



CATARINA DE JESUS MARTINS ALVES
Graduated in Cell and Molecular Biology

MULTITASK NBDS OF BACTERIAL ABC TYPE I IMPORTERS

CHARACTERIZATION OF PROTEIN-PROTEIN INTERACTIONS

Master in Biochemistry

NOVA University of Lisbon
December 2021

MULTITASK NBDS OF BACTERIAL ABC TYPE I IMPORTERS

CHARACTERIZATION OF PROTEIN-PROTEIN INTERACTIONS

CATARINA DE JESUS MARTINS ALVES

Graduated in Cell and Molecular Biology

Supervisor: Isabel Maria Godinho de Sá Nogueira,
Full Professor, NOVA University Lisbon

Co-supervisor: Lia Raquel Marques Godinho, Postdoctoral
Researcher, UCIBIO NOVA University Lisbon

Júri:

Presidente: Pedro António de Brito Tavares, Full Professor,
NOVA University Lisbon

Arguentes: Sofia Rocha Pauleta, Full Professor,
NOVA University Lisbon

MASTER IN BIOCHEMISTRY

NOVA University of Lisbon
December 2021

Multitask NBDs of Bacterial ABC type I importers: characterization of protein-protein interactions

Copyright © CATARINA DE JESUS MARTINS ALVES, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa.

A Faculdade de Ciências e Tecnologia e a Universidade NOVA de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

To my Grandparents

Acknowledgments

À minha professora e orientadora Isabel Sá Nogueira, quero agradecer por me ter recebido no Microbial Genetics Laboratory, por todo o apoio, orientação e conselhos prestados durante esta tese, para além da preocupação durante os confinamentos e isolamentos devido à Covid-19.

À Fundação para a Ciência e Tecnologia, agradeço o financiamento do projeto One4All_Bacterial ABC-type Multitask Energizers: one ATPase fits all? (PTDC / BIA-MIC / 30696/2017 – FCT_MCTES) e à UCIBIO e ao Departamento de Ciências da Vida (DCV) pelas instalações e ambiente.

À minha coorientadora, Lia Godinho, obrigada por todo o apoio, por toda a paciência, conselhos e disponibilidade para me explicares as minhas dúvidas. À Inês e à Joana, obrigada por toda a ajuda e bons momentos passados. Obrigada a todas pelo bom ambiente, pelas histórias e motivação. Quero também agradecer à vizinhança do Lab 333, pela companhia e boas horas de almoço e à Nicole pela boa disposição.

Às irmãs que a FCT me deu, Pipa, Rita, Maria, Alicia e Cat (M&P forever!). Sem vocês estes 5 anos nunca teriam sido a aventura que foram. Não vamos esquecer os dias em que apanhávamos a 158 e só parávamos de rir para lá de Alcântara (Pipa, Rita, obrigada por me sacrificarem enquanto se salvavam a vocês próprias!) sem conseguirmos respirar, as figuras (tristes), os concertos dados na máquina do café, as aulas de ginástica no edifício 4, os possíveis rituais que assustaram os nossos vizinhos na biblioteca ou as nossas viagens em que nos distraíamos a cantar e causávamos filas de trânsito. Obrigada por serem um apoio incondicional, por me serem a maior claqué tagarela (900 mensagens por ler é sempre bom...) e por me fazerem rir. Não podia ter pedido melhores pessoas na minha vida.

À minha maltinha do mestrado, Tininha, Nelson (ou será Nelsom?), Su e Rodrigo, por todas as horas passadas a jogar uno, a fingir que estudávamos ou a planear idas ao Starbucks (Tininha ainda estou à espera!). Vocês não só tornaram a minha estadia na FCT ainda mais fantástica como tornaram toda a situação das aulas online e da pandemia menos assustadora.

À minha família. Não tenho palavras para descrever o quão agradecida estou pela família que tenho. Mãe, Pai, nunca conseguirei agradecer todo o amor e apoio (mesmo quando já não me podiam ouvir) que me dão. Quero agradecer-vos por acreditarem sempre em mim, pelos conselhos e alegrias, por me ajudarem a concretizar os meus sonhos e por me acompanharem em todas as etapas da minha vida! Vocês fizeram de mim a pessoa que sou hoje e espero deixar-vos orgulhosos. Ao meu irmão e à Solange, obrigada por estarem sempre presentes para partilhar todos os momentos, vitórias e alegrias. Aos meus tios e aos meus primos pelo apoio e bons momentos que me proporcionaram ao longo dos anos. Por último, quero agradecer aos meus avós, que apesar de já não estarem aqui comigo sei que

estão a olhar por mim. Obrigada por me ajudarem a ir para a faculdade e a seguir os meus sonhos. Obrigada por todos os risos e memórias que me deram, vocês são os meus anjos da guarda.

“Science is just Magic that's been explained.”
(Stitchers)

Abstract

ATP-binding cassette (ABC) transporters are a family of membrane translocators ubiquitous to all domains of life. This supergroup is divided into importers and exporters, and both share a structure of two Nucleotide-Binding Domains (NBDs), responsible for the binding and hydrolysis of ATP and consequently powering the conformational changes in the two Transmembrane Domains (TMDs), that form the translocation pore. In the case of ABC type I importers, there is also a substrate-binding protein (SBP), responsible for the binding of substrate and its delivery to the TMDs. This subfamily of ABC transporters has been only identified in bacteria and plants, having an important part in pathogenesis and survival. Lately, several studies have shown the existence of a group of multitask ATPases, found in the Firmicutes phylum, capable of energizing several sugar transport systems. Intra- and inter-species exchangeability and exchangeability beyond the phylum were also found to be a characteristic of these multitask NBDs. To identify signature motifs associated with the multitasking ability of some NBDs, the AraNPQ-MsmX system of *Bacillus subtilis* was used as a model. A collection of mutant TMDs, AraP and AraQ was constructed by targeting several conserved residues in the EAA motifs and in the C-terminal tail of AraQ. Using the Bacterial Adenylate Cyclase Two-Hybrid (BACTH) system, the effect of these mutations on interaction with the NBD MsmX was quantified. Our results show that all targeted residues have different effects on MsmX-TMDs interaction, demonstrating the importance of these conserved motifs as key contact points with this ATPase. Moreover, the BACTH system was also used to characterize the interaction between previously identified multitask ATPases of the Firmicutes phylum and beyond it and the AraPQ translocators. Here, a correlation between NBD-TMD interaction measured in *E. coli* and the respective functionality of the complete transporter in *B. subtilis* was found, establishing the BACTH system as a powerful tool to study protein-protein interactions in bacterial ABC type I importers.

Keywords: ABC transporters, Multitask ATPases, AraNPQ-MsmX Importer, BACTH system, Firmicutes phylum

Resumo

Os transportadores de cassete de ligação de ATP (ABC) são uma família de transladores de membrana presentes em todos os domínios da vida. Este supergrupo está dividido em importadores e exportadores, e ambos possuem uma estrutura constituída por dois Domínios de ligação a Nucleótidos (NBD), responsáveis pela ligação e hidrólise do ATP e, consequentemente, energizam as mudanças conformacionais dos dois Domínios Transmembranares (TMD), que formam o poro de transporte. No caso dos importadores ABC tipo I, existe também uma proteína de ligação ao substrato (SBP), responsável pela ligação ao substrato e a sua entrega aos TMDs. Esta subfamília foi apenas identificada em bactérias e plantas, tendo um papel importante na sua patogénese e sobrevivência. Recentemente, vários estudos têm demonstrado a existência de um grupo de ATPases multitarefa, encontradas no filo dos Firmicutes, capazes de energizar diversos sistemas de transporte de açúcar. A permutabilidade intra e interespecie e a permutabilidade além do filo foi também encontrada nestas NBD multitarefas. De forma a identificar motivos típicos associados à capacidade multitarefa de alguns NBDs, o sistema AraNPQ-MsmX, do organismo modelo Gram-positivo *Bacillus subtilis*, foi usado como modelo. Foi construída uma coleção de mutantes AraPQ, tendo como alvo vários resíduos conservados nos motivos EAA e na cauda C-terminal do AraQ. O efeito destas mutações na interação com MsmX foi quantificado, usando o sistema Bacteriano de duplo híbrido baseado na enzima Adenilato Ciclase (BACTH). Os nossos resultados mostram que todos os resíduos selecionados têm diferentes efeitos na interação MsmX-TMDs, demonstrando a importância desses motivos conservados como pontos de contacto chave com esta ATPase. Além disto, o sistema BACTH foi também usado para caracterizar a interação entre várias ATPases multitarefa previamente identificadas do filo dos Firmicutes e outros e os transladores AraPQ. Uma correlação foi observada entre a interação NBD-TMD determinada em *E. coli* e a respetiva funcionalidade de um transportador completo em *B. subtilis*, estabelecendo o uso do sistema BACTH como uma ferramenta poderosa para estudar interações entre proteínas em importadores bacterianos ABC do tipo I.

Keywords: Transportadores ABC, ATPases multitarefa, Importador AraNPQ-MsmX, Sistema BACTH, Filo Firmicutes

Contents

Acknowledgments.....	v
Abstract.....	ix
Resumo	xi
Contents	xiii
Figures Index	xv
Tables Index.....	xvii
Abbreviations, Symbols and Notations.....	xix
Chapter 1 General Introduction.....	1
1.1. Transport Systems.....	3
1.2. ABC transporters	5
1.2.1. Transmembrane Domain (TMD)	7
1.2.2. Nucleotide-Binding Domain (NBD).....	9
1.2.3. Substrate-Binding Protein (SBP)	11
1.2.4. ABC type I importers: MalEFGK ₂ importer	12
1.2.4.1. Transport mechanism of ABC type I importers.....	15
1.3. Multitask Nucleotide-Binding Domains	17
1.3.1. MsmX ATPase and AraNPQ type I Importer	19
1.3.2. Other multitask ATPases	21
1.4. Study of protein-protein interactions	23
1.4.1. Methods of study.....	24
1.4.1.1. Bacterial Adenylate Cyclase Two-Hybrid System (BACTH)	25
1.5. Scope of Thesis.....	28
Chapter 2 Materials and Methods	29
2.1. Bioinformatic and statistical analysis	31
2.2. Competent cells	31
2.3. Phenolic extraction of chromosomal DNA	31
2.4. DNA manipulation	32
2.5. Plasmid construction for protein-protein interaction assays.....	32
2.6. Alkaline lysis for plasmid extraction.....	37
2.7. Construction of <i>E. coli</i> BTH 101 for qualitative and quantitative protein-protein interaction	37
2.8. Qualitative analysis of protein-protein interaction	37
2.9. Quantitative analysis of protein-protein interaction: β -Galactosidase assays	38
2.10. Protein extracts and Western Blot analysis	39

Chapter 3 Results and Discussion.....	41
3.1. Targeting key residues in TMDs AraPQ for interaction studies using BACTH system	43
3.1.1. Qualitative analysis of MsmX and AraPQ interaction	46
3.1.2. Quantitative analysis of MsmX and AraPQ interaction	48
3.1.3. Detection of <i>in vivo</i> accumulation of wild type and mutant TMDs by Western blot analysis.....	54
3.2. <i>In vivo</i> analysis of of MsmX homologs and AraPQ interaction using the BACTH system....	58
3.2.1. Qualitative analysis of MsmX homologs and AraPQ interaction.....	59
3.2.2. Quantitative analysis of MsmX homologs and AraPQ interaction.....	61
3.2.3. Detection of <i>in vivo</i> accumulation of multitask ATPases by Western blot analysis.	67
Chapter 4 Conclusions and Future perspectives	69
Chapter 5 References	73
Chapter 6 Appendices.....	85

Figures Index

Figure 1. Types of membrane transport.	3
Figure 2. Types of active transport.	4
Figure 3. Glucose transport by the bacterial PTS system.....	5
Figure 4. Types and Domain organization of ABC transporters.....	6
Figure 5. Schematic representation of NBDs and TMDs.....	8
Figure 6. Schematic representation of the relative positions of sequence motifs in NBDs.....	9
Figure 7. Spatial representation of motifs and domains of an NBD.....	10
Figure 8. Different arrangements of substrate-binding proteins.....	11
Figure 9. Structure of the maltose transporter in a catalytic intermediate conformation.....	13
Figure 10. The interface between TMDs and the MalK dimer.	14
Figure 11. Schematic representation of the Alternative access model.....	15
Figure 12. Second schematic representation of the Alternative access model.....	16
Figure 13. Interspecies exchangeability in Firmicutes phylum and beyond.	18
Figure 14. All carbohydrate transport systems energized by MsmX ATPase.	20
Figure 15. MsmX and MalK electrostatic potential at NBD-TMD interface.	21
Figure 16. Adenylate Cyclase activity.	26
Figure 17. Bacterial Two Hybrid system.	26
Figure 18. Schematic representation of two strategies of mutagenesis PCR.	34
Figure 19. AraP TMD from AraNPQ-MsmX system.....	43
Figure 20. Construction of a chimeric AraP protein.....	44
Figure 21. AraQ TMD from the AraNPQ-MsmX system.	44
Figure 22. Principle of BACTH System in qualitative assay using X-Gal plates as an indicator media.....	46
Figure 23. Principle of BACTH System in quantitative assays, using ONPG as substrate for β -galactosidase	49
Figure 24. MsmX+Mutant AraP interaction assay.	50
Figure 25. MsmX+Mutant AraQ interaction assay.....	52
Figure 26. MsmX_T18 detection by Western Blot analysis, from <i>E. coli</i> BTH101 growth.....	54
Figure 27. MsmX_T18 detection by Western Blot analysis, from <i>E. coli</i> BTH101 growth.....	55
Figure 28. MsmX_T18 detection by Western Blot analysis, from <i>E. coli</i> BTH101 growth.....	56
Figure 29. TMD_T18 detection by Western Blot analysis, from <i>E. coli</i> BTH101 growth.....	57
Figure 30. Putative multitask ATPases_T18 + AraP_T25 β -galactosidase assays..	63
Figure 31. MalK_T18 + AraP_T25 variants β -galactosidase assays..	64
Figure 32. Putative multitask ATPases_T18 + T25_AraQ β -galactosidase assays..	65

Figure 33. Schematic representation of the correlation found between multitask NBD - TMDs (β -Galactosidase activity (U/mg) and its doubling time (min)).....	67
Figure 34. Multitask ATPases_T18 detection by Western Blot analysis, from <i>E. coli</i> BTH101 growth	68

Tables Index

Table 1. List of used and constructed plasmids mentioned in this work for protein-protein interaction studies using BACTH system.....	35
Table 2. List of all oligonucleotides used for in this work for determining protein-protein interaction using BACTH system.....	36
Table 3. List of all <i>E. coli</i> BTH101 co-transformations constructed for protein-protein interaction studies using BACTH system.....	38
Table 4. Qualitative protein-protein interaction assay.	47
Table 5. Quantitative protein-protein interaction assays.	50
Table 6. List of MsmX homologs used in this work.	59
Table 7. Qualitative protein-protein interaction assay.	60
Table 8. Quantitative protein-protein interaction assays.....	62

Abbreviations, Symbols and Notations

aa – Amino acid	<i>lacZ</i> – β -galactosidase encoding gene
ABC – ATP-binding cassette	LB – Luria-Bertani medium
Abs - Absorbance	MCS – Multiple cloning site
ADP – Adenosine diphosphate	NBD – Nucleotide-binding domain
ATP – Adenosine triphosphate	O.D. – Optical Density
ATPase – Adenosine triphosphatase	ONPG – O-Nitrophenyl- β -galactoside
BACTH – Bacterial Adenylate Cyclase Two-Hybrid Essay	PCR – Polymerase Chain Reaction
BLAST – Basic Local Alignment Search Tool	Pi – Inorganic phosphate
bp – Base pair	PMSF – Phenylmethylsulphonyl fluoride
cAMP – Cyclic Adenosine monophosphate	PTS – Phosphotransferase system
CAP – Catabolite activator protein	RBS – Ribosome binding site
CRD – C-terminal regulatory domain	SBP – Substrate-binding protein
CUT1 – Carbohydrate uptake transporter 1	SDS – Sodium dodecyl sulphate
<i>cya</i> – Adenylate cyclase encoding gene	SDS-PAGE – Sodium dodecyl sulphate polyacrylamide gel electrophoresis
dNTP – Deoxyribonucleic acid	TAE – Tris-Acetate-EDTA
DNA – Deoxyribonucleic acid	TE – Tris-EDTA
ECF – Energy coupling factor	TMD – Transmembrane Domain
EDTA – Ethylenediaminetetraacetic acid	X-Gal – 5-bromo-4-chloro-3-indolyl- β -D-galacto-pyranoside
ITPG – Isopropyl- β -D-galactopyranoside	
<i>lac</i> operon – Lactose operon	

Amino acids – three and one letter code

Amino Acid	One letter code	Three letter code
Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartic Acid	D	Asp
Cysteine	C	Cys
Glutamic Acid	E	Glu
Glutamine	Q	Gln
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val

Nucleobases – one letter code

Base	Letter
Adenine	A
Cytosine	C
Guanine	G
Thymine	T

Chapter 1

General Introduction

1. General Introduction

1.1. Transport Systems

Since the beginning of life on earth, cells rely on a variety of processes to survive. One of these processes is the transport of matter, which occurs from the atomic scale to the macroscale (Kojic *et al.* 2011; Stillwell, 2016 and references therein). Since cells are separated from their environment through physical barriers, such as lipid membrane bilayers and peptidoglycan walls, regulation and selective transport of molecules are necessary for numerous purposes, like nutrient acquisition, metabolites excretion and regulatory functions (Stillwell, 2016 and references therein). The membrane bilayer has the capacity to be permeable to some simple molecules and impermeable to other more complex molecules, such as sugars, and for that reason, the membrane bilayer is classified as semipermeable. This characteristic allowed cells to develop a series of mechanisms of controlled transmembrane transport, which are usually performed by specific polypeptides coupled with the membrane and controlled by the interaction between macromolecules. As seen in Figure 1, these mechanisms are divided into passive transport and active transport (Higgins, 1992; Stillwell, 2016).

Passive transport is the simplest process of solute translocation. It relies on the energy of the solute's electrochemical gradient as the only energy supply and its consequence is the solute reaching equilibrium (Sanno *et al.* 1968; Stillwell, 2016). This effortless process of transport can be divided into osmosis, the diffusion of water across a semipermeable membrane, simple passive diffusion and

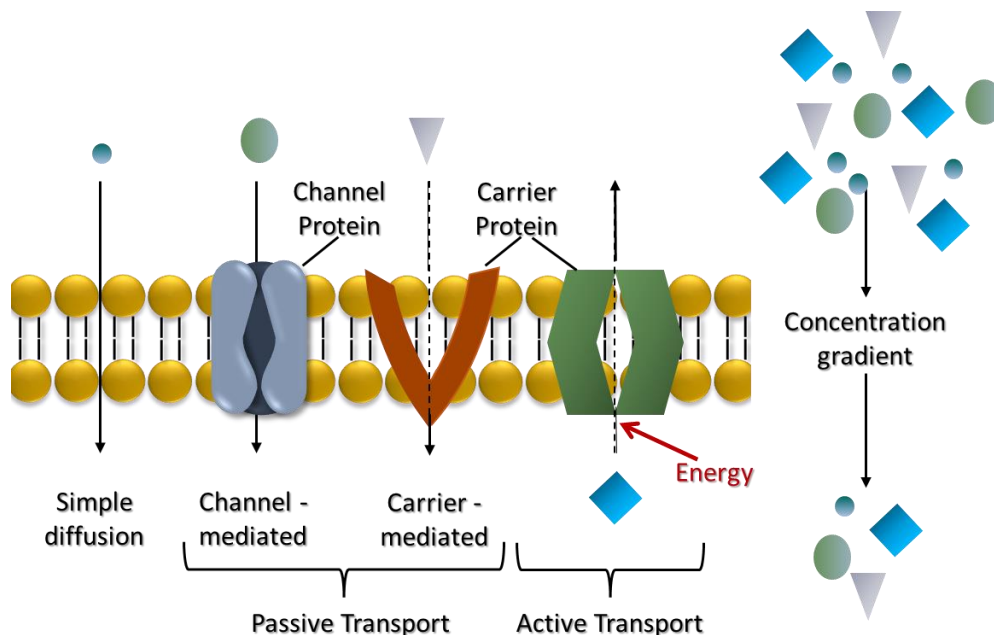


Figure 1. Types of membrane transport. Simple passive diffusion occurs according to gradient concentration, with the solute crossing the membrane without any intermediate protein. In passive transport, the solute is transported across a membrane, through channels or carrier proteins, according to gradient concentration. In active transport, solute is translocated against gradient concentration, requiring an energy source. Channel protein is represented in violet, carrier proteins are represented as orange and green and solute molecules are depicted as blue and green circles, triangles, and squares. Directions of translocation are depicted as black arrow. Adapted from Stillwell, 2016.

facilitated diffusion (Stillwell, 2016). In simple passive diffusion, a substance is translocated across the bilayer anywhere by diffusing through the membrane, while in facilitated passive diffusion, the molecules cross over the bilayer at specific locations, being aided by protein facilitators (Sanno *et al.*, 1968; Stillwell, 2016).

Active transport constitutes a more complex mechanism of solute translocation. This type of transport requires energy, usually adenosine triphosphate (ATP), to move substrate against its electrochemical gradient, resulting in a solute accumulation in one side of the bilayer and a chemical and electrical gradient far from the equilibrium (Scheer, 1958). Figure 2 shows the subdivisions of active transport, namely the division between uniport and cotransport, which can be subdivided into symport and antiport (Scheer, 1958; Stock & Roseman, 1971; West & Mitchell, 1974; Stillwell, 2016). The primary active transport also known as uniport allows the direct passage of substrate across the lipid bilayer in opposition to its electrochemical gradient by utilizing ATP as an energy source. This type of active transport is very common in bacteria, in particular the ATP-binding cassette (ABC) transporters. The co-transport or secondary active transport comprises two functions, an electrochemical gradient, which is formed by the movement of a solute across the membrane by facilitated diffusion, down to its electrochemical gradient, and then linked to the passage in the same direction of another substrate, called symport, or in the opposite direction, called antiport, against its gradient (Crane, 1962; West & Mitchell, 1974; Stillwell, 2016).

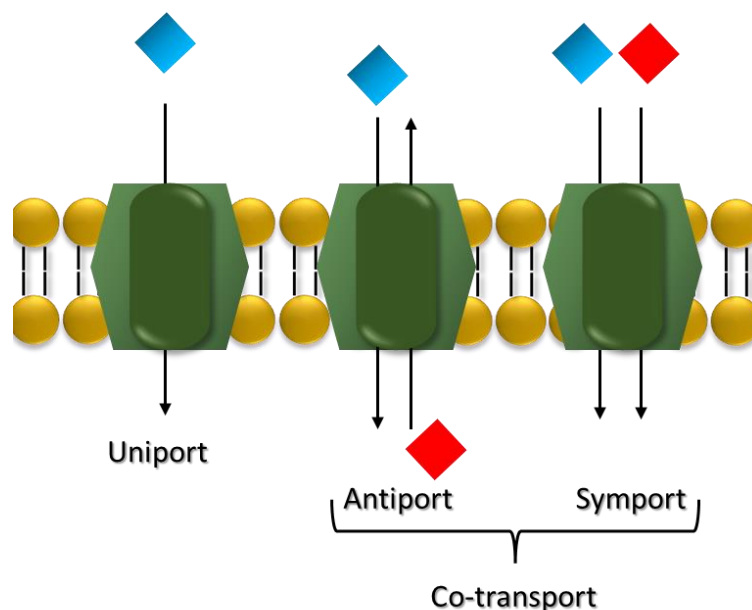


Figure 2. Types of active transport. Uniport translocate one molecule against its electrochemical gradient, while in cotransport, the translocation of one molecule in opposition to its electrochemical gradient is coupled to the translocation of another molecule by facilitated diffusion. In symport, molecules are transported in the same direction and in antiport, molecules are translocated in opposing directions. Carrier proteins are represented in green, directions of transportation are represented as black arrows and solute molecules are depicted as red and blue squares. Adapted from Stillwell, 2016.

Another type of membrane transport is group translocation, which alters the substrate to another product in the translocation process. The best-known example is depicted in Figure 3, the

phosphotransferase system (PTS), that allows the uptake of extracellular sugars in bacteria using the energy of phosphoenolpyruvate (PEP) (Werner Kundig *et al.*, 1964; W. Kundig & Roseman, 1971; W. Kundig, 1974; Stillwell, 2016; Jeckelmann & Erni, 2019).

This dissertation will focus on a family of active translocators, known as ATP-binding cassette (ABC) transporters, specifically ABC type I importers in charge for the uptake of sugars in Bacteria.

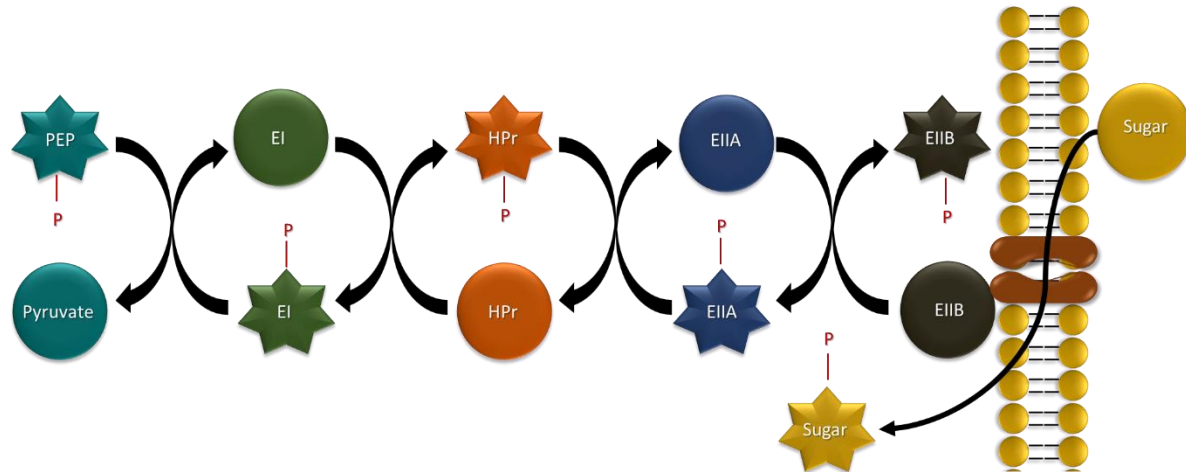


Figure 3. Glucose transport by the bacterial PTS system. PEP (cyan blue) is a high energy phosphorylated compound that provides a phosphate group (-P, red) to enzyme-I (EI, green). The heat-stable protein (HPr, orange) carries the -P from EI to EIIB (blue), which transports it to the high-energy group to EIIB (Brown). This protein receives glucose from EIIC (light brown) and phosphorylates it into glucose-6-phosphate. Adapted from Stillwell, 2016

1.2. ABC transporters

ATP-binding cassette (ABC) transporters are a family of primary active membrane translocators that has been receiving a lot of interest over the years. This family participates in numerous processes in prokaryotes and eukaryotes, namely in the transport of molecules, secretion, regulation, antigen presentation, antibiotic resistance and in clinical issues (Hyde *et al.*, 1990; Anderson *et al.*, 1991; Higgins, 1992; Davidson *et al.*, 2008; Mächtel *et al.*, 2019). Normally, these translocators are associated with the lipidic membrane, but there is a minor group of ABC proteins related with complexes, which participate in DNA repair, translation, and other genetic processes (Higgins, 1992; Swier *et al.* 2016; Gouridis *et al.*, 2019). This superfamily of primary active translocators is spread among all domains of life – Eukarya, Archaea and Bacteria – and can be divided into exporters and importers (Anderson *et al.*, 1991; Albers *et al.*, 1999; Ter Beek *et al.*, 2014; Swier *et al.*, 2016). Importers translocate molecules from the outside of the cell (or trans-side, in case of eukaryotic cells) to the inside of the cell (or cis side, in case of eukaryotic cells) where ATP molecules are bound and hydrolysed, while the exporters move the substrate from the inside to the outside of the cell (Saurin *et al.*, 1999; Rice *et al.*, 2014; Ter Beek *et al.*, 2014). This movement is possible because these transporters can couple hydrolysis of a phosphate bond between the γ - and the β -phosphate of the ATP molecule to the transport of a solute (Ter Beek *et al.*, 2014).

These two groups of ABC transporters have a core structure of two transmembrane domains (TMDs), that form the translocation pore in the lipid bilayer and two cytoplasmic nucleotide-binding domains (NBDs), accountable for the binding and hydrolysis of ATP and thus energizing the transport (Figure 4A). These two domains can have different levels of fusion, which means they can be expressed as individual proteins or fused together, as shown in Figure 4B (Theodoulou & Kerr, 2015). NBDs are also the hallmark of the family, displaying a high conserved structure and sequence among all ABC transporters. To this date, it is known that while exporters exist in all domains of life, importers are only present in bacterial and plant cells (Higgins, 1992; Higgins & Linton, 2004; Ter Beek *et al.*, 2014; Theodoulou & Kerr, 2015; Swier *et al.*, 2016).

In terms of structure and transport mechanism, ABC importers are subdivided into three groups, type I, II and III (also named Energy Coupling Factor, ECF). In bacteria, type I and II importers are

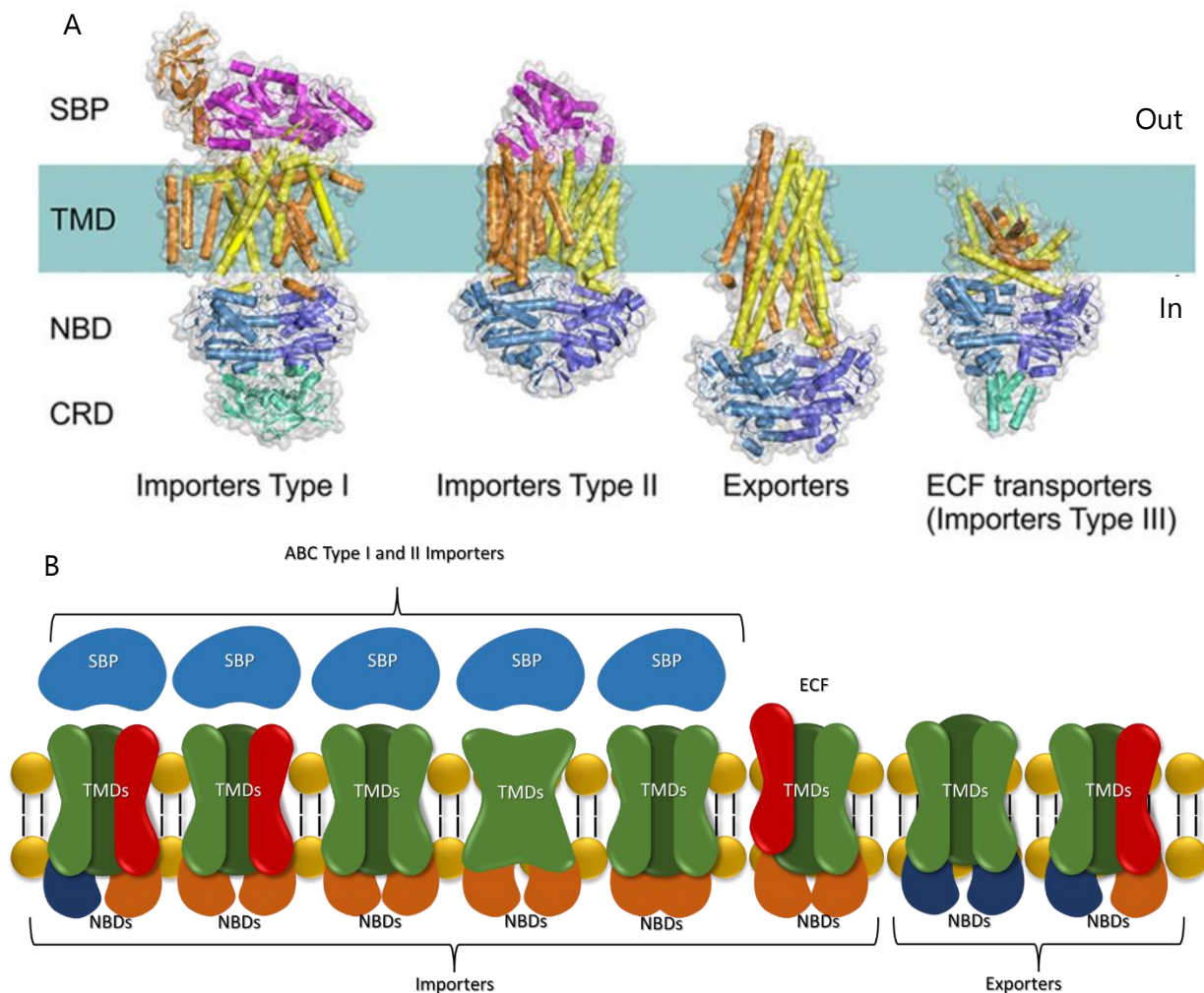


Figure 4. Types and Domain organization of ABC transporters. A) Four types of ABC transporters, sharing an architecture of two NBDs (blue and cyan blue) and two TMDs (yellow and orange). In type I and type II transporters, is also present a substrate-binding protein (SBPs; magenta). An additional domain, C-terminal domain (CRD, green) have regulatory functions. ECF, Energy coupling factor. From Ter Beek *et al.*, 2014. B) Domain organization in several ABC transporters from different organisms from Bacteria. Homodimers are represented with identical colour (green in case of TMDs and blue or orange in case of NBDs) while heterodimers are represented with different colours (green and red in case of TMDs and blue and orange in case of NBDs). SBP is represented in light blue. Different fusion levels are also depicted as different structures. Adapted from Theodoulou & Kerr, 2015.

quite similar, since they depend on a substrate-binding protein (SBP), responsible for solute binding and its delivery to transmembrane domains. These two categories also share the same nucleotide-binding domain fold (Ter Beek *et al.*, 2014; Tanaka *et al.*, 2018).

Regarding the solute transport, all types of ABC transporters translocate different substrates. Type I is responsible for translocation of small hydrophilic nutrients, such as amino acids, sugars or ions, while type II is accountable for the import of metal chelates, namely, B12, heme and metals. ECF transporters or type III is accountable for the translocation of vitamins, metal ions; exporters usually translocate hydrophobic compounds, such as antibiotics or drugs (Locker, 2002; Oldham *et al.*, 2007; Eitinger *et al.*, 2011; Ter Beek *et al.*, 2014; Swier *et al.*, 2016).

In bacterial genomes, elements of ABC transporters, TMDs, NBDs and SBPs (type I and II transporters), are usually encoded by genes clustered together or by genes in operons, allowing a coordinated regulation of their expression (Quentin *et al.*, 1999; Ter Beek *et al.*, 2014). In the genome of *Bacillus subtilis*, the Gram-positive model organism, genes encoding for proteins of ABC translocators are the most recurrent group of polypeptides coding genes (Quentin *et al.*, 1999).

1.2.1. Transmembrane Domain (TMD)

In all four types of ABC transporters the translocation pathway is comprised of two transmembrane domains, which are alternately available from the outside and inside of the lipid bilayer during the translocation process, allowing uptake or excretion of substrate. In contrast with the NBDs TMDs have low sequence similarity (Rice *et al.*, 2014; Ter Beek *et al.*, 2014; Swier *et al.*, 2016).

In type I importers, regardless of low sequence homology, transmembrane domains are typically homodimers or structurally similar heterodimers, having five transmembrane helices per domain. In other TMDs of the same group, eight transmembrane helices were identified and in others, an additional N-terminal helix is found to wrap around the helices of the opposing TMD and intertwine them, comprising twelve transmembrane helices (Ter Beek *et al.*, 2014). The translocation path is placed at the interface between two TMDs. In the archetype system of type I translocators, the maltose/maltodextrins importer MalEFGK₂ from *Escherichia coli*, the TMD MalF has a second and third binding sites at the centre of the bilayer, besides the binding site in their SBP, MalE (Oldham *et al.*, 2007; Oldham *et al.*, 2013; Ter Beek *et al.*, 2014; Abreu *et al.*, 2021).

Conformational modifications in TMDs, which lead to the alternating access hypothesis for the translocation mechanism, are induced by conformational changes in NBDs, through a coupling helix. This helix is a small α -helix present in an intracellular extension, usually found between transmembrane helix 3 and transmembrane helix 4 of the TMDs that is inserted into a cleft present in an ATPase monomer, connecting NBDs to TMDs (Figure 5). A very common and conserved sequence in coupling helices is the EAA motif, bearing the conserved sequence EAA-X(3)-G-X(9)-I-X-LP, constituting a site of interaction with the ATPase in the Q-loop region, situated at the interface between a RecA-like

subdomain and an α -helical subdomain (Dassa & Hofnung, 1985; Quentin *et al.*, 1999; Steinke *et al.*, 2001; Ter Beek *et al.*, 2014).

Another important location for TMDs-NBDs interaction, is a C-terminal portion of a TMD. This C-terminal segment appears to be partially inserted between ATPases dimers, interacting in the Q-loop region (Oldham *et al.*, 2007; Ter Beek *et al.*, 2014).

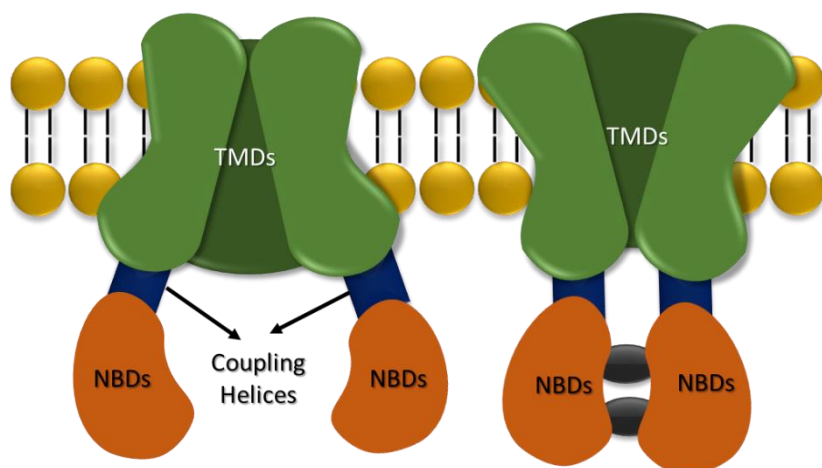


Figure 5. Schematic representation of TMDs and NBDs. Representation of an ABC transporter in two different configurations: On the left is the resting state, an inward open conformation, when substrate is not bound and on the right is the transition state, the outward open conformation, when substrate is bound. NBDs are represented in orange, while the TMDs are coloured in green. TMDs are attached to the NBDs by a coupling helices (coloured in blue) present in the cytoplasmic loops of the TMDs. ATP is represented in grey. Adapted from Ter Beek *et al.*, 2014.

In type II importers, for instance the BtuC₂D₂F, the vitamin B12 transporter, the translocation pore is formed by two identical TMDs having an opposite orientation in the membrane, resulting in ten transmembrane helices. They have analogous helical packing, and helices from an individual TMD are tightly packed together (Rice *et al.*, 2014; Ter Beek *et al.*, 2014). In this class of importers, the coupling helix is present in transmembrane helices 6 and 7. Additional interaction is also detected in type II importers, with an helical segment neighbouring the coupling helix responsible for interactions with NBDs (Locker, 2002; Rice *et al.*, 2014; Swier *et al.*, 2016).

In type III importers, also known as ECF-type ABC importers, the pore is composed of two structurally and functionally unrelated TMDs. The T-component or EcfT subunit is a TMD composed of five transmembrane helices, but some are predicted to have four to eight and is responsible for the interaction with NBDs through two long helices in an individual intracellular region. A second TMD binds the transported substrate with high affinity and is designated S-component. This hydrophobic integral membrane has six transmembrane helices and, in some cases, an additional N-terminal helix. In contrast with type I and type II translocators, the translocation of substrate is confined to the S-component and not present at the interface of TMDs (Eitinger *et al.*, 2011; Rice *et al.*, 2014; Santos *et al.*, 2018; Ter Beek *et al.*, 2014).

The translocation pore of ABC exporters is formed by two identical or structurally similar TMDs, each one consisting of six transmembrane helices. A key difference between importers and exporters is the distance between the membrane and the NBD: in importers the two are very close whereas in

exporters they are farther apart. The crystallized structures of exporters obtained so far show NBDs attached to TMDs and two helices of each TMD spanning to the other TMD (Ter Beek *et al.*, 2014). Exporters have the most extensive additional interaction areas having two coupling helices. One is in the cytoplasmic loop between transmembrane helices 4 and 5 and the other is present between transmembrane helices 2 and 3, interacting with the ATPase regions that confine the nucleotide adenine ring (Ter Beek *et al.*, 2014).

1.2.2. Nucleotide-Binding Domain (NBD)

In every functional ABC transporter, the Nucleotide-binding domain (NBD), also called ATPase, is present as a homo- or heterodimer having a high sequence similarity with other proteins from this domain. These NBDs are part of a diverse group of P-loop NTPases, whose catalysis is dependent on magnesium ions. These polypeptides comprises two subdomains: a large RecA-like domain, also present in different proteins of P-loop ATPase family, and an α -helical domain (Higgins, 1992; Vetter & Wittinghofer, 1999; Ter Beek *et al.*, 2014; Swier *et al.*, 2016). Each NBD dimer has one ATP binding site facing each subunit and it is described that ATP binding and hydrolysis is made by highly preserved amino acids present in motifs distinctive of this superfamily of ABC proteins. As seen in Figure 6, these motifs comprise the Walker A and B regions, the A, Q, H and D loops and the ABC motif (Ter Beek *et al.*, 2014; Swier *et al.*, 2016; Leisico *et al.*, 2020).

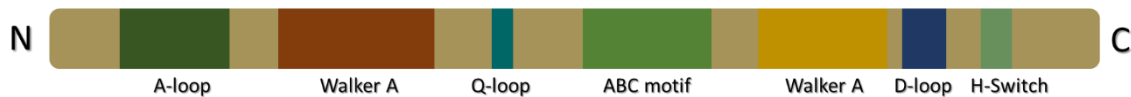


Figure 6. Schematic representation of the relative positions of sequence motifs in NBDs. N and C designate the N-terminus and C-terminus, respectively. Sequence motifs in NBDs are depicted as colourful boxes. From N-terminus to C-terminus: A-loop, green; Walker A – brown; Q-loop – cyan blue; ABC motif – light green; Walker A – yellow; D-loop – dark blue and H-switch – cyan green. Adapted from Ter Beek *et al.*, 2014.

Close to the N-terminal of this domain, is located the A-loop. This motif comprises an aromatic residue that facilitates the placing of ATP by packing with the adenine ring (Ter Beek *et al.*, 2014). Next to this motif is located a phosphate-binding loop, known as Walker A motif or P-loop (GXXGXXGK(S/T)). It includes a highly conserved lysine residue, whose ϵ -amino group and backbone amide nitrogen create an interaction web with β - and γ -phosphate of ATP (Ter Beek *et al.*, 2014). The Q-loop exists next to the Walker A region and is comprised of an eight residues sequence with a highly conserved glutamine at the N-terminal. This region is also placed at the interface between the α -helical and the RecA-like subdomains. During the hydrolysis cycle, the movement of a consensus glutamine residue in and out of the active site forms the active site when Mg-ATP is bound, interrupting it once ATP is hydrolysed which is caused by conformational changes in this loop. This motif is also an important interaction site with the coupling helices of TMDs (Figure 7) (Ter Beek *et al.*, 2014). The

hallmark of NBDs of ABC transporters is the LSGGQ motif or ABC signature motif. This motif appears as an α -helical subdomain and is present at the N-terminal of an elongated helix that manages positive charges of a helical dipole toward the γ -phosphate of the ATP molecule. To help to coordinate the magnesium ion, there is a conserved aspartate residue in Walker B motif ($\Phi\Phi\Phi\Phi$ DE, where Φ represents hydrophobic amino acids). This motif also has an acidic residue that polarizes the attacking water molecules. Following the Walker B region, there's the D-loop (SALD). When integrated on the assembled transporter, the D-loops of the two monomers go alongside each other and variations in their structure alters the architecture of the catalytic site and helps in the arrangement of the ATP hydrolysis site. The switch region or H-loop is placed near the C-terminal of the NBD and is known by its high consensus histidine residue, responsible for creating a hinge between an α -helix and a β -strand. This residue interacts in the D-loop with a conserved aspartate, in Walker B with a glutamate residue and with the γ -phosphate of the ATP molecule and facilitates the placement of the general base, the attacking water, and the magnesium ion (Ter Beek *et al.*, 2014). In exporters, an extra motif, known as the X-loop (TEVGERG), appears before the ABC motif and contacts with the coupling helices of both TMDs (Figure 7) (Higgins *et al.*, 1986; Schneider & Hunke, 1998; Daus *et al.*, 2007; Dawson *et al.*, 2007; Mächtel *et al.*, 2019).

In terms of functional conformation, ATPases in ABC transporters can adopt two forms, a closed conformation, when they are tightly packed against each other, or an open conformation, when they are dissociated. The closed conformation is adopted upon ATP binding since each ATP molecule contacts with both ATPases, causing the α -helical and the RecA-like subdomains within each NBD monomer to spin toward each other, allowing the hydrolysis of the ATP molecule. Once these reactions are completed, delivery of ADP and Pi allows NBDs to withdraw, causing a departure of the two

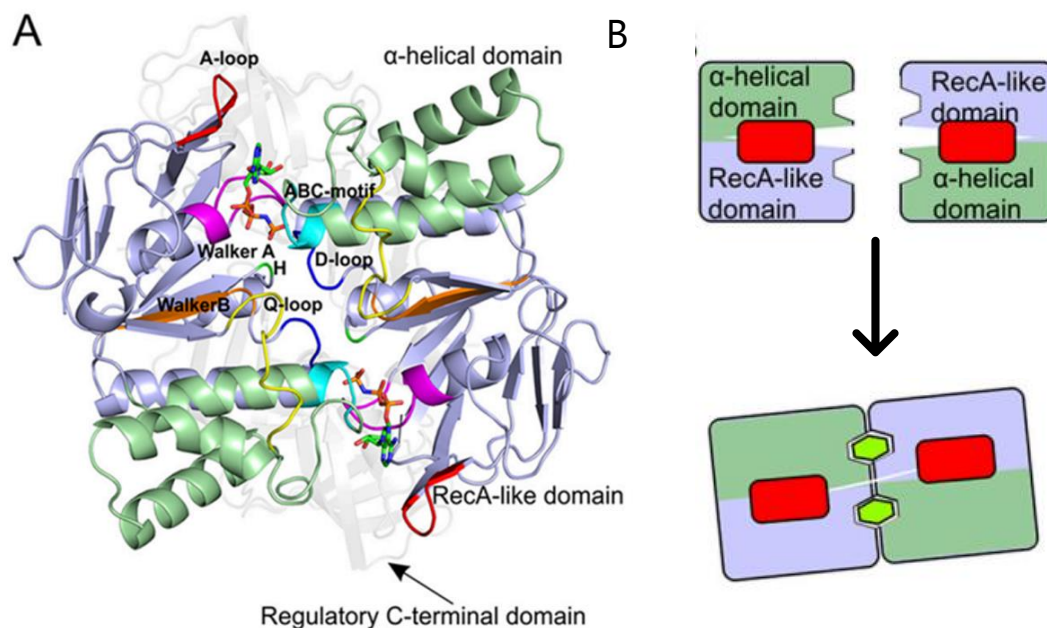


Figure 7. Spatial representation of motifs and domains of an NBD. A) Different conserved motifs of MalK dimer of an ABC transporter. From Ter Beek *et al.*, 2014. B) Top view of MalK with the coupling helices (coloured in red) and ATP molecules (bright green) represented. Adapted from Ter Beek *et al.*, 2014.

subdomains. The chemical energy released by ATP hydrolysis is converted into conformational energy, which will be conducted to the TMDs, allowing the alternating access of the translocation path to the two sides of the bilayer (Ter Beek *et al.*, 2014).

1.2.3. Substrate-Binding Protein (SBP)

In bacteria, type I and II importers require a high-affinity solute-binding protein (or a substrate-binding protein) for a fully operational uptake system. SBPs allow specificity and higher affinity for substances to translocation systems. In Gram-negative bacteria, for instance *E. coli*, SBPs are free and soluble in the periplasm, but in Gram-positive bacteria, like the model organism *B. subtilis*, SBPs are attached to the lipid bilayer through an N-terminal lipid extension (or a transmembrane peptide, in the case of Archaea) or bonded to the membrane (Figure 8) (Kellermann & Szmecman, 1974; Gilson *et al.*, 1988; Spurlino *et al.*, 1991; Albers *et al.*, 1999). In this case, it has been reported one or two SBPs per complex, but these proteins can also be fused in tandem, attached to TMDs, making a total of six SBPs per functional transporter (Poolman & van der Heide, 2002 and reference therein; Quentin *et al.*, 1999; Swier *et al.*, 2016). Various studies point out that when nutrients become low, cells express more SBPs or in some cases, related SBPs can interact with the same transmembrane proteins, allowing the translocation of different substrates in the same transporter. This suggests that growth conditions are correlated with the expression of these proteins (Swier *et al.*, 2016 and reference therein). In terms of stoichiometry, SBPs in type I importers are in excess over its transmembrane domains, allowing an

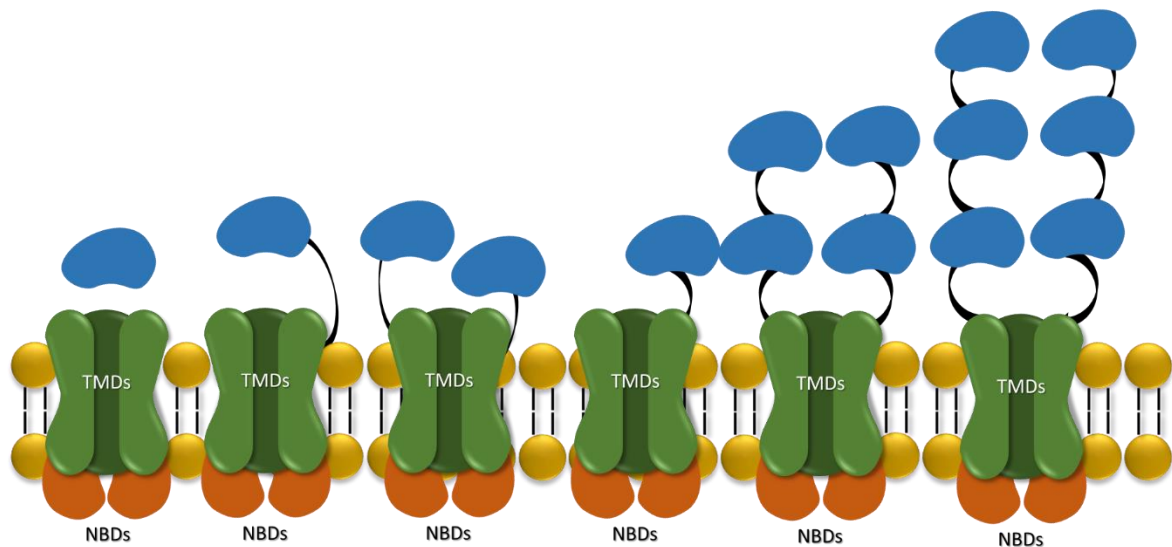


Figure 8. Different arrangements of substrate-binding proteins. In Gram-negative bacteria, SBP can be free in the periplasm, but in Gram-positive bacteria are fused to the lipid bilayer and TMDs. Here, SBPs may exist only one or two by complex or fused in tandem. TMDs are depicted in green and associated SBPs are shown in blue. NBDs are coloured in orange and SBP-TMD linker is depicted as black. Adapted from Swier *et al.*, 2016.

efficient uptake of substrates, while in type II translocators, SBPs seem to be stoichiometric with their correspondent TMDs (Swier *et al.*, 2016 and references therein).

SBPs are produced with a precursor signal peptide, which is cut by a lipoprotein signal peptidase II during the transport through the lipid bilayer (Quentin *et al.*, 1999; Ter Beek *et al.*, 2014). Regardless of if SBPs are present in the periplasm, attached to the membrane bilayer or linked to TMDs, the substrate-binding mechanism is similar. The substrate-binding protein is composed of two lobes linked by a linker and once the ligand gets attached to the substrate-binding site, the two lobes close and engulf the substrate. This mechanism is called Venus's Flytrap (Quioco & Ledvina, 1996; Swier *et al.*, 2016). Conformation of the substrate-binding protein has been resolved in various conformations, such as open-unliganded (O), close-unliganded (C), open-liganded (OL) and closed-liganded (CL), being the CL conformation the structure that probably interacts with the respective TMDs. Once ATP binds and induces conformational alterations in NBDs, these changes are transmitted to SBPs via TMDs, allowing the closed-liganded form to open, releasing the substrate to the translocation pore (Swier *et al.*, 2016 and references therein).

SBPs responsible for binding different substrates present low sequence similarity, although their architecture is described as highly conserved. According to their structural conformation these proteins are subdivided into six clusters (A to F). Type I importers use SBP from cluster D, characterised by two short hinges between two lobes, while the type II importers use exclusively SBP from cluster A, described as an individual rigid α -helix that links the two domains. It is likely that the use of SBPs from several groups is associated with the use of diverse TMDs folds (Ter Beek *et al.*, 2014; Swier *et al.*, 2016 and references therein;).

1.2.4. ABC type I importers: MalEFGK₂ importer

The best well-characterized ABC type I importer and considered the prototype of these transporters is the maltose/maltodextrin translocator MalEFGK₂, from *E. coli*. This translocator, allows the uptake of maltooligosaccharides up to seven glucose units, is comprised of a periplasmic maltose-binding protein, MalE, two transmembrane domains, MalF and MalG, present in the genome as single copies and is energized by an homodimer, two MalK ATPases (Kellermann & Szmelcman, 1974; Shuman & Silhavy, 1981; Gilson *et al.*, 1982; Dassa, 1990; Oldham *et al.*, 2007; Ter Beek *et al.*, 2014; Mächtel *et al.*, 2019; Narducci *et al.*, 2019; Abreu *et al.*, 2021).

Oldham *et al.* solved the structure of the complete maltose importer in a catalytic intermediate form (Figure 9). Here, TMDs MalF and MalG, have been shown to have low sequence identity (only 13%) and different lengths, 514 and 296 residues, respectively. While MalF is composed of eight transmembrane helices, MalG is comprised by six transmembrane helices and its C-terminal segment is partially introduced into MalK dimer, in the cytoplasmic side. Membrane-bound MalK exhibits a

very high structure similarity to isolated MalK, being the only significant difference the Q-loop region where these NBDs contact with the TMDs (Oldham *et al.*, 2007). This structure revealed two ATP molecules attached at the interface of the MalK dimer, contacting with the LSGGQ motif of one monomer and with residues of the Walker A and B from the opposite monomer. In this state a maltose molecule has already been translocated from MalE into the transmembrane domains, suggesting a stimulation of ATP hydrolysis by the system's SBP, MBP (Maltose-binding protein) by activating a conformational modification that causes the closure of NBDs (Oldham *et al.*, 2007).

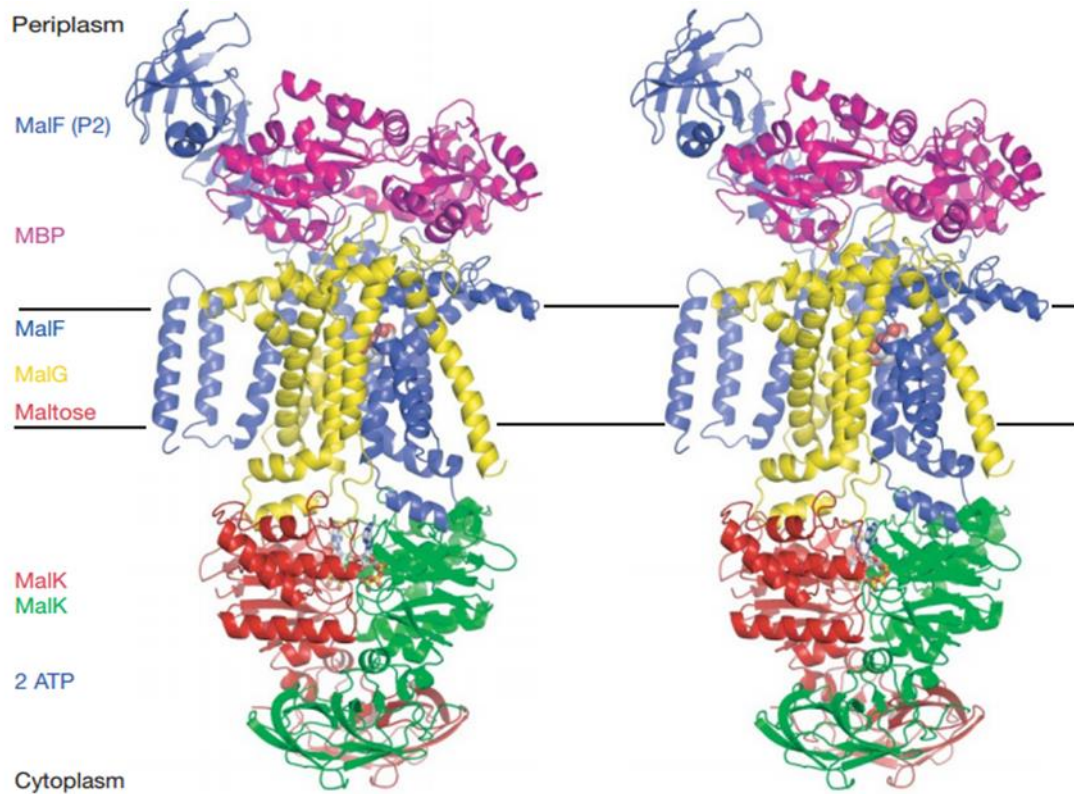


Figure 9. Structure of the maltose importer in a catalytic intermediate state. Magenta - MalE; Blue - MalF, yellow - MalG, red and green - MalK dimer. Maltose and ATP molecules are shown in CPK and Ball and stick models, respectively. From Oldham *et al.*, 2007.

This study also allowed to detect important contact points between TMDs and NBDs. In MalEFGK₂ importer, MalK subunits interact with MalF and MalG mainly by connections with the EAA motif (EAA-X(3)-G-X(9)-I-X-LP), a highly conserved region present in type I importers. This loop is present in both transmembrane domains, helices 6 and 7 in MalF and helices 4 and 5 in MalG, having a high similarity, each made up of two short helices. In both loops containing the EAA motif, the glutamate residue interacts with an arginine through a salt-bridge interaction. One of two helices, docks into a surface cleft on each MalK monomer (Figure 10A and B). This cleft is comprised of two helices from the helical subdomain, residues in the Q loop and the helix following the Walker A (Dassa & Hofnung, 1985; Saurin & Koster, 1994; Daus, Berendt, Wuttge, & Schneider, 2007; Oldham *et al.*, 2007; Ter Beek *et al.*, 2014).

The structure resolution also showed another interaction key point. Six terminal residues of the MalG C-terminal region are inserted into the MalK dimer interface, establishing three hydrogen bonds with the backbone atoms of the Q-loop of one monomer and packing along the Q-loop of the other monomer (Figure 10C). This tail is thought to not be necessary for the closing of a single MalK dimer but the coupling between two MalK monomers by this TMD seems to promote the ordering of the Q-loop, maybe leading to the formation of the catalytic intermediate structure of the complete importer (Dassa & Hofnung, 1985; Oldham *et al.*, 2007).

The resolved conformation also shows the substrate translocation path isolated from the lipid bilayer by transmembrane helices of MalG and MalF and being shielded, in the periplasm, by MalE. The translocation path is then available through several narrow solvent channels linking the transmembrane cavity with the periplasm (Oldham *et al.*, 2007).

Another particularity is the existence of substrate-binding sites inside the transmembrane subunits, MalF and MalG, with aromatic residues stacking and hydrogen bonding promoting sugar detection and binding. This finding corroborates the hypothesis of one maltose molecule per transport cycle (Oldham *et al.*, 2007; Abreu *et al.*, 2021).

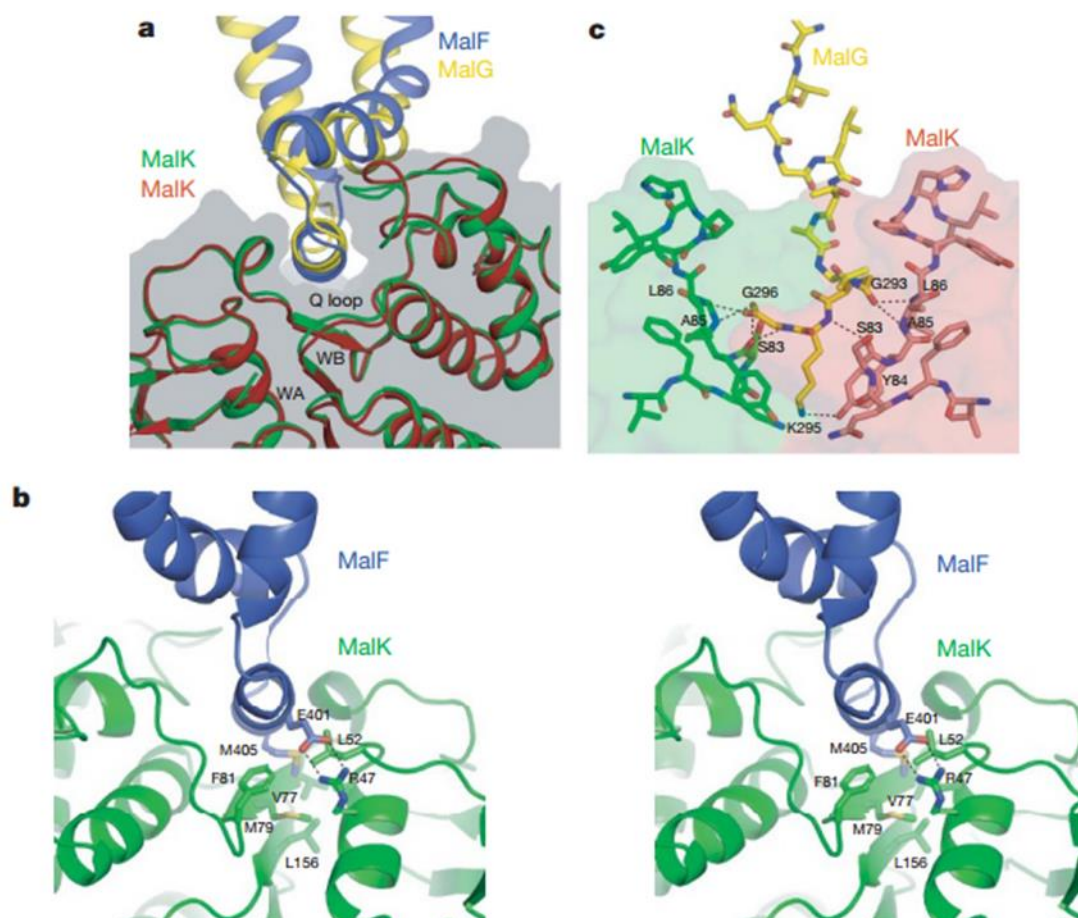


Figure 10. The interface between TMDs and the MalK dimer. A) Docking of the EAA motif of C3 loop into a surface cleft of MalK. MalG and MalF are compared by superposition of the two MalK subunits. Wa, Walker A and Wb, Walker B. B) Representation of MalF and MalK interactions. C) Insertion of the MalG C-terminal region into the MalK dimer interface. Black dashed lines indicate hydrogen bonds and salt bridges. From Oldham *et al.*, 2007.

The crystal conformation of the MalK dimer have been determined in several states, namely, open conformation or resting state, which corresponds to nucleotide-free form, characterized by well-separated NBDs, where the dimer is only kept by interactions of the carboxy-terminal regulatory domains, and in ATP-bound form, which due to the proximity of NBDs, two ATP molecules are engulfed in the dimer's interface, allowing numerous interactions with residues from both monomers. Another crystal structure determined was the ADP-bound conformation, where the ATPase monomers are detached in a similar way to the nucleotide-free structure, implying that ADP is not enough to stabilize NBD closure and thus, ATP hydrolysis may restart the translocator to the resting form (Oldham *et al.*, 2007).

1.2.4.1. Transport mechanism of ABC type I importers

As previously mentioned, ABC type I importers are involved in the translocation of small hydrophilic nutrients, for example sugars. These importers are comprised of two TMDs, that create the translocation pore, two NBDs, which power transport and an SBP, responsible for the delivery of substrate into the translocation pore. These translocators are widely spread in Bacteria, being the MalEFGK₂ importer, the maltose/maltodextrin importer from *E. coli*, the prototype for type I importers, as stated before. This importer permitted to understand several aspects of these proteins such as transport mechanism and its regulatory processes (Oldham *et al.*, 2007; Ter Beek *et al.*, 2014; Mächtel *et al.*, 2019; Narducci *et al.*, 2019; Abreu *et al.*, 2021).

To this date, two action mechanisms were proposed for ABC type I importers based on the crystal conformations of the maltose/maltodextrin's transporter in different states (Figure 11) or on its biochemical features (Figure 12). Both models are known as alternating access model mechanism with the only difference between them being the substrate binding by an SBP and its docking to the TMDs.

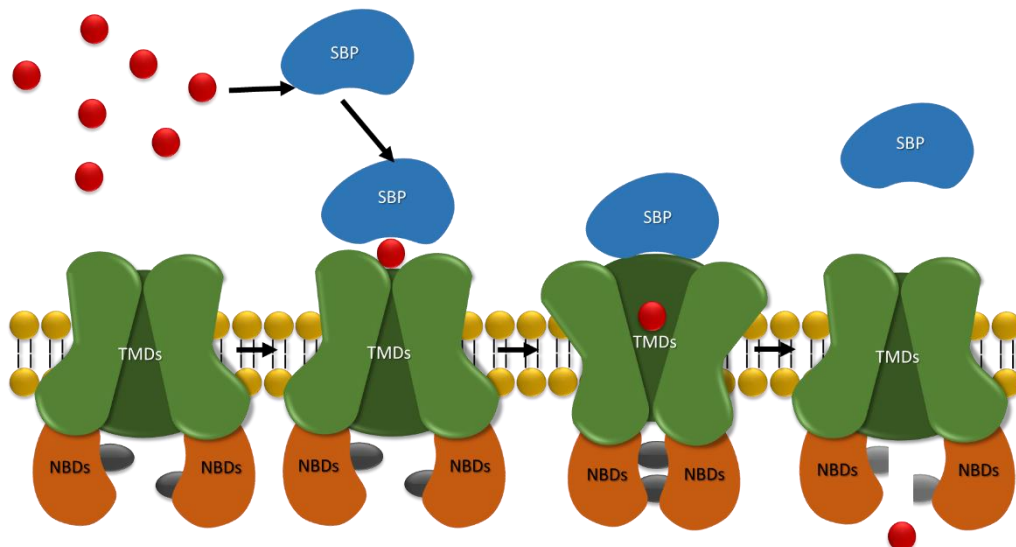


Figure 11. Schematic representation of the Alternative access model. In this model the substrate is capture by an SBP (blue) by the Venus flytrap mechanism, and it is delivered to the TMDs (green). After the hydrolysis of ATP (grey), the substrate (red), in this case maltose, is release inside the cell. NBDs are represented in orange. Adapted from Bao & Duong, 2013

In the first model, the transporter begins in the nucleotide-bound inward-facing state. Here, substrate is captured by an SBPs, which is thought to be by the induced fit model, where substrate binding stimulates bending of the middle beta-sheet hinge, resembling the Venus flytrap mechanism. This protein closes and becomes a close-ligand state (CL state), posteriorly docking onto TMDs, forming the pre-translocation state. In this state, biochemical studies show that MalK nucleotide-binding interface shuts only when MalE, in a close-ligand state, and Mg-ATP are present, inducing conformational modifications in the TMDs, that adopt the outward-facing state, allowing access of the binding site of TMDs. This model indicates that binding of close-ligand MalE to the translocator triggers ATP hydrolysis. After the passage of substrate through TMDs, hydrolysis of ATP and consequently release of inorganic phosphate and ADP, a substrate is then released inside the cell (Figure 11). Although this model is the most accepted translocation mechanism hypothesis, there are some aspects that remain unclear, such as what causes the transport cycle, why does MalE promote ATP hydrolysis even in the vacancy of substrate or how maltose is transported from MalE to MalFG (Oldham *et al.*, 2007; Bao & Duong, 2013; Mächtel *et al.*, 2019 and references therein; Abreu *et al.*, 2021).

The second model is based on biochemical studies of this prototype (Figure 12). Here, MalE in the open state interacts with TMDs in the outward-facing conformation, allowing binding of maltose molecule. After hydrolysis of ATP, the translocator changes to the inward-facing configuration and delivers substrate into the cytoplasm. Although being supported by various studies, stating that MalE does not bind to the transporter but detaches from the NBDs when it captures maltose, and open MalE binds to the outward-facing translocator creating a receptor for the maltose molecule, this model is controversial (Bao & Duong, 2013; Swier *et al.*, 2016). In this hypothesis, MalE in the close-ligand form does not influence the closing of MalK dimer or the change of MalFG to the outward facing conformation, suggesting that ATP alone is enough (Bao & Duong, 2013).

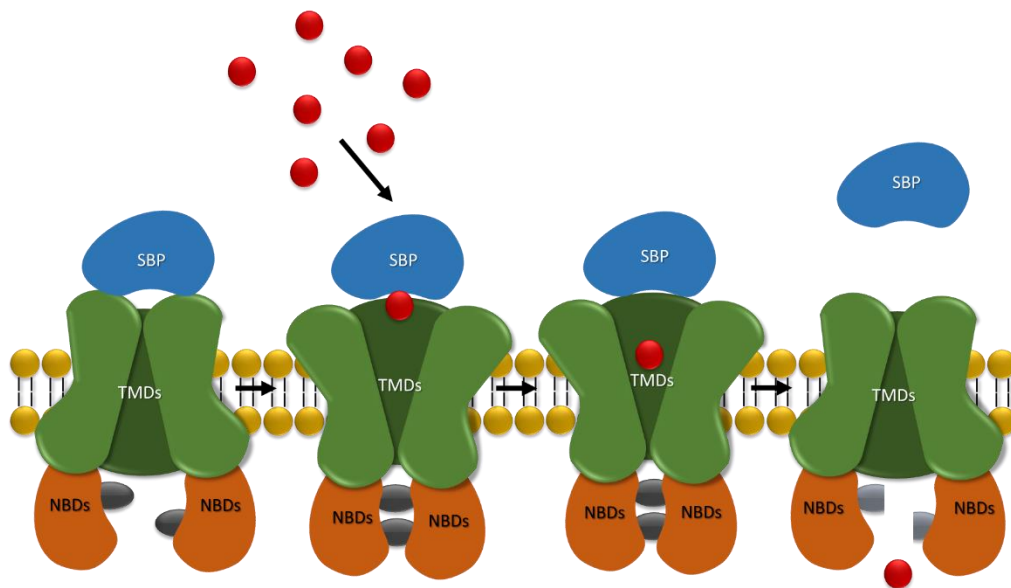


Figure 12. Second schematic representation of the Alternative access model. Here, the SBP (blue) interacts with the TMDs (green) in the open conformation allowing the binding of the substrate, maltose (red). After ATP (grey) hydrolysis, the substrate is release into the cytoplasm. NBDs are represented in orange. Adapted from Bao & Duong, 2013.

For any ABC transporter is estimated a mechanistic stoichiometry of two ATP molecules/one substrate molecule ratio, although this value is very controversial because this could drive to greater lipid membrane gradients of the translocated substrate (Patzlaff *et al.*, 2003; Ter Beek *et al.*, 2014). Selectivity of ABC type I importers was also studied in this prototype. Here, it was discovered that not only SBPs participate in substrate recognition, but also TMDs contribute to substrate specificity (Swier *et al.*, 2016), with two binding pockets being discovered in both MalF and MalG (Oldham *et al.*, 2007, 2013; Abreu *et al.*, 2021).

Activity control of the transporter can be made by additional domains fused to NBDs. In the maltose/maltodextrin's system, the C-terminal regulatory domain of MalK is engaged in carbon catabolite repression. This domain interacts with a component of phosphoenol-dependent sugar phosphotransferase System (PTS), IIA^{glc}. In this system, when glucose is present, IIA^{glc} is unphosphorylated and binds to the TMD, stabilizing the inward-facing conformation by wedging between the MalK dimer, preventing maltose uptake. When glucose becomes scarce, IIA^{glc} is phosphorylated, and inhibition is relieved (Chen *et al.*, 2013). Other modes of activity control include the medium osmolality, signalled as a rise in the cytoplasmic electrolyte concentration, as described for the OpuA transporter, that has two cystathionine- β -synthase (CBS) domains attached in tandem to the C-terminus of the ATPase and an anionic membrane surface, central to the gating of the translocator by ionic strength. In this model of inhibition, contacts of CBS with the anionic membrane secures the translocator in an inactive form (Biemans-Oldehinkel *et al.*, 2006; Karasawa *et al.*, 2011). Rise of intracellular ionic strength activates the transporter. Another hypothesis says that the transporter could have two like-charged surfaces, for example an anionic membrane and anionic protein, which with the rise of ionic strength, would interact (Swier *et al.*, 2016). Inhibition by pre-accumulation is another type of control of activity. In this type of regulation, when medium osmolality rises, ABC transporters are activated, diminishing inhibition, and preventing higher turgor pressure (Verheul *et al.*, 1997; Swier *et al.*, 2016).

1.3. Multitask Nucleotide-Binding Domains

NBDs constitute the hallmark of ABC translocators family. These proteins are responsible for hydrolysis of ATP, consequently powering substrate translocation. Although proteins belonging to this domain have a high sequence and structure similarity, a very surprising particularity was determined for some ATPases. The ability to energise more than one transport system has been proposed since the '90s in various transporters, such as MsiK in *Streptomyces lividans* (Schlosser, Kampers, & Schrempf, 1997), MalK in *Streptococcus suis* (Tan *et al.*, 2015), MalK in *Streptococcus pneumoniae* (Marion *et al.*, 2011) and in the recently studied multitask ATPase MsmX from *B. subtilis*, capable of energising six different sugar uptake systems (Ferreira & Sá-Nogueira, 2010; Ferreira *et al.*, 2017; Leisico *et al.*, 2020). To this day, multitask ATPases were only found in bacterial ABC sugar importers of the

Carbohydrate Uptake Transporter 1 (CUT1) family, responsible for the uptake of di-, tri-, or higher oligosaccharides and polyols (Leisico *et al.*, 2020).

Another surprising characteristic found in this protein and its homologs is the existence of intra- and inter-species exchangeability determined by several genetic studies on this multitask ATPases (Leisico *et al.*, 2020). Intra-species exchangeability was described for the first time in the Firmicutes phylum in 2017, when Ferreira *et al.* observed that a second orphan ATPase of *B. subtilis*, YurJ, now renamed FrLP, with high sequence similarity to MsmX, was able to replace this NBD in its role in the uptake of arabinooligosaccharides by the AraNPQ translocator (Ferreira *et al.*, 2017). Inter-species exchangeability was recently determined by Leisico and Godinho (Leisico *et al.*, 2020). Here, a genetic system was constructed to assess the capacity of proteins similar to MsmX from Gram-positive bacteria of the Firmicutes phylum and other Gram-negative Bacteria to replace this NBD in arabinotriose (transported by AraNPQ translocator) and galactan (transported by GanSPQ translocator) transport. The ABC_Cd ATPase, from the human pathogen *Clostridioides difficile* 630, was found to be an orphan and putative NBD, while ABC_Sa, also from a human pathogen, *Staphylococcus aureus*, exists in an operon encoding a complete ABC carbohydrate uptake system. These NBDs, alongside with *Bacillus thuringiensis* ABC_Bt ATPase and *Streptococcus pneumoniae* TIGR4 MsmK ATPase were shown to be able to replace MsmX in the previously mentioned sugar translocation with different efficiencies

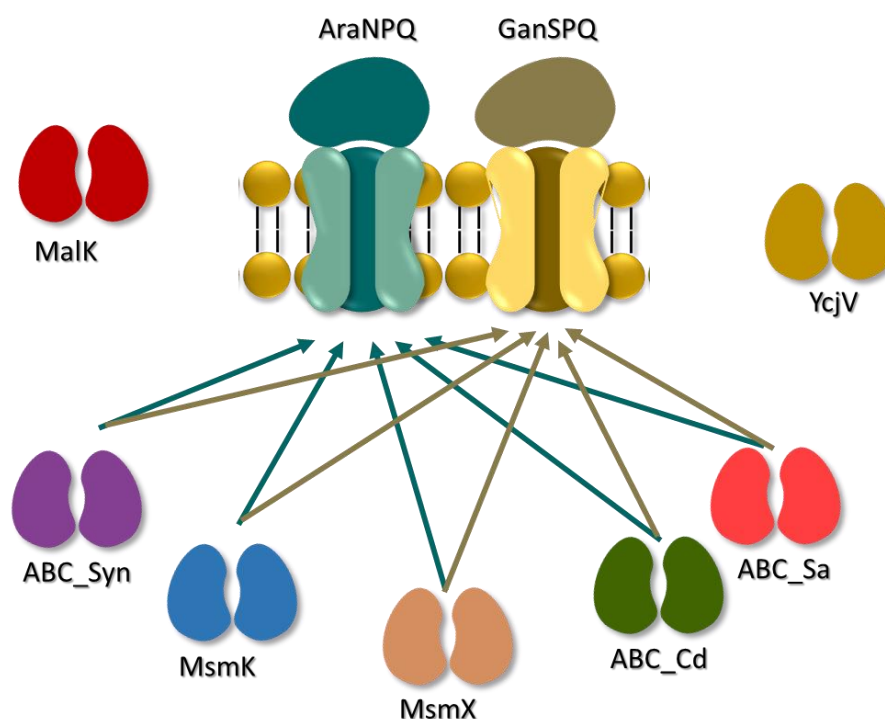


Figure 13. Interspecies exchangeability in Firmicutes phylum and beyond. MsmK (blue) from *S. pneumoniae*, ABC_Cd (green) from *C. difficile* 630, ABC_Sa (pink) from *S. aureus* were shown to be able to replace MsmX (pale pink) in arabinotriose (AraNPQ, cyan blue) and galactan (GanSPQ, pale yellow) transport, showing interspecies exchangeability in Firmicutes phylum. ABC_Syn (purple) from *Synechocystis* sp. was also shown to replace MsmX, showing interspecies exchangeability beyond Firmicutes phylum. MalK (red) and YcjV (yellow), from *E. coli* could not replace MsmX.

(Figure 13). However, despite also having high sequence similarity, other known ATPases from *E. coli*, YcjV and MalK, were not able to complement sugar transport uptake (Leisico *et al.*, 2020).

This capacity for exchangeability was also shown beyond Firmicutes phylum. In Cyanobacteria, a putative ATPase from *Synechocystis sp.* was also identified by its high sequence identity with MsmX and posteriorly, this NBD was also demonstrated to be able to replace MsmX in arabinotriose and galactan translocation, in *B. subtilis*. This implies that multitask ATPases not only can be exchangeable within the same species and between species in the Firmicutes phylum but also can be exchangeable beyond this phylum (Leisico *et al.*, 2020).

The existence of multitask ATPases and their spread across species and phyla could allow for different or additional levels of regulation over ABC sugar transporters, which can bring adaptive advantages to the organism, since it may not need to express different proteins for different systems (Leisico *et al.*, 2020). Since these ATPases are only found in Bacteria and studies in the MsmK ATPase from *Streptococcus pneumoniae* shows that inactivating this NBD, leads to the inactivation of multiple carbohydrate transporters, resulting in an attenuated virulence and consequently reducing the capacity of this pathogen to colonize airways and cause disease, it demonstrates the importance of these multitask proteins *in vivo* (Marion *et al.*, 2011).

1.3.1.MsmX ATPase and AraNPQ type I Importer

One of the most well-characterized multitask NBDs is MsmX, responsible for powering sugar type I importers in *B. subtilis*. This soil bacterium uses plant biomass as the main source of carbohydrates, especially complex sugars present in plant cell walls (Tam *et al.*, 2006). The genome of this Gram-positive model organism encodes two endo- α -1,5-arabinanases, AbnA and Abn2, that allow for the homopolysaccharide arabinan hydrolysis (Leal & Sá-Nogueira, 2004; Inácio & Sá-Nogueira, 2008). Then two intracellular α -L-arabinofuranosidases, Abf2 and AbfA, degrade arabinooligosaccharides, after their transport through two types of translocator systems (Inácio & Sá-Nogueira, 2008). One of these transport systems is the proton symporter AraE, in charge for the uptake of arabinose, α -1,5-arabinobiose, xylose and galactose and the other transport system is the ABC type I importer AraNPQ, responsible for the import of α -1,5-arabinobiose, α -1,5-arabinotriose, α -1,5-arabinotetraose (Krispin & Allmansberger, 1998; Sá-Nogueira & Ramos, 1997; Ferreira & Sá-Nogueira, 2010). This last importer is clustered in the operon *araABDLMNPQ-abfA*, which encodes enzymes involved in arabinose catabolism and degradation of arabinooligosaccharides. Here, *araN* encodes the high-affinity substrate binding protein (SBP) AraN, while AraQ and AraP are transmembrane domains (TMD) encoded by *araPQ* (Sá-Nogueira, Nogueira *et al.*, 1997; Sá-Nogueira & Ramos, 1997; Ferreira & Sá-Nogueira, 2010).

This operon does not encode for a nucleotide-binding domain (NBD) for the AraNPQ ABC translocator, however, five orphan ATPases in the *B. subtilis* genome were known (Quentin *et al.*, 1999). One of these NBDs is MsmX, encoded by a monocistronic gene and located in a distant locus without any ABC transport component close to it. This orphan ATPase is constitutively expressed and presents a high similarity to other well-known ATPases from different Bacteria (Quentin *et al.*, 1999; Ferreira & Sá-Nogueira, 2010). Later, it was proven the ability of this ATPase to complement six of seven incomplete carbohydrate transport systems, such as MdxEFG (Schönert *et al.*, 2006), involved in maltodextrins uptake, AraNPQ (Ferreira & Sá-Nogueira, 2010), involved in the arabinooligosaccharides uptake, GanSPQ (Ferreira *et al.*, 2017), involved in the galacto-oligosaccharides transport, YesOPQ and YtcPQ-YteP (Ferreira *et al.*, 2017), involved in the transport of rhamnose-galacturonic acid disaccharides and/or galacturonic acid oligomers, and MeIECD system, involved in the melibiose transport (Heravi *et al.*, 2019) (Figure 14). With MsmX involved in all these transport systems, it was designated a multitask ATPase (Schönert *et al.*, 2006; Ferreira & Sá-Nogueira, 2010; Ferreira *et al.*, 2017; Heravi *et al.*, 2019).

The structure of MsmX was recently solved in the open state (Leisico *et al.*, 2020) demonstrating a high structural similarity with MalK from *E. coli* and *Pyrococcus horikoshii*. MsmX shares its fold with other well-known ATPases of ABC type I translocators, namely, the general α/β type ATPase domain fold in the N-terminal domain. This subdomain is constituted by a RecA-like domain and an α -helical domain, and a C-terminal regulatory domain was also identified, containing a mixed barrel with 3 α -helices and 11 β -strands. Conserved functional motifs within ATPases are found in the N-terminal domain. Between the antiparallel $\beta 1$ and $\beta 2$ strands is located the A-loop motif, being the least

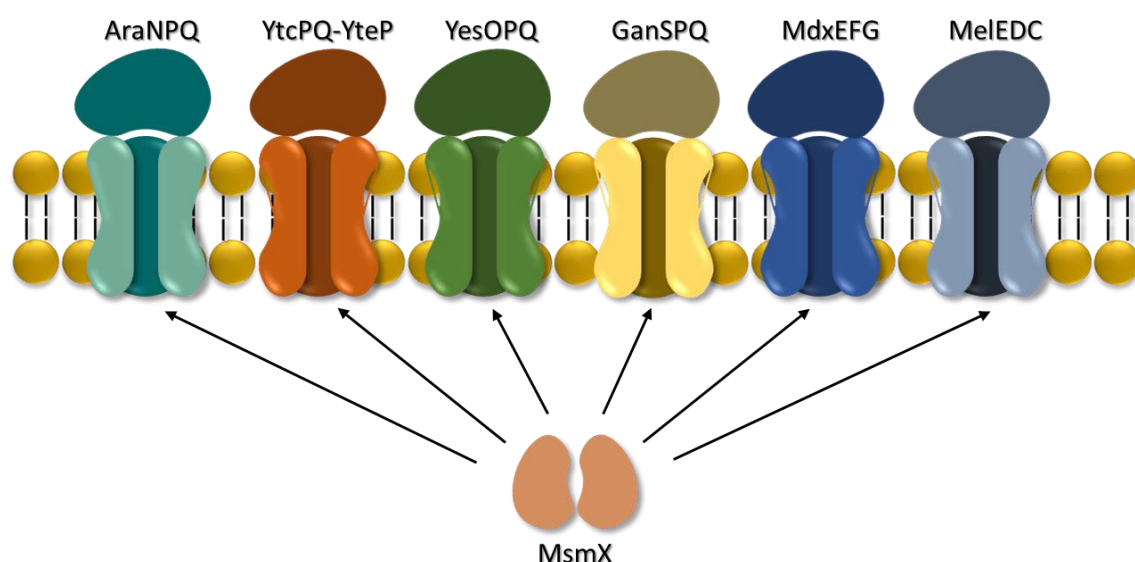


Figure 14. All carbohydrate transport systems energized by MsmX ATPase. MdxEFG (dark blue), involved in maltodextrins uptake, AraNPQ (cyan blue), involved in the arabinooligosaccharides uptake, GanSPQ (pale yellow), involved in the galacto-oligosaccharides transport, YesOPQ (green) and YtcPQ-YteP (orange), involved in the import of rhamnose-galacturonic acid disaccharides and/or galacturonic acid oligomers, and MeIECD (purple) system, involved in the melibiose transport. MsmX is represented as pale pink.

conserved and apparently conferring higher degree of flexibility to ATP binding. Depending on the absence or presence of ATP, it can adopt different conformations. (Leisico *et al.*, 2020). The Walker A motif is found between residues 37-44 and is conserved in this protein, while the Q-loop, which is involved in the hydrolysis cycle of ATP, is responsible for interaction with the α -helix of the TMDs. In MsmX the Q-loop is formed by residues 81-87. Another motif, Walker B, comprised by the conserved residues 155-160, forms strand β 7. This motif contains amino acids responsible for water polarization and Mg^{2+} coordination during ATP hydrolysis. Residues 192-194, forming the H-loop, are responsible for the interaction with the γ -phosphate of the ATP substrate (H193). Downstream of Walker B, there is the D-loop, comprised by residues 163-166. The characteristic ABC signature, in residues 135-139, is conserved and located between α -helices 5 and 6. The C-terminal regulatory domain is related with regulatory mechanisms and is the least conserved within ATPases of ABC translocators. In this region, MsmX has a helix α 11, between β 13 and β 14 (Leisico *et al.*, 2020).

When comparing this structure with MalK from *E. coli*, one of the most distinctive features is the larger surface area in the pocket responsible for interactions with the TMDs, the Q-loop (Figure 15). This characteristic may be one of the reasons for the ability to interact with TMDs of different ABC type I sugar importers (Leisico *et al.*, 2020).

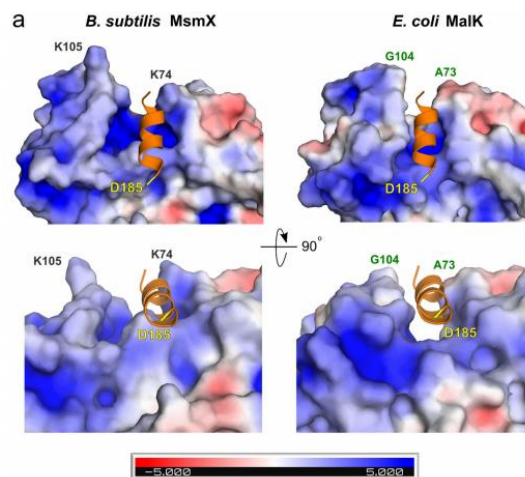


Figure 15. MalK and MsmX electrostatic potential at NBD-TMD interface. EAA helix from MalG protein is depicted as orange and its N-terminal residue (D185) is represented in yellow. In this representation it is possible to notice a wide interaction zone in MsmX when comparing with MalK, possibly contributing for the multitask ability. From Leisico *et al.*, 2020.

1.3.2. Other multitask ATPases

Currently there are other multitask ATPases identified, all of them associated with sugar transport. One of these ATPases belongs to the opportunistic respiratory, *Streptococcus pneumoniae*. Although this Gram-positive bacterium commonly colonizes airways asymptotically, it is a respiratory pathogen and the cause of various diseases, namely otitis and pneumonia (Bidossi *et al.*, 2012). To be able to grow and eventually develop the pathogenesis process, *S. pneumoniae* needs to

acquire carbohydrates, namely complex glycans because carbohydrates are limited in the airway. For the uptake of these molecules, it is estimated the existence of 21 phosphotransferase systems and 8 ABC transporters (Bidossi *et al.*, 2012). One of these carbohydrates is sialic acid, which can be cleaved by proteins present in these Gram-positive bacteria. The translocation of sialic acid is performed by proteins encoded by *satABC* genes, where *satA* encodes for the substrate-binding protein and *satBC* encodes for two permeases (Marion *et al.*, 2011). However, this system does not encode for an ATPase necessary to provide energy for the uptake. Later, MsmK was then proposed to energise this system (Marion *et al.*, 2011; Buckwalter & King, 2012).

Assessment of MsmK amino acid sequence indicates the existence of all typical features of an NBD, including Walker A, Walker B, Q-loop, D-loop, H-loop and ABC signature motifs. Uptake of carbohydrates by this system is most probably controlled by carbon catabolite repression (Marion *et al.*, 2011). Besides the SatABC system, MsmK was found to be necessary for growth on other two substrates, raffinose, uptaken by RafEFG system and maltooligosaccharides, uptaken by the MalXCD system. Both sugars are known to be transported by incomplete ABC transporters systems, lacking an ATPase. MsmK was thus classified as a multitask ATPase. Studies showed that inactivating this NBD leads to the deactivation of multiple carbohydrate transporters and consequently impacts the capacity of this bacterium to colonize the airway and provoke disease (Marion, Aten, *et al.*, 2011).

In other *Streptococcus* bacteria, namely, in *S. suis*, an emergent zoonotic bacterium that causes deadly infections in pigs and humans worldwide, an MsmK-like ATPase also energises the melibiose, raffinose, maltotriose, maltotetraose and glycogen transport systems. In this bacterium and like in *S. pneumoniae*, MsmK was also considered important for the *in vivo* survival and colonization (Tan *et al.*, 2015). In this species it was also found that another ATPase, MalK, was able to replace MsmK and vice-versa (Webb *et al.*, 2008).

Streptomyces species are Gram-positive bacteria that inhabits soil and like other soil bacteria, such as *B. subtilis*, secretes a variety of hydrolytic enzymes, namely cellulases, xylanases, mannanases and chitinases to degrade lignocellulosic biomass. In these bacteria, transport of disaccharides such as cellobiose, xylobiose, the monosaccharide xylose and (GlcNAc)₂, the degradation product of chitin, is dependent on an ATP-binding system. In *Streptomyces coelicolor*, it is described that *dasABC* gene cluster encodes subunits of ABC transporters for the uptake of (GlcNAc)₂. For this bacterium it was shown that the *dasA* gene encodes for an SBP, while *dasB* and *dasC* encode for transmembrane proteins. Other oligosaccharide transporter systems were found in streptomycetes, such as *cebEFG* for cellobiose and cellotriose in *Streptomyces reticuli*, *malEFG* for maltose in *S. coelicolor*, *ngcEFG* for GlcNAc and (GlcNAc)₂ in *Streptomyces olivaceoviridis* and *bxlEFG* for xylobiose in *Streptomyces thermoviolaceus*. All these ABC transporters encode genes for SBPs and TMDs but not for their ATPases. In *Streptomyces lividans* an orphan ATPase, known as MsiK, supports the cellobiose-, xylobiose and maltose-uptake systems as well the trehalose-uptake system in *S. reticuli*. This ATPase has high homology with MsmK of *S. mutans*, another multitask ATPase. So, therefore, MsiK is the putative

NBD for the incomplete systems described above and it was shown that the product of *msiK* gene is required for (GlcNAc)₂ uptakes in *S. coelicolor* (Hurtubise, Shareck, Kluepfel, & Morosoli, 1995; Schlösser *et al.*, 1997; Saito *et al.*, 2008).

Although multitask ATPases seem to be commonly spread among Gram-positive bacteria, there is also evidence of their existence in Gram-negative bacteria. In *Thermus thermophilus*, two ATPases, displaying homology with MalK from *T. litoralis* and *E. coli*, were named MalK1 and MalK2. MalK1 is an ATPase responsible for energising *malEFG1* system, allowing uptake of trehalose, maltose, sucrose and palatinose and another sugar ABC transporter, involved in the transport of mannose and glucose (Sampaio *et al.*, 2005; Chevance *et al.*, 2006;).

1.4. Study of protein-protein interactions

As described above, the existence of multitask ability of several ATPases could constitute a potential target for an antimicrobial therapy or other biotechnology/biomedical applications, but first it is necessary to understand what influences interaction between NBDs and TMDs and what contributes to the existence of several sugar ABC importer systems energized by a unique ATPases. For that, study of the interaction between proteins, also known as protein-protein interactions (PPIs) is fundamental to understand this phenomenon. PPIs are responsible for a variety of biological processes like metabolism, transport, or in the case of this dissertation, interaction between NBDs and TMDs are responsible for translocation of sugars by type I ABC importers. These interactions are due to noncovalent interactions between residue side chains (Ofrañ & Rost, 2003b). Nowadays, PPIs are an important theme for the development and progress of modern system's biology (Rao *et al.*, 2014; Miura, 2018; Lite *et al.*, 2020).

The study of PPIs allows to comprehend and infer about the molecular mechanism of cellular processes, biochemistry of the cell and aids in the identification of drug targets. PPIs can be either transient or permanent (Ofrañ & Rost, 2003a). Transient interactions are more diverse mechanistically, they are involved in all kinds of processes, like in receptor-ligand interaction or in signal transduction, and are present in most protein families while permanent interactions will form a stable protein complex, like the assembly of folded chains into multi-chain proteins (Ofrañ & Rost, 2003a; Rao *et al.*, 2014).

Detection methods for PPIs can have diverse goals and can be divided into three groups: *in vitro*, *in vivo* and *in silico*. In *in vivo* methods, procedures are performed on a living organism, like in the Two-Hybrid Systems (Bacterial and Yeast) and synthetic lethality. *In vitro* methods are performed outside a living organism in a controlled environment (Fields & Song, 1989; Karimova *et al.*, 1998; Ito *et al.*, 2017; Karimova *et al.*, 2017; Wong *et al.*, 2017; Klobucar & Brown, 2018; Bacon *et al.*, 2021). These techniques involve affinity chromatography, coimmunoprecipitation, NMR spectroscopy, protein fragment complementation, protein arrays, phage display, tandem affinity purification and X-ray crystallography (Meng *et al.*, 2005; Chavez *et al.*, 2011; Geetha *et al.*, 2012; Rao *et al.*, 2014;

Syafrizayanti *et al.*, 2014; Purslow *et al.*, 2020; Iacobucci *et al.*, 2021). *In silico* techniques are executed via computer simulation or on a computer and they are sequence, gene and structure-based approaches, gene fusion, chromosome proximity or phylogenetic tree (Meng *et al.*, 2005; Rao *et al.*, 2014)

1.4.1. Methods of study

In recent years biological assays have become essential tools for the study of PPIs, being the most used the Yeast, Bacterial or Mammalian Two-Hybrid approach (Karimova *et al.*, 1998, 2017; Saraon *et al.*, 2017; Bacon *et al.*, 2021). This methodology uses putative PPIs to reconstitute a functionally active molecule, normally a transcription activator that stimulates the expression of a determined reporter gene, of which product can be measured. The Yeast Two-Hybrid (Y2H) technique, the most famous Two-Hybrid method, uses the *lacZ* reporter gene under the control of the GAL4 transcriptional factor. In this method, one protein is fused to the GAL4-binding domain, named bait and the other protein is fused to the GAL4-activating domain, named as prey. If the two proteins under study interact, there is an activation of the *lacZ* gene, and it can be detected by measuring β -galactosidase activity. This methodology has high throughput and low sample consumption but has the setback of using fusion proteins, the interaction taking place in the nucleus and the fact that polypeptides that may come into contact are naturally expressed at different times or locations (Stynen *et al.*, 2012; Syafrizayanti *et al.*, 2014; Wong *et al.*, 2017; Miura, 2018). Y2H has been optimized by creating new yeast strains, harbouring new vectors and different reporter genes. Y2H, together with the Bacterial Two-Hybrid, which uses bacterial cells, allows not only a qualitative characterization but also a quantitative characterization. Other techniques that have been used to complement the Two-Hybrid assays are standard wide-field fluorescence microscope or a laser scanning confocal microscope because they are typically carried out in the context of living cells and since it is possible to “visualize” the interaction within the cell, providing spatial information lacking in the two-hybrid. These techniques include Fluorescence and Bioluminescence Resonance Energy Transfer (FRET and BRET, respectively), Fluorescence Correlation Spectroscopy (FCS) and Bimolecular Fluorescence Complementation (BiFC), which require expensive instrumentation (Meng *et al.*, 2005; Rao *et al.*, 2014 and references therein; Miura, 2018;)

Another biological approach is synthetic lethality, a type of *in vivo* genetic screening in which deletions or mutations are introduced in two or more genes, which are viable alone, but when combined cause cellular death. These interactions are called functional interactions since these associations do not need physical contact between polypeptides. This approach tries to identify conditions that allow phenotypic stability, despite environment changes, genetic variation or aleatory mutation events. (Bendert & Pringle, 1991; Rao *et al.*, 2014; Klobucar & Brown, 2018).

Another way to assess PPIs is by biochemical approaches, the most used methods nowadays. These procedures include co-immunoprecipitation, affinity purification, chemical crosslinking, and blot

assays. Affinity purifications involve the immobilization of a determined “bait” protein and after passing a protein extract, only polypeptides that establish contacts with a bait protein are immobilized and later eluted and may be identified by Mass Spectrometry (MS) or if the interacting proteins are already known, Western blot can be used. Co-immunoprecipitation has a similar principle to affinity purification and involves addition of antigen-antibody complex to the cell lysate and after antibody precipitation, any proteins forming a complex are eluted and like in co-immunoprecipitation, identified by MS or Western blot (Geetha *et al.*, 2012; Iacobucci *et al.*, 2021 and references therein). These two techniques have the advantages of being relatively low cost, no major infrastructure is needed and may capture low-affinity interaction, but like all techniques they have the disadvantage, like having low sensitivity, relying on the assumption that proteins involved exist in a complex that is stable enough to survive the cell fractional process and assay procedure and a high background might be present. In blot assays, lack of interaction is inconclusive, and any interaction detected must be verified by other assays because, in this technique, proteins are denatured and may or may not fully renature (Meng *et al.*, 2005; Kastelic, 2014;). Cross-linking is used to chemically link molecules by a covalent bond, so transient or weak PPIs can be identified that may otherwise evade detection (Chavez *et al.*, 2011; Iacobucci *et al.*, 2021).

There are other methodologies to study PPIs but the most reliable and successful means to learn about how proteins interact with one another is to combine various approaches (Meng *et al.*, 2005; Iacobucci *et al.*, 2021).

1.4.1.1. Bacterial Adenylate Cyclase Two-Hybrid System (BACTH)

In this dissertation the study of several mutations in TMDs was used to map important contact regions and to study the interaction between TMDs and multitask NBDs. For that, a Bacterial Adenylate Cyclase Two-Hybrid (BACTH) System was used. This methodology is a fast and simple approach to determine and characterize interaction between proteins that does not require specialized skills besides molecular biology and microbiology techniques. The BACTH system was designed in *E. coli* and uses the reconstitution of an enzyme, the calmodulin-dependent adenylate cyclase from *Bordetella pertussis*, to determine if two proteins of interest are capable of interacting or not with each other. This is possible because the catalytic domain of the adenylate cyclase protein is formed by two complementary fragments, T18 (residues 225-399) and T25 (residues 1-224), that interact when in presence of calmodulin, activating a regulatory cascade, leading to 3',5'-cyclic adenosine monophosphate (cAMP) production from ATP (Karimova *et al.*, 1998). Then, cAMP interacts with the pleiotropic regulator Catabolite Activator Protein (CAP), allowing activation of reporter genes, including genes involved in

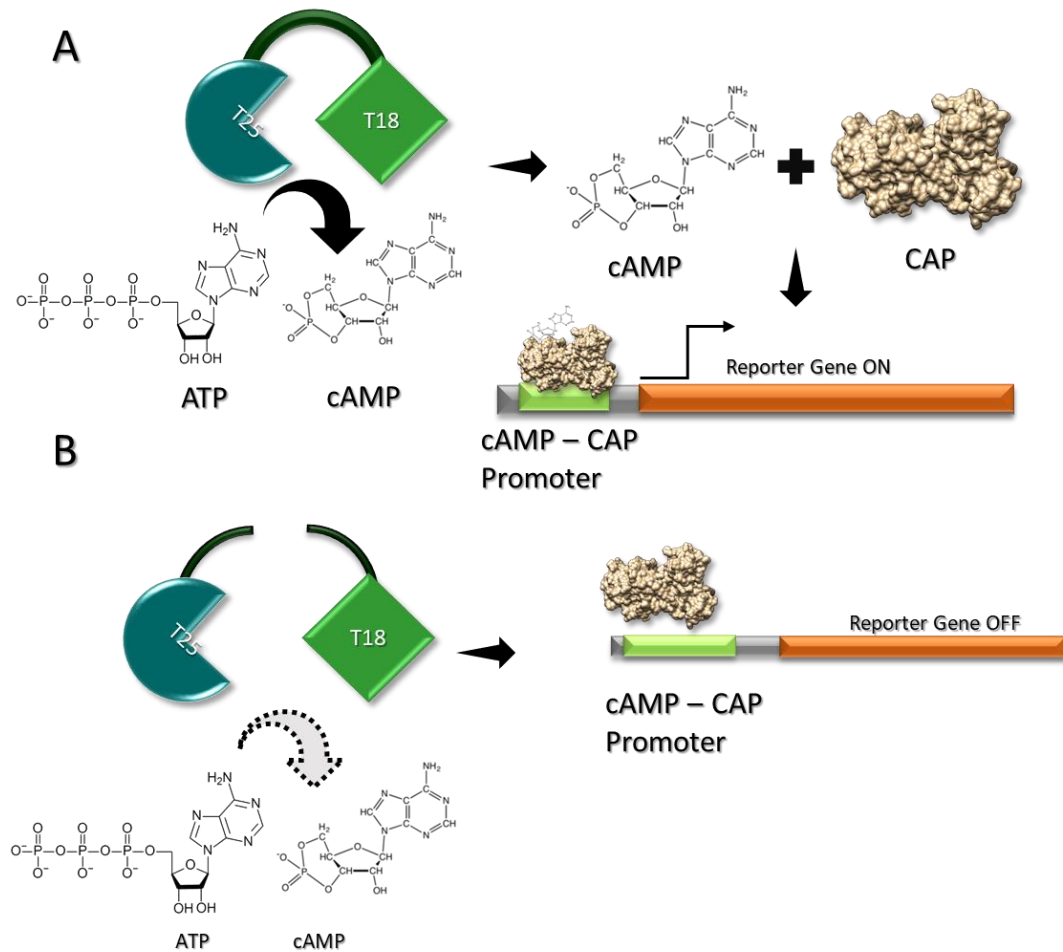


Figure 16. Adenylate Cyclase activity. A) If the two fragments T25 (cyan blue) and T18 (green) interact, the reconstitution of the catalytic domain of adenylate cyclase allows the conversion of ATP into cAMP. cAMP will interact with CAP protein and will activate a reporter gene. B) If these two fragments do not interact, there is no cAMP production and therefore no activation of the reporter gene. Adapted from Karimova *et al.*, 2017.

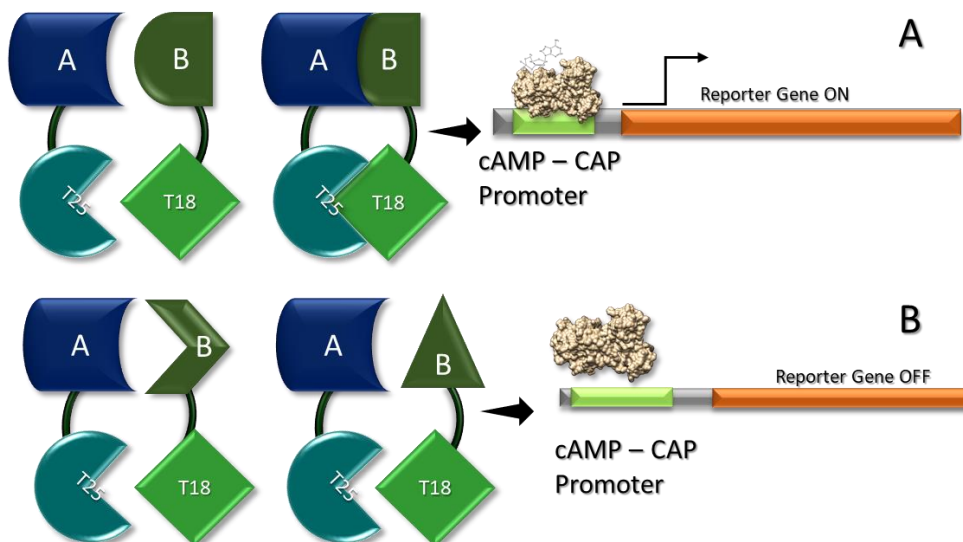


Figure 17. Bacterial Two Hybrid system. A) Two polypeptides in study, A (dark blue) and B (dark green), are fused to fragments T25 (cyan blue) and T18 (green), respectively. If these two hybrids interact with each other, there will be cAMP production and consequently activation of the reporter gene. B) However, if these hybrids do not interact, there will be no activation of the reporter gene. Adapted from Karimova *et al.*, 2017.

the catabolism of carbohydrates, such as maltose or lactose (Figure 16). The BACTH system uses T18

and T25, each fused to the polypeptides in study, expressed in an *E. coli* strain lacking an endogenous adenylate cyclase. If two polypeptides of interest interact (Figure 17A), they will bring the two modules close allowing heterodimerization, leading to complementation between the T18 and T25 fragments and consequently cyclic AMP production. This molecule will then interact with the pleiotropic regulator, CAP, allowing expression of reporter genes, including *lac* and *mal* operons, which can easily be detected by several easy and fast assays, using a medium indicator, detecting cAMP levels or by assaying β -galactosidase activity (Karimova *et al.*, 1998; Karimova *et al.*, 2017). If the proteins of interest are not interacting proteins (Figure 17B), heterodimerization does not occur between hybrid polypeptides, there is not cAMP production and, therefore, no reporter gene activation (Karimova *et al.*, 1998, 2017).

Another appealing characteristic of the BACTH system is the several methods available to detect PPI. A direct method to determine complementation efficiency is through measurement of cAMP levels inside the cell, that correspond to a measurement of the reconstructed adenylate cyclase activity. A colorimetric method to qualitatively assess interaction between two hybrid proteins uses a medium indicator, such as LB plates supplemented with X-Gal (5-bromo-4-chloro-2-indolyl- β -D-galactopyranoside), where colonies harbouring interacting polypeptides have blue coloration and not interacting polypeptides appear as white colonies, or MacConkey agar plates supplemented with maltose, in which lack of interaction is presented as colourless colonies and positive interaction displays a red colour. Another method is to determine adenylate cyclase reconstitution in an indirect way, by determining β -galactosidase activity. Here, interaction levels are correlated with cAMP levels produced in cells, since *lacZ* gene expression is positively regulated by cAMP/CAP complex. β -galactosidase activity can be determined using ONPG (ortho-nitrophenyl- β -galactoside) as substrate. This is a very simple procedure where cell lysates are incubated with a reaction buffer and ONPG, a galactoside analogue that can be cleaved by β -galactosidase, originating galactose and a chromophore o-nitrophenol, a yellow substance, that can be measured by a spectrophotometer, at 420 nm (Karimova *et al.*, 2017).

A clear sign of the versatility of BACTH system are the different types of proteins of different domains of life, from animals to bacteria, and even interaction between virus and host proteins that have been studied using this system through the years (Fransen *et al.*, 2002; Dautin *et al.*, 2003; Karimova *et al.*, 2017). Besides the advantages described before of being an easy and fast assay with several methodologies that can be used to study interaction between two proteins of interest, this system is especially fitting method to study interactions between membrane proteins, since these are not suitable to be tested with transcription-based two-hybrid system and since BACTH system uses a signalling cascade, it does not need to be close to the transcription machinery, as happens in other bacterial or yeast hybrid systems (Karimova *et al.*, 2017).

1.5. Scope of Thesis

With the discovery of multitask ATPases, the question arises of what contributes to these promiscuous interactions between TMDs and NBDs. Thus, the work presented in this dissertation aims to map interaction regions between TMDs-NBDs, using as main model the AraPQ TMDs and MsmX ATPase from the type-I importer AraNPQ belonging to the Gram-positive model organism, *B. subtilis*. Here, several amino acids conserved in all TMDs energized by MsmX were identified and were mutagenized. The effect of these mutations in interaction between AraPQ and MsmX was assessed by using the BACTH system, through β -galactosidase assays.

Previous results have also identified several multitask ATPases from Firmicutes phylum and *Synechocystis* sp. able to replace MsmX *in vivo*, using a genetic system in *B. subtilis* (Leisico *et al.*, 2020). This work also aims to study interactions between AraPQ and these multitask ATPases, using the previously mentioned method.

Chapter 2

Materials and Methods

2. Materials and Methods

2.1. Bioinformatic and statistical analysis

All nucleotide sequences were obtained from the National Center for Biotechnology Information at the National Institute of Health, Bethesda, Maryland. Amino acids sequences were obtained from Uniprot and all alignments were executed by the algorithm ClustalW2 (EMBL-EBI, UK). Oligonucleotides and plasmids constructed in this work were designed using SnapGene[®] software (from Insightful Science; available at snapgene.com). All graphics were constructed using GraphPad Prism version 6.0.0 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com, and statistical analysis of enzymatic activities were performed using Student's t-test for paired samples, adapted for multiple comparisons by applying the Bonferroni correction for false discovery rate, with the significant threshold being set at $\alpha = 0.05$. All statistical analyses were performed using R 3.5.1 software (R Core Team, 2018).

2.2. Competent cells

To prepare *E. coli* BTH101 competent cells, strain was streaked onto a LA plate supplemented with streptomycin (100 µg/mL) and incubated overnight (O/N) at 37 °C. A single colony was selected and was utilized to inoculate the pilot culture, LB supplemented with streptomycin (100 µg/mL), and was grown O/N at 37 °C, 180 rpm (Aquatron INFORS HT, Switzerland). Pilot culture was diluted 1:200 in the same medium and growth was followed until optical density (Ultrospec 2100 UV-Visible Spectrophotometer, Amersham Pharmacia Biotech, Chicago, USA) at 600 nm reached 0.6. Cells were recovered by centrifugation, 4 492 g for 5 min, at 4 °C, and then cells were resuspended in 1/2 of the initial volume in 0.1 M CaCl₂. After 2 h of incubation on ice, cells were centrifuged at 4 492 g, for 5 min at 4 °C and then resuspended in 1/15 of the initial volume in a 0.1 M of CaCl₂ and glycerol 15% (w/v). Aliquots were made and stored at -80 °C.

2.3. Phenolic extraction of chromosomal DNA

To extract the chromosomal DNA from *B. subtilis* 168T⁺, a single colony was selected and grown O/N, at 37 °C, 150 rpm in LB medium with the appropriate antibiotics. The O/N culture was centrifuge at 3427 g for 6 min and supernatant was discarded. Pellets were washed in the same volume of 50 mM Tris-HCl, 5 mM EDTA, pH 8 and were recentrifuged for 6 min at 6427 g. Supernatant was discarded, and pellets were resuspended in 150 µL of 50 mM Tris-HCl, 5 mM EDTA, pH 8, supplemented with lysozyme (1 mg/mL) and RNase (20 µg/mL) with posterior incubation at 37 °C for 30 min. After incubation, SDS was added to 1 % (w/v) and samples were incubated at 37 °C for 30 min.

100 µL of phenol was added (saturated with Tris-HCl pH 8.0) and mixed gently to complete homogenization. Samples were centrifuged for 10 min at 16060 *g*, and the upper phase was recovered. The phenolic extraction was repeated. 100 µL of chloroform: isoamyl alcohol (24:1) was added and samples were gently mixed to complete homogenization. After 10 min of centrifugation at 16 060 *g* (Heraeus Biofuge Pico, Massachusetts, USA), aqueous phase was recovered, and 2 volumes of absolute ethanol were added. Samples were centrifuge for 45 min at 15 493 *g* and 4 °C. The precipitate was washed with ethanol 70 %, dried at 45 °C and resuspended in 30 µL of water.

2.4. DNA manipulation

Routine DNA procedures were performed as described by Sambrook *et al.*, 1989. All restriction enzymes used in this work were purchased from Thermo Fisher Scientific TM (Massachusetts, USA) and used according to manufacturers' recommendations. Phusion High-Fidelity DNA Polymerase was used for PCR amplification (My Cyclor, Bio-Rad, California, USA). Colony PCRs were performed using NZYTaq II DNA polymerase (NZYTech, Lisbon, Portugal). IllustraTM GFX PCR DNA and Gel Band Purification kit (GE Healthcare Life Sciences, Illinois, USA) was used to purify DNA from enzymatic reactions and agarose gels. DNA ligations were carried out using T4 DNA Ligase (Thermo Fisher Scientific TM) and DNA dephosphorylation was performed with FastAP Thermosensitive Alkaline Phosphatase, also from Thermo Fisher Scientific TM. NZYMiniprep kit (NZYTech) was used to extract and purify Plasmid DNA. Primers used in this work were purchased from STAB VIDA or Metabion and plasmid DNA sequencing was performed at STAB VIDA, Lda using Sanger sequencing.

2.5. Plasmid construction for protein-protein interaction assays

All plasmids described in this work were constructed in *E. coli* XL-10 Gold (Stratagene, Tet^r Δ (*mcrA*)183 Δ (*mcrCB-hsdSMR-mrr*)173 *endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte* [F' *proAB lacI^qZ ΔM15 Tn10* (Tet^r) Amy Cam^r]), and derived from plasmids pUT18, pUT18C, pKNT25 or pKT25. Both pUT18 and pUT18C are high copy number vectors, encoding for T18 fragment of adenylate cyclase of *B. pertussis* and ampicillin resistance marker. However, in pUT18 plasmid, the MCS is upstream of this fragment, allowing the construction of chimeric proteins fused to the N-terminus of T18 and in pUT18C, the MCS is downstream of T18, allowing the construction of hybrid proteins fused to the C-terminus of this fragment. pKNT25 and pKT25 are low copy number vectors, expressing T25 fragment of adenylate cyclase of *B. pertussis* and kanamycin resistance marker. Similarly, pKNT25 has a MCS clone downstream of T25 fragment, allowing the design of hybrid proteins fused to the N-terminus of T25 and pKT25 has a MCS upstream of T25, allowing the creation of chimeric proteins fused to the C-terminus of this fragment. All vectors are expressed under the

transcriptional control of the *lac* promoter. Used or designed plasmids and oligonucleotides are displayed in Table 1 and Table 2, respectively.

Plasmid pCA1 was obtained by PCR amplification of the *malK* from chromosomal DNA of *E. coli* K-12, using oligonucleotides Ara858 and Ara974, containing the restriction sites HindIII and XbaI, respectively. The resulting fragment (1173 bp) was cloned into plasmid pUT18, between HindIII and XbaI sites. Plasmid pCA2 was obtained by overlap PCR. Chromosomal DNA of *B. subtilis* 168T⁺ was used as template for amplification of *araP* fragments, A and B. Fragment A (678 bp) was obtained using the oligonucleotides Ara934 and Ara976 and fragment B (337 bp) was obtained using oligonucleotides Ara935 and Ara975. Oligonucleotides Ara934 and Ara935 contain XbaI restriction site and oligonucleotides Ara976 and Ara975 allow the mutagenesis of amino acids A210 and D211 to S210 and A211. The resulting fragment (993 bp) was then cloned into plasmid pKNT25, between XbaI restriction sites. Plasmid pCA3 was obtained by overlap PCR (Figure 18). Chromosomal DNA of *B. subtilis* 168T⁺ was used as template for the amplification of two *araP* fragments, A and B. Fragment A was amplified using oligonucleotides Ara934 and Ara980 (647 bp), bearing XbaI restriction site and *malF* overhang sequence, respectively. Fragment B was amplified using oligonucleotides Ara935 and Ara979 (388 bp), bearing XbaI restriction site and a *malF* overhang sequence, respectively. For fragment C amplification (135 bp), chromosomal DNA from *E. coli* K-12 was used and oligonucleotides Ara977 and Ara978, containing *araP* overhang sequences. The resulting fragment (987 bp) of the overlap PCR was cloned into plasmid pKNT25, between XbaI restriction sites. Plasmid pCA4 was obtained by PCR amplification of *ycjV* from the pPS13 plasmid, containing the YcjV protein, from *E. coli* K-12, with a full length of 360 amino acids, using oligonucleotides Ara860 and Ara982, bearing HindIII and XbaI restriction sites. The resulting fragment (1144 bp) was cloned between HindIII and XbaI sites of plasmid pUT18. Plasmid pCA5 was obtained by site-directed mutagenesis, using plasmid pLG64, harbouring the *araQ* gene from *B. subtilis* 168T⁺ fused to the C-terminus of T25, as template and oligonucleotides Ara987 and Ara988, allowing the change of the first and second nucleotide of the codon 178 from AG to GC, changing an arginine into an alanine (R178A). Plasmid pCA7 was obtained by nested PCR. Chromosomal DNA of the pathogenic strain *Staphylococcus aureus* subsp. *aureus* ST398 was used as a template for amplification of *ugpC*, using oligonucleotides Ara746 and Ara747 (1215 bp). This PCR product was then used as template for amplification of the *ugpC*, using oligonucleotides Ara993 and Ara994 (1146 bp), both containing XbaI restriction site. The resulting fragment was cloned into plasmid pUT18, between XbaI restriction sites. Plasmid pCA8 was also obtained by nested PCR. Chromosomal DNA of the pathogenic strain *Clostridioides difficile* 630 was used as template for the amplification of *ABC_Cd*, using oligonucleotides Ara885 and Ara996 (1161 bp). This PCR product was then used as template for amplification of the *ABC_Cd*, using oligonucleotides Ara995 and Ara996 (1168 bp). The resulting fragment was cloned into plasmid pUT18, between XbaI restriction sites. Plasmid pCA9 was obtained by PCR amplification, using plasmid pIG3, harbouring the *ABC_Syn* from *Synechocystis* sp., as template and oligonucleotides

Ara741, bearing SalI restriction site, and Ara1017, bearing BamHI restriction site. The resulting fragment (1174 bp) was then cloned into plasmid pUT18, between SalI and BamHI restriction sites. Plasmid pCA11 was obtained by site-directed mutagenesis, using pLG64 plasmid as template, which harbours the *araQ* from *B. subtilis* 168T⁺. Oligonucleotides Ara671 and Ara672 allow the change in the second nucleotide of codon 180 from A to C, changing the amino acid aspartate to alanine (D180A). Plasmid pCA12 was obtained by overlap PCR. Chromosomal DNA from *B. subtilis* 168 T⁺ was used for amplification of fragments A and B. Fragment A (1129 bp) was amplified using oligonucleotides Ara680 and Ara674, containing a deletion of the last 4 C-terminal amino acids (Δ GVKG). Fragment B (827 bp) was amplified using oligonucleotides Ara675, containing the C-terminal deletion, and Ara676. Fragment C (1896 bp) was obtained by overlap PCR of fragments A and B, using oligonucleotides Ara680 and Ara676. Fragment C (872 bp) was used as template for *araQ* amplification with a deletion of 4 C-terminal amino acids, using oligonucleotides Ara951, bearing the restriction site for BamHI, and Ara674, bearing the C-terminal deletion mutation. The fragment was cloned into plasmid pKT25 between BamHI and SmaI restriction sites. Plasmid pCA16 was obtained by Site-Directed mutagenesis,

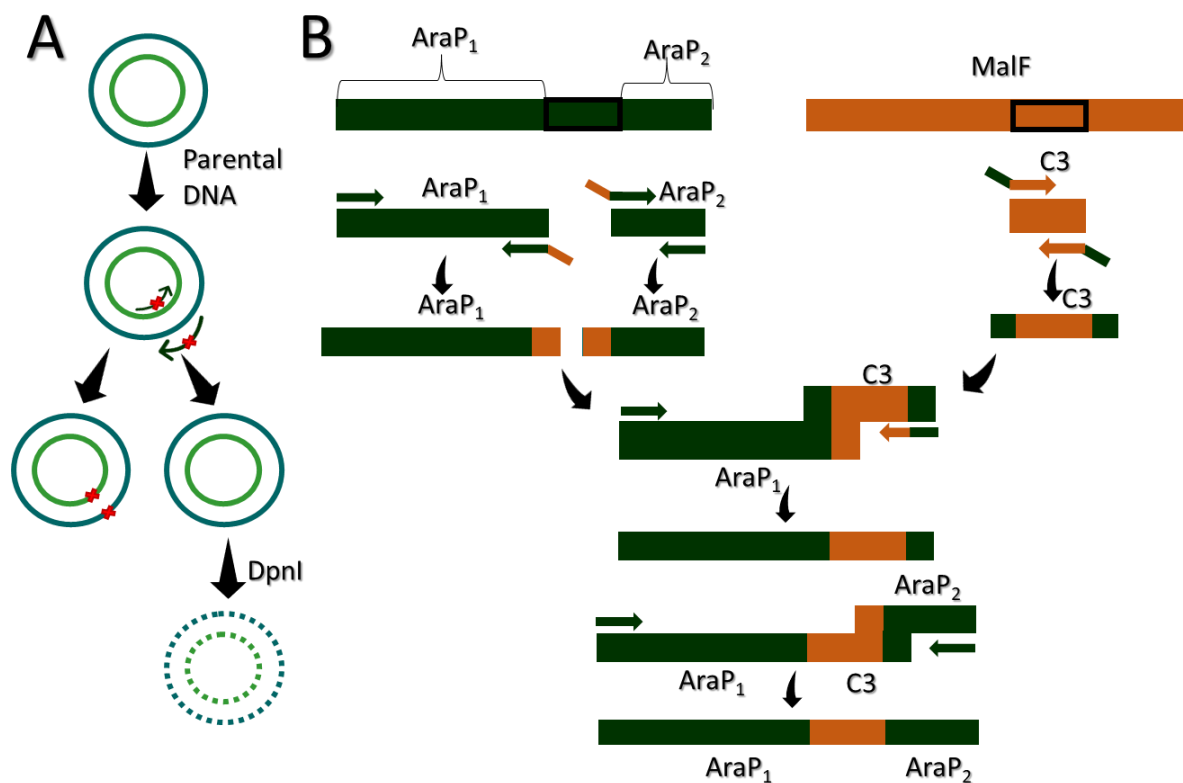


Figure 18. Schematic representation of two strategies of mutagenesis PCR. A) Site-Directed mutagenesis. In one PCR reaction, by adding mutating oligonucleotides (black arrows with red crosses), a specific mutation (red cross) is inserted into a DNA sequence (concentric green and blue circle). Parental DNA is then digested with DpnI restriction enzyme (dotted circles). B) Site-Directed mutagenesis performed by overlapping PCR (pCA3 construction). First, each site-directed mutagenesis is carried out by using a mutagenic oligonucleotide (green and orange arrows) and a flanking oligonucleotide (represented as green arrows, or in the case of *malF* sequence, represented in orange, two mutagenic oligonucleotides). Then, through overlapping PCR, two fragments, *araP* (represented in green) and *malF* are joined by its homology zone and amplified using one mutagenic primer and one flanking oligonucleotide. Then this fragment is joined by its homology zone with the second *araP* fragment and amplified by using two flanking oligonucleotides.

using plasmid pLG63 as template. This plasmid harbours the *araQ* from *B. subtilis* 168T⁺ fused to the C-terminal end of T18 fragment. Oligonucleotides Ara1014 and Ara1015 were used to introduce a stop codon in the 282nd amino acid, changing nucleotides GT to TA. Plasmid pCA17 was obtained using Fragment C from pCA12 as template, using oligonucleotides Ara950 and Ara674, bearing SalI restriction site and ΔGVKG, respectively. The resulting fragment (864 bp) was then cloned into plasmid pUT18C, between SalI and SmaI restriction sites.

Table 1. List of used and constructed plasmids mentioned in this work for protein-protein interaction studies using BACTH system. Map of these plasmids is displayed in Subchapter 6. Appendices.

Plasmids	Relevant Construction	Source or Reference
pUT18	Derivative of the high copy number vector pUC19, encoding the T18 fragment	BACTH System Kit (Euromedex)
pUT18C	Derivative of the high copy number vector pUC19, encoding the T18 fragment	BACTH System Kit (Euromedex)
pKNT25	Derivative of the low copy number vector pSU40, encoding the T25 fragment	BACTH System Kit (Euromedex)
pKT25	Derivative of the low copy number vector pSU40, encoding the T25 fragment	BACTH System Kit (Euromedex)
pUT18_Zip	pUT18 derivative, with the leucine zipper of GCN4 cloned in the N-terminal of T18	BACTH System Kit (Euromedex)
pKNT25_Zip	pKNT25 derivative, with the leucine zipper of GCN4 cloned in the N-terminal of T25	BACTH System Kit (Euromedex)
pLG55	pUT18 derivative, with <i>msmX</i> cloned in the N-terminal of T18	Godinho L. M. (unpublished work)
pLG56	pKNT25 derivative, with <i>msmX</i> cloned in the N-terminal of T25	Godinho L. M. (unpublished work)
pLG63	pUT18C derivative, with <i>araQ</i> cloned in the C-terminal of T18	Godinho L. M. (unpublished work)
pLG64	pKT25 derivative, with <i>araQ</i> cloned in the C-terminal of T25	Godinho L. M. (unpublished work)
pJA2	pUT18 derivative, with <i>araP</i> cloned in the N-terminal of T18	Almeida J. (unpublished work)
pJA4	pKNT25 derivative, with <i>araP</i> cloned in the N-terminal of T25	Almeida J. (unpublished work)
pJA20	pJA2 derivative, with D213A mutation in <i>araP</i> sequence cloned in the N-terminal of T18	Almeida J. (unpublished work)
pCA1	pUT18 derivative, with <i>malK</i> cloned in the N-terminal of T18	This work*
pCA2	pJA4 derivative, with A210S/D211A mutations in <i>araP</i> sequence	This work*
pCA3	pJA4 derivative, with the loop containing the EAA motif in <i>araP</i> sequence replaced with C3 loop of <i>malF</i>	This work*
pCA4	pUT18 derivative, with <i>ycjV</i> cloned in the N-terminal of T18	This work*
pCA5	pLG64 derivative, with R178A mutation in <i>araQ</i> sequence	This work*
pCA7	pUT18 derivative, with <i>ABC_Sa</i> cloned in the N-terminal of T18	This work*
pCA8	pUT18 derivative, with <i>ABC_Cd</i> cloned in the N-terminal of T18	This work*
pCA9	pUT18 derivative, with <i>ABC_Syn</i> cloned in the N-terminal of T18	This work*
pCA11	pLG64 derivative, with D180A mutation in <i>araQ</i> sequence	This work*
pCA12	pLG64 derivative, with ΔGVKG mutation in <i>araQ</i> sequence	This work*
pCA16	pLG63 derivative, with a codon stop inserted in the end of <i>araQ</i>	This work*
pCA17	pLG63 derivative, with ΔGVKG mutation in <i>araQ</i> sequence	This work*

Table 2. List of all oligonucleotides used for this work for determining protein-protein interaction using the BACTH system. Restriction sites are underlined, and mutations are in lowercase letters.

Oligonucleotides	Sequence 5'→3'
Ara671	CGTACAGCCGgCCATCCTTGC
Ara672	GCAAGGATGGcCGGCTGTACG
Ara674	CATCACACGTTTCCTCCTTACCTACCGTCAGGCCGGAGATAAA
Ara675	TTTATCTCCGGCCTGACGGTAGGGTAATGAAGGAGGAAACGTGTGATG
Ara676	GCCA <u>ACGCGT</u> TTGGCAGATTGT
Ara680	GTTTCTCAAGCCTGTAACCG
Ara741	CTTGTGTCGACAGGGAATTGCTG
Ara746	CGGGTCGACACGAAGTGTATTGC
Ara747	ATACGCATGCCACGGCTAACGTG
Ara858	CGGAAGCTTGCTGTGATGACAGG
Ara860	TGGCCAAGCTTATCGGCCTTCTG
Ara885	GGCGTCGACAAAATTTAAATATCAAGG
Ara934	TAGAATCCCAT <u>TCTAG</u> AAGTAAAG
Ara935	TGTGCCGCAACTCTAGACCCTCCCC
Ara950	TTAAGGGGGAGTCGACAATGTTGCGG
Ara951	TTAAGGGGGAGGATCCAATGTTGCGG
Ara974	TTTGTGGGGT <u>TCTAG</u> AACGCCCG
Ara975	TACGAAGCCtCaGcaATAGACGGCGC
Ara976	CGTCTATtgCtGaGGCTTCGTACAGCTC
Ara977	atgggaatcaacatcctttacttttaCCGGACGATTTGTATGAAGCC
Ara978	tacggttacaggcttgagaaaTTTAATCAGCAGCGGCAG
Ara979	ctgccgctgctgattaaaTTTCTCAAGCCTGTAACC
Ara980	ggcttcatacaaatcgctccgGTAAAAAGTAAAGGATGTTGAT
Ara982	TTCCCCCTCTAGATAGCGAATTGCTA
Ara987	TGGACTCTGCAGcGATGGACGGCTG
Ara988	GCCGTCCATCgTGcAGAGTCCAGC
Ara993	GCACAATCAAATTGTCTAGAAAGTATGGG
Ara994	ATACTCCCTCTAGAACGATACGATTTCC
Ara995	GATACTATATTCTAGAAATTTAAATATC
Ara996	AAATAAGATTATATCTAGACGTATAGCTAAAG
Ara1014	GAGTCAAAGGGtaACCGAGCTCGAA
Ara1015	AATTCGAGCTCGGTtaCCCTTTGAC
Ara1017	GGTGATGGGGATCCTCGAGTCG

2.6. Alkaline lysis for plasmid extraction

Cells were grown O/N in LB medium supplemented with the respective antibiotic and harvested by centrifugation, 3427 g for 6 min. After pellet resuspension, 100 μ L of solution I (50 mM glucose, 25 mM tris-HCl and 10 mM EDTA pH 8) were added. After 5 min of incubation at room temperature, 200 μ L of solution II (0.2 M NaOH and SDS 0.1 % (w/v)) were added and after a gentle mixture, samples were incubated for 5 min on ice. 150 μ L of solution III (5 M potassium acetate and acetic acid) were added and samples were incubated for 5 min on ice. After 10 min centrifugation at full speed, pellets were discarded. 2 volumes of 100% ethanol were added to supernatants and samples were incubated on ice for 10 min. Samples were centrifuged at full speed for 10 min and supernatants were discarded. Pellets were washed 3 times with 70% ethanol. After the pellets dried, they were resuspended in 30 μ L of TE 1X with RNase (20 μ L/mL).

Plasmid extraction was also performed using NZYMiniprep kit (NZYTech), according to manufacturer's recommendations.

2.7. Construction of *E. coli* BTH 101 for qualitative and quantitative protein-protein interaction

In order to quantitatively and qualitatively study protein-protein interactions using the BACTH system, *E. coli* BTH101 strains (Euromedex, F⁻, *cya*-99, *araD139*, *galE15*, *galK16*, *rpsL1* (*Str*^r), *hsdR2*, *mcrA1*, *mcrB1*) were constructed by performing bacterial transformation by heat shock (adapted from Sambrook *et al.*, 1989) using plasmids described in Table 1. For each co-transformation, one constructed plasmid encoding for the T18 fragment and one plasmid encoding for the T25 fragment were used in equal quantities (~50 ng). Since this *E. coli* strain lacks endogenous adenylate cyclase activity, the co-transformation with the two plasmids allows the restoration of adenylate cyclase enzymatic activity if interaction is present. All co-transformations are listed in Table 3.

2.8. Qualitative analysis of protein-protein interaction

To qualitatively test interaction between two proteins of interest, co-transformed *E. coli* BTH101 strains bearing two fusion proteins fused to T18 and T25 fragments, respectively, were grown O/N (37 °C, 150 rpm) in LB medium supplemented with kanamycin (30 μ g/mL), ampicillin (100 μ g/mL), streptomycin (100 μ g/mL) and IPTG (0.5 mM) in an orbital water-bath shaker (Aquatron INFORS HT, Switzerland). 3 μ L of each culture was used to perform spots in LA plates, supplemented with kanamycin (30 μ g/mL), ampicillin (100 μ g/mL), streptomycin (100 μ g/mL), IPTG (0.5 mM) and X-Gal (40 μ g/mL). Plates were incubated O/N at 30 °C, after each spot dried. Coloration of each spot was observed after incubation.

Table 3. List of all *E. coli* BTH101 co-transformations constructed for protein-protein interaction studies using the BACTH system.

Plasmid encoding T18 fragment	Plasmid encoding T25 fragment
pZip18 (Zip + pUT18)	pZip25 (Zip + pKNT25)
pLG55 (MsmX + pUT18)	pZip25 (Zip + pKNT25)
pLG55 (MsmX + pUT18)	pJA4 (AraP + pKNT25)
pLG55 (MsmX + pUT18)	pLG64 (pKT25 + AraQ)
pLG55 (MsmX + pUT18)	pCA2 (AraP_A210S/D211A + pKNT25)
pLG55 (MsmX + pUT18)	pCA3 (AraP_MalF_Chimera_pKNT25)
pLG55 (MsmX + pUT18)	pCA5 (pKT25_AraQ_R178A)
pLG55 (MsmX + pUT18)	pCA11 (pKT25_AraQ_D180A)
pLG55 (MsmX + pUT18)	pCA12 (pKT25_AraQ_ΔGVKG)
pCA16 (pUT18C + AraQ)	pLG56 (MsmX_pKNT25)
pCA17 (pUT18C + AraQ_ΔGVKG)	pLG56 (MsmX_pKNT25)
pCA1 (MalK + pUT18)	pJA4 (AraP + pKNT25)
pCA1 (MalK + pUT18)	pLG55 (pKT25 + AraQ)
pCA1 (MalK + pUT18)	pCA2 (AraP_A210S/D211A + pKNT25)
pCA1 (MalK + pUT18)	pCA3 (AraP_MalF_Chimera_pKNT25)
pCA4 (YcjV + pUT18)	pJA4 (AraP + pKNT25)
pCA4 (YcjV + pUT18)	pLG64 (pKT25 + AraQ)
pCA7 (ABC_Sa + pUT18)	pJA4 (AraP + pKNT25)
pCA7 (ABC_Sa + pUT18)	pLG64 (pKT25 + AraQ)
pCA8 (ABC_Cd + pUT18)	pJA4 (AraP + pKNT25)
pCA8 (ABC_Cd + pUT18)	pLG64 (pKT25 + AraQ)
pCA9 (ABC_Syn + pUT18)	pJA4 (AraP + pKNT25)
pCA9 (ABC_Syn + pUT18)	pLG64 (pKT25 + AraQ)

2.9. Quantitative analysis of protein-protein interaction: β -Galactosidase assays

To quantitatively assess interaction between two hybrid proteins, β -galactosidase assays were used to measure β -galactosidase activity in liquid cultures. Here, co-transformed *E. coli* BTH101 cells (see Table 3) were grown in 5 mL of LB supplemented with IPTG (0.5 mM), ampicillin (100 μ g/mL); kanamycin (30 μ g/mL) and streptomycin (100 μ g/mL) at 30 °C, 180 rpm in an orbital water-bath shaker (Aquatron INFORS HT, Switzerland) for 17 h. Liquid cultures were diluted 1:5 in M63 medium 1 X (M63 5X: 10 g/L (NH₄)₂SO₄, 68 g/L KH₂PO₄, 2.5 mg FeSO₄·7H₂O and pH adjusted to 7.0 with KOH) supplement with 5 mg/L thiamine and optical density at 600 nm was determined. 2 mL of the culture were used and the bacterial cells permeabilized by adding 24 μ L of toluene and 24 μ L of 0.1 % SDS

and tubes were vortexed for 10 sec. Tubes were kept open in a fume hood to allow toluene evaporation. After 10 min, tubes were shaken at 37 °C, for 30 min at 200 rpm in an orbital water-bath shaker (Aquatron INFORS HT, Switzerland). Meanwhile, PM2 buffer (70 mM Na₂HPO₄, 30 mM NaH₂PO₄, 1 mM MgSO₄ and 0.2 mM MnSO₄) and ONPG buffer were prepared. For the enzymatic reaction, 100 µL of permeabilized cells were added to 900 µL of PM2 buffer with β-mercaptoethanol (100 mM) and then equilibrated at 28 °C (FB15101 dry bath heating shock, Thermo Fisher Scientific TM) for 5 min. The enzymatic reaction was started by adding 250 µL of ONPG (4 mg/mL) in the PM2 buffer and after sufficient yellow colour was developed, the reaction was stopped by adding 500 µL of 1M Na₂CO₃ solution. Tubes were centrifuged at maximum speed for 5 min at room temperature and then absorbance was recorded at 420 nm. Enzymatic activity, A (units/mL) were calculated using the following formula:

$$Activity \left(\frac{U}{mL} \right) = 200 \times \frac{Abs_{420}}{time \text{ (min)}} \times dilution \text{ factor}$$

$$Activity \left(\frac{U}{mg} \right) = \frac{A \left(\frac{U}{mL} \right)}{mg \text{ of bacteria in the used volume}}$$

200 – Corresponds to the inverse of the absorption coefficient of o-nitrophenol, that is 0.005 per nmol/ml at pH 11.0 (i.e. after addition of Na₂CO₃) (Euromedex).

2.10. Protein extracts and Western Blot analysis

Immunoblot assays were performed using different techniques to optimize detection. First, total protein extracts were made to detect the accumulation of proteins fused to T18. Here, *E. coli* BTH101 containing T18 and T25 fusions (Table 3) were grown O/N (~17h) in LB medium supplemented with IPTG (0.5 mM), ampicillin (100 µg/mL); kanamycin (30 µg/mL) and streptomycin (100 µg/mL) at 30 °C, 180 rpm (Aquatron INFORS HT, Switzerland). O/N cultures were recovered by centrifugation at 16 060 g (Heraeus Biofuge Pico, Massachusetts, USA) for 1 minute. Pellets were resuspended in 100 µL Lysis Buffer (Saline Phosphate Buffer 20 mM pH 7.4, NaCl 62.5 mM, imidazole 100 mM and 10% glycerol) and lysozyme was added to a final concentration of 1 mg/mL. After incubation at 37 °C for 10 min, three cycles of freezing in liquid nitrogen and thawing at 37 °C for 5 min were performed. PMSF (10 mM) and Benzonase nuclease (2.5 U) (Sigma-Aldrich, Missouri, USA) were added, followed by 25 min of incubation at 37 °C. Protein quantification was performed as described in Bio-Rad Protein Assay. Five BSA dilutions (2.5 µg/mL; 5 µg/mL; 7.5 µg/mL; 10 µg/mL and 15 µg/mL) were used to construct a standard curve, in triplicates. Assays were performed by adding 200 µL of Bradford reagent to 800 µL of each standard or sample and after incubation at room temperature for 5 min, absorbance at 595 nm was measured.

For immunoblotting assays, 20 µg of cells extracts, cells grown O/N, or cells grown until O.D._{600nm} reached 0.3 - 0.5 were washed with Lysis Buffer (20 mM PBS pH 7.4, 62.5 mM NaCl, 100 mM imidazole and 10% glycerol) and PMSF (10 mM) and then resuspended in Running buffer (5X:

0.124 M Tris Base, 0.96 M glycine and 0.017 M SDS) and Loading Buffer (2X: 0.125 M Tris-HCl pH 6.8, 20% glycerol, 4% SDS, 10% β – mercaptoethanol, 0.004% bromophenol blue) (Sambrook *et al.*, 1989) were loaded in 12.5 % SDS-PAGE. Separated proteins were electrotransferred into a nitrocellulose membrane (Bio-Rad) in a mini Trans-Blot[®] cell (Bio-Rad, California, USA) for 1 h. To verify the transference quality, membranes were stained with Ponceau Red and subsequently incubated with blocking solution (1x TBS-Tween20 with 5% of low-fat powder milk – Molico Nestlé) for 30 min at room temperature or O/N at 4 °C. Membrane was washed for 15 min with 1x TBS-Tween20 at room temperature and later incubated with primary antibody solution [1x TBS-Tween20 + 0.5% powder milk + 1:1000; 1:2000; 1:4000 α -Cya (3D1 Kerafast)] O/N at 4 °C or for 2 h at room temperature. After two washes with 1X TBS-Tween20 for 10 min each, incubation with the secondary antibody (1x TBS-Tween20 + 0.5% powder milk + 1:10 000 HRP-conjugated goat anti-mouse IgG antibody [Jackson ImmunoResearch Europe Ltd.]) occurred for 30 min at room temperature followed by three washes of 10 min with 1x TBS-Tween. All subsequent steps were performed in a dark room. Blots were developed using ECL Western Blot Substrates (Thermo Fisher Scientific TM) and autoradiography plates (Cronex or GE Healthcare) were exposed with the adequate amount of time.

Chapter 3

Results and Discussion

3. Results and Discussion

3.1. Targeting key residues in TMDs AraPQ for interaction studies using BACTH system

To characterize and map interacting regions between multitask ATPases and TMDs and consequently, identify key residues in TMDs associated with the capacity to be energized by several multitask NBDs, ABC type I importer system AraNPQ-MsmX was used as model to study the influence of mutant and chimeric mutations in TMDs AraPQ on their interaction with MsmX. An *in silico* multiple sequence alignment of different TMDs from *B. subtilis* energized by MsmX and the MalFG TMDs of the prototype MalEFGK₂, from *E. coli*, was made (Figure 19 and Figure 21), with the alignment only showing the predicted MsmX-TMD interaction region.

In AraP, by site-directed mutagenesis the EAA signature of loop C3 was altered. As seen in Figure 19, all TMDs energized by MsmX have an EAA sequence, while MalF, which cannot be energized by MsmX has an EASA sequence. It is known that this motif plays an important role in interaction (Oldham *et al.*, 2007). With this in mind, the alanine in position 210 (A210) and the aspartic acid in position 211 (D211) were simultaneously change to a serine (S210) and to an alanine (A211), to mimic the sequence in MalF. Although D211, is not conserved in any TMDs energized by MsmX, they present either charged amino acids, have polar uncharged side chains or have aromatic side chains, while in MalF it is an alanine, a small, nonpolar and simple amino acid.

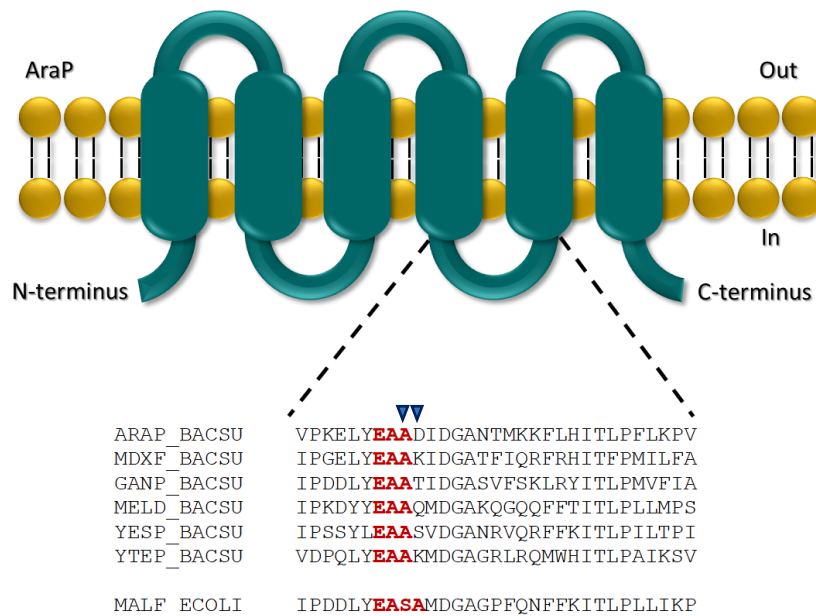


Figure 19. AraP TMD from AraNPQ-MsmX system. Alignments of TMDs from *B. subtilis* belonging to transporters energized by MsmX, Ara, Mdx, Gan, Mel, Yes, Yte and *E. coli* Mal ABC transporter were performed using Clustal W2 and are partially show. EAA sequence is display in red and schematic AraP representation in cyan blue. Residues targeted for mutagenesis are indicated with arrows. Adapted from Ferreira *et al.*, 2017

In addition, a chimeric AraP was also constructed (Figure 20). Here, the cytoplasmic loop containing the EAA motif of AraP was replaced by the cytoplasmic C3 loop of MalF, to assess the effect of the entire EAA motif change in its interaction with MsmX or MalK, *E. coli*'s ATPase (Daus *et al.*, 2007). These key residues were changed by site-directed mutagenesis performed by overlapping PCR (Figure 18).

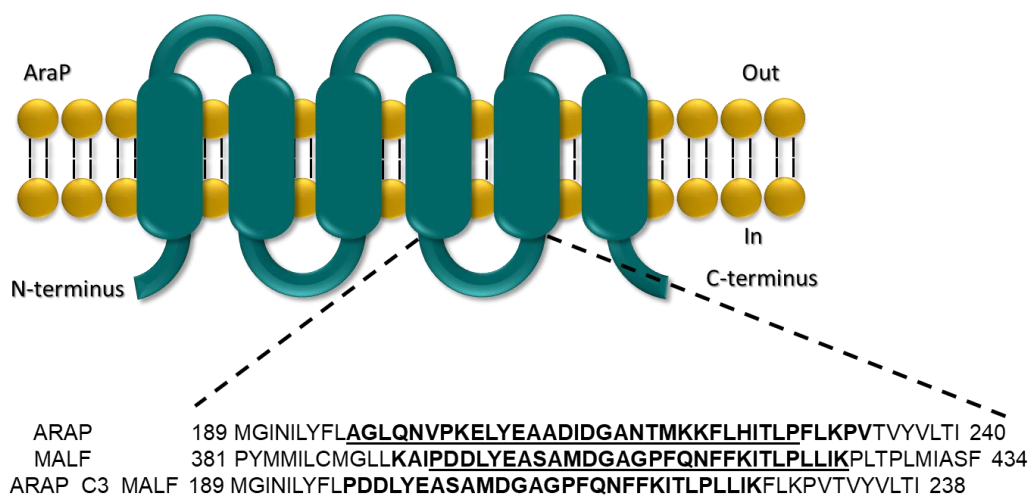


Figure 20. Construction of a chimeric AraP protein. The entire loop containing the EAA motif from AraP (*B. subtilis*) was replaced by the C3 loop from MalF protein (*E. coli*). All amino acids part of cytoplasmic loop are displayed in bold, while changed amino acids are underline.

Previous work had also identified several key residues in AraQ important for the function of the full AraNPQ-MsmX transporter (Ferreira *et al.*, 2017). Here, six amino acids were found of interest. The first amino acid was the aspartic acid in position 180 (D180) (Ferreira *et al.*, 2017). This amino

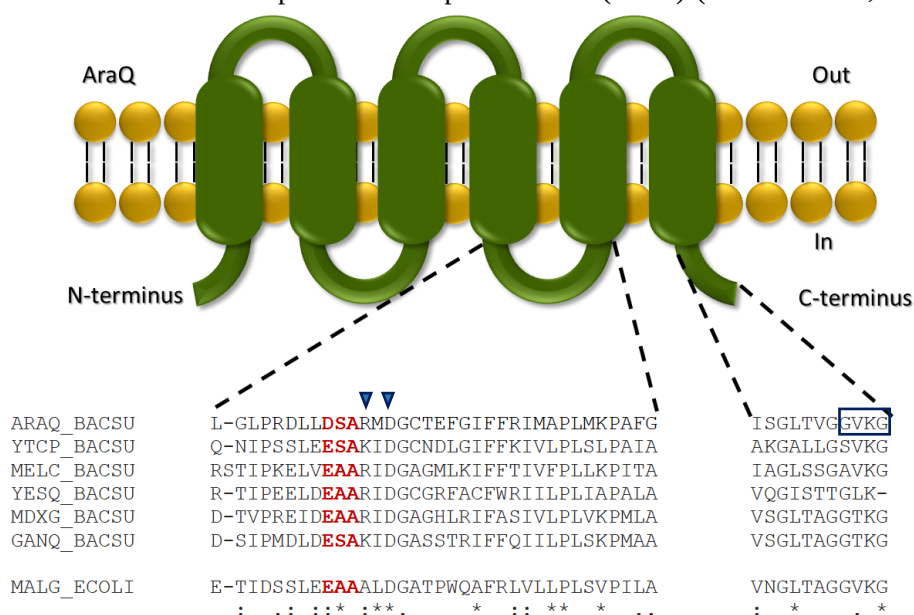


Figure 21. AraQ TMD from the AraNPQ-MsmX system. Alignments of TMDs from *B. subtilis* belonging to transporters energized by MsmX, Ara, Mdx, Gan, Mel, Yes, Yte and *E. coli* Mal ABC transporters were performed using Clustal W2 and are partially show. “EAA” sequence is display in red and schematic AraQ representation in green. Residues targeted for mutagenesis are indicated with arrows and residues deleted are boxed. Adapted from Ferreira *et al.*, 2017.

acid is conserved in all TMDs energized by MsmX and is also conserved in MalG. The last four amino acids of the C-terminal tail glycine-valine-lysine-glycine (GVKG) were also targeted. This tail was proposed to have important functions in the translocation mechanism in MalEFGK₂ (Oldham *et al.*, 2007) and was shown to be important for the function of AraNPQ-MsmX in *B. subtilis* (Ferreira *et al.*, 2017). As depicted in Figure 21 the C-terminal tail is highly conserved in all TMDs energized by MsmX. Thus, the study of these amino acids is of relevance to understand their role in the interaction between AraQ and MsmX. The last amino acid target was studied for the first time in this dissertation. The arginine in position 178 (R178) is conserved in all TMDs from *B. subtilis* energized by MsmX (Figure 21), except for GanQ and YtcP, which have a lysine (K) in that position, an amino acid with identical characteristics to arginine and positively charged. In MalG this amino acid corresponds to an alanine. To test the importance of these amino acids, they were targeted by site-directed mutagenesis or by site-directed mutagenesis performed by overlapping PCR (Figure 18), being the R178 and D180 changed to alanine (R178A and D180A), while the C-terminal tail was deleted (Δ GVKG), as previously described in Ferreira *et al.*, 2017.

In order to understand, *in vivo*, the impact of these mutations in the ability of MsmX to interact with AraP and AraQ mutants, the Bacterial Adenylate Cyclase Two-Hybrid (BACTH) system of Euromedex was used. All plasmids were constructed as described in subchapter 2.5 and are listed in Table 1, with all interactions being displayed in Table 3.

MsmX was cloned in plasmid pUT18, being replicated in high copy number (ColE1 origin) and AraP was cloned in plasmid pKNT25, being replicated in low copy numbers (p15A origin) in the cell. Since the C-terminal tail plays an important role in interaction for both, MalG (Oldham *et al.*, 2007) and AraQ (Ferreira *et al.*, 2017), all AraQ mutants were constructed in plasmid pKT25 (present in low copy number), fusing the T25 fragment to the N-terminus of AraQ. Posteriorly, wild-type AraQ and AraQ_ Δ GVKG were also constructed in plasmid pUT18C (present in high copy number). All plasmids share the same promoter, and each gene carries its own RBS (except for T18 and T25 expressed in pUT18C and pKT25, which harbour a ribosome binding site) (Ouellette, Karimova, Davi, & Ladant, 2017). In this work, only one combination of hybrid proteins was tested, since this methodology was already optimized in our laboratory, to allow detection of the effect of mutant proteins (Microbial Genetics Laboratory, unpublished results).

With the intent of studying protein-protein interactions (PPIs), qualitative and quantitative β -galactosidase assays were performed. For that, T18 and T25 fusion proteins were co-transformed by heat shock into *E. coli* BTH101, which lacks an endogenous adenylate cyclase activity. Since this assay is based on *lacZ* (reporter gene responsible for β -galactosidase production) expression, and it is positively regulated by cAMP/CAP complex, only when the two hybrid proteins interact (and consequently occurs T18 and T25 complementation), the catalytic domain of adenylate cyclase from *B. pertussis* is reconstructed, allowing the production of cAMP from ATP and eventually activation of *lacZ*. In this assay, IPTG (Isopropyl β -D-1-thiogalactopyranoside) must be added to prevent LacI

(lactose repressor) binding to the operator and blocking *lacZ* or hybrid proteins transcription. If the hybrid proteins do not interact, there are only basal levels of β -galactosidase being produced (Karimova *et al.*, 1998, 2017).

3.1.1. Qualitative analysis of MsmX and AraPQ interaction

A simple way to qualitatively assess the production of β -galactosidase is through an assay in solid medium. Here, *E. coli* BTH101 cells are co-transformed with both plasmids, one containing a fusion protein with T25 fragment and the other containing a fusion protein with the T18 fragment, as described in subchapter 2.7. These co-transformants are then plated in LA plates supplemented with X-Gal (5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside), and IPTG as described as subchapter 2.8. If the two fusion polypeptides interact, then adenylate cyclase is reconstructed, and cAMP is produced from ATP (Figure 22A). The cAMP-CAP complex will then activate the transcription of a reporter gene, in this case, *lacZ* gene, responsible for β -galactosidase production. β -galactosidase can cleave X-Gal into β -D-galactose and 5-bromo-4-chloro-3-hydroxyindole. This last component is then oxidized into 5,5'-dibromo-4,4'-dichloro-indigo, a blue component, which means that colonies, expressing interacting proteins can be distinguished by their blue colour (Juers *et al.*, 2012; Karimova *et al.*, 2017). However, if the two proteins in study do not interact, there is no cleavage of X-Gal, and no blue colouration can be detected (Figure 22B).

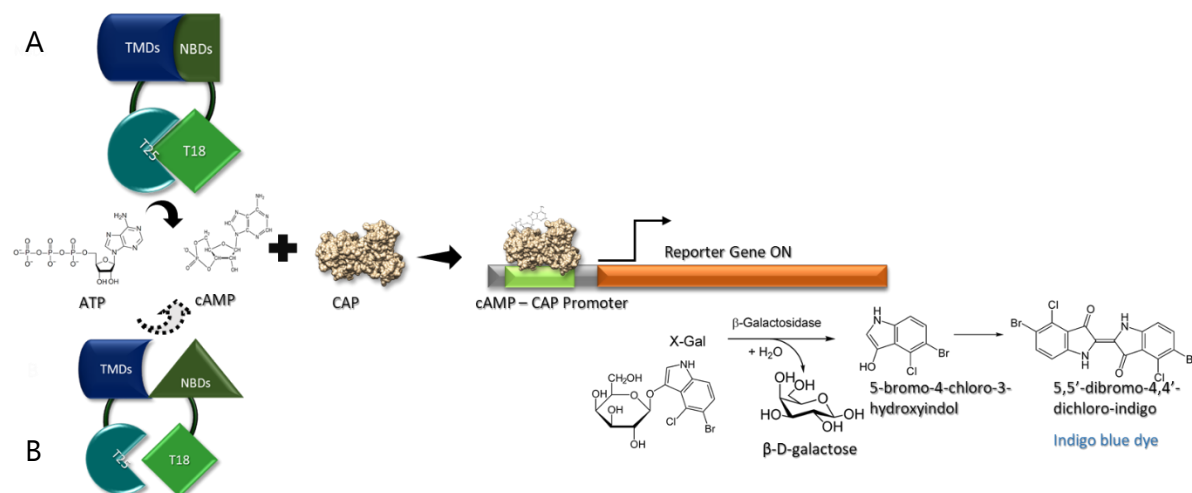


Figure 22. Principle of BACTH System in qualitative assay using X-Gal plates as an indicator media. A) When T25 and T18 fragments are co-expressed as fusions with two interacting polypeptides, there is association of these two chimeric proteins and consequently reconstitution of adenylate cyclase activity. This enzyme is going to produce cAMP that is going to interact with CAP protein and together are going to activate a reporter gene, in this case *lacZ*, responsible for β -galactosidase production. β -galactosidase will cleave X-Gal, a colourless compound, into β -D-galactose and 5-bromo-4-chloro-3-hydroxyindol, that will posteriorly be oxidise into 5,5'-dibromo-4,4'-dichloro-indigo. Thus, colonies will have blue coloration. B) When T25 and T18 fragments are co-expressed as fusion with two non-interacting polypeptides, there is no association and no reconstitution of adenylate cyclase. Since no β -galactosidase is produced, X-gal is not cleaved, and colonies will have their normal colouration.

Results are presented in Table 4. Here, the co-transformation Zip_T18 + Zip_T25 was used as positive control, displaying a strong blue colour, a phenotype expected for interacting hybrid proteins, while the co-transformation MsmX_T18 + Zip_T25, was used as negative control to establish the phenotype for non-interacting hybrid proteins, displaying an expected white colour. The wild-type phenotype was also assessed by studying the interaction between MsmX and AraPQ, through co-transformations MsmX_T18 + AraP_T25 and MsmX_T18 + T25_AraQ, both exhibiting a very strong blue colour, comparable to the Zip_T18 + Zip_T25 co-transformation, a phenotype of interacting proteins.

Table 4. Qualitative protein-protein interaction assays. Qualitative analysis of β -galactosidase activity using X-Gal as an indicator media. Blue colouration corresponds to a positive interaction while white colour corresponds to a non-interacting polypeptide. Zip_T18 + T25_Zip co-transformation and MsmX_T18 + T25_Zip were used as controls and MsmX_T18 + AraP_T25 and MsmX_T18 + T25_AraQ was used as controls for the wild-type interaction.

			T18 fragment			
			Plasmid pUT18		Plasmid pUT18C	
			Zip	MsmX	AraQ	AraQ (Δ GVKG)
T25 fragment	Plasmid pKNT25	Zip				
		MsmX				
		AraP				
		AraP EASA				
		AraP_MalF_Chimera				
	Plasmid pKT25	AraQ				
		AraQ D180A				
		AraQ R178A				
		AraQ (Δ GVKG)				

As can be seen in Table 4, all co-transformations of mutant AraP, MsmX_T18 + AraP_EASA_T25 (corresponding to A210S/D211A) and MsmX_T18 + AraP_MalF_Chimera_T25, present different colourations, suggesting different interaction levels. MsmX_T18 + AraP_EASA_T25

co-transformation presents a blue shade very similar to the wild-type co-transformation, MsmX_T18 + AraP_T25, indicating that this double mutation probably does not affect the contact NBD-TMD, while MsmX_T18 + AraP_MalF_Chimera_T25 presents a white colouration, equal to our negative control, suggesting that replacing the loop containing the EAA motif from *B. subtilis* with a C3 loop from *E. coli*, affects in a negative drastic way its interaction with MsmX. Nevertheless, β -galactosidase activities were quantified and are displayed in the next subchapter.

Similar to mutant AraP, all co-transformations of mutant AraQ, MsmX_T18 + AraQ_D180A, MsmX_T18 + AraQ_R178A and MsmX_T18 + AraQ_ Δ GVKG, display different colourations. Both single mutations, D180A and R178A display a white colour, comparable to the negative control, suggesting loss of interaction with MsmX. An unexpected result is the strong blue colour observed in mutation AraQ_ Δ GVKG. In *B. subtilis*, growth kinetic parameters were assessed in the presence of arabinotriose, a complex sugar transported by AraNPQ-MsmX system. Here, lower doubling times, suggest a higher capacity of the system to translocate the substrate and consequently grow and higher doubling times, suggest a lower transport capacity, leading to slower growth (above 500 min was considerate non-growth). This deletion showed a drastic increase in doubling time, suggesting a negative impact in arabinotriose transport that could be due to the loss of important interacting points between MsmX and AraQ (Ferreira *et al.*, 2017).

When MsmX is fused to the T25 fragment and AraQ and AraQ_ Δ GVKG are fused to T18, the co-transformations present a blue colouration, although lighter than the positive control and wild-type interaction, suggesting that interaction is present but not as strong. As observed in their reciprocal co-transformation, between these two co-transformations, T18_AraQ + MsmX_T25 and T18_AraQ_ Δ GVKG + MsmX_T25, the displayed blue shade is identical, suggesting similar levels of interaction.

These results give us an initial screening to which mutation may be more or less relevant, but to ensure that these differences are significant, all these mutations were studied by a quantitative analysis, by determining β -galactosidase activity in units/mg (U/mg) with results being present and discussed below.

3.1.2. Quantitative analysis of MsmX and AraPQ interaction

β -galactosidase activity was used to quantitatively assess levels of interaction between two hybrid proteins. For that, *E. coli* BTH101 cells were co-transformed (Table 3) and grown as described in subchapter 2.7 and 2.9. In this assay, cells were permeabilized using SDS (sodium dodecyl sulphate) and toluene and the enzymatic reaction started by adding ONPG (ortho-Nitrophenyl- β -galactoside). In the presence of β -Galactosidase, this substrate is hydrolysed into β -D-galactose and ortho-nitrophenyl (chromogenic yellow compound), which is produced in proportional quantities to β -galactosidase. The reaction is then stopped by the addition of Na₂CO₃, which elevates the pH to 11, inactivating the

enzyme. Ortho-nitrophenol is monitored at 420 nm and enzymatic activity, in U/mg (Figure 23), is calculated as explained in subchapter 2.9 (Juers *et al.*, 2012; Karimova *et al.*, 2017).

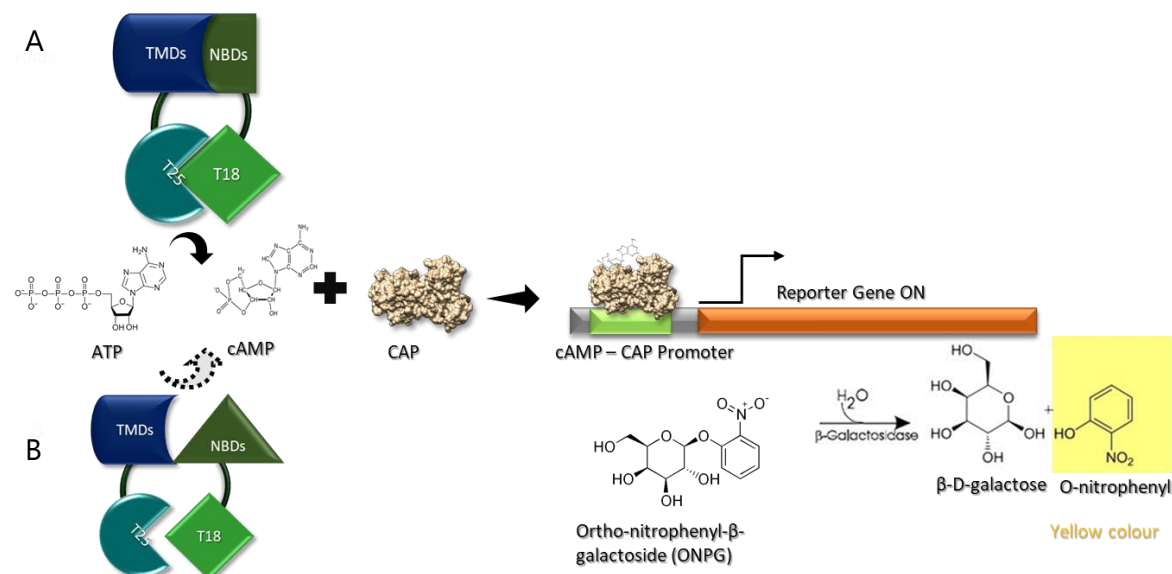


Figure 23. Principle of BACTH System in quantitative assays, using ONPG as substrate for β -galactosidase. A) When T25 and T18 fragments are co-expressed as fusions with two interacting polypeptides, there is association of these two chimeric proteins and consequently reconstitution of adenylate cyclase activity. This enzyme is going to produce cAMP that is going to interact with CAP protein and together are going to activate a reporter gene, in this case *lacZ*, responsible for β -galactosidase production. β -galactosidase will cleave ONPG, a colourless compound, into β -D-galactose and ortho-Nitrophenyl, a yellow compound that can be measured at 420 nm. B) When T25 and T18 fragments are co-expressed as fusion with two non-interacting polypeptides, there is no association and no reconstitution of adenylate cyclase. Since no β -galactosidase is produced, ONPG is not cleaved.

In this assay, the interaction between MsmX_T18 + Zip_T25 was used as negative control. Zip_T25 corresponds to the leucine zipper of GCN4, which does not interact with MsmX, having this co-transformation an enzymatic activity of 36.05 ± 3.66 U/mg. As positive control for NBDs-TMDs interaction, the combination Zip_T18 + T25_Zip (dimerization between leucine zippers of GCN4) was used, having an enzymatic activity of 2427.07 ± 50.63 U/mg. As controls for the wild type interaction between MsmX ATPase and AraPQ TMDs, co-transformations MsmX_T18 + AraP_T25 and MsmX_T18 + T25_AraQ were used, having β -galactosidase activities of 2228.57 ± 5.62 U/mg and 1113.75 ± 27.44 U/mg.

As described in Figure 24 and Table 5, both AraP_EASA and AraP_MalF_Chimera mutations show statistically different β -galactosidase activity when compared to the wild-type interaction, MsmX_T18 + AraP_T25.

A non-significant result when compared with the wild-type activity is the co-transformation MsmX_T18 + AraP_EASA_T25 that displays an enzymatic activity of 1843.38 ± 81.34 U/mg. In this chimeric mutation, the highly conserved residues EAA were changed to EASA (A210S/D211A) to mimic MalF. Here, the EAA motif (EASA in this protein) is a main point of contact with MalK being inserted within the monomer, in the Q-loop region, connecting NBDs and TMDs and allowing the

transmission of NBD conformational changes to the TMDs and consequently the alternating access mechanism (Oldham *et al.*, 2007; Ter Beek *et al.*, 2014).

Table 5. Quantitative protein-protein interaction assays. β -galactosidase activity (U/mg) was measured in each co-transformant (at least three independent experiments; ^a- Two independent experiments) to assess the impact of each TMD mutation on MsmX-TMDs interaction. Zip_T18 + T25_Zip was used as positive control and MsmX_T18 + T25_Zip was used as negative control. MsmX_T18 + AraP_T25 and MsmX_T18 + T25_AraQ were used as controls for the wild type interaction.

<i>E. coli</i> BTH101 co-transformation		
T18 Fusion	T25 Fusion	β -Galactosidase activity U/mg
Zip_T18	T25_Zip	2427.07 $\pm$ 50.63
MsmX_T18	T25_Zip	36.05 $\pm$ 3.66
MsmX_T18	AraP_T25	2228.57 $\pm$ 5.62 ^a
MsmX_T18	AraP_EASA_T25	1843.38 $\pm$ 81.34 ^a
MsmX_T18	AraP_MalF_Chimera_T25	30.78 $\pm$ 2.14
MsmX_T18	T25_AraQ	1113.75 $\pm$ 27.44
MsmX_T18	T25_AraQ_D180A	80.40 $\pm$ 15.13
MsmX_T18	T25_AraQ_R178A	103.13 $\pm$ 14.08
MsmX_T18	T25_AraQ_ΔGVKKG	713.59 $\pm$ 96.31 ^a
T18_AraQ	MsmX_T25	578.00 $\pm$ 71.16
T18_AraQ_ΔGVKKG	MsmX_T25	110.05 $\pm$ 6.33

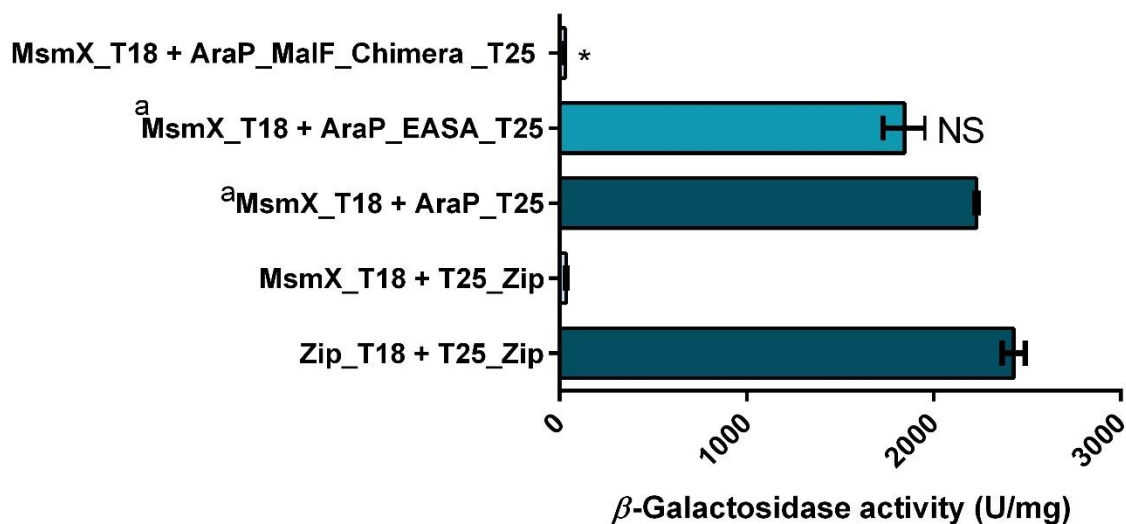


Figure 24. MsmX + Mutant AraP interaction assay. β -galactosidase activity (U/mg), BACTH system was used to determine the interaction between MsmX and chimeric AraP co-transformants. Zip_T18 + T25_Zip was used as positive control while T25_Zip + MsmX_T18 was used as negative control. MsmX_T18 + AraP_T25 was used as control for the wild-type interaction. At least three independent experiments (^a- Two independent experiments) were performed for each co-transformation and error bars indicate the standard mean deviation. Statistical significance between enzymatic activities determined for each co-transformant and co-transformant MsmX_T18 + AraP_T25 is indicated (*, $p < 0.05$; NS – Non-significant).

Since the EAA motif is a conserved sequence within some coupling helix, this motif in AraP most likely acts in a similar way to MalF, meaning, it will probably belong to a short α -helix, that is inserted in a MsmX cleft, allowing conformational changes and consequently the transport mechanism. Since the interaction of this region with the Q-loop of the NBD is maintained by a salt-bridge interaction, the non-significant reduction observed in the enzymatic activity of these mutations (A210S/D211A) is due to the insertion of the polar uncharged side chain of the serine and the removal of the negative charged side chain of the aspartic acid, leading to small changes in the electrostatic surface causing repulsions that may weaken this salt-bridge interaction. These results imply that alone these two amino acids are not sufficient to significantly disturb the loop containing the EAA motif.

The co-transformation MsmX_T18 + AraP_MalF_Chimera_T25 has the most significant change in β -galactosidase activity (30.78 ± 2.14 U/mg), being even less than the negative control. This lack of interaction could mean that sequence integrity of the wild-type loop (putative coupling helix) is necessary for this important interaction region between AraP and MsmX and its replacement for the MalF C3 loop impairs its contact, by the same reasons described in the previous paragraph. Another possible explanation for the lack of interaction may not be due to the interaction *per se* but due to the loop replacement. It may be possible that using MalF C3 loop could change the putative α -helix conformation or even alter the entire protein conformation leading to low expression levels, aggregation, misfolding leading to its signalling for proteolysis, consequently making it impossible for MsmX to interact.

AraQ mutants have different impacts in its interaction with MsmX, when compared with the wild-type interaction, MsmX_T18 + T25_AraQ (Table 5 and Figure 25).

MsmX_T18 + T25_AraQ_D180A co-transformant exhibits a β -galactosidase activity very similar to the negative control (80.40 ± 15.13 U/mg). This low enzymatic activity represents a ~75% loss in interaction. The loss of this highly conserved negative side chain, probably causes a change in electrostatic surface in contact with MsmX, leading to a severe negative impact, detected by the BACTH system. The effect of this mutation was also previously studied by our lab (Ferreira *et al.*, 2017), by determining the growth kinetic parameters in *B. subtilis*. Here, a negative impact on arabinotriose transport (a doubling time of 149.3 ± 12.8 min when compared with the wild-type 98.2 ± 10.0 min) was also determined; the difference between the mutant and the wild type doubling time is not as relevant as the difference observed between the two in the BACTH assay, suggesting that the effect of this mutation is in the context of the full transporter may have other surrounding interactions that stabilize the complex allowing the uptake of arabinotriose, only slightly impairing the transport. However, it is also possible that in the BACTH system, the effect of this mutation may be overemphasized due to only considering the interaction between one mutated AraQ and one wild type MsmX.

Another mutation assessed was the changing of the arginine in position 178 to an alanine, which causes the loss of a positive charged side chain. This co-transformation MsmX_T18 +

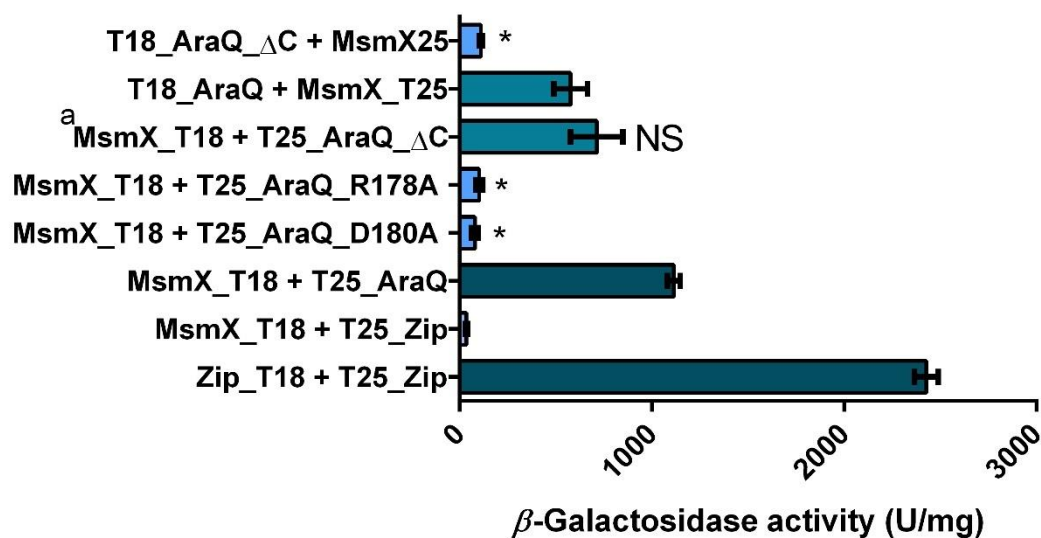


Figure 25. MsmX + Mutant AraQ interaction assay. β -galactosidase activity (U/mg), using BACTH system was used to determine the interaction between MsmX and mutant AraQ co-transformants. Zip_T18 + T25_Zip was used as positive control while MsmX_T18 + T25_Zip was used as negative control. MsmX_T18 + T25_AraQ and its reciprocal, T18_AraQ + MsmX_T25 were used as controls for the wild-type interaction. At least three independent experiments (^a- Two independent experiments) were performed for each co-transformation and error bars indicate the standard mean deviation. Statistical significance between enzymatic activities determined for each co-transformant and co-transformant MsmX_T18 + T25_AraQ is indicated (*, $p < 0.05$; NS – Non-significant). ΔC indicates AraQ_ΔGVKG mutation.

T25_AraQ_R178A exhibits a β -galactosidase activity of 103.13 ± 14.08 U/mg, which is also significantly lower than the wild-type co-transformation. Similarly, to this previous mutation, a lower enzymatic activity can be attributed to changes in electrostatic surface, causing weaker interactions or even repulsion or structural alteration, hampering the access to interaction points between TMD and NBD. These two mutations, D180A and R178A, may be responsible for more contacts with MsmX than expected and consequently can be important for maintaining the correct assembly of the system and since they are both located in the EAA motif of AraQ, they may have a role in the alternating access transport mechanism.

A deletion of the four C-terminal amino acids of AraQ was also constructed, having a β -galactosidase activity of 713.59 ± 96.31 U/mg, a statistically non-significant loss in interaction when compared with its wild type. This activity is higher than expected since in growth kinetics parameters determined by our lab in *B. subtilis* model (Ferreira *et al.*, 2017), showed a severely decrease in growth rate (472.8 ± 22.6 min, while wild type doubling time is 98.2 ± 10.0 min). This can be explained by looking at the MalEFGK₂ system (Oldham *et al.*, 2007). An unexpected feature revealed by Oldham *et al.* (2007) was the insertion of this MalG C-terminal region into MalK dimer interface, suggesting some important role in ABC transporter function. With this in mind, the deletion of these 4 amino acids, probably hampers arabinotriose transportation by interfering with the transport cycle, that cannot go from the outward state to the inward conformation and not by diminishing drastically the interaction,

as seen assessed in these β -galactosidase assays using the BATCH system. Removing AraQ C-terminal tail will probably not affect the closure of MsmX dimer, but the lack of hydrogen bonds with the Q-loops of both monomers, will probably affect the correct positioning of MsmX dimer and the intermediate conformation will be compromised. Although a low identity between the primary sequence of TMDs is present, this tail is highly conserved in MalG and AraQ but also in MdxG and GanP (Figure 21), so it is highly probable that the C-terminal tails of these last TMDs also play an essential role in the translocation's mechanism. To date, this C-terminal portion is not observed in other families of ABC transporters, showing its importance in the specific binding between in ABC sugar importers.

In order to understand if the mutations with lack of interaction phenotype, namely AraP_MalF_Chimera and point mutation in AraQ, occurred due to these modifications and not due to protein degradation or problems in protein expression, additional analysis is required. For this, doing detection of *in vivo* accumulation of these TMDs mutants by western blot was essential, but since all mutations were cloned into plasmids expressing the T25 fragment, and the only commercialized antibody was for T18 fragment, it made it impossible to detect these mutants. The natural solution was then, construct the same mutations into plasmids containing T18 fragment, and assess if the effect of the previous mutations is maintained.

Co-transformation T18_AraQ + MsmX_T25 exhibits a β -galactosidase activity of 578.00 ± 71.16 U/mg, about half of the activity determined for its reciprocal. This activity is also statistically low when compared with the positive control Zip_T18 + Zip_T25, which can possibly complicate the study of the impact of mutation in TMDs-NBDs Interaction. With this in mind, a previous studied mutation was also tested, the deletion of the C-terminal tail. For co-transformation T18_AraQ_ Δ GVKG + MsmX_T25, a β -galactosidase activity of 110.05 ± 6.33 U/mg was determined. This value is statistically different from its wild-type T18_AraQ + MsmX_T25. This shows that the effect of, at least, this mutation in this reciprocal combination, can still be seen, allowing the use for the detection of the mutated TMDs by Western Blot.

In sum, these results show that all mutations have a different negative impact in NBD-TMD interaction, with all mutations supporting the importance of the loop containing the EAA motif in both AraP and AraQ interaction with MsmX and consequently its possible role in the transport mechanism. The deletion of the C-terminal tail in AraQ also supports the importance of this highly conserved tail in ABC sugar importers. As previous and current results show, it remains to be seen if the data of some mutations, namely AraP_A210S/D211A, AraP_MalF_Chimera and AraQ_R178A correlates with growth kinetic parameters in *B. subtilis*. Using the reciprocal co-transformation, AraQ in T18 and MsmX in T25, also allows to assess the effect of mutations and allow a validation of this system by performing Western blot analysis, which was not possible when only considering the qualitative assays, since both co-transformations displayed the same colouration. Furthermore, the qualitative assays correlate to some degree with the quantitative assays, with the highest β -galactosidase activities being depicted with a strong blue colouration and β -galactosidase activities quantitatively close to the negative

control also appearing with a white colour, but due to the proximity of colouration, qualitative assays due not allow to distinguish co-transformation with proximal β -galactosidase activities.

3.1.3. Detection of *in vivo* accumulation of wild type and mutant TMDs by Western blot analysis

To determine if the observed phenotypes are a consequence of the introduction of the previously mentioned mutation or if they disturb normal protein production of TMDs, it is necessary to verify protein accumulation of mutant TMDs by Western Blot analysis. Because BACTH protein detection by Western Blot was not previously done in our lab, optimization was required. First, total cell extracts of O/N growth of *E.coli* BTH101 co-transformations Zip_T18 + Zip_T25, MsmX_T18 + AraP_T25, MsmX_T18 + AraP_EASA_T25, MsmX + AraP_MalF_Chimera_T25, MsmX_T18 + T25_AraQ, MsmX_T18 + T25_AraQ_D180A, MsmX_T18 + T25_AraQ_R178A and MsmX_T18 + T25_AraQ_ΔGVKG, were performed as described in subchapter 2.10, with subsequent steps being described in the same chapter. Here, co-transformation Zip_T18 + Zip_T25 (25.9 kDa) was used as experiment control and NZYcolour protein marker II (NZYTech), was used as molecular weight.

As seen in Figure 26A, the transfer was successful since the fractionated proteins were effectively transferred from the SDS-PAGE gel to the nitrocellulose membrane. Analysis of the radiographic plates (Figure 26B) reveals MsmX_T18 protein accumulation in all co-transformations but no other evidence can be inferred due to the high background, where non-specific bands are observed, except in the Zip_T18 lane. MsmX_T18 (based on its molecular weight, 61.8 kDa) and Zip_T25 (25.9 kDa) are marked with an arrow. This high background on the blot and non-specific bands below MsmX_T18 size are unexpected, since Zip_T18 has the same cell background, *E. coli* BTH101, and no bands are observed. There are many possible causes for this high background on the blot and for the non-specific

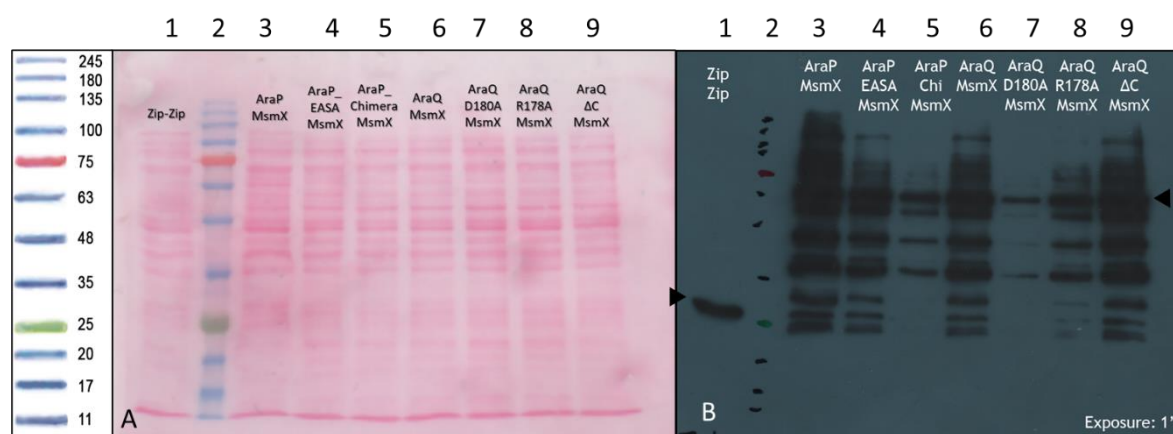


Figure 26. MsmX_T18 detection by Western Blot analysis, from *E. coli* BTH101 growth. 20 μ g of total cell extracts of Zip_T18 + Zip_T25 was loaded into the first lane. In lane 2, NZYcolour protein marker II was loaded and in lanes 3-9, 20 μ g of total cell extracts of MsmX_T18 + AraP_T25, MsmX_T18 + AraP_EASA_T25, MsmX_T18 + AraP_MalF_Chimera_T25, MsmX_T18 + T25_AraQ, MsmX_T18 + T25_AraQ_D180A, MsmX_T18 + T25_AraQ_R178A and MsmX_T18 + T25_AraQ_ΔGVKG were loaded. A) Nitrocellulose membrane stained with Ponceau Red, showing the fractionated proteins transferred from an SDS-PAGE gel (12.5%). B) Autoradiographic plate after detection with α -Cya antibody. MsmX_T18 (61.8 kDa) and Zip_T18 (25.9 kDa) are marked with an arrow.

bands, from non-specific binding of secondary antibody, high concentrations of antibodies, insufficient blocking, insufficient washing steps, protease activity in stationary phase of growth or degradation caused by freeze and thaw of samples of cell extracts, which were tested and described below.

First, cell extracts of co-transformations Zip_T18 + Zip_T25, MsmX_T18 + AraP_T25, MsmX_T18 + T25_AraQ and MsmX_T18 + MsmX_T25 were loaded into an SDS-PAGE gel, additionally to cells of the previous co-transformations recovered after O/N growth on the same medium and resuspended in Loading Buffer + Running Buffer to an O.D._{600nm} of 10, without being frozen or placed at 4 °C. As seen in Figure 27A, the fractionated proteins were successfully transferred for the nitrocellulose membrane, as Ponceau Red staining revealed. Antibody incubations were performed as described in subchapter 2.10, except for the primary antibody, which was used with a higher dilution of 1:2000. The respective autoradiographic plate is shown in Figure 27B. Here, accumulation in all lanes is seen, as expected but a higher primary antibody dilution does not improve the background. However, using cells resuspended in Loading + Running Buffer seems to slightly improve the

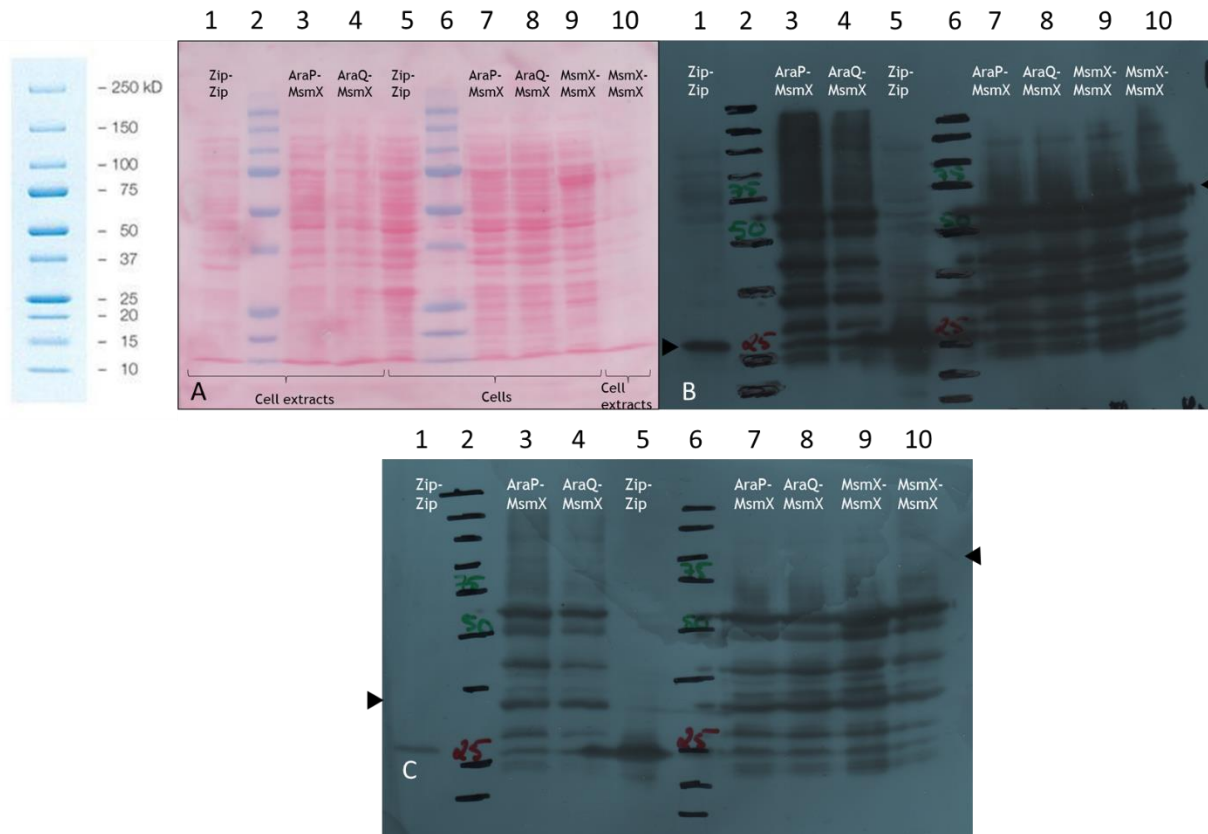


Figure 27. MsmX_T18 detection by Western Blot analysis, from *E. coli* BTH101 growth. 20 µg of total cell extracts of Zip_T18 + Zip_T25 was loaded into the first lane. In lane 2 and 6, Precision Plus Protein TM All Blue Prestained Protein Standards was loaded and in lanes 3, 4 and 10, 20 µg of total cell extracts of MsmX_T18 + AraP_T25, MsmX_T18 + T25_AraQ and MsmX_T18 + MsmX_T25 were loaded. In lanes 5, 7-9, cells of co-transformant Zip_T18 + Zip_T25, MsmX_T18 + AraP_T25, MsmX_T18 + T25_AraQ and MsmX_T18 + MsmX_T25 resuspended in Loading and Running Buffer were loaded. A) Nitrocellulose membrane stained with Ponceau Red, showing the fractionated proteins transferred from an SDS-PAGE gel (12.5 %). B) Autoradiographic plate after detection with α -Cya antibody using 1:2000 dilution. C) Autoradiographic plate after detection with α -Cya antibody using 1:2000 dilution for 2 h (blocking O/N at 4 °C). MsmX_T18 (61.8 kDa) and Zip_T18 (25.9 kDa) are marked with an arrow.

background, especially above the MsmX_T18 band. Since these improvements are not significant, this membrane was again used to test other conditions. Here, the blocking step was performed O/N, 4 °C, without shaking, while incubation with primary antibody 1:2000 was performed at RT, for 2 h. As seen in Figure 27C, doing the blocking step O/N significantly reduced the background, but since cell extract of Zip-Zip co-transformation was also significantly affected, these conditions may not be ideal. However, when using cells only resuspended in Loading + Running Buffer, Zip-Zip co-transformation appears not to be affected and its background was cleared. Several non-specific bands were still visible below the putative MsmX_T18 band, suggesting protease activity or protein degradation was occurring.

To test other conditions another membrane was used (Figure 28A), in which co-transformations Zip_T18 + Zip_T25 and MsmX_T18 + AraP_T25 were harvested at OD₆₀₀ of 0.5 (exponential phase of growth) with some cells being treated with PMSF, a serine protease inhibitor. These cells were never frozen or placed at 4 °C. Although the fractionated proteins were successfully transferred to the nitrocellulose membrane, it was possible to distinguish an air bubble (Figure 28A), which was considered when looking at the radiographic plate. Half of the membrane was incubated with 1:2000 primary antibody dilution while the other half was incubated with 1:4000 dilution, for 2 h with mild shaking. As seen in Figure 28B, there is no background and only a few non-specific bands remain. It appears to have no difference between the two antibody dilutions but using PMSF not only helped with the viscosity of samples when loading them into SDS-PAGE but when looking to the right plate (the left cannot be used for this analysis due to the air bubble) seems to help with the intensity of the non-specific bands. Using cells grown only in the exponential phase of growth treated with PMSF and resuspended in Loading + Running Buffer seems to be the ideal condition to perform Western blot analysis when using BACTH system to study NBD-TMD interactions.

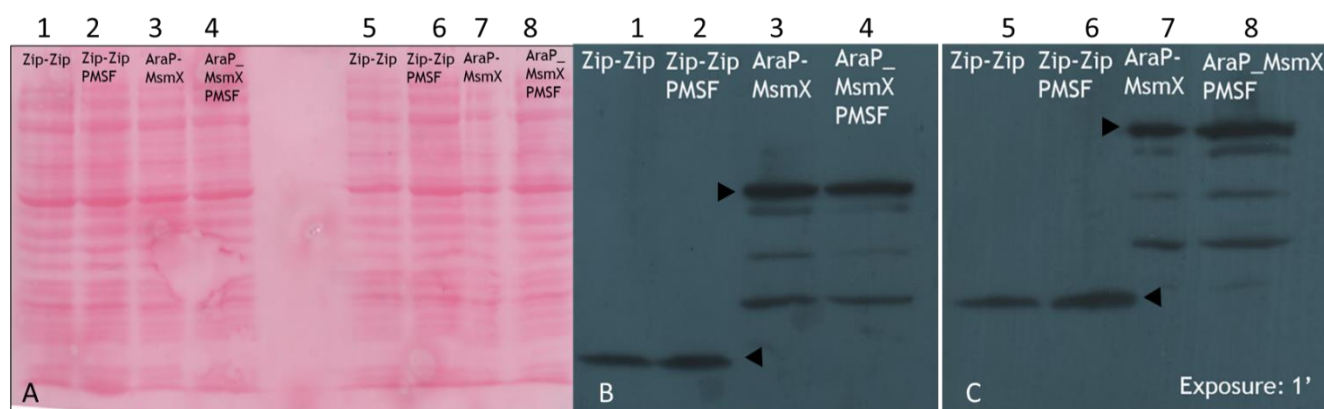


Figure 28. MsmX_T18 detection by Western Blot analysis, from *E. coli* BTH101 growth. In lanes 1 and 3 cells of co-transformant Zip_T18 + Zip_T25, MsmX_T18 + AraP_T25 resuspended in Loading and Running Buffer were loaded while in lanes 2 and 4 cells of co-transformant Zip_T18 + Zip_T25, MsmX_T18 + AraP_T25 treated with PMSF and then resuspended in Loading and Running Buffer were loaded. In lanes 6 to 10, the same samples were loaded. A) Nitrocellulose membrane stained with Ponceau Red, showing the fractionated proteins transferred from an SDS-PAGE gel (12.5 %). B1;2) Autoradiographic plate after detection with α -Cya antibody using 1:2000 dilution (B1) and 1:4000 dilution (B2). MsmX_T18 (61.8 kDa) and Zip_T18 (25.9 kDa) are marked with an arrow.

Since all our mutant TMDs are cloned into plasmids containing T25 fragment and currently no commercial antibody is available it is not possible to perform Western Blot analysis of these interactions. Since several TMDs energized by MsmX were previously constructed into plasmids containing T18 fragment (Microbial Genetics Laboratory, unpublished results), AraQ and AraQ_ΔGVKG were also fused to T18 fragment and after quantitatively assess its interaction with MsmX fused to T25 fragment, its accumulation was determined by Western Blot analysis. Here, *E. coli* BTH101 strains T18_AraQ + MsmX_T25, T18_AraQ_ΔGVKG + MsmX_T25, Gan_T18 + MsmX_T25 and Mel_T18 + MsmX_T25 were harvested at an OD_{600nm} of 0.3, as described in subchapter 2.10. Cells underwent the same treatment and were loaded into the gel without being boiled. After SDS-PAGE running was complete, a successful electrotransfer of proteins can be seen in a nitrocellulose membrane (Figure 29A) stained with Ponceau Red. Zip_T18 + Zip_T25 co-transformation was loaded into the first lane as experience control and NZY Protein Marker II was loaded into the second lane. AraQ, AraQ_ΔGVKG, GanP (TMD from GanSPQ system) and MelD (TMD from MelEDC system) were loaded into lanes 3 to 6, respectively.

This analysis shows accumulation of all TMDs *in vivo* (Figure 29B), since in each lane is visible a band with the expected size, 25.9 kDa, 52.2 kDa, 51.9 kDa, 57.8 kDa and 67.4 kDa, respectively. Apparent differences in accumulation are present, that could lead to the assumption that differences detected in β -galactosidase activity may be due to this aspect, but since there is no housekeeping gene that could be used as control, no conclusion can be taken. Thus, only can be acknowledged that all TMDs are accumulating *in vivo* and the AraQ_ΔGVKG mutation does not impair protein production.

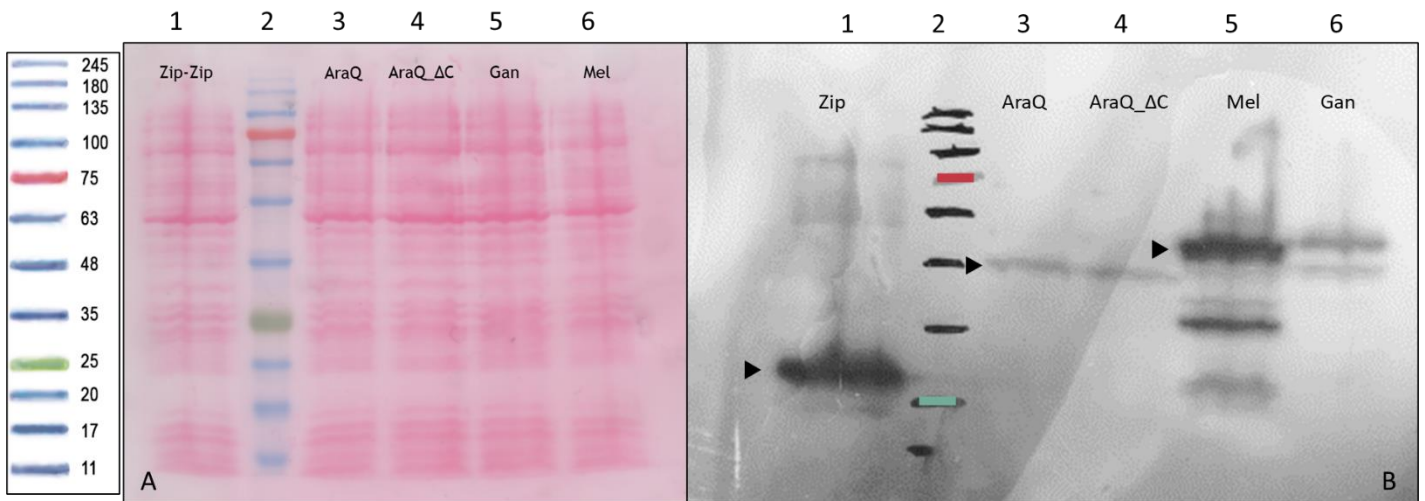


Figure 29. TMD_T18 detection by Western Blot analysis, from *E. coli* BTH101 growth. In lane 2, NZYcolour protein marker II was loaded. In lanes 1, 3-6 cells of co-transformants Zip_T18 + Zip_T25, T18_AraQ + MsmX_T25, T18_AraQ_ΔGVKG + MsmX_T25, Gan_T18 + MsmX_T25 and Mel_T18 + MsmX_T25 treated with PMSF and then resuspended in Loading and Running Buffer were loaded. A) Nitrocellulose membrane stained with Ponceau Red, showing the fractionated proteins transferred from an SDS-PAGE gel (12.5%). B) Autoradiographic plate after detection with α -Cya antibody. Zip_T18 (25.9 kDa), AraQ_T18 (52.2 kDa), AraQ_ΔGVKG (51.9 kDa), Mel_T18 (57.8 kDa) and Gan_T18 (67.4 kDa).

3.2. *In vivo* analysis of MsmX homologs and AraPQ interaction using the BACTH system

Previous work published by our lab, have identified several putative multitask ATPases based on their sequence identity in the predicted contact region NBD-TMD, between amino acids 60 to 179 of MsmX. From this analysis, ATPases from Gram-Positive of Firmicutes phylum and Gram-Negative bacteria were selected: FrlP (previously identified as YurJ from *B. subtilis*, 58%), ABC_Bt (from *B.thuringiensis*, 74%), MsmK (*Streptococcus pneumoniae*, 64%), ABC_Sa (*Staphylococcus aureus*, 66%), ABC_Cd (*Clostridioides difficile* 630, 57%), MalK and YcjV (from *E. coli*, with 45% and 64%, respectively) and beyond Firmicutes phylum, ABC_Syn (from *Synechocystis sp.*, 71%). Although MalK has low sequence identity, it was selected because it is part of the prototype translocator of ABC type I importers, MalEFGK₂ of *E. coli* and it is the most well characterized transporter (Leisico *et al.*, 2020).

These ATPases were introduced in *trans* in the *B. subtilis* 168T⁺ chromosome under the control of an inducible promoter. The efficiency of each ATPase to replace MsmX was assessed by determining growth kinetic parameters in the presence of the substrates, translocated by the specific transporter, arabinooligosaccharides and galactooligosaccharides. In these assays, lower doubling time reflects higher capacity of an NBD to energize the AraNPQ (arabinotriose) and GanSPQ (galactan) transporters and higher doubling times shows lower functional ability to power the transport. Doubling times above 500 min are considered non-growth. All these ATPases were shown to replace MsmX function *in vivo* with different efficiencies, except for MalK and YcjV, that are not able to energize AraPQ or GanPQ (Leisico *et al.*, 2020).

Western blot analysis confirmed that all ATPases accumulate *in vivo* and biochemistry characterization showed that both MalK and YcjV maintained their ATPase activity, suggesting another reason other than protein misfold or degradation for different levels of functionality. Since no relation seems to be found between primary sequence and its functional degree of exchangeability, studies about their interaction with TMDs can help to understand these promiscuous contacts and since these NBDs are potential targets for antibacterial therapy, studying this feature may allow drug design with the intent to disrupt carbohydrate uptake (Leisico *et al.*, 2020).

Thus, to conduct *in vivo* studies of the interaction between some of the previously mentioned ATPases (Table 6) and AraPQ translocator, the BACTH system was used. All plasmids constructed for these assays are described in subchapter 2.5 and listed in Table 1, with all interactions tested being displayed in Table 3.

Table 6. List of MsmX homologs used in this work. ^aAmino acid identity to *B. subtilis* MsmX.

Species	Gene name	Reference	Protein Identity ^a
<i>Staphylococcus aureus</i> subsp. <i>aureus</i> ST398	<i>ugpC</i> (<i>abc_Sa</i>)	WP_000818906.1	66%
<i>Clostridioides difficile</i> 630	<i>abc_cd</i>	WP_011861561.1	57%
<i>Escherichia coli</i> K-12	<i>malK</i> (or b4035)	NP_418459.1	45%
<i>Escherichia coli</i> K-12	<i>ycjV</i> (b4524)	P77481.2	64%
<i>Synechocystis</i> sp.	<i>malK</i> (<i>abc_Syn</i>)	WP_010874222.1	71%

All constructions were performed with the ATPases being cloned into a high copy number plasmid pUT18, with AraP and AraQ being constructed into low copy number plasmids, pKNT25 and pKT25 respectively, as described in 2.5. All ATPases and AraP are constructed into plasmids that allow their fusion to the N-terminal of the respective fragment, while AraQ is constructed on a plasmid allowing its fusion to the C-terminal of T25 fragment.

Protein-protein interactions were studied by qualitatively and quantitatively assessed β -galactosidase activity, using the BACTH system, as described in subchapter 3.1.

3.2.1. Qualitative analysis of MsmX homologs and AraPQ interaction

Qualitative analysis of protein-protein interaction between putative multitask NBDs and AraPQ TMDs was studied by performing solid medium assays, as described in subchapter 2.8 and discussed in subchapter 3.2.1. Results are presented in Table 7. The same co-transformations were used as controls as described previously in subchapter 3.2.1. Co-transformations Zip_T18 + Zip_T25 and MsmX_T18 + Zip_T25 were used as positive and negative control, displaying a strong blue colour and a white colouration, respectively, confirming the expected phenotype for interacting and non-interacting proteins. Co-transformations MsmX_T18 + AraP_T25 and MsmX_T18 + T25_AraQ were used as controls, since interaction between AraPQ and MsmX ATPase is our model for multitask NBDs-TMDs interaction and one of the best characterized systems. Here, both interactions display a strong blue colour, a typical phenotype of interacting proteins.

All co-transformations bearing MalK as ATPase, MalK_T18 + AraP_T25; MalK_T18 + AraP_EASA_T25; MalK_T18 + AraP_MalF_Chimera_T25 and MalK_T18 + T25_AraQ, present a white colour, suggesting lack of interaction. These results of the wild-type interaction are expected and correlate with the growth kinetic parameters (805.5 ± 188.4 min), suggesting that this higher doubling time is due to the inability of this ATPase to interact with the TMDs (Leisico *et al.*, 2020). The white

colour observed in co-transformants bearing a mutant AraP also suggest inability of MalK to interact with these mutant proteins.

Solid medium assays of co-transformations where the interaction of YcjV and AraPQ was tested, YcjV_T18 + AraP_T25 and YcjV_T18 + T25_AraQ, also correlate its growth kinetic parameters. All interactions display an expected white colour, indicating that no interaction is present and suggesting that this lack of interaction is responsible for the ATPase not being able to replace MsmX *in vivo*.

Table 7. Qualitative protein-protein interaction assay. Qualitative analysis of β -galactosidase activity using X-Gal as an indicator media. Blue colouration corresponds to a positive interaction while white colour corresponds to a non-interacting polypeptide. Zip_T18 + Zip_T25 co-transformation and MsmX_T18 + Zip_T25 were used as controls and MsmX_T18 + AraP_T25 and MsmX_T18 + T25_AraQ was used as controls for the wild-type interaction.

		T25 fragment				
		Plasmid pKNT25				Plasmid pKT25
		Zip	AraP	AraP_EASA	AraP_MalF_Chimera	AraQ
T18 fragment	Zip					
	MsmX					
	MalK					
	YcjV					
	ABC_Cd					
	ABC_Sa					
	ABC_Syn					

ATPases ABC_Cd, from *C. difficile*, ABC_Sa from *S. aureus* and ABC_Syn were all able to replace MsmX *in vivo* (Leisico *et al.*, 2020) with different degrees. It is no surprising to see a blue colouration in ABC_Cd_T18 + AraP_T25 and ABC_Sa_T18 + AraP_T25, although lighter than our wild type, suggesting that interaction is present but not as strong, but the white colouration in ABC_Cd_T18 + T25_AraQ and ABC_Sa_T18 + T25_AraQ is unexpected, since it reveal lack of interaction. All co-transformations bearing ABC_Syn as ATPase, ABC_Syn_T18 + AraP_T25 and ABC_Syn_T18 + T25_AraQ, also show an unexpected white colouration, suggesting that no interaction is present. However, this does not correlate with our previous results.

Since these analyses give us an initial screening to study the interaction of these ATPases with AraPQ and these results may help to explain the different degrees of complementation that these different ATPases have *in vivo*, a quantitative analysis, by determining β -galactosidase activity (U/mg) with results being present and discussed in subchapter 3.2.2. This quantitative analysis will also help to understand the lack of interaction detected in ATPases that can complement MsmX activity.

3.2.2. Quantitative analysis of MsmX homologs and AraPQ interaction

The quantitative analysis of PPIs to determine multitask NBD – TMDs interaction, was done by measuring β -galactosidase activity. The procedure is described in subchapter 2.9 and explained in subchapter 3.2.2 and Figure 23. Results are displayed in Table 8 and in Figure 30 and Figure 32. The controls for NBDs-TMDs interaction are the same described in subchapter 3.2.2: as positive control, the interaction between leucine zipper, Zip_T18 + T25_Zip, was measured as 2427.07 ± 50.63 U/mg and as negative control, the interaction between MsmX_T18 + T25_Zip was used measuring 36.05 ± 3.66 U/mg. Since these putative multitask ATPases are also being compared with the model, MsmX from *B. subtilis*, the interactions between this ATPase and the TMDs, AraP and AraQ, MsmX_T18 + AraP_T25 and MsmX_T18 + T25_AraQ, were also used as controls for the wild-type interaction, displaying β -galactosidase activities of 2228.57 ± 5.61 U/mg and 1113.75 ± 27.44 U/mg, respectively.

As previously observed (Leisico *et al.*, 2020), MalK from *E. coli* is unable to replace MsmX as an energizer of the AraNPQ transporter. This ATPase displayed a doubling time of 805.5 ± 188.4 min, which is higher than the 500 min defined as non-growth. This ATPase was shown to accumulate *in vivo* by Western Blot analysis, discarding the possibility of protein misfolding or degradation and maintained its ATPase activity, discarding the possibility of loss of energizing capacity (Leisico *et al.*, 2020). The authors hypothesized that MalK could either contact and interact with the transmembrane proteins, despite not being able to induce conformational changes necessary for the translocation of substrate, or that it did not contact the TMDs at all. To test this hypothesis, the BACTH system was used and the interaction between MalK, with both, AraP and AraQ, was determined. Both interactions, MalK_T18 + AraP_T25 and MalK_T18 + T25_AraQ, have enzymatic activities very close to the negative control, 39.10 ± 1.41 U/mg, and 44.00 ± 1.82 U/mg, respectively. These results show a total loss of interaction when compared with the wild-type interaction between MsmX and AraPQ, correlating with the white colouration determined in the qualitative assays and supporting the hypothesis of MalK not being able to replace MsmX and power the AraNPQ translocator due to the inability to interact with these TMDs. A possible hypothesis to what contributes to the multitasking ability of an ATPase was proposed by Leisico *et al.* 2020, when the comparison of the crystal structures of MalK (non-multitask) and MsmX (multitask) revealed a smaller surface area of the pocket in MalK, responsible for a contact point with the loop containing the EAA motif of a TMDs, a key interaction point that allows the conformational

changes in NBDs to be transmitted to the TMDs (Ter Beek *et al.*, 2014). This structural feature remains the best hypothesis to explain the multitask nature and in the case of MalK, the lack of interaction and consequently the inability to energize the arabinotriose transport system. Another possible explanation could be that the electrostatic potential of this pocket and other contact points may also not be ideal to allow some interaction with AraPQ.

Table 8. Quantitative protein-protein interaction assays. β -galactosidase activity (U/mg) was measured in each co-transformant (at least three independent experiments, ^a - Two independent experiments) to assess multitask NBDs-TMDs interaction. Zip_T18 + T25_Zip was used as positive control and MsmX_T18 + T25_Zip was used as negative control. MsmX_T18 + AraP_T25 and MsmX_T18 + T25_AraQ were used as controls for the wild type interaction. * Growth kinetic parameter determined in *B. subtilis* in liquid minimal medium (CSK) supplemented with α -1,5-arabinotriose. From Leisico *et al.*, 2020

T18 Fusion	T25 Fusion	β -Galactosidase activity U/mg	Doubling Time (min) *
MsmX_T18	T25_Zip	36.05 $\pm$ 3.66	ND
Zip_T18	T25_Zip	2427.07 $\pm$ 50.63	ND
MsmX_T18	AraP_T25	2228.57 $\pm$ 5.61 ^a	115.40 $\pm$ 9.30
	T25_AraQ	1113.75 $\pm$ 27.44	
MalK_T18	AraP_T25	39.10 $\pm$ 1.41	805.50 $\pm$ 188.40
	T25_AraQ	44.00 $\pm$ 1.82	
MalK_T18	AraP_EASA_T25	32.29 $\pm$ 0.77	ND
MalK_T18	AraP_MalF_Chimera_T25	37.88 $\pm$ 3.23	ND
YcjV_T18	AraP_T25	36.47 $\pm$ 1.35	569.00 $\pm$ 29.10
	T25_AraQ	41.98 $\pm$ 1.35	
ABC_Cd_T18	AraP_T25	1658.37 $\pm$ 40.31	225.70 $\pm$ 9.50
	T25_AraQ	216.06 $\pm$ 20.36	
ABC_Sa_T18	AraP_T25	1429.38 $\pm$ 7.71	373.70 $\pm$ 45.80
	T25_AraQ	65.49 $\pm$ 19.32	
ABC_Syn_T18	AraP_T25	37.18 $\pm$ 3.47	328.50 $\pm$ 7.40
	T25_AraQ	50.16 $\pm$ 3.45	

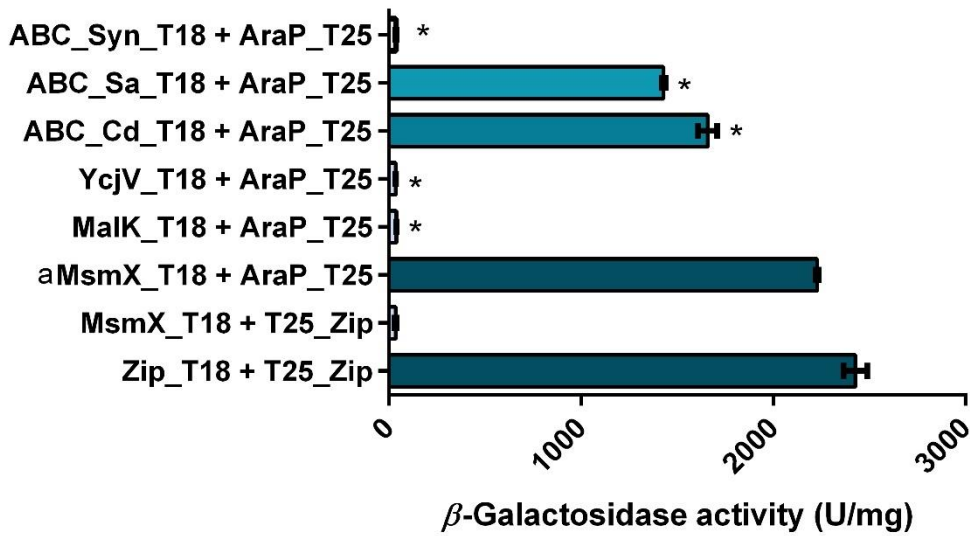


Figure 30. Putative multitask ATPases_T18 + AraP_T25 β -galactosidase assays. By using BACTH system, interaction between MsmX_T18 and AraP_T25 was quantified by determining β -galactosidase activity, expressed in units per mg of bacteria, and was used as control for wild-type interaction. Interaction between other putative multitask ATPases fused to T18 fragment and AraP_T25 was also determined by this method. As positive control Zip_T18 + Zip_T25 co-transformation was used as positive control and MsmX_T18 + Zip_T25 was used as negative control. At least three independent experiments (°- Two independent experiments) were performed for each co-transformation and error bars indicate the standard mean deviation. Statistical significance between enzymatic activities determined for each co-transformant and co-transformant MsmX_T18 + AraP_T25 is indicated (*, $p < 0.05$).

The effect of mutations in AraP described in subchapter 3.1 was also studied regarding their interaction with MalK. As seen in Table 8 and Figure 31, both co-transformations, MalK_T18 + AraP_EASA_T25 and MalK_T18 + AraP_MalF_Chimera_T25, have β -galactosidase activities of 32.29 ± 0.77 U/mg and 37.88 ± 3.23 U/mg, respectively. Since these mutations mimic regions of the coupling helix of MalF (EAA motif and C3 loop), the transmembrane domain energized by MalK, it would be expected to see an increase of enzymatic activity. However, these similar results to the negative control reveal that this conserved sequence, and its coupling helix alone are not sufficient to allow interaction with MalK, suggesting other key motifs or structural arrangements are involved in this specific interface.

β -galactosidase activities determined for the interaction between YcjV, from *E. coli*, and AraP and AraQ, also show values very close to the negative control, 36.47 ± 1.35 U/mg, and 41.98 ± 1.35 U/mg, respectively. Like MalK, these values show a complete loss of interaction when compared with the wild-type and these data also correlate with the growth kinetic assays determined, 569.0 ± 29.1 min (Leisico *et al.*, 2020) and with the qualitative assays described in subchapter 3.2.1. Although this time is considered non-growth, it is very close to the cut-off of 500 min and would be expected to determine some interaction. Western blot analysis also showed accumulation of YcjV, and this ATPase also retained its ATPase activity (Leisico *et al.*, 2020). Looking at the β -galactosidase assays, this lack of interaction is the reason for the inability to power the transport using only the AraNPQ translocator.

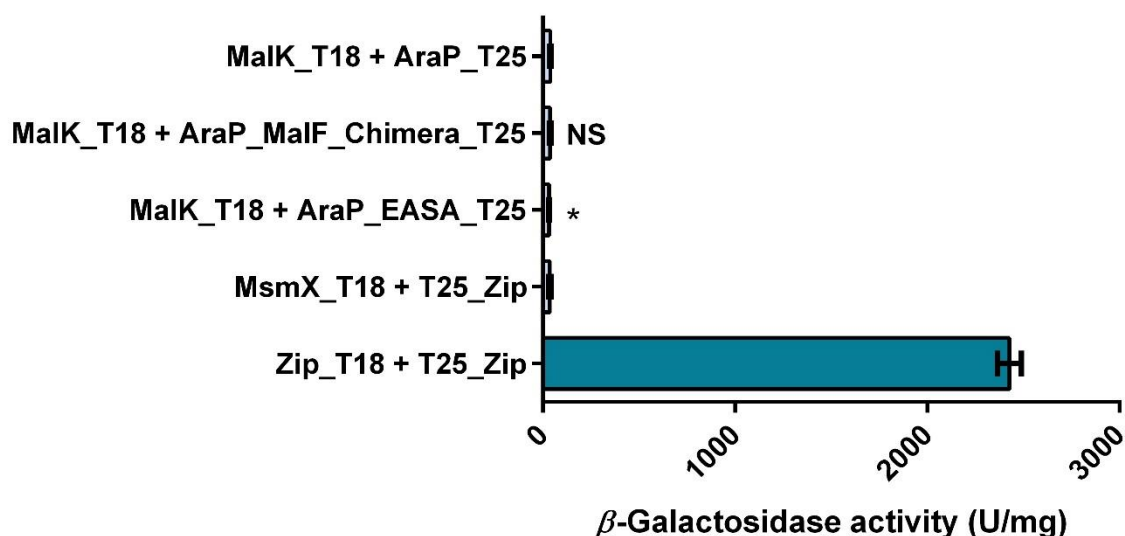


Figure 31. MalK_T18 + AraP_T25 variants β -galactosidase assays. By using BACTH system, interaction between MalK_T18 and AraP_T25 was quantified by determining β -galactosidase activity, expressed in units per mg of bacteria, and was used as control for wild-type interaction. Interaction between chimeric AraP protein and MalK ATPase were also determining by this method. As positive control Zip_T18 + Zip_T25 co-transformation was used as positive control and MsmX_T18 + Zip_T25 was used as negative control. At least three independent assays were performed for each co-transformation and error bars indicate the standard mean deviation. Statistical significance between enzymatic activities determined for each co-transformant and co-transformant MalK_T18 + AraP_T25 is indicated (*, $p < 0.05$).

Since YcjV is non-characterized ATPase and no structure has been resolved, a probable explanation is that the pocket responsible for the docking of the coupling helix is narrow and like MalK has a small surface area. The surface area of the contact regions may also not have the electrostatic potential adequate for interactions with TMDs. An important feature to consider about YcjV is its annotation as a pseudo-gene in *E. coli* K-12 genome due to a mutation that causes a truncated version of this ATPase, suggesting that maybe this protein is not functional in *E. coli*. For enzymatic activity determination the complete protein was used but it's interesting to think that throughout evolution, maybe this ATPase gene became a pseudogene due to its product toxicity or maybe this truncated version still maintain some function, but further studies are required (Goodhead & Darby, 2015).

Of all ATPases studied that were shown to replace MsmX and energize the AraNPQ translocator, ABC_Cd from the pathogen *Clostridioides difficile* 630 was the most efficient and like MsmX, it is classified as an orphan NBD since no other ABC component is found in its proximity (Leisico *et al.*, 2020). β -galactosidase assays of this ATPase with AraP (ABC_Cd_T18 + AraP_T25) and AraQ (ABC_Cd_T18 + T25_AraQ), display enzymatic activities of 1658.37 ± 40.31 U/mg and 216.06 ± 20.36 U/mg, respectively. These values, although statistically different from the wild-type interactions, demonstrate interaction between ABC_Cd and AraPQ TMDs, with interaction with AraP being 1.3-fold lower and interaction with AraQ being 5-fold Lower. This data also corroborates with the Growth Kinetic Parameters, displaying a doubling time of 225.7 ± 9.5 min, about 2-fold of the wild-type.

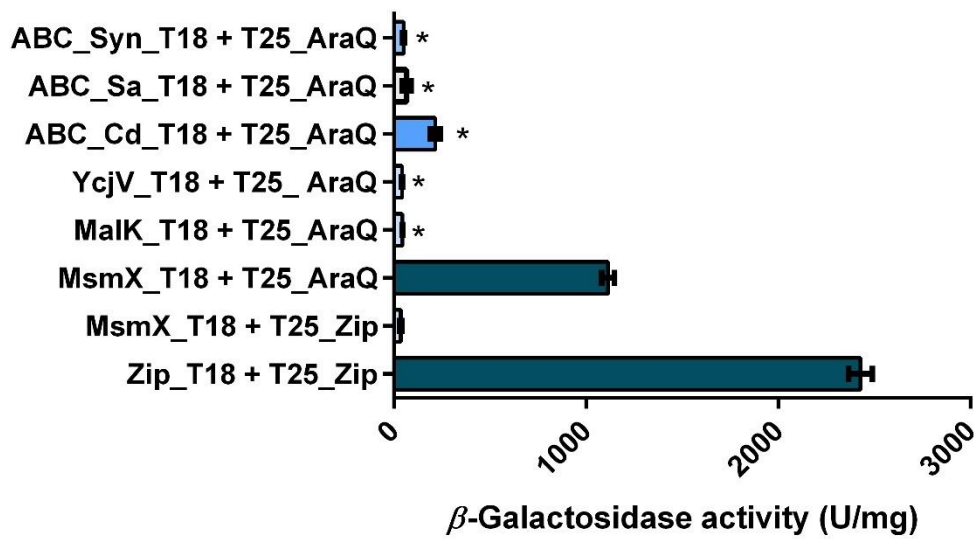


Figure 32. Putative multitask ATPases_T18 + T25_AraQ β -galactosidase assays. By using BACTH system, interaction between MsmX_T18 and T25_AraQ was quantified by determining β -galactosidase activity, expressed in units per mg of bacteria, and was used as control for wild-type interaction. Interaction between other putative multitask ATPases fused to T18 fragment and T25_AraQ was also determined by this method. As positive control Zip_T18 + Zip_T25 co-transformation was used as positive control and MsmX_T18 + Zip_T25 was used as negative control. At least three independent assays were performed for each co-transformation and error bars indicate the standard mean deviation. Statistical significance between enzymatic activities determined for each co-transformant and co-transformant MsmX_T18 + T25_AraQ is indicated (*, $p < 0.05$).

ABC_Cd is also a non-characterized ATPase and no structural or biochemical studies were published, but these results may suggest that the overall structure, especially the pocket responsible for the docking of the coupling helices may be similar to MsmX, in structure and in electrostatic potential.

ABC_Sa from the human pathogen *Staphylococcus aureus*, was also shown to replace MsmX *in vivo*, even though with less efficiency than ABC_Cd. This ATPase displayed the highest sequence identity with our model, but unlike MsmX and ABC_Cd, ABC_Sa appears to be part of an operon encoding a complete ABC carbohydrate uptake system (Leisico *et al.*, 2020). β -galactosidase assays were performed for interactions ABC_Sa_T18 + AraP_T25 and ABC_Sa_T18 + T25_AraQ, having β -galactosidase activities of 1429.38 ± 7.74 U/mg and 65.49 ± 19.32 U/mg, respectively. This represents a reduction of 1.6-fold when compared with the wild-type interaction with AraP and an almost total abolishment of the interaction with AraQ when compared with the wild-type. Considering that growth kinetic parameters *in B. subtilis* for ABC_Sa were 373.7 ± 45.8 min, higher time than MsmX and ABC_Cd, it is not surprising that the co-transformation ABC_Sa_T18 + AraP_T25 has less interaction than the wild-type interaction between MsmX and AraP, with ABC_Sa_T18 + T25_AraQ displaying an enzymatic activity very close to the negative control, suggesting ABC_Sa cannot interact with AraQ. Since ABC_Sa can energize the AraPQ translocator in *B. subtilis*, the lack of interaction measured with the BACTH system between the two proteins is unexpected. A possible explanation may rely on the type of test performed: while determining growth kinetic parameters, the ability of an ATPase dimer to energize the entire transporter is assessed but when using BACTH system to perform β -galactosidase

assays, only one ATPase *versus* one TMD is being studied, and it may be necessary for the ATPase be present as a dimer and integrated on the full transporter to interact with the TMDs. It is also known that the C-terminal tail of one of the TMDs, in this case AraQ, is necessary for the translocation cycle since this region interacts with the Q-loop from the ATPase to allow changes in conformation. Therefore, by looking at our results, they suggest that ABC_Sa cannot interact with AraQ alone and it only interacts with AraQ when integrated in the full transporter, probably with less efficiently than MsmX and ABC_Cd, which correlates with results obtained *in vivo* in *B. subtilis*.

Beyond the Firmicutes phylum, another NBD was identified with high similarity (71%) to MsmX. ABC_Syn, from *Synechocystis* sp., was shown to energize the AraNPQ system. This putative multitask ATPase, was shown to be an orphan NBD, like MsmX and ABC_Cd and like YcjV it is probably a pseudogene since it does not have an RBS (Leisico *et al.*, 2020). For β -galactosidase assays, the sequence of the RBS and ABC_Syn present in pPS7 (Leisico *et al.*, 2020) were cloned in pUT18, providing a functional Shine-Dalgarno sequence that allows translation of ABC_Syn. Interactions ABC_Syn_T18 + AraP_T25 and ABC_Syn_T18 + T25_AraQ were determined, yielding β -galactosidase activities of 37.18 ± 3.47 U/mg and 50.16 ± 3.45 U/mg. These values are practically the same as the negative control, suggesting that this ATPase cannot interact with AraPQ TMDs, which was not expected, since published results by Leisico *et al.*, 2020, show a doubling time of 328.5 ± 7.4 min, very close doubling time to the one determined for ABC_Sa, about 3-fold the wild-type doubling time. This can be due to the nature of the test, since when using the BACTH system only the interaction between one TMD and one NBD is being assessed, which can affect the interaction because these proteins are not integrated on a full transporter or maybe ABC_Syn can only interact with these TMDs when present as a dimer.

In summary, these results show that every ATPase previously shown to replace MsmX as energizers of the AraNPQ system interacts with the NBDs to different degrees, except for MalK, YcjV and ABC_Syn. It is also evident that AraP presents higher levels of interaction with the tested NBDs, which can probably be due to the nature of the test performed, as described previously, or to the less conserved EAA motif in AraQ. It was also shown a possible correlation between growth kinetic parameters determined in *B. subtilis* and these β -galactosidase activities in *E. coli* (Figure 33), since ATPases that lead to a higher doubling time, displayed lower enzymatic activities, meaning a lower interaction NBD-TMD. Differences in interaction level and consequently doubling time in these putative multitask NBDs, maybe be due to variations in highly conserved motifs within TMDs and NBDs and/or, as structural studies have shown, to wider or slender pocket in NBDs where coupling helices are inserted during the transport. This quantitative interaction characterization is also correlated with the previous results presented in subchapter 3.2.1. As previously stated, the strongest interaction

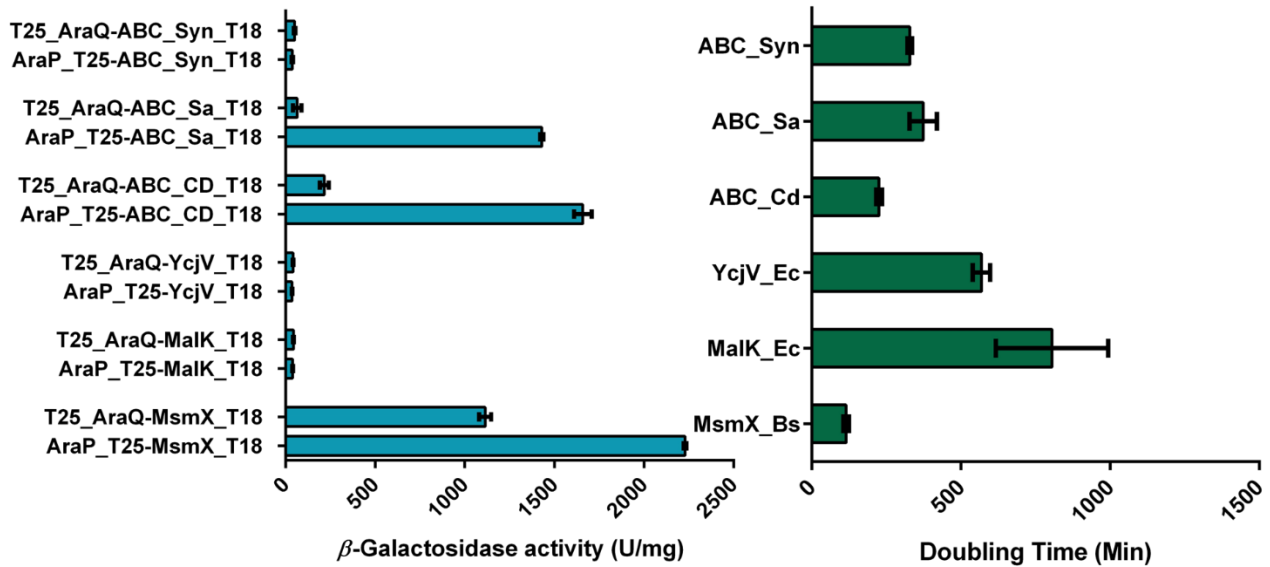


Figure 33. Schematic representation of the correlation found between multitask NBD - TMDs (β -Galactosidase activity (U/mg) and its doubling time (min). A correlation between NBD-TMDs interaction measured in *E. coli* and the respective functionality of the complete transporter in *B. subtilis* was found showing that higher β -galactosidase activities, suggesting higher protein interactions, correspond to lower doubling times, suggesting faster growth. Lower β -galactosidase activity, suggest lower or even lack of interaction between proteins and correspond to higher doubling time, suggesting slower growth.

displayed the higher β -galactosidase activity and the bluest colour, while non-interacting co-transformants displayed β -galactosidase activities similar to the negative control, as well as similar white colour.

3.2.3. Detection of *in vivo* accumulation of multitask ATPases by Western blot analysis

Western Blot analysis was performed to confirm accumulation of multitask ATPases *in vivo*, to determine if the observed phenotype of β -galactosidase assays is due to the interaction of different putative multitask ATPases or if the use of the BACTH system impairs the accumulation of these NBDs. Thus, *E. coli* BTH101 cells, bearing the co-transformations Zip_T18 + Zip_T25, MsmX_T18 + AraP_T25, MalK_T18 + AraP_T25, YcjV_T18 + AraP_T25, ABC_Cd_T18 + AraP_T25, ABC_Sa_T18 + AraP_T25 and ABC_Syn_T18 + AraP_T25, were harvested and treated as described in subchapters 2.10 and 3.1.3. Zip_T18 + Zip_T25 co-transformation was loaded into the first lane, as the positive control and NZYcolour protein marker II (NZYTech, Portugal), was loaded into the second lane. On lanes 3 to 8, co-transformations MalK_T18 + AraP_T25, YcjV_T18 + AraP_T25, ABC_Cd_T18 + AraP_T25, ABC_Sa_T18 + AraP_T25 and ABC_Syn_T18 + AraP_T25, were loaded. Results are summarized in Figure 34.

In Figure 34A, it described the successful electrotransfer of the fractionated proteins from the SDS-PAGE to the nitrocellulose membrane, stained with Ponceau Red and in Figure 34B is displayed the autoradiography plate. Even though optimization of Western blot analysis was previously performed (described in subchapter 3.1.3), some background is still present, especially on the MsmX lane.

However, other lanes were clear, and a band in every lane can be seen with the expected sizes of 25.9 kDa, 61.8 kDa, 61.4 kDa, 60.8 kDa, 62.4 kDa, 61.8 kDa and 61.3 kDa, meaning there is accumulation of every putative multitask ATPase. Although there are apparent differences in accumulation, especially between MsmX and other NBDs, that could explain differences in β -galactosidase activity, no certain conclusion can be drawn since no housekeeping gene was used as control. Thus, all NBDs studied in β -galactosidase assays using the BACTH system accumulate *in vivo*, and lack of interaction detected in this test cannot be attributed to protein degradation or misfolding. These results, along with published work by Leisico *et al.*, 2020, suggest that all ATPases are in fact multitask NBDs capable of replace MsmX with different efficiencies, that are correlated with NBD-TMD interaction level, with MalK and YcjV being the only ones that cannot replace MsmX, since they cannot interact with the AraNPQ translocator, even when accumulating *in vivo* and maintaining ATPase activity.

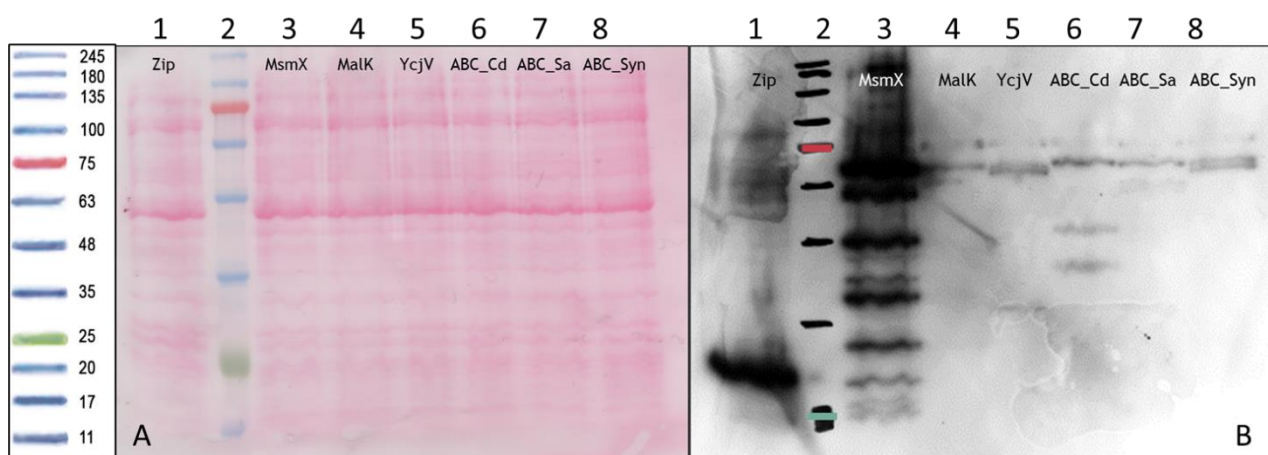


Figure 34. Multitask ATPases_T18 detection by Western Blot analysis, from *E. coli* BTH101 growth. In lane 2, NZYcolour protein marker II was loaded. In lanes 1, 3-8 cells of co-transformants Zip_T18 + Zip_T25, MsmX_T18 + AraP_T25, MalK_T18 + AraP_T25, YcjV_T18 + AraP_T25, ABC_Cd_T18 + AraP_T25, ABC_Sa_T18 + AraP_T25, ABC_Syn_T18 + AraP_T25 treated with PMSF and then resuspended in Loading and Running Buffer were loaded. A) Nitrocellulose membrane stained with Ponceau Red, showing the fractionated proteins transferred from an SDS-PAGE gel (12.5%). B) Autoradiographic plate after detection with α -Cya antibody. Zip_T18 (25.9 kDa), MsmX (61.8 kDa), MalK (61.4 kDa), YcjV (60.8 kDa), ABC_Cd (62.4 kDa), ABC_Sa (61.8 kDa) and ABC_Syn (61.3 kDa).

Chapter 4

Conclusions and Future perspectives

4. Conclusions and Future Perspectives

Since the discovery of ABC transporters, several surprising characteristics have been identified. First it was thought that each ABC type I system was complete, with every component being specific for their complex, but with the improvement of genomic analysis, it allowed to unveil the existence of orphan ATPases with the capacity to energise several systems (Ferreira & Sá-Nogueira, 2010; Marion *et al.*, 2011; Ferreira *et al.*, 2017). Recently, as other putative orphan or clustered multitask ATPases have been identified, their capacity to replace MsmX *in vivo* reveal intra- and interspecies exchangeability within the Firmicutes phylum (Leisico *et al.*, 2020), raising the question of what makes a multitask NBD and what contributes to these promiscuous interactions. Thus, in this work, not only NBD-TMD contact regions were characterize using the AraNPQ-MsmX system, from *B. subtilis*, as model and by targeting specific amino acids (A210, D211, loop containing the EAA motif in AraP and R178, D180 and the C-terminal tail GVKG in AraQ) by site-directed mutagenesis, but interaction between these recently identified ATPase (ABC_Sa, from *S. aureus*, ABC_Cd from *C. difficile* 630, MalK and YcjV from *E. coli*) and the AraPQ TMDs were characterized, using the BACTH system. Posteriorly, another putative multitask ATPase was identified beyond the Firmicutes phylum able to replace the function of MsmX *in vivo*, ABC_Syn from *Synechocystis* sp.

The results presented in this dissertation indicated that some mutations in both AraP and AraQ affect with different degrees the capacity of MsmX to interact with them. In AraP changing its entire loop containing the EAA motif to the C3 loop of MalF, from *E. coli*, leads to the complete abolition of interaction while only targeting residues A210 and D211 has a non-significant impact on interaction with MsmX, having interaction similar to the wild-type. In AraQ, both residues R178 and D180 have similar negative effects in interaction with MsmX, leading to a drastic reduction. Both these mutations and AraP mutations shows the importance of conserved residues within the EAA motif. Since this motif is normally an α -helix that is inserted within an NBD monomer and is involved in transmitting the conformational changes of the ATPase dimer to the TMDs, allowing the alternating access mechanism, it is probable that all these mutations cause disturbances in the electrostatic surface and conformational changes that will weaken this major interaction zone. Another interesting result was the effect of the deletion of the C-terminal tail of AraQ. Previous results have shown a severe negative impact on *B. subtilis* doubling time (Ferreira *et al.*, 2017), and here, testing exclusively protein-protein interactions between this mutant AraQ and MsmX revealed a non-significant reduction of interaction. This result, strongly support the hypothesis of Oldham *et al.*, 2007 concerning the role of the C-terminal region of MalG and by that in the formation of the catalytic intermediate state of the transporter. It is possible, that the hydrogen bonds between AraQ tail and Q-loop's MsmX as a key point for the intermediate state on the transport cycle and may represent a new signature feature of bacterial ABC sugar importers of the subfamily of carbohydrate uptake transporter 1 (CUT1).

By using the BACTH system, interaction between recently discovered ATPases and AraPQ TMDs were qualitatively and quantitatively characterized. Here, a correlation between NBD-TMD interaction measured in *E. coli* and the respective function of the whole transporter in *B. subtilis* was found. For lower doubling time in *B. subtilis*, that corresponds to faster growth due to higher degree of functionality of the hybrid transporters composed by the putative multitask ATPases and AraNPQ, higher values of β -galactosidase activity were determined, meaning higher interaction between TMD-NBD. For higher doubling times, like the ones determined for *E. coli* ATPases, MalK and YcjV, which meant poor or non-complementation between these ATPases and AraNPQ, lower β -galactosidase activities were revealed, which meant weaker or no interaction. This correlation was also found in other studies performed in our laboratory with mutants of the MsmX-AraNPQ system (Microbial Genetics Laboratory, unpublished results). Together, these results validate the use of the BACTH system as a valid method to study and complement studies about PPIs in ABC-type I importers. The qualitative assay is a simple method for an initial screening of interacting or non-interacting proteins, while the quantitative studies allow a more detailed characterization of the effect of both mutant TMDs and NBD.

In the future, structural and spectroscopic studies of both multitask ATPases and TMDs (mutant and wild-type) may help to shed light on these proteins and on the importance of these residues in TMD-NBD interaction. It is also important to assess the effect of the mutations studied in this dissertation *in vivo*, by introducing them into the *B. subtilis* chromosome, determine its growth kinetic parameters and posteriorly correlate the doubling times with the interaction studies described in this work. To help to validate the use of BACTH system as a method to study PPI, it would be also interesting to develop a method to also detect T25 fragment in Western Blot analysis, by producing an antibody against this polypeptide or by inserting a tag, like and his-tag tail. Also, it could be interesting to study and confirm interactions between TMDs-NBDs by a different methodology, namely, by pull-down assays or by co-immunoprecipitation, using *B. subtilis* cell extracts.

Throughout the years, it is increasingly evident the importance of sugar importers and sugar metabolism in bacteria of the Firmicutes phylum in colonization and pathogenesis, because it has been observed that inactivation of one multitask ATPase can lead to the inactivation of several sugar transporters (Marion *et al.*, 2011; Buckwalter & King, 2012; Leisico *et al.*, 2020). Since these ABC importers are only present in bacteria and not in humans, alongside with the capacity to be shared inter-species, it remains to see if they can constitute a target for the development of antimicrobial therapies, like antibiotics or vaccines. These therapies could lead to virulence attenuations or even colonization depletion (Marion *et al.*, 2011; Leisico *et al.*, 2020).

Chapter 5

References

5. References

- Abreu, B., Cruz, C., Oliveira, A. S. F., & Soares, C. M. (2021). ATP hydrolysis and nucleotide exit enhance maltose translocation in the MalFGK₂E importer. *Scientific Reports*, 11(1), 1–17. <https://doi.org/10.1038/s41598-021-89556-y>
- Albers, S. V., Elferink, M. G. L., Charlebois, R. L., Sensen, C. W., Driessen, A. J. M., & Konings, W. N. (1999). Glucose transport in the extremely thermoacidophilic *Sulfolobus solfataricus* involves a high-affinity membrane-integrated binding protein. *Journal of Bacteriology*, 181(14), 4285–4291. <https://doi.org/10.1128/jb.181.14.4285-4291.1999>
- Anderson, M. P., Berger, H. A., Rich, D. P., Gregory, R. J., Smith, A. E., & Welsh, M. J. (1991). Nucleoside triphosphates are required to open the CFTR chloride channel. *Cell*, 67(4), 775–784. [https://doi.org/10.1016/0092-8674\(91\)90072-7](https://doi.org/10.1016/0092-8674(91)90072-7)
- Bacon, K., Blain, A., Bowen, J., Burroughs, M., McArthur, N., Menegatti, S., & Rao, B. M. (2021). Quantitative Yeast-Yeast Two Hybrid for the Discovery and Binding Affinity Estimation of Protein-Protein Interactions. *ACS Synthetic Biology*, 10(3), 505–514. <https://doi.org/10.1021/acssynbio.0c00472>
- Bao, H., & Duong, F. (2013). ATP Alone Triggers the Outward Facing Conformation of the Maltose ATP-binding Cassette Transporter. *The Journal of Biological Chemistry*, 288(5), 3439–3448. <https://doi.org/10.1074/jbc.M112.431932>
- Bendert, A., & Pringle, J. R. (1991). Use of a Screen for Synthetic Lethal and Multicopy Suppressed Mutants To Identify Two New Genes Involved in Morphogenesis in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, 11(3), 1295–1305.
- Bidossi, A., Mulas, L., Decorosi, F., Colomba, L., Ricci, S., Pozzi, G., ... Oggioni, M. R. (2012). A functional genomics approach to establish the complement of carbohydrate transporters in *Streptococcus pneumoniae*. *PLoS ONE*, 7(3). <https://doi.org/10.1371/journal.pone.0033320>
- Biemans-Oldehinkel, E., Mahmood, N. A. B. N., & Poolman, B. (2006). A sensor for intracellular ionic strength. *Proceedings of the National Academy of Sciences of the United States of America*, 103(28), 10624–10629. <https://doi.org/10.1073/pnas.0603871103>
- Buckwalter, C. M., & King, S. J. (2012). Pneumococcal carbohydrate transport: food for thought. *Trends in Food Science and Technology*, 20(11), 517–522. <https://doi.org/10.1016/j.tim.2012.08.008>
- Chavez, J. D., Liu, N. L., & Bruce, J. E. (2011). Quantification of Protein–Protein Interactions with Chemical Cross-Linking and Mass Spectrometry. *J Proteome Res.*, 10(4), 1528–1537. <https://doi.org/10.1021/pr100898e>
- Chen, S., Oldham, M. L., Davidson, A. L., & Chen, J. (2013). Carbon catabolite repression of the maltose transporter revealed by X-ray crystallography. *Nature*, 499(7458), 364–368. <https://doi.org/10.1038/nature12232>

- Chevance, F. F. V, Erhardt, M., Lengsfeld, C., Lee, S., & Boos, W. (2006). Mlc of *Thermus thermophilus*: a Glucose-Specific Regulator for a Glucose/Mannose ABC Transporter in the Absence of the Phosphotransferase System. *Journal of Bacteriology*, 188(18), 6561–6571. <https://doi.org/10.1128/JB.00715-06>
- Crane, R. K. (1962). Hypothesis for mechanism of intestinal active transport of sugars. *Fed Proc.*, 21.
- Dassa, E. (1990). Cellular localization of the MalG protein from the maltose transport system in *Escherichia coli* K12. *MGG Molecular & General Genetics*, 222(1), 33–36. <https://doi.org/10.1007/BF00283019>
- Dassa, E., & Hofnung, M. (1985). Sequence of gene malG in *E. coli* K12: homologies between integral membrane components from binding protein-dependent transport systems. *The EMBO Journal*, 4(9), 2287–2293.
- Daus, M. L., Berendt, S., Wuttge, S., & Schneider, E. (2007). Maltose binding protein (MalE) interacts with periplasmic loops P2 and P1 respectively of the MalFG subunits of the maltose ATP binding cassette transporter (MalFGK₂) from *Escherichia coli*/*Salmonella* during the transport cycle. *Molecular Microbiology*, 66(5), 1107–1122. <https://doi.org/10.1111/j.1365-2958.2007.05982.x>
- Daus, M. L., Grote, M., Müller, P., Doebber, M., Herrmann, A., Steinhoff, H. J., ... Schneider, E. (2007). ATP-driven MalK dimer closure and reopening and conformational changes of the “EAA” motifs are crucial for function of the maltose ATP-binding cassette transporter (MalFGK₂). *Journal of Biological Chemistry*, 282(31), 22387–22396. <https://doi.org/10.1074/jbc.M701979200>
- Dautin, N., Karimova, G., & Ladant, D. (2003). Human Immunodeficiency Virus (HIV) Type 1 Transframe Protein Can Restore Activity to a Dimerization-Deficient HIV Protease Variant. *Journal of Virology*, 77(15), 8216–8226. <https://doi.org/10.1128/jvi.77.15.8216-8226.2003>
- Davidson, A. L., Dassa, E., Orelle, C., Chen, J., Ol, M. I. M., & Ev, B. I. O. L. R. (2008). Structure, Function, and Evolution of Bacterial ATP-Binding Cassette Systems. *Microbiology and Molecular Biology Reviews*, 72(2), 317–364. <https://doi.org/10.1128/MMBR.00031-07>
- Dawson, R. J. P., Hollenstein, K., & Locher, K. P. (2007). Uptake or extrusion: Crystal structures of full ABC transporters suggest a common mechanism. *Molecular Microbiology*, 65(2), 250–257. <https://doi.org/10.1111/j.1365-2958.2007.05792.x>
- Eitinger, T., Rodionov, D. A., Grote, M., & Schneider, E. (2011). Canonical and ECF-type ATP-binding cassette importers in prokaryotes: Diversity in modular organization and cellular functions. *FEMS Microbiology Reviews*, 35(1), 3–67. <https://doi.org/10.1111/j.1574-6976.2010.00230.x>
- Ferreira, M., Mendes, A. L., & Sá-Nogueira, I. de. (2017). The MsmX ATPase plays a crucial role in pectin mobilization by *Bacillus subtilis*. *PLoS ONE*, 12(12), 1–22. <https://doi.org/https://doi.org/10.1371/journal.pone.0189483>
- Ferreira, M., & Sá-Nogueira, I. de. (2010). A Multitask ATPase Serving Different ABC-Type Sugar Importers in *Bacillus subtilis*. *Journal of Bacteriology*, 192(20), 5312–5318.

<https://doi.org/10.1128/JB.00832-10>

- Fields, S., & Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Letters to Nature*, 340, 245–246.
- Fransen, M., Brees, C., Ghys, K., Amery, L., Mannaerts, G. P., Ladant, D., & Van Veldhoven, P. P. (2002). Analysis of mammalian peroxin interactions using a non-transcription-based bacterial two-hybrid assay. *Molecular and Cellular Proteomics*, 1(3), 243–252. <https://doi.org/10.1074/mcp.M100025-MCP200>
- Geetha, T., Langlais, P., Luo, M., Mapes, R., Lefort, N., Chen, C., ... Yi, Z. (2012). Label-Free Proteomic Identification of Endogenous, Insulin- Stimulated Interaction Partners of Insulin Receptor Substrate-1. *J Am Soc Mass Spectrom.*, 22(3), 457–466. <https://doi.org/10.1007/s13361-010-0051-2>
- Gilson, E., Alloing, G., Schmidt, T., Claverys, J. P., Dudler, R., & Hofnung, M. (1988). Evidence for high affinity binding-protein dependent transport systems in gram-positive bacteria and in Mycoplasma. *The EMBO Journal*, 7(12), 3971–3974. <https://doi.org/10.1002/j.1460-2075.1988.tb03284.x>
- Gilson, Eric, Nikaido, H., & Hofnung, M. (1982). Sequence of the malK gene in *E.coli* K12. *Nucleic Acids Research*, 10(22), 7449–7458. <https://doi.org/10.1093/nar/10.22.7449>
- Goodhead, I., & Darby, A. C. (2015). Taking the pseudo out of pseudogenes. *Current Opinion in Microbiology*, 23, 102–109. <https://doi.org/10.1016/j.mib.2014.11.012>
- Gouridis, G., Hetzert, B., Kiosze-Becker, K., de Boer, M., Heinemann, H., Nürenberg-Goloub, E., Tampé, R. (2019). ABCE1 Controls Ribosome Recycling by an Asymmetric Dynamic Conformational Equilibrium. *Cell Reports*, 28(3), 723-734.e6. <https://doi.org/10.1016/j.celrep.2019.06.052>
- Heravi, K. M., Watzlawick, H., & Altenbuchner, J. (2019). The *melREDCA* operon encodes a utilization system for the raffinose family of oligosaccharides in *Bacillus subtilis*. *Journal of Bacteriology*, 201(15), 1–12. <https://doi.org/10.1128/JB.00109-19>
- Higgins, C. F. (1992). ABC TRANSPORTERS: From Microorganisms to Man. *Annu. Rev. Cell Biol.*, 8, 67–113.
- Higgins, C. F., Hiles, I. D., Salmond, G. P. C., Gill, D. R., Downie, J. A., Evans, I. J., Hermodson, M. A. (1986). A family of related ATP-binding subunits coupled to many distinct biological processes in bacteria. *Nature*, 323(6087), 448–450. <https://doi.org/10.1038/323448a0>
- Higgins, C. F., & Linton, K. J. (2004). The ATP switch model for ABC transporters. *Nature Structural & Molecular Biology*, 11(10), 918–926. <https://doi.org/10.1038/nsmb836>
- Hurtubise, Y., Shareck, F., Kluepfel, D., & Morosoli, R. (1995). A cellulase/xylanase-negative mutant of *Streptomyces fividans* 1326 defective in cellobiose and xylobiose uptake is mutated in a gene encoding a protein homologous to ATP-binding proteins. *Molecular Microbiology*, 17(2), 367–377.

- Hyde, S. C., Emsley, P., Hartshorn, M. J., Mimmack, M. M., Gileadi, U., Pearce, S. R., Higgins, C. F. (1990). Structural model of ATP-binding proteing associated with cystic fibrosis, multidrug resistance and bacterial transport. *Nature*, 346(6282), 362–365. <https://doi.org/10.1038/346362a0>
- Iacobucci, I., Monaco, V., Cozzolino, F., & Monti, M. (2021). From classical to new generation approaches: An excursus of -omics methods for investigation of protein-protein interaction networks. *Journal of Proteomics*, 230, 103990. <https://doi.org/10.1016/j.jprot.2020.103990>
- Inácio, J. M., & De Sá-Nogueira, I. (2008). Characterization of *abn2* (*yxjA*), encoding a *Bacillus subtilis* GH43 arabinanase, Abn2, and its role in arabino-polysaccharide degradation. *Journal of Bacteriology*, 190(12), 4272–4280. <https://doi.org/10.1128/JB.00162-08>
- Inácio, J. M., Lopes Correia, I., & de Sá-Nogueira, I. (2008). Two distinct arabinofuranosidases contribute to arabino-oligosaccharide degradation in *Bacillus subtilis*. *Microbiology*, 154(9), 2719–2729. <https://doi.org/10.1099/mic.0.2008/018978-0>
- Ito, M., Morino, M., & Krulwich, T. A. (2017). Mrp Antiporters Have Important Roles in Diverse Bacteria and Archaea. *Frontiers in Microbiology*, 8(2325), 1–12. <https://doi.org/10.3389/fmicb.2017.02325>
- Jeckelmann, J., & Erni, B. (2019). Carbohydrate Transport by Group Translocation: The Bacterial Phosphoenolpyruvate: Sugar Phosphotransferase System. In A. Kuhn (Ed.), *Bacterial Cell Walls and Membranes. Subcellular Biochemistry* (pp. 223–274). Springer International Publishing. <https://doi.org/10.1007/978-3-030-18768-2>
- Juers, D. H., Matthews, B. W., & Huber, R. E. (2012). LacZ β -galactosidase: Structure and function of an enzyme of historical and molecular biological importance. *Protein Science*, 21(12), 1792–1807. <https://doi.org/10.1002/pro.2165>
- Karasawa, A., Erkens, G. B., Berntsson, R. P. A., Otten, R., Schuurman-Wolters, G. K., Mulder, F. A. A., & Poolman, B. (2011). Cystathionine β -synthase (CBS) domains 1 and 2 fulfill different roles in ionic strength sensing of the ATP-binding cassette (ABC) transporter OpuA. *Journal of Biological Chemistry*, 286(43), 37280–37291. <https://doi.org/10.1074/jbc.M111.284059>
- Karimova, G., Gaudiard, E., Davi, M., Ouellette, S. P., & Ladant, D. (2017). Protein – Protein Interaction : Bacterial Two-Hybrid. In *Bacterial Protein Secretion Systems: Methods and Protocols* (Vol. 1615, pp. 159–176). <https://doi.org/10.1007/978-1-4939-7033-9>
- Karimova, G., Pidoux, J., Ullmann, A., & Ladant, D. (1998). A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 95(10), 5752–5756. <https://doi.org/10.1073/pnas.95.10.5752>
- Kellermann, O., & Szmelcman, S. (1974). Active Transport of Maltose in *Escherichia coli* K12: Involvement of a “Periplasmic” Maltose Binding Protein. *European Journal of Biochemistry*, 47(1), 139–149. <https://doi.org/10.1111/j.1432-1033.1974.tb03677.x>
- Klobucar, K., & Brown, E. D. (2018). Use of genetic and chemical synthetic lethality as probes of complexity in bacterial cell systems. *FEMS Microbiology Reviews*, 42(1), 81–99.

<https://doi.org/10.1093/femsre/fux054>

- Kojic, M., Milosevic, M., Isailovic, V., & Kojic, N. (2011). Transport in biological systems. *Journal of the Serbian Society for Computational Mechanics*, 5(2), 101–128.
- Krispin, O., & Allmansberger, R. (1998). The *Bacillus subtilis* AraE protein displays a broad substrate specificity for several different sugars. *Journal of Bacteriology*, 180(12), 3250–3252. <https://doi.org/10.1128/jb.180.12.3250-3252.1998>
- Kundig, W. (1974). Molecular interactions in the bacterial phosphoenolpyruvate phosphotransferase system (PTS). *Journal of Supramolecular and Cellular Biochemistry*, 2(5–6), 695–714. <https://doi.org/10.1002/jss.400020514>
- Kundig, W., & Roseman, S. (1971). Sugar transport. Isolation of a phosphotransferase system from *Escherichia coli*. *Journal of Biological Chemistry*, 246(5), 1393–1406. [https://doi.org/10.1016/S0021-9258\(19\)76986-5](https://doi.org/10.1016/S0021-9258(19)76986-5)
- Kundig, Werner, Ghosh, S., & Roseman, S. (1964). Phosphate Bound To Histidine in a Protein As an Intermediate in a Novel Phospho-Transferase System*. *Proc Natl Acad Sci U S A*, 52(4), 1067–1074. Retrieved from https://search.proquest.com/docview/1763349336?accountid=17215%250Ahttps://limo.libis.be/services/KULeuven?url_ver=Z39.882004&rft_val_fmt=info:ofi/fmt:kev:mtx:journal&genre=article&sid=ProQ:ProQ%253Aeducation&atitle=Nobility%252C+Competence%252C+and+Disr
- Leal, T. F., & Sá-Nogueira, I. De. (2004). Purification, characterization and functional analysis of an endo-arabinanase (AbnA) from *Bacillus subtilis*. *FEMS Microbiology Letters*, 241(1), 41–48. <https://doi.org/10.1016/j.femsle.2004.10.003>
- Leisico, F., Godinho, L. M., Gonçalves, I. C., Silva, S. P., Carneiro, B., Romão, M. J., Nogueira, I. D. S. (2020). Multitask ATPases (NBDs) of bacterial ABC importers type I and their interspecies exchangeability. *Scientific Reports*, 10, 1–17. <https://doi.org/10.1038/s41598-020-76444-0>
- Lite, T. L. V., Grant, R. A., Nokedal, I., Littlehale, M. L., Guo, M. S., & Laub, M. T. (2020). Uncovering the basis of protein-protein interaction specificity with a combinatorially complete library. *ELife*, 9(ii), 1–57. <https://doi.org/10.7554/eLife.60924>
- Locker, K. P. (2002). The *E. coli* BtuCD structure: A framework for ABC Transporter Architecture and Mechanism. *Science*, 296(5570), 1091–1098. <https://doi.org/10.1126/science.1071142>
- Mächtel, R., Narducci, A., Griffith, D. A., Cordes, T., & Orelle, C. (2019). An integrated transport mechanism of the maltose ABC importer. *Research in Microbiology*, 170(8), 321–337. <https://doi.org/10.1016/j.resmic.2019.09.004>
- Marion, C., Aten, A. E., Woodiga, S. A., King, S. J., & Al, M. E. T. (2011). Identification of an ATPase, MsmK, Which Energizes Multiple Carbohydrate ABC Transporters in *Streptococcus pneumoniae*. *Infection and Immunity*, 79(10), 4193–4200. <https://doi.org/10.1128/IAI.05290-11>
- Marion, C., Burnaugh, A. M., Woodiga, S. A., & King, S. J. (2011). Sialic acid transport contributes to pneumococcal colonization. *Infection and Immunity*, 79(3), 1262–1269.

- <https://doi.org/10.1128/IAI.00832-10>
- Meng, J., Rojas, M., Bacon, W., Stickney, J. T., & Ip, W. (2005). Methods to Study Protein – Protein Interactions. In K. Turksen (Ed.), *Methods in Molecular Biology* (Vol. 289, pp. 341–357). Humana Press Inc.
- Miura, K. (2018). An Overview of Current Methods to Confirm Protein- Protein Interactions. *Protein & Peptide Letters*, 25(8), 728–733. <https://doi.org/10.2174/0929866525666180821122240>
- Narducci, A., Grif, D. A., & Cordes, T. (2019). An integrated transport mechanism of the maltose ABC importer. *Research in Microbiology*, 1–17. <https://doi.org/10.1016/j.resmic.2019.09.004>
- Ofran, Y., & Rost, B. (2003a). Analysing six types of protein-protein interfaces. *Journal of Molecular Biology*, 325(2), 377–387. [https://doi.org/10.1016/S0022-2836\(02\)01223-8](https://doi.org/10.1016/S0022-2836(02)01223-8)
- Ofran, Y., & Rost, B. (2003b). Analysing Six Types of Protein – Protein Interfaces. *J. Mol. Biol.*, 325, 377–387. [https://doi.org/10.1016/S0022-2836\(02\)01223-8](https://doi.org/10.1016/S0022-2836(02)01223-8)
- Oldham, M. L., Chen, S., & Chen, J. (2013). Structural basis for substrate specificity in the *Escherichia coli* maltose transport system. *Proceedings of the National Academy of Sciences of the United States of America*, 110(45), 18132–18137. <https://doi.org/10.1073/pnas.1311407110>
- Oldham, M. L., Khare, D., Quioco, F. A., Davidson, A. L., & Chen, J. (2007). Crystal structure of a catalytic intermediate of the maltose transporter. *Nature*, 450, 515–522. <https://doi.org/10.1038/nature06264>
- Ouellette, S. P., Karimova, G., Davi, M., & Ladant, D. (2017). Analysis of membrane protein interactions with a bacterial adenylate cyclase-based two-hybrid (BACTH) technique. *Current Protocols in Molecular Biology*, 2017(April), 20.12.1-20.12.24. <https://doi.org/10.1002/cpmb.36>
- Patzlaff, J. S., Van der Heide, T., & Poolman, B. (2003). The ATP/substrate stoichiometry of the ATP-binding cassette (ABC) transporter OpuA. *Journal of Biological Chemistry*, 278(32), 29546–29551. <https://doi.org/10.1074/jbc.M304796200>
- Poolman, B., & van der Heide, T. (2002). ABC transporters: one, two or four extracytoplasmic substrate-binding sites? *EMBO Reports*, 3(10), 938–943. Retrieved from <http://www.nature.com/embor/journal/v3/n10/abs/embor057.html>
- Purslow, J. A., Khatiwada, B., Bayro, M. J., & Venditti, V. (2020). NMR Methods for Structural Characterization of Protein-Protein Complexes. *Frontiers in Molecular Biosciences*, 7(January), 1–8. <https://doi.org/10.3389/fmolb.2020.00009>
- Quentin, Y., Fichant, G., & Joseph, C. (1999). Inventory , Assembly and Analysis of *Bacillus subtilis* ABC Transport Systems. *Journal of Molecular Biology*, 287, 467–484.
- Quioco, F. A., & Ledvina, P. S. (1996). Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: Variation of common themes. *Molecular Microbiology*, 20(1), 17–25. <https://doi.org/10.1111/j.1365-2958.1996.tb02484.x>
- Rao, V. S., Srinivas, K., Sujini, G. N., & Kumar, G. N. S. (2014). Protein-Protein Interaction Detection: Methods and Analysis. *International Journal of Proteomics*, 2014(ii), 1–12.

<https://doi.org/http://dx.doi.org/10.1155/2014/147648>

- Rice, A. J., Park, A., & Pinkett, H. W. (2014). Diversity in ABC transporters : Type I , II and III importers. *Critical Reviews in Biochemistry and Molecular Biology*, 49(5), 426–437. <https://doi.org/10.3109/10409238.2014.953626>
- Sá-Nogueira, I., Nogueira, T. V, Soares, S., & Lencastre, H. De. (1997). The *Bacillus subtilis* L-arabinose (ara) operon: Nucleotide Sequence , Genetic Organization and Expression. *Microbiology*, 143(1997), 957–969.
- Sá-Nogueira, I., & Ramos, S. S. (1997). Cloning, functional analysis, and transcriptional regulation of the *Bacillus subtilis* araE gene involved in L-arabinose utilization. *Journal of Bacteriology*, 179(24), 7705–7711. <https://doi.org/10.1128/jb.179.24.7705-7711.1997>
- Saito, A., Fujii, T., Shinya, T., Shibuya, N., Ando, A., & Miyashita, K. (2008). The *msiK* gene, encoding the ATP-hydrolysing component of N , N 9 -diacetylchitobiose ABC transporters , is essential for induction of chitinase production in *Streptomyces coelicolor* A3 (2). *Microbiology*, 154, 3358–3365. <https://doi.org/10.1099/mic.0.2008/019612-0>
- Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning: A laboratory manual*. (C. S. Harbor, Ed.) (2nd ed.). NY: Cold Spring Harbor.
- Sampaio, M., Henne, A., Bo, A., Gutzat, R., Boos, W., Costa, M. S., & Santos, H. (2005). The High-Affinity Maltose / Trehalose ABC Transporter in the Extremely Thermophilic Bacterium *Thermus thermophilus* HB27 Also Recognizes Sucrose and Palatinose. *Journal of Bacteriology*, 187(4), 1210–1218. <https://doi.org/10.1128/JB.187.4.1210>
- Sanno, Y., Wilson, T. H., & Lin, E. C. C. (1968). Control of perefation to glycerol in cells of *Escherichia coli*. *Biochemical and Biophysical Research Communications*, 34(2), 78–86.
- Santos, J. A., Rempel, S., Mous, S. T. M., Pereira, C. T., Beek, J. Ter, de Gier, J. W., Slotboom, D. J. (2018). Functional and structural characterization of an ECF-type ABC transporter for vitamin B12. *ELife*, 7, 1–16. <https://doi.org/10.7554/eLife.35828>
- Saraon, P., Grozavu, I., Lim, S. H., Snider, J., Yao, Z., & Stagljar, I. (2017). Detecting Membrane Protein-protein Interactions Using the Mammalian Membrane Two-hybrid (MaMTH) Assay. *Current Protocols in Chemical Biology*, 9(1), 38–54. <https://doi.org/10.1002/cpch.15>
- Saurin, W., Hofnung, M., & Dassa, E. (1999). Getting in or out: Early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters. *Journal of Molecular Evolution*, 48(1), 22–41. <https://doi.org/10.1007/PL00006442>
- Saurin, W., & Koster, W. (1994). Bacterial binding protein-dependent permeases: characterization of distinctive signatures for functionailly related integral cytoplasmic membrane proteins. *Molecular Microbiology*, 12(6), 993–1004.
- Scheer, B. T. (1958). Active Transport: Definitions and Criteria. *Bulletin of Mathematical Biophysics*, 20. Retrieved from <https://link.springer.com/content/pdf/10.1007%2F02478301.pdf>
- Schlosser, A., Kampers, T., & Schrempf, H. (1997). The *Streptomyces* ATP-Binding Component MsiK

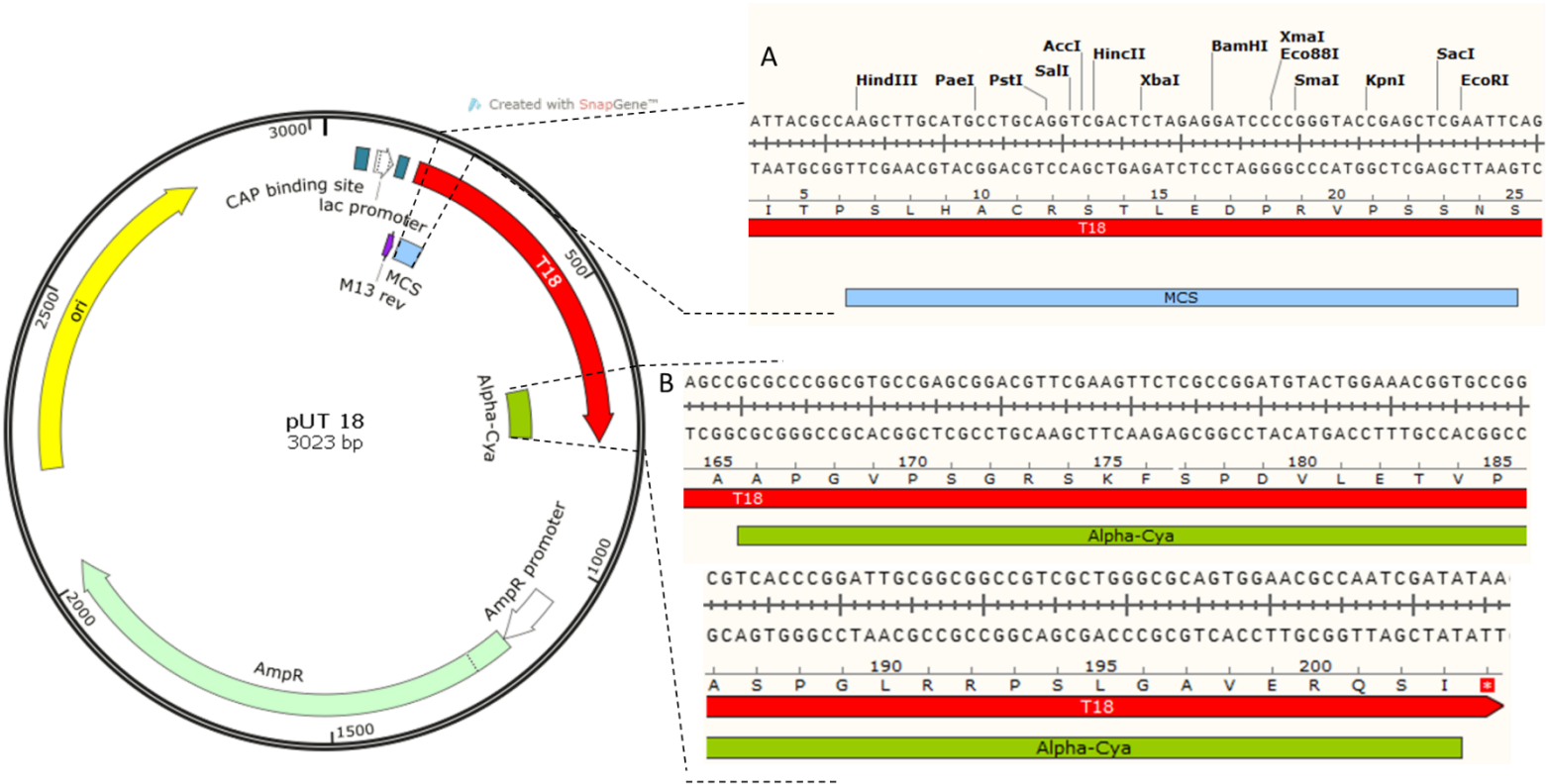
- Assists in Cellobiose and Maltose Transport. *Journal of Bacteriology*, 179(6), 2092–2095.
- Schneider, E., & Hunke, S. (1998). ATP-binding-cassette (ABC) transport systems: Functional and structural aspects of the ATP-hydrolyzing subunits/domains. *FEMS Microbiology Reviews*, 22(1), 1–20. [https://doi.org/10.1016/S0168-6445\(98\)00002-3](https://doi.org/10.1016/S0168-6445(98)00002-3)
- Schönert, S., Seitz, S., Krafft, H., Feuerbaum, E., Andernach, I., Witz, G., & Dahl, M. K. (2006). Maltose and Maltodextrin Utilization by *Bacillus subtilis*. *Journal of Bacteriology*, 188(11), 3911–3922. <https://doi.org/10.1128/JB.00213-06>
- Shuman, H. A., & Silhavy, T. J. (1981). Identification of the *malK* gene product. A peripheral membrane component of the *Escherichia coli* maltose transport system. *Journal of Biological Chemistry*, 256(2), 560–562. [https://doi.org/10.1016/s0021-9258\(19\)70005-2](https://doi.org/10.1016/s0021-9258(19)70005-2)
- Spurlino, J. C., Lu, G. Y., & Quioco, F. A. (1991). The 2.3-A resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. *Journal of Biological Chemistry*, 266(8), 5202–5219. [https://doi.org/10.1016/s0021-9258\(19\)67774-4](https://doi.org/10.1016/s0021-9258(19)67774-4)
- Steinke, A., Grau, S., Davidson, A. M. Y., Hofmann, E., & Ehrmann, M. (2001). Characterization of Transmembrane Segments 3, 4, and 5 of MalF by Mutational Analysis. *Journal of Bacteriology*, 183(1), 375–381. <https://doi.org/10.1128/JB.183.1.375>
- Stillwell, W. (Ed.). (2016). Membrane Transport. In *An Introduction to Biological Membranes* (2nd ed., pp. 423–451). Elsevier B.V. <https://doi.org/10.1016/B978-0-444-63772-7.00019-1>
- Stock, J., & Roseman, S. (1971). A sodium-dependent sugar co-transport system in bacteria. *Biochemical and Biophysical Research Communications*, 44(1), 132–138. [https://doi.org/10.1016/S0006-291X\(71\)80168-7](https://doi.org/10.1016/S0006-291X(71)80168-7)
- Stynen, B., Tournu, H., Tavernier, J., & Van Dijck, P. (2012). Diversity in Genetic In Vivo Methods for Protein-Protein Interaction Studies: from the Yeast Two-Hybrid System to the Mammalian Split-Luciferase System. *Microbiology and Molecular Biology Reviews*, 76(2), 331–382. <https://doi.org/10.1128/membr.05021-11>
- Swier, L. J. Y. M., Slotboom, D., & Poolman, B. (2016). ABC Importers. In A. George (Ed.), *ABC Transporters—40 Years on* (pp. 3–36). Springer International Publishing. <https://doi.org/10.1007/978-3-319-23476-2>
- Syafrizayanti, Betzan, C., Hoheisel, J. D., & Kastelic, D. (2014). Methods for analyzing and quantifying protein – protein interaction. *Expert Rev. Proteomics*, 11(1), 107–120. <https://doi.org/10.1586/14789450.2014.875857>
- Tam, N. K. M., Uyen, N. Q., Hong, H. A., Duc, L. H., Hoa, T. T., Claudia, R., Cutting, S. M. (2006). The Intestinal Life Cycle of *Bacillus subtilis* and Close Relatives. *JOURNAL OF BACTERIOLOGY*, 188(7), 2692–2700. <https://doi.org/10.1128/JB.188.7.2692>
- Tan, M., Gao, T., Liu, W., Zhang, C., Yang, X., & Zhu, J. (2015). MsmK, an ATPase, Contributes to Utilization of Multiple Carbohydrates and Host Colonization of *Streptococcus suis*. *PLoS ONE*,

- 10(7), 1–19. <https://doi.org/10.1371/journal.pone.0130792>
- Tanaka, K. J., Song, S., Mason, K., & Pinkett, H. W. (2018). Selective substrate uptake: The role of ATP-binding cassette (ABC) importers in pathogenesis. *Biochim Biophys Acta.*, 1860(4), 868–877. <https://doi.org/10.1016/j.bbamem.2017.08.011>. Selective
- Ter Beek, J., Guskov, A., & Slotboom, D. J. (2014). Structural diversity of ABC transporters. *The Journal of General Physiology*, 143(4), 419–435. <https://doi.org/10.1085/jgp.201411164>
- Theodoulou, F. L., & Kerr, I. D. (2015). ABC transporter research: going strong 40 years on. *Biochem. Soc. Trans.*, 43, 1033–1040. <https://doi.org/10.1042/BST20150139>
- Verheul, A., Glaasker, E., Poolman, B., & Abee, T. (1997). Betaine and L-carnitine transport by *Listeria monocytogenes* scott A in response to osmotic signals. *Journal of Bacteriology*, 179(22), 6979–6985. <https://doi.org/10.1128/jb.179.22.6979-6985.1997>
- Vetter, I. R., & Wittinghofer, A. (1999). Nucleoside triphosphate-binding proteins: Different scaffolds to achieve phosphoryl transfer. *Quarterly Reviews of Biophysics*, 32(1), 1–56. <https://doi.org/10.1017/S0033583599003480>
- Webb, A. J., Homer, K. A., & Hosie, A. H. F. (2008). Two Closely Related ABC Transporters in *Streptococcus mutans* Are Involved in Disaccharide and / or Oligosaccharide Uptake. *Journal of Bacteriology*, 190(1), 168–178. <https://doi.org/10.1128/JB.01509-07>
- West, I. C., & Mitchell, P. (1974). Proton/sodium ion antiport in *Escherichia coli*. *Biochemical Journal*, 144(1), 87–90. <https://doi.org/10.1042/bj1440087>
- Wong, J. H., Alfatah, M., Sin, M. F., Sim, H. M., Verma, C. S., Lane, D. P., & Arumugam, P. (2017). A yeast two-hybrid system for the screening and characterization of small-molecule inhibitors of protein-protein interactions identifies a novel putative Mdm2-binding site in p53. *BMC Biology*, 15(1), 1–17. <https://doi.org/10.1186/s12915-017-0446-7>

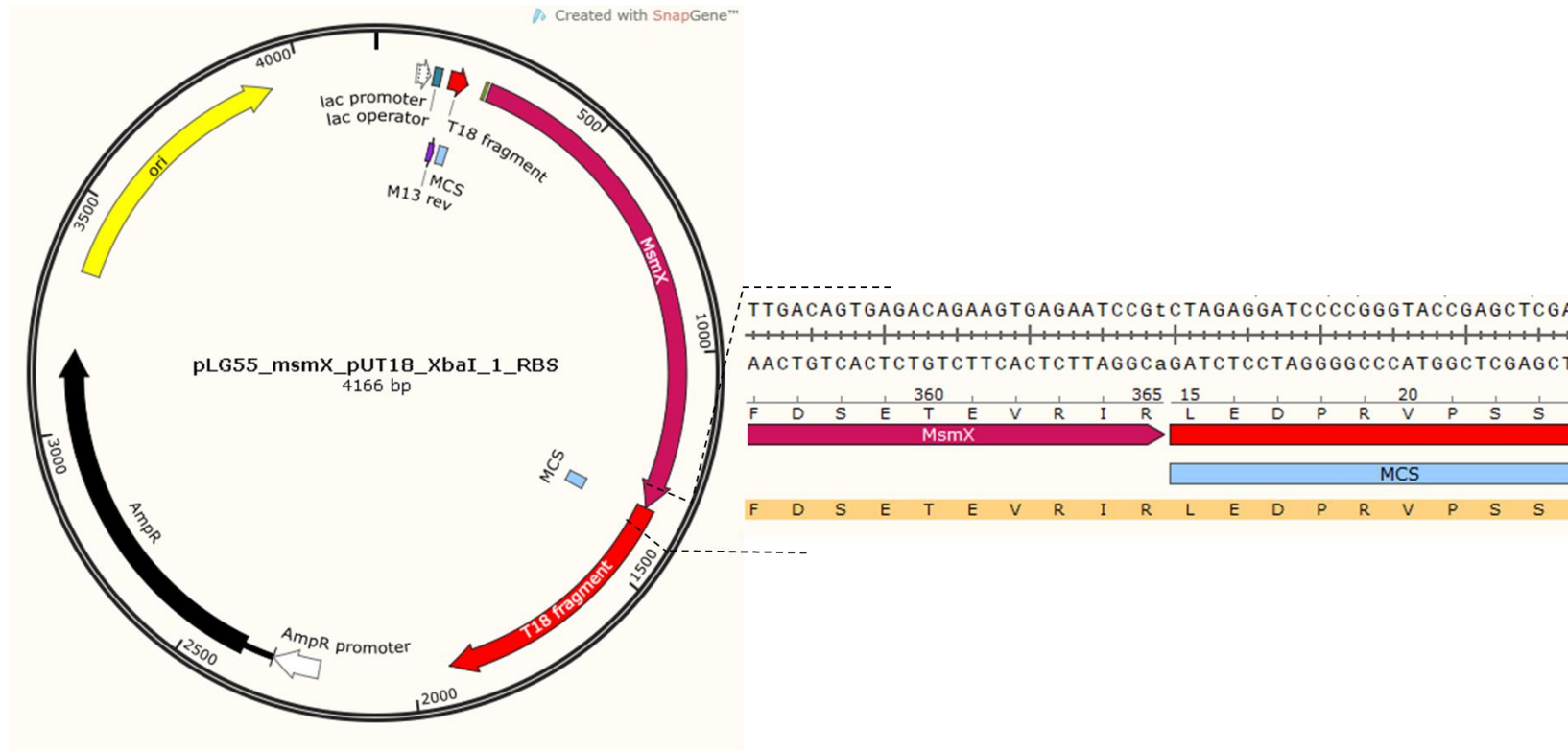
Chapter 6 |
Appendices

6. Appendices

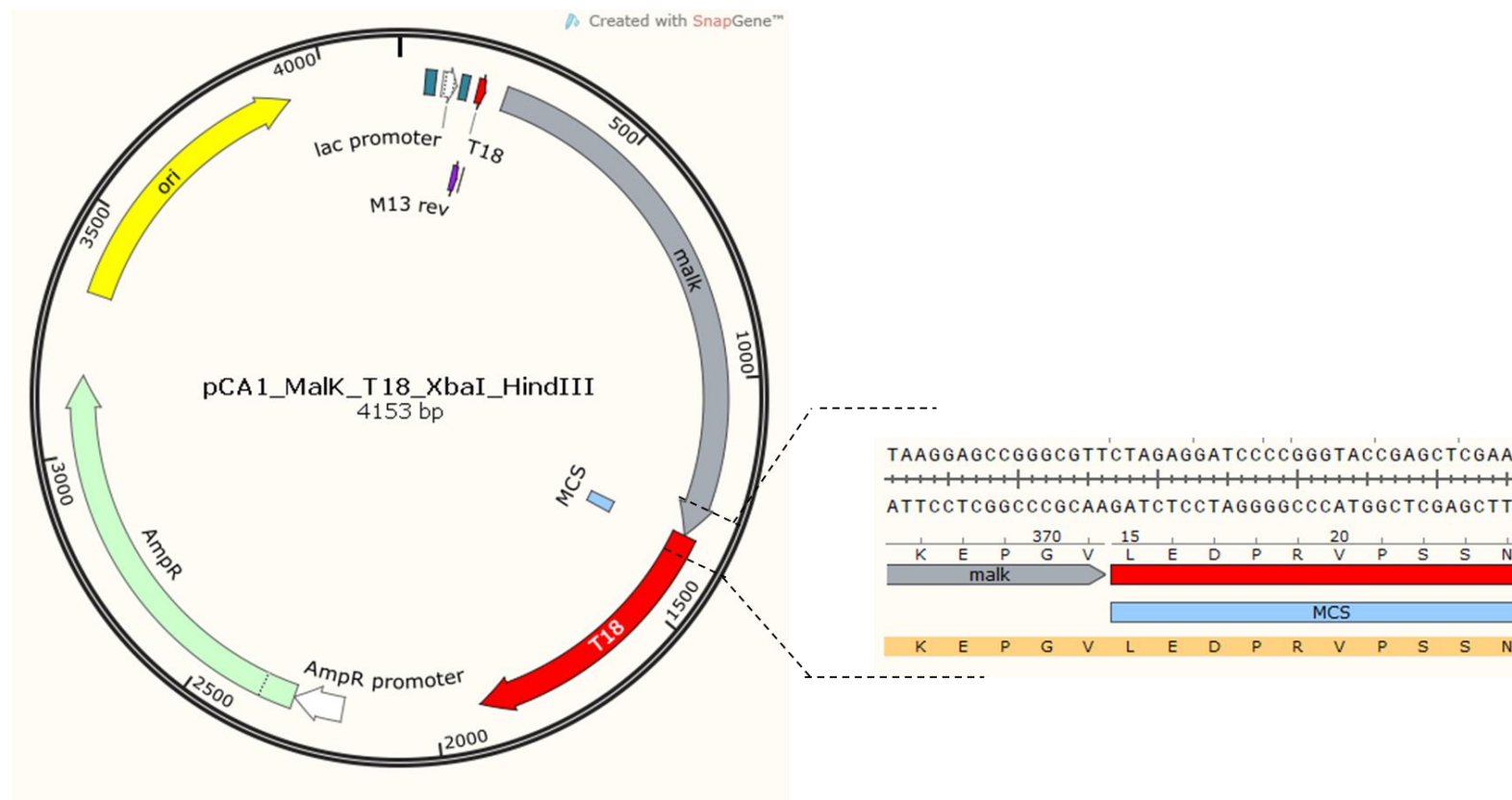
Appendix 1. Map of the high copy plasmid pUT18, encoding for the T18 fragment (corresponding to the 225-400 amino acids) of Adenylate Cyclase from *B. pertussis*. This plasmid expresses an ampicillin resistance selectable marker and is used to express chimeric proteins in which a heterologous polypeptide is fused to the N-terminal of T18. A) Detailed sequence of the multiple cloning site (MCS - sky blue). B) Detailed sequence of α -Cya antibody recognition site (green rectangle). T18 amino acid sequence is depicted in red.



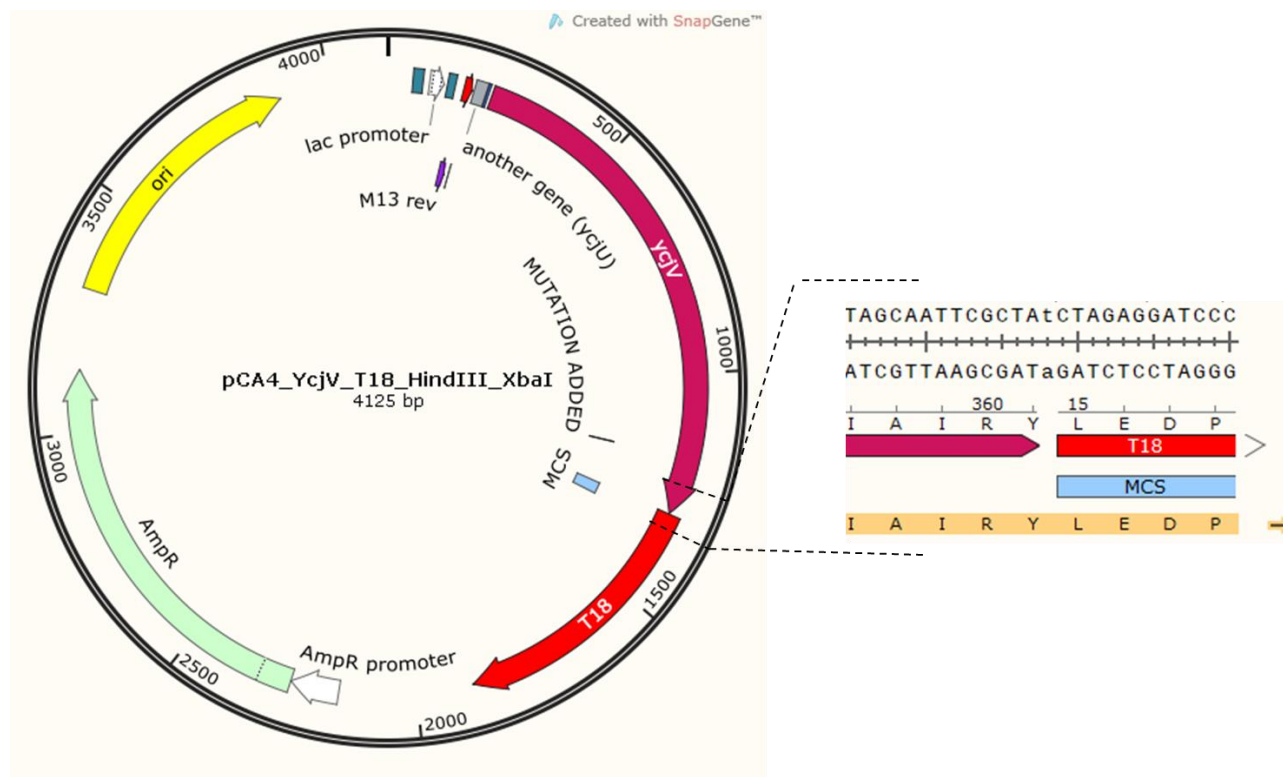
Appendix 2. Map of pLG55. Insertion of 1095 bp fragment containing *msmX* gene from *B. subtilis* 168T⁺ (amplified with primer Ara932 and Ara933 and digested with XbaI) between XbaI sites of pUT18. Detailed sequence of the fusion zone is shown in orange, MsmX amino acid sequence represented in pink, T18 amino acid sequence represented in red and MCS zone is depicted in sky blue.



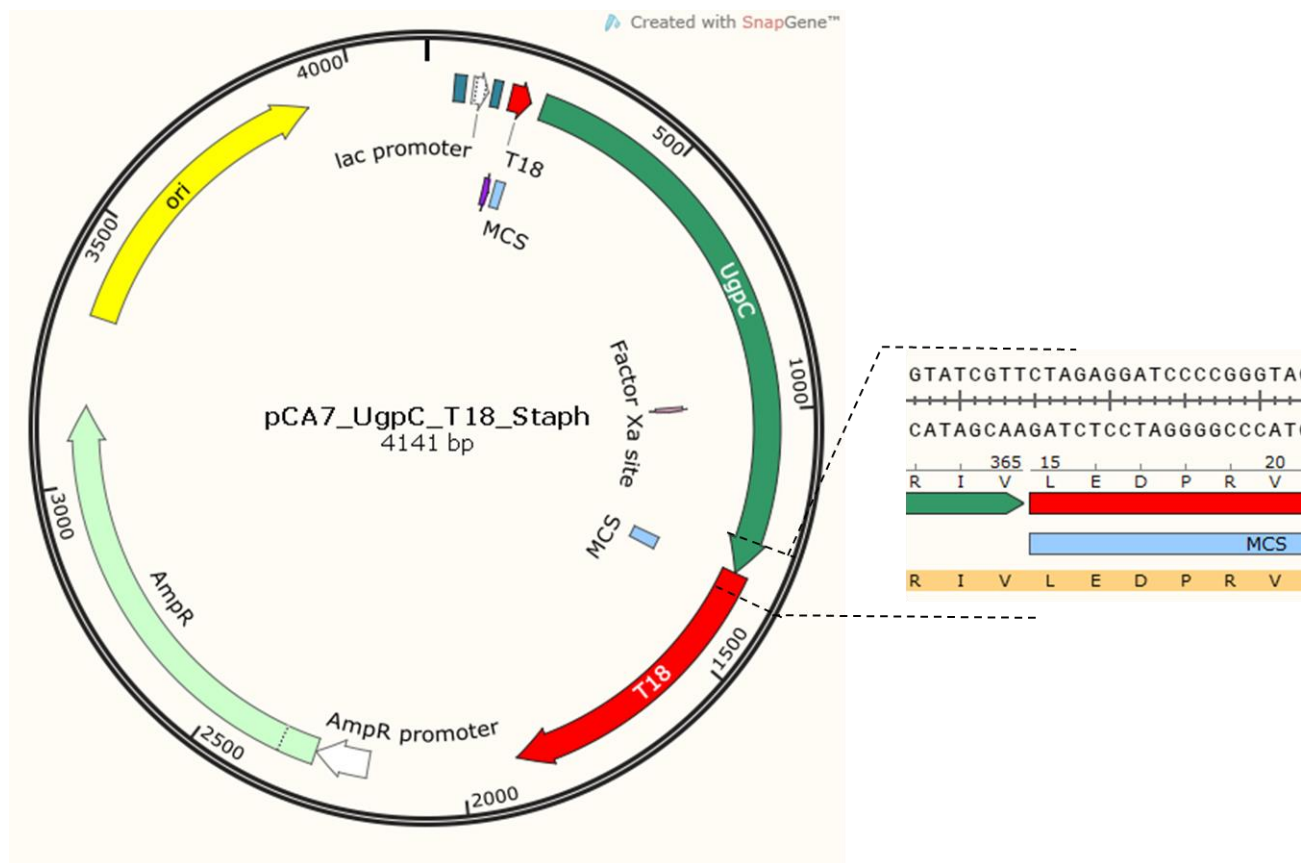
Appendix 3. Map of pCA1. Insertion of 1113 bp fragment containing *malK* gene from *E. coli* K-12 (amplified with primer Ara858 and Ara974 and digested with HindIII and XbaI) between the HindIII and XbaI sites of pUT18. Primers 166 and Ara958 were used for DNA sequencing. Detailed sequence of the fusion zone is shown in orange, MalK amino acid sequence represented in grey, T18 amino acid sequence represented in red and MCS zone is depicted in sky blue.



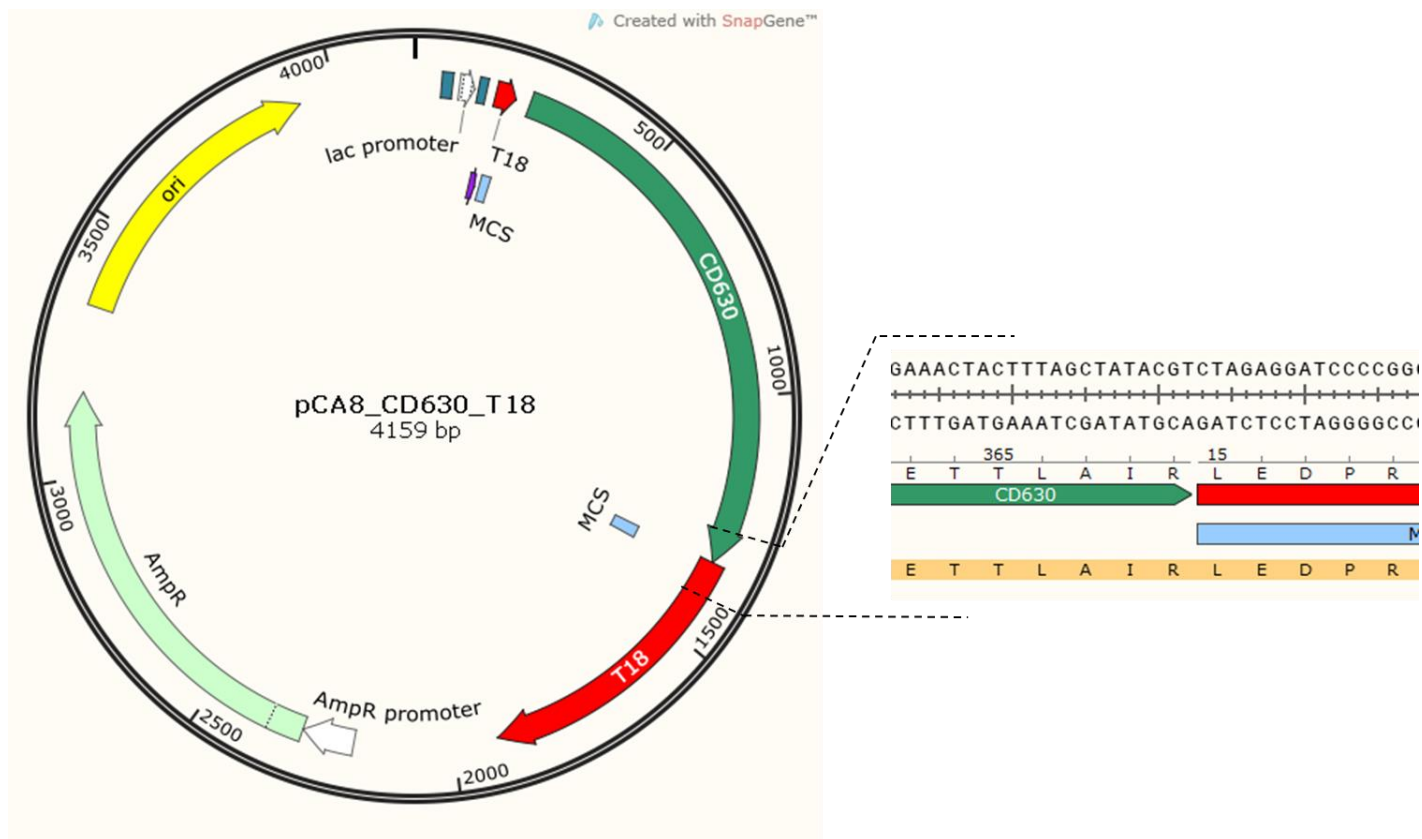
Appendix 4. Map of pCA4. Insertion of the wild-type phenotype of *ycjV* gene with the full length from *E. coli* K-12 (using primers Ara860 and Ara982 and digested with HindIII and XbaI) between the HindIII and XbaI sites of the pCA18 plasmid. Primers 166 and Ara958 were used for DNA sequencing. Detailed sequence of the fusion zone is shown in orange, YcjV amino acid sequence represented in pink, T18 amino acid sequence represented in red and MCS zone is depicted in sky blue.



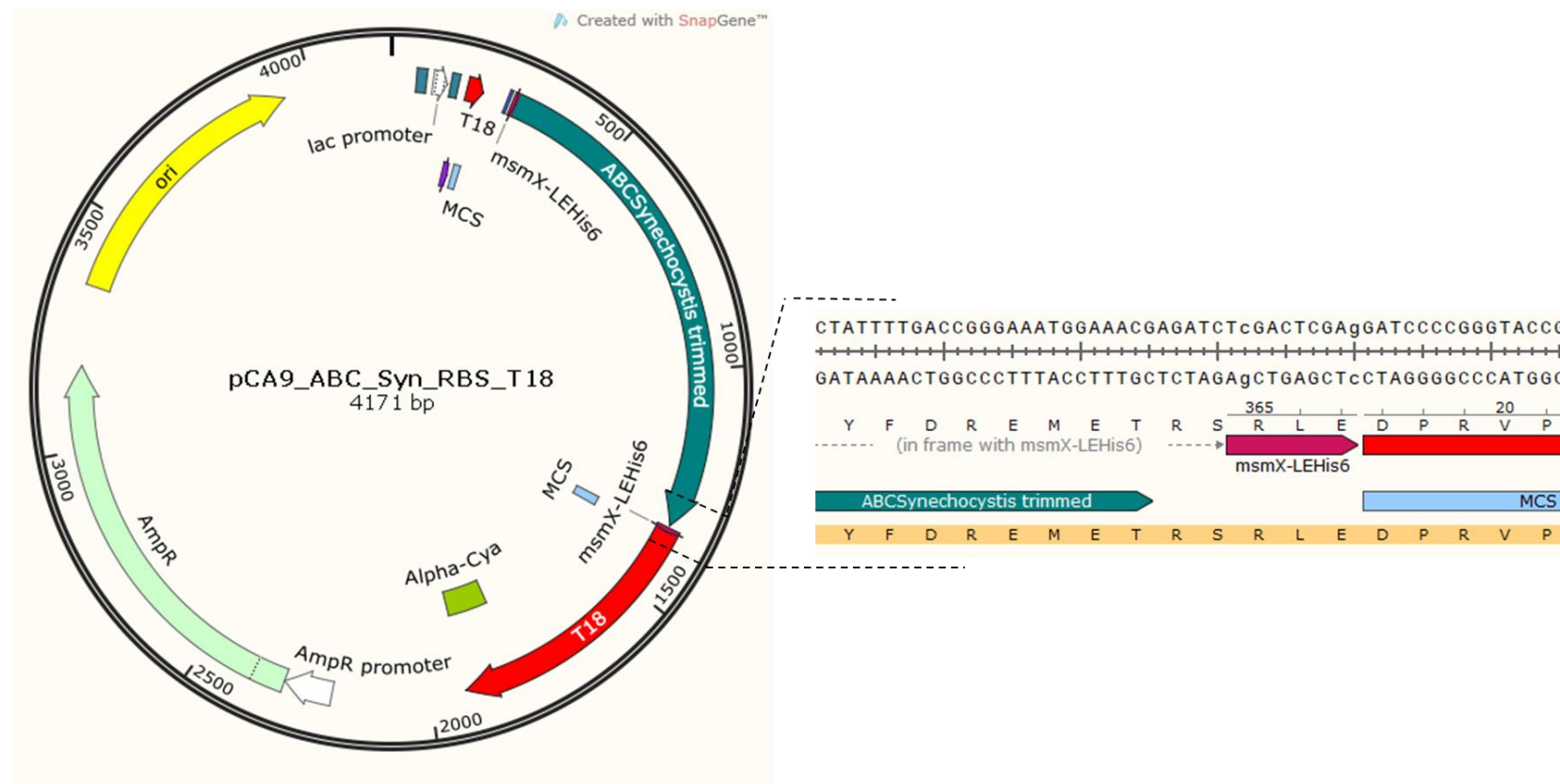
Appendix 5. Map of pCA7. Insertion of 1095 bp fragment containing *ugpC* gene from *S. aureus* (amplified with primer Ara993 and Ara994 and digested with XbaI) between XbaI restriction sites of pUT18. Primers M13rev and Ara958 were used for DNA sequencing. Detailed sequence of the fusion zone is shown in orange, UgpC amino acid sequence represented in green, T18 amino acid sequence represented in red and MCS zone is depicted in sky blue.



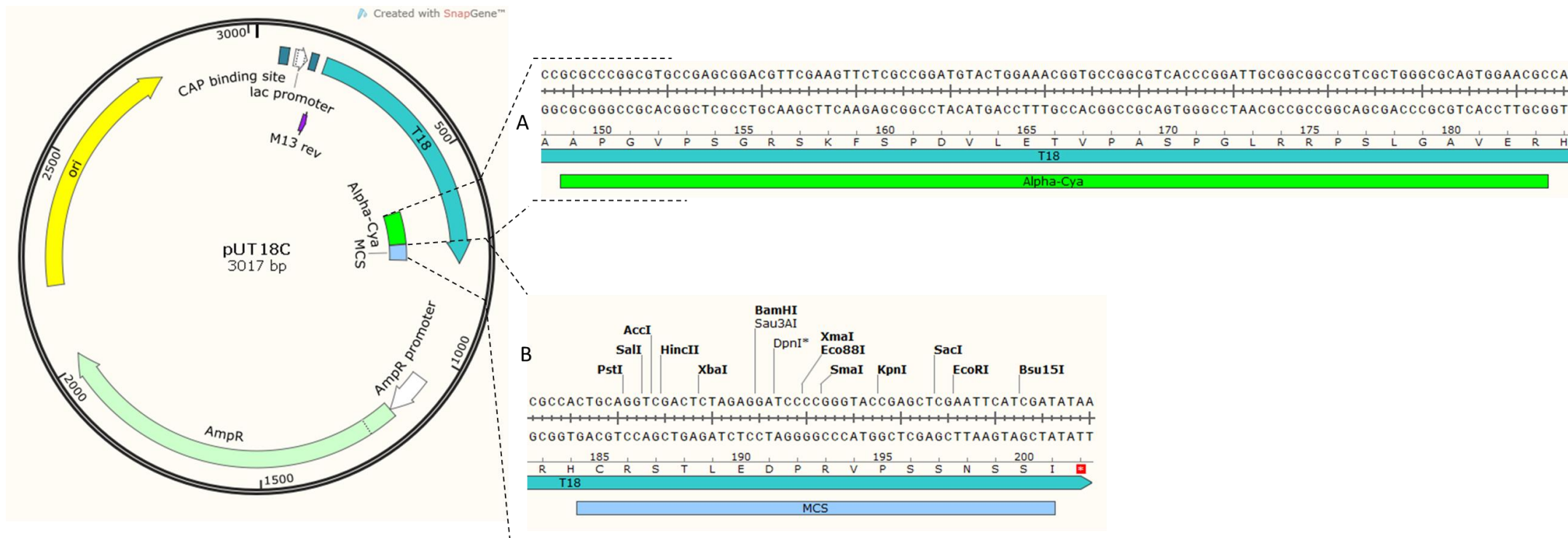
Appendix 6. Map of pCA8. Insertion of 1107 bp fragment containing *abc_cd* gene from *C. difficile* 630 (amplified with primer Ara995 and Ara996 and digested with XbaI) between XbaI restriction sites of pUT18. Primers M13rev and Ara958 were used for DNA sequencing. Detailed sequence of the fusion zone is shown in orange, ABC_Cd amino acid sequence represented in green, T18 amino acid sequence is represented in red and MCS zone is depicted in sky blue.



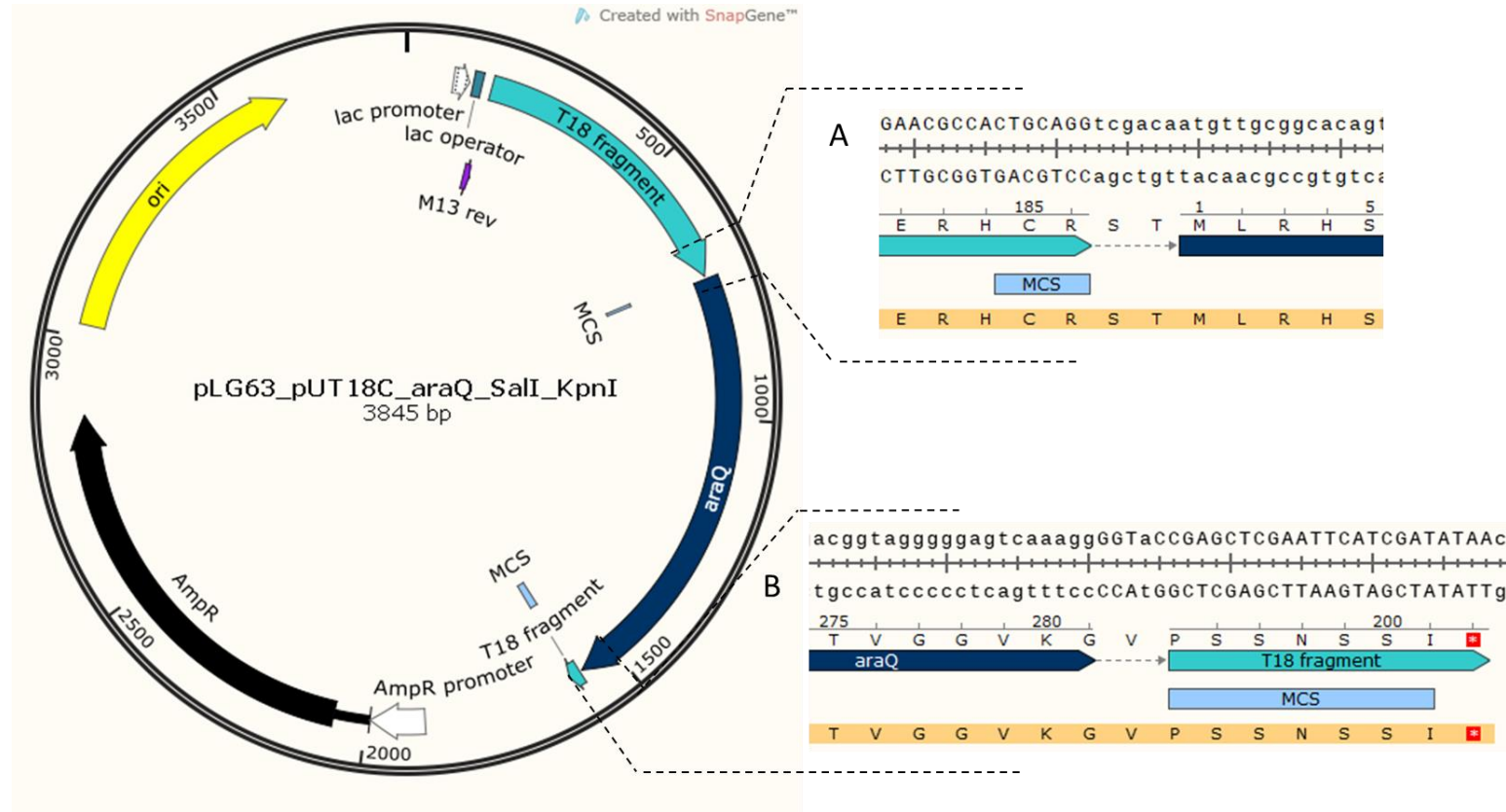
Appendix 7. Map of pCA9. Insertion of 1071 bp fragment containing *abc_syn* gene from *Synechocystis* sp. (amplified with primer Ara741 and Ara1017 and digested with Sall and BamHI) between Sall and BamHI restriction sites of pUT18. Primers M13rev and Ara958 were used for DNA sequencing. Detailed sequence of the fusion zone is shown in orange, ABC_Syn amino acid sequence is represented in blue, T18 amino acid sequence is represented in red and MCS zone is depicted in sky blue.



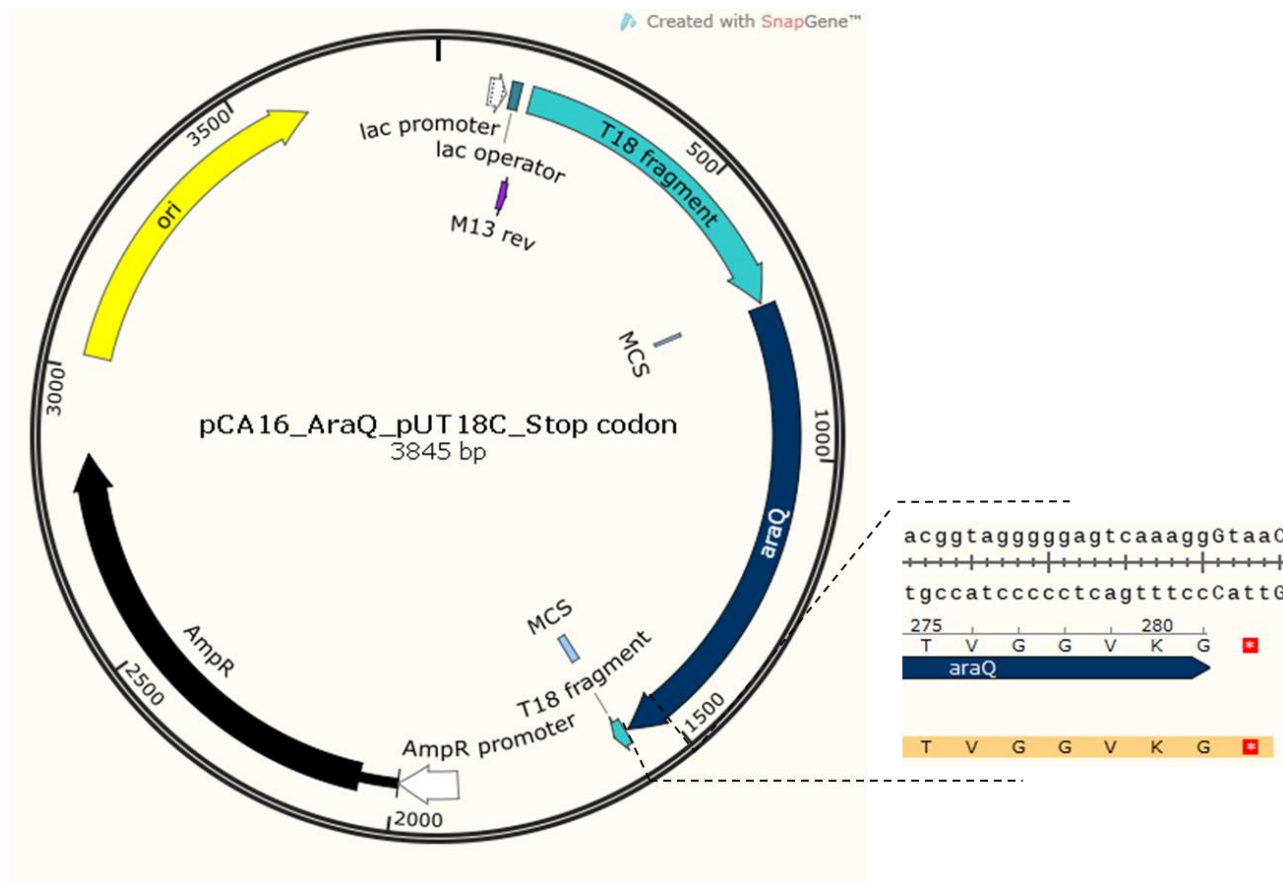
Appendix 8. Map of the high copy plasmid pUT18C, encoding for the T18 fragment (corresponding to the 225-400 amino acids) of Adenylate Cyclase from *Bordetella pertussis*. This plasmid expresses an ampicillin resistance selectable marker and is used to express chimeric proteins in which a heterologous polypeptide is fused to the C-terminal of T18. A) Detailed sequence of α -Cya antibody recognition site (green rectangle). T18 amino acid sequence is depicted in blue. B) Detailed sequence of the multiple cloning site (MCS - sky blue).



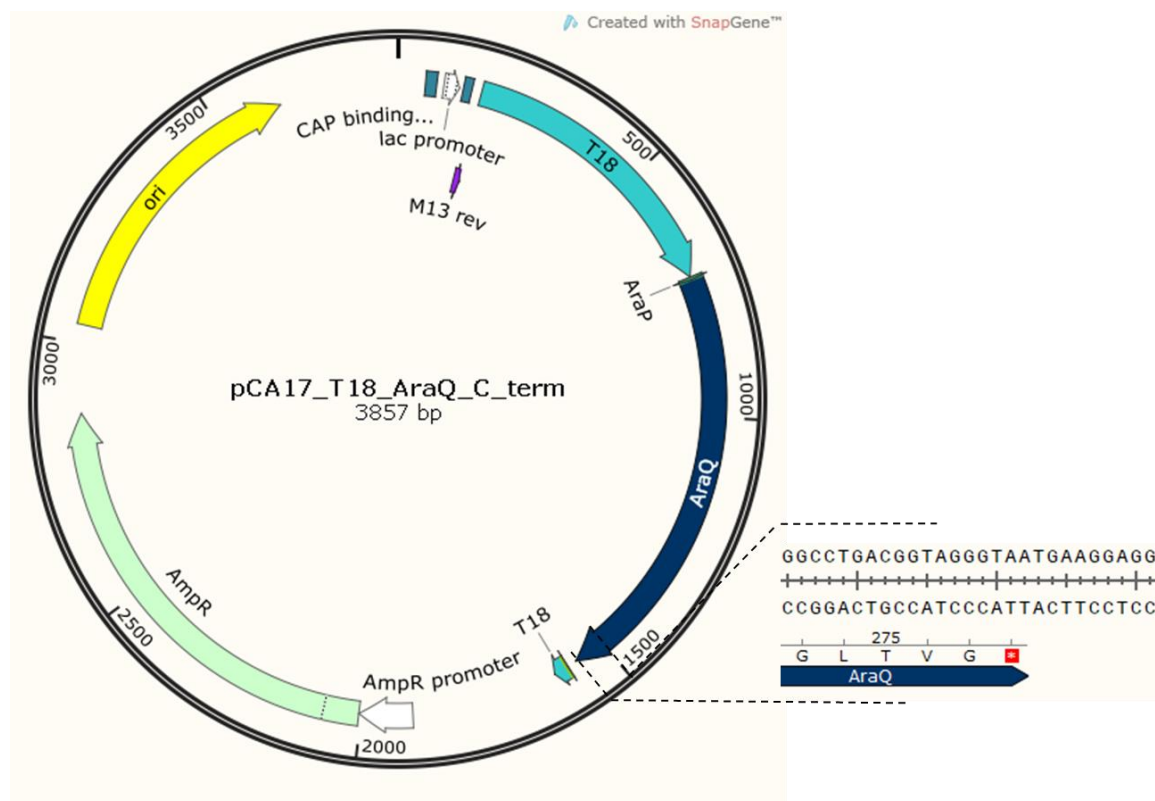
Appendix 9. Map of pLG63. Insertion of 842 bp fragment containing *araQ* gene from *B. subtilis* 168T⁺ (amplified with primer Ara950 and Ara945 and digested with SalI and KpnI) between SalI and KpnI restriction sites of pUT18C. A) Detailed sequence of the fusion zone is shown in orange, AraQ amino acid sequence represented in dark blue, T18 amino acid sequence represented in light blue and MCS zone is depicted in sky blue. B) Detailed sequence of the C-terminal region of AraQ fused to the T18 fragment. Codon stop is depicted as a red box with a white *.



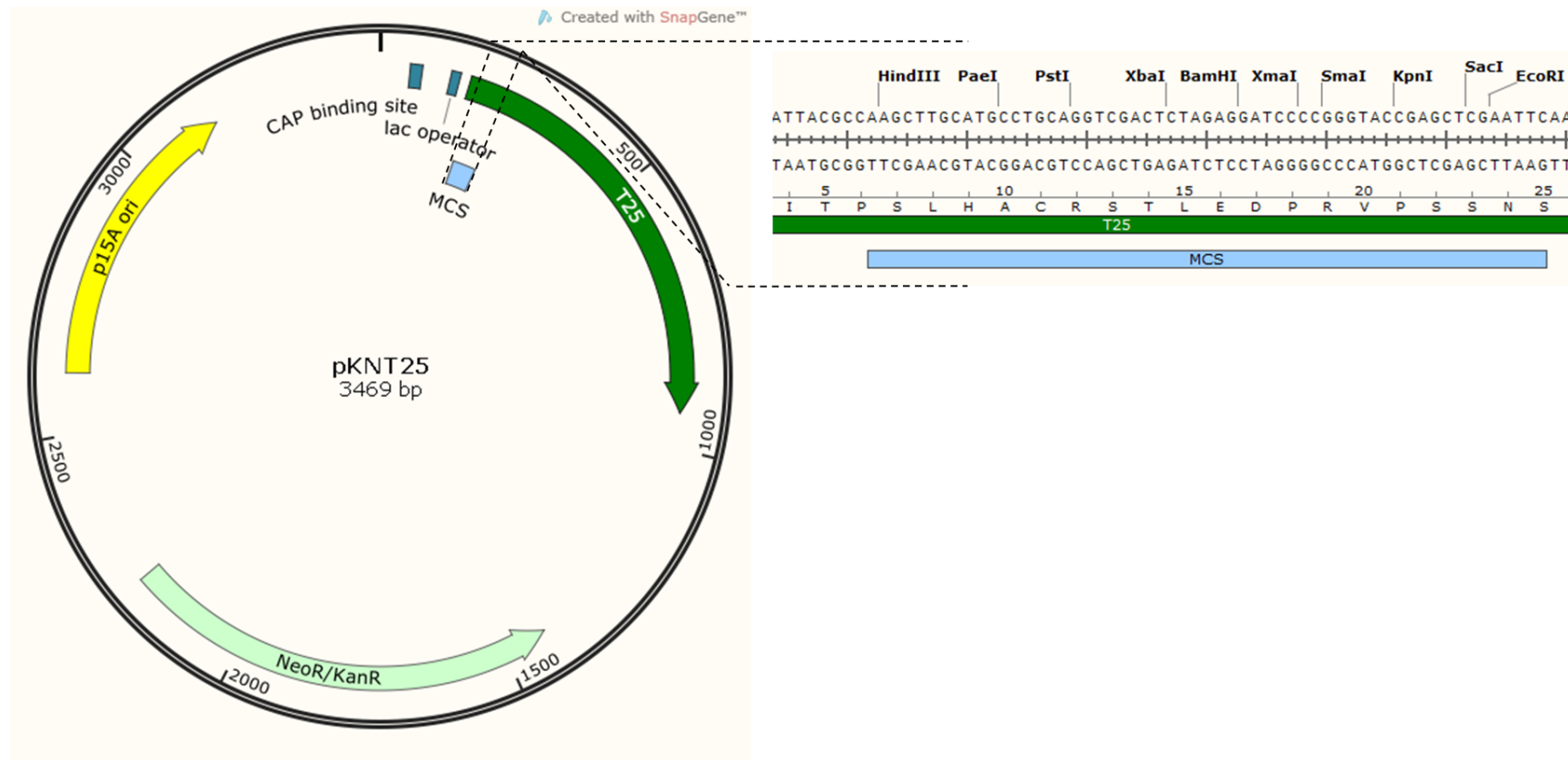
Appendix 10. Map of pCA16. Insertion of the native codon stop in pLG63 encoding for *araQ* gene from *B. subtilis* (using mutagenic primers Ara1014 and Ara1015). Primer Ara958 was used for DNA sequencing. Detailed sequence of the native C-terminal region of AraQ. Codon stop is depicted as a red box with a white *.



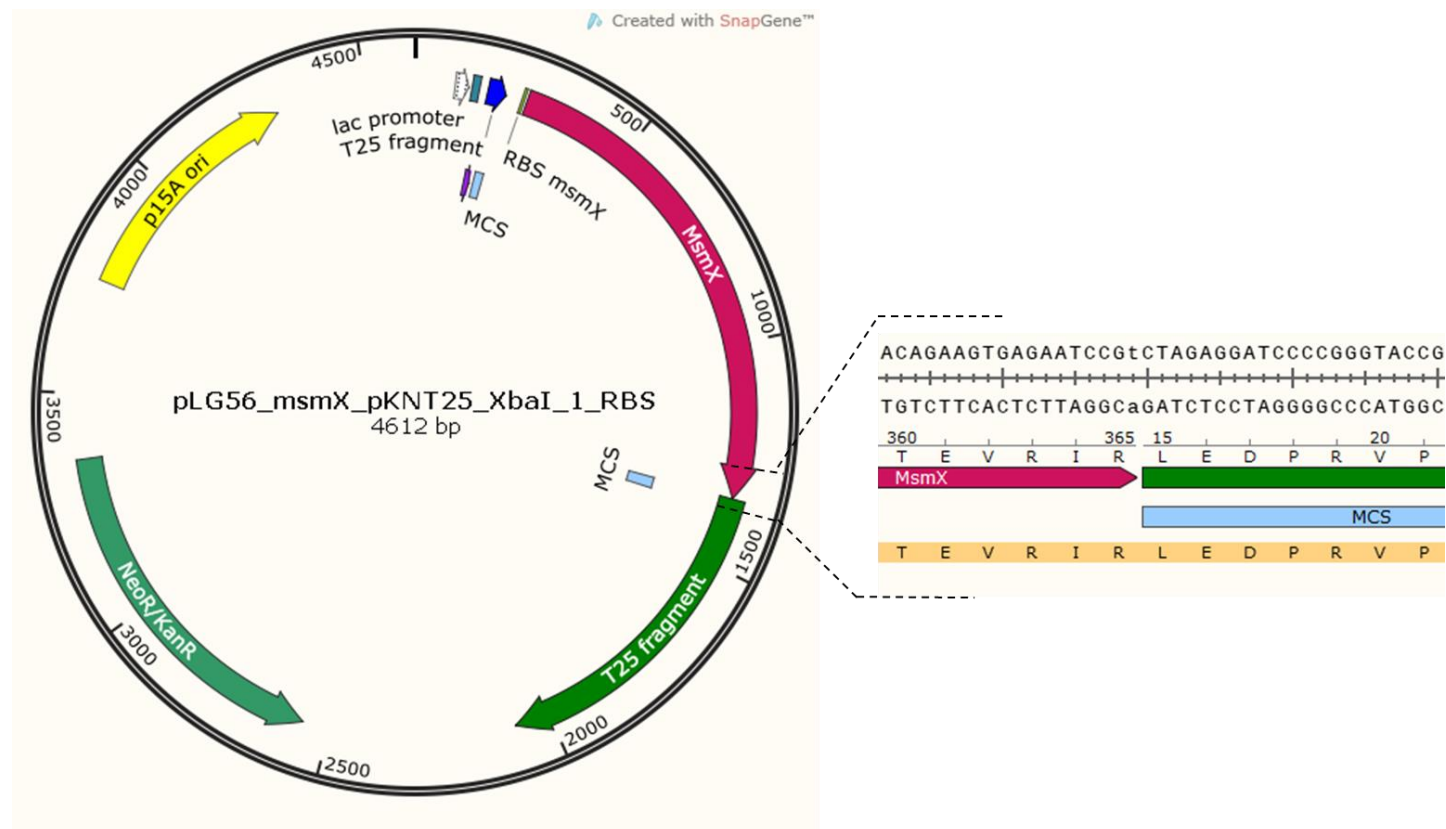
Appendix 11. Map of pCA17. Deletion of 12 bp of C-terminus of *araQ* gene from *B. subtilis* (using oligonucleotides Ara680, Ara674, Ara675, Ara676, Ara 951). Primers M13rev and Ara7 were used for DNA sequencing. Detailed sequence of the C-terminal region of AraQ is shown. Codon stop is depicted as a red box with a white *.



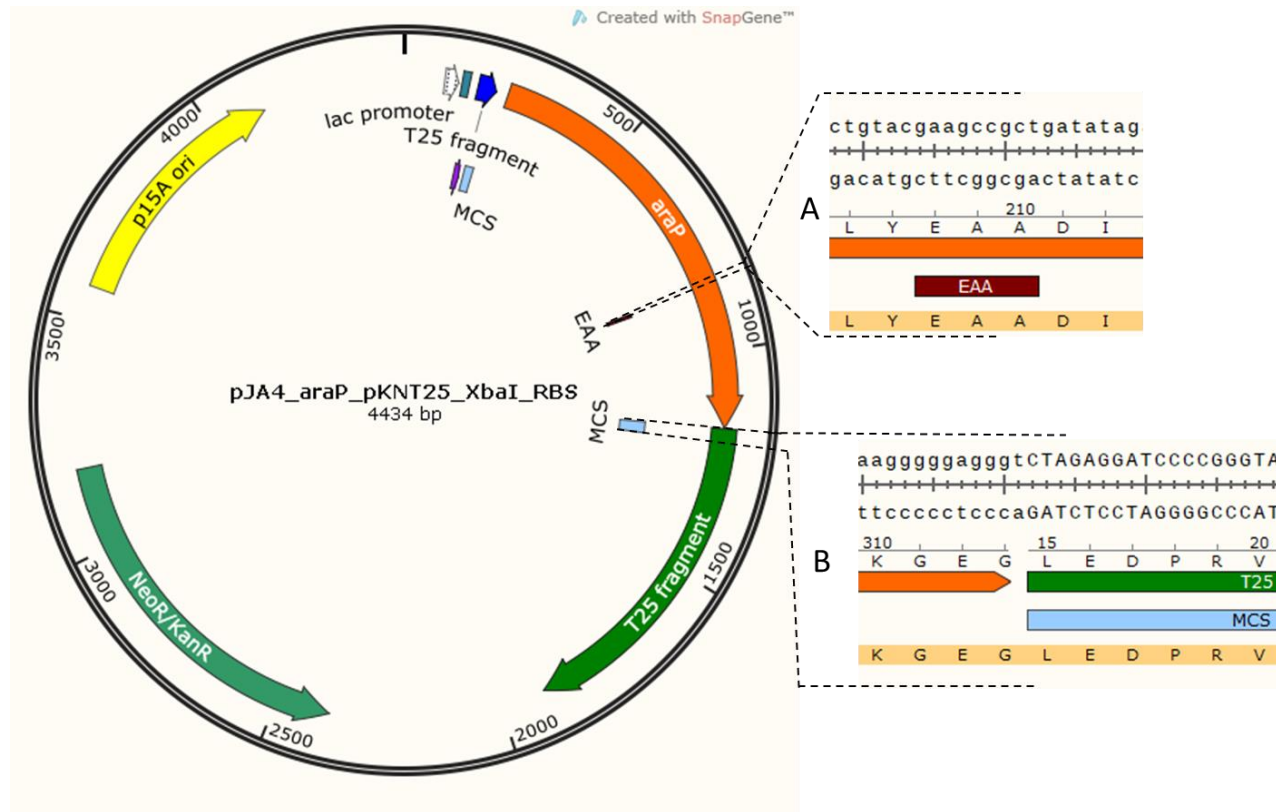
Appendix 12. Map of the low copy plasmid pKNT25, encoding for T25 fragment of Adenylate Cyclase from *Bordetella pertussis*. This plasmid expresses a kanamycin resistance selectable marker and is used to express chimeric proteins in which a heterologous polypeptide is fused to the N-terminal of T25. Detailed sequence of the multiple cloning site (MCS - sky blue) is shown. T25 amino acid sequence is depicted as a green box.



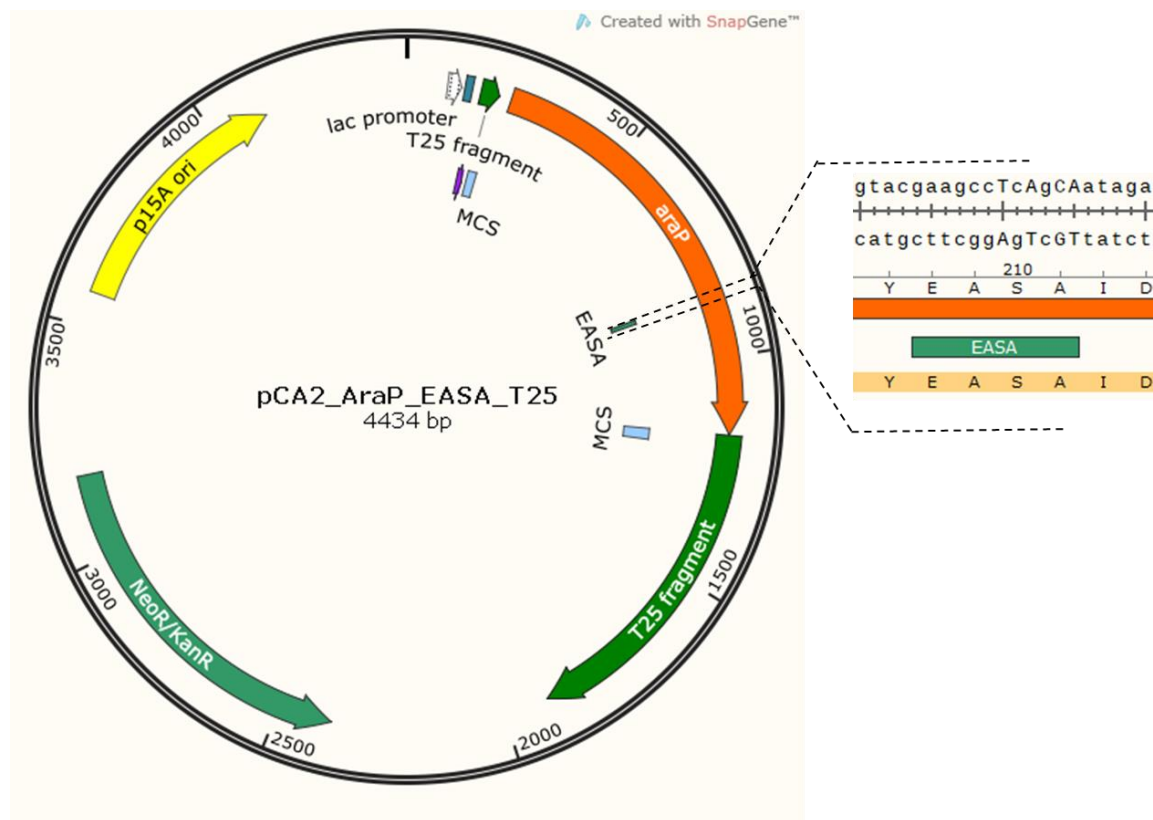
Appendix 13. Map of pLG56. Insertion of 1095 bp fragment containing *msmX* gene from *B. subtilis* 168T⁺ (amplified with primer Ara932 and Ara933 and digested with XbaI) between XbaI sites of pKNT25. Detailed sequence of the fusion zone is shown in orange, MsmX amino acid sequence is represented in pink, T25 amino acid sequence is represented in green and MCS zone depicted in sky blue.



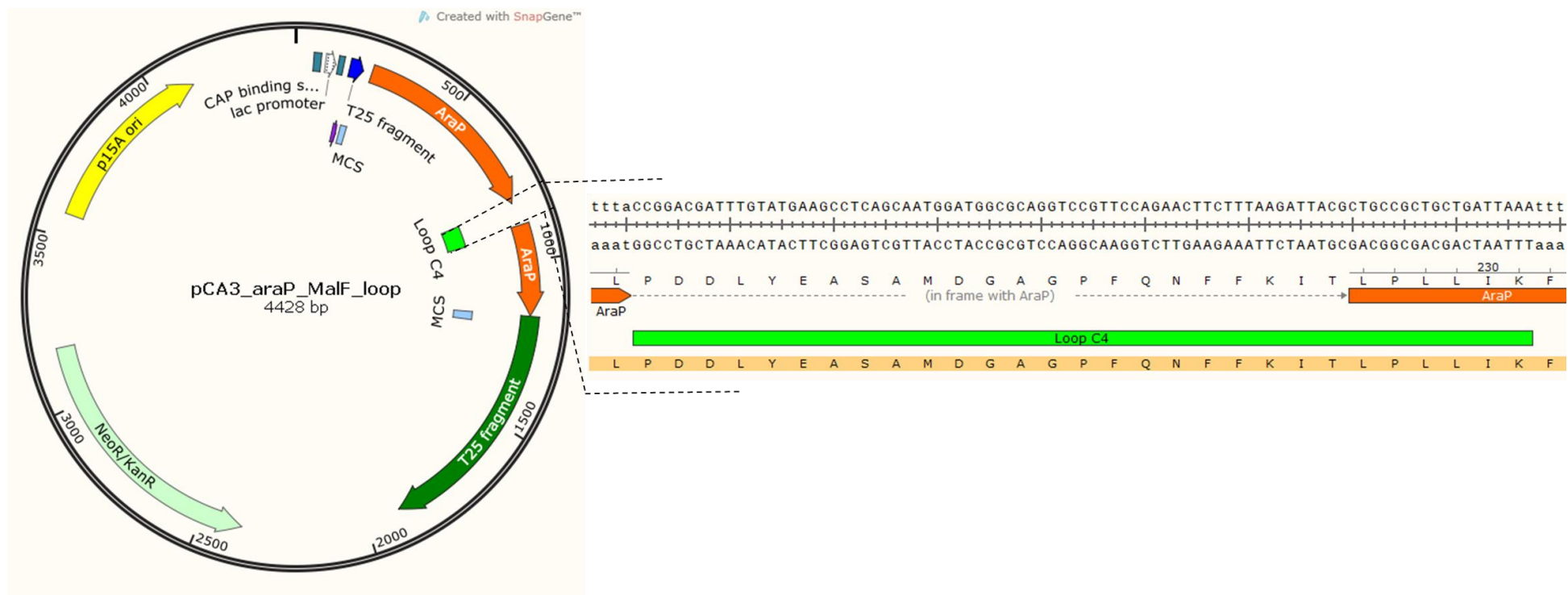
Appendix 14. Map of pJA4. Insertion of 843 bp fragment containing *araP* gene from *B. subtilis* 168T⁺ (amplified with primer Ara934 and Ara935 and digested with XbaI) between XbaI restriction sites of pKNT25. A) Detailed sequence of EAA motif. B) Detailed sequence of the fusion zone is shown in orange, AraP amino acid sequence represented in dark orange, T25 amino acid sequence represented in green and MCS zone depicted in sky blue.



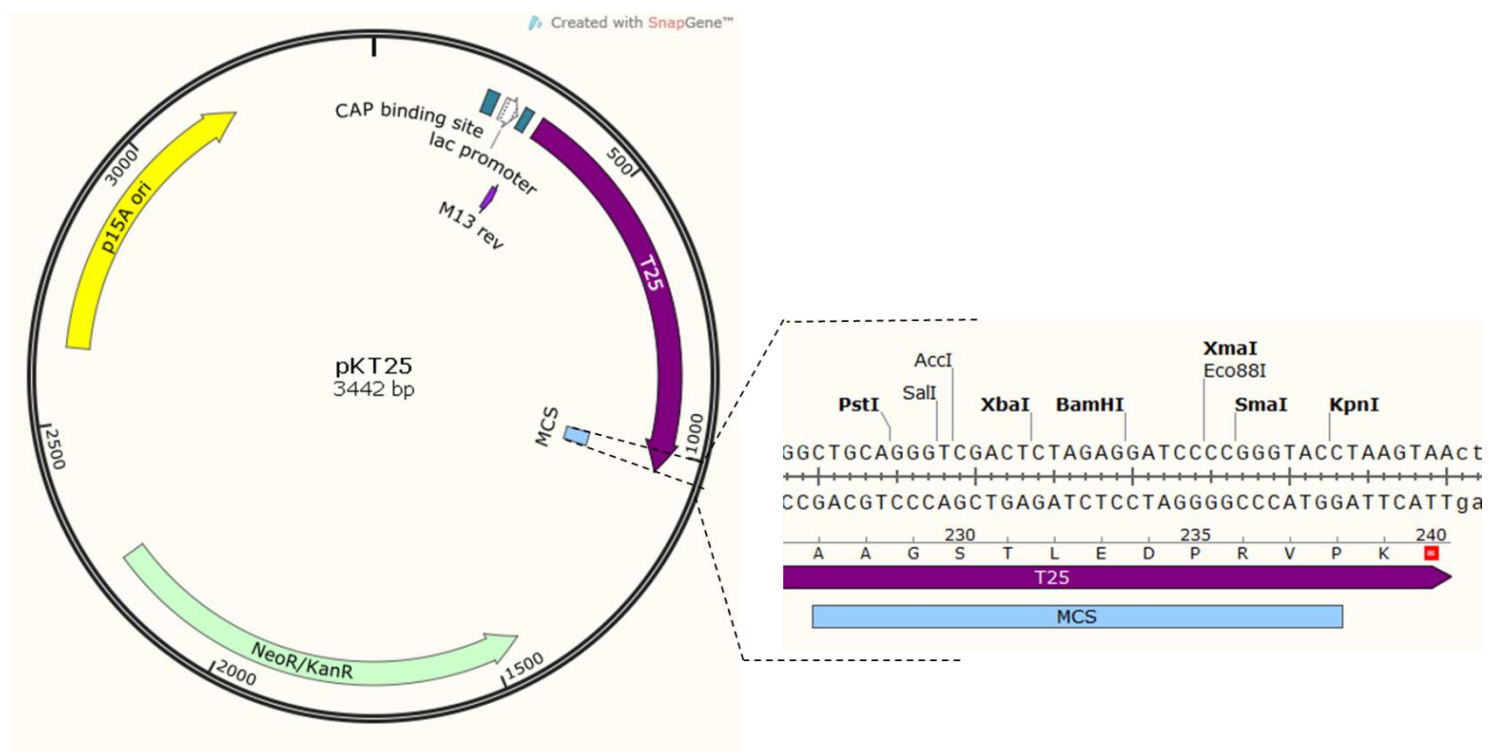
Appendix 15. Map of pCA2. Insertion of the mutation A210S/D211A by overlap PCR in *araP* gene from *B. subtilis* (mutagenic primers used were Ara975 and Ara976 and the external primers were Ara934 and Ara935). Primers M13rev and Ara958 were used for DNA sequencing. Detailed sequenced of mutations A210S/D211A in EAA motif (green). AraP amino acid sequence represented in dark orange.



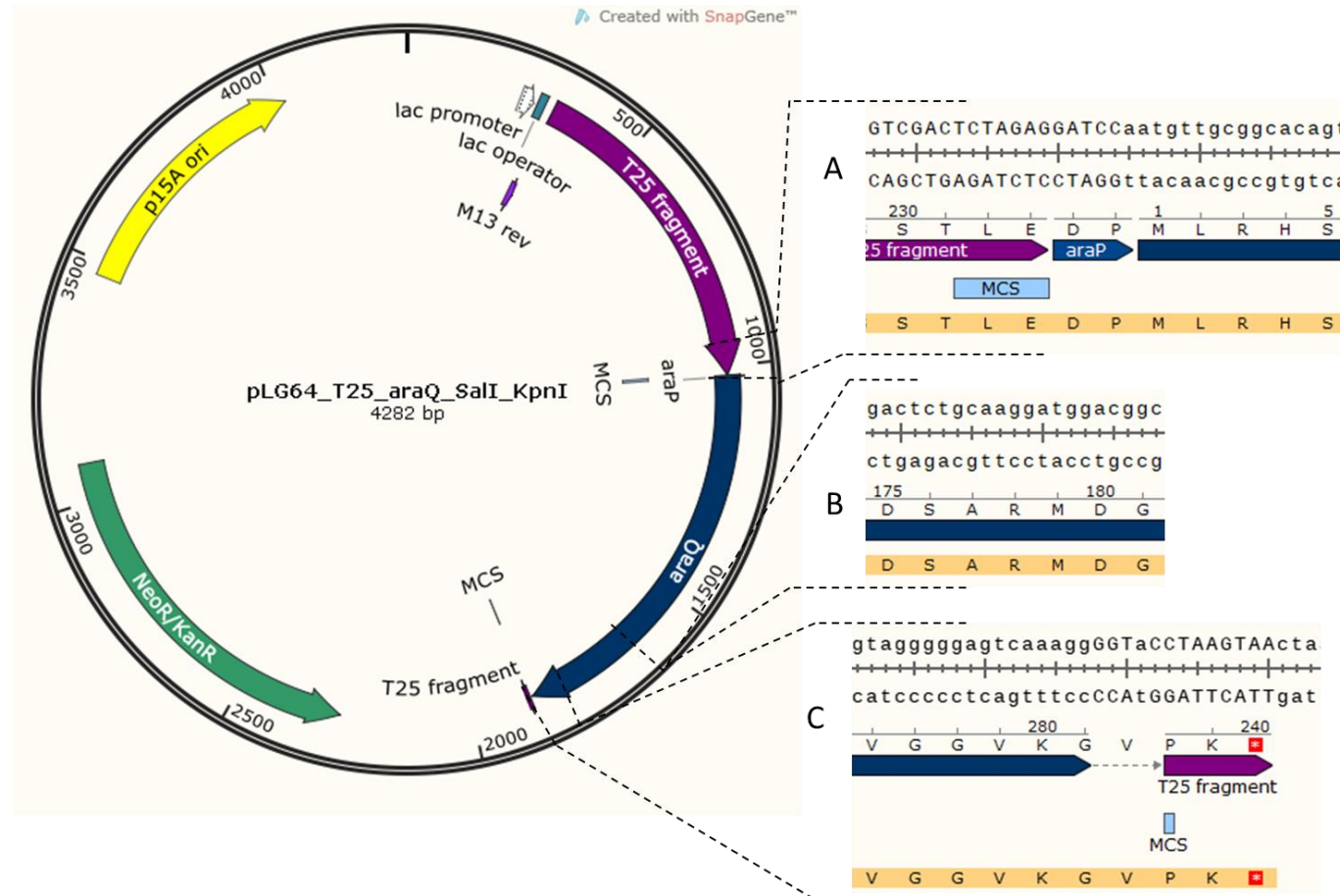
Appendix 16. Map of pCA3. Substitution of loop containing the EAA loop from *B. subtilis* by C3 loop from *E. coli* (using the primers Ara 934, Ara980, Ara979, Ara977, Ara 878 and Ara935) in *araP* gene from *B. subtilis*. Primers Ara958 and 166 were used for DNA sequencing. Detailed sequence of loop replacement (green). AraP amino acid sequence represented in dark orange.



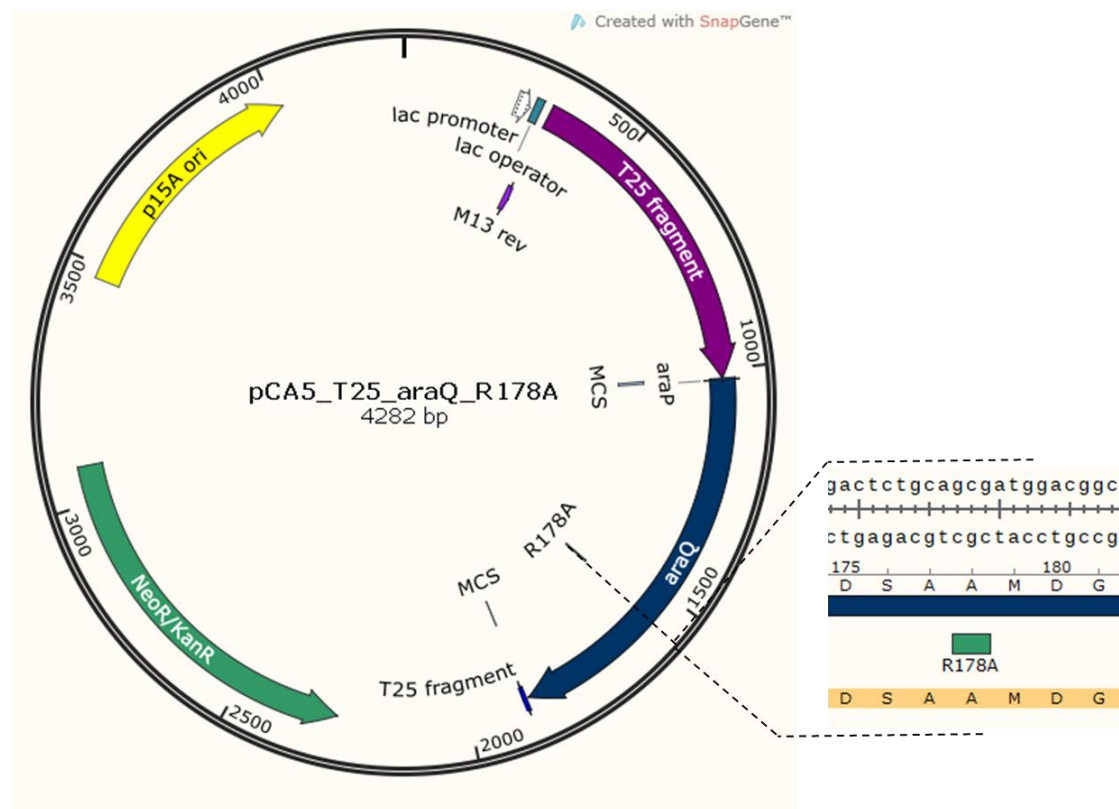
Appendix 17. Map of the low copy plasmid pKT25, encoding for the T25 fragment (corresponding to the first 224 amino acids) of Adenylate Cyclase from *B. pertussis*. This plasmid expresses a kanamycin resistance selectable marker and is used to express chimeric proteins in which a heterologous polypeptide is fused to the C-terminal of T25. Detailed sequence of the multiple cloning site (MCS - sky blue) is shown.



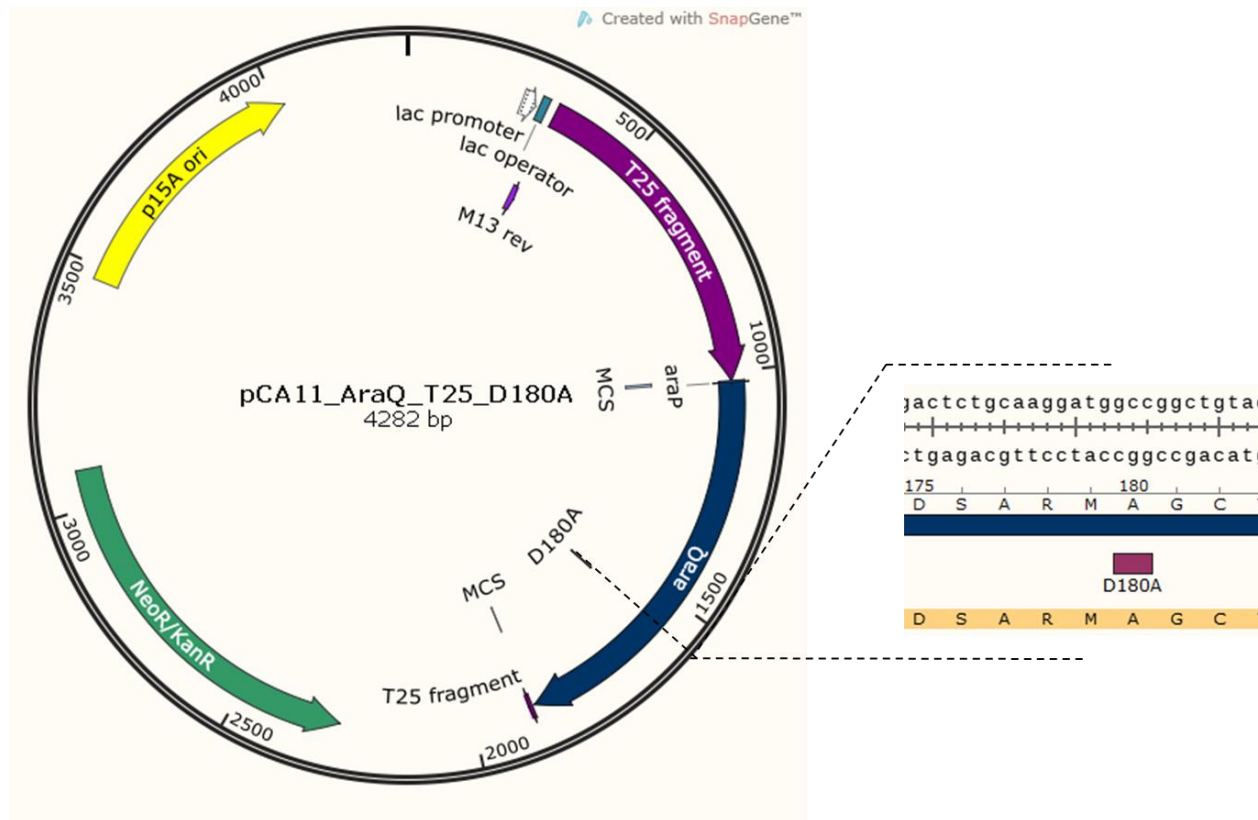
Appendix 18. Map of pLG64. Insertion of 842 bp fragment containing *araQ* gene from *B. subtilis* 168T⁺ (amplified with primer Ara951 and Ara945 and digested with BamHI and KpnI) between BamHI and KpnI restriction sites of pKNT25. A) Detailed sequence of the fusion zone is shown in orange, AraQ amino acid sequence represented in dark blue, T25 amino acid sequence represented in purple and MCS zone is depicted in sky blue. B) Detailed sequence of native amino acids targeted for mutagenesis. C) Detailed sequence of the C-terminal region of AraQ fused to the T25 fragment. Codon stop is depicted as a red box with a white *.



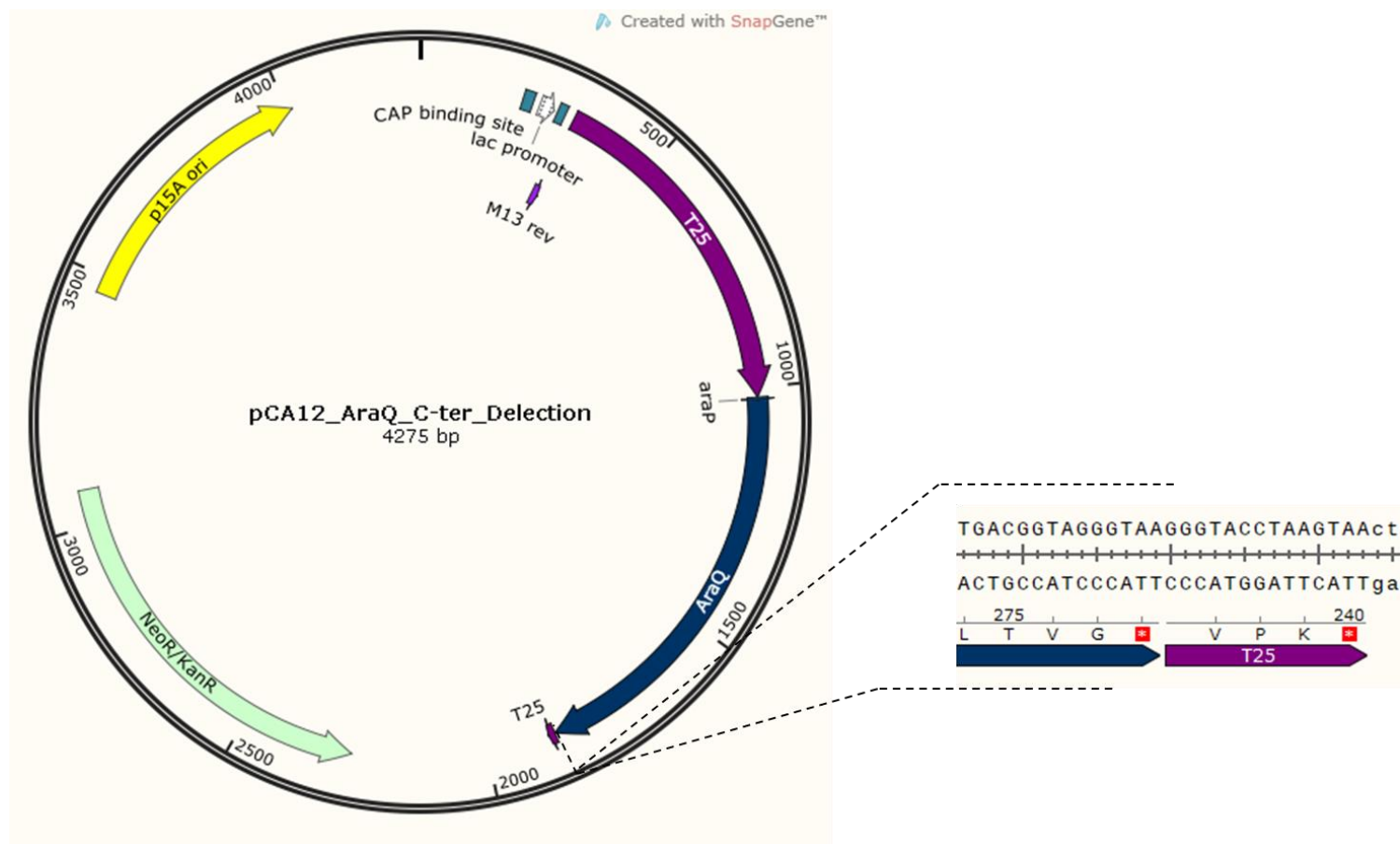
Appendix 19. Map of pCA5. Insertion of R178A mutation in *araQ* gene from *B. subtilis* (using mutagenic primers Ara987 and Ara988) in the pLG64 plasmid. Primers M13rev and Ara7 were used for DNA sequencing. Detailed sequence of the R178A mutation (green). AraQ is depicted in dark blue.



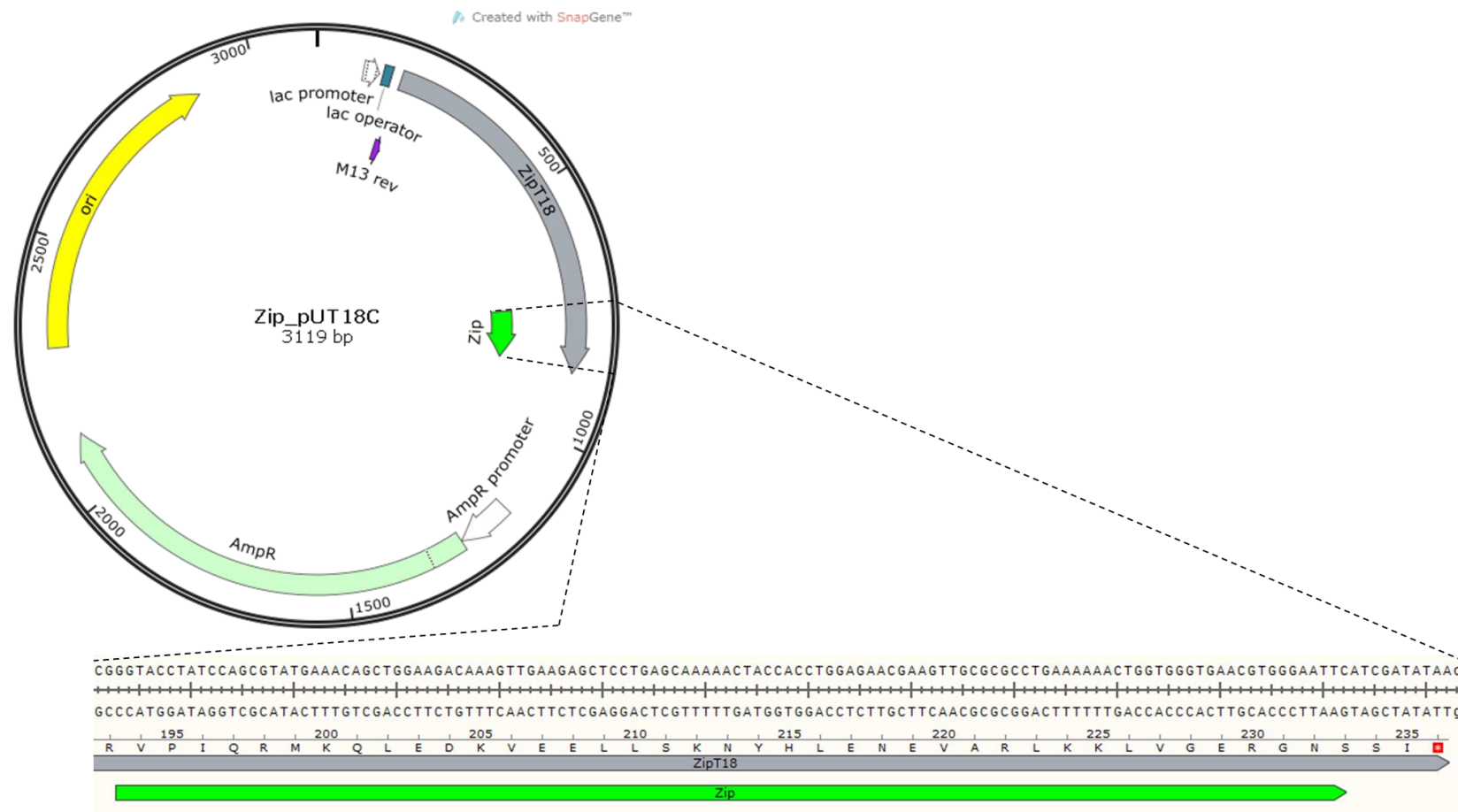
Appendix 20. Map of pCA11. Insertion of D180A mutation in pLG64 encoding for *araQ* gene from *B. subtilis* (using mutagenic primers Ara671 and Ara672). Primer Ara958 was used for DNA sequencing. Detailed sequence of the D180A mutation (pink). AraQ is depicted in dark blue.



Appendix 21. Map of pCA12. Deletion of 12 bp of C-terminus of *araQ* gene from *B. subtilis* (using oligonucleotides Ara680, Ara674, Ara675, Ara676, Ara 951). Primers M13rev and Ara7 were used for DNA sequencing. Detailed sequence of the C-terminal deletion of AraQ is shown. Codon stop is depicted as a red box with a white *.



Appendix 22. Map of Zip_pUT18C. Zinc finger (green) of CCN4 is fused to C-Terminus of T18 fragment of Adenylate Cyclase from *B. pertussis*. This plasmid expresses an ampicillin resistance selectable marker. Detailed sequence is shown.



Appendix 23. Map of Zip_pKT25. Zinc finger of CCN4 (green) is fused to C-Terminus of T25 fragment of Adenylate Cyclase from *B. pertussis*. This plasmid expresses a kanamycin resistance selectable marker. Detailed sequence is shown.

