Raquel Sofia Correia Cordeiro
Licenciada em Química Aplicada

Structural and functional characterization of ModA and TupA proteins belonging to the molybdate and tungstate transport system from *Desulfovibrio alaskensis* G20

Dissertação para obtenção do Grau de Mestre em Bioquímica

Orientadora: Doutora Márcia Correia, Pós-Doc, FCT/UNL

Co-orientadoras: Doutora Teresa Santos-Silva, Investigadora Auxiliar, FCT/UNL, Doutora Maria Gabriela Rivas, Investigadora Auxiliar, CONICET

Júri:

Presidente: Prof. Doutor Carlos Alberto Gomes Salgueiro
Arguente: Doutor Rui Miguel Lourenço Rocha de Almeida
Vogais: Doutora Márcia Alexandra Silva Correia

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Novembro de 2014

Raquel Sofia Correia Cordeiro

Licenciada em Química Aplicada

Structural and functional characterization of ModA and TupA proteins belonging to the molybdate and tungstate transport system from *Desulfovibrio alaskensis* G20

Dissertação para obtenção do Grau de Mestre em Bioquímica

Orientadora: Doutora Márcia Correia, Pós-Doc, FCT/UNL

Co-orientadoras: Doutora Teresa Santos-Silva, Investigadora Auxiliar, FCT/UNL, Doutora Maria Gabriela Rivas, Investigadora Auxiliar, CONICET

Júri:

Presidente: Prof. Doutor Carlos Alberto Gomes Salgueiro
Arguente: Doutor Rui Miguel Lourenço Rocha de Almeida
Vogais: Doutora Márcia Alexandra Silva Correia

Novembro 2014

Structural and functional characterization of ModA and TupA proteins belonging to the molybdate and tungstate transport system from *Desulfovibrio alaskensis* G20

Ao meu Pai,

Acknowledgments

À **Fundação para a Ciência e Tecnologia** por ter financiado e suportado este projecto: EXPL/BBB-BEP/0274/2012.

À **Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa**, instituição que me acolheu durante a licenciatura em Química Aplicada – Biotecnologia e no mestrado em Bioquímica onde me foi proporcionada toda a minha formação académica.

À **Professora Drª Maria João Romão** por me ter recebido mais uma vez no seu grupo de investigação e laboratório.

À **Drª Maria Gabriela Rivas**, minha co-orientadora e *PI* deste projecto, por me ter dado esta oportunidade, por ter e continuar a acreditar em mim e no meu trabalho e principalmente por me apoiar e fomentar o incentivo em novos projectos que contribuem para a minha formação académica e pessoal.

À minha orientadora, **Drª Márcia Correia**, por me ter recebido novamente como sua aluna e me ter deixado mostrar que sou capaz de ser autónoma, de desenvolver e ter pensamento crítico. Nunca é demais agradecer tudo o que me ensinaste, todo o apoio, confiança e motivação que me deste desde o primeiro dia, incondiconalmente: Obrigada!

À minha co-orientadora **Drª Teresa Santos-Silva**, pelo entusiasmo, ânimo e ajuda que me deu desde o dia em que iniciei esta tese de mestrado, mostrando-me que podemos sempre melhorar e motivando-me constamente.

I would like to thank to **Dr Rashmi Nair**, to would be with me in my firsts days in this project. Thank you for everything!

À minha (sub-chefe) **Ritinha**: foste verdadeiramente incansável comigo. Sem a tua ajuda e o teu apoio, principalmente nos dias em que o cansaço falava mais alto, não teria conseguido. Foste e és sem dúvida uma referência para mim.

Ao **Dr. Filipe Freire**, ao **Marino Santos** e ao **Francisco Leisico** por toda a ajuda na bancada, por me terem esclarecido as pontuais dúvidas experimentais e teóricas; por suportarem todo o meu stress pré, durante e pós-tese. Obrigado pela amizade e motivação ao longo deste árduo ano de trabalho.

Ao grupo de investigação do **X-TAL e RMN** que me receberam com todo o carinho: **Drª Ana Luísa**, **Drª Angelina Palma**, **Drª Benedita Pinheiro**, **Drª Catarina Coelho**, **Cecília Bonifácio**, **Diana Ribeiro**, **Hugo Correia**, Professor **Dr**. **Jorge Dias**, Professora **Drª Maria dos Anjos Macedo** e **Viviana Correia**. Obrigada pela amizade e apoio que me deram durante este ano.

Ao grupo de investigação de BioProt & BioIn, por me terem recebido e ajudado ao longo de vários meses de trabalho, mostrando-se sempre disponiveis e acessíveis: Drª Luísa Maia, Dr. Rui Almeida, Drª Sofia Pauleta, Drª Susana, Dr. Maiti Biplab,Drª Olga Mestre, Ana Teresa Lopes, Cíntia Carreia, Cláudia Nóbrega, Rui e Mariana.

Aos meus amigos e colegas de mestrado, **Ana Chocalheiro**, **Mariana Romão** e **Luís Fonseca**, por termos partilhado dois anos académicos, mas principalmente, por terem estado do meu lado nos momentos em que o desânimo queria vencer.

À minha melhor amiga, **Maria Inês**, que apesar da distância geográfica e embora não perceba muito do meu trabalho, partilhe comigo, todos os dias, desde que entrei para a faculdade, o meu entusiasmo e fascínio pela ciência e por ter estar do meu lado incondicionalmente, desde que me lembro. És uma verdadeira inspiração e um pilar na minha vida!

À minha grande amiga, colega de licenciatura e de casa, **Joana Marques**, por tudo o que fez e continua a fazer por mim a nivel pessoal e por ter estado (e continuar) sempre do meu lado, nos bons e maus momentos. Por toda a partilha e cumplicidade ao longo destes anos: foste e és imprescindivel. Obrigada pela motivação constante, pelo apoio e sobretudo pela amizade.

À minha amiga e madrinha académica, **Diana Freire**, pelo exemplo de vida e pela referência académica e pessoal que é para mim. Obrigada por esta amizade, por me ajudares e me incentivares um bocadinho todos os dias.

À minha amiga e afilhada académica, **Elisabeth Pires**, por toda a amizade, apoio e por acreditar todos os dias em mim e me incentivar igualmente. Por todas as palavras de conforto e amizade quando as coisas nem sempre corriam bem.

A toda a minha família, nomeadamente, aos meus afilhados, **Inês Santos** e **André Santos**, às minhas manas do coração, **Bruna Quintas** e **Ana Rita Quintas**, à minha madrinha **Vera Correia**, à minha tia **Laura Quintas**, ao meu primo **José Carlos** e aos meus tios **Mili** e **José Calé** por todo o incentivo presente ao longo destes anos, mas sobretudo pelo carinho, amizade e amor que me dão todos os dias, ajudando-me a acreditar que posso e consigo ir sempre mais além.

À **Dana**, minha fiel companheira, que esteve comigo em todos os momentos de estudo e de escrita, angústia e felicidade.

Ao avô **Eurico** e à avó **Olímpia**, aos tios **Paula** e **Artur Marques**, por terem confiado em mim e me terem recebido e acolhido da maneira mais calorosa e amiga possível. Nunca serei capaz de agradecer e expressar toda a minha gratidão: um obrigado nunca será suficiente.

À minha avó **Deolinda**, que está comigo desde que me conheço, embora não perceba o que faço e o que estudo e já não partilhe também os meus problemas e angústias diariamente; está sempre comigo em pensamento. Obrigada por todos os miminhos que me deste e por me motivares a lutar pelos meus sonhos.

Ao meu namorado, **Bernardo Moreira**, que me acompanhou ao longo de todo o meu percurso no mestrado, motivando-me e reconfortando-me quando nem tudo corria tão bem. Obrigado por estares a meu lado neste trajecto importante na minha vida e por poder partilhar contigo o meu entusiasmo e alegria, mas também as minhas frustrações diárias. O teu apoio é um enorme suporte.

Finalmente, e não menos importante, quero agradecer aos meus pais: **Luísa** e **José Cordeiro**, sem eles nada disto seria possível. Obrigado por lutarem pelo meu sonho comigo e de certa maneira me terem ajudado a conseguir chegar até aqui, acreditando sempre em mim e na minha força de vontade. Quero agradecer todos os valores que me transmitiram, nomeadamente, o espírito de sacrífico, o esforço e a dedicação. Como me ensinaram, mais tarde serei sempre recompensada.

Abstract

In this thesis was described the mutagenesis, expression, purification and crystallization of three mutants of TupA as well as the evaluation of the solved structure of TupA. The three mutants were crystallized in PEG8000; however, the diffraction experiment showed the bad quality of the crystals. Optimization of the crystallization conditions of the mutated forms and native protein in the presence and absence of ligands are under way. ModA had previously been expressed, purified, crystallized and the crystal diffracted up to 3.5 Å resolution, ongoing the optimization trials. ModA crystallized in PEG6000 and also with Benzamidine-HCl, diffracting up to 2.77 Å resolution. Crystallization in the presence of metals is underway for TupA and ModA proteins. By SAXS technique, we determined low resolution models for ModA in solution and in the presence and absence of ligands.

Biochemical characterization of the two proteins and corresponding mutants was developed through analysis of native gels, where TupA describes a higher conformational change when is binding to tungstate; however it can also binding molybdate. The conformational change of ModA is similar whether connected to the molybdate or tungstate. The mutation of R to K kills the binding with the molybdate but not with tungstate, while the other two mutants seem to have no binding for molybdate and decrease for tungstate. In order to defining the K_D for TupA, competitive studies were performed by ITC. The same strategy wasn't possible for ModA, since its high and equivalent affinity for both ligands. The results obtained for mutants by ITC also confirm those obtained for the native gels. The analysis of urea gels suggests a conformational change since TupA adopts a more compact conformation when bound tungstate but not to the molybdate. In ModA, the same behavior is observed, proposing that this protein is less selective than TupA. Concerning mutated forms of TupA, no conformational change was found.

Keywords: Molybdenum, Tungsten, X-ray crystallography, Isothermal titration calorimetry, ABC transporters, Small angle X-rays scattering

Resumo

Nesta dissertação foram descritas técnicas como mutagénese, expressão, purificação e cristalização de três mutantes da TupA bem como a avaliação da estrutura da proteína nativa. Os três mutantes cristalizaram em PEG8000, contudo a experiência de difracção demonstrou a má qualidade dos cristias. A optimização das condições de cristalização dos mutantes e da proteína nativa na presença e ausência de ligandos está a decorrer. A ModA foi previamente expressa, purificada, cristalizada, com uma difraação até 3.5 Å de resolução, optimizando-se as condições de cristalização. A ModA cristalizou em PEG6000 na presença de Benzamidina-HCl e difratou até 2.77 Å de resolução. A cristalização com metais está em curso para a TupA e ModA. Através da técnica de SAXS, determinaram-se modelos de baixa resolução para a ModA em solução, também na presença e ausência dos ligandos.

A caracterização bioquímica das duas proteínas e dos mutantes correspondentes foi desenvolvida através da análise de géis nativos, onde a TupA descreve uma elevada alteração conformacional quando ligada ao tungstato; contudo, consegue também ligar molibdato. A alteração conformacional da ModA é similar quando ligada ao molibdato ou tungstato. A mutação de R para K impede a ligação ao molibdato mas não ao tungstato, enquanto os outros dois mutantes parecem não ter qualquer ligação para o molibdato e apresentam um decréscimo para o tungstato. Para definir o K_D da TupA efectuaram-se estudos de competição por ITC. A mesma estratégia não foi possível para a ModA, uma vez que esta aprenta elevada e equivalente afinidade aos ligandos. Os resultados obtidos para os mutants por ITC confirmam também os obtidos para os géis nativos. A análise dos géis de urea sugerem uma alteração conformacional dado que a TupA adopta uma conformação mais compacta quando ligada ao tungstato mas não ao molibdato. Na ModA, o mesmo não foi observado, propondo-se que esta proteína é menos selectiva que a TupA. Considerando os mutantes da TupA, não foram encontradas alterações conformacionais.

Palavras-chave: Molibdénio, Tugsténio, Cristalografia de raios-X, *Isothermal titration calorimetry,* Transportadores ABC, *Small angle x-ray scattering*

Contents

List of figures

List of tables

Symbols and Notations

α – Wave's phase

α_{hkl} - Reflection

ε – Extinction Molar Coefficient

ΔG – Gibb's free energy

ΔH – Enthalpy

ΔS – Entropy

Δρ (r) – Contrast

λ – Wavelenght

ϕ - Angle about the N-Cα bound of polypeptide chain

ρ*xyz* - Eelectron density

σ*hkl* - Associated errors with each intensity

Ψ - Angle about the C-Cα bound of polypeptide chain

|**Fcalc**| **-** Amplitude of the matching structure factor calculated

|**Fhkl**| **-** Amplitude of structure factors

|**Fobs**| **-** Amplitude of the observed structure factor

ADP - Adenosine diphosphate

AMP- Adenosine monophosphate

AO – Aldehyde oxidase

AOR - Aldehyde oxirreductases

ATP - Adenosine triphosphate

Bis-Tris - Bis(2-hydroxyetyl)amino-tris(hydroxymetiy)mehtane

BSA - Bovine serum albumin

CCP4 – Collaborative Computational Project Number 4

cPMP - Cyclic pyranopterin monophosphate

D. alaskensis - *Desulfovibrio alaskensis*

***D. gigas* -** *Desulfovibrio gigas*

DMSO - Dimethyl sulfoxide

DNA - Deoxyribonucleic acid

***E. coli* -** *Escherichia coli*

EDTA – Ethylenediaminetetraacetic acid

ESRF - European Synchrotron Radiation Facility

F – Wave amplitude

FAD - Flavin adenine dinucleotide

Fcal – Calculated structure factor

FDH - Formate dehydrogenases

F*hkl* - Structure Factors

Fobs - Observed structure factor

FT - Fourier transform

FT-1 - Inverse Fourier transform

GMP - Guanosine monophosphate

GTP - Guanosine triphosphate

h,k,l - Miller's indices

HEPES - (4-(2-hydroxyethyl) - 1-piperazine ethanesulfonic acid)

ID - Insertion Device

$\mathbf{I_{hk}}$ - Intensities

I*hkl* – Intensity of reflection

IPTG - Isopropyl-β-D-thiogalactoside

ITC - Isothermal titration calorimetry

KD – Dissociation constant

kDa - kiloDalton

LB - Luria-Bertani

MAD - Multiple Wavelength Anomalous Dispersion

MCD - Molybdopterin cytosine dinucleotide

MGD - Molybdopterin guanine dinucleotide

MIR - Multiple Isomorphous Replacement

Moco - Molybdenum cofactor

MPT - Molybdopterin

MPT-AMP - Adenylylated Molybdopterin

MR - Molecular Replacement

n - Stoichiometry

NAD – Nicotinamide adenine dinucleotide

NBD - Nucleotide-binding domains

Ni-NTA - Nickel-nitrilotriacetic acid

NMR - Nuclear magnetic resonance

OD - Optical Density

P. furious - *Pyrococcus furiosus*

PCR - Polymerase Chain Reaction

PDB – Protein Data Bank

PEG – Polyethylene glycol

pH- Potential hydrogenionic

pI – isoelectric point

RMSD - Root Mean Square Deviation

SAD - Single Wavelenght Anomalous Dispersion

SAM - S-adenosylmethione

SAXS - Small-angle X-ray scattering

SBR - Sulfate reducing bacteria

SDH - Sulfite Dehydrogenase

SDS-PAGE - Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

SO - Sulfite oxidase

Tris - Tris(hydroxymetyl)aminomethane

VM – Matthews coefficient

XDH - Xanthine Dehydrogenase

XO - Xanthine oxidase

Wco – Cofactor of tungsten

W-MPT- Tungstopterin

W-bis-MPT- Tungsto-bispterin

W-bis-MGD - Tungsto-bispterin guanine dinucleotide

tag-His_6 - Six histidine residues

TMD - Transmembrane domains

UV - Ultraviolet

1. Introduction

1.1. Molibdoenzymes and Tungstoenzymes

1.1.1. The role of molybdenum and tungsten in nature

Molybdenum (Mo) and tungsten (W) are chemically analogous elements that are found in the environment and the only second and third row transition metals that have a known biological function. They are found as extremely soluble oxoanions with almost identical co-ordination chemistry.[1, 2] Both elements Mo and W are associated with several redox active enzymes under physiological conditions. The physiological roles of these elements are fundamental and include the catalysis of some main reactions in the metabolism of carbon, nitrogen and sulfur by microorganisms, plants and animals, besides a small number of species that do not require molybdenum use tungsten for the same purpose.[2-4]

Although tungsten not a universal bioelement, is broadly distributed in biology and is vital for some species. Nevertheless, in some cases tungsten seems be used because enviornmental restrictions. For the remaining species, tungsten is biochemical indifferent because they haven't developed a functional use of the element, nevertheless, upon its inadvertent intake, their physiology might be affected.[5]

Intake of Mo and W is achieved in the form of the corresponding oxoanions. This process should be regulated by the uptake system to select molybdate in the presence of competing anions, preventing toxycity symptoms or unavailability of Mo (lethal for the organism).[5-7]

1.1.2. The molybdenum cofactor

Once uptake into the cell, molybdate requires to be coordinated by a unique scaffold to become biologically active. This molecule is a pyranopterin named molybdopterin (MPT) coordinated with Mo, formed the Mo cofactor (Moco). Moco biosynthesis can be distributed into four steps: conversion of GTP in cyclic pyranopterin monophosphate (cPMP, previously indentified as percursor Z); transformation of cPMP into MPT; metal incoporation; and finally in some organisms maturation to an active cofactor (figure 1.1). [7, 8] In eukaryots the pyranopterin is found in the simplest monophosphate form MPT, as already mention. However, in prokaryotes it is conjugated to nucleosides, usually cytosine (MCD, molybdopterin cytosine dinucleotide) or guanosine (MGD, molybdopterin guanosine dinucleotide), so there is an adicional step that binds Moco to different nucleotides (figure 1.2).[9] The conversion of GTP to cPMP is catalyzed by two proteins; one of them (Cnx2 in plant, MOCS1A in humans and MoaA in *E.coli*) is a member of the superfamily of S- adenosylmethione (SAM)-dependent radical enzyme. Members of this protein family catalyze the formation of protein and/or substrate radicals by reductive cleavage of SAM involving the [4Fe-4S] cluster to generate the 5'-

deoxyadenosyl radical which subsquently initiates the transformation of 5'-GTP bound to the C-terminal [4Fe-4S] cluster. [10]

Formerly, the sulfur is transferred to percursor Z in order to generate MPT. This reaction is catalyzed by the enzyme MPTsynthase consisting of two small (plant Cnx7, human MOCS2B) and two large (plant Cnx6, human MOCS2A) subunits capable of cPMP conversion to MPT (the sulfur is bound as thiocarboxylate to the C-terminal of the small subunit); after transfering the two sulfurs to percursor Z, the protein has to be re-sulfurated by the MPT synhtase sulfurase in order to reactivate the enzyme for the next reaction cycle of percursor Z conversion. This resulfuration is catalyzed by MPT synthase sulfurase (Cnx5 in plants and MOCS3 in humans) and at this stage the sulfur transfer reaction in higher organisms seems to involve different protein components as the eukaryotic genes can not complement their bacterial counterparts. [7, 11, 12]

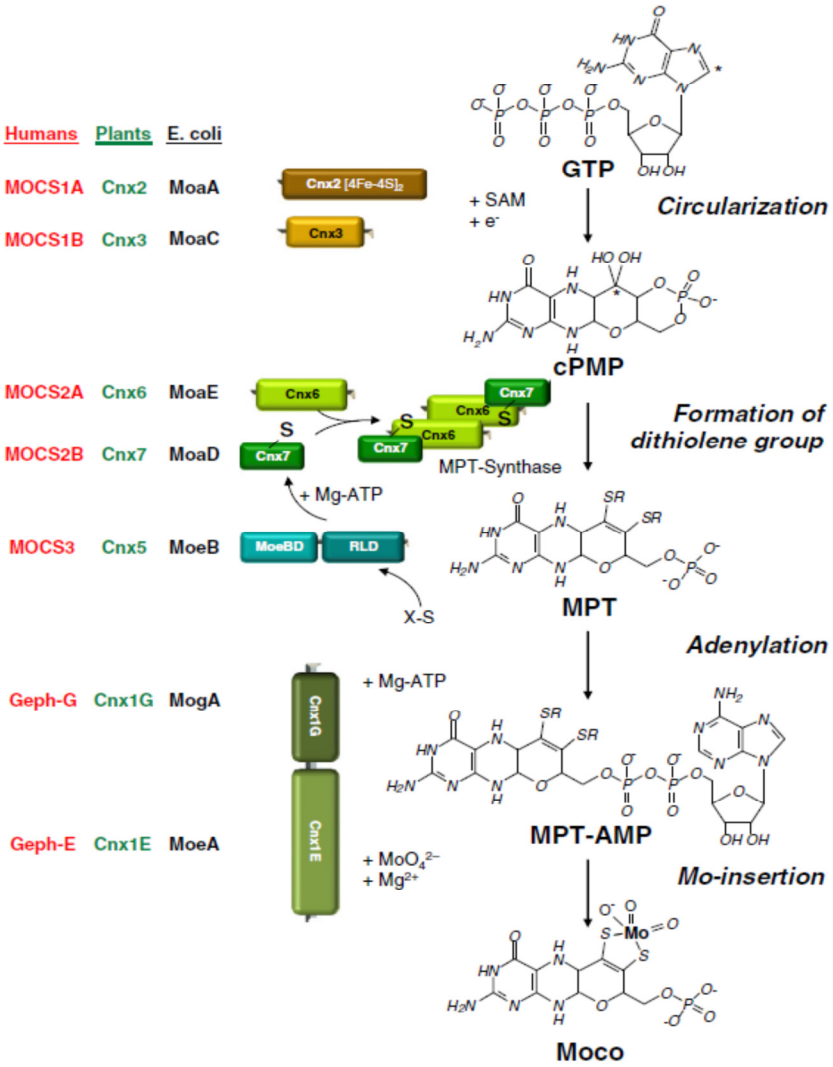


Figure 1.1 - Biosynthesis of eukaryotic molybdenum cofactor. The names for the proteins from plants (green), humans (red) and E. *coli* (black) catalyzing the respective steps are given. [7]

The insertion of Mo in MPT is not a spontaneous process; is catalyzed by a Mo-insertase. In bacteria, Mo insertion is carried out by two distinctly expressed proteins (MogA and MoeA) although higher organisms have fused these two proteins to a single two domain protein (Geph in humans and Cnx1 in plants). Moco-synthetic proteins bind MPT with high affinity and participate in the Mo insertion reaction. [7, 11]

In bacteria, additional modifications by covalent addition of GTP or CTP to MPT are necessary to produce both MGD and MCD. So, MGD is formed by the MobA protein which specifically binds GTP, while MCD is formed by MocA protein which acts specifically on CTP. In both cases, it has been shown that addition of the dinucleotide to the cofactor occurs after the insertion of Mo into MTP. [9, 13]

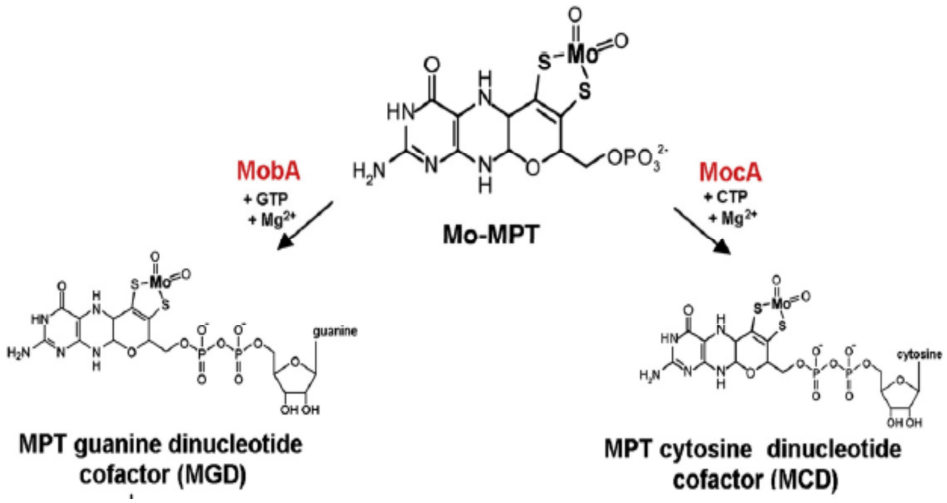


Figure 1.2 – Molybdopterin dinucleotide biosynthesis[13]

1.1.2.1. Importance of molybdenum containing enzymes

The Mo enzymes SO (sulfite oxidase) and DMSO (dimethylsulfoxide reductase) are important to the global sulfur cycle. The reduction of DMSO produces dimethyl sulfide which is photo-oxidized to methanesulfonic acid. This compound is eventually metabolized to sulfite by soil bacteria and then oxidized to sulfate by the sulfite oxidase. [14]

So, a deficiency of Mo enzymes or Moco can be dramatic, especially in humans. Moco insufficiency is a hereditary recessive disorder that affects neonates soon after birth. The major cause of the disease resides in the loss of sulfide oxidase, which effectively removes sulfite from human body. A damage in sulfide oxidase activity results in the accumulation of sulfite, ultimately causing severe neurological damage, disordered autonomic function, dysmorphic facial features, progressive cerebral palsy, seizures and death. [15, 16]

As it was already mentioned Moco catalyzes important redox reactions in the global carbon, sulfur and nitrogen cycles. So, on the basis of cofactor composition and catalytic function there are two different groups[11]:

I. Mo-dependent nitrogenase that contains an Fe-Moco in the active site;
II. All other molydoenzymes that bind Moco

Nitrogenase is essential for biological nitrogen fixation which is an essential step in the nitrogen cycle in the biosphere. It reduces atmospheric pressure with high-energy input in the form of ATP. However, four different types of nitrogenases have been identified, each of which has a particular combination of metals in the active site. The most abundant is the Fe-Mo-dependent nitrogenase which contains $MoFe_3S_3$ and Fe_4S_3.[8]

The second group of proteins which use Moco as cofactor involve SO, xanthine oxidase (XO) and DMSO families and the novel Moco-binding proteins. Each family includes a variety of subfamilies based on sequence similarity and substrate preferences (figure 1.3).[8]

In *E. coli,* the XO family contains the MCD. The SO family is characterized by a di-oxoMoco with an additional protein ligand, which usually is a cysteine. The DMSO family contains two MGDs connect to one Mo atom with additional ligands being an O/S, and a sixth ligand X which can be a serine, a cysteine, a selenocysteine, an aspartate or a hydroxide and/or water molecule.[13]

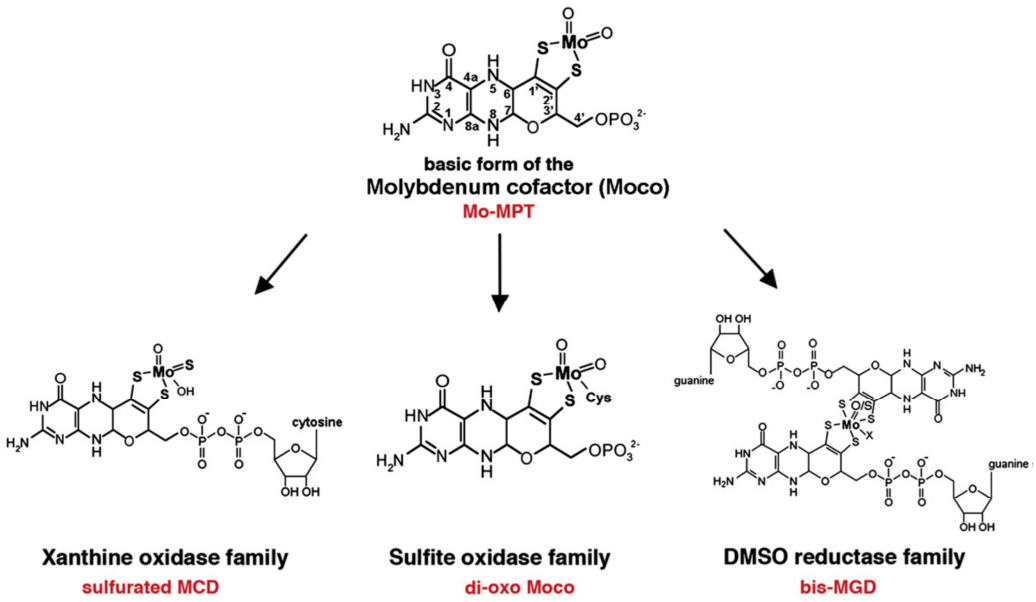


Figure 1.3 – Different structures of the molybdenum cofactor in *E. coli*. The basic form of the Moco is MPT with a unique dithiolene group coordinating the Mo atom. Moco exists in different variants and is divided into three molybdenum-containing enzyme families agreeing to the coordination at the Mo atom as it was said before.[14]

1.1.3. The tungsten cofactor

In contrast of Moco synthesis, the Wco synthesis has not been studied so far. However, it is expected that the basic mechanisms of Wco and Moco synthesis are similar since almost all genes that have an assigned function in the Moco biosynthetic pathway are also present in the genome of organisms that use tungsten in their metabolism. Main differences should occur in the metal insertion step as here a specific discrimination between Mo and W is crucial, as we will see later. [17, 18]

Tungsten is sequestered and transported into the cytoplasm as tungstate. It may then enter the biosynthetic machinery to be bound to dithiolene-functionality of a tricyclic pterin moiety to form tungstopterin (W-MPT) and finally end up in tungstoenzymes in the form of a complex metal-organic cofactor: tungsto-bispterin (W-bis-MPT) or tungsto-bispterin guanine dinucleotide (W-bis-MGD) (figure 1.4). [5]

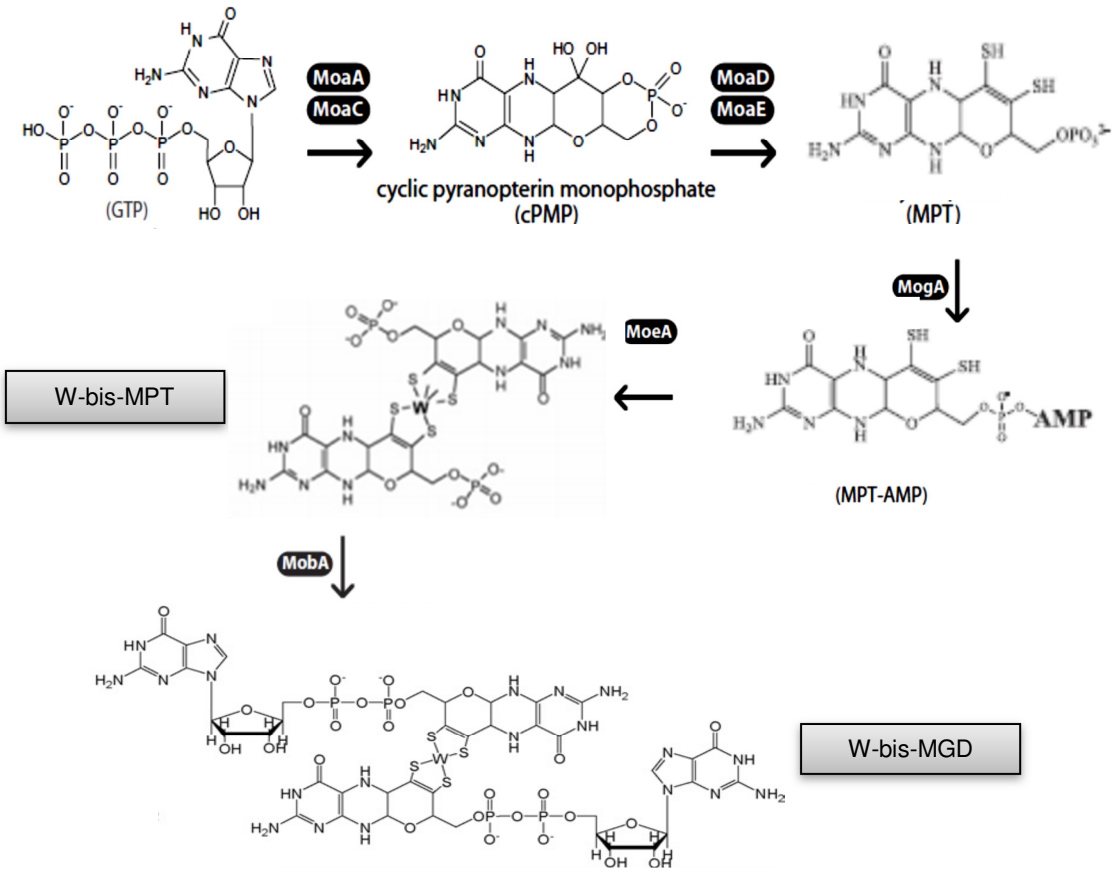


Figure 1.4 – Wco biosynthesis in E.*coli*. The conversion of GTP in MTP occurs like the Moco biosynthesis. Then, MPT is adenylylated by the action of the hexameric MoaB (in P.*furiosus*), a function which is identical to the trimeric MogA protein found in E. *coli*. Next, the metal is inserted by MoeA1 or MoeA2 or both. Finally, guanylylation has to be catalyzed by the MobA protein (adaptation).[17, 18, 19]

As it was said, there are main differences in the metal insertion step, for example, the tungstate transporter from P. *furiosus* selectively binds tungstate with a 1000-folder higher affinity over molybdate. MogA and MoeA are known to be important for *E.coli* Mo insertion, however, in *P.furiosus* and other archea only the MogA-homologous, MoaB proteins, are found. In addition, two different MoeA orthologs are present in these genomes. [17]

P.furiosus is strictly dependent on tungstate, so it is a model organism to study Wco biosynthesis. [17] In *E.coli* the biosynthesis of Wco begins as Moco, with the conversion called precursor Z or cPMP, catalyzed by two proteins MoaA and MoaC. Consequently, the MPT is synthesized from cPMP by MTP synthase which involves MoaD and MoaE proteins. The next step includes the insertion of the metal atom to the dithiolene sulfurs of one or two MPT moieties.[5, 19] Though, with the tungsten there are always two pterin moieties. The protein MogA plays a role in this step, and recently MoaB was found to be involved in this stage of the cofactor synthesis; these proteins catalyze the activation of MPT by adenylation with Mg-ATP. The trimeric MogA proteins are commonly found in bacteria and eukaryotes whereas the hexameric MoaB proteins are mostly found in archea.[17, 20]

MoeA is believed to bind the adenylylated MPT (MPT-AMP) and, in the presence of tungstate, the MPT-AMP complex is hydrolyzed, tungsten is incorporated through binding to the dithiolene sulfur resulting in tungstobispterin (W-bis-MPT) and AMP is released. This suggested role for MoeA in *E.coli* is based on the activity of its plant homologue: Cnx1E, which catalyzes the hydrolysis of MPT-AMP in the presence of molybdate.[12] When molybdate was replaced with tungstate the hydrolysis catalyzed by Cnx1E was considerably less efficient and this effect suggests that Cnx1E and homologues possibly will play a role in selectively incorporating either tungsten or molybdenum in MPT.[5, 12] Several bacterial genomes and all archea genomes sequenced until now, contain two different *moeA* genes which share nearly 40% sequence identity.[12] The presence of these two MoeA's can then explain how organisms are able to regulate and express two enzymes, one with Moco and the other with Wco in the active center.[17]

Another possible pathway is with the MobA protein, resulting in a tungsten center coordinated by four dithiolene sulfur ligands, the tungstobispterin guanine dinucleotide form (W-bis-MGD).[21]

1.1.3.1. Importance of tungsten containing enzymes

The AORs (Aldehyde oxirreductases) family can catalyze aldehyde oxidation to carboxylic acids and in some cases carboxylic acid reduction to aldehydes. The latter reaction is of particular interest for potential biotechnological applications of these enzymes, since aldehydes are highly valued compounds for the fine-chemical industry.[18]

Prokaryotic FDHs (Formate dehydrogenases), with W in the active site, have the capability to catalyze the reverse reaction, i.e. the reduction of CO_2 to formate. Consequently, tungsten-containing FDHs have a great potential as catalysts capable to fix CO_2 from the atmosphere to produce reduced carbon compounds that could be used as fuels. FDHs have the advantage over chemical catalysts as they work as homogeneous catalysts generating formate as sole product which is being known as an alternative energy source through fuel cells based on formic acid. Understanding the molecular features that control the substrate specificity and direction of the reaction (oxidation or reduction) in FDHs, would permit to design biocatalysts and improve procedures related to bioremediation and generation of alternative energy sources.[18, 22, 23]

In the following years, tungsten-containing enzymes were purified from bacteria and homologous genes are even found in the genomes of some organisms, proposing that tungsten enzymes are present in a much wider range of microorganisms. There have been no reports so far on any tungsten enzymes in eukaryotes.[5, 24] The tungsten-containing enzymes can be divided into two families:

I. Aldehyde oxirreductases (AORs) that contain a non-modified tungsto-bispterin cofactor;
II. Formate dehydrogenases (FDHs) which have a guanine monophosphate attached to each pterin moiety.[5]

The enzymes in the AOR family catalyze the oxidation of aldehydes to carboxylic acids, and they use ferredoxin as redox partner protein. They are generally oxygen sensitive and typically have broad substrate specificities with partial correspondence between enzymes from the same species.[5, 14] The enzymes of FDHs family catalyze the oxidation of formate to carbon dioxide. The enzymes exhibit a great variety of redox partners and in most cases the physiological redox partner has not been identified.[5, 14]

1.1.4. The uptake of molybdate and tungstate by ModA and TupA

ATP-binding cassette (ABC) transporter proteins bring different substrates through cell membranes. Bacterial ABC importers are clinically important ABC transporters implicated in various diseases or origin multidrug resistance of cancer cells. In contrast these transport systems are essential for the uptake of nutrients, including rare elements like molybdenum and tungsten .[25]

In prokaryotes, Mo and W are transported as soluble molybdate and tungstate oxoanions, respectively, by the high affinity Mod/Tup/WtpABC transport systems. The *mod/tup*ABC genes encode the proteins Mod/TupA, Mod/TupB and Mod/TupC, which together assemble into the Mod/TupABC transporter. The motif (cassette) of amino acids establishing

the ATP binding site on the protein Mod/TupC is positioned on the cytoplasmic site of native Mod/TupABC complex; Mod/TupB is in the transmembrane protein and finally, Mod/TupA is the molybdate-binding protein on the periplasmic site.[26, 27]

The WtpA protein integrates an ABC transporter specific for the tungstate uptake in presence of molybdate. The gene *wtpabc* which encodes the three different components as mentioned for TupABC and ModABC is present in several archea and some bacteria; the previous characterization of this transporter clarified the transport mechanism of these anions in organisms that lack the known uptake systems related with the ModA and also TupA proteins. Basically, the three transporters families differ to each other in protein sequence but also in activity for their periplasmic binding proteins and in the coordination chemistry of the anions.[28-30]

In this master thesis only the subunit A of the transport system ModABC and TupABC were studied and analyzed.

The periplasmic protein ModA binds specifically molybdate and also tungstate but it is not specific for other anions. The reason behind this is because the specificity for these anions and ModA are typically determined by the size of the binding pocket.[28] ModA proteins cannot differentiate between molybdate and tungstate in presence of both. On the other hand, the periplasmic tungsten protein TupA showed that binds tungstate with high affinity, however it can bind also molybdate but not with the same specificity. This periplasmic protein is able to discriminate between both anions.[28, 30] Molybdate or tungstate binds to the component A of the transporter, making this protein the first selection gate from which cells can should differentiate among Mo or W.[18]

The component B of the transport system is composed by two transmembrane domains (TMDs) that form the translocation pathway which are associated with the component C that include two nucleotide-binding domains (NBDs) that bind and hydrolyze ATP.[31, 32] The importers' resting state is inward-facing, where the NBD dimer interface is held open by the TMDs and facing outward. When the closed, substrate-loaded binding protein attaches near the periplasmic side of the transmembrane domains, the ATP binds and the NBD dimer closes. The resting state of the transporter changes into an outward-facing conformation in which the TMDs have reoriented themselves. So then, they are capable to accept the substrate from the binding protein. After the ATP has been hydrolyzed, the NBD dimer opens up and the substrate is released into the cytoplasm. The transporter then reverts back to its resting state upon the release of phosphate and ADP.[33, 34] Essentially, molybdenum as well tungsten goes into the cell as via ATP binding cassette and it is subsequently incorporated by complex biosynthetic machineries into metal cofactors.[16]

However, other proteins encoded by different operons have functions in intracellular molybdate and tungstate binding and regulation of gene expression. These *mod*ABC operon or gene cluster *tup*ABC are organized encoding additional proteins such as ModE, a repressor

DNA-binding protein in the case of ModABC operon.[27, 30, 35] The ModE transcription factor is encoded by the *mod*EF operon immediately upstream of *mod*ABC.[18] *E.coli* ModE controls gene expression of the *mod* by distinguishing molybdate concentrations. The expression of operons in molybdate uptake or molybdopterin (MPT) synthesis is regulated depending on the molybdate status. In the presence of high concentrations of molybdate the *mod* operon is repressed while the *moa* operon is induced in order to produce sufficient MPT scaffolds to bind molybdate, meaning that when ModE binds Mo undergoes conformational rearrangement and dimerization.[18, 36-39]

1.2. Sulfate reducing bacteria: *Desulfovibrio alaskensis* G20

The sulfate reducing bacteria (SBR) are anaerobic microorganisms that are general in anoxic environments where they use sulfate as a terminal electron acceptor for the degradation of organic compounds, causing the production of sulfide which is incorporated in amino acids and enzymes with sulfur.[40] These anoxic bacteria have high metabolic versatility, because they can use a wide range of other substrates and also play an important role in the cycle of different elements like sulfur and carbon.[41]

These kinds of microorganisms are also important in the bioremediation of heavy metals and in the elimination of sulfur compounds from the combustion waste water and gas. On the other hand, they are still involved in bio corrosion of ferrous metals in anaerobic environments, initiating several problems for some industries due to excessive production of sulfides, which are extremely toxic and corrosive.[40, 41]

In SBR, a tight regulation of intracellular concentrations of Mo and W is especially important because these metals suppress the SBR growth. The inhibition of SBR growth is frequently used for prevention and control of the microbial bio corrosion process. [42]

The analysis of the *D. alaskensis* G20 genome exposes two *moeA*-like genes sharing 32.5% sequence identity (*Dde_0230* and *Dde_3228*, respectively), proposing that this hypothesis may also be suggested for this organism. *D.alaskensis* G20 contains a Mo transport system encoded by the *mod* gene cluster and similarly, the proteins putatively involved in tungstate transport are encoded by the *tup* genes.[43]

One of the goals of this thesis is to structurally characterize the two transporters TupA and ModA and determine the interaction with the respective oxoanions. For this reason, a short introduction to the techniques used is going to be provided.

1.3. X-rays Crystallography

Several techniques are currently used to determine the structure of a protein, including X-ray crystallography, NMR spectroscopy and electron microscopy.[44]

X-ray crystallography is presently the main technique for the determination of protein structures. Knowledge of the three-dimensional structure of proteins can be used in various applications, namely, in biotechnology, biomedicine and drug design as well as a validation tool for protein modifications, ligand binding and also structural authenticity.[45] The extension of the technique to systems such as viruses, immune complexes and protein-nucleic acid complexes serves only to widen the appeal of crystallography.[46] Besides, the requirement for pure, homogeneous and stable protein solutions in crystallizations makes X-ray crystallography useful in other fields of protein research as well.[45]

In this master thesis we will be discuss the different steps from the crystallization conditions until crystals and finally the validation of the structure.

1.3.1. X-rays

In 1895, a german physicist, Wilhelm Rontgen discovered the X-rays being awarded the very first Nobel Prize in Physics in 1901. [47] X-rays have the proper wavelength (in the Angstrom range ~10^{-10}m) to be scattered by the electron cloud of an atom of comparable size. The X-rays radiation, in electromagnetic spectrum, is between UV and ɣ-rays (figure 1.5).[48]

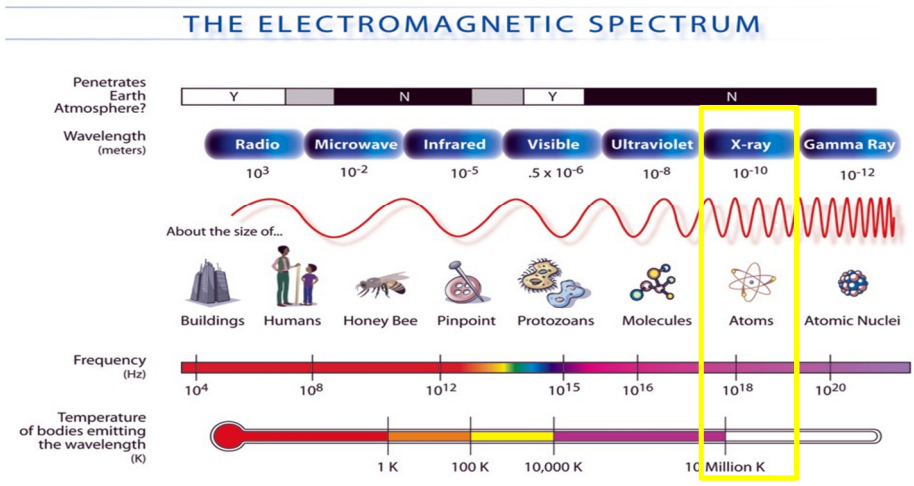


Figure 1.5- The electromagnetic spectrum. X-ray band highlighted in yellow over toward the right. The X-rays have shorter wavelengths, higher frequencies and higher energy than most other types of electromagnetic radiation and don't penetrate Earth's atmosphere. Their wavelengths are around the same scale as atomic sizes (adaptation)[49]

Fundamental electromagnetic theory states that when a charged particle is accelerated through a magnetic field, electromagnetic radiation is emitted. The introduction of synchrotron radiation facilities constitutes a revolution in X-ray macromolecular crystallography since a monochromatic and highly intense X-ray beam can be produced and used for diffraction experiments. [50] The X-rays can be generated from accelerating electrons in a synchrotron storage ring. A single X-ray wavelength is usually selected by absorption of unwelcome wavelengths in a procedure known as monochromation. The X-rays must be focused into a beam and then collimated to certify that the beam is parallel (figure 1.6).[46]

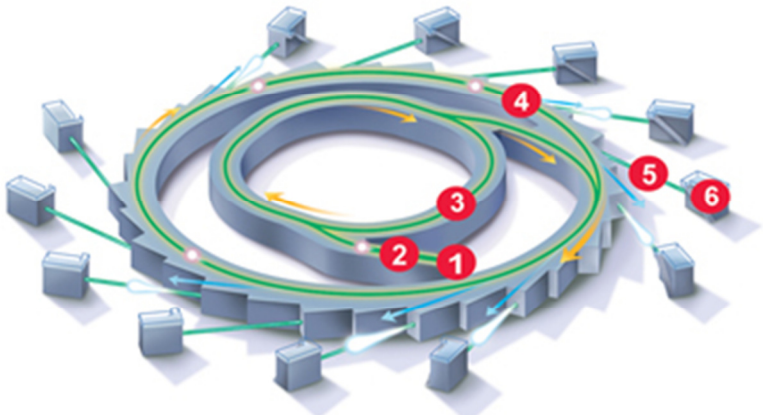


Figure 1.6- How is synchrotron light created. 1 – Electron gun, 2 – Linac, 3 - Booster ring, 4 – Storage ring, 5 – Beamline, 6 – End station. Electrons are generated in the center (1) and accelerated of 99.9997% of the speed of light by the linear accelerator (2). The electrons are transferred to the booster ring (3), where they are increased in energy. They are then transferred to the outer storage ring (4) (adaptation). [51]

1.3.2. Protein Crystallization

The determination of the structure of biological macromolecules by X-ray crystallography relies on previous steps as selection of the target molecule, cloning, expression, purification and crystallization; collection of diffraction data until the determination of atomic positions. Though, even when pure soluble protein is available, producing high-quality crystals remains a major block in structure determination. So, it is important to find appropriate crystallization conditions and optimize growth of diffraction-quality crystal as well.[52]

The main aspect of a crystal is its internal structure, three-dimensionally ordered and periodic, where the molecules are packed and interacting through non covalent interactions.[53] A crystal is characterized by a periodically arranged unit called unit cell that is distributed in the three dimensions as a lattice[48, 54]

The unit cell is defined by the unit cell parameters a,b and c, with respective angles α (between b and c), β (between a and c) and ɣ (between a and b) (figure 1.7).[55]

The unit cells parameters may diverge causing seven crystal systems, which represent different organizations inside the crystal (Appendix I). The crystal symmetry operations found in chiral molecules, only include translation and rotation, limitating the number of possible space groups of biomolecules form 230 to 65.[56]

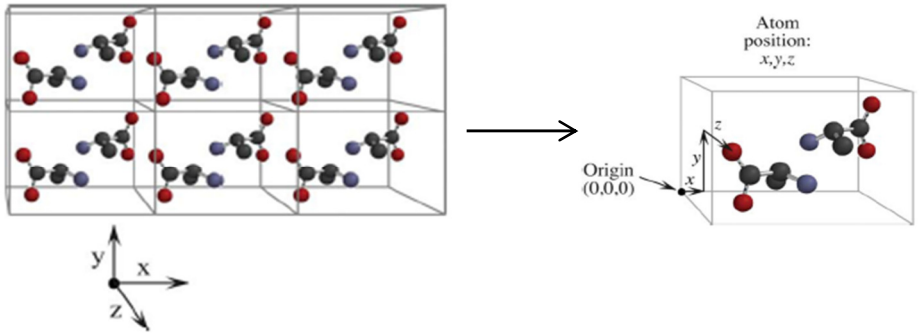


Figure 1.7 – Schematic representation of crystal by various unit cells. There are six unit cells in a crystalline lattice where each cell contains two molecules. The positions of an atom is in the unit cell can be specified by a set of spatial coordinates x, y, z. (adaptation)[56]

Usually protein crystals have very large solvent channels that occupy between 40-60% of the total volume. These solvent channels can be very useful for the study of protein-ligand interactions since they allow the access of the compound of interest in various areas of the protein and especially the active site.[53, 57]

Crystallization is a complex process involving three stages: nucleation, crystal growth and cessation of growth. During nucleation, molecules associate in three dimensions to form a thermodynamically stable aggregate called nucleus. These nucleuses provide surfaces appropriate for crystal growth, which can occur between nucleation and metastable zone. At the stage of nucleation higher protein concentrations are needed when comparing with metastable zone in order to occur a slow precipitation promoting the formation of protein aggregates. This leads to a reduction of protein concentration in solution and the entry into the metastable zone. Crystal growth ends when the solution is sufficiently depleted of protein molecules, deformation-induced strain destabilizes the lattice.[52, 56, 58]

Both crystal nucleation and growth occur in supersaturated solutions where the concentration of the protein exceeds its equilibrium solubility value. The area of metastable zone suitable for crystallization is generally represented on the phase diagram by the solubility curve (figure 1.8). The supersaturation requirements for nucleation and crystal growth are different. This is shown on the phase diagram where the supersaturation region is further divided into regions of higher supersaturation (the labile region), where both growth and nucleation occur, and lower supersaturation (the metastable region), where only growth is supported.[58]

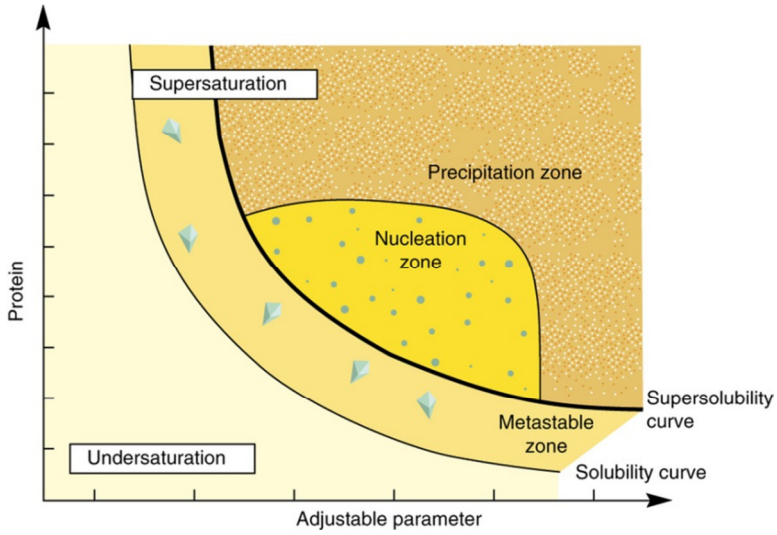


Figure 1.8 – Schematic illustration of a protein crystallization phase diagram. The adjustable parameter can be the precipitant or additive concentration, pH or temperature. The solubility is defined as the concentration of protein in the solute that is in equilibrium with crystals. The supersolubilty curve is defined as the line separating conditions where spontaneous nucleation occurs from conditions where the crystallizations solution remains clear if it is left intact.[59]

Despite the existence of many techniques to obtain crystals the most popular and used in this thesis is vapor diffusion.

In vapor diffusion, crystals are prepared either using the hanging or sitting drop systems. Both consist in a drop containing a mixture of precipitant and protein solutions that are sealed in a chamber with concentrated precipitant. Water vapor diffuses out of the drop until the osmolarity of the drop and the precipitant are equal. The dehydration of the drop causes a slow concentration of both protein and precipitant until equilibrium is completed, ideally in crystal nucleation zone of the phase diagram.[60] Basically in the hanging drop method the crystallization drop is hanging on top of the precipitant while in the sitting drop method it is laying on top of a small bridge (figure 1.9).

Usually, protein crystallization is a challenging process dependent on the type of protein in study. Determination of the crystallization conditions usually requires testing different screenings, namely, precipitant solutions by varying the nature and concentration of salts, pH, additives and temperature. Among all precipitants, PEG (polyethylene glycols) with a low molecular weight (400-20000) is the most used.[53]

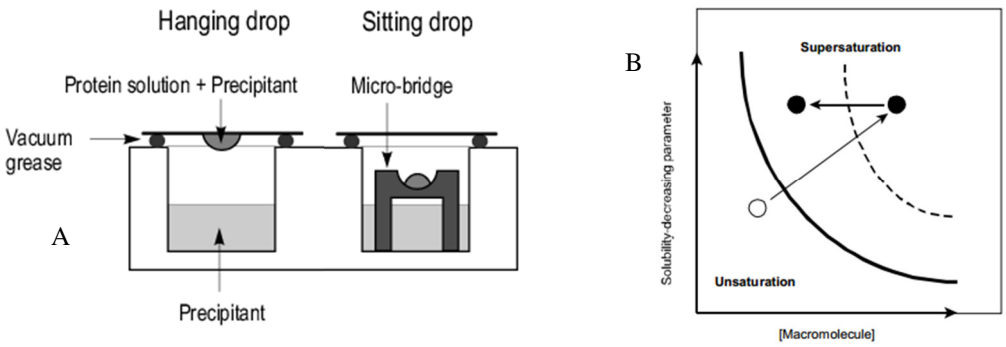


Figure 1.9 – Crystallization techniques by diffusion vapor. A – Schematic representation of hanging and sitting drop. B – Schematic representation of diffusion vapor method. The procedure begins in unsaturation zone and then progresses to nucleation zone because of water evaporation in the drop which incites an increase of protein concentration. (adaptation)[58, 61]

1.3.3. Diffraction, data-collection, phase problem and protein structure

After the crystals growth, the next step is the diffraction experiment; however, it is necessary to harvest the crystals in the presence of cryo-protective agents (glycerol or paratone) before the diffraction.

In a diffraction experiment the crystal is subjected to an X-ray beam. Through this radiation interaction with the electrons of the atoms that constitute the crystal, it originates a diffraction pattern with reflections which provide the internal organization of the crystalline lattice (figure 1.10).[53, 62]

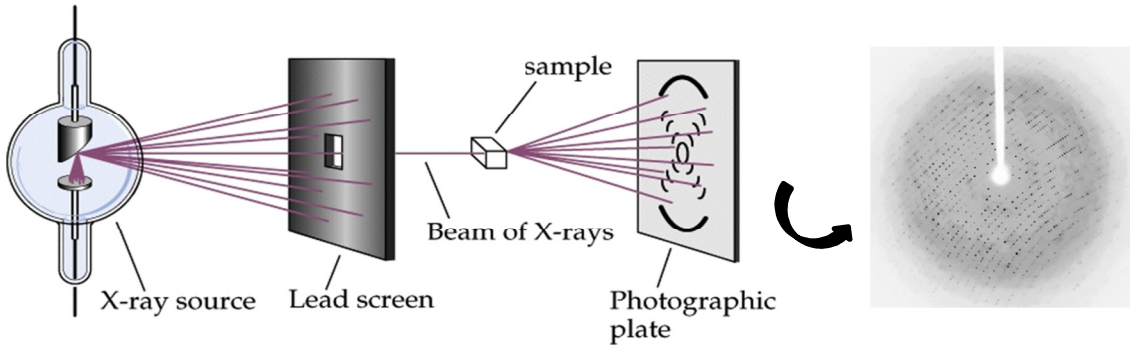


Figure 1.10 – Representation of X-rays diffraction. X-ray beams are shot through a crystal of the atom. The crystal causes the beam of X-ray to diffract in a predictable pattern based on their crystal lattice structure. The result is a diffraction pattern generally seen as the picture on the right. (adaptation)[63-65]

The use of low temperatures during the experiment aims at reducing the diffraction radiation damage, enabling the collection of a complete data set from a single crystal using powerful sources. The cryo-protective agents minimize the scattering of X-rays by water molecules, since they prevent the formation of ice crystals on the surface and inside of the crystal protein.[48, 62]

The diffraction experiment can be explain by Bragg's Law, which predicts that the strongest reflections off the crystal will occur for wavelengths such that the path difference between rays reflecting from consecutive layers in the lattice is equal to an integral number of wavelengths.[66] For some particular crystal orientations, constructive interference between the rays going out of the crystal will be observed and will give the reflections that constitute a diffraction pattern in which the contribution of all atoms is present (figure 1.11).[53, 56]

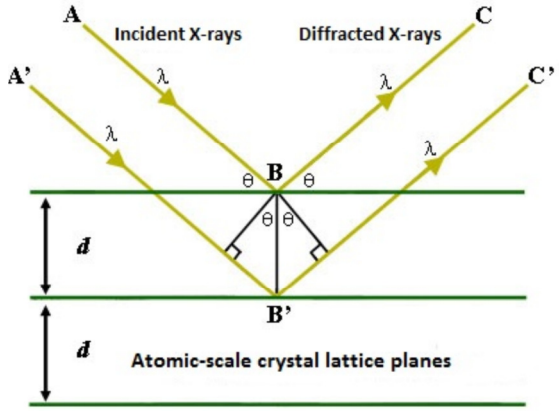


Figure 1.11 – Representation of Bragg's Law. A diagram of a light beam impinging on a crystal lattice. If the light meets the criterion nλ = 2d sinΘ, Bragg's law predicts that the waves reflecting off each layer of lattice interfere constructively, leading to a strong signal.[66]

As each reflection is constituted by a beam of X-rays, it can be defined as a periodic function, characterized by: amplitude, phase and wavelength. This characteristic allows the reflections to be defined with a Fourier transform using the structure factors, F_{hkl}.

In order to reconstruct the electron density of the molecule, two parameters need to be provided for each reflection: the F_{hkl} which is directly obtained through the experiment and is proportional to the square root of the measured intensities making possible to determine the amplitudes $|F_{hkl}|$; and the phase angle of each reflection α_{hkl} which is not directly observable making impossible to achieve the electron density map. This problem is commonly known as the famous Phase Problem in crystallography.[53, 56, 57]

Different methodologies are used to solve the phase problem, such as molecular replacement (MR), multiple isomorphous replacement (MIR) and multiple/ single wavelength anomalous dispersion (MAD/SAD).

The MR method is the most rapid and most recurrently used when a very closely related protein structure is accessible. The method comprises the crystallographic calculation in reverse: structure factors from the known coordinate file, it subsequent borrowing of the phases from the known model structure and their application to the new data set are used to calculate the new structure factor. This method has been used to solve the structure of TupA and will be used for the determination of the structure of ModA in the near future. [46, 67, 68]

When MR method cannot be used to solve the phase problem, other methods as multiple isomorphous replacement (which requires a derivative dataset from protein crystal with attached heavy atoms) or multiple/ single wavelength anomalous dispersion (where anomalous scattering of X-rays according to the wavelength is used), have been developed.

The determination of the initial phases allows the determination of the first map of electron density which can provide the position of all atoms in the crystal.

1.3.4. Refinement and validation

The quality of the electron density map may be upgraded by refinement. The geometrical operation that connects the subunits can be used to control the electron density of the subunits in a series of cycles in which the phases are continually improved by a set of mathematical parameters and manual corrections. The phases are checked by paralleling the observed structure factors with the calculated ones, and the difference is expressed as a percentage known as the R factor (equation 1.1).[46, 69]

$$R = \frac{\sum ||\mathbf{F}_{obs}| - |\mathbf{F}_{calc}||}{\sum |\mathbf{F}_{obs}|} \qquad 1.1$$

In this expression, each |**F**obs| is a consequence from measured reflection intensity and the |**F**calc| is the amplitude of the matching structure factor calculated from the current model.[56]

The R-factor combines the error inherent in the experimental data and the deviation of the model from reality. With increasingly better diffraction data, frequently characterized by R_{merge} of 4% or less, the crystallographic R-factor is efficiently a measure of model errors. Well-refined macromolecular structures are predictable to have R < 20%. When R approaches 30% should be considered with a high degree of reservation because at least, some parts of the model may be incorrect.[53, 69]

Another parameter that is considered is the R_{free}. This is calculated analogously to normal R-factor, but for only ~ 1000 randomly selected reflections which have never entered into model refinement, although they might have influenced model definition. R_{free} is an important validation parameter and should set a warning if it exceeds R by more than ~ 7%. Its high value may indicate over-fitting of the experimental data, or may result from a serious model defect.[69]

Refinement and validation is done based on the previous knowledge of protein models and several parameters such as the root-mean-square (RMSD) deviations of all the model's bond lengths and angles and Ramachandran diagram are considered.

At the end, a good model of a protein structure is obtained which can be a valuable tool for functional and other structural studies.

1.4. Small angle X-ray scattering

With recent advances in data analysis algorithms, X-ray detectors and synchrotron sources, small-angle X-ray scattering (SAXS) has become more accessible to the structural biology community.[70] This technique allows the analysis of the flexibility of the biological macromolecules in solution (nearly physiological conditions), which made this a major tool for the low-resolution structural characterization of biological macromolecules in solution. The most important limitation of current methods for reconstructing 3D models from SAXS is imposed by the requirement of solute monodispersity.[71, 72]

Though limited to ~10 Å resolution, SAXS can provide prosperity of structural information on biomolecules in solution and is compatible with a wide range of experimental conditions. SAXS is consequently an attractive alternative when crystallography is not possible. Furthermore, advanced use of SAXS can provide unique insight into bio molecular comportment that can only be detected in solution, such as large conformational changes and transient protein-protein interactions. Unlike crystal diffraction data, solution scattering data are subtle in appearance, highly sensitive to sample quality and experimental errors are easily misinterpreted.[70]

Experimentally, this technique consists of exposing the protein solution to a monochromatic beam of X-rays where the scattering intensity of radiation is detected as a function of scattering angles (figure 1.12). So, this dispersion is dependent on the protein concentrations and contrast Δρ (r). The last one is the difference between the electron density of the solute and the solvent.[73] The dispersion pattern obtained is indicated in the form of a curve. From these curves specific parameters which determine the size and general shape of protein can be obtained.[73, 74] The SAXS is a very useful and an excellent tool to study protein complexes in solution or attend important structural differences upon ligand binding.

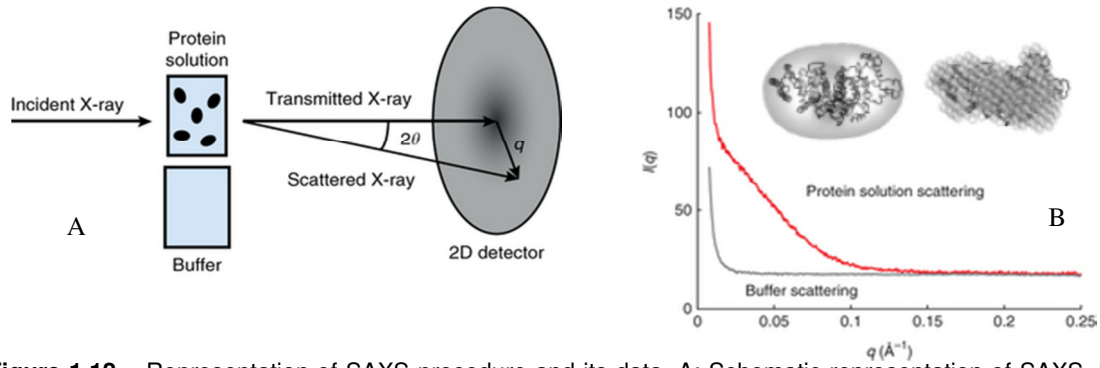


Figure 1.12 – Representation of SAXS procedure and its data. A: Schematic representation of SAXS. B: Data collection by SAXS of BSA protein showing the envelope after the processed data. (adaptation)[70, 74]

1.5. Isothermal Titration Calorimetry

The Isothermal titration calorimetry (ITC) measures directly the energy related with a chemical reaction triggered by the mixing of two components. The ITC instruments operate on the heat compensation principle, the instrumental response (measured signal) is the amount of power (micro calories per second) necessary to maintain constant the temperature difference between the reaction and reference cells.[75]

So, basically the ITC measures the heat absorbed or produced when molecules interact. [76] A typical ITC experiment is carried out by the addition of one of the reactants into the reaction cell containing the other reactant (figure 1.13 A). The chemical reaction generated by each injection either releases or absorbs a certain amount of heat proportional to the amount of ligand that binds to the protein in a particular injection and the characteristic binding enthalpy for the reaction. The heat after each injection is consequently obtained by calculating the area under each peak. Because the amount of uncomplexed protein available gradually decreases after each successive injection, the magnitude of the peaks becomes progressively smaller until complete saturation is achieved.[75]

Experimentally, the solution with the target biomolecule is placed in the sample cell, and a ligand solution in a corresponding buffer is placed in the syringe. When the ligand solution is injected into the cell, the ITC instrument detects heat that is released or absorbed as a result of the interaction (figure 1.13 B). Injections are performed constantly, and result in peaks that become smaller as the biomolecule becomes saturated. After titration is completed, the individual peaks are integrated by the instrument software (figure 1.13 C). An appropriate binding model is chosen and the isotherm is fitted to yield the binding enthalpy ΔH, the K_D, and the stoichiometry, n. From these data, Gibb's free energy, ΔG and entropy, ΔS are calculated.[76, 77]

For the reaction to be measured directly by ITC, the product of the protein concentration in the calorimeter cell and the binding constant, a parameter known as c, has to be lower than 1000. Accordingly with this restriction, the upper limit for the binding constant is between 10^8–10^9 M^{-1}. Further than a certain value, the titrations lose their curvature, becoming indistinguishable and it is not possible to determine the binding constant.[75, 78]

The key for this problem has been obtained by the design of competition experiments in which the high-affinity ligand is titrated into protein that is prebound to a weaker ligand.[75] There are some characteristics to concern in this type of competition assays; the advantages are: experimental conditions are not altered and, therefore, the binding affinity is determined at the planned experimental conditions; the thermodynamic stability of the macromolecule is not compromised by changing the experimental conditions. On the other hand, the limitations of this method are: the binding affinities of the weak ligand and the strong ligand must differ by a factor

of 10 or more as well as the binding enthalpies of the weak ligand and the strong ligand should be as different as possible, otherwise the measured heat signal will be small.[79, 80]

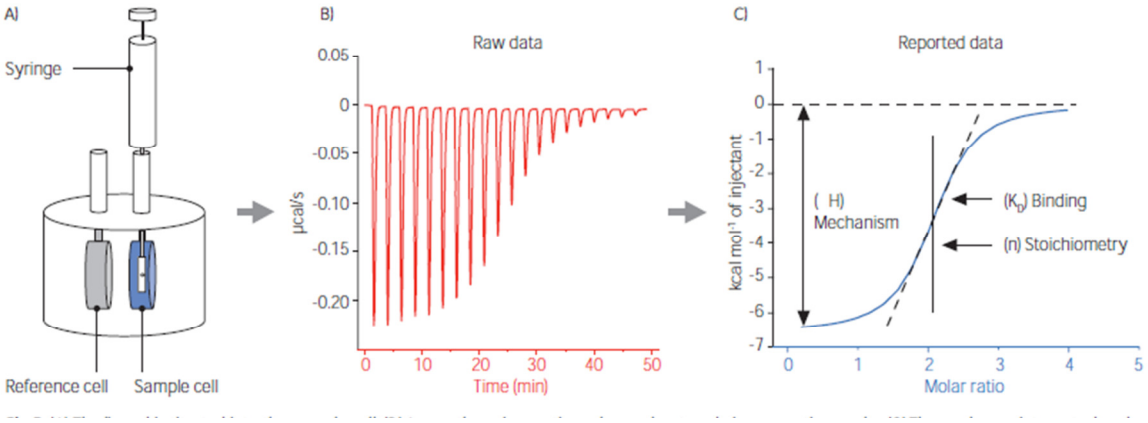


Figure 1.13 – The ITC experiment. A: The ligand titrated into the sample cell. B: An exothermic reaction releases heat and gives negative peaks. C: The peaks are integrated and presented in a reported data.[76]

2. Objectives

The main aim of this master thesis is to understand the specific interaction between the periplasmic component of ABC transporter systems with their ligands. Previous to this work, both ModA and TupA proteins were heterologously produced and purified. Consequently, this work focuses on determining the structural and functional characteristics that are responsible for the selectivity of proteins ModA and TupA from sulfate-reducing bacteria *Desulfovibrio alaskensis* G20 by molybdate and tungstate, respectively.

To achieve this main subject we point to solve the following specific objectives:

- Determine the crystallization conditions, X-ray structure and biochemical properties of ModA and TupA from *Desulfovibrio alaskensis* G20 and compare the metal coordination in holoproteins. In addition, basic biochemical characterization was necessary to determine the quaternary structure and UV-Vis properties;
- Evaluate the relevance of the highly conserved arginine (R118) in the TupA protein putatively involved in the coordination of the oxoanions*;*
- Site directed mutagenesis was performed on the positively charged arginine 118 (pI 10.1) residue. Three residues with different properties were chosen: Lysine (K), Glutamate (E) and Glutamine (Q);
- Study the interaction between each protein and both molybdate and tungstate;
- Analyze the conformational changes induced by metal-protein interaction.
  Both proteins were previously cloned, heterologously produced in *E.coli* and purified.

Although the structure of ModA was determined for several microorganisms, in this work we intend to structurally characterize TupA using protein X-ray crystallography. Furthermore, and to understand the importance of the highly conserved arginine (R118) putatively involved in the coordination of the ligand, site directed mutagenesis is applied. Three different mutants of TupA have been design and produced. To help understanding the selectivity of the proteins towards metals, biochemical characterization using different techniques is necessary. For a qualitative analysis, gel electrophoresis was performed while ITC enabled to quantify the affinity of the protein-ligand interaction. Optimization of the crystallization conditions of TupA was also carried out for the three variants and for the native protein in the presence of ligands (WO_4^{2-} and MoO_4^{2-}).

Despite the cloning, over-expression and purification conditions had already been determined; the biochemical description for ModA using the techniques mentioned above is also required. In this work we also intend to crystallize ModA in the presence and absence of both ligands for structural characterization.

Furthermore, conformational changes upon ligand binding are also of utmost importance and low resolution models of ModA, TupA and the three mutants of TupA in solution can be obtained by SAXS technique.

3. Experimental Procedure

The genes encoding TupA and ModA proteins were previously cloned in pET vectors, the constructs were used to transform *E. coli* cells (BL21 (DE3)) and the proteins were heterologously produced and purified.[30] This work was performed in the frame of the project "Molecular aspects of sulfate reducing bacteria inhibition by molybdate" (Ref. EXPL/BBB-BEP/0274/2012, Fundação para a Ciência e a Tecnologia) with Dr. Maria Gabriela Rivas as principal investigator. Crystallization assays were carried out in the laboratory of Prof. Maria João Romão by Ana Rita Cardoso and Dr. Teresa Santos-Silva. For TupA, a crystallization condition could be found where good diffracting crystals were obtained. The crystal structure of the protein was solved by Ana Rita Cardoso using the molecular replacement method.

This Master thesis work starts with the site directed mutagenesis of TupA.

3.1. Site directed mutagenesis of TupA protein

The *tupA* gene (locus tag Dde 0234) previously cloned in pET-46 Ek-LIC vector (Novagen) was used as template to mutate arginine 118 by PCR using the QuickChange® Site-Directed Mutagenesis Kit (Stratagene). Briefly, two complementary oligonucleotides containing the desired mutations (table 3.1) were designed to each mutation and used to amplify the pET-46-TupA plasmid. Then the PCR product obtained was incubated with DpnI to digest the template (not mutated plasmid) and the product of this reaction was used to transform competent cells.

The amplification reaction was performed on a final volume of 50 µL, using 1µL dNTP mix (concentration was not informed in the kit), 1x reaction buffer, 30 ng of the dsDNA template, 90 ng of each primer (forward and reverse), 2% of DMSO (Merck) and finally 2.5 U of *Pfu* Turbo DNA polymerase. Amplification program was executed on a thermocycler (Bio Rad My Cycler™) and included an initial denaturation step at 95ºC for 30 seconds. Then, 18 cycles composed by 30 seconds at 95ºC (denaturation) and 60 seconds at 68ºC (primer annealing) and 6 minutes at 68ºC (elongation) were performed. The amplified products were analyzed by 1% agarose gel (NZYTech) electrophoresis in 1X TAE Buffer using GeneRuler™ 1kb DNA Ladder (Fermentas) as DNA marker. To each 5 µL of PCR product, 2 µL of 5X sample buffer (Fermentas) were added. Electrophoresis was performed at 100 V for 30 minutes in Mini Run Ge-100 equipment. PCR products were first stained with SYBR Safe™ DNA gel staining (Invitrogen) observed and images captured on GenoSmart gel documentation system (VWR). A volume of 1 µL of the DpnI 10U/ µL was added to each amplification reaction to digest the parental supercoiled dsDNA. The reaction mixtures were spin down in a microcentrifuge (Minispin, Eppendorf) for 60 seconds and immediately incubated at 37ºC for 1 hour.

After the parental DNA digestion, the reactions were used to transform XL1-Blue supercompetent cells (QuickChange® Site-Directed Mutagenesis Kit, Stratagene) and the plasmids were isolated from a single colony (grown in Luria-Bertani media with ampicillin

(NZYTech, 100 µg/mL)) using the NZY-Tech Miniprep kit (NZY-Tech, Lisbon, Portugal). The recombinant plasmids were quantified in a Nanodrop (Thermo Scientific Nanodrop 2000c) and sequenced using an ABI3700 DNA analyzer (Perkin/Elmer/Applied Biosystems, Stabvida, Caparica, Portugal).The sequences were analyzed and aligned using the online tool: BLASTp[81] and ClustalW.[82]

Table 3.1 – Primers used to mutate arginine 118 to lysine (R118K), to glutamic acid (R118E) and to glutamine (R118Q)

Primer	Sequence
f-TupA_R118K (sense)	CCGTATTTGTAAGC**AAG**GGCGACAACTCGG
r- TupA_R118K (antisense)	CCGAGTTGTCGCC**CTT**GCTTACAAATACGG
f- TupA_R118E (sense)	CCGTATTTGTAAGC**GAG**GGCGACAACTCGG
r- TupA_R118E (antisense)	CCGAGTTGTCGCC**CTC**GCTTACAAATACGG
f- TupA_R118Q (sense)	CCGTATTTGTAAGC**CAG**GGCGACAACTCGG
r- TupA_R118Q (antisense)	CCGAGTTGTCGCC**CTG**GCTTACAAATACGG

3.2. Heterologous protein production

The first expression condition tested for the mutants was the one used for the recombinant TupA protein, it means *Escherichia coli* strain BL21 (DE3), transformed with one of the mutated plasmids, pET46-tupAR118K, pET46-tupAR118E or pET46-tupAR118Q, grown in Luria Bertani (LB) media till an OD_{600} ≈ 0.4 AU at 37 ºC, followed induction with 1mM of isopropyl β-D-galactopyranoside (IPTG) (NZYTech) during 3h at 30 ºC. Since these expression conditions yielded insoluble protein, different expression tests were performed for each mutant. Three different expression cells were used, namely, Rosetta2, Origami and Tuner and two different temperatures of induction were tested (19 ºC and 30 ºC). In addition, two different media, LB and autoinduction (Appendix II) at different concentrations of IPTG (0.1, 0.3, 0.5, 0.8 and 1 mM) were tested. The optimal condition for expression of each mutant was then used to produce the proteins. Note that the expressions tests were done in 10 mL of LB media.

TupA_R118K and TupA_R118E mutants were produced in E.*coli* Rosetta2 and Origami cells, respectively. Cells transformed with the plasmids were cultured in Luria-Bertani media supplemented with ampicillin (NZYTech, 100 µg/mL) at 180 rpm and 37 ºC. Cells were induced when OD_{600} reached 0.6 AU with 1 mM IPTG (NZYTech) and grown at 19 ºC overnight. The TupA_R118Q, construct was used to transform *E.coli* Tuner cells cultured in LB media containing ampicillin (NZYTech, 100 µg/mL) at 180 rpm and 37º till OD_{600} = 0.6 AU. The induction was performed by addition of 0.3 mM IPTG (NZYTech) and further incubation at 19 ºC overnight. All the cells were collected by centrifugation at 8000 rpm for 20 min (JA 10, Beckman

Coulter), resuspended in 50 mM sodium phosphate pH 8 (Sigma-Aldrich), 500 mM NaCl (Merck) buffer containing 5mM DNase (NZYTech) and 1 tablet of Protease Inhibitor Cocktail EDTA free (Sigma-Aldrich) using a ratio of 2 g cells/ mL.

The cells suspensions were broken by sonication (UP100H//Tip MS7/ Hielscher Ultrasonics), using a mild protocol where the ultrasounds repeatedly applied (5 times for 3 minutes with 0.5 cycles and 80% amplitude). Each crude extract was centrifuged at 10 000 rpm (F34-6-38, Eppendorf), for 30 min and the soluble fraction was filtered through a 0.45 µm membrane (VWR).

3.3. TupA mutants purification

The pET-46 Ek-LIC expression vector encoded a six-histidine tag at the *N*-terminal sequence enabling the purification by immobilized-metal affinity chromatography (IMAC). A His GraviTrap (GE Healthcare) equilibrated with 50 mM sodium phosphate pH 8 (Sigma-Aldrich), 500 mM NaCl (Merck) was used to inject the crude extract. Before injection the sample was dialyzed against the equilibration buffer. The proteins fractions were eluted with a step gradient of 20 mM, 50 mM, 100 mM and 250 mM imidazole (Sigma-Aldrich). The fractions containing the TupA mutants were identified by 10% SDS-PAGE (Tris-Tricine) and the gels were stained using Coomassie brilliant blue solution. The pure fractions containing the TupA_R118K, TupA_R118E and TupA_R118Q were pooled together and dialyzed against 5 mM Tris-HCl pH 7.6. The samples were concentrated using ultrafiltration devices (VIVASPIN TURBO 15. 10000 MWCO, Sartorius Stedim Biotech) and stored at -20 ºC until further use. All of the steps, including cell collection, soluble extracts preparation and purifications, were performed at 4 ºC.

3.4. Biochemical characterization of TupA and ModA

3.4.1. Extinction Coefficient Determination

The extinction coefficient (Ɛ) for TupA and ModA had already been previously determined, corresponding, respectively, to 29700 $M^{-1}cm^{-1}$ and 22500 $M^{-1}cm^{-1}$ at the 280 nm absorption peak. These values are in good agreement with the values deduced from the amino acid sequence of pure protein (30440 $M^{-1}cm^{-1}$ and 23295 $M^{-1}cm^{-1}$, respectively).[30] Nevertheless, the extinction coefficient of each TupA mutant was experimentally determined.

The concentration of the three mutants (TupA_R118K, TupA_R118E and TupA_R118Q) was determined by Bradford[83] using bovine serum albumin (0, 0.1, 0.15, 0.2, 0.25 and 0.3 mg/mL) as standard. Then, five different dilutions of each TupA mutant were performed and the absorbance at 280 nm was measured. The extinction coefficient was obtained from the slop of the absorbance at 280 nm vs protein concentration curve. The UV-

visible absorptions were measured in an Evolution 201 (Thermo Scientific) split beam spectrophotometer using 1 cm optical path cells. The values obtained were different from those determined using the bioinformatics tool, from JustBio.[84]

3.4.2. Protein Gel Shift Assay

TupA, ModA, TupA_R118K, TupA_R118E and TupA_R118Q protein samples (50 µM) were incubated with 500 µM of MoO_4^{2-} (Sigma-Aldrich) or WO_4^{2-}anions (Sigma-Aldrich) at room temperature for 25 min. Unbound anions were separated from each mutant with a PD Mini Trap G-25 desalting columns (GE Healthcare) following the manufacturer's instructions.[30, 85] The buffer used is the same where the proteins were stored: 5 mM Tris-HCl (pH7.6)

The electrophoresis was carried out at 120 V constant for 1h30min on a native 10% polyacrylamide gel with 1X of loading buffer (Appendix III). The mobility shift assay was analyzed after staining the gel with Coomassie Brilliant Blue solution.

3.4.3. Urea-polyacrylamide gel electrophoresis

The putative conformational changes upon metal ligand binding was analyzed for TupA and ModA by urea gel electrophoresis using the Novex 6% tris(hydroxymethyl) aminomethane (Tris)-borate (TBE)-urea minigels and a XCell SureLock™ Mini-Cell (Invitrogen). 5 µL of each sample mixed with 5 µL of 2x Novex sample buffer (without EDTA) were loaded in the gel. The electrophoresis was carried out for 2h30 min at 180 V and 40 mA. The proteins bands were examined by staining with Coomassie Brilliant Blue.[30, 86]

3.4.4. The Isothermal Titration Calorimetry

The isothermal titration calorimetry experiments were previously executed for TupA and ModA. The conditions described below (already optimized) were the same for the studies of the TupA, ModA and the mutated proteins.

The ITC experiments were performed using a VP-ITC calorimeter (MicroCal, GE Healthcare). The binding protein (10 µM) was equilibrated in reaction buffer at 30 ºC in the cell of the calorimeter, and subsequently, 25 injections of 10 µL of a 100 µM sodium tungstate (Sigma-Aldrich) or molybdate (Sigma-Aldrich) solution were performed and the heat response recorded. Once subtraction of the baseline, the integrated heat responses were fitted to the single binding site model using the ORIGIN software package provided with the calorimeter.[30]

3.5. Crystallization Studies

All the crystals obtained were observed in a microscope (Discovery V.12, Zeiss) and posteriorly stabilized with a solution of harvesting buffer (similar to the crystallization solution but where the concentration of precipitating agent is about 2% superior) and then they are transferred to a cyoprotectant solution which can be paratone (Hampton Research) or the harvesting buffer with 20-30% of glycerol (Sigma-Aldrich). Crystals were fished, frozen in liquid nitrogen and subsequently analyzed by synchrotron radiation.

3.5.1. TupA

TupA crystallized in only one condition of the *in-house* screen[87] containing 0.2 M magnesium chloride (Sigma-Aldrich), 0.1 M HEPES (pH 7.5) (Sigma-Aldrich) and 30% (*w*/*v*) polyethylene glycol 3350 (Sigma-Aldrich). Transparent plate-shaped crystals appeared in 4 days.[30] The crystallization was performed at 20 ºC using the sitting-drop vapor diffusion method, with 1 µL of protein (7.5 mg/mL plus 1 µL of precipitant solution) on 24-well crystallization plate (XRL Molecular Dimensions) with 700 µL of precipitant agent in reservoir.[30]

These crystals were used for data collection at beamline ID23-1 at the European Synchrotron Radiation Facility (ESRF, France) and the crystal diffracted up to 1.43 Å.[30]

3.5.2. ModA

The first crystallization studies were performed at 20ºC using the sitting-drop vapor diffusion method, with 0.5 µL of protein plus 0.5 µL of precipitant solution on 96-well crystallization plates (SWISSCI 'MRC' 2-Drop Crystallization Plates, Douglas Instruments). Different commercial screens were tested, namely, Midas (Molecular Dimensions), JCSG (Molecular Dimensions), an 80 conditions *in-house* screen (based on the screen of Jancarik *et al*)[87] and EWI/EWII (Hampton Research). Needle shaped crystals appeared in one condition of the *in-house* screen using (30% (*w*/*v*) polyethylene glycol 6000 (Sigma-Aldrich), 0.1M sodium cacodylate (pH 6.5) (Sigma), 0.2M ammonium sulfate (Panreac).

In order to obtain good crystals the precipitant concentration was optimized to 28% *(w/v)* and furthermore, an additive screen was tested (Additive Screen 1 from Hampton Research), at 20ºC using the hanging-drop vapor diffusion method. Crystals of ModA (at 10mg/mL) appeared in hanging-drops with 1 µL of protein plus 0.8 µL of precipitant solution and 0.2 µL of 20% (*w*/*v*) Benzamidine hydrochloride (Hampton Research) on a 24-well crystallization plate (XRL Molecular Dimensions) with 700 µL of precipitant solution in reservoir. The colorless crystals appeared within 5-6 days.

Using the synchrotron radiation, all the crystals obtained were tested. The needle shaped crystals diffracted up to 3.5 Å resolution and a complete dataset was collected at beamline ID23-1 at the European Synchrotron Radiation Facility (ESRF, Grenoble, France).The larger crystals obtained in the presence of the additive diffracted up to 2.77 Å resolution and were used to collect a complete dataset at beamline ID I03 at Diamond Light Source (DLS, England).

3.5.3. TupA_R118K, TupA_R118E and TupA_R118Q

The first crystallization test was performed at 20 ºC using the sitting-drop vapor diffusion method, with 0.5 µL of protein plus 0.5 µL of precipitant solution on 96-well crystallization plates (SWISSCI 'MRC' 2-Drop Crystallization Plates, Douglas Instruments). Different commercial screens were used, namely the EWI/EWII (Hampton Research) and the *in-house* screen (based on the screen of Jancarik *et al*)[87].

Both mutants crystallized in the same condition of the *in-house* screen containing 0.1 M magnesium chloride (Sigma-Aldrich), 0.1 M MES (pH 6.5) (Sigma-Aldrich)and 30% (*w*/*v*) polyethylene glycol 8000 (Sigma-Aldrich) on a 24-well crystallization plate (XRL Molecular Dimensions). Transparent plate-shaped crystals appeared after 3 days of crystallization set-ups.

Scale-up and optimization experiments were performed and new crystals appeared in hanging-drops with 1 µL of protein (TupR118K (14 mg/mL), TupR118E, (10 mg/mL) or TupR188Q (11 mg/mL) plus 1 µL of precipitant solution on a 24-well crystallization plate (XRL Molecular Dimensions) with 700 µL of precipitant solution in reservoir. These crystals were also used for data collection at beamline ID I03 at the Diamond Light Source (DLS, England). However the crystals didn't diffract.

3.6. Small Angle X-ray scattering

SAXS was performed for ModA in the absence and presence of ligands (MoO_4^{2-} and WO_4^{2-}).

In order to study the influence of the ligands, the protein solutions were incubated with MoO_4^{2-} and WO_4^{2-}anions (Sigma-Aldrich) (1:10) at room temperature for 25 min. Unbound anions were separated with a PD Mini Trap G-25 desalting columns (GE Healthcare) following the manufacturer's instructions. The buffer used is the same where the proteins were stored: 5 mM Tris-HCl (pH7.6). The protein-ligand solutions were to a final concentration of 21, 23 and 20 mg/mL for ModA, ModA+Mo and ModaA+W, respectively.

For each sample, 5 serial dilutions were used for scattering measurements. The experiment was carried out using synchrotron radiation at the beamline ID BM29 at the

European Synchrotron Radiation Facility (ESRF, Grenoble, France). The data collected was analyzed using the ATSAS[88] tool and PyMOL.[89]

4. Results and Discussion

4.1. Primary sequence analysis of TupA and ModA

The genome of all the *Desulfovibrio* species published to date encodes a tungstate ABC uptake systems belonging to the Tup family of proteins. Regarding the 3D structure of TupA proteins, two structures from *Geobacter sulfurreducens* and *Eubacterium acidaminophilum* were announced in the last years and the structure of Gs is available in the protein data bank (PDB code 3LR1). The TTTS motif is the signature of TupA proteins where Thr9 and Ser11 (*G. sulfurreducens* numbering) are suggested as residues able to bind the oxoanion through hydrogen bonds. In addition, the Thr124 and the positively charged Arg118 (both highly conserved in the primary sequences analyzed, figure 4.1) could bind the oxoanion also by hydrogen bonds. Arg118 is suggested as the structural element that gives the high selectivity of the TupA proteins. The *Da*G20 TupA contains all the conserved residues putatively involved in the oxoanion coordination (figure 4.1).[18, 26, 30]

Regarding to ModA proteins from *Desulfovibrio* species, multiple sequences alignment (figure 4.2.) were analyzed, showing that Ser12 (*E. coli* numbering) is conserved, interacting with the oxygen of MoO_4^{2-} through the amine group. At position 152 isoleucines and valines, are found, both capable of binding molybdate. In addition, the conserved Tyr170 is sometimes replaced by an Ala, probably coordinating molybdate through an H-bond.[18, 26, 90] In fact, the alignment shows that *Ec* ModA from strain K12 is also very similar to the corresponding sulfate reducing bacteria.

In order to study the oxoanion coordination by TupA and the relevance of arginine 118 in the selectivity process, this residue was mutated to lysine (a basic and also positively charged amino acid), glutamic acid (an acidic and negatively charged residue) and glutamine (a polar neutral amino acid that can form hydrogen bonds).

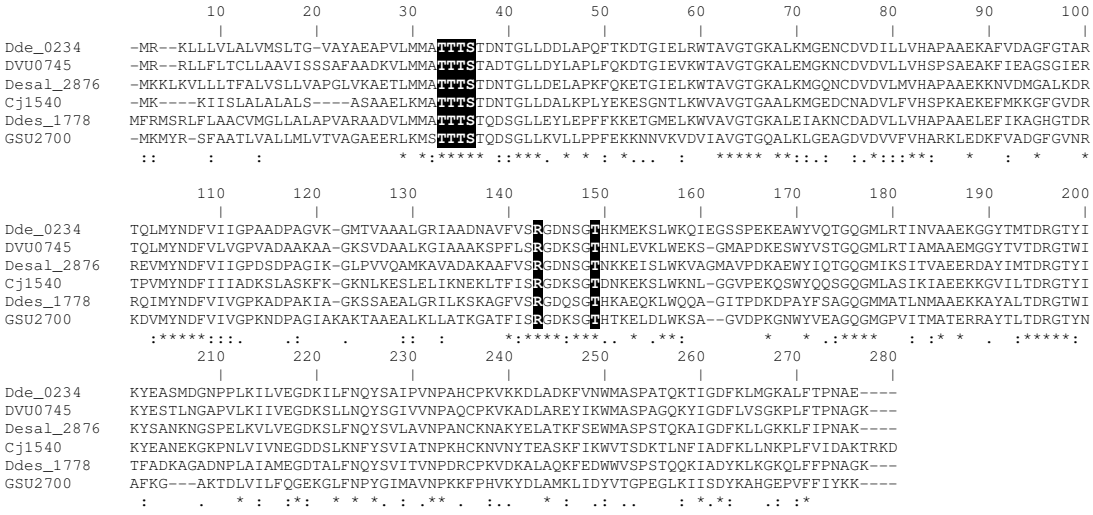


Figure 4.1 - Multiple sequence alignment of TupA proteins. Dde_0234, *D. alaskensis* G20; DVU0745, *D.vulgaris* Hildenborough; Dde_2876, *D.salexigens*; Cj1540, *C. jejuni* strain NCTC 11168; Dde_1778, *D.desulfuricans* ATCC 27774; GSU2700, *G.sulfurreducens*. Residues putatively involved in the coordination of tungstate are highlighted in black. Symbols: (*) identity, (:) strongly similar and (.) weakly similar.[30]

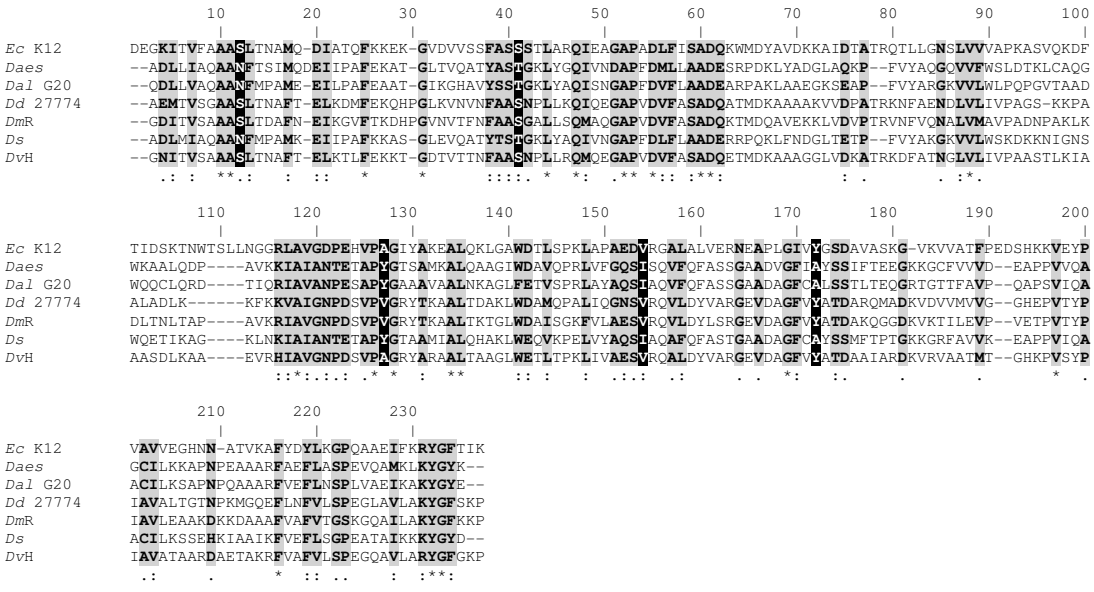


Figure 4.2 - Multiple primary sequence alignment of ModA proteins. *E. coli* K12 (Ec K12, locus tag b0763), *D.aespoeensis* (Daes, locus tag Daes 2623), *D.alaskensis* G20 (Dal G20, locus tag Dde 3518), D.desulfuricans ATCC 27774 (Dd 27774, locus tag Dde 0168), *D.magneticus* RS-1 (DmR, locus tag DMR 17460), *D.salexigens* (Desal, locus tag Dde 3518) and *D.vulgaris Hildenborough* (DvH, locus tag DVU0177). Residues involved in coordination of molybdate are highlighted in black. Symbols: (*) identity, (:) strongly similar, (.) weakly similar.

4.2. Site directed mutagenesis of TupA and optimization of the mutants expression

As mentioned in section 3.1., the *tupA* gene was previously cloned in pET-46 Ek-LIC vector (Appendix IV) and the construct was used as template to mutate Arg118 by PCR. The vector pET46-TupA was used to clone each mutant using two complementary oligonucleotides (primers) containing the desired mutation (R118K, R118E and R118Q).The PCR products obtained were electrophoresed in agarose gels to confirm the amplification (figure 4.3) and the non-mutated DNA was digested by DpnI. The digestion reaction was used to transform competent cells, the mutant plasmids were isolated from a single colony by miniprep and the mutation was confirmed by sequencing the obtained *tupA* gene.

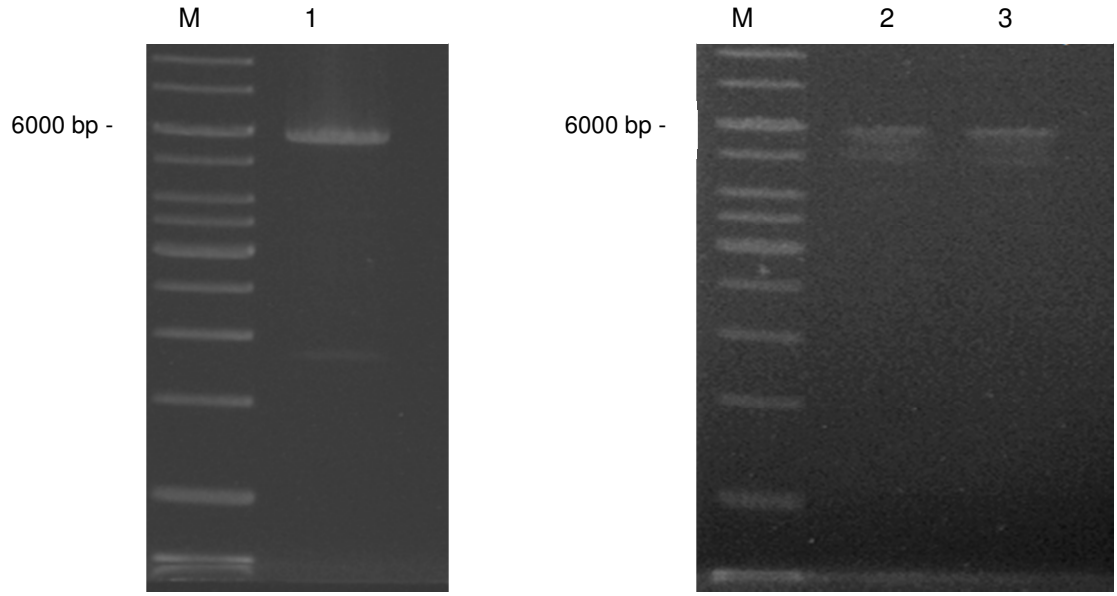


Figure 4.3 – 1% agarose Gel electrophoresis of the PCR product obtained (before DpnI digestion). M: Marker; 1:TupA_R118K; 2: DNA of TupA_R118E and 3: DNA of TupA_R118Q. The size of the fragment (6000 bp) is in line with the expected value, it means, the vector (5200 bp) plus and the *tupA* gene (756 bp).

The first condition tested for the expression of TupA mutants was the one used for the recombinant TupA protein[30], however the three mutants were produced in insoluble forms (Appendix V). Therefore, several expression conditions were tested including four different *E.coli* strains (BL21, Rosetta2, Origami and Tuner), two culture media with different composition (LB and autoinduction) and different induction conditions (at 19ºC and 30ºC using several IPTG concentrations). The expression of the mutants was evaluated by SDS-PAGE.

As shown in figure 4.4 each mutant is produced under different conditions. TupA_R118K mutant can be expressed by Rosetta2 and TupA_R118E by Origami cells, both induced with 1mM of IPTG. In the case of TupA_R118Q, the best expression condition was found using Tuner cells induced with 300 µM IPTG. In all the cases, cells were grown in LB medium and proteins were induced overnight at 19 ºC.

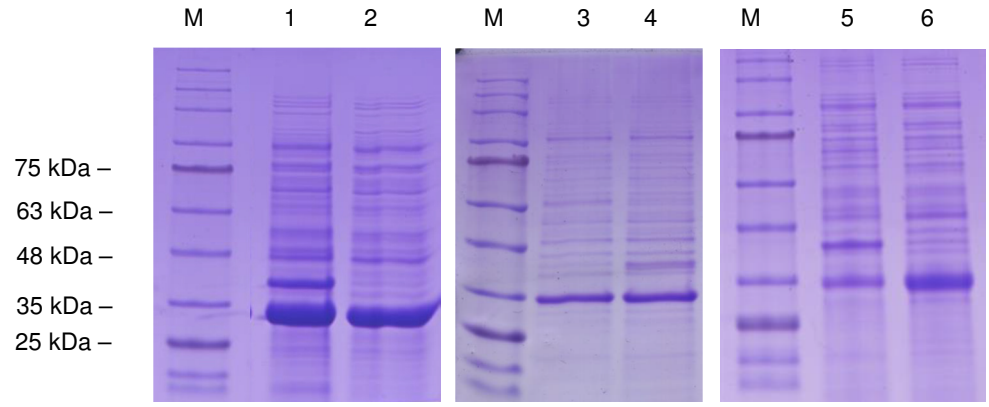


Figure 4.4 - Expression of TupA mutants evaluated by electrophoresis in polyacrylamide gels (12.5%) under denaturing conditions (SDS-PAGE), stained with coomassie brilliant blue: M) molecular weight markers; (1) Soluble protein fraction of cells transformed with the pET46-TupA_R118K vector; (2) Insoluble protein fraction of cells transformed with the pET46- TupA_R118K vector; (3) soluble protein fraction of cells transformed with the pET46- TupA_R118E vector, (4) insoluble protein fraction of cells transformed with the pET46- TupA_R118E vector (5) the insoluble protein fraction of TupA_R118Q vector and (6) soluble protein fraction of cells transformed with the pET46- TupA_R118Q vector.

The migration of the recombinant TupA protein correspond to approximately 35 kDa which is in agreement with the expected size of the protein (29 kDa –TupA- plus 1 kDa -His_6-tag- and 2 kDa -signal peptide-). Although the mutant proteins are produced in both soluble and insoluble fractions (figure 4.4), the amount in the soluble fraction was enough to perform the studies proposed in this work. The production yield obtained was around 12 mg of TupA_R118K, 13 mg of TupA_R118E and 35 mg of TupA_R118Q per liter of cell culture.

4.3. Purification of TupA mutants

The purification of recombinant TupA includes two steps: an anionic exchange and a size exclusion chromatography. TupA elutes from the anionic exchange resin at approximately 200 mM Tris–HCl (pH 7.6), which is in agreement with the isoelectric point calculated for the protein (pI 5.69) by the bioinformatics JustBio tool.[84] This purification protocol was tested for the three mutants but with unsuccessful results (Appendix VI). Instead, an immobilized metal ion affinity chromatography was performed (IMAC) using a Ni-NTA resin.

The three mutated proteins and the recombinant TupA have a His_6-tag, however only the mutants showed affinity for the Ni-NTA resin .Therefore, the three mutants were suspended in 50 mM sodium phosphate (pH 8) plus 500 mM NaCl and a gradient of 50-250 mM imidazole was used in order to differentially elute mutants and contaminants (SDS-PAGE of the eluted fractions shown in figure 4.5). After this step, imidazole was immediately removed by dialysis using 5 mM Tris–HCl (pH 7.6) buffer.

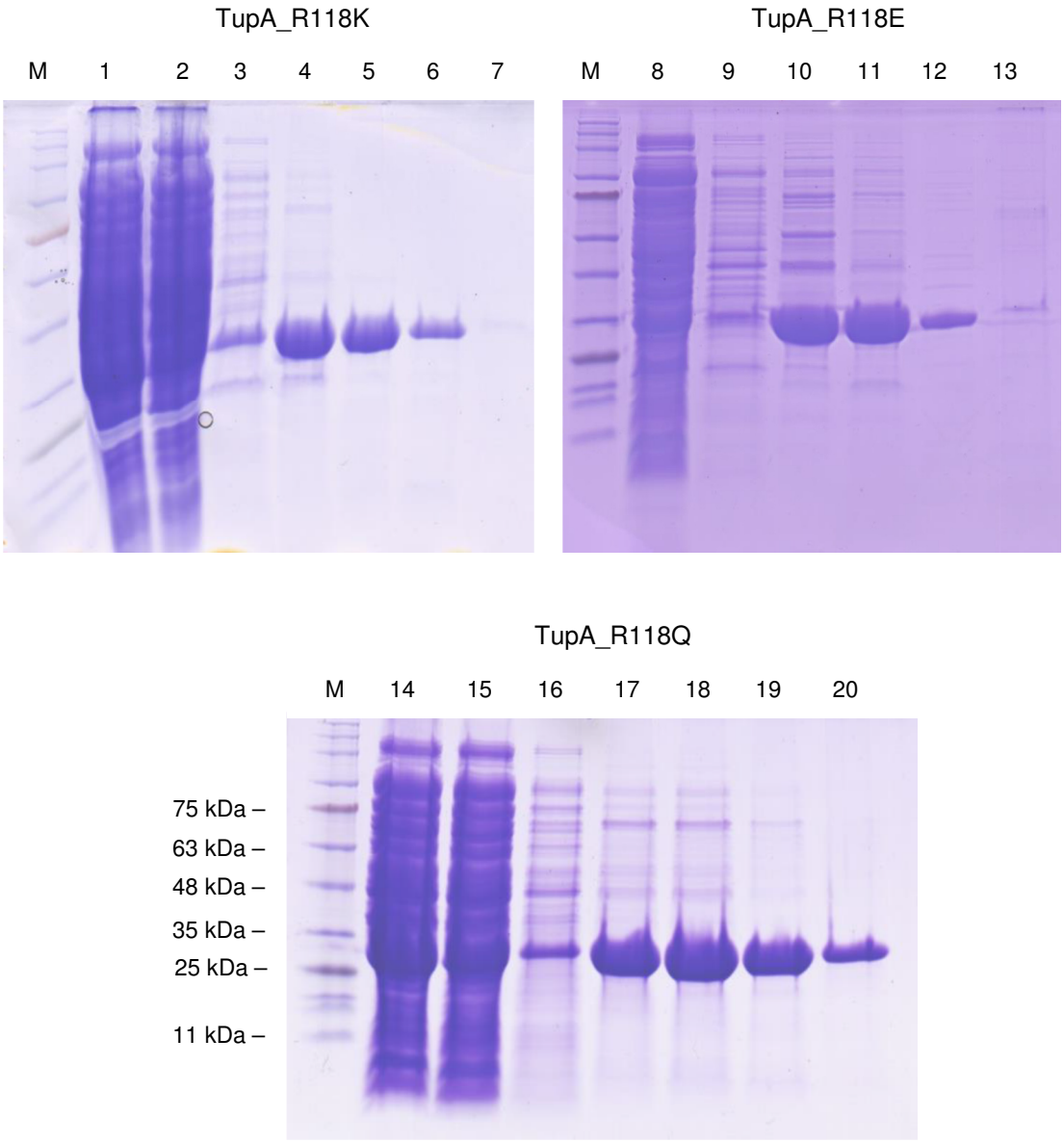


Figure 4.5 - Purification of TupA mutants evaluated by SDS-PAGE (12.5%), stained with coomassie brilliant blue: (M) molecular weight markers; 1) Soluble fraction of cells transformed with pET-TupA_R118K vector; (2 to 7 correspond to TupA_R118K mutant) 2) Flow through; 3) Washing step; 4) Elution with 20 mM imidazole; 5) Elution with 50 mM imidazole; 6) Elution with 100 mM imidazole; 7) Elution with 250 mM imidazole; (8 to 13 correspond to TupA_R118E mutant) 8) Flow through; 9) Washing step; 10) Elution with 20 mM imidazole; 11) Elution with 50 mM imidazole; 12) Elution with 100 mM imidazole; 13) Elution with 250 mM imidazole; (14 to 20 correspond to TupA_R118Q mutant) 14) Soluble fraction of cells transformed with pET- TupA_R118Q vector; 15) Flow through; 16) Washing step; 17) Elution with 20 mM imidazole; 18) Elution with 50 mM imidazole; 19) Elution with 100 mM imidazole; 20) Elution with 250 mM imidazole.

As shown in figure 4.5, during the first step of purification the resin was probably saturated with the sample and the protein was diffuse along the fractions.

4.4. Biochemical characterization

4.4.1. Extinction Coefficient Determination

As mention in section 3.4.1 the extinction coefficient at 280 nm (ε_{280}) of TupA and ModA were already determined showing values of 29.7 $mM^{-1}cm^{-1}$ and 22.5 $mM^{-1}cm^{-1}$, respectively. Using the bioinformatics tool, JustBio[84], it was found 30.44 $mM^{-1}cm^{-1}$ as the ε_{280} value for the three mutated proteins.

Experimentally, the ε_{280} was calculated by determining the mutant concentration using BSA as standard. After that, a new calibration curve was performed with different dilutions of each mutant. The extinction coefficient at 280 nm was determined from the slope of the curve obtained by linear regression of the absorbance at 280 nm versus protein concentration (Appendix VII). The ε_{280} for TupA_R118K, TupAR118_E and TupAR118_Q were 32 ±1 $mM^{-1}cm^{-1}$, 27 ±1 $mM^{-1}cm^{-1}$ and 24 ±1 $mM^{-1}cm^{-1}$, respectively. Despite the differences between the experimental ε_{280} values, they are in good agreement with the theoretical ones.

4.4.2. Protein Gel Shift Assay

Sequence analysis suggests that *Da*G20 TupA is a tungstate-binding protein that is able to bind tungstate as well as molybdate ions.[18] To test the affinity and specificity of the three mutants as well as ModA with the two different anions, native polyacrylamide gel electrophoresis of protein samples pre-incubated with MoO_4^{2-} and WO_4^{2-} was carried out as described in section 3.4.2. The samples were submitted to a size-exclusion PD-10 minitrap G-25 column prior to loading on native polyacrylamide gel in order to separate the unbound ions and ensure that differences in mobility were only due to binding anions to the proteins.

As seen in figure 4.6, TupA can bind both metals since a mobility shift is observed for the protein incubated with the metals when compared to the free form. In the case of the mutants, only TupA_R118K showed a mobility difference upon binding tungstate. This result evidences the importance of a positively charged residue like arginine at this position for the coordination of WO_4^{2-} oxoanion and that this residue can be only replaced by another with similar characteristics, as is the case of lysine. For ModA, the native gel shows that the protein has comparable affinity for both metals, similar to TupA

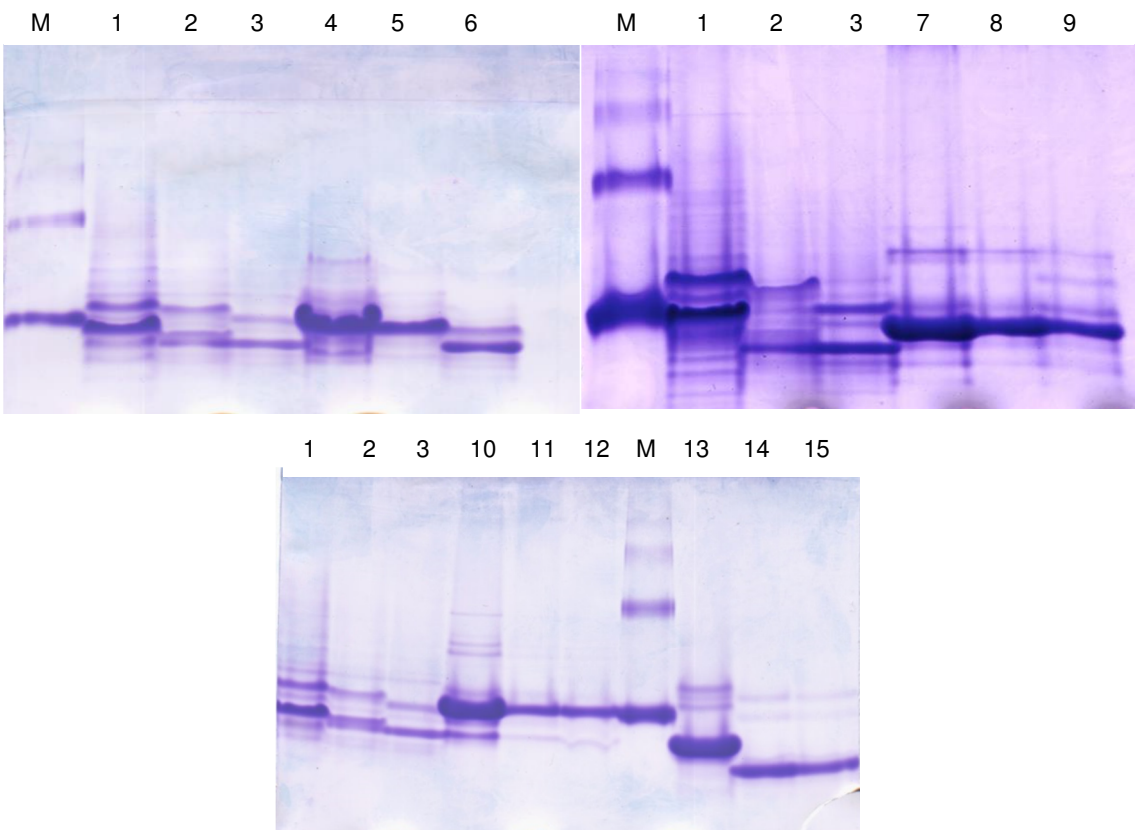


Figure 4.6 - Ligand-dependent mobility shift assays for the three mutated proteins in the presence of two different oxoanions (10-fold excess); (M): BSA protein as control, (1): TupA, (2): TupA+MoO_4^{2-}; (3): TupA+WO_4^{2-}; (4) TupA_R118K, (5): TupA_R118K+MoO_4^{2-}, (6): TupA_R118K+ WO_4^{2-}, (7): TupA_R118E; (8): TupA_R118E+WO_4^{2-}; (9): TupA_R118E+MoO_4^{2-}, (10): TupA_R118Q; (11): TupA_R118Q+MoO_4^{2-}, (12): TupA_R118Q+ WO_4^{2-}, (13): ModA, (14): ModA+ MoO_4^{2-} and (15): ModA+ WO_4^{2-}

4.4.3. Denaturation pattern of recombinant ModA, TupA and mutants evaluated by Urea-polyacrylamide gel electrophoresis

Urea-polyacrylamide gel electrophoresis studies have been used before to evaluate the interaction between the transferrin (hTF) transport system and Fe^{3+}. In this system it was observed that when iron is bound, each domain rearranges to form the *closed conformation*. Correspondingly, upon release of iron, the metal-binding domains move apart. In the *closed* form, hTF can be recognized by the hTF-cell receptors and iron is internalized by the cell.[86] In the apo (*open)* form, hTF has a moderate gel migration, opposite to the holo protein (*closed conformation*) which migrates very extensively in the gel. In general, migration through these gels is influenced by the pH, time and temperature of the experiment together with shape, size and charge.

Analogously, the influence of the ligands (molybdate and tungstate) on both proteins (TupA and ModA) was studied by this technique since conformational changes are expected upon ligand binding. To determine the effect of molybdate and tungstate binding on the conformation of each mutant as well as the wild type protein, gel electrophoresis experiments were carried out using a TBE Urea gel system. To ensure that all the protein molecules interact with metal oxoanions, samples were prepared using a 10-fold excess of metal and the unbound metal was then separated by size exclusion chromatography. Figure 4.7 shows the migration pattern of samples prepared in the presence and absence of ligands.

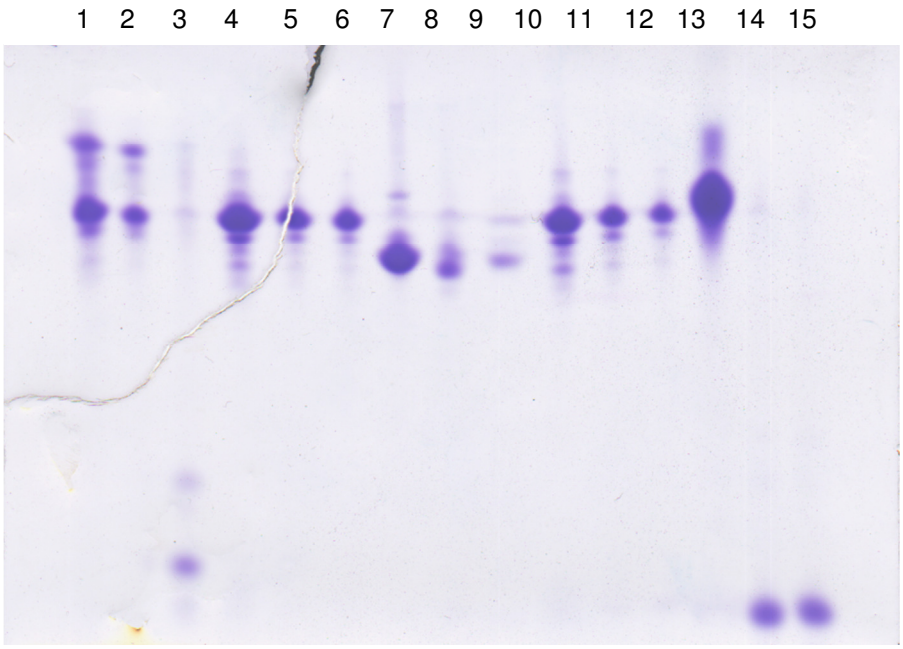


Figure 4.7 - Urea polyacrylamide gel electrophoresis of samples containing (1): TupA, (2): TupA+MoO_4^{2-}; (3): TupA+WO_4^{2-}; (4) TupA_R118K, (5): TupA_R118K+MoO_4^{2-}, (6): TupA_R118K+ WO_4^{2-}, (7): TupA_R118E; (8): TupA_R118E+MoO_4^{2-}; (9): TupA_R118E+WO_4^{2-}, (10): TupA_R118Q; (11): TupA_R118Q+MoO_4^{2-} and (12): TupA_R118Q+ WO_4^{2-}, (13): ModA, (14): ModA+ MoO_4^{2-} and (15): ModA+ WO_4^{2-}. All samples applied to the gel except (1), (4), (7),(10) and (13) were first passed through a size-exclusion PD-10 minitrap G-25 columns to eliminate the excess of ligand.

Analysis of the gel suggests that TupA+WO_4^{2-}, ModA+WO_4^{2-} and ModA+MoO_4^{2-} adopt a very compact structure migrating further than the other samples under increasing denaturing conditions. Accordingly, wild type TupA, TupA_R118K, TupA_R118Q and ModA are probably in the open form and denature more rapidly in the 6M urea gradient. The same results are obtained for two of the mutants, in the presence of molybdate or tungstate (TupA_R118K+ MoO_4^{2-} , TupA_R118K+ WO_4^{2-}, TupA_R118Q+ WO_4^{2-} and TupA_R118Q+ MoO_4^{2-}). The reason for this is probably because none of the metals bound tightly to these proteins, which were in the open form. The results obtained for the TupA_R118E mutant also indicate that in the presence or absence of any of the metals, the protein is in the open form. The fact that this mutated protein samples migrate a little further than the others might be due to charge effects

and not ligand binding. The results support the major role of arginine 118 in metal coordination, also observed by native gel analysis.

In order to quantify the protein metal interaction, isothermal titration calorimetry (ITC) binding studies were performed.

4.4.4. The Isothermal Titration Calorimetry

ITC is a sensitive method to determine affinity/dissociation constants for protein-ligand binding, namely, tungstate and molybdate to the corresponding transporting proteins, TupA and ModA, in nanomolar and subnanomolar ranges.[1, 30] This technique has the advantage that nearly all interactions create a heat change that can be monitored with a high-sensitivity calorimeter, allowing the determination of the binding enthalpy (ΔH_{obs}) and dissociation constant.[91] It was reported by *Otrelo-Cardoso et al* that the observed behavior of TupA at 30 °C is consistent with an exothermic process, with a single binding site model. However, the thigh nature of these bindings precluded an accurate fit to determine the K_D values. [30] In this work we performed displacement titrations in order to obtain the correct affinity constants. ITC of TupA and ModA showed that the proteins exothermically bind tungstate and molybdate; however the TupA_R118K only binds exothermically tungstate. The TupA_R118E and TupA_R118Q mutants are not able to bind either tungstate or molybdate.

Figure 4.8 and table 4.1 show the ITC data collected for TupA with an exothermic interaction between the protein and the metals and a 1:1 stoichiometry, as deduced from the heat release upon addition of tungstate or molybdate to the protein solution. Direct titration of molybdate against TupA produced an exothermic binding with a K_D value of 6.1 ± 0.9 nM. The value of ΔH_{obs} (~6.6 kcal/mole of injectant) is also significantly less favorable when compared with the tungstate binding. In contrast the binding of tungstate to TupA is much more exothermic with ΔH_{obs} being increased to ~14 kcal/mole of injectant.[30]

To determine a K_D value for tungstate a binding competition strategy was adopted. A displacement titration of the molybdate-saturated protein with tungstate clearly showed that the protein favors the binding of tungstate even when the binding site is occupied with a molybdate molecule. The apparent binding constant depends on the concentration of free molybdate which was 0.5 mM, and therefore K_D for tungstate was determined to be 6.30 ± 0.02 pM (figure 4.8 C, Table 4.1) when the protein is saturated with molybdate. The displacement titration and the particularly low K_D value for tungstate indicate that the latter should be the physiological substrate of TupA as expected.[18, 30]

Table 4.1 - Data for the ITC analysis of oxoanion binding to TupA protein at 30 ℃

Protein (+oxoanion)	Ligand	N	K_A (M^{-1})	K_D (nM)	ΔH (kcalmol^{-1})
TupA	WO_4^{2-}	0.842±0.001	$2 \times 10^9 \pm 2 \times 10^9$	0.5 ± 0.5	−13.500±0.005
	MoO_4^{2-}	0.868±0.002	$16 \times 10^7 \pm 2 \times 10^7$	6.3 ± 0.8	−6.600 ± 0.003
TupA+0.5mM MoO_4^{2-}	WO_4^{2-}	0.845±0.003	$1600 \times 10^8 \pm 6 \times 10^8$	$6.30 \times 10^{-3} \pm 0.02 \times 10^{-3}$	−14.60 ± 0.04
TupA+0.5mM WO_4^{2-}	MoO_4^{2-}	No displacement			

n = measured stoichiometry of binding

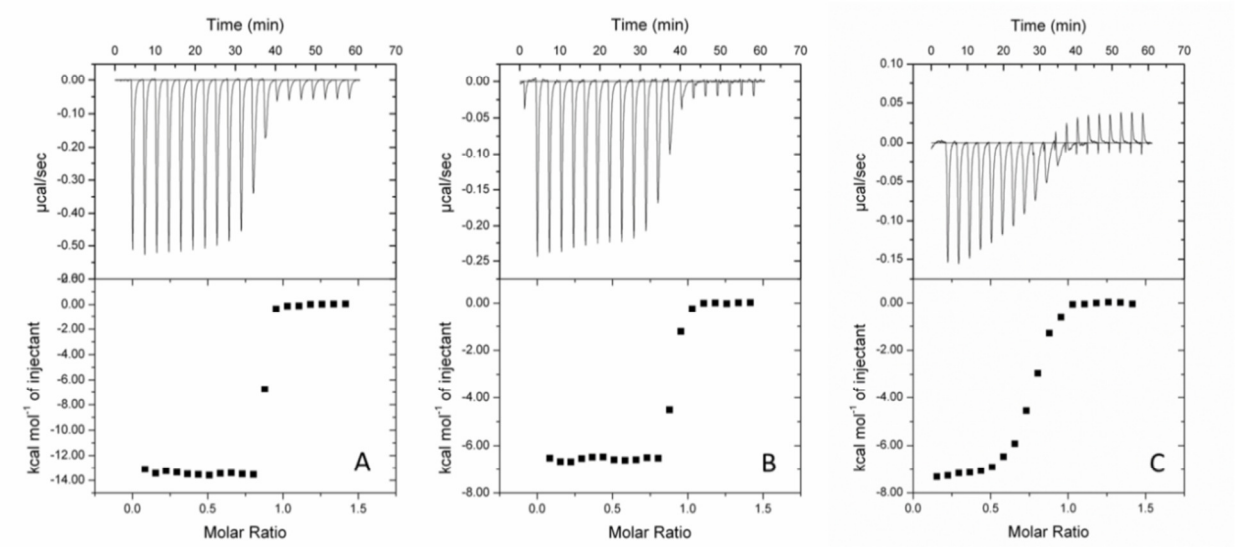


Figure 4.8: Isothermal titration calorimetry of ligand binding to TupA. 10 µM of TupA was titrated with injections of 100 µM tungstate (A) and 100 µM molybdate (B); (C) Displacement titration of 10 µM TupA incubated with 0.5 nM molybdate, with injections of 100 µM tungstate. Data was fitted with ORIGIN software. The raw ITC data is shown in the top graphs.[30]

The ITC assay in ModA aims to investigate if the reaction is exothermic and the interaction behavior between both ligands and ModA is similar to the previously described TupA (figure 4.9, table 4.2). However, the heat release of ModA is much lower than TupA and together with the tight binding nature of these isotherms we were unable to have an accurate fit to determine the K_D value for both molybdate and tungstate. Displacement titrations were done with the protein saturated with 0.1 mM of each ligand and we could conclude that ModA binds both molybdate and tungstate with high affinities and also that it is unable to replace tungstate for molybdate or vice versa. In this case it was not possible to use the binding competition assay to calculate the K_D as it was adopted for TupA.

As molybdate and tungstate are very similar in their geometry and charge it can be proposed that transport systems might not be able to discriminate both anions. [92]

Table 4.2 - ITC analysis data of oxoanion binding to ModA protein at 30 °C

Protein (+oxoanion)	Ligand	n	K_A (M^{-1})	K_D (nM)	ΔH (kcalmol^{-1})
ModA	WO_4^{2-}	0.820±0.01	$9\times10^7 \pm 8\times10^7$	12±10	–1.900±0.007
	MoO_4^{2-}	0.781±0.02	$2\times10^7 \pm 1\times10^7$	43± 25	–2.70± 0.01
ModA+0.1mM MoO_4^{2-}	WO_4^{2-}		No displacement		
ModA+0.1mM WO_4^{2-}	MoO_4^{2-}		No displacement		

n = measured stoichiometry of binding

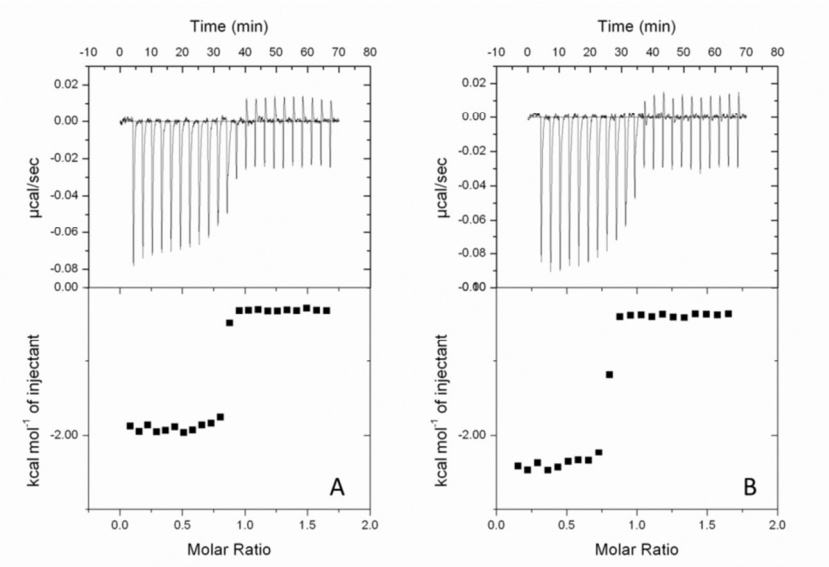


Figure 4.9: Isothermal titration calorimetry of ligand binding to ModA. ModA (10 µM)was titrated with injections of 100 µM tungstate (A) and 100 µM molybdate (B). Data was fitted with ORIGIN software. The raw ITC data is shown in the top figures.

In case of TupA_R118K the stoichiometry is around one mole oxoanion per mole of protein, as deduced from the heat release upon the addition of tungstate or molybdate to the protein solution (figure 4.10). Direct titration of tungstate against TupA_R118K produced an exothermic binding with a K_D value of 1.8 ± 0.7 nM. The value of ΔH_{obs} (~15.9 kcal/mole of injectant) is also more satisfactory when compared to the molybdate binding experiment. However, the extremely high affinity of the protein for tungstate resulted in a very steep binding curve which hampers the determination of the correct value of the K_D. Analyzing the data from the molybdate binding we can see that the affinity between TupA_R118K and molybdate is lower than the tungstate and the equivalent binding for the wild type (figure 4.10 B; table 4.3). The data analysis shows clearly that TupA_R118K has a high affinity for tungstate and a lower affinity for molybdate. In addition, comparison of the data obtained for the recombinant wild type TupA protein and for TupA_R118K mutant shows that the mutant is more selective for tungstate.

To determine a reliable K_D value for tungstate a binding competition strategy is required as mention in 4.5.4.1.However we still need to redo the titrations with molybdate to have a more reliable K_D required for the displacement titration with tungstate.

Table 4.3 –ITC analysis data of oxoanion binding to TupA_R118K protein at 30 ℃

Protein	Ligand	N	K_A (M^{-1})	K_D (nM)	ΔH (kcalmol^{-1})
TupA_R118K	WO_4^{2-}	1.160±0.002	$6 \times 10^8 \pm 2 \times 10^8$	1.7±0.6	–15.9±0.1
	MoO_4^{2-}	1.04±0.03	$6 \times 10^6 \pm 3 \times 10^6$	167±75	-0.78± 0.04

n = measured stoichiometry of binding

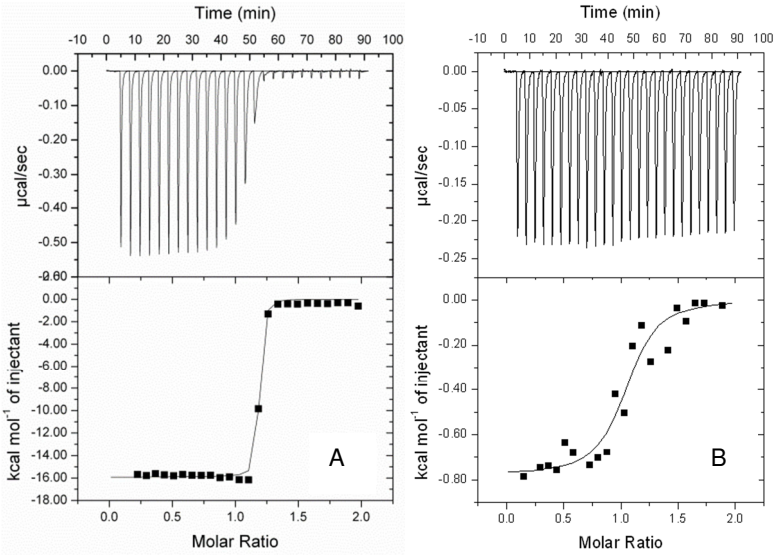


Figure 4.10: Isothermal titration calorimetry of ligand binding to TupA_R118K (raw and processed data). TupA_R118K (10 µM) was titrated with injections of 100 µM tungstate (A) and 100 µM molybdate (B). Data was fitted with ORIGIN software. The raw ITC data is shown in the top figures.

Direct titration of tungstate against TupA_R118E produced a lower exothermic binding with a meaningless K_D and stoichiometry value demonstrating there is no protein-metal binding. The same is true for molybdate (figure 4.10 B).

The data analysis shows clearly that TupA_R118E has drastic decrease in affinity for tungstate and probably a nonspecific affinity for molybdate. Possibly to obtain a linear response between TupA_R118E and MoO_4^{2-} we might have to use different protein and ligand concentration.

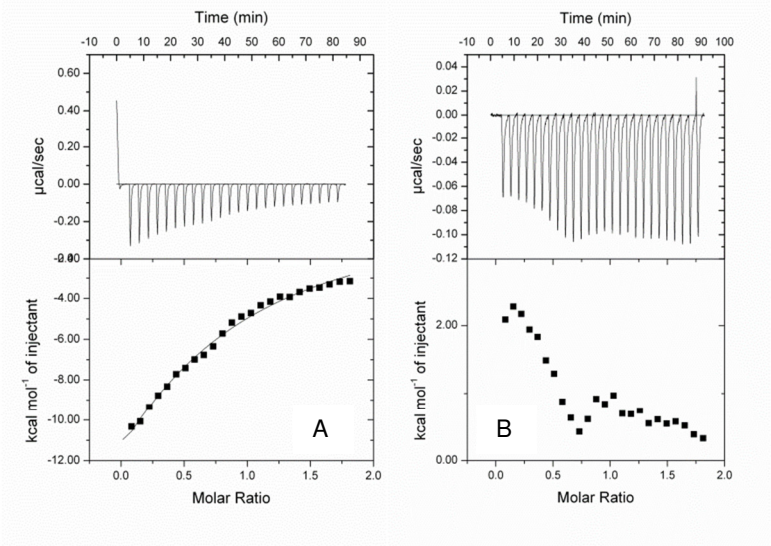


Figure 4.11: Isothermal titration calorimetry of ligand binding to TupA_R118E. TupA_R118E (10 µM) was titrated with injections of 100 µM tungstate (A) and 100 µM molybdate (B). Data was fitted with ORIGIN software. The raw ITC data is shown in the top graphs**.**

Despite the results obtained from native polyacrylamide gel, the ITC revealed that TupAR118Q can bind tungstate (figure 4.12, table 4.4). This discrepancy could be related to the lower sensitive of the native polyacrylamide gel compared to the ITC experiment and this binding is not strong enough to be evident in the gel. The binding between TupA_R118Q and the metal is however lower than the binding of wild type TupA or TupA_R118K suggesting that the charge of the residue is very important for the coordination.

Again, the result of the interaction between the protein and molybdate shows that there is no binding between them and the K_D of the reaction could not be determined.

Table 4.4 - ITC analysis data of oxoanion binding to ModA protein at 30 ℃

Protein	Ligand	N	K_A (M^{-1})	K_D (nM)	ΔH (kcalmol^{-1})
TupA_R118Q	WO_4^{2-}	0.647±0.002	139x10^5± 6×10^5	72±3	−2.22±0.09
	MoO_4^{2-}	No determined			

n = measured stoichiometry of binding

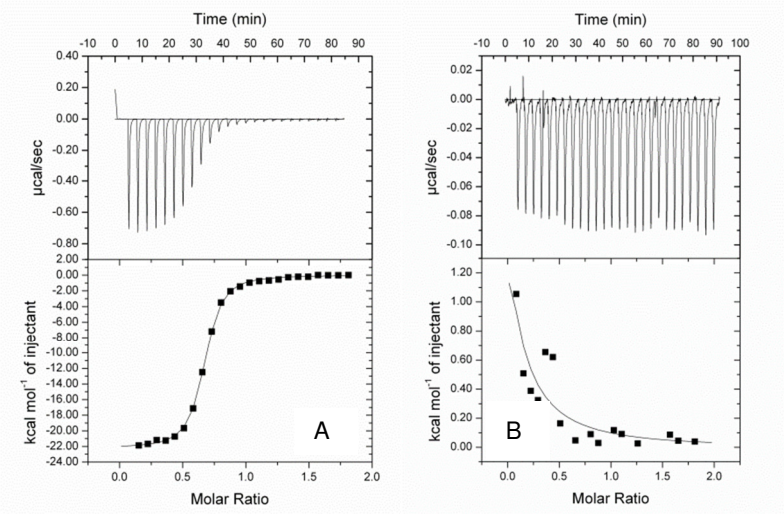


Figure 4.12: Isothermal titration calorimetry of ligand binding to TupA_R118Q. 10 µM of TupA_R118Q was titrated with injections of 100 µM tungstate (A) and 100 µM molybdate (B). Data was fitted with ORIGIN software. The raw ITC data is shown in the top graphs

4.5. X-ray crystallography

The first approach to crystallize the proteins under study consisted in testing several precipitating solutions for different protein concentrations and ratio of protein + precipitating solution in the crystallization drops in 96-well plates by the vapor diffusion sitting drop method. In addition to the commercial screens, an 80 conditions in-house screen (based on the screen of Jancarik *et al*)[87] was tested as well as EWI/EWII and Additive Screen 1 and 2.

4.5.1. TupA

The crystallization of recombinant TupA as well as data processing and structure resolution were done by the PhD student Ana Rita Cardoso. Nevertheless, the work was repeated so that the acquired know how could be applied to other datasets collected in the time course of this thesis.

4.5.1.1. Crystallization studies and diffraction experiment

To crystallize TupA from *Da*G20 several commercial screens were tested in 96-well plates using the sitting drop by vapor diffusion method. Plate shaped crystals appeared about four days after crystallization setup when using a solution of 0.2 M magnesium chloride, 0.1 M HEPES pH 7.5 and 30% (*w*/*v*) polyethylene glycol 3350 as the precipitating agent, as mentioned in section 3.5.1 (figure 4.13).[30] The scale-up optimization was done by changing the protein: precipitant ratio in the crystallization drop.

The same crystallization conditions were repeated with the main goal of producing more crystals for protein-ligand structure determination. Co-crystallization and soaking techniques were performed with tungstate and molybdate but with no success since in the presence of ligands the crystals no longer diffract the X-rays.

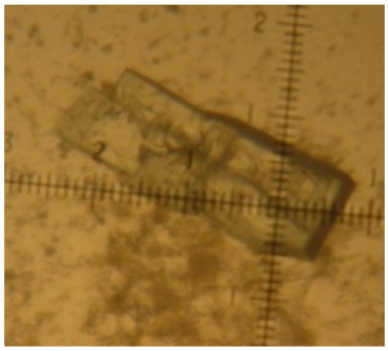


Figure 4.13: TupA crystal grown in 0.2 M magnesium chloride, 0.1 M HEPES, pH 7.5 and 30% (*w*/*v*) polyethylene glycol 3350.

Before harvesting the crystals a stabilizing solution known as harvesting buffer was added to the drops. As stated in section 3.5, this solution has the same composition of the crystallization solution except for a higher concentration of precipitant agent (32% of PEG 3350), to avoid the crystals dissolution during handling. Before freezing in liquid nitrogen, the crystals were transferred to a cryoprotectant solution containing glycerol. This prevents the formation of ice which appears in the diffraction pattern in the form of rings, hence ensuring good data collection at cryogenic temperatures.

The diffraction experiment was performed at the ESRF, at beamline ID 23-1 and the crystals obtained diffracted up to 1.43 Å resolution (figure 4.14, table 4.5).

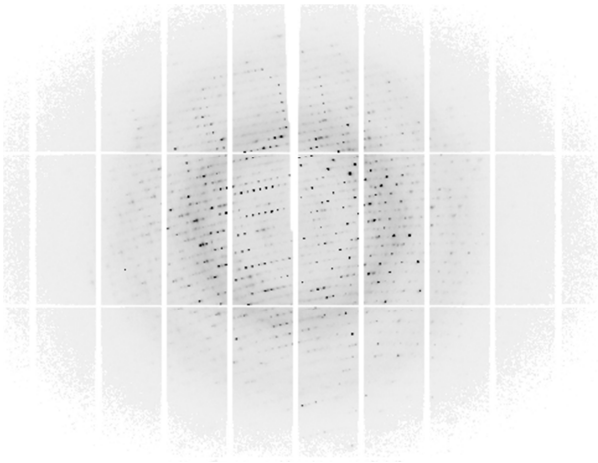


Figure 4.14: Diffraction pattern of the TupA crystal obtained at 20ºC. The experiment was carried out at beamline ID23-1 (ESRF - France) and the data collected have a maximum resolution of 1.43 Å (edge of the diffraction pattern)

The images (reflections) were indexed using MOSFLM program[93] in order to determine the Miller indices *hkl*, the intensity (I_{hkl}) and their associated errors (σhkl). It is also possible to characterize the unit cell by determining the constants a, b, c and the α, β, γ angles. Thus, TupA crystal belongs to a monoclinic crystal system, with the cell constants: *a* = 52.25 Å; *b* = 42.50 Å; *c* = 54.71 Å; β = 95.43º.

Through the analysis of the reflection intensities based on AIMLESS (part of the CCP4 program) (Collaborative Computational Project Number 4) the most likely space group can be determined. In this case we concluded that the TupA crystal belongs to $P12_11$ space group.

During the diffraction experiment the crystal may be damaged (radiation damage), which can affect the reflection intensities. Thus it is necessary to scale the data because it is not possible to compare intensities that are not on the same scale. For TupA, data collection statistics are presented in table 4.5.

Table 4.5: Data collection and processing statistics for the TupA crystal. Values in parentheses correspond to the highest resolution shell.

$$^{*}R_{merge} = \frac{\sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|}{\sum_{hkl}\sum_i I_i(hkl)} \qquad ^{+}R_{pim} = \frac{\sum_{hkl}\left[\frac{1}{N-1}\right]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|}{\sum_{hkl}\sum_i I_i(hkl)}$$

$$^{\S}R_{meas} = \frac{\sum_{hkl}\left[\frac{N}{N-1}\right]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|}{\sum_{hkl}\sum_i I_i(hkl)}$$

Data collection parameters	
X-ray source	ID23-1 (ESRF, Grenoble)
Detector	PILATUS 6M-F
Wavelength (Å)	0.954

Processing statistics	
Unit-cell parameters (Å, º)	a = 52.25; b = 42.50; c = 54.71; β = 95.43
Space group	$P12_11$
Molecules per AU	1
Matthews coefficient (Å^3, Da)	2.09
Mosaicity (º)	0.22
Resolution range (Å)	42.50–1.43 (1.45–1.43)
$<I/\sigma I>$	10.3 (2.1)
R_{merge} (%) *	4.1 (33.5)
R_{pim} (%) +	2.7 (23.4)
R_{meas} (%) §	5.0 (4.1)
Multiplicity	3.0 (2.8)
No. of observed reflections	132,115 (6,040)
No. of unique reflections	43,950 (2,151)
Completeness (%)	99.1 (98.8)

The statistics presented in table 4.5 were thoroughly analyzed in order to attest the quality of the collected data. Each parameter is going to be described in detail in the following paragraphs.

The mosaicity corresponds to the angular measure of the disorder degree of the unit cells in the crystal. Lower values of mosaicity (bellow 1) indicate better ordered crystals and as a result better diffraction patterns.[94] Thus the mosaicity average of the crystal measured was 0.22º indicating a good packing of the unit cells within the crystal.

Another parameter to consider is the average of the ratio between the intensities and the associated error, $<I/\sigma I>$. In this case the intensities are ten times higher than the error associated ($<I/\sigma I>$. = 10.3) reflecting the high signal to noise ratio.[69]

The resolution of 1.45-1.43 Å allows the determination of the protein structure, obtaining a model at an atomic level. Three other factors that also reveal the good quality of the data are the values of R_{merge}, R_{pim} and R_{meas}.

The R_{merge} is a measure of the agreement among multiple measurements of the same reflections in different frames or different datasets.

Useful data for structure determination should present values for the overall R_{merge} ranging from 5% to 10% and, in this case, the overall R_{merge} value obtained was 4.1%. The R_{merge} for the highest resolution shell was 33.5% which is an acceptable value.[95] The R_{pim} value is equivalent to the previous but decreases with the increase of redundancy (multiplicity) and it will reflect the increasing accuracy (precision) of the intensities as more and more observations are merged[57] This expresses the precision of the intensity average with a value of 2.7% for the experimental data. The R_{meas} is introduced as the equivalent robust indicator for data consistency. It should accurately reflect the reliability of individual measurements, independently from multiplicity.[96]

A diffraction experiment involves measuring a large number of reflection intensities. As crystals have intrinsic symmetry, some reflections are expected to be equivalent and consequently have equal intensity. The average number of measurements per individual symmetrically unique reflection is called redundancy or multiplicity. Because every reflection is measured with a certain degree of error, the higher the redundancy the more accurate the final estimation of the averaged reflection intensity will be.[69] For this dataset, the multiplicity is 3.0.

The completeness measures all the reflections that are collected and it should be as near as possible to 100%.[57] In this case the global completeness was 99.1% and 98.8% for the highest resolution shell, showing that the data is appropriate for structure determination.

As a final remark, the prior analysis revealed that the collected dataset was of good quality and could be used for structure determination, model construction and refinement.

Prior to structure determination it is important to know the number of molecules in the asymmetric unit. The Matthews's coefficient can provide this information since it is related to the solvent content.[97] Both are calculated from the unit cell and the molecular weight of the molecules present. The Matthews coefficient calculation (2.09 $Å^3 \cdot Da^{-1}$) suggests the presence of one molecule of TupA per asymmetric unit and a solvent content of 40.84%

4.5.1.2. Structure determination

The available structures deposited in the Protein Data Bank (www.rscb.org) from the families of transporters ModA and TupA have a low sequence identity (16.5%)[98], however they have a high degree of three-dimensional homology with very few structural differences.

In order to solve the TupA structure sequence alignments were performed, allowing to find the best homology models that could be used for molecular replacement (MR). Structure determination was performed with PHASER[99] using as molecular models: a protein with unknown function from *Vibrio parahaemolyticus* (PBD code 3MUQ) and the *Gs* TupA (PDB code 3LR1). [30] In the first trials to solve the phase problem the two homology models were superposed (figure 4.15). Nonetheless, a MR solution could only be acquired when searching for small sections of the protein individually. This procedure is generally used for large proteins, where a high degree of flexibility is expected, suggesting that *Da*G20 TupA is also a flexible protein that can assume multiple conformations.[30]

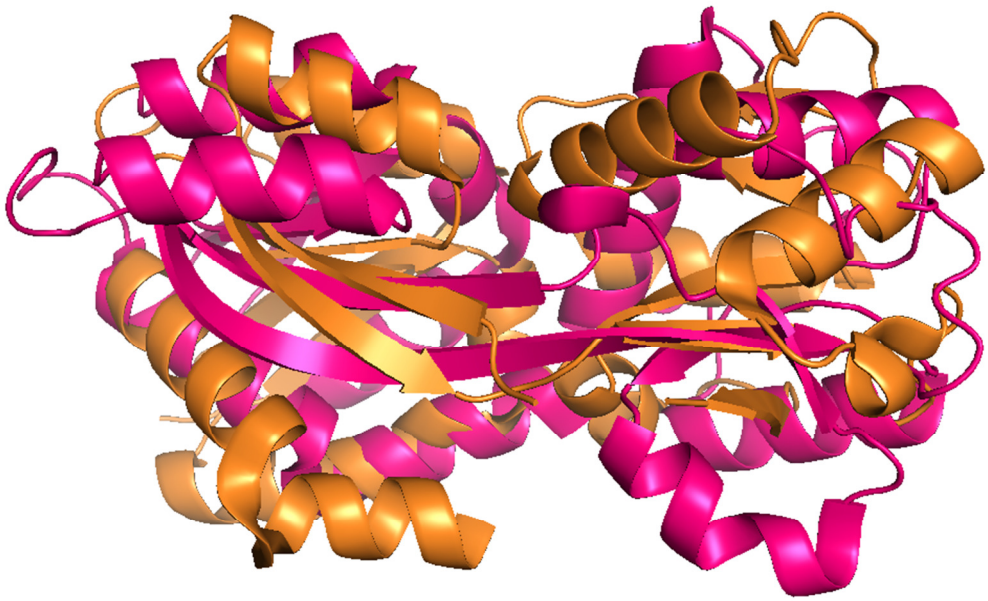


Figure 4.15: PDB of 3LR1 (orange) and 3MUQ (pink) superposed in PyMOL program

The TupA structure (figure 4.16 and figure 4.17) was determined using the data from table 4.5 and the structures presented in figure 4.15 after the structural superposition. It is possible to see the similarities of the envelope and also of some of elements of secondary structure between the two models (PDB: 3LR1 and PDB: 3MUQ) used to solve the structure of the protein TupA. The structure of *Gs* TupA was solved in presence of tungsten. The coordination of the metal is not well described and in fact it is in the form of W^{6+} and not the corresponding oxoanion, as expected. Nevertheless, it is clear that the folding of this structure with the ligand adopts a more compact conformation that the other free protein. This result corroborates the idea that TupA can have a close and open form and is in good agreement with the biochemical characterization performed by urea polyacrylamide gel electrophoresis for *Da*G20 TupA.

Figures 4.16 and 4.17 show the superposition of the solved structure of TupA protein (purple) with the two models used for molecular replacement to solve the TupA structure (data not published).

The structure is in the last steps of refinement and a manuscript is under preparation.[100]

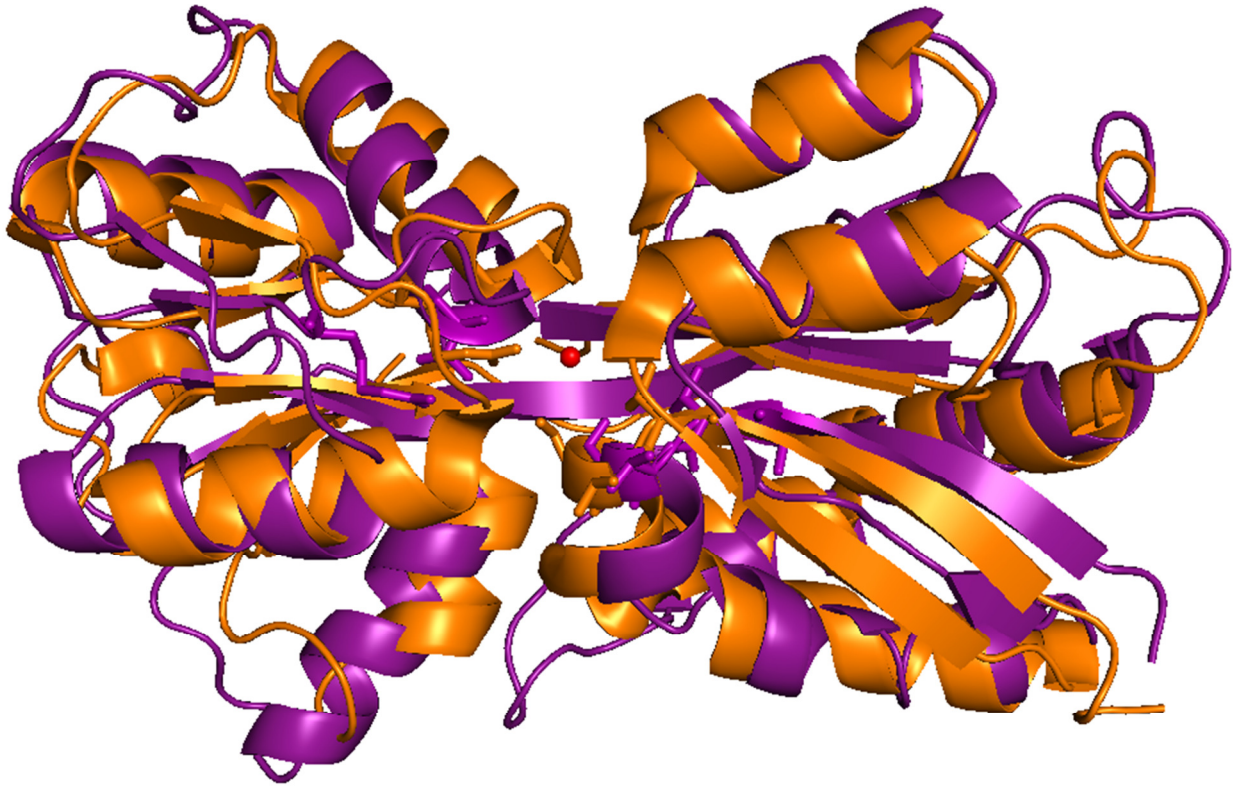


Figure 4.16: Three dimensional structure of *Da*G20 TupA (purple) (data not published) superposed with the PDB 3LR1 (orange); the sphere in red represents the tungsten metal (W^{6+}). The residues with lateral chains represent the conserved amino acids involved in the possible coordination

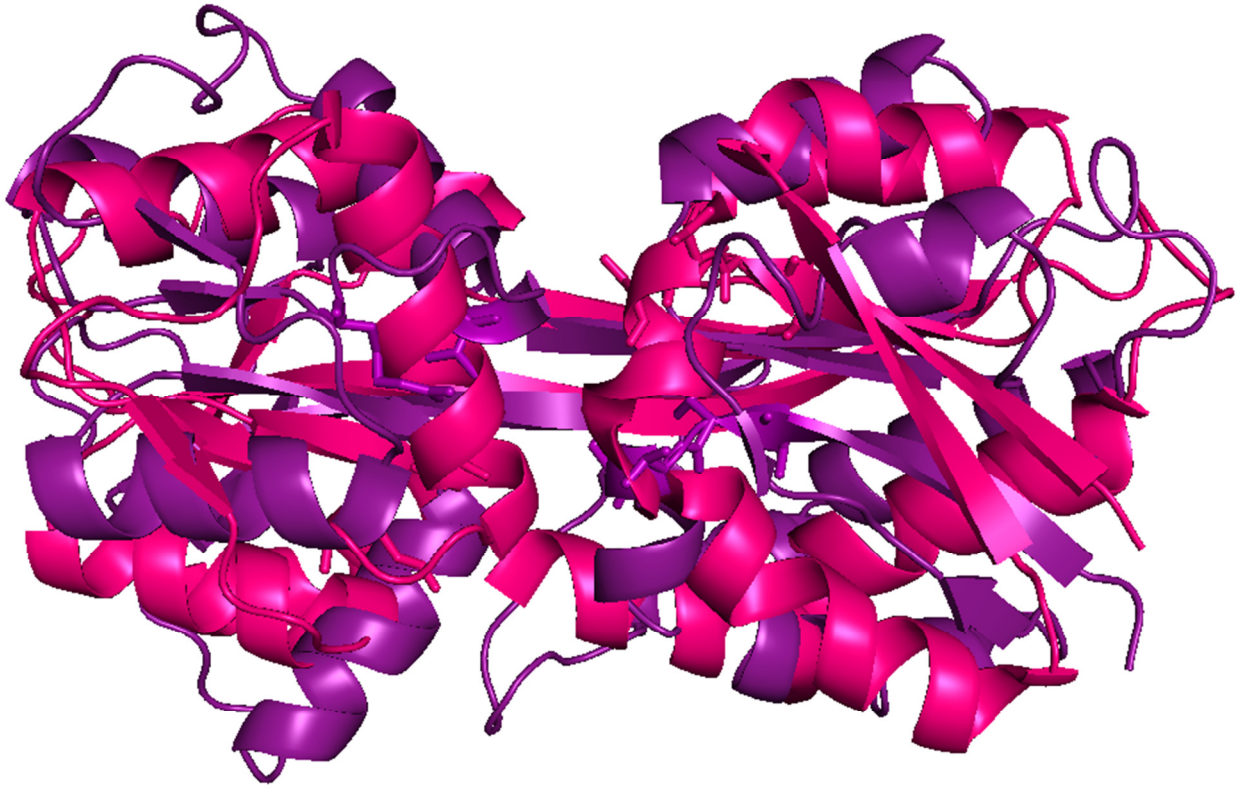


Figure 4.17: Three dimensional structure of *Da*G20 TupA (purple) (data not published) superposed with the PDB: 3MUQ (pink). The residues with lateral chains represent the conserved amino acids involved in the possible coordination

4.5.2. TupA_R118K, TupA_R118E and TupA_R118Q

4.5.2.2. Crystallization studies and diffraction experiments

With the purpose of crystallizing the three TupA mutants, numerous commercial screens were tested in 96-well plates using the sitting drop vapor diffusion method. Crystals of the three protein mutants appeared about two days after crystallization setup when using a solution of 0.1 M magnesium chloride, 0.1 M MES (pH 6.5) and 30% (*w*/*v*) polyethylene glycol 8000 as the precipitating agent as mentioned in section 3.5.3 (figure 4.18).

The scale-up optimization was then performed using the exact same conditions with only some adjustments in the ratio of the protein+precipitant volume in the drop (1µL+1 µL or 2 µL +1 µL).

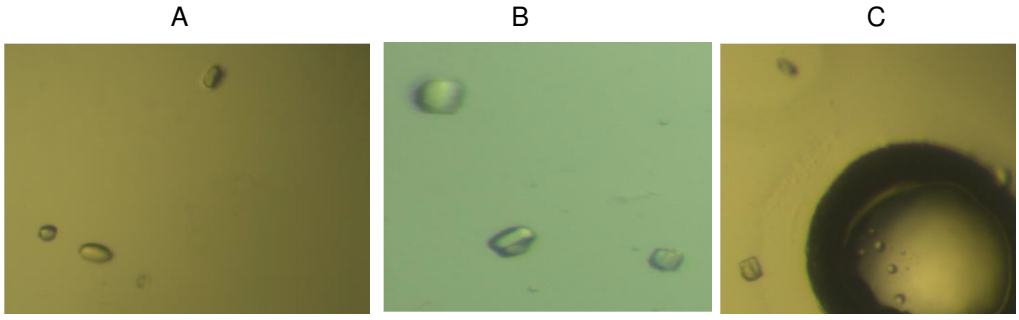


Figure 4.18: Crystals of the three mutated proteins. A: TupA_R118K (14 mg/mL); **B:** TupA_R118E (10 mg/mL) and C: TupA_R118Q (11mg/mL). The three mutated forms in 0.1 M magnesium chloride, 0.1 M MES (pH 6.5) and 30% (*w*/*v*) polyethylene glycol 8000 with ~ 0.05x0.02x0.03 mm.

The harvesting buffer containing 32% of PEG 8000 was added to the drops prior to crystal harvesting, following the approach described for the crystal of TupA.

The diffraction experiment occurred in Diamond Light Source at beamline ID I03, however no diffraction of the crystals was observed (figure 4.19A, 4.19B and 4.19C). Therefore new crystallization assays will be carried out with the main goal of obtaining good diffracting crystals and complete datasets for each of the mutants.

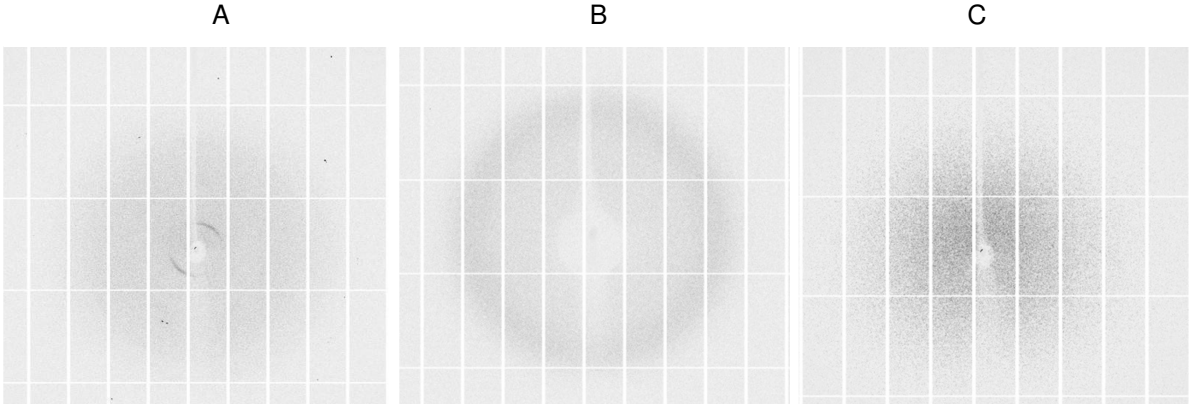


Figure 4.19: Diffraction pattern of the A: TupA_R118K, B: TupA_R118E and C: TupA_R118Q crystals obtained at 20ºC. The experiments were carried out at beamline ID03 (DLS – England)

4.5.3. ModA

4.5.3.1. Crystallization Studies and diffraction experiment

For ModA some commercial screens were also used to test in 96-well plates using the sitting drop by vapor diffusion method. Some needles appeared about five days after crystallization setup when using a solution of 28% (*w*/*v*) polyethylene glycol 6000, 0.1M sodium cacodylate (pH 6.5) and 0.2 M ammonium sulfate. To the scale-up optimization we changed the protein: precipitant ratio in the crystallization drops. New crystals were obtained again with a needle shape (figure 4.20 A). A diffraction experiment of the needles revealed poor resolution of the diffraction data (3.5 Å) (figure 4.21) and the crystallization condition was optimized using the additive screen 1 and 2 (Hampton Research). The crystals present in figure 4.20 B were obtained using the additive benzamidine hydrochloride at a final concentration of 20% (*w*/*v*) from the Additive Screen 1. These are plate shape crystals, larger than the previous and easier to manipulate with, probably, a better diffraction (figure 4.22).

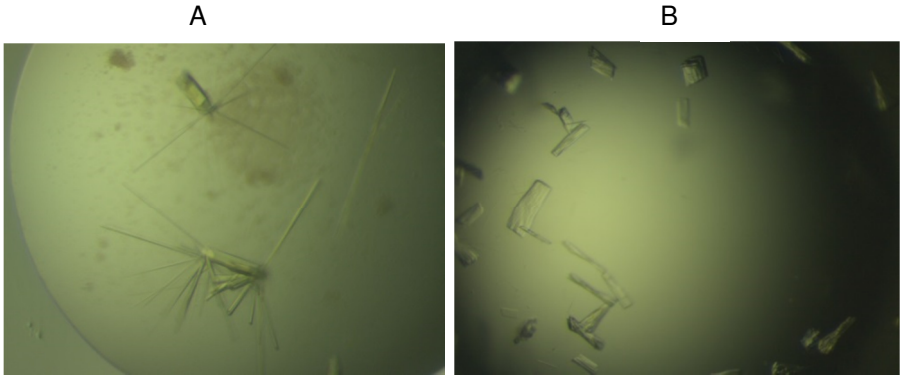


Figure 4.20: ModA crystals. A: Crystals growth in 28% (*w*/*v*) polyethylene glycol 6000, 0.1 M sodium cacodylate (pH 6.5) and 0.2 M ammonium sulfate. B: Crystals growth in 28% (*w*/*v*) polyethylene glycol 6000, 0.1 M sodium cacodylate (pH 6.5) and 0.2 M ammonium sulfate and 20% (*w*/*v*) of benzamidine hydrochloride.

To harvest the crystals, the harvesting buffer, containing 30% of PEG 6000, was added to the drops following the methodology described above for the TupA crystal.

The crystals shown in figure 4.19A were used to collect a complete dataset in the ESRF at beamline ID 23-1, diffracting up to 3.5 Å (figure 4.21, table 4.6). The plate crystals presented in figure 4.20B were also used to collect a dataset in the DLS at ID I03, where one of the crystals diffracted up to 2 Å resolution (figure 4.22. table 4.7).

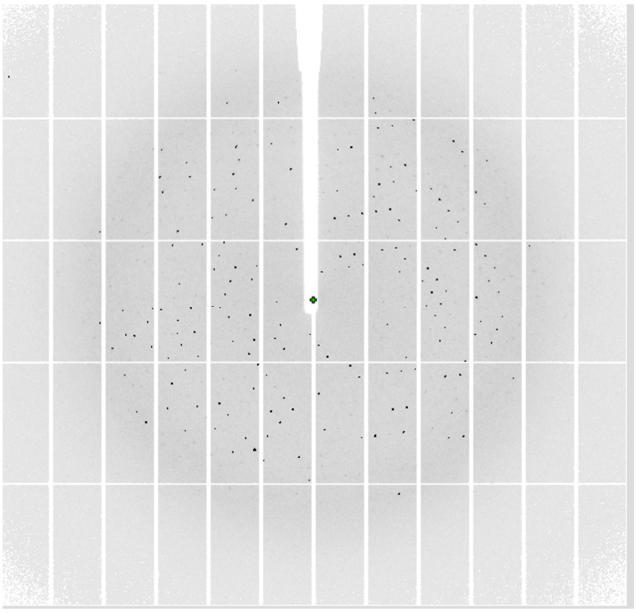


Figure 4.21: Diffraction pattern of the ModA crystal obtained at 20ºC. The experiment was carried out at the beamline ID23-1 (ESRF - France) and the data collected has a maximum resolution of 3.5 Å.

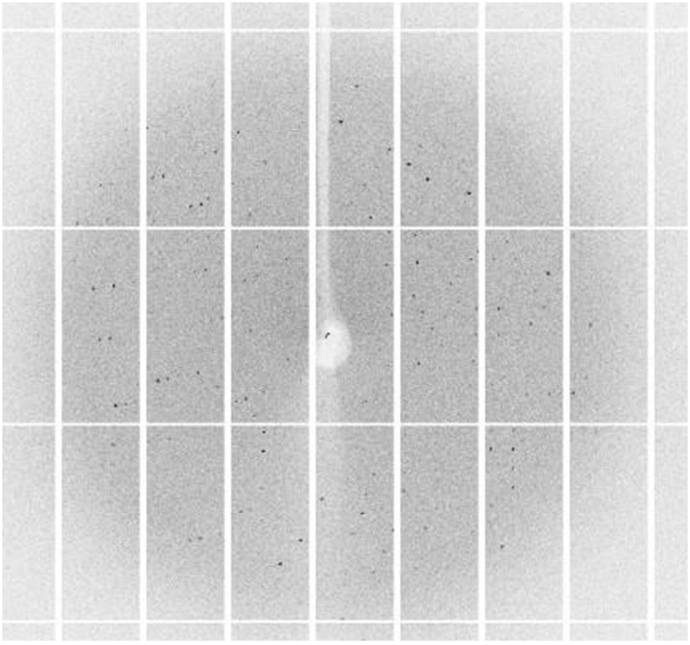


Figure 4.22: Diffraction pattern of the ModA crystal obtained at 20ºC. The experiment was carried out at beamline ID03 (DLS - England) and the data collected has a maximum resolution of 2.77 Å.

The data were indexed using MOSFLM.[93]. The crystals belong to two different space groups: C21, with cell constants *a* = 178.31; *b* = 103.22; *c* = 61.53; β = 90.05º for the crystals with maximum resolution of 3.5 Å and P21, with cell constants *a* = 56.98; *b* = 103.9; *c* = 91.89; β = 105.3º, for the better diffracting crystals. In spite of the differences in cell units and space groups, the crystalline system is the same for the two crystal forms, monoclinic. This difference can result from the packing of the crystals.

After these steps, we used the XDS package and AIMLESS from CCP4 package to scale and evaluate the data quality, resulting in the statistical parameters present in tables 4.6 and 4.7. The obtained values show the quality of the collected data for both crystal forms**.**

Table 4.6: Data collection and processing statistics for the ModA crystal (figure 4.19A). Values in parentheses correspond to the highest resolution shell

$$^{*}R_{merge} = \frac{\sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|}{\sum_{hkl}\sum_i I_i(hkl)} \qquad ^{+}R_{pim} = \frac{\sum_{hkl}\left[\frac{1}{N-1}\right]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|}{\sum_{hkl}\sum_i I_i(hkl)}$$

$$^{\S}R_{meas} = \frac{\sum_{hkl}\left[\frac{N}{N-1}\right]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|}{\sum_{hkl}\sum_i I_i(hkl)}$$

Data collection parameters		
Parameters	ModA (without additive)	ModA (with additive)
X-ray source	ID 23-1 (ESRF, France)	ID I03 (Diamond, England)
Detector	PILATUS 6M-F	PILATUS3 6M
Wavelength (Å)	0.954	0.969

Processing statistics		
Parameters	ModA (without additive)	ModA (with additive)
Unit-cell parameters (Å, °)	*a* = 178.31; *b* = 103.22; *c* = 61.53; β = 90.05	*a* = 56.98; *b* = 103.9; *c* = 91.89; β = 90.05
Space group	$C12_11$	$P12_11$
Molecules per AU	4	4
Matthews coefficient (Å^3, Da)	2.62	2.43
Mosaicity (°)	0.20	0.20
Resolution range (Å)	44.67–3.50 (3.69–3.50)	48.59–2.77 (2.92–2.77)
$<I/\sigma I>$	7.0(3.5)	8.3(2.2)
R_{merge} (%) *	11.7(34.5)	17.8(72.8)
R_{pim} (%) $^+$	9.2 (28.2)	10.6 (51.1)
R_{meas} (%) §	15.0 (4.48)	20.7 (98.6)
Multiplicity	2.2 (2.2)	3.7 (3.6)
No. of observed reflections	27081 (3938)	96383(13573)
No. of unique reflections	12080 (1767)	26279(3804)
Completeness (%)	85.2 (86.8)	99.6 (99.6)

Data analysis shows that the two collected datasets are of good quality and as a final remark, the prior analysis revealed that the second dataset collected was of good and with better quality that the first one, and could be used for model construction and refinement.

4.5.3.2. Structural insights on ModA from DaG20

The structure of ModA from *Da* G20 with the 2.77 Å resolution has not solved yet. However, it is possible to speculate about the likely structure, based on the similar proteins of ModA from other organisms.

The ModA protein from *E.coli* K12 was one of the proteins used to perform the alignment with the other species of *Desulfovibrio* (figure 4.1.) with the main goal of determining the amino acids most likely involved in the coordination of ligands. This protein from *E.coli* K12 has already been solved and the structure is available at the PDB.

In order to understand the structural/conformational behavior of ModA protein from *Da* G20 in presence and absence of its ligands, it was performed a Homology detection & structure prediction by HHpred tool and three structures of ModA with some identity were found.

The ModA binding WO_4 from *Azotobacter vinelandii* (PDB: 1ATG), ModA binding ReO_4^- from *E.coli* K12 (PDB: 3R26) and ModA from *Clostridium difficile* (PDB: 4KD5). The superposition of the three structures is present in figure 4.23.

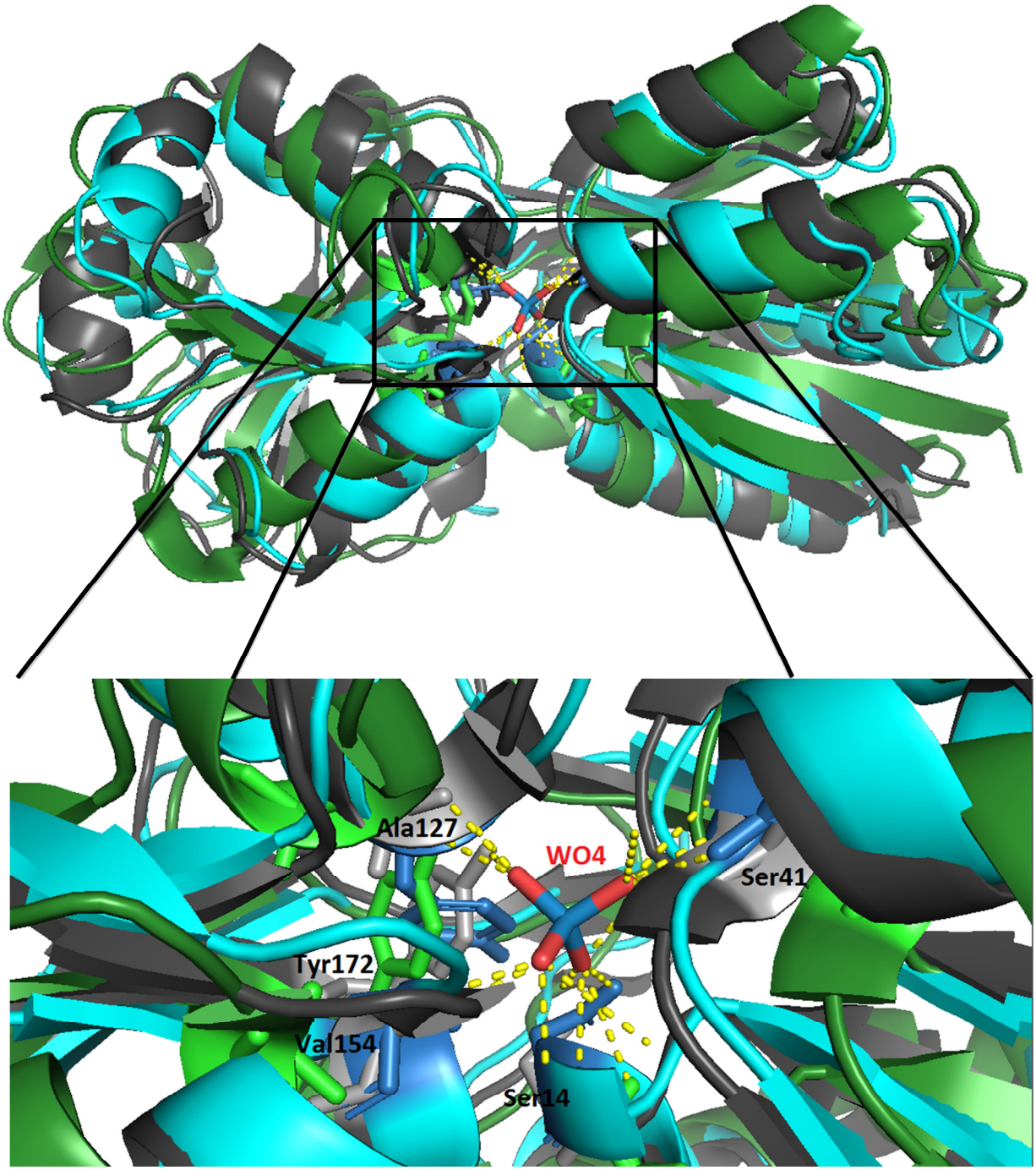


Figure 4.23: Three dimensional structures of ModA superposed (*E.coli* K12 numbering): ModA from *Azotobacter vinelandii* with WO_4 interactions (PDB: 1ATG) in blue, ModA from *Clostridium difficile* (PDB: 4KD5) in green and ModA binding ReO_4^- from *E.coli* K12 (PDB 3R26) in grey. The residues with different shades represent the conserved amino acids involved in the possible coordination.

The ModA from *E.coli* K12 shares 25% identity and 0.418 similarity with ModA from *Da* G20, the ModA from *Azotobacter vinelandii*, 35% of identity and 0.594 similarity and ModA from *Clostridium difficile* shares 26% of identity and 0.417 of similarity. The crystal structures of ModA from *Av* and also from *Ec* K12 have been solved complexed with metals. On the other hand, the uncomplexed ModA from *Clostridium difficile* exhibits an *open conformation* meaning that the secondary and tertiary structures between the ModA from *Av* and ModA from *Ec* K12 are very similar. Once again, this result is in good agreement with the biochemical characterization of ModA protein.

In order to understand the binding and metal transport of ModA and TupA a superposition of TupA from Da G20 and the ModA from *Clostridium difficile* was performed (figure 4.24).

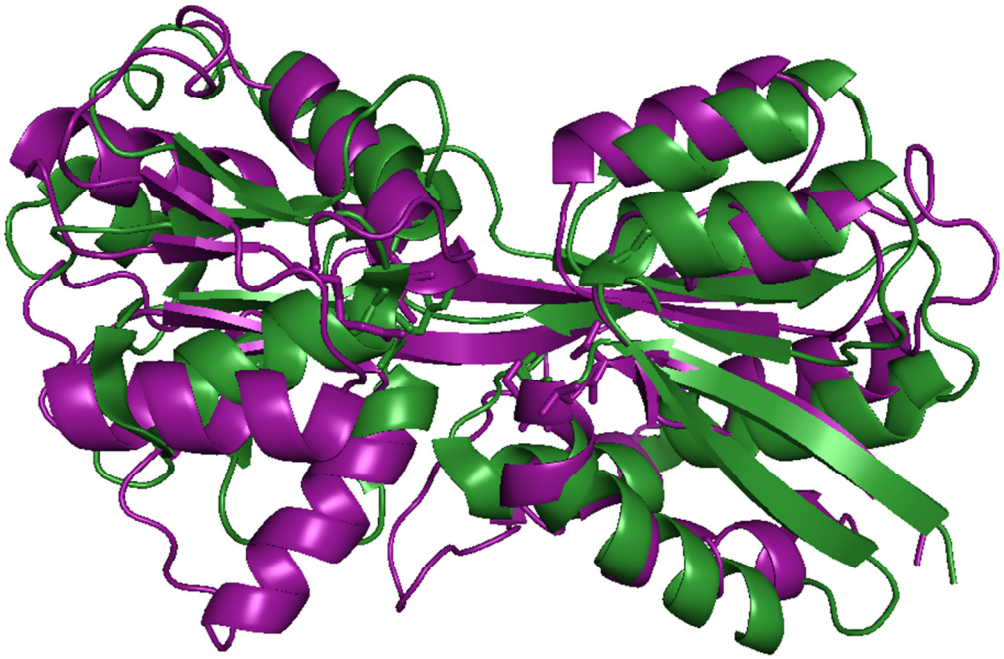


Figure 4.24: Three dimensional structures of TupA (purple) from *Da* G20 (not published) superposed with ModA PDB 4KD5 (green)

In the superposition one can see that the two proteins share the same envelope although differences are found in the second and tertiary structure elements of the proteins. The position of the amino acids involved in the coordination of the metals is also different which could be explained by the binding specificity. However our biochemical data shows that the proteins are capable of binding the two metals and more experiments are required to better understand this question.

4.5.4. Small angle X-ray scattering

This technique can be applied to biological molecules of different molecular weights and different natures. In the case of proteins, this technique is a very important tool for conformational analysis and structural changes evaluation, in solution. SAXS provides a model of the protein envelope which can be used to complement the high resolution data obtained by X-ray crystallography, including the more flexible regions of the protein.[101]

SAXS was applied to study *Da* ModA in the absence and in the presence of molybdate and tungstate ligands. The datasets were processed through the ATSAS software[88] which subtracts each buffer contribution to each protein concentration. This software is able to define the Guinier analysis which determines the radius of gyration (R_g) and the zero scattering angle intensity, *I* (0). In case of monodisperse samples, the plot is supposed to be linear, whereby the line slope gives us the R_g and its intersection with the ordinate provides the scattering intensity. The *I* (0) value normalized to the solute concentration is proportional to the molecular mass of the solute.[102] For this determination it is important to perform an assay with a standard protein, BSA, in order to determinate proportionality constant.

Once the R_g is determined as well as the *I* (0), the next step is to use the GNOM program to determine the maximum linear dimension, D_{max} and the interatomic distance distribution functions p(r). The last one is calculated from an indirect Fourier transform of the scattering between all of the electrons within the macromolecular structure. [103]

To the shape determination we used the DAMMIN program. This algorithm represents the shape by densely packed beads in a controlled search volume. The final results (figure 4.25, figure 4.26 and figure 4.27) are PDB files that have been superposed in PyMOL with the PDB file of the protein with the highest identity resulted from the ModA alignment in HHpred tool. This corresponds to ModA of *Azotobacter vinelandii* (PDB code 1ATG)**.** The superposition shows that the two proteins are quite similar, with the same folding. Nevertheless some differences are found, namely a few loops at the surface of the protein. These can be derived from the not so high sequence homology (35%) or by the fact that, in solution, the proteins have more degrees of freedom.

The Guiner analysis as well the GNOM can be seen in the figures 4.25, 4.26, 4.27 and also in the table 4.8 for ModA, ModA+Mo and ModA+W, respectively. These tables include the radius of gyration, the maximum particle dimension and the excluded volume of the hydrated particle. The molecular mass (MM) of the solute estimated from the scattering intensity are close to the ModA sequence (26761.9 Da).

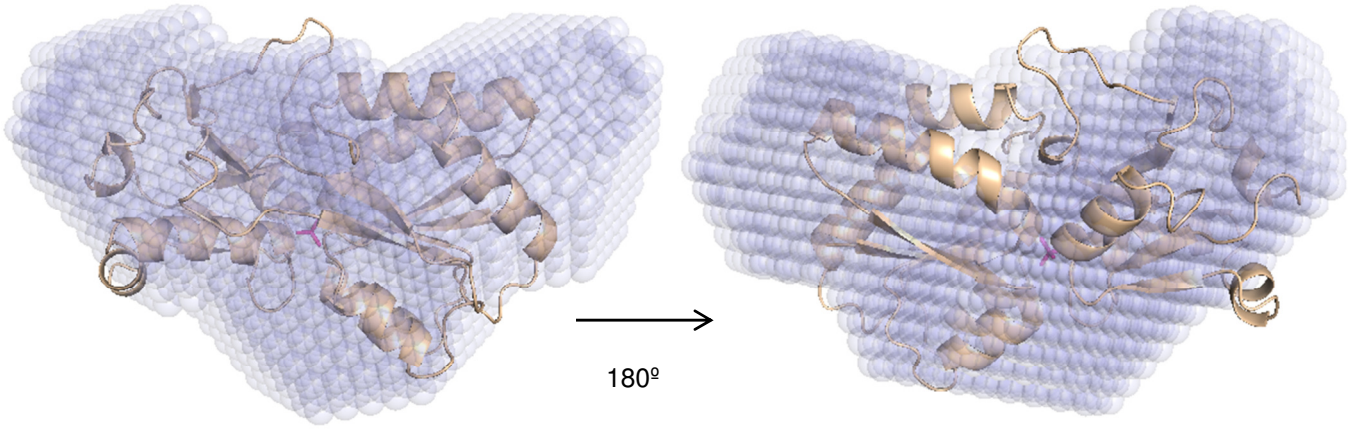


Figure 4.25.: Superposition of PDB 1ATG and the model resulted from SAXS experiment of ModA. The WO_4 ligand is in red

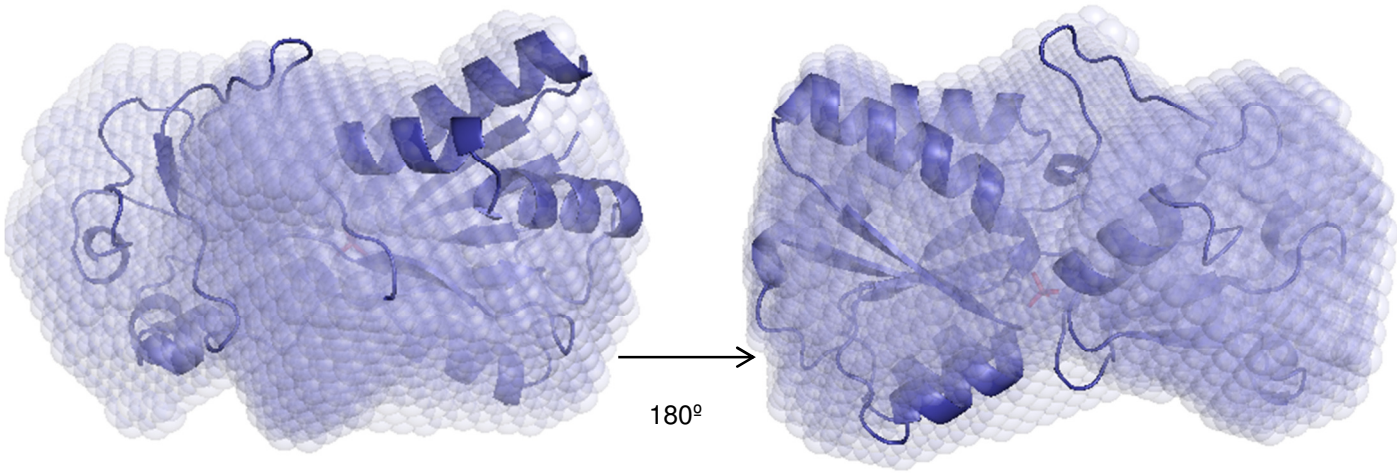


Figure 4.26.: Superposition of PDB 1ATG and the model resulted from SAXS experiment of ModA with molybdate. The WO_4 ligand is in red

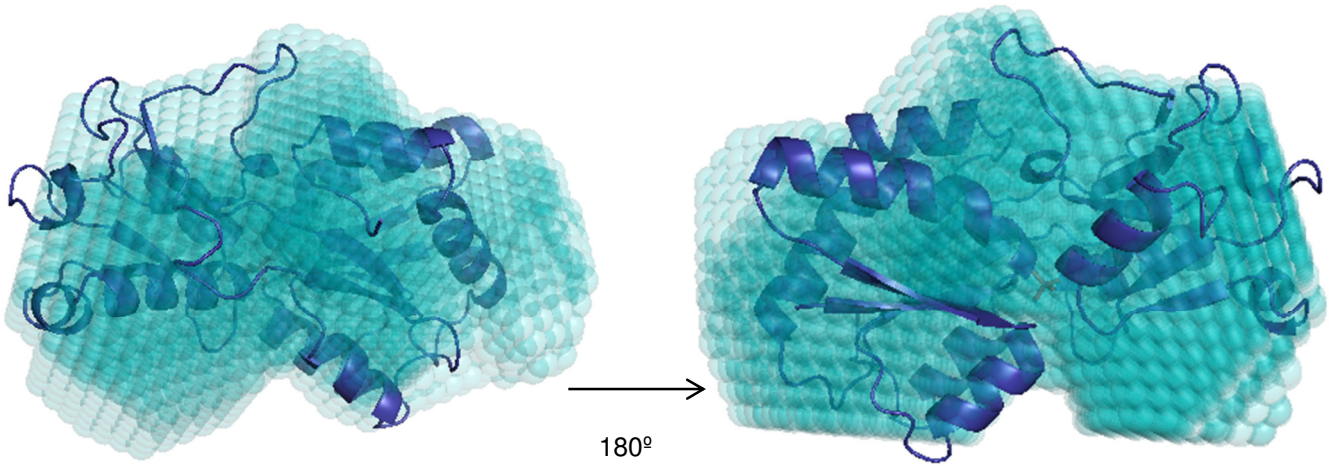


Figure 4.27.: Superposition of PDB 1ATG and the model resulted from SAXS experiment of ModA with tungstate. The WO_4 ligand is in red

Table 4.7 – Data analysis of ModA by SAXS experiment

SAXS analysis				
	Guinier		**GNOM**	
ModA	Rg (nm)	2.040 ± 0.017	Quality estimate	63.20%
	I_0	28.400±0.102	Reciprocal Space Rg/I_0	2.11 26.54
	Fidelity	0.989	Real Space Rg/I_0	2.11 26.54
	MM	28.52 KDa	Porod Volume	40.31
			D_{max}	6.73
ModA+Mo	Rg (nm)	1.910 ± 0.008	Quality estimate	72.60%
	I_0	27.700±0.077	Reciprocal Space Rg/I_0	1.93 26.05
	Fidelity	0.925	Real Space Rg/I_0	1.93 26.05
	MM	27.8 KDa	Porod Volume	41.23
			D_{max}	6.01
ModA+W	Rg	1.820 ± 0.016	Quality estimate	73.40%
	I_0	27.500±0.089	Reciprocal Space Rg/I_0	1.91 28.42
	Fidelity	0.987	Real Space Rg/I_0	1.91 28.42
	MM	27.6 KDa	Porod Volume	37.56
			D_{max}	5.77

Analyzing the apo and holoproteins scattering profiles (figure 4.28), some differences are found mainly when ModA is bound to tungstate suggesting that probably in solution ModA adopts a more close conformation when interacting with tungstate. All this small differences between the apo and the holoproteins will be reflected in the final shape/envelope of the protein calculated by DAMMIN as we can see in figures 4.25, 4.26 and 4.27.

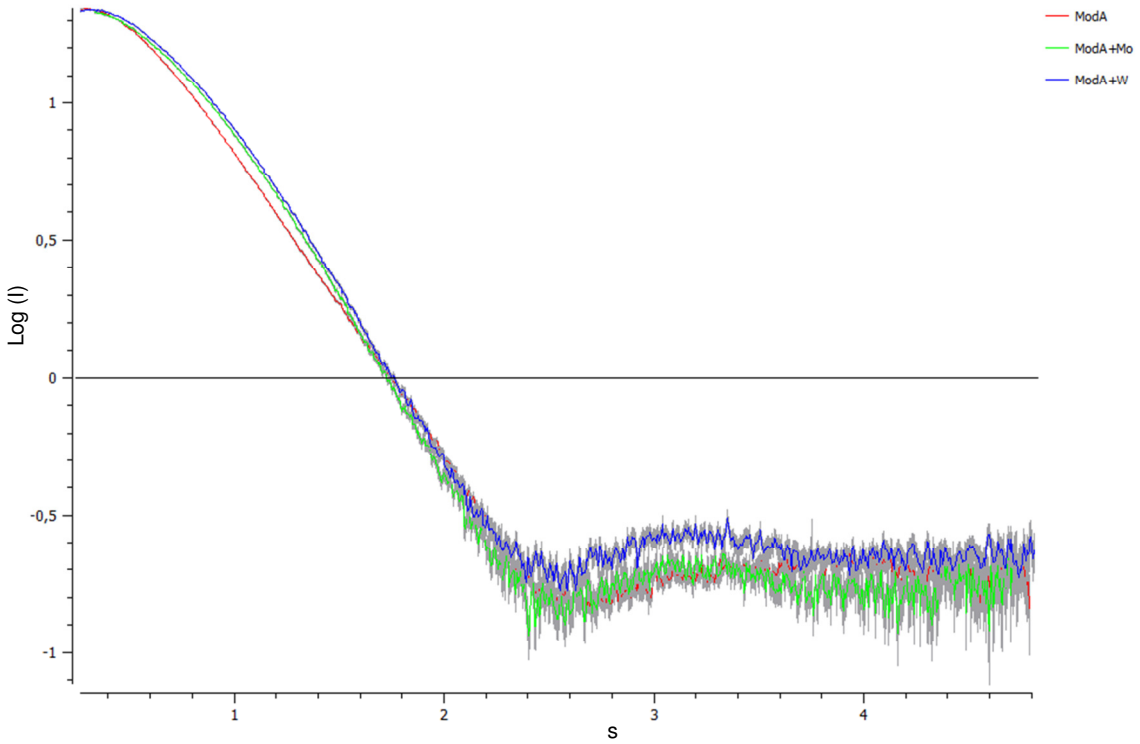


Figure 4.28.: Scattering profile of ModA (red), ModA+Mo (green) and ModA+W (blue) by SAXS.

These results are in agreement with the biochemical data here described where open and closed forms were speculated. Other techniques as NMR or circular dichrohism (CD) can be useful to detect conformational alterations and the mass spectroscopy (MS) to confirm the presence of the metal.

5. Final Remarks

Molybdenum and tungsten enzymes are extensively distributed in biology and can be found in all domains of life.[1, 3, 18, 29] These are crucial elements, which play key roles in numerous metabolic pathways, catalyzing essential reactions of the biochemical cycles of the more abundant elements on the earth. Usually, both Mo and W are part of enzymes that catalyze redox processes involving the transfer of an atom from the substrate to the metal ion or vice versa.[18] They catalyze some of the main reactions in the metabolism of carbon, nitrogen and sulfur by microorganisms, plants and animals, besides a small number of species that do not require molybdenum uses tungsten for the same purpose.

In order to be physiologically relevant and enable enzymes to perform its function, both Mo and W must be bound to organic cofactors, Moco or Wco, respectively. Despite Mo and W share several chemical characteristics, unspecific incorporation of the metal atom in the catalytic site of the proteins sometimes yields less active and even inactive enzymes[14, 18, 19]. The Mo and W cellular uptake is performed through specific transport systems (ModABC and TupABC) in which the component A (ModA/TupA) is proposed as the first selection gate to these elements. Which structural characteristics make the component A able to discriminate between these homologous elements still as an open question.

To give insights about the features that could contribute to the specificity of these proteins, in the present master thesis were studied both TupA and ModA from a functional and structural point of view.

Initially, it was mutated the positively charged R118 of TupA that was proposed as one of the responsible for the oxoanion coordination and to afford the high selectivity of the TupA proteins. Arginine 118 was replaced by three amino acids with different charges: lysine (positively charged, pI=10.76), glutamine (neutral, pI=3.22) and glutamic acid (negatively charged, pI=5.65). The expression conditions were optimized and a purification protocol was established for the three mutants. When this process was accomplished, it was performed a biochemical characterization which includes the experimental determination of the extinction coefficient, binding studies with metal oxoanions and determination of stoichiometry, association constant and enthalpy (n, K_D and ΔH) by ITC.

The value of the experimental extinction coefficients of mutants were in good agreement with the theoretically calculated values.

The native polyacrylamide gels showed that TupA and ModA can bind ligands, molybdate and tungstate, the mutation of R118 to R118K avoids the molybdate binding but not the tungstate coordination. The gels also showed that the mutation of R118 to neutral or negatively charged residue blocks ligand binding, which means that the arginine is an important amino acid involved in the coordination of metals.

The analysis of the urea-polyacrylamide gel suggests that there is a conformational change of the proteins upon ligand binding. In case of TupA and the TupA_R118K, the proteins probably adopt a more compact conformation when bound to tungstate but not for molybdate.

However, for ModA, the urea-polyacrylamide gel showed that this protein bind both ligands with very high and similar affinity, adopting also, a compact conformation.

After the competition assays, the ITC data showed an extremely low K_D value (6.30×10^{-3} nM) for TupA from *Da*G20 saturated with molybdate and titrated with tungstate. This is in agreement with the homologous protein WtpA from *P.furiosus.* Using the ITC technique was observed that WtpA bind both anions, tungstate ($K_D=17\times10^{-3}$ nM) and molybdate (K_D=11 nM). The apparent K_D for tungstate when the protein is saturated with molybdate was determined to be 15 nM [28]. A displacement titration of the tungstate-saturated TupA with molybdate was also done and no displacement occurred. This suggests that TupA binds very tightly both ligands, however, with higher affinity to tungstate than to molybdate. The same strategy could not be adopted for ModA since this protein binds with high affinity/specificity both ligands, meaning that the K_D determined has a high error associated. However the stoichiometry determined (n=1) is in good agreement to that expected for both proteins. Related to the mutated forms of TupA, only the TupA_R118K showed affinity for tungstate (K_D = 1.7 nM) with a very low binding for molybdate and a stoichiometry of approximately 1. The TupA_R118Q exhibits the lowest binding with tungstate (K_D = 72 nM), however with the stoichiometry below expected (n=0.647). Only the TupA_R118E does not exhibit any binding for molybdate or tungstate, which is in agreement with the native gel electrophoresis studies.

Previous to this master thesis work, the TupA was crystallized, and the tridimensional structure was solved and under refinement. After this, co-crystallization and soaking techniques were performed in order to understand the conformational behavior of the protein in presence of ligands, molybdate and tungstate. However, the several conditions tested so far not allowed crystallizing any of the protein-ligand complexes under study. Regarding the mutated forms of TupA, these were successfully crystallized, but their tridimensional structure could not be obtained because the crystals did not diffract properly. Crystallization screening of the mutant using new crystallization conditions and in the presence of ligands is under way. SAXS experiments are also planned to the near future where conformational changes of the protein in the presence and absence of the two different ligands can be addressed.

Regarding to the crystal structure of ModA, a complete dataset had been previously collected at 3.5 Å resolution. To improve the resolution optimization trials were carried out and a diffraction pattern with higher resolution (2.77 Å) was successfully obtained. The structure is currently being solved and superposition of some reported ModA structures show similarities in the envelopes and in some elements of secondary structure.[104-106] In addition, crystallization of

ModA in presence of both ligands is also currently being performed. A low resolution structural model for this protein was obtained using SAXS experiments. The scattering data obtained provides the envelope of the protein, which in fact is very similar to the periplasmic protein ModA from *Azotobacter vinelandii*, also responsible for the transport of molybdate.[105]

Finally, in order to understand the differences/similarities between both proteins (TupA and ModA) from the two transporter systems, the structure solved of TupA from *Da*G20 was superposed with the ModA from *Clostridium difficile*.[106] Here, the envelope of the two proteins are also analogous, however the secondary and tertiary structures show some differences, as expected.

The preliminary functional and structural characterization of the component A of both ABC transport systems present in this work contributes to the deeper understanding of these kind of systems. Particularly regarding to the TupA protein, it was not only proved the relevance of the arginine 118 of TupA on the oxoanion coordination but also was obtained a mutant (TupA_R118K) with higher specificity towards tungstate.

Besides the crystallization experiments that are under way, other techniques for structural and biochemical characterization could be used to better understand this class of transporters (namely circular dichroism or Nuclear Magnetic Resonance). For now, establishment of a metal specific variant of TupA is already quite an achievement that opens the way for future biological applications.

6. References

1. Smart, J. P., Cliff, M. J., and Kelly, D. J. (2009) A role for tungsten in the biology of Campylobacter jejuni: tungstate stimulates formate dehydrogenase activity and is transported via an ultra-high affinity ABC system distinct from the molybdate transporter, *Mol Microbiol 74*, 742-757.
2. McMaster, J., Enemarkt, J. H.,. (1998) The active sites of molybdenum and tungsten containing enzymes, pp 201-207, Current Opinion in Chemical Biology.
3. Aguilar-Barajas, E., Díaz-Pérez, C., Ramírez-Díaz, M. I., Riveros-Rosas, H., and Cervantes, C. (2011) Bacterial transport of sulfate, molybdate, and related oxyanions, *Biometals 24*, 687-707.
4. Hille, R. (2002) Molybdenum and tungsten in biology, pp 360-367, TRENDS in Biochemical Sciences.
5. Bevers, L. E., Hagedoorn, P., and Hagen, W. R. (2009) The bioinorganic chemistry of tungsten, pp 269-290, Coordination Chemistry Reviews.
6. Mendel, R. R., and Bittner, F. (2006) Cell biology of molybdenum, *Biochim Biophys Acta 1763*, 621-635.
7. Mendel, R. R., and Kruse, T. (2012) Cell biology of molybdenum in plants and humans, *Biochim Biophys Acta 1823*, 1568-1579.
8. Zhang, Y., Rump, S., and Gladyshev, V. N. (2011) Comparative Genomics and Evolution of Molybdenum Utilization, *Coord Chem Rev 255*, 1206-1217.
9. Romão, M. J. (2009) Molybdenum and tungsten enzymes: a crystallographic and mechanistic overview, *Dalton Trans*, 4053-4068.
10. Hänzelmann, P., and Schindelin, H. (2004) Crystal structure of the S-adenosylmethionine-dependent enzyme MoaA and its implications for molybdenum cofactor deficiency in humans, *Proc Natl Acad Sci U S A 101*, 12870-12875.
11. Schwarz, G., and Mendel, R. R. (2006) Molybdenum cofactor biosynthesis and molybdenum enzymes, *Annu Rev Plant Biol 57*, 623-647.
12. Llamas, A., Otte, T., Multhaup, G., Mendel, R. R., and Schwarz, G. (2006) The Mechanism of nucleotide-assisted molybdenum insertion into molybdopterin. A novel route toward metal cofactor assembly, *J Biol Chem 281*, 18343-18350.
13. Iobbi-Nivol, C., and Leimkühler, S. (2013) Molybdenum enzymes, their maturation and molybdenum cofactor biosynthesis in Escherichia coli, *Biochim Biophys Acta 1827*, 1086-1101.
14. Majumdar, A., and Sarkar, S. (2011) Bioinorganic chemistry of molybdenum and tungsten enzymes: A structural–functional modeling approach pp 1039-1054, Coordination Chemistry Reviews.
15. Mendel, R., and Scharz, G. (2011) Molybdenum cofactor biosynthesis in plants and humans, pp 1145-1158, Coordination Chemistry Reviews.
16. Schwarz, G., Mendel, R. R., and Ribbe, M. W. (2009) Molybdenum cofactors, enzymes and pathways, *Nature 460*, 839-847.

17. Bevers, L. E., Hagedoorn, P. L., Santamaria-Araujo, J. A., Magalon, A., Hagen, W. R., and Schwarz, G. (2008) Function of MoaB proteins in the biosynthesis of the molybdenum and tungsten cofactors, *Biochemistry 47*, 949-956.
18. Gonzalez, P. J., Rivas, Maria G., Mota, C., Brondino, C.D, Moura, Isabel, Moura, José J. (2013) Periplasmic nitrate reductases and formate dehydrogenases: Biological control of the chemical properties of Mo and W for fine tuning of reactivity, substrate specificity and metabolic role, *Elsevier 257*, 315-331.
19. Regulski, E. E., Moy, R. H., Weinberg, Z., Barrick, J. E., Yao, Z., Ruzzo, W. L., and Breaker, R. R. (2008) A widespread riboswitch candidate that controls bacterial genes involved in molybdenum cofactor and tungsten cofactor metabolism, *Mol Microbiol 68*, 918-932.
20. Havarushka, N., Fischer-Schrader, K., Lamkemeyer, T., and Schwarz, G. (2014) Structural Basis of Thermal Stability of the Tungsten Cofactor Synthesis Protein MoaB from Pyrococcus furiosus, *PLoS One 9*, e86030.
21. Stevenson, C. E., Sargent, F., Buchanan, G., Palmer, T., and Lawson, D. M. (2000) Crystal structure of the molybdenum cofactor biosynthesis protein MobA from Escherichia coli at near-atomic resolution, *Structure 8*, 1115-1125.
22. Moura, J. J., Brondino, C. D., Trincão, J., and Romão, M. J. (2004) Mo and W bis-MGD enzymes: nitrate reductases and formate dehydrogenases, *J Biol Inorg Chem 9*, 791-799.
23. da Silva, S. M., Pimentel, C., Valente, F. M., Rodrigues-Pousada, C., and Pereira, I. A. (2011) Tungsten and molybdenum regulation of formate dehydrogenase expression in Desulfovibrio vulgaris Hildenborough, *J Bacteriol 193*, 2909-2916.
24. L'vov, N. P., Nosikov, A. N., and Antipov, A. N. (2002) Tungsten-containing enzymes, *Biochemistry (Mosc) 67*, 196-200.
25. Hollenstein, K., Frei, D. C., and Locher, K. P. (2007) Structure of an ABC transporter in complex with its binding protein, *Nature 446*, 213-216.
26. Hagen, W. R. (2011) Cellular uptake of molybdenum and tungsten, pp 1117-1128, Coor. Chemistry Reviews.
27. Balan, A., Santacruz, C. P., Moutran, A., Ferreira, R. C., Medrano, F. J., Pérez, C. A., Ramos, C. H., and Ferreira, L. C. (2006) The molybdate-binding protein (ModA) of the plant pathogen Xanthomonas axonopodis pv. citri, *Protein Expr Purif 50*, 215-222.
28. Bevers, L. E., Hagedoorn, P. L., Krijger, G. C., and Hagen, W. R. (2006) Tungsten transport protein A (WtpA) in Pyrococcus furiosus: the first member of a new class of tungstate and molybdate transporters, *J Bacteriol 188*, 6498-6505.
29. Bevers, L. E., Schwarz, G., and Hagen, W. R. (2011) A molecular basis for tungstate selectivity in prokaryotic ABC transport systems, *J Bacteriol 193*, 4999-5001.
30. Otrelo-Cardoso, A. R., Nair, R. R., Correia, M. A. S., Rivas, M. G., and Santos-Silva, T. (2014) TupA: A Tungstate Binding Protein in the Periplasm of *Desulfovibrio alaskensis* G20, pp 1-17, Int. J. Mol. Sci.

31. Schneider, E., Eckey, V., Weidlich, D., Wiesemann, N., Vahedi-Faridi, A., Thaben, P., and Saenger, W. (2012) Receptor-transporter interactions of canonical ATP-binding cassette import systems in prokaryotes, *Eur J Cell Biol 91*, 311-317.
32. Weidlich, D., Wiesemann, N., Heuveling, J., Wardelmann, K., Landmesser, H., Sani, K. B., Worth, C. L., Preissner, R., and Schneider, E. (2013) Residues of a proposed gate region in type I ATP-binding cassette import systems are crucial for function as revealed by mutational analysis, *Biochim Biophys Acta 1828*, 2164-2172.
33. Yang, D. C., Peters, N. T., Parzych, K. R., Uehara, T., Markovski, M., and Bernhardt, T. G. (2011) An ATP-binding cassette transporter-like complex governs cell-wall hydrolysis at the bacterial cytokinetic ring, *Proc Natl Acad Sci U S A 108*, E1052-1060.
34. Eikim. (2009) Proposed transport mechanism for ABC importers, (ABC_importer.jpg, Ed.), http://en.wikibooks.org/wiki/Structural_Biochemistry/Membrane_Proteins/ATP-Binding_Cassette_Proteins.
35. Grunden, A. M., and Shanmugam, K. T. (1997) Molybdate transport and regulation in bacteria, *Arch Microbiol 168*, 345-354.
36. Leimkühler, S., Wuebbens, M. M., and Rajagopalan, K. V. (2011) The History of the Discovery of the Molybdenum Cofactor and Novel Aspects of its Biosynthesis in Bacteria, *Coord Chem Rev 255*, 1129-1144.
37. Self, W. T., Grunden, A. M., Hasona, A., and Shanmugam, K. T. (1999) Transcriptional regulation of molybdoenzyme synthesis in Escherichia coli in response to molybdenum: ModE-molybdate, a repressor of the modABCD (molybdate transport) operon is a secondary transcriptional activator for the hyc and nar operons, *Microbiology 145 (Pt 1)*, 41-55.
38. Schwarz, G., Hagedoorn, P.-L., and Fischer, K. (2007) Molybdate and Tungstate: Uptake, Homeostasis, Cofactors, and Enzymes *Microbiol Monogr 6*, 423-445.
39. Wagner, U. G., Stupperich, E., and Kratky, C. (2000) Structure of the molybdate/tungstate binding protein mop from Sporomusa ovata, *Structure 8*, 1127-1136.
40. Muyzer, G., and Stams, A. J. (2008) The ecology and biotechnology of sulphate-reducing bacteria, *Nat Rev Microbiol 6*, 441-454.
41. Hansen, T. A. (1994) Metabolism of sulfate-reducing prokaryotes, *Antonie Van Leeuwenhoek 66*, 165-185.
42. Kazakov, A. E., Rajeev, L., Luning, E. G., Zane, G. M., Siddartha, K., Rodionov, D. A., Dubchak, I., Arkin, A. P., Wall, J. D., Mukhopadhyay, A., and Novichkov, P. S. (2013) New family of tungstate-responsive transcriptional regulators in sulfate-reducing bacteria, *J Bacteriol 195*, 4466-4475.
43. Mota, C. S., Valette, O., González, P. J., Brondino, C. D., Moura, J. J., Moura, I., Dolla, A., and Rivas, M. G. (2011) Effects of molybdate and tungstate on expression levels and biochemical characteristics of formate dehydrogenases produced by Desulfovibrio alaskensis NCIMB 13491, *J Bacteriol 193*, 2917-2923.

44. (last access 14.07.2014) Looking at Structures: Methods for Determining Atomic Structures, http://www.rcsb.org/pdb/101/static101.do?p=education_discussion/Looking-at-Structures/methods.html.
45. Papageorgiou, A. C., and Mattsson, J. (2014) Protein structure validation and analysis with X-ray crystallography, *Methods Mol Biol 1129*, 397-421.
46. Smyth, M. S., and Martin, J. H. (2000) x ray crystallography, *Mol Pathol 53*, 8-14.
47. *Lectures, N.* Wilhelm Conrad Röntgen - Biographical, http://www.nobelprize.org/nobel_prizes/physics/laureates/1901/rontgen-bio.html.
48. Wiener, M. (last access 14.07.2014) Two Lectures in X-ray Crystallography, http://www.its.caltech.edu/~heathgrp/Courses/Lectures_2014/wiener_503_xray.pdf.
49. NASA. (last access 15.07.2014) Electromagnetic Spectrum Diagram, (EM_Spectrum3-new, Ed.), http://mynasadata.larc.nasa.gov/science-processes/electromagnetic-diagram/.
50. C., L., MacLean, W., Beauchemin, S., and Rasmussen, P. E. (2011) Application of Synchrotron X-ray Techniques for the Determination of Metal Speciation in (House) Dust Particles, In *Urban Airborne Particulate Matter*, pp 193-216, Springer Berlin Heidelberg.
51. (last access 15.07.2014) How is synchrotron light created?, (create_example1, Ed.), https://www.synchrotron.org.au/synchrotron-science/how-is-synchrotron-light-created.
52. Chayen, N. E., and Saridakis, E. (2008) Protein crystallization: from purified protein to diffraction-quality crystal, *Nat Methods 5*, 147-153.
53. Romão, M. J. (1996) Cristalografia de Proteínas: metodologias e aplicações em Bioquímica, pp 18-36, Boletim de Biotecnologia.
54. Attwood, D. (2007) *Soft X-Rays and Extreme Ultraviolet Radiation: Principles and Applications*, Cambridge University Press.
55. Woolfson, M. M. (1997) *An Introduction to X-ray Crystallography*, Second ed., Cambridge University Press.
56. Gale, R. (2006) *Crystallography made crystal clear - A guide for users of macromolecular models*, Third ed., Academic Press.
57. Rupp, B. (2009) *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*, First ed., Garland Science.
58. Leunissen, M. (2001) An essay on several aspects of protein crystallization research, In *Department of solid state chemistry*, p 44.
59. Chayen, N. E. (2004) Turning protein crystallisation from an art into a science, *Curr Opin Struct Biol 14*, 577-583.
60. Dessau, M. A., and Modis, Y. (2011) Protein crystallization for X-ray crystallography, *J Vis Exp*.
61. (last access 21.07.2014), http://www.xray.bioc.cam.ac.uk/xray_resources/whitepapers/xtal-in-action/img11.gif.

62. (last access 21.07.2014) Introduction to Protein Crystallography: X-ray Diffraction and Data Collection, http://www.proteinstructures.com/Experimental/Experimental/protein-crystallography.html.
63. (last access 21.07.2014) DNA, http://course1.winona.edu/kbates/bio241/images/figure-11-04.jpg.
64. (last access 21.07.2014) Guide to Understanding X-ray Crystallography, http://www.chem.ucla.edu/harding/ec_tutorials/tutorial73.pdf.
65. (last access 21.07.2014) Diffraction Pattern, http://departments.colgate.edu/chemistry/images/rowlett_diffpattern.jpg.
66. Piburn, G., and Barrow, A. R. (2013) An introduction to energy dispersive x-ray spectroscopy, pp 90-98, *Connexions/Rice University, Houston*, *Physical methods in chemistry and nano science*.
67. Terwilliger, T. C. (2004) Using prime-and-switch phasing to reduce model bias in molecular replacement, *Acta Crystallogr D Biol Crystallogr 60*, 2144-2149.
68. Taylor, G. (2003) The phase problem, *Acta Crystallogr D Biol Crystallogr 59*, 1881-1890.
69. Wlodawer, A., Minor, W., Dauter, Z., and Jaskolski, M. (2008) Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures, *FEBS J 275*, 1-21.
70. Skou, S., Gillilan, R. E., and Ando, N. (2014) Synchrotron-based small-angle X-ray scattering of proteins in solution, *Nat Protoc 9*, 1727-1739.
71. Petoukhov, M. V., Billas, I. M., Takacs, M., Graewert, M. A., Moras, D., and Svergun, D. I. (2013) Reconstruction of quaternary structure from X-ray scattering by equilibrium mixtures of biological macromolecules, *Biochemistry 52*, 6844-6855.
72. Geerlof, A., Brown, J., Coutard, B., Egloff, M. P., Enguita, F. J., Fogg, M. J., Gilbert, R. J., Groves, M. R., Haouz, A., Nettleship, J. E., Nordlund, P., Owens, R. J., Ruff, M., Sainsbury, S., Svergun, D. I., and Wilmanns, M. (2006) The impact of protein characterization in structural proteomics, *Acta Crystallogr D Biol Crystallogr 62*, 1125-1136.
73. Mertens, H. D., and Svergun, D. I. (2010) Structural characterization of proteins and complexes using small-angle X-ray solution scattering, *J Struct Biol 172*, 128-141.
74. Svergun, D. I., and Koch, M. H. (2002) Advances in structure analysis using small-angle scattering in solution, *Curr Opin Struct Biol 12*, 654-660.
75. Leavitt, S., and Freire, E. (2001) Direct measurement of protein binding energetics by isothermal titration calorimetry, *Curr Opin Struct Biol 11*, 560-566.
76. (last access 22.07.2014) The ITC experiment, http://www.cif.iastate.edu/sites/default/files/uploads/Other_Inst/ITC/ITC%20Explained.pdf.
77. (last access 22.07.2014) Isothermal Titration Calorimetry, http://structbio.vanderbilt.edu/wetlab/designing.itc.v7.expts.pdf.

78. Wiseman, T., Williston, S., Brandts, J. F., and Lin, L. N. (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter, pp 131-135, Anal Biochem.
79. Velazquez-Campoy, A., and Freire, E. (2006) Isothermal titration calorimetry to determine association constants for high-affinity ligands, *Nat Protoc 1*, 186-191.
80. Todd, M. J., and Gomez, J. (2001) Enzyme kinetics determined using calorimetry: a general assay for enzyme activity?, *Anal Biochem 296*, 179-187.
81. Altschul, S. F., and Lipman, D. J. (1990) Protein database searches for multiple alignments, pp 5509-5513, *Proc. Nat. Acad. Sci. USA*.
82. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res 22*, 4673-4680.
83. Sigma-Aldrich. (2013) Bradford Reagent: Technical Bulletin, sigma-aldrich.com.
84. (last access 3.09.2014) JustBio: Bioiformatics at the tip of your fingers, justbio.com.
85. Rech, S., Wolin, C., and Gunsalus, R. P. (1996) Properties of the periplasmic ModA molybdate-binding protein of Escherichia coli, *J Biol Chem 271*, 2557-2562.
86. Mehtab, S., Gonçalves, G., Roy, S., Tomaz, A. I., Santos-Silva, T., Santos, M. F., Romão, M. J., Jakusch, T., Kiss, T., and Pessoa, J. C. (2013) Interaction of vanadium(IV) with human serum apo-transferrin, *J Inorg Biochem 121*, 187-195.
87. Jancarik, J., and Kim, S.-H. (1991) Sparse matrix sampling: A screening method for crystallization of proteins, pp 409-411, *J. Appl. Cryst*.
88. EMBL. (last access 9.09.2014) Data analysis software ATSAS 2.5.2, http://www.embl-hamburg.de/biosaxs/software.html.
89. (last access 5.09.2014) PyMOL, http://www.pymol.org/.
90. Hu, Y., Rech, S., Gunsalus, R. P., and Rees, D. C. (1997) Crystal structure of the molybdate binding protein ModA, *Nat Struct Biol 4*, 703-707.
91. Saboury, A. A. (2006) A review on the ligand binding studies by isothermal titration calorimetry, pp 1-21, Journal of the Iranian Chemical Society.
92. Callis, G. E., and Wentworth, R. A. (1977) Tungsten vs. Molybdenum in models for biological systems, *Bioinorg Chem 7*, 57-70.
93. Leslie, A. G. W. (1992) Recent changes to the MOSFLM package for processing film and image plate data, Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography.
94. University, Q. s. (last access 28.08.2014) *Protein Crystallography*, Protein, http://pldserver1.biochem.queensu.ca/~rlc/work/teaching/definitions.shtml.
95. (last access 28.08.2014) Stamp Collecting Project, http://www.stampcollectingproject.org/crystallography/diffraction/r-symm-and-r-merge.
96. Diederichs, K., and Karplus, P. A. (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography, *Nat Struct Biol 4*, 269-275.

97. Matthews, B. W. (1968) Solvent content of protein crystals, *J Mol Biol 33*, 491-497.

98. Expasy. (last access 10.10.2014) LALIGN, http://www.ch.embnet.org/software/LALIGN_form.html.

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

100. Nair, R. R., Otrelo-Cardoso, A. R., Cordeiro, R. S. C., Correia, M. A. S., Santos-Silva, T., and Rivas, M. G. Insights of the Tungsten binding protein TupA: structural and biochemical characterization, Metallomics.

101. Bernadó, P., and Svergun, D. I. (2012) Analysis of intrinsically disordered proteins by small-angle X-ray scattering, *Methods Mol Biol 896*, 107-122.

102. Blanchet, C. E., and Svergun, D. I. (2013) Small-angle X-ray scattering on biological macromolecules and nanocomposites in solution, *Annu Rev Phys Chem 64*, 37-54.

103. Synchrotron, T. E. (last access 9.09.2014) SAXS data analysis, http://www.esrf.eu/UsersAndScience/Experiments/MX/About_our_beamlines/BM29/Users/data-collection-information/saxs-data-analysis.

104. Aryal, B. P., Brugarolas, P., and He, C. (2012) Binding of ReO4(-) with an engineered MoO4(2-)-binding protein: towards a new approach in radiopharmaceutical applications, *J Biol Inorg Chem 17*, 97-106.

105 Lawson, D.M., Williams, C.E., Mitchenall, L.A. and Pau, R.N. (1998), Ligand size is a major determinant of specificity in periplasmic oxyanion-binding proteins: the 1.2 A resolution crystal structure of Azotobacter vinelandii ModA, *Structure*, 6: 1529,1539;

106 Maltseva, N., Kim, Y., Grimshaw, S., Anderson, W.F. and Joachimiak, A., Substrate binding domain of putative molybdenum ABC transporter from Clostridium difficile, http://rcbs.org/pdb/explore/explore.do?structureId=4KD5

7. Appendixes

Appendix I

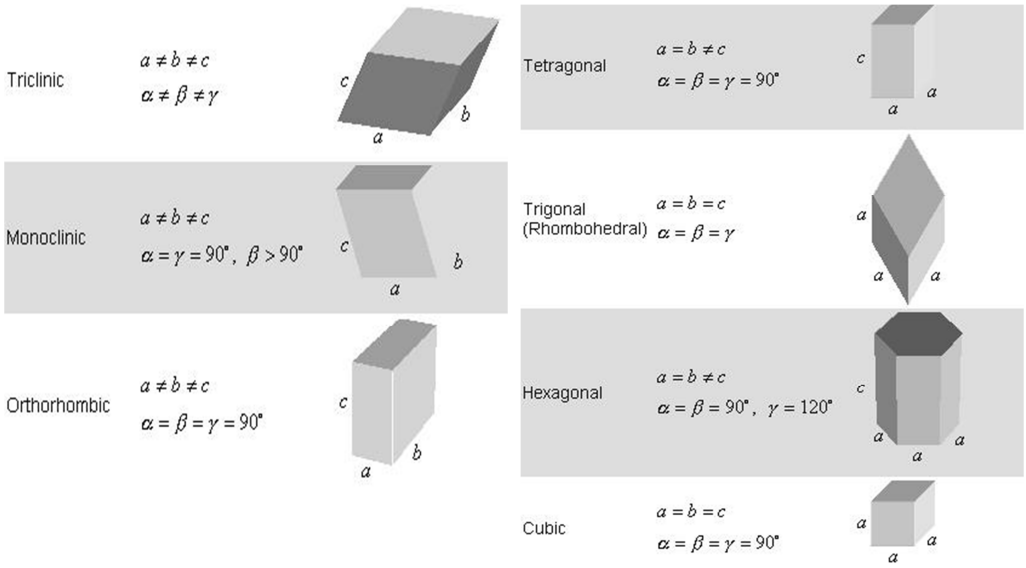


Figure 7.1: Representation of crystalline systems

Appendix II

Table 7.1: LB media composition for 1 L for the growth of *E.coli* which is prepared with distilled water and it is submitted to the autoclave at 120º for 20 minutes

LB media	
NaCl (Scharlau)	5 g
Tryptone (Bacto™)	10 g
Yeast Extract (Cultimed)	5 g

Table 7.2: Autoinduction media composition for 100 mL for the growth of *E.coli* which is prepared with distilled water and it is submitted to the autoclave at 120º for 20 minutes

Autoinduction media	
$MgSO_4$ (Sigma-Aldrich)	100 µL of 1M
50x52 Solution*	2 mL
20xNPS Solution**	5 mL
LB media	93 mL

*50x52 Solution STOCK (no autoclave)	
Glycerol (Scharlau)	25% (v/v)
Glucose (Sigma-Aldrich)	2.5% (v/v)
α-Lactose (Sigma-Aldrich)	10% (v/v)

**20xNPS Solution STOCK	
$(NH_4)_2SO_4$ (Panreac)	25 mM
KH_2PO_4 (Sigma-Aldrich) Na_2HPO_4 (Sigma-Aldrich)	50 mM

Appendix III

Table 7.3.: Composition of the solutions employed

Stock Solutions	Concentration Gel (mL)	Running Gel (mL)
Solution I	-	0.166
Solution II	0.9	-
Solution III	0.45	1.49
H_2O	2.15	3.34
PSA 10% (w/v) (µL)	27	25
TEMED (µL)	4	2.5

Table 7.4.: Composition of the solutions employed

Solution	Composition
I	Tris-HCl 1 M pH 8.8
II	Tris-HCl 0.5 M pH 6.8
III	Acrylamide/Bis Solution, 40:30%

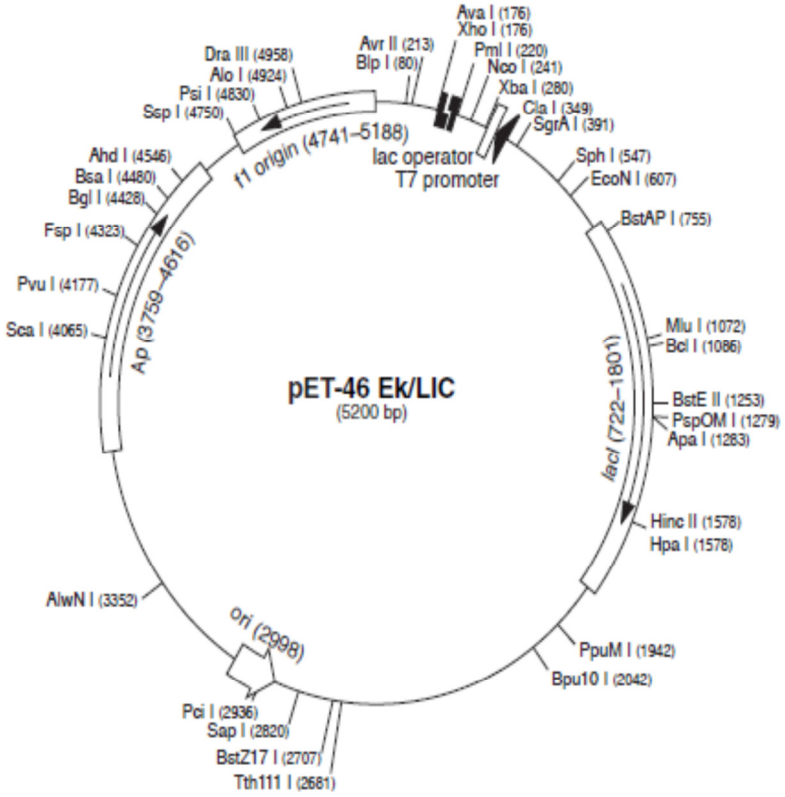
Table 7.5.: Composition of the sample buffer for 1X.

	Concentration
Tris-HCl pH 6.8	125 mM
Glycerol	10% (v/v)
Bromophenol Blue	0.02% (w/v)

Table 7.6.: Composition of the loading buffer for 1X.

	Concentration
Tris-HCl pH 8.3	25 mM
Glycine	192 mM
SDS *	0.1%

Appendix IV



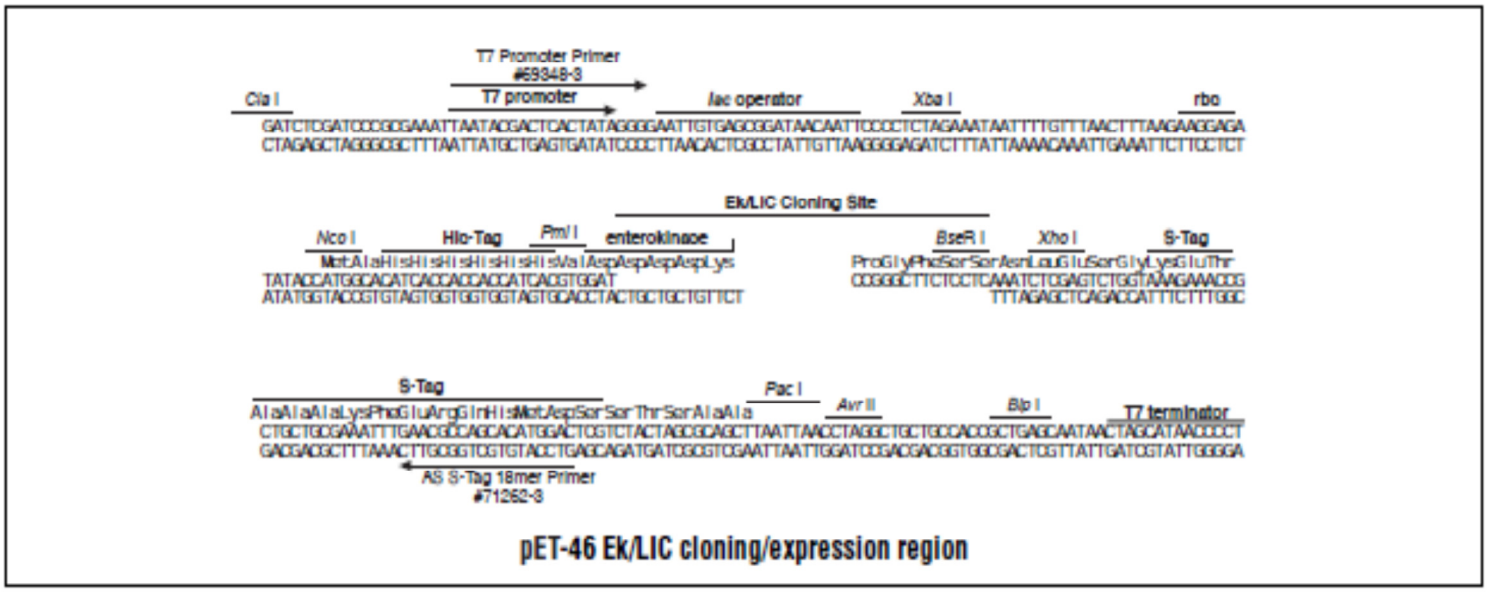


Figure 7.2.: pET-46 Ek/LIC vector map and sequence previously used to clone the *tupA* gene and consequently used for the mutagenesis procedure

Appendix V

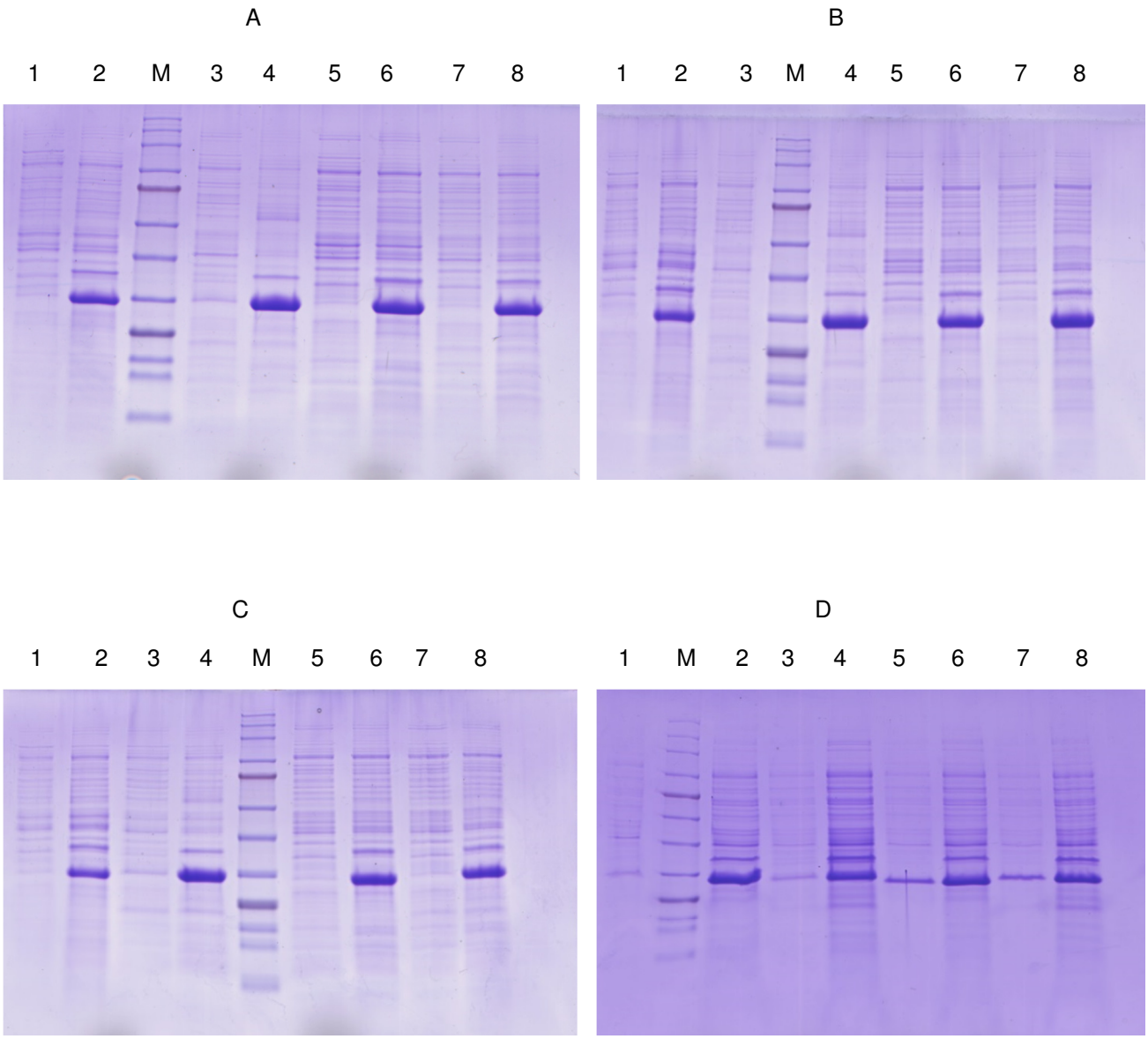


Figure 7.3.: Expression tests in E.coli with 100 mL of media of each mutant of TupA. **A**: TupA_R118K ; M: Marker, (1-4): Induction 30ºC, (1): BL21 soluble fraction, (2): BL21 insoluble fraction, (3): Ros soluble fraction, (4): Ros insoluble fraction; (5-8): Autoinduction 30ºC, (5): BL21 soluble fraction, (6): BL21 insoluble fraction, (7): Ros soluble fraction, (8): Ros insoluble fraction; **B**: TupA_R118E; M: Marker, (1-4): Induction 30ºC, (1): BL21 soluble fraction, (2): BL21 insoluble fraction, (3): Ros soluble fraction, (4): Ros insoluble fraction; (5-8): Autoinduction 30ºC, (5): BL21 soluble fraction, (6): BL21 insoluble fraction, (7): Ros soluble fraction, (8): Ros insoluble fraction; **C**: TupA_R118Q M: Marker, (1-4): Induction 30ºC, (1): BL21 soluble fraction, (2): BL21 insoluble fraction, (3): Ros soluble fraction, (4): Ros insoluble fraction; (5-8): Autoinduction 30ºC, (5): BL21 soluble fraction, (6): BL21 insoluble fraction, (7): Ros soluble fraction, (8): Ros insoluble fraction; **D**: Induction 19ºC, M: Marker, (1-4): TupA_R118E, 1): BL21 soluble fraction, (2): BL21 insoluble fraction, (3): Ros soluble fraction, (4): Ros insoluble fraction, (5-8): TupA_R118Q, (5): BL21 soluble fraction, (6): BL21 insoluble fraction, (7): Ros soluble fraction, (8): Ros insoluble fraction.

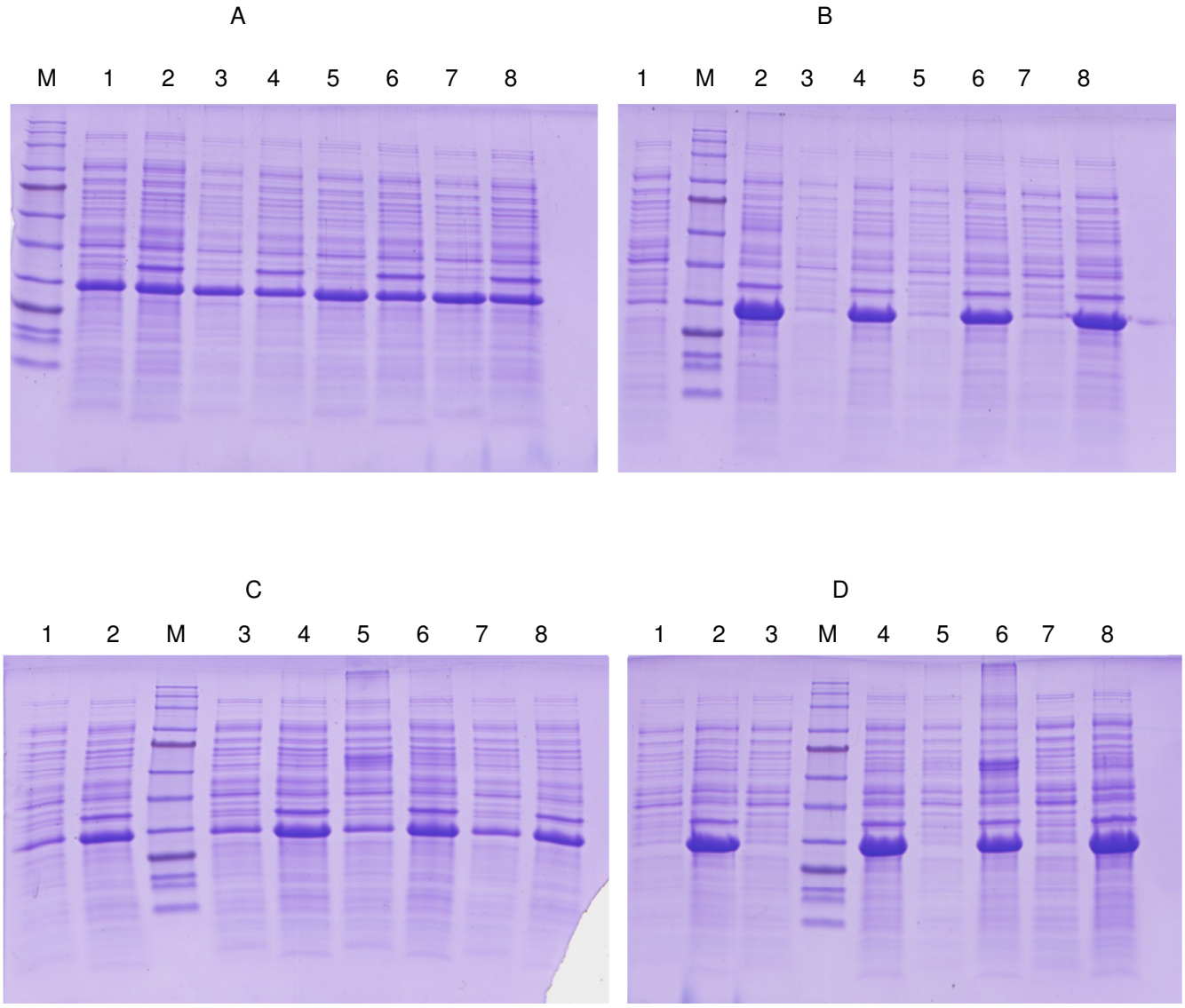


Figure 7.4.: Expression tests in E.coli with 100 mL of media of each mutant of TupA. **A**: TupA_R118K, Induction 19ºC and Tuner cells ; M: Marker, (1-2): 100 µM of IPTG, (1): soluble fraction, (2): insoluble fraction, (3-4): 300 µM of IPTG, (3): soluble fraction, (4): insoluble fraction; (5-6): 500 µM of IPTG (5): soluble fraction, (6): insoluble fraction, (7-8): 800 µM of IPTG, (7): soluble fraction, (8): insoluble fraction; **B**: TupA_R118K, Induction 30ºC and Tuner cells; M: Marker, (1-2): 100 µM of IPTG, (1): soluble fraction, (2): insoluble fraction, (3-4): 300 µM of IPTG, (3): soluble fraction, (4): insoluble fraction; (5-6): 500 µM of IPTG (5): soluble fraction, (6): insoluble fraction, (7-8): 800 µM of IPTG, (7): soluble fraction, (8): insoluble fraction, **C**: TupA_R118E, Induction 19ºC and Tuner cells, M: Marker, (1-2): 100 µM of IPTG, (1): soluble fraction, (2): insoluble fraction, (3-4): 300 µM of IPTG, (3): soluble fraction, (4): insoluble fraction; (5-6): 500 µM of IPTG (5): soluble fraction, (6): insoluble fraction, (7-8): 800 µM of IPTG, (7): soluble fraction, (8): insoluble fraction; **D**: TupA_R118E, Induction 30ºC, M: Marker, (1-2): 100 µM of IPTG, (1): soluble fraction, (2): insoluble fraction, (3-4): 300 µM of IPTG, (3): soluble fraction, (4): insoluble fraction; (5-6): 500 µM of IPTG (5): soluble fraction, (6): insoluble fraction, (7-8): 800 µM of IPTG, (7): soluble fraction, (8): insoluble fraction.

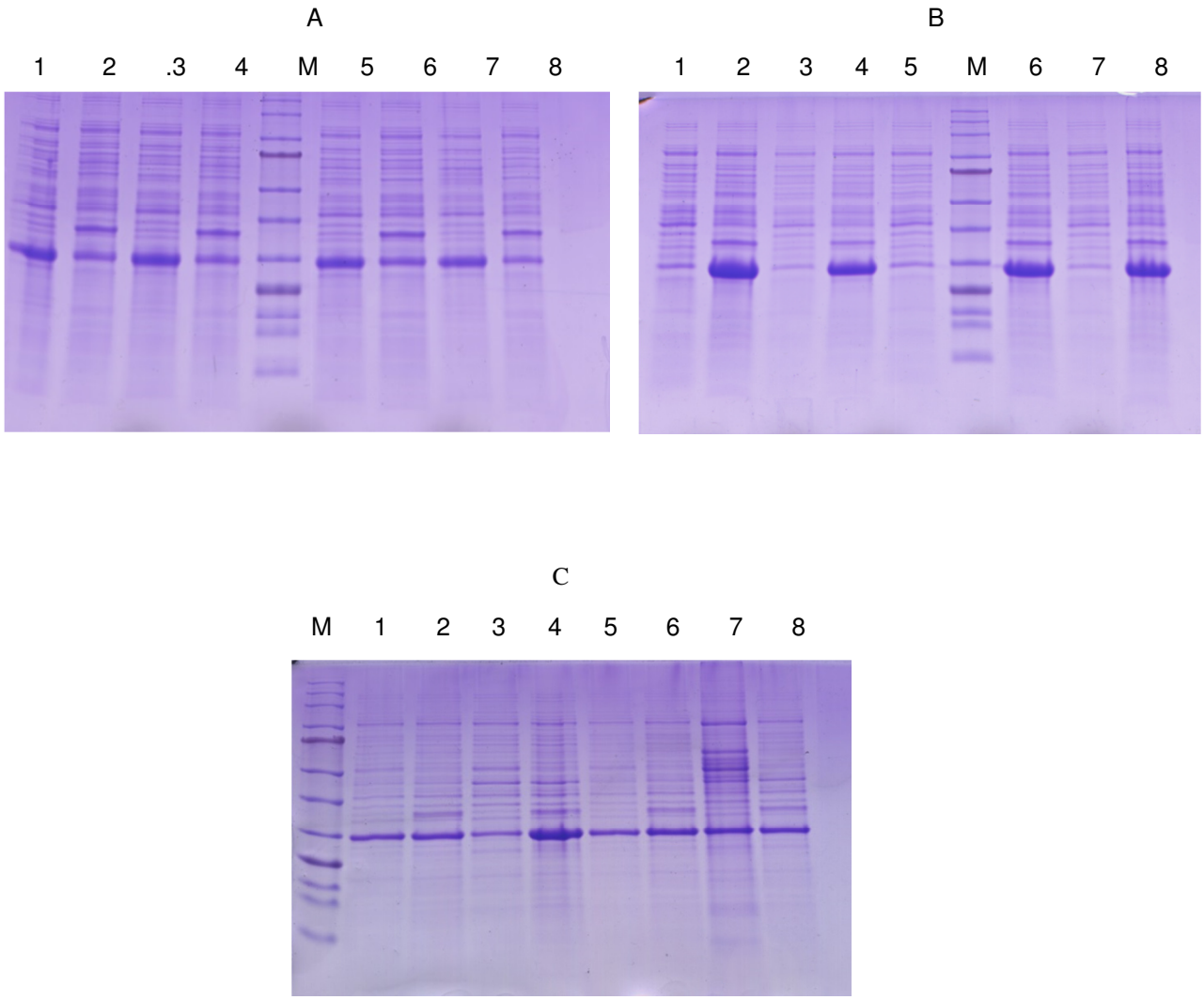


Figure 7.5.: Expression tests in E.coli with 100 mL of media of each mutant of TupA. **A**: TupA_R118Q, Induction 19ºC and Tuner cells ; M: Marker, (1-2): 100 µM of IPTG, (1): soluble fraction, (2): insoluble fraction, (3-4): 300 µM of IPTG, (3): soluble fraction, (4): insoluble fraction; (5-6): 500 µM of IPTG (5): soluble fraction, (6): insoluble fraction, (7-8): 800 µM of IPTG, (7): soluble fraction, (8): insoluble fraction; **B:** TupA_R118Q, Induction 30ºC and Tuner cells; M: Marker, (1-2): 100 µM of IPTG, (1): soluble fraction, (2): insoluble fraction, (3-4): 300 µM of IPTG, (3): soluble fraction, (4): insoluble fraction; (5-6): 500 µM of IPTG (5): soluble fraction, (6): insoluble fraction, (7-8): 800 µM of IPTG, (7): soluble fraction, (8): insoluble fraction, **C:** TupA_R118E and TupA_R118Q, Origami cells, M: Marker, (1-2): TupA_R118E, Induction 19ºC, (1): soluble fraction, (2): insoluble fraction, (3-4): TupA_R118E, Induction 30ºC, (3): soluble fraction, (4): insoluble fraction; (5-6): TupA_R118Q, Induction 19ºC, (5): soluble fraction, (6): insoluble fraction, (7-8): TupA_R118Q, Induction 30ºC, (7): soluble fraction, (8): insoluble fraction.

Appendix VI

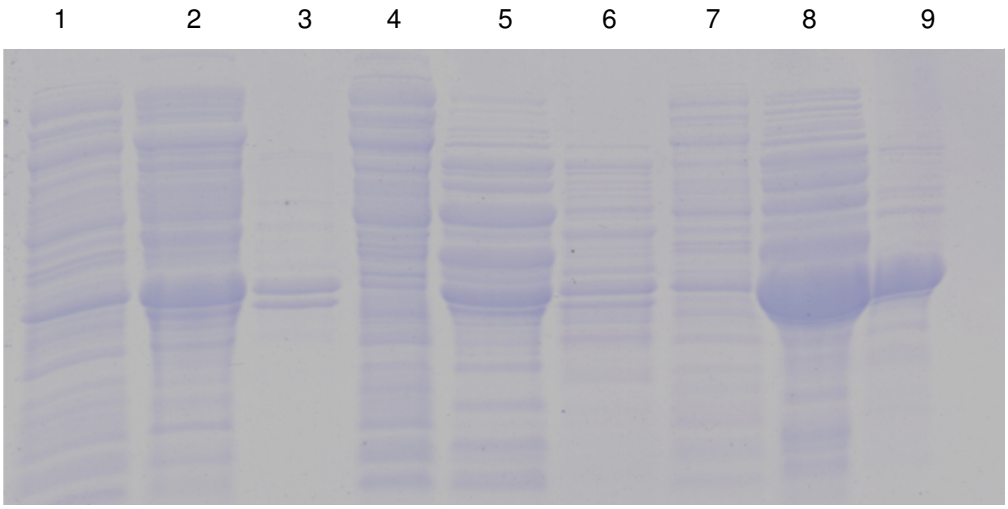


Figure 7.6.: Purification protocol used for mutated forms of TupA, according with TupA purification. (1-3): TupA_R118K, (1): Crude, (2): Ionic exchange chromatography (DEAE), (3): Gel filtration chromatography (Superdex 75), (4-6): TupA_R118E, (4): Crude, (5): Ionic exchange chromatography (DEAE), (6): Gel filtration chromatography (Superdex 75), (7-9): TupA_R118Q, (7): Crude, (8): Ionic exchange chromatography (DEAE) and (9): Gel filtration chromatography (Superdex 75)

Appendix VII

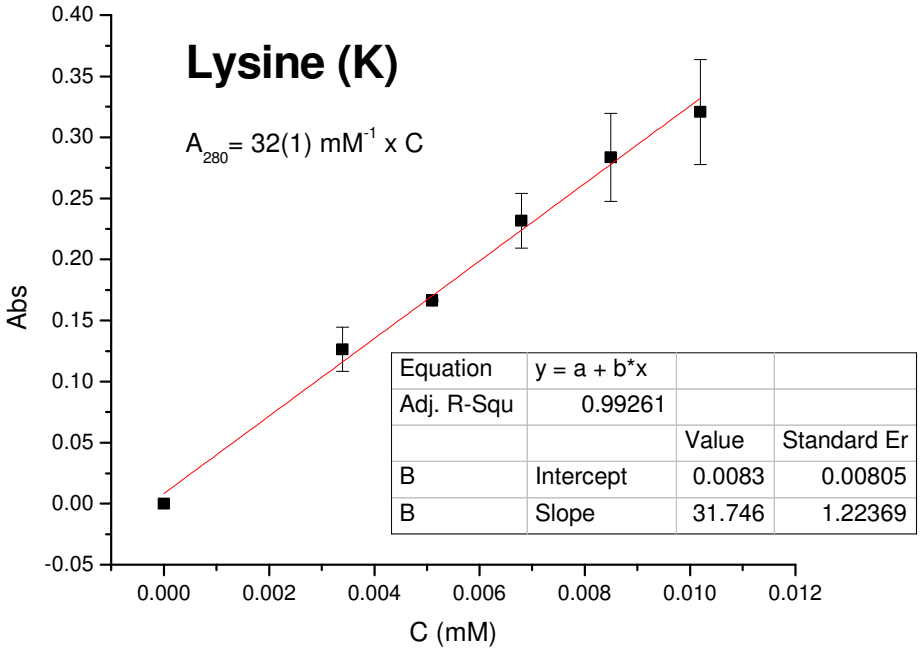


Figure 7.7.: Extinction coefficient for TupA_R118K determined from the slope of the curve obtained by linear regression of the absorbance at 280 nm *vs* protein concentration

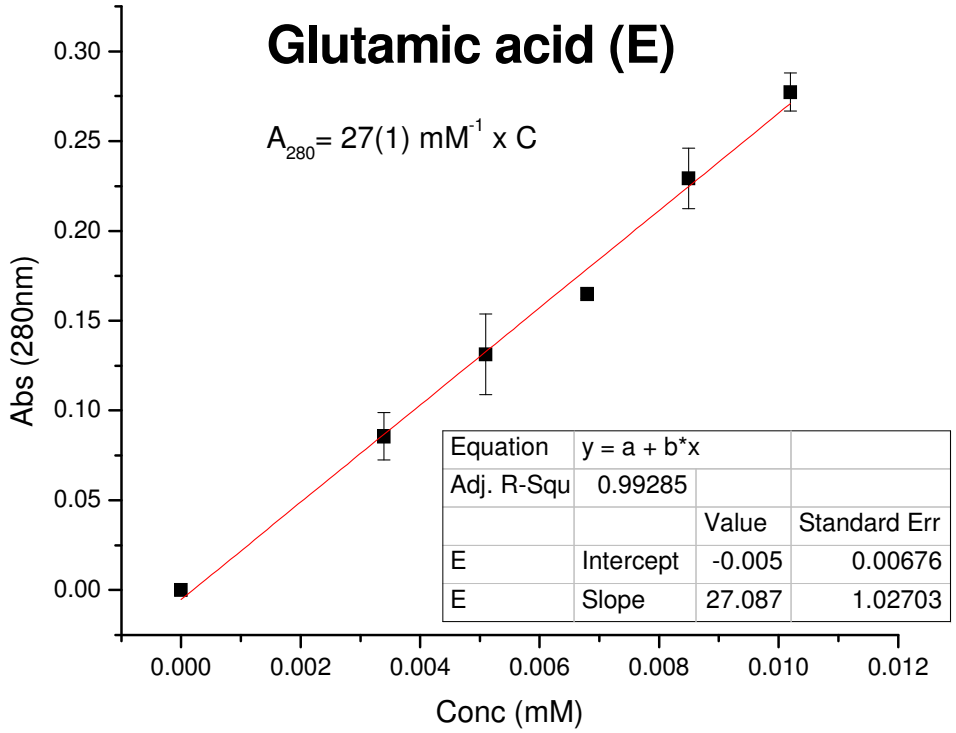


Figure 7.8.: Extinction coefficient for TupA_R118E determined from the slope of the curve obtained by linear regression of the absorbance at 280 nm *vs* protein concentration

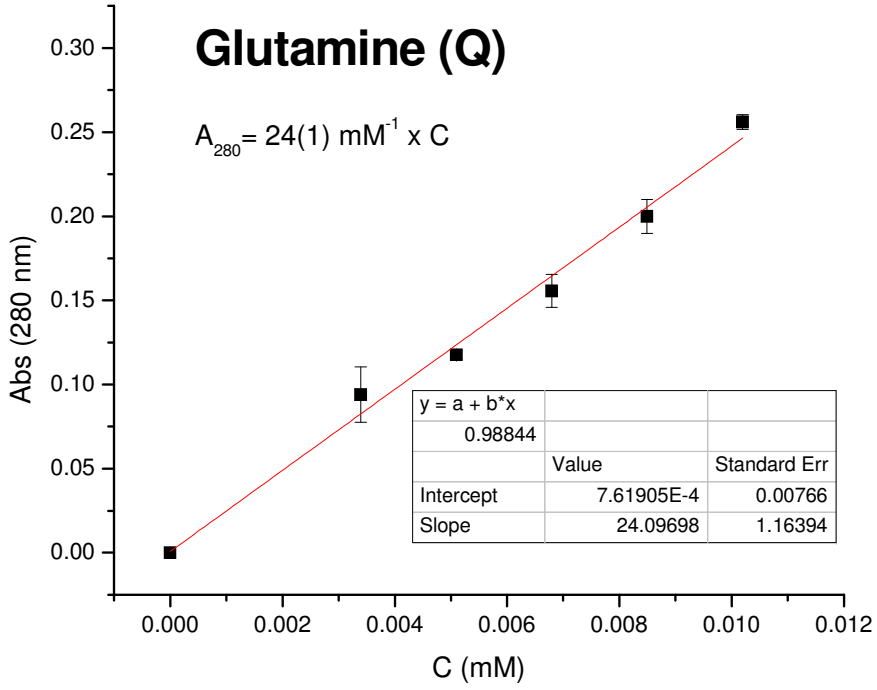


Figure 7.9.: Extinction coefficient for TupA_R118EQ determined from the slope of the curve obtained by linear regression of the absorbance at 280 nm *vs* protein concentration.

Appendix VIII – Screening based on Jancarik *et al*

1. 0.1 M Acetate pH=4.5; 0.2 M $CaCl_2$; 30% MPD

2. 0.1 M MES pH=6.5; 1 M $K_2/Na_2C_4H_4O_6$;

3. 0.4 M $(NH_4)_2PO_4$

4. 0.1 M Tris-HCl pH=8.5; 3 M $(NH_4)_2SO_4$

5. 0.1 M HEPES pH=7.5; 0.2 M $Na_3C_6H_5O_7$; 30% MPD

6. 0.1 M Acetate pH=4.5; 0.2 M $MgCl_2$; 30% PEG 4K

7. 0.1 M HEPES pH=7.5; 1.2 M $Na_3C_6H_5O_7$

8. 0.2 M $Na_3C_6H_5O_7$

9. 0.1 M Citrate pH=5.5; 0,2 M CH_3COONH_4; 30% PEG 400

10. 0.1 M Acetate pH=4.5; 1.5 M NH_4HPO_4

11. 0.1 M HEPES pH=7.5; 1.5 M K_2HPO_4; 1.5 M NaH_2PO_4; 0.2 M $(NH_4)_2SO_4$

12. 0.1 M Tris-HCl pH=8.5; 0.2 M $Na_3C_6H_5O_7$; 20% PEG 400

13. 0.1 M HEPES pH=7.5; 0.2 M $CaCl_2$; 25% PEG 4K

14. 0.1 M MES pH=6.5; 0.1 M $MgCl_2$; 30% PEG 8K

15. 0.1 M Citrate pH=5.5; 0.2M Li_2SO_4; 30% PEG 4K

16. 0.1 M Acetate pH=4.5; 1 M Li_2SO_4

17. 0.1 M Tris-HCl pH=7.5; 0.2 M NH_4HPO_4; 30% MPD

18. 0.1 M Tris-HCl pH=7.5; 0.2 M CH_3COONH_4; 1.5 M K_2HPO_4; 1.5 M NaH_2PO_4

19. 0.1 M Citrate pH=5.5; 0.1 M $(NH_4)_2SO_4$; 30% PEG 8K

20. 0.1 M MES pH=6.5; 30% MPD

21. 0.1 M HEPES pH=7.5; 0.2 M $MgCl_2$; 30% PEG 4K

22. 0.1 M Tris-HCl pH=8.5; 0.2 M $NaCH_3COO$; 30% PEG 4K

23. 0.1 M Tris-HCl pH=7.5; 1 M $K_2/Na_2C_4H_4O_6$;

24. 0.1 M Tris-HCl pH=8.5; 0.2 M $CaCl_2$

25. 0.1 M Citrate pH=5.5; 0.5 M NH_4CH_3COO; 30% MPD

26. 0.1 M MES pH=6.5; 2 M CH_3COONa

27. 0.1 M MES pH=6.5; 0.2 M $K_2/Na_2C_4H_4O_6$; 30% PEG 8K

28. 0.1 M HEPES pH=7.5; 1 M $K_2/Na_2C_4H_4O_6$

29. 0.1 M Acetato pH=4.5; 0.2 M $(NH_4)_2SO_4$; 30% PEG 400

30. 0.1 M HEPES pH=7.5; 0,1 M $(NH_4)_2SO_4$; 20% PEG 4K

31. 0.1 M MES pH=6.5; 2 M $(NH_4)_2SO_4$

32. 0.1 M MES pH=6.5; 0.2 M NaCl; 30% CH_3CH_2OH

33. 0.1 M HEPES pH=7.5; 0.2 M $MgCl_2$; 30% CH_3CH_2OH

34. 0.1 M Tris-HCl pH=8.5; 0.2 M NH_4CH_3COO; 30% CH_3CH_2OH

35. 0.1 M Acetate pH=4,5; 0.2 M $CaCl_2$; 30% CH_3CH_2OH

36. 0.1 M HEPES pH=7.5; 0.2 M $NaCH_3COO$; 30% CH_3CH_2OH

37. 0.1 M HEPES pH=7.5; 0.2 M $MgCl_2$; 30% C_3H_8O

38. 0.1 M Cacodylate pH=6.5; 30% MPD

39. 0.1 M Acetate pH=4.5; 2 M NaHCOO

40. 0.1 M Cacodylate pH=6.5; 0.2 M $Na_3C_6H_5O_7$; 40% C_3H_8O

41. 0.1 M HEPES pH=7.5; 20% PEG 400; 10% C_3H_8O

42. 0.1 M HEPES pH=7.5; 1 M Li_2SO_4

43. 0.1 M Tris-HCl pH=8.5; 0.2 M Li_2SO_4; 30% PEG 4K

44. 0.1 M Cacodylate pH=6.5; 0.2 M $(NH_4)_2SO_4$; 30% PEG 6K

45. 0.1 M Acetate pH=4.5; 1.5 M $NaCH_3COO$

46. 0.1 M $Na_3C_6H_5O_7$; 1 M NH_4 H_2PO_4

47. 4 M NaHCOO

48. 0.1 M HEPES pH=7.5; 1.2 M $Na_3C_6H_5O_7$;

49. 0.4 M $K_2/Na_2C_4H_4O_6$

50. 0.1 M Tris-HCl pH=8.5; 0.2 M $MgCl_2$; 30% PEG 4K

51. 0.1 M Cacodylate pH=6.5; 1.4 M $NaCH_3COO$

52. 0.1 M Citrate pH=5.5; 0.2 M NH_4CH_3COO; 30% PEG 4K

53. 0.1 M Citrate pH=4.5; 0.2 M NH_4CH_3COO, 30% PEG 4K

54. 0.1 M HEPES pH=7.5; 0.2 M $CaCl_2$; 28% PEG 400

55. 0.1 M Cacodylate pH=6.5; 0.2 M $(NH_4)_2SO_4$; 30% PEG 8K

56. 0.1 M Cacodylate pH=6.5; 0.2 M $Mg(CH_3COO)_2$; 30% PEG 8K

57. 0.1 M Tris-HCl pH=8.5; 0.2 M NH_4CH_3COO; 30% C_3H_8O

58. 0.1 M Acetate pH=4.5; 0.2 M $(NH_4)_2SO_4$; 25% PEG 4K

59. 0.1 M Cacodylate pH=6.5; 0.2 M $Mg(CH_3COO)_2$; 30% MPD

60. 0.1 M Acetate pH=4.5; 0.2 M $CaCl_2$; 20 % C_3H_8O

61. 0.1 M Imidazol pH=7; 1 M $NaCH_3COO$;

62. 0.1 M Cacodylate pH=6.5; 0,2 M $Na_3C_6H_5O_7$; 20 % C_3H_8O

63. 0.1 M Cacodylate pH=6.5; 0.2 M $NaCH_3COO$; 30% PEG 8K

64. 0.2 M $(NH_4)_2SO_4$; 30% PEG 8K

65. 0.2 M $(NH_4)_2SO_4$; 30% PEG 4K

66. 0.1 M HEPES pH=7.5; 1.6 M K_2HPO_4; 1.6 M NaH_2PO_4

67. 0.1 M Tris-HCl pH=8.5; 8% PEG 8K

68. 0.1 M Acetate pH=4.5; 8% PEG 4K

69. 0.1 M HEPES pH=7.5; 1.8 M NH_4 H_2 PO_4; 2% PEG 400

70. 0.1 M Citrate pH=5.5; 20% PEG 4K; 20 % C_3H_8O

71. 0.05 M K_2HPO_4; 20% PEG 8K

72. 30% PEG 1K

73. 0.2 M $Mg(HCOO)_2$

74. 0.1 M Cacodylate pH=6.5; 0.2 M $Zn(CH_3COO)_2$; 18% PEG 8K

75. 0.1 M Cacodylate pH=6.5; 0.2 M $Ca(CH_3COO)_2$; 18% PEG 8K

76. 0.1 M Acetate pH=4.5; 2 M $(NH_4)_2SO_4$

77. 0.1 M Tris-HCl pH=8.5; 2 M $(NH_4)_2SO_4$

78. 1 M Li_2SO_4; 2% PEG 8K

79. 0.5 M Li_2SO_4; 18% PEG 8K

80. 0.1 M Citrate pH=5.5; 0.2 M NH_4CH_3COO; 20 % PEG 4K; 20 % C_3H_8O