASA proteins, could have a role in protein/protein interaction; for instance the STAS domains of SCL26A3, -A4, -A6 and -A9 interact with the R domain of CFTR (Cystic Fibrosis Transmembrane conductance Regulator), the transmembrane protein involved in cystic fibrosis disease. So far three 3D structures of STAS domains from different species are available in literature, two from bacteria and one from mammalian, the latter solved during my Master Degree Thesis in the same laboratory where I've attended the PhD. The structural characterization of the full-length

SulP/SLC26 transporters and of their STAS domains is fundamental for the comprehension of their mode of action and it is an essential step for the understanding of the functional consequences of the mutations responsible for related pathologies.

To address this issue, one part of my PhD project focused on the production and the structural characterization of STAS domains from different species, and mutants of the STAS domain whose 3D structure have been solved, in order to study the anion-binding site and the possible role of the STAS domain in the transport. We identified a fundamental residue for the proper function of the transporter, probably implicated in the anion translocation within the transmembrane domain.

The other part of the project dealt with the production of a selection of full-length SulP/SLC26 transporters from different orthologs, both Prokaryotes and Eukaryotes. To this aim, in collaboration with Prof. Frank Bernhard at the Johann Wolfgang Goethe University of Frankfurt (Germany), I used the cell-free (CF) expression method, an emerging technique for the large-scale production of membrane proteins for structural studies. Sample properties after post-translational solubilization have been analyzed by evaluation of homogeneity and protein stability. This is the first quality evaluation of the SulP/SLC26 transporters produced by CF expression mode in quantities appropriate for structural approaches.

Riassunto

Il soggetto di questa tesi è una famiglia di trasportatori di anioni nota come famiglia SulP/SLC26 (*Sulfate Permease/Solute Carrier* 26), una grande ed ubiquitaria famiglia di proteine di membrana in grado di trasportare un'ampia varietà di anioni monovalenti e divalenti, i cui membri sono stati trovati in eubatteri, piante, funghi, e mammiferi. La rilevanza clinica della famiglia genica SulP/SLC26 è stata evidenziata dall'identificazione di mutazioni patogene connesse a malattie umane genetiche ed ereditarie, con diversi sintomi che sorgono come risultato delle differenti specificità di substrato e localizzazioni tissutali dei differenti trasportatori, come la displasia distrofica (SLC26A2), la diarrea cloridrica congenita (SLC26A3) e la sindrome di Pendred (SLC26A4). La famiglia SulP/SLC26 appartiene alla superfamiglia APC (*Amino Acid-Polyamine-Organocation*), una delle più grandi superfamiglie di trasportatori secondari. Mentre alcuni membri di altre famiglie della superfamiglia APC sono stati caratterizzati strutturalmente, si sa molto poco riguardo l'organizzazione molecolare delle proteine SulP/SLC26 e non è disponibile nessuna struttura tridimensionale ad elevata risoluzione delle intere sequenze. I trasportatori di anioni SulP/SLC26 condividono un'organizzazione strutturale simile: un dominio transmembrana altamente conservato ed una porzione C-terminale meno conservata principalmente composta da un dominio STAS. Il nome STAS (*Sulfate Transporter and Anti-Sigma factor antagonist*) è dovuto ad una similarità di sequenza remota ma statisticamente significativa con le proteine batteriche ASA (*Anti-Sigma factor Antagonist*) (Aravind and Koonin, 2000). Le proteine batteriche ASA sono state ben caratterizzate funzionalmente e strutturalmente nella loro struttura 3D sia mediante la spettroscopia NMR sia mediante cristallografia a raggi X. A differenza di queste proteine, i domini STAS presenti nei trasportatori di anioni sono stati poco caratterizzati sia in termini della loro funzione, sia della loro struttura. Nonostante il fatto che il loro preciso ruolo non sia chiaro, i domini STAS svolgono un ruolo fondamentale nella funzione/regolazione dei trasportatori di anioni SulP/SLC26. In particolare, è stato proposto che il dominio STAS, come le proteine ASA, potesse svolgere un ruolo nell'interazione proteina/proteina; per esempio, domini STAS di SCL26A3, -A4, -A6 e -A9 interagiscono con il dominio R di CFTR (*Cystic Fibrosis Transmembrane conductance Regulator*), la proteina transmembrana coinvolta nella fibrosi cistica. Finora tre strutture 3D di domini STAS provenienti da specie diverse sono disponibili in

letteratura, due da batteri ed una da mammifero, quest'ultima risolta durante la mia tesi di laurea specialistica nello stesso laboratorio dove ho frequentato il dottorato. La caratterizzazione strutturale degli interi trasportatori SulP/SLC26 e dei loro domini STAS è fondamentale per la comprensione del loro modo di azione ed è una fase essenziale per comprendere le conseguenze funzionali delle mutazioni responsabili delle patologie collegate.

Per raggiungere questo obiettivo, una parte del mio progetto di dottorato si è focalizzata sulla produzione e caratterizzazione strutturale dei domini STAS provenienti da diverse specie, e mutanti del dominio STAS la cui struttura 3D è stata risolta per studiare il sito di legame dell'anione ed il possibile ruolo del dominio STAS nel trasporto. È stato identificato un residuo fondamentale per il corretto funzionamento del trasportatore, probabilmente implicato nella traslocazione dell'anione all'interno del dominio transmembrana.

L'altra parte del progetto riguarda la produzione di una selezione di trasportatori SulP/SLC26 interi provenienti da diversi ortologhi, sia Procarioti che Eucarioti. Per questo scopo, in collaborazione con il Prof. Frank Bernhard presso l'università Johann Wolfgang Goethe di Francoforte (Germania), utilizzai il metodo di espressione *cell-free* (CF), una tecnica emergente per la produzione a larga scala di proteine di membrana per studi strutturali. Le proprietà dei campioni dopo la solubilizzazione post-traduzionale sono state analizzate mediante la valutazione di omogeneità e della stabilità della proteina. Questa è la prima valutazione della qualità dei trasportatori SulP/SLC26 prodotti mediante il modo di espressione CF in quantità appropriate per approcci strutturali.

1. Introduction

1.1. The APC superfamily

The Amino Acid-Polyamine-Organocation (APC) superfamily appears to be the second largest superfamily of secondary carriers and it is composed of 11 families displayed in the table 1.1 (Wong et al., 2012).

Table 1.1 Families within the APC superfamily

Families within the APC superfamily			**3D structure**
2.A.3	Amino Acid-Polyamine-Organocation (APC) family	SLC7, SLC12	AdiC, ApcT
2.A.15	Betaine/Carnitine/Coline Transporter (BCCT) family		CaiT, BetP
2.A.18	Amino Acid/Auxin Permease (AAAP) family	SLC32, SLC36, SLC38	
2.A.21	Solute:Sodium Symporter (SSS) family	SLC5	vSGLT
2.A.22	Neurotransmitter:Sodium Symporter (NSS) family	SLC6	LeuT
2.A.25	Alanine or Glycine:Cation Symporter (AGCS) family		
2.A.30	Cation-Chloride Cotransporter (CCC) family		
2.A.39	Nucleobase:Cation Symporter-1 (NCS1) family		Mhp1
2.A.40	Nucleobase-ascorbate transporter (NAT) / Nucleobase:Cation Symporter-2 (NCS2) family	SLC23	UraA
2.A.42	Hydroxy/Aromatic Amino Acid Permease (HAAAP) family		
2.A.53	Sulfate permease (SulP) family	SLC26	

Based on www.TCDB.org. In the last column there are the transporters that have been structurally characterized.

The APC superfamily includes members that function as solute:cation symporters and solute:solute antiporters; they occur in bacteria, archaea, yeast, fungi, unicellular eukaryotic protists, slime molds, plants and animals (Saier et al., 2000). They are predicted to possess 10-14 transmembrane α-helical spanners and they vary in length, being as small as 350 residues and as large as 850 residues; the smaller proteins are generally of prokaryotic origin while the larger ones are of eukaryotic origin.

Some members of some families of the APC superfamily have been structurally characterized [LeuT (Yamashita et al., 2005), Mhp1 (Weyand et al., 2008), vSGLT (Faham et al., 2008), BetP (Ressl et al., 2009), AdiC (Gao et al., 2009), ApcT (Shaffer et al., 2009), CaiT (Schulze et al., 2010), UraA (Lu et al., 2011)], while very little is known

about the molecular organization of the SulP/SLC26 proteins and no high-resolution 3D structure of full-length sequences is available.

According to the phylogenetic analysis (figure 1.1), the closest member to the SulP/SLC26 family whose 3D structure has been solved is the *Escherichia coli* uracil/H^+ symporter UraA (belonging to the NAT/NCS2 family), that shows a novel structural fold with 14 transmembrane segments divided into two inverted repeats (Lu et al., 2011).

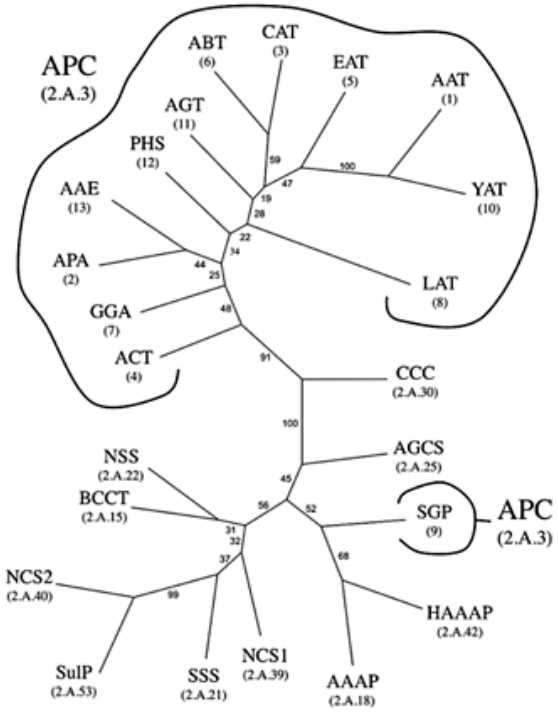


Figure 1.1 *Phylogenetic tree for the APC superfamily using the proteins belonging to this superfamily in TCDB (Transporter Classification Database). Family abbreviations are presented with TC (Transporter Classification) family numbers in parentheses. Small numbers adjacent to the branches present the "bootstrap" values, indicating the reliability of the branching order (Wang et al., 2012).*

1.2. The SulP/SLC26 family

The SulP/SLC26 family is a large and ubiquitous family of anion transporters, whose members were found in archea, bacteria, plants, fungi, and animals. The eukaryotic SulP/SLC26 transporters are large proteins comprised of 700-1000 amino acids, while prokaryotic are smaller (400-600 amino acids) and the individual family members have 21-43% amino acid identity. Many bacteria and eukaryotes possess multiple SulP/SLC26 family paralogues. The relatedness of the family members within and among species is depicted in the evolutionary tree shown in figure 1.2. Many of these

proteins have been functionally characterized as anion transporters or anion: anion exchangers that transport a wide variety of monovalent and divalent substrates, including sulfate (SO_4^{2-}), chloride (Cl^-), bicarbonate (HCO_3^-), iodide (I^-), oxalate, formate, hydroxyl, mannose and fructose (Felce and Saier, 2004; Detro-Dassen et al., 2008; Dorwart et al., 2008).

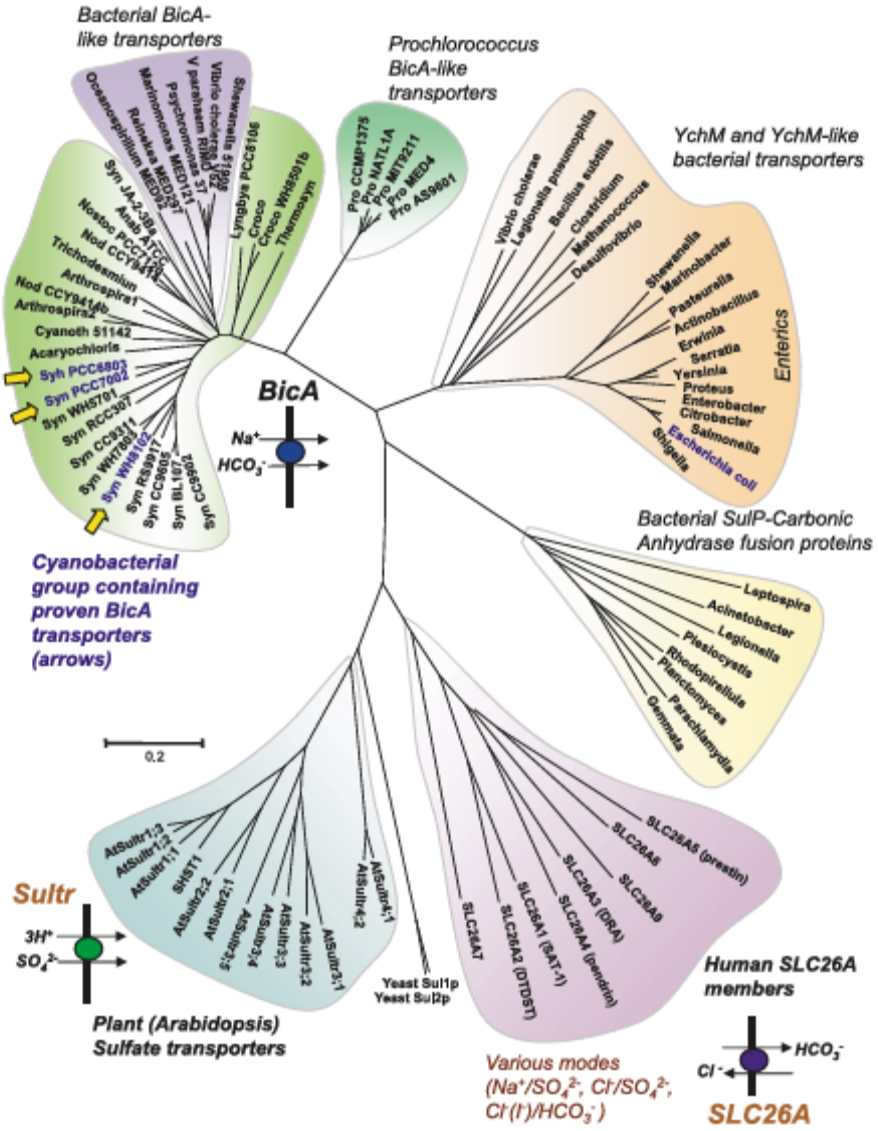


Figure 1.2 *Phylogenetic tree showing the relatedness of selections of SulP/SLC26 family members from diverse groupings (Price and Howitt, 2011).*

1.2.1. The transport function of the human SulP/SLC26 transporters

The human family of SulP/SLC26 transporters consists of 11 members that are expressed in polarized cells in organs such as the kidney, pancreas, intestine, and liver,

where they mediate SO_4^{2-}, Cl^- and HCO_3^- transport across the plasma membrane. SLC26A1 and -A2 transport SO_4^{2-} (Markovich, 2012; Ohana et al., 2012), whereas SLC26A3, -A4, -A6 exchange Cl^- for a wide range of anions, including HCO_3^-, I^-, oxalate, formate and SO_4^{2-}. SLC26A9 can function in three discrete physiological modes: n(Cl^-)/HCO_3^- exchanger, Cl^- channel, and Na^+-anion cotransporter (Chang et al., 2009). SLC26A7 is a recently identified Cl^--base exchanger and/or Cl^- transporter (Xu et al., 2009). Recently also SLC26A11 has been identified as an electrogenic n(Cl^-)/HCO_3^- exchanger (Xu et al., 2011). To date, the protein products of SLC26A8 have poorly understood transport capabilities, and SLC26A10 is a pseudogene. The mammalian SLC26A5 is the motor protein in the outer hair cells of the cochlea (Zheng et al., 2000).

Table 1.2 The Human SulP/SLC26 Gene Family

Human Gene Name	**Protein name**	**Predominant substrates**	**Tissue distribution and cellular /subcellular expression**	**Disease association**
SLC26A1	Sat-1	SO_4^{2-}, HCO_3^-, oxalate	Liver, kidney (basolateral), pancreas, brain	
SLC26A2	DTDST	SO_4^{2-}, Cl^-	Widespread	Diastrophic dysplasia, atelosteogenesis type II, achondrogenesis type IB, multiple epiphyseal dysplasia type 4
SLC26A3	DRA	Cl^-, SO_4^{2-}, HCO_3^-, oxalate	Intestine, sweat gland, pancreas, prostate	Congenital chloride diarrhoea (CCD)
SLC26A4	pendrin	Cl^-, HCO_3^-, I^-, formate	Inner ear, kidney, thyroid	Pendred syndrome, non-syndromic deafness (DFNB4)
SLC26A5	prestin		Inner ear	
SLC26A6	PAT-1	SO_4^{2-}, Cl^-, HCO_3^-, oxalate, formate	Widespread	
SLC26A7		SO_4^{2-}, Cl^-, oxalate, HCO_3^-	Kidney (basolateral)	
SLC26A8	Tat1	SO_4^{2-}, Cl^-, oxalate	Testis (sperm)	
SLC26A9		SO_4^{2-}, Cl^-, oxalate	Lung, stomach	
SLC26A10			pseudogene	
SLC26A11		SO_4^{2-}	Placenta, kidney, brain	

Based on www.bioparadigms.org/slc/menu.asp; original version (Mount and Romero 2004).

1.2.2. Structural features of the SulP/SLC26 transporters

The topology of SulP/SLC26 family members is not known and 10-14 transmembrane helices have been predicted from hydropathy plots and predictive algorithms with intracellular N- and C-terminus (Figure 1.3).

Much of the homology between SulP/SLC26 exchangers is found within the hydrophobic core of transmembrane domains. One region of homology encompasses the 22 amino-acid "sulfate transport" consensus signature (Prosite, PS01130), which was initially defined by the comparison of the first mammalian family members with homologs in lower organisms. There is a second cluster of invariant residues at the C-terminal end of the hydrophobic core of the proteins, in a conserved segment defined as Saier motif (Saier et al., 1999). This region includes the triplet -NQE-, which is conservatively variable only in Slc26a8 (-NQD-). Many of the SulP/SLC26 proteins end with a class I PDZ (Post-synaptic density 95/Discs large/Zona occludens 1) interaction motif; the exceptions include human SLC26A1, -A2, -A4, -A5, and -A11 (Mount and Romero, 2004).

The less conserved cytoplasmic C-terminal portion of the SulP/SLC26 anion transporters is mainly composed of a STAS (Sulfate Transporter and Anti-Sigma factor antagonist) domain, whose name is due to a remote but statistically significant sequence similarity with the bacterial AntiSigma-factor Antagonists (ASA) proteins, typified by Bacillus subtilis SpoIIAA (Aravind and Koonin, 2000). This homology suggests that the SpoIIAA proteins and the SulP/SLC26 STAS domains have similar structures. The physiological and/or mechanistic roles of the STAS domains in the SulP/SLC26 exchangers are not still well known, whereas the function of SpoIIAA is well characterized (chapter 1.3.1). However the existence of disease-associated mutations in this domain underscores its importance.

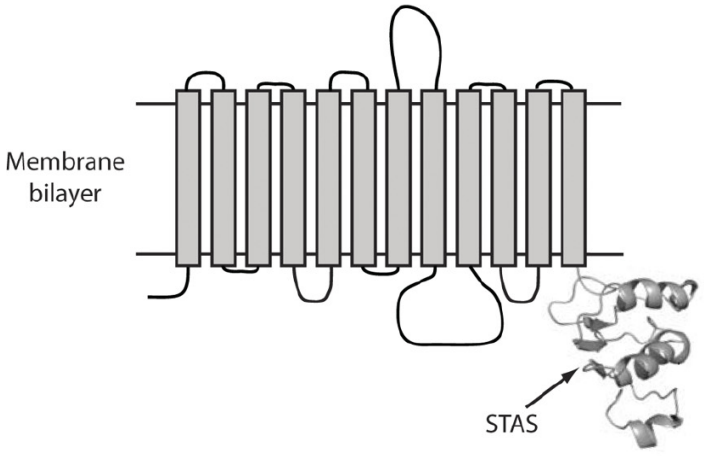


Figure 1.3 *Predicted topology of a general SulP/SLC26 transporter (Compton et al., 2011).*

All SulP/SLC26 transporters appear to be assembled as dimers or tetramers composed of identical subunits (Detro-Dassen et al., 2008; Mio et al., 2008; Compton et

al., 2011). The so far unique low-resolution structures of SulP/SLC26 anion transporters available in literature are that of rat SLC26A5/prestin, solved by TEM (Transmission Electron Microscopy), and that of *Yersinia enterocolitica* Slc26A2 protein, solved by SANS (Small Angle Neutron Scattering). The first structure of purified prestin (Figure 1.19) exhibits a bullet-shaped fourfold symmetric molecule with an inner cavity and suggests a tetrameric arrangement of subunits (Mio et al., 2008). While the structure of bacterial Slc26A2 protein (Figure 1.4) suggests that the protein forms a dimer stabilized via its transmembrane core and that the cytoplasmic STAS domain projects away from the transmembrane domain and is not involved in dimerization. It was hypothesized that large movements of the STAS domain underlie the conformational changes that occur during transport (Compton et al., 2011).

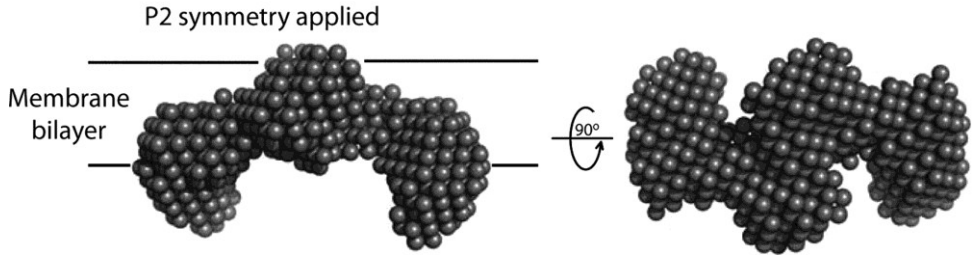


Figure 1.4 *Low-resolution model of Yersinia enterocolitica Slc26A2 generated by the program DAMMIN from the 5 mg/ml SANS data at the contrast match point of the Fos-choline-12 detergent. The main bulk of the protein, encompassing the dimerization domain, lies mostly within a plane that corresponds to the depth of a membrane bilayer (30 Å) with two globular domains extending from this plane, thus revealing a multidomain organization. The model on the right is generated from the one on the left by a rotation of 90° around their long axes (Compton et al., 2011).*

1.2.3. The SulP/SLC26 family in human diseases

Mutations in some SulP/SLC26 members, including their STAS domains, result in human diseases with diverse symptoms that arise as a result of the different substrate specificities and tissue localizations of the different transporters. Mutations in SLC26A2 lead to four autosomal recessively inherited chondrodysplasias (diastrophic dysplasia, atelosteogenesis type II, achondrogenesis type IB and multiple epiphyseal dysplasia type 4). SLC26A3/DRA is associated with congenital chloride diarrhoea (CCD), a rare autosomal recessive disorder with defective Cl^-/HCO_3^- exchange in the ileum and colon. SLC26A4 is associated with Pendred syndrome, a common form of hereditary hearing loss that manifests primarily as profound sensorineural deafness which results from disruption of cochlear endolymphatic fluid HCO_3^- buffering, loss of the endocochlear

potential and subsequent oxidative stress; and non-syndromic deafness, DFNB4. These disorders highlight the important roles of SulP/SLC26 transporters, including their STAS domains, in normal human physiology (Dawson and Markovich, 2005). The above human disease phenotypes have been replicated in knockout mice. Mouse-specific pathological phenotypes have been observed in mice genetically deficient in other SulP/Slc26 gene products. These include severe hearing impairment observed in Slc26a5-/- mice, urolithiasis in Slc26a1-/- and Slc26a6-/- mice, gastric achlorhydria in Slc26a7-/-and Slc26a9-/- mice, distal renal tubular acidosis in Slc26a7-/- mice, and male infertility in Slc26a8-/- mice (Sharma et al., 2011).

1.2.3.1. SulP/SLC26 and CFTR

Cystic fibrosis is a chronic recessive disease caused by mutations in the gene encoding the CFTR, a member of the ABC (ATP-Binding Cassette) family of membrane transporters. CFTR functions as a cAMP regulated Cl^- channel that is regulated by PKA (Protein Kinase A) phosphorylation and it is expressed mainly in the apical plasma membrane of epithelial tissues, where it has a crucial role in regulating fluid secretion in the airways, salivary glands, intestine and genital tract. When CFTR fails to work properly, it produces an imbalance in salt and fluid transport that results in thickened secretions in all these tissues. In the airways there is an accumulation of sticky mucus that eventually results in respiratory failure (Gray, 2004).

The primary defect in cystic fibrosis is a problem with HCO_3^--driven fluid secretion, caused by the inability of mutant forms of CFTR to activate Cl^-/HCO_3^- exchange (Choi et al., 2001). Subsequent studies from the same group (Ko et al., 2002) extended the link between CFTR and HCO_3^- by demonstrating that CFTR specifically up-regulated the activity of select SulP/SLC26 transporters (SLC26A3, -A4 and -A6). Later work (Ko et al., 2004) localized the relevant interacting regions to the R-domain of CFTR and the STAS domain of SulP/SLC26 transporters. This interaction is enhanced by phosphorylation of the R-domain by PKA and is modulated by PDZ binding scaffold proteins that tether the two transporters into a multimeric complex with other regulatory proteins (Figure 1.5). Furthermore, while the interaction between R-CFTR and the STAS domain of SLC26A3, -A4, -A6 increases their transport activity; the same interaction with SLC26A9 STAS domain inhibits transport (Chang et al., 2009 (2)). Recently it has

been reported that also prestin interacts with CFTR, but the functional role of this interaction and the interacting regions aren't clear yet (Homma et al., 2010).

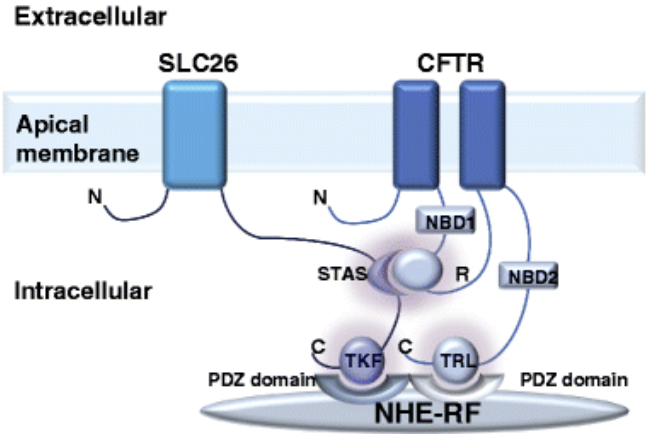


Figure 1.5 *Cartoon showing hypothesized apical macromolecular signaling complexes promoted by PDZ protein-facilitated interactions between the CFTR and SulP/SLC26 transporters. PDZ domain binding motif of SLC26A3, TKF, is shown in this figure; the counterpart for SLC26A6 is TRL. For CFTR, the motif is TRL as shown. This motif is lacking in SLC26A4, but the potential for regulation by STAS domain interactions with the R domain of CFTR is preserved (Fong 2012).*

1.3. The STAS domain

The STAS domain is present throughout phylogeny from eubacteria onward as a conserved fold encoded by highly divergent amino acid sequences. STAS domain proteins include small STAS domain-only polypeptides and larger, multidomain polypeptides. The intramolecular partner domains present in these bacterial multidomain STAS proteins are remarkably diverse. In bacteria, STAS domain function has been intensively studied in the context of the regulation of the large family of sigma factors that bind to RNA polymerase to confer transcriptional target gene specificity (chapter 1.3.1). Additional investigations have focused on STAS domain function in various signaling contexts such as blue light phototransduction (LOV-STAS, Light-Oxygen-Voltage sensing, domains, chapter 1.3.2). More recently, STAS domains of SulP/SLC26 anion transport proteins from bacteria and plants have been studied genetically and structurally, leading to new insights on STAS function. The first structure of an engineered mammalian SulP/SLC26 STAS domain (chapter 1.3.4.1) and growing numbers of SulP/SLC26 STAS domain structures (chapter 1.3.4.2 and 1.3.4.3) and interacting proteins has provided new contexts in which to evaluate STAS domain function. Most importantly, connections between intrinsic activities or interactions of

STAS domains and the regulation of anion transport function in SulP/SLC26 anion transporter polypeptides are emerging (Sharma et al., 2011).

1.3.1. STAS domains as bacterial anti-σ factor antagonists: SpoIIAA

One of the most extensively studied models of biological stress response is the sporulation of the soil bacterium *Bacillus subtilis* in response to nutrient deprivation.

Early in sporulation, the cell divides asymmetrically to form two unequal compartments, the mother cell and the prespore (Figure 1.6 b). Differential gene expression occurs in these two compartments through the activation of specific transcription factors known as sigma factors. The prespore-specific sigma factor, σ^F or SpoIIAC, is the first of these to be activated, leading to a cascade of downstream activation of forespore-specific gene expression. The three proteins that regulate σ^F are SpoIIAA, which is a STAS domain protein, SpoIIAB and SpoIIE. SpoIIAB is an anti-σ factor, which can either bind to and inhibit σ^F (Figure 1.6 a.1), or act as a specific kinase for SpoIIAA, the anti-anti-σ factor, by transferring the γ-phosphate of ATP to Ser58 of SpoIIAA (Figure 1.6 a.2). Before asymmetric septation, SpoIIAA is in a phosphorylated state, which, under physiological conditions, does not interact with SpoIIAB, leaving the latter free to inhibit σ^F. However, at the time of asymmetric septation, SpoIIE, the specific phosphatase for SpoIIAA-phosphate, becomes active in the prespore. The dephosphorylation of SpoIIAA (Figure 1.6 a.3) allows it to interact with SpoIIAB, thereby releasing σ^F activity (Figure 1.6 a.4).

SpoIIAA binds and hydrolyzes GTP and, to a lesser degree, ATP. However, the GTP-binding and hydrolase activities of SpoIIAA have unclear physiological functions (Clarkson et al., 2003).

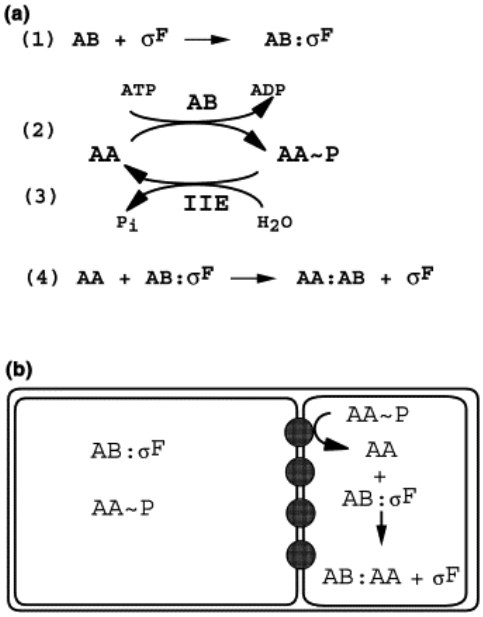


Figure 1.6 *The interactions among SpoIIA (AA), SpoIIAB (AB) and SpoIIE (IIE and shaded circles in b) proteins in the regulation of σ^F activity (Seavers et al., 2001).*

The structures of SpoIIAA have been solved by heteronuclear NMR (Kovacs et al., 1998) and X-ray crystallography (Figure 1.7) (Seavers et al., 2001) in phosphorylated and unphosphorylated forms, and in complex with SpoIIAB bound to either ADP or ATP (Masuda et al., 2004). Published structures of additional STAS domain proteins include the NMR solution structure of *Thermotoga maritima* putative anti-σ antagonist TM1442 in phosphorylated and unphosphorylated states (Etezady-Esfarjani et al., 2006), the NMR and crystal structures of *T. maritima* putative anti-σ antagonist TM1081 (Serrano et al., 2010), and the crystal structure of the putative stressosome component RsbS from *Moorella thermoacetica* (Quin et al., 2008).

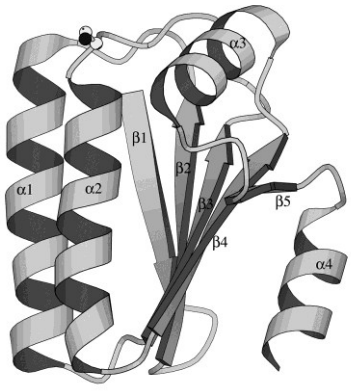


Figure 1.7 *3D crystal structure of SpoIIAA from Bacillus sphaericus (PDB code = 1H4Y). The molecule contains five β strands (β1-β5) and four α helices (α1-α4) in the order β1-β2-α1-β3-α2-β4-α3-β5-α4 (Seavers et al., 2001).*

The SpoIIAA structure has a central β-sheet that is flanked by two α-helices on one side and by two small α-helices at the end of the β-sheet. An important difference between the SpoIIAA protein and the human SulP/SLC26 transporter STAS domains is a highly variable loop or InterVening Sequence (IVS) between the first predicted α-helix and the third β-strand of the SulP/SLC26 STAS domains (secondary structure nomenclature for ASAs, Figure 1.8). Secondary structure prediction algorithms suggest that the IVS is likely to be disordered in solution. This variable loop is the site of significant insertions in the mammalian SulP/SLC26 proteins, of as much as 150 amino acids in the case of SLC26A8. No such insertion is present in bacterial SulP/SLC26 homologues, and this loop is also much shorter in SLC26A11 and in the *Drosophila* and *C. elegans* paralogs; how this variable loop contributes to paralog-specific function and/or regulation is not known. The functional significance of this region is currently unknown (Mount and Romero, 2004; Dorwart et al., 2008).

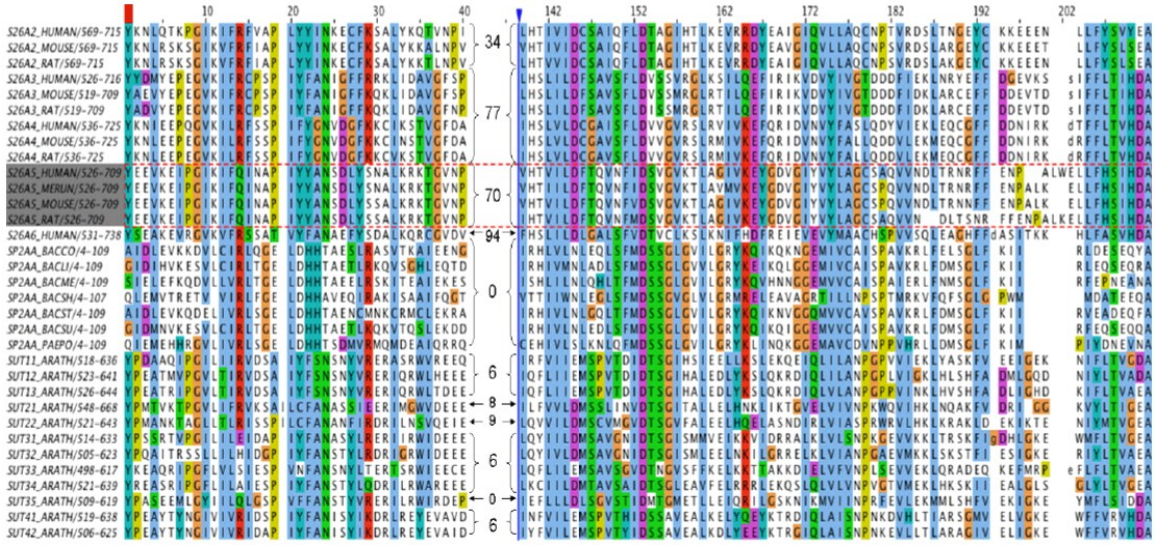


Figure 1.8 *Multiple alignment of a selected set of anti-sigma factor antagonists and SulP/SLC26 anion transporters STAS domains. 14 mammalian STAS domains from human, rat and mouse (S26A) were aligned with 7 anti-sigma factor antagonist (SP2AA) from different bacteria and 12 paralogous STAS domains of Sultr from A. thaliana (SUT). The sequences of the inserts in the variable loop are replaced by the number of amino acid residues (between curly brackets or arrows). The alignment was obtained with the program Jalview, using colour matrix ClustalW.*

1.3.2. STAS domains in phototransduction

Sensing of (blue) light is essential for many organisms in order to be successful in their struggle for survival. Evolution has led to the emergence of four blue-light absorbing photoreceptor protein families, of which the LOV domain containing photoreceptors are

the most widely distributed. YtvA from *Bacillus subtilis* is a blue-light responsive, flavin-binding photoreceptor, built of a light-sensing LOV domain and an NTP-binding STAS domain. Like other regulators of the stress responsive network in *B. subtilis*, YtvA bears a STAS domain that is supposed to be the effector part of the protein or a secondary switch. Both domains are connected by a linker polypeptide. Blue light is absorbed by the flavin mononucleotide chromophore of the LOV domain, triggering a local conformational change that is believed to be transmitted through the agency of the J linker to the STAS domain, which binds BODIPY-GTP (Nakasone and Hellingwerf, 2011). The near complete solution NMR backbone assignment of holo-YtvA in the dark state (Jurk et al., 2011) has demonstrated a dimeric structure with LOV-LOV and STAS-STAS interactions, partially consistent with the proposed mechanism of light-induced conformational change (figure 1.9).

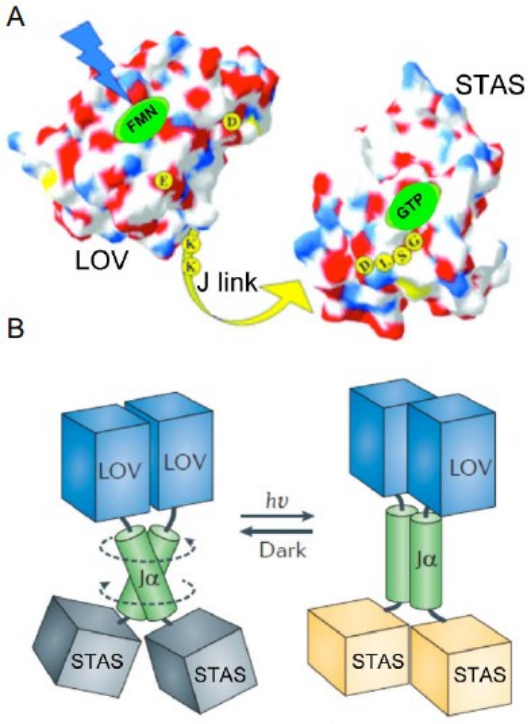


Figure 1.9 A) *B. subtilis YtvA electrostatic surface image from LOV domain crystal structure and from STAS domain (modelled on crystal structure of B. subtilis SpoIIAA).* ***B)*** *Light-induced conformational change of YtvA as imagined from holoprotein structure (Sharma et al., 2011).*

1.3.3. STAS domains in SulP/SLC26 anion transporters: function

In bacteria, ASA proteins function and structure have been intensively studied. More recently, STAS domains of SulP/SLC26 anion transport proteins have been functionally and structurally studied, leading to new insights on STAS function.

The function of the STAS domain is not well understood but, like SpoIIAA, the SulP/SLC26 transporter STAS domains have been hypothesized to be protein-protein interaction domains. Indeed, the STAS domain of SLC26A8 can bind to MgcRacGAP (Tourè et al., 2001), that of SLC26A5 can bind to MAP1S (Bai et al., 2010), the STAS domains of SLC26A3, -A4, -A6, -A8 and -A9 interact with CFTR R-domain (Fong, 2012) and the STAS domain of Sultr1;2 from *A. thaliana* interacts with OASTL (O-AcetylSerine (Thiol)Lyase) (Shibagaki and Grossman, 2010).

Additional information regarding the function of the STAS domains has come from genetic studies on the *A. thaliana* SulP/SLC26 transporters (Sultr). The Sultr1;2 STAS domain has been well functionally characterized as essential for activity, biosynthesis/stability of the transporter and for proper transporter localization (Shibagaki and Grossman, 2004; 2006; Rouached et al., 2005). As it has been shown for Sultr1;2, molecular truncation experiments have indicated that STAS domain integrity is required for optimal surface trafficking and functional expression also for SLC26A3 (Chernova et al., 2003).

Trafficking impairment is noted with most of the examined disease mutations of the SulP/SLC26 STAS domains, which constitute a minority of all SulP/SLC26 human disease mutations. The largest number of distinct disease associated variants in STAS domains, as well as in holoproteins, is found in pendrin. Notably, among IVS sequences of SulP/SLC26 human disease genes, only the IVS of pendrin has been shown to date to harbor disease genes.

1.3.4. STAS domains in SulP/SLC26 anion transporters: structure

Structural information is available for STAS domains of SulP/SLC26 anion transporters of SLC26A5/prestin from *Rattus norvegicus* (Pasqualetto et al., 2010), YchM from *Escherichia coli* (Babu et al., 2010) and of Rv1739c from *Mycobacterium tuberculosis* (Sharma et al., 2010); their structures revealed a common global fold but there are significant differences at the N-terminus where SLC26A5 has an original structure.

1.3.4.1. SLC26A5/prestin STAS domain from Rattus norvegicus

The X-ray crystal structure of the prestin STAS domain from *Rattus norvegicus* with an engineered central deletion of the IVS sequence was solved during my master degree thesis in the same laboratory where I attended the PhD, with the collaboration of Dr. Bellanda, at the Department of Chemical Sciences, University of Padova. Some rat prestin STAS domain constructs turned out to be insoluble when expressed in *E. coli*, and others had a strong tendency to aggregate in solution (Pasqualetto et al., 2008) while prestin STAS constructs without the variable loop, which is predicted to be mostly disordered, showed a better behaviour in solution, more suitable for structural studies.

The core of the structure is composed of a β-sheet of six β-strands, named from β0 to β5. The four C-terminal strands from β2 to β5 are surrounded by five α-helices (α1-α5). The structure is stabilized by two extensive networks of hydrophobic residues, connecting each side of the β-sheet with the nearby helices. The main differences between the structure of prestin STAS and those of bacterial ASA are a long rigid insertion between β1 and β2 composed of a short β-strand (β0) and a series of tight β-turns in the N-terminal portion and in the direction of helix α1. Despite these differences in the N-terminal portion and in helix α1, secondary structure elements β1, β2, β3, β4, β5, α2, α3, α4, and α5 of prestin STAS are structurally and topologically similar to those of bacterial ASA proteins, justifying also from a structural point of view the notion of STAS domains, linking anion transporters and SpoIIAA bacterial proteins.

As the prestin STAS domain is located in the cytosolic portion of a membrane protein, in close proximity to the lipid bilayer, we have tried to identify which part of the molecule faces the membrane surface using a computational approach (PPM program) (Lomize et al., 2006). The program revealed the existence of a non-polar surface region, composed of residues Pro543, Tyr545, Tyr546, Phe651, and Val655 (highlighted in the figure 1.10) that shows a possibility of weak association with a lipid bilayer or with the hydrophobic surface of the protein transmembrane domain (Pasqualetto et al., 2010).

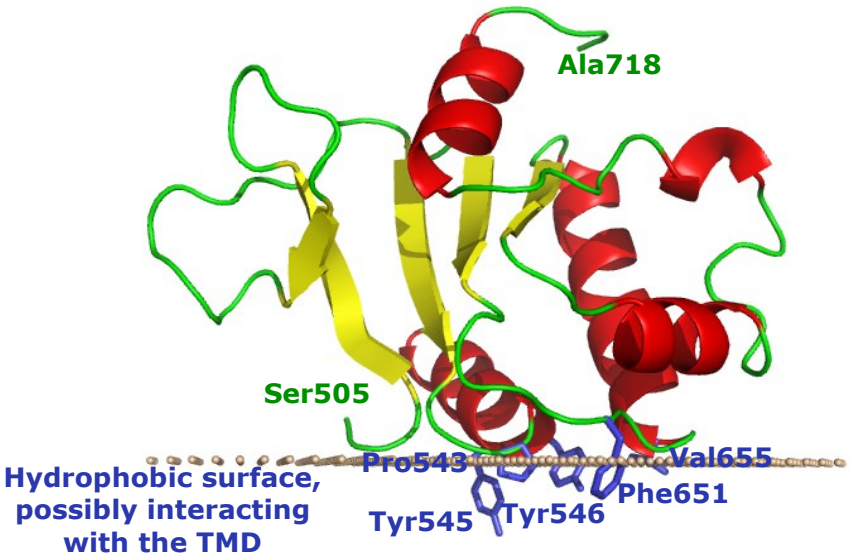


Figure 1.10 *Crystal structure of SLC26A5 STAS domain from Rattus norvegicus (PDB ID code 3LLO) exhibiting a secondary structural element sequence of β1-β0-β2-α1-β3-α2-β4-α3-α4-β5-α5 and the proposed orientation of prestin STAS domain with respect to the membrane (small brown spheres). The side chains of the amino acids which have been detected interacting the lipid bilayer by the PPM program are in blue (Pasqualetto et al., 2010).*

1.3.4.2. YchM STAS domain from Escherichia coli

In *Escherichia coli*, YchM is a C(4)-dicarboxylic acid transporter, an unexpected new function for a prokaryotic member of this transporter family (Karinou et al., 2012). The structure of the YchM STAS domain has been solved using X-ray crystallography. The STAS domain expressed in *E. coli* was found to copurify and cocrystallize in a 1:1 complex with Acyl Carrier Protein (ACP), an essential component of the fatty acid biosynthetic apparatus. The ACP in the crystal was covalently attached to a malonyl-coA moiety which was situated at the ACP-STAS interface. The cocrystal structure showed a specific interaction between the STAS domain of YchM and ACP, suggesting that the STAS domain of YchM could play a role in sequestering ACP to the inner membrane, the site of phospholipid biosynthesis.

The structure the STAS domain begins with β strand β1 antiparallel to the second β strand β2 followed by α1, β3, α2, β4, α3, β5 and α4. In comparison with SpoIIAA, the STAS domain of YchM contains an additional β strand onto its core four-stranded β sheet, with residues 532-541 of the YchM STAS domain creating part of a binding pocket to accommodate the 4'- PPa prosthetic group from ACP. The potential binding site for the SpoIIAB kinase on SpoIIAA does not overlap with its ACP interaction site on STAS

domain of YchM. This binding surface on the YchM STAS domain may provide an additional docking site for other interacting proteins. The structural core features of the STAS domain are likely common to highly related bacterial homologs and shared by other homologous members of the SulP/SLC26 family, including the human proteins that contain larger STAS domains. This conserved core structure is evident in the prestin STAS domain structure that retains a similar overall architecture to the STAS domain of YchM. The main differences with rat prestin STAS are its higher angle of divergence between helices α1 and α2 (Babu et al., 2010; Sharma et al., 2011).

Babu and colleagues suggest that perhaps the reason for which the STAS domains from human SulP/SLC26 transporters have been difficult to crystallize and they were unable to crystallize the STAS domain of YchM on its own, is in part that STAS domains have intrinsically disordered regions in the absence of binding partners, a common theme in many signalling proteins (Babu et al., 2010).

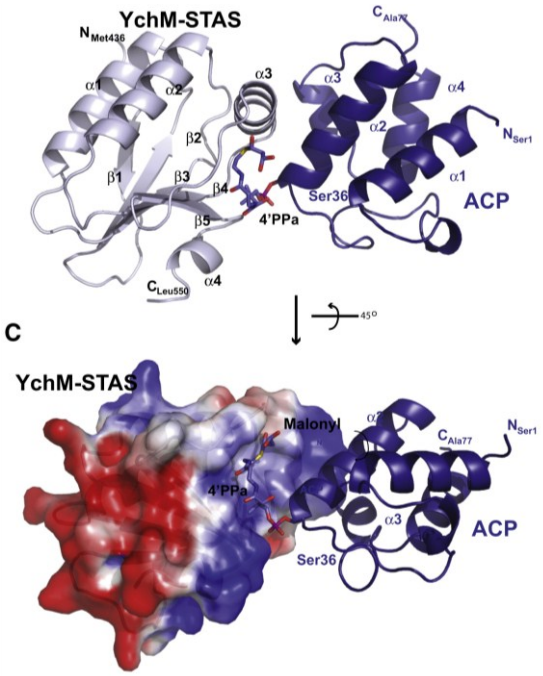


Figure 1.11 *Crystal structure of YchM STAS domain from E. coli in complex with ACP, exhibiting a secondary structural element sequence of β1-β2-α1-β3-α2-β4-α3-β5-α4 (PDB ID code 3NY7) (Babu et al., 2010).*

1.3.4.3. Rv1739c STAS domain from Mycobacterium tuberculosis

The structure of Rv1739c STAS domain from *M. tuberculosis* has been solved using heteronuclear multidimensional NMR spectroscopy. The structure revealed a four-stranded β-sheet with five interspersed α-helices, with the pattern β1-α1-β2-α2-β3-α3-α4-β4-α5. The topological fold of Rv1739c STAS closely resembles that of the anti-σ factor antagonist and that of the rat prestin construct. A notable exception to this similarity is the absence of a β strand within the most N-terminal residues 1-14, observed in the anti-σ factor antagonist structures. The Rv1739c STAS domain construct lacks the 15-aa N-terminal extension present in the crystal structure of the C-terminal truncated, IVS-deleted rat prestin STAS structure. This 15-aa extension includes a "β1 strand" at its extreme N-terminus and a "β0 strand" in the immediately juxtaposed loop. Rat prestin STAS also differs from the STAS domains and Rv1739c in its higher angle of divergence between helices α1 and α2. Rv1739c differed significantly in its secondary structure as compared to YchM with relative reversal of α1 and β2 in the linear sequence and presence of helix α5 that is instead present in prestin STAS domain.

NMR CSP studies and docking calculations suggest two possible nucleotide binding sites in Rv1739c STAS involved in binding GDP and/or GTP. Therefore Rv1739c STAS domain binds and maybe hydrolyzes guanine but the functional importance of these activities remains obscure for Rv1739c as well as for anti-anti-σ SpoIIAA of *B. subtilis* (Sharma et al., 2010; Sharma et al., 2011).

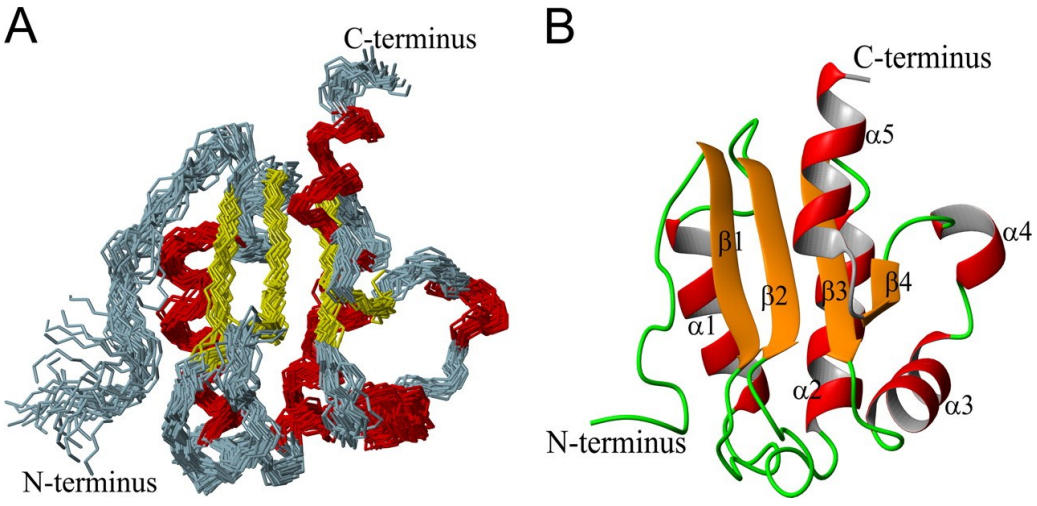


Figure 1.12 *Solution structure of Rv1739c STAS domain from M. tuberculosis exhibiting a secondary structural element sequence of β1-α1-β2-α2-β3-α3-α4-β4-α5 (PDB ID code 2KLN) (Sharma et al., 2010).*

1.4. Selected SulP/SLC26 STAS domains and full-length transporters

This PhD thesis focused on a selection of SulP/SLC26 transporters from different orthologs:

- SLC26A5/prestin from *Rattus norvegicus*, *Gallus gallus* and *Danio rerio* (chapter 1.4.1),
- SCL26A4/pendrin from *Homo sapiens* sapiens (chapter 1.4.2),
- Sultr 1;2 from *Arabidopsis thaliana* (chapter 1.4.3),
- Rv1739c from *Mycobacterium tuberculosis* (chapter 1.4.4),
- BicA from *Synechocystis sp. strain PCC 6803*, *Trichodesmium erythraeum IMS101* and *Thermosynechococcus elongatus BP-1* (chapter 1.4.5).

1.4.1. SLC26A5/prestin

Mammalian prestin is a motor protein abundantly expressed in the lateral membrane of outer hair cells (OHCs) of the mammalian inner ear but it is very different from all the other molecular motors. In fact the linear motors, such as myosin, kinesin and dynein, generate motion by using nucleoside triphosphate hydrolysis to produce conformational changes in proteins. Other motors, such as helicases, ribosomal motors, chaperonins, etc., also require ATP hydrolysis for energy. Prestin, however, is a direct voltage-to-force converter and it is probably unique in the animal kingdom. Moreover prestin is implicated in the mechanism responsible for mammalian sound amplification (Dallos et al., 2006).

1.4.1.1. SLC26A5/prestin and OHCs

Mechanical amplification of acoustic signals is apparently a common feature of vertebrate auditory organs. In non-mammalian vertebrates amplification is produced by stereociliary processes, related to the mechanotransducer channel complex and probably to the phenomenon of fast adaptation (Dallos, 2008). Though these standard stereocilia-based mechanisms to promote amplification persist in mammals, an additional radically different mechanism, which resides in the organ of Corti, evolved: the so called somatic electromotility which refers to the elongation/contraction of the OHCs' cylindrical cell body in response to membrane hyperpolarization/depolarization cycles (Elgoyhen and Franchini, 2011). The organ of Corti, the sense organ of hearing in mammals, is a cellular

matrix that incorporates various supporting cells and two types of sensory receptor cell, the IHCs (inner hair cells) and the OHC (Figure 1.13).

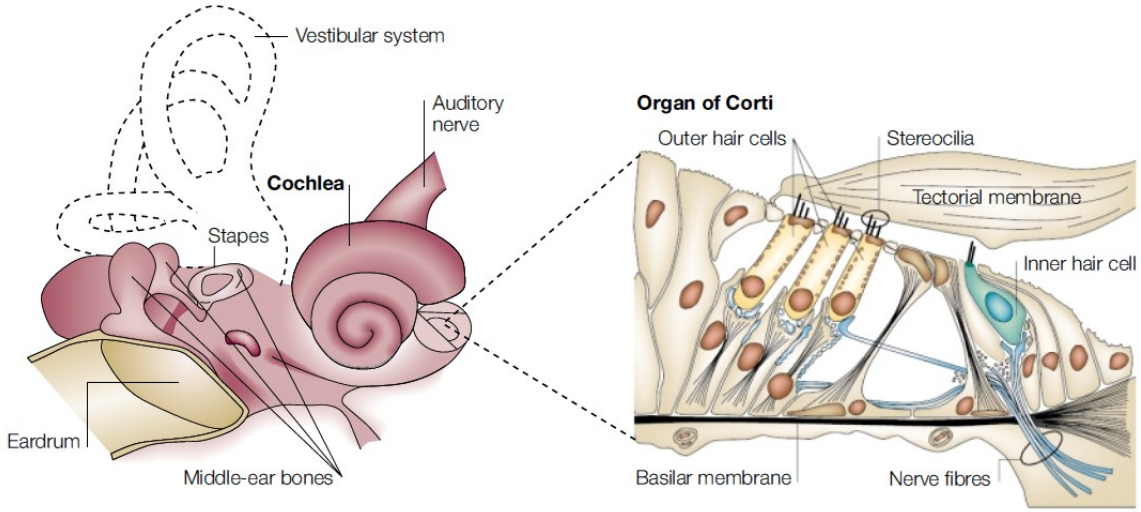


Figure 1.13 *A cross section of the mammalian inner ear and of the cochlea illustrating the organ of Corti. The organ of Corti is a cellular matrix that incorporates various supporting cells and two types of sensory receptor cell. The latter consist of a single row of IHCs and 3-4 rows of OHCs (Dallos and Fakler, 2002).*

IHCs function as the sensory receptors of the hearing organ and convey essentially all auditory information to the brain. OHCs are cylindrical cells 15-70 μm long and have a distinctive hair (stereocilia) bundle, which is the mechanosensory input organelle of these cells. When mechanically stimulated by incoming sound waves, the ciliary bundle is deflected, and thereby triggers the opening and closing of mechanosensitive ion channels in the stereocilia membrane. Then OHCs translate the resulting changes in membrane potential into macroscopic changes in the length of their cylindrical cell bodies. Depolarization triggers cell contraction, whereas hyperpolarization results in cell elongation. This electrically driven cell motility generates the mechanical energy that is required for amplifying the sound-induced vibrations in the cochlea.

Somatic electromotility in OHCs, as the basis for cochlear amplification, is a mammalian novelty and it is largely dependent upon the properties of the unique motor protein prestin. Prestin is thought to act as an area motor by alternating between two major conformations that occupy different cross-sectional areas within the membrane and were, therefore, termed “long” and “short states” (Dallos and Fakler, 2002). Joint conformational transitions of the tightly packed prestin motors in the plasma membrane of the OHC lead to length changes of the whole cell (Dallos and Fakler, 2002; Ashmore, 2008; Dallos, 2008). The distribution between long and short conformations is

immediately voltage dependent, which inevitably requires a mobile charged particle acting as a voltage sensor. The protein was named "prestin" because it was able to confer on cells the ability to move presto (fast in Italian) (Zheng et al., 2000). Mechanistically, the electromotility of OHCs is similar to piezoelectric materials that change dimensions under the influence of voltage (Ashmore, 2008).

Although we are still left without a clear picture of how these two mechanisms interact, stereocilia and somatic motility could provide alternative or complementary ways for cochlear amplification in mammalian hair cells (Elgoyhen and Franchini, 2011).

1.4.1.2. SLC26A5/ prestin's voltage sensor

In any voltage-operated membrane system, such as voltage-gated ion channels or skeletal muscle, a rapid change in membrane potential is accompanied by transient current flow through the membrane. This current results from electrical charges that are translocated across the membrane under the influence of the electrical field. The moving charge is called the "gating charge"; the resulting current is known as the "gating current". In ion channels, this gating current arises from the movement of an α-helix (S4 helix) that contains several positively charged amino acids and serves as a voltage sensor. In OHCs, the gating current was also thought to arise from the charged moiety of prestin. Movement of this charged moiety (voltage sensor) is manifested as a measurable capacitance known as NonLinear Capacitance (NLC), which is voltage-dependent.

Capacitance is defined as the derivative of charge with respect to voltage (see equation in figure 1.14.c). The total charge moved during the transient gating current is given as the integral of current over time (Figure 1.14.a). Ashmore (Ashmore, 2008) discovered that OHCs have NLC that shows a bell-shaped dependence on membrane potential, as in the bottom graph. The shape of the NLC curve reflects the probability of movable charges being translocated between two positions at opposite sides of the membrane. Therefore, capacitance peaks at the voltage that is most effective in producing a motile response.

As NLC measurements can be easily and accurately made experimentally and they are highly correlated to motility (Figure 1.14.b), they are often used to characterize electromotility instead of directly measuring changes in cell length. Therefore it is generally accepted that NLC can serve as a "signature" of the electromotility of OHCs. (Dallos and Fakler, 2002).

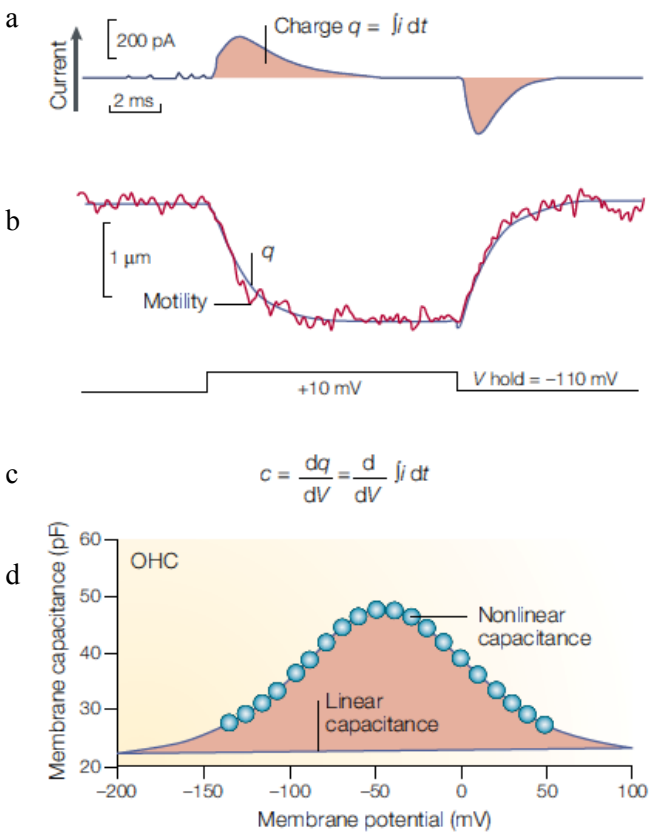


Figure 1.14 a) *When the current (i) is integrated (area under the curve, shown shaded), the total charge moved (q) is obtained.* ***b)*** *When the voltage- and time-dependence of the gating charge (in blue) are compared to that of electromotility in OHCs (in red), excellent correlation is found.* ***c)*** *Capacitance (c) is defined as the derivative of charge (q) with respect to voltage (V).* ***d)*** *OHCs have NLC that shows a bell-shaped dependence on membrane potential (Dallos and Fakler, 2002).*

The nature of the voltage sensor was determined by Oliver and colleagues (Oliver et al., 2001). To identify the sensor, all the charged residues that are not conserved between prestin and SLC26A6, the SulP/SLC26 protein with the closest homology to prestin, were neutralized either alone or in combination and the resulting molecules were probed for NLC. As NLC was not abolished in any of these cases, the concept that charged particles that are extrinsic to the protein act as the voltage sensor was tested. Neither cations nor external anions had any effect on NLC. However, NLC and, concomitantly, electromotility were reversibly eliminated by removing Cl^- ions from the cytoplasm of cells containing prestin while they were unaffected by removal of external Cl^-. Detailed analysis showed that the half-activating Cl^- concentration was 6 mM, matching the normal intracellular amount of this anion. Other monovalent anions were also found to be effective in promoting NLC, in the order $I^- \approx Br^- > NO_3^- > Cl^- > HCO_3^- > F^-$. Divalent anions (for example, SO_4^{2-}) were ineffective in functionally replacing their monovalent

counterparts. On the basis of these results, it was concluded that prestin uses an extrinsic voltage sensor (monovalent anions) available in the cytoplasm (Dallos and Fakler, 2002).

Later contrasting experimental observations showed that a residual NLC remained in the absence of internal Cl^-, which led to the proposal that the NLC was generated by an intrinsic voltage-sensor, and the dependence of the NLC on the concentration of the internal Cl^- was due to an allosteric action of Cl^- on prestin (Rybalchenko and Santos-Sacchi, 2003; Muallem and Ashmore, 2006).

Furthermore, Bai and co-workers identified an intrinsic voltage sensor in prestin: they mutated every charged residue in prestin that lies within or in very close proximity to the predicted transmembrane segments and identified a large number of charged residues within prestin's transmembrane domains that are also conserved in SLC26a6, and that contribute to the generation of NLC (Bai et al., 2009).

As yet, the nature of the NLC-generating transition is still not well known: NLC may derive from shuttling of Cl^- through the electric field (Oliver et al., 2001) or from movement of an intrinsic voltage sensor (Bai et al., 2009) or a combination of both.

1.4.1.3. SLC26A5/ prestin's molecular mode of action

There are various models for prestin's molecular mode of action, but the predominant view is that the conformational changes of the molecule depend on partial transmembrane movements of Cl^- ions. Unlike the other SulP/SLC26 transporters, in the case of mammalian prestin no significant unidirectional transport has been detected when monovalent (HCO_3^-, Cl^-) or divalent (SO_4^{2-}) anions were tested (Oliver et al., 2001; Schaechinger and Oliver, 2007). In mammalian prestin monovalent anions (HCO_3^-, Cl^-) have been proposed as forming the extrinsic voltage sensor in these proteins when they act as "incomplete" transporters, so that movement of these anions from the intracellular surface triggers changes in molecular conformation (Figure 1.15).

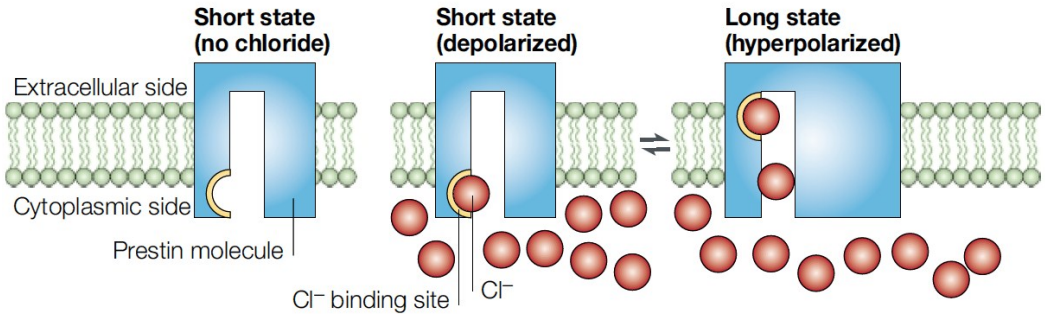


Figure 1.15 *Model of the control of prestin by internal Cl^- ions.* ***Left:*** *In the absence of internal Cl^-, the molecule is in its "short" state.* ***Middle:*** *The cell membrane is depolarized. Cl^- is bound to the molecule but remains at the cytoplasmic face of the membrane.* ***Right:*** *The cell membrane is hyperpolarized. Bound Cl^- is translocated across the molecule towards the external face (Dallos and Fakler, 2002).*

After binding to a site with millimolar affinity, these anions are translocated across the membrane in response to changes in the transmembrane voltage. They move towards the extracellular surface following hyperpolarization and towards the cytoplasmic side in response to depolarization. As a consequence, this translocation triggers conformational changes in the protein that ultimately alter its surface area in the plane of the plasma membrane. The area decreases when the anion is near the cytoplasmic face of the membrane ("short state"), and increases when the ion has crossed the membrane to the outer surface ("long state"). So, prestin acts as an incomplete transporter. It swings anions across the membrane, but does not allow these anions to dissociate and escape to the extracellular space. When monovalent anions are not present in the cytoplasm, all prestin molecules are in their "short" state, as the OHC is maximally contracted (Dallos and Fakler, 2002).

Subsequently, a theoretical work suggests that many experimental data could be better explained if one assumes that prestin acts as an electrogenic anion exchanger, exchanging one Cl^- ion for one divalent or two monovalent anions. The authors suggested that prestin functions as a Cl^-/ SO_4^{2-} antiporter (Figure 1.16) (Muallem and Ashmore 2006).

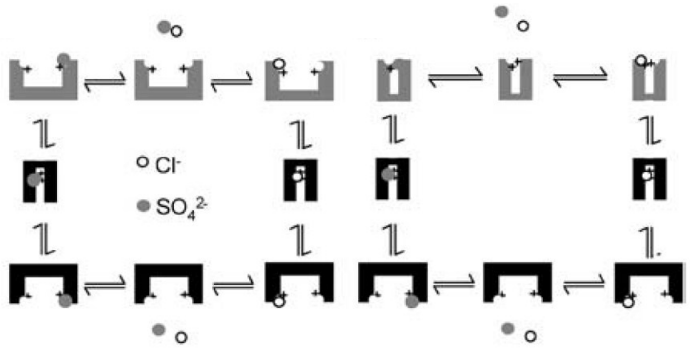


Figure 1.16 *Anion antiporter models associated with intrinsic charge movement (Muallem and Ashmore 2006).*

Some recent experiments, in contrast to previous observations, suggest that prestin is able to transport anions, although not necessarily those implicated in normal physiological processes, such as formate and oxalate (Bai et al., 2009), SCN^- (pseudohalide Thiocyanate) (Schanzler and Fahlke, 2012), HCO_3^- and Cl^- with a stoichiometry of 2 HCO_3^-:1 Cl^- (Mistrík et al., 2012). The physiological significance of this supposed transport is still not clear.

By creating recombinant prestins it has been shown that regions of the prestin amino acid sequence are implicated in transport and motility (Schaechinger et al., 2011; Tan et al., 2011); these regions were not located in the same region, but it is possible that there may be several areas in the prestin molecule that are important for electromechanical conversion.

Schaechinger and co-workers also proposed a model in which the NLC-generating transition is embedded within an alternating-access transport cycle (Figure 1.17).

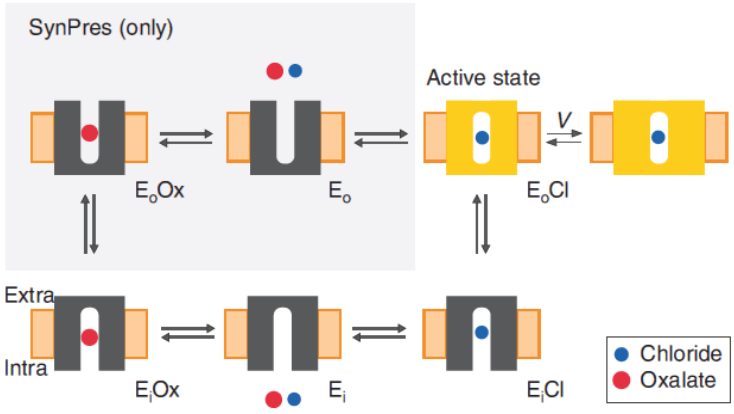


Figure 1.17 *Alternating-access model. Mammalian prestin may adopt only states in the non-shaded area. NLC arises from fast voltage-dependent transitions between the states highlighted in yellow, E_oCl (V, membrane potential) (Schaechinger et al., 2011).*

The transport process may determine the occupancy of a structural conformation ("active state") that enables the voltage-fuelled rapid elongation/contraction transition generating NLC and fast electromotility. This active state corresponds to a conformation in which a monovalent anion (Cl^-) occupies a binding site within prestin's permeation pathway, the Cl-bound state E_oCl. The model predicts the suppressive effect of intracellular oxalate on NLC generated by rat prestin: because NLC is completely independent of extracellular Cl^- (Oliver et al., 2001), dissociation from and binding of Cl^- to the outward-facing conformation must be impossible or greatly disfavoured (shaded in Figure 1.17). This is also consistent with the predominant view of the lack of transport activity in rat prestin (Schaechinger and Oliver, 2007; Tan et al., 2011). Hence, when binding of intracellular oxalate increases occupancy of E_iOx, occupancy of E_oCl and consequently NLC must decrease. Furthermore, loss of NLC upon removal of intracellular Cl^- is explained by accumulation of prestin in the E_i state that cannot promote NLC (Schaechinger et al., 2011).

As yet, short of a tertiary structure for prestin, the relationship between electromotility and transport remains unclear and the prestin's mode of action is still not well understood.

1.4.1.4. SLC26A5/ prestin's orthologs and evolution

If prestin-driven electromotility was a key step for the evolution of sound amplification (and tuning) in the mammalian ear, prestin should either be a mammalian novelty or mammalian prestin should have acquired new functional features to serve as the OHC motor. A phylogenetic and evolutionary analysis showed that prestin orthologs are present in all vertebrate species analyzed, indicating that prestin is not a mammalian exclusive protein. Comparing the percentage of protein identity (% ID), defined as the percentage of identical amino acids, among *Homo Sapiens* prestin with placental mammal and non-mammalian vertebrate prestin orthologs it is obtained that, whereas human prestin sequence shares a very high % ID (95% and higher) with mammal vertebrate, there is a significant drop in identity (< 59%) when it is compared to those of non-mammalian vertebrates. In addition, the analysis indicated that mammalian prestin % ID values are the highest among all members of the SulP/SLC26 family. Moreover, in the other SulP/SLC26 members the % ID values between mammals and non-mammalian vertebrate are much higher than for prestin. These results indicate that mammalian prestin

had a somewhat different evolutionary history when compared to other members of the family; it accumulated more amino acid changes in its primary sequence, most likely to accommodate a novel function (Elgoyhen and Franchini, 2011).

Recent evidence suggests that prestin orthologs from zebrafish and chicken are electrogenic antiporters, exchanging sulfate or oxalate for chloride in a strictly coupled manner with a 1:1 stoichiometry (Schaechinger and Oliver, 2007), with no motor function (Albert et al., 2007). Moreover several aspects of mammalian prestin's function show striking similarities when compared with non-mammalian prestin.

1. Voltage sensitivity and the resultant conformational rearrangements of mammalian prestin depend on the presence of millimolar concentration of internal Cl^-. Binding of the monovalent anion could be to the same binding sites involved in Cl^- transport by non-mammalian prestin (Oliver et al., 2001).
2. Oliver and co-workers (Oliver et al., 2001) showed that salicylate, the active component of aspirin, acts as a competitive antagonist at the anion-binding site of prestin. The anionic salicylate is able to induce NLC. The amplitude of this NLC, however, is more than an order of magnitude lower than that found for the small Cl^- ion. Nevertheless, the binding affinity of salicylate is around 300-fold higher than that found for Cl^-. This result provides a possible explanation for the significantly reduced OHC electromotility that probably underlies the hearing loss induced by large doses of aspirin (Dallos and Fakler, 2002). Salicylate block of mammalian prestin quantitatively equals the inhibition of non-mammalian transport.
3. Prestin from *D. rerio* generates voltage-dependent charge movements (Albert et al., 2007) resembling the gating charge associated with electromotility. The presence of a small NLC in zebrafish and chicken prestins may suggest a common mechanism of charge movement seen in the mammalian prestins. However, lack of motor function in zebrafish and chicken prestins indicates that the charge movements and motor function may evolve independently, although the two are believed to be fully coupled in mammalian prestin (Tan et al., 2011).
4. Mammalian and chicken isoforms share a substantial degree of sequence conservation, especially in the hydrophobic core region that presumably provides the structural basis for ion transport.
5. Although zebrafish prestin is expressed in hair cells, it is distributed throughout the hair cells (Albert et al., 2007). This subcellular staining pattern contrasts with

the exclusive basolateral membrane localization of prestin in mammalian OHCs. The lack of membrane localization of zebrafish prestin most likely precludes its function as a molecular motor (Franchini and Elgoyhen, 2006; Elgoyhen and Franchini, 2011).

These studies suggest that prestin's unique function evolved as a modification of an anion exchange mechanism.

1.4.1.5. SLC26A5/ prestin structure and quaternary organization

The topology for SLC26A5/prestin is not known, but it is agreed that an even number of helices span the membrane as both NH_2 terminal and COOH terminal ends lie within the cytoplasm (Zheng et al., 2000).

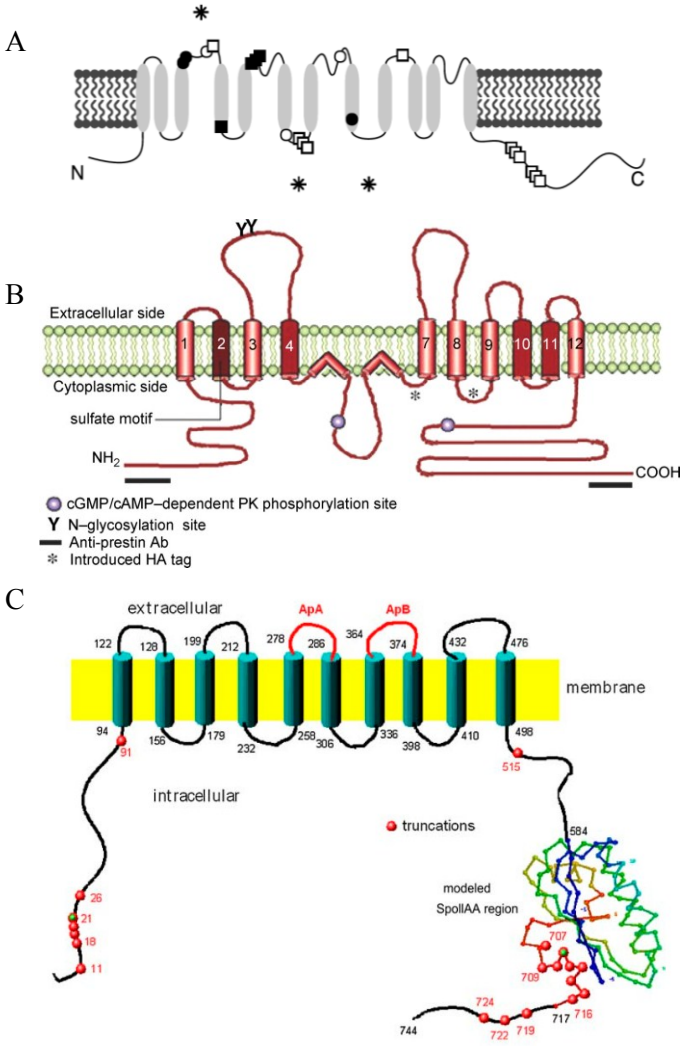


Figure 1.18 *Models of prestin's transmembrane topology.* ***A:*** *12 transmembrane α-helices (Zheng et al., 2000; Oliver et al., 2001);* ***B:*** *10 α-helices inserting across the membrane with 2 nonspanning helices present in the set of helices (Deák et al., 2005).* ***C:*** *10 transmembrane α-helices (Navaratnam et al., 2005).*

Subsequent reports have been more definite and identified 12 transmembrane α-helices (Figure 1.18.A) (Oliver et al., 2001), 10 α-helices inserting across the membrane with 2 nonspanning helices present in the set of helices (Figure 1.18.B) (Deák et al., 2005) or 10 transmembrane helices alone (Figure 1.18.C) (Navaratnam et al., 2005).

Morphological studies show that OHC membrane contains large protein particles with a diameter between 9 and 13 nm (Le Grimellec et al., 2002; Mio et al., 2008; Murakoshi et al., 2008; Kumano et al., 2009; Sinha et al., 2010). Correlation between the development of motility and particle density in developing OHCs suggests that the majority of the intramembrane particles (IMPs) are the motor protein prestin (He et al., 2010). Prestin has 744 amino acids with a predicted molecular mass of 81.4 kDa. The predicted amino acid sequence and molecular mass of prestin monomer appear to be inadequate to account for the size of IMPs in the membrane. In fact the particles are 4-5 times the expected size of prestin, but some biochemical studies suggest that prestin forms oligomers. FRET analyses of prestin coupled to fluorescent proteins suggested the existence of homo-oligomeric interaction, but did not provide any information on stoichiometry (Greeson et al., 2006; Navaratnam et al., 2005). A Western blot analysis of prestin purified from expression systems suggested dimer states and tetramer states formed from dimers connected via hydrophobic bonds (Zheng et al., 2006).

A low-resolution electron density map of purified prestin exhibited a bullet-shaped fourfold symmetric molecule of 77 x 77 x 115 Å in size with an inner cavity (Figure 1.19), suggesting a tetrameric arrangement of subunits (Mio et al., 2008).

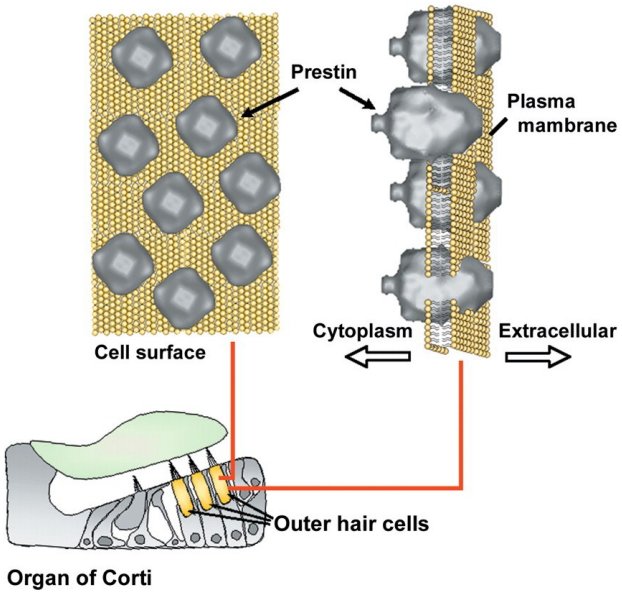


Figure 1.19 *Images of prestin particles embedded in the plasma membrane of the OHCs, viewed from outside the cell (left) and across the membrane (right) (Mio et al., 2008).*

On the other hand, Le Grimellec and colleagues (Le Grimellec et al., 2002) found structures with a central depression in the cytoplasmic side of the OHC plasma membrane by AFM. Murakoshi and co-workers (Murakoshi et al., 2008) detected prestin in the plasma membrane of prestin-transfected CHO cells using Qdots as topographic markers and observed ring-like structures, possibly prestin, by AFM. In agreement with these results also Kumano and colleagues (Kumano et al., 2010) demonstrated that prestin might form a ring-like structure with a diameter of about 11 nm (Figure 1.20). Thus, the structure of prestin is still a controversial issue.

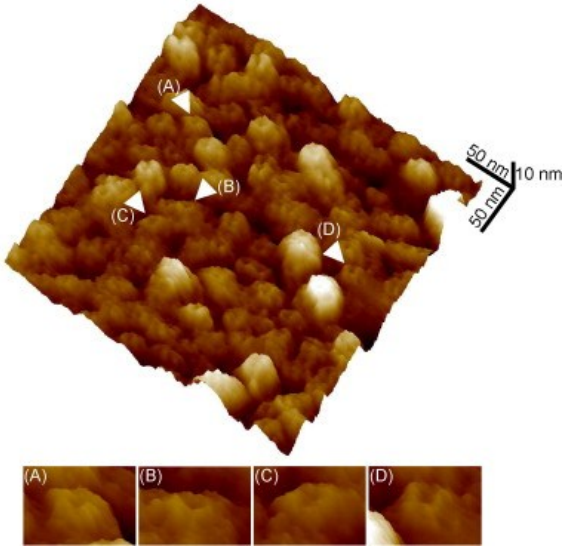


Figure 1.20 *3D AFM height image of the prestin-reconstituted lipid bilayer. Representative examples of ring-like structures were digitally magnified and are shown (A and B) (Kumano et al., 2010).*

More recently other two works demonstrated that in cell membranes prestin oligomerizes to a tetramer (Wang et al., 2010; Hallworth and Nichols 2011). However, another study suggested that dimers were the normal configuration of prestin and some other SulP/Slc26 family members (Detro-Dassen et al., 2008). Therefore, it is still not clear whether prestin forms tetramers, trimers, or dimers, but it is possible that all these high-order oligomers can co-exist and are functional (He et al., 2010).

1.4.1.6. SLC26A5 binding proteins

It has been reported that prestin binds to and is potentiated by CFTR coexpression in HEK-293 cells, but the functional role of this interaction isn't still clear (Homma et al., 2010).

It has been discovered, through yeast two-hybrid interaction, that MAP1S bound to prestin. MAP1S interacts through a linker region between light and heavy chain domains with the proximal region of the STAS domain. The interaction leads to marked stimulation of prestin-mediated NLC in transiently transfected CHO cells in parallel with increased surface expression (Bai et al., 2010).

The integral membrane protein VAPA (or VAP-33) was discovered as a prestin binding protein through a membrane-based yeast two-hybrid screen, and validated by co-immunoprecipitation and immuno-colocalization in outer hair cells. OHCs from Slc26a5-/- mice expressed VAPA at decreased abundance, suggesting a possible role for VAPA in vesicular trafficking of prestin (Sengupta et al., 2010).

Additional, incompletely validated prestin-binding proteins identified in this membrane-based yeast two-hybrid screen include integrin-interacting proteins tetraspanins 6 and 29 (CD9) and β2-microglobulin, glycosylphosphoinositide-linked costimulatory molecule CD52/CAMPATH-1, the small conductance Ca^{2+}-activated K^+ channel KCNN2/SK2, fatty acid binding protein FABp3, Wnt receptor frizzled-3, bone γ-carboxyglutamate protein I, and dynein light chain Tctex-type I (Zheng et al., 2009).

Prestin has also been shown to interact indirectly with βV-Spectrin, likely through the cytoskeletal proteins F-actin and protein 4.1. However, these intermediate proteins have themselves not yet been shown to bind directly to prestin (Legendre et al., 2008).

1.4.1.7. SLC26A5 and genetic diseases

Autosomal recessive deafness was initially attributed to biallelic mutations in the SLC26A5 gene (Liu et al., 2003) encoding prestin. However, subsequent studies established that the very few known prestin coding sequence variants found to date are variants also present in individuals with normal hearing (Minor et al., 2009), despite the severe hearing impairment observed in Slc26a5-/- mice.

1.4.2. SLC26A4/ pendrin

Pendrin is the fourth member of the SulP/SLC26 anion transporter family, which is involved in the transport of anions, including Cl^-, HCO_3^-, OH^-, I^-, or formate, within the ear, thyroid, and kidney (Mount and Romero, 2004). In thyrocytes, I^- and Na^+ are brought into the cells via the basolaterally located Na^+-I^- symporter. Apically located pendrin

seems to be responsible for the efflux of iodide into the follicular lumen (Yoshida et al., 2002). In the kidney, pendrin is suspected to mediate Cl^-/HCO_3^- exchange in the acid-base regulating β- and non-α-non-β-intercalated cells (Soleimani et al., 2001). Similarly, in the inner ear, pendrin is thought to mediate Cl^-/HCO_3^- exchange, and is therefore involved in the conditioning of endolymphatic fluid, presumably due to HCO_3^- secretion (Wangemann et al., 2007). Malfunction of pendrin leads to Pendred syndrome (PS). PS (OMIM#274600) is an autosomal recessive disorder accounting for 4-10% of inherited hearing losses and involves two organ systems: the ear, and the thyroid gland. As regards the ear, impaired pendrin function i) promotes a progressive increase in endolymph volume followed by an enlargement of the membranous labyrinth and surrounding osseous structures, and ii) leads to degeneration of inner ear sensory cells (Everett et al., 2001). The resulting phenotype is a severe/profound sensorineutral hearing loss (SNHL). While impaired pendrin function at the thyroid level can result in goiter, defects in iodide organification, and hypothyroidism.

Structural information regarding membrane topology of pendrin is limited and controversial. There are different topology models in literature, among which 11 TM, 12 TM and 13 TM models, but in a recent paper Dossena and co-workers proposed a new membrane topology model in which SLC26A4 would be formed by up to 15 TM helical segments (Figure 1.21) (Dossena et al., 2009).

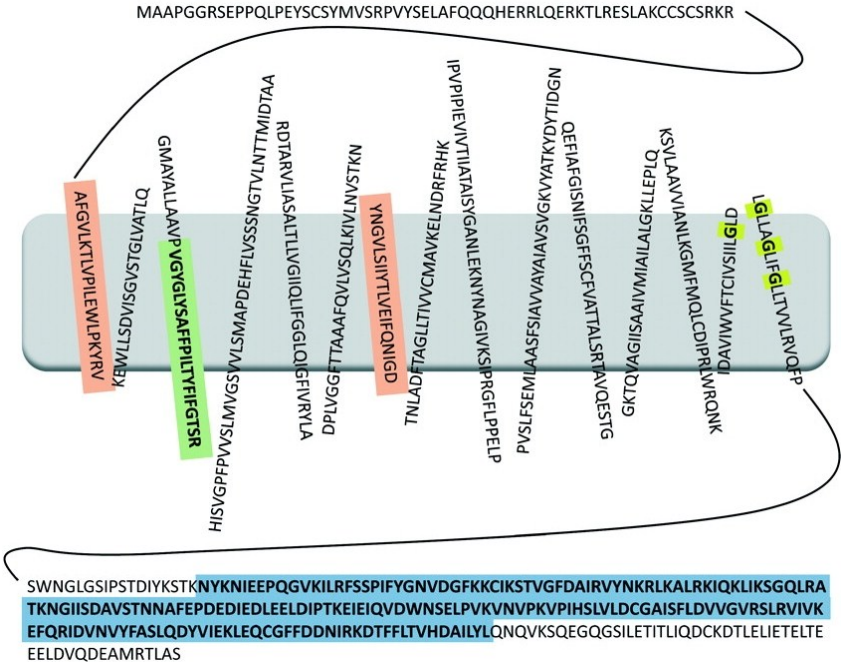


Figure 1.21 *Putative 15 TM-segment model topology of human pendrin proposed by Dossena and co-workers. The amino terminus in this model would be located on the extracellular side, and the carboxy terminus would be located at the cytoplasmic side (Dossena et al., 2009).*

1.4.3. Sultr1;2

Most environmental sulfur exists in the oxidized form, primarily SO_4^{2-}, which is reduced and assimilated by plants. The sulfate ion (SO_4^{2-}) is transported into plant root cells by SO_4^{2-} transporters and then mostly reduced to sulfide (S^{2-}). The S^{2-} is then bonded to O-acetylserine (OAS) through the activity of cysteine synthase or OASTL to form cysteine, the first organic molecule of the SO_4^{2-} assimilation pathway.

In the plant *Arabidopsis thaliana*, at least 14 distinct SulP/Sultr genes in 5 classes encode plasmalemmal and intracellular membrane sulfate transporters. These transporters are expressed in distinct subcellular membranes in specific tissues and with developmental stage-specific patterns. Although most are SO_4^{2-} transporters, some preferentially exhibit nM affinity uptake of molybdate and at least one has turned out to regulate cellular levels of phytic acid.

The two best studied SulP/SLC26 proteins in plant are Sultr1;2 and Sultr1;1, two high affinity, root-localized SO_4^{2-} transporters, that take up SO_4^{2-} from the rhizosphere. Sultr1;2 transcript levels are high in sulfur-rich conditions and increase slightly upon sulfur deprivation, whereas the Sultr1;1 transcript levels increase by >100-fold during sulfur deprivation. In a recent paper Shibagaki and Grossman demonstrated that the STAS domain of Sultr1;2 physically interacts with OASTL, decreasing SO_4^{2-} transport activity of Sultr1;2 and increasing OASTL activity. Based on these results they proposed a model for the regulation of SO_4^{2-} assimilation by roots (Figure 1.22).

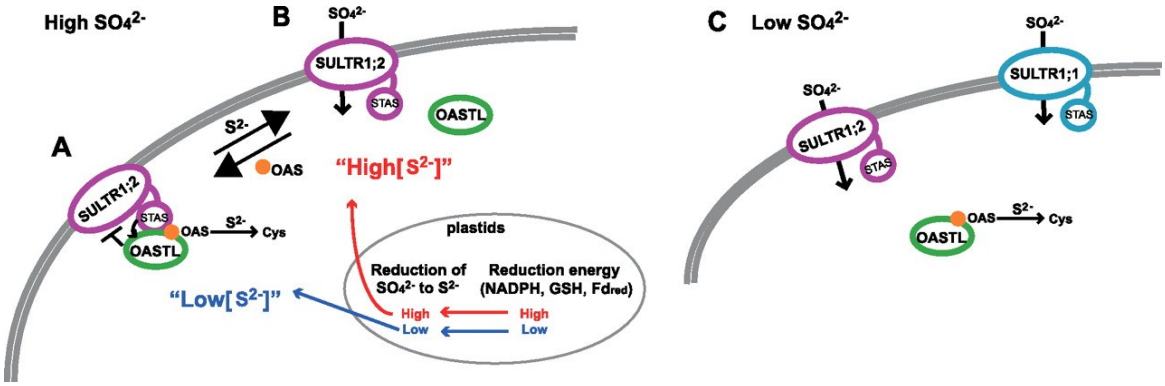


Figure 1.22 *A potential novel regulatory mechanism controlling SO_4^{2-} transport into root cells that involves OAS/S^{2-}-modulated interaction of OASTL with SULTR1;2 (Shibagaki and Grossman, 2010).*

At high SO_4^{2-} concentrations (Figure 1.22 A and B), when Sultr1;2 is the dominant transporter responsible for SO_4^{2-} uptake from the soil, direct interaction of Sultr1;2 and OASTL helps coordinate SO_4^{2-} transport with S^{2-} production and cysteine synthesis. Availability of S^{2-} in the cytoplasm of the root cells would be a major control feature of

this model. OASTL would associate with the STAS domain of Sultr1;2 when OAS rises, and this interaction inhibits transporter activity. As OAS is consumed in the synthesis of cysteine, under conditions where S^{2-} production begins to keep pace with OAS generation, OASTL would dissociate from Sultr1;2 releasing the transporter from inhibition. When SO_4^{2-} is low (Figure 1.22 C), a conformational change in Sultr1;2 may lessen the STAS-OASTL interaction, allowing higher rates of SO_4^{2-} transport activity despite potentially high OAS levels (Shibagaki and Grossman, 2010).

1.4.4. Rv1739c

Growth and virulence of mycobacteria requires sulfur uptake. Although genetic growth experiments in vitro suggest that mycobacterial sulfate assimilation is mediated entirely by the ABC-type sulfate permease CysTWA, with possible contributions of transport systems for thiosulfate and sulfur-containing amino acids, the sulfate assimilation systems employed by mycobacteria sequestered inside lysosomes are less well understood. Among candidate sulfate or anion transporters in the *M. tuberculosis* genome are three SulP/SLC26 genes of unknown function, among which Rv1739c.

Zolotarev and co-workers demonstrated that induction of Rv1739c expression in *E. coli* increased bacterial uptake of sulfate. The increased sulfate uptake occurs by a mechanism requiring the cytoplasmic CysA subunit of the ABC sulfate permease. The transmembrane domain of Rv1739c suffices for this activity, therefore the function of the Rv1739c STAS domain remains unknown (Zolotarev et al., 2008).

1.4.5. BicA

Marine cyanobacteria generate as much as 25% of global primary productivity through efficient inorganic carbon fixation mechanisms. Cyanobacteria possess an environmental adaptation known as a CO_2 concentrating mechanism (CCM) that evolved to improve photosynthetic performance, particularly under CO_2-limiting conditions. The CCM functions to actively transport dissolved inorganic carbon species (Ci; HCO_3^- and CO_2) resulting in accumulation of a pool of HCO_3^- within the cell that is then utilized to provide an elevated CO_2 concentration around the primary CO_2 fixing enzyme, ribulose bisphosphate carboxylaseoxygenase (Rubisco). Rubisco is encapsulated in unique micro-compartments known as carboxysomes and also provides the location for elevated CO_2

levels in the cell. Five distinct transport systems for active Ci uptake are known, including two types of Na^+-dependent HCO_3^- transporters (BicA and SbtA), one traffic ATPase (BCT1) for HCO_3^- uptake and two CO_2 uptake systems based on modified NADPH dehydrogenase complexes (NDH-I3 and NDH-I4) (Price 2010).

BicA is a low- to medium affinity, high-flux, Na^+-dependent HCO_3^- transporter (probably a Na^+/ HCO_3^- symporter) belonging to the SulP/SLC26 family (Price et al., 2004). Recently the transmembrane topography of BicA (from *Synechococcus PCC 7002*) has been studied by progressive C-terminal truncation and fusion with alkaline phosphatase and β-lactamase, identifying 12 transmembrane helices and cytoplasmic N- and C-termini (figure 1.23) and yielding an experimentally derived topographical model for SulP/SLC26 anion transporter transmembrane domains (Shelden et al., 2010, Price and Howitt, 2011).

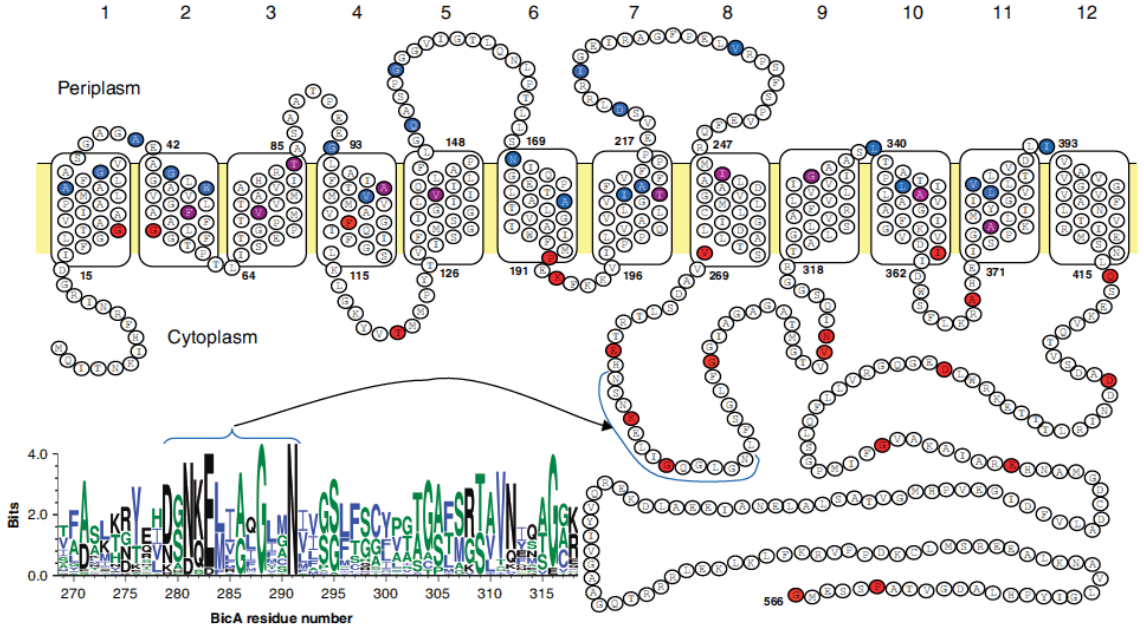


Figure 1.23 *The topology map for the Synechococcus PCC 7002 Na^+-dependent HCO_3^- transporter, BicA. The inset shows a WebLogo (weblogo.berkeley.edu) representation of the second signature sequence identified by Saier et al., for the SulP/SLC26 family (Saier et al., 2000); the most highly conserved subset, NSNKELIGQGLGN (279-291), is highlighted (Shelden et al., 2010).*

Concerning the cyanobacterial BicA proteins analyzed in this thesis, BicA from *Synechocystis sp. strain PCC 6803* is a Na^+-dependent HCO_3^- transporter (Xu et al., 2008); while the functions of BicA from *Trichodesmium erythraeum IMS101* and from *Thermosynechococcus elongatus BP-1*, are still unknown, but maybe they act as sulfate transporters.

2. The project

2.1. Aim of the project

Most of the knowledge related to the SulP/SLC26 members has come from studies within the last decade and the interest in these transporters has expanded, as shown by the increasing number of publications within recent years. Despite the increasing interest in the SulP/SLC26 members, a substantial amount of research is needed to understand the roles of these transporters in human physiology. In particular, very little is known about the structural organization of the SulP/SLC26 transporters and of their STAS domains and no 3D structure of the full-length sequences is available. The structural characterization is fundamental for the comprehension of the mode of action of a protein and it is an essential step for the understanding of the functional consequences of the mutations responsible for related pathologies. In this context, the long-term task is to derive a model for the structural organization of these transporters, which will provide insights into the functional consequences of mutations linked to genetic diseases. Another task is to analyze the prestin STAS domain whose 3D structure has been solved during my Master Degree Thesis from a functional point of view, in order to understand its possible role in the transport.

To this purpose the first part of the PhD project focused on the production and the structural characterization of STAS domains from different species and of mutants of the prestin STAS domain whose 3D structure have been solved. The second part of the project was more ambitious and concerned the production of the full-length SulP/SLC26 transporters by cell-free expression system.

2.2. The strategy

The strategy followed during the PhD can be outlined in this way:

1) Selection of full-length SulP/SLC26 transporters and STAS domains from different distance-related species.
2) Production and characterization of different orthologs and mutants of the STAS domain:

a) Cloning of the coding regions for the STAS domains into the pET SUMO expression vector.
b) Protein expression in bacterial systems (*E. coli).*
c) Protein purification by chromatographic techniques as affinity chromatography (IMAC, immobilized-metal affinity chromatography) and size exclusion chromatography (SEC).
d) Structural characterization by classical biophysical methods: circular dichroism spectroscopy (CD), dynamic light scattering (DLS), differential scanning fluorimetry (DSF).
e) Crystallization screening with the most promising constructs.
f) 3D structure determination through X-ray crystallography.
g) Structure-function analysis.

3) Production and characterization of full-length SulP/SLC26 anion transporters:
   a) Cloning of the coding regions for the full-length SulP/SLC26 anion transporters into the pET21cHx expression vector.
   b) Cell-free (CF) expression in the continuous-exchange (CECF) configuration in the precipitate-forming (P-CF) mode.
   c) Protein purification by chromatographic techniques as IMAC and SEC.
   d) Detergent screening, stability and size homogeneity analysis (SEC, CD, DSF).
   e) Crystallization screening with the most promising constructs.

The structural information is then integrated with functional data that come from literature and international collaborations to understand the role of the STAS domains and of the full-length SulP/SCL26 transporters and the relationship between their mutations and the associated diseases.

1) Selection of full-length SulP/SLC26 transporters and STAS domains from different distance-related species

The following full-length SulP/SLC26 transporters from different distance-related species and their STAS domains were selected for the structural and biophysical characterization:

- SLC26A5/ prestin from mammal (*Rattus norvegicus*) and non-mammals (*Gallus gallus* and *Danio rerio*) to have information about the protein evolution and their different functions in orthologous proteins.

- SLC26A4/ pendrin from *Homo sapiens*, interesting because mutations in this protein and in its STAS domain can cause genetic human diseases.
- Sultr1;2 from *Arabidopsis thaliana* is an eukaryotic protein, but it is smaller than mammalian members and could constitute a reliable model system for the mammalian ones; furthermore it is well functionally characterized.
- Bacterial members of the SulP/SLC26 family (Rv1739c from *Mycobacterium tuberculosis*, BicA from *Synechocystis sp. PCC 6803*, *Trichodesmium erythraeum* and *Thermosynechococcus elongatus BP-1*) that often provide good model systems for the study of structure and function because they could be more accessible to experimental techniques that are either not available or more difficult for eukaryotic proteins.

This kind of approach has two main reasons: first, differences in the structures can shed light on the evolution of these proteins and on the different roles played by the various transporters; second, this approach can be instrumental to the final aims since it is well known that even similar polypeptides can have substantially different properties in terms of propensity to crystallize.

2) Production and characterization of different orthologs and mutants of the STAS domain

The accurate selection of the N- and C-termini is more critical for domains that are part of a larger protein, as it is the case of the SulP/SLC26 STAS domains, whose boundaries are not clearly defined by sequence alignments. To address this issue, constructs' ends were designed according to multiple sequence analyses between the construct whose 3D structure was solved, that corresponds to a compact single domain, and the C-terminal portion of selected SulP/SLC26 transporters from different species (figure 2.1). The main reason is the sequence similarity among the STAS domains of these transporters; therefore it is very likely that the ends of these domains correspond. The percent amino acid identity between the mammalian prestin STAS domain (*R. norvegicus*) and that of non-mammalian orthologs is rather high (*G. gallus* =60%, *D. rerio* =55%), while is lower in comparison with homologous transporters (SLC26A4 from *H. sapiens* =36%; Sultr1;2 from *A. thaliana* =26%). The main difference between STAS domains from animals and that from lower organisms has been found in a long insertion, known as variable loop. No such insertion is present in bacterial SulP/SLC26 homologues, and this loop is also much shorter in *A. thaliana* paralogs. The evolutionary

and functional role of this loop is still unknown; secondary structure predictions suggest that it is mostly unstructured in solution. To investigate the role of the variable loop, we designed the STAS domains from animals with and without variable loop (in the square, Figure 2.1), while the STAS domain from plant was designed with loop because it is very small, ~10 residues.

```
A5 R.norvegicus 505 -SPSYTVLGQLPDTDVYIDIDAYEEVKEIPGIKIFQINAPIYYANSDLYSSALKRKTGVN 563
A5 G.gallus     511 SRPQYRILGQIPDTDIYCDVEEYEEVKEYPGIKIFQANTSLYFANSESYTSALKKKTGVD 569
A5 D.rerio      508 SRPKNVVLGQIPDTGLYFDVDEYEEAEECSGIKIFQSNSSIYFANSELYVKALKAKTGID 566
A4 H.sapiens    515 SFPSWNGLGSIPSTDIYKSTKNYKNIEEPQGVKILRFSSPIFYGNVDGFKKCIKSTVGFD 573
Sultr1;2        502 SRPRTAVLGNIPRTSVYRNIQQYPEATMVPGVLTIRVDSAIYFSNSNYVRERIQRWLHEE 560
                      *    **.:* *.:* . . * :     *:  :: .:.:::.* :   . ::     :

A5 R.norvegicus 563 PAIIMGARRKAMRKYAKEVGNAN---------IANATVVKVDAEVDGENATKPEEEDDEV 616
A5 G.gallus     569 PCAILAARRKAQKKHAREIKKANKVKKKAVLKLVNSSTNDVEASVKHEIANDGLPANGKF 629
A5 D.rerio      566 PEKLLDAKKLQLKYAKRDTEGTKTVNQGSLLKKNAVVLLDMELGVTHEVLNGPQKPKHVH 626
A4 H.sapiens    573 AIRVYNKRLKALRKIQKLIKSGQ---------LRATKNGIISDAVSTNNAFEPDEDIEDL 624
Sultr1;2        560 EEKVKAASLP---------------------------------------------------- 570
                       :

A5 R.norvegicus 616 KFPPIVIKTTFPE-ELQRFLPQG-----ENIHTVILDFTQVNFMDSVGVKTLAGIVKEYG 671
A5 G.gallus     629 AFVDAGVQDGSPD-ELEHFVEPK-----TNVHSLILDFAPVNFVDSVGAKTLKSVIKEYN 683
A5 D.rerio      626 TNGQMTEKHIESESEDEFFLQRL-----TPIHSVILDFTPVNFIDSVGAKTIKSVIKEYA 681
A4 H.sapiens    624 EELDIPTKEIEIQVDWNSELPVKVNVPKVPIHSLVLDCGAISFLDVVGVRSLRVIVKEFQ 684
Sultr1;2        570 ----------------------------RIQFLIIEMSPVTDIDTSGIHALEDLYKSLQ 601
                                                :: ::::   :. :*  * :::  : *.

A5 R.norvegicus 671 DVGIYVYLAGCSAQVVNDLTSNRFFENPALKELLFHSIHDAVLGSQ-VREA-- 718
A5 G.gallus     683 EVGVCVCIASCSGPVMNELTRLNFFDNTVTRELLFHSIHDAVLACQ-GKDR-- 733
A5 D.rerio      681 TVDVKVVLAGCSRTLLSELRTLQFFSEPVTPDLIFPTIHDAVLQCKRWRDLP- 733
A4 H.sapiens    684 RIDVNVYFASLQDYVIEKLEQCGFFDDNIRKDTFFLTVHDAILYLQNQVKSQ- 736
Sultr1;2        601 KRDIQLILANPGPLVIGKLHLS-HFADMLGQDNIYLTVADAVEACCPKLSNEV 653
                      .: : :*.    :: .*    .* :    : :: :: **:       .
```

Figure 2.1 *Multiple sequence alignment of selected STAS domains using ClustalW tool, EMBL-EBI. Variable loop is highlighted in the squares.*

In recent structures of the prestin STAS domain, whose 3D structure has been solved, in our laboratory (unpublished data), in presence of heavy atoms as I^- and Br^- an anion-binding site was identified. In order to study the anion-binding site and to investigate the possible role of the STAS domain in the transport, we designed some STAS domains' mutants of single amino acids implicated in the anion-binding site (Figure 2.2 and 2.3).

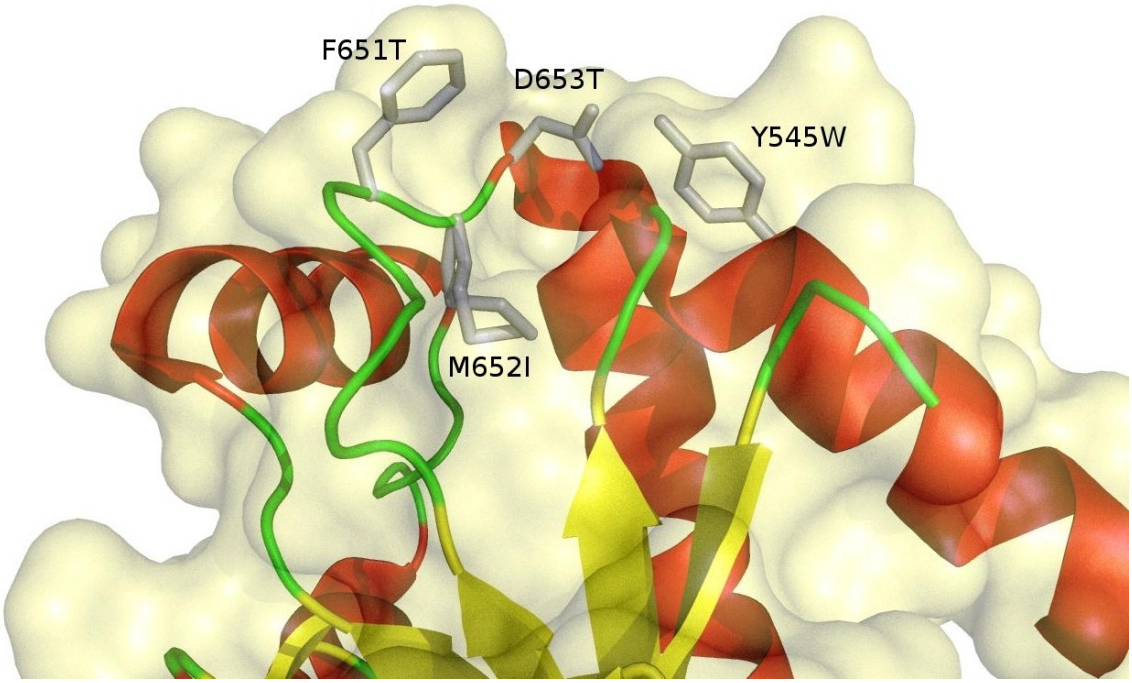


Figure 2.2 *Detail of the prestin STAS domain anion-binding site and of the residues mutated.*

```
Wild-type  505 SPSYTVLGQLPDTDVYIDIDAYEEVKEIPGIKIFQINAPIYYANSDLYSSALKRKTGVN 563 GS
F651T      505 SPSYTVLGQLPDTDVYIDIDAYEEVKEIPGIKIFQINAPIYYANSDLYSSALKRKTGVN 563 GS
M652I      505 SPSYTVLGQLPDTDVYIDIDAYEEVKEIPGIKIFQINAPIYYANSDLYSSALKRKTGVN 563 GS
D653T      505 SPSYTVLGQLPDTDVYIDIDAYEEVKEIPGIKIFQINAPIYYANSDLYSSALKRKTGVN 563 GS
Y545W      505 SPSYTVLGQLPDTDVYIDIDAYEEVKEIPGIKIFQINAPIWYANSDLYSSALKRKTGVN 563 GS

Wild-type  637 ENIHTVILDFTQVNFMDSVGVKTLAGIVKEYGDVGIYVYLAGCSAQVVNDLTSNRFFEN 695
F651T      637 ENIHTVILDFTQVNTMDSVGVKTLAGIVKEYGDVGIYVYLAGCSAQVVNDLTSNRFFEN 695
M652I      637 ENIHTVILDFTQVNFIDSVGVKTLAGIVKEYGDVGIYVYLAGCSAQVVNDLTSNRFFEN 695
D653T      637 ENIHTVILDFTQVNFMTSVGVKTLAGIVKEYGDVGIYVYLAGCSAQVVNDLTSNRFFEN 695
Y545W      637 ENIHTVILDFTQVNFMDSVGVKTLAGIVKEYGDVGIYVYLAGCSAQVVNDLTSNRFFEN 695

Wild-type  696 PALKELLFHSIHDAVLGSQVREA 718
F651T      696 PALKELLFHSIHDAVLGSQVREA 718
M652I      696 PALKELLFHSIHDAVLGSQVREA 718
D653T      696 PALKELLFHSIHDAVLGSQVREA 718
Y545W      696 PALKELLFHSIHDAVLGSQVREA 718
```

Figure 2.3 *Multiple sequence alignment among mutant and wild-type prestin STAS domains (the single amino acid mutated are highlighted in bold type and undelined).*

The derived structural information will help in a better understanding of the role played by the STAS domain in the activity of the transporters and of the mode of action of the full-length prestin.

3) Production and characterization of full-length SulP/SLC26 anion transporters

The membrane proteins are one of the most difficult classes of protein to produce because of their hydrophobic nature, toxic effects and specific requirements for targeting and translocation systems. In particular their overexpression in *E. coli* can result into the accumulation of insoluble aggregated material, and indeed this is what we found for SulP/SLC26 proteins we preliminary tested.

CF expression system has recently emerged as a promising and highly versatile technique for the general production of membrane proteins. CF expression reduces the high complexity of protein production known from living organisms to the basic translation process. Most toxic effects of synthesized proteins to the expression host are virtually eliminated and the speediness of CF expression is highly competitive. Those features render CF approaches often more reliable and reproducible if compared with many cell-based expression system.

As for the STAS domain, some full-length SulP/SLC26 from different species were selected for CF expression; no truncated constructs were designed. The plan was to set-up an appropriate procedure for the production of selected SulP/SLC26 members by CF technique in sufficient amount for a further characterization. After CF production, samples were been evaluated for their quality, both in their biophysical and biochemical characteristics and in their activity and functional properties. When produced in sufficient amounts and quality, the transporters were been submitted to crystallization trials.

3. The STAS domain

3.1. Methods

3.1.1. Design of STAS domains variants

Eight STAS domain variants from different distance-related species were designed according to the sequence of the construct whose 3D structure was solved (PDB code = 3LLO, first row in the Table 3.1), that corresponds to a compact single domain. STAS domains of SulP/SLC26 transporters from animals were designed both with and without variable loop, while the STAS domain from plant was designed with loop because it is very small, ~10 residues.

Table 3.1 Selection of SulP/SLC26 anion transporters STAS domains

Protein	**Source organism**	**Construct**	**Loop**	**Coding sequence (bp)**	**Aa (residues)**	**MW (kDa)**
SLC26A5/ prestin	*Rattus norvegicus*	[505-563] GS[637-718]	without	429	143	15.7
		[505-718]	with	642	214	23.7
SLC26A5/ prestin	*Gallus gallus*	S[511-569] GS[652-733]	without	432	144	16.0
		S[511-733]	with	672	224	24.8
SLC26A5/ prestin	*Danio rerio*	S[508-566] GS[650-733]	without	438	146	16.2
		S[508-733]	with	681	227	25.5
SLC26A4/ pendrin	*Homo sapiens*	S[515-573] GS[653-736]	without	438	146	16.5
		S[515-736]	with	669	223	25.3
Sultr1;2	*Arabidopsis thaliana*	S[502-653]	with	816	272	30.6

STAS domain mutagenesis of 3LLO construct was performed in order to study the anion-binding site and the possible role of the STAS domain in the transport. In recent structures of wt (wild type) STAS domain in presence of heavy atoms, such as I^- and Br^-, a binding site has been identified (Figure 3.1). These anions were chosen, even if they are not implicated in normal physiological processes, because they have a higher affinity to prestin in comparison with Cl^- and HCO_3^- ($I^- \approx Br^- > NO_3^- > Cl^- > HCO_3^- > F^-$) (Oliver et al., 2001) and, furthermore, they can be easily identified by diffraction measurements at the appropriate wavelength using the anomalous diffraction effect. This kind of strategy is widely used to identify the binding of physiological anions to protein sites, for instance in

the cases of the glutamate-gated chloride channel α GluCl (Hibbs and Gouaux, 2011), and the *Escherichia coli* Cl^-/H^+ exchange-transporter CLC-ec1 (also denoted EcCLC) (Accardi et al., 2006; Lobet and Dutzler, 2006).

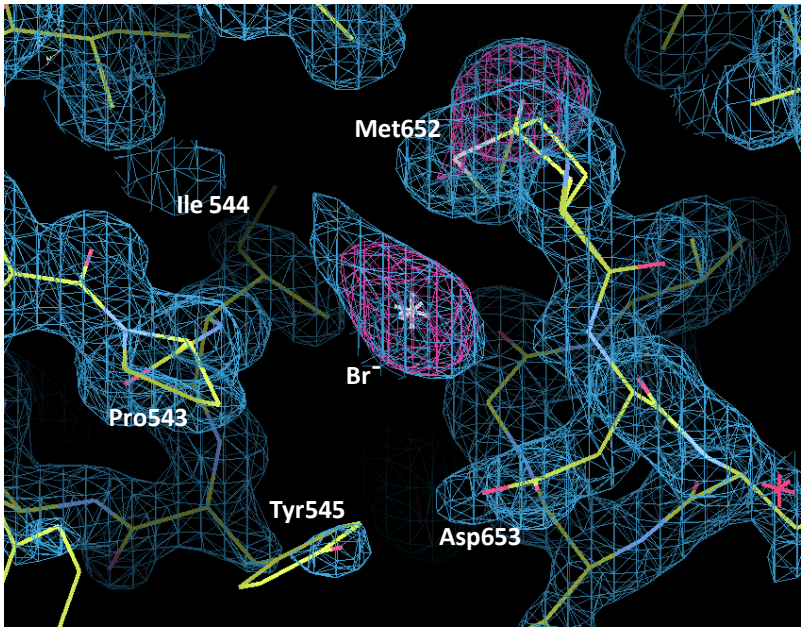


Figure 3.1 *Electron density map (blue lattice) and atomic model of the anion-binding site. The electron density map of the anomalous signal (pink lattice) detects the anion Br-'s position.*

The Br^- anion occupies a cavity presents on the protein surface (Figure 3.2.a) predicted to face the membrane surface. The amino acids implicated in the binding site (highlighted in the figure 3.2.b) are: methionine 652, flexible and in double conformation, phenylalanine 651, tyrosine 545, proline 543, isoleucine 544 and aspartic acid 653, the only charged residue that forms a H-bond with tyrosine 545.

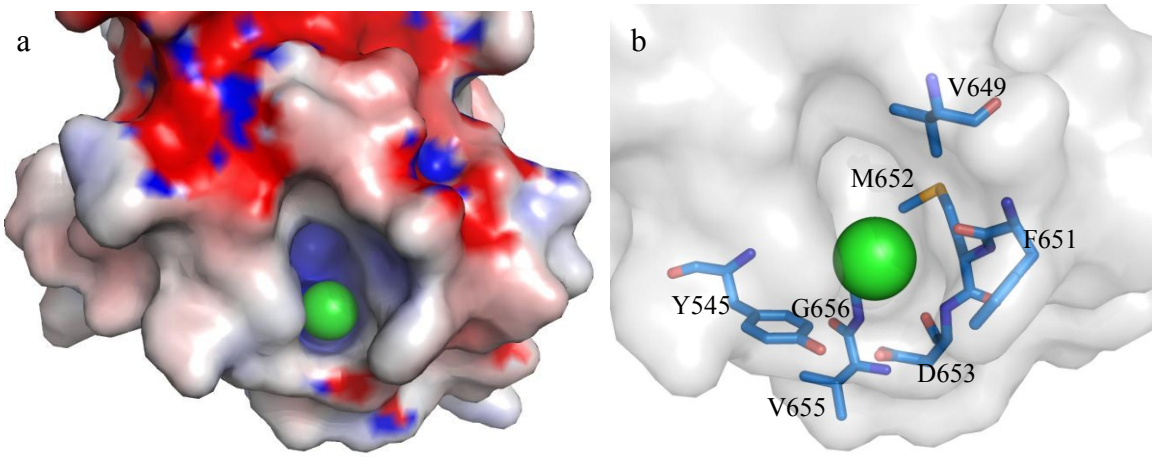


Figure 3.2 a) *The Br- anion (green sphere) occupies a cavity presents on the protein surface colored on the basis of the electrostatic potential.* ***b)*** *Detail of the anion-binding site and of the residues implicated.*

Single-site mutations on amino acids implicated in the anion-binding site were analyzed by PoPMuSiC V2.1 tool (Prediction of Protein Mutations Stability Changes) that evaluates the changes in stability of a given protein under single-site mutations, on the basis of the protein's structure (Table 3.2) (Dehouck et al., 2009).

Table 3.2 Stability changes resulting from given mutations predicted by PoPMuSiC tool

Mutation	ΔΔG (kcal/mol)	Notes
P543D	0.45	Negative charge near halogens
Y545W	0.04	H-bond with D653 abolished
Y545F	0.36	H-bond with D653 abolished
F646E	4.02	Negative control
F651T	-0.28	Stabilizing mutation but likely it disturbs membrane interaction
M652I	0.45	Different shape of cavity because Met side chain is flexible while that of Ile is not
D653T	0.23	Negative charge and H-bond with Y545 abolished
V655R	-0.33	Stabilizing mutation but likely it disturbs membrane interaction
G656I	0.21	Smaller cavity
L676E	4.79	Negative control

Most of these mutations were tested in vitro on the full-length SLC26A5 transporter by Prof. Dominik Oliver's laboratory at the Department of Neurophysiology, Institute of Physiology and Pathophysiology, Philipps-University, Marburg (Germany), in order to characterize them in terms of membrane localization and electromotility function, measuring NLC (Table 3.3).

Table 3.3 Results on full-length SLC26A5 transporter mutants tested in vitro

Mutation	TM localization	NLC
P543D	-	-
Y545F	+	+
F646E	-	-
F651T	+	-
D653T	+	-
V655R	+	+
G656I	-	-
L676E	-	-

According to achieved results, it is possible to develop the following considerations:

- F646E and L676E were predicted completely destabilizing the STAS structure (ΔΔG > +4) and indeed they cause mistrafficking probably by inducing misfolding. They are a sort of control of the reliability of the prediction ability of PoPMuSiC.

- Y545F and D653T: the negative charge of D653 is important for NLC while it is not the H-bond with Y545 (abolished by the Y545F mutation that retains localization and NLC). This supports the hypothesis that the D653 negative charge can have an important "gating" role on the anions binding or movement.
- F651T: this stabilizing mutation ($\Delta\Delta G$ = -0.28) was predicted not to alter anion binding but possibly interfere with the interaction of the STAS with the membrane (or transmembrane portions of prestin). This seems the case, the membrane localization is correct but there is no NLC, indicating that this is a residue important for the proper functional interactions of prestin.
- G656I and P543D: they were predicted to strongly interfere with anion binding (therefore with NLC in general) hopefully without affecting the protein localization ($\Delta\Delta G$ values not too destabilizing). However they cause more severe damages indicating that they are even more important than predicted.
- V655R: it was predicted as a STAS stabilizing mutation that, however, could interfere with important functional interactions with the membrane. The last one is not the case because Arg can be accommodated in the lipid bilayer, possibly interacting with the negative charges of the membrane.

Four STAS domain mutants were chosen: F651T, M652I, D653T and Y545W.

The choice of F651T and D653T comes from Prof. Oliver's data, since full-length SLC26A5 transporter's folding in the transmembrane is correct, while its function is abolished. Therefore we expected that these mutations do not invalidate the overall 3D structure because of the correct membrane location, but that they determine some variations around the mutation in the anion-binding site that cause loss of function.

M652I and Y545W mutations weren't functionally analyzed, but the first could introduce less flexibility in the anion-binding site; while the second, since the loss of H-bond with D653 doesn't seem essential for function and mutation effect doesn't seem destabilizing ($\Delta\Delta G$ = 0.04), it could be useful for fluorescence analysis. We expected that these mutations would not invalidate the overall 3D structure because they are predicted to be not destabilizing ($\Delta\Delta G_{M652I}$ = 0.45; $\Delta\Delta G_{Y545W}$ = 0.04).

3.1.2. Cloning

To design the STAS domain variants with the variable loop, the nucleotide sequences codifying the proteins of interest were amplified by PCR starting from the cDNA of the entire protein using the primer sets indicated in table 3.4. The final DNA sequences were inserted into the pET SUMO (small ubiquitin-like modifier; Invitrogen) vector in frame with an N-terminal poly(His)$_6$ tag flanked by the SUMO protein, using a pET SUMO Cloning kit (Invitrogen).

To design the STAS domain variants devoid of the variable loop, the nucleotide sequences before and after the loop were separately amplified by PCR starting from the cDNA of the entire protein using the primer sets indicated in the table 3.4 in order to insert a BamHI site. The two PCR products were ligated by means of a BamHI restriction site encoding for the dipeptide GlySer. This fragment was phosphorylated using T4 Polynucleotide Kinase (New England BioLabs) and inserted into the pBluescript II SK (+/−) storage phagemid (Agilent Technologies) previously digested to create blunt ends by EcoRV restriction enzyme. The final DNA sequences were inserted into the pET SUMO (Invitrogen) vector in frame with an N-terminal poly(His)$_6$ tag flanked by the SUMO protein, using a pET SUMO Cloning kit (Invitrogen).

SUMO Protein expression system offers the following advantages:

- may increase expression and solubility of recombinant fusion proteins;
- allows generation of native protein using SUMO Protease with no extra amino acids added between the cleavage site and the start of the desired protein;
- the tertiary structure of the SUMO protein is recognized by a cysteine protease, SUMO Protease, which specifically cleaves conjugated SUMO from target proteins;
- SUMO fusion protein and SUMO Protease can be removal easily after cleavage by affinity chromatography on a nickel-chelating resin (Malakhov et al., 2004).

To increment the efficiency of SUMO protease, a Serine residue was introduced at the N-terminal of all constructs.

Table 3.4 Oligonucleotide primers used for the indicated constructs

Protein	Source organism	Construct	Primer	Sequence
SLC26A5/ prestin	*Rattus norvegicus*	[505-718]	5' 3'	**agt**ccgagctacacagtcc ctaattcgcctcacggacttgg
SLC26A5/ prestin	*Gallus gallus*	S[511-733]	5' 3'	**agt**agacctcaatatagaatccttg ctactacctgtctttcccttgg
		S[511-569] GS[652-733]	5' 3' 5' 3'	**agt**agacctcaatatagaatccttg atggatccgtccactccagtct atggatccacaaatgtccactctt ctactacctgtctttcccttgg
SLC26A5/ prestin	*Danio rerio*	S[508-733]	5' 3'	**agt**cgccccaagaatgtcgttc ctactaaggaagatccctccaac
		S[508-566] GS[650-733]	5' 3' 5' 3'	**agt**cgccccaagaatgtcgttc atggatccgtcaattccagtctt atggatccacccccattcactc ctactaaggaagatccctccaac
SLC26A4/ pendrin	*Homo sapiens*	S[515-736]	5' 3'	**agt**tttccttcttggaatggc ctatcattgagatttcacttggt
		S[515-573] GS[653-736]	5' 3' 5' 3'	**agt**tttccttcttggaatggc taggatccatcaaatccaactg taggatccgtgccaatccata ctatcattgagatttcacttggt
Sultr1;2	*Arabidopsis thaliana*	S[502-653]	5' 3'	**agt**agacctagaactgcagttc ttatcagacctcgttggagag

The Bam H1 restriction site is underlined; N-terminal Serine residue is in bold font.

3.1.3. Mutagenesis

Oligonucleotide-directed mutagenesis uses short mutagenic oligonucleotides that incorporate one or a few substitutions, short insertions or deletions. Single-site mutagenesis was performed starting from the construct whose 3D structure was solved, using QuickChange Site-Directed Mutagenesis kit (Stratagene) and PfuTurbo DNA polymerase.

3.1.4. Protein expression

The STAS domains were produced as fusion proteins with the SUMO system, which guarantees proteolytic cleavage without extraneous amino acids at the N-terminal end.

E. coli BL21(DE3) cells, harboring the plasmid pET SUMO-STAS of the desired variant, were grown overnight (ON) at 37 °C in LB medium (10 g/l tryptone, 5 g/l yeast extract, and 10 g/l NaCl) supplemented with 50 μg/ml kanamycin. LB medium was

inoculated with this ON culture (ratio 1:50) and grown at 37 °C, in a suitable shaker. Protein expression was induced at an OD_{600} of 0.6 by adding 1 mM IPTG and prolonged for 4-5 h at 30 °C under vigorous shaking. Bacteria were harvested by centrifugation at 8500 g for 15'.

As regards the optimized expression protocol for SLC26A5/prestin STAS domain without variable loop from *D. rerio*, the protein expression was induced at an OD_{600} of 0.6 by adding 0.5 mM IPTG and 1% glucose, and prolonged ON at 16 °C under vigorous shaking. Bacteria were harvested by centrifugation at 8500 g for 15'.

SLC26A5/prestin STAS domain without variable loop from *G. gallus* was also purified as a ^{15}N-labeled protein for NMR studies. *E. coli* BL21(DE3) cells, harboring the plasmid coding for this domain, were grown overnight (ON) at 37 °C in ^{15}N-labeled M9 minimal medium (Na_2HPO_4 47.75 mM, KH_2PO_4 22.04 mM, NaCl 8.55 mM, $^{15}NH_4Cl$ 18.69 mM, $C_6H_{12}O_6$ 222 mM, $MgSO_4$ $2 \cdot 10^{-3}$ mM, $CaCl_2$ $0.1 \cdot 10^{-3}$ mM. Microelement mix in HCl 0.1 M: H_3BO_3 $12.8 \cdot 10^{-6}$ mM, $MnCl_2 \cdot 4H_2O$ $23.6 \cdot 10^{-6}$ mM, $ZnSO_4 \cdot H_2O$ $17.4 \cdot 10^{-6}$ mM, $Na_2MoO_4 \cdot 2H_2O$ $8.3 \cdot 10^{-6}$ mM, $CoCl_2$ $7.7 \cdot 10^{-6}$ mM, $FeSO_4 \cdot 7H_2O$ $26.3 \cdot 10^{-6}$ mM, $CuCl_2 \cdot 2H_2O$ $2.8 \cdot 10^{-6}$ mM. Vitamin mix: D-pantothenate$\cdot 1/2Ca^{2+}$ $0.42 \cdot 10^{-3}$ mM, Choline$\cdot Cl^-$ $0.72 \cdot 10^{-3}$ mM, Folic Acid$\cdot 2H_2O$ $0.21 \cdot 10^{-3}$ mM, myo-Inositol $1.1 \cdot 10^{-3}$ mM, Nicotinamide $0.82 \cdot 10^{-3}$ mM, Pyridoxine$\cdot$HCl $0.49 \cdot 10^{-3}$ mM, Riboflavin $26.6 \cdot 10^{-6}$ mM, Thiamine$\cdot$HCl $0.29 \cdot 10^{-3}$ mM) supplemented with 50 μg/ml kanamycin. ^{15}N-labeled M9 minimal medium was inoculated with this ON culture (ratio 1:10) and grown at 37 °C, in a suitable shaker. Protein expression was induced at an OD_{600} of 0.9 by adding 1 mM IPTG and prolonged for 4 h at 37 °C under vigorous shaking. Bacteria were harvested by centrifugation at 8500 g for 15'.

3.1.5. Protein purification

Bacteria were suspended in buffer A [50 mM Na_2HPO_4, 300 mM NaCl, and 10 mM β-mercaptoethanol (pH 8.0)] containing 10 mM imidazole, supplemented with protease inhibitors (Roche) and lysed with a French press (Thermo Spectronic) at high pressure.

The lysate was centrifuged to remove cell debris at 27000 g for 30' and applied onto a His-Trap column (GE Healthcare) equilibrated with buffer A containing 10 mM imidazole. After extensive washing with buffer A containing 20 mM imidazole, the fusion protein was eluted in a single peak increasing the concentration gradient of imidazole. Fractions containing the protein were pooled, diluted into buffer A, and

concentrated by ultrafiltration. For proteolytic cleavage of the $(His)_6$-SUMO-tag, the sample was incubated with SUMO protease 1 (LifeSensors) ON at 4 °C. The proteolytic product was applied onto the His-Trap column and immediately recovered to separate the purified protein from the $(His)_6$-SUMO-tag, the uncleaved fusion protein and the protease, retained in the column. The cleaved protein was further purified by SEC performed on an Äkta FPLC chromatographic system (GE Healthcare) using a Superdex 75 prep-grade 16/60 column or a Superdex 75 10/300 GL column (GE Healthcare) equilibrated with 20 mM Tris-HCl, 150 mM NaCl, and 5 mM DTT (pH 7.5).

The purification of the ^{15}N-labeled SLC26A5/prestin STAS domain without variable loop from *G. gallus* follows the same steps for unlabeled proteins except for the buffer of SEC that is: K_2HPO_4 27 mM, KH_2PO_4 5 mM, NaCl 50 mM, EDTA 1 mM, DTT 5 mM (pH 6.5).

3.1.6. Dynamic light scattering

DLS data were recorded on a Zetasizer NanoS instrument (Malvern Instruments Ltd.) at 20 °C, using a quartz cuvette and 20 µl of sample. Protein solutions were filtered using 0.22 µm filters. The data were recorded and analyzed with the Dispersion Technology Software (Malvern).

3.1.7. CD spectroscopy

CD data were recorded on a Jasco J-715 spectropolarimeter, using quartz cuvettes of 0.02 cm path-length. The spectra were determined as an average of 10 scans. The protein concentration was 1 mg/ml in 50 mM Na_2HPO_4 and 150 mM NaCl (pH 7.5). The data were recorded and analyzed with Spectra Manager Software (JASCO).

3.1.8. Differential Scanning Fluorimetry

To monitor thermal protein unfolding, the fluorescent dye Sypro orange (Sigma-Aldrich) was used. Thermofluor assay was conducted on a Mini Opticon Real-Time PCR detection system (Bio-Rad). Solution of 1 µl of 10 mg/ml protein, 1 µl of Sypro orange 100x and 18 µl of test compound were added to the 48-well low-profile plates (Bio-Rad).

The plates were spun and heated from 20 to 90 °C, with a heating rate of 0.5 °C/min. The data were recorded and analyzed with the CFX Manager Software (Bio-Rad).

3.1.9. Protein crystallization

Crystallization trials using commercial kits (Qiagen, Molecular Dimensions and Emerald Biosystems) based on sparse matrix, grid screen, and/or ionic sampling, were performed by vapor diffusion (with the sitting drop method) techniques, using the Oryx8 automatic system (Douglas Instrument).

Since the optimal solution conditions for nucleation of the crystals are not the ideal ones to support their subsequent growth, we also used the most popular technique for the separation of nucleation and growth, the so-called 'seeding'. Seeding experiments fall in four general categories: (i) macroseeding (or seed transfer), (ii) microseeding, (iii) streak and (iv) cross-seeding and epitaxial (Bergfors, 2003; D'Arcy et al., 2007; Benvenuti and Mangani, 2007). Microseeding experiments, using microcrystalline precipitate, needles and spherulite, were performed both manually and automatically. As regards the rat STAS domain with variable loop, also cross-seeding experiments using crystals of the same STAS domain but devoid of the loop were tried.

STAS domain mutants D653T and Y545W were concentrated to 17.9 mg/ml and 15.5 mg/ml respectively for crystallization purposes. High-quality hexagonal crystals were grown at 20 °C, using the following precipitant solution: 2 M ammonium sulfate, 5% (w/v) polyethylene glycol (PEG) 400, and 0.1 M 4-morpholineethanesulfonic acid (MES, pH 6.5) (Ammonium Sulfate Suite no. 87; Qiagen) adding 0.1% (w/v) octyl-β-D-glucopyranoside (OG, Sigma). For anion-binding site studies were also added 180 mM KBr.

3.1.10. Data collection, structure determination, and refinement

Datasets at 0.91 Å wavelength (corresponding to the absorption edge of bromine) were collected at the ELETTRA-Synchrotron beamline XDR1 (Trieste, Italy). Datasets were measured at 100 K using the precipitant solution including 16% (v/v) glycerol as cryoprotectant. Crystals belong to space group $P3_121$, with unit cell parameters reported in Table 3.5. Diffraction data were processed with XDS (Kabsch 2010) and with SCALA from the CCP4 suite (Evans 2005).

For mutant Y545W in absence of bromine atoms, structure was solved by molecular replacement (MR). For mutants Y545W and D653T in presence of bromine atoms, structures were solved both by MR and by SAD (Single-wavelength anomalous dispersion), in the last case using the anomalous signal of bromines. It is well known that, in general, structures solved by MR can be biased towards the search "phasing" model. In our case, structures from MR and SAD do not show any relevant difference once refined, reflecting the minimal sequence difference with the search model (only one single amino acid).

For molecular replacement, Phaser (McCoy et al., 2007) (CCP4) rotation and translation searches were performed using as search model the coordinates of the wt STAS domain taken from PDB ID 3LLO. For SAD, Autosol and Autobuild (Phenix) (Adams et al., 2002) were used to identify the initial phases from the anomalous signal.

Models were then refined alternating several cycles of automatic refinement and manual model building with Coot (Emsley and Cowtan, 2004). The refinement (with isotropic atomic B-factors) was performed in parallel with phenix.refine (Phenix) and with refmac5 (Murshudov et al., 1997), used independently, with almost identical final results.

During refinement water molecules were added to the model, both automatically and manually, and those with final high B-factors were excluded. One molecule of OG for each structure was also introduced and refined. Br atoms were added and refined according to their positions clearly visible in anomalous difference maps. Finally also the TLSs (Translation/Libration/Screw parameterization) were added. Statistics of the refinements and final models are reported in table 3.5.

Table 3.5 Data collection and refinement statistics for D653T and Y545W mutants

	D653T		**Y545W**		
Phase problem resolution	MR	SAD	MR	SAD	MR
150 mM KBr	YES	YES	YES	YES	NO
Visible residues	[505-553] [638-718]	[506-550] [638-718]	[506-553] [638-718]	[506-550] [638-718]	[506-550] [638-718]
Data collection statistics	ELETTRA beamline XDR1, λ=0.91 Å				
Space group	$P3_121$				
Cell dimensions a, b, c (Å) α, β, γ (°)	 62.02 62.02 66.73 90.0 90.0 120.0		 62.04 62.04 66.70 90.0 90.0 120.0		 61.67 61.67 66.68 90.0 90.0 120.0
Total number of observations	241,327 (33,977)		303,448 (44,948)		417,847 (61,195)
Total number of unique	12,485 (1,777)		15,198 (2,152)		21,771 (3,137)
Resolution (Å)	41.84-1.88 (1.98)		41.84-1.75 (1.84)		41.68-1.55 (1.63)
R_{merge} (%)	9.5 (88.7)		4.8 (84.3)		4.6 (148.0)
R_{meas} (%)	10.0 (93.5)		4.8 (88.4)		4.8 (151.9)
Mn[I/σ(I)]	20.2 (3.9)		40.0 (4.1)		32.5 (2.4)
Completeness (%)	100 (100)		99.0 (99.3)		100 (100)
Anomalous completeness (%)	100 (100)		99.3 (98.6)		-
Multeplicity	19.3 (19.1)		20.0 (20.9)		19.2 (19.5)
Anomalous multeplicity	10.3 (10.0)		10.6 (10.8)		-
Refinement statistics					
R_{work} (%)	20.60	19.65	20.56	19.04	23.09
R_{free} (%)	24.74	23.30	24.40	21.00	25.68

$$R_{merge} = \frac{\sum_{hkl}\sum_{i}\left|I_{(hkl)i} - \left\langle I_{(hkl)}\right\rangle\right|}{\sum_{hkl}\sum_{i}\left\langle I_{(hkl)}\right\rangle}\ ; \quad R_{meas} = \frac{\sum_{hkl}\left(\frac{N}{N-1}\right)^{1/2}\sum_{i}\left|I_{(hkl)i} - \left\langle I_{(hkl)}\right\rangle\right|}{\sum_{hkl}\sum_{i}\left\langle I_{(hkl)}\right\rangle}\ ; \quad R_{work} = \frac{\sum_{hkl}\left\|F_{obs}\right| - \left|F_{calc}\right\|}{\sum_{hkl}\left|F_{obs}\right|}\ .$$

3.2. Results

3.2.1. STAS domains overview

Eight STAS domain variants from different related-species and four STAS domain mutants of STAS domain whose 3D structure was solved (PDB code = 3LLO, first row in the Table 3.6) were cloned successfully in the pET SUMO expression vector. Table 3.6 summarizes cloning, expression, purification, characterization and crystallization results for each of the selected constructs.

Table 3.6 Summary of SulP/SLC26 anion transporters production and characterization

Protein/ Source organism	**Sequence**	**Loop**	**Mutant**	**Cloning**	**Expression level/ Solubility**	**Purifi-cation**	**Characte-rization**	**Crystal-lization**
SLC26A5/ *Rattus norvegicus*	[505-563] GS[637-718]	W/O	-	YES	High/ High	YES	SEC, CD, DLS, DSF	YES
			F651T	YES	High/ High	YES	SEC, CD, DLS, DSF	In progress
			M652I	YES	High/ High	YES	SEC, CD, DLS, DSF	In progress
			D653T	YES	High/ High	YES	SEC, CD, DLS, DSF	YES
			Y545W	YES	High/ High	YES	SEC, CD, DLS, DSF	YES
	[505-718]	With	-	YES	Medium /High	YES	SEC, CD, DLS, DSF	In progress
SLC26A5/ *Gallus gallus*	S[511-569] GS[652-733]	W/O	-	YES	High/ High	YES	SEC, CD, DLS, DSF	In progress
	S[511-733]	With	-	YES	-	-	-	-
SLC26A5/ *Danio rerio*	S[508-566] GS[650-733]	W/O	-	YES	Medium/ Medium	YES	SEC	-
	S[508-733]	With	-	YES	Medium/ Low	-	-	-
SLC26A4/ *Homo sapiens*	S[515-573] GS[653-736]	W/O	-	YES	Medium/ Low	YES	-	-
	S[515-736]	With	-	YES	Medium/ Low	YES	-	-
Sultr1;2/ *Arabidopsis thaliana*	S[502-653]	With	-	YES	-	-	-	-

Some variants of the STAS domain, among which SLC26A5 from *D. rerio* with variable loop and SLC26A4 from *H. sapiens* both with and without variable loop, showed low solubility levels (Figure 3.3.A, lanes 3 and 4), despite the presence of the SUMO-tag that may increase the solubility of recombinant fusion proteins. Furthermore, after the

proteolytic cleavage, almost all the protein precipitated as insoluble aggregates (Figure 3.3.B, lanes 2 and 3).

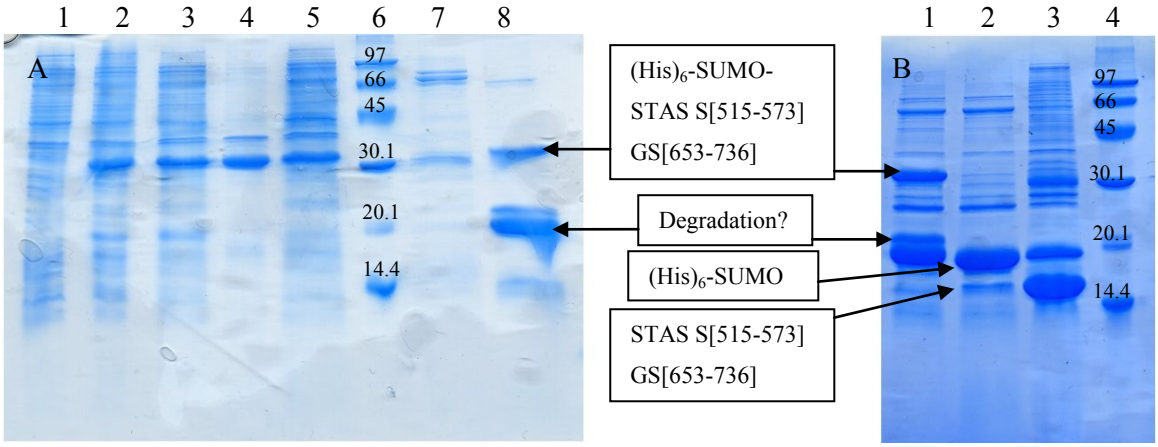


Figure 3.3 *Coomassie-stained SDS-PAGE of expression in BL21(DE3) of SLC26A4 STAS domains S[515-573]GS[653-736] from H. sapiens.* ***(A)*** *Lane 1: control, not induced bacterial cells. Lane 2: IPTG induced cells. Lane 3: soluble portion of bacterial lysate. Lane 4: insoluble fraction of bacterial lysate. Lane 5: flow through. Lane 6: low molecular weight protein markers in kDa. Lanes 7 and 8: fusion proteins eluted from the IMAC column.* ***(B)*** *Lane 1: fusion proteins eluted from the IMAC column before the proteolytic cleavage: $(His)_6$-SUMO-STAS S[515-573]GS[653-736] (29 kDa). Lane 2: soluble portion of the sample after the proteolytic cleavage by the SUMO protease: STAS S[515-573]GS[653-736] (16 kDa) and $(His)_6$-SUMO (13 kDa). Lane 3: insoluble portion of the sample after the proteolytic cleavage: STAS S[515-573]GS[653-736] and $(His)_6$-SUMO. Lane 4: protein markers (kDa).*

The amount and the quality of these proteins were not sufficient for structural studies, despite we tried different protocols to increase the solubility, by lowering induction temperature and inductor concentration, by adding agents that may promote protein solubility like detergents, sugars, amino acids, salts, etc. (Bondos and Bicknell, 2002; Vuillard et al., 1995; Hamada et al., 2009). We also tried to purify them from inclusion bodies following different refolding protocols, among which the on-column refolding procedure used for Rv1739c STAS domain from *M. tuberculosis* (Oganesyan et al., 2004; Sharma et al., 2009).

Instead, as regards the SLC26A5 STAS domain from *D. rerio* without variable loop, the solubility level increased by lowering induction temperature and inductor concentration (Figure 3.5) but, despite this, the final yield was not sufficient for structural studies (this purification is reported in section 3.2.2).

Furthermore, as regards the STAS domains of SLC26A4 from *H. sapiens,* both with and without variable loop, after IMAC purification and before the proteolytic cleavage, two main bands, a little higher than the SUMO tag, appeared in the SDS-PAGE (Figure

3.3.A, lane 8 and Figure 3.3.B, lane 1). These bands disappeared after the cleavage with SUMO protease (Figure 3.3.B, lanes 2 and 3) therefore they were supposed to be owed to protein degradation maybe for the presence of a flexible loop between the STAS domain and SUMO.

Other variants of the STAS domain, among which SLC26A5 from *R. norvegicus* with variable loop, SLC26A5 from *G. gallus* without variable loop and SLC26A5 mutants from *R. norvegicus* without variable loop, were purified in milligrams quantities and characterized in solution by classical biophysical methods (SEC, CD, DLS, DSF). Two mutants of the STAS domain were successfully crystallized and the structural information has been integrated with functional data that come from international collaborations. The purification and characterization of SLC26A5 STAS from *R. norvegicus* with variable loop and from *G. gallus* without variable loop are reported in section 3.2.2. The purification, characterization and X-ray analysis of SLC26A5 STAS domain mutants of from *R. norvegicus* are described in details in section 3.2.3.

3.2.2. SLC26A5 STAS domains from R. norvegicus, G. gallus and D. rerio

The STAS domain corresponding to that whose 3D structure was solved, but supplied with the variable loop, was purified and compared by different biophysical techniques (CD, DSF, DLS) with the STAS domain variant devoid of the variable loop in order to structurally characterize the variable loop and to understand its possible role in SLC26A5 oligomerization and function.

The SLC26A5 STAS domains from *G. gallus* and *D. rerio* devoid of the variable loop were purified and characterized to have information about STAS domain evolution and its different functions in different orthologs.

3.2.2.1. Expression and purification

SLC26A5 STAS domain with variable loop from *R. norvegicus* and that from its ortholog *G. gallus* but devoid of variable loop, cloned in the pET SUMO expression vector, were successfully overexpressed mainly in soluble forms in *E. coli* BL21(DE3) (Figure 3.4 lanes 3).

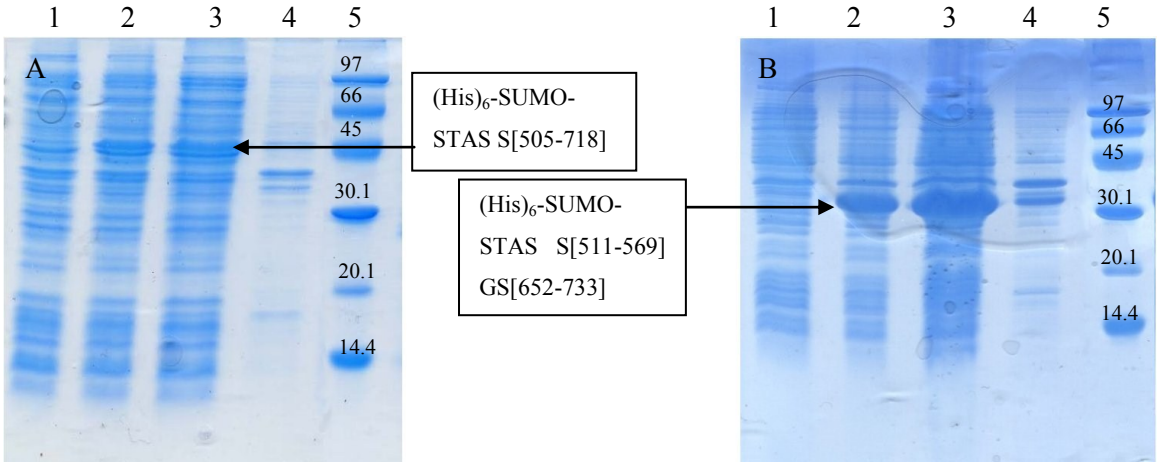


Figure 3.4 *Coomassie-stained SDS-PAGE of expression in BL21(DE3) of* ***(A)*** *$(His)_6$-SUMO-STAS domain [505-718] of SLC26A5 from R. norvegicus (37 kDa) and of* ***(B)*** *$(His)_6$-SUMO-STAS domain S[511-569]GS[652-733] of SLC26A5 from G. gallus (29 kDa). Lane 1: control, not induced bacterial cells. Lane 2: IPTG induced cells. Lane 3: soluble portion of bacterial lysate. Lane 4: insoluble fraction of bacterial lysate. Lane 5: protein markers (kDa). The proteins of interest are indicated by arrows.*

As regards the SLC26A5 STAS domain from *D. rerio* without variable loop, cloned in the pET SUMO expression vector, the solubility level of the protein expressed in *E. coli* BL21(DE3) was low (Figure 3.5.A lanes 3 and 4), but it was increased by lowering induction temperature and inductor concentration (Figure 3.5.B lanes 4 and 5).

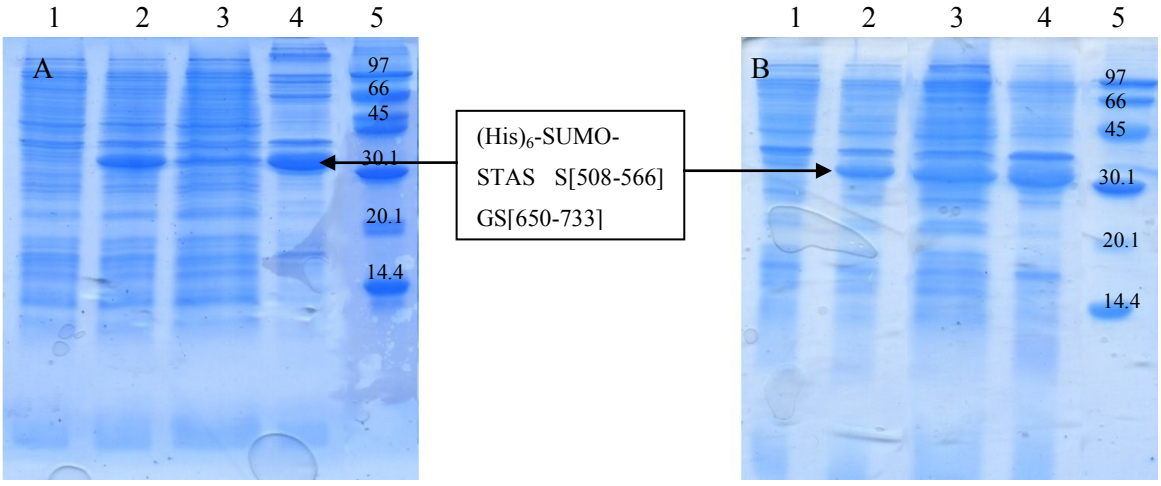


Figure 3.5 *Coomassie-stained SDS-PAGE of expression in BL21(DE3) of $(His)_6$-SUMO-STAS domains S[508-566]GS[650-733] of SLC26A5 from D. rerio (29 kDa).* ***(A)*** *Standard expression.* ***(B)*** *Expression with lower induction temperature and inductor concentration. Lane 1: control, not induced bacterial cells. Lane 2: IPTG induced cells. Lane 3: soluble portion of bacterial lysate. Lane 4: insoluble fraction of bacterial lysate. Lane 5: protein markers (kDa). The protein of interest is indicated by arrows.*

The soluble fractions of the $(His)_6$-SUMO fusion proteins were purified by an IMAC affinity step followed by the proteolytic cleavage of the $(His)_6$-SUMO-tag. After

an incubation at 4 °C ON, the cleavage with SUMO protease was very efficient for all the constructs, as shown in figure 3.6.

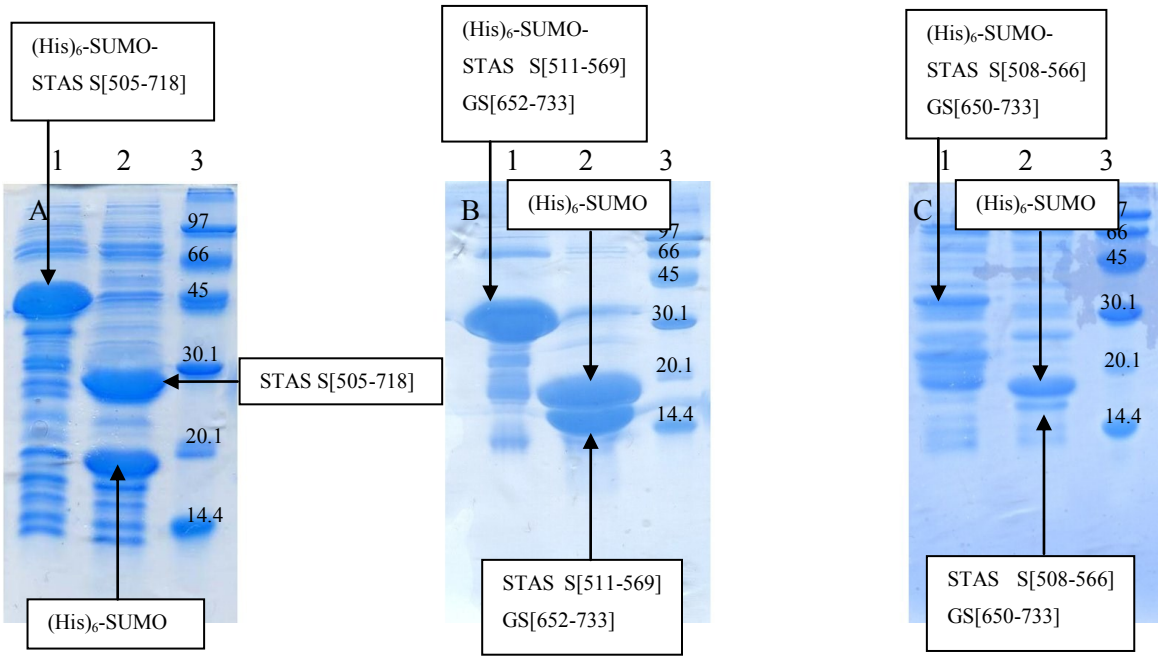


Figure 3.6 *Coomassie-stained SDS–PAGE of the proteolytic cleavage of the $(His)_6$-SUMO tag from* ***(A)*** *SLC26A5 STAS domain [505-718] from R. norvegicus,* ***(B)*** *SLC26A5 STAS domain S[511-569]GS[652-733] from G. gallus and* ***(C)*** *SLC26A5 STAS domain S[508-566]GS[650-733] from D. rerio. Lane 1: fusion proteins eluted from the IMAC column before the proteolytic cleavage; $(His)_6$-SUMO-STAS [505-718] (37 kDa), $(His)_6$-SUMO-STAS S[511-569]GS[652-733] (29 kDa), $(His)_6$-SUMO-STAS S[508-566]GS[650-733] (29 kDa). Lane 2: samples after the proteolytic cleavage by the SUMO protease; STAS [505-718] (24 kDa), $(His)_6$-SUMO (13 kDa), STAS S[511-569]GS[652-733] (16 kDa), S[508-566]GS[650-733] (16 kDa). Lane 3: protein markers in kDa.*

The proteolytic products were purified by another nickel affinity chromatography to separate the STAS domain from $(His)_6$-SUMO, the uncleaved fusion protein and the protease, followed by a gel permeation step. The elution profiles and final results of proteins concentration are shown in figure 3.7.

Fractions corresponding to the main peak (Figure 3.7.A and 3.7.B) were collected and concentrated for the following characterization. The final yield of purified proteins were around 2 mg and 10 mg per liter of culture medium for STAS domain [505-718] from *R. norvegicus* and S[511-569]GS[652-733] from *G. gallus*, respectively.

As regards the STAS domain from *D. rerio* (Figure 3.7.C), high molecular weight aggregates were present in solution, as pointed out by the gel permeation elution profile. In the SEC elution profile there are also two peaks, both corresponding to the STAS domain (data not shown). Maybe they correspond to a dimeric (MW predicted= 24 kDa) and a monomeric (MW predicted= 15 kDa) form. The final yield of purified protein was around <0.2 mg per liter of culture medium, too low for structural characterization.

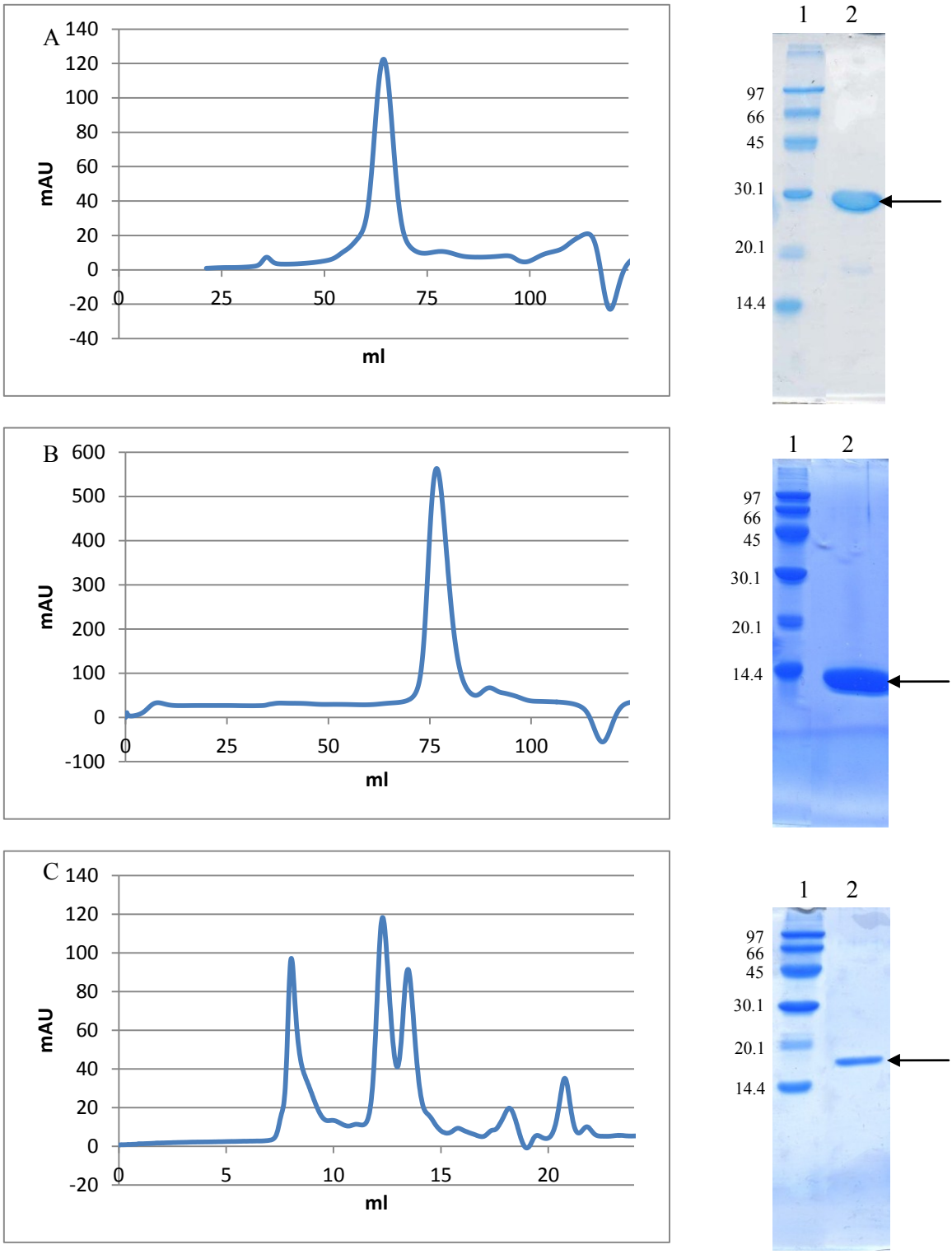


Figure 3.7 *SEC elution profiles and Coomassie–stained SDS–PAGEs after gel permeation chromatography of* ***(A)*** *SLC26A5 STAS domain [505-718] from R. norvegicus and* ***(B)*** *SLC26A5 STAS domain S[511-569]GS[652-733] from G. gallus. Column: HiLoad 16/60 Superdex 75 prep grade (GE Healthcare) equilibrated with 20 mM Tris-HCl, 150 mM NaCl and 5 mM DTT (pH 7,5).* ***(C)*** *SLC26A5 STAS domain S[508-566]GS[650-733] from D. rerio. Column: Superdex 75 10/300 GL (GE Healthcare) equilibrated with 20 mM Tris-HCl, 150 mM NaCl and 5 mM DTT (pH 7,5). Coomassie–stained SDS–PAGEs; Lane 1: protein markers (kDa). Lane 2: purified proteins (A) STAS [505-718] (24 kDa), (B) STAS S[511-569]GS[652-733] (16 kDa) and (C) STAS S[508-566]GS[650-733] (16 kDa).The proteins are indicated by arrows.*

3.2.2.2. *Circular dichroism*

To verify if the purified STAS domains have defined structures, they were characterized by circular dichroism spectroscopy, and the resulting CD spectra were compared with that of STAS domain crystallized.

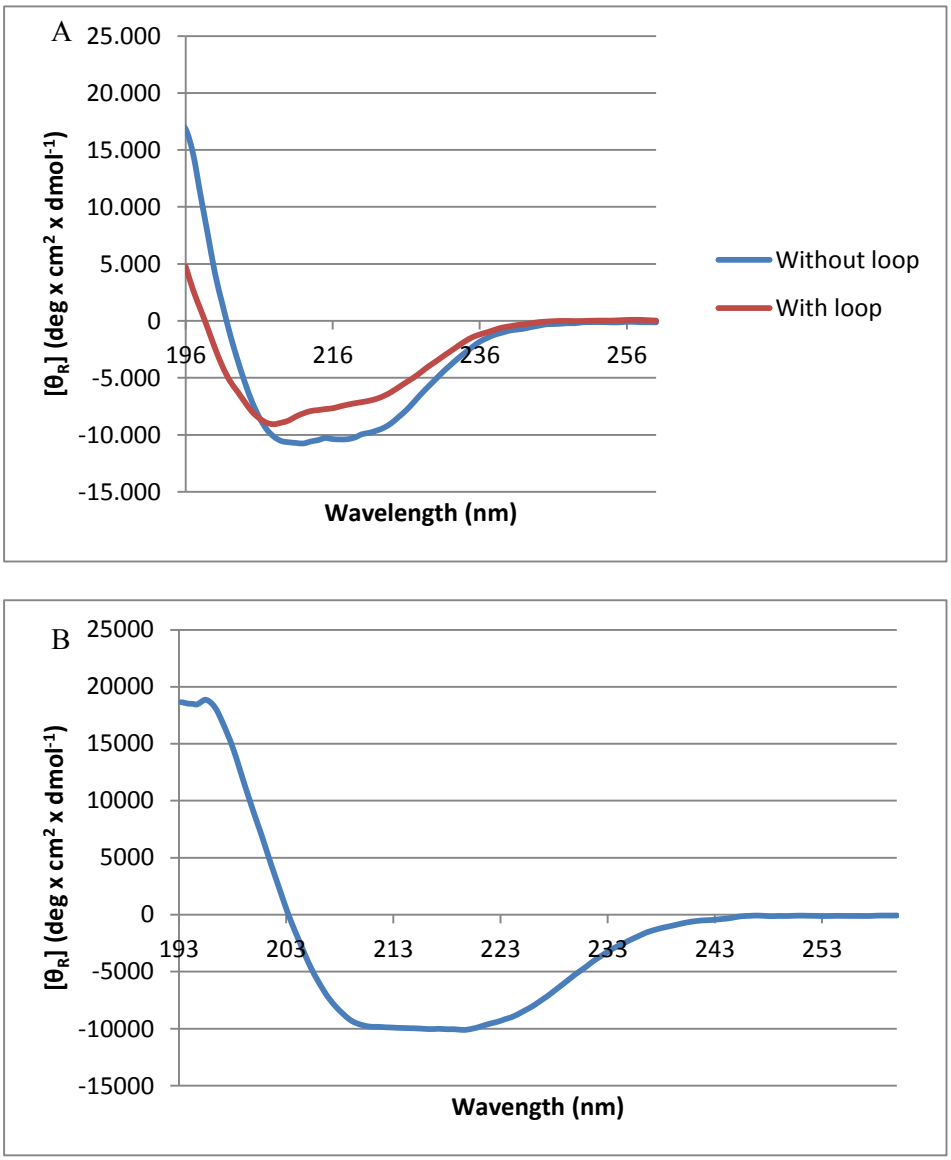


Figure 3.8 *Far-UV circular dichroism spectra of* ***(A)*** *SLC26A5 STAS domains from R. norvegicus with variable loop (in red) and without variable loop (in blue), and* ***(B)*** *SLC26A5 STAS domain S[511-569]GS[652-733] from G. gallus (in green). The proteins concentration was 1 mg/ml in 50 mM Na_2HPO4 and 150 mM NaCl (pH 7.5).*

Figure 3.8.A shows the comparison between STAS domains from *R. norvegicus* with (in red) and without (in blue) variable loop. Both domains have a mixed α/β secondary structure, but it is evident that the STAS domain devoid of loop has a higher

percentage of secondary structure, according to the secondary structure predictions that predict the variable loop as mostly unstructured in solution.

The CD spectrum of the STAS domain from *G. gallus* devoid of loop (Figure 3.8.B) shows two negative bands around 209 and 222 nm, indicative of the presence of α-helices. The overall shape of the spectra and the difference in the intensity of the two negative bands indicate the presence also of a certain amount of β-structure. The secondary structure experimentally determined is in accordance with that of its ortholog without loop.

3.2.2.3. Differential scanning fluorimetry (DSF)

The fluorescence-based thermal stability assay (or Thermofluor, DSF) was developed by Pantoliano and coworkers (US patent 6,020,141) as a high-throughput screen for buffer optimization and ligand-induced stabilization of proteins. This assay can be used to estimate the temperature of melting (Tm) and to identify stabilizing additives to obtain increased success rates in crystallization experiments. The Tm of SLC26A5 STAS domain from *R. norvegicus* without loop is higher (53 °C) than that of STAS domain with loop (40 °C) in the same buffer (20 mM Tris, pH 7.5), indicating that the deletion of this unstructured and flexible region makes more stable and structured the STAS devoid of loop. The Tm of SLC26A5 STAS domain from *G. gallus* is a little lower (44 °C) than that of its ortholog. Different buffers and salts were tested in order to improve the thermal stability, but no buffer or salt tested improved the Tm markedly.

3.2.2.4. Dynamic light scattering

It was reported that mammalian full-length SLC26A5 aggregates in living cells at the level of plasma membrane, forming stable dimers or tetramers that are supposed to be essential for the physiological function (Zheng et al., 2006; Detro-Dassen et al., 2008; Mio et al., 2008). The STAS domain of Sultr1;2 from *A. thaliana* homodimerizes, unlike that of the bacterial SulP/SLC26 transporters (Shibagaki et al., 2006). To investigate if the oligomerization is caused because of STAS domains interaction both STAS domains with loop and that devoid of loop were analyzed by Dynamic Light Scattering (DLS). The particle size distributions by number of SLC26A5 STAS domain from *R. norvegicus* with loop is showed in figure 3.9.A. The main form in solution is the monomeric one, as well

as for the STAS domain devoid of loop (data not shown), indicating that neither the loop, nor the STAS domain seem to cause the SLC26A5 oligomerization.

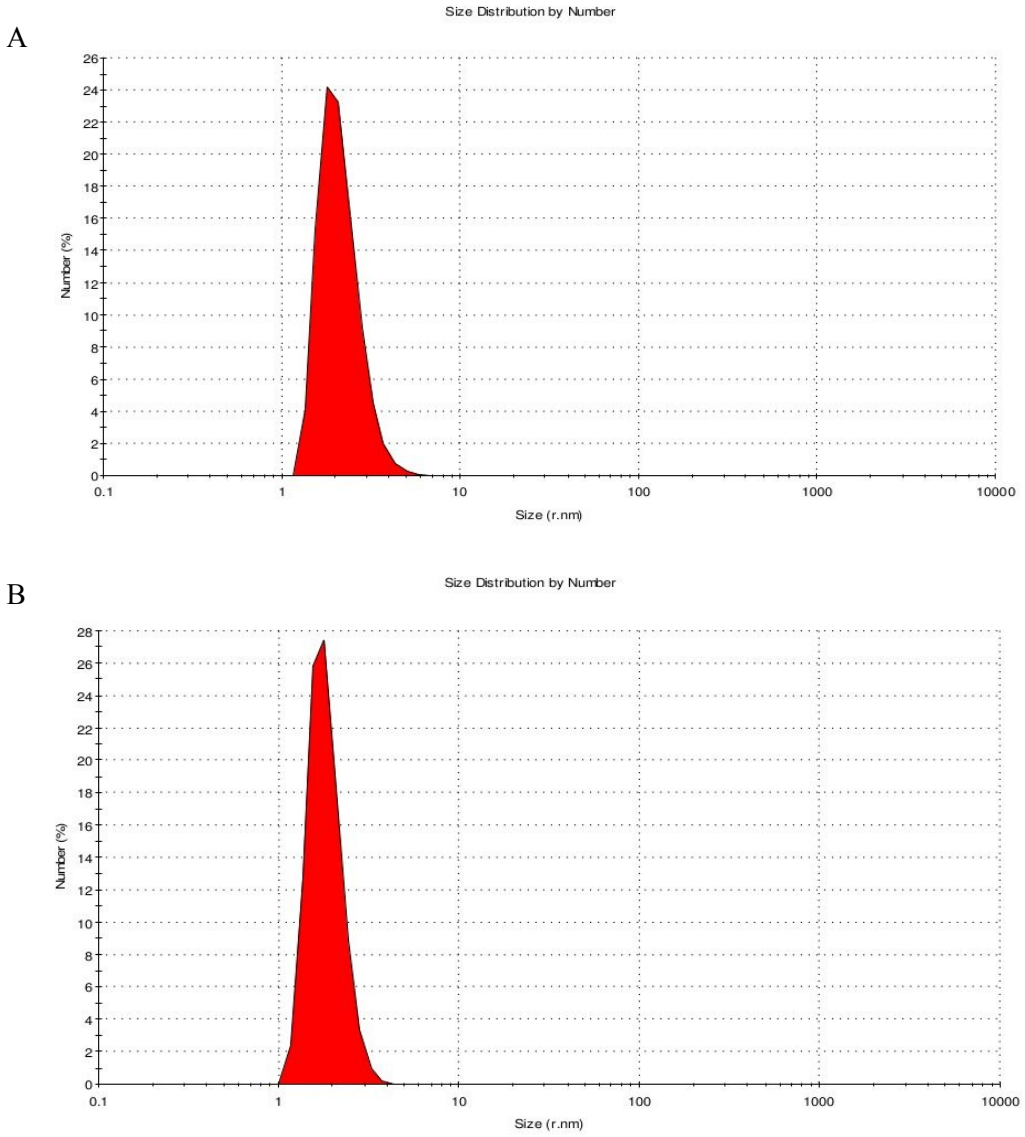


Figure 3.9 *Particle size distribution by number of* ***(A)*** *SLC26A5 STAS domains [505-718] from R. norvegicus and of* ***(B)*** *SLC26A5 STAS domain S[511-569]GS[652-733] from G. gallus. The proteins concentration was 5 mg/ml in 20 mM Tris-HCl, 150 mM NaCl and 5 mM DTT (pH 7.5).*

Zulauf and D'Arcy (Zulauf and D'Arcy, 1992) reported that monodispersity of proteins measured using DLS correlated well with crystallizability, in fact monodisperse preparations of macromolecules can be crystallized, whereas samples that aggregate randomly cannot. The STAS domain that was crystallized has a polydispersion of 32.2%; the polydispersion of the same STAS domain but with loop is higher (38.1%), while ortholog STAS domain has a lower polydispersion (23.1%, Figure 3.9.B). None of these

samples has high-molecular weight aggregates and the polydispersion is not too much high, therefore these constructs seem to be good candidates for crystallographic studies.

3.2.2.5. NMR studies

The biophysical analysis of the STAS domain from *G. gallus* was also integrated by NMR studies on the sample expressed in ^{15}N-labeled M9 minimal medium. The protocol of purification of ^{15}N-labeled STAS domain was the same of the unlabeled one, but with a less yield of protein purified, 5 mg per liter of culture medium. The two-dimensional HSQC spectrum obtained in collaboration with Dr. Bellanda, at the Department of Chemical Sciences, University of Padova, shows several peaks well separated, confirming that the protein is well folded and not aggregated.

3.2.2.6. Crystallization trials

Looking at the results of the characterization of the SLC26A5 STAS domains, we can affirm that they are good candidates for crystallographic studies. In fact they are well folded, structured, not aggregated and in conditions that assure a good thermal stability. Therefore several crystallization screenings were tested: The JCSG+ Suite, The MbClass Suite, The MbClass II Suite, The $AmSO_4$ Suite, The PACT Suite (Qiagen) and Structure screen 1 and 2 (Molecular Dimensions) for the STAS domain from *R. norvegicus* and moreover also The Anions Suite, The Cations Suite, The MPD Suite, The PEGs II Suite (Qiagen), Precipitant Synergy (Emerald Biosystems) and Morpheus™ (Molecular Dimensions) for the STAS domain from *G. gallus*. Different protein concentrations were also tested. Many small crystals were obtained to diffraction analysis to ELETTRA synchrotron (Trieste, Italy) or ESRF synchrotron (Grenoble, France) but images had few and very intense spots, typically generated by the diffraction of salts. Despite they were good candidates for crystallographic studies, it has not been possible to obtain crystals suitable for X-ray diffraction analysis. Probably the problem is due to the intrinsic low propensity of these STAS domains to build ordered aggregation forms, from which single crystals grow.

3.2.3. SLC26A5 STAS domain mutants

STAS domain mutants of 3LLO construct were performed to study the anion-binding site and the possible role of the STAS domain in the transport. Four STAS domain mutants of single amino acids involved in the anion-binding site were produced. Table 3.7 summarizes the characteristics of STAS domain mutants.

Table 3.7 Characteristics of STAS domain mutants

	Functional analysis	**Prediction**
F651T	Membrane localization, loss of function	Stabilizing mutation
M652I	-	Not destabilizing
D653T	Membrane localization, loss of function	Not destabilizing
Y545W	-	Not destabilizing

Functional analysis was performed in collaboration with Prof. Oliver on the full-length SLC26A5 transporter. Prediction of stability changes was performed by PoPMuSiC V2.1 tool.

All these four mutations were predicted to be not destabilizing by PoPMuSiC V2.1 tool, therefore they shouldn't invalidate the overall 3D structure and we expected that 3D structures of mutants resembled that of wt STAS domain. Furthermore F651T and D653T full-length SLC26A5 transporter mutants have a correct folding in the transmembrane while their function is abolished; therefore we hypothesized the presence of some variations causing loss of function in the anion-binding site around the mutation.

3.2.3.1. Expression and purification

The four SLC26A5 STAS domain mutants were produced following the protocol of QuickChange™ Site-Directed Mutagenesis kit (Stratagene). The proteins were successfully overexpressed mainly in soluble form and were purified using the same protocol described in the chapter 3.2.2.1: an IMAC affinity step followed by the SUMO protease proteolytic cleavage, another IMAC affinity step and finally size exclusions chromatography. The elution profiles are shown in figure 3.10.

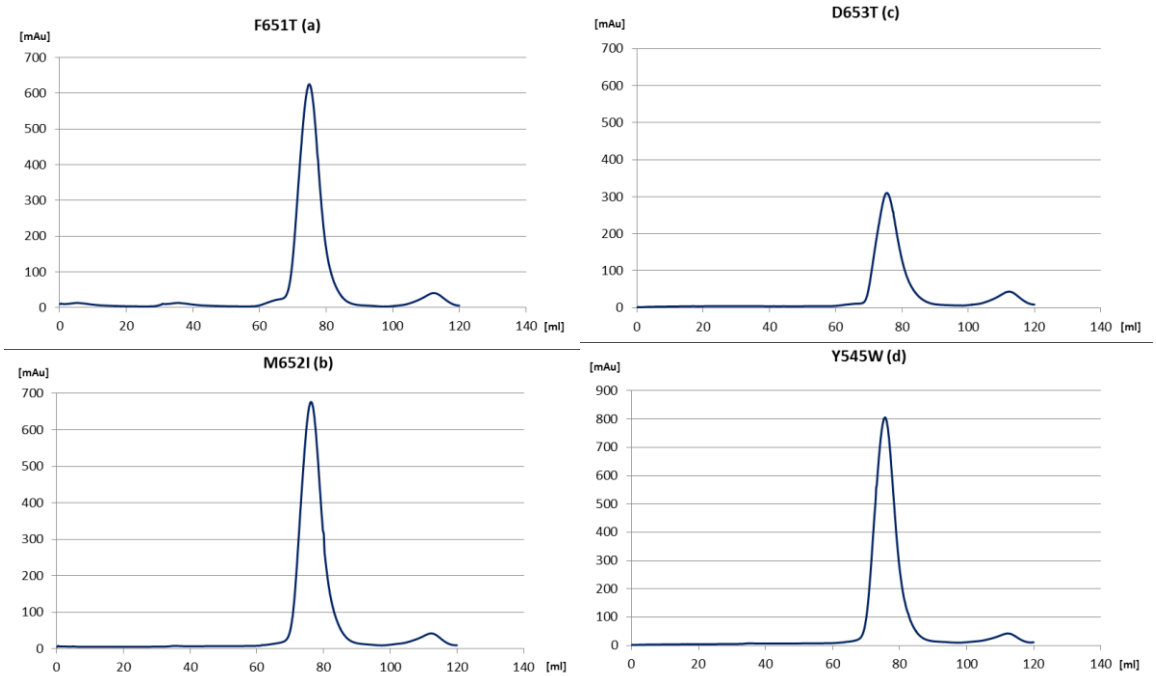


Figure 3.10 *SEC elution profiles of SLC26A5 STAS domain [505-563]GS[637-718] mutants from R. norvegicus. Column: HiLoad 16/60 Superdex 75 prep grade (GE Healthcare) equilibrated with 20 mM Tris-HCl, 150 mM NaCl and 5 mM DTT (pH 7.5).* ***(a)*** *F651T,* ***(b)*** *M652I,* ***(c)*** *D653T,* ***(d)*** *Y545W.*

The final yields of purified proteins were around 9 mg per liter of culture medium for F651T mutant, 10 mg for M652I, 5 mg for D653T and 9 mg for Y545W.

3.2.3.2. Circular dichroism spectroscopy

To verify if the mutations don't affect the secondary structure, the purified mutants were characterized by circular dichroism in the far-UV region and the CD spectra were compared to that of the wt STAS domain. The comparison is showed in figure 3.11 and it is clearly evident that wt STAS domain and mutants have similar amount of mixed α/β secondary structure.

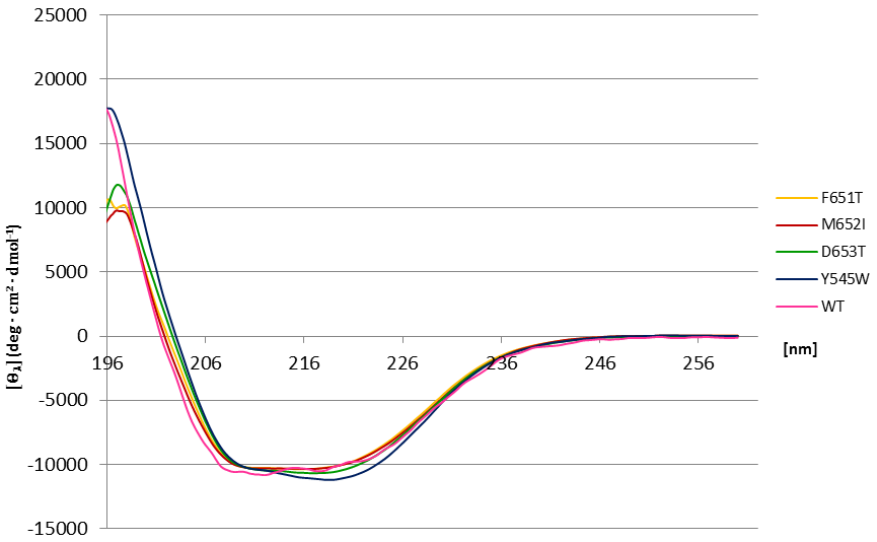


Figure 3.11 *Far-UV circular dichroism spectra of SLC26A5 STAS domain [505-563]GS[637-718] mutants [F651T (yellow), M652I (red), D653T (green), Y545W (blue)] and wt (pink). Proteins concentration was 1 mg/ml in 50 mM Na_2HPO_4 and 150 mM NaCl (pH 7.5).*

3.2.3.3. Differential scanning fluorimetry

To evaluate the effects of the mutations on the thermal stability, the temperatures of melting of the STAS domain mutants and that of the wt were measured by differential scanning fluorimetry in the same conditions (Figure 3.12). F651 mutant is more stable than wt, while M652I and D653T are less stable. Y545W mutant's Tm resembles that of wt.

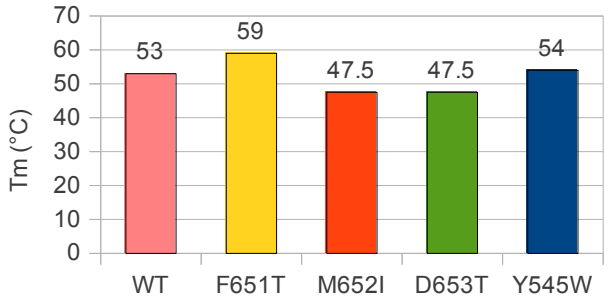
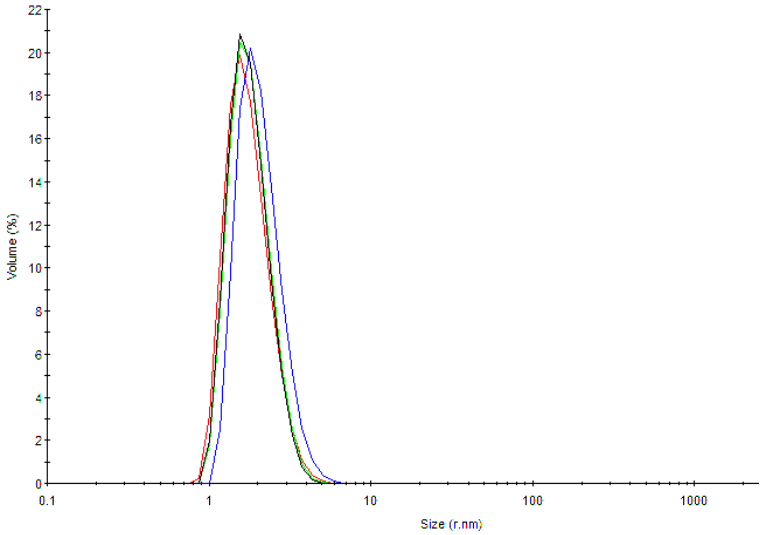


Figure 3.12 *Results of differential scanning fluorimetry of STAS domain mutants [F651T (yellow), M652I (red), D653T (green), Y545W (blue)] and wt (pink). The temperature of melting (Tm) values are the mean of three measurements. Protein concentration was 0.5 mg/ml in 20 mM Tris-HCl (pH 7.5).*

3.2.3.4. Dynamic light scattering

In order to analyze the monodispersity of the STAS domain mutants, they were characterized by Dynamic Light Scattering measurements and the results are shown in figure 3.13. Since STAS domain mutants have a similar low polydispersion and there are no high molecular weight aggregates, we can conclude that they are in a monomeric form and that they are suitable for crystallization trials.



		Pd (%)
	WT	32.2
	F651T	33.1
	M652I	24.1
	D653T	48.4
	Y545W	26.7

Figure 3.13 *Particle size distribution by volume and summarizing table of polydispersion (Pd) percentage of SLC26A5 STAS domains mutants: F651T (red), M652I (green), D653T (blue), Y545W (black). The proteins concentration was 5 mg/ml in 20 mM Tris-HCl, 150 mM NaCl and 5 mM DTT (pH 7.5).*

3.2.3.5. Crystallization trials

The first crystallization test was to try the same conditions wherein wt STAS domain crystallizes: The $AmSO_4$ Suite n° 87 (1 M MES pH 6.5, 2 M Ammonium Sulfate, 5% (w/v) PEG 400; Qiagen) adding 0.1% (w/v) OG (Sigma). In this condition, high-quality hexagonal crystals, resembled those of wt, were grown for two mutants, D653T and Y545W (Figure 3.14). On the other hand, F651T and M652I didn't crystallize in this condition and in any other condition, even if several crystallization screenings were tested: The JCSG+ Suite, The $AmSO_4$ Suite, The MPD Suite (Qiagen), Structure Screen 1 and 2 and Morpheus™ (Molecular Dimensions).

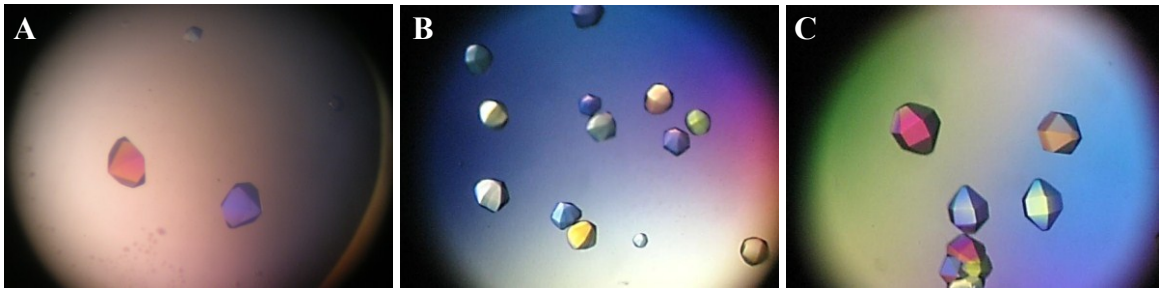


Figure 3.14 *Crystals of STAS domain wt* ***(A)****, D653T* ***(B)*** *and Y545W* ***(C)*** *mutants in the same crystallization conditions.*

Despite we expected that the 3D structure of STAS domain mutants could resemble that of wt, two of them (F651T and M652I) didn't crystallize in the same condition of wt and in any other condition.

Looking at the results of the structural characterization of the STAS domain mutants, by means of CD, DLS and DSF, we can affirm that there are no many differences between mutants and wt. To understand the reason of the failure of the crystallization in the same condition of wt, we looked at the position of these mutations in the crystal lattice (Figure 3.15).

Regarding F651T mutation, Phenylalanine is in the contact's area between a molecule and its symmetric one, according to the symmetry of the crystal space group. Modifying this residue, their interactions change and probably these changes make the packaging between molecules harder and hinder the crystallization. As regards M652I mutation, Methionine isn't exposed directly toward the crystallographic contact but likely by changing it into Isoleucine, with a larger steric hindrance, this produces a movement of the closer amino acids, among which some exposed toward the crystallographic contact, making the crystal grown unfavorable.

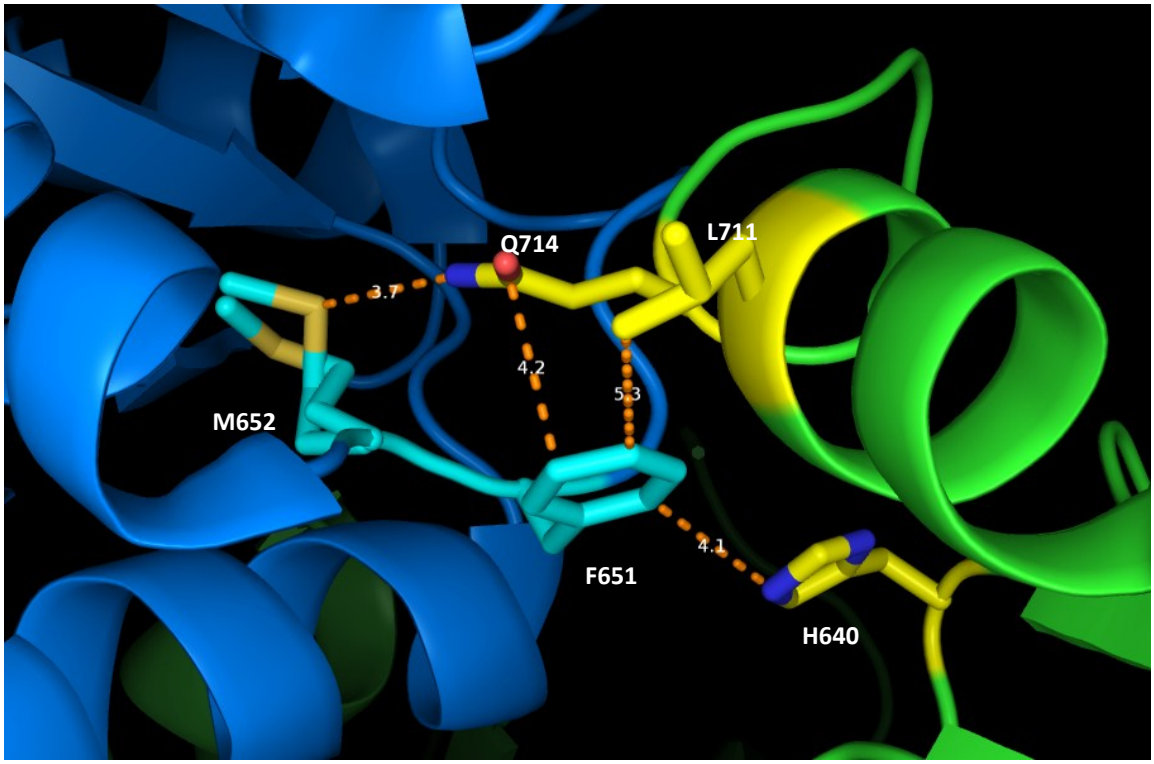


Figure 3.15 *Crystallographic contacts between two STAS domains wt (blue and green) according to the symmetry of the crystal space group $P3_121$. F651 and M652 residues of a molecule (blue) interact with H640, L711 and Q714 residues of its symmetric one (green).*

3.2.3.6. Structure determination

The structures were solved both by molecular replacement starting from the coordinates of the wt STAS domain taken from PDB ID 3LLO, and by Single-wavelength Anomalous Dispersion (SAD) using the anomalous signal of Bromine (when present in the crystal). Statistics on data collections and refinement are reported in table 3.5. As wt, mutant proteins crystallized in the space group 3_121 and some amino acids preceding the position of the variable loop were not visible in the electron density because disordered and were not included in the final model. The final crystallographic R-factors are listed in the two last rows of the table 3.5.

3.2.3.7. D653T mutant

As expected, the overall structure of D653T mutant resembles that of wt STAS domain (Figure 3.16.a) and also in the anion-binding site there are no significant differences in the position of the side chains of the residues around the mutation (Figure 3.16.b).

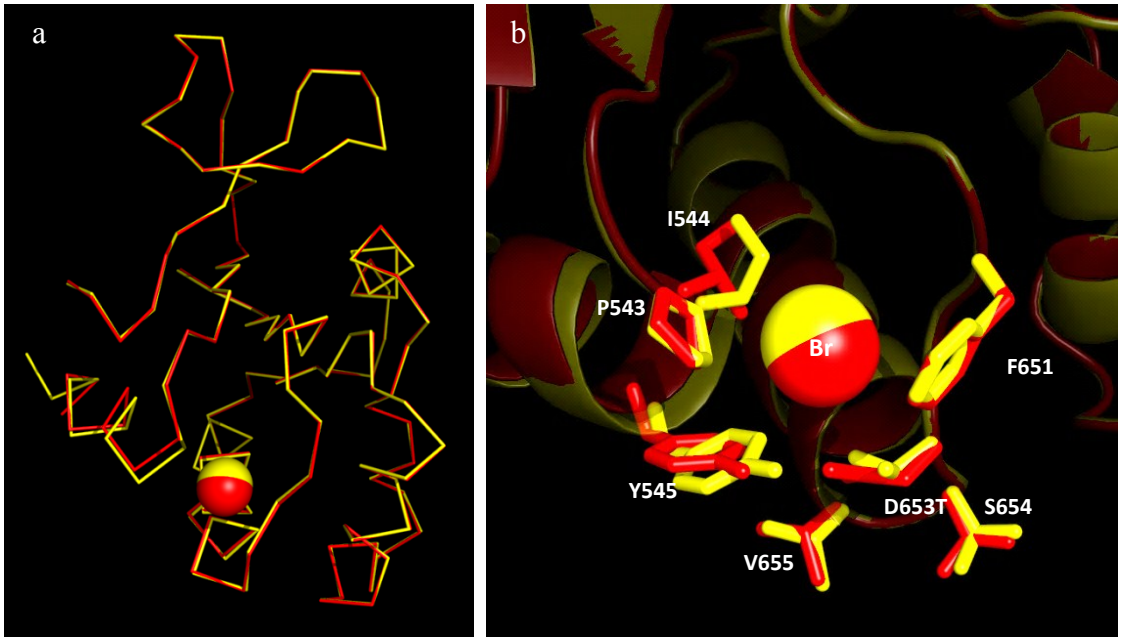


Figure 3.16 a) *3D structure superposition of wt STAS domain (red) and D653T mutant (yellow).* ***b)*** *Detail of the anion-binding site and residues involved.*

D653T mutation of SLC26A5 STAS domain causes loss of OHCs electromotility in vitro. The electromotility is due by voltage-dependent conformational changes of prestin after two preliminary steps: (1) anion binding in the cytosolic portion and (2) the translocation of this voltage sensor within the transmembrane domain (Figure 3.17). The loss of function can be caused by either the loss anion binding or the mutation doesn't influence the anion binding but the anion translocation within the transmembrane domain.

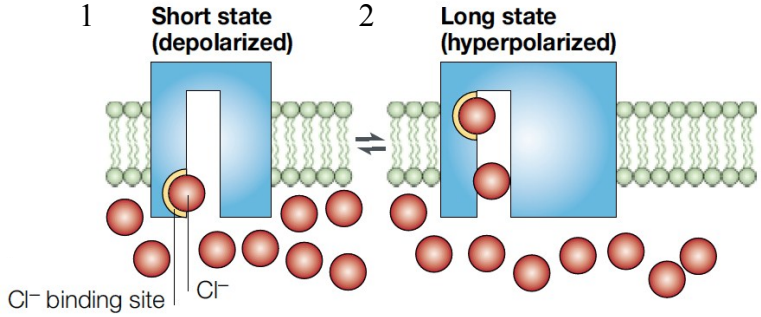


Figure 3.17 *Model of SLC26A5/prestin's molecular mode of action (Dallos and Fakler 2002).*

Since the Br^- is still present in the anion-binding site whose characteristics are very similar to those of the wild-type STAS, the mutation doesn't seem to significantly affect the anion binding, even if we cannot completely exclude a different affinity; so we can suppose that this mutation affects the anions translocation within the transmembrane domain. The loss of the negative charge of the Aspartic acid 653 residue could prevent

the correct translocation of the anion within the transmembrane domain and could increase its permanence in the binding site. This hypothesis is supported by the predicted orientation of the STAS domain with respect to the membrane. In this model (Figure 3.18) the residues involved in the binding are close to the lipid bilayer and to the transmembrane domain.

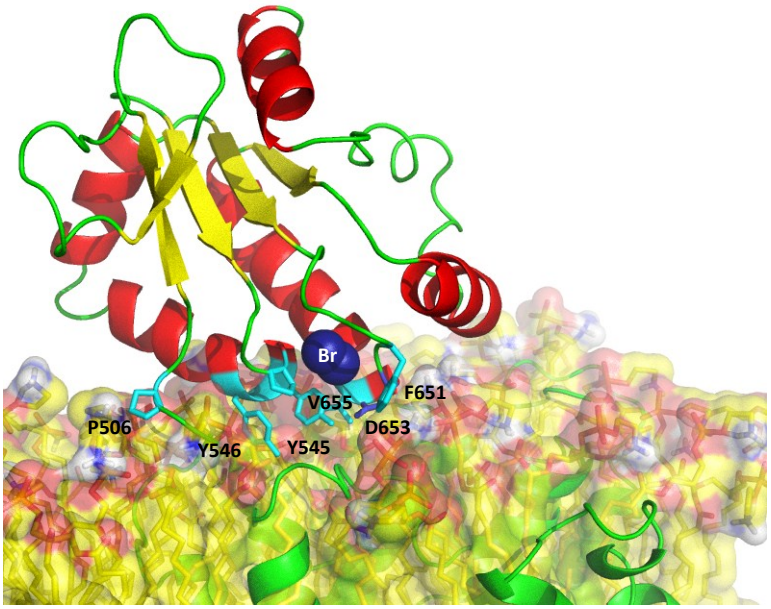


Figure 3.18 *Proposed orientation of SLC26A5 STAS domain with respect to the membrane. The side chains of the amino acids that have been predicted interacting the lipid bilayer and the transmembrane domain are shown (light blue).*

3.2.3.8. Y545W mutant

As expected, there are no significant differences between the overall structure of Y545W mutant and that of the wt STAS domain (Figure 3.19.a). Looking at the anion-binding site, while most of the residues doesn't change, the Tryptophan 545 is flexible and its position seems to be different in comparison with that of Tyrosine in the wt STAS domain (Figure 3.19.b). The structure around the mutation is not well defined neither when it has been solved by MR nor by SAD.

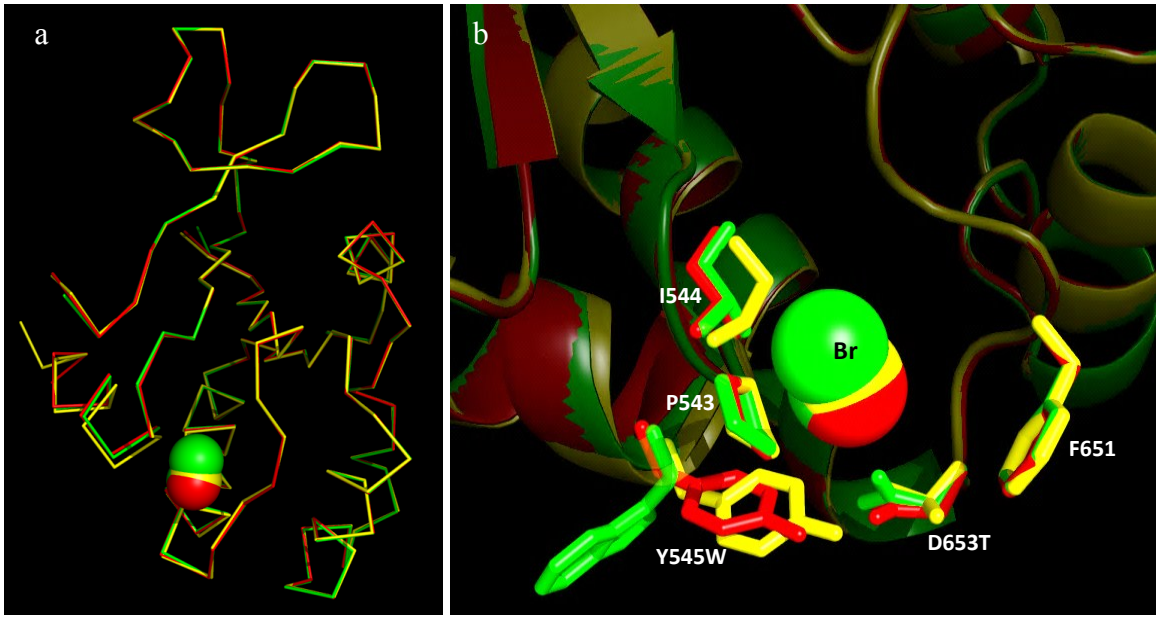


Figure 3.19 a) *3D structure superposition of wt STAS domain (red), D653T mutant (yellow) and Y545W mutant (green).* ***b)*** *Detail of the anion-binding site and residues involved.*

In order to understand if the flexibility of Trp is due to a larger steric hindrance than Tyr or to the interaction with Br^-, we prepared crystals of Y545W without KBr. In absence of Br^-, the electronic density around the mutation is not as well defined as in the wild-type STAS; therefore the disordering effect around the binding site is caused by the mutation itself rather than by the anion binding.

It could be interesting to evaluate the effect of this mutation in the full-length SLC26A5.

3.3. Conclusions

With the final purpose to obtain structural and functional information on the SulP/SLC26 STAS domains, STAS domains of different transporters, from distance-related species, were selected, produced in *E. coli* and characterized: SLC26A5/prestin from *Rattus norvegicus*, *Gallus gallus* and *Danio rerio*; SLC26A4 from *Homo sapiens* and Sultr 1;2 from *Arabidopsis thaliana*. STAS domains of SulP/SLC26 transporters from animals were designed with and without variable loop, while the STAS domain from plant was designed with loop. Four mutants of the STAS domain whose 3D structure has been solved were produced in order to study the anion-binding site and the possible role of the STAS domain in the transport. Among these, two variants of prestin STAS domain from *R. norvegicus* and *G. gallus* and all four STAS domain mutants were obtained in a soluble form. For each of these constructs, an optimized three-step purification protocol allowed obtaining a yield (2-10 mg of protein per liter of culture) sufficient for structural studies with high purity level. The proteins were characterized in solution by size exclusion chromatography and by classical biophysical methods, such as circular dichroism spectroscopy, dynamic light scattering and differential scanning fluorimetry. The purified constructs were submitted to extensive crystallization trials and two of them (D653T and Y545W mutants) were successfully crystallized and it was possible to determine their crystal structures by means of the SAD and MR technique. These structures can help to understand the important role of the STAS domain in the transport of full-length prestin. The structure-function analysis indicates that the Asp653 residue is fundamental for NLC and, consequently, for OHCs electromotility probably because its negative charge repulses the negative charge of the anion in the binding site and permits the anion translocation within the transmembrane domain. Instead, the loss of Asp653's negative charge could increase the permanence of the anion in the binding site, preventing the prestin's function.

To conclude, some STAS domains had a strong tendency to aggregate, preventing the purification in sufficient amount for structural studies, despite the presence of a SUMO-tag that may increase the solubility of recombinant fusion proteins. Also Rv1739c STAS domain precipitated as insoluble aggregates and Sharma and co-workers purified it from inclusion bodies following an on-column refolding procedure (Sharma et al., 2009). Furthermore, as regards the most of the purified STAS domains, it has not been possible to obtain crystals suitable for X-ray diffraction analysis despite they were good candidates

for crystallographic studies (well-folded and structured single domain, not aggregated, and in conditions that assure a good thermal stability). Moreover, Babu and colleagues were unable to crystallize the STAS domain of YchM on its own, but only in complex with a protein (Babu et al., 2010). And we crystallized rat prestin STAS domain (mutants and wt) only in the presence of OG, visible in the 3D structure (Pasqualetto et al., 2010). Altogether these data indicate that isolated STAS domains excised from full transporters have generally a poor tendency to give origin to well diffracting crystals.

4. Full-length SulP/SLC26 transporters

4.1. Introduction

4.1.1. Membrane proteins

Integral membrane proteins (MPs) play crucial roles in many aspects of biology by mediating the transfer of material and signals between cells and their environment. They comprise more than a quarter of all sequenced genomes, and a majority of the targets of currently marketed drugs. A detailed knowledge of their structure, function and dysfunction is essential to a wide range of biomedical and biotechnological applications but, yet (at January 2013), only around 380 unique polytopic structures have been elucidated (http://blanco.biomol.uci.edu/). These low MPs number is a reflection of the technical difficulties of working with membrane proteins (Sonoda et al., 2011).

Efficient production of high quality samples in adequate quantities (in the order of milligrams) is a major bottleneck in structural approaches of MPs because of their hydrophobic nature, toxic effects and specific requirements for targeting and translocation systems. Particularly in *Escherichia coli*, the most commonly used heterologous expression host, overexpression of MPs can result into the accumulation of aggregated material.

4.1.2. Cell-free expression

Cell-free (CF) expression system has recently emerged as a promising and highly versatile technique for the general production of membrane proteins. A number of intrinsic key characteristics set CF expression clearly apart from any other traditional MP expression system. (1) Reaction volumes are small and ranging from µls to mls in preparative scales. Production efficiencies are very high, resulting in yields of mg protein per ml reaction. (2) The operator always has unlimited access to the reaction and stabilizing compounds, such as detergents or lipids, can be supplemented at any time point. This feature makes CF expression the most versatile expression system available. (3) Most intrinsic general problems of conventional cellular MP preparation, such as toxicity or efficient targeting, translocation and membrane extraction, are virtually

eliminated and the speediness of CF expression is highly competitive. (4) The user can switch between different expression backgrounds by using cell extracts from either prokaryotic or eukaryotic sources (Junge et al., 2010; Schwarz et al., 2007; Schneider et al., 2010).

Two different CF reaction configurations are commonly used: the one compartment batch reaction and the two compartments continuous-exchange (CECF) configuration (Figure 4.1.A). In the second configuration two compartments are separated by a semipermeable membrane of molecular weight cut-offs of 12–14 kDa. One compartment holds the reaction mixture (RM) containing the cell extract and all other high molecular weight compounds like enzymes and nucleic acid templates. Supplies of low-molecular weight precursors such as NTPs and energy sources are present in a feeding mixture (FM) in the second compartment. Vigorous shaking or stirring ensures an efficient exchange of substances between the two compartments. CF expression offers the opportunity to produce MPs in three new ways: in the precipitate forming (P-CF) mode, the detergent based (D-CF) mode or in the lipid based (L-CF) mode (Figure 4.1.B).

In the P-CF mode, MP precipitation occurs immediately after translation as CF extracts are almost completely devoid of membranes. These precipitates can usually be solubilized efficiently by the addition of particular detergents within a few hours, without applying classical refolding procedures. In the D- CF mode, the MPs are inserted into micelles provided during or shortly after translation. After the reaction, the RM containing the MP proteomicelles can be applied instantly to the first purification column without any further treatment. In the L-CF mode, defined lipids or lipid mixtures are supplied into the RM and translated MPs may directly insert into the bilayers (Figure 4.1.C) (Junge et al., 2010).

CF expression and the first steps of production and functional characterization of the SulP/SLC26 transporters were performed in the laboratory of Prof. Dötsch, under supervision of Dr. Frank Bernhard at the Institute of Biophysical Chemistry, Centre of Biomolecular Magnetic Resonance, Johann Wolfgang Goethe University, Frankfurt/Main (Germany), where I spent six months of my PhD. Then I've continued CF expression, purification and structural characterization at the University of Padova.

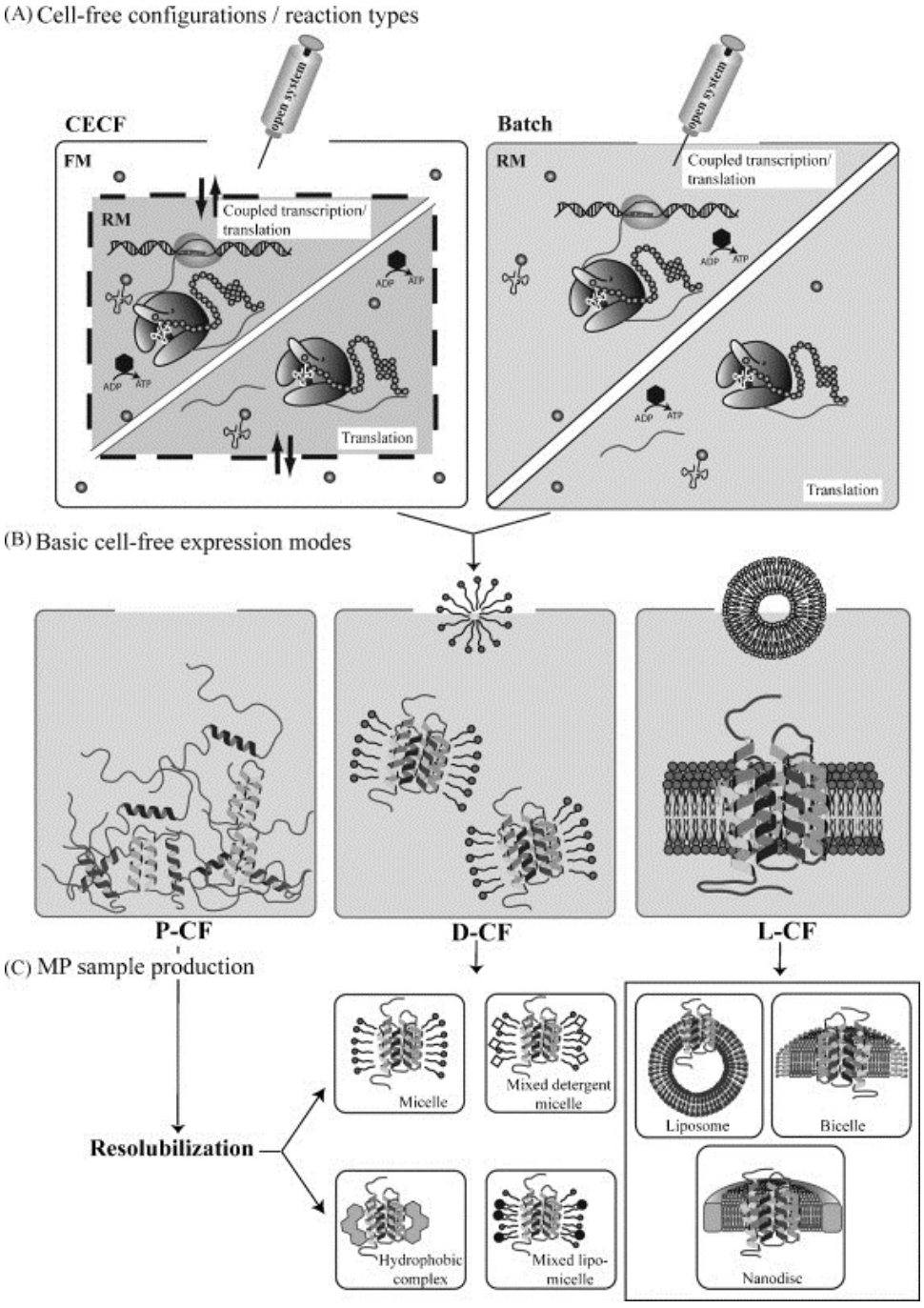


Figure 4.1 *Versatility of the CF expression technique.* ***(A)*** *CF reaction configurations (CECF or batch).* ***(B)*** *Basic expression modes for the CF synthesis of MPs; the protein is either expressed without any hydrophobic agents (P-CF mode) or in the presence of detergents (D-CF mode) or lipids (L-CF mode).* ***(C)*** *Variety of primary MP samples resulting from the basic expression modes (Junge et al., 2010).*

4.2. Methods

4.2.1. Selection of the SulP/SLC26 proteins for CF expression

The selected SulP/SLC26 proteins for CF expression are: SLC26A5/prestin from *Rattus norvegicus*, *Gallus gallus* and *Danio rerio*; Sultr 1;2 from *Arabidopsis thaliana*; Rv1739c from *Mycobacterium tuberculosis* and BicA from *Synechocystis sp. PCC 6803*, *Trichodesmium erythraeum* and *Thermosynechococcus elongatus BP-1*. For all the proteins only the full-length sequence were tested (Table 4.1).

Table 4.1 Selection of SulP/SLC26 anion transporters for CF expression

Protein	**Source organism**	**Coding sequence (bp)**	**aa (residues)**	**MW (kDa)**	**N-terminal tag**	**C-terminal tag**
SLC26A5	*Rattus norvegicus*	2232	744	81.28	t7	poly(His)$_{10}$
SLC26A5	*Gallus gallus*	2226	742	81.11	t7	poly(His)$_{10}$
SLC26A5	*Danio rerio*	2217	739	81.40	t7	poly(His)$_{10}$
Sultr1;2	*Arabidopsis thaliana*	1959	653	71.71	t7	poly(His)$_{10}$
Rv1739c	*Mycobacterium tuberculosis*	1680	560	59.36	t7	poly(His)$_{10}$
BicA	*Synechocystis sp. PCC 6803*	1692	564	59.61	optimized	poly(His)$_{10}$
BicA	*Trichodesmium erythraeum*	1722	574	60.43	optimized	poly(His)$_{10}$
BicA	*Thermosynechococcus elongatus BP-1*	1683	561	59.92	optimized	poly(His)$_{10}$

Optimized N-terminal tag =AAATATTATAAATATTAT

4.2.2. Plasmid design and cloning of the SulP/SLC26 genes

The cloning of BicA genes from different cyanobacteria was performed in collaboration with Dr. Elisabetta Bergantino, at the Department of Biology A. Vallisneri, University of Padova.

The coding regions for BicA proteins were amplified by standard polymerase chain reaction (PCR) using Taq polymerase (Genespin) starting from genomic DNA for *Trichodesmium erythraeum* and *Thermosynechococcus elongatus BP-1* and from supernatant of bacterial lysate for *Synechocystis sp. PCC 6803*. To be more precise, firstly, a region larger than the coding region was amplified using highly specific

oligonucleotide primers, that don't had to have affinity with other genomic regions (external primers, Table 4.2); secondly, the coding region was amplified using primers containing restriction sites necessary for the following cloning into the vector (cloning primers, Table 4.2). To facilitate the digestion by restriction enzymes (NdeI and XhoI), the purified and phosphorylated PCR fragments were inserted into a linearized blunt-end vector (pBSK), subsequently digested by restriction enzymes. The digested and purified fragments were then inserted into the corresponding cloning site of the expression vector pET-21cHx, digested with the same enzymes. The plasmid is a derivative of the standard T7 promoter based vectors pET-21a(+) (Merck Biosciences) encoding for an extended C-terminal poly(His)$_{10}$ tag. The extended poly(His)$_{10}$ tag at the C-terminus not only helps for efficient binding during purification but also to verify the full length expression. The low yield can be countered with an efficiently translated expression tag, such as T7 tag or an optimized tag at the N-terminal end. Haberstock and colleagues (Haberstock et al., 2012) demonstrated the importance of high AT content and presence of a triple A sequence as second codon for the improvement of CF production efficiencies of membrane proteins. According to these findings, an optimized N-terminal tag (AAATATTATAAATATTAT) was added to the DNA fragments by PCR with suitable oligonucleotide primers (cloning primers, Table 4.2).

Table 4.2 Oligonucleotide primers used for cyanobacteria SulP/SCL26 transporters, BicA

External primers		
Constructs	**Primer**	**Sequence**
Synechocystis sp. PCC 6803	5' 3'	gccaagctcgacagcagagcaccg agctgccagtcggccagcgg
Trichodesmium erythraeum	5' 3'	ggctgccgtgacaagccactgt acgcatcaaggctttgcaacagaca
Thermosynechococcus elongatus BP-1	5' 3'	cctcagggcgaggggttccccc ctcagcccgcccttgccgcc
Cloning primers		
Constructs	**Primer**	**Sequence**
Synechocystis sp. PCC 6803	5' 3'	aaacatatgAAATATTATAAATATTATcaaataactaacaaaattcattttag cttccgtccagaccacatacctccgaaaa
Trichodesmium erythraeum	5' 3'	aaacatatgAAATATTATAAATATTATgctacgcaagttttcaataaaatac caatctgggattgataatattatcaaactccgaaaa
Thermosynechococcus elongatus BP-1	5' 3'	aaacatatgAAATATTATAAATATTATttaagcaatttagttaatcgagttc gagctgacaacgggtagcctccgaaaa

NdeI and XhoI restriction sites are underlined; optimized N-terminal tag is in upper-case letter.

The others SulP/SLC26 transporters, SLC26A5/prestin from *Rattus norvegicus*, *Gallus gallus* and *Danio rerio*, Sultr 1;2 from *Arabidopsis thaliana* and Rv1739c from

Mycobacterium tuberculosis, were already been cloned into a pET-21cHx vector with an N-terminal T7 tag and a C-terminal poly(His)$_{10}$ tag.

The resulting plasmids were isolated with commercial kits (Qiagen) and used as templates for the CF expression. A high quality of template DNA at a concentration of ≥ 0.15 µg/µl is of extreme importance.

4.2.3. Western blot analysis

For western blot analysis, the SDS-PAGE gels were transferred on a 0.45 µm Immobilion-P poly(vinylidene difluoride) membrane (Millipore) in a Hoefer TE22 (GE Healthcare) wet western blot apparatus for 35 min at 340 mA. The membrane was then blocked for 1 h at RT in blocking buffer containing PBS (8 mM Na_2HPO_4, 15 mM KH_2PO_4, 0.137 mM NaCl, 3 mM KCl, pH 7.4), 4% (w/v) skim milk powder and 0,05% (w/v) Tween. The first antibody, the anti His-tag antibody from mouse (Qiagen) diluited 1:2000, was incubated ON at 4 °C with the membrane. The second antibody, the anti-mouse antibody (Sigma Aldrich) diluted 1:5000 was incubated for 1 h at RT with the membrane. After extensive washing with PBS, 0.05% (w/v) Tween, the blots were analyzed by chemiluminescence in a Lumi-imager F1 (Roche Diagnostics).

4.2.4. Preparation of CF lysates

The cell extract preparation is one of the key steps in obtaining successful and reproducible results using the CF protein synthesis. Bacterial CF extracts were prepared from the *E. coli* strain A 19 according to the following the protocol optimized in the laboratory of Prof. Dötsch (Schneider et al., 2010; Schwarz et al., 2007).

The cells were grown with good aeration until mid-log phase (OD_{600} of approximately 4-4.5) at 37 °C in 2x YTPG medium (per liter: 2.99 g KH_2PO_4, 6.97 g K_2HPO_4, 19.82 g glucose, 16 g tryptone, 10 g yeast extract, 5 g NaCl), chilled down rapidly and harvested by centrifugation. The cell pellet was resuspended and washed three times in ice cold S30-A buffer [10 mM Tris-acetate, 14 mM $Mg(OAc)_2$, 0.6 mM KCl, 6 mM β-mercaptoethanol (pH 8.2)] and it was finally suspended in S30-B buffer [10 mM Tris-acetate, 14 mM $Mg(OAc)_2$, 0,6 mM KCl, 1 mM DTT, 0.1 mM PMSF (pH 8.2)] pre-cooled at 4 °C. The cells were disrupted by passing through a pre-cooled French-Press. Cell-debris was removed by centrifugation at 30000 g at 4 °C for 30' and the

supernatant was transferred into a fresh vial. The centrifugation step and transfer of supernatant was repeated once. The supernatant was adjusted to a final concentration of 400 mM NaCl follow by incubation at 42 °C for 45' in a water bath. This step is used to get rid of endogenous mRNA; the solution will become turbid after incubation. The turbid solution was filled into a dialysis tube (MWCO 14 kDa) and dialyzed at 4 °C against 100 volumes of S30-C buffer [10 mM Tris-acetate, 14 mM $Mg(OAc)_2$, 0.6 mM KOAc, 0.5 mM DTT (pH 8.2)] with gentle stirring. After one further exchange of the dialysis buffer ON at 4 °C the *E. coli* S30-extract was harvested by centrifugation at 30000 g at 4 °C for 30'. The clear supernatant was transferred in suitable aliquots and frozen in liquid nitrogen.

4.2.5. CF expression: precipitate forming (P-CF) mode

CF reactions are performed in analytical scales for optimization studies and in preparative scales for high-level production. The volume ratio of reaction mixture (RM): feeding mixture (FM) was generally kept at 1:14 for analytical scale reactions and 1:17 for preparative scale reactions. For 55 µl analytical scale reactions, self-made Mini-CECF reactors were used with regenerated cellulose membranes with 14 kDa cut-off. The Mini-CECF reactors were incubated in standard 24-well microplates and the cavities of the microplates were used as a FM compartment. Preparative scale reactions of 1 ml were performed with self-made Maxi-CECF reactors and with commercial Slide-A-Lyzer units (Pierce) as a RM container. Optimal concentration of Mg^{2+} for every S30 extract batch was analyzed in order to obtain higher yields. The reaction protocol is given in the table 4.3. CF reactions were performed in the continuous exchange cell-free (CECF) configuration and incubated for 16 h at 30 °C with gentle shaking. CF expressed protein was harvested as precipitate (P-CF) by centrifugation; the synthesis of the protein was verified by immunodetection with antibodies directed against the $poly(His)_{10}$ tag.

Table 4.3 Cell-free expression protocol

	Final concentration in RM	**Final concentration in FM**
S30 extract	35%	-
S30-C Buffer	-	35%
Plasmid	15-20 μg/ml	-
Rnasin	0.3 U/μl	-
T7-RNA Polymerase	0.05 U/μl	-
tRNA *E. coli*	0.5 mg/ml	-
Pyruvate kinase	0.04 mg/ml	-
Amino acids (RCWMDE)	1 mM	1 mM
Amino acids mix	0.55 mM	0.55 mM
Acetyl phosphate	20 mM	20 mM
Phosphoenol pyruvate	20 mM	20 mM
ATP	1.2 mM	1.2 mM
CTP, GTP, UTP	0.8 mM	0.8 mM
DTT	2 mM	2 mM
Folinic acid	0.1 mg/ml	0.1 mg/ml
Complete protease inhibitor	1x	1x
Hepes-KOH pH 8 EDTA	100 mM	100 mM
$Mg(OAc)_2$	17 mM	17 mM
KOAc	270 mM	270 mM
PEG 8000	2%	2%
NaN_3	0.05%	0.05%

Abbreviations: RM= reaction mixture; FM= feeding mixture.

4.2.6. Detergent solubilization of precipitated proteins

After expression in the absence of detergent, the precipitated protein was harvested from the RM by centrifugation. The pellet was suspended in one volume of washing buffer [20 mM Tris-HCl, 150 mM NaCl (pH 7.5)] and centrifuged. The washing step was repeated twice. The precipitate was then re-suspended in resuspension buffer [20 mM Tris-HCl, 150 mM NaCl (pH 7.5) supplemented with appropriate detergent]. SulP/SLC26 solubilization was achieved by incubation for 1 h at 30 °C with gentle shaking; residual precipitate was removed by centrifugation. The protein in both fractions was quantified separately by western blot analysis. Routinely analyzed detergents for resolubilization are: 1% (w/v) FOS-12, 1% (w/v) LMPC, 1% (w/v) LPPG, 1% (w/v) LMPG, 1% (w/v) DDM, 1% (w/v), 1% (w/v) DHPC, 1% (w/v) FOS-16, 1% (w/v) OG, 5% (w/v) Nvoy, 1% (w/v) A8-35 (Table 4.4). A panel of non-ionic detergents (DDM, OG), zwitterionic (LMPC, FOS-12, FOS-16, DHPC) and negatively charged ones (LMPG, LPPG) was chosen. Many of them display typical phospholipid polar heads (LMPC, LMPG, LPPG, DHPC, FOS-12).

Table 4.4 Detergents for resolubilization of P-CF products

Detergent	Short name	Detergent class	CMC (H_2O)
n-dodecylphosphocholine	FOS-12/ DPC	zwitterionic	~ 1.5 mM (0.047%)
1-myristoyl-2-hydroxy-sn-glycero-3-phosphocholine	LMPC	zwitterionic	0.043 - 0.090 mM
1-palmitoyl-2-hydroxy-sn-glycero-3-[phospho-RAC-(1-glycerol)]	LPPG	ionic	0.6 mM
lyso-myristoylphosphatidylglycerol	LMPG	ionic	0,05 mM
n-dodecyl β-D-maltoside	DDM	non ionic	~ 0.17 mM (0.0087%)
l,2-diheptanoyl-sn-glycero-3-phosphocholine	DHPC	zwitterionic	1.4 mM
n-hexadecylphosphocholine	FOS-16	zwitterionic	~ 0.013 mM (0.00053%)
n-Octyl-β-D-glucopyranoside	OG	non ionic	~ 18-20 mM (0.53%)
Amphipol A8-35	A8-35		NA

CMC= critical micelle concentration, the concentration of detergent above which monomers self-assemble into non-covalent aggregates (called micelles). NA= not available.

4.2.7. Protein purification

Immobilized Metal Affinity Chromatography (IMAC) was applied for purification of the poly$(His)_{10}$-tagged transporters and for the optional exchange of the first detergents used for solubilization against second detergents used for analysis.

Resolubilized protein produced in the P-CF mode was centrifuged in order to remove precipitates. The supernatant containing the detergent-solubilized membrane protein was diluted in equilibration buffer [20 mM Tris–HCl, 150 mM NaCl, 10 mM imidazole, detergent at >2x CMC (pH 7.5)] and mixed with pre-equilibrated NTA-agarose beads loaded with Ni^{2+} ions. The mixture was incubated for 1 h at RT followed by the application to an empty gravity flow column and washed with 2 washing buffers: washing buffer 1 [20 mM Tris–HCl, 150 mM NaCl, 50 mM imidazole, detergent at >2x CMC (pH 7.5)] and washing buffer 2 [20 mM Tris–HCl, 150 mM NaCl, 80 mM imidazole, detergent at >2x CMC (pH 7.5)]. Immobilized protein was finally eluted with 300 mM imidazole in elution buffer [20 mM Tris–HCl, 150 mM NaCl, 300 mM imidazole, detergent at >2x CMC (pH 7.5)].

Size exclusion chromatography (SEC) was used in order to assess the homogeneity and the stability of SulP/SLC26 samples produced, solubilized and stored under different detergent conditions. Samples of solubilized transporter were separated on a pre-

equilibrated analytical Superdex 200 5/150 GL column (GE Healthcare) at a flow rate of 0.3 ml/min (bed volume = 3 ml). Routinely detergents for crystallization screenings are: 0.1% (w/v) FOS-12, 0.05% (w/v) DDM, 0.06% (w/v) LDAO, 0.1% (w/v) UM, 0.2% (w/v) DM, 0.02% (w/v) DM NG, 0.4% (w/v) NG, 1% (w/v) OG (Table 4.5).

Table 4.5 Detergents for crystallization screenings

Detergent	**Short name**	**Detergent class**	**CMC (H_2O)**	**Micelle size**
n-dodecylphosphocholine	FOS-12	zwitterionic	~ 1.5 mM (0.047%)	~ 19 kDa
n-dodecyl beta-D-maltoside	DDM	non ionic	~ 0.17 mM (0.0087%)	~ 40-76 kDa
n-Dodecyl-N,N-Dimethylamine-N-Oxide	LDAO	zwitterionic	~ 1-2 mM (0.023%)	17 -21.5 kDa
n-Undecyl-β-D-Maltopyranoside	UM	non ionic	~ 0.59 mM (0.029%)	~ 35 kDa
n-Decyl-β-D-Maltopyranoside	DM	non ionic	~ 1.8 mM (0.087%)	~ 33 kDa
Decyl Maltose Neopentyl Glycol	DM NG	non ionic	~ 0.036 (0.0034%)	NA
n-Nonyl-β-D-Glucopyranoside	NG	non ionic	~ 6.5 mM (0.20%)	~ 41 kDa
n-Octyl-β-D-glucopyranoside	OG	non ionic	~ 18-20 mM (0.53%)	~ 8-29 kDa

4.2.8. CD spectroscopy

CD data were recorded on a Jasco J-715 spectropolarimeter, using quartz cuvettes of 0.02 cm path-length. The spectra were determined as an average of 10 scans. The protein concentration was 1 mg/ml in 50 mM Na_2HPO_4, 150 mM NaCl, pH 7.5. The data were recorded and analyzed with Spectra Manager Software (JASCO).

4.2.9. Protein crystallization

Crystallization trials using commercial kits for membrane proteins (The MbClass Suite and The MbClass II Suite, Qiagen) based on sparse matrix were performed by vapor diffusion (with the sitting drop method) techniques, using the Oryx8 automatic system (Douglas Instrument).

4.2.10. CF expression in presence of lipids

Lipids can be supplied into CF reactions in different formulations such as preformed liposomes of defined compositions, isolated fractions of cell membranes, detergent solubilized lipomicelles providing a mixed environment for MPs, bicelles consisting of planar bilayers surrounded by detergents, and nanolipid particles or nanodiscs (NDs) providing highly soluble and confined bilayer areas. The ND is a non-covalent assembly of phospholipid and a genetically engineered membrane scaffold protein (MSP) which itself is based upon the sequence of human serum apolipoprotein AI. The phospholipid associates as a bilayer domain while two molecules of MSP wrap around the edges of the discoidal structure in a belt-like configuration, one MSP covering the hydrophobic alkyl chains of each leaflet (Figure 4.2) (Bayburt and Sligar, 2009).

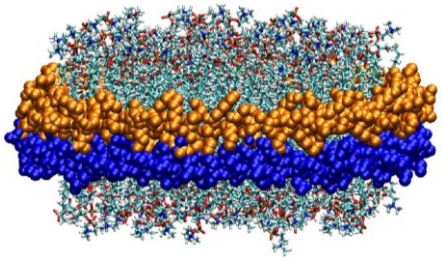


Figure 4.2 *Illustration of the NDs structure shown. The two MSPs are colored orange and blue (Bayburt and Sligar, 2009).*

For functional studies, lipids and NDs were added to the reaction mixture at the final concentrations: phosphocholine (4 mg/ml), *E. coli* polar lipids (4 mg/ml), NDs (30-40 μM).

4.2.11. Stock lipids preparation: phosphocholine and polar lipids

Chloroform solubilized lipids were applied to a rotary evaporator (Rotavapor RE120, Büchi) to remove chloroform. Each lipid was reconstituted in water or buffer at final concentrations of 40 mg/ml, and solubilized by vortexing and incubating at 37 °C in an ultrasonic water bath to form multilamellar vesicles. The multilamellar vesicle solution was passed at least 21 times through an Avanti Polar Lipids mini extruder holding a 0.45 μm Whatman polycarbonate membrane filter (Florham Park, NJ) sandwiched with two filter supports on each side. The resulting unilamellar liposome solution were stored at 4 °C or, for time long storage, at -80 °C and thawed on ice before usage.

4.2.12. Preparation of empty nanodiscs

The membrane scaffold protein MSP1E3D1 has the follow features: the first α-helix is truncated (Δ1-11, indicated by D1) to give more stable discs and an additional three α-helices (α-helix 4, 5, and 6) inserted into the sequence between α-helix 3 and α-helix 4 to make larger discs (~12.1 nm bilayer discs). MSP1E3D1 was expressed with an N-terminal poly$(His)_6$ tag and purified following the protocol of Bayburt and colleagues (Bayburt et al., 2002). Stocks of purified MSP1E3D1s were stored at -80 °C and thawed on ice before usage. Chloroform solubilized lipids were applied to a rotary evaporator (Rotavapor RE120, Büchi) to remove chloroform. Each lipid was resuspended in water at final concentrations of 50 mM and solubilized with Na^+-cholate supported by vortexing or incubating at 37 °C in an ultrasonic water bath. Solubilization was complete when the suspension turned clear. Required Na^+-cholate concentrations for complete solubility were 100 mM for DMPC and 300 mM for DMPG. MSP1E3D1 was combined with the selected lipids at a defined molar stoichiometry (MSP1E3D1:DMPC 1:115; MSP1E3D1:DMPG 1:110). The mixtures of MSP1E3D1, lipid and DPC 0.1% (w/v) were incubated at RT for 30'. NDs formation was initialized by dialysis against 5 l of dialysis buffer [40 mM Tris-HCl, 100 mM NaCl (pH 8.0)]. The first dialysis was performed over weekend at 4 °C, or for 2 h at RT and ON at 4 °C, followed by buffer exchange and further dialysis at RT for additional 3 h. This step was repeated twice.

After dialysis, NDs were centrifuged (22000 g for 20') to remove residual aggregates and stored on ice before concentrating. The supernatant of ND assemblies were applied to Centriprep concentrator devices (10 kDa MWCO, Millipore, Merck) equilibrated with dialysis buffer. Centripreps were centrifuged at 4 °C and 2000 g and NDs were concentrated up to final MSP concentrations of 240 µM. ND stocks were centrifuged (22000 g for 20') and stored on ice before usage. For long time storage, NDs were flash frozen in liquid nitrogen and stored at -80 °C.

4.2.13. Restriction-free (RF) cloning

RF cloning is inspired by oligonucleotide-directed mutagenesis (QuickChange™, Stratagene) and provides a simple, universal method to precisely insert a DNA fragment into any desired location within a circular plasmid, independent of restriction sites, ligation, or alterations in either the vector or the gene of interest. The technique uses a PCR fragment encoding a gene of interest as a pair of primers in a linear amplification

reaction around a circular plasmid. The RF cloning was performed following a published protocol (van den Ent and Loewe, 2006). SLC26A5/prestin from *Danio rerio* and BicA from *Synechocystis sp. PCC 6803* were inserted into a pET-21cHX plasmid coding for GFP, in frame with GFP (Table 4.6). The resulting plasmids were prepared for functional studies.

Table 4.6 Oligonucleotide primers used for RF cloning

Constructs	**Primer**	**Sequence**
BicA from *Synechocystis sp. PCC 6803*	5'	TAACTTTAAGAAGGAGATATACATATGAAATATTATAAATATTATCAAAT
	3'	AGTGAAAAGTTCTTCTCCTTTGCTCTCGAGGTATGTGGTCTGGACGGA
SLC26A5/prestin from *Danio rerio*	5'	TAACTTTAAGAAGGAGATATACATATGGCTAGCATGACTGGTGGACAGCA
	3'	AGTGAAAAGTTCTTCTCCTTTGCTCTCGAGGTGGATGTTTGGGTGGAC

4.3. Results

4.3.1. Full-length SulP/SLC26 transporters

Eight genes encoding selected SulP/SLC26 proteins were successfully cloned into a pET-21cHx plasmid and these proteins were expressed with a C-terminal poly(His)$_{10}$ tag to facilitate subsequent detection and purification.

4.3.2. CF expression screening of SulP/SLC26 transporters

The primary problem of membrane protein production is the generation of sufficient protein yields in order to have enough starting material for subsequent structural analysis. The most productive reaction conditions are best evaluated in the P-CF expression mode.

Purified plasmids DNA were added as template into the RM and the proteins were produced in the CECF configuration in the P-CF mode at 30 °C without added detergents. Critical parameter for CF protein production is optimal ion concentrations of Mg^{2+}. After screening of Mg^{2+} ion concentrations in the P-CF expression mode in a range between 16-22 mM, optima were determined at 17 mM Mg^{2+}. All proteins, but SLC26A5 from *G. gallus*, could be detected by immunodetection of the terminal poly(His)$_{10}$ tag after Western blotting. The expression level of SLC26A5 from *R. norvegicus* was low (≤ 200 µg/ml), therefore the other six candidates were selected for further studies (Figure 4.3). In contrast, the conventional in vivo expression using different *E. coli* strains (BL21(DE3), C41/C43(DE3)) yielded no detectable expression of all SulP/SLC26 transporters by SDS-PAGE and western blot analysis (data not shown).

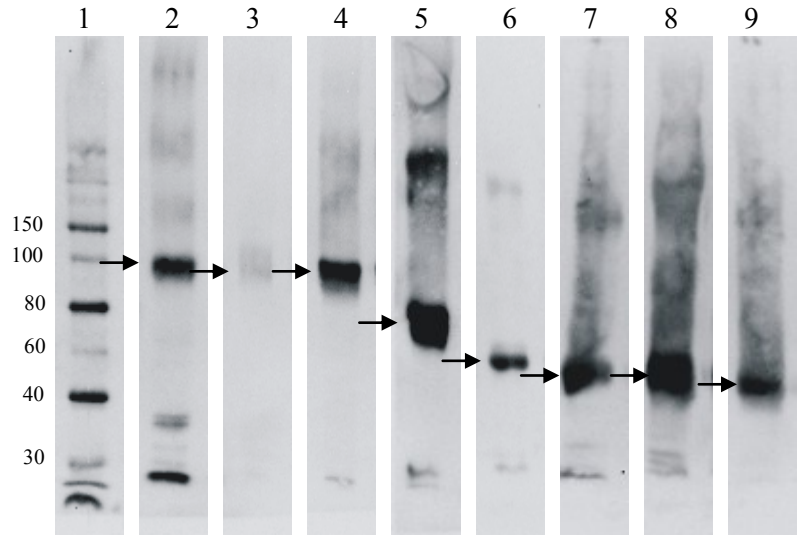


Figure 4.3 *P-CF expression screening of SulP/SLC26 transporters followed by Western blotting. P-CF produced precipitated SulP/SLC26 transporters were washed and suspended in H_2O. Sample volumes of 2 µl were separated by 12% SDS-PAGE and identified by Western blotting with anti-His antibodies. Arrows indicate the synthesized SulP/SLC26 proteins. Lane 1: Marker proteins in kDa; Lane 2: SLC26A5 from R. norvegicus (81 kDa); Lane 3: SLC26A5 from G. gallus (81 kDa); Lane 4: SLC26A5 from D. rerio (81 kDa); Lane 5: Sultr 1.2 from A. thaliana (71 kDa); Lane 6: Rv1739c from M. tuberculosis (59 kDa); Lane 7: BicA from Synechocystis sp. PCC 6803 (60 kDa); Lane 8: BicA from Trichodesmium erythraeum (60 kDa); Lane 9: BicA from Thermosynechococcus elongatus BP-1 (60 kDa).*

4.3.3. Detergent resolubilization screening

After insoluble expression of SulP/SLC26 transporters, selected detergents were tested for their efficiency in the resolubilization of the CF produced precipitates. The western blot analysis was used to quantify the protein in the soluble and insoluble fractions, after incubation with detergents for 1 h at 30 °C. An example of resolubilization screening result is displayed in figures 4.4 and 4.5.

The solubilization of this cyanobacterial SulP/SLC26 protein (BicA from Synechocystis sp. PCC 6803) was apparently 100% efficient in LMPC. Solubilization in the other detergents was 95% in DPC and LMPC, 90% in LPPG and 85% in FOS-16, all the other detergents had no or minor resolubilization effects. The yield in presence of the long-chain phosphocholine FOS-16 seems to be higher than the other tested detergent (row 15, figure 4.4) therefore FOS-16 has been chosen as the best resolubilization detergent.

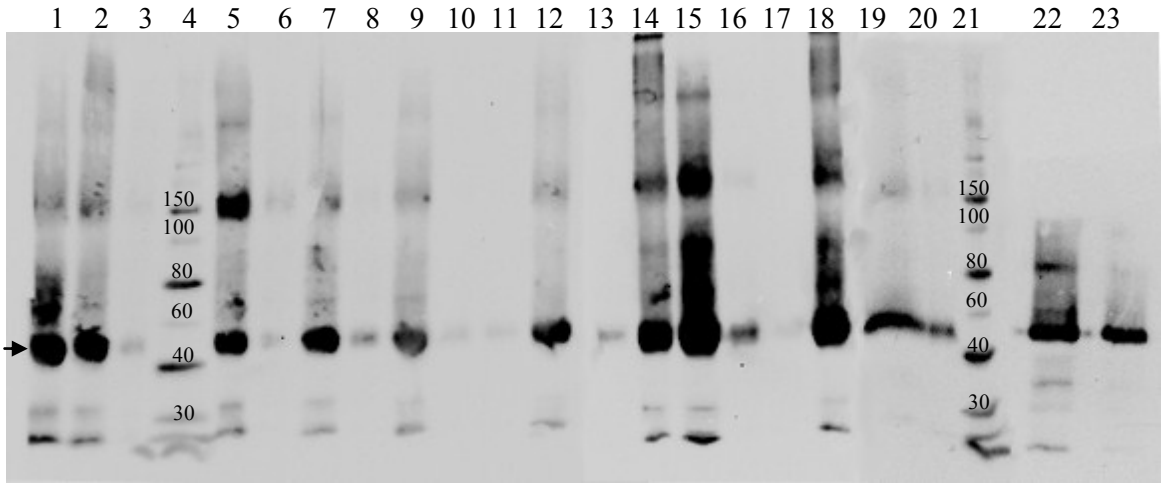


Figure 4.4 *Resolubilization screening of a P-CF-produced cyanobacterial SulP/SLC26 protein. Sample volumes of 2 µl were analyzed by 12% SDS-PAGE. The solubilization efficiencies were determined by densitometry after immunoblotting using anti-His antibodies. Lane 1: Pellet not solubilized; Lane 2: SN DPC; Lane 3: PL DPC; Lane 4 Marker proteins in kDa; Lane 5: SN LMPC; Lane 6: PL LMPC; Lane 7: SN LPPG; Lane 8: PL LPPG; Lane 9: SN LMPG; Lane 10:PL LMPG; Lane 11: SN DDM; Lane 12: PL DDM; Lane 13: SN DHPC; Lane 14: PL DHPC; Lane 15: SN FOS-16; Lane 16: PL FOS-16; Lane 17: SN OG; Lane 18: PL OG; Lane 19: SN NVoy; Lane 20: PL Nvoy; Lane 21: Marker proteins in kDa; Lane 22: SN A8-35; Lane 23: PL A8-35. The protein is indicated by an arrow (MW= 60 kDa).*

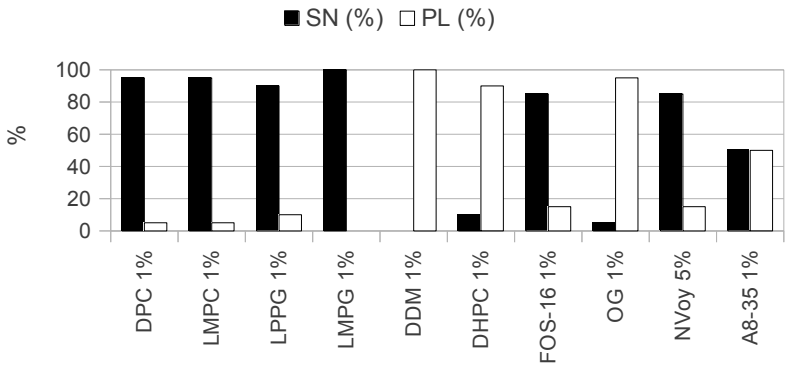


Figure 4.5 *Summary of resolubilization screening (SN= supernatant, PL= pellet).*

The western blot in figure 4.4 indicates a prominent ~50 kDa signal as the cyanobacterial SulP/SLC26 protein, BicA, monomer. In presence of some detergents (LMPC, DHPC, FOS-16, OG, A8-35), additional protein bands were detected by immunoblotting, that could correspond to BicA multimeric complexes (dimers, trimers and tetramers).

4.3.4. Purification of SulP/SLC26 transporters

The resolubilized P-CF precipitates were purified by IMAC based on binding of the C-terminal poly(His)$_{10}$ tag at Ni^{2+} loaded NTA-agarose beads, as described in chapter 4.2.7. Furthermore, the immobilization of SulP/SLC26 proteins upon IMAC purification was used to modify the initially selected hydrophobic environment either by decreasing detergent concentrations or by exchanging the primary detergent with a second detergent having better properties for downstream processes.

An example of IMAC result (of BicA from Synechocystis sp. PCC 6803) is displayed in figure 4.6, where the SulP/SLC26 transporter eluted from the column (fractions 25-30) is almost pure.

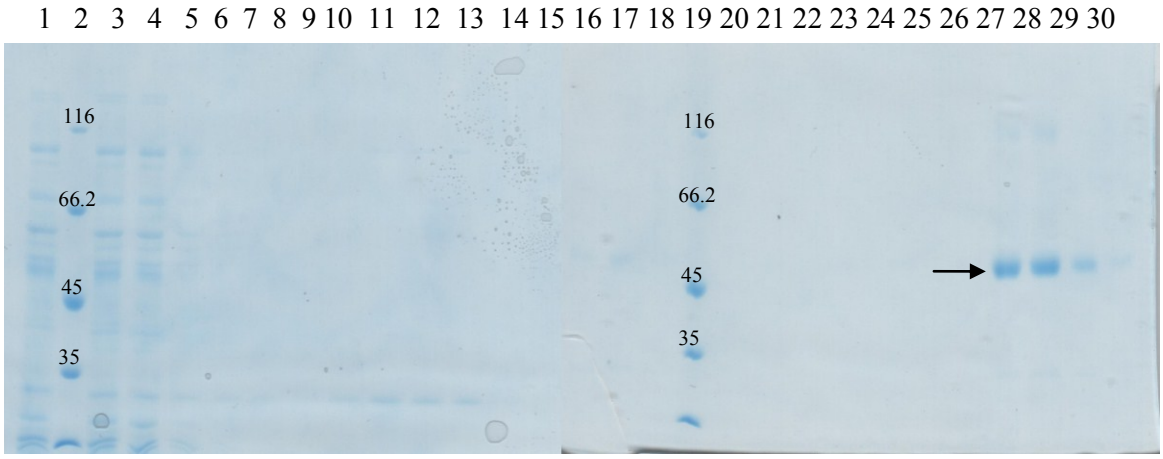


Figure 4.6 *SDS-PAGE of cyanobacterial transporter IMAC purification. Lane 1: PL; Lane 2: FT (flow through); Lane 3: Marker proteins in kDa; Lanes 4-10: equilibration buffer; Lanes 11-16: washing buffer 1; Lanes 17-24: washing buffer 2; Lane 19: Marker proteins in kDa; Lanes 25-30: elution buffer. The protein is indicated by an arrow (MW= 60 kDa).*

4.3.5. Evaluation of SulP/SLC26 transporters quality

The homogeneity of SulP/SLC26 samples purified by Ni^{2+}-NTA affinity chromatography was evaluated by SEC elution profiling to detect potential residual protein aggregates. Analyzing elution profiles of size exclusion columns is a first and general tool in order to detect aggregates or apparent heterogeneities in protein samples. Thus, a selection of detergents frequently used in crystallization screenings of membrane proteins was evaluated with respect to their effects on SulP/SLC26 transporters homogeneity; there is, in fact, a correlation between monodispersity and crystallizability. A single peak on the resulting chromatogram corresponds to a size-homogeneous protein and is often considered most suitable for crystallization. Additional peaks in the

chromatogram are usually an indicator of undesirable multimeric or aggregated target membrane protein or contaminants (Procedure 28-9490-13 AA, GE Healthcare, 2008).

It has been reported that most transporters have been crystallized in DDM, that has a long alkyl chain, and consequently is relatively mild, but those structures solved in DDM are, on average, of lower resolution compared to short-chain detergents as DM, NM and OG (Sonoda et al., 2011).

In presence of 0.05% (w/v) DDM a main peak was obtained close to the V_0 (indicated by arrows in the figure 4.7), indicating severe oligomerization or aggregation. Furthermore in presence of this detergent, Sultr1;2 and Rv1739c precipitated in very high molecular weight aggregates.

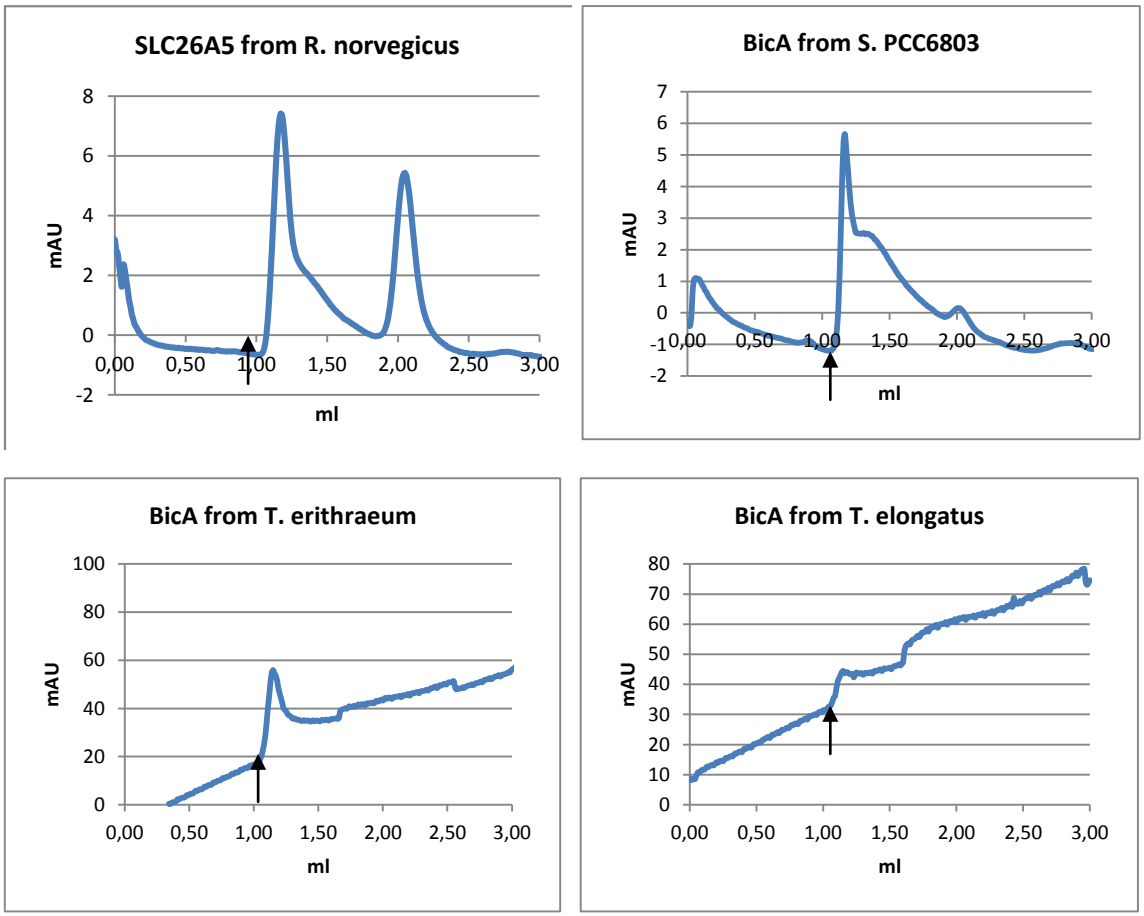


Figure 4.7 *SEC elution profiles of selected SulP/SLC26 transporters in presence of DDM. Column: Superdex 200 5/150 GL (GE Healthcare) equilibrated with 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% (w/v) DDM. The arrows indicate the void volume of the column (V_0).*

SulP/SLC26 transporters in 0.06% LDAO (Figure 4.8) showed a broad elution profile indicating a high degree of proteins heterogeneity. Furthermore, in presence of this

detergent BicA from *Synechocystis sp. PCC 6803* and from *Thermosynechococcus elongatus* precipitated in very high molecular weight aggregates.

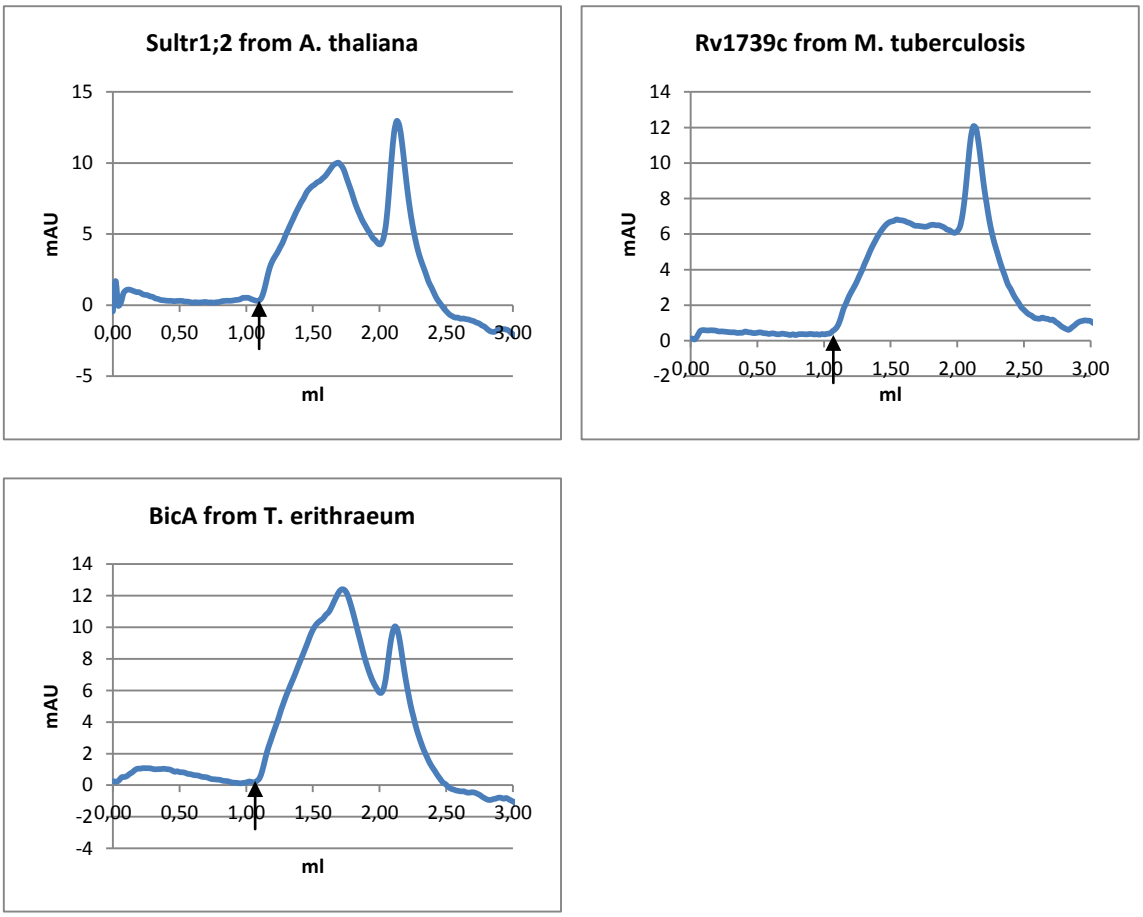


Figure 4.8 *SEC elution profiles of selected SulP/SLC26 transporters in presence of LDAO. Column: Superdex 200 5/150 GL (GE Healthcare) equilibrated with 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.06% (w/v) LDAO. The arrows indicate the void volume of the column (V_0).*

FOS-12 is not one of the detergents most frequently used in crystallization, but it was tested because we wanted to analyze the structure of selected SulP/SLC26 transporters by SANS measurements in collaboration with Dr. Arnaud Javelle, at the Division of Molecular Microbiology, College of Life Sciences, University of Dundee, United Kingdom. Javelle and colleagues published the first low-resolution structure of a bacterial SulP/SLC26 transporter in presence of FOS-12 (Compton et al., 2011) and this detergent was preferred because they already knew the contrast match point (%D_2O use to purify the protein). The final yield of the purified protein was 0.2 mg per 1 ml of CF expression, too low for SANS measurements (1-2 mg), but with a bigger preparation this analysis could be done.

In presence of 0.1% (w/v) FOS-12 (Figure 4.9), a small peak appeared close to the void volume (V_0), indicating that oligomerization or aggregation appeared to a limited extent. Cyanobacterial transporters have a better behavior in comparison with eukaryotic proteins and Rv1739c, with only a small shoulder close to V_0.

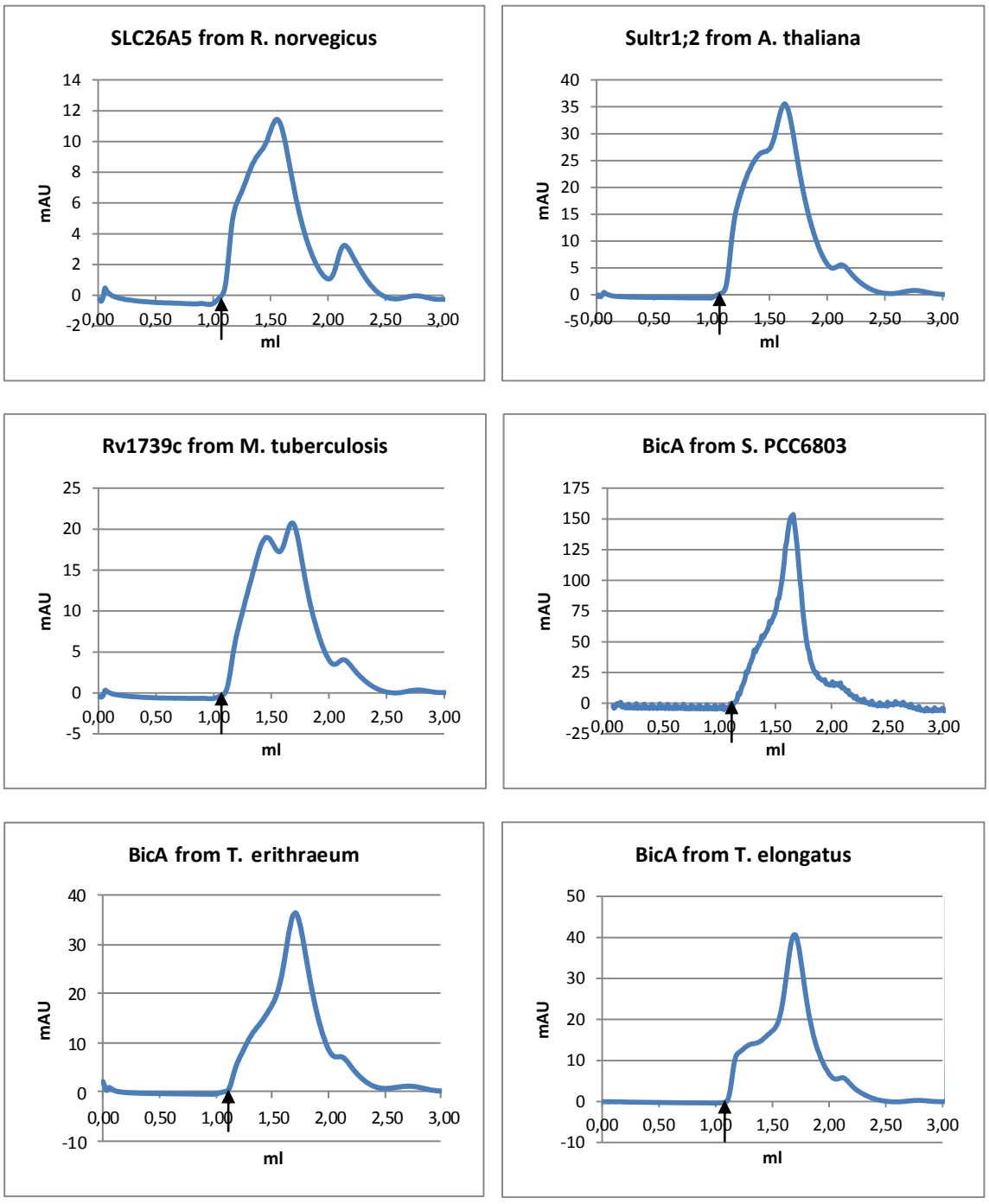


Figure 4.9 *SEC elution profiles of selected SulP/SLC26 transporters in presence of FOS-12. Column: Superdex 200 5/150 GL (GE Healthcare) equilibrated with 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% (w/v) FOS-12. The arrows indicate the void volume of the column (V_0).*

DM NG detergent belongs to Neopentyl Glycol class detergents that are particularly beneficial in the crystallization process due to some unique properties conferred by a revolutionary new architecture. The amphiphilic molecule consists of a central quaternary carbon with two hydrophilic heads and two lipophilic tails, generating subtle constraints on overall conformational flexibility that allows the molecule to pack densely when forming a micelle. This dense packing increases thermal stability of the detergent/protein complex and most importantly, produces exceptionally low critical micelle concentrations and extreme water solubility.

But, as in presence of DDM (Figure 4.7), in presence of 0.02% (w/v) DM NG (Figure 4.10) a main peak was obtained close to the V_0, indicating severe oligomerization or aggregation. Furthermore, in presence of this detergent Sultr1;2 and Rv1739c had a strong tendency to aggregate.

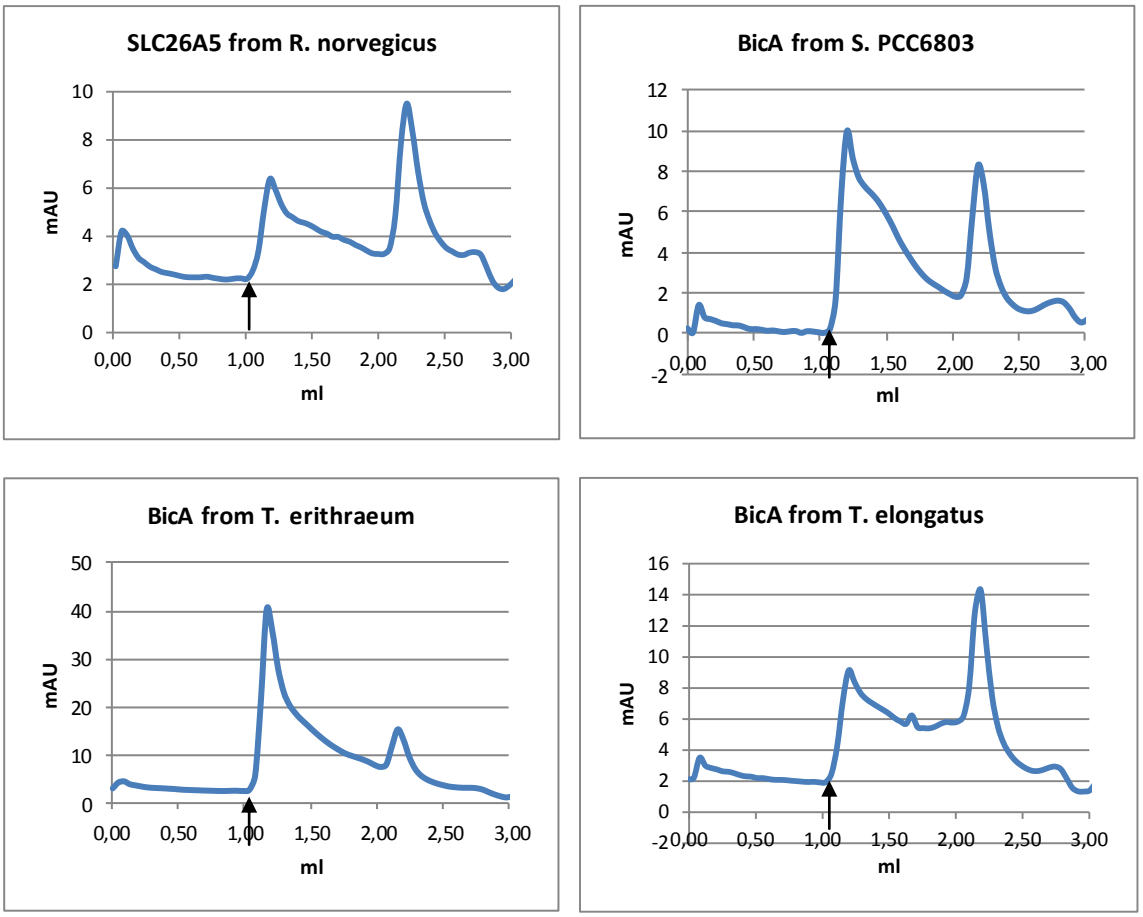


Figure 4.10 *SEC elution profiles of selected SulP/SLC26 transporters in presence of DM NG. Column: Superdex 200 5/150 GL (GE Healthcare) equilibrated with 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.02% (w/v) DM NG. The arrows indicate the void volume of the column (V_0).*

In presence of 0.4% NG the membrane protein UraA, a member of the NAT/NCS2 family, the closer family to the SulP/SLC26 transporters belonging to the APC superfamily (Figure 1.1), was crystallized (Lu et al., 2011).

A symmetric peak shape was obtained with the samples in 0.4% (w/v) NG (Figure 4.11), indicating a homogenous protein under these conditions. SLC26A5, Sultr1;2 and Rv1739c haven't been analyzed in this detergent yet.

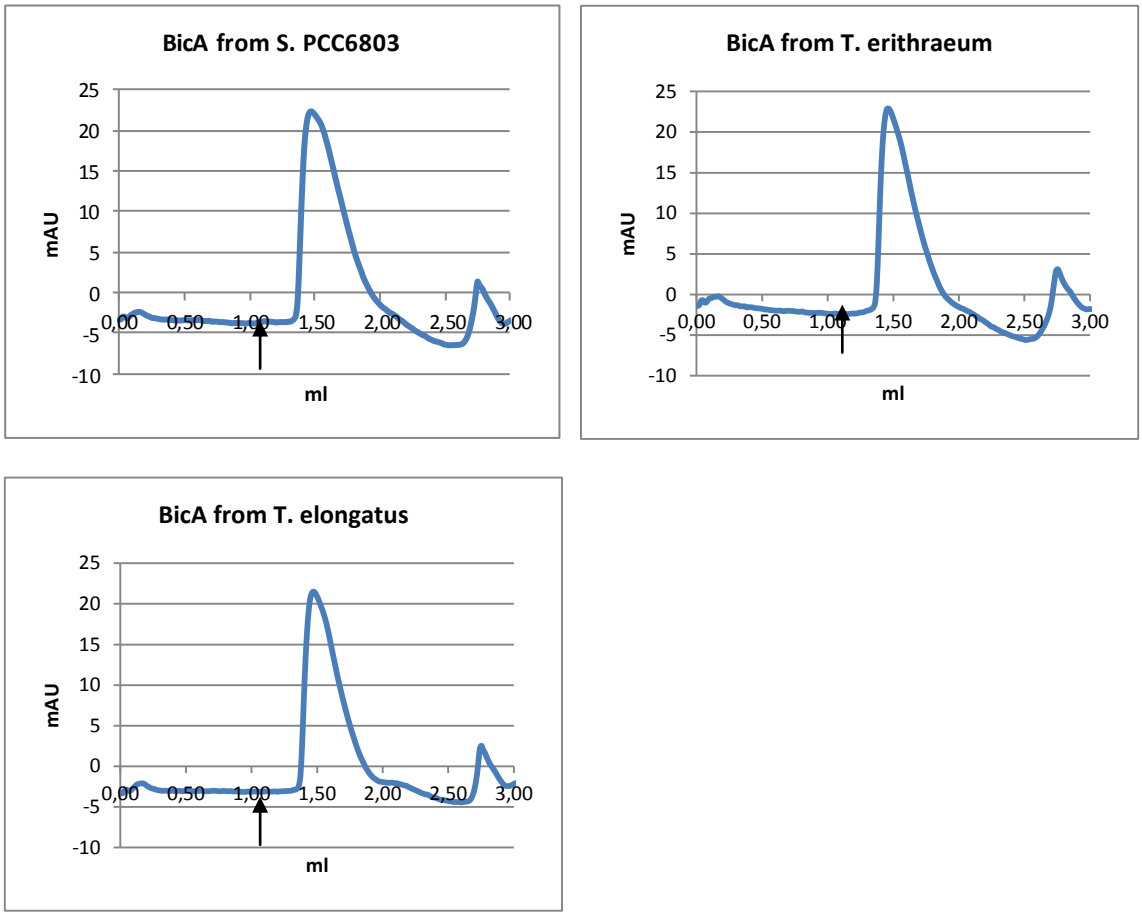


Figure 4.8 *SEC profiles of selected SulP/SLC26 transporters in presence of NG. Column: Superdex 200 5/150 GL (GE Healthcare) equilibrated with 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.4% (w/v) NG. The arrows indicate the void volume of the column (V_0).*

The best elution profiles with the highest yield and the lowest degree of apparent transporter aggregation were achieved by cyanobacterial transporters BicA in NG. According to its SEC elution profile, 0.4% (w/v) NG was therefore chosen as most promising detergent for crystallization trials.

In order to verify if the SulP/SLC26 transporters purified have defined structures, they were characterized by circular dichroism spectroscopy. Figure 4.12 displays an

example of a CD spectrum of a bacterial Sulp/SLC26 transporter, BicA from Synechocystis sp. PCC 6803, that shows two negative bands around 208 and 222 nm, indicative of a high content in α-helical structure, according to the secondary structure prediction (Shelden et al., 2009).

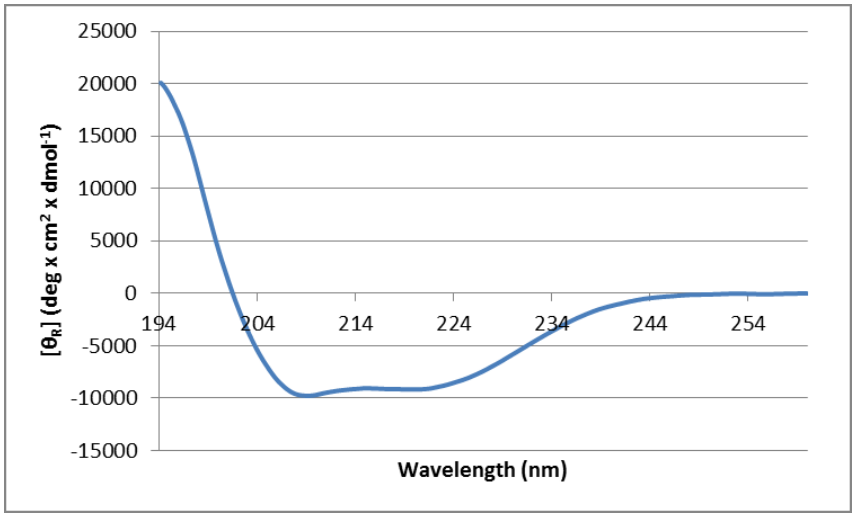


Figure 4.9 *Far-UV circular dichroism spectra of a cyanobacterial SulP/SLC26 transporter, BicA. The protein concentration was 1 mg/ml in 50 mM Na_2HPO_4, 150 mM NaCl, pH 7.5, 0.1% (w/v) FOS-12.*

To understand how stable is a membrane protein in the most successful detergents selected by SEC profile, it is possible to measure its thermal stability, using a fluorescence-based thermal stability assay (or Differential Scanning Fluorimetry, DSF). In this assay the dye N-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]-maleimide (CPM) principally becomes fluorescent upon reacting with free sulfhydryl groups. As most cysteines are predominantly located within transmembrane segments, cysteine accessibility is a good measure of protein unfolding (Alexandrov et al., 2008). Sonoda and co-workers demonstrated that membrane proteins with an unfolding rate longer than approximately 17 min at 40 °C are sufficiently stable for crystallization trials in that detergent (Sonoda et al., 2011). Some preliminary experiments have been tested out, but more work has to be done.

4.3.6. Crystallization screenings

SulP/SLC26 transporters in presence of the most promising detergent for crystallization, according to the SEC elution profile, have to be screened under different

crystallization conditions. Very preliminary results are small crystals of probably not well-ordered protein, unusable for structure determination.

4.3.7. Functional analysis

Lipids are the natural environment of MPs and often are essential modulators of folding, stability and function. Defined lipid environments are thus beneficial for functional studies. The L-CF expression mode in the presence of supplied lipids offers a new and unique option for the co-translational insertion of MPs. In order to functionally characterize the L-CF-produced SulP/SLC26 transporters two main techniques can be used: the patch clamp analysis of transporters reconstituted in proteoliposomes and the Solid Supported Membrane (SSM) based electrophysiology.

4.3.7.1. Patch clamp analysis of proteoliposomes

As regards prestin from *Danio rerio*, a patch clamp analysis of L-CF-produced proteoliposomes was performed in the Prof. Dominik Oliver laboratory at the University of Marburg (Germany), where the transport mode and stoichiometry of this transporter have been identified (Schaechinger and Oliver, 2007). But patch clamp analysis didn't work because proteoliposomes were too small for analysis and/or there were not enough protein inserted into liposomes. Possible solutions are the use of giant liposomes and to increase the insertion of protein into liposomes, following it by GFP fusion protein fluorescence. Selected SulP/SLC26 transporters were cloned successfully in frame with GFP coding sequence.

Another approach to increase the insertion of the protein into liposomes is to produce the SulP/SLC26 transporters in P-CF mode and then insert them into lipids using the in-vitro reconstitution.

4.3.7.2. SSM based electrophysiology

Electrophysiological measurements based on solid supported membranes (SSM) have been used for the functional characterization of ion pumps and transporters. The SSM consists of an alkanethiol monolayer (Thiol) with a lipid monolayer (PC) on top. Proteoliposomes, membrane vesicles, or membrane fragments containing the transport

protein under investigation are adsorbed to an SSM and are activated using a rapid substrate concentration jump. Then transient currents corresponding to the electrogenic translocation of the substrates are measured (Schulz et al., 2008). We tried a new approach of this technique using nanodiscs (NDs) instead of proteoliposomes; nanodiscs are planar phospholipid bilayers surrounded by a protein belt, termed membrane scaffold protein. Nanodiscs were added to RM of CF expression system and the solubilization efficiency of selected SulP/SLC26 transporter was determined by densitometry after immunoblotting. In the example of NDs screening displayed in the figure 4.13, the best NDs concentration was evaluated between 30-40 µM (lanes 6 and 8).

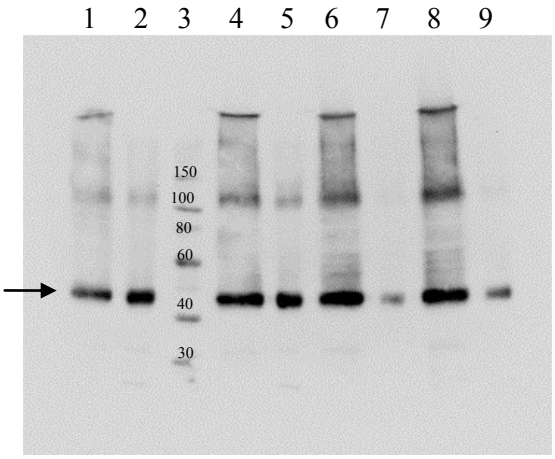


Figure 4.10 *NDs screening of a cyanobacterial transporter, BicA. Sample volumes of 2 µl were analyzed by 12% SDS–PAGE and stained by immunoblotting using anti-His antibodies. Lane 1: SN 10 µM NDs; Lane 2: PL 10 µM NDs; Lane 3: Marker proteins in kDa; Lane 4: SN 20 µM NDs; Lane 5: PL 20 µM NDs; Lane 6: SN 30 µM NDs; Lane 7: PL 30 µM NDs; Lane 8: SN 40 µM NDs; Lane 9: PL 40 µM NDs. The NDs used were assembled with DMPC. The protein is indicated by an arrow (MW= 60 kDa).*

Firstly, we tried this new approach using a model protein, LacY, a lactose permease from *Escherichia coli* well functionally characterized. SSM based electrophysiology analysis was performed in the Prof. Fendler laboratory, at the Department of Biophysical Chemistry, Max-Planck-Institut für Biophysik, Frankfurt. But this approach didn't seem to work, and then the standard SSM-based electrophysiology could be tested out.

4.4. Conclusions

With the final aim to obtain structural and functional information on the full-length SulP/SLC26 anion transporters, different proteins, from distance-related species, were selected, produced by CF expression system based on *E. coli* extracts, and characterized: SLC26A5/prestin from *Rattus norvegicus*, *Gallus gallus* and *Danio rerio*; Sultr 1;2 from *Arabidopsis thaliana*; Rv1739c from *Mycobacterium tuberculosis* and BicA from *Synechocystis sp. PCC 6803*, *Trichodesmium erythraeum* and *Thermosynechococcus elongatus BP-1*. Among these, the expression levels of only SLC26A5/prestin from *Rattus norvegicus* and *Gallus gallus* were considered too low for structural analysis. After the insoluble expression of SulP/SLC26 transporters, selected detergents were tested for their efficiency in the resolubilization of the CF produced precipitates. Once identified the best detergents for resolubilization, SulP/SLC26 transporters were purified by affinity chromatography. Analysis of SEC elution profiles is an established and widely accepted monitoring tool for initial estimation of protein aggregation and heterogeneity. Conditions were found that resulted in symmetrically shaped SulP/SLC26 elution peak profiles, indicating a relatively uniform particle size distribution. Resolubilization with FOS-16 followed by exchange to NG was best suitable for P-CF produced cyanobacterial BicA precipitates.

SulP/SLC26 transporters in presence of the most promising detergent for crystallization, according to the SEC elution profiles, have to be submitted to more extensive crystallization trials, also trying another useful method for membrane proteins crystallization, the *in meso* method, that makes use of a lipidic mesophase, the cubic phase, formed by homogenizing lipid, typically monoolein, and water in approximately equal parts (Caffrey et al., 2012).

The most critical test for the proper folding of reconstituted membrane proteins is functionality; therefore, it will be important to confirm the transport activity of the P-CF-produced SulP/SLC26 transporters by means of functional analysis or binding assays.

To conclude, this is the first quality evaluation of a SulP/SLC26 transporters production by CF expression system and we demonstrate that these transporters can be produced by CF protocols, resulting in qualities apparently adequate for further functional and structural approaches.

5. Bibliography

A. Accardi, S. Lobet, C. Williams, C. Miller and R. Dutzler (2006). Synergism between halide binding and proton transport in a CLC-type exchanger. *Journal of molecular biology* **362**: 691-699

P. D. Adams, R. W. Grosse-Kunstleve, L. W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter, T. C. Terwilliger (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **58**(11): 1948-54

J. T. Albert, H. Winter, T. J. Schaechinger, T. Weber, X. Wang, D. Z. Z. He, O. Hendrich, H. Geisler, U. Zimmermann, K. Oelmann, M. Knipper, M. C. Göpfert and D. Oliver (2007). Voltage-sensitive prestin orthologue expressed in zebrafish hair cells. *J Physiol* **580**(2): 451-461

A. I. Alexandrov, M. Mileni, E. Y. T. Chien, M. A. Hanson, and R. C. Stevens (2008). Microscale Fluorescent Thermal Stability Assay for Membrane Proteins. *Structure* **16**: 351-359

L. Aravind and E. V. Koonin (2000). The STAS domain - a link between anion transporters and antisigma-factor antagonists. *Curr. Biol.* **10**: 53-55

J. Ashmore (2008). Cochlear Outer Hair Cell Motility. *Physiol Rev* **88**: 173-210

M. Babu, J. F. Greenblatt, A. Emili, N. C. J. Strynadka, R. A. F. Reithmeier, and T. F. Moraes (2010). Structure of a SLC26 Anion Transporter STAS Domain in Complex with Acyl Carrier Protein: Implications for E. coli YchM in Fatty Acid Metabolism. *Structure* **18**: 1450-1462

J. Bai, A. Surguchev, S. Montoya, P. S. Aronson, J. Santos-Sacchi, and D. Navaratnam (2009). Prestin's Anion Transport and Voltage-Sensing Capabilities Are Independent. *Biophysical Journal* **96**: 3179-3186

J. Bai, A. Surguchev, Y. Ogando, L. Song, S. Bian, J. Santos-Sacchi, and D. Navaratnam (2010). Prestin surface expression and activity are augmented by interaction with MAP1S, a microtubule associated protein. *The journal of Biological Chemistry* **285***:* 20834-20843

T. H. Bayburt, Y. V. Grinkova, and S. G. Sligar (2002). Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Lett* **2**(8): 853-856

T. H. Bayburt and S. G. Sligar (2002). Membrane protein assembly into Nanodiscs. FEBS Letters **584**: 1721-1727

M. Benvenuti and S. Mangani (2007). Crystallization of soluble proteins in vapor diffusion for x-ray crystallography. *Nature Protocols* **2**(7): 1633-1651

T. Bergfors (2003). Seed to crystals. *Journal of structural Biology* **142**: 66-76

S. E. Bondos and A. Bicknell (2003). Detection and prevention of protein aggregation before, during, and after purification. *Analytical Biochemistry* **316**: 223-231

M. Caffrey, D, Li, and A. Dukkipati (2012). Membrane Protein Structure Determination Using Crystallography and Lipidic Mesophases: Recent Advances and Successes. *Biochemistry* **51**: 6266-6288

M. Chang, C. Plata, K. Zandi-Nejad, A. Sinđić, C. R. Sussman, A. Mercado, V. Broumand, V. Raghuram, D. B. Mount, and M. F. Romero (2009). Slc26A9 - anion exchanger, channel and Na^+ transporter. *J Membr Biol.* **228**(3): 125-140

M. Chang, C. Plata, A. Sindic, W. K. Ranatunga, A. Chen, K. Zandi-Nejad, K. W. Chan, J. Thompson, D. B. Mount, and M. F. Romero (2009 b). Slc26a9 Is Inhibited by

the R-region of the Cystic Fibrosis Transmembrane Conductance Regulator via the STAS Domain. *J Biol Chem* **284**(41): 28306-28318

M. N. Chernova, L. Jiang, B. E. Shmukler, C. W. Schweinfest, P. Blanco, S. D. Freedman, A. K. Stewart and S. L. Alper (2003). Acute regulation of the SLC26A3 congenital chloride diarrhoea anion exchanger (DRA) expressed in Xenopus oocytes. *J Physiol* **549**(1): 3-19

J. Y. Choi, D. Muallem, K. Kiselyov, M. G. Lee, P. J. Thomas and S. Muallem (2001). Aberrant CFTR-dependent HCO_3^- transport in mutationsassociated with cystic fibrosi. *Nature* **410**: 94-97

J. Clarkson, I. D. Campbell and M. D. Yudkin (2003). Phosphorylation induces subtle structural changes in SpoIIAA, a key regulator of sporulation. *Biochem. J.* **372**:113-119

E. L. R. Compton, E. Karinou, J. H. Naismith, F. Gabel, and A. Javelle (2011). Low Resolution Structure of a Bacterial SLC26 Transporter Reveals Dimeric Stoichiometry and Mobile Intracellular Domains. *J. Biol. Chem.* **286**(30): 27058-27067

P. Dallos and B. Fakler (2002). Prestin, a new type of motor protein. *Nature reviews, Molecular cell biology* **3**: 104-111

P. Dallos, J. Zheng and M. A. Cheatham (2006). Prestin and the cochlear amplifier. *J Physiol* **576**(1): 37-42

P. Dallos (2008). Cochlear amplification, outer hair cells and prestin. *Current Opinion in Neurobiology* **18**: 370-376

A. D'Arcy, F. Villarda and M. Marshb (2007). An automated microseed matrix-screening method for protein crystallization. *Acta Cryst.* D**63**: 550-554

P. A. Dawson and D. Markovich (2005). Pathogenetics of the Human SLC26 Transporters. *Current Medicinal Chemistry* **12**: 385-396

L. Deák, J. Zheng, A. Orem, G. Du, S. Aguiñaga, K. Matsuda and P. Dallos (2005). Effects of cyclic nucleotides on the function of prestin. *J Physiol* **563**(2): 483-496

Y. Dehouck, A. Grosfils, B. Folch, D. Gilis, P. Bogaerts and M. Rooman (2009). Fast and accurate predictions of protein stability changes upon mutations using statistical potentials and neural networks: PoPMuSiC-2.0. *Structural Bioinformatics* **25**(19): 2537-2543

S. Detro-Dassen, M. Schanzler, H. Lauks, I. Martin, S. M. zu Berstenhorst, D. Nothmann, D. Torres-Salazar, P. Hidalgo, G. Schmalzing, C. Fahlke (2008). Conserved Dimeric Subunit Stoichiometry of SLC26 Multifunctional Anion Exchangers. *Journal of biological chemistry* **283**: 4177-4188

M. R. Dorwart, N. Shcheynikov, D. Yang, S. Muallem (2008). The Solute Carrier 26 Family of Proteins in Epithelial Ion Transport. *Physiology* **23**: 104-114

S. Dossena, S. Rodighiero, V. Vezzoli, C. Nofziger, E. Salvioni, M. Boccazzi, E. Grabmayer, G. Botta, G. Meyer, L. Fugazzola, P. Beck-Peccoz and M. Paulmichl (2009). Functional characterization of wild-type and mutated pendrin (SLC26A4), the anion transporter involved in Pendred syndrome. *Journal of Molecular Endocrinology* **43**: 93-103

A. B. Elgoyhen, and L. F. Franchini (2011). Prestin and the cholinergic receptor of hair cells: positively-selected proteins in mammals. *Hear Res* **273**(1-2): 100-108

P. Emsley, K. Cowtan (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**(12 1): 2126-2132

T. Etezady-Esfarjani, W. J. Placzek, T. Herrmann and K. Wuthrich (2006). Solution structures of the putative anti-σ-factor antagonist TM1442 from Thermotoga maritima in the free and phosphorylated states. *Magn. Reson. Chem.* **44**: S61-S70

P. Evans (2005). Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**(1): 72-82

L. A. Everett, I. A. Belyantseva, K. Noben-Trauth, R. Cantos, A. Chen, S. I. Thakka, S. L. Hoogstraten-Miller, B. Kachar, D. K. Wu, E. D. Green (2001). Targeted disruption of mouse Pds provides insight about the inner-ear defects encountered in Pendred syndrome. *Human Molecular Genetics* **10**: 153-161

S. Faham, A. Watanabe, G. M. Besserer, D. Cascio, A. Specht, B. A. Hirayama, E. M. Wright, J. Abramson (2008). The Crystal Structure of a Sodium Galactose Transporter Reveals Mechanistic Insights into Na^+/Sugar Symport. *Science* **321**: 810-814

J. Felce, M. H. Saier, Jr. (2004). Carbonic Anhydrases Fused to Anion Transporters of the SulP Family: Evidence for a Novel Type of Bicarbonate Transporter. *Mol Microbiol Biotechnol* **8**: 169-176

P. Fong (2012). CFTR-SLC26 transporter interactions in epithelia. *Biophys Rev* **4**:107-116

L. F. Franchini, A. B. Elgoyhen (2006). Adaptive evolution in mammalian proteins involved in cochlear outer hair cell electromotility. *Molecular Phylogenetics and Evolution* **41**: 622-635

X. Gao, F. Lu, L. Zhou, S. Dang, L. Sun, X. Li, J. Wang, Y. Shi (2009). Structure and Mechanism of an Amino Acid Antiporter. *Science* **324**: 1565-1568

M. A. Gray (2004). Bicarbonate secretion: it takes two to tango. *Nature cell biology* **6**(4): 392-394

J. N. Greeson, L. E. Organ, F. A. Pereira, R. M. Raphael (2006). Assessment of prestin self-association using fluorescence resonance energy transfer. *Brain research* **1091**: 140-150

S. Haberstock, C. Roos, Y. Hoevels, V. Dötsch, G. Schnapp, A. Pautsch, F. Bernhard (2012). A systematic approach to increase the efficiency of membrane protein production in cell-free expression systems. *Protein Expression and Purification* **82**: 308-316

R. Hallworth and M. G. Nichols (2011). Prestin in HEK Cells is an Obligate Tetramer. *J Neurophysiol* **107**(1): 5-11

H. Hamada, T. Arakawa and K. Shiraki (2009). Effect of Additives on Protein Aggregation. *Current Pharmaceutical Biotechnology* **10**: 400-407

D. Z. Z. He, S. Jia, T. Sato, J. Zuo, L. R. Andrade, G. P. Riordan, and B. Kachar (2010). Changes in plasma membrane structure and electromotile properties in prestin deficient outer hair cells. *Cytoskeleton (Hoboken)* **67**(1): 43-55

R. Hibbs, and E. Gouaux (2011). Principles of activation and permeation in an anion-selective Cys-loop receptor. *Nature* **474**: 54-60

K. Homma, K. K. Miller, C. T. Anderson, S. Sengupta, G. Du, S. Aguiñaga, M. Cheatham, P. Dallos, J. Zheng (2010). Interaction between CFTR and prestin (SLC26A5). *Biochim. Biophys. Acta* **1798**(6):1029-1240

F. Junge, S. Haberstock, C. Roos, S. Stefer, D. Proverbio, V. Dötsch and F. Bernhard (2010). Advances in cell-free protein synthesis for the functional and structural analysis of membrane proteins. *N. Biotechnol* **28**(3): 262-71

M. Jurk, M. Dorn, and P. Schmieder (2011). Blue Flickers of Hope: Secondary Structure, Dynamics, and Putative Dimerization Interface of the Blue-Light Receptor YtvA from *Bacillus subtilis*. *Biochemistry* **50**: 8163−8171

W. Kabsch (2010) XDS. *Acta Cryst.* D**66:** 125-132

E. Karinou, E. L. Compton, M. Morel, A. Javelle (2012). The Escherichia coli SLC26 homologue YchM (DauA) is a C(4) -dicarboxylic acid transporter. *Mol Microbiol.*

S. B. H. Ko, N. Shcheynikov, J. Y. Choi, X. Luo, K. Ishibashi, P. J.Thomas, J. Y. Kim, K. H. Kim, M. G. Lee, S. Naruse and S. Muallem (2002). A molecular mechanism for aberrant CFTR-dependent HCO_3^- transport in cystic fibrosis. *The EMBO Journal* **21**(21): 5662-5672

S. B. H. Ko, W. Zeng, M. R. Dorwart, X. Luo, K. H. Kim, L. Millen, H. Goto, S. Naruse, A. Soyombo, P. J. Thomas and S. Muallem (2004). Gating of CFTR by the STAS domain of SLC26 transporters. *Nature cell biology* **6**(4): 343-350

H. Kovacs, D. Comfort, M. Lord, I. D. Campbell, and M. D. Yudkin (1998). Solution structure of SpoIIAA, a phosphorylatable component of the system that regulates transcription factor σ^F of Bacillus subtilis. *Proc. Natl. Acad. Sci. USA* **95**: 5067-5071

S. Kumano, X. Tan, D. Z. Z. He, K. Iida, M. Murakoshi, H. Wada (2009). Mutation-induced reinforcement of prestin-expressing cells. *Biochemical and Biophysical Research Communications* **389**: 569-574

S. Kumano, M. Murakoshi, K. Iida, H. Hamana, H. Wada (2010). Atomic force microscopy imaging of the structure of the motor protein prestin reconstituted into an artificial lipid bilayer. *FEBS Letters* **584**: 2872-2876

K. Legendre, S. Safieddine, P. Küssel-Andermann, C. Petit and A. El-Amraoui (2008). αII-βV spectrin bridges the plasma membrane and cortical lattice in the lateral wall of the auditory outer hair cells. *Journal of Cell Science* **121**: 3347-3356

C. Le Grimellec, M. C. Giocondi, M. Lenoir, M. Vater, G. Sposito, and R. Pujol (2002). High-resolution three-dimensional imaging of the lateral plasma membrane of cochlear outer hair cells by atomic force microscopy. *J. Comp. Neurol.* **451**: 62- 69

X. Z. Liu, X. M. Ouyang, X. J. Xia, J. Zheng, A. Pandya, F. Li, L. L. Du, K. O. Welch, C. Petit, R. J. Smith, B. T. Webb, D. Yan, K. S. Arnos, D. Corey, P. Dallos, W. E. Nance, Z. Y. Chen (2003). Prestin, a cochlear motor protein, is defective in non-syndromic hearing loss. *Hum Mol Genet* **12**(10): 1155-62

S. Lobet, and R. Dutzler (2006). Ion-binding properties of the ClC chloride selectivity filter. *The EMBO journal* **25**: 24-33

M. A. Lomize, A. L. Lomize, I. D. Pogozheva, and H. I. Mosberg (2006). OPM: Orientations of Proteins in Membranes Database. *Bioinformatics* **22**: 623-625

F. Lu, S. Li1, Y. Jiang, J. Jiang, H. Fan, G. Lu, D. Deng, S. Dang, X. Zhang, J. Wang, N. Yan (2011). Structure and mechanism of the uracil transporter UraA. *Nature* **472**: 243-247

M. P. Malakhov, M. R. Mattern, O. A. Malakhova, M. Drinker, S. D. Weeks, T. R. Butt (2004). SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *Journal of Structural and Functional Genomics* **5**: 75-86

D. Markovich (2012). Slc13a1 and Slc26a1 KO Models Reveal Physiological Roles of Anion Transporters. *Phisiology* **27**: 7-14

S. Masuda, K. S. Murakami, S. Wang, C. A. Olson, J. Donigian, F. Leon, S. A. Darst and E. A. Campbell (2004). Crystal Structures of the ADP and ATP Bound Forms of the Bacillus Anti-σ Factor SpoIIAB in Complex with the Anti-anti-σ SpoIIAA. *J. Mol. Biol*. **340**: 941-956

A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read (2007). Phaser crystallographic software. *J Appl Crystallogr* **40**(4): 658-674

J. S. Minor, H. Tagng, A. Pereira, R. L. Alford (2009). DNA Sequence Analysis of SLC26A5, Encoding Prestin, in a Patient-Control Cohort: Identification of Fourteen Novel DNA Sequence Variations. *PloS ONE* **4**(6): e5762

K. Mio, Y. Kubo, T. Ogura, T. Yamamoto, F. Arisaka, C. Sato (2008). The Motor Protein Prestin Is a Bullet-shaped Molecule with Inner Cavities. *J Biol Chem* **283**(2): 1137-1145

P. Mistrìk, N. Daudet, K. Morandell and J.F. Ashmore (2012). Mammalian prestin is a weak Cl-/HCO3- electrogenic antiporter. *J Physiol* **590**(22): 5597-5610

D. B. Mount, M. F. Romero (2004). The SLC26 gene family of multifunctional anion exchangers. *Pflugers Arch – Eur J Physiol* **447**: 710-721

D. Muallem and J. Ashmore (2006). An Anion Antiporter Model of Prestin, the Outer Hair Cell Motor Protein. *Biophysical Journal* **90**: 4035-4045

M. Murakoshi, K. Iida, S. Kumano, H. Wada (2008). Immune atomic force microscopy of prestin-transfected CHO cells using quantum dots. *Pflugers Arch – Eur J Physiol* **457**: 885-898

G. N. Murshudov, A. A. Vagin, E. J. Dodson (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**(3): 240-255

Y. Nakasone and K. J. Hellingwerf (2011). On the Binding of BODIPY-GTP by the Photosensory Protein YtvA from the Common Soil Bacterium Bacillus subtilis. *Photochemistry and Photobiology* **87**: 542-547

D. Navaratnam, J. Bai, H. Samaranayake, and J. Santos-Sacchi (2005). N-Terminal-Mediated Homomultimerization of Prestin, the Outer Hair Cell Motor Protein. *Biophysical Journal* **89**: 3345-3352

N. Oganesyan, S. Kim and R. Kim (2004). On-column Chemical Refolding of Proteins. *Pharmagenomics* **4**: 22-25

E. Ohana, N. Shcheynikov, M. Park, S. Muallem (2012). Solute Carrier Family 26 Member a2 (Slc26a2) Protein Functions as an Electroneutral SO_4^{2-}/OH^-/Cl^- Exchanger Regulated by Extracellular Cl^-. *J. Biol. Chem*. **287**: 5122-5132

D. Oliver, D. Z. Z. He, N. Klocker, J. Ludwig, U. Schulte, S. Waldegger, J. P. Ruppersberg, P. Dallos, B. Fakler (2001). Intracellular Anions as the Voltage Sensor of Prestin, the Outer Hair Cell Motor Protein. *Science* **292**: 2340-2343

M. W. Pantoliano, A. W. Rhind, F. R. Salemme. Microplate thermal shift assay for ligand development and multi-variable protein chemistry optimization. *US Patent* 6,020,141 (Issued 2/1/2000).

E. Pasqualetto, A. Seydel, A. Pellini, R. Battistutta (2008). Expression, purification and characterisation of the C-terminal STAS domain of the SLC26 anion transporter prestin. *Protein Expression and Purification* **58**: 249-256

E. Pasqualetto, R. Aiello, L. Gesiot, G. Bonetto, M. Bellanda and R. Battistutta (2010). Structure of the Cytosolic Portion of the Motor Protein Prestin and Functional Role of the STAS Domain in SulP/SLC26 Anion Transporters. *J. Mol. Biol.* **400**: 448-462

G. D. Price, F. J. Woodger, M. R. Badger, S. M. Howitt, and L. Tucker (2004). Identification of a SulP-type bicarbonate transporter in marine cyanobacteria. *PNAS* **101**(52): 18228-18233

G. D. Price (2011). Inorganic carbon transporters of the cyanobacterial CO_2 concentrating mechanism. *Photosynth Res* **109**(1-3): 47-57

G. D. Price and S. M. Howitt (2011). The cyanobacterial bicarbonate transporter BicA: its physiological role and the implications of structural similarities with human SLC26 transporters. *Biochem. Cell Biol.* **89**: 178-188

Procedure 28-9490-13 AA (2008). Combining detergent screening and size homogeneity analysis of histidine-tagged membrane protein using His MultiTrap FF and Superdex 200 5/150 GL. *GE Healthcare*

M. Quin, J. Newman, S. Firbank, R. J. Lewis, and J. Marles-Wright (2008). Crystallization and preliminary X-ray analysis of RsbS from Moorella thermoacetica at 2.5 Å resolution. *Acta Cryst.* **F64**: 196-199

S. Ressl, A. C. Terwisscha van Scheltinga, C. Vonrhein, V. Ott, C. Ziegler (2009). Molecular basis of transport and regulation in the Na^+/betaine symporter BetP. *Nature* **458**: 47-53

H. Rouached, P. Berthomieu, E. El Kassis, N. Cathala, V. Catherinot, G. Labesse, J. Davidian, and P. Fourcroy (2005). Structural and Functional Analysis of the C-terminal STAS (Sulfate Transporter and Anti-sigma Antagonist) Domain of the *Arabidopsis thaliana* Sulfate Transporter SULTR1;2. *J. Biol. Chem.* **280**(16): 15976 -15983

V. Rybalchenko and J. Santos-Sacchi (2003). Cl^- flux through a non-selective, stretch-sensitive conductance influences the outer hair cell motor of the guinea-pig. *J Physiol* **547**(3): 873-891

M. H. Saier Jr., B. H. Eng, S. Fard, J. Garg, D. A. Haggerty, W. J. Hutchinson, D. L. Jack, E. C. Lai, H. J. Liu, D. P. Nusinew, A. M. Omar, S. S. Pao, I. T. Paulsen, J. A. Quan, M. Sliwinski, T. Tseng, S. Wachi, G. B. Young (1999). Phylogenetic characterization of novel transport protein families revealed by genome analyses. *Biochimica et Biophysica Acta* **1422**: 1-56

M. H. Saier, Jr (2000). Families of transmembrane transporters selective for amino acids and their derivatives. *Microbiology* **146**: 1775-1795

T. J. Schaechinger and D. Oliver (2007). Nonmammalian orthologs of prestin (SLC26A5) are electrogenic divalent/chloride anion exchangers. *PNAS* **104**(18): 7693-7698

T. J Schaechinger, D. Gorbunov, C. R. Halaszovich, T. Moser, S. Kugler, B. Fakler and D. Oliver (2011). A synthetic prestin reveals protein domains and molecular operation of outer hair cell piezoelectricity. *The EMBO Journal* **30**: 2793-2804

M. Schänzler and C. Fahlke (2012). Anion transport by the cochlear motor protein prestin. *J Physiol* **590**(2) 259-272

B. Schneider, F. Junge, V. A. Shirokov, F. Durst, D. Schwarz, V. Dötsch, and F. Bernhard (2010). Membrane Protein Expression in Cell-Free Systems. I. Mus-Veteau (ed.), H*eterologous Expression of Membrane Proteins, Methods in molecular Biology* **601**: 165-186

P. Schulz, J. J. Garcia-Celma, K. Fendler (2008). SSM-based electrophysiology. *Methods* **46**: 97-103

S. Schulze, S. Koster, U. Geldmacher, A. C. Terwisscha van Scheltinga, W. Kuhlbrandt (2010). Structural basis of Na^+-independent and cooperative substrate/product antiport in CaiT. *Nature* **467**: 233-237

D. Schwarz, F. Junge, F. Durst, N. Frolich, B. Schneider, S. Reckel, S. Sobhanifar, V. Dötsch and F. Bernhard (2007). Preparative scale expression of membrane proteins in Escherichia coli-based continuous exchange cell-free systems. *Nature Protocols* **2**: 2945-2957

P. R. Seavers, R. J. Lewis, J. A. Brannigan, K. H. G. Verschueren, G. N. Murshudov, and A. J. Wilkinson (2001). Structure of the Bacillus Cell Fate Determinant SpoIIAA in Phosphorylated and Unphosphorylated Forms. *Structure* **9**: 605-614

S. Sengupta, K. K. Miller, K. Homma, R. Edge, M. A. Cheatham, P. Dallos, J. Zheng (2010). Interaction between the motor protein prestin and the transporter protein VAPA. *Biochim Biophys Acta* **1803**(7): 796-804

P. Serrano, B. Pedrini, M. Geralt, K. Jaudzems, B. Mohanty, R. Horst, T. Herrmann, M. Elsliger, I. A. Wilson, and K. Wuthrich (2010). Comparison of NMR and crystal structures highlights conformational isomerism in protein active sites. *Acta Cryst.* **F66**: 1393-1405

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

A. K. Sharma, L. Ye, A. S. Zolotarev, S. L. Alper, A. C. Rigby (2009). NMR assignment and secondary structure of the STAS domain of Rv1739c, a putative sulfate transporter of Mycobacterium tuberculosis. *Biomol NMR Assign* **3**(1): 99-102

A. K. Sharma, L. Ye, C. E. Baer, K. Shanmugasundaram, T. Alber, S. L. Alper, and A. C. Rigby (2010). Solution Structure of the Guanine Nucleotide-binding STAS Domain of SLC26-related SulP Protein Rv1739c from *Mycobacterium tuberculosis*. *J. Biol. Chem.* **286**(10): 8534-8544

A. K. Sharma, A. C. Rigby and S. L. Alper (2011). STAS Domain Structure and Function. *Cell Physiol Biochem* **28**: 407-422

M. C. Shelden, S. M. Howitt, G. D. Price (2010). Membrane topology of the cyanobacterial bicarbonate transporter, BicA, a member of the SulP (SLC26A) family. *Molecular Membrane Biology* **27**(1): 12-23

N. Shibagaki and A. R. Grossman (2004). Probing the Function of STAS Domains of the Arabidopsis Sulfate Transporters. *J. Biol. Chem.* **279**(29): 30791-30799

N. Shibagaki and A. R. Grossman (2006). The Role of the STAS Domain in the Function and Biogenesis of a Sulfate Transporter as Probed by Random Mutagenesis. *J. Biol. Chem.* **281**(32): 22964-22973

N. Shibagaki and A. R. Grossman (2010). Binding of Cysteine Synthase to the STAS Domain of Sulfate Transporter and Its Regulatory Consequences. *J. Biol. Chem.* **285**(32): 25094 -25102

G. P. Sinha, F. Sabri, E. K. Dimitriadis, K. H. Iwasa (2010). Organization of membrane motor in outer hair cells: an atomic force microscopic study. *Pflugers Arch - Eur J Physiol* **459**: 427-439

M. Soleimani, T. Greeley, S. Petrovic, Z. Wang, H. Amlal, P. Kopp and C. E. Burnham (2001). Pendrin: an apical $Cl^-/OH^-/HCO_3^-$ exchanger in the kidney cortex. *Am J Physiol Renal Physiol* **280**: F356-F364

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

X. Tan, J. L. Pecka, J. Tang, O. E. Okoruwa, Q. Zhang, K. W. Beisel, and D. Z. Z. He (2011). From Zebrafish to Mammal: Functional Evolution of Prestin, the Motor Protein of Cochlear Outer Hair Cells. *J Neurophysiol* **105**: 36-44

A. Touré, L. Morin, C. Pineau, F. Becq, O. Dorseuil, and G. Gacon (2001). Tat1, a Novel Sulfate Transporter Specifically Expressed in Human Male Germ Cells and Potentially Linked to RhoGTPase Signaling. *J. Biol. Chem.* **276**(23): 20309 -20315

F. van den Ent, J. Lowe (2006). RF cloning: A restriction-free method for inserting target genes into plasmids. J. Biochem. Biophys. *Methods* **67**: 67-74

L. Vuillard, C. Braun-Bretont and T. Rabilloud (1995). Non-detergent sulphobetaines: a new class of mild solubilization agents for protein purification. *Biochem. J.* **305**: 337-343

X. Wang, S. Yang, S. Jia, D. Z. Z. He (2010). Prestin forms oligomer with four mechanically independent subunits. *Brain Res* **1333**: 28-35

P. Wangemann, K. Nakaya, T. Wu, R. J. Maganti, E. M. Itza, J. D. Sanneman, D. G. Harbidge, S. Billings, D. C. Marcus (2007). Loss of cochlear HCO_3^- secretion causes deafness via endolymphatic acidification and inhibition of Ca^{2+} reabsorption in a Pendred syndrome mouse model. American Journal of Physiology. *Renal Physiology* **292**: F1345-F1353

S. Weyand, T. Shimamura, S. Yajima, S. Suzuki, O. Mirza, K. Krusong, E. P. Carpenter, N. G. Rutherford, J. M. Hadden, J. O'Reilly, P. Ma, M. Saidijam, S. G. Patching, R. J. Hope, H. T. Norbertczak, P. C.J. Roach, S. Iwata, P. J. F. Henderson, and A. D. Cameron (2008). Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science* **322**: 709-713

F. H. Wong, J. S. Chen, V. Reddy, J. L. Day, M. A. Shlykov, S. T. Wakabayashi, M. H. Saier, Jr (2012). The Amino Acid-Polyamine-Organocation Superfamily. *J Mol Microbiol Biotechnol.* **22**:105-113

M. Xu, G. Bernat, A. Singh, H. Mi, M. Rogner, H. B. Pakrasi, and T. Ogawa (2008). Properties of Mutants of Synechocystis sp. Strain PCC 6803 Lacking Inorganic Carbon Sequestration Systems. *Plant Cell Physiol* **49**(11): 1672-1677

J. Xu, P. Song, S. Nakamuraţ, M. Miller, S. Barone, S. L. Alper, B. Riederer, J. Bonhagen, L. J. Arend, H. Amlal, U. Seidler, and M. Soleimani (2009). Deletion of the Chloride Transporter Slc26a7 Causes Distal Renal Tubular Acidosis and Impairs Gastric Acid Secretion. *J. Biol. Chem.* **284**: 29470 -29479

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

A. Yoshida, S. Taniguchi, I. Hisatome, I. E. Royaux, E. D. Green, L. D. Kohn, and K. Suzuki (2002). Pendrin Is an Iodide-Specific Apical Porter Responsible for Iodide Efflux from Thyroid Cells. *The Journal of Clinical Endocrinology & Metabolism* **87**(7): 3356-3361

J. Zheng, W. Shen, D. Z. Z. He, K. B. Long, L. D. Madison and P. Dallos (2000). Prestin is the motor protein of cochlear outer hair cells. *Nature* **405**: 149-155

J. Zheng, G. Du, C. T. Anderson, J. P. Keller, A. Orem, P. Dallos, and M. Cheatham (2006). Analysis of the Oligomeric Structure of the Motor Protein Prestin. *J. Biol. Chem.* **281**(29): 19916-19924

J. Zheng, C. T. Anderson, K. K. Miller, M. Cheatham and P. Dallos (2009). Identifying components of the hair-cell interactome involved in cochlear amplification. *BMC Genomics* **10**: 127

A. S. Zolotarev, M. Unnikrishnan, B. E. Shmukler, J. S. Clark, D. H. Vandorpe, N. Grigorieff, E. J. Rubin, S. L. Alper (2008). Increased sulfate uptake by E. coli overexpressing the SLC26-related SulP protein Rv1739c from Mycobacterium tuberculosis. *Comparative Biochemistry and Physiology, Part A* **149**: 255-266

M. Zulauf and A. D'Arcy (1992). Light scattering of proteins as a criterion for crystallization. *J Cryst. Growth* **122**: 102-106

Acknowledgments

First of all I would like to thank my parents for supporting all my choices, including that of doing the PhD.

Then I thank Alessandro for the support, the helps, and the advice and for making the work days happier.

A special thanks to Roberto for giving me the opportunity to attend the PhD in his research laboratory. Furthermore I thank his group, especially Graziano, Elisa and Rosa, for their helps and advice, and Michele for the structural studies on mutant STAS domains.

Then I want to thank Frank and Prof. Dötsch for giving me the opportunity to spend six months of my PhD in their research group. I thank also their group, especially Davide, Sigrid, Birgit, Christian, Lei, Erik, Edith and Aisha.

Furthermore I thank the people with whom I collaborated: Massimo and Meri for NMR studies, Prof. Oliver for functional studies on mutant full-length transporters, Prof. Bergantino and Vanessa for their help during the cloning of the cyanobacterial transporters and Prof. Fendler for functional studies on the transporters produced by means of the CF expression system.

Finally I thank Prof. Oliver and Prof. Roberto Steiner for evaluating my thesis.

Ringraziamenti

Prima di tutto vorrei ringraziare i miei genitori per aver sostenuto ogni mia scelta, tra cui quella di fare il dottorato.

Ringrazio poi Alessandro per il sostegno, gli aiuti e i consigli e per aver reso i giorni di lavoro più allegri.

Un grazie particolare a Roberto per avermi dato l'opportunità di frequentare il dottorato nel suo laboratorio di ricerca. Ringrazio inoltre il suo gruppo, specialmente Graziano, Elisa e Rosa, per i loro aiuti e i consigli, e Michele per gli studi strutturali sui domini STAS mutati.

Voglio poi ringraziare Frank e il Prof. Dötsch per avermi dato l'opportunità di trascorrere sei mesi del mio dottorato nel loro gruppo di ricerca. Ringrazio anche il loro gruppo, specialmente Davide, Sigrid, Birgit, Christian, Lei, Erik, Edith e Aisha.

Ringrazio inoltre le persone con cui ho collaborato: Massimo e Meri per gli studi NMR, il Prof. Oliver per gli studi funzionali sui trasportatori interi mutati, la Prof. Bergantino e Vanessa per l'aiuto durante il clonaggio dei trasportatori di cianobatterio e il Prof. Fendler per gli studi funzionali sui trasportatori prodotti tramite il sistema di espressione CF.

Ringrazio infine il Prof. Oliver e il Prof. Roberto Steiner per aver valutato il mio lavoro di tesi.