UNIVERSIDADE NOVA DE LISBOA

Faculdade de Ciências e Tecnologia

Departamento de Química

Heavy-metal resistance in *Marinobacter aquaeolei* 617

Insights into copper resistance

Franklin Luzia de Nóbrega

Dissertação apresentada na Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa para obtenção do grau de Mestre em Biotecnologia

Orientador: Doutora Sofia Rocha Pauleta

Lisboa

2009

Aos meus Pais

ACKNOWLEDGEMENTS

Durante a realização da minha tese de mestrado, muitos foram os que de uma forma directa ou indirecta tiveram um papel importante. Gostaria de referir alguns:

Quero aqui deixar expresso o meu profundo agradecimento à minha orientadora, a Doutora Sofia Pauleta, por todo o conhecimento que me transmitiu, pelas inúmeras discussões científicas, bem como pela a sua amizade. Foi um ano fantástico, aprendi muito contigo e espero continuar a poder aprender.

À Doutora Maria Gabriela Almeida pelos ensinamentos transmitidos na Proteómica, pela abertura de horizontes numa área até então desconhecida, bem como pelas discussões e a sua pronta ajuda para melhorar o trabalho.

To Professor Bart Devreese from the University of Ghent, Mass Spectrometry Proteomics group (L-ProBE), for the mass analysis and helpful advices.

À Professora Isabel Moura e ao Professor José Moura por me terem aceite nos seus grupos de investigação, possibilitando-me fazer parte de um grupo extraordinário, bem como pelas sempre pertinentes pequenas grandes discussões científicas.

À Isabel Ribau, companheira de luta na "nautica"; à Célia Silveira, por toda a ajuda, por aturar a frustação na Proteómica, boa disposição e amizade; à Luisa Maia, pelos conselhos e excelentes conversas tardias; à Marta Carepo pela boa disposição e amizade; à Ana Teresa que muito foi o seu apoio, bem como prontidão na resposta a qualquer pedido; ao Pablo González pela sua ajuda com o EPR; e por último a ti Rui Duarte que nunca serás esquecido. Agradeço ainda a todo o grupo BioProt e BioIn, pela boa camaradagem durante a realização da minha tese.

Ainda, a todos aqueles que apesar de não terem estado relacionados directamente com o trabalho, o seu apoio e amizade tornou os momentos difíceis em pequenos obstáculos, são eles: Catarina Baptista, Jorge Dias, Inês Osório, amigos o meu Muito Obrigado por me aturarem. Finalmente, à Raquel Barros por ter sido um verdadeiro pilar durante todo o tempo.

Por fim, quero agradecer a minha família, especialmente aos meus Pais por sempre acreditarem em mim e terem possibilitado este sonho ser uma realidade.

ABSTRACT

Heavy metal resistance in *Marinobacter aquaeolei* (*Ma.aq*) 617 in aerobic conditions was studied for three different ions, cadmium, cobalt and copper. The main aim of this work was the study of a putative copper resistance operon, *copSRXAB*, located in the chromosome of Marinobacter and the biochemical characterization of a unique copper binding protein CopX (proposed designation), associated with the copper resistance system.

Growth under heavy metal ion stress was performed for those three heavy metals and the Minimum Inhibitory Concentration (MIC) / Maximum Tolerant Concentration (MTC) was determined using two different approaches, solid artificial sea water (ASW) plates and liquid ASW medium, supplemented with lactate and yeast extract, as carbon sources. The MIC/MTC of cadmium, cobalt and copper ions was found to be 200 μM, 4-6 mM, and 1.6 mM, respectively. These values classify *Ma.aq* strain 617 as cadmium, cobalt and copper resistant strain. Moreover, during the cobalt resistance studies we observed the production of an unknown protein or compound, which is proposed be a cobalamine containing protein and/or cobalamine itself.

Under the scope of copper resistance, preliminary proteomics analysis of the *Ma.aq* periplasmic fraction was performed. CopX, identified by MALDI TOF-TOF mass spectrometry, was shown to be differentially expressed under copper stress. This demonstrated that the proposed copper operon, *copSRXAB*, has a role in the *Ma.aq* copper resistance.

CopX was successfully heterologously expressed in *Escherichia coli* (*Es.coli*), and purified for the first time using usually two chromatographic steps (anionic exchangeand size exclusion) with a yield of 5.7 mg or 1.8 mg of purified CopX, per L of LB or M9 medium, respectively. Mass spectrometry Electron Spray Ionisation (ESI) and N-terminus analysis revealed that the signal peptide of CopX comprises 21 residues, and is efficiently processed by the Sec system of *Es.coli*.

Biochemical characterization of CopX proved that it is a periplasmic monomeric type 1 copper protein, with a molecular weight of 17253.25 ± 0.30 Da, determined by mass spectrometry (ESI), that binds approximately 1 copper ion per polypeptide chain. The apparent molecular weight of CopX, 20.4 kDa, determined by size-exclusion chromatography does not depend on the ionic strength.

Spectroscopic characterization showed that it presents an intense charge transfer (S_{cys} – Cu ion) band at 440 nm and 580 nm and 720 nm. The extinction coefficient at 580 nm was found to be 3.8 $mM^{-1}cm^{-1}$, according to the copper content. CopX EPR spectrum is axial. The ^{15}N HSQC NMR spectra of CopX confirms that it is folded, with 131 out of 147 backbone amide resonances identified, showing that it is amenable to NMR solution structure determination.

CopX presents some unique features, such as, a ratio between A_{440nm} and A_{580nm} of 0.94 and a high hyperfine coupling constant, 170 G. Taking into account the biochemical properties, CopX is proposed to be part of a new class of the type 1 copper proteins, shown preliminarily for the first time to be associated with copper resistance.

RESUMO

Neste trabalho, efectuou-se o estudo de resistência a três metais pesados, cádmio, cobalto e cobre em *Marinobacter aquaeolei* (*Ma.aq*) 617 em condições aeróbias. O principal objectivo deste trabalho foi o estudo de um operão, localizado no cromossoma, putativamente envolvido na resistência ao cobre, *copSRXAB,* e o estudo bioquímico de uma proteina deste sistema que liga cobre, CopX (designação proposta).

Dois métodos distintos foram utilizados para efectuar os crescimentos em elevadas concentrações de metais pesados, crescimentos em placa (meio sólido) e em meio líquido. O meio de crescimento utilizado foi a água do mar sintética, suplementada com lactato e extracto de levedura como fontes de carbono. As concentrações minimas inibitórias (MIC) / concentrações máximas toleradas (MTC) obtidas para o ião cádmio, cobalto e cobre foram 200 μM, 4-6 mM e 1.6 mM, respectivamente. Estes dados permitiram classificar a bactéria *Ma.aq* estirpe 617 como resistente ao cádmio, cobalto e cobre. Crescimentos na presença de elevadas concentrações de $CoCl_2$ revelaram a produção de uma proteína desconhecida ou composto, que se pensa ser uma proteína contendo cobalamina e/ou cobalamina.

Em relação à resistência ao cobre, a análise do proteoma foi efectuada por técnicas de electroforese 2D, permitindo a identificação preliminar da CopX por espectroscopia de massa MALDI-TOF-TOF. Foi ainda possível observar que esta proteína só é expressa em elevadas concentrações de cobre, indicando preliminarmente que a transcrição do operão proposto de resistência ao cobre, *copSRXAB*, está envolvido no mecanismo de resistência.

CopX foi expressa heterologamente em *Escherichia coli* e purificada pela primeira vez, usando normalmente dois passos cromatográficos (coluna de troca aniónica e uma coluna de exclusão molecular), obtendo-se 5.8 mg e 1.8 mg de CopX por L de meio LB ou meio M9, respectivamente.

A caracterização bioquímica da CopX demonstrou que esta é uma proteína periplasmática de cobre tipo 1, com uma massa molecular de 17253.25 ± 0.30 Da, determinada por espectrometria de massa – ionização por electrão (IES), contendo aproximadamente um cobre por cadeia polipeptídica. A massa molecular aparente, determinada por cromatografia de exclusão molecular, é de 20.4 kDa não sendo dependente da força iónica.

A caracterização espectroscópica da CopX mostrou que o seu espectro de UV-visível apresenta uma intensa banda de transferência de carga (S_{cys} – ião cobre) a 440 nm, 580 nm e 720 nm. O coeficiente de extinção molar determinado a 580 nm, por cobre, é de 3.8 $mM^{-1}cm^{-1}$. CopX apresenta um espectro de EPR do tipo axial. A análise preliminar do espectro ^{15}N HSQC, confirma o enrolamento da proteína, tendo sido identificadas 131 das 147 ressonâncias das amida da ligação peptídica. Este dado confirma a possível determinação de estrutura por esta técnica.

A proteína apresenta características únicas, como a razão entre A_{440nm} e A_{580nm} de 0.94 e a constante hiperfina de acoplamento de 170 G. Tendo em conta as suas propriedades bioquímicas, é proposto que a proteína CopX faça parte de uma nova classe de proteínas de cobre tipo 1. Pela primeira vez, de uma forma preliminar, uma proteína deste tipo é associada a um mecanismo de resistência ao cobre.

ABBREVIATIONS

Ø: Diameter

µ: Specific growth rate

Abs: Absorbance

Amp: Ampicillin

APS: Ammonia Persulphate

Asc: Ascorbate

ASW: Artificial Sea Water

ATP: Adenosine triphosphate

BCA: Bicinchoninic acid

BLAST: Basic Local Alignment Search Tool

bp: Base pair

BSA: Bovine Serum Albumin

CARA: Computer Aided Resonance Assignment

cbp: copper-binding protein

cDNA: Complementary DNA

cp: copper protein

DNA: Deoxyribonucleic acid

EPR: Electron Paramagnetic Resonance

ESI: Electron Spray ionisation

GL: Glass

HK: Histidine kinase

HSQC: Heteronuclear Single Quantum Coherence NMR experiment

ICP: Inductively couple plasma

IEF: Isoelectric Focusing

IPTG: Isopropyl β-D-thiogalactopyranoside

Kav: Gel-phase distribution coefficient

L: Liter

LB: Luria-Bertani broth

LMW: Low Molecular Weight

mAU: mili Absorbance Unit

MBD: Metal Binding Domain

MIC: Minimum Inhibitory Concentration

min: minutes

MEGA: Molecular Evolutionary Genetic Analysis

MTC: Maximum Tolerant Concentration

MW: Molecular Weight

NCBI: National Center for Biotechnology Information

NMR: Nuclear Magnetic Resonance

No: Number

nr: Non-redundant protein sequence database

OD: Optical Density

ORF: Open Reading Frame

P: Phosphate

PA: Alkaline phosphatase

PAGE: Polyacrylamide gel electrophoresis

PCR: Polymerase Chain Reaction

PHYLIP: Phylogeny Inference Package

Pi: Sodium Phosphate

pI: Isoelectric point

pmf: Peptide mass fingerprinting

REC: Signal Receiver domain

rpm: Rotation per minute

RT: Room Temperature

Sec pathway: Secretory pathway

SDS: Sodium dodecyl sulphate

SH: Sulfhydryl or Thiol group

TAE: Tris-acetate-EDTA buffer

td: mean generation time or doubling time

TEMED: N - N - N'- N' – tetramethyl - 1,2 diaminemetane

Tm: Melting temperature

UV: Ultraviolet

Vo: Void volumn

Amino acid one- and three-letter code

A or Ala: Alanine

C or Cys: Cysteine

D or Asp: Aspartic acid

E or Glu: Glutamic acid

F or Phe: Phenylalanine

G or Gly: Glycine

H or His: Histidine

I or Ile: Isoleucine

K or Lys: Lysine

L or Leu: Leucine

M or Met: Methionine

N or Asn: Asparagine

P or Pro: Proline

R or Arg: Arginine

Q or Gln: Glutamine

S or Ser: Serine

T or Thr: Threonine

V or Val: Valine

W or Trp: Tryptophan

Y or Tyr: Tyrosine

Heavy metal resistance systems, mechanisms and transporters

Cad: Cadmium resistance system

CBA: Three polypeptide chemiosmotic antiporter

CDF: Cation Diffusion Facilitator

Cnr: Cobalt and Nickel resistance system

Cop: Copper operon resistance system

Coe: Copper efflux regulon (resistance mechanism)

Cus: Copper sensing locus (resistance mechanism)

Czc: Cadmium, Zinc and Cobalt resistance system

Czr: Cadmium and Zinc resistance mechanism

HoxN: High-affinity nickel transport

MFP: Major Facilitator Protein

MFS: Major Facilitator Superfamily

MIT: Metal inorganic transport

Ncc: Nickel, Cobalt and Cadmium resistance mechanism

Pco: Plasmid copper resistance system

RND: Resistance-Nodulation-Cell Division

TAT-pathway: Twin-arginine translocation pathway

yohLMN: Cobalt and Nickel resistance system

Microorganisms

sp.: Species

str.: Strain

Al. borkumensis or *Al. b.*: *Alcanivorax borkumensis*

Ac.ferrooxidans: *Acidithiobacillus ferrooxidans*

Ag.tumefaciens: *Agrobacterium tumefaciens*

Ba.licheniformis: *Bacillus licheniformis*

Br.melitensis: *Brucella melitensis*

Br.suis: *Brucella suis*

Es. coli: *Escherichia coli*

En. hirae: *Enterococcus hirae*

H.influenzae: *Haemophilus influenzae*

K.pneumoniae: *Klebsiella pneumoniae*

M.: Marinobacter

Ma.alg.: *Marinobacter algicola*

Ma.aq: *Marinobacter aquaeolei*

My.bovis: *Mycobacterium bovis*

Methanotroph: Methane oxidizing bacterium

Me.loti: *Mesorhizobium loti*

O. Antarcticus: *Octadecabacter antarcticus*

Ps.aeruginosa or *Ps.a*: *Pseudomonas aeruginosa*

Ps.alcaligenes: *Pseudomonas alcaligenes*

Ps.c: *Pseudomonas chlororaphis*

Ps. fluorescens: *Pseudomonas fluorescens*

Psd.haloplanktis: *Pseudoalteromonas haloplanktis*

Ps.mendocina: *Pseudomonas mendocina*

Prot.mirabilis: *Proteus mirabilis*

Pa.multocida: *Pasteurella multocida*

Ps. nautica: *Pseudomonas nautica*

Ps. putida or *Ps.p*: *Pseudomonas putida*

Prov.rettgeri: *Providencia rettgeri*

Pseudomonad: Any of the gram-negative bacterium of the genus *Pseudomonas*

Ps. syringae or *Ps.sy: Pseudomonas syringae*

Ps.stutzeri: *Pseudomonas stutzeri*

Ra. metallidurans: *Ralstonia metallidurans*

Ro. nubinhibens: *Roseovarius nubinhibens*

Rh.leguminosarum: *Rhizobium leguminosarum*

Ra.solanacearum: Ralstonia solanacearum

Strec.agalactiae: *Streptococcus agalactiae*

Sta.aureus: *Staphylococcus aureus*

Stret.coelicolor: *Streptomyces coelicolor*

Sy.elongatus: Synechococcus elongatus

Si. medicae: *Sinorhizobium medicae*

Si.meliloti: *Sinorhizobium meliloti*

Sh.oneidensis: *Shewanella oneidensis*

Sa.typhi: *Salmonella typhi*

Sa.typhimurium: *Salmonella typhimurium*

Sh.woodyi: *Shewanella woodyi*

V.cholerae: *Vibrio cholerae*

V.parahaemolyticus: *Vibrio parahaemolyticus*

V.vulnificus: *Vibrio vulnificus*

Y.pestis: *Yersinia pestis*

TABLE OF CONTENTS

CHAPTER 1 - INTRODUCTION

CHAPTER 2 - MATERIAL AND METHODS

CHAPTER 3 – RESULTS AND DISCUSSION

FIGURE INDEX

APPENDIX FIGURE INDEX

TABLE INDEX

APPENDIX TABLE INDEX

INTRODUCTION

1. General Introduction

It was previously thought that resistance to heavy metals appeared as a result of the recent year's pollution. However, nowadays it is believed that it has been present since the beginning of life, in an already metal-polluted world [1]. This idea is in line with the antibiotic resistance determinants that are preexistent to recent human antibiotic handling [2, 3]. Although this is a clear conclusion for the antibiotic resistance, it requires caution when referring to heavy metal resistance since it lacks experimental validation [1].

The heavy metals are 53 out of the 90 naturally occurring elements. This designation is determined by the density value above 5 g/cm^3. Almost, for each and every inorganic cation or anion that is found in natural environments, organisms have mechanisms to regulate their metabolic movement [3]. This specific complex homeostasis mechanisms control their acquisition from the environment. Inside the cell, they control the sequestration of the metal ions (that cannot be free inside the cell), by controlling the expression of metal chaperones or other specific proteins that either store or expell the excess of metal ions (heavy metals or metalloids) [4].

Since the origin of life, microorganisms have been exposed to heavy metal cations and oxyanions, and the fact that these metal ions are able to form complex compounds (redox-active or not) makes them essential trace elements, a feature that is related with the unfilled *d* orbital of most heavy metals [3, 5].

Nevertheless, at higher concentrations most heavy metals are toxic. The toxicity is related with the formation of unspecific complex compounds inside the cell that has toxic effects. This can turn highly required trace elements into strong toxic complexes, which makes them too dangerous for any physiological function. Therefore, the intracellular concentration of heavy metals has to be tightly controlled [5].

Heavy metal resistance is just a specific component of the heavy metal homeostatic system in each living cell. Throughout the years, several studies have discovered specific genes and systems for transport of essential metals and for resistance to toxic heavy metal ions [2, 3, 6].

The study of these systems is important because the insights into the interactions between microorganisms and metals will be useful to understand the relation of toxic metals with higher organisms, such as plants and mammals [7]. This knowledge has a vast application, such as, in the study of copper transport diseases, like the Menkens and Wilson's diseases. These diseases are due to mutations in ATPases pumps, ATP7A in the case of Menkes disease [8] and ATP7B in the case of Wilson disease [9], and these diseases are related with copper starvation and copper excess, respectively. To strength this view, it was found that those proteins are homologous to bacterial copper transport proteins. Therefore, the understanding the bacterial versions of these ATPases pumps will contribute to a better understanding of those diseases and possibly help finding a "cure" for them [10, 11].

Another example of an application is biotechnology. There are three main areas where this could be achieved. First, adding metal resistance to a microorganism may facilitate a biotechnological process, which may or may not be linked to heavy metals. Second, heavy metal resistant bacteria may be used for bio-mining, either directly or by recovering from effluents of industrial process. Third, use of heavy metal resistant bacteria to bioremediation of metal-contaminated environments [5].

2. Biological important heavy metals

Although the definition of ***heavy metal*** is not consensus, there are 53 elements considered as that, but not all of these are toxic. This is simply because some of the heavy metals are not available in the ecosystems to the living organism, due to its low quantity in the earth's crust or its low solubility. To summarize these two factors, and distinguish the trace elements from the 53 existing heavy metals, sea water is used as an "average environment". Depending on the concentration of these elements in sea water, they can be grouped in four classes: **frequent elements** with concentration between 100 nM and 1 µM, such as Fe, Zn and Mo; elements with concentrations between 10 nM and 100 nM, such as Ni, Cu, As, V, Mn, Sn and U; **rare elements**, such as Co, Ce, Ag and Sb, and finally the ones that exist at 1 nM level, such as Cd, Cr, W, Ga, Zr,Th, Hg and Pb. The remaining 31 elements/heavy metals occur below 1 nM, and are very unlikely to ever be essential or toxic.

The importance of the mentioned 22 heavy metals is directly related with the solubility of their ions under physiological conditions, which eliminates Sn, Ce, Ga, Zr and Th (**insoluble**), and to their toxicity which involves its affinity to sulfur and their interaction with macrobioelements. From the 17 elements/heavy metals that are **soluble**, Fe, Mo and Mn are important trace elements with low toxicity; Zn, Ni, **Cu**, V, **Co**, W and Cr are **toxic elements** with high to moderate importance; and As, Ag, Sb, **Cd**, Hg, Pb and U, have limited beneficial functions and are considered **highly toxic** [5, 12].

3. Heavy metal resistant bacteria

In bacteria, the heavy metal ions have to enter the cell in order to cause toxicity. When inside the cell, the heavy metal cations, especially those with high atomic numbers, tend to bind to -SH groups. The **Minimal Inhibitory Concentration** (**MIC**) or **Maximum Tolerant Concentration** (**MTC**) of the metal ions is directly related with the complex dissociation constant of the respective sulphides. The binding of these metal ions to the sulfide groups of proteins/enzyme can inactivate them and as a consequence inhibit metabolic pathways [5].

In other cases, the cations of these heavy-metals can interact with the physiologic ions inhibiting its function [5]. In gram-negative bacteria, cations can bind to glutathione resulting in bisglutathionate complexes that react with oxygen species, and causes oxidative stress [13]. Oxyanions can also interfere with related non-metal structures producing radicals [5].

When the cell is in an environment with high concentrations of heavy metals, these ions are transported to the cytoplasm, through unspecific transporters (constitutively expressed) even when there are no metabolic needs. This “open gate” event is one of the reasons why heavy metal ions are toxic. Indeed, it was observed that mutations in genes coding for these transporters decreases the unspecific transport, and the organisms become metal-tolerant [14].

However, many organisms have molecular systems for heavy metal resistance. One is the ion efflux “pumping”, the most well-known bacterial heavy metal resistance system, which can decrease the accumulation of heavy metal ions inside the cell. Another is the segregation of thiol-containing molecules that bind heavy-metal cations, precipitating them

into complex compounds rendering them insoluble and less toxic. Thirdly, the organism can have molecular mechanisms for enzymatic detoxification (oxidation, reduction, methylation, and demethylation) modifying the heavy metal ion to a less toxic or less available form. Many heavy metal homeostasis and resistance system can involve two or three of these mechanisms [2, 5, 15].

Another mechanisms that can be considered important for resistance is the bio-accumulation (intracellular or at the cell surface) or bio-sequestration (conversion of the toxic ion through biological processes – "indirect" resistance system) [2].

Addressing the heavy metal metabolism, detoxification by reduction can only occur if the redox potential of the heavy metal is between the hydrogen/proton couple (- 421 mV) and the oxygen/hydrogen couple (+ 808 mV) (at 30°C and pH 7.0, which is the physiological redox range for most aerobic cells [12]. After reduction, the metal ion has to diffuse out of the cell to avoid re-oxidation. Therefore, the organism must present an efflux system to export the reduced metal ion. For a fast-growing organism, the cost of complexation for efflux is very high. Thus, complexation occurs mainly at low metal concentrations. Therefore, detoxification through efflux systems, alone or in combination, is the only possible option for a fast-growing organism in an environment contaminated with high concentrations of heavy metals. Thus, heavy metal metabolism can be described as **transport metabolism** [16].

3.1 Heavy metal transporters

The ion uptake by the cell is directly dependent of its structure. Looking at the ions structure, all mono- or divalent cations are small and structurally very similar, when compared with oxyanions that differ widely in structure and are bigger. Thus, uptake systems have to bind the metal ions tightly to differenciate them, which is possible at the cost of energy [17].

The transport of toxic cations or oxyanions found in bacterial resistance systems are mainly ATPase membrane pumps and chemiosmotic membrane pumps. These systems can be divided into seven major types:

- Two types of ATPases:

(1) P-type ATPases (also known as CPx-type ATPases) are generally single polypeptide chains, that can catalyze auto-phosphorylation by ATP intermediate. These ATPases are

located in the inner (cytoplasmic) membrane and they can transport ions to the cytoplasm or can operate as expulsion systems, removing ions from the cytoplasmatic space. Nevertheless, these ATPases need to be associated to outer membrane proteins to expel the ions to the exterior. The members of the P-type ATPases present eight transmembrane helices (TM), a large citoplasmatic region that binds to ATP (ATPBD, composed of N and P domains) and also several cysteines residues that "trap" the ions that are being expelled (actuator A domain) (Figure 1). Between the transmembrane helix six and seven there is a proline residue preceded by a probable metal-binding residue, the CPX motif that is essencial for its function. At the N-terminal, there is a soluble metal binding domain (MBDs) (Figure 1) [7, 18, 19].

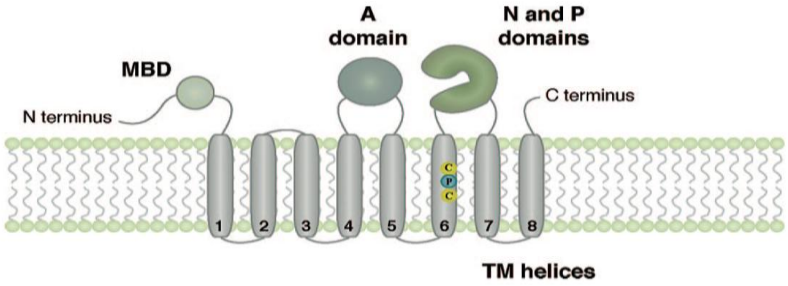

Figure 1 – Overall architecture of Cu (I) P-type ATPases. The number of MBDs at the N-terminus ranges from one to six depending on the organism [19].

(2) ABC ATPases (**A**TP **b**inding **c**assette without phosphorylated intermediates) consists of a cytoplasmatic membrane associated with an ATPase subunit, one or two membrane embeded pump channels, an ion-binding protein, and an outer membrane porin protein. In bacteria, the ABC ATPases are organized in four domains, requiring additionally a binding protein (Figure 2). The pumps of this family do not function as metal ions efflux systemm, only uptake [20, 21].

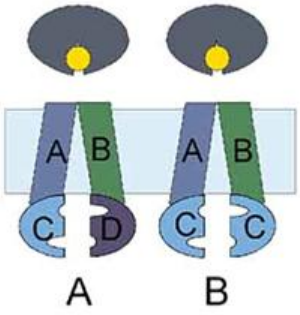

Figure 2 – Bacterial ABC transporters. The models show the diversity of the domain organization in these transporters. **A)** All four domains are encoded by separate genes. **B)** The transmembrane domains are encoded by separate genes while the nucleotide-binding domains are present in two copies, encoded by the same gene. Both cases require a binding protein. Adapted from [21].

- Five efflux pump classes that are chemiosmotic ion/proton exchangers:

(3) the single membrane polypeptides of the **Major Facilitator Superfamily** (MFS), generally a single polypeptide that transverses the membrane 12 or 14 times primarily in an alpha helical structure (Figure 3) [22, 23].

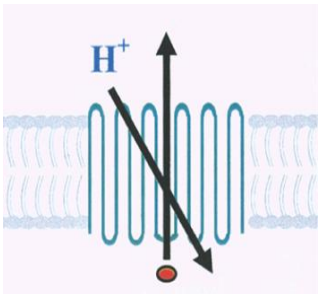

Figure 3 – Schematic model of the membrane topology of MFS proton-driven pumps in Gram-negative bacteria. Adapted from [23].

(4) the **Cation Diffusion Facilitator** (CDF) family [24] are medium size polypeptides (less than 400 residues), with six transmembrane segments, that function as homodimer in the inner membrane and are H^+ antiporter (for metal efflux) or H^+ symporter (for metal uptake). In this chemiosmotic ion transport process there is the participation of histidine (HRR – histidine-rich region), aspartate and glutamate residues (Figure 4) [7, 25].

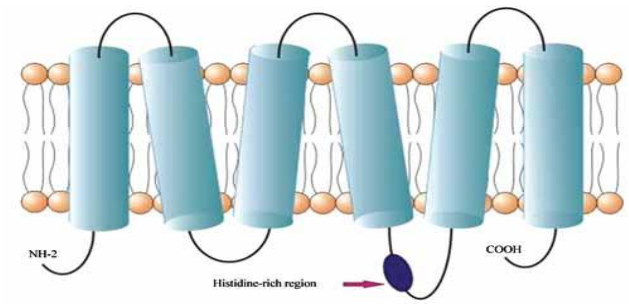

Figure 4 – Model of an CDF family protein [25].

(5) The **Resistance-Nodulation-cell Division** (RND) superfamily are transporters that have different roles in bacteria, namely resistance, nodulation and cell division. The inner membrane segment is composed of around 1000 residues, that expelles the ions. The auxiliary segments are a small protein from the outer membrane, and a protein that connects the outer and inner membrane segments, forming a complex that spans over the periplasmic space (Figure 5) [26-28].

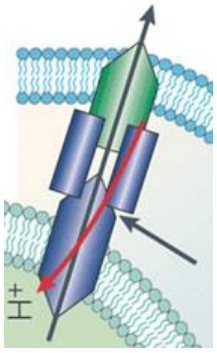

Figure 5 - Schematic representation of a RND efflux system in gram-negative bacteria. Adapted from [28].

- The last two chemiosmotic pumps have narrower substrate range:

(6) the **Chromate efflux system** (CHR) [29], and

(7) the **Arsenite** and **Antimonite efflux systems** that can function alone or with energy from ATP, being named as A-type ATPases [30].

There are two other transport system that differ from the ones presented above, the **mercury uptake system** [31] and the **arsenic efflux system** [32] which constitute their own group [15, 33].

3.2 Heavy metal repressors and activators

Bacterial chromosomes have genes for transport and/or for confering resistance to inorganic nutrient cations and oxyanions (such as NH_4^+, K^+, Mg^{2+}, Co^{2+}, Fe^{3+}, Cu^{2+}, Mn^{2+}, Zn^{2+}), to other trace cations, to PO^{3-}_4 and SO^{2-}_4, and to other less abundant oxyanions [3]. There are also bacterial vectors that present genes for resistance to metals ions, such as Ag^+, AsO_2^-, AsO_4^{3-}, Cd^{2+}, Co^{2+}, CrO_4^{2-}, Cu^{2+}, Hg^{2+}, Ni^{2+}, Sb^{3+}, TeO_3^{2-}, Tl^+ and Zn^{2+}. Plasmid resistance to $B_4O_7^{2-}$, Pb^{2+} and organo-tin compounds has also been reported [2] [3]. Together, this genetically-defined resistance accounts for a few hundred genes [3].

All bacterial heavy metal resistance systems are constituted by regulons, sometimes as operons (units of mRNA transcription), or clusters of contiguous genes (from two to more than ten) regulated by a protein that binds to the upstream operator DNA regions. Metal homeostasis relies on the ability of metalloregulatory proteins to coordinate the expression

of transport and storage proteins. In general, there are four classes of metal-responding transcriptional regulators:

(1) the **Mer family** of homo-dimeric activator proteins that bind to usually long 10-35 RNA polymerase-binding operator regions, with 19 or 20 nucleotide spacing [34];

(2) the **ArsR/CadC/SmtB family**, which act negatively, binding to the 17-spaced – 10-35 operator region preventing the binding of the RNA polymerase. When the protein binds to the co-repressor ion, the repressor is released from the DNA allowing transcription of the genes controlling the resistance to the metal ions [35];

(3) the **two-component systems**, in which the transmembrane histidine kinase membrane sensor (S) autophosphorylates at a conserved His residues, in response to an extracellular stimulus. The phosphate is then transferred to a conserved aspartate on the second responder (R) regulatory protein, which is then either activated or inactivated for DNA-binding transcriptional control [36];

(4) the **YXH complexes**, in which the H proteins are sigma factor subunits of RNA polymerase and respond to environmental cation addition, and the Y and X proteins bind to anti-sigma factors. Multiple promoter sites are found, one for the regulatory genes and another for genes determining the efflux pump [37].

In addition to the first level of regulatory proteins there is a second level. These proteins are directly involved in gene regulation and are at the same time chaperones: intracellular carrier proteins with the role of picking up the toxic ions, reducing their free cellular levels, and specifically transfering the inorganic ion to the transcriptional regulator or to other specific protein that binds metals ions as a cofactor [38].

4. Cadmium and Cobalt bacterial mechanisms

Cadmium has almost no known biological function, and can lead to cell damage and death even at low concentrations [39].

On the other hand, **cobalt** is an essential micronutrient for many bacteria, as a cofactor in different enzymes that catalyze a diverse array of reactions, such as production of adenosylcobalamin (one of the two natural cofactor forms of vitamin B_{12}). Cobalt is present mostly in the Co^{2+} form, and the Co^{3+} form is only stable in complex compounds.

This heavy metal is considered to have medium toxicity [40]. As mention, cobalt is present mainly as part of the cofactor cobalamine, vitamine B_{12}, but there is also another group of cobalt containing enzymes, such as nitrile hydratases [41].

Nevertheless, at high concentrations both transition metal ions are toxic to the cells, and some prokaryotes present sophisticated homeostatic mechanisms to regulate their uptake or efflux. Interestingly, both metals, together with zinc and nickel, are often recognized by the same membrane transporters and have the same type of gene regulators. This is due to their similar chemical properties [15,42].

4.1 Cadmium resistance

Toxic effects of cadmium ions can be resumed to "thiol-binding and protein denaturation", "interaction with calcium metabolism and membrane damage" and "interaction with zinc metabolism", or loss of a protective function The ion is transported typically by RND- and P-type transport systems in cyanobacteria, by RND-driven systems (Figure 5) in gram-negative bacteria and by P-type transport systems (Figure 1) in gram-positive bacteria. In all these cases the uptake systems are not exclusively for cadmium ion [5].

On the contrary to what is observed for other metal resistance systems, the molecular systems involved in cadmium resistance do not present a high homology between different bacterial families [42]:

- efflux ATPases in gram-positive bacteria. The plasmid born CadCA, CadDX and CadBX operons comprise metal efflux ATPase that differs in sequence and function. These systems are induced by Cd^{2+}, Pb^{2+}, Bi^{3+}, Zn^{2+}, and Co^{2+} at different levels [43, 44];

- metallothionines, cysteine-rich proteins that bind Zn^{2+} and Cd^{2+}, found in cyanobacterium and some Pseudomonads [45, 46]); and

- unrelated chemiosmotic cation-proton antiporters of gram-negative bacteria: as the **Czc** (**cadmium, zinc and cobalt resistance system**) and the **Czr** (**cadmium and zinc resistance system**) operons [47, 48].

4.2 Cobalt resistance

Cobalt ion, Co^{2+}, is rapidly accumulated in the cell by the fast and unspecific CorA system in most bacterial cells. This "pump" is commonly referred as a Co^{2+} and Ni^{2+} transporter, thus Co^{2+} transporters are unspecific [49]. This integral membrane protein belongs to the Metal Inorganic Transport (MIT) family [50]. At a low cobalt ion concentration, no ATP-driven uptake system has been found, but exceptionally a nickel transporter system, HoxN-type, can supply Co^{2+} for non-B_{12}-cobalt proteins [51].

In gram-negative bacteria, cobalt detoxification occurs, typically, by RND-driven systems [30] and in gram-positive bacteria by CDF transporters (Figure 4) [14, 52].

As previous described cobalt resistance appears always associated to other heavy metal, either nickel; **Ncc** operon (nickel, cobalt, and cadmium resistance system) [53], and **Cnr** operon (cobalt and nickel resistance system), which are mainly a nickel exporter [37, 54], or Zinc: **Czc** (cadmium, zinc and cobalt resistance system), mainly a zinc exporter [47] and **yohLMN**, a single polypeptide efflux protein, mainly found in gram-positive bacteria [55].

4.3 Cadmium and Cobalt resistance systems in Gram-negative bacteria

The Czc system

The Czc, **cadmium** (II), **zinc** (II) and **cobalt** (II) resistance system, is a efflux pump that function as a chemiosmostic divalent cation/proton antiporter [56]. This system lacks the ATPases motifs, demonstrating, that not all plasmid-resistance systems, with the function of ion efflux, are ATPases [6]. The proteins involved gave origin to a new family of three component chemiosmotic exporters called CBA systems [57], from the RND superfamily (Figure 5) [7, 58].

The Czc system from *Ralstonia metallidurans* (*R. metallidurans*) is used as a model for other gram-negative bacteria [47]. This system comprises usually, *czcNICBARSE*; **CzcN** is a membrane protein; **CzcI** serves a regulatory function and contains metal binding sites as well as leader peptidase signal sequence; **CzcC** is an outer membrane associated protein required to increase the substrate range of CzcAB; **CzcB** is a membrane fusion protein that mediates, with CzcA, the zinc resistance; **CzcA** is a membrane-bound cation-proton-antiporter that can mediate zinc and cobalt transport alone, but inefficiently. The junction of the *czcCBA* gene products comprises the three-component system, constituting the chemiosmotic efflux pump, Figure 6 [59];

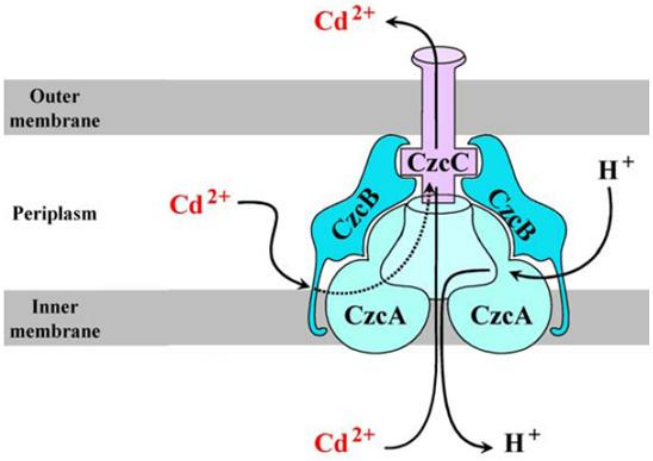

Figure 6 – Molecular model of the efflux RND chemiosmotic antiporter, czcCBA with cadmium (II) being transported [15].

CzcD is a membrane-bound sensor for periplasmic cations and a member of the CDF family [60]; **CzcR** is a response regulator of *czc* and interacts with **CzcS,** which is a histidine kinase sensor, homologous to the two-component regulatory system [14, 36, 57]; and **CzcE** is a periplasmic protein [61].

The Czr system

The Pseudomonad Czr, **cadmium** (II) and **zinc** (II), system is encoded by the operon *czrSRCBA*. The predicted gene products of *czrRCBA* showed a significant similarity with the proteins encoded by the plasmid metal resistant determinants CzcCBA (Figure 6). The predicted CzrS and CzrR are a two component regulatory system [36, 48].

The Cnr system

The inducible Cnr, **cobalt** (II) and **nickel** (II) resistance, operon from the model organism *R. metallidurans, cnrYXHCBAT* [37], is a chemiosmostic divalent cation/proton antiporter resistance system another example of non-ATPases toxic ion efflux [54, 62-64]. The Cnr system has high sequence homology to the Czc system, and thus might function in a similar manner. However, it presents some differences, since the Cnr system lacks the *czcR* and *czcD* genes, and instead three unrelated upstream genes *cnrYXH*, of the YXH complex family, are present [56, 62].

The *cnr* genes are organized in two gene clusters *cnrYXH* and *cnrCBA.* The product of the genes *cnrA, cnrB,* and *cnrC* are propose to form a membrane-bound protein complex catalyzing an energy-dependent efflux of Ni^{2+} and Co^{2+}, and the mechanism of action of the CnrCBA complex may be that of a proton/cation antiporter (Figure 6). CnrY is a trans-membrane signal transducer, forming part of the CnrYX anti-sigma-factor complex for CnrH. CnrX is a periplasmic nickel sensor, which is part of the CnrYX. CnrH is an extracytoplasmic sigma factor, which is required for transcription initiation at the promoter *cnrYp* and *cnrCp.* **CnrC** is an outer membrane associated protein, probably required for nickel export across the outer membrane. **CnrB** is a membrane fusion protein, which is probably required for transport of nickel ions across the periplasmic space. **CnrA** is a proton/cation antiporter of the RND family and is required for nickel export across the cytoplasmic membrane. **CnrT** is a nickel transport protein [63].

The Ncc system

The *ncc* operon, coding for another non-ATPase cation efflux system for **nickel** (II), **cobalt** (II), and **cadmium** (II) transport, is present mostly in *R. metallidurans*-like bacteria. The operon *nccYXHCBAN* codes for a high level resistance system that is non-inducible. [53]. Ncc is regulated by the YXH complex family [15]. The protein **NccY** affects the downregulation of the *ncc* operon; **NccX** contains many histidine residues that can bind Ni^{2+}, and is proposed to be involved in the regulation; **NccH** probably functions as an extracytoplasmic sigma70*;* **NccC** is involved in cation binding or substrate specificity, since its N-terminal contains small hydrophobic regions (15 residues), which may be embedded in the membrane at the N-terminus and functions in conjunction with the export permeases; **NccB** belongs to a novel MFP family of cytoplasmic proteins; **NccA** contains a double set of 6 extended hydrophobic regions possibly spanning the membrane, and it is thought to form a membrane channel to allow ion transport. There is no homology to ATP-binding motifs or to known metal ion-translocating ATPases; **NccN** is apparently involved in metal specificity, since it only affects nickel resistance and together with its hydrophobic properties, it suggests that the protein is located in the membrane and possibly involved in Ni^{2+} transport [53].

In general, a bacterium strain is consider to be cadmium resistant when it can grow in the presence of at least 150 μM cadmium ions, and it is cobalt resistant when it tolerates concentrations as high as 250 μM cobalt ions [65].

5. Copper bacterial mechanisms

5.1 Copper homeostasis

Copper is widely distributed in the environment and it is a key element in cellular metabolism. Its cellular role is related to the fact that copper is a cofactor of several key enzymes. Nevertheless, copper ions can also be very toxic, when unbound inside the cell, since it can easily catalyze Fenton reactions [66], producing hydroperoxide radicals [67-69]. This highly reactive hydroxyl radicals damage biomolecules, such as DNA, proteins and lipids. Copper ion toxicity can also be due to its binding to adventitious sites in proteins, nucleid acids, polysaccharides and lipids, causing the displacement of native metal ions, as well as, changes in the protein structure and/or function [66, 70].

The reduced form, Cu^{+}, is a closed shell $3d^{10}$ transition metal ion and hence diamagnetic. As a soft Lewis acid, it tends to bind to soft bases, such as, thiols, hybrids, alkyl groups, cyanide, and phosphines. On the other hand, the oxidized form, Cu^{2+}, has a $3d^{9}$ configuration and is paramagnetic. As an intermediate Lewis acid, it forms complexes with additional ligands to the ones of Cu^{+}, including sulfate and nitrate [71].

Cu^{+} is susceptible to disproportionation to give Cu^{2+} and Cu^{0}, while Cu^{0} is instantaneously oxidized to Cu^{+} and subsequently to Cu^{2+}, by atmospheric oxygen. Copper reduction can occur at the cell surface, prior to the transport. However, Cu^{2+} is more abundant than Cu^{+} in aerobic environments and is less toxic [15, 71].

In anaerobic conditions copper is much more toxic, due to the reduction of Cu^{2+} to Cu^{+}, which can diffuse through the cytoplasmatic membrane causing increased copper accumulation [72]. In normal conditions, the cytoplasm form of copper is exclusively Cu^{+}, and never as a free cation. Indeed, all copper ions must be compartmentalized or protein bound to avoid its toxic effects [73].

Therefore, bacteria have to maintain copper levels in a precise concentration to avoid toxicity, but maintaining a supply for copper-requiring proteins. In order to achieve this, transcriptional regulators have to distinguish copper from the other metal ions, and in response to levels below or above a certain biological threshold trigger a physiological

response, such as copper-import, -export, or -detoxification (copper ion oxidation). Furthermore, copper-chaperones can be used to divert copper ions from adventitious sites [74].

The models of copper homeostasis and resistance involve the control of cellular copper uptake and efflux by proteins encoded by chromosomal genes, as well as plasmid systems for additional resistance in high copper environments [3]. Bacterial copper-binding proteins appear to be extracytoplasmatic and are found mainly in the periplasm or embedded in the cytoplasmic membrane [75].

All the proteins responsible for the copper homeostasis present metal-binding motifs with atypical coordination chemistry, relative to enzymes that incorporate copper as a cofactor [76]. This atypical coordination is observed in cytoplasmatic proteins with the use of cysteines to bind copper ion, in contrast, with the extracellular or periplasmatic copper proteins that use the side-chain of methionine and histidine or even tryptophan residues, since the oxidizing environment of the periplasm is not appropriated for the use of cysteines [76-78].

The most widely distributed copper-trafficking proteins/pumps in both prokaryotes and eukaryotes for copper homeostasis are the CPx-type ATPases, functioning as energy dependent efflux of toxic ions (Figure 1) [5, 74, 75]. These ATPases can be separated in a subfamily called P_{1B}-type ATPases that have an invariant aspartate residue in the conserved DKTGT motif, which becomes transiently phosphorilated during the reaction cycle [79].

5.2 Copper acquisition and resistance

Analysis of bacterial genomes shows the presence of multiple sets of copper-related pathways [80]. Copper-resistance genes are organized into a regulatory network, which senses and responds differently to copper concentration, depending on the organism [75]. Nevertheless, the known mechanisms can be grouped in gram-positive and gram-negative bacteria occurring systems.

In gram-positive bacteria, the *cop* operon from *Enterococcus hirae* (*En. hirae*) is the most studied, which comprises four genes that encode a repressor, CopY, a copper chaperone, CopZ, and two P_{1B}-type copper ATPases, CopA and CopB [81]. Gram-negative bacteria systems can be separated in two major classes: the *Es.coli*-like bacteria systems that comprises the *cue* (Cu-efflux) regulon, *cus* (Cu-sensing) locus, and *pco* plasmid copper

resistance systems, and the Pseudomonads-like bacteria systems that comprises the *cop* operon, plasmid or chromosomal copper resistance, and the *cue* system [75].

There is another copper homeostasis system that is unique to Methanotrophs, and controls the copper uptake for methane monooxygenase (pMMO, that catalyzes the conversion of methane to methanol). This system comprises the segregation of a small siderophore-like compound called methanobactin, with the role in copper-uptake and/or copper detoxification. This compound binds to Cu^{+} and Cu^{2+}, with the ability to reduce Cu^{2+} to Cu^{+}, and supplies the periplasm of the cell with copper ions [82-84].

5.3 Gram-negative bacteria – Copper resistance systems

5.3.1 *Es.coli*-like bacteria

In *Es.coli* there are multiple systems for the safe handling of copper. A central component in copper homeostasis of *Es.coli* is the integral inner membrane copper-transporter P-type ATPase, CopA, which ensures removal of excess Cu^{+} from the cytoplasm. In the periplasmic space, both the multi-copper oxidase, CueO, and the multicomponent copper efflux system CusCFBA are responsible for the control of copper levels. Additionally, various *Es.coli* strains harbor plasmid-encoded genes that increase considerably the copper tolerance, allowing growth in copper-rich environments (with more than 1 mM of copper ions) [85, 86].

Cue (<u>Cu</u>-<u>e</u>fflux) regulon

The highly competitive nature of copper imposes that the cell somehow restricts protein access to this metal ion [86]. For this purpose there is a multicopper oxidase, **CueO** (responsible for copper detoxification) within the *Es.coli* periplasm and a cytosolic copper sensor, **CueR**, and an P_{1B}-type ATPase, **CopA** (Figure 7). This system solely detects "stray" copper atoms, even though this element is not-essential within the *Es.coli* cytosol. Thus, when copper ions enter the cell by porins in the outer membrane or by passive diffusion, the first system to respond is the ***cue*** (**<u>cu</u>-<u>e</u>fflux**) regulon, a primary copper homeostasis system, that clears the cytosol from excess copper ions, while the excess is kept in the periplasm [72].

CopA functions under aerobic (and anaerobic) conditions, maintaining a low intracellular copper level. CueO can also oxidize catechol siderophores that can sequester

copper [15, 86]. As other multi-copper oxidases, this enzyme binds four copper ions [75] and its activity is oxygen dependent [72, 75].

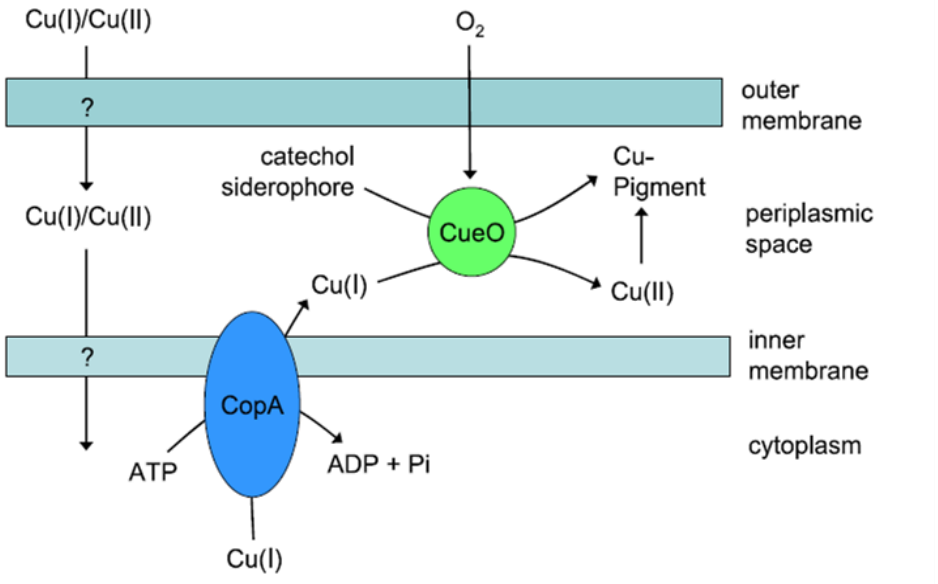

Figure 7 – CopA ATPase and CueO multicopper oxidase from *Es.coli*. These systems are activated at 50 µM copper sulfate under aerobic conditions. Adapted from [87] and [72].

Cus (Cu-sensing) locus

At near-lethal copper ion concentration (higher than 200 µM copper ions), or under anaerobic conditions, *Es.coli* has to rely on another copper homeostasis system, the *cus* (**Cu-sensing**) locus [72].

This chromosomally encoded copper resistance system comprises the *cusCBFA* operon, and is induced by the CusR/S two-component regulatory system, which detects surplus Cu^{+} in the periplasm, mediating export of copper across the outer membrane (Figure 8) [88-90].

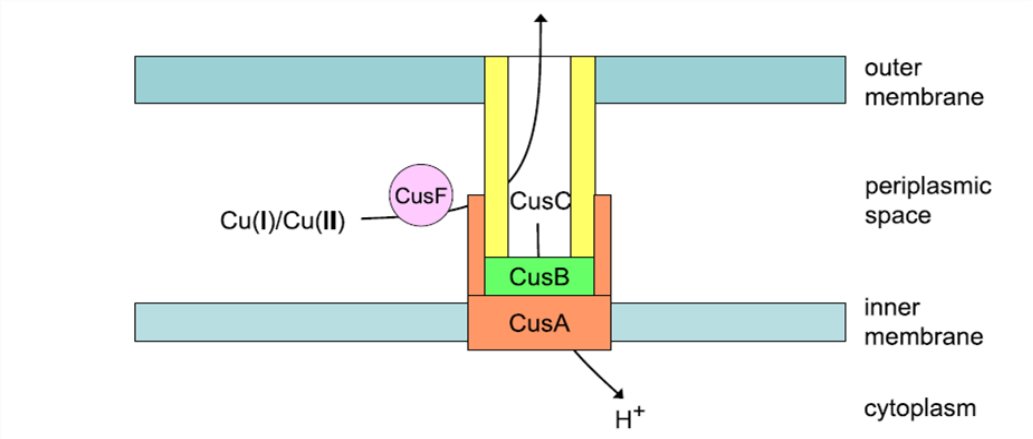

Figure 8 – The Cus system of *Es.coli.* Adapted from [87]

The main component of this system is the RND efflux chemiosmotic carrier, CusCBA. Recent experimental evidence suggests that the Cus system transports copper from the periplasmic space, rather than from the cytoplasm, to the extracellular environment [88] [89]. The components are an outer membrane efflux, CusC, a periplasmatic coupling protein, CusB, and an inner porin protein, CusA, respectively. It is assumed that CusA functions as a secondary transporter energized by proton-substrate antiport and is responsible for the substrate specificity. The *cusB* codes for a membrane fusion protein, which serves as an adaptor to link CusA with CusC [86].

Another component of this system is CusF, a 10 kDa periplasmic protein, that is proposed to interact with CusB and CusC and is a novel periplasmic copper chaperone (Figure 8). It is suggested that the entry of copper to the CusCBA occurs from the periplasm via this copper chaperone [88]. CusF binds Cu^{+} and Cu^{2+} and exhibits an absorption maximum at 510 nm, exhibiting an unusual pink color for a copper protein [88, 91]. This unusual feature is a consequence of the unusual structure and geometry of the copper center: CusF forms a five-stranded *β*-barrel, classified as an OB-fold, which is a unique topology for a copper-binding protein. NMR chemical shift mapping experiments suggested that Cu^{+} is bound by conserved residues His-36, Met-47, and Met-49 located in the *β*-strands 2 and 3 [92].

Pco - plasmid copper resistance

In addition to the chromosomally-determined copper homeostasis gene products, some *Es.coli* have the ***pco***, **plasmid-copper resistance system**, which is activated when the

concentration of copper ions in the environment is extremely high (1 mM copper ions) [86]. The *pco* gene operon, *pcoABCDRSE*, is a copper inducible system and is thought to mediate periplasmatic binding and oxidation of excess copper, carrying out transmembrane ions transport [85, 86]. The *Es.coli* PcoABCDRS proteins are closely related to the *cop* products of Pseudomonad *cop* system (see below). In addition, there is a gene, only present in *Es.coli*, *pcoE*, downstream from the regulatory genes [93]. The regulation of this operon is performed by a two-component regulatory system: PcoS is a membrane-bound Cu^{2+} sensor, and PcoR is a DNA-binding soluble regulator [36].

PcoA, a 66 kDa protein, is a multicopper oxidase enzyme [94] similar to CueO. PcoA is proposed to be involved in copper detoxification by oxidizing Cu^{+} to less toxic Cu^{2+} and/or by oxidizing catechol siderophores, such as enterobactin, which can then sequester copper. This later function may be related to the formation of brown colonies by *Es.coli*, and other bacteria that have Pco-like systems, when grown in the presence of high copper ions concentrations. These strains have shown no periplasmatic copper storage [95].

PcoB is a 33 kDa protein, predicted to be located in the outer membrane. PcoB together with PcoA confer higher copper resistance than PcoA alone, indicating that they act in concert, but the exact function of PcoB, remains to be identified. Nevertheless, it is known that PcoB contains four imperfect repeats of the octapeptide Asp-His-Ser-Gln/Lys-Met-Gln-Gly-Met, which are proposed to be involved in copper binding. Similar motifs are also present in PcoA [93].

PcoC is a 12 kDa periplasmic protein, that binds simultaneously Cu^{+} and Cu^{2+} [96]. It is suggested that PcoC is a copper chaperone and is postulated to interact with PcoA. PcoC-like proteins differ from copper chaperones in their copper binding ligands (CXXC motif, Cu^{+}, and $H(X)_nE(X)_mDXH$, Cu^{2+}, see below "Small copper proteins"). Instead of a CXXC motif, PcoC-like proteins contain conserved methionine-rich $[M(X)_nM]_m$ sequences, for Cu^{+} binding, as well as, conserved histidine residues, for Cu^{2+} binding (at equatorial positions) [19, 97].

PcoD is a 34 kDa inner membrane protein with eight predicted transmembrane helices and **PcoE**, which is transcribed from its own promoter, is a small periplasmic protein that is required for full copper resistance, and it is also proposed to be a periplasmic copper chaperone [86, 88, 98].

Copper ions enter in the periplasmic space, possible by cell porins or by passive diffusion, and PcoC shuttles these ions to PcoD, which can provide the required copper to be

incorporated into apo–PcoA prior to its transport to the periplasm via the TAT-pathway (twin arginine translocation pathway). PcoA, the multi-copper oxidase, is located in the periplam, where it can detoxify the copper ions and oxidize catechol siderophores, transported by PcoB. PcoE shuttles Cu^{+} to PcoA for oxidation to a less toxic form, Cu^{2+} (Figure 9)[100].

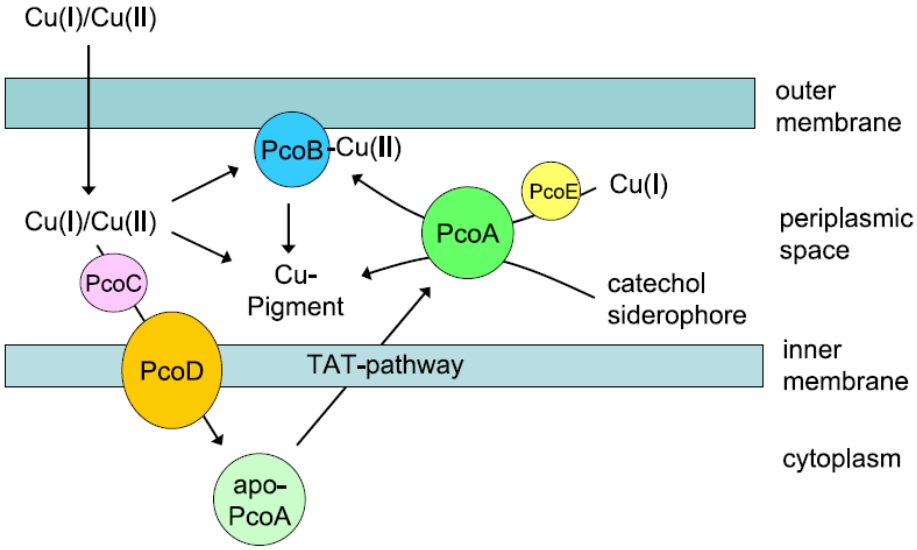

Figure 9 - The Pco systems of *Es.coli*. The Cop system from *Ps.syringae* is homologous to Pco, with the expection of PcoE that is not present in this bacterium. Adopted from [87]

In *Es.coli* the strains that resist to a value higher than 0.8 mM copper ions are considered copper resistant strains [75]

5.3.2 Pseudomonads-like bacteria

Cop operon

In the Pseudomonads-like bacteria, the plasmid copper resistance operon, Cop, (associated with resistance to high copper concentration) is characterized by four structural genes, *copABCD,* and two regulatory genes, *copRS*, a two-component regulatory system. This system is homologous to the Pco system found in *Es.coli*-like bacteria [36].

The *cop* system is induced by copper ions, in a copper concentration-dependent manner, in which the regulatory genes are transcribed separately by a upstream promoter of *copR* [99]. Similarly, to any other two-component regulatory system CopR is located in the cytoplasm and CopS in the inner membrane. As for the other proteins, CopD in located in the

inner membrane, CopB in the outer membrane, and CopA and CopC are present in the periplasm (Figure 9) [100-102].

The role of **CopA**, a type 1 (or blue) multiple copper-oxidase that binds 11 Cu^{2+} ions, is to oxidise Cu^{+} to Cu^{2+}, by accepting electrons from a mononuclear copper centre and transferring them to a trinuclear copper centre; dioxygen binds to the trinuclear centre and, following the transfer of four electrons, is reduced to two water molecules. CopA contains four conserved motifs, two MGGM and two MAGM motifs. This protein has a leader peptide with two arginines, which are responsible for export through the twin-arginine translocation pathway, similarly to what was described for PcoA [102].

CopB, as PcoB, is associated with copper sequestration in the periplasm and outer membrane, and it may prevent toxic levels of copper from entering in the cytoplasm [100] [101]. **CopC** is a small soluble copper-binding protein that binds 1 Cu^{+} ion and 1 Cu^{2+} ion, and it is proposed to function as copper carrier in the periplasm. The protein presents a β-barrel with two separated binding sites specific for Cu^{+} and Cu^{2+}: $[Cu^{+}(His)(Met)_{2,3}]$ and $[Cu^{2+}(His)_2(N\text{-term})(OH_2)]$, respectively. The two copper binding sites are independent with the clear distinction of the methionine-rich motif that is involved in Cu^{+} binding [102]. **CopD**, is proposed to be involved in copper transport [103].

In Pseudomonad*s,* the *cop* operon is involved in periplasmatic sequestration of Cu^{2+}. In addiction, these bacteria present chromosomally encoded P_{1B}-type ATPases, which provide partial resistance by effluxing Cu^{2+} or Cu^{+} out of the cytoplasm [3]. Sequestering copper ions in the periplasmatic space and cytoplasmic membrane is a protective measure for the cell. This fact explains why some colonies of the copper-resistant Pseudomonads turn bright blue when grown in high copper-containing media, while the supporting agar become colorless, as the copper salts are taken up by the cells [101, 104].

The study of copper resistant strains shows that the presence of the *cop* operon does not necessarily lead to higher levels of copper tolerance. Furthermore, studies of mutant strains demonstrate that the number of copies of the *cop* operon or the lack of certain genes of this operon does not change its function. However, if the mutation occurs at the level of a protein with important metal-binding motifs (MBF), such as CopC, the organism changes from copper resistant to copper sensitive (decreasing its MIC/MTC up to 8–fold). Thus, there is a strong correlation between the number of such domains and the level of copper tolerance of the bacterium [80, 100, 105].

Some *cop* operons and their homologues have been found encoded in the chromosomal DNA [106-108]. These genes impare relatively low levels of copper tolerance when compared to strains carrying plasmid-borne *cop* genes [107].

Many Pseudomonads strains have been reported to be tolerant to different levels of copper ions [106, 108, 109]. The copper-sensitive group, with a lower tolerance to copper ions, presents MIC/MTC values of 0.4 to 0.6 mM of copper ions, while the copper-resistant group has MIC/MTC higher than 0.6 mM [105].

Gene clusters similar to *Es.coli pcoABCDRSE* and to *Ps.syringae copABCRS* operons can be identified in the genomes of many bacteria. Often only homologues of PcoA/CopA and PcoB/CopB are encoded on a genome, suggesting that PcoC/CopC and PcoD/CopD are an accessory system required for maximal resistance [75].

Cue operon

In some *Pseudomonas* species, a second operon is present, *cueAR*, similar to the one found in *Es.coli* (Figure 7). This operon codes for a P_{1B}-type ATPase exporter of copper, CueA, and a MerR-type helix-turn-helix regulator, CueR. This operon is involved in copper homeostasis, rather than confering resistance to high copper concentrations. The *Pseudomonas* species that present this operon have a MIC/MTC value of 0.4 mM of copper ions [105].

The chromosomally located operon is autoregulated by the presence of copper ions, with the CueR being positively regulated. This operon, in opposition to the one previously described (*cop* operon), is located in the chromosome, and when it is present the *cop* operon lacks the two-component regulator, CopRS. This was observed in copper sensitive strains. Nevertheless, in some strains, in the presence of higher copper concentration, the *cop* operon is regulated by a homologous two-component regulatory system present in the chromosome [105, 106].

6. Small-copper binding proteins

Copper is needed for living systems and the intracellular trafficking of this metal to copper-dependent proteins is fundamental for the normal cellular metabolism. The copper chaperones have the dual role of trafficking and preventing the cytoplasmic exposure to copper ions when it is in transit to proteins that require copper as a cofactor (during copper routing) [110].

Each copper-dependent protein in the cell is served by a specific copper chaperone. Most of the copper chaperones characterized up-to-date have an "open-faced β-sandwich" fold with a conserved metal-binding motif (Cu^{+} binding site). The specificity for a given copper-dependent protein appears to be mediated by the residues around the copper-binding site. These copper chaperones bind to such proteins as Cu^{+} in a trigonal planar geometry coordinated by three sulfur ligands, MXCXXC; others can also bind Cu^{2+} in the motif, $H(X)_nE(X)_mDXH$ (Cu^{2+} binding site) [110, 111].

There are other copper-binding proteins, which have an intense blue color and present the so-called type 1 copper center, that in the oxidized state, exhibit unusual electronic paramagnetic resonance (EPR) spectra [112]

The blue copper proteins, such as pseudoazurins, plastocyanin, ceruloplasmin, azurin, and stellacyanin, presents different copper binding motifs, being $Cu(N\delta_{His})_2S\gamma_{Cys}$-R, the most common, with R differing for each protein family: R-$S\delta_{Met}$, R- $O\varepsilon_{Glu}$, and R-H_2O [102]; The intense blue color is due to an electronic transition that has a sulfur-metal charge transfer character [112].

7. *Marinobacter aquaeolei* case study

Marinobacter aquaeolei (*Ma.aq*) VT8 is a gram-negative halophilic and mesophilic bacterium. This bacterium is non-spore-forming, rod-shaped and presents mobility by a polar flagellum. It can grow aerobic and anaerobically, in the presence of nitrate on succinate, citrate or acetate, but not on glucose. Several organic acids and amino acids can be used as sole carbon and energy sources [113].

Phylogenetic analysis of 16S rDNA indicates that this bacterium is closely related to *Marinobacter* sp. strain CAB (99.8% similarity), *Marinobacter hydrocarbonoclasticus* (99.4% similarity), and *Pseudomonas nautica* (*Ps.nautica*) (99.7% similarity) [113, 114]. A preliminary phylogenetic study indicates that *Ps.nautica* 617 is misclassified and also belongs to the genus *Marinobacter*, but the specie remains to be determined [113-115]. During this work it was considered that *Ps.nautica* 617 is misclassified and was considered that it is a strain of the species *Ma.aq*, which justifies the use of *Ma.aq* genome to perform the bioinformatic studies presented, as well as for the primer desing.

7.1 Copper Regulon

In the last years, there was an increase in the number of total sequenced genomes. These projects followed the practice of the initial annotation of the open reading frames (ORF) by the attribution of the same name as the closest related protein studied. So, several of these projects uncovered thousand of systems by functionally equivalent comparative genomics [3]. *Ma.aq* was sequenced as an ideal model for the study of microbial iron oxidation, and as in other sequenced projects several molecular systems were uncovered [113].

Searching for copper resistance proteins annotated in the genome of *Ma.aq* VT8 it was found a putative copper regulon, *copSRXAB-copA-cusCBA* (Figure 10), which is located in a distinctive locus, 150214 – 163665, in the chromosome with ORFs annotated as:

- Putative operon for the copper resistance, *copSRXAB* (150214 – 155496)

(1) Two-component regulatory system, *copSR*, associated with the regulation of heavy-metal resistance (non-specific);

(2) Copper-binding protein, *copX*, with no relation with known proteins from copper resistance systems;

(3) Copper resistance proteins, *copAB*, similar to the *cop* system of Pseudomonads-like bacteria and *pco* plasmid copper resistance *Es.coli*-like bacteria;

- The other two components of the *locus*, *copA-cusCBA*

(4) Copper binding P_{1B}-type ATPases, CopA,associated with copper transport, similar to CueA of Pseudomonads and CopA in *Es.coli*-like bacteria;

(5) Proposed copper chemiosmotic diffusion facilitator, cusCBA, similar to related proteins in the *cus* system in *Es.coli*-like bacteria.

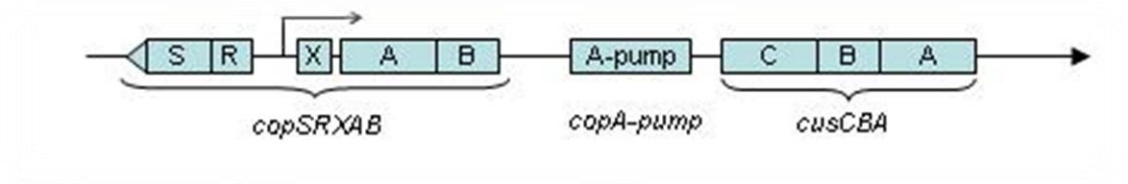

Figure 10 – *Loci* (150214 – 163665) organization of the putative copper regulon in *Ma.aq copSRXAB* operon, copper binding P_{1B}-type ATPases, CopA, and *cusCBA.* Arrow highlights the transcription orientation.

The organization of this regulon implies that *Ma.aq* presents a unique copper resistance system in gram-negative bacteria, which might have evolved differently from the ones present in *Pseudomonads*-like and *Es.coli*-like bacteria.

The work presented here aims at determining whether *Ma.aq* is a copper resistant bacterium, and gaining insights into the relation of the copper operon with this copper resistance, with special focus on CopX (proposed designation for the annotated copper-binding protein). CopX was heterologously isolated and biochemical characterized in order to determine if this protein has the features of copper chaperones from the CopC/PcoC family or of other copper-binding proteins.

7.2 Cobalt and Cadmium system

There is no cobalt or cadmium system annotated in the genome of *Ma.aq*. Indeed, in this genome there are genes annotated as belonging to a *czc* system, which are not organized as an operon. Nevertheless, near the copper *cusCBA* locus there is an annoted gene product as CzcD, and some databases associate the *cusCBA* locus to a *czcCBA* system, as a cation diffusion facilitator of cadmium (II), zinc (II) and cobalt (II). There is also some genes annoted as *cnr* system related proteins, but again they are not organized in as an operon.

The aim of this work was to establish whether *Ma.aq* is resistant to these two heavy metal ions.

MATERIAL AND METHODS

1. Bioinformatic Analysis

1.1 Operon Analysis

1.1.1 Genome comparison

A complete genome analysis was performed using the graphic interface Gene Ortholog Neighborhoods at the Integrated Microbial Genome system [116]. The search was performed against all genomes from the DOE Joint Genome Institute databank using the *Ma.aq* VT8 genome, particularly the proposed copper operon (with the *copX* has target). The alignment (Figure 1 App.A) and score Table (Table 1 App.A) can be found in Appendix A.

1.1.2 Protein Homology study

The aminoacid and nucleotide sequences of the proposed copper operon of *Ma.aq* VT8 were aligned with homologous copper resistance systems. The study was performed in the online software Basic Local Alignment Search Tool (BLAST) [117] using the graphic interface – Protein Blast and Nucleotide Blast. The selected database was the non-redundant protein sequence (nr) and the algorithm used was blastp (protein-protein BLAST) for amino acid sequences; and for the gene sequences the selected database was the nucleotide collection (nr/nt) and the algorithm used was blastn (somewhat similar sequences) or discontiguos megablast (more dissimilar sequences).

Two-component regulatory system, CopR and CopS

The proteins and DNA sequences of the target regulatory proteins, CopR and CopS, of the putative copper operon of *Ma.aq* VT8 were aligned with other known two-component regulatory system, using the ClustalX multiple alignment software (version 2.0) [118]. All protein and DNA sequences were obtained in the FASTA format [119] from the National Center for Biotechnology Information (NCBI) using the graphic interface ENTREZ – protein sequence database and nucleotide sequence database. Accession numbers for all the protein sequence and nucleotide sequence used in this study (Table 2 and 3 App.B.1(A/B) and Table 4 and 5 App.B.2(A/B) can be found in Appendix B.

A novel copper protein

The amino acid sequence of the target copper binding protein, CopX, from the putative copper operon of *Ma.aq* VT8 was aligned with other known copper proteins using the ClustalX multiple alignment software (version 2.0) [118]. All protein sequences were obtained in FASTA format from the National Center for Biotechnology Information (NCBI) using the graphic interface ENTREZ – protein sequence database. Accession numbers for all protein sequence used in this study (Table 6 App.B.3) can be found in Appendix B

1.2 Protein phylogeny

CopR

To perform the phylogenic inference of the protein that encodes the regulator protein, CopR, of the two-component system that regulates the proposed copper operon of *Ma.aq*, selected protein sequences were collected from the NCBI graphic interface ENTREZ – protein sequence database, which widely represent the bacterial metal resistance system regulators in Eubacteria [120, 121]. Almost all correspondent protein sequences from these studies were used, and other sequences were added, such as an outgroup sequences from related regulators but with far relationship (determined by protein homology studies), from new sequenced metal resistant bacteria, and sequences from the previously gene ortholog neighborhoods studied. All these protein sequences were retrieved from the databank in FASTA format, aligned and converted to PHYLIP format using the ClustalX multiple alignment software (version 2.0) [118]. The phylogenic construction was obtained using a software package to infer phylogeny, Phylogeny Inference Package (PHYLIP) (version 3.68) [122]. The phylogenic analysis was performed using the following algorithms: Maximum Likelihood (ML), Neighbor-Joining (NJ) and Parsimony (P). These methods were applied to the PHYLIP sequences base using PROML, PROTDIST+NEIGHBOR and PROTPARS, with SEQBOOT programmed to 1000 replicates to bootstrap and number seed 9. Bootstrap values were calculated using CONSENSE [122]. Phylogenic tree obtained from PHYLIP were represented and arranged, without being altered, using the TreeExplorer (version 2.12) inside the Molecular Evolutionary Genetic Analysis (MEGA) software (version 4.0) [123] [124]. Accession numbers for all protein sequences used can be found in Appendix C (Table 7 App.C.1).

CopX

A similarly procedure, as the one described for CopR, was used to perform the phylogenic inference of the copper binding protein, CopX, from the proposed copper operon of *Ma.aq* VT8. Selected protein sequences were collected from the NCBI graphic interface ENTREZ – protein sequence database, which widely represent copper binding proteins and copper proteins related to copper resistance in bacteria [125]. Accession numbers for all protein sequences used can be found in Appendix C (Table 8 App.C.3).

2. Growth of bacterial cells

2.1 *Marinobacter aquaeolei* 617

The bacterial strain was maintained on Artificial Sea Water (ASW) agar plates, corresponding to a modified version of marine bacterium medium presented by Baumann & Baumann [126], containing 0.8 % agar, 0.2 % lactate and 0.1% yeast extract as a carbon source. Before every inoculum, ASW medium was supplement with three solutions, modified version Starkey solution [127], phosphate and iron solutions, according to Table 2. The bacteria were grown in aerobic conditions in liquid and solid medium with the composition presented in Table 1. In liquid medium, the growth was performed in three steps: day one – 5 mL pre-inoculum inoculated from the agar plate; day two – from 5mL→100mL; and day three from 100mL→500mL. All inoculums were performed in a laminar flow chamber (Heraeus-SB48) and the growth stages were performed in an orbital shaker (230 rpm) at 30°C (Infors HT- AK62). The first two growths had the duration of 24h, while the last had the duration of 24 h to 48 h.

Table 1 – Composition of ASW medium. This solution is adjusted to pH 7.5 with HCl. The yeast extract was only added after the pH adjustment. Solution was sterilized by autoclaving. To prepare solid medium add 8 g/L of agar to the medium before autoclaving.

Compound	Concentration (g/L)
NaCl	11.7
$MgSO_4.7H_2O$	12.3
KCl	0.75
Tris-Base	6.05
NH_4Cl	3.00
$CaCl_2$	1.49
Yeast extract	1.00
Lactate	1.0%

Table 2 – Compositon of Starkey, phosphate and iron solutions. The phosphate solution was sterilized by autoclave process and the Starkey and iron solutions were sterilized by filtration using 0.2 µM diameter filters (Sarstedt).

Solution	Compound	Concentration (g/L)
Starkey	$FeSO_4.7H_2O$	6.20
	$ZnSO_4.7H_2O$	1.44
	$MnSO_4.4H_2O$	1.12
	$CuSO_4.5H_2O$	0.25
	$CoSO_4.7H_2O$	0.90
	BO_3H_3	0.06
	$Mo_7(NH_4)_6O_{24}.4H_2O$	1.00
	$NaSe0_3$	0.02
	$Ni(NO_3)_3.6H_2O$	0.04
	HCl 99%	51.4 mL
Iron	$FeSO_4.7H_2O$	1.00
	$H2SO_4$	1 to 2 drops
Phosphate	K_2HPO_4	18.6

The OD_{600nm} of the growth was measured, at hourly intervals, to follow the growth curve. Before inoculating a fresh culture or harvesting, the growth was checked at an optical microscope (Olympus-BX51) for culture purity, using white light, at an amplification of 400x and 1000x (Figure 11). After 24 h or 48 h of growth, cells were harvested in a Beckman centrifuge, model Avanti J-25 (rotor JA-10), at 10000 rpm, during 10 min at 4°C. A 1 mL sample from the medium was collected and stored at -80°C, and the rest discarded. The cell pellet was resuspended in 50 mM Tris-HCl buffer pH 7.6, in a 2 mL of buffer/1 g of wet cell. The cell suspension was storaged at -80°C until further use.

2.2 Heavy metal resistance

A defined ASW solid and liquid medium was used to examine the response of *Ma.aq* to cadmium, copper and cobalt ions (Cd^{2+}, Cu^{2+}, and Co^{2+}). During these tests, and before inoculums, plates were controlled for pure cultures, Figure 11. Other methodologies were identical to the ones mentioned in 2.1.

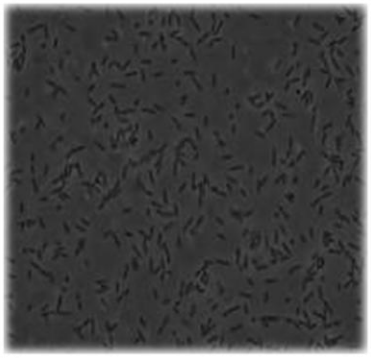

Figure 11 – *Ma.aq* grown in liquid medium for 24 h. The culture was visualized using an optic microscope (Olympus-BX51), amplification 1000x.

Figure 11 shows an image of *Ma.aq* grown in liquid medium for 24 h in aerobic conditions. Identical image was obtained from an ASW plate inoculated with *Ma.aq* and incubated for 24 - 120 h. It is possible to recognize the *Ma.aq* bacteria from possible contaminants by the non-spore-forming rod-shaped form and flagellum mobility [113] [115].

2.2.1 Copper ion stress

The effect of copper ion in the growth of *Ma.aq* 617 was studied both in solid and liquid medium. For this, ASW medium was supplemented with different concentrations of $CuSO_4.5H_2O$, in the third step of the growth (day three). The solid medium was supplemented before sterilization and the liquid medium after sterilization. The copper concentrations tested were: 0.2 (control), 1, 2, 7, 10, 100 nM; 0.5, 1.0, 10, 20, 100, 200, 300, 400, 500, 600, 700, 800, 900 µM; 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 3.0 mM. The growth conditions were the ones described in 2.1. A control growth was always performed alongside these copper stress growths. All concentrations were streaked in triplicate and in liquid medium concentrations were tested at least twice. The registered OD_{600nm} values were corrected for the contribution of the copper sulfate at 600 nm for

growth curve representation (a sample of each medium was taken before inoculation and the OD_{600nm} values obtained were subtracted at each measured OD_{600nm} during the growth).

In solid media, non-inoculated copper supplemented plates were also incubated up to 120h, and no change in the medium was observed (Appendix E, Figure 11 App. E.2)

It is important to mention that the standard ASW solid or liquid media already has 0.2 nM copper ion concentration (present in the Starkey solution, see Table 2 - 2.1). All concentrations were corrected for that initial value.

2.2.2 Cadmium and Cobalt ion stress

The effect of cadmium and cobalt ion in the *Ma.aq* 617 growth was also studied using solid and liquid ASW medium. The stress was induced by adding $CdCl_2$ and $CoCl_2$, respectively, to the growth medium. The addiction moment was the same as in 2.2.1, growth and harvesting followed the procedure described in 2.1. Cadmium and cobalt ion concentrations tested were for: $CdCl_2$ – 0, 0.3, 1, 10, 30, 50, 75, 100, 150, 160, 170, 190, 200, 210, 220, 230, 240, 250 μM; and for $CoCl_2$ - 0.75 nM; 0.1, 0.25, 0.30, 0.50, 0.75, 1.0, 1.5, 2.0, 4,0, 6.0, 8.0, 10.0, 15.0 mM. A control growth, in ASW medium, was also performed alongside these two heavy metal ions stress growths.

All concentrations were streaked in triplicate and in liquid medium concentrations tested at least twice. Although some of the growth curves are not shown the patterns obtained were identical. The reason for this is related with different starting inoculums percentage invalidating the direct comparison (the same happened in the copper stress 2.2.1). OD_{600nm} values presented are directly registered values because cadmium and cobalt ion do not contribute to the absorption at 600 nm.

It is important to mention that the standard ASW solid or liquid media already has 0.75 nM cobalt ion concentration (present in the Starkey solution, see Table 2 - 2.1). All concentrations were corrected for that initial value.

2.3 Preparation of spheroplasts

Spheroplasts were obtained using the modified Goodhew, *et al.* and Hunter, *et al*. procedure [128, 129]. The cell pellet from any of the growths were resuspended in 50 mM Tris-HCl buffer pH 7.6 and stored at -80°C. For the preparation of the spheroplasts the cells were thawed at room temperature (RT), and diluted 1:5 (v/v) in cold distilled water and

incubated at RT for 30 min. Centrifugating the cell suspension for 30 min, 12000 rpm at 4°C in Sigma centrifuge model 3K30 to separate the periplasm from the spheroplasts. The spheroplasts were resuspended in 10 mM Tris-HCl buffer pH 7.6. Both fractions were storage at -80°C and samples were taken and kept frozen at the same temperature for SDS-PAGE and 2D electrophoresis. DNAse (Sigma) was added to this fraction whenever needed (to avoid viscosity due to the presence of DNA).

3. Sodium dodecyl sulphate – Polyacrylamide gel electrophoresis (SDS-PAGE)

The SDS-PAGE method was performed according to the Schagger and von Jagow System , with a slightly alteration of using 2X Tris-Tricine-SDS cathode tank buffer with the standard Laemmli gel/buffer system [130]. This method was based in acrylamide/bisacrylamide tris-trycine gels. Typically the gels were 0.75 mm 12.5% acrylamide/bysacrylamide proportion, with the polymerization agents, ammonia persulphate (APS) and N - N - N'- N' – tetrametil - 1,2 diaminemetane (TEMED). The polymerization agents are the last to be added to the solution. Samples were prepared in loading buffer (0.1 M Tris-base pH 6.8, 0.25 M β-mercaptoethanol, 2.6 M glycerol and $2.9x10^{-3}$ M bromophenol blue) and denaturation was carry out by heating samples to 100°C for 5 min before loading into the gel.

Usually the gels were prepared on the day before its use in order to allow peroxides and other radicals that could affect the normal running of electrophoresis to decay. Electrophoresis was performed at 150 V for 75 min. The anode buffer was 0.2 M Tris-HCl at pH 8.9, and cathode buffer was 0.2 M Tris- HCl, 0.1 M Tricine and 0.1% SDS at pH 8.25.

3.1 Staining procedure

Gels were stained for the presence of protein. The staining used was Coomassie blue, Coomassie Blue R-250 5 g/L in a 45 % methanol, 7.5% acetic acid solution, which has a detection limit of nmol (amount of proteins). Gels were left staining for approximately 30 min, and distained in 45 % methanol, 7.5 % acetic acid until coomassie blue background disappeared from gel. The distaining solution is substitute for fresh solution to fasten the procedure. Both solutions can be reused; distaining solution is recycled with activated charcoal.

4. Determination of protein concentration

4.1 BCA method

The total protein content was determined using the Bicinchoninic Acid Protein (BCA) methodology, commercially available in a kit from Sigma. Procedure was followed according to the commercial protocol [131], with some modifications, using bovine serum albumin (BSA) as a standard protein (from 0 to 300 µg/mL). The modifications consisted in using 1 mL of BCA working reagent, and 50 µL of sample and standard proteins. Samples were left to incubate at 37°C during 30 min, and the OD_{562nm} was measured in a UV-spectrometer Shimadzu UV-160A. A typical calibration curve obtained with BSA is showed in Figure 12.

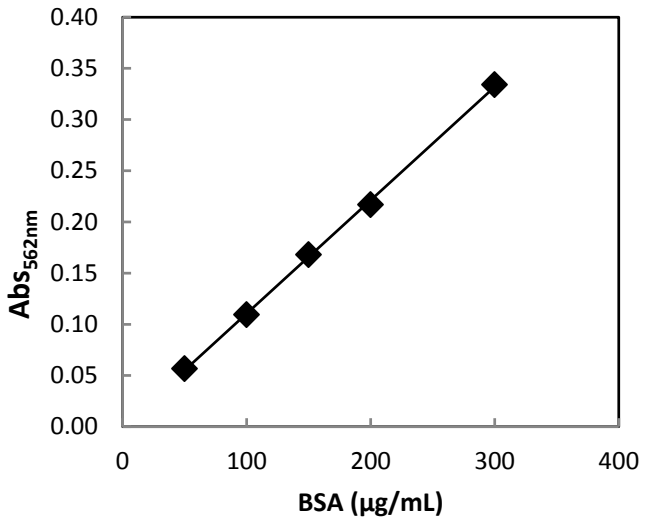

Figure 12 – Calibration curve used in the BCA method for total protein quantification. Calibration curve obtained was Abs 562 = 0.001 [Protein] +0.0107 with an R value of 0.996 (the zero value was not used for equation determination). Similar curves were obtained in other assays. The standard protein was BSA.

This assay was also used to determine the total protein concentration of the samples used in the SDS-PAGE normalizations, in the proteomic analysis, and also in the biochemical characterization of CopX.

5. Proteomics analysis

5.1 Isoelectric focusing

The protein concentration of the periplasmatic fractions obtained from the growths under copper stress (1.0 mM $CuSO_4$) and from the control (0.2 nM $CuSO_4$) were determined using the BCA procedure described in 4.1. The samples were re-hydrated for sample solubilization using a sample buffer constituted by 7 M urea, 2 M thiourea, 2% CHAPS, and 1% DTT. IPG *ImmobilineDrystrip* (GE Healthcare) used were pH 3-10 non-linear gradient (NL) 13 cm and pH 4-7 linear gradient, 7 and 13 cm. Re-hydration solution applied to the Dry strips were supplemented with 0,5% of the corresponding pH IPG buffer and bromophenol blue to help the isoelectric run visualization.

At least 2 h before loading in the IPG strip the samples had to be incubated with the re-hydration solution containing 0.28% DTT for protein denaturation and reduction at 4°C, before applying using the Reswelling Tray for *ImmobilineDrystrip* (GE Healthcare). Volume ratio and concentration was used according to the 2D electrophoresis handbook from GE Healthcare for size and staining procedure. The applied gradient was the one recommended by GE Healthcare [132].

5.2 2D electrophoresis

SDS-PAGE gels were prepared using the Laemmli method [133, 134], only with two modifications, 1% of SDS and no stacking gel is prepared. This method was based in acrylamide/bisacrylamide tris-glicine gels. Typically the gels were 0.75, 1.0 or 1.5 mm 12.5% and 10 % acrylamide/bysacrylamide proportion, with the polymerization agents, APS and TEMED. The gel components are the same as in 3.SDS-PAGE, only differing in 40% Acrylamide/3% Bisacrylamide, and SDS quantity already mentioned. The electrophoresis buffer used was only one (for anode and cathode) from the anode buffer in 3.SDS-PAGE. Before applying the IPG strip to the gel, two equilibration steps were performed, the first was the use of 10 mg/mL of DTT for 10 min, and the second 25 mg/ mL iodoacetamide for 10 min, to reduce and alkylate the proteins for SDS-PAGE run. The gradient was applied at 15°C, according to the 2D electrophoresis handbook from GE Healthcare [132].

Knowing that protein markers tend to have migration problems associated with loading difficulties in 2D electrophoresis (using either loading papers or well loading).

Electrophoretic runs were performed to determine the relative mobility of proteins markers, using gels with the same composition and size (data not shown).

5.3 Staining Procedure and Gel analysis

Silver staining and Coomassie Blue were used for gel staining. For qualitative analysis, Silver staining method was used according to the protocol from Pharmacia (PlusOne Silver Staining kit) [135]. For quantitative analysis, Coomassie blue was used, since it is (more) compatible with mass spectrometry analysis, and in this case EZBlue from Sigma was used, following the protocol of the manufacter [136].

Scanning of the gels for analysis was performed using the GE ImageScanner. Gel analysis was performed using GE ImageMaster 7.0 [137] and Ludesi REDFIN3 [138]. Periplasm protein position was determined using Jvirgel2.0 [139, 140].

5.4 Mass spectra spot analysis

Spots were removed manually and sent to the Laboratory of Protein Biochemistry and Protein Enginneering (L-ProBE) in State University of Gent, Belgium. Samples were analyzed by Dr. Bart Devreese, who performed MALDI-TOF-TOF-MS in a 4700 Proteomics analyzer (Applied Biosystems).

6. Molecular biology techniques

Molecular biology procedures were carried out following the manufactures instructions, with some modifications, which are indicated in each case.

6.1 Preparation of *CopX* cloning

6.1.1 Genomic DNA extraction

For the heterologous expression of gene maqu_0125 from *Ma.aq* 617, named *copX*, the genomic DNA was extracted from a growth harvested after 24 h, performed following the procedure described in 2.1 and stop after 24 h of step three. DNA extraction was performed using a distilled water boiling method [141]. The cells were resuspended in

distilled water and incubated for 15 min at 100°C. The DNA sample was stored at -20°C until further use. For the PCR amplification 1µL of this genomic DNA suspension was used directly.

6.1.2 PCR amplification and primers used

The gene *copX* was cloned with its signal peptide in order to direct the protein to the periplasmatic space of *Es.coli*. The gene with 534 bp, and located in the *Ma.aq* 617 chromosome between 152394-152927 nucleotide position, was amplified by PCR (Biometra) using a HighFidelity DNA polymerase (Fermentas). The PCR reaction mixture comprised, 1X PCR buffer, 0.2 mM dNTP'S, 50 pmol of primers (Forward and Reverse), 1.25 U DNA polymerase, and 1 µL of genomic DNA fraction.

The PCR program was the following: for the DNA denaturation 1 min, at 94 °C, for the annealing of the primers 50 °C during 1 min and 2 min at 72°C for DNA extension, followed by 35 cycles. Primers used for amplification of *copX* were produced by SigmaGenosys, and are showed in Figure 13. The sequence of the gene *copX* is presented in Appendix D (Figure 7 App.D.1).

*Nde*I

CopX Forward Primer - 5' T AAG TAA **CAT**│ATG AAT GCA TCA AAC CTG 3'

*Not*I

CopX Reverse Primer - 5' TTA CGA ATT CGC│GGC CGG **TTA** TTC GAA ATT CAC GTC 3'

Figure 13 – Primers used for the amplification and isolation of *copX***.** The forward primer presents a melting temperature (Tm) of 50°C, and the reverse a Tm of 48°C. Legend: Restriction sites, showed underlined were included in the primers to enable the cloning into the expression vector. Nucleotides in bold correspond to the initiation codon, forward primer, and STOP codon, reverse primer. Strikes indicate the restriction nuclease action, in the forward primer *Nde*I and in the reverse primer *Not*I.

PCR products were purified using the modified QIAquick PCR Purification Kit Protocol (QIAGEN) [142]. Instead of using the PB buffer, as described in the protocol for step 1 and step 5, it was used QC buffer from the QIAquick Gel Extraction Kit (QIAGEN) [142].

6.1.3 Agarose gels

DNA electrophoresis was performed to check the purity, PCR amplification and verification of complete restriction enzyme reaction. Typically, 0.8% agarose gel was used to check the plasmid DNA and 1% agarose to check the PCR products. The agarose gel was prepared in the running buffer 1X TAE (0.04 M Tris-HCl, 0.02 M acetic acid (100%) and 0.001 M EDTA at pH 8.5). Electrophoresis was run at a constant voltage 100 V, during 20 min. The samples were loaded using Loading solution (Fermentas). Visualization of the electrophoresis results was done using an Ethidium Bromide free method, Sybr Safe (Invitrogene) prepared in 1X TAE buffer [143]. This method allows the visualization of the DNA by nucleotide binding, and the staining agent has fluorescence excitation maxima at 280 and 502 nm, and an emission maximum at 530 nm. Therefore, the gel was left in the staining solution for 30 min and was visualized using a blue-light transiluminator (Invitrogen - S3710).

6.2 Construction of over-expression plasmid

6.2.1 Double digestion

The *copX* gene was cloned into the expression vector pET 21-c(+) (5441bp) [144], Appendix D.2 (Figure 8 App.D.2). For that, both the vector and the pure PCR product (not quantified, 15 µL were used) were doubled digested with 10 U of the enzymes *Not*I (Invitrogen) (at the C-terminal) and *Nde*I (Invitrogen) (at the N-terminal), in 1X React3 buffer (Invitrogen). The reaction mixtures were incubated during 4h at 37°C.

After digestion, the vector was dephosphorylated with 1 U of alkaline phosphatase (Promega) in 1X Multicare Buffer, during 30 min at RT, and deactivated at 65°C during 15 min. Digestion of both vector and PCR product were check in a 0.8% agarose gel and purified from solution using the QIAquick PCR Purification Kit according to the manufactor's Protocol (QIAGEN) [142], with the same modification as described in 6.1.2.

6.2.2 Ligation

The amplified DNA insert*, copX*, was cloned into pET21-c(+), in the multiple cloning site, between *Nde*I (located at 238 bp) and *Not*I (located at 166 bp). The ligation reaction was performed using T4 DNA ligase (Roche), according to the manufactors instructions,

during 15 min at RT. After incubation the resulting vector, pETCopX (5975bp), was transformed into a non-expression host, *Es.coli* NovaBlue (Novagen) [145].

6.2.3 Transformation

The protocol of transformation was the one provided by the company [145], using 4 μL of the ligation mixture, with the remaining being stored at -20°C. Transformation was plated in Luria-Bertani (LB) broth [146] agar plates supplemented with 100 μg/mL ampicilin. Plates were incubated overnight at 37°C.

6.2.4 DNA sequencing

From the colonies obtained, 10 were selected to isolate its plasmidic DNA and be sent for DNA sequencing. Each of these colonies were used to inoculate 5 ml of LB medium supplemented with 100 μg/ml amplicillin, and were incubated overnight in an orbital shaker (230 rpm), at 37°C. The plasmid was isolated using the Miniprep Kit (Nzytech), following the procedure of the company [147]. The isolated plasmid DNA was eluted with 30 μl of AE (10 mM Tris-HCl, pH 8.0) buffer, in order to obtain a more concentrated plasmid DNA sample.

The purity and integrity of the plasmids were verified in a 0.8% agarose gel. The presence in the isolated expression plasmids, of the fragment corresponding to *copX* gene, was verified by double digestion with the enzymes used in the cloning *Nde*I (Fermentas, 10U) and *Not*I (Fermentas, 10U), in 1X Tango buffer, and verified in a 1% agarose gel. The isolated expression vectors (3 out of 6 that contained) *copX* were sent for DNA sequence verification. Sequencing was performed by StabVida (www.stabvida.net).

6.3 Production of recombinant CopX in *Es.coli*

6.3.1 Transformation

One of the positive clones was transformed into the expression host *Es.coli* BL21(DE3) (Nzytech). Transformation procedure protocol was the one provided by NzyTech [148], in which incubation time was reduced by half. The transformed cells were plated in Luria-Bertani (LB) broth [146] agar plates supplemented with 100μg/mL ampicilin. Plates were incubated overnight at 37°C.

6.3.2 Growth conditions

Production of the protein was performed in two different growth media: a rich medium, LB medium supplemented with 0.5 mM $CuSO_4$, and a minimum medium, a modified version of M9 medium (1.0 g/L NH_4Cl, 3.0 g/L KH_2PO_4, 6.0 g/L $Na_2HPO_4.7H_2O$, 1.2% Glucose, 1 mM $MgSO_4$, 0.5 g/L NaCl, 0.01 mg/mL B1 Vitamin, 0.01 mM $CuSO_4$, 18.5 µM $FeCl_3$ $6H_2O$ and 0.1 mM $CaCl_2$) [149], in order to uniformly label CopX with ^{15}N for the preliminary NMR studies. Both media were supplemented with 100 µg/mL of ampicilim. Before large scale production of the protein, small scale tests (50 mL) were performed for both media. The tests consisted in different incubation times (0 to overnight) for both media using 1 mM IPTG to induce the transcription of the cloned gene.

In the growths in rich medium, the $CuSO_4$ was added at different growth times: no addition of $CuSO_4$, addition in the beginning of the growth or at the time of induction. The conditions were incubation in an orbital at shaker (Infors HT- AK62) at 210 rpm, 37°C, for 7h. The gene expression was induced with 1 mM IPTG, when the culture OD_{600nm} reached 0.6 - 0.8. Samples (1 mL) were taken during the growth to measure the OD_{600nm} and observe the protein CopX expression in a SDS-PAGE. The samples collected were normalized to an OD_{600nm} of 1.2, assuming the loading of 20 µL samples.

In order to uniformely label CopX with ^{15}N, *Es.coli* BL21(DE3) carrying the expression vector pETCopX, was grown in M9 minimum medium with the same composition as before, but adding 1 g/L ^{15}N (Cambridge Isotope Laboratories). The induction conditions were the ones described before.

The recovery of the wet cell weight for the two mediums was: M9, 4.3 g/L; LB – no copper added, 11.5 g/L, LB – copper added in the inducing moment, 10.5 g/L, LB – copper added in the inoculum moment, 11.0 g/L.

6.4 Cell fractionation

Cells were harvested by centrifugation using a Beckman centrifuge, model Avanti J-25 (rotor JA-10), at 8000 rpm, 10 min and 4°C. The cell pellet was resuspended in 50 mM Tris-HCl buffer pH 7.6, in a 2 mL of buffer/1g of wet cell weight. The cell suspension was stored at -80°C until further use. Since the protein was cloned with its own signal peptide to direct it to the *Es.coli* periplasm, spheroplasts were prepared by inducing a termic shock to the cellular fraction as a method to recover the recombinant CopX. So, the frozen cellular

fraction was thawed (RT) and frozen (-80ºC) 5 times. Protease inhibitors cocktail (Roche Complete EDTA-free) was added after the first thawing cycle at a concentration of one complete tablet for 50 mL extraction volume. The periplasmic extract was obtained by centrifugation at 6000 using a Sigma centrifuge 3K30.

The soluble cellular extract was obtained by breaking the spheroplasts, obtained previously, by three passages in a French Press (Thermo-FA-080A using the 40K cell) between 10000 and 15000 psi. Cellular debris were separated from the soluble cellular extract by ultracentrifugation using Beckman (Optima L-70) at 43000 rpm (rotor 45Ti), for 1h at 6°C. The cell debris pellet was resuspended in 10 mM Tris-HCl pH 7.6 and stored at -80ºC until further use. A small aliquot was kept aside for SDS-PAGE. DNAse was added before passing through the French-Press to reduce the viscosity of the sample.

The presence of CopX in the periplasm, soluble cellular extract and cell debris pellet was check by SDS-PAGE. The soluble fractions that contained CopX were combined and used in the purification.

7. Protein purification

The recombinant CopX was purified from the periplasmic fraction or from the combined periplasmic and soluble cellular extract obtained as described in 6.4. The purification started by an anion-exchange chromatography using a Source-Q 15 column (GE HealthCare) (2.6 cm diameter (Ø) x 23 cm). This column was equilibrated before sample loading with 10 mM Tris-HCl pH 7.6, with 3 - 5 total column volumes. After, the sample was loaded onto the columun and the unbound proteins were eluted using the same buffer. The elution gradient used was a linear gradient in 10 mM Tris-HCl pH7.6, from 0 to 500 mM NaCl at 2.0 mL/min during 110min. The fractions containing this protein were collected and concentrated using a diaflow with an YM3 44.5 diameter membrane (Millipore), with a molecular weight cut off of 3000 Da, up to 5 mL.

The concentrated fraction was loaded onto a gel filtration column Superdex 75 (GE HealthCare) (1.6 cm Ø x 70 cm), equilibrated with 50 mM Tris-HCl, 150 mM NaCl, pH 7.6 and eluted in the same buffer at 1.5 mL/min. The selection of the tubes containing CopX was as described before for the first chromatographic step. The tubes containing CopX with the

higher purity index were concentrated using a VIVASPIN 4 (Sartorius Stedim Biotech), with a cut off of 5000 Da.

In some cases, another chromatographic step was required. In this case, another anion-exchange chromatography using a pre-pack ResourceQ 6 mL or 1 mL was used. The column was equilibrated before sample loading with 20 mM Tris-HCl pH 7.6 and the protein was eluted using a linear gradient in 20 mM Tris-HCl pH7.6, from 0 to 500 mM NaCl at 1.0 mL/min. In Figure 14, a flowchart of the CopX purification is shown.

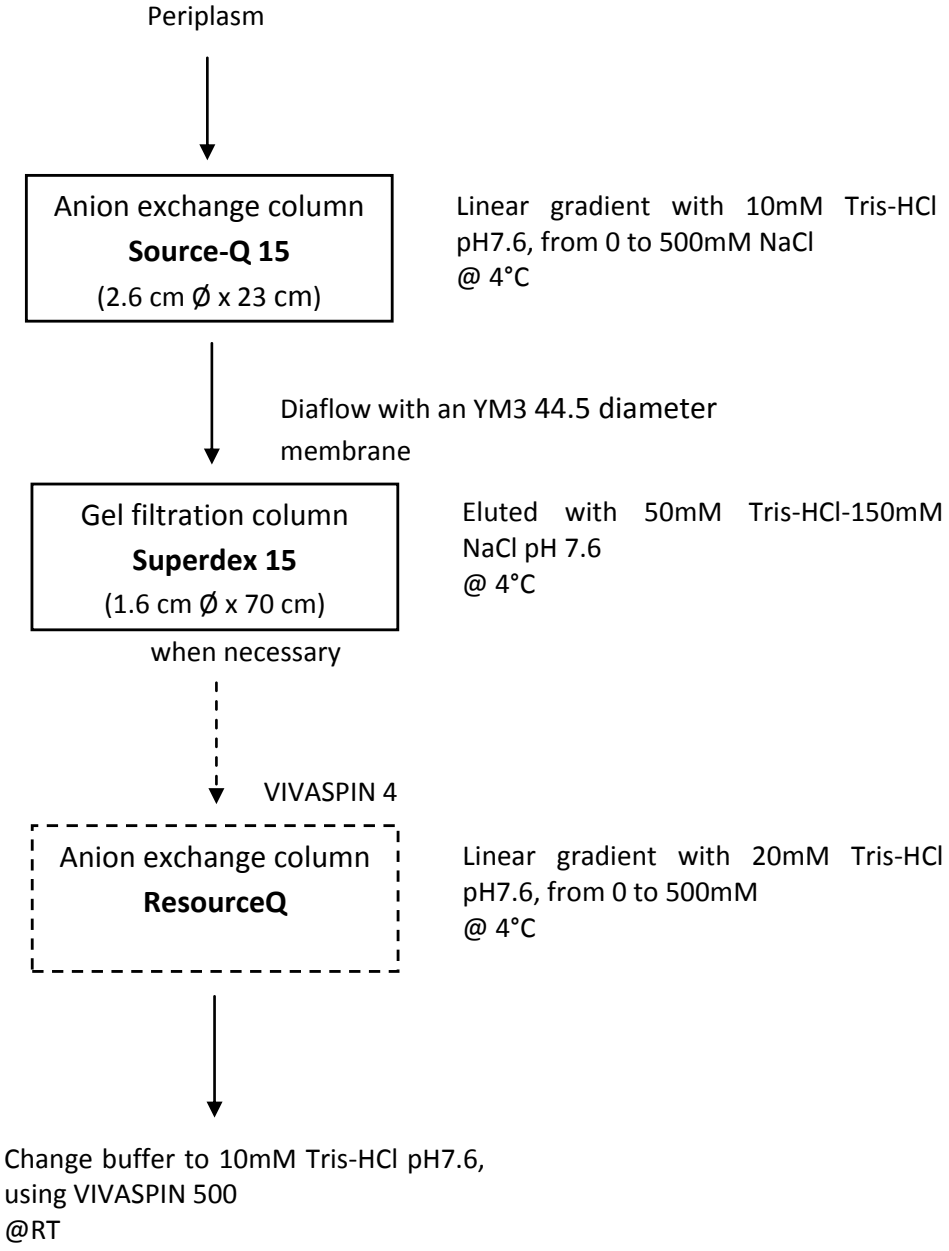

Figure 14 – Purification flowchart of CopX. Step 3 depends on purity ratio of CopX after gel filtration column.

The pure CopX was concentrated using a VIVASPIN 500 (Sartorius Stedim Biotech), with a cut off of 3000 Da.The buffer of the pure protein was exchanged to 10 mM Tris-HCl, pH 7.6 and stored at -80°C.

All elution buffers were prepared in deionized H_2O and filtrated before use. Purification was performed at 4°C. Fractions containing the CopX during purification steps were kept at 4°C.

The fractions throughout the purification were monitored at 280nm (for protein content). SDS-PAGE was used to select the fractions containing CopX. The UV-visible (Shimadzu UV-160A) spectrum of the resulting samples was performed. CopX was obtained in the reduced form, so samples were oxidized by potassium ferricyanide.

Heterologous expressed CopX is in the reduced state. In this form, it was not possible to visualize the protein in the column. A way to change this was to oxidize the loaded sample before the purification, creating a slightly blue ring

7.1 Protein sequencing – N-terminus

The heterologously obtained CopX was identified by N-terminal analysis by a protein sequencer (Applied Biosystems - LC 491).

7.2 Preparation of samples for mass spectra

Purified protein was sent to the Laboratory of Protein Biochemistry and Protein Enginneering (L-PROBE) in State University of Gent, Belgium. Samples were analyzed by Dr. Bart Devreese, who performed Mass Spectrometry analysis. Electrospray mass spectrometry was performed on a 4700 Proteomics analyzer (Applied Biosystems). Purified 300 µM CopX, in 10 mM Tris-HCl pH7.6 was sent for analysis. The protein was freeze-dried for transport.

8. Biochemical characterization

8.1 Spectroscopic techniques

8.1.1 UV-Visible spectroscopy

The UV-visible spectra were acquired on a Shimadzu UV-1800 spectrometer. All spectra were recorded at RT using 1 cm path lenght quartz cuvettes between 250 nm and 900 nm with absorbance being measured each 0.5 nm.

The heterologous expressed CopX samples were diluted in 10 mM Tris-HCl pH 7.6. The oxidized spectrum was obtained by adding aliquots of a potassium ferricyanide solution, until full oxidation was achieved. The reduced spectrum was obtained by adding aliquots of a 1.0 mM sodium ascorbate solution, untill full reduction was achieved.

For the cobalt cellular fractions the reducing agents used were sodium dithionite and sodium borohydride (data not shown), and all samples were normalized by total cellular content.

8.1.2 Nuclear Magnetic Resonance

8.1.2.1 Sample preparation

Reduced ^{15}N-labeled CopX was prepared in sodium phosphate buffer at pH 7.6, 7.0 and 6.0, containing 10 mM sodium ascorbate and 10% D_2O. The pH values of the NMR samples were not corrected for the deuterium isotope effect.

8.1.2.2 Data acquisition and processing

All spectra were acquired on a Bruker AvanceIII 600 NMR spectrometer equipped with a TCI cryoprobe. NMR spectra were processed with Topspin software provided by BRUKER. The water was removed by a Watergate pulse. The ^{1}H-^{15}N HSQC were acquired at different temperatures, 287 K or 293 K. Spectral widths were for ^{1}H 14 ppm and for ^{15}N 45 ppm. For the data analysis, Computer Aided Resonance Assignment (CARA) software [150] was used.

8.1.3 Electron Paramagnetic Resonance

The untreated EPR spectrum of CopX was recorded on an X-Band Bruker EMX spectrophotometer equipped with an Oxford Instruments liquid helium flow cryostat.

Experimental conditions were: temperature 70K, 1 scan, microwave frequency 9.65 GHz, gain 10^5 and modulation amplitude 30 mT. The simulation and data was obtained using WIN-EPR software (version 2.11), from Bruker. The sample was 0.83 mM CopX in 10 mM Tris-HCl pH 7.6.

8.2 Copper and Zinc determination

The total amount of copper and zinc of CopX was determined by inductively couple plasma (ICP) (Ultima - Jobin-Yovn). Samples used in the quantification were diluted in 10 mM Tris-HCl. The total copper and zinc content of the dilution buffer was also determined.

8.3 Determination of the extinction coefficient of CopX

The extinction coefficient was determined using the fully oxidized UV-visible spectrum of CopX. The protein concentration was determined using the BCA method, procedure in 4.1, and the copper concentration obtained by ICP (procedure described in 8.2). Extinction coefficient of CopX was determined for total protein and for total Cu.

8.4 Gel Filtration

8.4.1 Calibration of size-exclusion column

A gel filtration chromatographic column, a pre-pack Superdex 75 10/300, was used to determine the apparent molecular weight of CopX at different NaCl concentrations. The column was equilibrated with 50 mM Tris-HCl, 0, 50, 150, 300, 500, and 1000 mM NaCl, pH 7.6. The column was calibrated using the Low Molecular Weight (LMW) Gel Filtration Calibration Kit (GE Healthcare), following the procedure provided with the kit [151]. One of the chromatograms obtained is presented in Figure 15.

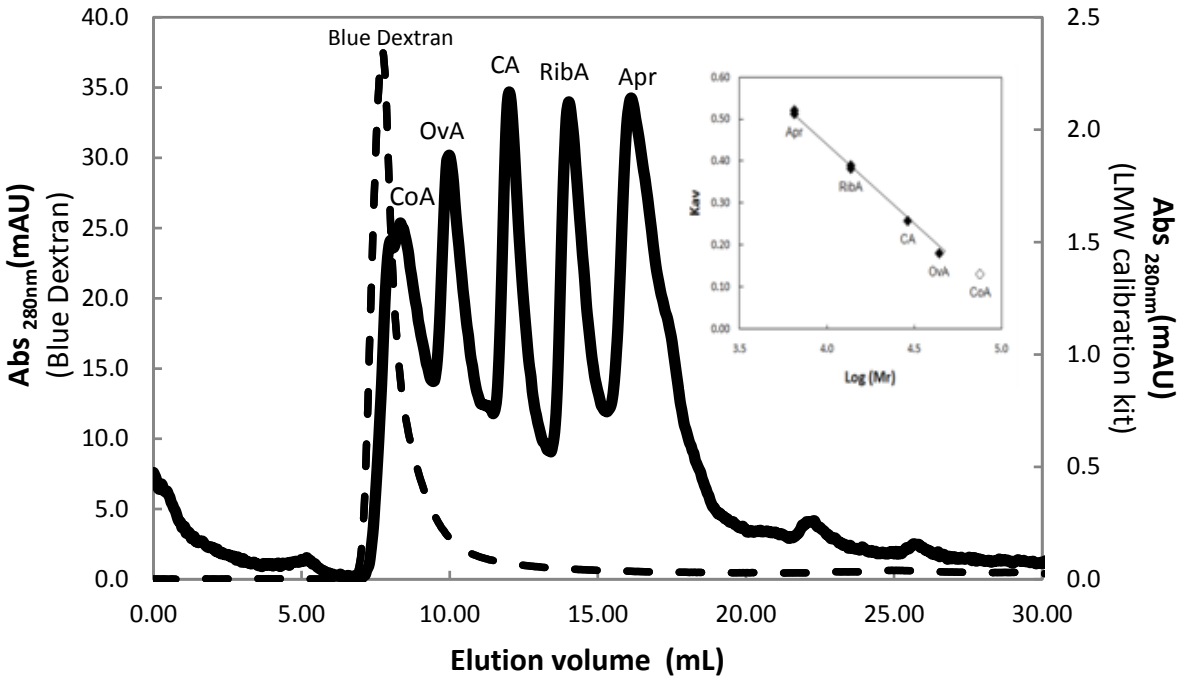

Figure 15 – Elution profile of LMW Gel Filtration Calibration Kit from a Superdex 75 10/300 GL. Column eluted with 50 mM Tris-HCl-150 mM NaCl pH 7.6 at RT. Legend: Apr – Aprotinin, RibA – RibonucleaseA, CA – Carbonic Anhydrase, OvA – Ovalbumin, CoA – Conalbumin, Eluation profile of Blue Dextran (Dashed line), and combined Mixture A and B (Solid line). See LMW Gel Filtration Calibration Kit protocol for details [151]. **Insert** - Calibration curve for the molecular weight determination using LMW Gel Filtration Calibration Kit. Kav = - 0.3771 Log (MW) + 1.9482. CoA was discarded as its molecular weight is higher than the higher molecular weight limit of Superdex 75 10/300 (70000 Da). Double Apr and RibA points corresponds to the same proteins in MixtureA and MixtureB for calibration of Superdex 75 10/300 GL.

8.4.2 Molecular exclusion chromatography

8.4.2.1 Molecular weight determination and Ionic strength equilibrium

The calibration curve is obtained by representing the partition coefficient that is the fraction of the stationary phase that is available for diffusion of a given molecular species (kav) as a function of the log apparent molecular weight (MW), as shown in the insert of Figure 15.

A calibration curve was obtained for the different NaCl concentration for the ionic strength equilibrium study (the same concentrations as in 8.4.1). CopX was equilibrated in the same buffer as the one in which the column was equilibrated. All calibrations were similar to the one presented in the insert of Figure 15.

RESULTS AND DISCUSSION

1. Bioinformatic Analysis

The increased number of complete sequenced genomes uncovered a series of different molecular systems that might be involved in heavy-metal homeostasis. Even if, these annotations are based on similar motifs and sites encounter in protein families, the Microbiology Societies estimates that 5-10% of the gene annotations are incorrect, and about 40% are unspecificly annotated [152]. Thus, although experimental validation is always required for confirming positively the annotation, other bioinformatic methods rather than those from database annotation can be used to validate an experimental strategy.

1.1 Copper operon analysis

The putative chromosomal copper resistance operon in *Ma.aq* is *copSRXAB*. This system is proposed to be regulated by a two-component system, *CopSR* located upstream the putative copper operon, *copXAB*, and annotated as an unspecific heavy-metal regulator (gene *maqu_0123* and *maqu_0124* in *Ma.aq* VT8 genome). Moreover, the operon *copXAB* presents a gene, *copX* annotated as coding for a copper-binding proteins (gene *maqu_0125* in *Ma.aq* VT8 genome) with no relation with known proteins of copper resistance systems.

The bioinformatic analysis of the *Ma.aq* copper operon was divided into two parts. The first consisted in the analysis of the two-component regulatory system, CopSR. This was performed by protein sequence analysis and phylogenic distance studies of the regulator (CopR) with known related heavy metal regulators. The second was the analysis of the copper-binding protein, CopX, which consisted in the analysis of the protein sequence by alignment with known copper binding proteins and copper proteins associated with resistance, and also by phylogenic distances studies.

1.1.1 Two-component regulatory system, CopR and CopS

The two-component regulatory family regulates different molecular systems that respond to stress events, such as temperature shock, antibiotics, heavy-metal ions (our

study). Taking into account that usually, but not always, there is a two component system regulating copper resistance operons, as mentioned in Introduction, Chapter 1, a bioinformatic study was conducted to assess the hypothesis that CopRS is involved in the regulation of CopXAB. The genes that code for these proteins are encoded by *copRS*, located upstream of *copXAB* and transcribed in the opposite direction. In these studies, the protein sequence of CopS and CopR was used, since the homology at the DNA level is usually much smaller. The DNA sequence alignments of *copR* and *copS* can be found in Appendix C, Figure 2 App. B.1(B) and 4 App.B.2(B), and illustrate this idea.

First a sequence alignment was performed with the more related sequences (Figure 16), showing that CopR presents, as expected two domains, the receiver (REC) domain between residue 1 and 112, and the transcriptional regulatory domain (DNA-binding site) between residue 146 and 217. This alignment identifies the most conserved residues and sequence motifs in *Ma.aq* CopR sequence.

```
Ps.putida(CopR)          ----------------MKLLVAEDEPKTGAYLQQGLAEAGFTVDRVLTGT  34
Ps.stutzeri(CopR)        MRRCHTCAVHHRWNIRMKLLVAEDEPKTGTYLQQGLSEAGFTVDRVTSGT  50
Ps.aeruginosa(CopR)      ----------------MKLLIVEDEPRIGQYLRQGLAEAGFAVDLSDDGN  34
Es.coli(PcoR)            ---------------MQRILIVEDEQKTGRYLQQGLVEEGYQADLFNNGR  35
Ma.aq(CopR)              ----------------MRLLLVEDDRLLAEGLVRQLEKAGFSIDHTSSAR  34
                                        ::*:.**:   .  * : * : *:  *    .

Ps.putida(CopR)          DALQHALSESYDLLILDVMMPGLEGWEVLRMVRAAGKDVPVLFLTARDGV  84
Ps.stutzeri(CopR)        DALQHVLSVPYDLLILDVMMPGLDGWEVLRLVRASGNEVPVLFLTARDRV 100
Ps.aeruginosa(CopR)      EGEQLALGGDYDLLILDVMLPGRDGWQILRSVRDAGMTVPVLFLTARDAV  84
Es.coli(PcoR)            DGLGAASKGQYDLIILDVMLPFLDGWQIISALRESGHEEPVLFLTAKDNV  85
Ma.aq(CopR)              EAQILGEQEDYRAAVLDLGLPDGNGLEVLKRWRSKLVECPVLVLTARGDW  84
                         :.        *   :**: :*  :* :::   *      ***.***:.

Ps.putida(CopR)          DDRVKGLELGADDYLIKPFAFSELLARVRTLLRRGTSTAAQTTMKMADLE 134
Ps.stutzeri(CopR)        EDRVKGLELGADDYLVKPFAFSELLARVRTLLRRGNATALQTQLQVADLQ 150
Ps.aeruginosa(CopR)      EDRVRGLEQGADDYLVKPFAFVELLARVRTLLRRGSQQLQETTLQLADLE 134
Es.coli(PcoR)            RDKVKGLELGADDYLIKPFDFTELVARVRTLLRR-ARSQAATVCTIADMT 134
Ma.aq(CopR)              QDKVNGLKAGADDYLAKPFQTEELIARINALIRR-SEGRVHSQVKAGGFE 133
                          *:*.**: ****** **   **:**:.:*:**       :    ..:

Ps.putida(CopR)          VDLLKR-RATRSGKRIDLTAKEFSLLELLMRRRGEVLPKSLIASQVWDMN 183
Ps.stutzeri(CopR)        VDLLKR-RATRAGQRIDLTAKEFALLELLLRRRGEVLPKSLIASQVWDMN 199
Ps.aeruginosa(CopR)      LDLLRR-RVQRQGKRIDLTAKEFALLELLLRRSGEVLPKSLIASQVWDMN 183
Es.coli(PcoR)            VDMVRR-TVIRSGKKIHLTGKEYVLLELLLQRTGEVLPRSLISSLVWNMN 183
Ma.aq(CopR)              LDENRQSLRTEEGAEHALTGTEFRLLRCLMSRPGHIFSKEQLMEQLYNLD 183
                         :*  ::    . * .  **..*: **. *: * *.::.:. : . :::::
```

```
Ps.putida(CopR)        FDSDTNVIEVAIRRLRAKIDDDFAPKLIHTARGMGYMMDAPE-------- 225
Ps.stutzeri(CopR)      FDSDTNVIEVAIRRLRAKIDDGFEPKLIHTARGMGYMLDEPDPS------ 243
Ps.aeruginosa(CopR)    FDSDTNVIEVAIRRLRAKVDDDYPQRLIHTVRGMGYVLEERDE------- 226
Es.coli(PcoR)          FDSDTNVIDVAVRRLRSKIDDDFEPKLIHTVRGAGYVLEIREE------- 226
Ma.aq(CopR)            ESPSENVIEAYIRRLRKLVGN----ETITTRRGQGYMFNAQR-------- 221
                         ... ***:. :****  :.:    . * * ** **:::
```

Figure 16 – Protein alignment of CopR with known copper resistance regulators. Alignment was performed using Clustal X (version 2.0) [118]. Legend: Green – Dimerization interface; Dark green - Intermolecular recognition site; Yellow – Phosphorylation site; Turquoise – REC site; Gray – DNA binding site. Asterisk – Similar residues; Colon – One residue is different; Dot – Two residues are different. For the accession numbers see Appendix B – Table 2 App.B.1(A).

Similarly, *Ma.aq* CopS protein sequence was analyzed in terms of domains and conserved sequence motifs, showing that there are three domains, the HAMP domain for histidine kinase (HK) between residue 165 and 233, the dimerisation and phosphoacceptor domain (found in HK) between residue 237 and 298, and the HK-like ATPases between residue 340 and 442. The sensor domain (between residue 37 and 164) presents low sequence homology to known metal sensors (Figure 17). In addition, a transmembrane helix is predicted between residue 163 and 185, separating the sensor domain from the HAMP domain, and it was also identified a signal peptide, between residue 1 and 34.

```
Ps.putida(CopS)        ---MILTRSSLVKRLTLMIMFAVIAVLVVAGISFNMLSQHHFRMLDEQAL  47
Ps.stutzeri(CopS)      -----MTRLSLTARLSLMFMLAVTGVLAAAGYLFHQLSERHFNELDQHTL  45
Ps.aeruginosa(CopS)    MSAGFGSRMSLGVRLSLLFAACTAAVSLIAGLIFSRAIDEHFVELDHMAM  50
Es.coli(PcoS)          ----MRFKISLTTRLSLIFSAVMLTVWWLSSFILISTLNGYFDNQDRDFL  46
Ma.aq(CopS)            -MPNVKRPASVKGMLLVLLLPAGIALMGVAWFVHGLLLDRMSREFLESRL  49
                                *:   * :::      :   :        :       .  :

Ps.putida(CopS)        SEKLESTRHILS--IQAPGTGMEELRPQLRALLGAHQDLTAEILTSEGEV  95
Ps.stutzeri(CopS)      HEKLRAGEALLG--ELQADDDFERLRPRLQALLGGHSEIRGFVFDAAGRQ  93
Ps.aeruginosa(CopS)    SAKLAVFRDELR--GLGSEQQMRRREAELLRELARHPDLGLRLNGPDGNL  98
Es.coli(PcoS)          TGKLQLTEEFLKTETFRNKTDIKSLSEKINDAMVGHNGLFISIKNMENEK  96
Ma.aq(CopS)            KDEAAFLEHQIR--EVRGQVETLQTGDYFQDVFHHAFAIRTPDRTIISPK  97
                         :    .  :                 :   :     :        .

Ps.putida(CopS)        LFSDLKAVQIPDKYKRADKEEMWEWQDESH-------NFRGITTDVEIQG 138
Ps.stutzeri(CopS)      LFPEPLPGGADPSLVALAAERDGELSLDGR-------LWRAEVTQVPV-G 135
Ps.aeruginosa(CopS)    WFERLP----QPAHPGLPANRELGAPLEP--------------------G 124
Es.coli(PcoS)          IVELYAKNSVVPAVLLNKSGDILDYMIQTEENNTVYRSISRRVAVTPEQG 146
Ma.aq(CopS)            AWEPLLAPLINHEQNGTLRLEGRQAPDSPS---------DILAYRHSFQV 138
                                                .

Ps.putida(CopS)        QPAPGTVMLMLDITSHAHFFETLQRWFAIGLVISALVSAGIGWLVAKSGL 188
```

```
Ps.stutzeri (CopS)      QTAL-TLLILLDVTVHKAFFHTFSGWLWAALVLCALLSGLLGWLLVRSGL 184
Ps.aeruginosa (CopS)    NDASPRLTVILDISHHQHFLQRMRQLIWLTMSLSALATALLGAWATGASL 174
Es.coli (PcoS)          KSKHVIITVATDTGYHTLFMDKLSTWLFWFNIGLVFISVFLGWLTTRIGL 196
Ma.aq (CopS)            NGSPIVVVVSEDLEALKRSQAELHAWTAVVSVLLIALLVAVIWFGITLSL 188
                        :     : :  *          :                :       .*

Ps.putida (CopS)        RPVEQITKVATSISARSLQERIPLEPVPLELQTLILSFNGMLARLEDAFV 238
Ps.stutzeri (CopS)      RPLREVTQVAASVSAKSLRERIPDDSTPAELQQLVQAFNAMLARLEDAFV 234
Ps.aeruginosa (CopS)    APLRRMREVAARVSANSLTTRLDASRMPEELRGLAGELNAMLARLEEAFQ 224
Es.coli (PcoS)          KPLREMTSLASSMTVHSLDQRLNPDLAPPEISETMQEFNNMFDRLEGAFR 246
Ma.aq (CopS)            RPVVTLKAALKRLQDGEISR--INAPSPEEFQPLVMQLNQLLDSLDKRLV 236
                         *:  :      :   .:         * *:      :* ::  *:  :

Ps.putida (CopS)        RLSNFSADIAHELRTPVSNLLTHTEVVLTRKRDVDAYEENLYSNLEDLKR 288
Ps.stutzeri (CopS)      RLSNFSADIAHELRTPLSNLMTHTEVALTRGRSLEQYQDNLHSNLEELQR 284
Ps.aeruginosa (CopS)    RLSAFSADIAHELRTPLTSLLTQTQVVLSRPRSLEDYREALHGNLEELER 274
Es.coli (PcoS)          KLSDFSSDIAHELRTPVSNLMMQTQFALAKERDVSHYREILFANLEELKR 296
Ma.aq (CopS)            RSRDALANLSHSVKTPIAAVR-----QILEDMDRPLPSDLRIQMAARLSD 281
                        :     ::::*.::**:: :       : .  .     :        *.

Ps.putida (CopS)        MSRMIDDMLFLAKSDNGLIIPEQVDIQLHDLVSKLFEYYQLLADDRGIRL 338
Ps.stutzeri (CopS)      MSRMIDDMLFLAKADNGLIVPDAKPVALEQLCTQLLDYYQLPADERGVRF 334
Ps.aeruginosa (CopS)    LTAMVNDMLLLAKADHGLLAPSRQALDLGAEVDSLLEFYQPLAEDRDIRL 324
Es.coli (PcoS)          LSRMTSDMLFLARSEHGLLRLDKHDVDLAAELNELRELFEPLADETGKTI 346
Ma.aq (CopS)            IDRQLEAEMRRSRFAGPQVGKSAYPVKQARDLLWMLGRLYPEKS-FELSS 330
                        :    .  :  ::     :  .   :        :        .

Ps.putida (CopS)        TLQGKGVISGDRLMIDRAVSNILSNAVRYTPDGSGISVEIQQAADKVTLT 388
Ps.stutzeri (CopS)      ELSGAGTIQGDLLMLRRALSNLLSNALRYTPDGERIRVAIERGTGQVTLE 384
Ps.aeruginosa (CopS)    LREGSLSLPGDRGMLRRVLANLLDNALRFTADGGEIRIRLGDRR----LS 370
Es.coli (PcoS)          TVEGEGVVAGDSDMLRRAFSNLLSNAIKYSPDNTCTAIHLERDSDCVNVM 396
Ma.aq (CopS)            SLPEDTRWPIEEHDLNEVLGNLLDNAGKWSSR--CVELSLKQDNNSRQIV 378
                                  :   : ....*:*.** :::.      : :        :
Ps.putida (CopS)        IKNGGATIDPQHINKIFDRFYRADPARREGGPSNAGLGLAITRSIVKAHD 438
Ps.stutzeri (CopS)      VSNPGATIAPEHLERLFDRFYRVDPARREGSPSNAGLGLAITRSIVQAHQ 434
Ps.aeruginosa (CopS)    VENQGPAIPPERLPRLFDRFYRADPARREGQGEHAGLGLAICRSIVQAHG 420
Es.coli (PcoS)          ITNTMSGQVPANLERLFDRFYRADSSRFYN-TEGAGLGLSITRSIIHAHG 445
Ma.aq (CopS)            VSDDGPGVNGDDLSSLGQRGLRLDEQ-----TPGHGLGLAIVREIVARYE 423
                        : :  .      :  : :*  * *           ***:* *.*:  :

Ps.putida (CopS)        GKVWCTSSEGV-TAFNFIFPAAHRTAHIRTRH------------------ 469
Ps.stutzeri (CopS)      GRIHCTSAQGR-TSFVLEFPN---------------------------- 454
Ps.aeruginosa (CopS)    GEIRCESADGW-TRFVIDFARPRR------------------------- 443
Es.coli (PcoS)          GELSAEQQGRE-IVFKVRLLMD---------------------------- 466
Ma.aq (CopS)            GNISFSTGPGSGLRVTIEF------------------------------ 442
                        *.:           . . :
```

Figure 17 – Alignment of the protein sequences obtained by the BLAST query of CopS and known copper resistance sensor proteins. Alignment was performed using Clustal X (version 2.0) [118]. Legend: Green – ATP binding site; Dark green – G-X-G motif; Yellow – Mg^{2+} binding site; Turquoise – Phosphorylation; Red – Dimer interface. Underlined – Transmembrane region. Asterisk – Similar residues; Colon – One residue is different; Dot – Two residues are different. For the accession numbers see Appendix B – Table 4 App. B.2 (B).

The present alignments recognize the sequence motifs that classify these two proteins as a two-component system [36]. However, the homology is lower in the homologous sites between these proteins than the homology between Cop-like two-component regulatory systems from gram-negative bacteria.

The second part of the study consisted in a comparative analysis of different heavy-metal resistance regulatory systems present in Eubacteria that were classified as associated with resistance to different heavy metal ions, Cu, Zn, Cd, Ni and Pb [120] (in most cases this classification is based on experimental evidence).

This analysis was performed for the regulatory component, CopR, with more than one bioinformatic algorithm, including two far related outgroups, and was validated by bootstrap values. This strategy enables the better separation on the phylogenic inference. The results obtained are presented in Figure 18. All trees obtained bootstrap values above 80% and the different algorithms used presented the same resulting tree, indicating a good phylogenic inference (not altering either the bootstrap values significantly, nor the node locations).

The studied regulator, CopR, fits into the copper responsive regulators group and at the same time constitutes a single clade, giving the idea of a new class within the copper related regulators. This hypothesis explains the low homology found in the protein sequence shown in Figure 16 and 17. The other proteins present in this new class are the ones that were found to be more homologous in the BLAST analysis (data not shown). Nevertheless, experimental validation is always required to confirm the specifity of these proteins toward copper response.

When comparing the tree presented in Figure 18 with the one from E.A Permina *et al*. 2005 study (Figure I, presented in Appendix C, Figure 6 App.C.2), it is possible to observe a new position for copper affinity regulators, such as *Sy.elongatus* Q8DHQ7, *Synechoscystis* Q55963 and *Sh.oneidensis* Q8EGB7, which could be related with the importance of copper as a cofactor (see Introduction, Chapter 1). In addition, the lead and cadmium regions are mixed, which is in accordance with the fact that most of these heavy metal resistance systems respond to more than one heavy metal [153].

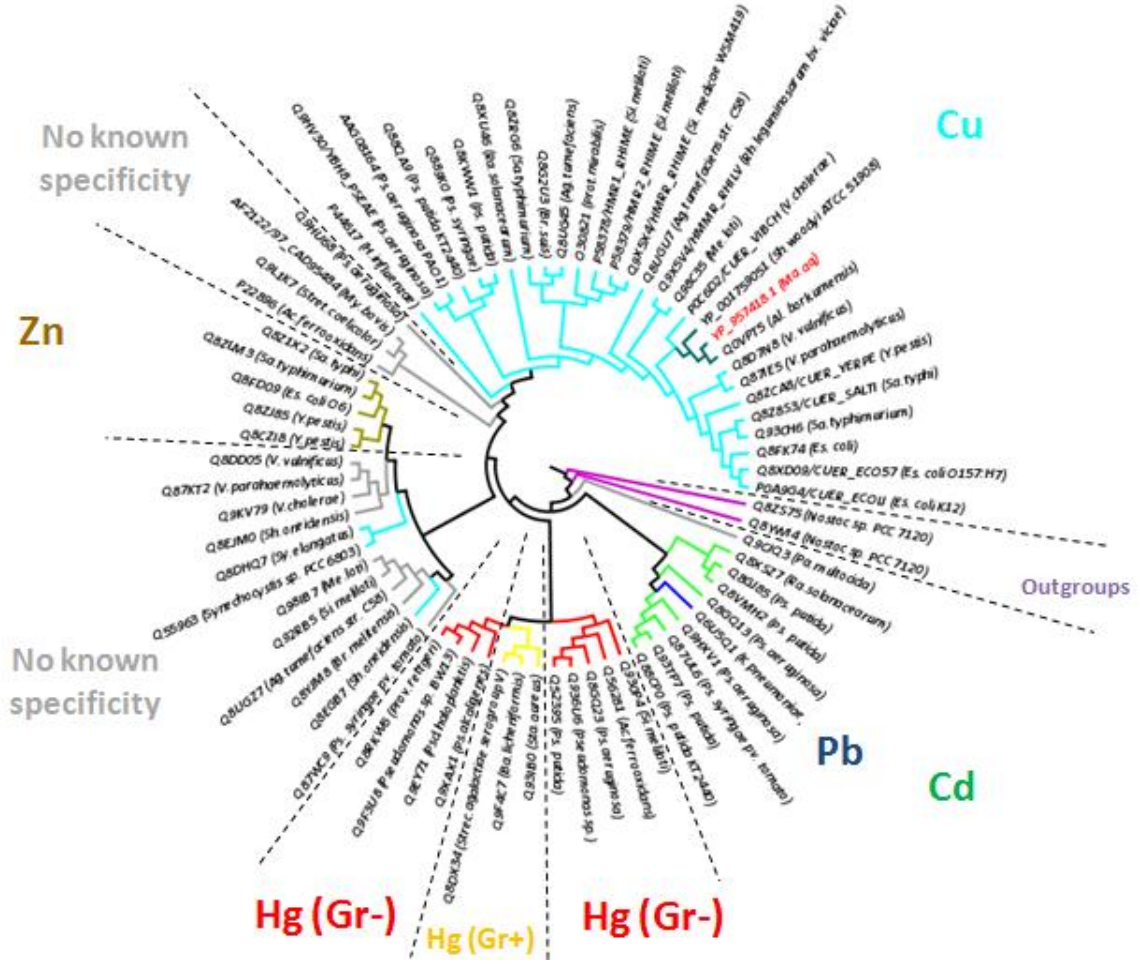

Figure 18 – Neighbor-joining algorithm regulators cladogram tree (proteins accessions numbers – Appendix C – Table 7 App.C.1). Polar representation. Legend: Light blue - Cu ion region; Green - Cd; Dark blue - Pb; Yellow - Hg Gr+; Red - Hg Gr-; Gray - No known specificity; Teal - *Ma.aq* (studied sequences). Bootstrap values above 80% (data not shown for figure clarity). Clade distances relates to evolutionary distances. Identical to the other algorithms used, see Material and Methods, Chapter 2 –1.2 (data not shown).

Considering the proposed copper regulon (Figure 10), there are no other putative regulators, so it is possible that the proposed copper pump system CusCBA (similar to the *cus* system in *Es.coli*-like bacteria) might be regulated by this two component system. This hypotesis is supported by the sequence alignment between *Es.coli* CusR/S and *Ma.aq* CopR/S shown in Appendix B, Figure App.B.1(3) and App.B.2(3) Indeed, in a BLAST search using *Es.coli* CusR/S sequences against *Ma.aq* genome the most homologous system identified was the CopRS (*maqu_0123* and *maqu_0124*).

1.1.2 - A novel copper protein, CopX

The annotated biological function of CopX [154], classifies this protein as a copper-binding protein. The copper resistance-like operon, Cop or Pco, has a copper-trafficking protein, CopC, that binds two copper ions, Cu^{2+} and Cu^{+} (see Introduction, Chapter 1). Nevertheless, CopX does not have homology to those proteins, as can be seen by the sequence analysis and alignment, in Figure 19.

```
Al.b(cbp)          -------------------------MTVSKFLVAAATIL---LSATTFAAGTHGGGHDHG 32
Ma.aq(CopX)        -------------------------MNASNLIVAGLAFS---MSVSAFAAGTHGGGHGHG 32
Ps.p(cp-type1)     ---------------------------MKRLFIAAALAL---VSLPTLAG------TGKS 24
Ps.a(cbp)          MLPTASRQAQRHRIGDINPMEIPR-MFPRRLLPASLIVLGVLFGASAQAS------PAHG 53
Ps.c(cbp)          -------------------MFLPKGVLPRSLTLAACLLA---LGAPAWAS------PAQN 32
(1)                                           :  *   .     . .: *.        :
Ps.sy(copC)        -------------------MLLN-------------------------------------R 5
(2)

Al.b(cbp)          AASGEPGKASEASRTITVEMYDNYYEPESIIVSPGETVRFVVENKGNLVHEFNIGTPEMH 92
Ma.aq(CopX)        ASIGEPGKASEASRTITVEMHDNYYEPEEIRVKPGETVRFVVQNKGNLVHEFNIGTPGMH 92
Ps.p(cp-type1)     FSFGEPAPAAKATRTVEVVLKDIAFEPRSLQVEAGETVRFVLINEGKLPHEFNLGDKAMH 84
Ps.a(cbp)          QAFGKPAQAAQASRSIEVVLGDMYFKPRAIEVKAGETVRFVLKNEGKLLHEFNLGDAAMH 113
Ps.c(cbp)          YDFGQAASAAKATRTVELVMGDMSFSPKTLDIKAGETVRFVLVNKGQLLHEFNLGDAAMH 92
(1)                   *:.. *::*:*:: : : *  :.*. : :..*******: *:*:* ****:*   **
Ps.sy(copC)        TRFVTLFAAGMLVSALAQAHPKLVSSTPAEGSEGAAPGKIELHFSENLVTQFSGAKLVMT 65
(2)                        *.    ::     .   ..     . . . :: :  . :*  :*. .   *

Al.b(cbp)          EGHQEEMMMMVEHGVIQGG--------KLNHDMMEMDMGNGQSMKHDDPNSVLLEPGQRQ 144
Ma.aq(CopX)        EAHQKEMRMMVEHGVIQGN--------KLNHDMMNMDMGNGHSMKHDDPNSVLLEPGQSR 144
Ps.p(cp-type1)     AQHQKEMISMQGKLFTAG----------MNHEGMDHGQMDHGAHGHDAANTVLLQPGQRA 134
Ps.a(cbp)          AEHQKEMLEMQQSGMLTPTGMASMDHSQMGHGMAGMDHGRM--MKHDDPNSVLVEPGKSA 171
Ps.c(cbp)          AKHQQEMLKMQQSGMLTSTAVKDMPAGAMDH--AAMGHGAMPGMKHDDPNSVLVEPGKTA 150
(1)                  **::*  * .  .             . *             .**. *:**: **:
Ps.sy(copC)        AMPGMEHSPMGVKTAVSGG--------------------------GDPKTMIITPASP- 97
(2)                     *   *                                    . .::::: *..

Al.b(cbp)          EVIWKFSDNSNIEFACNVPGHYQAGMYGDVKVQ-- 177
Ma.aq(CopX)        EVVWTFANQGNIEFACNVPGHYQSGMYGDVNFE-- 177
Ps.p(cp-type1)     ELTWTFRQSAPIEFACNVPGHYQAGMVGPLTIE-- 167
Ps.a(cbp)          ELTWTFAKATRLEFACNIPGHYQAGMVGQLTVQP- 205
Ps.c(cbp)          ELTWTFSKAGSLEFACNIPGHYQAGMVGKLT---- 181
(1)                *: *.* .   :*****:*****:** * :...
Ps.sy(copC)        ------LTAGTYKVDWRAVSSDTHPITGSVTFKVK 126
(2)                           :.  .  .     : * :.
```

Figure 19 – Alignment of CopX with other blue type 1 copper protein (1) (proteins accessions numbers Appendix B –Table 6 App.B.3) and known copper resistant function protein (2), CopC from *Ps. syringae*. Legend: Gray residues correspond to type 1 blue copper protein motif. Asterisk – Identical residues in all the alignment; Colon – Conserved substitutions; Dot – Semi-conserved substitutions.

The analysis of CopX sequence shows clearly that CopX presents the four type 1 copper ligands, Cys160, His165 and Met170 at the C-terminal with another H82 towards the

N-terminal (see Introduction, Chapter 1) [112]. The proposed C-terminal ligand-containing loop of CopX, (**C**X_2PG**H**X_3G**M**), is similar to the one found in rusticyanin, a type 1 copper protein isolated from *Thiobacillus ferrooxidans*, which has a high redox potential (+ 680 mV) and is stable in very acid conditions (up to pH < 2.0) [155].

Another important difference is the size, when compared to other blue proteins associated with copper resistance, CopX is larger (Figure 19). Other differences can be inferred by the phylogenic tree (Figure 20).

The BLAST analysis showed that CopX is not related to known small copper proteins associated with copper resistance systems, CopC-like, which is also highlighted by the phylogenetic analysis, Figure 20. Indeed, CopX does not present either Cu^{+} or $Cu2^{+}$ binding sites that are characteristic of those small proteins: MXCXXC for Cu(I) or $H(X)_nE(X)_mDXH$ for Cu(II) (see Introduction) [97]. In addition, the analysis of the protein sequence of CopX shows that another sequence identified as a Cu^{+} binding site, $[M(X)_nM]_m$, is also absent [102].

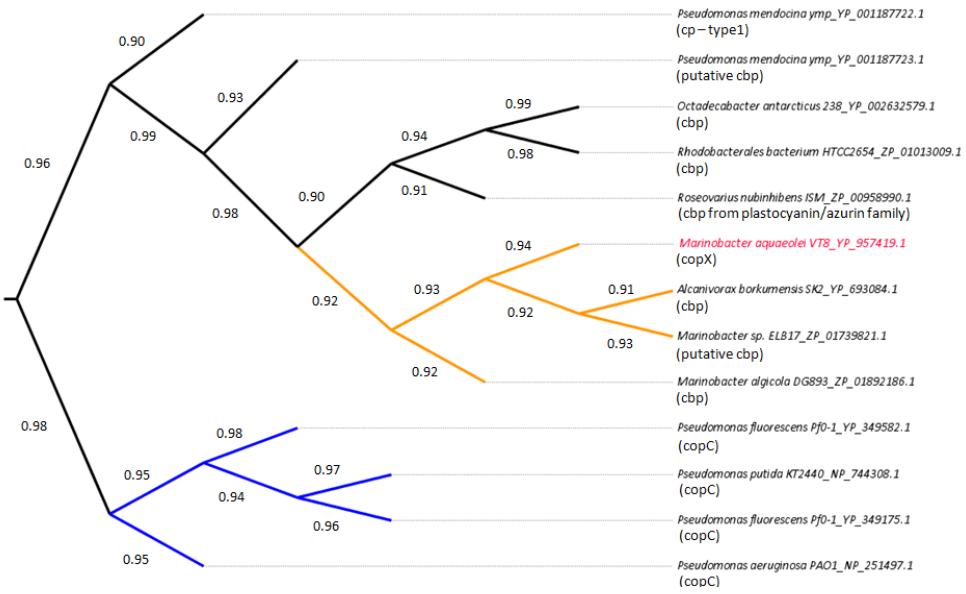

Figure 20 – Neighbor-joining algorithm CopX phylogenic inference (proteins accessions numbers Appendix C – Table 7 App.C.3). Legend: Blue - Known copper associated resistance proteins. Yellow - CopX (and homologous ORF and proteins). Black - Copper proteins. Clade distances do not relate to evolutionary distances. Homologs to the other algorithms used, see Material and Methods, Chapter 2 –2.1 (data not shown).

The tree presented above was performed by three algorithms (data not shown) and no difference was encountered, not even in bootstrap values above 90%. In the CopX tree no

outgroups was used, as the known blue copper proteins associated with copper resistance differ widely from CopX.

This tree clearly shows that copper binding proteins associated with known copper resistance systems are separated from CopX. Furthermore, looking closely to the node distribution, we can see that CopX seems to have evolved when compared to other known copper resistance systems, similarly, the clade length highlights the evolutionary distance (data not shown). This can explain the different ORF position. The group associated with our protein has an ORF organization similar to *Ma.aq* genome but in opposition to this bacterium, no proposed copper operon is annotated. The phylogenetic analysis confirms the protein sequence analysis presented before.

2. Determination of MIC/MTC values

Solid and Liquid media assays

The Minimum Inhibitory Concentration (MIC) / Maximum Tolerance Concentration (MTC) evaluates de susceptibility of a bacterium to a specific substance that has bactericidal activity or inhibits the growth (bacteriostatic activity), such as, antimicrobial agents (antibiotics) and heavy metals. The MIC/MTC can be determined using both solid and liquid media assays.

Solid media have been proven useful in qualitative studies of morphogenetic effects of heavy metals, and also to observe the relationship between metal uptake and toxicity in microorganisms [156]. This method has been used in the screening for resistant and sensitive species to heavy metals.

However, defining a bacterium susceptibility to metal ions only with qualitative data would not be precise. Thus in this work, plates with a range of metal ions concentration were used to determine the MIC / MTC of *Ma.aq* to cadmium, cobalt and copper ions.

Solid assays illustrate well the differences between organisms and their interactions with different heavy metal ions, but it must be stressed that usually the metal concentrations quoted are total values and the biologically available concentrations may be lower. This is due to the problematic of metal ion complexation by the physical and chemical nature of the growth medium: binding or complexation of metals by organic compounds,

precipitation, pH and ionic interactions. This means that components of the solid media, such as agar or yeast extract are capable of binding significant amounts of metal ions, reducing, apparently, their toxicity.

Although, the problems outline for solid media are also applied to liquid media assays, the later can be better controlled, reducing the negative effects. Therefore, the maximum metal concentrations tolerated by a bacterium are likely to be lower in liquid medium than in solid medium, as the amount of metal ion biologically available is higher in liquid medium. For this reason, liquid medium assays are more precise and are used to confirm the results obtained in solid medium assays.

In the present study, these two assays were used to determine the susceptibility of *Ma.aq* to cadmium, cobalt and copper ions.

2.1. Cadmium Stress

2.1.1 Solid Media

Ma.aq cadmium ion stress was induced with $CdCl_2$. ASW Agar medium plates were prepared as described in Materials and Methods, Chapter 2 – 2.1. Plates were streaked and left incubating at 30°C during 5 days (120h). Plates were check for colonies and changes in the color/aspect of the solid media – qualitative analysis. The concentrations tested ranged from 0 to 250 µM $CdCl_2$ – quantitative analysis. The results are summarized in Table 3 and Figure 21.

The qualitative results indicate that no variation was observed in the color or aspect of the colonies (in both cases the colonies were white, Figure 21-A and B). *A priori* this was expected as cadmium ion compartmentalization is not expected, usually the detoxification mechanisms involves "pumps" that expel the ions to the medium. Moreover, changes in colony color have never been reported in cadmium resistance bacteria (see Introduction, Chapter 1).

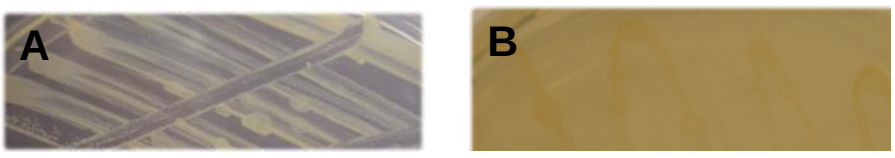

Figure 21 – A) *Ma.aq* control growth in ASW agar after 48 h incubation. **B)** *Ma.aq* growth in ASW agar with 100 µM $CdCl_2$ after 48 h incubation.

The MIC/MTC concentration for Cd^{2+} ions was determined to be 200 μM $CdCl_2$, since at this concentration no colonies were observed (Table 3). It is also important to highlight that at concentrations of $CdCl_2$ higher than 100 μM, the growth was only observed after 48 h. This can be explained attending to the fact that the response of the resistance system is slower due to the high intracellular concentration of free ion, and only after 48 h the organism was able to lower this concentration to acceptable levels to start growing. This might correspond to the time needed for a certain resistance system to be expressed in response to cadmium stimuli. Thus, we assume that during the first 24 h, the colonies were very small and undetectable to the human eye.

This phenomenon was confirmed by the omission of the "*lag* phase" by streaking 100 μM $CdCl_2$ plates with adapted cadmium stress *Ma.aq* (inoculum from a 190 μM $CdCl_2$ plate grown for 48 h). In these experiments, visible growth was observed after 24 h incubation. However, when the inoculum was a colony from a 190 μM $CdCl_2$ plate grown for 120 h, it was possible to observe a loss of viability after prolonged incubation, which might be due to the accumulation of toxic quantities of metal ions.

Table 3 – Results obtained in the determination of the MIC value for cadmium ions using solid media. Growth was followed during 5 days to register qualitative changes in the colony and media**.** Legend: V – Visible Growth; - No growth; + Growth; C – Colonies Color; W- White; n/a - Not applicable. Control ASW plates correspond to a non-stress growth of *Ma.aq*.

†Heavy metal ion Cd^{2+}	Days (hours)									
	24		48		72		96		120	
[$CdCl_2$] (μM)	V	C	V	C	V	C	V	C	V	C
ASW control (0)	+	W	+	W	+	W	+	W	+	W
0.1 to 75	+	W	+	W	+	W	+	W	+	W
100 to 190	-	n/a	+	W	+	W	+	W	+	W
200 to 250	-	n/a	-	n/a	-	n/a	-	n/a	-	n/a

† ASW Agar medium plates with $CdCl_2$ and control results were obtained in triplicates.

2.1.2 Liquid Media

In order to confirm and accurately determine the MIC value obtained using solid medium, the bacteria growth kinetics was followed during the stress induced by 75 μM, 100 μM and 250 μM $CdCl_2$. The growth was prepared as described in Material and Methods,

Chapter 2 –2.1, and was followed at OD_{600nm} during 48 h (Figure 22 and Appendix E – Figure App. E.1).

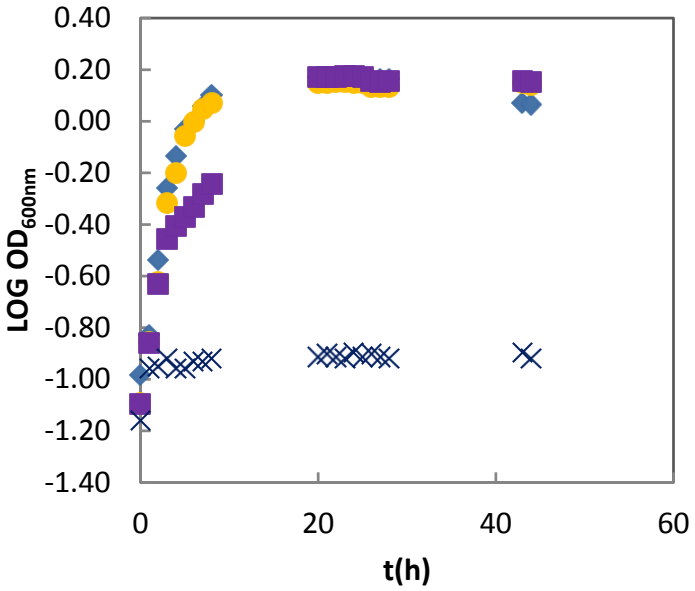

Figure 22 – *Ma.aq* growth curve in the presence of different cadmium ion concentration. The growths were followed for nearly 48 h. Legend: ◆ 0 $CdCl_2$ (ASW control growth); ● 75 µM $CdCl_2$; ■ 100 µM $CdCl_2$ and ✕ 250 µM $CdCl_2$.

The results presented in Figure 22 and Table 4, show that the growths in the presence of up to 100 µM $CdCl_2$, have an exponential phase with duration identical to the control growth and there is no *lag* phase. It is also possible to observe a slower exponential phase with the increase of cadmium ion concentration in the medium (see below and Table 4) and that only at 250 µM $CdCl_2$ no growth is observed (inhibitory concentration). However, to accurately confirm the MIC/MTC values determined using solid media, growths in the presence of $CdCl_2$ between 100 µM and 250 µM need to be performed.

In the presence of 100 µM $CdCl_2$, the growths performed in solid media presents no visible growth after 24h incubation, but in liquid media the exponential phase started after 1 hour. This can be explained by the lower specific growth rate observed in the growth with 100 µM $CdCl_2$, (Table 4).

Specific growth rate (µ) and the generation time (td) were determined for each concentration, using semilog plot of the exponential phase (1^{st} to the 5^{th} hour) of the growth and the results can be found below (Table 4).

Table 4 – Specific *Ma.aq* 617 growth rate (μ) and generation time (td) of *Ma.aq* growths in the presence of difference $CdCl_2$ concentrations. Values were determined from an exponential phase (1st to the 5th hour) semilog plot using logarithms to the base 10. Legend: n/r – not-recovered.

[$CdCl_2$] (μM)	μ (h^{-1})	td (h)	Wet weight of cells (g/L)
0.0 (control)	0.50	1.39	4.2
0.3	0.49	1.34	n/r
1.0	0.47	1.47	4.0
10	0.46	1.50	3.6
30	0.45	1.54	n/r
50	0.41	1.71	n/r
75	0.35	1.97	n/r
100	0.34	2.04	3.0

The specific growth rate decreases with the increase of $CdCl_2$ concentration present in the medium. However, it should be pointed out that those values come from a single experiment. So the decrease in the specific growth rate in the growths in the presence of 0.3 to 10 μM $CdCl_2$ may not be significant, although the amount of cells recovered decreases.

In the growths performed in the presence of $CdCl_2$, the nutrients take more time to be exhausted (lower growth rate) from the media and therefore, these cultures enter in the death phase later than the control growth. This can be seen by comparing the td values of the growths with increasing cadmium concentrations with the control growth, Table 4.

Corroborating this, the marginal higher value of the OD_{600nm} after 40h in higher cadmium concentrations, might indicate that the growths are in different cellular stages and the dead phase is reached later for higher cadmium concentrations (Figure 22 and Appendix E Figure App.E.1).

The MIC/MTC of 200 μM $CdCl_2$ determined for *Ma.aq* 617 is higher than the one found for cadmium resistant bacteria (150 μM cadmium ions) [153], thus establishing *Ma.aq* 617 as a cadmium resistant bacteria. The *Ma.aq* genome was searched for a molecular system that could be responsible for this resistance, but no annotated system was found. Nevertheless, there is the hypothesys that the annotated *cusCBA* found in the copper loci, might be a *czcCBA* system, due to its similarities (see Introduction, Chapter 1) and be responsible for the cadmium resistance.

2.2. Cobalt Stress

2.2.1 Solid Media

Ma.aq cobalt ion stress was induced using $CoCl_2$. ASW Agar medium plates were prepared as described in Material and Methods, Chapter 2– 2.1. The $CoCl_2$ concentration tested ranged from 0.75 nM to 8 mM of $CoCl_2$. The results are summarized in Table 5 and Figure 23.

These qualitative assays showed a significantly change in the color of the colonies upon increase of $CoCl_2$ concentration in the medium: from white color to light pink (starting at 0.25 mM $CoCl_2$) and to dark pink color in the presence of higher concentrations (above 0.75 mM $CoCl_2$) (Figure 23). The change in the colony color is concominant with the change in the color of the medium. The plates with cobalt chloride prior to inoculation presented a pink color. After inoculation, this pink color disappears and the solid medium becomes transparent after 48 h.

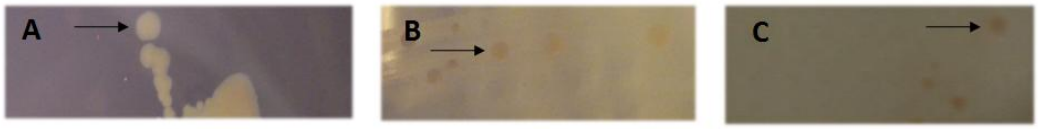

Figure 23 – A) *Ma.aq* control growth in ASW media agar plate, after 72 h incubation. **B)** *Ma.aq* growth in ASW agar plate, in the presence of 0.5 mM $CoCl_2$, after 72 h incubation. **C)** *Ma.aq* growth in ASW agar plate, in the presence of 4.0 mM $CoCl_2$, after 96 h incubation. Arrows highlight isolated colonies for comparison.

This color change might be an indication that the cobalt ions are being removed from the solid medium, which seems to be linked with the production of a cobalt derived “compound” inside the cell (as the colonies become pink).

This “compound” might be a cobalt-dependent enzyme that is being induced by the high concentration of this ion in solution. The most common cobalt cofactor is the cobalamin, vitamin B_{12}, and proteins that bind this type of cofactor present a characteristic pink color [157]. Nevertheless, although *Ma.aq* presents genes to endeavor the metabolic pathway of vitamin B_{12} production [158], qualitative assays are not enough to attribute the change in the colonies color to this specific group of enzymes.

The growth performed in solid media also shows that there is a strong *lag* phase at higher $CoCl_2$ concentrations (Table 5). At low concentrations (until 0.25 mM $CoCl_2$), after 24h incubation only very small colonies were visible. At medium concentrations (from 0.30 to 1.5 mM $CoCl_2$), only after 48 h it was possible to observe the appearance of the small colonies. At high concentrations (from 1.5 to 6 mM $CoCl_2$), there was a lag phase of 72 h for the observation of the first colonies in the plates.

Table 5 – Results obtained in the determination of the MIC value for cobalt ions using solid media. *Ma.aq g*rowth was followed during 5 days to register qualitative changes of colony and media**.** Legend: V – Visible Growth; - No growth; + Growth; C – Colonies Color; sp – Slightly pink; P – Pink; n/a - Not applicable. Control ASW plates correspond to a non-stress growth of *Ma.aq*.

†Heavy metal ion Co^{2+}	Time of Incubation (h)									
	24		48		72		96		120	
[$CoCl_2$] (mM)	V	C	V	C	V	C	V	C	V	C
ASW control (0.75 nM)	+	W	+	W	+	W	+	W	+	W
0.1	±	W	+	W	+	W	+	W	+	W
0.25	±	W	+	sp	+	sp	+	sp	+	sp
0.30 - 0.5	-	n/a	±	W	+	sp	+	sp	+	sp
0.75 – 1.5	-	n/a	±	W	+	P	+	P	+	P
2.0 – 6.0	-	n/a	-	n/a	±	W	+	P	+	P
8.0 – 15.0	-	n/a	-	n/a	-	n/a	-	n/a	-	n/a

†ASW Agar medium plates with $CoCl_2$ and control results were obtained in triplicates.

The *lag* phase observed could be reduced by 24 h by using cobalt ion stress adapted *Ma.aq* (48 h growth from a plate with 1.5 mM $CoCl_2$ and 72 h growth from a plate with 2.0 mM $CoCl_2$). Although reduced, the *lag* phase is maintained above medium concentrations (higher then 0.30 mM). The MIC/MTC value for $CoCl_2$ was found to be between 4 and 6 mM $CoCl_2$, thus more assays are required to accurately determine this value.

2.2.2 Liquid Media

Cobalt ion stress in *Ma.aq* was also studied in liquid media, in the range of 0.75 nM to 1.0 mM $CoCl_2$. Bacteria growth kinetics was followed during 48 h. The growth was prepared as described in Material and Methods, Chapter 2 –2.1, and the OD was measured at 600 nm. The results obtained are presented in Figure 24, and the exponential phase of each growth, 1^{st} to the 6^{th} hour, was analyzed to determine μ, and td (Table 6).

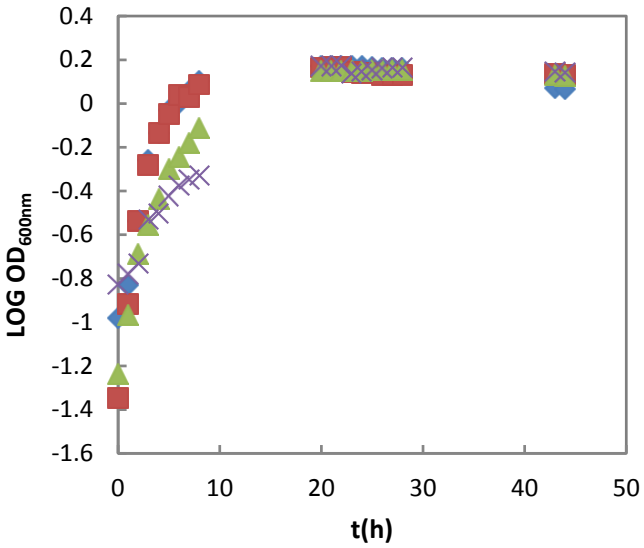

Figure 24 – *Ma.aq* growth curve in the presence of different cobalt concentrations. The growths were followed for nearly 48 h. Legend: ◆ 0.75 nM $CoCl_2$ (ASW control growth); ■ 100 µM $CoCl_2$; ▲750 µM $CoCl_2$; × 1.0 mM $CoCl_2$.

It was observed a slower growth with increasing $CoCl_2$ concentration, corroborated by the µ values (Table 6), indicating an adaptation to increased cobalt ion concentration (Figure 24). In addition, the OD_{600nm} after 40 h incubation increases with increasing $CoCl_2$ in the medium, which can be explained by the slower growth (Table 6), as nutrients are available for longer periods of time. The µ, and the td, determined explain the two observations stated before, Table 6.

Table 6 – Specific *Ma.aq* 617 growth rate (µ) and generation time (td) of cobalt ion stress concentration range tested by liquid media in *Ma.aq*. Values were determined by an exponential phase (1st to the 6th hour) semilog plot using logarithms of base 10.

[$CoCl_2$]	µ (h^{-1})	td (h)	Wet weight of cells (g/L)
0. 75 nm (Control)	0.52	1.34	5.0
100 µM	0.41	1.70	3.2
750 µM	0.38	1.85	2.6
1.0 mM	0.19	3.66	2.2

Similarly, to what was observed in the solid media assays, the growths in liquid media also show a change of color, becoming more pinkish with increasing cobalt chloride concentration in the medium. It was observed that the cell pellet becomes pink while the medium loses the pink color. This observation strengths the idea of cobalt-depend enzymes

or a cobalt containing compounds being produced, and thus of cobalt sequestration inside the cell.

It is important to mention, that the results obtained in liquid media cannot be used to confirm the MIC/MTC range found in solid media, since the maximum concentration of $CoCl_2$ tested was 1 mM.

The MIC/MTC of $CoCl_2$ for the *Ma.aq* was determined to be between 4 and 6 mM, which is well above the MIC/MTC value found for cobalt resistant bacteria, 250 μM $CoCl_2$ [42]. This result establishes *Ma.aq* 617 as a cobalt resistant bacterium. A possible molecular system responsible for the cobalt resistance is the *cnr* (see Introduction, Chapter 1) or the *czc* refered in the cadmium stress, but a homologous system was not found in the genome of *Ma.aq.* The other hypothesis is the cobalt ions being used in the synthesis of a cobalt-containing compound or being bound to an unknown protein inside the cell, which would explain the strong pink color of the cells.

2.3 Copper stress

2.3.1 Solid Media

The *Ma.aq* copper ion stress was induced using different concentrations of $CuSO_4$, ranged from 0.2 nM to 3.2 mM $CuSO_4$. The results are summarized in Table 7.

Table 7 – Results obtained for the determination the *Ma.aq* MIC value for copper ions using solid media. The plates were incubated during 5 days to register changes in media and colony color and aspect. Legend: V – Visible Growth; - No growth; ± Small colonies; + Growth; C – Colonies Color; W- White; b – Slightly brown; B – Brown; n/a – Not applicable. Control ASW plates present 0.2 nM $CuSO_4$.

†Heavy metal ion Cu^{2+}	Days (hours)									
	24		48		72		96		120	
[$CuSO_4$]	V	C	V	C	V	C	V	C	V	C
ASW control (0.2 nM)	+	W	+	W	+	W	+	W	+	W
1 nM to 900 μM	+	W	+	W	+	W	+	W	+	W
1.0 mM	±	b	+	B	+	B	+	B	+	B
1.1 to 1.5 mM	-	n/a	+	b	+	B	+	B	+	B
1.6 to 2.0 mM	-	n/a	-	n/a	-	n/a	-	n/a	-	n/a

†ASW Agar medium plates with $CuSO_4$ and control results were obtained in triplicates.

According to the results presented in Table 7, the MIC/MTC for the copper ion is 1.6 mM $CuSO_4$. In addition, it is important to point out that the growth in the plates containing 1.1 mM to 1.5 mM $CuSO_4$ only appeared after 24 h. However, when inoculated from one plate adapted to the copper stress (1.0 mM $CuSO_4$ 48 h plate), this "adaptation" or "lag" phase is no longer observed. This can be explained by the time needed for the putative copper resistance system to respond to the copper stimuli and be transcribed. It was also observed a lost of viability when using plates with long incubation time, particularly in cells submitted to higher toxic concentrations, above 1.0 mM $CuSO_4$, which may be due to the toxic effects of copper ions that permanently damage part of the cell machinery, most probably at the DNA level.

Another observation was the color of the colonies at higher $CuSO_4$ concentration (Figure 25). After 24 h incubation, in the presence of 1.0 mM – 1.5 mM $CuSO_4$, the colonies started to appear slightly brown, and after 48 h the colonies became dark brown (Figure 25). This had already been observed in similar studies of copper resistance in *Es.coli* and *Xhantomonas* strains [87], which was considered to be an indication of sequestration of copper ions from the growth media. A similar event might be occurring in *Ma.aq*.

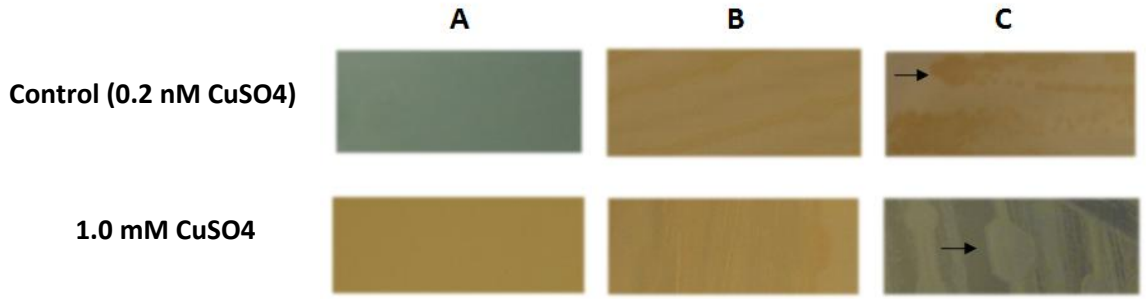

Figure 25 – ASW agar control plates and copper plates with a $CuSO_4$ concentration of 1.0mM. **(A)** Control plate after 24h incubation. Plate containing 1 mM $CuSO_4$ before inoculation **(B)**, after inoculation and incubation for 24 h **(C)**, and 120 h (control with dark background for picture clarity). Arrows highlight colonies differences.

In addition, the initial intense green/blue color of the plates containing $CuSO_4$ starts to disappear upon inoculation and with the incubation time (Figure 25). This change in the color of the medium is not due to the temperature (no change was observerd after 120 h of incubation in the absence of growth), thus it is a result of the growth (Appendix E – Figure 11

App. E.2). This change in the color of the medium can be attributed to copper being complexed and/or reduced from Cu^{2+} to Cu^{+}.

2.3.2 Liquid Media

The stress induced by the copper ions was also studied in liquid media, in order to determine the bacteria growth kinetics under these conditions. The growth was prepared as described in Material and Methods Chapter 2 – 2.1, and the OD was followed at 600 nm during 24 h (Figure 26). The $CuSO_4$ concentrations studied ranged from 0.2 nM to 2.0 mM, and the results obtained for the μ, and td, for each growth are presented in Table 8.

In Figure 26, are presented the growth curves of *Ma.aq* in the presence of 100 μM, 1.0 mM and 1.6 mM $CuSO_4$, followed during 24 h. The analysis of this Figure in combination with the values listed in Table 8, shows that higher copper ion concentration, up to 1.0 mM $CuSO_4$, affects the growth slightly, by decreasing the specific growth rate. The decrease in the growth rate with increasing copper concentration might explain why in the solid medium it takes 48 h to visualize growth. This "*lag* phase" would correspond to the time required to induce the transcription of the putative copper resistance operon. The summarized OD_{600nm} value in Appendix E (Figure 12 App. E.3) also demonstrates these observations.

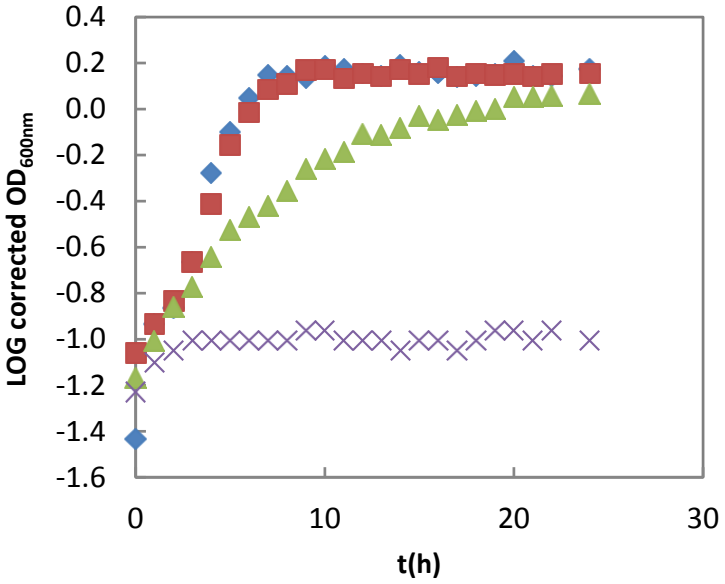

Figure 26 - *Ma.aq* growth curve of copper ion stress followed during 24h. Legend: ◆ 0.2 nM $CuSO_4$ (ASW control growth); ■ 100 μM $CuSO_4$ ▲ 1.0 mM $CuSO_4$ × 1.6 mM $CuSO_4$.

In addition, the MIC/MTC concentration determined in liquid media matches the one determined previously using solid media, 1.6 mM $CuSO_4$. Indeed, the cells that were incubated in a media containing 1.6 mM $CuSO_4$ were observed at the optic microscope demonstrating that there was no mobility, and only cell debris were present (data not shown).

Table 8 – Specific *Ma.aq* 617 growth rate (μ) and generation time (td) of the growth under copper ion stress. Values were determined by an exponential phase (1st to the 6th hour) semilog plot using logarithms of base 10. Legend: n/r – non-recovered.

[$CuSO_4$]	μ (h^{-1})	td (h)	Wet weight of cells (g/L)
0.2 nM (control)	0.59	1.08	4.6
2.0 nM	0.59	1.09	n/r
100 μM	0.52	1.24	3.6
300 μM	0.50	1.28	n/r
700 μM	0.41	1.56	n/r
900 μM	0.40	1.61	n/r
1.0 mM	0.39	1.63	2.8
1.5 mM	0.38	1.69	2.0

Qualitative results were also registered. The prepared media has an increasing blue color with increasing copper concentrations ($CuSO_4$). After harvesting the cells the resulting media had lost its dark blue color, presenting a similar incolor/whitish, characteristic of the ASW control *Ma.aq* growth. The cell pellet from the copper stress growth (1.0 and 1.5 mM $CuSO_4$) had a more opaque color but not as brown as was expected. This strengths the hypothesis of copper being reduced rather than sequester. Indeed, the copper content of the medium was determined by the Cu^+ assay of blue copper proteins, and the amount of copper present before and after the growth is identical, within the errors of the method.

The MIC/MTC of $CuSO_4$ for *Ma.aq* 617 was determined to be 1.6 mM. The bacteria that are more related with *Ma.aq*, the Pseudomonads-like bacteria, are considered to be copper resistant with a MIC/MTC above 0.6 mM $CuSO_4$ [159], thus *Ma.aq* 617 is a copper resistant strain. Moreover, when searching in the Pseudomonads strains the MIC/MTC value of 1.5 mM $CuSO_4$ is one of the highest.

The proposed operon *copXAB*, has less genes than the Pseudomonads homologs, *cop* operon (see Introduction, Chapter 1). This characteristic together with the different gene

composition (three instead of four and the presence of an unrelated copper binding protein, CopX) and the higher MIC/MCT value determined, might point out to a new copper resistance mechanism.

3. Expression of the Resistance System in *Marinobacter aquaeolei* 617

In an attempted to identify the proteins that might be associated with the resistance mechanism of the different heavy metal ions studied cellular fractions were analyzed. Only the results obtained for cobalt and copper are presented, as the equivalent study for cadmium ion stress in *Ma.aq* was not performed. In the case of cobalt stress, the cellular fractions were analysed by SDS-PAGE and UV-visible spectroscopy (see Material and Methods, Chapter 2 – 3 and 8.1), while for copper stress a preliminary proteomic approach by 2D electrophoresis was performed (see Material and Methods, Chapter 2 – 5).

3.1. Stress induced by Co^{2+} - Analysis of the cell content

After the growth in the presence of 750 μM, 1.0 mM and 2.0 mM $CoCl_2$, the cells were harvested, as described in Chapter 2 Material and Methods – 2.1. The total protein content is summarized in Table 9.

Table 9 – Total protein content of cellular fractions obtained of the growths in the presence of different cobalt chloride concentrations.

[$CoCl_2$]	Total protein content (mg/mL)		
	Periplasm	Citoplasm	Membrane
0.75nM (control)	1.3	2.1	5.3
1.0 mM	2.0	0.9	3.4
2.0 mM	2.6	0.8	3.3

The cellular fractions were prepared as described in Material and Methods, Chapter 2 – 2.2, and was analysed by SDS-PAGE (Figure 27).

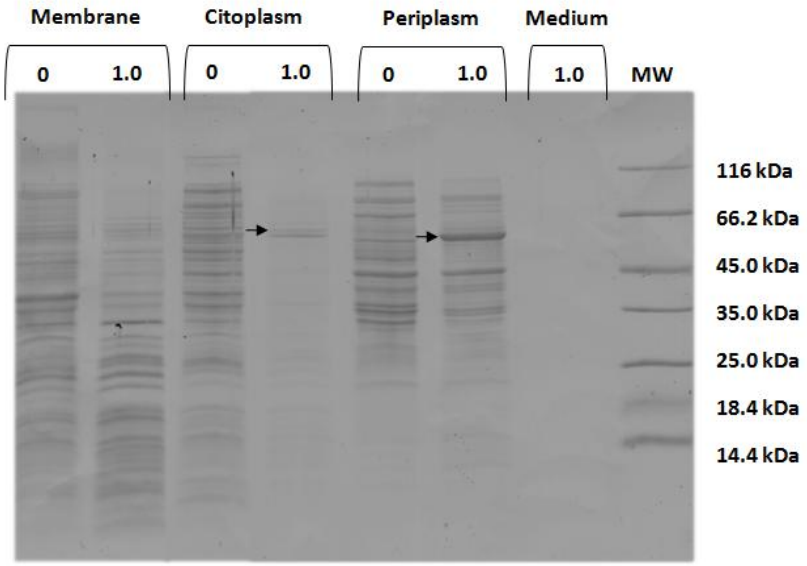

Figure 27 – SDS-PAGE analysis of the cellular fractions of *Ma.aq* 617 grown under cobalt ion stress. Legend: 0 – control, 0.75 nM $CoCl_2$; 1.0 – 1.0 mM $CoCl_2$; MW – Low molecular weight marker. The samples were normalized by total protein content (10 µg). Arrow highlights possible differential proteins. 12.5 % polyacrylamide gel stained with Coomassie blue.

The comparative analysis of the electrophoretic profile of the periplasmic fraction of the control growth and the one obtained in the presence of 1 mM $CoCl_2$ (Figure 27), shows that there are some protein bands that become more intense in 1 mM $CoCl_2$ periplasm (such as the one close to 66.2 kDa), while others are less intense (such as near 35 kDa). [1]

The electrophoretic profile of the cytoplasm cannot be analyzed as the normalization of the loaded samples has failed. However, it shows that the most abundant protein in the cytoplasm, obtained in the presence of 1 mM $CoCl_2$, would be one with a molecular weight close to 66.2 kDa.

It is important to point out that these are preliminary data, and that it is not enough to identify the protein(s) responsible for the cobalt ion resistance. In the future the periplasmic fraction, as well as, the cytoplasmic and membrane fraction will be analyzed in a 2D gel electrophoresis and the differentially expressed proteins will be identified by peptide –mass fingerprint.

[1] It is not possible to say whether these protein bands correspond to a single protein, and most probably more than one with similar molecular weight are being over-expressed.

3.1.1 UV-visible spectra of cellular fractions

The periplasmic fraction was analyzed by UV-visible spectroscopy in order to identify the compound/protein responsible for the change of color between the control and the growth in the presence of 1 mM $CoCl_2$ (Figure 28). A similar analysis was performed for the membrane and cytoplasmic fraction, but since there was cross-contamination between the membrane and cytoplasmic fraction, no conclusions can be taken from them (these results are presented in Appendix F, Figure 13 App.F.1 and Figure 14 F.2). The medium after cell harvesting was also analyzed but no control spectrum was performed, thus no conclusion can be taken (data not shown).

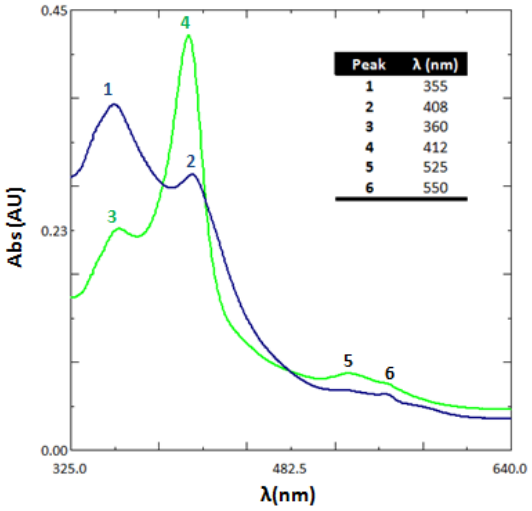

Figure 28 – UV-visible spectra of the periplasmatic fraction of the control growth, 0.75 nM $CoCl_2$ (green line), and of the growth in the presence of 1.0 mM $CoCl_2$ (blue line).

The visible spectrum of the periplasm of the control growth shows a strong 410 nm band that can be attributed to the Soret band of cytochromes, while in the case of the one from 1 mM $CoCl_2$ growth, the most intense band is at 355 nm [160]. This band is characteristic of cyanocobalamin compounds, which present λ_{max} at 360 nm, 548 nm and 518 nm. In the spectrum it is also possible to observe two other absorption bands at 520 nm and 550 nm, which would match well with the other maxima of cyanocobalamin, but these are also close to the α and β bands of cytochromes in the reduced form [157] (this spectrum also has a maximum absorbance at 412 nm, corresponding to the Soret band of reduced cytochromes).

The UV-visible spectrum of proteins with cobalamin as a cofactor also present maxima absorbance bands at 360 nm, 548 nm and 518 nm [161]. Thus, the visible spectrum of the periplasm indicates that vitamin B_{12} is being synthesized at high rates to decrease the amount of the cobalt that would be toxic for the cell, which can be or not concomitant with the induction of a protein that binds cobalamine as a cofactor.

Analysing the *Ma.aq* genome and searching in the identified metabolic pathways [158], it is possible to propose that the induced pathway is the one encoded by the adenosylcobalamin biosynthesis II (late cobalt incorporation) operon, chromosome *locus* 1.291.095 – 1.291.472. This pathway is proposed to be responsible for the synthesis of vitamin B_{12} under aerobic conditions.

Taking this hypothesis into consideration the protein present in the SDS-PAGE with a molecular weight close to 66.2 kDa could be the reactivating factor for ethanolamine ammonia lyase (*maqu_1239* in the chromosome *locus* 1396779 - 1398212), which is a periplasmic protein with a theoretical molecular weight, without the signal peptide, of 50 kDa (expasy/protparam). This protein has an important role in the adenosylcobalamin pathways activation of ethanolamine to acetaldehyde, functioning as a chaperone [157, 158], but under the growth conditions used this protein is not expected to be required. Further studies, as the ones mentioned before, are needed to identify the proteins involved in the proposed metabolic pathways, as well as, responsible for the cobalt resistance.

3.2. Stress induced by Cu^{2+} Preliminary proteomic analysis

In order to confirm that the putative copper resistance operon, CopXAB, is being transcribed under copper stress conditions, a preliminary proteomic analysis of the periplasmic fraction was conducted to identify CopX and CopA. The proteomic analysis focused on the periplasmic fraction because the translated gene sequence of those two proteins presents a signal peptide to direct them to the periplasm.

The preliminary proteomic analysis was carried out by comparative analysis of the 2D electrophoretic profile of the periplasmic fraction of the control growth (0.2 nM $CuSO_4$) and copper stress growth (1 mM $CuSO_4$).

The theoretical molecular weight and pI of CopA and CopX were determined using the translated gene sequence in the program Protparam [162], after removing the signal

peptide identified in SignalP [163] CopA is proposed to have an isoelectric point (pI) of 4.95 and a molecular weight of 61.6 kDa, and CopX a pI of 5.57 and a molecular weight of 17.2 kDa. This information will be used to locate their spots in the 2D gel.

3.2.1 2D Electrophoresis

The 2D electrophoresis procedure is described in Material and Methods, Chapter 2 – 5.2. The periplasmatic fraction obtained from the control growth, in 0.2 nM $CuSO_4$, was compared with the one obtained from a *Ma.aq* growth in the presence of 1.0 mM $CuSO_4$. The total protein concentration of the periplasmic fraction of the first growth was 0.8 mg/mL while in the second was 1.0 mg/mL.

3.2.1.1 3-10 pH range

The first 2D gel analysis was performed using a 3-10 non-linear (NL) pH 13 cm Immobiline DryStrip in the first dimension. The 2D maps obtained are presented in Figure 29.

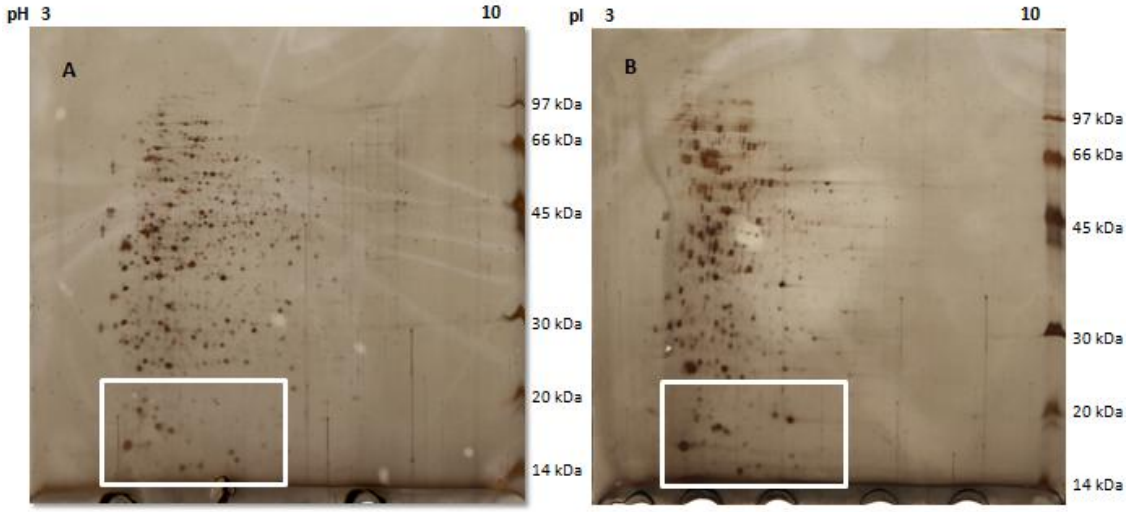

Figure 29 – 2D electrophoresis of the periplasmatic proteome of *Ma.aq* in 3-10 NL pH 13 cm Immobiline DryStrip **A)** Control growth with 0.2 nM $CuSO_4$ (30 µg protein applied); **B)** Copper stress growth with 1.0 mM $CuSO_4$ (30 µg protein applied). 12.5% Polyacrylamide gel stained by silver staining. Legend: White squares highlights region with visible differences.

The first conclusion that we can take from Figures 29A and 29B is that there are not many protein spots in the pH range 3-4 and 7-10 (according to the GE guidelines for 3-10 pH NL 13 cm Immobiline DryStrip). This is in agreement with the observation that most

periplasmic bacterial proteins are within the 4-7 pH range [164]. Therefore in the following experiments a narrowed pH range can be used (4-7), in order to improve the resolution.

Moreover, in both gels, it is possible to observe vertical streaks that can be related to poor cleaning conditions of the gel support system or to an exceeding Isoelectric Focusing (IEF). In the copper stress gel it is possible to observe a "shadow" effect, which can be attributed to three phenomenons:

i) poor IEF, due to either the difficulty in solubilization of high molecular weight proteins or to a possible high ionic strength (periplasmatic fraction, either from the growth in the presence of 0.2 nM or 1.0 mM $CuSO_4$, were resuspended in 10 mM Tris-HCl, so no treatment with a clean up kit was performed);

ii) polyacrylamide gel preparation, either the gel might have not been aged enough to create a homogenous network or the solution was not completely homogeneous;

iii) the temperature of electrophoresis cooling might have been too cold, interfering with the current voltage being applied during the run.

The coloration shadows that are observed are due to a poor silver staining procedure.

It is important to point out that, even without any image analysis, it is possible to observe a major difference between the two conditions tested at the low molecular weight region (the expected CopX region, attending to its pI and MW). In this region there are protein spots that appear or become more intense in the copper stress periplasm. Narrowed pH and linear pI gradients experiments were performed in order to further study these protein spots.

3.2.1.2 4-7 pH range

Narrowed pH and linear IEF experiments were performed using 4-7 pH range using a 13 cm Immobiline DryStrip. The results obtained with the control and copper stress periplasm fractions are shown below, in Figure 30.

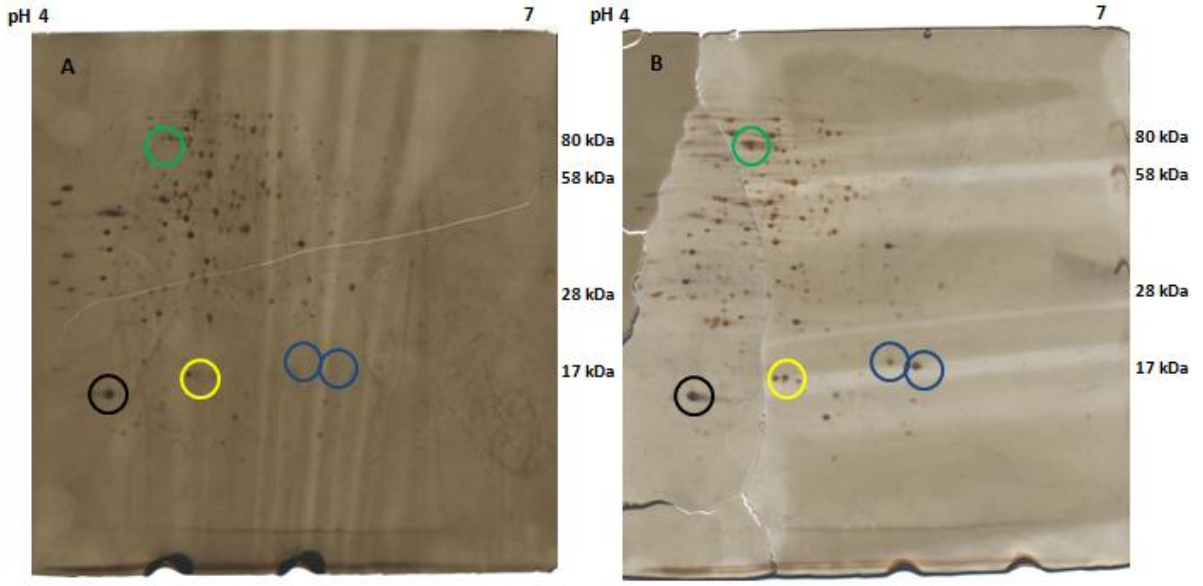

Figure 30 – 2D periplasmatic proteome electrophoresis of *Ma.aq* in 4-7 pH 13 cm Immobiline DryStrip. **A)** Control growth with 0.2 nM $CuSO_4$ (30µg applied); **B)** Copper stress 1.0mM $CuSO_4$ (30µg applied). 12.5% Polyacrylamide gel. Visualization method – silver staining. Legend: Circles highlights regions with main changes comparison.

The visual comparison of these 2D gels enabled the identification of some protein spots that correspond to overexpressed proteins in the presence of copper ions excess. This is the case of two spots located in the low molecular weight region, that appear between 28 and 17 kDa and around a pI of 5.6 (CopX has a pI of 5.57 and a molecular weight of 17.2 kDa, according to its aminoacid sequence), which are not present in the control periplasmic fraction (blue circles). In addition, in the low molecular weight region close to pI 4.5 there are protein spots that seem to be upregulated upon copper stress (yellow circle). Another important spot is the one highlighted with a black circle, which is one of the major spots in both gels and presents a blue color when compared to other protein spots. In the future, this spot could be used as an internal marker.

In the region where it would be expected to detect CopA (pI 4.95 and 61.6 kDa) it is also observed some differential spots (green circle). This observation was corroborated with proteomic analysis using 10% Polyacrylamide gel, to enable a better separation of the high molecular weight proteins. These gels present polymerization problems and are shown in Appendix G (Figure 15 App. G.2). One of these protein spots might correspond to CopA.

The 2D gel electrophoresis of the control periplasmic proteome (0.2 nM $CuSO_4$) (reference gel) and of the copper stress periplasmic proteome (1.0 mM $CuSO_4$) were analyzed with two programs GE 2D ImageMaster 7.0 and Ludesi REDFIN3.

This preliminary comparative analysis identified 346 protein spots in the control gel and 320 protein spots in the copper stress gel. Match spot analysis (with R around 80% - data not shown) found 162 similarities, which covers 47 % of the reference proteome and 51 % of the copper stress proteome. According with the 2.0 fold factor change spot intensity, 74 spots had differential expression upon copper stress, with 28 of them up-regulated and 46 down-regulated.

The low molecular weight region of these gels, analyzed using the 3D view, Figure 31, presents two main differential protein spots in the CopX expected region (black circles): no spot in this region is detected in the reference gel and those two protein spots are only present in the copper stress periplasmic fraction Appendix G (Figure 14 G.2). In this area, there are other differential protein spots that are only present either in the copper stress periplasm proteome (blue arrows) or in the control periplasmic proteome (green arrows).

The R value being around 80% is due to the impossibility of stretching the gel to account for the different gel sizes, which does not allow a perfect gel match. This is the principal problem in achieving quality gels for class analysis and invalidates as well the gel analysis of a large number of spots, such as in the CopA area. So, these preliminary results cannot be further analised, and only when more biological replicates are obtained will it be possible to do the class analysis with statistic significance.

Although, good replicates were not obtained, these other gels corroborate this analysis, specially the differential expression of the two spots in the region of the gel where CopX spot is expected to be (Appendix G (Figure 14 G.2)).

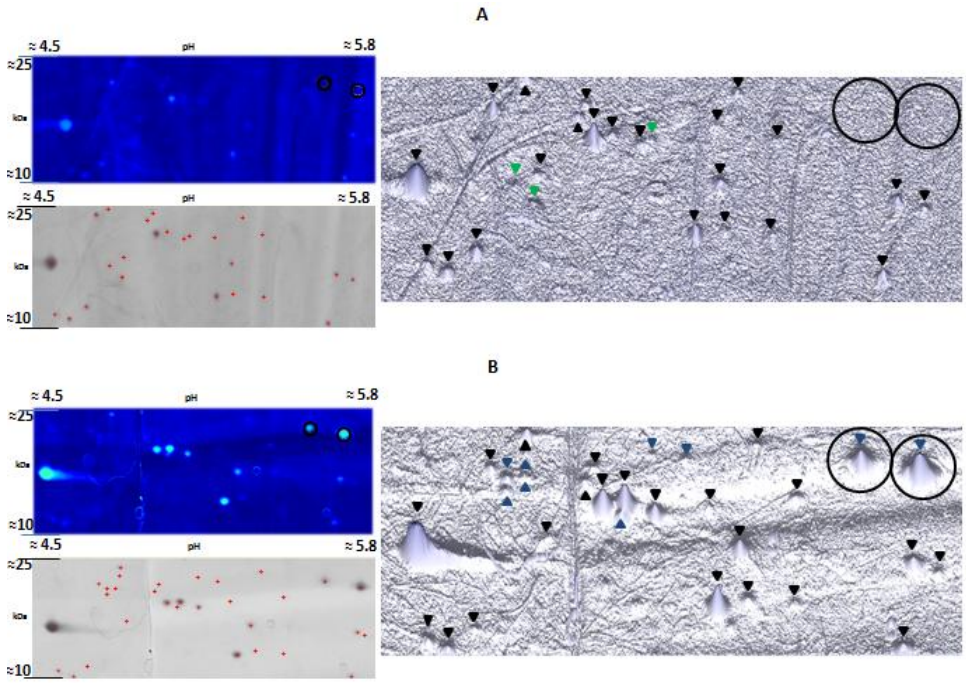

Figure 31 – 3D view analysis of the low molecular weight area of the 2D gels presented in Figure 30. Images were obtained using GE 2D ImageMaster (Gray images) 7.0 and Ludesi REDFIN3 (Blue images). **A)** 0.2 nM $CuSO_4$ periplasmic proteome. **B)** 1.0 mM $CuSO_4$. Legend: Blue arrows correspond to protein spots only present in copper stress periplasmic proteome. Green arrows correspond to protein spots only present in control periplasm. Red plus and Black arrow indicate protein spot. Black circles – expected CopX area.

3.2.1.3 MALDI-TOF-TOF-MS identification

In order to identify the protein spots in the gels, the 2D gel electrophoresis analysis of the copper stress growth in 1.0 mM $CuSO_4$ was repeated this time staining with Coomassie blue, which is compatible with mass spectrometry analysis. An image of this gel is presented in Figure 32.

The protein spots chosen for MS analysis are identified by a red circle, and the mass spectrometry results are presented in Appendix G (Table 9 App. G.3). The protein spots were removed manually from the gel and analyzed by peptide-mass fingerprint as described in Materials and Methods, Chapter 2 –5.4.

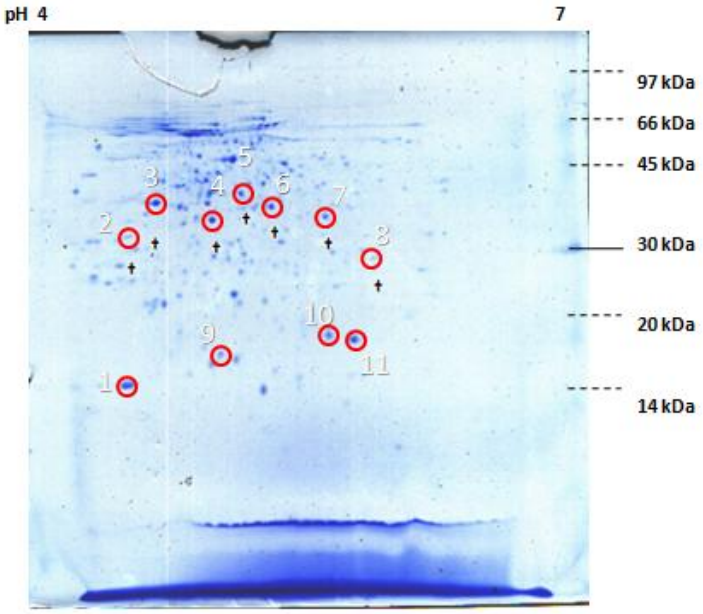

Figure 32 - 2D gel electrophoresis of the periplasmatic proteome of *Ma.aq* grown under copper stress (in the presence of 1 mM $CuSO_4$; 60 µg of protein were applied to a 4-7 pH 13 cm Immobiline DryStrip. 12.5 % polyacrylamide gel. Legend: Dashed lines correspond to molecular weight migration experiments. Red circles indicate the protein spots analyzed. † Indicates protein spots analyzed from another gel (periplasmic fraction of a control growth, 0.2 nM $CuSO_4$ but with 6 times more carbon source) and spot matching was performed using GE ImageMaster and Jvirgel 2.0 approach (Material and Methods, Chapter 2 –5.3). Visualization method – Coomassie blue.

From the 11 spots analyzed, only 9 were identified (Appendix G, Table 9 App. G.3). For spot 1 and 9 no match was obtained, possibly due to keratin contamination (indeed, the analysis of a blank spot found keratin contamination – data not shown). The cellular location of the 9 identified protein spots indicates that no or only small cross-fraction contamination occurred, since all the proteins are periplasmic and one is extra-cellular (spot 5).

In the low molecular weight region, two spots were identified as the product of the gene *copX*, spot 10 with 40 % coverage and spot 11 with 21% coverage. First, these results confirm our hypotesis that CopX is a periplasmic protein only expressed under copper stress. Second, considering that *copX* is part of the *copXAB* operon, it is expected that the product of the two other genes are also present under these growth conditions. However, this needs to be further explored by periplasmatic fraction validation with more biological replicas and further MS/MS analysis, RT-PCR or qRT-PCR experiments.

The presence of two spots both identified as CopX might be due to proteolysis, since no protease inhibitor was added to this cellular extract, or natural processing in *Marinobacter*. In the case of spot 11, the keratin contamination might explain the low score and coverage in the MS/MS identification, and lead to wrong identification.

4. Heterologous expression and purification of CopX

In order to obtain large amounts of CopX to biochemical characterize this protein, the gene that codes for it was cloned into an expression vector and the protein was obtained heterologously in *Es.coli*.

4.1. Construction of the expression plasmid

4.1.1 PCR amplification

The *copX* gene from *Ma.aq* was isolated by PCR. The primers were designed based on the DNA sequence of *Ma.aq* strain VT8 (close related organisms – see Introduction) to amplify the complete gene, including the signal peptide to direct the protein to the periplasm of *Es.coli*, and simplifying the purification procedure. The primers also included restriction enzymes sites in order to clone this gene into the chosen expression vector, pET-21c(+) (for primers see Figure 13, Chapter 2 – 6.1.2).

The PCR products obtained were visualized in a gel electrophoresis using a 1% agarose gel as shown in Figure 33.

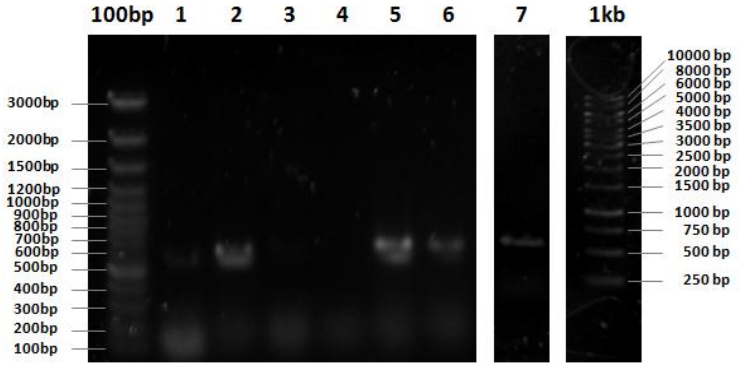

Figure 33 – PCR amplification of the CopX gene from *Ma.aq* using the primers and PCR program described in Material and Methods, Chapter 2 – 6.1.2. Legend: Lane 1-7 loaded with 5 µL of PCR reaction. 100bp and 1kb - DNA ladders. 1% agarose gels. Staining with Sybr Safe visualized under blue light.

The analysis of Figure 33 shows that the DNA amplification was successful since there is a band between 500 bp and 600 bp corresponding to the *copX* gene, which has 534 bp. In the low molecular weight region there is a smear corresponding to primers dimers and primers that were not used in the PCR reaction.

The PCR product was purified for further manipulation.

4.1.2 Digestion with the restriction enzymes

The purified PCR product and the chosen expression vector were digested with the selected restriction enzymes, *Nde*I (N-terminal) and *Not*I (C-terminal) that only cut the vector once in the multi-cloning site. The selected restriction enzymes were also used because they do not cut inside the *copX* gene, and places *copX* gene under the control of the T7 RNA polymerase, in frame with the ATG site (Appendix D Figure 8 App.D.2 an Figure 9 App.D.3).

A digestion test was performed to determine the number of hours required for a complete digestion without the observation of star activity. The results obtained in the digestion test (Figure 34A) show that after 1 hour the digestion is completed, by comparing the control lane, undigested supercoiled pET21c – between 3500 and 4000 bp , with the digestion of pET21c at different incubation times (between 5000 and 6000 bp - linear form (5441 bp). Nevertheless, to ensure a complete digestion both samples were incubated during 4 h, since the time required for the complete digestion of a PCR fragment is larger than the one for a circular DNA. After the 4 h incubation, both the digested PCR fragment and the digested pET21c vector were purified (Figure 34B and C). Afterwards the digested purified vector was treated with alkaline phosphatase to avoid recirculation.

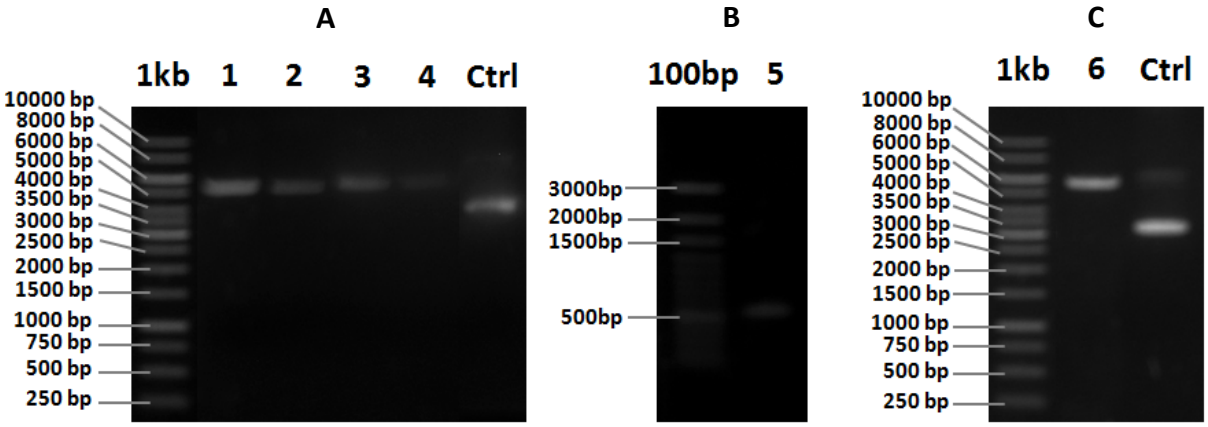

Figure 34 – A) Double digestion test of pET21c over time**,** with *Not*I and *Nde*I. Legend: 1kb – DNA ladder; 1 – Incubation time 1h; 2 – Incubation time 2h; 3 – Incubation time 3h; 4 – Incubation time 4h; Ctrl – Control undigested pET-21c(+). **B)** Purified double digested PCR product, *copX***,** with *Not*I and *Nde*I. Legend: Incubation time 4h. 100bp – DNA ladder; 5 – *copX* double digested. **C)** Double digestion of pET-21c(+) with *Not*I and *Nde*I. Legend: Incubation time 4h. 1kb – DNA ladder; 6 –pET-21c(+) double digested; Ctrl – Control undigested pET-21c(+). All gel electrophoresis were run in 0.8 % agarose gel, stained with SybrSafe and visualized under blue light.

4.1.3 Ligation

The digested PCR product *copX* was inserted into the linear pET21-c(+) using T4 DNA ligase (see Material and Methods - Chapter 2 – 6.2.2 for conditions). The resulting vector was pET-CopX with 5975 bp, that is shown schematically in Figure 35.

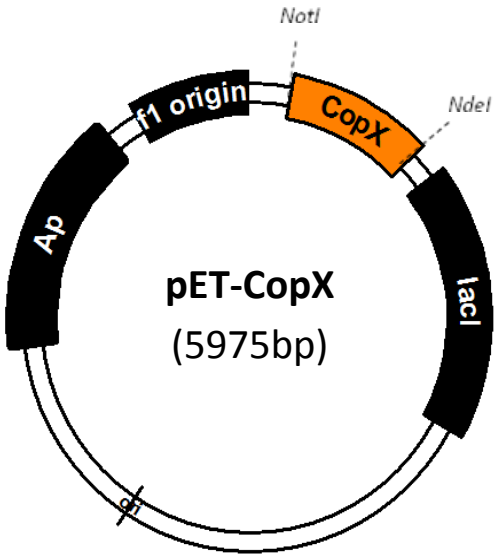

Figure 35 – Cloning vector pET-CopX (5975bp). Legend: DNA insert in orange, CopX (534bp). Cloning at the multiple cloning site, between *Not*I (located at 166 bp in pET21-c(+)) and *NdeI* (located at 238 bp in pET21-c(+)). Image was created using pDRAW32 (version 1.1.106) [165].

4.1.4 Isolation of expression vector pET-CopX

The resulting vector, pET-CopX, was transformed into a non-expression host, *Es. coli* Gigablue, and 10 of the colonies obtained were used to inoculate LB medium supplemented with 100 μg/mL ampicillin for plasmid amplification and isolation (Figure 36).

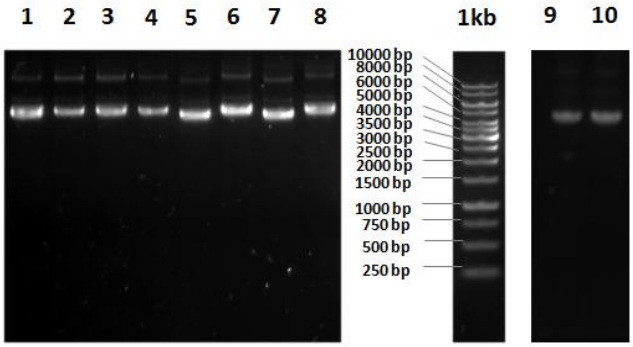

Figure 36 - Screening pET-CopX clones. Stained with SybrSafe and visualized under blue light. Legend: 1 to 10 – Colony number. 1kb - DNA ladders. 1% Agarose gel (image correspond to two different gels with colonies of the same plate).

The plasmids were digested with the restriction enzymes used in the cloning, *Nde*I and *Not*I, in order to check if the DNA fragment corresponding to *copX* was successfully inserted into the expression plasmid (Figure 37). The analysis of Figure 37 shows that the clones 2 to 5, 7 and 10 were positive clones, as a DNA band between 500 and 600 bp is observed; this band corresponds to *copX* (534 bp).

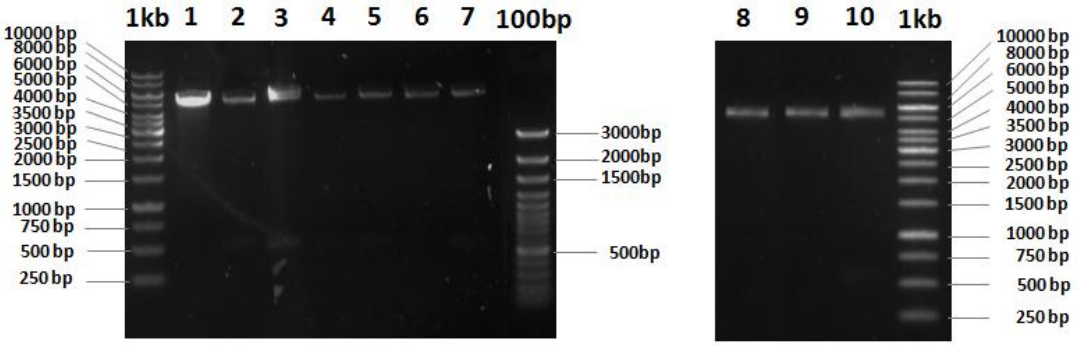

Figure 37 – Positive clones screening. Stained with SybrSafe and visualized under blue light. Legend: 1 to 10 – Colony number. 100bp and 1kb - DNA ladders. 1% Agarose gel (image correspond to two different gels with colonies of the same plate).

The selected positive clones (2, 4 and 10) were sent for DNA sequencing to check the DNA sequence of the inserted DNA fragment, *copX*.

4.1.5 DNA sequencing

The DNA sequence of the three selected clones is shown in Figure 38 and the translated sequence in Figure 39.

```
1 ATGAATGCATCAAACCTGATTGTGGCCGGCCTGGCTTTTTCGATGTCGGTATCGGCGTTT 60
2 ATGAATGCATCAAACCTGATTGTGGCCGGCCTGGCTTTTTCGATGTCGGTATCGGCGTTT 60
3 ATGAATGCATCAAACCTGATTGTGGCCGGCCTGGCTTTTTCGATGTCGGTATCGGCGTTT 60
X ATGAATGCATCAAACCTGATTGTGGCCGGCCTGGCTTTTTCGATGTCGGTATCGGCGTTT 60
  ************************************************************

1 GCGGCTGGCACCCATGGCGGTGGTCACGGCCATGGTGCTTCCATTGGCGAACCGGGAAAA 120
2 GCGGCTGGCACCCATGGCGGTGGTCACGGCCATGGTGCTTCCATTGGCGAACCGGGAAAA 120
3 GCGGCTGGCACCCATGGCGGTGGTCACGGCCATGGTGCTTCCATTGGCGAACCGGGAAAA 120
X GCGGCTGGCACCCATGGCGGTGGTCACGGCCATGGTGCTTCCATTGGCGAACCGGGAAAA 120
  ************************************************************

1 GTCTCAGAAGCCAGCCGGACCATAACGGTTGAAATGCACGACAACTACTACGAACCCGAA 180
2 GTCTCAGAAGCCAGCCGGACCATAACGGTTGAAATGCACGACAACTACTACGAACCCGAA 180
3 GTCTCAGAAGCCAGCCGGACCATAACGGTTGAAATGCACGACAACTACTACGAACCCGAA 180
X GCCTCAGAAGCCAGCCGGACCATAACGGTTGAAATGCACGACAACTACTACGAACCCGAA 180
  * ************************************************************

1 GAAATCCGTGTGAAGCCGGGCGAAACCGTTCGGTTTGTGGTGCAGAACAAGGGCAACCTT 240
2 GAAATCCGTGTGAAGCCGGGCGAAACCGTTCGGTTTGTGGTGCAGAACAAGGGCAACCTT 240
3 GAAATCCGTGTGAAGCCGGGCGAAACCGTTCGGTTTGTGGTGCAGAACAAGGGCAACCTT 240
X GAAATCCGTGTGAAGCCGGGCGAAACCGTTCGGTTTGTGGTGCAGAACAAGGGCAATCTT 240
  ************************************************************ ***

1 GTGCACGAGTTCAACATCGGGACTCCGGGTATGCATGAGGCCCACCAGAAAGAAATGAGA 300
2 GTGCACGAGTTCAACATCGGGACTCCGGGTATGCATGAGGCCCACCAGAAAGAAATGAGA 300
3 GTGCACGAGTTCAACATCGGGACTCCGGGTATGCACGAGGCCCACCAGAAAGAAATGAGA 300
X GTGCACGAGTTCAACATCGGGACTCCGGGTATGCATGAGGCCCACCAGAAAGAAATGAGA 300
  ********************************** ************************

1 ATGATGGTCGAGCATGGCGTGATCCAGGGTAACAAGTTGAATCATGACATGATGAACATG 360
2 ATGATGGTCGAGCATGGCGTGATCCAGGGTAACAAGTTGAATCATGACATGATGAACATG 360
3 ATGATGGTCGAGCATGGCGTGATCCAGGGTAACAAGTTGAATCATGACATGATGAACATG 360
X ATGATGGTCGAGCATGGCGTGATCCAGGGTAACAAGTTGAATCATGACATGATGAACATG 360
  ************************************************************


1 GACATGGGCAACGGCCACTCGATGAAACATGACGATCCCAACAGCGTCCTGCTGGAGCCG 420
2 GACATGGGCAACGGCCACTCGATGAAACATGACGATCCCAACAGCGTCCTGCTGGAGCCG 420
3 GACATGGGCAACGGCCACTCGATGAAACATGACGATCCCAACAGCGTCCTGCTGGAGCCG 420
X GACATGGGCAACGGCCACTCGATGAAACATGACGATCCCAACAGCGTCCTGCTGGAGCCG 420
  ************************************************************


1 GGCCAGAGCCGGGAAGTGGTCTGGACGTTTGCCAATCAGGGCAATATCGAATTCGCATGT 480
2 GGCCAGAGCCGGGAAGTGGTCTGGACGTTTGCCAATCAGGGCAATATCGAATTCGCATGT 480
3 GGCCAGAGCCGGGAAGTGGTCTGGACGTTTGCCAATCAGGGCAATATCGAATTCGCATGT 480
X GGCCAGAGCCGGGAAGTGGTCTGGACGTTTGCCAATCAGGGCAATATCGAATTCGCATGT 480
  ************************************************************
```

```
1 AACGTACCAGGGCACTACCAGTCTGGCATGTACGGTGACGTGAATTTCGAATAA 534
2 AACGTACCAGGGCACTACCAGTCTGGCATGTACGGTGACGTGAATTTCGAATAA 534
3 AACGTACCAGGGCACTACCAGTCTGGCATGTACGGTGACGTGAATTTCGAATAA 534
X AACGTACCAGGGCACTACCAGTCTGGCATGTACGGTGACGTGAATTTCGAATAA 534
  **************************************************
```

Figure 38 – Alignment of the nucleotide sequences obtained by the DNA sequencing (sequencing was performed by StabVida (www.stabvida.net)) from the 3 selected clones (1, 2 and 3) against the reference gene, CopX (X) retrived from the *Ma.aq* VT8 chromosomal genome. Alignment was performed using Clustal X (version 2.0) [116]. Legend: Asterisk – Identical nucleotides in all the alignment.

4.1.5.3. Protein translation

```
1 MNASNLIVAGLAFSMSVSAFAAGTHGGGHGHGASIGEPGKVSEASRTITVEMHDNYYEPE 60
2 MNASNLIVAGLAFSMSVSAFAAGTHGGGHGHGASIGEPGKVSEASRTITVEMHDNYYEPE 60
3 MNASNLIVAGLAFSMSVSAFAAGTHGGGHGHGASIGEPGKVSEASRTITVEMHDNYYEPE 60
X MNASNLIVAGLAFSMSVSAFAAGTHGGGHGHGASIGEPGKASEASRTITVEMHDNYYEPE 60
  ************************************.******************

1 EIRVKPGETVRFVVQNKGNLVHEFNIGTPGMHEAHQKEMRMMVEHGVIQGNKLNHDMMNM 120
2 EIRVKPGETVRFVVQNKGNLVHEFNIGTPGMHEAHQKEMRMMVEHGVIQGNKLNHDMMNM 120
3 EIRVKPGETVRFVVQNKGNLVHEFNIGTPGMHEAHQKEMRMMVEHGVIQGNKLNHDMMNM 120
X EIRVKPGETVRFVVQNKGNLVHEFNIGTPGMHEAHQKEMRMMVEHGVIQGNKLNHDMMNM 120
  ************************************************************

1 DMGNGHSMKHDDPNSVLLEPGQSREVVWTFANQGNIEFACNVPGHYQSGMYGDVNFE 177
2 DMGNGHSMKHDDPNSVLLEPGQSREVVWTFANQGNIEFACNVPGHYQSGMYGDVNFE 177
3 DMGNGHSMKHDDPNSVLLEPGQSREVVWTFANQGNIEFACNVPGHYQSGMYGDVNFE 177
X DMGNGHSMKHDDPNSVLLEPGQSREVVWTFANQGNIEFACNVPGHYQSGMYGDVNFE 177
  ********************************************************
```

Figure 39 – Alignment of the amino acid sequences obtained by the translation of the sequenced DNA sequencing selected (Expasy translate tool [162]) from the 3 selected clones (1, 2 and 3) against the reference gene, CopX (X) retrived from the *Ma.aq* VT8 chromosomal genome. Alignment was performed using Clustal X (version 2.0) [116]. Legend: Asterisk – Identical residues in all the alignment. Dot – Semi-conserved substitutions.

The alignment of the translated DNA sequence with the one of CopX, retrieved from the *Ma.aq* VT8 genome, shows that as expected the gene is inserted in the correct orientation and the DNA sequence of the cloned *cop*X was correct. However, it was detected 2 silent mutations (cytosine to thymine in positions 237 and 257) and a non silent mutation (thymine to cytosine in position 122) that corresponds to a mutation in residue 41, an

alanine was substituted by a valine. These changes were observed in the DNA sequence of all the clones.

The differences can be attributed to the fact that we are comparing the DNA sequence of *Ma.aq* 617 *copX* with the one of *Ma.aq* VT8. Nevertheless, these results strength the proposal that *Ma.aq* 617 (also known as *Ps.nautica* 617) should be classified as a strain of *Marinobacter aquaeolei*.

4.2. Heterologous expression

One of the sequenced expression vectors pETCopX, was used to transform the expression host *Es.coli* BL21(DE3). Two types of growth media were used for protein production, a rich medium Luria-Broth and a minimum medium M9.

In order to establish the growth condition for maximum protein production, different induction time (0 to overnight) were tested using 1 mM IPTG, as *copX* gene is under the control of a T7 lac promoter (Appendix D, Figure 8 D.2). Both rich and minimum media were tested, and the growth was performed in small scale (500 mL), as described in Material and Methods Chapter 2 - 6.3.2. Expression profile in both media is summarized in Figure 40.

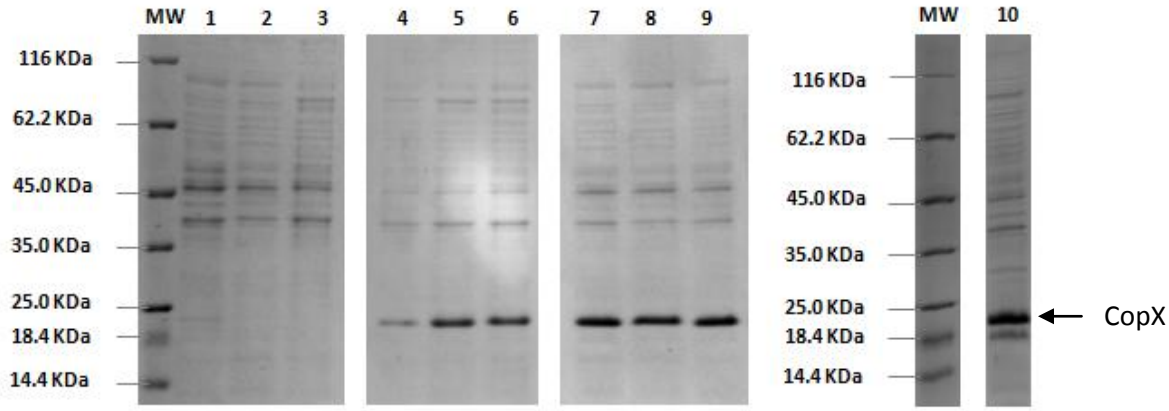

Figure 40 – SDS-PAGE analysis of the expression profile of CopX in two growth media (LB and M9). Legend: MW – Molecular weight marker; 1 – Inoculum; 2 – LB before induction; 3 – M9 before induction; 4 – M9 2h IPTG induction; 5 – M9 4h IPTG induction; 6 – M9 7h IPTG induction; 7 - LB 2h IPTG induction; 8 - LB 4h IPTG induction; 9 - LB 7h IPTG induction; 10 – LB overnight IPTG induction. 12.5 % Polyacrylamide gel, stained with Coomassie blue.

The protein production was induced with IPTG and no protein expression was observed before induction (Figure 40, lane 2,3), although in the cells grown overnight it is

observed a low intensity band at the same molecular weight as the one observed after induction (between 18.4 kDa and 25 kDa) (Figure 40, lane 1).

Since the amount of cells loaded onto the gel was normalized, Figure 40 shows that in the growth performed in minimum medium, the amount of CopX does not increase after 5 h induction, while in the case of rich medium it seems that the amount of CopX is larger after 20 h induction. However, under these conditions another protein band appears at approximately 18.4 kDa (below the one corresponding to CopX), which might correspond to proteolyzed or truncated CopX.

Therefore, the duration of the induction was chosen to be 7h. Another reason to chose this induction time for both growth media, is that it is known that IPTG can fragile the cell wall, releasing the protein to the growth medium [148].

It is also important to point out that the protein band that appears in the gel, after induction with IPTG, has a molecular weight higher than 18.4 kDa, which is higher than expected, as the molecular weight of CopX after processing the signal peptide is 17.2 kDa, and in case the signal peptide is not processed 19.3 kDa. This will be further discussed after protein purification and mass spectrometry analysis, since this result seems to indicate that the signal peptide is not being processed even if CopX is present in the periplasm of *Es.coli*, as shown below.

4.3. Protein purification

The gene coding for CopX was cloned with its own signal peptide to direct the protein to the periplasm. Therefore, CopX was purified from the periplasmic extract, in order to simply the purification, avoiding the contamination with all the cytoplasmic proteins. The periplasmic extract was obtained by 5 freeze-thaw cycles, as described in Material and Methods, Chapter 2 – 6.4.

In order to verify whether this procedure was enough to release all the periplasm and to verify whether there was protein either in the cytoplasm or in insoluble membrane fraction, the cells were disrupted in a French Press after obtaining the periplasm.

The periplasmic, cytoplasmic extract and insoluble membrane fraction were analyzed by SDS-PAGE (Figure 41). The Figure shows that almost all protein is in the soluble periplasmatic fraction, and CopX is not present in the cytoplasmatic extract. Nevertheless, it

is possible to observe some CopX in the insoluble fraction (either membrane or inclusion bodies).

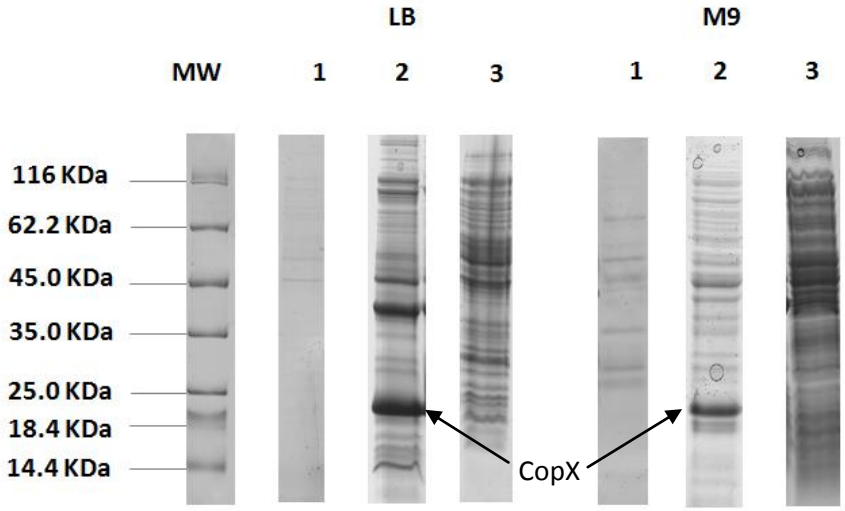

Figure 41 - SDS-PAGE analysis of the periplasmic, cytoplasmic and membrane/insoluble fractions prepared from growths in rich and minimum media. Legend: MW – Molecular weight marker; 1 – Cytoplasmatic fraction; 2 – Periplasmatic fraction; 3 – Insoluble fraction. 12.5 % Polyacrylamide gel, stained with Coomassie blue

The presence of CopX in the periplasmic fraction was also verified through UV-visible spectroscopy, attending to the fact that it might bind copper ions, and small copper proteins have a characteristic absorbance band around 600 nm (Figure 42).

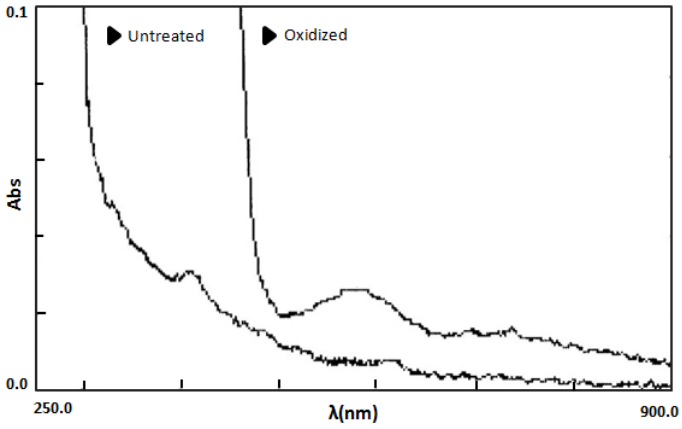

Figure 42 –UV-visible spectra of the periplasmatic fraction obtained from a growth in LB medium with 0.5 mM $CuSO_4$ added in the inoculums moment, untreated fraction and oxidized with potassium ferricyanide.

This spectra show that CopX is in the reduced state and an absorption band appears at 580 nm after addition of potassium ferricyanide, characteristic of small blue copper binding proteins. It is important

to point out that the protein is in the reduced form through out the purification, and only during the fraction analysis it was oxidized with potassium ferricyanide.

CopX was purified from the periplasmatic fraction following the three-step procedure described in Material and Methods Chapter 2 – 6. The first step was an anionic exchange chromatography, as CopX pI determined by its aminoacid sequence is 5.57 (expasy protparam), which means that it will have a negative charge at pH 7.6. An example of the profile is presented below (Figure 43). The fractions obtained were analyzed by SDS-PAGE showing that CopX was eluted at 285 mM NaCl.

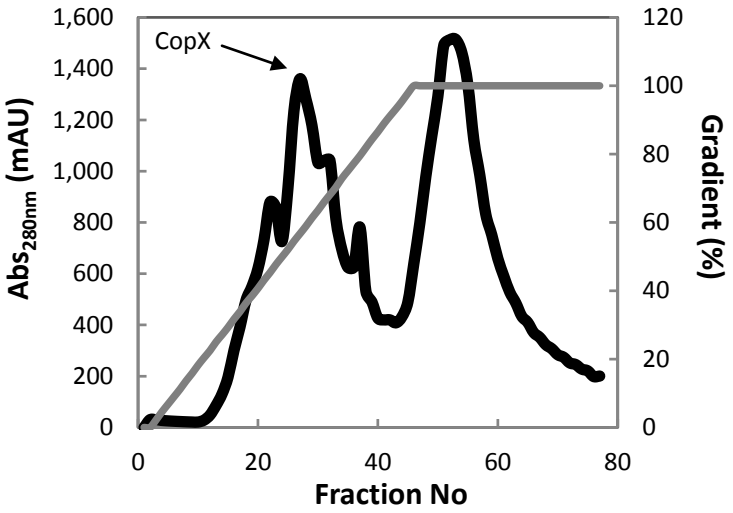

Figure 43 — Elution profile of the first step of CopX purification, Source-Q 15. Legend: Black line - Protein content followed at 280nm. Gray line - Linear gradient in 10 mM Tris-HCl, pH7.6, from 0 to 500 mM NaCl. CopX eluted at 285 mM NaCl (highlighted with an arrow).

After the first step the CopX is contaminated with other proteins, mainly with higher molecular weight protein (Figure 46, lane S15) and therefore, the second chromatographic step was a size-exclusion chromatography. The chromatogram obtained can be found in Figure 44. This second purification step was shown to be very efficient as CopX purity improved considerably, as can be observed by the analysis of the SDS-PAGE (Figure 46, lane S75) and in the purification Table 10.

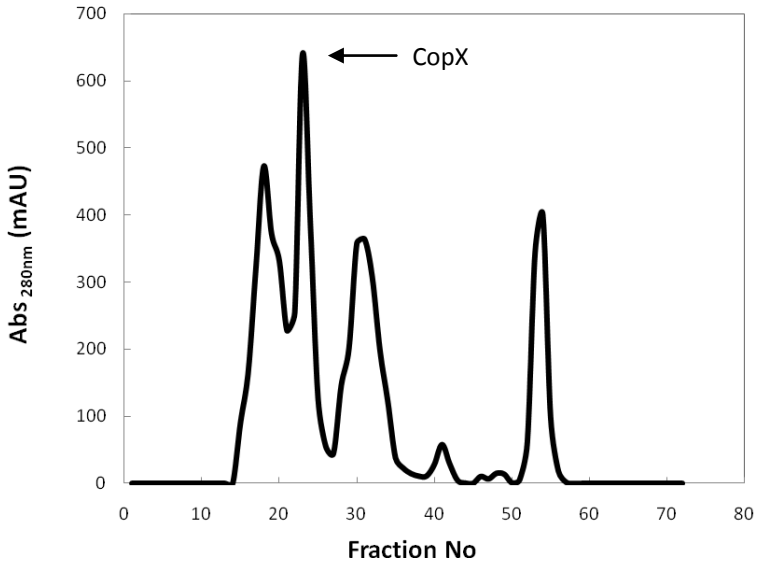

Figure 44 – Elution profile of the second chromatographic step of CopX purification, Superdex 75, eluted with 50 mM Tris-HCl, pH7.6, 150 mM NaCl. Protein content followed at 280nm. CopX peak is highlighted with an arrow.

In some of the purifications that were performed during this work, the protein was pure after this second chromatographic step. However, in some cases, like the one presented here a third chromatographic step was performed, using a pre-packed anionic exchange column. The elution profile obtained is the one shown in Figure 45, which shows that CopX was eluted at 285 mM NaCl. The protein was considered to be pure by the analysis of the SDS-PAGE (Figure 46, lane RQ).

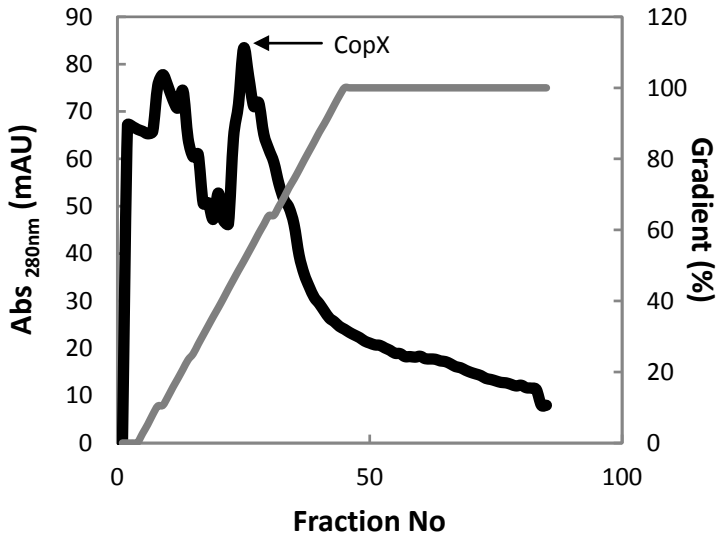

Figure 45 – Elution profile of the third step (optional) of CopX purification, ResourceQ 1 mL. CopX peak is highlighted with an arrow. Legend: (Gray line) linear gradient in 20 mM Tris-HCl, pH7.6, from 0 to 500 mM. (Black line) Protein content followed at 280 nm. CopX eluted at 285 mM NaCl.

The SDS-PAGE results obtained during the purification are summarized in Figure 46.

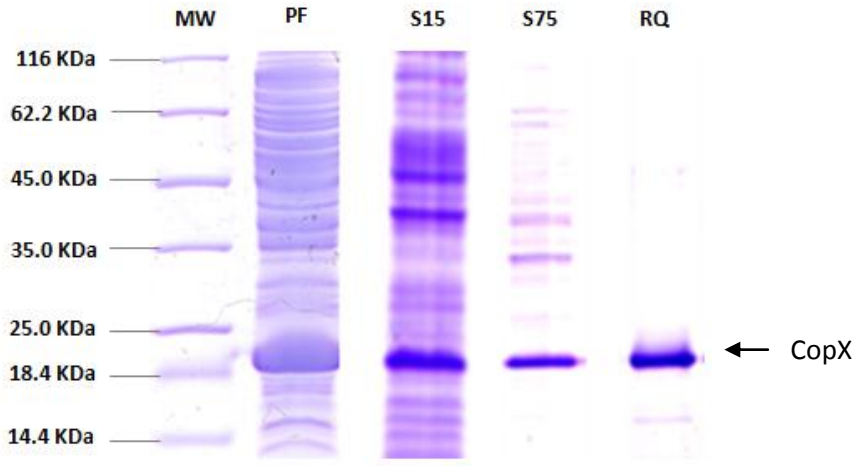

Figure 46 – SDS-PAGE analysis of the different purification steps. Legend: MW – Molecular weight marker. PF – Periplasmatic fraction contained after expression of pETCopX in *Es.coli* Bl21(DE3). S15 – First chromatographic step, Source -15 column; S75 – Second chromatographic step, Superdex 75 column; RQ – Third chromatographic step, resource-Q column. 12.5% Polyacrylamide gel, stained with Coomassie blue.

The purification Table is related to a specific purification done in this work, with the $CuSO_4$ added in the inoculum moment, demonstrating how the purification ratio evolves during the purification stages. In the case of minimum media (M9) the third chromatographic step, resource-Q, was never performed (Table 10). The recovery was around 30% in both cases which is low, but normal since three chromatographic steps were used and the purification started with a small amount of protein.

Table 10 – Purification Table with the amounts of CopX recovered at the different stages of purification, for both media per L of growth with the $CuSO_4$ added in the inoculum moment. Legend: n/p – not performed.

	LB			M9		
Stages of purification	**CopX† (nmol)**	**Purity Ratio A_{580nm}/A_{280nm}**	**Recovery (%)**	**CopX† (nmol)**	**Purity Ratio A_{580nm}/A_{280nm}**	**Recovery (%)**
Periplasm	970	0.05	100	680	0.02	100
Source-15	738	0.10	76	464	0.04	68
Superdex 75	463	0.32	48	211	0.11	31
Resource-Q	298	0.40	30	n/p	n/p	n/p

†These values were determined after establishing the extinction coefficient at 580 nm.

The amount of pure CopX purified from the periplasmic extract was between 210 nmol and 470 nmol per L of growth in LB medium. In minimum medium the range of purified periplasmic extract was between 190 and 211 nmol per L of CopX.

The expression of CopX in mininum medium is lower than the one in LB medium as expected by the expression tests (Figure 40). The protein obtained in minimum medium has a low purity ratio when compared with the one obtained from LB growth, but was considered pure in the SDS-PAGE (not shown). This can be explained if we considered that there is apo-CopX, although $CuSO_4$ was added to the growth upon inoculation. The presence of apo-CopX is considered due to the low Cu/protein ratio of this sample (see also below).

5. Biochemical characterization

CopX was isolated to study its biochemical properties: mass spectrum, extinction coefficient, U.V.-visible spectra, EPR spectrum, and quaternary structure. Preliminary NMR studies were also performed to determine whether the protein was amenable to solution structure determination.

5.1 N-terminal sequence

In order to confirm that the purified protein was CopX, the N-terminal of the heterologously isolated CopX was sequenced (Figure 47).

```
     1   5     10     15
     |   |      |      |
 i)  AGTHG  MMHGH  GANXX
ii)  KHNDP  NSNLD  MPS
       DV     I     ENK
                    Y
```

Figure 47 – N-terminal sequence of the purified CopX. Legend: Underlined residues correspond to uncertained residues and X - amino-acids correspond to failed sequenced residue.

The N-terminal sequencing identified the presence of two polypeptides. The N-terminal sequence i) corresponds to the CopX without the predicted signal peptide, indicating that this sequence was correctly processed by *Es.coli* machinery (Figure 48).

MNASNLIVAG LAFSMSVSAF AAGTHGGGRG HGASIGEPGK ASEASRTITV EMHDNYYEPE 60
EIRVKPGETV RFVVQNKGNL VHEFNIGTPG MHEAHQKEMR MMVEHGVIQG NKLNHDMMNM 121
DMGNGHSMKH DDPNSVLLEP GQSREVVWTF ANQGNIEFAC NVPGHYQSGM YGDVNFE 177

Figure 48 – CopX residue sequence. Legend: Gray highlights predicted signal peptide identified using SignalP 3.0 server [163]. Green highlights the N-terminal sequence i), in red bold correspondent residues; Brown bold highlights N-terminal sequence ii).

The N-terminal sequence ii) corresponds to a peptide in the middle of CopX sequence (Figure 48), identified by a BLAST search against CopX residue sequence. This result implies that there was proteolysis during the purification (even though protease inhibitors were present during all the purification).

Attending to the fact that CopX suffers proteolysis at that residue 128 (numbering using the full length), the two expected polypeptides are: one with 107 residues, from residue 21 to residue 128 and another with 48 residues, from residue 129 to 177. The expected molecular weight of these two polipeptides is 12 kDa and 5.2 kDa, respectively. Indeed, in the SDS-PAGE of Figure 46, lane RQ, is possible to observe the presence of a second protein band that corresponds to a protein with approximately 15 kDa.

CopX has an abnormal mobility on a 12.5% SDS-Tricine-PAGE (higher than 18.2 kDa), since the expected molecular weight of this protein, without its signal peptide, is 17.2 kDa. Similarly, the truncated CopX at the C-terminal (with an expected 12 kDa), migrates as a 15 kDa protein. This abnormal migration might be due to conformation or acidic nature of part of the protein.

In conclusion, the N-terminus analysis indicates that the purification was successful and that the signal peptide of CopX is composed of 21 residues and is recognized and processed by the Sec pathway of *Es.coli*.

5.2 UV-Visible spectra

UV-visible spectroscopy is a valuable biochemical technique for characterizing metalloproteins. It observes electronic transitions between energy levels, transitions that are relevant to metalloproteins are mostly the ones between the d orbitals of the metal and those between orbitals of the ligand and the metal (ligand to metal and metal to ligand

charge transfer transitions). Of those, the former is not an allowed transition and is very weak, while the later are allowed and are characterized by very high extinction coefficients [166]. Metalloproteins that exhibit such transitions can be identified in a UV-visible spectrum, even at low concentrations and can be used to classify them in different families.

The CopX UV-visible spectrum of the as-isolated protein presents bands at 440 nm and 580 nm and 720 nm with very low absorbance (Figure 49 - green line). The absorbance of this bands increases upon oxidation with potassium ferricyanide and disappear upon reduction with sodium ascorbate, demonstrating that purified CopX (untreated) is mainly in the reduced state (73.2%).

The three absorbance bands, at 440 nm, 580 nm and 720 nm, are characteristic of type 1 copper protein with a tetrahedric copper center. The bands at 440 nm and 580 nm have a charge transfer character (S_{cys} – Cu ion; 580 nm: CysS $\pi \geqslant$ Cu $x^2 - y^2$ and 440 nm: CysS $\sigma \geqslant$ Cu x^2-y^2) [167].

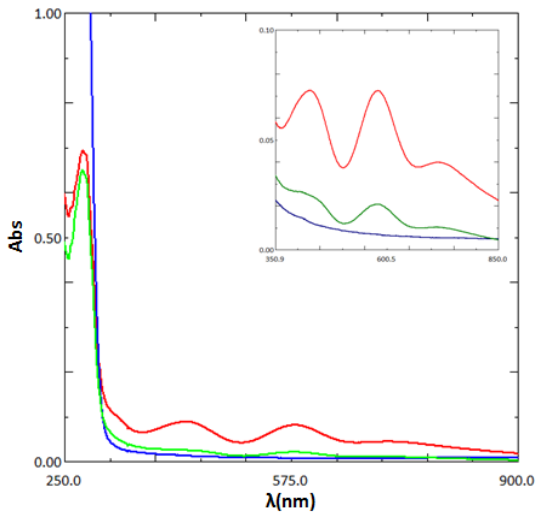

Figure 49 – UV-visible spectra of 27.3 µM heterologously expressed CopX, in 10 mM Tris-HCl, pH 7.6. Legend: (Green line) Untreated CopX; (Blue line) Reduced CopX with sodium ascorbate; and (Red line) Fully oxidized with potassium ferricyanide.

The UV-visible spectrum also revealed a new feature in type 1 blue copper proteins, the A_{440nm}/A_{580nm} of 0.94, which is higher than the value reported for pseudoazurin (0.38), a type 1 copper protein, which has a similar visible spectrum (with absorption bands at 450 nm, 590 nm and 720 nm) [112, 168]. This feature might indicate that the geometry of the copper center is different.

The pure CopX has a purity ratio between the absorbance at 580 nm, of the fully oxidized protein, and the one at 280 nm, of the untreated protein, above 0.4 for the fully holo-CopX.

5.2.1 Determination of the extinction coefficient of CopX

The extinction coefficient of CopX was determined by total protein and by copper ion, at 580 nm. This also enables the determination of the number of copper ions bound to the polypeptide chain of CopX.

The total protein, copper and zinc concentration were determined using the BCA kit and atomic absorption, respectively, as described in Material and Methods, Chapter 2 – 4.1 and 8.2. The values obtained for a pure CopX with a purity ratio of 0.2 are presented in Table 11.

Table 11 – Values of the extinction coefficient (mM^{-1} cm^{-1}) at 580 nm for heterologous expressed CopX. The A_{580nm} of the sample was 0.087.

	Concentration (mM)	ε_{580nm} (mM^{-1} cm^{-1})
Zn (ICP)	0.003	**-**
Cu (ICP)	0.023	**3.8**
Total Protein (BCA)	0.030	**2.9**

The results presented show that the purified CopX has 0.90:0.10 Cu:Zn bound to the polypeptide chain. Therefore, even with the addition of copper ions in the beginning of the growth, it was not possible to obtain CopX completely in the holo-form.

Those values also indicate that CopX binds 1 copper per polypeptide chain, and we could explain the presence of 10% apo-protein as the need of a chaperone to deliver copper to this protein, which is not present in Es.coli. The CopX purified from minimum media has a purity ratio of 0.11, but is considered pure by its SDS-PAGE, hence this low value can be attributed to higher amount of CopX in the apo-form.

The value obtained for the extinction coefficient at 580 nm, determined by copper is similar to the ones reported for other type 1 copper proteins [168].

5.3 EPR spectroscopy

The EPR spectrum of CopX is an axial spectrum (Figure 50), that can be simulated with the parameters presented in Table 12.

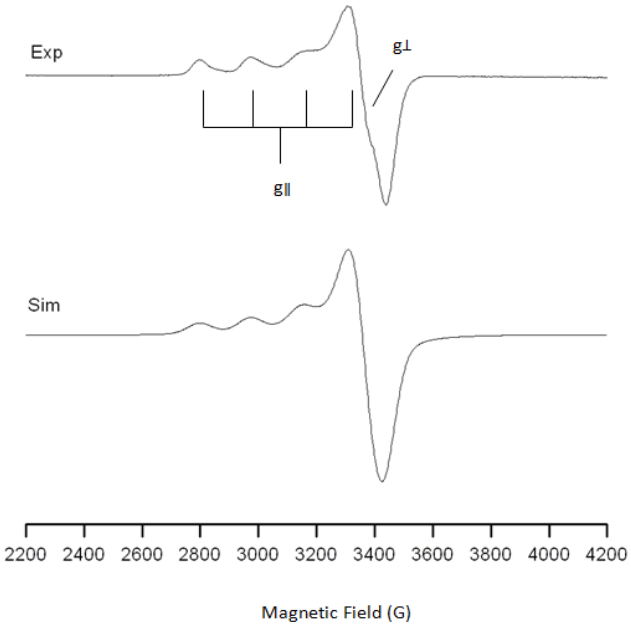

Figure 50 – EPR spectra of 0.83 mM heterologous expressed oxidized CopX in 10 mM Tris-HCl pH 7.6. Legend: Exp – Experimental result obtained using WinEPR (version 2.11) [169]; Sim – Simulated EPR spectrum obtained using and SimFonia (version 1.25) [170]. Experimental conditions: temperature 70K, 1 scan, microwave frequency 9.65 GHz, gain 10^5, attenuation 40 dB and modulation amplitude 30 mT.

The interesting feature of this spectrum is its hyperfine coupling constant (A = 170G), which is higher than the one reported for other type I copper proteins, such as plastocyanins, pseudoazurins, azurins or amicyanins [168].

Table 12 – Values used to simulate the CopX EPR spectrum.

g value	A value (G)	Line width (G)
$g_{\parallel}$ = 2.256	$A_{\parallel}$ = 170	80
$g_{\perp}$ = 2.052	-	80

Type 1 copper proteins can be classified into two groups based on their EPR spectrum, the rhombic – perturbed blue copper site, and axial – classic blue copper site.

Another feature is the small hyperfine-coupling constant. The axial spectrum is a characteristic of the classic blue proteins, which present a small UV-visible band in the 450 nm region. The perturbed blue copper proteins, present a rhombic spectrum related with an increase of the 450 nm region and the S_{Met}-copper bond [168].

Taking into account that the A_{440nm}/A_{580nm} ratio of CopX is 0.94, the EPR spectrum of CopX was not expected to be axial, but rhombic. Moreover, the hyperfine-coupling constant is higher than the one reported for type 1 copper proteins. These features strength the idea that CopX belongs to a new class of type 1 copper proteins.

Up-to-date, there are two other isolated proteins that present a A_{440nm}/A_{580nm} ratio close to 1. One is a small blue copper protein isolated from the methylotrophic *Hyphomicrobium denitrificans* [171], which has an EPR spectrum similar to other pseudoazurins and with low sequence homology to CopX. The other is a *Pseudomonas putida* methionine-rich azurin-like protein heterologously isolated, with a high redox potential (456 ± 4 mV), which is proposed to be the electron donor of a queuosine reductase [172]. The EPR spectrum of this azurin-like protein is not shown, but it has a high protein sequence homology to CopX (46 % similar), and was identified in the gene ortholog neighborhoods shown in Appendix A – Figure 1 App.A.1. However, it is important to point out that the other gene present in the operon of the azurin-like protein has no homology to any of the genes *copA* or *copB* (Appendix A – Figure 1 App.A.1).

5.4. Molecular weight determination

The molecular weight of CopX was determined by mass spectrometry and molecular exclusion chromatography. The expected molecular weight based on its aminoacid sequence (expasy protparam) is 17226.1 Da, considering that the signal peptide was processed, as it was determined by N-terminus analysis.

5.4.1 Mass spectrometry

The mass spectrum of CopX was determined in denaturating conditions by Electrospray ionization mass. This spectrum shows the presence of two species, one with a molecular weight of 17253.0 ± 0.3 Da and another with 17315.50 ± 0.1 Da (Figure 51).

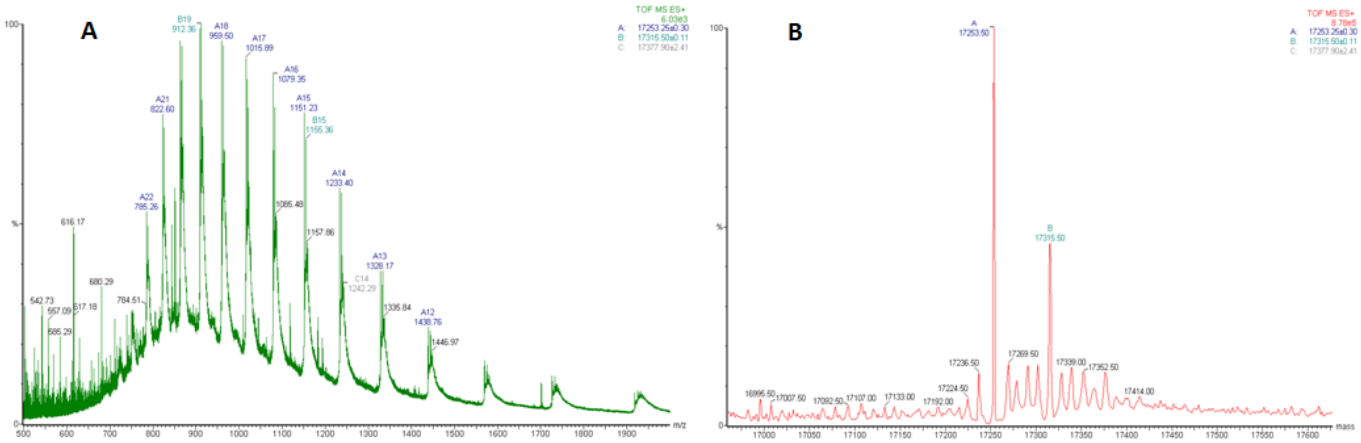

Figure 51 – Electro Spray ionization mass spectra of hetereologous expressed CopX in 50% ACN, 50% Water and 0.1% HCOOH. **A)** Mass-to-charge spectrum, the code at the top of each peak contains the number of positive charges for that m/z peak **B)** Deconvoluted spectrum. (The mass spectra was performed by Dr. Bart Devreese, in Belgium).

The protein with a molecular weight of 17253 Da corresponds to apo-CopX, without its signal peptide. The raise of 27 Da could be related with formylation reactions. This result confirms that the signal peptide was correctly processed as already expected by the N-terminus sequence analysis.

The two species present in the mass spectrum have a molecular weight difference of 62.5 Da, which matches the atomic mass of copper (63). Thus, the higher molecular weight protein, which corresponds to 40% of the sample, might be holo-CopX. Taking into account that CopX sample sent to mass spectra analysis had a 0.90 Cu/protein ratio, this 40% holo-CopX indicates that this protein has a high affinity for copper ions (does not lose copper ions even in the denaturating conditions used).

5.4.2 Molecular exclusion chromatography

The apparent molecular weight was determined using a pre-packed Superdex 75 10/300 GL, as described in Material and Methods Chapter 2 – 7.4. The cromatogram is shown in Figure 52 and the calibration curve that was used for apparent molecular weight determination is presented in the insert of Figure 52.

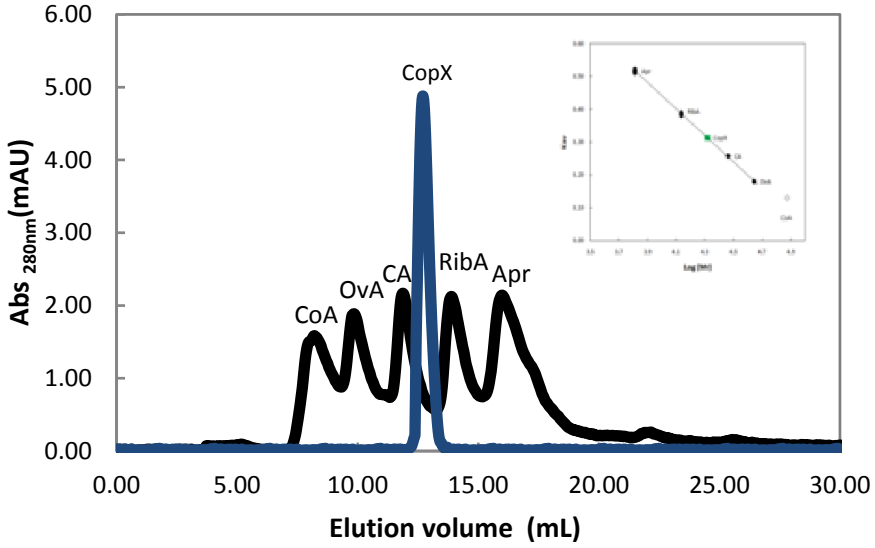

Figure 52 - Elution profile of heterologous expressed CopX in a Superdex 75 10/300 GL (Blue line). Protein was eluted with 50mM Tris-HCl-150mM NaCl pH 7.6 at 4°C. Elution profile of using LMW Gel Filtration Calibration Kit from a Superdex 75 10/300 GL Column eluted with 50mM Tris-HCl-150mM NaCl pH 7.6 at 4°C (Black line). Legend: Apr – Aprotinin, RibA – RibonucleaseA, CA – Carbonic Anhydrase, OvA – Ovalbumin, CoA – Conalbumin. **Insert -** Calibration curve for the molecular weight determination using LMW Gel Filtration Calibration Kit in Superdex 7510/300 GL. Trendline equation obtained Kav = - 0.3591 Log (MW) +1.8352. CoA was discarded as its molecular weight is higher than the higher molecular weight limit of Superdex 75 10/300 (70000 Da). Double Apr and RibA points corresponds to the same proteins in MixtureA and MixtureB for calibration of Superdex 75 10/300 GL. Double Apr and RibA points corresponds to the same proteins in MixtureA and MixtureB for calibration, see LMW Gel Filtration Calibration Kit protocol for details [151]. The profiles correspond to independent injections.

The apparent molecular weight was determined to be 20.8 kDa, in 50 mM Tris-HCl, 150 mM NaCl pH 7.6. The conalbumin was not used in the calibration curve due to its molecular weight, 75 kDa, which is outside of the exclusion limits of Superdex 75 10/300 GL (70 kDa).

The CopX apparent molecular weight is higher than the one determined by mass spectrometry, which could be due to monomer-dimer equilibrium or just to CopX shape (Stokes volume). In order to explore the first hypothesis, the apparent molecular weight was also determined as a function of the ionic strength (as described in Material and Methods, Chapter 2 - 8.4.2.1), and the results obtained are presented in Figure 53.

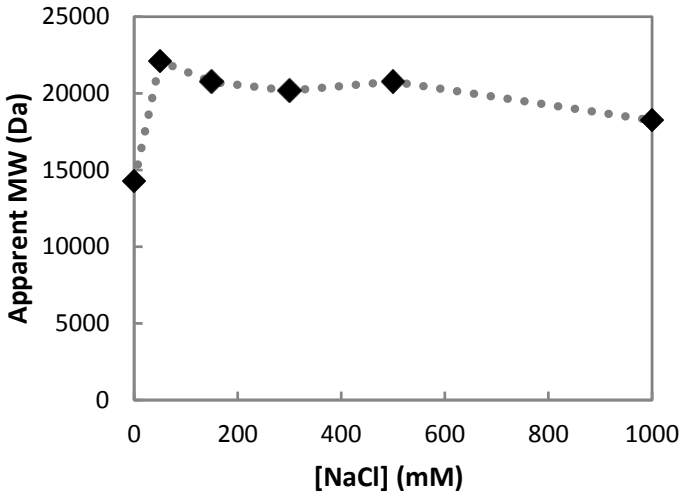

Figure 53 – Ionic strength dependence of the apparent molecular weight of CopX determined using Superdex 75 10/300 GL.

The analysis of the results obtained shows that the apparent molecular weight does not change with the [NaCl] between 50 mM and 1 M. However, in the absence of added NaCl, in 50 mM Tris-HCl, pH 7.6, the apparent molecular weight is smaller, 14.3 kDa (see Appendix H – Table 10 App.H.1), even than the one obtained by mass spectrometry. This could be due to problems with column calibration, which is not supposed to be eluted at such low ionic strengths, or CopX might be interacting with the matrix, thus eluting with a higher elution volume, as if it was a smaller protein.

As for the monomer-dimer equilibrium, it cannot be fully evaluated: or there is a monomer-dimer equilibrium, with the protein being only partially dimer that is not dependent on the ionic strength, or it is always a monomer, but it has a Stokes volume (the volume it would occupy during molecular exclusion) higher than expected, since the molecular exclusion behavior is influenced by shape, as well as, size and CopX might not have a simple spherical shape.

5.5 Preliminary 2D NMR analysis - ^{1}H-^{15}N HSQC spectra

^{15}N HSQC NMR spectra can be considered as fingerprint of a protein at a specific pH, temperature and ionic strength. This spectrum correlates ^{1}H and ^{15}N of the amide of the peptide bond, giving peaks for each amino acid of the protein, except for prolines. The chemical shift of each resonance depends on its environment, and because of the large

diversity of primary and terciary protein structures each protein has a unique ^{15}N HSQC spectra.

A preliminary NMR analysis was performed in order to assess whether CopX was folded and amenable to solution structure determination. The NMR spectra were recorded for the reduced CopX (reduced with sodium ascorbate), in order to avoid the paramagnetic effect of Cu^{2+} (Figure 54).

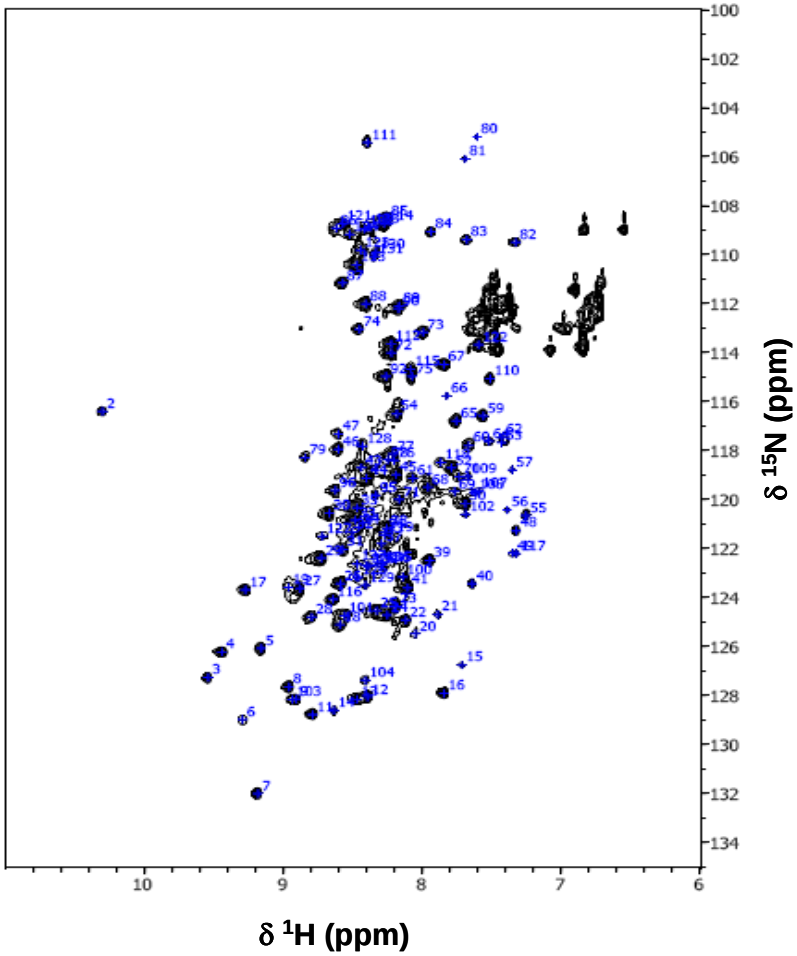

Figure 54 – ^{15}N HSQC spectrum of 0.1 mM CopX in 10 mM NaPi, pH 6.0, 298 K. The spectrum was acquired on a Bruker AvanceIII 600 NMR spectrometer equipped with a TCI cryoprobe.

This spectrum shows that recombinant CopX is folded since there is dispersion of the HN resonances between 10 and 6 ppm.

Attending to the fact that CopX has 156 residues (without the signal peptide), 8 being prolines, the number of expected HN resonances is 147 (as the first residue is also not observed). In addition, CopX has 13 asparagine and 6 glutamine residues, each with 2 HN in

their sidechain, and 1 Trp residue that has 1 HN in its sidechain. Therefore, 39 additional HN resonances should be observed, making in total 186 resonances that are expected to be observed in the ^{1}H-^{15}N HSQC spectrum of CopX.

In order to optimize the conditions at which the ^{1}H-^{15}N HSQC should be recorded, so that the number of observed amide resonances matches the expected one, 186, the ^{15}N HSQC was recorded at different pH conditions, 6.0, 7.0 and 7.6 and at two temperatures, 287 K and 298 K. The total number of amide resonances, 131 (backbone HN), did not change in the different conditions. The missing HN resonances, 16, might correspond to residues present in a mobile region, with the HN in exchange with the bulk solvent.

CONCLUSION

In this work, the *Marinobacter aquaeolei* 617 (also known as *Pseudomonas nautica* 617) resistance to cadmium, cobalt and copper ions was evaluated using both solid and liquid assays. The MIC/MCT determined was 200 µM of $CdCl_2$, 4 - 6 mM of $CoCl_2$, and 1.6 mM of $CuSO_4$. The comparison of these values with the ones obtained for other bacteria allowed the classification of *Marinobacter aquaeolei* as resistant to these heavy metal ions [173, 174] The MIC/MCT values obtained for copper and cadmium ions are within the range reported in the literature, while the range determined for cobalt ions is in the high-level for cobalt resistance [175, 176]. Moreover, during the growth of *Ma.aq* in the presence of 1 mM $CoCl_2$, was identified the synthesis of a compound that might be either cobalamine or a cobalamine-containing protein.

The copper resistance was expected, because the analysis of the *Ma.aq* sequenced genome revealed the presence of a copper resistance regulon, *copSRXAB_copA-pump_cusABC*. As for the cadmium and cobalt ion resistance it was not possible to identify putative genetic determinants in the genome of this bacterium responsible for this resistance, due to the absence of homologous genes associated with the resistance to these heavy metals.

Under the scope of copper resistance, the main focus of this work, the bioinformatic analysis unraveled the uniqueness of the proposed copper operon resistance in *M.aq*, *copSRXAB* and highlighted its association with copper resistance. The uniqueness of the system is directly related with the number of genes present (3, *copXAB*), and with *copX*. The protein sequence encoded by that gene does not present the characteristic copper binding motifs that are found in the small copper chaperones usually associated with copper resistance systems. Indeed, the protein encoded by *copX* has the type 1 copper binding motif (**His** towards the N-terminal and 60 residues after the **Cys**-X_2-ProGly**His**-X_3-Gly**Met** at the C-terminal) and is a much larger protein.

Preliminary proteomic analysis demonstrated that the copper resistance operon is expressed only under copper ion stress, as CopX was identified in periplasm of *Ma.aq* by MALDI TOF-TOF analysis of the 2D electrophoresis.

In order to biochemical characterize CopX, the gene that codes for it was successfully heterologously expressed in *E.coli* and purified for the first time. The expression conditions were induction at an OD_{600nm} of 0.6-0.8, with 1 mM IPTG during 7 h at 30 °C, and requires the supplementation with 0.5 mM $CuSO_4$ upon inoculation to obtain mainly holo-CopX. The protein was purified from the periplasmic extract using two chromatographic steps, an anionic exchange followed by a size-exclusion chromatography. In some cases, a third chromatographic step was performed using stronger anionic exchange chromatography.

The purified CopX presents 0.90 copper per polypeptide chain, and its signal peptide was correctly processed by the Sec machinery of *Es.coli*. The molecular weight was determined to be 17253 $\pm$ 0.3 Da by mass spectrometry (ESI), and the N-terminus analysis showed that the signal peptide is composed by the first 21 residues. In addition, CopX behaves as a monomer in size-exclusion chromatography with an apparent molecular weight of 20.4 kDa, behavior that is not dependent on ionic strength.

The UV-visible spectra of CopX has the characteristic absorption bands of type 1 copper proteins, such as, an intense charge transfer (S_{cys} – Cu ion) band at 440 nm, 580 nm and 720 nm, as well as, some distinctive characteristics, such as a high absorbance ratio between the 440 nm and 580 nm band (close to 1). The extinction coefficient at 580 nm was determined to be 3.8 mM^{-1} cm^{-1} by Cu ion.

CopX EPR spectrum is an axial spectrum, but has a hyperfine-coupling constant, $A_{\parallel}$, of 170 G, higher than the one reported for blue copper binding proteins [177].

Preliminary ^{15}N HSQC NMR experiments showed that the protein is folded; as 131 disperse amide resonances out of the 147 expected backbone amide resonances are observed. This makes CopX amenable to solution structure determination.

FUTURE PERSPECTIVES

The studies presented here show that it is clear that the proposed copper resistance system is indeed being expressed under copper ion stress in *Ma.aq*, which was also classified as a copper resistance bacterium. The future studies can be divided into three different directions to unravel this molecular system.

The first is gene expression analysis. In order to determine the involvement of the *copSRXAB* locus in the *Ma.aq* copper resistance. The focus of this analysis will be the genes coding for the periplasmatic proteins, *CopA* and *CopX*, and two-component regulatory system, *copRS*. The increase of *copA, copX*, *copR* and *copS* will be analyzed by the related increase in the level of its transcripts using Real-Time PCR under copper ion induced stress. This technique will also be used to determine whether the system has a specific response to copper or responds to other heavy ions, as cadmium, cobalt, zinc and silver ion. In addition, the reverse transcription will be used to confirm that *copXAB* are transcribed as an operon. Relatively to *CopRS*, the DNA binding sequence of CopR in the promoter region of copXAB operon will be identified by DNA binding studies, such as electrophoretic mobility shift assays and DNase I footprinting. Metal ion specificity will also be tested using lacZ activity for metal induction of the operon.

Second, the proteomic approach. Our preliminary proteomic analysis of the periplasm clearly shows that CopX is being expressed under copper stress, and unraveled a series of changes in the proteome of *Ma.aq* under copper ion stress. Therefore, there is the need to perform more 2D electrophoresis experiments (identical biological replicates) and a full statistical analysis, as well as, mass spectrometry analysis to identify the proteins associated with the metabolic changes under copper stress. Non-gel based proteomics will be used to analyze the membrane fraction, which usually is poorly represented in the 2D-electrophoresis.

The third part are the structural studies of copX that encodes the uncharacterized copper-binding protein, with low sequence homology to other copper-containing proteins,

and that has never been characterized nor reported to be associated with copper tolerance systems. NMR preliminary studies showed a good perspective in achieving the assignment, has well has the structure of CopX. Copper titration studies should also be performed to unravel the copper binding locations in CopX, as well as the use of other biochemical techniques, such as Raman spectroscopy to characterize this copper center. Interaction with CopA, the multicopper oxidase, will be characterized using NMR spectroscopy and microcalorimetry.

The other two heavy metals studied in this work, cadmium and cobalt, constitute two major independent areas of research. *Ma.aq* was classified in this work has a cadmium and cobalt resistant bacterium, but the resistance system is unknown. In both cases, a proteomic analysis can be used to identify the molecular mechanisms responsible for this resistance. Concerning the cobalt resistance studies, further studies are needed to identify the protein/compound produced by *Ma.aq*, which could be accomplished by using GC-MS and also by purifying the proteins from the cellular fractions.

REFERENCES

1. Koch, A.L. and S. Silver, *The first cell.* Adv Microb Physiol, 2005. **50**: p. 227-59.
2. Silver, S. and L.T. Phung, *Bacterial heavy metal resistance: new surprises.* Annu Rev Microbiol, 1996. **50**: p. 753-89.
3. Silver, S., *Genes for all metals--a bacterial view of the periodic table. The 1996 Thom Award Lecture.* J Ind Microbiol Biotechnol, 1998. **20**(1): p. 1-12.
4. Silver, S., *Transport of inorganic cations*, in *(Neidhardt F.C. et al) Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. 1996, ASM Press: Washington, DC. p. 1091-1102.
5. Nies, D.H., *Microbial heavy-metal resistance.* Appl Microbiol Biotechnol, 1999. **51**(6): p. 730-50.
6. Silver, S. and G. Ji, *Newer systems for bacterial resistances to toxic heavy metals.* Environ Health Perspect, 1994. **102 Suppl 3**: p. 107-13.
7. Cervantes, C., et al., *[Heavy Metal Microbial Interactions].* Rev Latinoam Microbiol, 2006. **48**(2): p. 203-210.
8. Goodyer, I.D., et al., *Characterization of the Menkes protein copper-binding domains and their role in copper-induced protein relocalization.* Hum Mol Genet, 1999. **8**(8): p. 1473-8.
9. Cater, M.A., et al., *Intracellular trafficking of the human Wilson protein: the role of the six N-terminal metal-binding sites.* Biochem J, 2004. **380**(Pt 3): p. 805-13.
10. Danks, D.M., *Disorders of copper transport*, in *(Scriver, C.R.; Beaudet, A.L; Sly, W.M.; Valle, D.) The Metabolic and Molecular Basis of Inherited Disease*, McGraw-Hill, Editor. 1995: New York. p. 2211-2235.
11. Monaco, A.P. and J. Chelly, *Menkes and Wilson diseases.* Adv Genet, 1995. **33**: p. 233-53.
12. Weast, R.C., *CRC handbook of chemistry and physics* 64th ed. 1984, Boca Raton, Florida: CRC.
13. Kachur, A.V., C.J. Koch, and J.E. Biaglow, *Mechanism of copper-catalyzed oxidation of glutathione.* Free Radic Res, 1998. **28**(3): p. 259-69.

14. Nies, D.H. and S. Silver, *Ion efflux systems involved in bacterial metal resistances.* J Ind Microbiol, 1995. **14**(2): p. 186-99.

15. Silver, S. and T. Phung le, *A bacterial view of the periodic table: genes and proteins for toxic inorganic ions.* J Ind Microbiol Biotechnol, 2005. **32**(11-12): p. 587-605.

16. Cervantes, C., K. Chavez, and S. Vaca, *[Mechanisms of bacterial resistance to heavy metals].* Rev Latinoam Microbiol, 1991. **33**(1): p. 61-70.

17. Agranoff, D.D. and S. Krishna, *Metal ion homeostasis and intracellular parasitism.* Mol Microbiol, 1998. **28**(3): p. 403-12.

18. Fagan, M.J. and M.H. Saier, Jr., *P-type ATPases of eukaryotes and bacteria: sequence analyses and construction of phylogenetic trees.* J Mol Evol, 1994. **38**(1): p. 57-99.

19. Boal, A.K. and A.C. Rosenzweig, *Structural biology of copper trafficking.* Chem Rev, 2009. **109**(10): p. 4760-79.

20. Fath, M.J. and R. Kolter, *ABC transporters: bacterial exporters.* Microbiol Rev, 1993. **57**(4): p. 995-1017.

21. Fetsch, E.E. and A.L. Davidson, *Maltose transport through the inner membrane of E. coli.* Front Biosci, 2003. **8**: p. d652-60.

22. Pao, S.S., I.T. Paulsen, and M.H. Saier, Jr., *Major facilitator superfamily.* Microbiol Mol Biol Rev, 1998. **62**(1): p. 1-34.

23. Borges-Walmsley, M.I., K.S. McKeegan, and A.R. Walmsley, *Structure and function of efflux pumps that confer resistance to drugs.* Biochem J, 2003. **376**(Pt 2): p. 313-38.

24. Haney, C.J., et al., *New developments in the understanding of the cation diffusion facilitator family.* J Ind Microbiol Biotechnol, 2005. **32**(6): p. 215-26.

25. Florianczyk, B. and T. Trojanowski, *Zinc transporting proteins.* Journal of Chinese Clinical Medicine, 2008. **3**(11).

26. Paulsen, I.T., M.H. Brown, and R.A. Skurray, *Proton-dependent multidrug efflux systems.* Microbiol Rev, 1996. **60**(4): p. 575-608.

27. Saier, M.H., Jr., *Tracing pathways of transport protein evolution.* Mol Microbiol, 2003. **48**(5): p. 1145-56.

28. Krulwich, T.A., et al., *Do physiological roles foster persistence of drug/multidrug-efflux transporters? A case study.* Nat Rev Microbiol, 2005. **3**(7): p. 566-72.

29. Nies, D.H., et al., *CHR, a novel family of prokaryotic proton motive force-driven transporters probably containing chromate/sulfate antiporters.* J Bacteriol, 1998. **180**(21): p. 5799-802.

30. Saier, M.H., Jr., et al., *Two novel families of bacterial membrane proteins concerned with nodulation, cell division and transport.* Mol Microbiol, 1994. **11**(5): p. 841-7.

31. Barkay, T., S.M. Miller, and A.O. Summers, *Bacterial mercury resistance from atoms to ecosystems.* FEMS Microbiol Rev, 2003. **27**(2-3): p. 355-84.

32. Esquivel, J.A., et al., *Arsenic resistance determinants from environmental bacteria.* Rev Latinoam Microbiol, 1998. **40**(1-2): p. 45-52.

33. Saier, M.H., Jr., *Computer-aided analyses of transport protein sequences: gleaning evidence concerning function, structure, biogenesis, and evolution.* Microbiol Rev, 1994. **58**(1): p. 71-93.

34. Brown, N.L., et al., *The MerR family of transcriptional regulators.* FEMS Microbiol Rev, 2003. **27**(2-3): p. 145-63.

35. Eicken, C., et al., *A metal-ligand-mediated intersubunit allosteric switch in related SmtB/ArsR zinc sensor proteins.* J Mol Biol, 2003. **333**(4): p. 683-95.

36. Stock, A.M., V.L. Robinson, and P.N. Goudreau, *Two-component signal transduction.* Annu Rev Biochem, 2000. **69**: p. 183-215.

37. Mergeay, M., et al., *Ralstonia metallidurans, a bacterium specifically adapted to toxic metals: towards a catalogue of metal-responsive genes.* FEMS Microbiol Rev, 2003. **27**(2-3): p. 385-410.

38. Solioz, M. and J.V. Stoyanov, *Copper homeostasis in Enterococcus hirae.* FEMS Microbiol Rev, 2003. **27**(2-3): p. 183-95.

39. Silver, S. and M. Walderhaug, *Bacterial plasmid-mediated resistances to mercury, cadmium, and copper*, in *(Goyer, R.A.; Cherian, M.G.) Toxicology of Metals, Biochemical Aspects* 1995, Springer: Berlin. p. 435-458.

40. Nemery, B., C.P. Lewis, and M. Demedts, *Cobalt and possible oxidant-mediated toxicity.* Sci Total Environ, 1994. **150**(1-3): p. 57-64.

41. Kobayashi, M. and S. Shimizu, *Metalloenzyme nitrile hydratase: structure, regulation, and application to biotechnology.* Nat Biotechnol, 1998. **16**(8): p. 733-6.

42. Nies, D.H., *Resistance to cadmium, cobalt, zinc, and nickel in microbes.* Plasmid, 1992. **27**(1): p. 17-28.

43. Yoon, K.P. and S. Silver, *A second gene in the Staphylococcus aureus cadA cadmium resistance determinant of plasmid pI258.* J Bacteriol, 1991. **173**(23): p. 7636-42.

44. Chen, Y.Y., et al., *cadDX operon of Streptococcus salivarius 57.I.* Appl Environ Microbiol, 2008. **74**(5): p. 1642-5.

45. Morby, A.P., et al., *SmtB is a metal-dependent repressor of the cyanobacterial metallothionein gene smtA: identification of a Zn inhibited DNA-protein complex.* Nucleic Acids Res, 1993. **21**(4): p. 921-5.

46. Blindauer, C.A., et al., *Multiple bacteria encode metallothioneins and SmtA-like zinc fingers.* Mol Microbiol, 2002. **45**(5): p. 1421-32.

47. Grosse, C., et al., *Identification of a regulatory pathway that controls the heavy-metal resistance system Czc via promoter czcNp in Ralstonia metallidurans.* Arch Microbiol, 2004. **182**(2-3): p. 109-18.

48. Hassan, M.T., et al., *Identification of a gene cluster, czr, involved in cadmium and zinc resistance in Pseudomonas aeruginosa.* Gene, 1999. **238**(2): p. 417-25.

49. Hebbeln, P. and T. Eitinger, *Heterologous production and characterization of bacterial nickel/cobalt permeases.* FEMS Microbiol Lett, 2004. **230**(1): p. 129-35.

50. Paulsen, I.T., et al., *Unified inventory of established and putative transporters encoded within the complete genome of Saccharomyces cerevisiae.* FEBS Lett, 1998. **430**(1-2): p. 116-25.

51. Komeda, H., M. Kobayashi, and S. Shimizu, *A novel transporter involved in cobalt uptake.* Proc Natl Acad Sci U S A, 1997. **94**(1): p. 36-41.

52. Paulsen, I.T. and M.H. Saier, Jr., *A novel family of ubiquitous heavy metal ion transport proteins.* J Membr Biol, 1997. **156**(2): p. 99-103.

53. Schmidt, T. and H.G. Schlegel, *Combined nickel-cobalt-cadmium resistance encoded by the ncc locus of Alcaligenes xylosoxidans 31A.* J Bacteriol, 1994. **176**(22): p. 7045-54.

54. Liesegang, H., et al., *Characterization of the inducible nickel and cobalt resistance determinant cnr from pMOL28 of Alcaligenes eutrophus CH34.* J Bacteriol, 1993. **175**(3): p. 767-78.

55. Rodrigue, A., G. Effantin, and M.A. Mandrand-Berthelot, *Identification of rcnA (yohM), a nickel and cobalt resistance gene in Escherichia coli.* J Bacteriol, 2005. **187**(8): p. 2912-6.

56. Nies, D.H., et al., *Expression and nucleotide sequence of a plasmid-determined divalent cation efflux system from Alcaligenes eutrophus.* Proc Natl Acad Sci U S A, 1989. **86**(19): p. 7351-5.

57. Diels, L., et al., *The czc operon of Alcaligenes eutrophus CH34: from resistance mechanism to the removal of heavy metals.* J Ind Microbiol, 1995. **14**(2): p. 142-53.

58. Tseng, T.T., et al., *The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins.* J Mol Microbiol Biotechnol, 1999. **1**(1): p. 107-25.

59. Nies, D.H., *Efflux-mediated heavy metal resistance in prokaryotes.* FEMS Microbiol Rev, 2003. **27**(2-3): p. 313-39.

60. Nies, D.H., *CzcR and CzcD, gene products affecting regulation of resistance to cobalt, zinc, and cadmium (czc system) in Alcaligenes eutrophus.* J Bacteriol, 1992. **174**(24): p. 8102-10.

61. von Rozycki, T. and D.H. Nies, *Cupriavidus metallidurans: evolution of a metal-resistant bacterium.* Antonie Van Leeuwenhoek, 2009. **96**(2): p. 115-39.

62. Collard, J.M., et al., *A new type of Alcaligenes eutrophus CH34 zinc resistance generated by mutations affecting regulation of the cnr cobalt-nickel resistance system.* J Bacteriol, 1993. **175**(3): p. 779-84.

63. Grass, G., C. Grosse, and D.H. Nies, *Regulation of the cnr cobalt and nickel resistance determinant from Ralstonia sp. strain CH34.* J Bacteriol, 2000. **182**(5): p. 1390-8.

64. Tibazarwa, C., et al., *Regulation of the cnr cobalt and nickel resistance determinant of Ralstonia eutropha (Alcaligenes eutrophus) CH34.* J Bacteriol, 2000. **182**(5): p. 1399-409.

65. Collard, J.M., et al., *Plasmids for heavy metal resistance in Alcaligenes eutrophus CH34: mechanisms and applications.* FEMS Microbiol Rev, 1994. **14**(4): p. 405-14.

66. Liochev, S.I. and I. Fridovich, *The Haber-Weiss cycle -- 70 years later: an alternative view.* Redox Rep, 2002. **7**(1): p. 55-7; author reply 59-60.

67. Rodriguez-Montelongo, L., et al., *Membrane-associated redox cycling of copper mediates hydroperoxide toxicity in Escherichia coli.* Biochim Biophys Acta, 1993. **1144**(1): p. 77-84.

68. Suwalsky, M., et al., *Cu2+ ions interact with cell membranes.* J Inorg Biochem, 1998. **70**(3-4): p. 233-8.

69. Fraústo da Silva, J.J.R. and R.J.P. Williams, *The Biological Chemistry of the Elements: The Inorganic Chemistry of Life*. 1993, Oxford: Oxford University Press.

70. Outten, C.E. and T.V. O'Halloran, *Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis.* Science, 2001. **292**(5526): p. 2488-92.

71. Crichton, R.R. and J.L. Pierre, *Old iron, young copper: from Mars to Venus.* Biometals, 2001. **14**(2): p. 99-112.

72. Outten, F.W., et al., *The independent cue and cus systems confer copper tolerance during aerobic and anaerobic growth in Escherichia coli.* J Biol Chem, 2001. **276**(33): p. 30670-7.

73. Changela, A., et al., *Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR.* Science, 2003. **301**(5638): p. 1383-7.

74. Tottey, S., D.R. Harvie, and N.J. Robinson, *Understanding how cells allocate metals using metal sensors and metallochaperones.* Acc Chem Res, 2005. **38**(10): p. 775-83.

75. Osman, D. and J.S. Cavet, *Copper homeostasis in bacteria.* Adv Appl Microbiol, 2008. **65**: p. 217-47.

76. Finney, L.A. and T.V. O'Halloran, *Transition metal speciation in the cell: insights from the chemistry of metal ion receptors.* Science, 2003. **300**(5621): p. 931-6.

77. Davis, A.V. and T.V. O'Halloran, *A place for thioether chemistry in cellular copper ion recognition and trafficking.* Nat Chem Biol, 2008. **4**(3): p. 148-51.

78. Xue, Y., et al., *Cu(I) recognition via cation-pi and methionine interactions in CusF.* Nat Chem Biol, 2008. **4**(2): p. 107-9.

79. Axelsen, K.B. and M.G. Palmgren, *Evolution of substrate specificities in the P-type ATPase superfamily.* J Mol Evol, 1998. **46**(1): p. 84-101.

80. Canovas, D., I. Cases, and V. de Lorenzo, *Heavy metal tolerance and metal homeostasis in Pseudomonas putida as revealed by complete genome analysis.* Environ Microbiol, 2003. **5**(12): p. 1242-56.

81. Odermatt, A., et al., *Primary structure of two P-type ATPases involved in copper homeostasis in Enterococcus hirae.* J Biol Chem, 1993. **268**(17): p. 12775-9.

82. Kim, H.J., et al., *Methanobactin, a copper-acquisition compound from methane-oxidizing bacteria.* Science, 2004. **305**(5690): p. 1612-5.

83. Kim, H.J., et al., *Purification and physical-chemical properties of methanobactin: a chalkophore from Methylosinus trichosporium OB3b.* Biochemistry, 2005. **44**(13): p. 5140-8.

84. Choi, D.W., et al., *Spectral, kinetic, and thermodynamic properties of Cu(I) and Cu(II) binding by methanobactin from Methylosinus trichosporium OB3b.* Biochemistry, 2006. **45**(5): p. 1442-53.

85. Rouch, D., et al., *Inducible plasmid-mediated copper resistance in Escherichia coli.* J Gen Microbiol, 1985. **131**(4): p. 939-43.

86. Rensing, C. and G. Grass, *Escherichia coli mechanisms of copper homeostasis in a changing environment.* FEMS Microbiol Rev, 2003. **27**(2-3): p. 197-213.

87. Brown, N.L., B.T.O. Lee, and S. Silver, *Bacterial transport of and resistance to copper*, in *(Sigel, H.; Sigel, A.) Metal Ions in Biological Systems*. 1994, Dekker: New York. p. 405-34.

88. Franke, S., et al., *Molecular analysis of the copper-transporting efflux system CusCFBA of Escherichia coli.* J Bacteriol, 2003. **185**(13): p. 3804-12.

89. Zgurskaya, H.I. and H. Nikaido, *Multidrug resistance mechanisms: drug efflux across two membranes.* Mol Microbiol, 2000. **37**(2): p. 219-25.

90. Munson, G.P., et al., *Identification of a copper-responsive two-component system on the chromosome of Escherichia coli K-12.* J Bacteriol, 2000. **182**(20): p. 5864-71.

91. Astashkin, A.V., et al., *Characterization of the copper(II) binding site in the pink copper binding protein CusF by electron paramagnetic resonance spectroscopy.* J Biol Inorg Chem, 2005. **10**(3): p. 221-30.

92. Koronakis, V., et al., *Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export.* Nature, 2000. **405**(6789): p. 914-9.

93. Brown, N.L., et al., *Molecular genetics and transport analysis of the copper-resistance determinant (pco) from Escherichia coli plasmid pRJ1004.* Mol Microbiol, 1995. **17**(6): p. 1153-66.

94. Huffman, D.L., et al., *Spectroscopy of Cu(II)-PcoC and the multicopper oxidase function of PcoA, two essential components of Escherichia coli pco copper resistance operon.* Biochemistry, 2002. **41**(31): p. 10046-55.

95. Tetaz, T.J. and R.K. Luke, *Plasmid-controlled resistance to copper in Escherichia coli.* J Bacteriol, 1983. **154**(3): p. 1263-8.

96. Zhang, L., et al., *Intermolecular transfer of copper ions from the CopC protein of Pseudomonas syringae. Crystal structures of fully loaded Cu(I)Cu(II) forms.* J Am Chem Soc, 2006. **128**(17): p. 5834-50.

97. Arnesano, F., et al., *A redox switch in CopC: an intriguing copper trafficking protein that binds copper(I) and copper(II) at different sites.* Proc Natl Acad Sci U S A, 2003. **100**(7): p. 3814-9.

98. Rouch, D.A. and N.L. Brown, *Copper-inducible transcriptional regulation at two promoters in the Escherichia coli copper resistance determinant pco.* Microbiology, 1997. **143 (Pt 4)**: p. 1191-202.

99. Mills, S.D., C.A. Jasalavich, and D.A. Cooksey, *A two-component regulatory system required for copper-inducible expression of the copper resistance operon of Pseudomonas syringae.* J Bacteriol, 1993. **175**(6): p. 1656-64.

100. Cha, J.S. and D.A. Cooksey, *Copper resistance in Pseudomonas syringae mediated by periplasmic and outer membrane proteins.* Proc Natl Acad Sci U S A, 1991. **88**(20): p. 8915-9.

101. Cooksey, D.A., *Molecular mechanisms of copper resistance and accumulation in bacteria.* FEMS Microbiol Rev, 1994. **14**(4): p. 381-6.

102. Monchy, S., et al., *Transcriptomic and proteomic analyses of the pMOL30-encoded copper resistance in Cupriavidus metallidurans strain CH34.* Microbiology, 2006. **152**(Pt 6): p. 1765-76.

103. Cha, J.S. and D.A. Cooksey, *Copper Hypersensitivity and Uptake in Pseudomonas syringae Containing Cloned Components of the Copper Resistance Operon.* Appl Environ Microbiol, 1993. **59**(5): p. 1671-1674.

104. Cooksey, D.A., *Copper uptake and resistance in bacteria.* Mol Microbiol, 1993. **7**(1): p. 1-5.

105. Adaikkalam, V. and S. Swarup, *Molecular characterization of an operon, cueAR, encoding a putative P1-type ATPase and a MerR-type regulatory protein involved in copper homeostasis in Pseudomonas putida.* Microbiology, 2002. **148**(Pt 9): p. 2857-67.

106. Cooksey, D.A., *Plasmid-Determined Copper Resistance in Pseudomonas syringae from Impatiens.* Appl Environ Microbiol, 1990. **56**(1): p. 13-16.

107. Lim, C.K. and D.A. Cooksey, *Characterization of chromosomal homologs of the plasmid-borne copper resistance operon of Pseudomonas syringae.* J Bacteriol, 1993. **175**(14): p. 4492-8.

108. Vargas, E., et al., *Chromosome-encoded inducible copper resistance in Pseudomonas strains.* Antonie Van Leeuwenhoek, 1995. **68**(3): p. 225-9.

109. Bender, C.L. and D.A. Cooksey, *Indigenous plasmids in Pseudomonas syringae pv. tomato: conjugative transfer and role in copper resistance.* J Bacteriol, 1986. **165**(2): p. 534-41.

110. Harrison, M.D., C.E. Jones, and C.T. Dameron, *Copper chaperones: function, structure and copper-binding properties.* J Biol Inorg Chem, 1999. **4**(2): p. 145-53.

111. Puig, S., E.M. Rees, and D.J. Thiele, *The ABCDs of periplasmic copper trafficking.* Structure, 2002. **10**(10): p. 1292-5.

112. Ehrenstein, D. and G.U. Nienhaus, *Conformational substates in azurin.* Proc Natl Acad Sci U S A, 1992. **89**(20): p. 9681-5.

113. Huu, N.B., et al., *Marinobacter aquaeolei sp. nov., a halophilic bacterium isolated from a Vietnamese oil-producing well.* Int J Syst Bacteriol, 1999. **49 Pt 2**: p. 367-75.

114. Cabrito, I.M.C.G., *[Structural Studies and Mechanisms of Nitrous Oxide Reductase]* in *Chemistry Department*. 2004, Nova de Lisboa - Faculdade de Ciências e Tecnologia: Lisbon. p. 303-22.

115. Marquez, M.C. and A. Ventosa, *Marinobacter hydrocarbonoclasticus Gauthier et al. 1992 and Marinobacter aquaeolei Nguyen et al. 1999 are heterotypic synonyms.* Int J Syst Evol Microbiol, 2005. **55**(Pt 3): p. 1349-51.

116. Markowitz, V.M., et al., *The integrated microbial genomes (IMG) system.* Nucleic Acids Res, 2006. **34**(Database issue): p. D344-8.

117. Altschul, S.F., et al., *Basic local alignment search tool.* J Mol Biol, 1990. **215**(3): p. 403-10.

118. Larkin, M.A., et al., *Clustal W and Clustal X version 2.0.* Bioinformatics, 2007. **23**(21): p. 2947-8.

119. Pearson, W.R. and D.J. Lipman, *Improved tools for biological sequence comparison.* Proc Natl Acad Sci U S A, 1988. **85**(8): p. 2444-8.

120. Permina, E.A., et al., *Comparative genomics of regulation of heavy metal resistance in Eubacteria.* BMC Microbiol, 2006. **6**: p. 49.

121. Giedroc, D.P. and A.I. Arunkumar, *Metal sensor proteins: nature's metalloregulated allosteric switches.* Dalton Trans, 2007(29): p. 3107-20.

122. Felsenstein, J., *Phylogeny Inference Package (PHYLIP)*. 2009.

123. Tamura, K., et al., *MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0.* Mol Biol Evol, 2007. **24**(8): p. 1596-9.

124. Kumar, S., et al., *MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences.* Brief Bioinform, 2008. **9**(4): p. 299-306.

125. Andreini, C., et al., *Occurrence of copper proteins through the three domains of life: a bioinformatic approach.* J Proteome Res, 2008. **7**(1): p. 209-16.

126. Baumann, P. and L. Baumann, *The prokaryotes. A Handbook on Habitats, Isolation and Identification of Bacteria*. Vol. 2nd. 1986. 1302-1331.

127. Starkey, R.L., *A study of spore formation and other morphological characteristics of Vibrio desulfuricans.* Arch Microbiol, 1938(8): p. 268-304.

128. Goodhew, C.F., et al., *The cellular location and specificity of bacterial cytochrome c peroxidases.* Biochem J, 1990. **271**(3): p. 707-12.

129. Hunter, D.J., K.R. Brown, and G.W. Pettigrew, *The role of cytochrome c4 in bacterial respiration. Cellular location and selective removal from membranes.* Biochem J, 1989. **262**(1): p. 233-40.

130. Schagger, H. and G. von Jagow, *Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa.* Anal Biochem, 1987. **166**(2): p. 368-79.

131. *Bicinchoninic Acid Protein Assay Kit - Product Code: BCA1 AND B 9643*, Sigma.

132. *2-D Electrophoresis, Principles and Methods Handbook*. 2008, GE Healthcare.

133. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.

134. Laemmli, U.K. and M. Favre, *Maturation of the head of bacteriophage T4. I. DNA packaging events.* J Mol Biol, 1973. **80**(4): p. 575-99.

135. *PlusOne Silver Staining kit*. 2006, Pharmacia.

136. *EZBlue protocol*. 2008, Sigma.

137. *ImageMaster 2D Platinum* 2008, GE Healthcare.

138. *Ludesi REDFIN3*. 2008, Ludesi

139. Hiller, K., et al., *JVirGel: Calculation of virtual two-dimensional protein gels.* Nucleic Acids Res, 2003. **31**(13): p. 3862-5.

140. Hiller, K., et al., *JVirGel 2.0: computational prediction of proteomes separated via two-dimensional gel electrophoresis under consideration of membrane and secreted proteins.* Bioinformatics, 2006. **22**(19): p. 2441-3.

141. Lench, N., P. Stanier, and R. Williamson, *Simple non-invasive method to obtain DNA for gene analysis.* Lancet, 1988. **1**(8599): p. 1356-8.

142. *QIAquick Spin Handbook*. 2002, QIAGEN.

143. *SYBR Safe™ DNA Gel Stain - Product code: 62923 TC562-1*. 2008, Invitrogen.

144. *pET System Manual*. 2005, Novagen.

145. *NovaBlue Singles™ Competent Cells - Cat. No. 70181* 2008, Novagen.

146. Bertani, G., *Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli.* J Bacteriol, 1951. **62**(3): p. 293-300.

147. *NZYMiniprep Manual - Cat. MB01001*. 2008, Nzytech.

148. *BL21(DE3) Competent Cells Manual - Cat. MB00601*. 2008, NZytech.

149. Sambrook, J., E.F. Fritsch, and T. Maniatis, *Molecular Cloning: a Laboratory Manual*. 2nd ed. 1989, Cold Spring Harbor, New York.

150. Keller, R.L.J., *Optimizing the process of Nuclear Magnetic Resonance Spectrum Analysis and Computer Aided Resonance Assignment* 2004.

151. *Gel Filtration LMW Calibration Kit Handbook - Cat. 28-4038-41*. 2008, GE Healthcare's.

152. Carter, K., et al., *Bioinformatics issues for automating the annotation of genomic sequences.* Genome Inform, 2001. **12**: p. 204-11.

153. Lee, S.W., E. Glickmann, and D.A. Cooksey, *Chromosomal locus for cadmium resistance in Pseudomonas putida consisting of a cadmium-transporting ATPase and a MerR family response regulator.* Appl Environ Microbiol, 2001. **67**(4): p. 1437-44.

154. Edwards, K., *Marinobacter aquaeolei VT8 complete genome*. 2005, DOE Joint Genome Institute

155. Dennison, C., *Ligand and loop variations at type 1 copper sites: influence on structure and reactivity.* Dalton Trans, 2005(21): p. 3436-42.

156. Tinnell, W.H., B.L. Jefferson, and R.E. Benoit, *Manganese-mediated morphogenesis in Penicillium claviforme and Penicillium clavigerum.* Can J Microbiol, 1977. **23**(2): p. 209-12.

157. Roth, J.R., J.G. Lawrence, and T.A. Bobik, *Cobalamin (coenzyme B12): synthesis and biological significance.* Annu Rev Microbiol, 1996. **50**: p. 137-81.

158. Caspi, R., et al., *The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases.* Nucleic Acids Res, 2008. **36**(Database issue): p. D623-31.

159. Teitzel, G.M., et al., *Survival and growth in the presence of elevated copper: transcriptional profiling of copper-stressed Pseudomonas aeruginosa.* J Bacteriol, 2006. **188**(20): p. 7242-56.

160. Pettigrew, G.W., A. Echalier, and S.R. Pauleta, *Structure and mechanism in the bacterial dihaem cytochrome c peroxidases.* J Inorg Biochem, 2006. **100**(4): p. 551-67.

161. Skjeldal, L., et al., *Spectroscopic studies of cobalt-substituted ferredoxin from Clostridium pasteurianum.* Biol Met, 1989. **2**(3): p. 135-41.

162. Gasteiger, E., et al., *ExPASy: The proteomics server for in-depth protein knowledge and analysis.* Nucleic Acids Res, 2003. **31**(13): p. 3784-8.

163. Emanuelsson, O., et al., *Locating proteins in the cell using TargetP, SignalP and related tools.* Nat Protoc, 2007. **2**(4): p. 953-71.

164. Chou, K.C. and H.B. Shen, *Large-scale predictions of gram-negative bacterial protein subcellular locations.* J Proteome Res, 2006. **5**(12): p. 3420-8.

165. AcaClone, *pDRAW32 DNA analysis*. 2009 - 1.1.106 revision.

166. Henry, Y. and A. Guissani, *[Metalloproteins, cellular targets of nitric oxide: a concise review].* C R Seances Soc Biol Fil, 1995. **189**(6): p. 1039-57.

167. Kyritsis, P., T. Kohzuma, and A.G. Sykes, *Redox reactivity of the type 1 copper protein amicyanin from Thiobacillus versutus with its physiological partner cytochrome C550 and inter-protein cross-reaction studies.* Biochim Biophys Acta, 1996. **1295**(2): p. 245-52.

168. Belford, R.L., *Understanding the blue copper proteins.* Biophys J, 1993. **64**(1): p. 3-4.

169. Bruker, *WinEPR System*. 1990-96.

170. Bruker, *SimFonia*. 1994-96.

171. Hira, D., et al., *Identification of a blue copper protein from Hyphomicrobium denitrificans and its functions in the periplasm.* J Biochem, 2007. **142**(3): p. 335-41.

172. Quaranta, D., et al., *The copper-inducible cin operon encodes an unusual methionine-rich azurin-like protein and a pre-Q0 reductase in Pseudomonas putida KT2440.* J Bacteriol, 2007. **189**(14): p. 5361-71.

173. Hu, Q., et al., *Detection, isolation, and identification of cadmium-resistant bacteria based on PCR-DGGE.* J Environ Sci (China), 2007. **19**(9): p. 1114-9.

174. Roane, T.M. and S.T. Kellogg, *Characterization of bacterial communities in heavy metal contaminated soils.* Can J Microbiol, 1996. **42**(6): p. 593-603.

175. Seidametova, E.A., et al., *[Isolation of cobalt-resistant strains of propionic acid bacteria, potent producers of vitamin B12].* Prikl Biokhim Mikrobiol, 2004. **40**(6): p. 645-8.

176. Srinath, T., et al., *Chromium (VI) biosorption and bioaccumulation by chromate resistant bacteria.* Chemosphere, 2002. **48**(4): p. 427-35.

177. Pauleta, S.R., *Electron transfer complexes between Paracoccus cytochrome c peroxidase and its electron donors*. 2003, UNL/FCT: Lisbon.

178. Vincze, T., J. Posfai, and R.J. Roberts, *NEBcutter: A program to cleave DNA with restriction enzymes.* Nucleic Acids Res, 2003. **31**(13): p. 3688-91.

APPENDIX

APPENDIX A - Gene Ortholog Neighborhoods Analysis

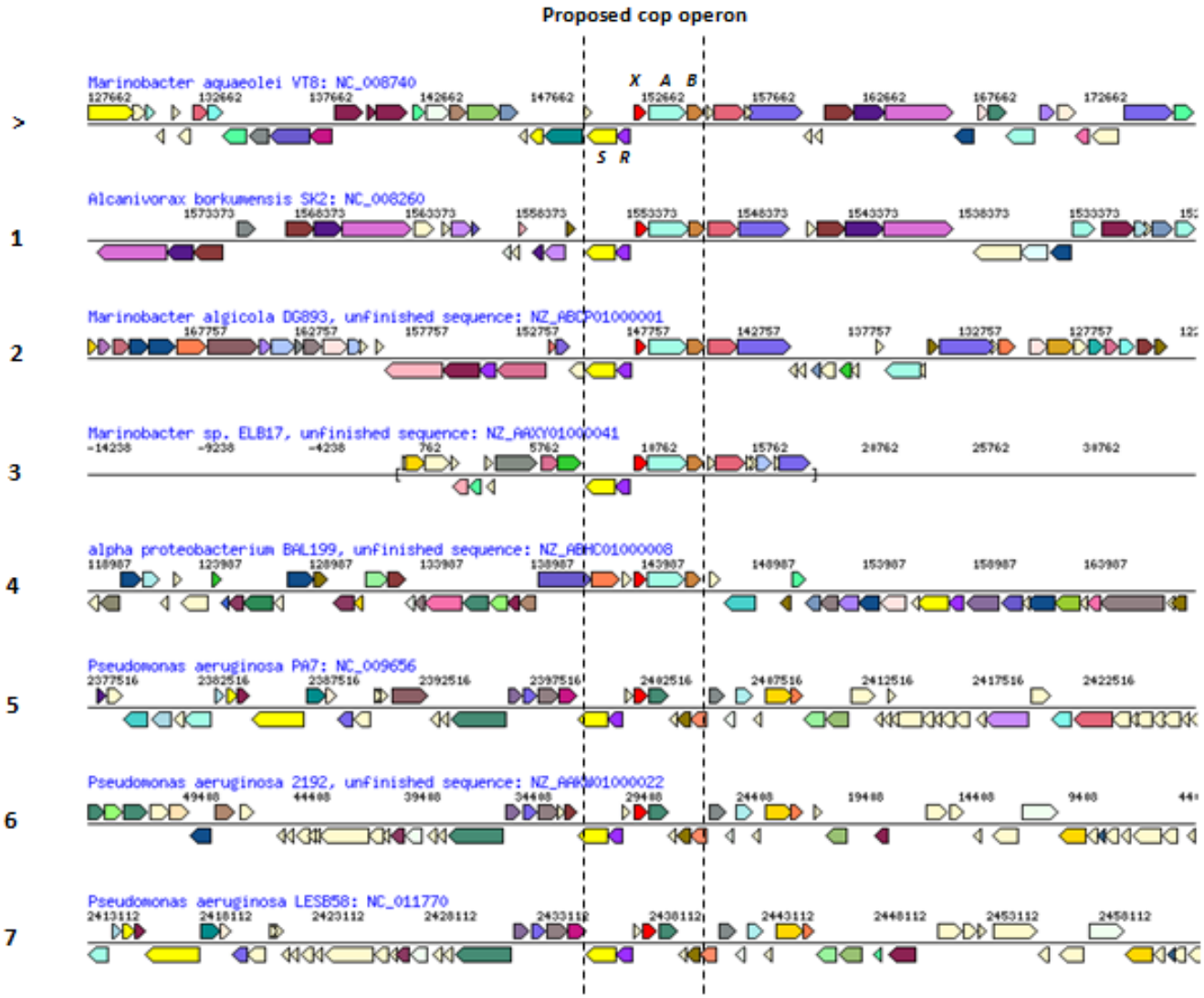

Figure 1 App.A. – Gene Ortholog Neighborhoods analysis of the putative copper operon, *copSRXAB*, from the *Ma.aq* chromosome using the *copX* (>) as the neighborhood reference (between dashed lines). Query was performed in Eubacteria (against genomes present at the DOE Joint Genome Institute databank). First, highly similar, 7 obtained results are shown with the correspondent accession numbers. Legend: Same color genes belong from the same ortholog group, except light yellow that correspond to no ortholog assignment. White color corresponds to pseudo gene. Image obtained from the graphic interface Gene Ortholog Neighborhoods using the Integrated Microbial Genome system [116].

Table 1 App.A – BLAST score of the target gene for the ortholog study, *copX* (>). The algorithm used was blastn (somewhat similar sequences) [117]. For the corresponding accession numbers see Figure App.1.

Query	Score	E value	Similarity (%)
>	987	0.0	100
1	263	$4x10^{-69}$	76
2	256	$5X10^{-67}$	83
3	261	$1x10^{-68}$	75
4	222	$1X10^{-56}$	58
5	175	$2x10^{-42}$	46
6	170	$4x10^{-41}$	46
7	170	$4x10^{-41}$	46

APPENDIX B - Protein Homology study

B.1 - Two-component regulatory system, CopR

A. Data from the BLAST protein query

Table 2 App.B.1(A) – Accession numbers of the National Center for Biotechnology Information (protein sequence database) from the BLAST protein query of CopR (†), first 4 most similar sequences of known copper resistance regulators. Algorithm used was blastp (protein-protein BLAST) [117].

Organism	Score	E value	Similarity (%)	Accession number
***Ma.aq*†**	449	0.0	100	YP_957418.1
Ps.stutzeri	212	$2x10^{-53}$	51	YP_001173204.1
Ps.aeruginosa	202	$1x10^{-5}$	49	YP_002439866.1
Ps.putida	200	$6x10^{-5}$	47	YP_001748660.1
Es.coli	169	$2x10^{-42}$	42	YP_001481477.1

B. Alignment from the nucleotide sequence query result

Table 3 App.B.1(B) - Accession numbers of the National Center for Biotechnology Information (nucleotide sequence database) from the BLAST nucleotide query of *copR* (†), first 4 most similar sequences of known copper resistance regulators. Algorithm used was blastn (somewhat similar sequences) [117].

Organism	Score	E value	Similarity (%)	Accession number
***Ma.aq*†**	1298	0.0	100	VIMSS3521613
Ps.aeruginosa	134	$8x10^{-13}$	48	FM209186.1
Ps.stutzeri	50	$1x10^{-10}$	20	CP000304.1
Es.coli	54	$1x10^{-4}$	24	DQ517526.1
Ps. putida	48	$1x10^{-4}$	18	AE015451.1

```
Ps.putida        ATGAAAC---TACTCGTAGCTGAGGATGAACCGAAAACTGGTGCTTATTT  47
Es.coli          ATGCAGCGTATTTTAATCGTTGAAGACGAACAAAAAACAGGTCGTTACCT  50
Ps.aeruginosa    ---TCATTCGTCGCGCTCTTCGAGAACGTAGCCCATGCCGCGCACGGTAT  47
Ps.stutzeri      TCATGATGGGTCCGGCTCGTCGAGCATGTAGCCCATGCCGCGCGCGGTAT  50
Ma.aq            ---ATGCGTTTATTGCTCGTTGAAGACG-ACCGTTTGCTGGCCGAGGGGC  46
                                *     *    **  * * *       * *

Ps.putida        ACAACAGGGC-CTTGCAGAGGCGGGATTCACTGTTGACCGGGTCCTCACG  96
Es.coli          GCAGCAGGGA-CTGGTTGAGGAAGGCTATCAGGCCGATCTCTTTAATAAT  99
Ps.aeruginosa    GGATCAGGCG-CTGCGGGTAGTCGTCGTCGACCTTGGCGCGCAGGCGGCG  96
Ps.stutzeri      GGATCAGTTT-AGGCTCGAAGCCATCGTCGATCTTCGCCCGTAGCCGGCG  99
Ma.aq            TGGTCAGACAGCTGGAAAAAGCGGGGTTTAGCATCGACCACACGTCCAGT  96
```

```
                    ***            *

Ps.putida           GGTACCGATGCCCTTCAACATGCACTCAGCGAAAGCTATGACCTGCTGAT 146
Es.coli             GGCCGCGATGGTCTCGGGGCCGCGTCGAAGGGACAGTATGATTTGATAAT 149
Ps.aeruginosa       GATGGCCACCTCGATGACGTTGGTATCGCTGTCGAAGTTCATGTCCCAGA 146
Ps.stutzeri         GATCGCCACCTCGATGACGTTGGTGTCGCTGTCGAAGTTCATGTCCCACA 149
Ma.aq               GCCCGTGAGGCCCAGATTCTGGGGGAGCAGGAGGATTATCGTGCCGCTGT 146
                    *      *             *        *       *

Ps.putida           ACTCGACGTGATGATGCCGGGGCTAGAAGGATGGGAAGTCTTGCGAATGG 196
Es.coli             ACTGGACGTGATGCTGCCTTTCCTCGACGGGTGGCAAATCATCAGCGCAC 199
Ps.aeruginosa       CCT--GCGAGGCGAT-CAGCGACTTGGGCAGCACCTCGCCGCTGCGCCGC 193
Ps.stutzeri         CCT--GCGAGGCGAT-CAGCGATTTGGGCAGCACCTCGCCGCGCCGGCGC 196
Ma.aq               TCTCGATCTCGGCCTGCCGGATGGCAACGGGCTGGAGGTTTTGAA---AC 193
                     **           * *

Ps.putida           TCCGCGCAGCCGGCAAAGACGTCCC----GGTATTGTTTCTGAC-TGCCA 241
Es.coli             TGAGGGAGTCCGGGCACGAAGAACC----GGTCCTGTTTTTAAC-CGCAA 244
Ps.aeruginosa       AGCAGCAGTTCGAGCAGGGCGAACTCCTTGGCGGTGAGGTCGATCCGCTT 243
Ps.stutzeri         AACAGCAGTTCCAGCAGGGCGAACTCCTTGGCGGTCAGGTCGATGCGCTG 246
Ma.aq               GATGGCGATCGAAGCTCGTCGAGTGCCC-GGTATTGGTCCTCAC-CGCCA 241
                                    *  *        **   *       *   **

Ps.putida           GGGATGGGGTAGATGATCGAG-TGAAAGGCCTTGAGCTGGGAGCCGATGA 290
Es.coli             AGGACAACGTGCGGGACAAAG-TGAAAGGACTGGAGCTTGGCGCAGATGA 293
Ps.aeruginosa       GCCCTGGCGTTGCACCCGGCGGCGCAGCAAATCGAGTTCGAGGTCGGCCA 293
Ps.stutzeri         ACCGGCGCGCGTGGCGCGACGCTTGAGCAGGTCGACCTGCAGGTCGGCCA 296
Ma.aq               GAGGCGACTGGCAGGACAAGG-TCAATGGACTGAAAGCAGGGGCGGATGA 290
                                       *    *     *  *       *  *   *

Ps.putida           CTACCTGATCAAGCC-CTTCGCTTTCTCAGAGCTACTGGCTCGGGTGCGG 339
Es.coli             CTACCTGATTAAGCC-CTTTGATTTTACGGAGCTGGTTGCACGTGTAAGA 342
Ps.aeruginosa       GTTGCAGGGTGGTCTCCTGCAACTGCTGGCTGCC-CCGGCGCAGCAGGGT 342
Ps.stutzeri         CCTGCAGCTGGGTCTGCAGCGCCGTCGCATTGCC-GCGCCGCAGCAGGGT 345
Ma.aq               CTATCTGGCCAAACCGTTCCAGACA-GAAGAACTGATCGCCCGCATCAAT 339
                        * *      *                  *      * *

Ps.putida           ACCTTGCTTCGCCGA-GGCACCAGCACAGCAGCCCAGACCACCATGAAAA 388
Es.coli             ACCCTACTGCGCCG--GGCACGCTCGCAG--GCCGCAACAGTCTGCACCA 388
Ps.aeruginosa       GCGCACCCGGGCGAGCAGTTCGACGAAGGCGAACGGCTTCACCAGGTAAT 392
Ps.stutzeri         GCGTACCCGCGCCAGCAGCTCGGAAAAGGCGAACGGCTTGACCAGGTAGT 395
Ma.aq               GCACTCATACGCCGCAGCGAAGGGCGAGTGCACTCCCAGGTTAAAGCCGG 389
                     *        **

Ps.putida           TGGC--GGATCTTGAGGTAGATCTCCTCAAGCGAAGAGCTACTCGAAGTG 436
Es.coli             TCGC--CGATATGACCGTTGATATGGTGCGCCGGACCGTGATCCGTTCGG 436
Ps.aeruginosa       CGTC--GGCGCCTTGTTCGAGGCCGCGCACGCGGTCCTCTACGGCATCGC 440
Ps.stutzeri         CATC--CGCTCCCAGCTCCAGCCCCTTGACCCGGTCCTCGACACGATCCC 443
Ma.aq               TGGCTTCGAACTGGACGAAAATCGCCAGAGCCTGCG---GACAGAAGAAG 436
                       *   *                        *       *

Ps.putida           GCAAGCGAATCGACCTGACCGCTAAAGAGTTTTCACTTCTGGAGCTTTTG 486
Es.coli             GGAAGAAGATCCATCTCACCGGTAAAGAATACGTTCTGCTTGAGTTGCTG 486
Ps.aeruginosa       GGGCGGTCAGGAACAGCACCGGTACGGTCA-TCCCGGCGTCGCGCACGCT 489
Ps.stutzeri         GCGCGGTGAGGAACAGCACCGGCACCTCGT-TGCCCGAGGCACGCACCAG 492
Ma.aq               GAGCAGAACACGCCCTGACGGGTACTGAGTTCCGCCTGCTACGATGCCTG 486
                    *                ** *  *
```

```
Ps.putida       ATGCGCCGGCGCGGTGAGGTGCTCCCCAAATCTCTGATTGCGTCCCAGGT 536
Es.coli         CTGCAACGCACCGGAGAAGTGTTACCCAGGAGTCTTATCTCGTCCCTGGT 536
Ps.aeruginosa   GCGCAGGATCTGCCAGCCGTCACGGCCGGGCAGC--ATGACGTCGAGGAT 537
Ps.stutzeri     TCGCAGCACTTCCCAGCCGTCCAGACCCGGCATC--ATCACATCAAGGAT 540
Ma.aq           ATGAGCCGTCCGGGCCACATCTTTTCCAAGGAACAGCTAATGGAGCAGCT 536
                  *                 *     **      *   *         * *

Ps.putida       GTGGGACATGAACTTCGACAGCGACACCAATGTGATCGAGGTTGCGATAC 586
Es.coli         CTGGAACATGAATTTTGACAGTGATACGAATGTGATTGATGTCGCCGTGA 586
Ps.aeruginosa   CAGCAGGTCGTAGTCGCCGCCCAGGGCCAGT-TGCTCGCCCTCGTTGCCA 586
Ps.stutzeri     CAGCAGGTCGTAGGGCACGCTCAACACGTGC-TGCAGCGCATCGGTGCCG 589
Ma.aq           ATACAACCTGGATGAGAGCCCCAGCGAAAACGTGATTGAGGCGTATATTC 586
                         * *                    **

Ps.putida       GACGGCTACGGGCAAAGATCGATGATGACTTTGCCCCTAAGCTGATTCAC 636
Es.coli         GACGTCTGAGAAGTAAAATTGATGATGACTTTGAGCCAAAACTGATCCAT 636
Ps.aeruginosa   TCGTCGCTCAGGTCGACGGCGAAGCCTGCCTCGGCGAGCCCCTGGCGCAG 636
Ps.stutzeri     CTGGTGACGCGGTCGACGGTGAAGCCGGCCTCGCTCAGCCCCTGCTGCAG 639
Ma.aq           GGCGCTTGAGAAAGCTGGTCGGCAACGA----AACGATCACCACACGC-C 631
                               *                      *     *

Ps.putida       ACAGCTCGGGGAATGGGCTACATGATGGATGC---ACCGGAGTGA----- 678
Es.coli          ACCGTTCGCGGTGCCGGATATGTCCTGGAGATCAGAGAAGAGTGA----- 681
Ps.aeruginosa   GTACTGGCCGATGCGCGGTTCGTCTTCGACGATC-AGCAGTTTCAT---- 681
Ps.stutzeri     GTAGGTGCCGGTTTTCGGCTCGTCTTCGGCGACC-AGCAGTTTCATGCGG 688
Ma.aq           GTGGCCAGGGATACATGTTCAATGCCCAACGTTAA--------------- 666
                         *     *     *
Ps.putida       --------------------------------------------------
Es.coli         --------------------------------------------------
Ps.aeruginosa   --------------------------------------------------
Ps.stutzeri     ATATTCCATCGGTGGTGGACGGCGCAAGTGTGGCAGCGCCGCAC------ 732
Ma.aq           --------------------------------------------------
```

Figure 2 App.B.1(B) - Alignment of the nucleotide sequences obtained by the BLAST query of *copR* and known copper resistance regulators. Alignment was performed using Clustal X (version 2.0) [118]. Legend: Asterisk – Identical nucleotides in all the alignment. For the accession numbers see Table App.3.

C. Alignment with CueR

```
CusR  MKLLIVEDEKKTGEYLTKGLTEAGFVVDLADNGLNGYHLAMTGDYDLIILDIMLPDVNGW  60
CopR  MRLLLVEDDRLLAEGLVRQLEKAGFSIDHTSSAREAQILGEQEDYRAAVLDLGLPDGNGL  60
      *:**:***::  .* *.: * :*** :* :... :.  *.   **   :**: *** **

CusR  DIVRMLRSANKGMPILLLTALGTIEHRVKGLELGADDYLVKPFAFAELLARVRTLLRRGA 120
CopR  EVLKRWRSKLVECPVLVLTARGDWQDKVNGLKAGADDYLAKPFQTEELIARINALIRR-S 119
      ::::  **     *:*:*** *  :.:*:**: ******.***   **:**:.:*:** :

CusR  AVIIESQFQVADLMVDLVSR-KVTRSGTRITLTSKEFTLLEFFLRHQGEVLPRSLIASQV 179
CopR  EGRVHSQVKAGGFELDENRQSLRTEEGAEHALTGTEFRLLRCLMSRPGHIFSKEQLMEQL 179
         :.**.:...: :*   :   *..*:. :**..** **. :: : *.::.:. : .*:
CusR  WDMNFDSDTNAIDVAVKRLRGKIDNDFEPKLIQTVRGVGYMLEVPDGQ 227
CopR  YNLDESPSENVIEAYIRRLRKLVGN----ETITTRRGQGYMFNAQR-- 221
      :::: ... *.*:. ::***  :.*    : * * ** ***::.
```

Legend in the next page

Figure 3 App.B.1(C) - Alignment of the peptides sequences obtained by the NCBI query of cueR (NP_752588.1) in *Es.coli* and *Ma.aq* CopR. Alignment was performed using Clustal X (version 2.0) [117]. Legend: Asterisk – Identical residues in all the alignment, Colon – Conserved substitutions; Dot – Semi-conserved substitutions.

B.2 - Two-component regulatory system, CopS

A. Data from the BLAST protein query

Table 4 App.B.2(A) – Accession numbers of the National Center for Biotechnology Information (protein sequence database) from the BLAST protein query of CopS (†), first 4 most similar sequences of known copper resistance sensors. Algorithm used was blastp (protein-protein BLAST) [117].

Organism	Score	E value	Similarity (%)	Accession number
Ma.aq†	890	0.0	100	ZP_00817991.1
Ps. putida	161	$1x10^{-37}$	33	AAN70949.1
Ps.aeruginosa	170	$3x10^{-40}$	30	CAW26989.1
Ps.stutzeri	152	$6x10^{-35}$	30	ABP80361.1
Es.coli	143	$3x10^{-32}$	30	ABF67807.1

B. Alignment from the nucleotide sequence query result

Table 5 App.B.2(B) - Accession numbers of the National Center for Biotechnology Information (nucleotide sequence database) from the BLAST nucleotide query of *copS* (†), first 4 most similar sequences of known copper resistance sensors. Algorithm used was discontiguos megablast (more dissimilar sequences) [117].

Organism	Score	E value	Similarity (%)	Accession number
Ma.aq†	2397	0.0	100	VIMSS3521612
Ps.stutzeri	57.2	$2x10^{-4}$	4	CP000304.1
Ps.aeruginosa	48.2	0.12	4	FM209186.1
Ps. putida	44.6	1.40	2	AE015451.1
Es.coli	36.2	1.80	1	DQ517526.1

```
Ps.aeruginosa  -------------TCATCGGCGCG-GGCGAGCGAAA--TCGATGAC-GAA 33
Ps.stutzeri    ----------------------TC-AGTTCGGGAAT--TCCAGCAC-GAA 24
Ps.putida      ------------GTGATTTTGACC-CGTTCCTCACTCGTCAAGCGC-CTG 36
Es.coli         GTGAGGTTCAAAATTTCCCTGACCACACGCCTGAGCC-TGATTTTT-TCT 48
Ma.aq          -------------------------ATGCCCAACGTTAAAAGGCCAGCA 24
                                                 *
```

```
Ps.aeruginosa  ACGGGTCCAGCCGTCGGC--AGACTCGCAGC-GGATCTCGCCGCCGTGGG 80
Ps.stutzeri    AGACGTCCGGCCCTGCGC--GGAAGTGCAAT-GGATACGACCCTGATGGG 71
Ps.putida      ACGCTGATGATCATGTTC--GCGGTCATCGC-TGTTCTTGTGGTCGCAGG 83
Es.coli        GCGGTGATGCTTACGGTATGGTGGTTATCAAGTTTTATCCTGATTAGCAC 98
Ma.aq          TCCGTCAAAGGCATGCTGCTGGTGTTATTGCTGCCCGCCGGCATCGCTCT 74


Ps.aeruginosa  CCTGGACGATGGAGCGGCAGATCGCCAGGCCGAGCCCGGCGT-GTTCACC 129
Ps.stutzeri    CCTGAACGATCGAACGGGTAATCGCCAGCCCCAGGCCGGCATTGCTCGGG 121
Ps.putida      CATCAGCTTCAACATGCTGAGCCAGCATCACTTCCGCATGCTCGATGAAC 133
Es.coli        CCTTAATGGCTATTTCGATAATCAGGACCGCGATTTTCTGACAGGTAAAC 148
Ma.aq          GATGGGCG-TAGCCTGGTTTGTCCACGGCCTGCTGCTGGATCGAATGTCC 123
                 *                 *                     *

Ps.aeruginosa  CTGGCCTTCGCGGCGCGCCGGA-TCGGCACGATAGAACCGGTCGAACAGG 178
Ps.stutzeri    CTG-CCTTCGCGGCGCGCCGGA-TCGACCCGATAGAAGCGGTCGAACAAT 169
Ps.putida      AAGCGCTGTCCGAAAAGCTGGAGTCAACGAGACACATCCTGTCAATCCAG 183
Es.coli        TTCAGCTCACCGAAGAGTTTCT-TAAAACAGAGACGTTCAGGAACAAAAC 197
Ma.aq          CGGGAATTCCTTGAAAGCCGCCTTAAAGACGA----AGCTGCCTTTCTGG 169
                     *         *      *      **      * *

Ps.aeruginosa  CGCGGCAAGCGCTCGGGAGGAATCGCCG-----GACCCTGGTTCTCCACG 223
Ps.stutzeri    CGCTCCAGATGCTCGGGCGCGATGGTCG-----CGCCGGGATTGCTCAC- 213
Ps.putida      -GCACCTGGAACGGGAATGGAAGAGCTGA----GACCACAACTCCGCGCA 228
Es.coli        GGATATTAAGTCATTATCAGAAAAAATAAACGATGCGATGGTGGGGCACA 247
Ma.aq          AGCACCAGATTCGCGAGGTCAGAGGTCAA-----GTCGAAACCCTGCAGA 214
                *        *                                 *

Ps.aeruginosa  CT--CAGGCGTCGGTCG-------CCGAGACGGATACGTATCTCG----- 259
Ps.stutzeri    CT--CAAGCGTCACCTGACCCGTGCCGCGCTCGATGGCCACCCGGATGCG 261
Ps.putida      TTGCTGGGTGCTCATCAGGATCTGACGGCGGAGATTCTGACCTCAGAGGG 278
Es.coli        ATGGCTTATTCATTTCTATAAAAAACATGGAAAATGAAAAAATTGTTGAA 297
Ma.aq          CCGGTGATTACTTCCAG------GACGTCTTCCATCATGCCTTCGCCATA 258
                                      *       **

Ps.aeruginosa  ----CCGCCATC--GGCGG--TGAAGCGCAGGGCGTTGT--CCAG-CAGG 298
Ps.stutzeri    C---TCGCCGTC--GGGGG--TGTAGCGCAGCGCATTGG--ACAG-CAGG 301
Ps.putida      TGAGGTACTATT--TTCAGATCTCAAAGCAGTACAAATT--CCAGATAAA 324
Es.coli        CTCTATGCCAAAAATTCTGTTGTTCCAGCGGTCCTGCTTAATAAGTCGGG 347
Ma.aq          CGC-ACGCCCGA---TCGAACCATCATATCGCCAAAGGC---CTGGGAAC 301
                      *                   *           *

Ps.aeruginosa  T-TGGCCAGCACC--CGCCGCAGCATGCCGCGGTCGCCGGGCAGGCTCAG 345
Ps.stutzeri    T-TCGACAACGCC--CGGCGCAGCATCAACAGATCGCC---CTGAATCGT 345
Ps.putida      T-ACAAGCGCGCCGACAAGGAAGAAATGTGGGAATGGCAGGACGAATCCC 373
Es.coli        TGATATTCTCGACTATATGATCCAGACGGAAGAAAATAACACCGTGTACC 397
Ma.aq          CCTTGCTGGCACCACTGATCAACCATGAGCAGAATGGCACGCTTCGTCTT 351
                        *  *                 *             *

Ps.aeruginosa  GCTGC-CTTCGCGCAGC---AGGCGGATGTCGCGGTCTTCCGCCAGC-GG 390
Ps.stutzeri    ACCGG-CACCGGACAGCTCGAAGCGCACGCCGCGCTCGTCCGCCGGC-AG 393
Ps.putida      ATAAT-TTCCGTGGTAT-CACCACTGACGTTGAGATCCAGGGCCAACCAG 421
Es.coli        GCAGTATCTCGCGGCGGGTTGCCGTGACGCCGGAA--CAGGGTAAAAGCA 445
Ma.aq          GAAGGCCGTCAGGCGCC--GGACAGTCCGTCCGATATCCTGGCATACCGT 399
                        *                 *           *
```

```
Ps.aeruginosa  CTGGTAGAA--CTCCAGCAGC-GAGTCGACTTCCGCGCCCAGGTCC-AGG 436
Ps.stutzeri    CTGGTAGTA--GTCGAGCAGTTGCGTACAGAGCTGCTCCAGCGCCA-CCG 440
Ps.putida      CTCCCGGCACCGTCATGCTGATGC-TGGATATCACCAGTCATGCGC-ACT 469
Es.coli        AACATGTCATCATTACGGTTG-CCACGGATACTGGGTATCACACCC-TGT 493
Ma.aq          CACTCTTTTCAGGTGAACGGG-TCACCGATCGTCGTGGTTGTATCCGAGG 448
                                       *

Ps.aeruginosa  GCCTGGCGGCTGGGCGCCA---GCAGGCCGTGGT-CGGCCTTGGCCAGCA 482
Ps.stutzeri    GCTTGGCATC-GGGGACGA---TCAGGCCGTTGT-CGGCCTTGGCGAGGA 485
Ps.putida      TCTTCGAAACCCTGCAACG---GTGGTTTGCGAT-TGGT-TTGGTTATCA 514
Es.coli        TTATGGACAAACT----CA---GTACCTGGCTGT-TCTGGTTCAATATCG 535
Ma.aq          ATCTGGAGGCCCTGAAACGCAGTCAGGCCGAGCTGCACGCCTGGACAGCC 498
                  * *                       *   *       *    *

Ps.aeruginosa  ACAGCATGTCGTTGACCATCGCGGTGAGCCGTTCCAGCTCCTCGAGGTTG 532
Ps.stutzeri    ACAGCATGTCGTCGATCATCCGCGACATGCGCTGCAGTTCCTCAAGGTTG 535
Ps.putida      GTG------CGTTGGTCAGTGCTGGTATAGGCTG--GCTGGTCGCGAAAA 556
Es.coli        GTC---TGGTCTTTATTTCTGTTTTTCTGGGCTG--GCTGACCACACGTA 580
Ma.aq          GTGGTGTCGGTGCTGTTGATCGCACTTCTGGTTGCCG-TCATCTGGTTTG 547
                                          * *   * *   *

Ps.aeruginosa  CCGTGCAGGGCCTCGCGATAATCCTCCAGCGAGCGTGGGCGGGACAATAC 582
Ps.stutzeri    GAATGCAGGTTGTCCTGGTACTGCTCCAGCGAACGCCCGCGCGTCAACGC 585
Ps.putida      GTGGTCTACGCCCCGTCGAACAGATC--ACGAAGGTTGCTACGTCCATAT 604
Es.coli        TTGGTCTGAAACCGCTACGGGAAATG--ACCAGTCTGGCTTCCTCCATGA 628
Ma.aq          GCATCACGCTATCGCTGCGGCCGGTGGTGACACTGAAAGCGGCGTTAAAA 597
                                    *     *             *

Ps.aeruginosa  CACCTGGGTC--TGGGTCAGCAAACTGGTGAGCG-GGGTGCGCAACTCGT 629
Ps.stutzeri    CACCTCGGTG--TGGGTCATCAGGTTGCTCAGTG-GCGTGCGCAGCTCGT 632
Ps.putida      CTGCCAGGTCACTTCAAGAGCGCATCCCTCTGGA-GCCTGTACCACTCGA 653
Es.coli        CCGTACACAGCCTGGATCAGCGTCTAAATCCCGA-TCTGGCTCCGCCGGA 677
Ma.aq          CGATTGCAGGATGGAGAGATTTCCCGGATCAATGCACCATCACCCGAAGA 647
               *                 *         *             *     *

Ps.aeruginosa  G--GGCGATGTCGGCGGAGAACGCCGAGAGG-CGCTGGAAGGCTTCTTCC 676
Ps.stutzeri    G--GGCGATGTCGGCCGAGAAGTTCGACAGC-CGCACGAAGGCGTCTTCC 679
Ps.putida      ACTGCAAACACTGATCTTGTCCTTCAACGGGATGCTTGCA--CGCCTGGA 701
Es.coli        AATCTCTGAGACCATGCAGGAGTTCAATAATATGTTTGAT--CGCCTGGA 725
Ma.aq          ATTTCAACCGCTCGTGATGCAGTTAAACCAGTTGCTTGA---CTCCCTCG 694
                           *       *      *   *    *  *

Ps.aeruginosa  AGGCGCGCCAGCATGGCGTTCAGCTCGCCGGCCAGGCCGCGCAGCTCCTC 726
Ps.stutzeri    AGACGCGCCAGCATCGCGTTGAAGGCCTGTACCAGCTGCTGCAGCTCTGC 729
Ps.putida      AGATGCCTTCGTTCGTCTATCCAATTTCTCAGCTGATATCGCGCATGAGC 751
Es.coli        GGGGGCATTCCGGAAACTGTCAGATTTCTCGTCTGACATCGCGCATGAGC 775
Ma.aq          ACAAGCGACTGGTGCGTTCCCGGGATGCGCTCGCAAATCTGTCACACAGT 744
                   **                                  *

Ps.aeruginosa  TGGCATCCGGCTGGCGTCGAGGCGGG--TGGTGAGGGAGTTGGCGGAAAC 774
Ps.stutzeri    CGGCGTCGAGTCATCGGGAATGCGCT--CGCGCAGCGACTTGGCCGAGAC 777
Ps.putida      TG-CGAACACCTGTCAGTAACCTGTTAACGCACACCGAGGTGGTGCTGAC 800
Es.coli        TG-CGCACACCAGTCAGTAATCTGATGATGCAGACGCAGTTTGCACTGGC 824
Ma.aq          GTCAAAACGCCCATC-GCAGCCGTCCGGCAGATACTGGAG-GATATGGAT 792
                           *                 *
```

```
Ps.aeruginosa  CCGCGC-GGCGACTTCGCGCATCCGTCGCAGCGGCGCCAGGCTGGCGCCG 823
Ps.stutzeri    CGAGGC-CGCCACCTGGGTGACCTCGCGCAGCGGCCGCAGCCCGCTGCGC 826
Ps.putida      CCGGAAACGCGACGTCGACG--CCTATGAAGAAAACCTTTATTCAAACCT 848
Es.coli        TAAGGAAAGGGATGTT-TCG-CATTACCGCGAAATTTTATTCGCTAACCT 872
Ma.aq          CGCCCATTGCCTAGTGATCTAAGAATCCAAATGGCCGCCCGTCTCAGTGA 842
                       *     *

Ps.aeruginosa  GTGGCCCAGGCGCCGAGCAGCGCCG-TGGCCAGCGCCGACAGGCTCATGG 872
Ps.stutzeri    ACCAGGAGCCAGCCGAGCAGCCCGC-TGAGCAGTGCGCACAGCACCAGCG 875
Ps.putida      GGAAGATCTGAAGCGAATGTCACGCATGATCGATG---ACATGCTCTTCC 895
Es.coli        GGAAGAACTGAAAAGGTTGTCACGAATGACCAGTG---ACATGCTTTTTC 919
Ma.aq          CATCGACAGACAATTGGAAGCGGAAATGCGCCGC----AGCCGCTTTGCC 888
                               *     **  *       *

Ps.aeruginosa  TCAGCCAGA-TCAACTGGCGCATGCGTTGCAGGAAGTGCTGGTGATGAC- 920
Ps.stutzeri    CCGCCCAGAGCCAGCCGCTGAAGGTATGGAAGAACGC-CTTGTGCACGG- 923
Ps.putida      TCGCCAAG--TCAGACAACGGCCTCAT-------CATCCCTGAGCAGGT- 935
Es.coli        TGGCACGT--TCAGAGCATGGTCTGCT--------GCGGCTGGATAAACA 959
Ma.aq          GGGCCCCAGGTCGGGAAAAGCGCTTAT--------CCCGTCAAACAGGCG 930
                          *       *      *

Ps.aeruginosa  TGATGTCGAGG-ATCACGGTAA-GTCGCGGCGAAGCGTCG---TTGCCGG 965
Ps.stutzeri    TGACATCGAGC-AGGATCAGCA-GCGTCAGCGCCGTCTGA---CCGACCG 968
Ps.putida      TGACATCCAGC-TTCATGACCTCGTTTCCAAGTTATTTGAGTACTACCAG 984
Es.coli        TGATGTGGATC-TGGCAGCCGAACTGAATGAATTACGTGAGTTGTTCGAG 1008
Ma.aq          CGGGATCTTTTGTGGATGCTGGGGCGGCTGTATCCGGAAAAATCCTTCGA 980
                *   *

Ps.aeruginosa  GCTCCAGCG--------GGGCGCCGAGTTCGCG-GTTGGCGGGCAGGC-C 1005
Ps.stutzeri    GCACCTGCGTCACCTCCGCGCGCCAGAGGCGCCCATCCAGTGACAGCT-C 1017
Ps.putida      TTGCTGGCGGACGATCGAGGCATACGACTCACGCTTCAAGGGAAAGGTGT 1034
Es.coli        CCCCTGGCAGACGAAACAGGAAAGACAATCACGGTTGAAGGAGAGGGCGT 1058
Ma.aq          ACTATCAAGCTCACTGCCGGAA-GACACCCGCTGGCCGATAGAGGAGCAT 1029
                              *         * *

Ps.aeruginosa  GGGATGCGCAGGCTGCGGCAGGCG-CT-CGAACCAGAGGTTG----CCAT 1049
Ps.stutzeri    GCCAT-CGCGCTCAGCGGCCAGCG-CGACCAGGCTCGGGTCGGC-CCCGC 1064
Ps.putida      CATCTCCGGAGACAGGCTCATGAT-CGACCGAGCTGTATCCAATATTCTC 1083
Es.coli        TGTTGCCGGAGACAGCGATATGCT-CCGACGTGCTTTCAGTAACCTGCTT 1107
Ma.aq          GACTT--GAATGAAGTGCTAGGCAACCTGCTCGATAATGCCGGCAAATGG 1077
                      *      *      *   *

Ps.aeruginosa  CCGG--GCCGTTCAG---GCGCAGGCCGAGGTCGGGATG--------GCG 1086
Ps.stutzeri    CTGGCAGCGGCTCGG---GAAACAGCTGGCGCCCCGCTGCATCGAAGACG 1111
Ps.putida      TCGAATGCCGTTCGATACACGCCGGACGGAAGCGGGATC--TCGGTAGAA 1131
Es.coli        TCCAATGCAATCAAGTATTCTCCCGATAACACCTGTACAG--CGATACAC 1155
Ma.aq          TCGT-CGCGGTGCGTAGAACTCTCGCTGAAACAAGACAACAACAGCAGAC 1126
                     **              *

Ps.aeruginosa  TGCCAGTTCGCGCAGCAGTTCC----------GCTTCGCGGCGGCGCAT- 1125
Ps.stutzeri    AACCCCCTTATCTCGCTGTGCCCACCGAGCAGGGCCTGCAGCCGCGGACG 1161
Ps.putida      ATTCAGCAAGCCGCAGACAAGGTAACCCTCACGATCAAAAACGGCGGGGC 1181
Es.coli        CTTGAGCGTGACAGTGACTGTGTGAACGTGATGATTACGAAT-ACGATGT 1204
Ma.aq          AAATCGTTGTTTCTGATGATGGT----CCTGGAGTCAATGGCGACGACT- 1171
                                                          **
```

```
Ps.aeruginosa  --CTGTTGTTCCGAGCCGAGGCCG---CGCAGCTCATCGCGAAACACCGC 1170
Ps.stutzeri    --CAGCCGCTCGAAGTCATCATCGGCCTGCAGCTCGCCGAGCAGCGCCTC 1209
Ps.putida      AACCATTGACCCACAACA-CATCAACAAGATTTTTGATCGCTTCTACCGG 1230
Es.coli        CCGGCCAGGTTCCCGCTAATCTGGAACGTTTGTTTGACCGGTTCTATCGC 1254
Ma.aq          --TGTCCAGTCTGGGGCAACGAGGGC-TGCGACTTGACGAGCAGACCCCT 1218
                                                 *           *

Ps.aeruginosa  G---------AGCTTGGCGGACATCGCCATATGGTCGAGCTCG--ACGAA 1209
Ps.stutzeri    GCCGGCGCGAAGCTTCTCGTGCAGGGTGTGCTGGTCCAGCTCG--TTGAA 1257
Ps.putida      GCCGACCC--GGCAAGGCGAGAGGGGGGACCTAGCAATGCTGGGCTTGGT 1278
Es.coli        GCAGACTC--ATCAAGGTTCTACAACA---CGGAAGGCGCGGGGCTGGGA 1299
Ma.aq          GGCCACG---GTCTGGGCCTCGCGATTGTCCGGGAGATCGTCG-CTCGCT 1264
               *           *                               *    *

Ps.aeruginosa  ATGCTCGTCGA-TGGCCCTGCTGAAGATCAATCCGGCGA-----TCAGCG 1253
Ps.stutzeri    GTGGCGTTCGC-TGAGCTGGTGAAAGAGGTACCCCGCGG-----CCGCCA 1301
Ps.putida      CTGGCCATTAC-CCGCTCCATCGTCAAAGCACATGATGGAAAAGTCTGGT 1327
Es.coli        TTATCAATTAC-AAGGTCGATCATTCATGCTCACGGCGG-----CGAGCT 1343
Ma.aq          ATGAGGGCAACATTAGTTTTTCGACGGGG--CCCGGTAG------CGGCT 1306
                *

Ps.aeruginosa  AGACCGCCGCCGTGCAGGCGGCGAACAGCAGGCTCAG----GCGGACGCC 1299
Ps.stutzeri    GCACGCCGGTCACCGCGAGCATGAACATCAGGCTCAG----CCGTGCGGT 1347
Ps.putida      GCACCTCGTCCGAAGGGGTTACAGCATTCAATTTCATTTTTCCTGCCGCC 1377
Es.coli        GTCAGCAGAACAGCAGGGGCGTGAAATTGTGTTCAAAGT------GCGCC 1387
Ma.aq          TGCGCGTAACCATAGAGTTTTAA--------------------------- 1329
                         *     *

Ps.aeruginosa  GAGGGA-CATCCGCGAGCCGAACCCGGCGCTCAT---------------- 1332
Ps.stutzeri    CAGCGA-CAGCCGCGTCAT------------------------------ 1365
Ps.putida      CACCGAACAGCTCACATTCGCACGCGGCATTAA---------------- 1410
Es.coli        TGTTAATGGATTAA----------------------------------- 1401
Ma.aq          -------------------------------------------------
```

Figure 4 App.B.2(B)– Alignment of the nucleotide sequences obtained by the BLAST query of *copS* and known copper resistance sensors. Alignment was performed using Clustal X (version 2.0) [118]. Legend: Asterisk – Identical nucleotides in all the alignment. For the accession numbers see App. Table 5.

C.Alignment with CueS

```
CusS  MVSKPFQRPFSLATRLTFFISLATIAAFFAFAWIMIHSVKVHFAEQDIN-DLKEISATLE  59
CopS  --MPNVKRPASVKG-MLLVLLLPAGIALMGVAWFVHGLLLDRMSREFLESRLKDEAAFLE  57
           .:** *:   : :.: *.:  *::..**::   :  :::.: ::  **: :* **

CusS  RVLNHPDETQARRLMTLEDIVSGYSNVLISLADSHGKTVYHSPGAPDIREFTRDAIPDKD 119
CopS  HQIREVRG-QVETLQTGDYFQDVFHHAFAIRTPDR------TIISPKAWEPLLAPLINHE 110
      : :..    *.. * * : : . : :.:   : .:      :  :*.  *    .: :::

CusS  AQGGEVYLLSGPTMMMPGHGHGHMEHSNWRMINLPVGPLVDGKPIYTLYIALSIDFHLHY 179
CopS  QNG---------TLRLEGR---QAPDSPSDILAYRHSFQVNGSPIVVV-VSEDLEALKRS 157
       :*         *: : *:   :  .*   ::    .  *:*.** .: :: .::   :

CusS  INDLMNKLIMTASVISILIVFIVLLAVHKGHAPIRSVSRQIQNITSKDLDVRLDPQTVPI 239
CopS  QAELHAWTAVVSVLLIALLVAVIWFGITLSLRPVVTLKAALKRLQDGEISRINAPS--PE 215
```

```
        :*     :.: ::  *:* :: :.:  .  *: ::.  ::.: . ::.    *.  *

CusS  ELEQLVLSFNHMIERIEDVFTRQSNFSADIAHEIRTPITNLITQTEIALSQSRSQKELED 299
CopS  EFQPLVMQLNQLLDSLDKRLVRSRDALANLSHSVKTPIAAVR-----QILEDMDRPLPSD 270
      *:: **:.:*:::: ::. :.*. :  *:::*.::***: :       : :. .:   .*
CusS  VLYSNLEELTRMAKMVSDMLFLAQADNNQLIPEKKMLNLADEVGKVFDFFEALAEDRGVE 359
CopS  LRIQMAARLSDIDRQLEAEMRRSRFAGPQVG---KSAYPVKQARDLLWMLGRLYPEKSFE 327
      :  .   .*: : : :.  :  ::  . *:    *    ..:. .:: ::  *  ::..*

CusS  LR-FVGDECQVAGDPLMLRRALSNLLSNALRYTPTGETIVVRCQTVDHLVQVTVENPGTP 418
CopS  LSSSLPEDTRWPIEEHDLNEVLGNLLDNAGKWSSR--CVELSLKQDNNSRQIVVSDDGPG 385
      *   : :: : . :   *...*.***.** :::.    : :  :  ::  *:.*.: *.

CusS  IAPEHLPRLFDRFYRVDPSRQRKGEGSGIGLAIVKSIVVAHKGTVAVTSDVRGTRFVIIL 478
CopS  VNGDDLSSLGQRGLRLDEQTP----GHGLGLAIVREIVARYEGNISFSTGP-GSGLRVTI 440
      :  :.*. * :*  *:* .      * *:*****:.**. ::*.::.::.  *: : : :

CusS  PA 480
CopS  EF 442
```

Figure 5 App.B.2(C)- Alignment of the peptides sequences obtained by the NCBI query of *Es.coli* CueS (NP_752587.1) and *Ma.aq* CopR. Alignment was performed using Clustal X (version 2.0) [118]. Legend: Asterisk – Identical residues in all the alignment, Colon – Conserved substitutions; Dot – Semi-conserved substitutions.

B.3 A novel copper protein, CopX

Table 6 App.B.3 - Accession numbers of the National Center for Biotechnology Information (protein sequence database) from the BLAST protein query of CopX (†), first 4 most similar sequences of copper binding proteins. Algorithm used was blastp (protein-protein BLAST) [117]. Last protein was added for comparison with alignment *Ps.sy* CopC.

Organism	Score	E value	Similarity (%)	Accession number
***Ma.aq*†**	365	$8x10^{-100}$	100	YP_957419.1
Al.b	263	$4x10^{-69}$	76	YP_693084.1
Ps.a	175	$2x10^{-42}$	46	YP_002439868.1
Ps.c	147	$4x10^{-34}$	46	ABC68313.1
Ps.p	142	$1x10^{-32}$	44	YP_001748550.1
Ps.sy	-	-	-	YP_234583.1

Appendix C - Protein phylogeny

C.1 Regulator, CopR

Table 7 App.C.1 – Accession numbers of the National Center for Biotechnology Information (protein sequence database) for proteins used in the phylogenic inference of *Ma.aq* CopR (†). Protein regulators are Eubacteria related heavy metal resistance regulators [120, 121].

Organism	Accession number
Ma.aq†	YP_957418.1
H.influenzae	P44617
Sy.elongatus	Q8DHQ7
Ra.solanacearum	Q8XU46
Synechocystis sp. PCC 6803	Q55963
Ps.putida	Q8KWW1
Ps.syringae pv. tomato	Q889K0
Ps.putida KT2440	Q88QA9
Ps.aeruginosa PAO1	AAG08164
Ps.aeruginosa	Q9HV30/YBH8_PSEAE
Sa.typhimurium	Q8ZRG6
Prot.mirabilis	O30821
Ag.tumefaciens str. C58	Q8UG45
Br.suis	Q8G2U3
Si.meliloti	P58379/HMR2_RHIME
Si.meliloti	P58378/HMR1_RHIME
Si. medicae WSM419	Q9X5X4/HMRR_RHIME
Ag.tumefaciens str. C58	Q8UGU7
Me.loti	Q98C35
Rh.leguminosarum bv. viciae	Q9X5V4/HMMR_RHILV
Sh.oneidensis	Q8EGB7
V.cholerae	P0C6D2/CUER_VIBCH
V.vulnificus	Q8D7N8
V.parahaemolyticus	Q87IE5
Sa.typhimurium	Q93CH6
Sa.typhi	Q8Z8S3/CUER_SALTI
Es.coli O157:H7	Q8XD09/CUER_ECO57
Es.coli O6	Q8FK74
Es.coli K12	P0A9G4/CUER_ECOLI
Y.pestis	Q8ZCA8/CUER_YERPE
Ps.putida	Q8VMH2
Ps.putida	Q8GJ85
Ra.solanacearum	Q8XSZ7
Ps.aeruginosa	Q8GQ13
Ps.syringae pv. tomato	Q87UL6
Ps.putida	Q93TP7
Ps.putida KT2440	Q88CP0
Ps.aeruginosa	Q9HXV1
K.pneumoniae	Q6U5Q1
Pa.multocida	Q9CJQ3
Ac.ferrooxidans	P22896
Y.pestis	Q8ZJ85
Y.pestis	Q8CZI8
Sa.typhimurium	Q8ZLM3
Es.coli O6	Q8FD09
Sa.typhi	Q8Z1X2
V.cholerae	Q9KV79
V.vulnificus	Q8DD05
V.parahaemolyticus	Q87KT2
Sh.oneidensis	Q8EJM0
Stret.coelicolor	Q9L1K7
My.bovis AF2122/97	CAD95484
Ps.aeruginosa	Q9HU68
Nostoc sp. PCC 7120	Q8YWI4
Nostoc sp. PCC 7120	Q8ZS75
Br.melitensis	Q8YJM8
Me.lotis	Q98IB7
Si.meliloti	Q92RB5
Ag.tumefaciens str. C58	Q8UGZ7
Ps.syringae pv. tomato	Q87WC9
Prov.rettgeri	Q8RKW6
Pseudomonas sp. BW13	Q9F3U8
Psd.haloplanktis	Q9EY71
Ps.alcaligenes	Q9XAX1
Ps.aeruginosa	Q8GQ23
Pseudomonas sp.	Q936U6
Ps.putida	Q52395
Ac.ferrooxidans	Q56281
Si.meliloti	Q930P4
Sta.aureus	Q93IB0
Sh.woodyi ATCC 51908	YP_001759051
Strec.agalactiae serogroup V	Q8DX34
Ba.licheniformis	Q9F4C7
Al.borkumensis	Q0VPT5

C.2 –Regulators tree from E.A. Permina *et al.* 2005 study

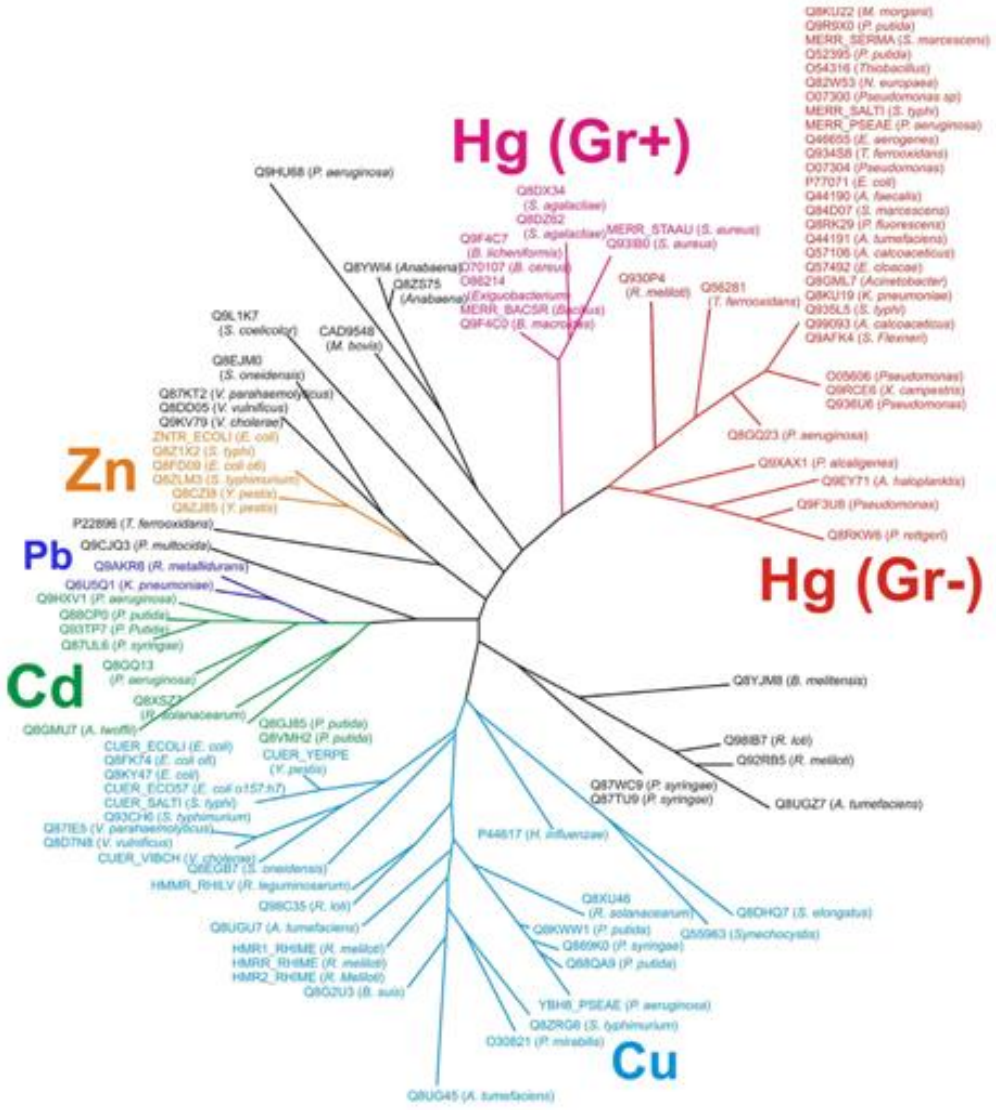

Figure 6 App.C.2 – The tree of regulators from the E.A. Permina *et al.* 2005 study (Figure I). Different specificity is shown by the color code. Legend: Red and magenta are for Gram-negative and Gram-positive members of MerR subfamily, respectively. Light blue is for members of CueR and HmrR subfamily, green and deep blue are for members of CadR and PbrR subfamilies and orange is for members of ZntR subfamily. The identificators are given according to SWISSPROT Database. Black denotes regulators, whose specificity could not be specified. Figure from [120].

C.3 - CopX

Table 8 App.C.3 – Accession numbers of the National Center for Biotechnology Information (protein sequence database) for proteins used in the phylogenic inference of *Ma.aq* CopX (†). Sequences of copper-biding proteins and small copper proteins associated with copper resistance were used [120].

Organism	Accession number
***Ma.aq*†**	YP_957419.1
Ma.alg.	ZP_01892186.1
Al.b.	YP_693084.1
M.sp ELB17	ZP_01739821.1
Ps.mendocina	YP_001187723.1
Ps.mendocina	YP_001187722.1
Ro. nubinhibens	ZP_00958990.1
O. Antarcticus	YP_002632579.1
Rhodobacterales	ZP_01013009.1
Ps. aeruginosa	NP_251497.1
Ps. putida	NP_744308.1
Ps. fluorescens	YP_349582.1
Ps. fluorescens	YP_349175.1

Appendix D – Strain and vectors

D.1. *CopX* sequence

ATGAATGCATCAAACCTGATTGTGGCCGGCCTGGCTTTTTCGATGTCGGT
ATCGGCGTTTGCGGCTGGCACCCATGGCGGTGGTCACGGCCATGGTGCTT
CCATTGGCGAACCGGGAAAAGCCTCAGAAGCCAGCCGGACCATAACGGTT
GAAATGCACGACAACTACTACGAACCCGAAGAAATCCGTGTGAAGCCGGG
CGAAACCGTTCGGTTTGTGGTGCAGAACAAGGGCAATCTTGTGCACGAGT
TCAACATCGGGACTCCGGGTATGCATGAGGCCCACCAGAAAGAAATGAGA
ATGATGGTCGAGCATGGCGTGATCCAGGGTAACAAGTTGAATCATGACAT
GATGAACATGGACATGGGCAACGGCCACTCGATGAAACATGACGATCCCA
ACAGCGTCCTGCTGGAGCCGGGCCAGAGCCGGGAAGTGGTCTGGACGTTT
GCCAATCAGGGCAATATCGAATTCGCATGTAACGTACCAGGGCACTACCA
GTCTGGCATGTACGGTGACGTGAATTTCGAA**TAA**

Figure 7 App.D.1 – *CopX* sequence was retrived from the *Ma.aq* VT8 chromosomal genome, locus coordinates 152394-152927. The sequence has 534 bp. Green highlight correspond to the START codon and red highlight correspond to the STOP codon.

D.2. pET21-c(+)

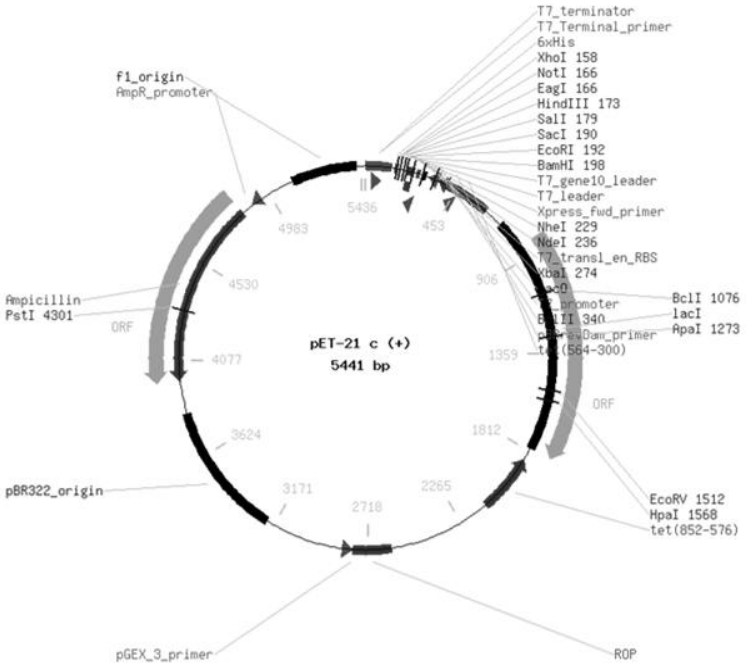

Figure 8 App.D.2 – Graphical backbone representation of the expression plasmid, pET – 21c (+) (5441bp) used to clone *copX*. Image was retrived from EMB Biosciences – Novagen pET system manual [144].

D.3 – Restriction Map of *copX*

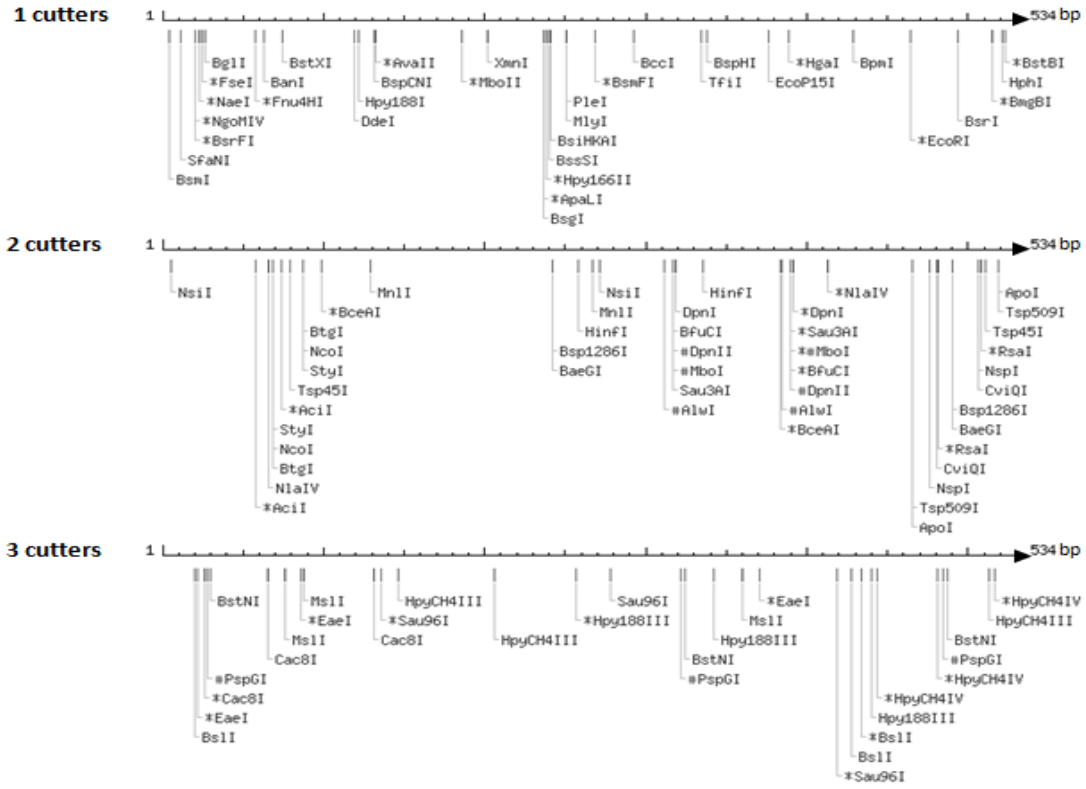

Figure 9 App. D.3 – Restriction map of CopX, obtained using NEBcutter (version 2.0) [178].

Appendix E – *Ma.aq* 617 growth

E.1 Cadmium stress – Liquid medium

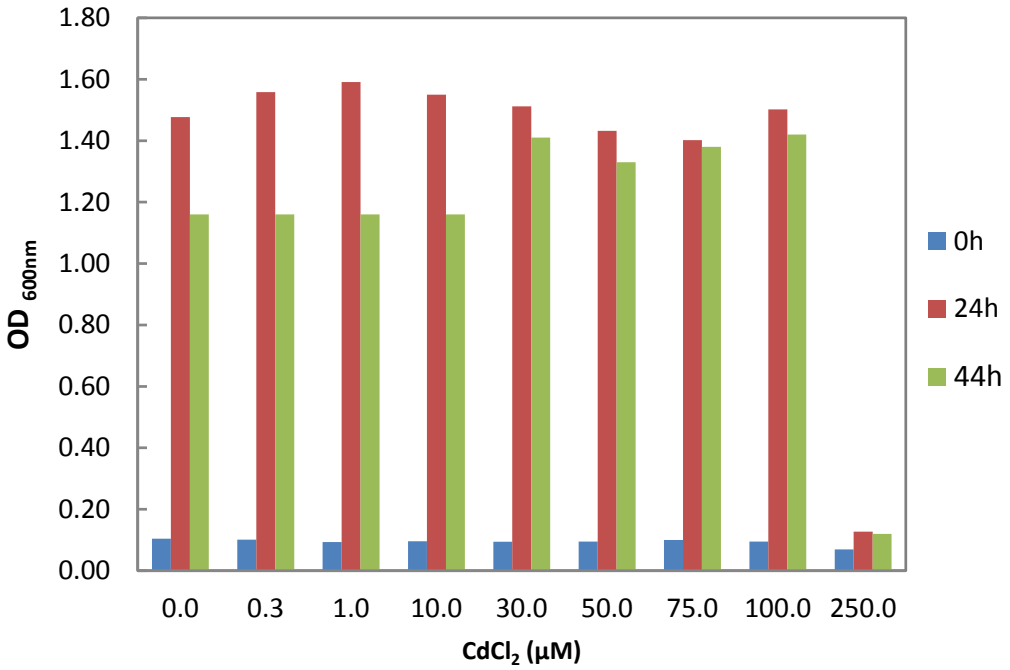

Figure 10 App. E.1 – Selected OD values at 600nm of *M.aq* growth curve of cadmium ion stress followed during two days (nearly 48h). 0 h (inoculums moment), 24 h and 48 h of incubation.

E.2 Copper stress – Solid medium

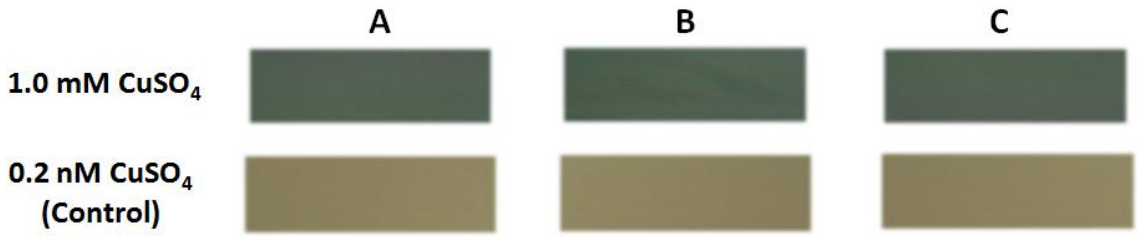

Figure 11 App. E.2 – Temperature effect in plastes used in the solid assays. Control plates – 0.2 nM $CuSO_4$ and stress plates 1.0 mM CuSO4. **A)** 24h incubation; **B)** 72 h incubation, **C)** 120 h incubation at 30 °C.

E.3 Copper stress – Liquid medium

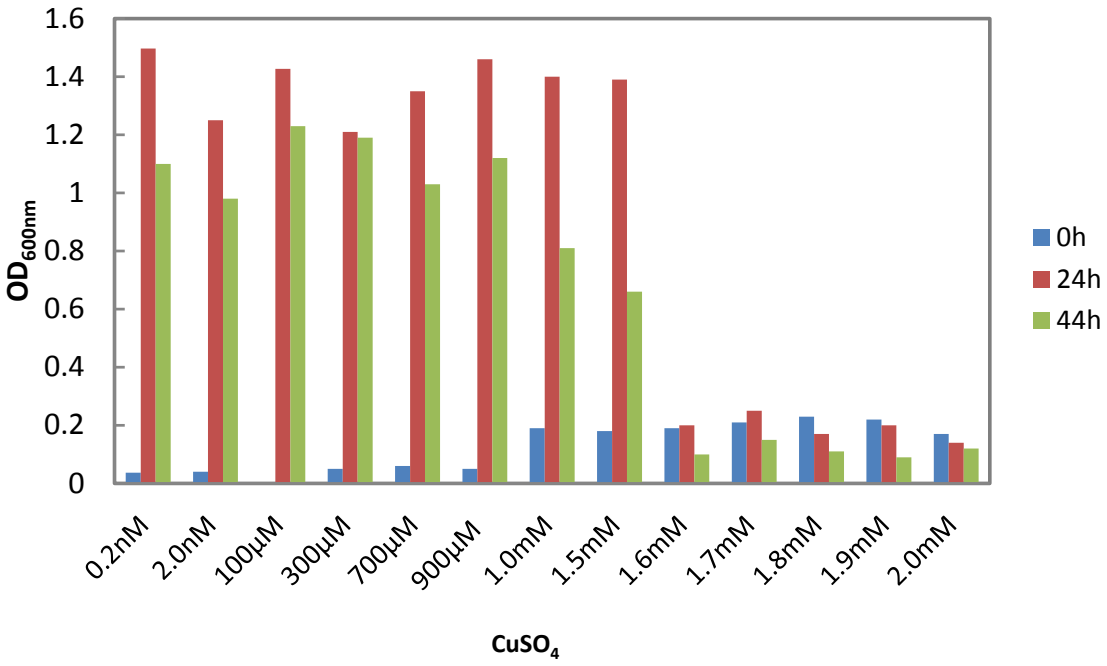

Figure 12 App. E.3– Selected low and high copper ion concentrations OD values at 600nm of *Ma.aq* growth curve for copper ion stress. The growth was followed during two days (nearly 48h). 0 h (inoculums moment), 24 h and 48 h incubation at 30 °C.

Appendix F – Stress induced by Co^{2+} - UV-visible spectra of cellular fractions

F.1 - Citoplasmatic fraction

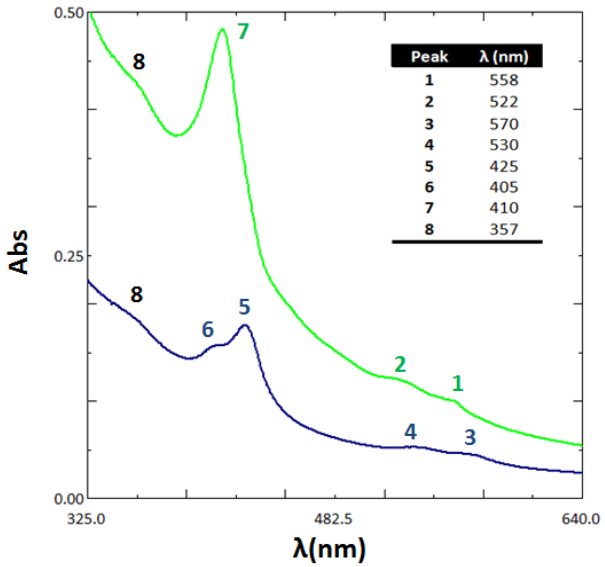

Figure 13 App. F.1– UV–visible spectra of the *Ma.aq* citoplasmatic fraction of the control growth, 0.75 nM $CoCl_2$ (green line), and of the growth in the presence of 1.0 mM CoCl2 (blue line).

F.2 - Membrane fraction

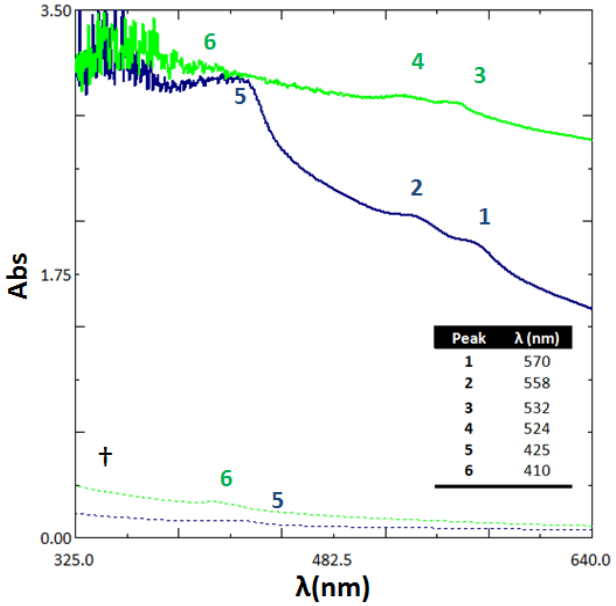

Figure 14 App.F.2 – UV–visible spectra of the *Ma.aq* membrane fraction of the control growth, 0.75 nM $CoCl_2$ (solid green line), diluted 0.75 nM CoCl2 (dashed green line) and of the growth in the presence of 1.0 mM CoCl2 (solid blue line), diluted 1.0 mM CoCl2.(dashed blue line).

Appendix G – Preliminary proteomic analysis

G.1 - Low molecular weight detail – Supposed CopX

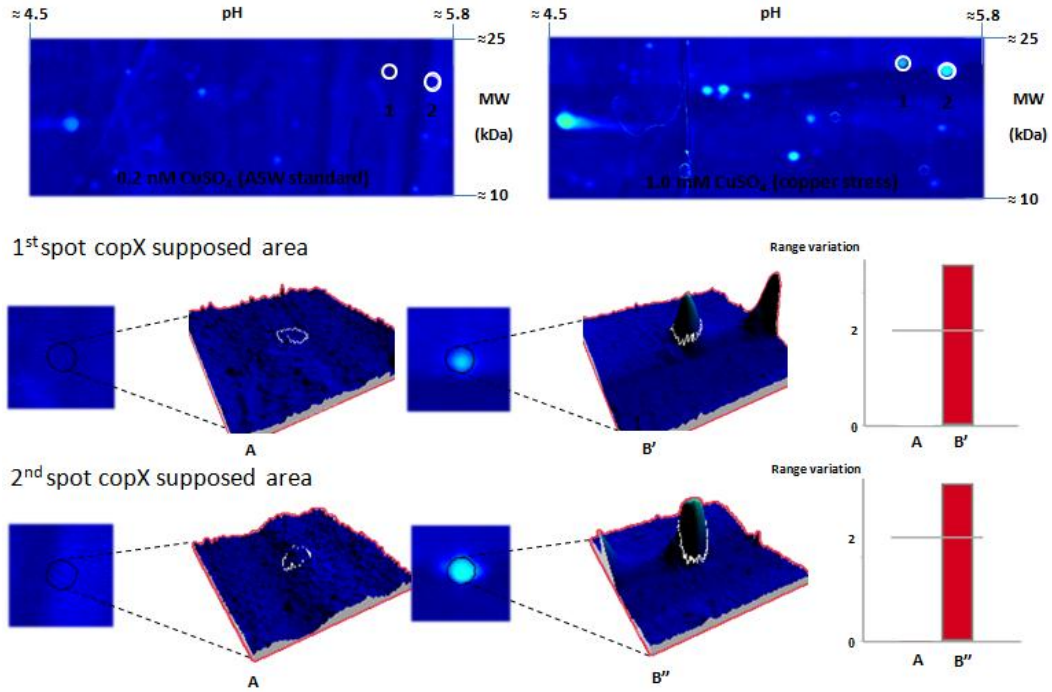

Figure 15 App G.1– 3D view and gel range variation analysis of the 12.5% polyacrylamide gel of 4-7 pH 13 cm Immobiline DryStrip run CopX possible area, see Figure 30. **A)** Reference gel; **B)** Copper induce stress (1.0mM $CuSO_4$). Analysis performed using Ludesi REDFIN3 [138].

G.2 – High molecular weight proteins analysis

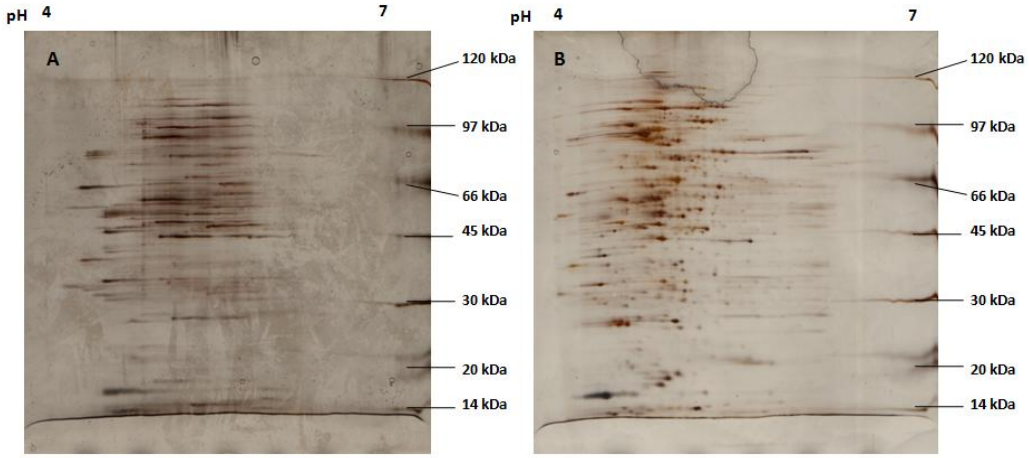

Figure 16 App G.2 – 2D periplasmatic proteome electrophoresis of *Ma.aq* in 4-7 pH 13 cm Immobiline DryStrip. **A)** Control growth with 0.2 nM $CuSO_4$ (30µg applied); **B)** Copper stress 1.0mM $CuSO_4$ (30µg applied). 10% Polyacrylamide gel. Visualization method – Silver staining.

G.3 - Maldi-TOF-TOF-MS analysis

Table 9 App. G.3 – Maldi-TOF-TOF-MS analysis of the protein spots removed from the 2D gel electrophoresis, Figure 32. Legend: † - Protein spots were analyzed from a different gel (periplasmic fraction of a control growth, 0.2 nM $CuSO_4$ but with 6 times more carbon source).

Spot No.	Protein Name	†Molecular Weight (kDa)	†Predicted celular Location	Score	Coverage (%)	†Expression (Range variation)
1	-	-	-	no significant pmf score	-	-
2†	electron transfer flavoprotein, alpha subunit	32.21	periplasm	124	31	1.8
3†	phosphate ABC transporter periplasmic binding protein	34.90	periplasm	83	55	1.5
4†	TRAP transporter solute receptor TAXI family protein	32.30	periplasm	64	10	6.1 Up-regulated upon copper stress
5†	bacterial extracellular solute-binding protein	25.00	extracellula	68	10	3.0 Down-regulated upon copper stress
6†	TRAP dicarboxylate transporter-DctP subunit-	38.63	periplasm	95	38	3.6 Down-regulated upon copper stress
7†	succinyl-CoA synthetase, alpha subunit	26.75	periplasm	94	32	1.1
8†	electron transfer flavoprotein beta-subunit	26.58	periplasm	95	34	2.1 Down-regulated upon copper stress
9	-	-	-	no significant pmf score	-	-
10	copper-binding protein	17.20	periplasm	91	40	Only present under copper stress
11	copper-binding protein+keratin	17.20	periplasm	81	21	Only present under copper stress

†Cellular location predited from the aminoacid sequence without signal peptide analysis and expression analysis performed by GE ImageMaster 7.0.

Appendix H – Biochemical characterization

H.1 - CopX ionic strength equilibrium

Table 10 App.H.1 – Apparent molecular weight values of heterologous expressed CopX determined by size-exclusion chromatography at different [NaCl] in 50 mM Tris-HCl, pH 7.6.

[NaCl] (mM)	Apparent MW (kDa)
0	14.3
50	22.1
150	20.7
300	20.2
500	20.8
1000	18.2