Filipa Sofia Silva Engrola

Licenciatura em Bioquímica

Bacterial arsenite oxidation at the molecular level

Dissertação para obtenção do Grau de Mestre em Bioquímica

Orientador: **Doutora Teresa Sacadura Santos-Silva, Professora Auxiliar, FCT-NOVA**

Co-orientador: **Doutora Maria João Lobo de Reis Madeira Crispim Romão, Professora Catedrática, FCT-NOVA**

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Outubro, 2019

Bacterial arsenite oxidation at the molecular level

Agradecimentos

Espero que estes poucos parágrafos de expressão de gratidão, consigam minimamente espelhar o contributo que cada uma das pessoas abaixo descritas imprimiram neste documento, mas mais relevante que isso, na minha formação académica e pessoal.

À Professora Doutora Maria João Romão, por ter possibilitado a realização do meu trabalho de tese no grupo X-tal. Apesar dos inumeráveis cargos que acumula, teve a disponibilidade, paciência e vontade efusiva de me coorientar e ensinar. Fico muito grata de ter tido o privilégio de aprender através de alguém com a sua experiência.

À Professora Doutora Teresa Santos-Silva por dar significado à palavra orientadora. O seu gosto pelo que faz, capacidade comunicativa e espírito positivo, fomentaram em mim, sem dar conta, o gosto pela cristalografia e a descoberta interior desta paixão que é a ciência e o fazer ciência. Obrigada por me incentivar sempre, a mim e aos meus colegas, a pensar fora da caixa, por nos abrir os horizontes e por ser um exemplo.

À Doutora Márcia Correia. Obrigada pela paciência e disponibilidade incansável que demonstrou ter para comigo. Obrigada por me deixar sempre testar todas as minhas ideias (até as menos felizes), por me ensinar que tudo tem solução (até recuperar proteína da bancada) e por ter aturado as minhas mil perguntas (até as mais ridículas). Tenho a certeza de que os meus alicerces enquanto bioquímica foram em grande parte cimentados graças a si.

Aos meus colegas de mestrado, Andreia Fernandes e João Paquete, pela companhia, ajuda e partilha durante os momentos muito bons e aqueles de mais desespero, próprios da idade e inexperiência. Vou ter saudades das nossas peripécias, duvidas "totós" e até das arrelias. Mas findado o ano já sabemos, pelo menos, como não destruir o *Akta* (por enquanto)!

Ao Marino Santos, Muthu e Francisco Leisico pelos valiosíssimos conselhos, apoio moral, incentivo e ensinamentos básicos.

A todos os restantes membros do grupo X-tal, não menos importantes nesta jornada, que tiveram sempre uma postura de enorme generosidade, paciência, disponibilidade e alegria laboral, o que contribuiu para o sucesso do meu percurso.

Aos meus queridos e incansáveis pais e irmão. De todas as sortes do mundo não há melhor do que o amor incondicional. Obrigada por fazerem tudo por mim, nunca duvidarem das minhas capacidades e confiarem nas minhas decisões. PS: Espero também justificar o investimento monetário.

Um beijinho muito grande à restante família, mas em especial ao meu tio Rui por todas as vezes que me foi buscar ao comboio e ao meu primo Edgar, por fingir não se entediar com estas coisas da ciência e ouvir atenciosamente os meus disparates (para o ano és tu).

Aos meus amigos, pelos momentos maravilhosos que fazem valer tudo a pena e pela força e felicidade que me transmitem. Daqui a uns tempos podemos pedir *ubereats* sem pensar na taxa de entrega e nos cupões que não funcionam! Um especial agradecimento ao Luís, por ser o melhor, o mais atencioso e especialmente por cozinhar por mim e fazer as tarefas que não suporto!

À Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, ao Departamento de Química, e a todos os Professores que me acompanharam ao longo desta trajetória pelo ensino superior, por me transmitirem valores como o espírito crítico ou o rigor científico, entre tantos outros.

Abstract

According to the WHO, arsenic is one of the top 10 chemical contaminants in drinking-water worldwide and affects more than 140 million people. The arsenite oxidising enzyme (Aio), from microorganisms *Rhizobium* sp. NT-26 (NT-26_Aio) and *Alcaligenes faecalis* (*A.f.*_Aio), and their final electron acceptors – cytochrome c_{552} (NT-26_cytc_{552}) and azurin (*A.f.*_azu), respectively – are currently being studied for their use as biosensors and in bioremediation processes. Both Aio enzymes share high structural similarity (948 matching residues with an r.m.s.d. of 1.84 Å for Cα atoms) and are composed of a large subunit (AioA) which contains a molybdenum centre and a [3Fe-4S] cluster, and a small subunit (AioB) that possess a Rieske [2Fe-2S] cluster.

Aiming to elucidate the catalysis mechanism of the enzymes, and their electron transfer to the final electron acceptors, a combination of expression and purification of the proteins, crystallisation, structural analysis, enzyme kinetics and affinity tests were conducted.

A 1.84 Å resolution structure of *A.f.*_Aio in complex with a substrate analogue - Sb^{V} oxoanion - was determined using molecular replacement (PDB: IG8K). Additionally, a previously obtained 1.89 Å resolution structure of NT-26_Aio, containing a Sb^{III} oxoanion near the active site, was investigated and used for comparison. Analysis of bond lengths and geometry of the ligands at the Mo active site suggests that both crystallized enzymes reveal different reaction intermediates, corresponding to different stages of the mechanism. The specific activity of two active site mutants of NT-26_Aio – D169A and E453A – determined by UV-Vis spectroscopy, revealed that these only uphold 46 and 8% of the WT enzyme efficiency, respectively. This information, together with structural analysis, strongly suggest that both amino acid residues play an essential role in substrate orientation through a complex network of hydrogen-bonds.

Microscale thermophoresis assays revealed that *A.f.*_Aio interacts with the same affinity with its physiological and non-physiological electron acceptors, *A.f.*_azu and NT-26_cytc_{552}, respectively. On the other hand, NT-26_Aio interacts with *A.f.*_azu, although with slightly less affinity than with NT-26_cytc_{552}, showing that electron acceptor diversity is also plausible in this enzyme, unlike what was previously expected.

Keywords: Arsenic, antimony, arsenite oxidase, molybdoenzymes, X-ray crystallography, microscale thermophoresis, enzyme kinetic assays.

Resumo

Segundo a OMS, o arsénio é um dos 10 maiores contaminantes de águas potáveis em todo o mundo, afetando mais de 140 milhões de pessoas. A enzima oxidante de arsenito (Aio), proveniente dos microrganismos *Rhizobium* sp. NT-26 (NT-26_Aio) e *Alcaligenes faecalis* (*A.f.*_Aio), e os seus respetivos aceitadores de eletrões fisiológicos – citocromo c_{552} (NT-26_cytc_{552}) e azurina (*A.f.*_azu) – têm vindo a ser estudados com perspetivas de poderem ser utilizados em técnicas de deteção ou biorremediação de arsenito.

Tanto *A.f.*_azu como NT-26_Aio, partilham bastantes semelhanças estruturais (r.m.s.d de 1.84 Å, entre os Cα de 948 resíduos), sendo compostas por uma grande subunidade (AioA), que alberga um centro de Molibdénio e um centro de [3Fe-4S], e uma pequena subunidade (AioB) composta por um centro Rieske [2Fe-2S].

Visando elucidar o mecanismo de catálise da reação das enzimas e a sua transferência eletrónica para com os seus aceitadores finais de eletrões, foram realizados vários passos de expressão e purificação das proteínas, cristalização, análise estrutural, cinética e testes de afinidade.

Através do método de substituição molecular (PDB: 1G8K), foi possível obter uma estrutura de *A.f.*_Aio em complexo com um análogo do substrato – um oxoanião de Sb^{V} – a uma resolução de 1.84 Å. Foi também realizada a análise de uma estrutura previamente obtida de NT-26_Aio que contém um oxoanião de Sb^{III} ligado ao centro ativo. O comprimento e geometria dos ligandos sugere que ambas as estruturas contêm distintos intermediários de reação, em diferentes fases dos mecanismos reacionais. A atividade específica de dois mutantes do centro ativo de NT-26_Aio – D169A e E453A – revelaram que os mutantes apenas apresentam 46 e 8% da atividade da enzima nativa, respetivamente. Estes dados, aliados ao conhecimento adquirido por análise estrutural, indiciam que estes resíduos de aminoácidos tomam um papel importante na orientação do substrato, através de pontes de hidrogénio.

Os ensaios de MST revelam que *A.f.*_Aio interage com a mesma afinidade com os aceitadores de eletrões fisiológico e não fisiológico, *A.f.*_azu e NT-26_cytc_{552}. NT-26_Aio, por outro lado, interage com ligeiramente menor afinidade com *A.f.*_azu do que com NT-26_cytc_{552}, demonstrando que esta enzima também é capaz de transferir eletrões para diferentes proteínas, ao contrário do que era esperado.

Palavras-chave: Arsénio, Antimónio, Arsenito oxidase, Molibdenoenzimas, cristalografia de raio-X, termoforese em microescala, cinética enzimática.

Table of contents

List of figures

List of tables

Abbreviations and symbols

A.f. – *Alcaligenes faecalis*

***A.f._*Aio –** Arsenite oxidase from *Alcaligenes faecalis*

***A.f._*azu –** Azurin from *Alcaligenes faecalis*

Aio – Arsenite oxidase

ARMs – Arsenate-resistant microbes

As^{III} – Arsenite

As^{V} – Arsenate

DARPs – Arsenate-reducing prokaryotes

DCPIP – 2,6-dichlorophenolindophenol

DFT – Density functional theory

DMAA – Dimethyl arsenic acid

DMSOR – Dimethylsulfoxide reductase

E. coli – *Escherichia coli*

EXAFS – extended X-ray absorbance fine structure

IMAC – Immobilized metal affinity chromatography

IPTG – Isopropyl β-D-1-thiogalactopyranoside

IR – Infrared

K_d – Dissociation constant

LB – Luria-Bertani

MCD – Molybdopterin cytosine dinucleotide

MGD – Molybdopterin guanosine dinucleotide

MMAA – Monomethyl arsenic acid

MST – Microscale thermophoresis

native-PAGE – Native polyacrylamide gel electrophoresis

NT-26 – *Rhizobium* sp. NT-26

NT-26_Aio – Arsenite oxidase from *Rhizobium* sp. NT-26

NT-26_AioA_D169A – Mutant D169A of arsenite oxidase from *Rhizobium* sp. NT-26

NT-26_AioA_E453A – Mutant E453A of arsenite oxidase from *Rhizobium* sp. NT-26

NT-26_cyt*c*$_{552}$ – Cytochrome c_{552} from *Rhizobium* sp. NT-26

O.D. – Optical density

pDNA – Plasmid DNA

r – Ionic radius

r.m.s.d – Root mean square deviation

r.p.m – Rotations per minute

Sb^{III} – Antimonite

Sb^{V} – Antimonate

SDS-PAGE – Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SEC – Size Exclusion Chromatography

SO – Sulfite oxidase

tat leader sequence – Twin arginine translocation peptide signal

TEV – Tobacco Etch Virus

T_m – Melting temperature

TSA – Thermal shift assay(s)

WHO – World health organisation

WT – Wild-type

XO – Xanthine oxidase

ε – Molar extinction coefficient

1. Introduction

1.1. Arsenic contamination: a global problem

Arsenic (As), is a well-known chemical element that has been ubiquitously used by humankind throughout the ages in metallurgy, warfare, pigmentation, medicine, and perhaps more infamously, by virtue of its poisoning properties[1,2]. Due to its toxicity, exposure entails numerous hazardous health effects such as high blood pressure, gangrene, melanosis, diabetes mellitus, cerebrovascular diseases or ultimately cancer[3–5], to name a few.

Nowadays, As is considered one of the top 10 chemicals of primary public health concern by the World Health Organization (WHO)[6]. The leading cause of human contact to arsenic, especially its most dangerous inorganic forms, comes from drinking contaminated water, while inhalation and skin absorption represent minor routes of exposure[7]. The maximum recommended concentration of this pollutant in drinking water, by the WHO, is only of 10 $\mu g.L^{-1}$[6], however, various countries in the world have arsenic contaminated water and do not fulfil this safety guard; currently around 300 million people, worldwide, are estimated to be at risk of gradual arsenic poisoning: arsenicism[8]. The situation is especially worrying in South and Southeast Asian countries like Bangladesh, India or Nepal, were populations heavily rely on groundwater supplies and concentrations as high as 4730 $\mu g.L^{-1}$ have been reported[9]. Widespread nations, such as Japan, China, Argentina, Brazil, United States of America, Canada, Germany, Greece or Hungary, also have documented cases of arsenic contamination, making this a worldwide issue[8]. Figure 1.1[10] helps to set the scale of the problem.

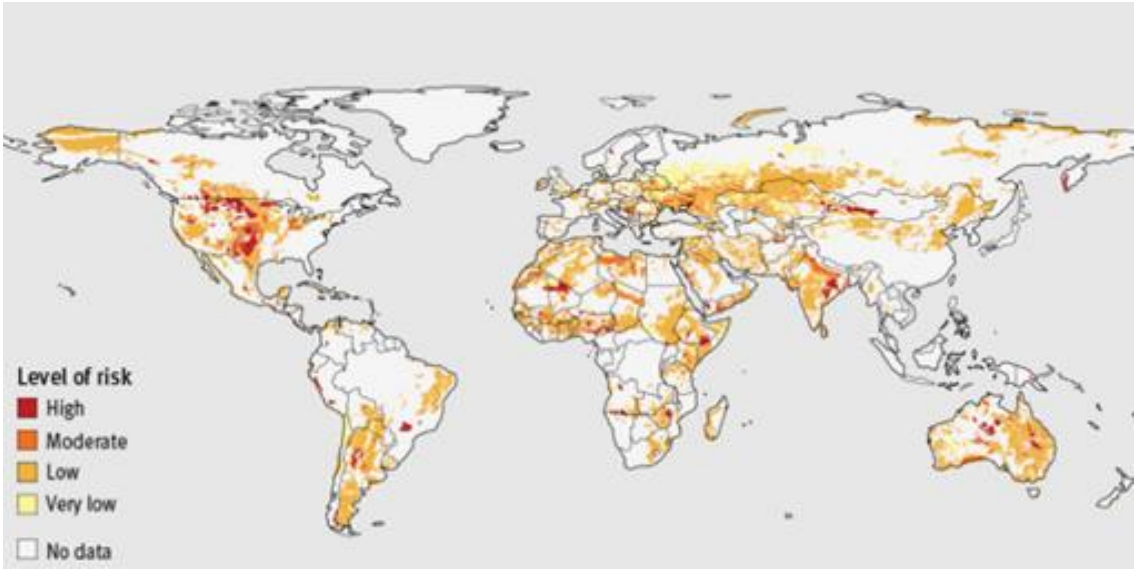


Figure 1.1. Modelled global probability of geogenic arsenic contamination in groundwater. Retrieved from [10].

A sole cause of the problem cannot be isolated; instead, an intricate network of reasons contributes to the widespread of arsenic in the environment. Elevated concentrations can come from natural sources, i.e. leaching and weathering of naturally arsenic-enriched minerals and rocks[11,12], microbial metabolism of such matter[13] or geothermal activity[14], but also from anthropogenic contribution, i.e. coal combustion, use of As-enriched insecticides and herbicides, mining and smelting, among others[15]. Understanding arsenic chemistry is imperative to tackle this environmental crisis.

1.2. Understanding arsenic chemistry

Arsenic is the 20th most abundant natural trace element on earth[16], 99% of which is present in the mineral form, mostly as arsenopyrite (FeAsS), while the remaining 1% can be found in waters, the atmosphere or sediments. This metalloid, belongs to the 15th group of the periodic table, lying between phosphorus and antimony, and having only one stable isotope, ^{75}As, with an average molecular mass of 74.92 g.mol^{-1}[17].

Depending on the pH, the occurrence of complexing ions, and redox potential, arsenic can exist in multiple oxidation states: -III (arsine), 0 (arsenic), and more frequently as +III (arsenite) and +V (arsenate). In water, although the organic forms of arsenic - monomethyl arsenic acid (MMAA) and dimethyl arsenic acid (DMAA) - can be produced via biological activity in scarce amounts, they are irrelevant as a pollutant; instead, the focus should remain on the prevalent inorganic species: arsenite, likely to be present in anaerobic groundwaters, and arsenate, found in aerobic surface waters[18,19]. Arsenic in these oxidation states is present in water as arsenious acid, $H_3As^{III}O_3$ – the most common form at neutral pH – and the corresponding arsenite ions ($H_2As^{III}O_3^-$; $HAs^{III}O_3^{2-}$; $As^{III}O_3^{3-}$), or as arsenic acid ($H_3As^{V}O_4$) and its arsenate ions ($H_2As^{V}O_4^-$; $HAs^{V}O_4^{2-}$; $As^{V}O_4^{3-}$) – at pH 7, the first two charged species prevail in equilibrium[20] (Figure 1.2).

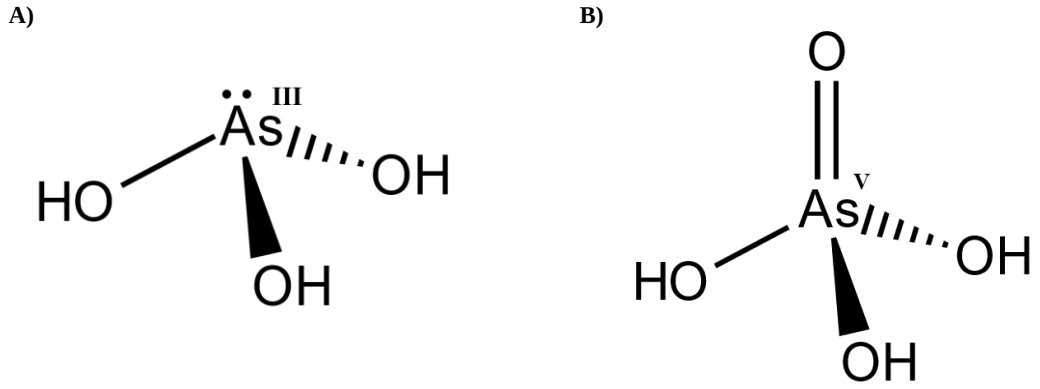


Figure 1.2. Structural depiction of the predominant arsenic species A) arsenite (III), $pKa_{1,2,3}$ = 9.2, 12.1 and 13.41; B) arsenate (V), $pKa_{1,2,3}$ = 2.19, 6.94 and 11.5. Adapted from [20].

As^{III}, due to its size (ionic radius, r, of 0.58 Å) and polarizability, is prone to react with softer S- and N-donating species, resulting in a high reactivity towards cysteine thiols and imidazolium nitrogen of histidine protein residues. Therefore, this form of arsenic can interact with numerous enzymes as well as antioxidants, such as glutathione or lipoic acids, causing oxidative stress and interfering with several metabolic pathways. On the other hand, As^{V} (r = 0.46 Å), has higher stability with harder *O*-donating species, which results in its pervasiveness in aqueous solution. Arsenate is also a structural analogue of orthophosphate (H_3PO_4) with similar pK_a's (2.3, 7.0, and 11.5 for arsenate *vs* 2.1, 7.2, and 12.7 for orthophosphate), which means that arsenate can replace phosphate in energy transfer phosphorylation; nevertheless, it can be easily adsorbed to minerals, rendering it less bioavailable and thus less toxic[18,21–24].

In theory, this As^{V} high affinity to mineral surfaces could serve as an opportunity to explore new oxidation/precipitation techniques and provide effective remediation for the harmful arsenic environmental problem.

1.3. Arsenic remediation and detection

Arsenic remediation has been attempted using several abiotic and biotic techniques. Phytoremediation and membrane technology are not economically viable and have low success rates; ion exchange resins only function at optimal pH and produce low yield results when other competitive ions are present in solution. Methods such as adsorption and precipitation are preferred, but these require a pre-oxidation step of As^{III} into As^{V} due to the previously mentioned low affinity of arsenite to adsorb onto mineral surfaces. Oxidation is currently achieved by a diverse panoply of strategies, such as:

- Addition of chemical compounds like Cl_2, O_3, H_2O_2, NH_2Cl or FeO_4^{2-}, among others, that either react with unsatisfactory time-rates or produce harmful by-products, themselves identified as pollutants and some as carcinogens;
- Photochemical reactions that require unsustainable high-intensity light sources that have not proven to be completely efficient;
- Photocatalytic reactions that fail to produce satisfactory outcomes due to the interference of organic matter or other competitive ions in solution;
- *In-situ* oxidation by pumping oxygenated water into groundwaters aquifers, a slow process which requires complicated engineering systems that are not feasible in most situations;
- Biological oxidation, which is currently gaining attention because of the advantages it provides[25,15].

Another concern is the ability to detect arsenic in water. This is accomplished by the use of highly sensitive laboratory methods like fluorescence, mass spectrometry or inductively coupled plasma optical emission[26,27] that, although providing accurate measurements in a vast range of concentrations, require specialised personal and are not routinely available. Colorimetric-based tests are also accessible in the market[28–30], but these too do not come without disadvantages, involving the handling of dangerous chemicals, the use of specialised equipment or ultimately presenting high market prices.

The exploitation of microorganisms capable of this oxidation step has been attempted to a small extent[31], so the knowledge of these systems and their respective enzymes is crucial for further engineering of highly specific, sustainable and commercially alluring methods for detection and remediation of arsenic-polluted waters.

In the present dissertation, two arsenite oxidase enzymes (Aio), from two distinct bacteria, *Rhizobium* sp. NT-26 (NT-26) and *Alcaligenes faecalis* (*A.f.*), will be studied in detail. Growing evidence shows that this type of enzymes can also catalyse the oxidation reaction from antimonite (Sb^{III}) to antimonate (Sb^{V})[32–34].

1.4. Antimony and the similarities with arsenic

Antimony (Sb) exploitation and industrial emissions have increased in the last decades for its diverse use in semiconductors, fire retardants, batteries or ceramics, alongside other products[35–37], to the point where it was considered by the European Union and the United States Environmental Protection Agency as a priority pollutant[38,39].

Many resemblances can be drawn between arsenic and antimony chemistry. Sb can exist in the same oxidation states as As, (-III, 0, +III and +V), with antimonite (Sb^{III}) and antimonate (Sb^{V}) as the most common; likewise, the first is thermodynamically favoured in anoxic environments and the later under oxygenated conditions[40]. The reduced form of the metalloid has also been reported to have 10-fold higher toxicity than its oxidised counterpart[41]. In aqueous solution, the speciation of antimonite has correspondingly been proposed to result in the prevalence of the uncharged species ($H_3Sb^{III}O_3$) and negatively charged antimonate ($H_2Sb^{V}O_4^-/H_6Sb^{V}O_6^-$) at pH 2-11 range, although an exact depiction of this Sb-O complexes remains controversial, due to the difficulty in isolating them[40,42].

Another similarity between both elements resides in their relatively high redox potential, $E^{0\prime}$ (at pH 7, 25 ^{o}C), assuming the values of +140 mV and +94 mV for the As^{III}/As^{V} and Sb^{III}/Sb^{V} pairs, respectively; this explains the existence of both the oxidised and reduced forms under environmental and biological conditions (unlike phosphate, $E^{0\prime} = -690$ mV)[23,43].

Microorganisms play a significant role in the geochemical cycle of As, affecting its mobility, toxicity and bioavailability and may play an equivalent part in respect to Sb[33,44].

1.5. Microorganisms and their role in arsenic and antimony transformation

Arsenic transformations by microorganisms are relatively well-studied systems, whereas antimony information remains scarce[45].

Four bacterial operons are reported as being involved in uptake, resistance and metabolism of As: *ars, arr, aio, arx*. The first two concern the reduction of arsenate and the later arsenite oxidation. In the *ars* operon, As^{V} is reduced by a periplasmic enzyme, ArsC, leading to its cellular extrusion by an As^{III} efflux pump, ArsB, that also removes Sb^{III}; microorganisms that harbour these genes are called arsenate-resistant microbes (ARMs)[46,47]. The anaerobic respiratory system that uses As^{V} as the final electron acceptor is controlled by the *arr* operon and is found in dissimilatory arsenate-reducing prokaryotes (DARPs); the reaction, catalysed by arsenate reductase,

Arr, provides energy and gives rise to the formation of As^{III}[18]. Moreover, arsenite oxidation can be implied in detoxification mechanisms in heterotrophs or as a source of energy by chemoautotrophs[48]; the *arx* operon system uses the ArxA enzyme to drive the reaction under anaerobic conditions[49–51], while the arsenite oxidase AioAB (or Aio, for short), encoded by the *aio* operon (formerly referred as *aro*, *aso* or *aox*[52]), performs the same task in aerobic settings. Aio was also identified as being able to oxidise Sb^{III}[53,54].

Although more than 50 Sb^{III}-oxidising bacteria have been identified, only recently, an antimonite oxidase – AnoA – belonging to the short-chain dehydrogenase/reductase family of enzymes in *Agrobacterium tumefaciens* GW418, was discovered and characterised[53].

Many of these proteins belong to the molybdopterin-dependent family, a large, ancient group of enzymes responsible for various metabolic steps in all kingdoms of life.

1.6. Molybdenum enzymes and the DMSOR family

Almost all forms of life require the trace element molybdenum (Mo) to thrive; in alternative, some microorganisms use the element that lies immediately below in the periodic table, tungsten (W)[55]. Due to their chemical versatility, they are, for example, essential to the functioning of several enzymes involved in vital physiological steps of sulfur, carbon or nitrogen metabolism. Both metals, under physiological conditions, can transition between VI and IV oxidation states, with the V state also accessible, which enables oxidation-reduction catalysis between obligatory two-electron and one-electron systems[55,56].

The categorisation of Mo and W families is achieved by grouping enzymes based on their active site structures and subsequent chemistry, as well as in respect to the metal coordination to the pterin moiety (one or two pterins) and additional ligands. Molybdenum enzymes are usually assembled in three classes – xanthine oxidase (XO), sulfite oxidase (SO) and dimethylsulfoxide reductase (DMSOR). The DMSO reductase family is the only one that includes W containing enzymes such as Mo/W-containing formate dehydrogenases and Mo-containing nitrate reductases. The DMSOR family includes those Mo and W enzymes that have two pterins coordinated to the metal, while in the XO and SO families the Mo atom is bound to one single pterin. What distinguishes the latter two families is the covalent linkage to an amino acid side chain in SOs, as opposed to xanthine oxidases[57] (Figure 1.3). Pterins can exist in the monophosphate form (MPT) observed in eukaryotes or, as seen in prokaryotes, conjugated to nucleosides: molybdopterin cytosine dinucleotide (MCD) or molybdopterin guanosine dinucleotide (MGD). Their noteworthy role consists in modulating the redox potential of the active centre metal and mediating the

electron transfer between prosthetic groups; this is intricately correlated with the interactions with the surrounding amino acids side chains[55,58].

Arsenite oxidases described in the following sections belong to the DMSOR family (Mo/W-bis MGD) family and are unique in the lack of direct coordination of the Mo atom to an amino acid side chain. Enzymes of this family typically catalyse the dehydrogenation of a substrate or transfer of an oxygen atom, as in the case of Aio[57,59].

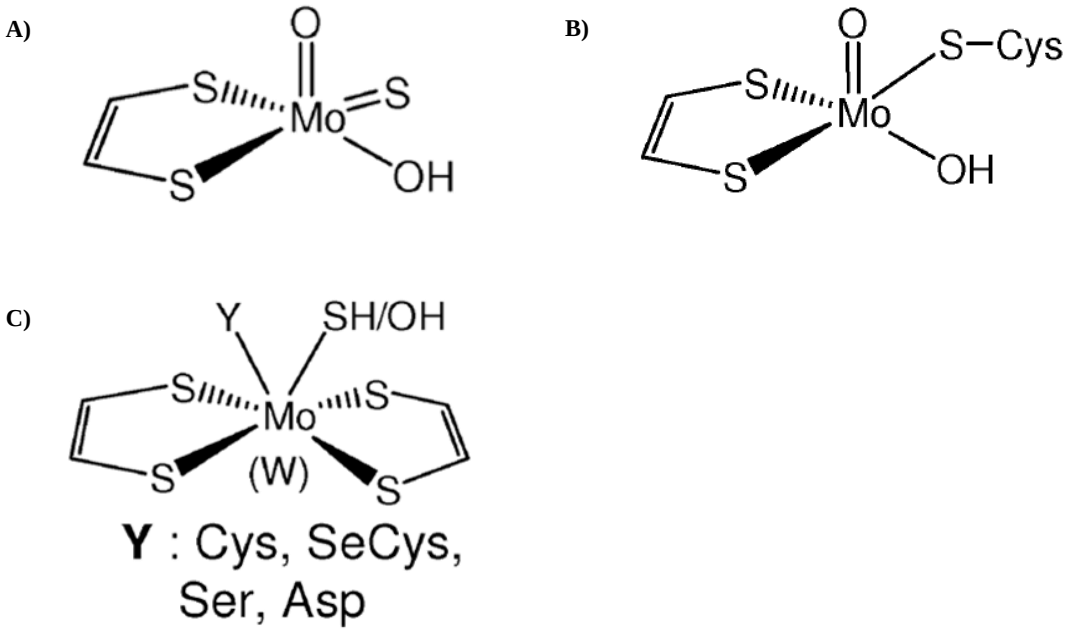


Figure 1.3. Generic core of A) the xanthine oxidase family which possesses only one pyranopterin, an oxo ligand, a hydroxyl ligand and a sulfido ligand; B) the sulphite oxidase family - Mo is coordinated by a single pyranopterin, an oxo ligand, a sulfido ligand and the side-chain of a cysteine residue; C) the DMSOR family where the central metal atom can be either Mo or W bonded to two pyranopterins, and at least one oxo ligand, the final ligand is either an oxo ligand, a hydroxyl, a sulfido ligand or an amino acid side chain. Adapted from [55].

1.7. Arsenite oxidase

Arsenite oxidase is thought to be an ancient bioenergetic enzyme that originated before the evolutionary split of Archaea and Bacteria[60]. Its presence is widespread among prokaryotes, and it has been purified and characterised from several microorganisms[61–65], however, only the arsenite oxidases approached in this dissertation, from *Rhizobium* sp. NT-26 (NT-26_Aio) and *Alcaligenes faecalis* (*A.f.*_Aio) have had their structures solved, by X-ray crystalgraphy[66,67]. Aio catalyses the following oxidation reaction of arsenite to arsenate[68]:

$$2H_3As^{III}O_3 + O_2 \rightarrow HAs^{V}O_4^{2-} + H_2As^{V}O_4^{-} + 3H^+$$

It is composed of two different subunits, α-AioA (~ 90 kDa), that contains the catalytic molybdenum centre and a [3Fe-4S] cluster and β-AioB (~ 15 kDa), which harbours a Rieske [2Fe-2S] cluster. Depending on the microorganism and the purification conditions, it presents different quaternary structures: heterodimeric, tetrameric and hexameric. NT-26_Aio shows an $(\alpha\beta)_2$ arrangement, while *A.f.*_Aio is present as $\alpha\beta$ or $(\alpha\beta)_2$, depending on if it was homologous or heterogeneously expressed, respectively. Its cellular location also varies from periplasmatic, in the case of NT-26_Aio, to cytoplasmatic-membrane bound, in the case of *A.f.*_Aio[65–67,69].

1.7.1. A protein of exceptions

Crystallographic structures of *A.f.*_Aio and NT-26_Aio were solved up to a maximum resolution of 1.64 and 2.70 Å, (PDB: 1G8K and 4AAY[66,67]), respectively. Although both proteins only share 48% homology in sequence, their overall structure is very similar, as well as the arrangement of their co-factors in the overall structure[66] (see Figure 1.4). The first was crystallised as a heterodimer ($\alpha\beta$), and the second as a heterotetramer $[(\alpha\beta)_2]$; in both cases, the α and β subunits are held together by an intricate network of hydrogen-bonds as do the two heterodimers in the case of NT-26_Aio, with additional salt bridges favouring the interaction[66,67].

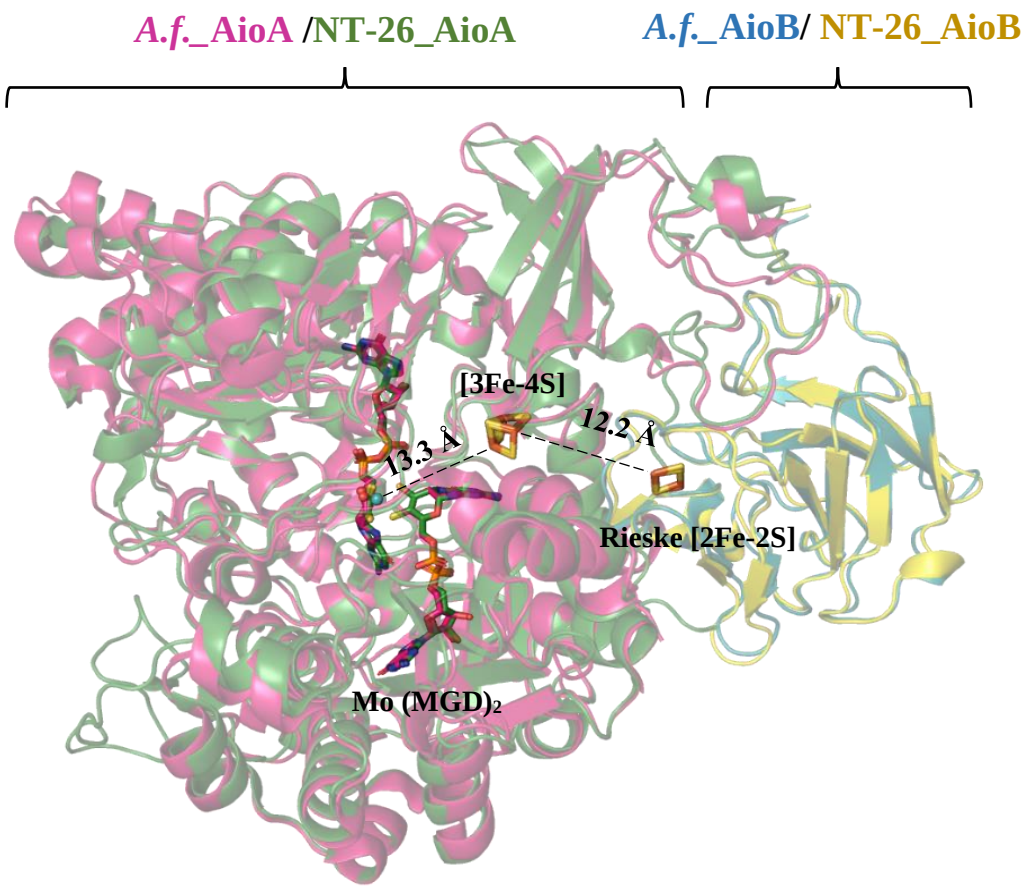


Figure 1.4. Structure of NT-26_Aio superimposed on that of *A.f.*_Aio (948 matching residues with an r.m.s.d. of 1.84 Å for Cα atoms): NT-26 AioA in green and AioB in yellow; *A.f.* AioA in pink and AioB in blue. The cofactors are shown as sticks and the molybdenum atom as space filling mode; iron - orange, sulphur - yellow, carbon - bright green, molybdenum - blue, phosphorus - bright orange, oxygen - red, nitrogen - deep blue. Distances were calculated using PyMOL[105].

The large subunit, AioA, shows the same conserved architecture as other enzymes from the DMSOR family. Four domains (I-IV), not continuous in primary sequence, are arranged in a configuration were the co-factors are buried deep inside the centre of the enzyme[57,66,67] (Figure 1.5). In both deposited structures, the reduced Mo atom (Mo^{IV}) - seemingly due to photoreduction by X-ray, during data acquisition - lies at the bottom of a hydrophilic flattened funnel-like entrance, coordinated by 4 sulphurs from two MGD moieties (P_{MGD} and Q_{MGD}), and one single oxo-ligand in a square pyramidal geometry[66,67] (Figure 1.6). The oxidised state of the metal (Mo^{VI}), is usually not observed by crystallography. Nonetheless, X-ray absorption and Raman spectroscopy suggest that Mo is bound to a second O ligand, either a Mo-O, a Mo-OH, or a Mo=O bond[70]. Recently, protein film voltammetry and extended X-ray absorbance fine structure (EXAFS) results have been interpreted as the Mo^{VI} atom having two Mo=O bonds[71,72].

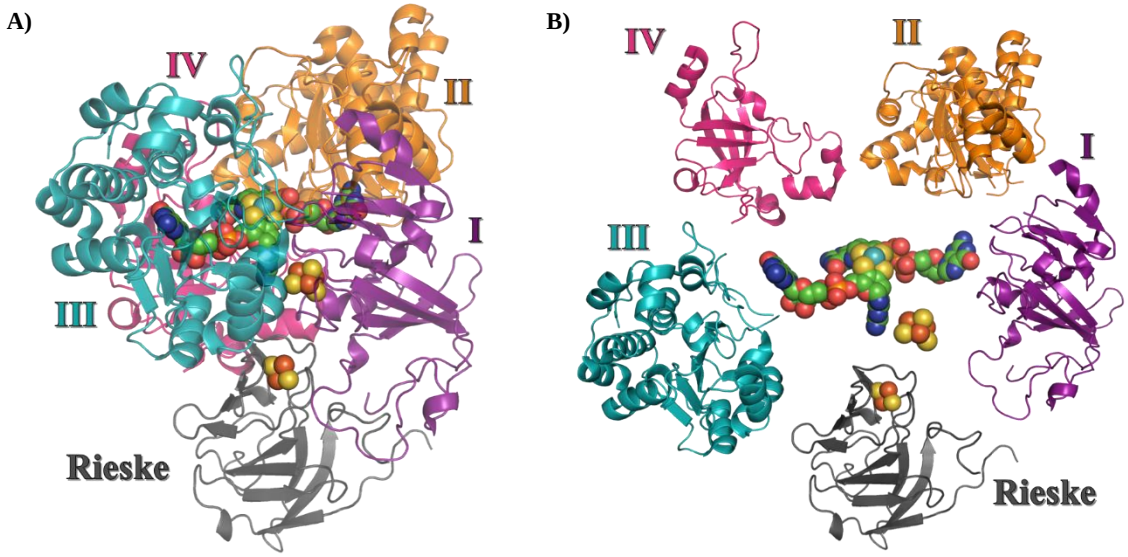


Figure 1.5. Representation of *A.f.*_Aio domains a) domains I-IV of the catalytic unit AioA are colored in purple, orange, blue and pink, respectively; Rieske domain (AioB) is colored in grey; b) an exploded view of arsenite oxidase showing the individual domains, for clarity. Co-factors are shown as space filling mode. Images were designed using PyMOL software[105].

In comparison to all the other Mo isolated proteins that own double MGD coordination, Aio is unique due to the lack of any covalent bond between the metal and the polypeptide chain, resembling the W-containing enzyme aldehyde:ferrodoxin oxidoreductase. While in all other members of the DMSOR family, Mo is coordinated to the side chain of serine, cysteine or selenocysteine, in arsenite oxidase the corresponding residue is a non-bonding Alanine (Ala199 in *A.f.*_Aio numbering). This small difference alters the folding of the *loop* containing this amino acid and significantly exposes the Mo centre[56,66,67]. The catalytic subunit of Aio also harbours a [3Fe-4S] cluster of unusual high redox potential, coordinated by three cysteines and one serine; once again, Aio fails to fit precisely into the DMSO reductase category, since those enzymes usually present a [4Fe-4S] cluster of lower redox potential[57].

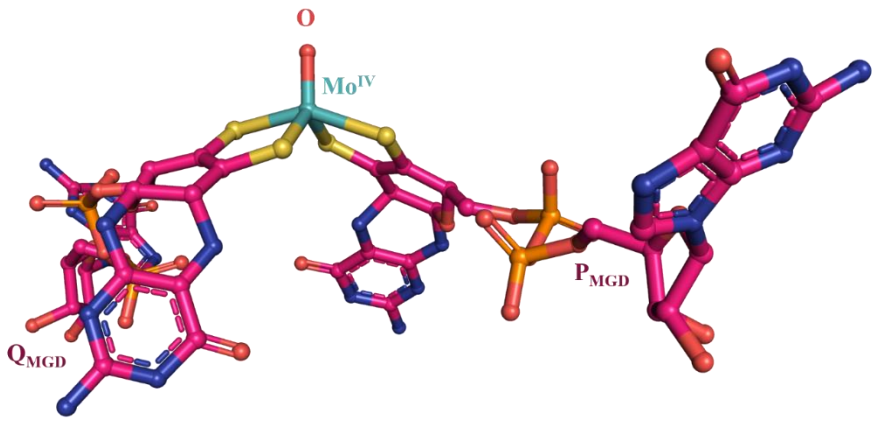


Figure 1.6. *A.f.*_Aio catalytic centre showing a central Mo^{IV} ion coordinated by four sulfurs from the molybdenum guanine dinucleotide pterins and an oxo-ligand. Oxygen (red); Mo (blue); phosphorous (orange); nitrogen (deep blue); carbon (pink); sulphur (yellow). Images were designed using PyMOL software[105]. PDB code: IG8K[66].

The smaller subunit, AioB, possess a high potential Rieske [2Fe-2S] cluster. This type of cluster is coordinated by 2 histidines and 2 cysteines, differing from typical [2Fe-2S] clusters that are bound by 4 cysteines. This feature creates a pH dependency on the redox potential of Rieske centres since Cys endures pH changes and remains negatively charged, while His varies its protonation state[73]. *A.f.*_Aio has a Rieske [2Fe-2S] $E^{0'}$= +130/160 mV and owns a disulphide bridge near this redox centre, between Cys65 and Cys80; NT-26_Aio has an appreciable different corresponding redox potential of +225 mV, and a Phe and a Gly in the position of the bridging cysteines[66]. Both proteins present the same spectral properties, and mutagenesis assays indicate that the disulphide bridge is not responsible for the different redox potentials. However, Phe108 present in NT-26_Aio may influence the protein's affinity towards electron acceptors, since this large hydrophobic aromatic residue may block solvent accessibility[66,74,75] (Figure 1.7). Another notable exception of Aio within the Mo/W-bis MGD family is that this enzyme possesses a twin arginine translocation peptide signal (tat leader sequence) on the β subunit, unlike its relatives, where this sequence is usually present in the α subunit[57].

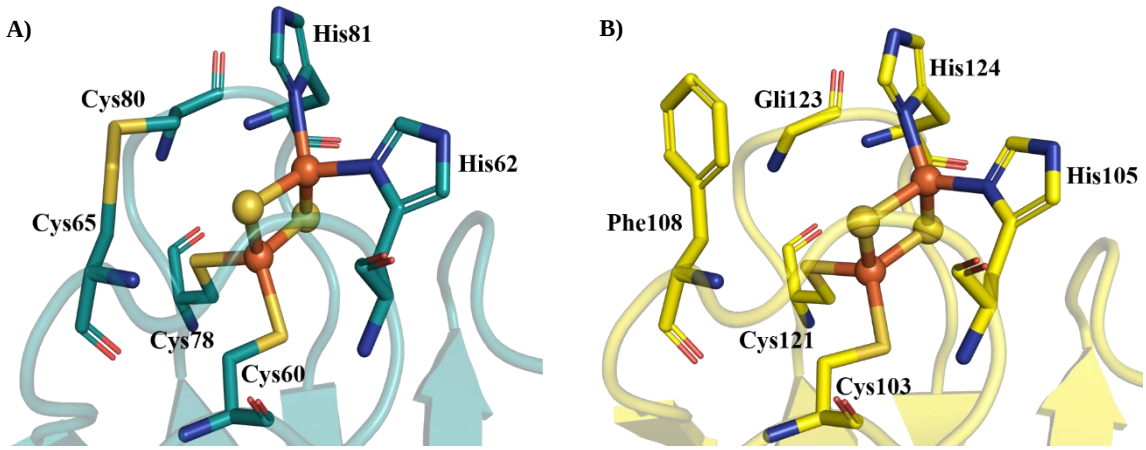


Figure 1.7. Aio Rieske [2Fe-2S] centre representation, showing the iron atoms coordinated by two cysteines and two histidines. A) *A.f.*_Aio cluster is enclosed by a dissulphide bridge (Cys65-Cys80); B) NT-26_AioB entrance is sheltered by a Phe108 residue. Important residues are drawn as stick. Images were designed using PyMOL software[105].

1.7.2. Catalytic mechanism and electron transfer pathway

A proposed mechanism of arsenite oxidation at the molybdenum catalytic centre, interpreted by EXAFS and density functional theory (DFT) calculations is depicted in Figure 1.8.

It has been postulated that one of the oxo-groups of the Mo atom suffers a nucleophilic attack by the nonbonding electron pair of arsenite; As^{III} loses two electrons to Mo^{VI} resulting in As^{V} and reduced Mo^{IV}. Supposedly, the geometry of the oxidised state of Mo gives rise to a relatively high energy conformation due to lack of peptide covalent bond and positioning of the pterins, so the MGD co-factors are thought to twist upon substrate binding to accommodate a transition state bonded to Aio by an oxygen. When the product is released, the Mo oxo-ligand is replaced by reduction of water, resulting in its re-oxidation, and the electrons are transferred to the [3Fe-4S] and Rieske [2Fe-2S] clusters, that become reduced; these centres are thought to only harbour one electron at a time[71,72].

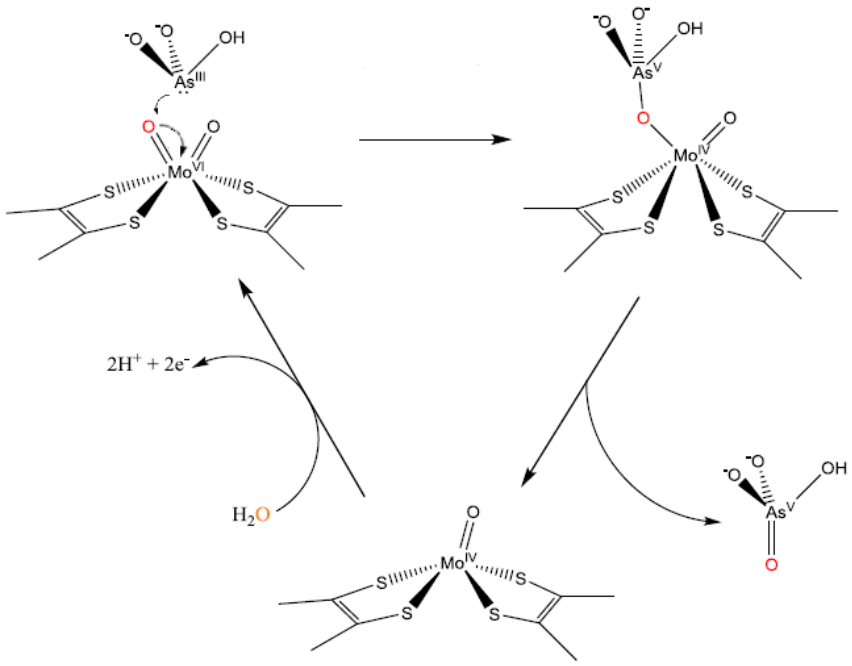


Figure 1.8. Proposed active site mechanism of the oxidation reaction of arsenite to arsenate by the molybdenum active centre of Aio. Adapted from [72].

From the structural point of view, the 2nd coordination sphere of Mo is mostly conserved between Aio structures; the residues His195, Glu203, Arg419 and His423 (NT-26_Aio numbering; Figure 1.9) have been suggested as necessary for catalysis since they make direct contact with Mo=O group and are bound to three water molecules that seem to occupy the substrate-binding site[66]. Even before crystal structures of the arsenite oxidases were available, histidines have been suggested as being implicated in the mechanism due to the nitrogen atoms resulting in a favourable but relatively weak interaction with arsenite (as opposed to the strong binders sulphur atoms of cysteines), which theoretically would facilitate the turnover of the enzyme and dissociation of arsenate[76].

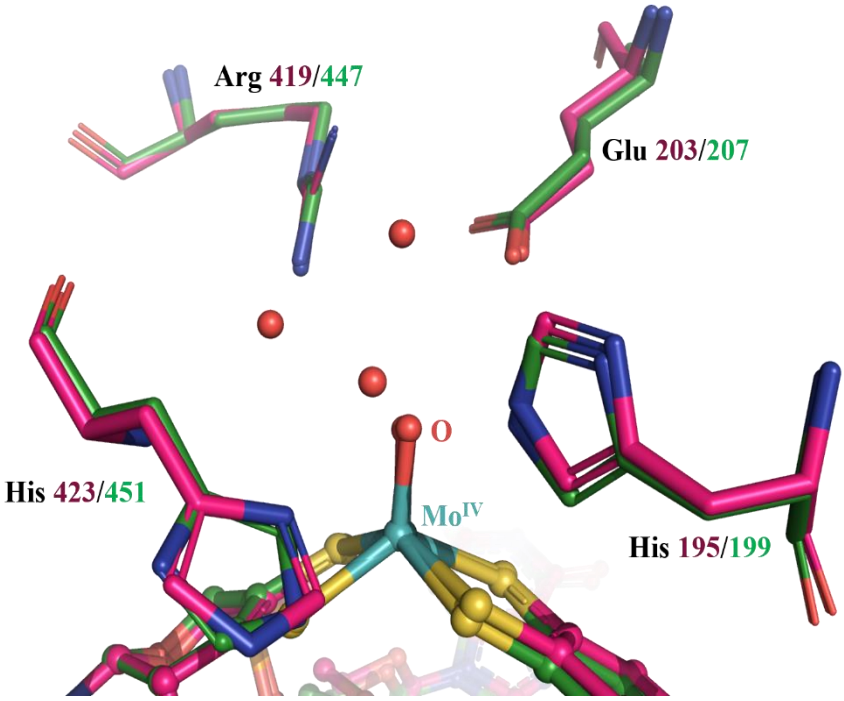


Figure 1.9. Structural superposition of the NT-26_Aio and *A.f.*_Aio conserved 2nd coordination sphere of the Mo active site. Important residues are drawn as stick, and water molecules allegedly occupying the substrate binding site represented as red spheres. Images were designed using PyMOL software[105].

The [3Fe-4S] redox centre sits approximately equidistant from the Mo centre, at ≈ 12 Å, and the Rieske [2Fe-2S] cluster, at ≈ 13 Å. All the redox mid potentials (E_m) of the redox-active centres have been determined, for both enzymes, using protein film voltammetry and electron paramagnetic resonance and are summarised in Table 1.1. The rates of the reaction have been determined only for NT-26_Aio: molybdenum active site reduction rate by As^{III} was >4000 s^{-1}, and the simultaneous reduction of the Fe-S clusters was 564 ± 46 s^{-1}[175].

Table 1.1. Comparison of the redox mid potential values of the redox-active centres of *A.f.*_Aio and NT-26_Aio.

Redox-active centres	E_m (mV) *A.f.*_Aio	E_m (mV) NT-26_Aio	Ref.
Mo	292	367	[71,77]
[3Fe-4S]	260	270	
Rieske [2Fe-2S]	115	225	

The steady state kinetics parameters of the enzymes also afford valuable information about substrate interaction: V_{max} (µmol.min^{-1}.mg^{-1}) measures the maximum catalysis rate of the enzyme; K_M (µM) corresponds to half of the substrate concentration needed to achieve V_{max} and the lower its value, the higher the affinity towards the substrate; finally, k_{cat} (s^{-1}), or *turnover number*, defines the number of substrate molecules that can be turned into product every second by each enzyme molecule. Besides catalysing the oxidation of arsenite, it has been demonstrated that Aio, was also able to oxidise Sb^{III}, in the form of $C_8H_4K_2O_{12}Sb_2 \cdot 3H_2O$, around 6500 times slower than with As^{III}[34]; the kinetic values of the enzymes for these reactions, using different electron acceptors are depicted in Table 1.2.

Table 1.2. Summary of the kinetic parameters of *A.f.*_Aio and NT-26_Aio with several substrates and electron acceptors.

		V_{max} (µmol.min^{-1}.mg^{-1})	K_{cat} (s^{-1})	K_M (µM)	Ref.
Enzyme	NT-26_Aio	4.9	9.3	68.4	[71]
Substrate	As^{III}				
Electron acceptor	DCPIP				
Enzyme	NT-26_Aio	120.4	228	9.3	[34,75]
Substrate	As^{III}				
Electron acceptor	cyt *c* from horse heart				
Enzyme	NT-26_Aio	0.0184	34.8	3	[34]
Substrate	Sb^{III}				
Electron acceptor	cyt *c* from horse heart				
Enzyme	*A.f.*_Aio	-	27	8	[64]
Substrate	As^{III}				
Electron acceptor	*A.f.*_azu				
Enzyme	*A.f.*_Aio	3	-	2	[64,78]
Substrate	As^{III}				
Electron acceptor	DCPIP				

The electron transfer pathway proposed for Aio, Mo → [3Fe-4S] → [2Fe-2S], culminates in the interaction with an electron acceptor. Allegedly the electron transport is facilitated by the solvent-exposed His of the Rieske centre, and physiologically, this role is attributed to a small protein (Figure 1.10). In the case of *A.f.*, it can be either the copper cupredoxin, azurin (*A.f.*_azu), or a cytochrome *c*, although this has never been isolated[79]; as for NT-26, cytochrome c_{552} (NT-26_cytc_{552}) is known to accept electrons from Aio but is believed not to be the only one[80]. *In vitro*, other exogenous electron acceptors have been used: 2,6-dichlorophenolindophenol (DCPIP), which is reduced by both enzymes; azurin from *Pseudomonas aeruginosa*, which can be reduced by *A.f.*_Aio, and equine cytochrome c, only reduced by NT-26_Aio[79,80] (Table 1.3). This selective electron transfer has been verified in other Aio proteins and appears to be strain-dependent[65,69,80–82] and unrelated with the reduction potential (all of the electron acceptors above mentioned have $E^{0'}$>240 mV)[80].

Table 1.3. Comparison of electron acceptor specificity between *A.f.*_Aio and NT-26_Aio.

Electron acceptor	***A.f.*_Aio**	**NT-26_Aio**
*A.f.*_azu	+	N.A.
NT-26_cytc_{552}	N.A.	+
Azurin from *Pseudomonas aeruginosa*	+	-
Cytochrome *c* from horse heart	-	+
DCPIP	+	+

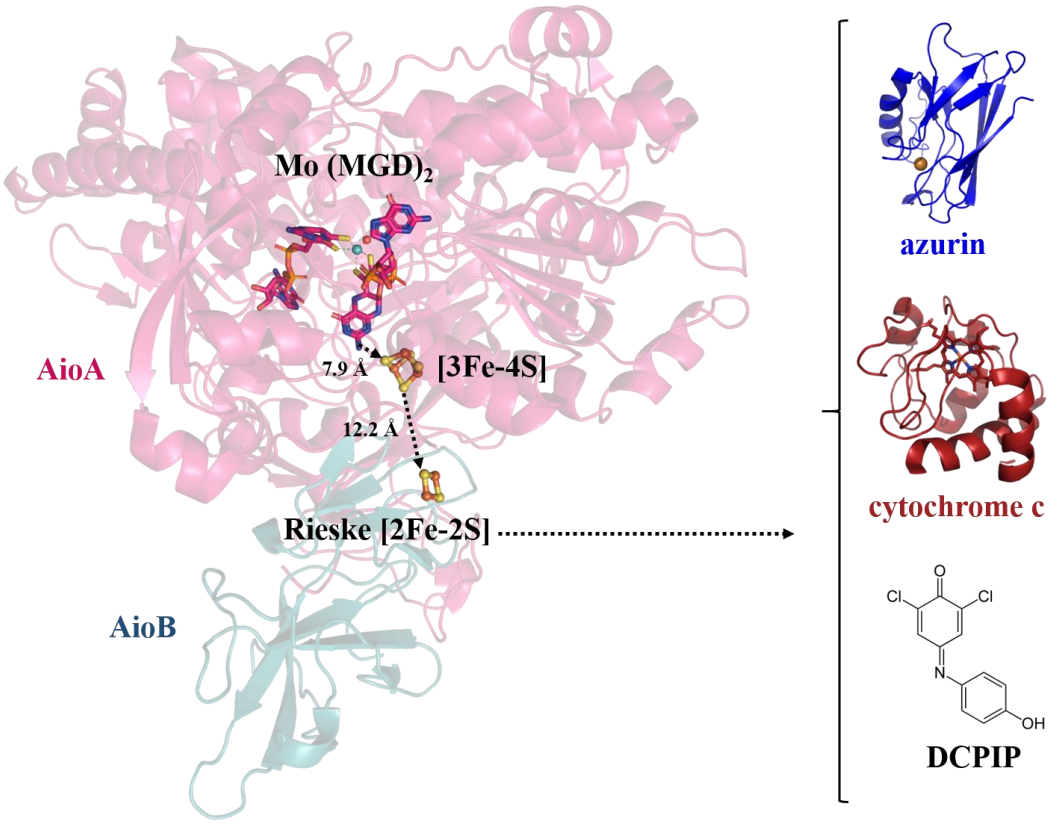


Figure 1.10. Electron transfer pathway between Aio (*A.f.*_Aio, as an example) and some of the possible final electron acceptors – *A.f.*_azu (PDB code: 2H3X), cytochrome *c* from horse heart (PDB code: 1HRC), DCPIP – that are transferred from the pterins to the [3Fe-4S] cluster and later to the Rieske [2Fe-2S] centre, culminating in the final electron acceptor. Dashed arrows represent the direction of the flow of electrons. Images were designed, and distances were calculated using PyMOL software[105].

2. Objectives

The present dissertation aims to clarify the catalytic mechanism and electron transport to electron acceptors of two arsenite oxidases, from two distinct microorganisms: *Rhizobium* sp. NT-26 and *Alcaligenes faecalis* and has the following objectives:

- Optimisation steps of previously *E. coli* heterologous expression and purification protocols to improve expression yield and sample quality.
- Structural characterisation by X-ray crystallography, of the *A.f.*_Aio structure complexed to a substrate analogue in the active centre and comparison with an NT-26_Aio previously obtained structure, also showing the presence of a substrate analogue in the active centre. This will possibly allow the visualisation of reaction intermediates and enlighten the reaction mechanism.
- X-ray crystallography and UV-Visible spectroscopic studies of two active site mutations – NT-26_AioA_D169A and NT-26_AioA_E453A – intenting to elucidate the role of these amino acids in the reaction with the substrate.
- Microscale thermophoresis assays of both arsenite oxidases and their electron acceptors to analyse further the interaction between these protein pairs.

3. Experimental work

3.1. Protein expression and purification

The molecular biology component in the following sections results from a collaboration with Professor Joanne Santini from University College of London (UCL), England. All plasmids containing the genes of interest (*aioAB* and *cytC* from *Rhizobium* sp. NT-26, as well as *aioAB* and *azu* from *Alcaligenes faecalis*) were previously provided. DNA cloning and site-directed mutagenesis protocol, where the genes *aioA* and *aioB* (*aioAB*) from *Rhizobium* sp. NT-26 were used as a template to produce two mutants - NT-26_AioA_D169A and NT-26_AioA_E453A - was performed by student Marta Vieira, under the supervision of Professor Teresa Santos-Silva and Professor Maria João Romão from UCIBIO, FCT-UNL, and by Professor Joanne Santini and Doctor Cameron Watson from UCL.

Bacterial liquid cultures were always grown in sterile Luria-Bertani (LB) medium supplemented with the adequate antibiotic (100 $\mu g.mL^{-1}$ of ampicillin in the case of arsenite oxidases and *A.f.*_azu or 100 $\mu g.mL^{-1}$ ampicillin + 60 $\mu g.mL^{-1}$ chloramphenicol for NT-26_cytc_{552}), except for transformation protocol.

3.1.1. Transformation of competent cells

Expression of the recombinant proteins requires a first transformation step using previously competent *E. coli* cells from different strains - Rosetta 2(DE3)pLysS, BL21 (DE3), DH5α and Tuner (DE3). For each case, the same thermal shock method was applied, under sterile environment, where 1 µL of plasmid DNA (pDNA), was added to a previously cooled *Eppendorf* containing 100 µL of competent cells. The mixture was incubated on ice for 30 minutes, followed by 40 seconds at 42 ºC, and then, 5 minutes on ice. Subsequently, fresh LB medium was added to the cells suspension and incubated for 1 hour at 37 ºC (Orbital Shaker-Incubator ES-20, Grant-bio). The cells were recovered by centrifugation at 1000 x g for 1 minute; 1 mL of excess culture medium was removed, and the pellets resuspended. The resulting cell suspension was plated onto solid LB medium supplemented with the adequate antibiotic and incubated at 37 ºC, overnight (Orbital Shaker-Incubator ES-20, Grant-bio).

3.1.2. Heterologous expression of NT-26_Aio and its mutants: NT-26_AioA_D169A and NT-26_AioA_E453A

The reported wild-type (WT) heterologous expression conditions[66] were attempted for both the WT and mutant enzymes of NT-26. Several expression hosts from different strains of *Escherichia coli* (*E. coli*) - Rosetta 2(DE3)pLysS, BL21 (DE3) and Tuner (DE3) (from Merck Millipore) - as well as different isopropyl β-D-1-thiogalactopyranoside (IPTG; NZYtech) concentrations - 0.1, 0.25, 0.5, 0.75 and 1 mM - were tested (Appendix 7.1). All cells containing the respective plasmid of interest were transferred from a petri dish and grown, initially, overnight, at 37 ºC, 200 r.p.m. (Orbital Shaker-Incubator ES-20, Grant-bio) in 50 ml falcon tubs containing 10 ml of culture medium. The next day, 50% of the inoculum was transferred to fresh media (500 ml in a 2000 ml flask) supplemented with 1 mM $Na_2MoO_4.2H_2O$ and incubated at 37 ºC under 180 r.p.m. agitation (Shaker Gallenkamp) for 5 hours. When optical density (O.D.) of the cultures reached 0.7, protein expression was induced for 24 hours at 21 ºC with 180 r.p.m agitation (Shaker Gallenkamp), by addition of IPTG.

3.1.3. Heterologous expression of *A.f.*_Aio

The expression of *A.f.*_Aio was adapted from the protocol previously described[83]. An isolated colony of *E. coli* DH5α containing the plasmid with recombinant *aioBA* genes for *A.f.*_Aio was inoculated in a 50 ml falcon containing 3 ml of media at 37 ºC under 200 r.p.m. agitation (Orbital Shaker-Incubator ES-20, Grant-bio). After 7 hours, 2 ml of the inoculum were transferred to 100 ml of media supplemented with 1 mM $Na_2MoO_4.2H_2O$, in a 500 ml flask, and incubated overnight using the previous temperature and rotation conditions. Lastly, the totality of the culture was transferred to a 2 Litre flask containing 400 ml of fresh LB; the media was also supplemented with 1 mM of $Na_2MoO_4.2H_2O$ and 40 µM of IPTG for induction, that lasted for 24 hours at 21 ºC, 180 r.p.m (Shaker Gallenkamp).

3.1.4. Cell harvesting, purification and concentration determination of arsenite oxidase proteins

All arsenite oxidase proteins (Aio_NT-26, NT-26_AioA_D169A, NT-26_AioA_E453A, and *A.f.*_Aio) were subject to the same (adapted) protocol, except if explicitly specified.

Subsequently to the induction period, cells were harvested at 6 ºC by centrifugation at 6.000 r.p.m. for 15 minutes (Rotor JA-10, Avanti J-26 XPI, Beckman Coulter). The supernatant was discarded, and the cell *pellets* resuspended with binding buffer (50 mM HEPES, 500 mM NaCl, 20 mM imidazole, pH 7; Sigma Aldrich) at 10 $ml.g^{-1}$ of wet weight cells. To disrupt the cells, 6 cycles of ultrasounds that lasted for 1 minute with 1-minute intervals (UP100H/Tip MS7, Hielscher Ultrasonics; 50% cycle and 80% amplitude) were applied; the lysate was clarified by centrifugation: 16000 r.p.m for 30 minutes at 6 ºC (Rotor JA25-50, Avanti J-26 XPI, Beckman Coulter). Incubation of the viscous soluble fraction occurred for about 40 minutes, with 5 mM $MgCl_2$ and DNase I and the insoluble cellular debris discarded. Finally, the samples were filtered using a 0.40 mm filter (Millipore).

The first purification step of Aio proteins was an Immobilized Metal Affinity Chromatography (IMAC). The soluble fraction containing the protein was loaded into a nickel His-Trap™ column (GE-Healthcare) pre-equilibrated with binding buffer, using an Äkta Purifier system and a UNICORN™ start 1.0 module to monitor the process. Using an elution buffer (50 mM HEPES, 500 mM NaCl, 500 mM imidazole, pH 7), elution of the proteins that presented affinity to the column was possible by two steps of 44 and 500 mM imidazole. The flux was kept constant throughout the assay, at 2 $ml.min^{-1}$.

To know which collected fractions contained Aio and assess their purity, these were analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) with 10% polyacrylamide, followed by protein Coomassie Brilliant Blue staining. Each well was loaded with 12 µL of a sample consisting of 4 µL of loading buffer + 14 µL of protein solution, previously boiled for 5 minutes. The fractions that contained both subunits of Aio protein were pooled and dialysed overnight (3.5 kDa membrane *cutoff*) at 4 ºC with 1 mg of Tobacco Etch Virus enzyme (TEV) against 1 L of binding buffer, with 1 mM EDTA and 1 mM DTT.

The sample was subjected to a 2nd IMAC and analysed by SDS-PAGE, in the same manner as described previously. For buffer exchange to 50 mM Tris-HCl pH 7.8, samples containing Aio were loaded onto a PD-10 desalting column (GE Healthcare), according to the manufacturer's instructions. The protein was then concentrated (Centricons Vivaspin® Turbo 15, 30 kDa, Sartorius) up to 45 $mg.ml^{-1}$. Protein concentration was determined by spectroscopic absorbance readings (Spectra Max 190, Molecular Devices), using molar extinction coefficient values (ε) calculated by the bioinformatic tool ProtParam from the ExPASy portal, Swiss Institute of Bioinformatics[84]: ε_{280nm}(NT-26_Aio)=142100 $M^{-1}.cm^{-1}$; and ε_{280nm}(*A.f.*_Aio)=157110 $M^{-1}.cm^{-1}$. The proteins were divided into aliquots (20 µL) and stored at -80 ºC.

3.1.5. Heterologous expression of NT-26_cytc_{552} and *A.f.*_azu

The protocol for heterologous expression of both electron acceptors was conducted similarly apart for some differences noted further.

E. coli BL21 (DE3) cells containing the plasmids with the genes of interest (*azu* or *cytC*), were inoculated in 20 mL of LB medium (100 mL flasks) and incubated, overnight, under 180 r.p.m agitation at 37 ºC in the case of *A.f.*_azu and 30 ºC for NT-26_cytc_{552} (Orbital Shaker-Incubator ES-20, Grant-bio). After this step, the cultures were transferred to 2 L flasks containing 1 L of medium supplemented with 10 mL of trace element solution (Appendix 7.2) or 0.2 M $CuCl_2$, for NT-26_cytc_{552} or *A.f.*_azu, respectively. The proteins were incubated for approximately 4 hours under 180 r.p.m agitation, at 28 ºC for NT-26_cytc_{552} or 30 ºC for *A.f.*_azu until their cell suspension reached $O.D._{600nm}$=0.6. Then, NT-26_cytc_{552} expression was induced by addition of 20 µM of IPTG, overnight, at 180 r.p.m and 30 ºC, while *A.f.*_azu induction was achieved with 250 µM of IPTG, for 6 hours, under the same agitation but at 28 ºC (Shaker Gallenkamp).

3.1.6. Cells harvesting, purification and concentration determination of NT-26_cytc_{552} and *A.f.*_azu

For both *A.f.*_azu and NT-26_cytc_{552}, cell harvesting, lysis and soluble extract treatment before purification, were conducted in the same way as described for Aio but using 20 mM K_3HPO_4, 500 mM NaCl, 20 mM imidazole, pH 7.4 as a binding buffer (reagents purchased from Sigma Aldrich).

The first purification step, for both proteins, was an IMAC performed under the constant flux of 2 $ml.min^{-1}$.The soluble fraction of each lysate was loaded onto the column and oxidised by 25 mL of a binding buffer solution containing 1 mM $K_3[Fe(CN)_6]$. Elution of the proteins happened by a 44 mM imidazole step followed by an imidazole gradient, ranging from 44 to 444 mM, using an elution buffer consisting in 20 mM K_3HPO_4, 500 mM NaCl, 500 mM imidazole, pH 7.4.

The proteins were concentrated (Centricons Vivaspin® Turbo 15, 3 kDa *cutoff*, Sartorius) up to 300 uL; during this procedure, some endogenous *E. coli* proteins precipitated and were removed by centrifugation at 4 ºC at 13000 r.p.m.

The second purification assay consisted of a size exclusion chromatography (SEC), using a superdex-75 10/300 column (GE-Healthcare) and monitoring the process with LabSolutions software (Shimadzu) through absorbance changes at 280 nm. The column was equilibrated with 50 mM Tris-HCl pH 7.8, before each injection of 10 mg of protein. Flow rate was set to 0.7

ml.min^{-1} and 500 µL fractions were collected and analysed using SDS-PAGE, as described for arsenite oxidases. The ones that contained protein with a high degree of purity were concentrated again (Centricons Vivaspin® Turbo 15 de 3 kDa, Sartorius) up to 15 mg.ml^{-1} in the case of NT-26_cytc_{552}, and 66 mg.ml^{-1} for *A.f.*_azu. Protein concentration was determined by spectroscopic absorbance readings, with molar extinction coefficients calculated by the bioinformatic tool Prot-Param from the ExPASy portal[84]: ε_{280nm}(*A.f.*_azu)= 7042 M^{-1}.cm^{-1}; ε_{280nm}(NT-26_cytc_{552})= 25680 M^{-1}.cm^{-1}.

3.2. Size exclusion chromatography of NT-26_AioA_D169A

Previously stored NT-26_AioA_D169A was subject to an SEC assay, using a superdex-200 10/300 column (GE-Healthcare) and monitoring the process with LabSolutions software (Shimadzu), through absorbance changes at 280 nm. The column was equilibrated with 50 mM Tris-HCl pH 7.8, before each injection of 10 mg of protein. The flow rate was set to 0.7 ml.min^{-1}, and 500 µL fractions were collected and analysed using 10% SDS-PAGE and 10% native polyacrylamide gel electrophoresis (native-PAGE). Samples were prepared as described for arsenite oxidases, apart from temperature denaturation, when performing native-PAGE.

3.3. Thermal shift assays of NT-26_AioA_D169A

To evaluate and improve arsenite oxidases stability, several thermal shift assays (TSA) were conducted: first, NT-26_AioA_D169A was screened against 96 buffer solutions (Jena Biosciences), and later, using an additive screen (Jena Biosciences), in order to find more suitable buffers for all arsenite oxidases from NT-26.

The solution assays contained a mixture of 2 µL NT-26_AioA_D169A (10 µM final concentration), 3 µL of protein thermal shift dye (3 µM final concentration; ROXTM, Biotium), either 10 µL of buffer or 2.2 µL of additive screen solution, and water or the selected buffer up to a final volume of 20 µL. The solutions were centrifuged at 13000 r.p.m for 1 minute at 25 ºC before added to MicroAmpa™ fast 96-well reaction plates (Applied Biosystems™, ThermoFisher Scientific). The experiments were conducted in a StepOnePlus™ Real-Time PCR System (Applied Biosystems™, ThermoFisher Scientific) applying a temperature ramp, from 25 to 95 ºC at a rate of 2 ºC.min^{-1}.

3.4. X-ray crystallography

3.4.1. Crystallisation, soaking and data collection of *A.f._*Aio

*A.f._*Aio had been previously crystallised by the hanging-drop vapour diffusion method at 4 ºC in two different crystallisation conditions: i) PEG 10000, cobalt(II) chloride, calcium chloride, MES, mercury(II) chloride, pH 6.4; ii) PEG 6000, tris-HCl, pH 8.5[67]. In order to obtain crystals of Aio in complex with the substrate analogue - antimony - and possibly of the Aio-*A.f._*azu electron transfer complex, the proteins were co-crystallized at 4 ºC using the commercial screen JBScreen Classic 1-5 (Jena Bioscience), since it contained many conditions that used different PEG's as the precipitating agents or additives, as in the assays mentioned beforehand. Using a crystallisation robot (Oryx8; Douglas Instrument) the drops were prepared with 0.33 µL of *A.f._*azu at 130 µM or 1.78 mg.ml^{-1}, 0.33 µL of *A.f._*Aio at 130 µM or 13.95 mg.ml^{-1} and 0.33 µL of a precipitant solution; while the reservoirs contained 50 µL of the screen solution. Several crystals with the same morphology, thin brownish plates that appeared within one to four weeks, from different crystallisation conditions, are summarised in Table 3.1.

All crystals were harvested in a harvesting solution containing the same reagents as the crystallisation solution but with the precipitant 2% more concentrated and were later cryo-preserved in the same harvesting solution with 30 % glycerol. Some of the drops were selected for soaking experiments: i) using a harvesting solution containing 50 mM of $K_3[Fe(CN)_6]$ for 5 minutes, and later with a harvesting solution supplemented with 10 mM $C_8H_4K_2O_{12}Sb_2.3H_2O$ for a period of time that varied between 15 and 45 minutes ii) using a harvesting solution containing 4 mM of *A.f._*azu, also for 15 to 45 minutes. All crystals were flash-frozen into liquid nitrogen and exposed to synchrotron X-ray radiation at beamline Biomax, MAX IV (Lund, Sweden) and beamline I04, Diamond Light Source (DLS - Didcot, Oxfordshire, United Kingdom).

Table 3.1. Summary of the crystallisation conditions, soaking procedures and X-ray diffraction results of 13.95 mg.ml^{-1} *A.f.*_Aio and 1.78 mg.ml^{-1} *A.f.*_azu crystals.

Crystallisation conditions	Soaking agent/ Complex	Soaking agent: time (min)	Maximum resolution of the obtained crystals (Å)
30% (v/v) PEG 4 K; 0.2 M ammonium sulphate	$C_8H_4K_2O_{12}Sb_2.3H_2O$	20	NA
20% (v/v) PEG 4 K; 0.1 M trisodium citrate; 5% (v/v) isopropanol	-	-	2.31
20% (v/v) PEG 4 K; 0.1 M trisodium citrate; 5% isopropanol	$C_8H_4K_2O_{12}Sb_2.3H_2O$	30	2.18
20% (v/v) PEG 3 K; 0.1 M HEPES pH 7.5; 0.2 M sodium acetate	-	-	1.96
20% (v/v) PEG 3 K; 0.1 M HEPES pH 7.5; 0.2 M sodium acetate	$C_8H_4K_2O_{12}Sb_2.3H_2O$	15	2.44
20% (v/v) PEG 3 K; 0.1 M HEPES pH 7.5; 0.2 M sodium acetate	$C_8H_4K_2O_{12}Sb_2.3H_2O$	30	2.84
15% (v/v) PEG 4 K; 0.1 M sodium citrate pH 5.5; 0.2 M ammonium acetate	-	-	2.01
15% (v/v) PEG 4 K; 0.1 M sodium citrate pH 5.5; 0.2 M ammonium acetate	$C_8H_4K_2O_{12}Sb_2.3H_2O$	30	1.84
10% (v/v) PEG 4 K; 0.1 M HEPES pH 7.5; 10% isopropanol	-	-	NA
10% (v/v) PEG 4 K; 0.1 M HEPES pH 7.5; 10% isopropanol	$C_8H_4K_2O_{12}Sb_2.3H_2O$	45	NA
25% (v/v) PEG 4 K; 0.1M MES pH 8.5; 0.2 M magnesium chloride	$C_8H_4K_2O_{12}Sb_2.3H_2O$	30	2.81
25% (v/v) PEG 4 K; 0.1 M MES pH 8.5; 0.2 M magnesium chloride	-	-	3.02
30% (v/v) PEG 4 K; 0.1M Tris pH 8.5; 0.2M lithium sulphate	*A.f.*_azu	30	1.93
30% (v/v) PEG 4 K; 0.1 M Tris pH 8.5; 0.2 M lithium sulphate	*A.f.*_azu	35	1.78

15% (v/v) PEG 1.5 K	-	-	9.03
15% (v/v) PEG 1.5 K	*A.f.*_azu	35	2.47
15% (v/v) PEG 1.5 K	*A.f.*_azu	35	2.15
10% (v/v) PEG 4 K; 0.1 M HEPES pH 7.5; 5% isopropanol	*A.f.*_azu	40	2.42
10% (v/v) PEG 4 K; 0.1 M HEPES pH 7.5; 5% isopropanol	*A.f.*_azu	40	2.03
10% (v/v) PEG 8K; 0.05 M magnesium acetate; 0.1 M sodium acetate	-	-	1.50

3.4.2. Structure determination and refinement of *A.f.*_Aio crystal

Data were automatically processed at the beamline, using XDS program package[85], and scaled with Aimless from the CCP4 package[86,87]. Structure determination was accomplished via Molecular Replacement, using Phaser[88], with PDB code 1G8K serving as the searching model. An anomalous electron density map was generated using CAD and FFT[89]. Manual model building and interactive cycles of refinement were conducted using COOT[90] and Refmac5[91] as well as PHENIX[92], respectively. Structure quality was assessed with MolProbity[93].

3.4.3. Protein crystallisation attempts for arsenite oxidase mutants from *Rhizobium* sp. NT-26

Two structures from arsenite oxidase of NT-26, one of an AioB_F108A mutant and other of the WT protein are deposited in the Protein Data Bank (5NQD and 4AAY), and both crystals were obtained by sitting drop vapour diffusion method, using 2% PEG 400, 2 M ammonium sulphate, and 0.1 M HEPES Sodium pH 7.5 as crystallization condition (temperature not described[66,75]).

The first approach to try to obtain crystals of NT-26_AioA_D169A and NT-26_AioA_E453A mutants was, therefore, to mimic the same procedures and perform further optimisations. Variation of numerous parameters, such as protein concentration (5, 10, 12, 15, 20 $mg.ml^{-1}$), protein/precipitant drop ratio (1:1; 1:2; 1:3; 2:1; 2:2), precipitant concentration (1-3 M

ammonium sulphate), pH range (6.5-8.5), temperature (4 ºC, 20 ºC and a temperature ramp - 3 ºC per hour from 4 ºC to 30 ºC and back to 4 ºC, in the next day) were conducted using the vapor diffusion method in both hanging and sitting drop techniques, the latter performed manually and using the crystallisation robot (Oryx8; Douglas Instrument).

Manual drops were prepared by sitting drop technique with the protein expressed with K_2HPO_4 pH 7.4 as the purification buffer, and *E. coli* DH5α as the expression host cells. Protein mutants at 10 mg.ml^{-1} were mixed with a precipitant solution in a 2:1 ratio, and the pH was varied between 7 and 8 and ammonium sulphate between 1.5-2.5 M.

Several commercial screens were also tested to try to find new crystallisation conditions, namely JBScreen Classic 1-10 (Jena Bioscience), Morpheus (Molecular Dimensions), JCSG+ (Molecular Dimensions) and 80!. The drops were set-up automatically by the crystallisation robot and stored at 4 ºC; the protein used was expressed and purified as described in sections 3.1.2 and 3.1.4, with 0.66 µL of protein at 5 and 15 mg.ml^{-1}, and 0.33 µL of precipitant solution. JBScreen 6-10 was also tested using the same drop proportion and storage temperature, by co-crystallisation of mutant/electron-acceptor in a 1:1 ratio (130 µM of each protein); NT-26_AioA_D169A at 13.95 mg.ml^{-1} was mixed with NT-26_cytc_{552} 1.73 mg.ml^{-1} and separately with *A.f.*_azu at 1.78 mg.ml^{-1} concentration.

3.5. Kinetic assays of arsenite oxidases from *Rhizobium* sp. NT-26 via UV-Vis Spectroscopy

To assess the effect of the mutations of arsenite oxidase from NT-26, kinetic assays were conducted for the WT enzyme as well as for the mutants, aiming to determine the specific activity of the proteins.

The experimental procedure was conducted as previously described[75] in SpectraMax 190 Microplate Reader (Molecular Devices), using 96 plastic wells (Molecular Devices), at 25 ºC. The proteins were previously oxidised for 10 minutes by 10 mM $K_3[Fe(CN)_6]$, and the excess of the oxidising agent was removed using a PD-10 desalting column (GE Healthcare), according to the manufacturer's instructions, with 50 mM MES pH 5.5 serving as the elution buffer. The reactions were monitored at 600 nm to follow the reduction of the artificial electron acceptor: DCPIP - $\Delta\varepsilon_{red\text{-}ox}$=8.2 mM^{-1}.cm^{-1}[75] - using $Na_3As^{III}O_3$ as the substrate and 50 mM MES pH 5.5 as the reaction buffer.

First, 2 µL of enzyme - [NT-26_Aio] = 10 µM; [NT-26_AioA_D169A] = 17 µM; [NT-26_AioA_E453A] = 12 µM - and 50 µL of 1.2 mM DCPIP were incubated in 98 µL of reaction

buffer; when absorbance stabilised, the reaction was started by the addition of 50 µL of $Na_3As^{III}O_3$ at 2.5 mM.

3.6. Microscale Thermophoresis assays to assess the binding affinity of the two arsenite oxidases and *A.f.*_azu or NT-26_cyt*c*$_{552}$ complexes

The assays for determining the binding dissociation constants (K_d), were conducted for the physiological and non-physiological electron-transfer complexes: NT-26_Aio and *A.f.*_azu; NT-26_Aio and NT-26_cyt*c*$_{552}$; *A.f.*_Aio and *A.f.*_azu; *A.f.*_Aio and NT-26_cyt*c*$_{552}$.

During purification, arsenite oxidases are reduced by the presence of DTT in one of the buffers and NT-26_cyt*c*$_{552}$ and *A.f.*_azu are oxidised by $K_3[Fe(CN)_6]$. Since the proteins uphold their oxidation state after storage, no additional redox step was required.

3.6.1. Labelling of *A.f.*_azu and NT-26_cyt*c*$_{552}$

To selectively monitor the interaction between each pair of proteins, *A.f.*_azu and NT-26_cyt*c*$_{552}$, were marked with the fluorescent dye RED-tris-NTA (NanoTemper Technologies). Both the dye and the proteins were diluted in 50 mM Tris-HCl pH 7.8 buffer, to the concentration of 200 nM and 100 nM, respectively, and incubated for 30 minutes at room temperature, followed by centrifugation at 15000 g for 10 min.

3.6.2. Microscale thermophoresis (MST) measurements

The measurements were made on the Monolith™ NT.115 instrument (NanoTemper Technologies), under medium MST power, using standard capillaries that hold a reaction volume of around 10 µL. The excitation power using the Nano-RED filter set was 90% for *A.f.*_azu and 60% for NT-26_ cyt*c*$_{552}$ assays. All dilutions were performed in 50 mM Tris-HCl pH 7.8 buffer.

For binding studies, the concentration of NT-26_cyt*c*$_{552}$ or *A.f.*_azu was kept constant at 25 nM and arsenite oxidases were titrated in a 1:1 ratio by serial dilution throughout 16 capillaries (with a concentrations range of 202000-6 nM for NT-26_Aio and 72000-2 nM for *A.f.*_Aio). For each case, data of three independently pipetted measurements were analysed using the MO Affinity Analysis Software™ (NanoTemper Technologies) from the MST-on time signal at 5 seconds.

4. Results and discussion

4.1. Protein expression and purification

4.1.1. Expression and purification of arsenite oxidase proteins

As described in sections 3.1.2 and 3.1.3, the arsenite oxidases were subjected to different expression protocols that used different expression host cells, as well as different IPTG concentrations; all proteins were successfully expressed resulting in bright brownish cell *pellets*, which allowed proceeding with the purification assays after cell lysis and separation of the soluble and insoluble portions.

Regarding the protein's purifications, all were subjected to the same protocol, as described in section 3.1.4, based on affinity chromatography as the purification method. IMAC is a highly specific, easy and widely utilised method of separating complex biomolecular mixtures. Since the protein of interest, Aio, was recombinantly expressed with an His-tag in the N-terminus of the B subunit, it could be separated from other protein contaminants through reversible affinity of the histidine's side chain to the Ni^{2+} ions present in the column matrix. Imidazole was used to recover the protein; this histidine analogue acts by competitively interacting with the stationary phase. A 10% SDS-PAGE, followed by protein staining, was performed to evaluate the eluted samples quality and purity; this technique allows protein separation only by molecular weight.

Since the elution profiles for all arsenite oxidases are quite similar, NT-26_Aio will be taken as a representative example (a complete portrayal of the purification results for all proteins can be found in Appendix 7.3).

The first IMAC had the primary goal of separating AioAB-His-tag from the bulk of the host cell soluble proteins. Through analysis of the chromatogram depicted in Figure 4.1, it is possible to observe the 1st absorbance peak, where a low concentration of imidazole was used (20 mM); this peak corresponds to the elution of proteins with residual affinity to the column. Gel analysis (lane 3, 4.1 B), suggests that this sample contained mainly proteins from the host cells, but also a small amount of AioA (91 kDa), that is either expressed in a higher quantity than AioB (20 kDa) or, more likely, is not interacting with AioB-His-tag, due to the high salt concentration of the buffer (0.5 M NaCl) that disrupts the AioAB complex. The 2nd absorbance peak, containing proteins eluted at a 44 mM imidazole concentration, also shows the presence of contaminants as well as AioA (lane 4, Figure 4.1 B). When the imidazole concentration is increased up to 500 mM, AioAB is eluted from the column, along with other *E. coli* endogenous proteins (lane 5, Figure 4.1 B), corresponding to the 3rd absorbance peak.

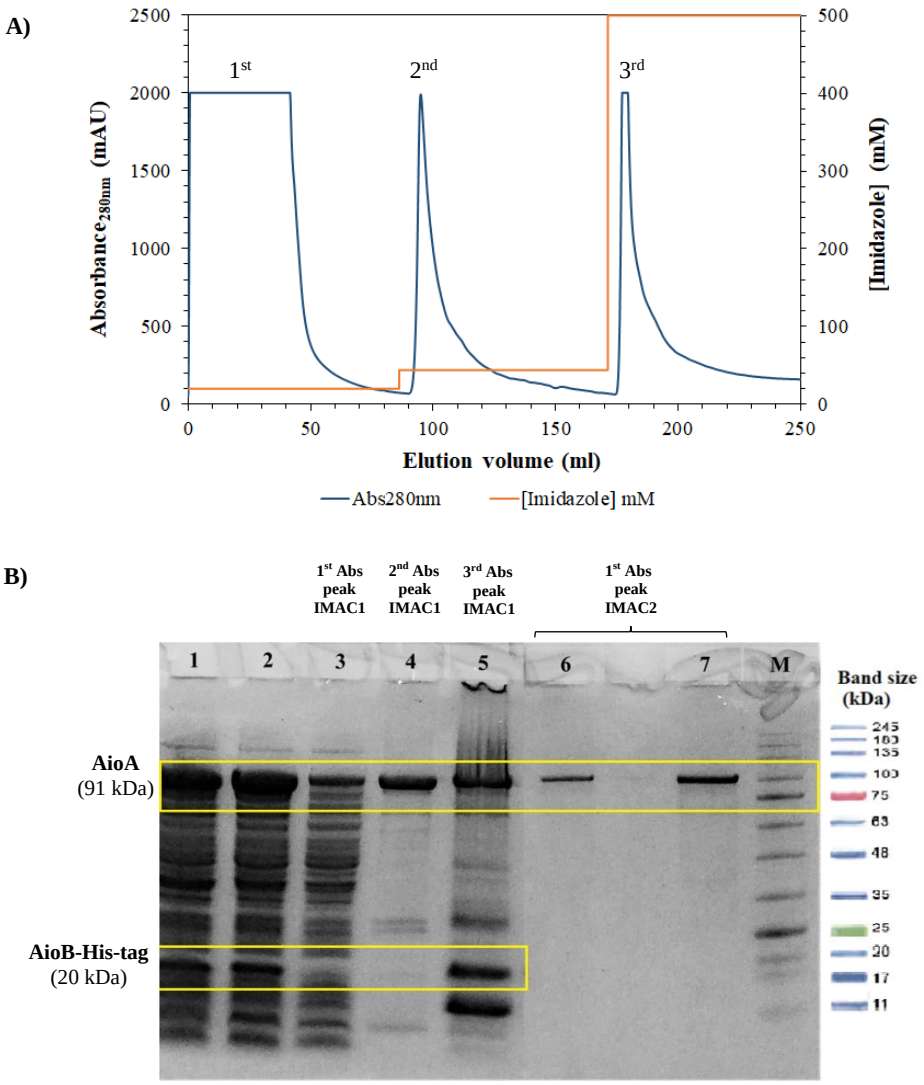


Figure 4.1. Elution profile of the recombinant His-tagged NT-26_Aio, purified by Ni^{2+}-based IMAC. A) chromatogram resulting from the protein purification; the blue line corresponds to absorbance at 280 nm and the orange line to the concentration of imidazole. B) 10% SDS-PAGE containing samples corresponding to the expression and purification fractions; 1) Insoluble fraction from cell lysis; 2) soluble fraction from cell lysis; 3) fraction 0-70 mL of the elution volume from the 1st IMAC (1st peak), corresponding to the proteins with minimal affinity for the column; 4) fraction 90-150 mL of the elution volume from the 1st IMAC (2nd peak), proteins eluted at 44 mM of imidazole, including AioA; 5) fraction 175-250 mL of the elution volume from the 1st IMAC (3rd peak), proteins eluted at 500 mM of imidazole, including AioAB and other contaminants; 6,7) fraction 7-9 mL of the elution volume from the 2nd IMAC (1st shoulder of the 1st peak), corresponding to proteins with minimal affinity for the column: AioA uncomplexed; M) Molecular weight marker Protein Marker II (NZYTech).

Considering that the arsenite oxidase sample still contained a reasonable number of other contaminants and the fact that the His-tag is undesirable for crystallisation assays, the protein was dialysed, overnight, with an endoprotease, TEV, to cleave it. A second IMAC was performed; the

resulting chromatogram can be seen in Figure 4.2. This time, since the protein lost the moiety that provides affinity to the column, it was eluted in the flow-through (9-45 mL in 20 mM imidazole); the SDS-PAGE analysis showed that the 1st shoulder of the 1st absorbance peak corresponds to some uncomplexed AioA (lanes 6,7 Figure 4.1 B and lane 1 Figure 4.2 B) and the rest of the peak contained the AioAB complex (lane 2-6 Figure 4.2 B). Lastly, the remaining contaminants and a small amount of uncomplexed the AioA, remained attached to the column, being only eluted at a 500 mM imidazole gradient (2nd chromatogram peak Figure 4.2 A; lane 7 and 8, Figure 4.2 B). The highly pure fractions, that contained both subunits, corresponding to lanes 2-6 of Figure 4.2 B, were concentrated up to 45 $mg.ml^{-1}$ in 50 mM Tris-HCl pH 7.8 buffer. The overall expression and purification yield of each protein are summarised in Table 4.1.

.

Table 4.1. Resulting expression optimisation conditions and respective yields for Aio proteins from NT-26.

Aio protein	***E. coli* strain**	**IPTG (µM)**	**Expression yield** **(mg of pure protein per litre of culture)**
NT-26_Aio	BL21 (DE3)	250	3.8
NT-26_AioA_D169A	Tuner (DE3)	250	7.2
NT-26_AioA_E453A	Tuner (DE3)	250	5.1
*A.f.*_Aio	DH5α	40	1.9

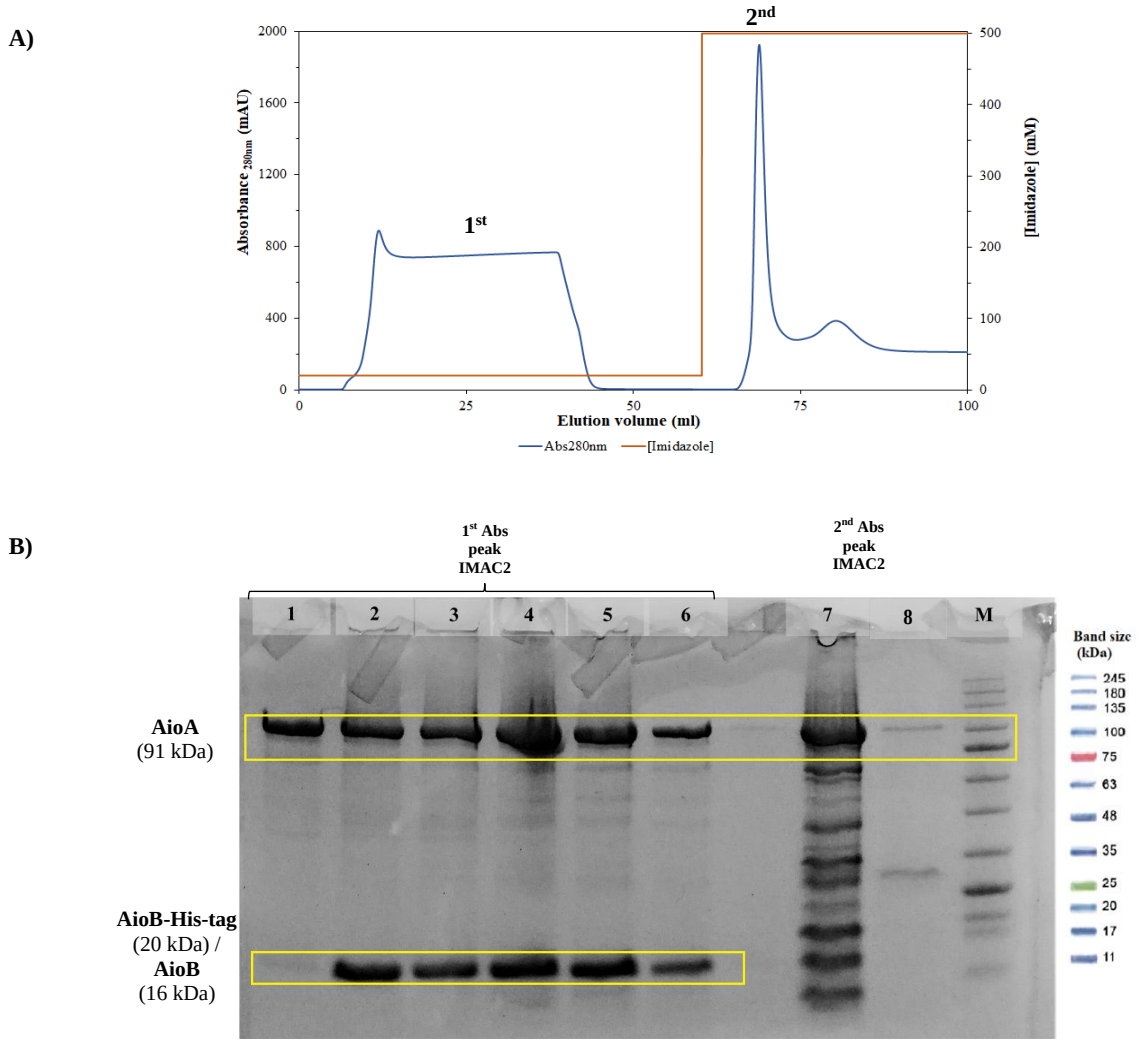


Figure 4.2. Second Ni^{2+}-based IMAC elution profile of the recombinant NT-26_Aio. A) Chromatogram resulting from the protein purification; the blue and orange lines correspond to absorbance at 280 nm and concentration of imidazole, respectively. B) 10% SDS-page containing samples correspondent to the purification samples; 1) fraction 9-12 mL of the elution volume from the 2nd IMAC (1st shoulder of the primary peak), corresponding to proteins with minimal affinity to the column: AioA uncomplexed; 2-6) fraction 12-45 mL of the elution volume, corresponding to proteins minimal affinity to the column: AioAB with cleaved His-tag; 7, 8) fraction 65-75 mL eluted at 500 mM imidazole from the 2nd IMAC (2nd peak), protein contaminants, including TEV and some remaining uncomplexed AioA; M) molecular weight marker Protein Marker II (NZYTech).

4.1.2. Expression and purification of NT-26_cytc_{552}

Heterologous expression protocol of NT-26_cytc_{552} in *E. coli* BL21 (DE3) cells resulted in red cell pellets that were later disrupted by sonication. The soluble part of the lysate was red, suggesting the presence of the protein of interest, and was subject to the purification steps.

As Aio enzymes, NT-26_cytc_{552} owns a His-tag in the N-terminal portion of its amino acid sequence, which allowed for an IMAC as the first purification technique. Figure 4.3 depicts the resulting chromatogram.

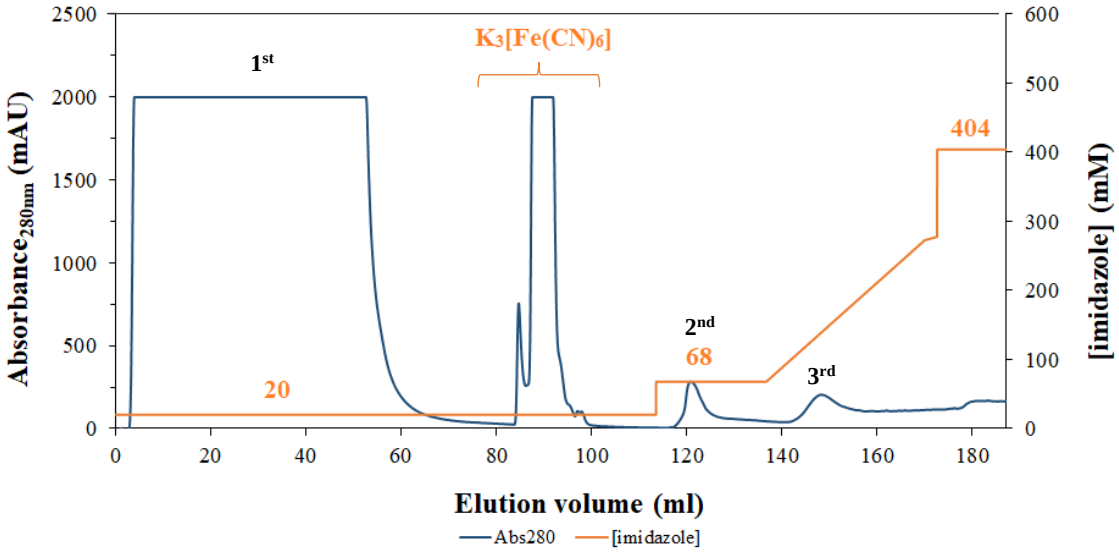


Figure 4.3. Ni^{2+}-based IMAC elution profile of the recombinant NT-26_cytc_{552}. The blue and orange lines correspond to absorbance at 280 nm and concentration of imidazole, respectively.

The 1st saturated peak of absorbance in the initial 60 mL of elution volume, in a buffer containing only 20 mM of imidazole, corresponds to the bulk of endogenous *E. coli* proteins from the soluble cellular extract, with minimal affinity to the column. When the absorbance had stabilised, a washing step was conducted with a solution of binding buffer containing 1 mM of $K_3[Fe(CN)_6]$. This step aimed to oxidise NT-26_cytc_{552} in the early stages of the purification, which was visually confirmed by a change of the colour of the protein bound to the column, from bright red to a more brownish colour. The 2nd absorbance peak, located at around 120 mL, corresponds to endogenous cellular proteins eluted at 68 mM imidazole. A gradient of imidazole promoted the elution of NT-26_cytc_{552}, which corresponds to the 3rd absorbance peak at approximately 150 ml of the elution volume. Cytochrome assays involved interaction with arsenite oxidases, so it was convenient that the protein was oxidised in an earlier stage in order to participate in redox reactions and receive electrons.

To further eliminate remaining contaminants, a size exclusion chromatography was performed (Figure 4.4.); the elution buffer was Tris-HCl pH 7.8, to match the arsenite oxidases buffer, allowing uniformity in subsequent assays.

The 1st peak in absorbance, centred at approximately 8.5 mL of the elution volume corresponds to *E. coli* protein contaminants of higher molecular weight than NT-26_cytc_{552}. The 10% SDS-PAGE gel lanes show the 0.5 mL collected fractions, sequentially, between 11 to 15 mL of eluted volume (2nd absorbance peak), that contains a highly pure solution of the protein. Samples 2-11 were pooled and concentrated up to 15 mg.mL^{-1}, providing the final yield of 0.4 mg of protein per litre of culture.

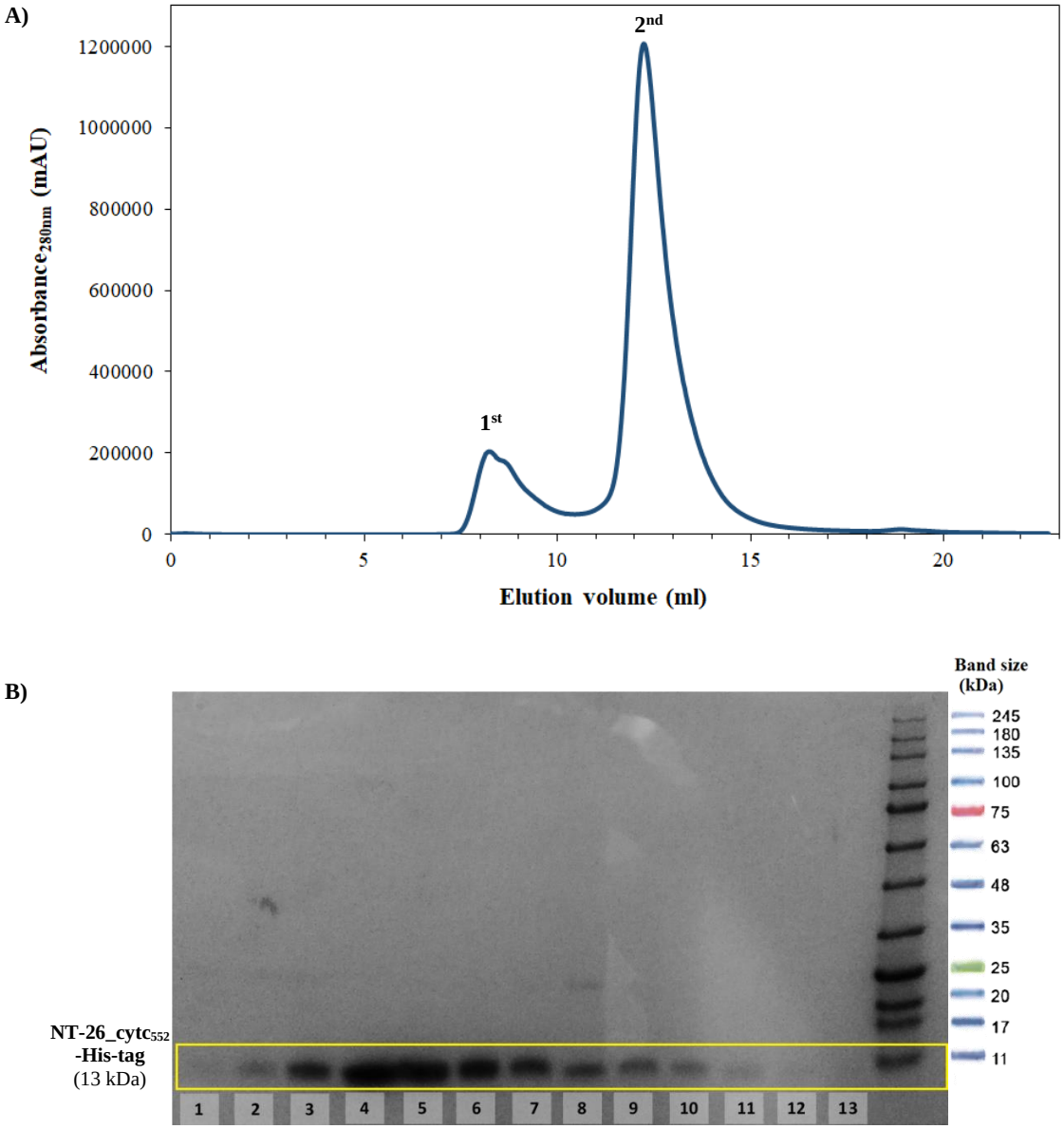


Figure 4.4. Size exclusion elution profile of recombinant NT-26_cytc_{552}. A) SEC resulting chromatogram; blue line corresponds to absorbance at 280 nm. B) 10 % SDS-PAGE containing samples correspondent to the 2nd peak of absorbance from 10 to 15 mL of elution volume, containing the protein of interest.

4.1.3. Expression and purification of *A.f._azu*

Recombinant expression of *A.f._azu* in *E. coli* BL21 (DE3) cells resulted in white cell pellets that were later subject to sonication. As for the case of NT-26_cytc_{552}, the soluble part of the cell extract lysate was purified in two steps, first an IMAC, and later an SEC, with an intermediate oxidation step using 1 mM of $K_3[Fe(CN)_6]$. The IMAC resulting chromatogram is depicted in Figure 4.5, where the absorbance peak corresponding to the elution of *A.f._azu* is centred at about 220 mL of the elution volume (imidazole gradient, between 68 and 500 mM).

SEC elution profile (Figure 4.6), shows three absorption peaks. The 1st peak (7-10 mL elution volume; lanes 1-5, Figure 4.6 B), contained the highest molecular weight proteins, corresponding to some endogenous *E. coli* contaminants and *A.f._azu* aggregates, as confirmed by analysis of 10% SDS-PAGE. The 2nd peak (10-12 mL of elution volume; lanes 6-9, Figure 4.6 B) contained *A.f._azu* with a high degree of purity, and the 3rd peak (12-15 mL; lanes 10-18, Figure 4.6 B) corresponds to *A.f._azu* in the presence of some lower weight protein contaminants. Samples 6-9, corresponding to the 2nd absorbance peak were pooled and concentrated up to 60 mg.mL^{-1}, providing the final yield of 2 mg of protein per litre of culture.

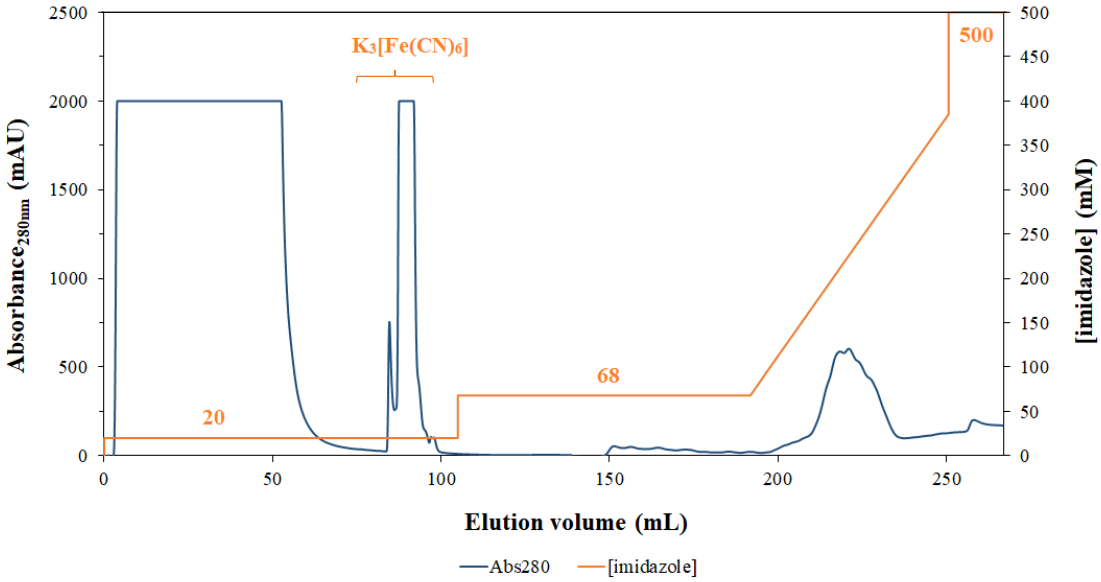


Figure 4.5. Ni^{2+}-based IMAC elution profile of the recombinant *A.f._azu*. The blue and orange lines correspond to absorbance at 280 nm and concentration of imidazole, respectively.

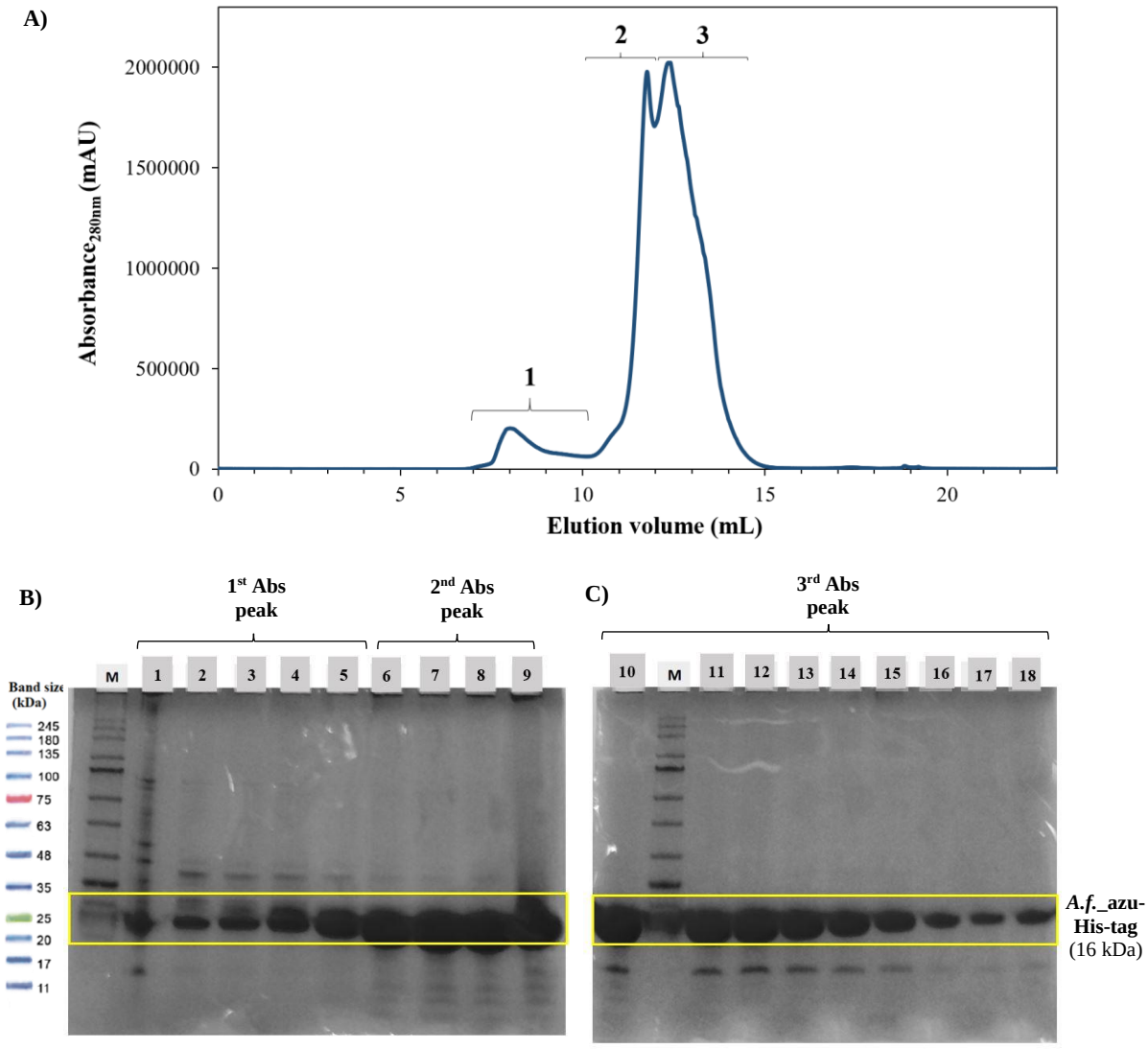


Figure 4.6. Size exclusion elution profile of recombinant *A.f.*_azu. A) SEC resulting chromatogram; blue line corresponds to absorbance at 280 nm. B) 10% SDS-PAGE containing samples correspondent to SEC absorbance peaks; 1-5) fractions eluted between 7 and 10 mL (1st peak), containing protein contaminants and *A.f.*_azu aggegates; 6-9) fractions eluted at 10 to 12 mL (2nd peak), containg *A.f.*_azu and some minor contaminants; 10-18) fractions eluted from 12 to 15 mL (3rd peak), containg *A.f.*_azu and lower weigth contaminants; M) molecular weight marker Protein Marker II (NZYTech).

4.2. Detection of oligomeric forms of NT-26_Aio_D169A by size exclusion chromatography

Due to experimental constraints, SEC of arsenite oxidase could not be performed as part of the purification protocol, as an appropriate column capable of separating proteins of arsenite oxidase molecular weight, was only available in the latter stages of experimental work. The chromatogram corresponding to the size exclusion elution profile of NT-26_AioA_D169A and the corresponding gels, performed under denaturating and non-denaturating conditions, are shown in Figure 4.7. The main goal of this assay was to check for sample homogeneity and discriminate between monomers, oligomers and aggregated forms of the protein. The elution profile of the sample should present only one peak, representative of one oligomeric state of the protein – expected to correspond to the $(\alpha\beta)_2$ NT-26_Aio form (214 kDa) – since the active site mutation should not influence the interaction between functional subunits.

Several peaks are depicted in the chromatogram, indicating a complex mixture of population states. The first 3 peaks - samples 9 to 21 - corresponding to high molecular weight proteins, present a pronounced band of AioA subunit and a faint lower molecular weight band of AioB in SDS-PAGE. Native-PAGE analysis suggests that different populations are present in the solution; in this technique, proteins are not subject to denaturation and migrate according to their charge, size and packing, so a uniform sample should present only one band. The 4th peak – samples 22 to 25 – shows the presence of AioAB in SDS-PAGE, and only one band in the native gel, apart from sample 22 that shows two bands, that might still be contaminated with the samples of the overlapping previous peak. The lower molecular weight proteins, eluted in samples 26 to 29, seem to contain only AioB.

Only a small fraction of the sample contained AioAB protein in a homogeneous oligomeric state, so further TSA studies were conducted to find a stabilising buffer for the protein (see section 4.3.).

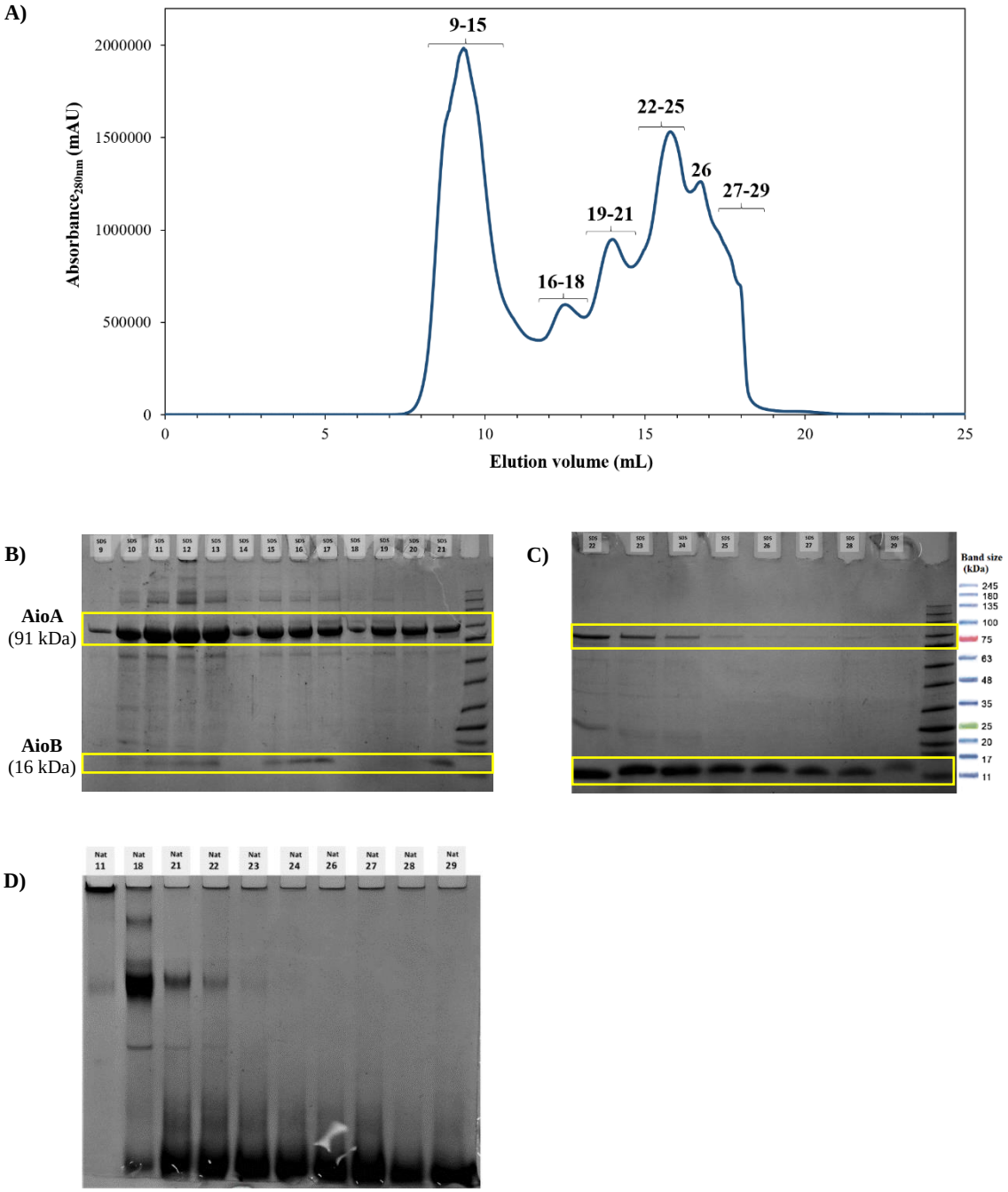


Figure 4.7. Size exclusion elution profile of NT-26_AioA_D169A. A) SEC resulting chromatogram; blue line corresponds to absorbance at 280 nm. B and C) 10% SDS-PAGE containing samples correspondent to SEC absorbance peaks. D) 10% native-PAGE containing samples correspondent to SEC absorbance peaks.

4.3. Aio stability monitored by Thermal Shift Assay

4.3.1. Purification buffer

To reduce Aio tendency to unfold or aggregate during purification, the AioA_D169A protein was tested in a wide range of buffers, using thermal shift assay to find a better purification buffer. In this technique, the changes in the fluorescence of a dye interacting with the protein are measured; the dye is quenched in an aqueous environment but strongly fluoresces when bound to exposed protein hydrophobic groups. A stabilising buffer causes an increase in the unfolding transition midpoint or melting temperature (T_m) of the protein, upon thermal denaturation[94,95]. The screen contained buffers with high (1 M), medium (0.3 M) or no NaCl content, so, in a first stage, only the effects of high ionic strength, necessary to avoid interspecific interactions during purification, were investigated. The first derivative results plotted against temperature, where the local minima represent the inflexion point of the fluorescence melting curves, are depicted in Figure 4.8 (see Appendix 7.4 for detailed information of all protein-buffer T_m).

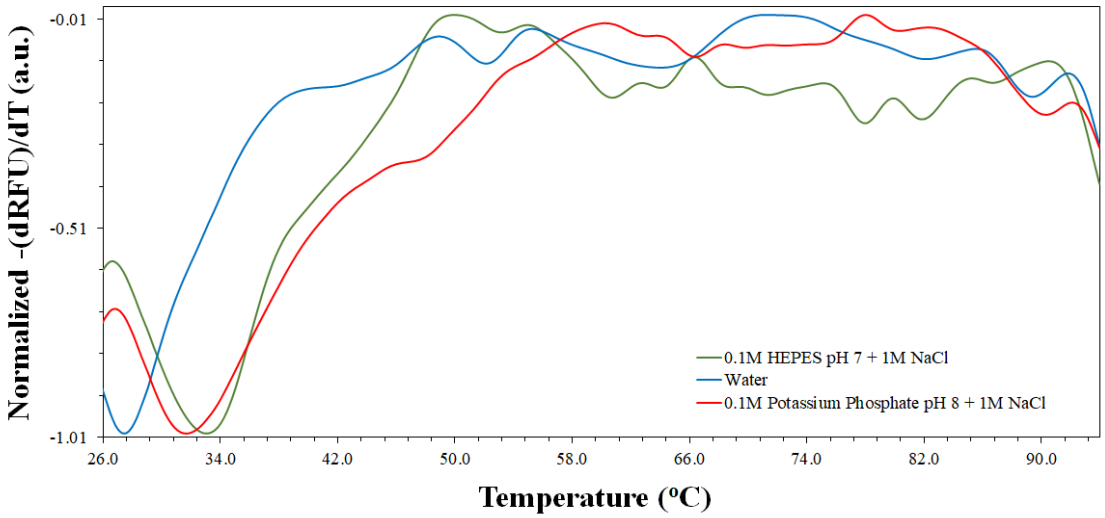


Figure 4.8. First derivative TSA results of NT-26_AioA_D169A in the presence of 0.1 M HEPES pH 7 + 1 M NaCl (green), 0.1 M potassium phosphate pH 8 + 1 M NaCl (red) and water (blue). The results were normalized for better visualisation.

The T_m of the protein in the previously used buffer, 0.1 M K_2HPO_4, 1 M NaCl pH 8, had a value of 31.5 ºC. From TSA results, the best hit corresponds to 0.1 M HEPES, 1 M NaCl pH 7 buffer, that resulted in a positive shift of ΔT_m = 6.1 ºC in comparison with water (used as a control), and $+\Delta T_m$ = 2.0 ºC regarding the buffer used initially. NT-26_Aio is allegedly present in $(\alpha\beta)_2$ oligomeric state, in solution[96], so one could expect several destabilisations (or denaturation)

events to occur: the heterodimers separation, the α and β subunits separation, and the unfolding of each subunit. Even though it is not possible to pinpoint them, several peaks are present in the TSA curves, in contrast to a smooth single peak (corresponding to a unique transition) of a monomeric protein. Although the difference in T_m was small, the phosphate buffer was abandoned, and HEPES was adopted as the purification buffer for all Aio proteins.

4.3.2. Final protein buffer

After numerous attempts of crystallising both mutants of Aio from NT-26 (AioA_D169A and AioA_E453A; see section 3.4.3) had proved to be unsuccessful, a size exclusion chromatography assay of AioA_D169A revealed that the protein had a complex mixture of oligomeric states in solution (see section 4.2.). TSA of the same protein was used to find a new stabilising final protein buffer. This had been avoided up to this point, to minimise changes to the protocol that led to the appearance of crystals of other forms of this protein.

This time, the focus was on buffers with no salt added. Salt has been proven to lower the activity of the protein with electron acceptors and also not to be physiologically imperative, since the protein is periplasmatic, which means its pH and salt conditions are in equilibrium with the surroundings, and NT-26 was isolated from a low salt environment[75,97]. Also, additives to further stabilise the protein were tested in a separate TSA.

Analysis of the protein thermogram depicted as the first derivative results plotted against temperature, reveals that NT-26_AioA_D169A is not stable in the current buffer (Figure 4.9). One can observe that the curve corresponding to Tris-HCl pH 7.5 buffer closely resembles the curve of the protein in water, presenting several bumps indicative of different oligomeric species in solution and, most likely, aggregates. From these two data sets, a prominent minimum in the graphic cannot be identified and consequently, neither a correct estimate of the protein T_m.

MES pH 5.8 buffer stood out as a better choice for the protein than the previously used one; investigation of the protein melting curve in this buffer, discloses two emerging well-defined minima, allowing for a T_m estimate of 53.1 ºC. Although a smooth line before the first peak is not observed, its relative intensity shows that most of the protein population underwent the same thermal phase transformation, meaning high sample uniformity.

A new TSA was conducted, using the same protein, but this time to search for a buffer additive that could additionally stabilise it. Addition of MES pH 5.8 to every well, during sample preparation, intended to enable the observation of additive effect together with the buffer. The experiment control revealed somewhat reproducible results, regarding the protein in the previously chosen final buffer; T_m was calculated from the first pronounced peak, assuming the value of 53.9 ºC, however, this time, only a small peak around 80 ºC could be seen. From all the tested

additives, $MgCl_2$ had the main positive shift in Aio T_m, that assumed the value of 55.6 ºC; its curve its quite similar to the control curve, indicating that it could slightly improve protein stability without significantly altering its dynamics in solution.

These results might be further improved if the assays were to be repeated with freshly purified protein instead of stored protein in Tris-HCl pH 7.8 buffer. Nonetheless, a buffer containing 50 mM MES pH 5.8 and 1 mM $MgCl_2$ seems to have a promising effect in protein stabilisation and oligomeric state uniformity.

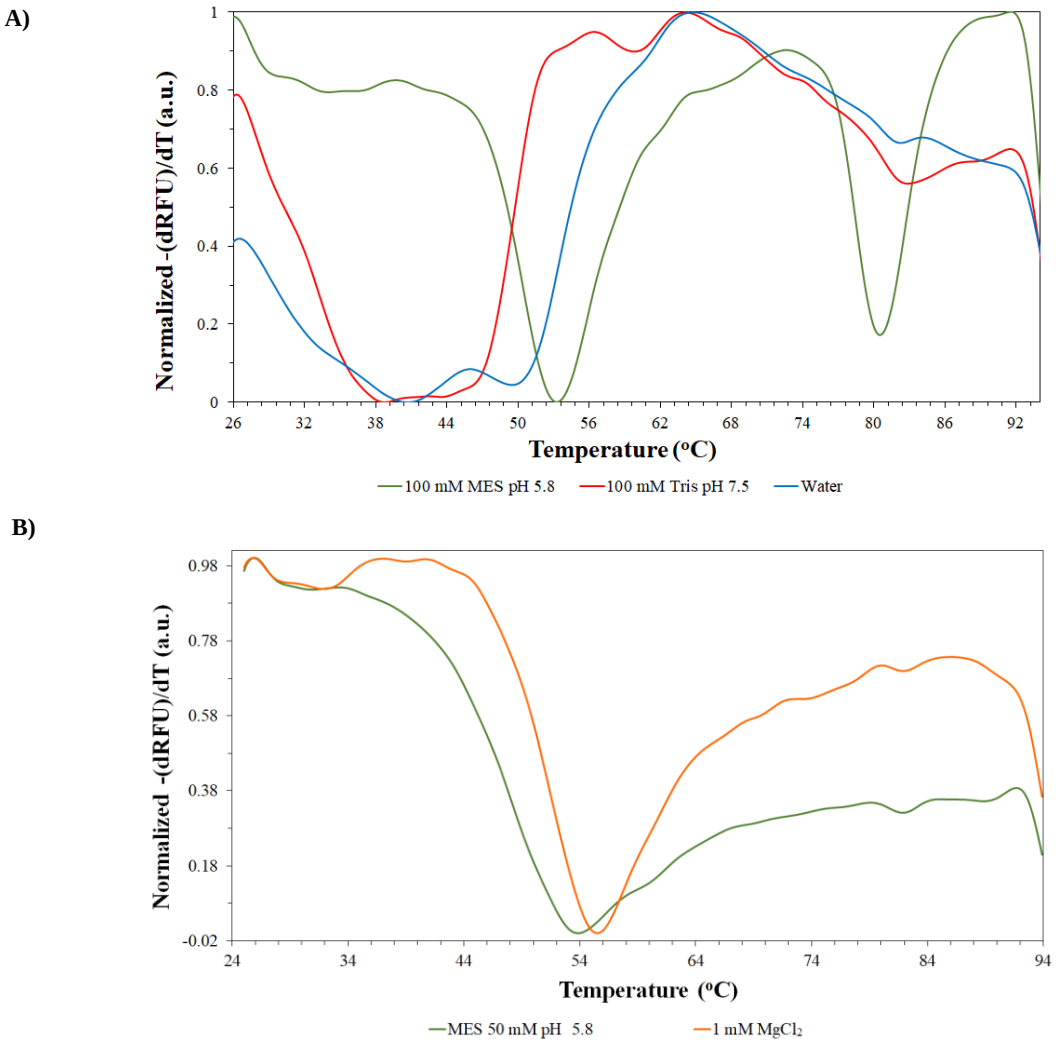


Figure 4.9 First derivative TSA results of NT-26_AioA_D169A in the presence of: A) 100 mM MES pH 5.8 (green), 100 mM Tris pH 7.5 (red) and water (blue). B) 50 mM MES pH 5.8 (green) and 50 mM MES pH 5.8 + 1 mM $MgCl_2$ (orange). The results were normalized for better visualisation.

4.4. Unravelling arsenite oxidases selectivity towards electron transfer partners

4.4.1. Microscale thermophoresis (MST) assays

Microscale thermophoresis assays were conducted to infer the binding dissociation constants of all possible combinations of electron transfer complexes between Aio proteins (from *A.f.* and NT-26), and their physiological and non-physiological partners: *A.f.*_azu and NT-26_cytc_{552}.

This technique is used to quantify a wide variety of biomolecular interactions, namely affinity between two proteins since it is a fast and highly sensitive method that requires only a small amount of sample. It consists on monitoring the changes in the fluorescence of interacting molecules upon induction of a temperature gradient by an Infrared (IR) laser; these variations occur, among other factors, due to fluctuations on the size, charge, hydration shell and conformation of the species in solution[98].

The titration assays required the labelling of the protein. We selected RED-tris-NTA since it binds to His-tag proteins, which, in this case, were the electron acceptor proteins. A problem could arise due to signal interference of the natural fluorophores of the proteins (tryptophan, phenylalanine, and tyrosine) however, the wavelength used to excite the molecules ($\lambda_{ex} \approx 625$ nm) does not have enough energy to promote significant absorption by the amino acids ($\lambda_{ex} \approx 275$ nm).

The MST-traces for each capillary (Figure 4.10), decreased linearly as a function of the ligand concentration, Aio, in all cases; the initial fluorescence, before IR irradiation (-3 to 0 s), remained constant which means that the samples did not suffer photobleaching; also, the smooth lines are indicative of a good signal to noise ratio, excluding the presence of significant aggregates.

In order to plot each dose-response curve and derive K_d (nM), the relative fluorescence value, F_1, measured when the sample had been heated for 5 seconds was divided by F_o, the relative fluorescence value before exposing the sample to a temperature gradient (time 0 s), giving rise to F_{norm} (‰). Automatic fitting of the bound fraction (normalised amplitude of F_{norm}) *vs* ligand concentration, allowed K_d extraction. From this point, it was possible to draw a dose-response curve that best fitted the experimental data (4.11). The results show that all pairs interact within the same range of affinity (1μM > K_d > 100 nM), with slight differences (Table 4.2.). This low to medium binding values are commonly seen in electron-transfer complexes due to the nature of the reactions that need to occur in a high-turnover manner, typically with multiple partners[99]. The behaviour of Aio_ *A.f.* seems to be promiscuous towards both azu_ *A.f.* and NT-26_cytc_{552} since

the K_d values are very similar (693±23 nM for azu_ *A.f.* and 660±40 nM for NT-26_cytc_{552}). This result is in agreement with the literature since this enzyme is known to reduce azurin as well as an unidentified cytochrome, *in vivo*[100]. On the other hand, the analysis of NT-26_Aio affinity profile suggests a small preference towards its physiological electron acceptor than the non-physiological one (K_d of 261±23 nM for NT-26_cytc_{552} and 341±29 for *A.f.*_azu).

It has been reported that *A.f.*_Aio can interact with multiple azurin proteins, *A.f.*_azu and azurin from *Pseudomonas aeruginosa*, as well as a physiological cytochrome – although it is not able to reduce cytochrome *c* from horse heart. Based on our findings, we propose a novel interaction with NT-26_cytc_{552}, K_d of 660±40 nM, which is in line with the capability of these enzymes to transfer electrons to different types of partners. On the contrary, NT-26_Aio has only been described to reduce cytochromes – NT-26_cytc_{552} and cyt *c* from horse heart (K_d of 2.3 ± 0.7 µM for NT-26_Aio - cyt *c* horse heart complex) – and could not interact with azurin from *P. aeruginosa*[75,80]. On the contrary, our results suggest that NT-26_Aio interacts with *A.f.*_azu, with a K_d of 341±29 nM, showing that electron acceptor diversity is also plausible in this enzyme.

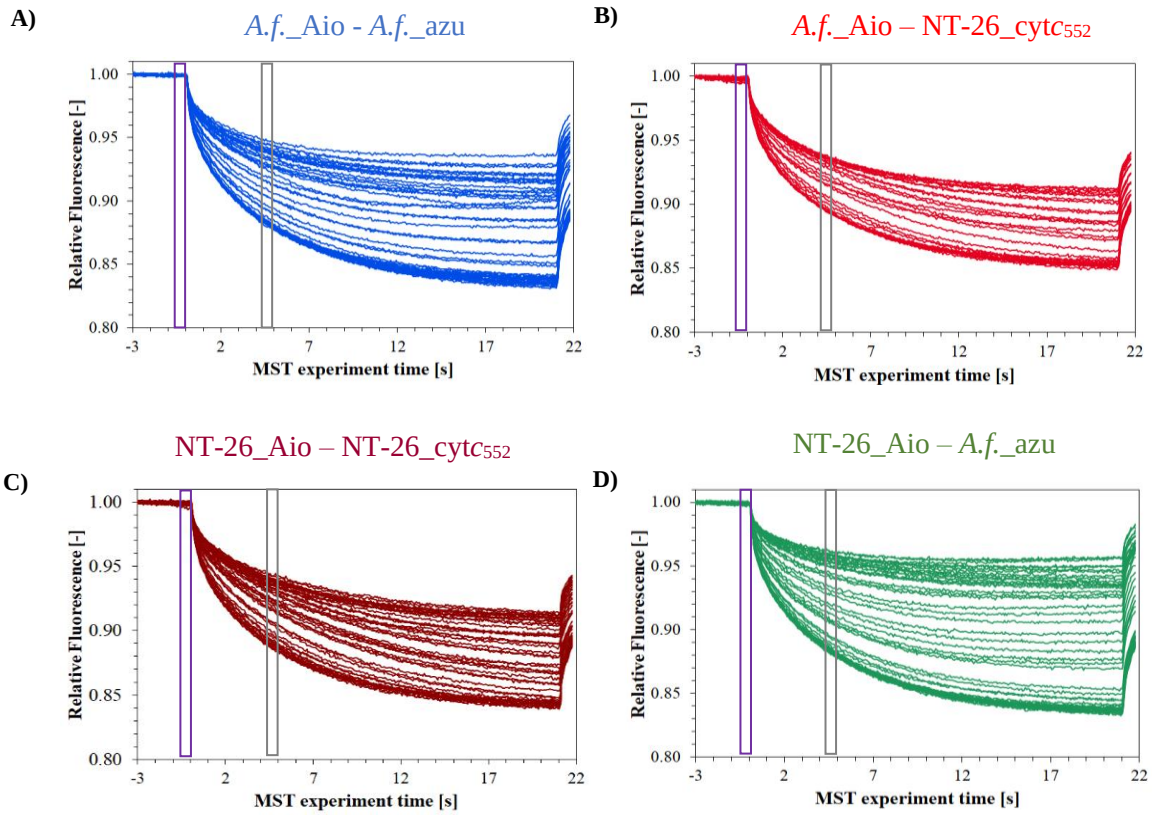


Figure 4.10. Thermographs of the interaction of arsenite oxidases with *A.f.*_azu or NT-26_cytc_{552} in 50 mM Tris-HCl pH 7.8: A) *A.f.*_Aio – *A.f.*_azu; B) *A.f.*_Aio – NT-26_cytc_{552}; C) NT-26_Aio – NT-26_ cytc_{552}; D) NT-26_Aio – *A.f.*_azu. Each graphic contains data from 3 independently prepared assays .The cold region is set to 0 seconds (purple) and the hot region to 5 seconds (grey).

This selectivity/promiscuity dichotomy has been noticed in other arsenite oxidases, like Aio from *Ralstonia* sp. 22[65]. Several authors pointed to the fact that all the exploited electron partners have similar redox potentials[65,75],so the different affinities may exist due to differences in the surface of the proteins.

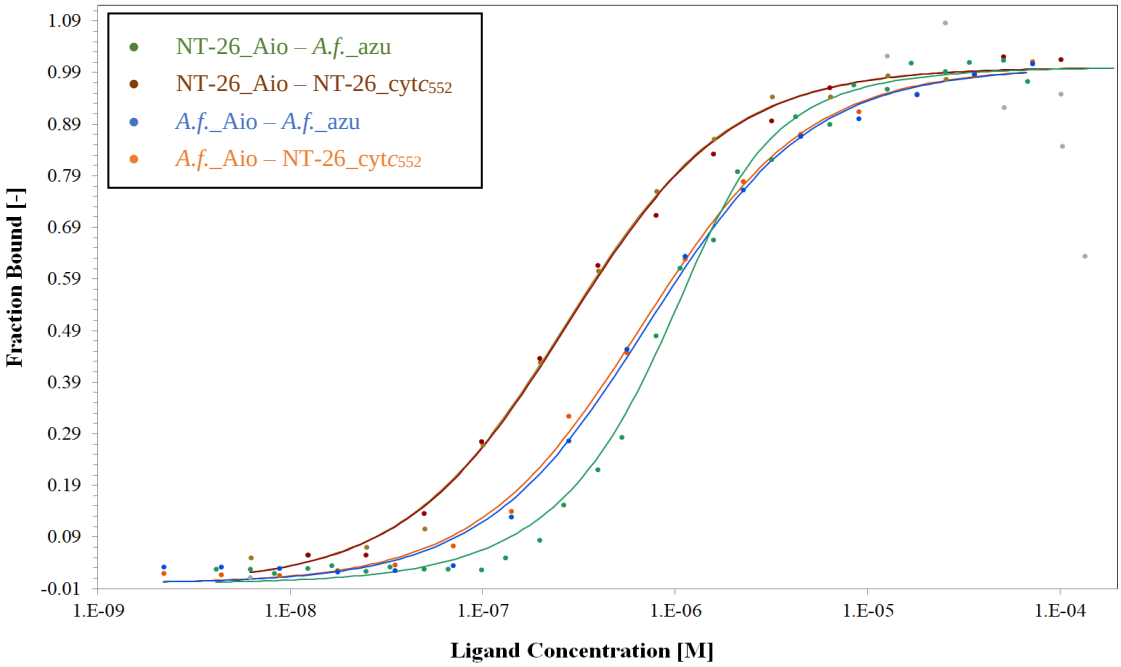


Figure 4.11. Microscale thermophoresis binding measurements of Aio enzymes ([202 to 2.19x10^{-3} µM]) and their final electron acceptors: NT-26_cyt*c*$_{552}$-RED-Tris-NTA and *A.f.*_azu-RED-Tris-NTA ([25 nM]). The fraction bound [-] is plotted as a function of Aio enzymes concentration (M), and curves are fitted using the "K_d fit model" of the NanoTemper Analysis software. MST power was set to medium and excitation power was 90% for *A.f.*_azu and 60% for NT-26_cyt*c*$_{552}$.

Table 4.2. Binding dissociation constants of arsenite oxidases and electron transfer partner pairs.

Arsenite oxidase	Electron acceptor	K_d (nM)
NT-26_Aio	NT-26_cyt*c*$_{552}$	261 ± 23
NT-26_Aio	*A.f.*_azu	341 ± 29
*A.f.*_Aio	*A.f.*_azu	693 ± 23
*A.f.*_Aio	NT-26_cyt*c*$_{552}$	660 ± 40

4.5. Protein Crystallography

4.5.1. Crystallisation attempts of NT-26_AioA_D169A and NT-26_AioA_E453A

Both NT-26_Aio mutants, NT-26_AioA_D169A and NT-26_AioA_E453A, were subject to several crystallisation attempts (see section 3.4.3). These resulted in drops often with heavy to light precipitates, sometimes colourless, that did not follow any trend regarding precipitant, protein concentration or pH. The only evident tendency was that the protein precipitated a few hours after setting up the drops when the temperature was above 4 ºC.

TSA (section 4.3) and SEC (section 4.2) analysis of NT-26_AioA_D169A in 50 mM Tris-HCl pH 7.8, expressed in *E. coli* Tuner (DE3) cells and stored at -80 ºC, revealed that the protein sample had a complex mixture of multiple oligomeric states and protein aggregates, in solution. This sample heterogeneity might explain the failure of the crystallisation assays since it affects the ability of the protein to pack in an organised manner. TSA results also suggest that changing the final protein buffer greatly improves NT-26_AioA_D169A stability and lowers the number of oligomeric states present in solution. Sample quality needs to be further investigated before continuing crystallisation trials.

4.5.2. *A.f.*_Aio•Sb crystal structure

Heterologously expressed *A.f.*_Aio crystallises in several different newfound conditions, distinct from the ones reported in the literature for the homologous expressed protein[67]; the common denominator to drive crystal formation seems to be PEG, both as an additive or a precipitating agent, as well as the 4 ºC temperature.

Some crystals were soaked with $C_8H_4K_2O_{12}Sb_2.3H_2O$, a source of Sb^{III}, in order to obtain a structure of the enzyme bound with this substrate analogue in the active site and help to clarify the reaction mechanism. Using an As^{III}-based compound would seem like the obvious choice, however, since the catalysis is about 6500 times faster than with Sb^{III}, as reported for NT-26_Aio[34], it would be hard to visualise clear structural differences, or even, possibly, reaction intermediaries. Enzyme activation required a previous addition of an oxidising agent so that the redox centres could be available to harbour the electrons from Sb^{III}; in this case, the commonly used $K_3[Fe(CN)_6]$ was chosen since its redox reaction does not involve changes in the metal's coordination state.

Another set of crystals was also subject to a different soaking assay, with *A.f.*_azu protein as the soaking agent. This approach could provide structural enlightenment of the interaction between the transient complex since azurin is a small protein (dimensions of 38 x 22 Å; see Appendix 7.5) that might, in principle, be able to enter the solvent channels and interact with the redox partner near the solvent-exposed Rieske [2Fe-2S] centre.

The crystal subject to examination in the following sections (Figure 4.12) appeared after 1 week in 15% (v/v) PEG 4 K, 0.1 M sodium citrate pH 5.5, and 0.2 M ammonium acetate; was soaked with $K_3[Fe(CN)_6]$ for 5 min, $C_8H_4K_2O_{12}Sb_2.3H_2O$ for 30 min, and diffracted up to 1.84 Å resolution (Figure 4.13).

Data were processed in the space group, P1, with the cell constants a = 90.4 Å, b = 109.2 Å, c = 117.3 Å, α = 97.7º, β = 90.0º, γ = 96.3º (Figure 4.14), as the *A.f.*_Aio structure deposited at the PDB (IG8K).

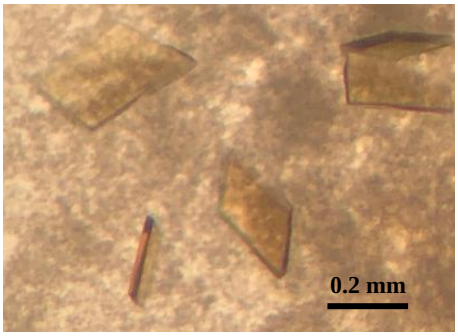


Figure 4.12. Crystals of *A.f.*_Aio protein in 15% (v/v) PEG 4 K, 0.1 M sodium citrate pH 5.5 and 0.2 M ammonium acetate before soaking experiments.

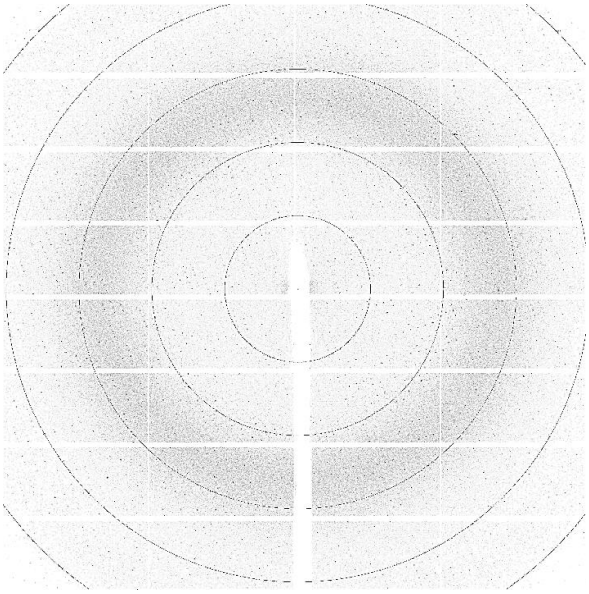
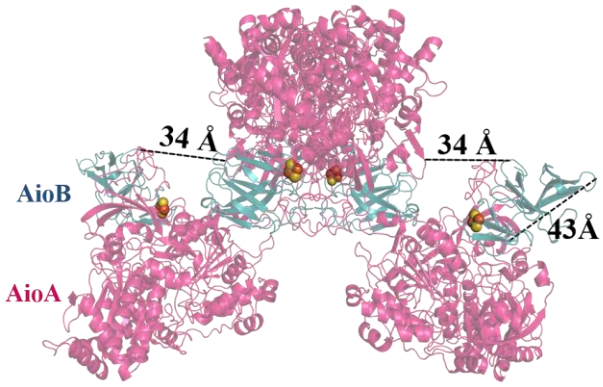


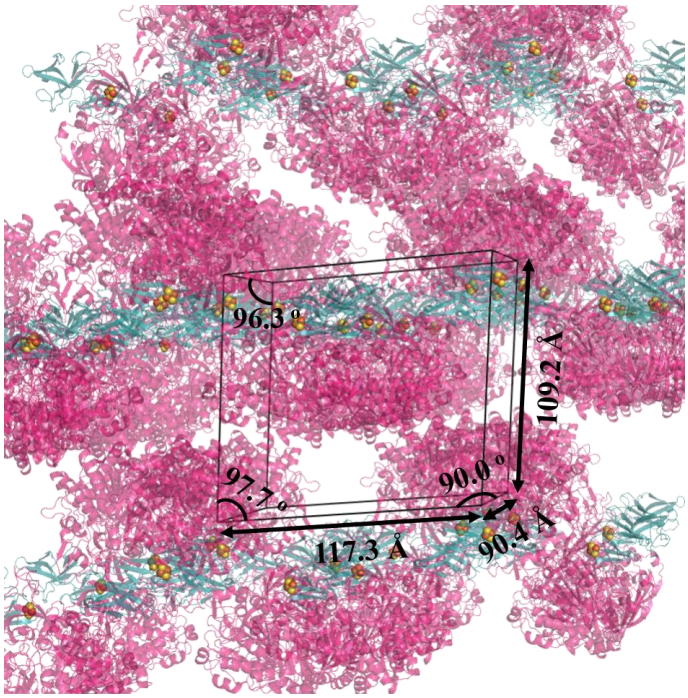
Figure 4.13. Difraction pattern of the *A.f.*_Aio•Sb adduct crystal, up to 1.84 Å.

A)



Asymmetric unit

B)



Crystal lattice

Figure 4.14. Structural arrangement of *A.f.*_Aio molecules in the crystal lattice: A) asymmetric unit composed of 4 arsenite oxidase molecules; B) crystal lattice molecular packing, showing the unit cell and its constants - a = 90.4 Å, b = 109.2 Å, c = 117.3 Å, α = 97.7º, β = 90.0 º, γ = 96.3º. All ligands, cofactors and solvent molecules were omitted for clarity, apart from the Rieske [2Fe-2S] centre, represented in space filling mode. Images were designed in PyMOL software.

Examination of some statistic parameters, summarised in Table 4.3, assured the quality of the diffraction data. The average ratio of reflection intensity to its estimated error, $\langle I/\sigma(I)\rangle$, is about 4, for all the reflections used, and 1.7 in the outer resolution shell, which indicates meaningful detection/measurement of the spots in the detector. *CC ½* is a correlation coefficient between the average intensities of two randomly selected data sets, each consisting in half of the measurements of each unique reflection; it can vary between 1, at excellent resolution, and 0, at poor resolution, and it is an important indicator of signal/noise and a good parameter for resolution *cut-off* [101,102]; the value of 0.848 for the last resolution shell is highly satisfactory. A complete data set should have as many experimentally measured unique reflections as the number of theoretically proposed, for a particular crystal system. In this case, the completeness is 95.3 % on average, and 94.8 % for the outer shell; this value is considered acceptable for the low symmetry space group – P1[103]. R_{pim} describes the precision of the averaged merged intensity measurements in relation to their symmetry mates, so the discrepancy of 6.1% for the overall, and 2.7% in the highest resolution shell shows high agreement within the data.

In order to solve the phase problem, the polypeptide chain of the deposited *A.f.*_Aio structure was used as a search model for molecular replacement (MR) (ligands and cofactors were removed). It was possible to verify the presence of 4 Aio_AB molecules in the asymmetric unit, designated AB, CD, EF, GH, and 54.13% solvent content, as in the deposited structure.

Several steps of refinement were conducted, where continuous adjustment of the model to the electron density allowed for a better agreement between the calculated structure factors (F_{calc}) and the observed structure factors (F_{obs}). During the process, the difference map, F_{obs}-F_{calc}, showed positive ($F_{obs}>F_{calc}$) and negative ($F_{obs}<F_{calc}$) unmodeled density regions. Water molecules, other compounds present in the crystallisation condition – Tris, PEG, glycerol – and the protein's cofactors – P_{MGD}, Q_{MGD}, [3Fe-4S] and Rieske [2Fe-2S] centres – were added to the model, taking into consideration the presence of density in the $2F_{obs}$-F_{calc} map and positive density in the F_{obs}-F_{calc} map, and analysing their temperature and occupation factors. Generation of the anomalous map showed anomalous signal of the high electron density atoms: Mo, Fe and S in some well-ordered residues. These are capable of absorbing and diffracting X-rays in a way that their phases are unique from lower density atoms such as carbon, oxygen or nitrogen.

In the same manner, a blob in the active site of all AioA chains, showing anomalous signal at roughly 3.5 Å of the Mo atom, was modelled as an Sb analogue. Additionally, 5 oxygen atoms were added to the model near Sb and Mo atoms.

The model is in the last stages of refinement, nearly ready for deposition, and refinement statistics were analysed in detail. R_{work} is a global statistical parameter that shows how well the calculated model (F_{calc}) compares to the experimental data (F_{obs}); R_{work} = 15.5 % validates the model since this value is considered very good when compared to structures of the same

resolution. R_{free} is similar to R_{work}, but it was calculated using only 5% of the data not included in the refinement, in order to avoid model *bias*; the value of 18.9 % is also in agreement with structures at the same resolution. The discrepancy between R_{work} and R_{free} is smaller than 5%, which additionally ensures a well-refined model.

The grand majority of protein residues are found in the allowed region of the Ramachandran plot, with only 0.13% in the disallowed region. R.m.s.d. of bond lengths and angles are calculated in comparison with ideal values from other structures; these should remain under 2 Å and 4°, respectively, a condition which was fulfilled.

In sum, good quality data allowed for the construction of a model of *A.f.*_Aio bound to an Sb substrate analogue, that successfully explained the diffraction data.

Table 4.3. Data collection and refinement statistics for *A.f.*_Aio•Sb adduct crystal. Values in parentheses correspond to the highest resolution shell.

Data collection and refinement parameters	
Beamline	Biomax (Max IV)
X-ray wavelength (Å)	0.918
Resolution (Å)	48.47-1.87 (1.87-1.84)
Space Group	P1
Unit cell parameters *a, b, c* (Å) *α, β, γ* (°)	 90.4, 109.2, 117.3 97.7, 90.0, 96.3
No. of molecules per asymmetric unit	4
No. of observed reflections	737465 (36892)
No. of unique reflections	365490 (17998)
Completeness (%)	95.3 (94.8)
$\langle I/\sigma(I) \rangle$	4.2 (1.7)
R_{pim} (%)*	6.1 (2.7)
CC ½	0.993 (0.848)
Multiplicity	2
Solvent content (%)	54.13

Refinement	
R_{work} (%)**	15.5
R_{free} (%)***	18.9
R.m.s.d. deviations	
Bond lengths (Å)	0.017
Bond angles (°)	1.954
Ramachandran plot (%)	
Residues in favoured regions	96.64
Residues allowed regions	3.23
Residues in disallowed regions	0.13

*R_{pim}= $\Sigma_{hkl}[1/(N-1)]^{1/2}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)$, where N is the multiplicity measured.

**$R_{work} = \Sigma\ ||F_{calc}|-|F_{obs}|| / \Sigma|F_{obs}| \times 100$, where F_{calc} and F_{obs} are the calculated and observed structure factor amplitudes, respectively.

***R_{free} is calculated for a randomly chosen 5% of the reflections for each dataset.

4.5.3. Active centre of arsenite oxidase bound to Sb

To understand the implications of the Aio•Sb adduct complex in the catalytic mechanism, both *A.f.*_Aio and NT-26_Aio structures in complex with antimonite will be analysed (Figure 4.15).

NT-26_Aio was previously crystallised, as described earlier[75], and soaked for 30 min with 50 mM $K_3[Fe(CN)_6]$, and later with 10 mM $C_8H_4K_2O_{12}Sb_2.3H_2O$, for 2h. Diffraction data was collected up to 1.89 Å, and the structure was solved in the same space group, with the same cell constants as the deposited NT-26_Aio model (PDB code: 4AAY; space group: $P22_12_1$; unit cell constants: a = 141.5 Å, b = 148.3 Å, c = 232.6 Å, $\alpha = \beta = \gamma = 90^o$). Each asymmetric unit harbours 2 dimers of heterodimers (AB-CD, EF-GH) with the substrate analogue bound to the active site, surrounded by 5 oxygen atoms. These assays were carried out by student Marta Vieira and Professor Teresa Santos Silva[104].

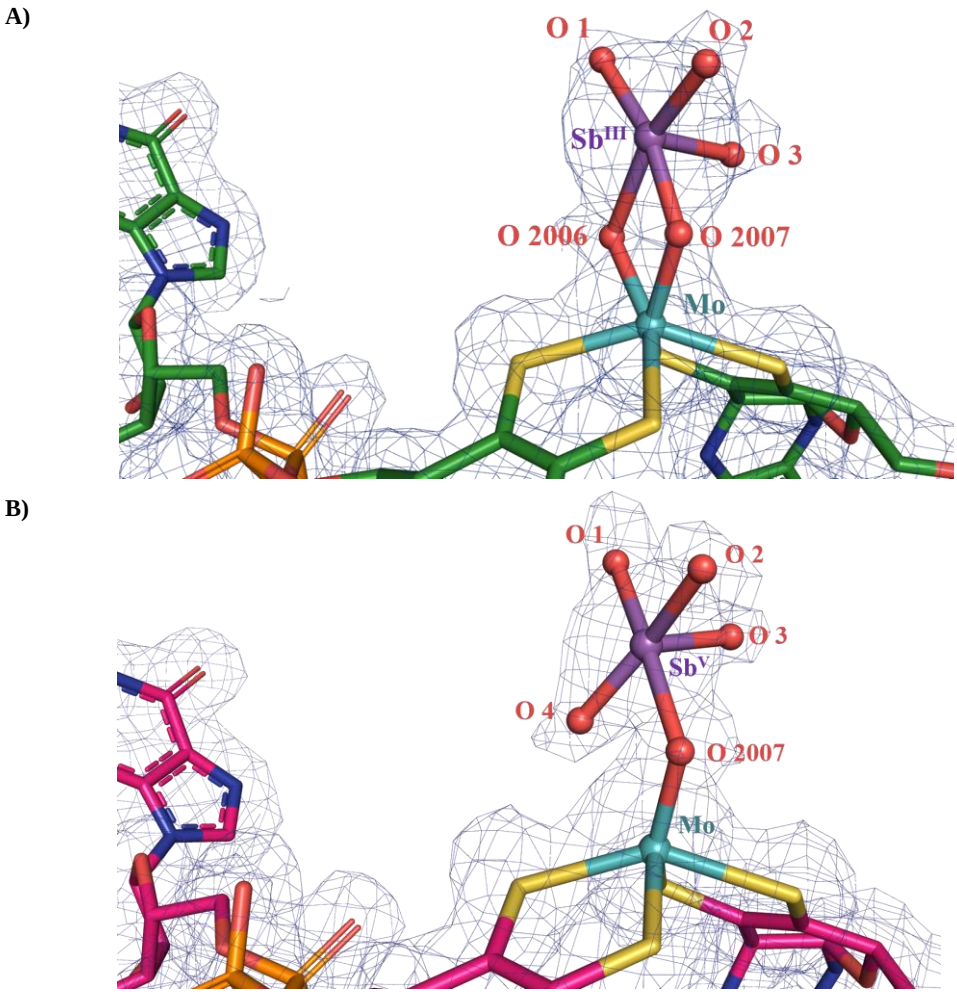


Figure 4.15. Overall representation of arsenite oxidases active site in complex to an Sb oxoanion: A) NT-26_Aio•Sb^{III}; B) *A.f.*_Aio•Sb^{V}. Electron density map $2F_o$-F_c contoured at 1σ.

As mentioned, both structures showed Mo and Sb near 5 oxygens atoms, that could belong to a Sb^{III} or Sb^{V} oxo analogue, so a thorough investigation of the geometry of the ligand and bond lengths was conducted. The results are summarised in Table 4.5. Past EXAFS studies of both enzymes in the presence of As^{III}[70,72] have revealed some of the chemical parameters that will be taken into consideration in the next paragraphs; NT-26_Aio EXAFS assays in the presence of Sb^{III}, conducted by our collaborators from the UCL, have revealed an As-similar behaviour of the substrate analogue (data not published). For the reduced state of the enzyme, in the presence of As^{III}, the literature suggests that Mo is bound to an oxo group at 1.75 Å and to an oxo/hydroxo group at 1.81 - 2.02 Å also interacting with As, at 3.25 Å, *via* the oxo group. Our results show that the Sb atom sits vertically to the Mo atom, at 3.5 Å and 3.3 Å, in *A.f.*_Aio or NT-26_Aio, respectively. The bond lengths within Sb oxo anions are very similar among the different molecules in the asymmetric unit and between the two protein structures (*ca* 2 Å, Table 4.4).

Table 4.4. Bond length analysis of all the atoms involved in Moco and ligand.

	***A.f.*_Aio polypeptide chain**				**NT-26_Aio polypeptide chain**			
Bond length (Å)	A	C	E	G	A	C	E	G
Sb - O 1	2.05	1.92	2.03	1.91	2.00	1.97	1.95	1.99
Sb - O 2	2.02	2.00	1.98	1.81	2.00	1.98	2.01	2.06
Sb - O 3	1.88	1.99	1.88	1.99	1.97	2.00	2.03	1.91
Sb - O 4 / Sb - O 2006	1.78	1.81	1.82	1.99	2.15	2.34	2.21	2.26
Sb - O 2007	2.10	2.12	2.16	2.04	2.08	1.84	1.90	2.05
Mo - Sb	3.54	3.55	3.52	3.51	3.25	3.34	3.34	3.28
Mo - O 2007	1.82	2.00	1.94	1.95	2.07	2.16	1.97	2.26
Mo - O 4 / Mo - O 2006	3.22	2.74	2.67	3.05	2.29	1.87	1.94	2.10

In the case of NT-26_Aio, the Mo atom is *cis* coordinated to two oxygen atoms (named O2006 and O2007). The distance between the metal and the oxygens suggests that these correspond to hydroxo groups which can arise from X-ray photo-reduction. Sb is slightly closer to O2007 than O2006, with bond lengths comprehended between 1.84-2.08 Å for Sb-O2007, and 2.15-2.34 Å, for Sb-O2006. The Mo atom of *A.f.*_Aio has a similar geometry, also coordinated to a hydroxo group (O2007) with similar bond distances as in NT-26 _Aio. The second oxygen (O4) is also found in the Mo vicinity, but much closer to antimony than to the metal: the distance between Mo and O4 is about 2.67-3.22 Å while between Sb and O4 is only 1.78-1.99 Å. This result suggests that O4 is not bound to *A.f.*_Aio active centre. The bond angles of the Sb oxo compounds seems to suggest a distorted tetrahedral geometry in the case of *A.f.*_Aio structure and a distorted trigonal planar geometry when considering NT-26_Aio.

These results suggest that different reaction intermediates, bound through a μ-oxo bridge to the active enzyme site, could be present in each structure. In the case of the NT-26_Aio structure, it appears that we have captured the Sb^{III} oxoanion before the nucleophilic attack of the metalloid to the Mo-oxo group (O2006), while for *A.f.*_Aio, the chemical species close to the active site seems to be an Sb^{V} oxoanion formed before product displacement. Both Sb-reaction intermediates are interacting with the polypeptide chain by an intricate network of hydrogen bonds (Figure 4.16; Table 4.5).

Table 4.5. Conserved residues between both arsenite oxidases interacting with the substrate through a network of H-bonds.

Conserved protein residue	**Residue numbering**	
	NT-26_Aio	***A.f.*_Aio**
Glu	453	425
Asp	169	165
Glu	207	203
Arg	201	197
Asn	200	196
His	451	423
Gli	449	422
Arg	447	419
Lys	413	385

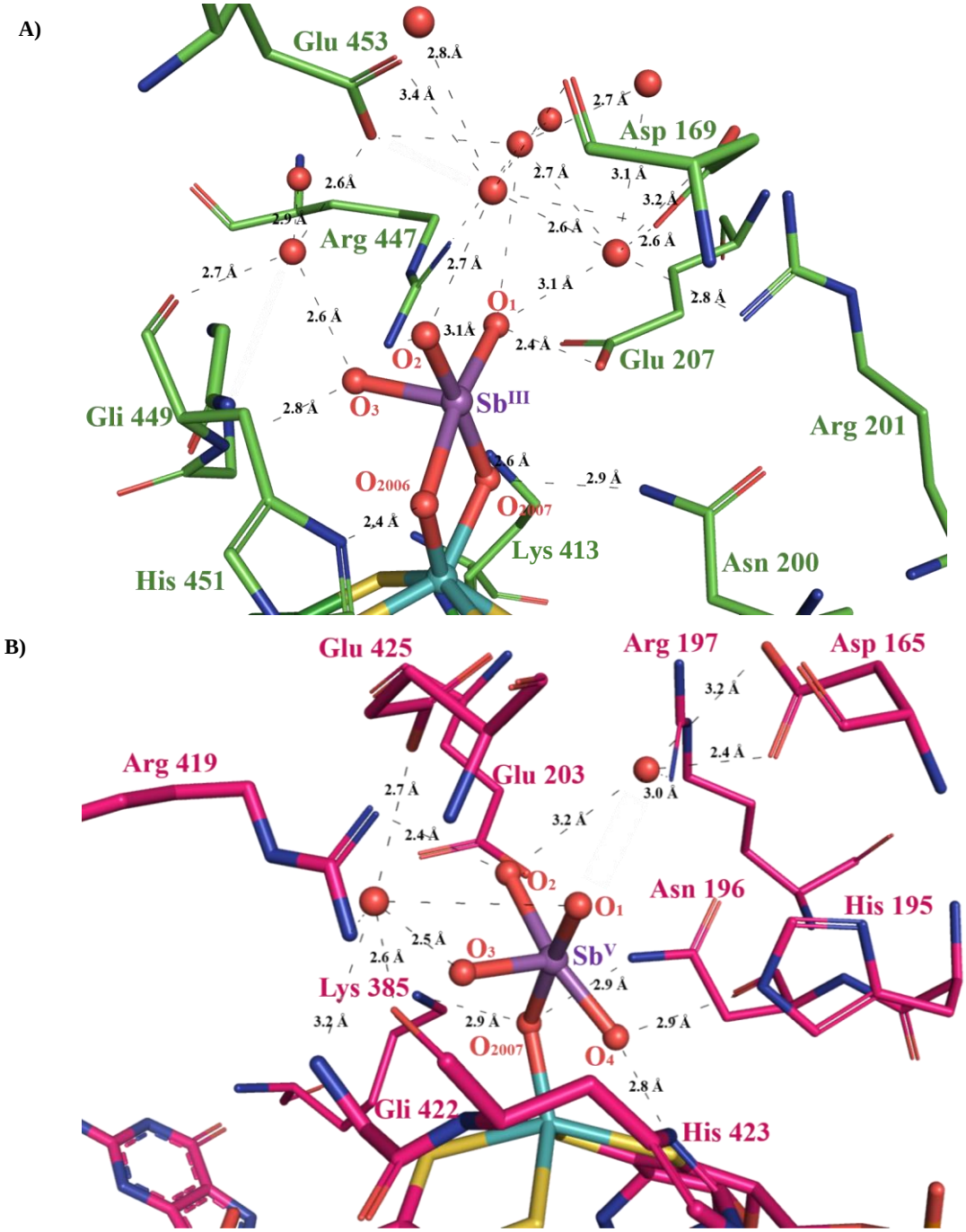


Figure 4.16. Schematic representation of the hydrogen-bonding interactions of the polyptide chain to water molecules and the ligands near the active centre. A) NT-26_Aio•Sb^{III} oxoanion; B) *A.f.*_Aio•Sb^{V} oxoanion.

*A.f.*_Aio kinetic parameters have not been fully characterized, so the fact that this enzyme could be faster than NT-26_Aio in the catalysis of Sb^{III} might explain why the reaction underwent quicker in this soaking experiment, despite the shorter soaking time.

The occupancy of the ligand atoms is also different when comparing the two structures (see Table 4.6). Sb^{V} has an occ of *ca* 0.4 in all AioA chains of the *A.f.* enzyme, which indicates

that the metalloid is only present in about 40% of the crystal molecules. The oxygen ligands, however, were refined with an occ value of 1 because these are in the same position as water molecules found in the ligand-free deposited structure. This means that 40% of the O atoms arise from the $SbO(OH)_3^-$, while the other 60% belong to free waters. Analysis of the B factors corroborates this hypothesis since the values observed for the oxygens of the ligands (28.91-41.90 $Å^2$) are in the same range as the vicinal atoms and the waters molecules in the structure (average B factor of 31.43 $Å^2$). In the case of Sb^{III}, present in NT-26_Aio structure, all ligand atoms were refined with an occ value of 0.8, meaning that 80% of the crystal molecules are bound to the ligand. The average B factor of the oxygen ligands (21.40 - 33.23 $Å^2$) is also in the same range as the B factors for vicinal atoms and water molecules (39.47 $Å^2$).

The discrepancy in occ values of Sb between both enzymes could be explained by the fact that Sb^{III} becomes less soluble as the pH decreases. The concentrations of the soaking agents were common to both assays, but the crystallisation pH was 5.5 in *A.f.*_Aio, and 7.5 in the case of NT-26_Aio, so more Sb^{III} molecules would be available for the later. Moreover, some product may already have been released from the active site.

Table 4.6. B factor analysis of all the atoms involved in Moco and ligand, and corresponding occupancy values in parenthesis.

	***A.f.*_Aio polypeptide chain**				**NT-26_Aio polypeptide chain**			
B (Å^2) (occ)	A	C	E	G	A	C	E	G
Sb	27.65 (0.38)	29.94 (0.39)	25.41 (0.37)	29.85 (0.39)	24.79 (0.8)	27.33 (0.8)	26.22 (0.8)	25.66 (0.8)
O 1	35.02 (1)	41.89 (1)	33.44 (1)	38.04 (1)	30.24 (0.8)	27.08 (0.8)	29.60 (0.8)	28.57 (0.8)
O 2	28.91 (1)	33.80 (1)	28.38 (1)	37.45 (1)	22.19 (0.8)	27.50 (0.8)	27.61 (0.8)	26.88 (0.8)
O 3	35.81 (1)	38.58 (1)	33.44 (1)	34.86 (1)	22.45 (0.8)	29.82 (0.8)	26.42 (0.8)	27.58 (0.8)
O 4 / O 2006	36.32 (1)	41.90 (1)	37.37 (1)	37.72 (1)	29.98 (1)	25.76 (1)	28.50 (1)	33.23 (1)
Mo	19.37 (1)	18.39 (1)	19.20 (1)	19.66 (1)	21.40 (1)	24.56 (1)	27.75 (1)	22.86 (1)
O 2007	25.85 (1)	28.22 (1)	26.56 (1)	25.85 (1)	25.11 (1)	19.45 (1)	26.25 (1)	27.84 (1)

4.6. Effect of the active site NT-26_AioA mutations on the catalytic enzyme mechanism monitored by UV-Visible spectroscopy

Structural analysis suggested a possible important implication of NT-26_AioA Glu 453 and Asp 169 on substrate/product interaction in the active site; both are located at the end of a highly solvated funnel-like cavity, interacting with water molecules and facing the Mo atom. Site-directed mutagenesis was performed to help clarify the role of these amino acids, resulting in two mutants: NT-26_AioA_D169A and NT-26_AioA_E453A (cloned prior to this thesis).

DCPIP and $Na_3As^{III}O_3$ were used as non-physiological electron acceptor and substrate, respectively, as described in section 3.5. The reductive reaction of DCPIP (see Figure 4.17), was followed by a decrease in absorbance at 600 nm, after adding $Na_3As^{III}O_3$ at *ca* 400 s – initial time required for absorbance stabilisation. The spectroscopic profile of each reaction is depicted in Figure 4.18.

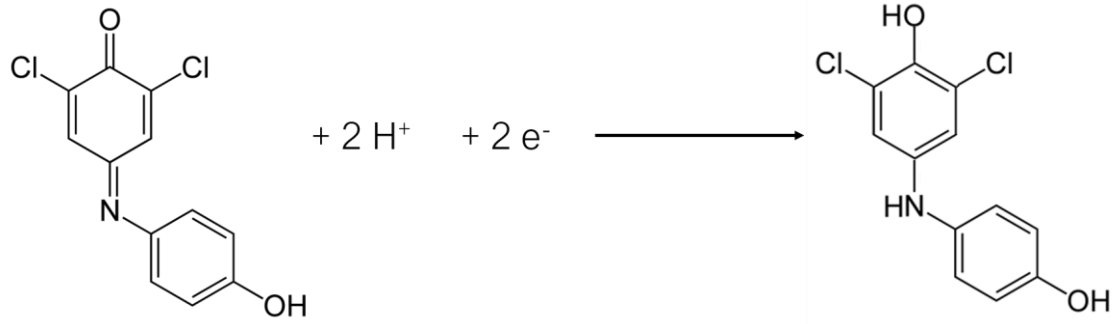


Figure 4.17. Redox reaction of DCPIP, where the electron acceptor becomes reduced by accepting two electrons prevenient from arsenite oxidase activity.

The specific activity of each enzyme, summarised in Table 4.7, was calculated considering the final enzyme concentration and the slope of the linear maximum initial velocity, at 400 to 500 s, before product accumulation. The specific activity of the WT enzyme obtained - 2.4 $\mu mol.min^{-1}.mg^{-1}$ - is relatively similar to the previously obtained value of 1.7 $\mu mol.min^{-1}.mg^{-1}$, for NT-26_Aio heterogeneously expressed in *E. coli* DH5α cells, under the same reaction conditions. The enzyme activity of the two mutants is lower than the WT: only 46% in the case of NT-26_AioA_D169A and 8% for NT-26_AioA_E453A. From the structural analysis, both amino acids seem to participate in a complex network of hydrogen-bonds involved in substrate interaction (Figure 4.19). Upon mutation to Ala, a neutral amino acid with an aliphatic side chain, these interactions can no longer occur, thus, possibly explaining the decrease in enzyme activity. It

should be stressed, however, that these are preliminary results, corresponding to only one technical assay and further assays are required to confirm the results.

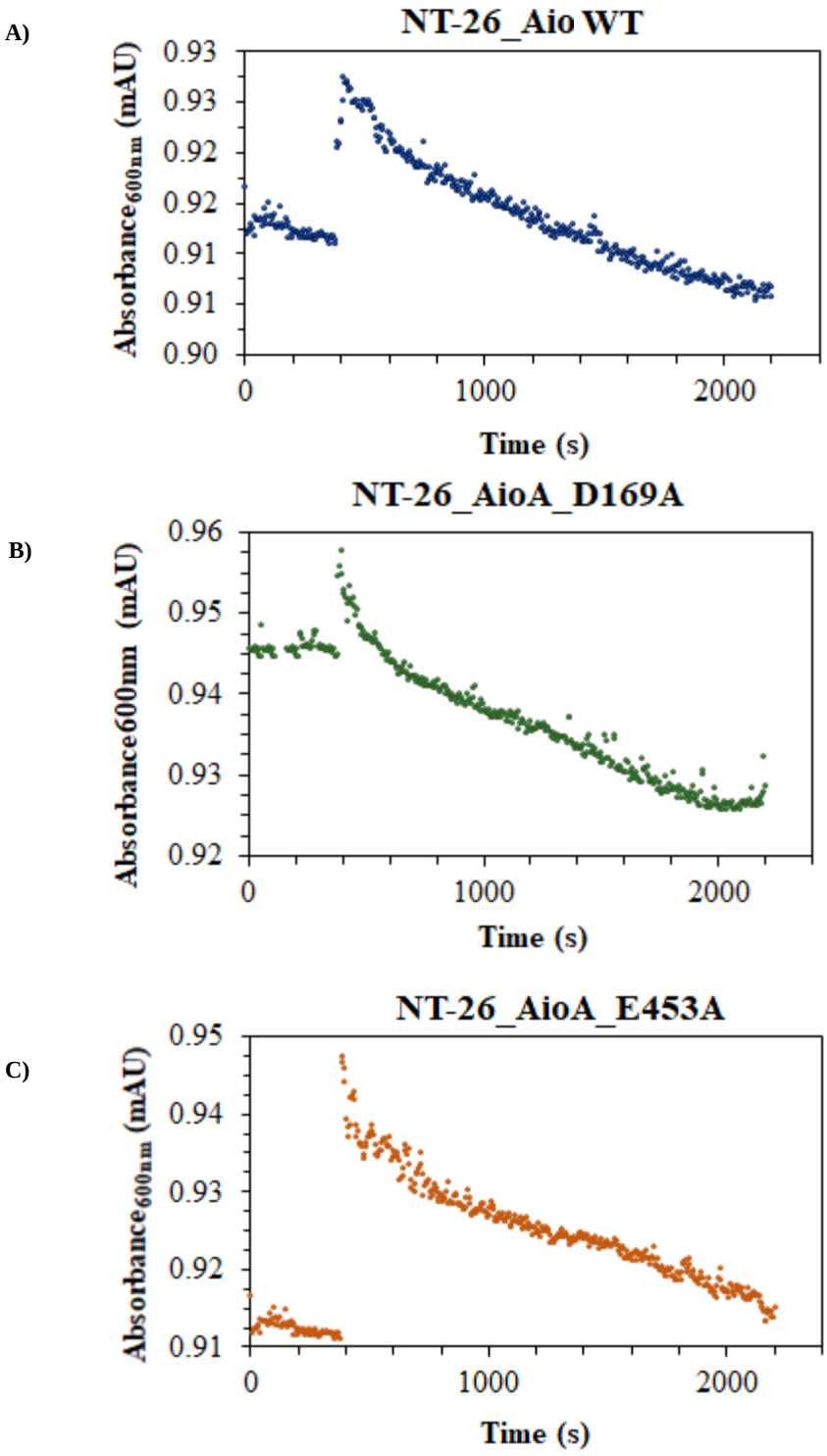


Figure 4.18. Kinetic profile of the reaction of arsenite oxidases with 0.3 mM of DCPIP in 50 mM MES pH 5.5, monitored at 600 nm, at 25 ºC: A) NT-26_Aio, B) NT-26_AioA_D169A, C) NT-26_AioA_E453A. Upon absorbance stabilization, after ≈ 400 s, 2.5 mM of $NaAsO_3$ was added to the reaction mixture.

Table 4.7. Summary of specific activity assays of arsenite oxidases enzymes. Trials were conducted at 25 ºC in 50 mM MES pH 5.5, using 0.3 mM of DCPIP as the electron acceptor and 2.5 mM of $NaAs^{III}O_3$ as the substrate

	NT-26_Aio WT	NT-26_Aio_D169A	NT-26_Aio_E453A
Enzyme final concentration ($mg.mL^{-1}$)	1.81	10.92	12.99
Slope (min^{-1})	0.0018	0.0042	0.0012
Specific activity ($\mu mol.min^{-1}.mg^{-1}$)	2.4	0.9	0.2
Specific activity (%)	100	46	8

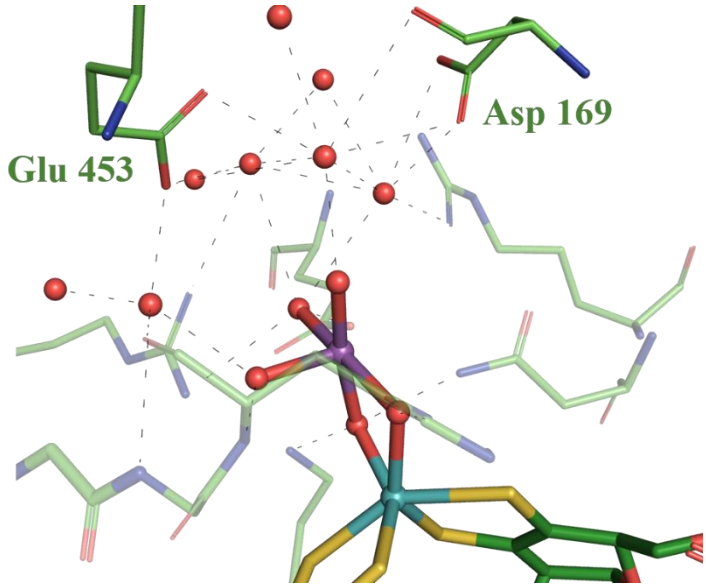


Figure 4.19. Aio_NT-26•Sb^{III} active site structure, showing Glu 453 and Asp 169 complex H-bond interactions with the substrate analogue, mediated by water molecules.

5. Conclusions and Future Perspectives

Arsenic is a ubiquitous environmental threat that dangerously continuous to harm populations worldwide. Perfectly efficient and sustainable detection or remediation techniques are not currently available, so the use of microorganisms such as *Rhizobium* sp. NT-26 or *Alcaligenes faecalis,* that possess arsenite oxidase enzymes could pose as an alternative. In addition to catalysing As^{III} to As^{V}, these enzymes can also oxidise the Sb^{III} analogue, a characteristic that might increase the value of these biomolecules.

NT-26_Aio WT and mutant enzymes – NT-26_AioA_D169A and NT-26_AioA_E453A – were expressed and purified in different conditions. Their yields (3.8, 7.2 and 5.1 $mg.L^{-1}$, respectively) were successfully improved compared to previous work, through the use of *E. coli* Tuner (DE3) expression host cells, instead of *E. coli* DH5α cells, and using different IPTG concentrations: 250 µM *vs* 40 µM. TSA and SEC results suggest that the proteins present several oligomerisation states in solution, as well as aggregates. This oligomerisation/low stability issue was slightly improved when using 50 mM MES pH 5.5, 1 mM $MgCl_2$ buffer; however, more tests need to be conducted, using freshly purified protein and considering new storage conditions.

The two NT-26_Aio mutants were subject to crystallisation assays. To this date, no crystals of the mutant forms were produced, possibly due to the inadequate state of the samples subjected to the crystallisation trials. Upon sample quality improvement, future crystallisation strategies should contemplate repeating the previously used conditions but also, new crystallisation/optimisation methods, such as micro or macro-seeding, as well as cross-seeding with *A.f.*_Aio crystals.

Several crystals of *A.f.*_Aio enzyme were obtained using the vapour diffusion method, sitting drop technique, at 4 ^{o}C, in numerous PEG-based crystallisation conditions. Since the structure of the protein had been already determined, the focus of this study resided on how this protein interacts with the substrate analogue – Sb^{III} – and its physiological electron acceptor – *A.f.*_azu. To pursue this goal, several soaking trials were performed, using different soaking times and agents, namely *A.f.*_azu and $C_8H_4K_2O_{12}Sb_2.3H_2O$; the enzyme was beforehand oxidised by the addition of $K_3[Fe(CN)_6]$. Regarding the crystals treated with *A.f.*_azu, some cracked during soaking, possibly suggesting the entrance of the small protein in the crystal solvent channels; SDS-PAGE or X-ray fluorescence should be carried out to confirm this hypothesis and help to improve soaking conditions.

At the time of writing, only the structure of *A.f.*_Aio was solved and refined, presenting the final R_{work} and R_{free} values of 15.5% and 18.9%, respectively. The model corresponds to *A.f.*_Aio soaked with $C_8H_4K_2O_{12}Sb_2.3H_2O$ for 30 min and revealed extra electron density as well as anomalous signal near the active centre of the enzyme; this was modelled as an Sb oxoanion analogue, with the Sb vertically to the Mo atom. Additionally, a previously obtained structure of

NT-26_Aio, also containing an Sb oxoanion near the active site, was investigated and used for comparison. Analysis of the geometry and bond lengths of the ligand atoms, suggested that both ligands were present in the form of reaction intermediates of the proposed mechanism: in the case of NT-26_Aio, the ligand is in the form of a Sb^{III} derivate [$Sb(OH)_3$], which might correspond to the substrate before attacking the non-spectator oxygen atom of Mo; regarding *A.f.*_Aio, an Sb^{V} derivate [$SbO(OH)_3$] is coordinated to Mo, through the spectator oxygen ligand (preceding product displacement). In *A.f.*_Aio model, the low occupancy of the antimony atom is perhaps a consequence of the acidic pH of the crystallisation condition, that lowers the solubility of the metalloid or the partial product release. To overcome this issue, a higher concentration of the substrate could be used, or other diffraction data sets from crystals that appeared at more basic crystallisation conditions could provide results with a higher occupancy value for this atom. The question of why *A.f.*_Aio shows a reaction intermediate in a latter reaction stage, even though it was soaked with the substrate for a shorter period when compared to NT-26_Aio, remains to be elucidated. Future kinetic studies of *A.f.*_Aio and Sb^{III} could help to interpret these results. In sum, two substrate analogue reaction intermediaries were successfully isolated and contributed to elucidating the reaction mechanism; moreover, structural data of the oxidised enzyme should be pursued in order to understand the steps before substrate interaction and confirm the proposed mechanism.

Kinetic inspection of the NT-26_AioA_D169A and NT-26_AioA_E453A mutants, *via* UV-Visible spectroscopy revealed a significant decrease in enzyme activity. The values of specific activity, 0.9 and 0.2 $\mu mol.min^{-1}.mg^{-1}$, for D169A and E453A mutations, respectively, show that only 46% or 8% of the efficiency of the enzyme is maintained, when compared to the WT (2.4 $\mu mol.min^{-1}.mg^{-1}$). This information, together with structural analysis, implies that both amino acid residues play an essential role in substrate orientation through a complex network of hydrogen-bonds. These are only preliminary results, and technical replicates should be performed.

*A.f.*_Aio has been reported to interact with multiple azurin proteins, as well as a physiological, non-characterised, cytochrome; MST results reveal a novel interaction of this enzyme with NT-26_cytc_{552}, with the same affinity as for *A.f.*_azu. NT-26_Aio is capable of transferring electrons to *A.f.*_azu, with slightly less affinity than for NT-26_cytc_{552}, showing that electron acceptor diversity is also plausible in this enzyme, unlike what was excepted. These variances may exist due to differences in the surfaces of the proteins, so further detailed protein surface inspections should be pursued. The crystallisation of the protein complex should allow for structural characterisation trough X-ray crystallography; in alternative, nuclear magnetic resonance spectroscopy of the protein transient complexes could reveal which residues are involved in electron transfer and small-angle X-ray scattering how the proteins interact in solution.

In sum, this work contributed for the clarification of the enzymatic reaction mechanism showing, at the molecular level, several transient protein complexes, opening new routes for the future development of new technologies in the remediation and identification of two emergent pollutants: arsenite and antimonite.

6. Bibliography

(1) Nriagu, J. O. Arsenic Poisoning Through the Ages. In *Environmental Chemistry of Arsenic*; William T., Eds.; Marcel Dekker: New York, 2002; pp 1-26.

(2) Hong, Y. S.; Song, K. H.; Chung, J. Y. Health Effects of Chronic Arsenic Exposure. *J. Prev. Med. Public Heal.* **2014**, *47* (5), 245–252.

(3) Chen, C. J.; Hsueh, Y. M.; Lai, M. S.; Shyu, M. P.; Chen, S. Y.; Wu, M. M.; Kuo, T. L.; Tai, T. Y. Increased Prevalence of Hypertension and Long-Term Arsenic Exposure. *Hypertens.* **1995**, *25* (1), 53–60.

(4) Bates, M. N.; Smith, A. H.; Hopenhayn-Rich, C. Arsenic Ingestion and Internal Cancers: A Review. *Am. J. Epidemiol.* **1992**, *135* (5), 462–476.

(5) Meliker, J. R.; Wahl, R. L.; Cameron, L. L.; Nriagu, J. O. Arsenic in Drinking Water and Cerebrovascular Disease, Diabetes Mellitus, and Kidney Disease in Michigan: A Standardized Mortality Ratio Analysis. *Environ. Health.* **2007**, *6*, 4.

(6) WHO. Chemical Aspects. *Guidelines for Drinking-Water Quality; 4th Edition*; World Health Organisation: Geneva, 2011, 178.

(7) Shi, H.; Shi, X.; Liu, K. J. Oxidative Mechanism of Arsenic Toxicity and Carcinogenesis. *Mol. Cell. Biochem.* **2004**, *255* (1–2), 67–78.

(8) Hassan, M. M. Groundwater Arsenic Catastrophe: The Global Scenario. *Arsenic in Groundwater: Poisoning and Risk Assessment*. CRC Press: 2018; pp 25-68.

(9) Chakraborty, M.; Mukherjee, A.; Ahmed, K. M. A Review of Groundwater Arsenic in the Bengal Basin, Bangladesh and India: From Source to Sink. *Current Pollution Reports*. Springer: 2015; pp 220–247.

(10) Schwarzenbach, R. P.; Egli, T.; Hofstetter, T. B.; von Gunten, U.; Wehrli, B. Global Water Pollution and Human Health. *Annu. Rev. Environ. Resour.* **2010**, *35* (1), 109–136.

(11) Slotnick, M. J.; Meliker, J.; Nriagu, J. Natural Sources of Arsenic in Southeastern Michigan Groundwater. *J. Phys. IV.* **2003**, *107*, 1247 - 1250.

(12) Das, D.; Samanta, G.; Mandal, B. K.; Roy Chowdhury, T.; Chanda, C. R.; Chowdhury, P. P.; Basu, G. K.; Chakraborti, D. Arsenic in Groundwater in Six Districts of West Bengal,

India. *Environ. Geochem. Health.* **1996**, *18* (1), 5–15.

(13) Nickson, R. T.; Mcarthur, J. M.; Ravenscroft, P.; Burgess, W. G.; Ahmed, K. M. Mechanism of Arsenic Release to Groundwater, Bangladesh and West Bengal. *Applied Geochem.* **2000**, *15* (4), 403-413.

(14) Welch, A. H.; Lico, M. S.; Hughes, J. L. Arsenic in Ground Water of the Western United States. *Groundwater*. **1988**, *26* (3), 333-347.

(15) Singh, R.; Singh, S.; Parihar, P.; Singh, V. P.; Prasad, S. M. Arsenic Contamination, Consequences and Remediation Techniques: A Review. *Ecotoxicol. Environ. Saf.* **2015**, *112*, 247–270.

(16) Mandal, B. K.; Suzuki, K. T. Arsenic Round the World: a Review. *Talanta*. **2002**, *58*, 201–235.

(17) Greenwood, N. N.; Earnshaw A. Arsenic, Antimony and Bismuth. *Chemistry of the Elements*; Elsevier: 1997; pp 547–599.

(18) Oremland, R. S.; Stolz, J. F. Arsenic, Microbes and Contaminated Aquifers. *Trends in Microbiology*. Elsevier: 2005; pp 45–49.

(19) Dhuldhaj, U. P.; Sharma, N. K.; Singh, S. Microbial Removal of Arsenic: An Overview. *Bioremediation Pollut.* IK International Publishing House: 2012; pp 113-125.

(20) Ropp, R. C. Group 15 (N, P, As, Sb and Bi) Alkaline Earth Compounds. In *Encyclopedia of the Alkaline Earth Compounds*; Elsevier: 2013; pp 199–350.

(21) Hughes, M. F. Arsenic Toxicity and Potential Mechanisms of Action. *Toxicol. Lett.* **2002**, *133* (1), 1–16.

(22) Sharma, V. K.; Sohn, M. Aquatic Arsenic: Toxicity, Speciation, Transformations, and Remediation. *Environment International*. Elsevier: 2009; pp 743–759.

(23) Van Hue, N. Bioremediation of Arsenic Toxicity. In *Arsenic Toxicity*; CRC Press: 2015; pp 155–166.

(24) O'Day, P. A. Chemistry and Mineralogy of Arsenic. *Elements* **2006**, *2* (2), 77–83.

(25) Zouboulis, A. I.; Katsoyiannis, I. A. Recent Advances in the Bioremediation of Arsenic-Contaminated Groundwater. *Environ. Int.* **2005**, *31* (2), 213–219.

(26) Komorowicz, I.; Barałkiewicz, D. Arsenic and Its Speciation in Water Samples by High Performance Liquid Chromatography Inductively Coupled Plasma Mass Spectrometry-

Last Decade Review. *Talanta* **2011**, *84* (2), 247–261.

(27) Anawar, H. M. Arsenic Speciation in Environmental Samples by Hydride Generation and Electrothermal Atomic Absorption Spectrometry. *Talanta* **2012**, *88*, 30–42.

(28) Sanger, C. R. The Quantitative Determination of Arsenic by the Gutzeit Method. *J. Am. Chem. Soc.* **1908**, *30* (6), 1041–1042.

(29) Babu Christus, A. A.; Panneerselvam, P.; Ravikumar, A. Novel, Sensitive and Selective Colorimetric Detection of Arsenate in Aqueous Solution by a Fenton-like Reaction of Fe_3O_4 Nanoparticles. *Anal. Methods* **2018**, *10* (36), 4378–4386.

(30) Sirawatcharin, S.; Saithongdee, A.; Chaicham, A.; Tomapatanaget, B.; Imyim, A.; Praphairaksit, N. Naked-Eye and Colorimetric Detection of Arsenic(III) Using Difluoroboron-Curcumin in Aqueous and Resin Bead Support Systems. *Anal. Sci.* **2014**, *30* (12), 1129–1134.

(31) Lièvremont, D.; N'negue, M. A.; Behra, P.; Lett, M. C. Biological Oxidation of Arsenite: Batch Reactor Experiments in Presence of Kutnahorite and Chabazite. *Chemosphere* **2003**, *51* (5), 419–428.

(32) Terry, L. R.; Kulp, T. R.; Wiatrowski, H.; Miller, L. G.; Oremland, R. S. Microbiological Oxidation of Antimony(III) with Oxygen or Nitrate by Bacteria Isolated from Contaminated Mine Sediments. *Appl. Environ. Microbiol.* **2015**, *81* (24), 8478–8488.

(33) Lu, X.; Zhang, Y.; Liu, C.; Wu, M.; Wang, H. Characterization of the Antimonite- and Arsenite-Oxidizing Bacterium *Bosea* sp. AS-1 and Its Potential Application in Arsenic Removal. *J. Hazard. Mater.* **2018**, *359*, 527–534.

(34) Wang, Q.; Warelow, T. P.; Kang, Y. S.; Romano, C.; Osborne, T. H.; Lehr, C. R.; Bothner, B.; McDermott, T. R.; Santini, J. M.; Wang, G. Arsenite Oxidase Also Functions as an Antimonite Oxidase. *Appl. Environ. Microbiol.* **2015**, *81* (6), 1959–1965.

(35) Krachler, M.; Emons, H.; Zheng, J. Speciation of Antimony for the 21st Century: Promises and Pitfalls. *TrAC - Trends Anal. Chem.* **2001**, *20* (2), 79–90.

(36) Fu, Z.; Wu, F.; Mo, C.; Liu, B.; Zhu, J.; Deng, Q.; Liao, H.; Zhang, Y. Bioaccumulation of Antimony, Arsenic, and Mercury in the Vicinities of a Large Antimony Mine, China. *Microchem. J.* **2011**, *97* (1), 12–19.

(37) He, M.; Wang, X.; Wu, F.; Fu, Z. Antimony Pollution in China. *Sci. Total Environ.* **2012**, *421–422*, 41–50.

(38) United States Environmental Protection Agency, 1979. Water Related Fate of the 129 Priority Pollutants, vol. 1. USEPA,Washington, DC, USA, EP-440/4-79-029A.

(39) Council of the European Union, 1998. Council Directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption. Official Journal L 330, 05/12/1998, pp. 32–54.

(40) Filella, M.; Belzile, N.; Chen, Y. W. Antimony in the Environment: A Review Focused on Natural Waters II. Relevant Solution Chemistry. *Earth-Science Rev.* **2002**, *59* (1–4), 265–285.

(41) Li, J.; Zheng, B. H.; He, Y.; Zhou, Y.; Chen, X.; Ruan, S.; Yang, Y.; Dai, C.; Tang, L. Antimony Contamination, Consequences and Removal Techniques: A Review. *Ecotoxicology and Environmental Safety*. Academic Press: 2018; pp 125–134.

(42) Kang, M.; Kawasaki, M.; Tamada, S.; Kamei, T.; Magara, Y. Effect of PH on the Removal of Arsenic and Antimony Using Reverse Osmosis Membranes. *Desalination* **2000**, *131* (1–3), 293–298.

(43) Abin, C. A.; Hollibaugh, J. T. Dissimilatory Antimonate Reduction and Production of Antimony Trioxide Microcrystals by a Novel Microorganism. *Environ. Sci. Technol.* **2014**, *48* (1), 681–688.

(44) Hudson-Edwards, K. A.; Santini, J. M. Arsenic-Microbe-Mineral Interactions in Mining-Affected Environments. *Minerals*. MDPI: 2013, pp 337–351.

(45) Stolz, J. F.; Basu, P.; Santini, J. M.; Oremland, R. S. Arsenic and Selenium in Microbial Metabolism. *Annu. Rev. Microbiol.* **2006**, *60*, 107–130.

(46) Cai, J.; Salmon, K.; DuBow, M. S. A Chromosomal Ars Operon Homologue of *Pseudomonas aeruginosa* Confers Increased Resistance to Arsenic and Antimony in *Escherichia coli*. *Microbiology.* **1998**, *144 (10)*, 2705–2713.

(47) Meng, Y.-L.; Liu, Z.; Rosen, B. P. As(III) and Sb(III) Uptake by GlpF and Efflux by ArsB *Escherichia coli*. *J. Biol. Chem.* **2004**, *279* (18), 18334–18341.

(48) Oremland, R. S.; Stolz, J. F. The Ecology of Arsenic. *Science*. **2003**, *300* (5621), 939-944.

(49) Hoeft, S. E.; Kulp, T. R.; Han, S.; Lanoil, B.; Oremland, R. S. Coupled Arsenotrophy in a Hot Spring Photosynthetic Biofilm at Mono Lake, California. *Appl. Environ. Microbiol.* **2010**, *76* (14), 4633–4639.

(50) Zargar, K.; Conrad, A.; Bernick, D. L.; Lowe, T. M.; Stolc, V.; Hoeft, S.; Oremland, R.

S.; Stolz, J.; Saltikov, C. W. ArxA, a New Clade of Arsenite Oxidase within the DMSO Reductase Family of Molybdenum Oxidoreductases. *Environ. Microbiol.* **2012**, *14* (7), 1635–1645.

(51) Zargar, K.; Hoeft, S.; Oremland, R.; Saltikov, C. W. Identification of a Novel Arsenite Oxidase Gene, ArxA, in the Haloalkaliphilic, Arsenite-Oxidizing Bacterium *Alkalilimnicola ehrlichii* sp. MLHE-1. *J. Bacteriol.* **2010**, *192* (14), 3755–3762.

(52) Lett, M. C.; Muller, D.; Lièvremont, D.; Silver, S.; Santini, J. Unified Nomenclature for Genes Involved in Prokaryotic Aerobic Arsenite Oxidation. *J. Bacteriol.* **2012**, *194* (2), 207–208.

(53) Mukhopadhyay, R.; Rosen, B. P.; Phung, L. T.; Silver, S. Microbial Arsenic: From Geocycles to Genes and Enzymes. *FEMS Microbiol. Rev.* **2002**, *26* (3), 311–325.

(54) Li, J.; Yang, B.; Shi, M.; Yuan, K.; Guo, W.; Wang, Q.; Wang, G. Abiotic and Biotic Factors Responsible for Antimonite Oxidation in *Agrobacterium tumefaciens* GW4. *Sci. Rep.* **2017**, *7*.

(55) Romão, M. J. Molybdenum and Tungsten Enzymes: A Crystallographic and Mechanistic Overview. *Dalton Trans.* **2009**, *21*, 4053–4068.

(56) Hille, R. Molybdenum and Tungsten in Biology. *Trends Biochem. Sci.* **2002**, *27* (7), 360–367.

(57) Grimaldi, S.; Schoepp-Cothenet, B.; Ceccaldi, P.; Guigliarelli, B.; Magalon, A. The Prokaryotic Mo/W-BisPGD Enzymes Family: A Catalytic Workhorse in Bioenergetic. *Biochim. Biophys. Acta*. **2013**, *1827* (8-9), 1048–1085.

(58) Basu, P.; Burgmayer, S. J. N. Pterin Chemistry and Its Relationship to the Molybdenum Cofactor. *Coord Chem Rev*. **2011**, *255* (9-10), 1016–1038.

(59) Pushie, M. J.; George, G. N. Spectroscopic Studies of Molybdenum and Tungsten Enzymes. *Coord Chem Rev*. **2011**, *255* (9-10), 1055–1084.

(60) Lebrun, E.; Brugna, M.; Baymann, F.; Muller, D.; Lièvremont, D.; Lett, M.-C.; Nitschke, W. Arsenite Oxidase, an Ancient Bioenergetic Enzyme. *Mol. Biol. Evol.* **2003**, *20* (5), 686–693.

(61) Vanden Hoven, R. N.; Santini, J. M. Arsenite Oxidation by the Heterotroph *Hydrogenophaga* sp. str. NT-14: The Arsenite Oxidase and Its Physiological Electron Acceptor. *Biochim. Biophys. Acta.* **2004**, *1656* (2–3), 148–155.

(62) Teoh, W. K.; Salleh, F. M.; Shahir, S. Characterization of *Thiomonas delicata* Arsenite Oxidase Expressed in *Escherichia coli*. *3 Biotech.* **2017**, *7* (2), 97.

(63) Osborne, T. H.; Heath, M. D.; Martin, A. C. R.; Pankowski, J. A.; Hudson-Edwards, K. A.; Santini, J. M. Cold-Adapted Arsenite Oxidase from a Psychrotolerant *Polaromonas* Species. *Metallomics.* **2013**, *5* (4), 318–324.

(64) Prasad, K. S.; Subramanian, V.; Paul, J. Purification and Characterization of Arsenite Oxidase from *Arthrobacter* sp. *BioMetals.* **2009**, *22* (5), 711–721.

(65) Lieutaud, A.; Van Lis, R.; Duval, S.; Capowiez, L.; Muller, D.; Lebrun, R.; Lignon, S.; Fardeau, M. L.; Lett, M. C.; Nitschke, W.; et al. Arsenite Oxidase from *Ralstonia* sp. 22: Characterization of the Enzyme and Its Interaction with Soluble Cytochromes. *J. Biol. Chem.* **2010**, *285* (27), 20433–20441.

(66) Warelow, T. P.; Oke, M.; Schoepp-Cothenet, B.; Dahl, J. U.; Bruselat, N.; Sivalingam, G. N.; Leimkühler, S.; Thalassinos, K.; Kappler, U.; Naismith, J. H.; et al. The Respiratory Arsenite Oxidase: Structure and the Role of Residues Surrounding the Rieske Cluster. *PLoS One* **2013**, *8* (8), e72535.

(67) Ellis, P. J.; Conrads, T.; Hille, R.; Kuhn, P. Crystal Structure of the 100 kDa Arsenite Oxidase from *Alcaligenes faecalis* in Two Crystal Forms at 1.64 Å and 2.03 Å. *Structure.* **2001**, *9* (2), 125-132.

(68) Santini, J. M.; Sly, L. I.; Schnagl, R. D.; Macy, J. M. A New Chemolithoautotrophic Arsenite-Oxidizing Bacterium Isolated. *Appl Env. Microbiol.* **2000**, *66* (1), 92–97.

(69) Santini, J. M.; Vanden Hoven, R. N. Molybdenum-Containing Arsenite Oxidase of the Chemolithoautotrophic Arsenite Oxidizer NT-26. *J. Bacteriol.* **2004**, *186* (6), 1614–1619.

(70) Conrads, T.; Hemann, C.; George, G. N.; Pickering, I. J.; Prince, R. C.; Hille, R. The Active Site of Arsenite Oxidase from *Alcaligenes faecalis*. *J. Am. Chem. Soc.* **2002**, *124* (38), 11276–11277.

(71) Hoke, K. R.; Cobb, N.; Armstrong, F. A.; Hille, R. Electrochemical Studies of Arsenite Oxidase: An Unusual Example of a Highly Cooperative Two-Electron Molybdenum Center. *Biochemistry.* **2004**, *43* (6), 1667–1674.

(72) Warelow, T. P.; Pushie, M. J.; Cotelesage, J. J. H.; Santini, J. M.; George, G. N. The Active Site Structure and Catalytic Mechanism of Arsenite Oxidase. *Sci. Rep.* **2017**, *7* (1), 1757.

(73) Zu, Y.; Couture, M. M. J.; Kolling, D. R. J.; Crofts, A. R.; Eltis, L. D.; Fee, J. A.; Hirst, J. Reduction Potentials of Rieske Clusters: Importance of the Coupling between Oxidation State and Histidine Protonation State. *Biochemistry.* **2003**, *42* (42), 12400–12408.

(74) Van Lis, R.; Nitschke, W.; Warelow, T. P.; Capowiez, L.; Santini, J. M.; Schoepp-Cothenet, B. Heterologously Expressed Arsenite Oxidase: A System to Study Biogenesis and Structure/Function Relationships of the Enzyme Family. *Biochim. Biophys. Acta - Bioenerg.* **2012**, *1817* (9), 1701-1708.

(75) Watson, C.; Niks, D.; Hille, R.; Vieira, M.; Schoepp-Cothenet, B.; Marques, A. T.; Romão, M. J.; Santos-Silva, T.; Santini, J. M. Electron Transfer through Arsenite Oxidase: Insights into Rieske Interaction with Cytochrome *c*. *Biochim. Biophys. Acta - Bioenerg.* **2017**, *1858* (10), 865–872.

(76) McNellis, L.; Anderson, G. L. Redox-State Dependent Chemical Inactivation of Arsenite Oxidase. *J. Inorg. Biochem.* **1998**, *69* (4), 253–257.

(77) Bernhardt, P. V; Santini, J. M. Protein Film Voltammetry of Arsenite Oxidase from the Chemolithoautotrophic Arsenite-Oxidizing Bacterium NT-26. *Biochemistry.* **2006**, *45* (9), 2804–2809.

(78) Philips, S. E.; Taylor, M. L. Oxidation of Arsenite to Arsenate by *Alcaligenes faecalis*. *Appl. Environ. Microbiol.* **1976**, *32* (3).

(79) Anderson, G. L.; Williams, J.; Hille, R. The Purification and Characterization of Arsenite Oxidase from *Alcaligenes faecalis*, a Molybdenum-Containing Hydroxylase. *J. Biol. Chem.* **1992**, *267* (33), 23674–23682.

(80) J.M., S.; U., K.; S.A., W.; M.J., H.; R.N., vanden H.; Santini, J. M.; Kappler, U.; Ward, S. A.; Honeychurch, M. J.; vanden Hoven, R. N.; et al. The NT-26 Cytochrome c_{552} and Its Role in Arsenite Oxidation. *Biochim. Biophys. Acta - Bioenerg.* **2007**, *1767* (2), 189-196.

(81) Vanden Hoven, R. N.; Santini, J. M. Arsenite Oxidation by the Heterotroph *Hydrogenophaga* sp. str. NT-14: The Arsenite Oxidase and Its Physiological Electron Acceptor. *Biochim. Biophys. Acta - Bioenerg.* **2004**, *1656* (2–3), 148–155.

(82) Branco, R.; Francisco, R.; Chung, A. P.; Morais, P. V. Identification of an Aox System That Requires Cytochrome *c* in the Highly Arsenic-Resistant Bacterium *Ochrobactrum tritici* SCII24. *Appl. Environ. Microbiol.* **2009**, *75* (15), 5141–5147.

(83) Osborne, T. H.; Heath, M. D.; Martin, A. C. R.; Pankowski, J. A.; Hudson-Edwards, K.

A.; Santini, J. M. Cold-Adapted Arsenite Oxidase from a Psychrotolerant *Polaromonas* Species. *Metallomics.* **2013**, *5* (4), 318-324.

(84) Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M. R.; Appel, R. D.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*; Humana Press: 2005; pp 571–607.

(85) Kabsch, W. XDS. *Acta Crystallogr. Sect. D.* **2010**, *66* (2), 125–132.

(86) Evans, P. Scaling and Assessment of Data Quality. In *Acta Crystallographica Section D: Biological Crystallography*; 2006; Vol. 62, pp 72–82.

(87) Dodson, E. J. A Collaborative Computational Project, Number 4: Providing Programs for Protein Crystallography. *Methods Enzymol;* 1997; Vol. 227, pp 620-33.

(88) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40* (4), 658–674.

(89) Read, R. J.; Schierbeek, A. J. A Phased Translation Function. *J. Appl. Crystallogr.* **1988**, *21* (5), 490–495.

(90) Emsley, P.; Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **2004**, *60* (Pt 12 Pt 1), 2126–2132.

(91) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC5 for the Refinement of Macromolecular Crystal Structures. *Acta Crystallogr. D. Biol. Crystallogr.* **2011**, *67* (Pt 4), 355–367.

(92) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; et al. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (2), 213–221.

(93) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (1), 12–21.

(94) Bai, N.; Roder, H.; Dickson, A.; Karanicolas, J. Isothermal Analysis of ThermoFluor Data Can Readily Provide Quantitative Binding Affinities. *Sci Rep*. *9* (1), 2650.

(95) Huynh, K.; Partch, C. L. Analysis of Protein Stability and Ligand Interactions by Thermal

Shift Assay. *Curr. Protoc. protein Sci.* **2015**, *79* (28), 1-14.

(96) Santini R.N., J. M. A.-V. H. Molybdenum-Containing Arsenite Oxidase of the Chemolithoautotrophic Arsenite Oxidizer NT-26. *J. Bacteriol.* **2004**, *186* (6), 1614–1619.

(97) Santini, J. M.; Sly, L. I.; Wen, A.; Comrie, D.; De Wulf-Durand, P.; Macy, J. M. New Arsenite-Oxidizing Bacteria Isolated from Australian Gold Mining Environments - Phylogenetic Relationships. *Geomicrobiol. J.* **2002**, *19* (1), 67–76.

(98) Jerabek-Willemsen, M.; André, T.; Wanner, R.; Roth, H. M.; Duhr, S.; Baaske, P.; Breitsprecher, D. MicroScale Thermophoresis: Interaction Analysis and Beyond. *J. Mol. Struct.* **2014**, *1077*, 101-113.

(99) Crowley, P. B.; Ubbink, M. Close Encounters of the Transient Kind: Protein Interactions in the Photosynthetic Redox Chain Investigated by NMR Spectroscopy. *Acc. Chem. Res.* **2003**, *36* (10), 723-730.

(100) Day, A. G.; Withers, S. G. The Purification and Characterization of a β-Glucosidase from *Alcaligenes faecalis*. *Biochem. Cell Biol.* **2009**, *64* (9), 914–922.

(101) Karplus, P. A.; Diederichs, K. Linking Crystallographic Model and Data Quality. *Science.* **2012**, *336* (6084), 1030–1033.

(102) Karplus, P. A.; Diederichs, K. Assessing and Maximizing Data Quality in Macromolecular Crystallography. *Current Opinion in Structural Biology*. Elsevier: 2015, pp 60–68.

(103) Dauter, Z. Collection of X-Ray Diffraction Data from Macromolecular Crystals. In *Methods Mol Biol.*; Humana Press Inc: 2017; Vol. 1607, pp 165–184.

(104) Vieira, M. Arsenito Oxidase Bacteriana: Caracterização a Nível Molecular. M.S. Thesis, NOVA Univestity Lisbon, Portugal, 2016.

(105) DeLano, W. L.. Pymol: An open-source molecular graphics tool. CCP4 Newsletter On Protein Crystallography, **2002**. 40, 82-9.

7. Appendix

7.1. Appendix 1

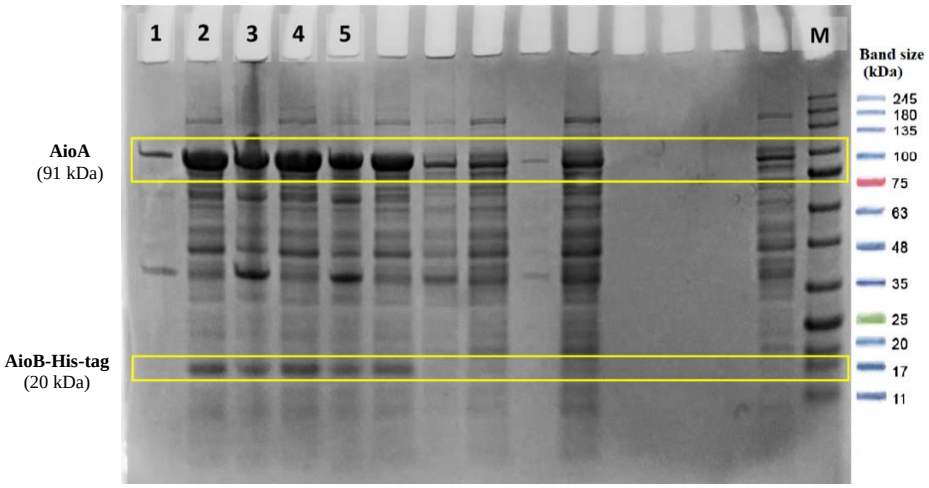


Figure 7.2. 10% SDS-PAGE containing samples correspondent to the expression tests of arsenite oxidase from NT-26 in different expression hosts *E.coli* strains. 1) Insoluble fraction from BL21 (DE3) cells; 2) soluble fraction from BL21 (DE3) cells; 3) insoluble fraction from Rosetta 2(DE3)pLysS cells; 4) soluble fraction from Rosetta 2(DE3)pLysS cells. 5) insoluble fraction from Tuner (DE3) cells; 6) soluble fraction from Tuner (DE3) cells; M) Protein Marker II (NZYTech).

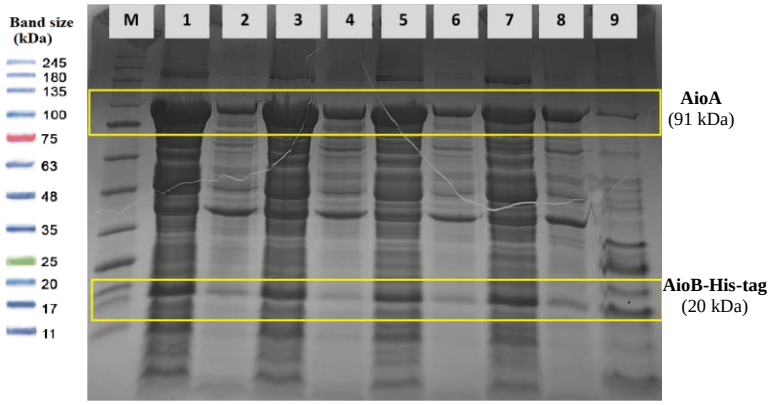


Figure 7.1. 10% SDS-PAGE containing samples correspondent to the expression tests of arsenite oxidase mutant NT-26_AioA_D169A in *E. coli* Tuner (DE3) cells using different IPTG induction concentrations: M) Protein Marker II (NZYTech); 1) soluble fraction, 0.25 mM; 2) insoluble fraction, 0.25 mM; 3) soluble fraction, 0.50 mM; 4) insoluble fraction, 0.50 mM 5) soluble fraction, 0.75 mM; 6) insoluble fraction, 0.75 mM; 7) soluble fraction, 1 mM; 8) insoluble fraction, 1 mM. 9) Mutant in DH5α expression host, for comparison.

7.2. Appendix 2

Table 7.1. Trace elements solution composition.

Reagent	Concentration ($g.L^{-1}$)
$MgCl_2.6H_2O$	13.00
$FeCl3.6H_2O$	7.70
$MnCl_2.4H_2O$	1.10
$CuSO_4.5H_2O$	0.15
$CoCl_2.6H_2O$	0.13
H_3BO_3	0.12
$NaMoO_4.2H_2O$	1.00

7.3. Appendix 3

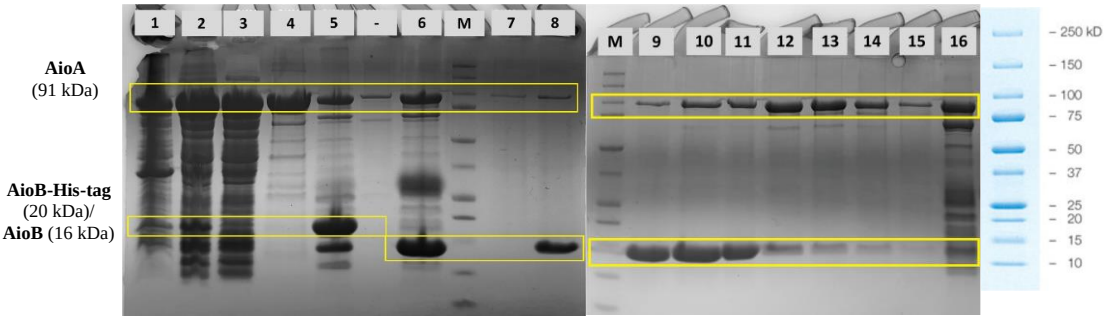


Figure 7.3. 10% SDS-PAGE of the expression and IMAC purification of recombinant NT-26_AioA_D169A protein in *E.coli* Tuner (DE3) cells. M) Precision Plus Protein™ (Bio-Rad). Samples of 1) lysate insoluble fraction; 2) lysate insoluble fraction; 3) proteins eluted at 20 mM imidazole, host cell proteins and excess AioA; 4) proteins eluted at 43 mM imidazole, host cell proteins and excess AioA; 5) 500 mM imidazole elution volume, host cell proteins and AioAB-His-tag; 6) sample 5 after TEV cleavege, includes host cell proteins and AioAB. 7-15) AioAB eluted at 20 mM imidazole after TEV cleavage; 16) proteins eluted at 500 mM imidazole, host cell proteins, TEV enzyme and uncomplexed AioA.

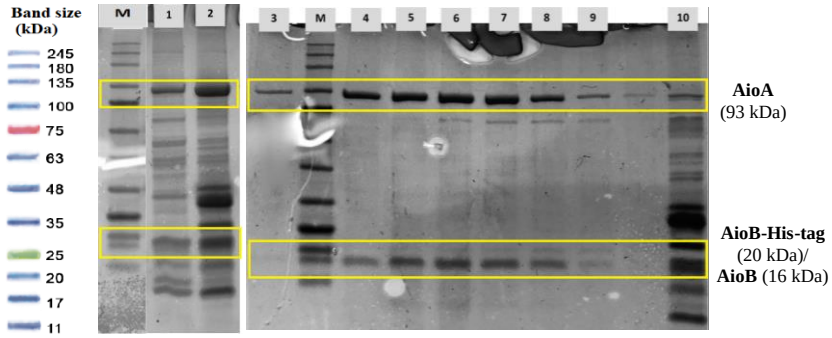


Figure 7.4. 10% SDS-PAGE of the expression and purification of recombinant *A.f._*Aio in *E. coli* DH5α cells. M) Protein Marker II (NZYTech); 1) proteins eluted at 20 mM imidazole, host cell proteins and excess AioA; 2) 500 mM imidazole elution volume, host cell proteins and AioAB-His-tag; 3-8) AioAB eluted at 20 mM imidazole after TEV cleavage; 10) proteins eluted at 500 mM imidazole, host cell proteins, TEV enzyme and uncomplexed AioA.

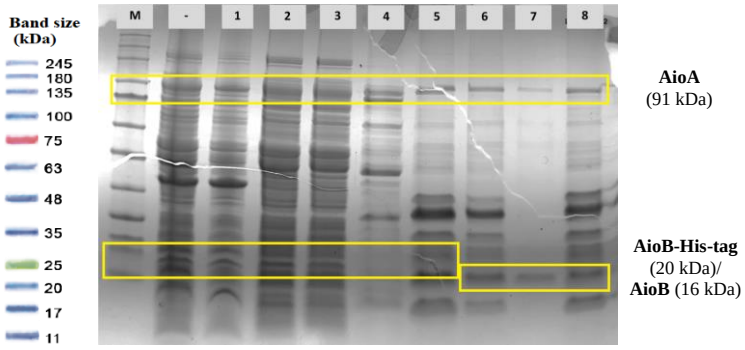


Figure 7.5. 10% SDS-PAGE of the expression and IMAC purification of recombinant NT-26_AioA_E453A protein in *E.coli* Tuner (DE3) cells. M) Protein Marker II (NZYTech). Samples of 1) lysate insoluble fraction; 2) lysate insoluble fraction; 3) proteins eluted at 20 mM imidazole, host cell proteins and excess AioA; 4) proteins eluted at 43 mM imidazole, host cell proteins and excess AioA; 5) 500 mM imidazole elution volume, host cell proteins and AioAB-His-tag; 6) sample 5 after TEV cleavege, includes host cell proteins and AioAB. 7) AioAB eluted at 20 mM imidazole after TEV cleavage; 8) proteins eluted at 500 mM imidazole, host cell proteins, TEV enzyme and uncomplexed AioA.

7.4. Appendix 4

Table 7.2. TSA 1st melting temperature results of NT-26_AioA_D169A screened with Jena Biosciences buffer solutions.

	Condition	1st T_m (°C)
control	Water	27.4
No salt buffers	citric acid pH 4.0	34.6
	sodium acetate pH 4.5	35.4
	sodium citrate pH 4.7	33.8
	sodium acetate pH 5.0	48.7
	potassium phosphate pH 5.0	44.7
	sodium phosphate pH 5.5	39.0
	sodium citrate pH 5.5	45.0
	MES pH 5.8	53.1
	potassium phosphate pH 6.0	NA
	MES pH 6.2	42.3
	sodium phosphate pH 6.5	NA
	sodium cacodylate pH 6.5	41.6
	MES pH 6.5	43
	PIPES pH 6.7	35
	MOPS pH 7.0	NA
	HEPES pH 7.0	49
	ammonium acetate pH 7.3	39
	sodium phosphate pH 7.5	NA
	Tris pH 7.5	38.6
	imidazole pH 8.0	NA
	HEPES pH 8.0	30
	Tris pH 8.0	NA
	bicine pH 8.5	35.9
	Tris pH 8.5	NA
	bicine pH 9.0	NA
	CAPS pH 9.0	35

	glycine-HCl pH 9.5	36.6
	glycine-HCl pH 10	33
	CAPS pH 10.5	NA
	CAPS pH 11	29.9
1 M salt buffers	citric acid pH 4.0	31.1
	sodium acetate pH 4.5	30.1
	sodium citrate pH 4.7	29.9
	sodium acetate pH 5.0	30.2
	potassium phosphate pH 5.0	NA
	sodium phosphate pH 5.5	NA
	sodium citrate pH 5.5	29.9
	MES pH 5.8	32.3
	potassium phosphate pH 6.0	31.5
	MES pH 6.2	31.1
	sodium phosphate pH 6.5	NA
	sodium cacodylate pH 6.5	NA
	MES pH 6.5	32.2
	PIPES pH 6.7	32.0
	MOPS pH 7.0	31.1
	HEPES pH 7.0	31.0
	ammonium acetate pH 7.3	30.3
	sodium phosphate pH 7.5	30.4
	Tris pH 7.5	31.6
	imidazole pH 8.0	NA
	HEPES pH 8.0	33.5
	Tris pH 8.0	32.3
	bicine pH 8.5	31.2
	Tris pH 8.5	29.2
	ammonium acetate pH 7.3	30.3
	sodium phosphate pH 7.5	30.4
	Tris pH 7.5	31.6

7.5. Appendix 5

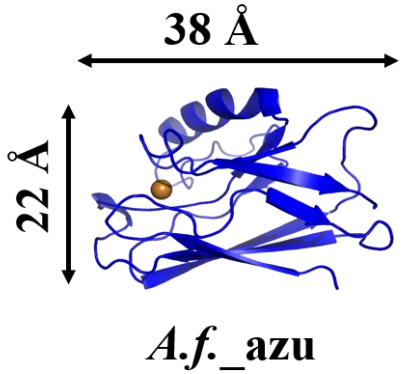


Figure 7.5. *A.f.*_azu structure measurements. PDB: 2H3X. Distances were calculated using PyMOL software.