Tomás Onofre de Almeida Frazão
Licenciado em Bioquímica

Insight into the biophysical properties of *As*P2Ox, a candidate enzyme for health diagnosis biosensors

Dissertação para obtenção do Grau de Mestre em
Bioquímica para a Saúde

Orientador: Patrícia Borges, Post-Doc Researcher, ITQB
Co-orientador: Lígia Martins, Assistant Professor, ITQB

Novembro 2021

Tomás Onofre de Almeida Frazão
Licenciado em Bioquímica

Insight into the biophysical properties of *As*P2Ox, a candidate enzyme for health diagnosis biosensors

Dissertação para obtenção do Grau de Mestre em
Bioquímica para a Saúde

Orientador: Patrícia Borges, Post-Doc Researcher, ITQB
Co-orientador: Lígia Martins, Assistant Professor, ITQB

Júri:

Presidente: Prof. Doutor Pedro Matias
Arguente (s): Doutor Filipe Folgosa
Vogal (ais): Prof. Doutora Ana Varela Coelho

Instituto de Tecnologia Química e Biológica António Xavier

Novembro 2021

Insight into the biophysical properties of *As*P2Ox, a candidate enzyme for health diagnosis biosensors

Acknowledgements

Terminando o meu mestrado em Bioquímica para a Saúde, e em específico a sua tese final, quero primeiro agradecer de modo geral a todos aqueles que direta ou indiretamente participaram ou demonstraram interesse neste processo de aprendizagem.

Em primeiro lugar, agradeço à minha orientadora, Patrícia Borges, por todo o conhecimento, disponibilidade e tempo dedicado ao longo deste ano.

Em segundo lugar, e com particular atenção, um enorme obrigado à professora Lígia Martins pela oportunidade de participar neste projeto e fazer parte da equipa do laboratório Microbial and Enzyme Technology, toda a sua preocupação, dedicação e exigência demonstrada. Agradeço também a todos os outros investigadores e alunos do laboratório, em especial ao André Taborda pela enorme disponibilidade, orientação e preocupação. Um obrigado também à unidade de Macromolecular Crystallography, em especial ao professor Carlos Frazão, pelos recursos e espaço fornecido para a realização de grande parte deste trabalho.

Um especial obrigado à Professora Teresa Catarino pela preocupação demonstrada ao longo destes dois anos de mestrado, assim como na licenciatura, por toda a sua dedicação ao ensino e alunos.

Pelo interesse e apoio apresentado em relação ao meu projeto, agradeço particularmente a dois colegas e amigos, Teresa Pissarro e João Costa.

Finalmente, mas não menos importante, agradeço a toda a minha família, em especial pais e irmã pelo enorme apoio e interesse demonstrado, não só neste projeto, mas também ao longo de todo o meu percurso escolar e futuro. Um enorme obrigado à Beatriz, por todos os esclarecimentos existenciais e linguísticos, apoio incondicional e incansável alegria transmitida. Los Cotons, obrigadinho.

Abstract

Pyranose-2-oxidases (P2Oxs) are flavoenzymes capable of oxidizing a wide range of sugars, in particular D-glucose, with concomitant reduction of molecular oxygen to hydrogen peroxide. The vast majority of P2Oxs were identified in fungi, but recently, bacterial P2Oxs were characterized showing distinctive and interesting properties to be explored in biotechnology. This work aims to unveil structural details of the bacterial P2Ox from *Pseudoarthrobacter siccitolerans* (*As*P2Ox). Protein crystals were obtained in space group $C222_1$ space group and diffracted X-rays to a resolution of 2.01 Å, and a crystal structure of the *As*P2Ox-Glucose complex at 2.35 Å resolution was also obtained. The typical P2Oxs substrate and FAD-binding domains were identified, and a flexible loop (substrate loop) found in fungal P2Oxs with a role of gating and accommodating the substrate in the active site is also present in *As*P2Ox. This substrate loop shows a very poor electronic density in the glucose complex crystal structure, suggesting alterations of conformation, after interactions with the substrate, required for catalytic functions. *As*P2Ox is a monomer, lacking the oligomerization domains present in fungal homologous structures, and showing the active site in an exposed cavity. The FAD cofactor is not covalently bonded, in contrast to the fungal enzymes, possibly due to interactions at the FAD binding-motif. Mutants were designed and constructed using site-directed and site-saturation mutagenesis that could force the conserved and putative histidine-ligand (H127) to rotate to a conformation more likely to covalently link FAD to the enzyme. However, all attempts resulted in variant enzymes that are either apo-enzymes, or enzymes binding FAD non-covalently. Interestingly, variant N120C, showed almost 2-fold higher FAD incorporation than wild-type. Overall, this study contributed to increase our understanding of *As*P2Ox structural and functional features, and consequently to boost the application of protein engineering to improve its features e.g., for application in third-generation biosensors for D-Glucose monitoring.

Keywords: Pyranose 2-oxidases, flavoenzymes, biotechnology, structural details, biosensors

Resumo

As piranose-2-oxidases, P2Oxs, são flavo-enzimas capazes de oxidar uma vasta variedade de monossacarídeos, como D-Glucose, reduzindo consequentemente oxigénio molecular a peróxido de hidrogénio. Estas enzimas estão identificadas maioritariamente em fungos, mas, recentemente, P2Oxs bacterianas foram caracterizadas, com propriedades distintas e interessantes para serem exploradas em biotecnologia. Este trabalho tem como principal objetivo revelar determinantes estruturais da P2Ox bacteriana de *Pseudoarthrobacter siccitolerans* (*As*P2Ox). Foram obtidos cristais desta proteína no grupo espacial $C222_1$, os quais difrataram raios-X até uma resolução de 2.01 Å e uma estrutura cristalográfica do complexo *As*P2Ox-Glucose foi também obtida, com uma resolução de 2.35 Å. Os domínios típicos de ligação ao substrato e ao cofator FAD, das P2Oxs, foram identificados, assim como um *substrate loop*, identificado em P2Oxs fúngicas como sendo responsável por regular o acesso ao centro ativo. Este *loop* apresenta uma densidade eletrónica muito fraca na estrutura cristalográfica do complexo com glucose, sugerindo uma alteração conformacional após interação com o substrato, necessária para as funções catalíticas. *As*P2Ox é um monómero, não tendo os domínios de oligomerização presentes nas estruturas homólogas fúngicas e, contém o centro ativo numa cavidade consideravelmente exposta ao solvente. O seu FAD não está ligado covalentemente, ao contrário das enzimas fúngicas. Foram construídas variantes através de mutagénese dirigida ou de saturação para tentar "forçar" a H127, possível ligando do FAD, a adotar uma conformação favorável à ligação covalente. Todas as tentativas resultaram em variantes sem FAD incorporado ou ligado de forma não covalente. Ainda assim, uma variante, N120C, apresentou uma incorporação de FAD duas vezes maior que a enzima nativa. Em suma, este estudo contribuiu para o aumento do conhecimento sobre as características estruturais e funcionais da *As*P2Ox e irá, consequentemente, facilitar a aplicação de engenharia de proteínas para melhorar as suas características, por exemplo para aplicação em biossensores de terceira geração para monitorização de D-Glucose.

Termos chave: Piranose-2-oxidase, flavo-enzima, biotecnologia, estrutura, determinantes estruturais

List of Contents

Figures Index

Tables Index

Abreviations List

1,5-AG	1,5-anhydro-D-glucitol
AAP	4-aminophenazone
ABTS	2,2'-azino-bis (3- ethylbenzothiazoline-6-sulfonic acid)
a.d.p	Atomic displacement parameters
***As*P2Ox**	*Pseudorthrobacter siccitolerans* pyranose 2-oxidase
CDH	Cellobiose dehydrogenase
CNT	Carbon nanotubes
DCHBS	3,5-dichloro-2-hydroxybenzene-sulfonic acid
DET	Direct electron transfer
dNTPs	Deoxyribonucleotide triphosphate
ESRF	European Synchrotron Radiation Facility
epPCR	Error prone polymerase chain reaction
FAD	Flavin adenine dinucleotide
GDH	Glucose dehydrogenase
GMC	Glucose-methanol-choline
GOx	Glucose oxidase
HRP	Horseradish peroxidase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Luria-Bertani
MAD	Multiple anomalous diffraction
MIR	Multiple isomorphous replacement
NMR	Nuclear magnetic resonance spectroscopy
OD_{600nm}	Optical density at 600 nm
P2Ox	Pyranose 2-oxidase
PCR	Polymerase chain reaction
PDB	Protein data bank
PDH	Pyranose dehydrogenase
Pink chromogen	N-(4-antipyryl)-3-chloro-5-sulfonate-p-benzoquinone-monoimine
rmsd	Root mean square deviation
SAD	Single anomalous diffraction
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SIR	Single isomorphous replacement
TLSMD	Translation, Libration, Screw Motion Determination
TRIS	Tris(hydroxymethyl)aminomethane
WT	Wild-type

1. Introduction

1.1 The GMC oxidoreductases superfamily

The glucose-methanol-choline (GMC) superfamily of enzymes is a group of FAD-dependent oxidoreductases that share a surprisingly similar primary sequence despite having distinct substrates: electron donors, from small alcohols to sugar moieties, and various electron acceptors such as molecular oxygen and small chemical compounds. GMC superfamily includes glucose dehydrogenases (GDH), alcohol (methanol) oxidases (AOx), glucose oxidases (GOx), choline dehydrogenases (CHD), cellobiose dehydrogenases (CDH), pyranose oxidases (P2Ox) and pyranose dehydrogenases (PDH) [1].

It is known that these groups of homologous enzymes also share a common three-dimensional structural fold [2]. The FAD binding domain is strictly conserved on fungal P2Oxs, allowing a covalent or non-covalent binding of the prosthetic group and, not surprisingly, the substrate-binding domain, shows several variations, supporting the disparate substrates used by this group of enzymes. The conserved catalytic residues are a pair of His/His or His/Asn, which are located in the neighborhood of the FAD isoalloxazine ring [2]. These enzymes share similar reaction mechanisms, with a reductive half-reaction followed by an oxidative half-reaction The final electron acceptor defines the enzyme classification: it is an oxidase if oxygen is one of the substrates used as electron acceptor, being reduced to hydrogen peroxide, or a dehydrogenase if O_2 is not the final electron acceptor [2], [3].

Fungal GMC oxidoreductases can presumably help lignocellulose degradation by the native microorganisms, producing H_2O_2 for lignin degrading peroxidases, and by directly reducing phenoxy radicals and quinones, crucial steps in ligninolysis [4], [3]. For this reason, there has been a special interest in investigating fungal GMC oxidoreductases for their potential application in biomass utilization and in the food industry [2]. Other applications like biosensors will be addressed later.

The P2Ox from the bacteria *Pseudoarthrobacter siccitolerans* belongs to the GMC oxidoreductase family and is the focus of this work. Recently, phylogenetic analysis of fungal GMC sequences demonstrated that P2Ox shows a widespread taxonomic distribution when compared with the other enzymes of the GMC superfamily. It was proposed that P2Ox was introduced into fungi by horizontal gene transfer from bacteria with an apparent gene loss in many fungi caused by a redundant function of P2Ox [2].

1.2 Fungal Pyranose 2-Oxidases

In spite of the vast taxonomic distribution and high sequence variation within P2Ox, when compared to other GMC oxidoreductases, most of the P2Ox biochemically studied to date have fungal origins [2], [5]. These enzymes preferentially oxidize D-glucose at the C2 position, despite also being capable of oxidizing other monosaccharides such as D-xylose, D-galactose and L-sorbose, yielding the corresponding 2-keto sugar, with subsequent molecular oxygen reduction to hydrogen peroxide [5], [6]. P2Oxs were found in the extracellular and periplasmic space, the latter localized at fungi hyphae, therefore being potentially able to produce hydrogen peroxide to assist lignocellulose degrading peroxidases [7]. These latter enzymes produce phenoxy radicals by depolymerization of lignin, and these radicals can be further repolymerized hampering lignin degradation. P2Ox can use these radicals as electron acceptors and inhibit the repolymerization of lignin fragments. Since these non-oxygenic electron acceptors frequently result in higher catalytic efficiencies, they might be used by P2Ox in a later stage of lignocellulose degradation due to its increase in concentration, while at an initial stage, the enzyme uses O_2, and produces H_2O_2. P2Ox is also able to reduce quinones, and metal complexes of some peroxidases, for example manganese peroxidase [8] [9]. For these reasons, P2Ox is a member of auxiliary activity family 3 (AA3) of the carbohydrate-active enZymes database (CAZy: http://www.cazy.org/) that comprises enzymes which assist lignocellulose degradation [5]. In some fungi, the product 2-keto-glucose can be further metabolized to the antibiotic cortalcerone [10].

The P2Ox active site has the typical catalytic His/Asn pair found in the GMC family, positioned below the FAD isoalloxazine ring (Figure 1) [11], [5]. In all three fungal P2Ox structures studied to date, (Table 1), namely from the white-rot fungus *Peniophora sp.* (*Pe*P2Ox) [11], *Trametes multicolor* (*Tm*P2Ox) [12], and *Phanerochaete chrysosporium* (*Pc*P2Ox) [13], the conserved histidine residue acts as the catalytic base, responsible for the deprotonation of the monosaccharide [14], while the asparagine participates in important hydrogen bonds with the sugar.

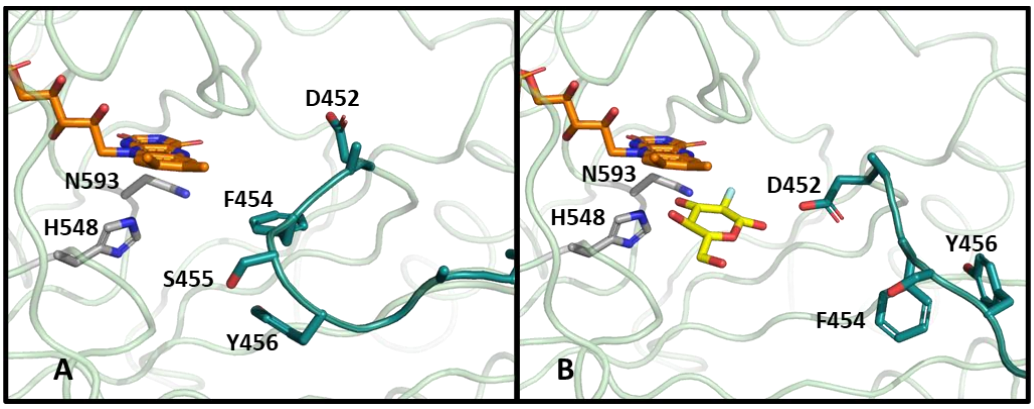

Figure 1 – Closed (A) and open (B) loop conformations of *T. multicolor* P2Ox (A - PDB 1TT0; B – PDB 2IGO). The loop can be seen in blue, the FAD in orange and the catalytic residues His/Asn in white. Oxygen and nitrogen atoms are represented in red and blue, respectively.

As in other GMC oxidoreductases, the reaction mechanism consists of two half-reactions: a reductive half-reaction where FAD is reduced to $FADH_2$ by the electron donor (such as D-Glucose) and an oxidative half-reaction where $FADH_2$ reduces O_2 to H_2O_2 or other electron acceptors. The mechanism is suggested to be of the Ping Pong Bi Bi type where the keto-sugar product leaves before the oxidative half-reaction (oxygen reduction). In Figure 2, we can see that the sugar binds to the active site in a two-step binding process: complex formation and isomerization [14], [15]. The active site has a dynamic substrate loop gating the active site (residues 452–457 in *Tm*P2Ox: PDB 1TT0) and allowing the entry and exit of electron donors. This loop can adopt both open and closed conformations, depending on the half-reaction (Figure 1). In the reductive half-reaction, the sugar substrate can enter the active site when the substrate loop is in an open conformation (PDB 2IGO *Tm*P2Ox). In the oxidative half-reaction, after the keto-sugar release, the substrate loop can be found in a closed conformation (PDB 1TT0 *Tm*P2Ox). This loop can also be accountable for defining regioselective catalysis [12]. The first step of the sugar oxidation at C2 position is a hydride transfer (C-H bond breakage) from the monosaccharide to the flavin N5 atom, generating a protonated ketone and a reduced flavin. The protonated ketone is shortly after deprotonated by the conserved catalytic base (histidine), generating the keto-sugar product that is released before initiation of the oxidative half-reaction (Figure 2).

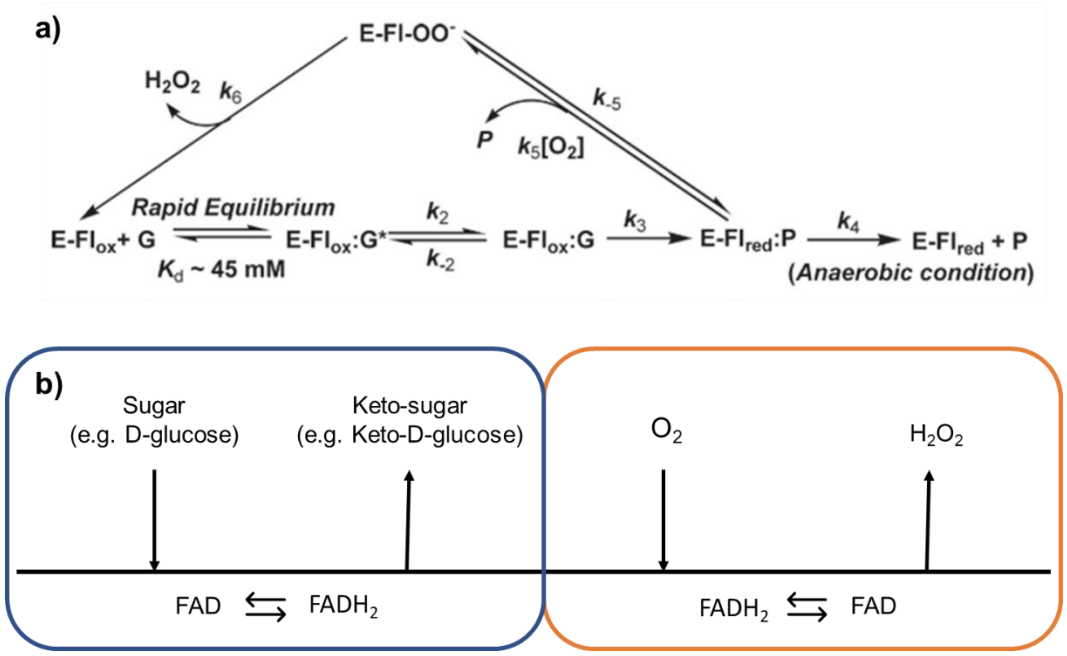

Figure 2 – Proposed catalytic scheme for P2Oxs. a) Detailed kinetic mechanism of the reaction of P2Ox taken from [15], showing a two-step binding of the sugar. E-Fl_{ox}, oxidized form of P2Ox; E-Fl_{red}, reduced form of P2Ox; E-Fl_{ox}:G*, P2Ox:D-glucose complex; E-Fl_{ox}:G, P2O:D-glucose complex after isomerization of D-glucose; E-Fl_{red}:P P2Ox:Keto-D-glucose complex; E-Fl-OOH, C4a-hydroperoxy-FAD intermediate **b) Simplified Cleland notation for the ping pong bi bi type mechanism of the reaction of** P2Ox**.** Besides D-glucose, other sugars like D-galactose, D-frutose and D-xylose can be reduced. Other electron acceptors can be found (quinones, complexed metal ions, radicals). The reductive half reaction is outlined in blue, while the oxidative half reaction is outlined in orange.

The enzyme is also capable of oxidizing sugars at C3, such as 2-deoxy-D-glucose and 2-keto-glucose, [16]. The lack of the C2 hydroxyl group seems to be the main factor for C3 oxidation binding mode. However, the C2 binding mode of sugars seems to be promoted by interactions with the substrate loop, since its deletion in computational modelling, allowed D-glucose to be modeled in both C2 and C3 oxidation binding modes [17]. In the C2 oxidation binding mode of *Tm*P2Ox, the substrate's C6 hydroxyl group interacts with residues (D452 and Y456) that are part of the substrate loop (PDB 3PL8). However, these interactions in a C3 binding mode are not present (PDB 2IGO), reinforcing the idea that the preference of C2 over C3 oxidation could be controlled by the substrate loop.

Despite its similarities to other GMC oxidoreductases, P2Ox is the first flavoprotein oxidase for which a C4a-hydroperoxy-FAD intermediate was reported in the oxidative half-reaction [18]. In 2009, Warintra Pitsawong and her colleagues suggested that there is an important threonine residue (T169 in *Tm*P2Ox) near the active site that can form an H-bond with the FAD^{N5}, which can affect flavin reduction. This residue is also responsible for stabilizing the C4a-hydroperoxy-FAD intermediate in the reduction of O_2 [19].

The preferential substrate of all fungal P2Ox described to date is D-glucose with a Michaelis constant (K_m) ranging from 0.7 to 5.0 mM and k_{cat} values from 1.5 to 111 s^{-1}. Despite the presumed "natural" electron acceptor being oxygen as previously mentioned, P2Ox can reduce a vast group of other molecules such as quinones (1,4-benzoquinone), ABTS and complexed metal ions, some of them resulting in better catalytic efficiencies than O_2 (lower K_m and higher k_{cat}) [5], [6], [20].

All P2Oxs structurally characterized from fungal sources (Table 1) are homotetrameric containing one FAD molecule per monomer (Figure 3). The tetramer forms a large central cavity or void, connected to all active sites.

Table 1 – Fungal P2Oxs with available crystal structure, corresponding resolution and PDB code. There are also variants and structures complexed with substrates of these P2Ox that have their structure determined [13], [21].

Organism	Resolution (Å)	PDB code	Ref
Peniophora sp.	2.35	1TZL	[11]
Trametes multicolor	1.8	1TT0	[12]
Phanerochaete chrysosporium	1.8	4MIF	[13]

The active center is located at the three-ring structure of the flavin, called isoalloxazine [11], [5], and substrates can only enter through narrow glucose-sized channels, making the oxidation of polysaccharides impossible. All structures studied to date share a common chain fold with the rest of GMC superfamily enzymes and have a FAD covalently bonded via its 8a-methyl group to N3 of a histidine residue (H167 in *T. multicolor*), resulting in an 8a-(N3)-histidyl flavinylation of the enzyme. The flavinylation in fungal P2Ox is usually associated with the amino acid motif STHW (^{165}STHW^{168} in *Tm*P2Ox). Fungal P2Oxs have a flavin-binding domain (Rossmann fold or βαβ mononucleotide-binding motif) with around 200 residues and a substrate-binding domain with

some extent of sequence variation, in accordance with the different substrates used by P2Oxs from different sources. This latter domain has a six-stranded central β sheet and three α helices, having around 370 residues [5], [12]. On top of these two major domains, fungal P2Oxs have an oligomerization arm domain and oligomerization loop, responsible for their homotetrameric form, and a head domain, thought to be responsible for the association with polysaccharides from the cell-wall for immobilization at the fungal periplasmic space.

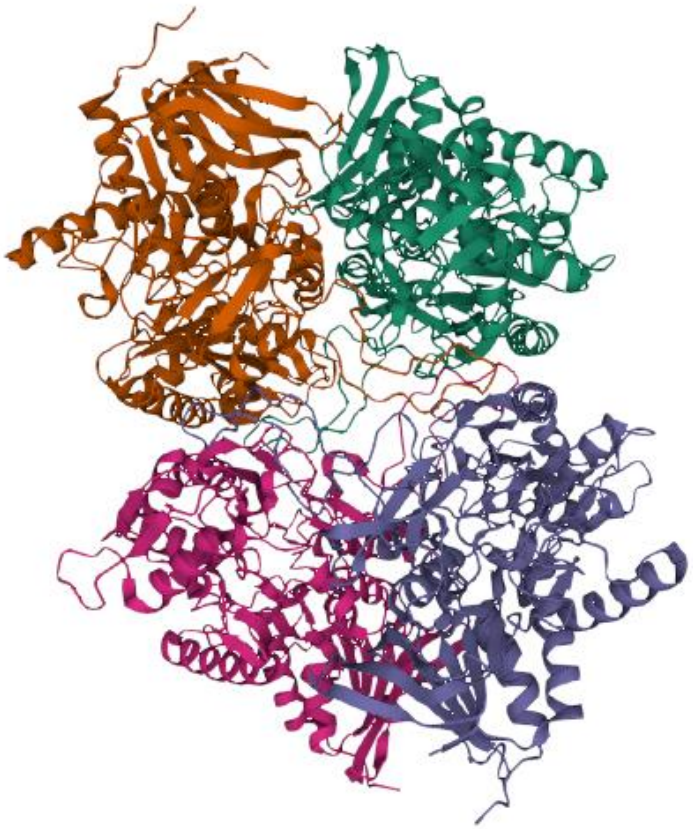

Figure 3 – Crystal structure from the (fungal) *Trametes multicolor* homotetrameric P2Ox at 1.80 Å resolution, taken from the Protein Data Bank (code: 1TT0) and described by B. Martin Hallberg [12].

1.3 Bacterial Pyranose 2-Oxidase

P2Ox can also be found in actionobacteria, proteobacteria and bacilli, being the only GMC oxidoredutase present in both bacteria and fungi [2], [8]. As stated above, there are evidences that the fungal P2Ox gene has bacterial origins. To confirm this later statement, a phylogenetic tree (Figure 4) was built by P. Herzog and colleagues [8]. In figure 4 one can confirm the close relationship between fungal and bacterial P2Ox, since for example *Kitasatospora aerofaciens* P2Ox (bacteria) is more related to many fungi than to other actinobacteria like *Pseudoarthrobacter siccitolerans* P2Ox. Three bacterial P2Oxs have been studied to date, none of them having its structure determined. In contrast to its homotetrameric fungal counterparts, bacterial P2Ox quaternary structure seems to vary between bacterial organisms: monomers, dimers and even tetramers [4], [8], [22].

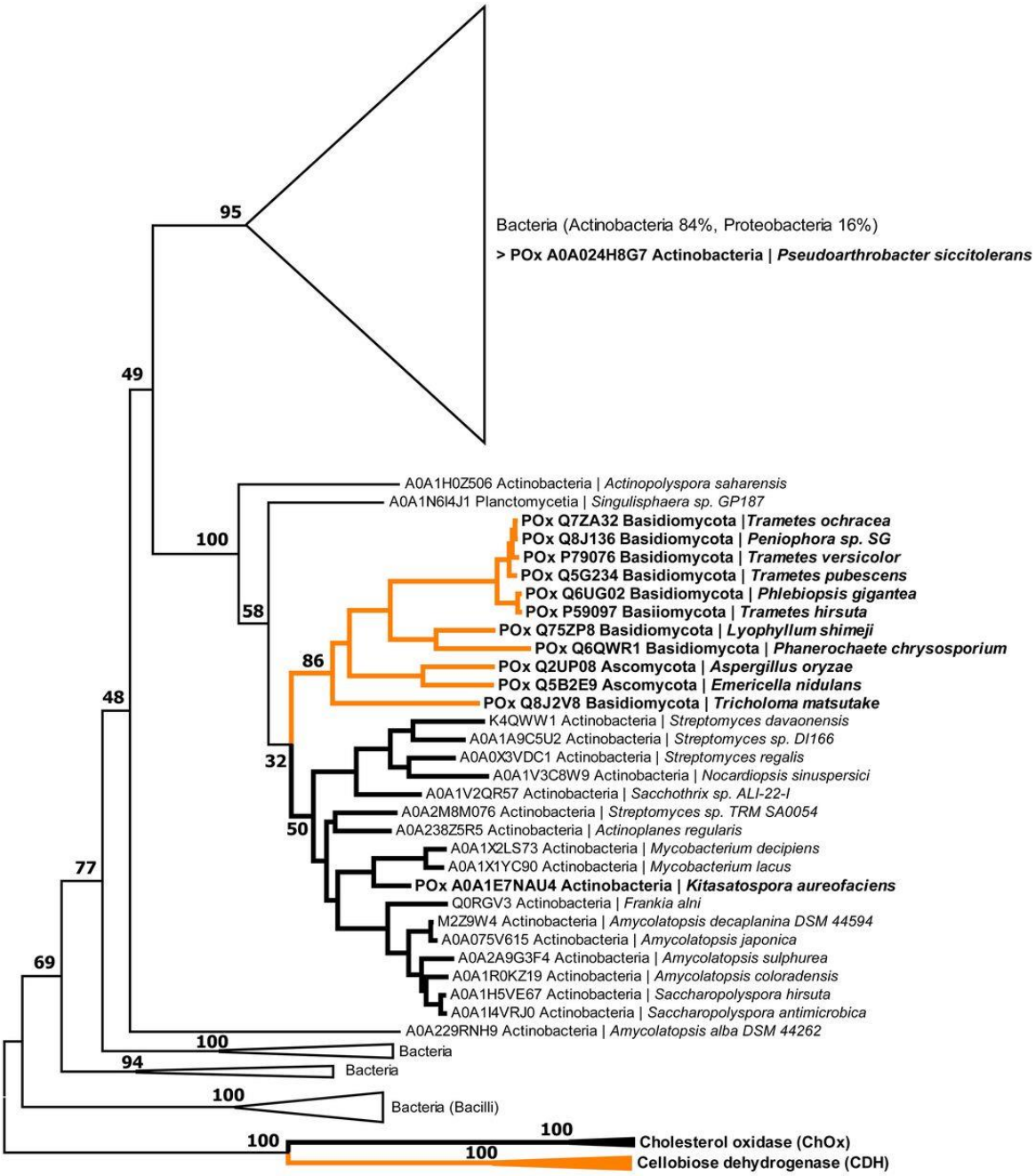

Figure 4 – Phylogenetic tree of P2Ox from several bacterial (black) and fungal (orange) origin, taken from [8]**.** Numbers in the graph represent bootstraps coefficients, expressed as percentages. Sequences from characterized enzymes are represented in bold.

The first bacterial P2Ox characterized was from the bacterium *Pseudoarthrobacter siccitolerans* (*As*P2Ox), a Gram-positive desiccation-tolerant species found in common soils and extreme environments such as dry soils [4], [23]. This enzyme showed some different and interesting properties, when compared to its fungal counterparts studied to date. FAD is not covalently bonded, and *As*P2Ox is monomeric in solution, with 64 kDa. Similarly to fungal P2Ox, *As*P2Ox also preferentially oxidizes D-glucose (K_m and k_{cat} with O_2 as electron acceptor of 370 mM and 0.51 s^{-1} [24]), being also able to oxidize D-galactose, D-xylose, L-arabinose, and D-

ribose. Besides molecular oxygen, it can also use 1,4-benzoquinone as electron acceptor, the latter being a preferred electron acceptor. *As*P2Ox has an optimum pH of 6.5, an optimum temperature of 37°C, and a relatively low thermal stability, when compared to fungal P2Oxs, with a melting temperature of 43°C. The first two residues of the flavinylation motif from fungal P2Ox (^{165}STHW^{168} in *Tm*P2Ox) are replaced by two alanine residues, ^{125}AAHW^{128}, in *As*P2Ox. Despite the differences, *As*P2Ox has a histidine residue (His440) as the catalytic base in the active site, as in fungal organisms [4]. In order to reach a more stable and catalytic efficient enzyme for future industrial applications a directed evolution approach in *As*P2Ox has been performed in MET Lab (Microbial & Enzyme Technology Lab) throughout the years [24].

The second bacterial P2Ox characterized was from *Pantoea ananatis* (*Pa*P2Ox) [22]. Unlike *As*P2Ox, *Pa*P2Ox is homotetrameric. This enzyme has a steroid-binding domain that may act as a substrate-binding domain. When compared to *As*P2Ox, it has a higher optimal temperature of 50°C. *Pa*P2Ox can oxidize D-glucose (highest activity), D-mannose, D-galactose, and DL-arabinose. Reduction of phenoxy radicals by *Pa*P2Ox was also reported.

Finally, the characterization of a P2Ox from *Kitasatospora aureofaciens* (*Ka*P2Ox) was reported [8]. This P2Ox, unlike the other bacterial enzymes, has a covalently bonded FAD, and has a homodimeric structure in solution. KaP2Ox is not an exception and displays a preference for D-glucose oxidation. The investigation of this GMC oxidoreductase reinforces the idea that P2Ox can reduce quinoid intermediates produced in lignocellulose degradation, thus preventing their repolymerization. *Ka*P2Ox is able to reduce Mn(III), which can be found in some lignin degrading peroxidases as a complexed metal [8].

In general, the bacterial P2Oxs characterized to date, despite their lower catalytic activities, convey the impression of sharing general similar structural characteristics and presumably fulfilling similar roles as those exhibited by P2Ox from lignin-degrading fungi.

1.4 Applications of P2Oxs – biofuel cells and biosensors

Pyranose oxidases have promising applications in synthetic carbohydrate chemistry, production of new drugs, rare sugars and sweeteners, clinical, industrial and pharmaceutical monitoring, food safety, agricultural control and energy generation (biofuel cells) [5], [25], [26].

1.4.1 Biosensors

A biosensor can be defined as an analytical device for qualitative and quantitative identification of a specific target analyte. In these devices, there is a biological component, responsible for recognition of the analyte, and a physicochemical detector, the transducer. The recognition element is immobilized onto the transducer's surface and can interact with the target analyte, producing changes that can be recognized by the transducer. The transducer is then able to convert these changes into measurable signals that relate to the concentration of analyte. To date, several examples of cells, enzymes, nucleic acids, and antibodies were identified, acting

as biological components to make the bridge between the analyte and the transducer. There are also different types of identified transducers: electrochemical, optical, thermal, and piezoelectric (mass based) [25].

Enzyme-based biosensors are the most used and studied type of biosensors, being applied in both clinical and industrial applications. Clear advantages are their high specificity and sensitivity, low cost, miniaturization, portability, and mass-production, all of these making point-of-care diagnostics possible [25], [26]. Using enzymes as the biological component, high affinity enzyme-substrate interactions can be achieved. As mentioned above, although there are four categories of transducers, electrochemical biosensors are the most used ones. In this type of biosensors, changes in current, voltage, resistance or superficial charge occur due to a redox reaction catalyzed by the enzyme immobilized onto the transducer surface [25].

One of the most common and extensively used electrochemical enzyme-based biosensors is the amperometric biosensor, which can measure the current and is extensively used in glucose levels measurement. Indeed, the redox reaction produces a change in magnitude of the current, which is proportional to the concentration of the analyte. There are three generations of amperometric enzyme biosensors, defined by the method of electron transfer (figure 5). In first-generation biosensors, the analyte is monitored with coenzymes by measuring the product of the redox reaction or the consumption of the co-substrate, for example oxygen as electron acceptor. Drawbacks of these first-generation biosensors are the need of coenzymes and the dependency on electron acceptor concentration like oxygen. For this reason, second-generation biosensors were created (oxygen-independent), being able to use other electron acceptors besides oxygen as mediators for electron transport between the enzyme and the electrode. Despite their advantages, second-generation biosensors experience mediator leaching and interferences. Third-generation biosensors use direct electron transfer (DET) between the enzyme and the electrode. However, not all enzymes can be applied in third-generation biosensors, since not all are able to transfer electrons directly to the electrode, this being one of the main issues for the successful enzyme application in biosensors [25], [27]. For this reason, new approaches to enzyme biosensors have been studied in the past years, from new immobilization techniques to enzyme engineering.

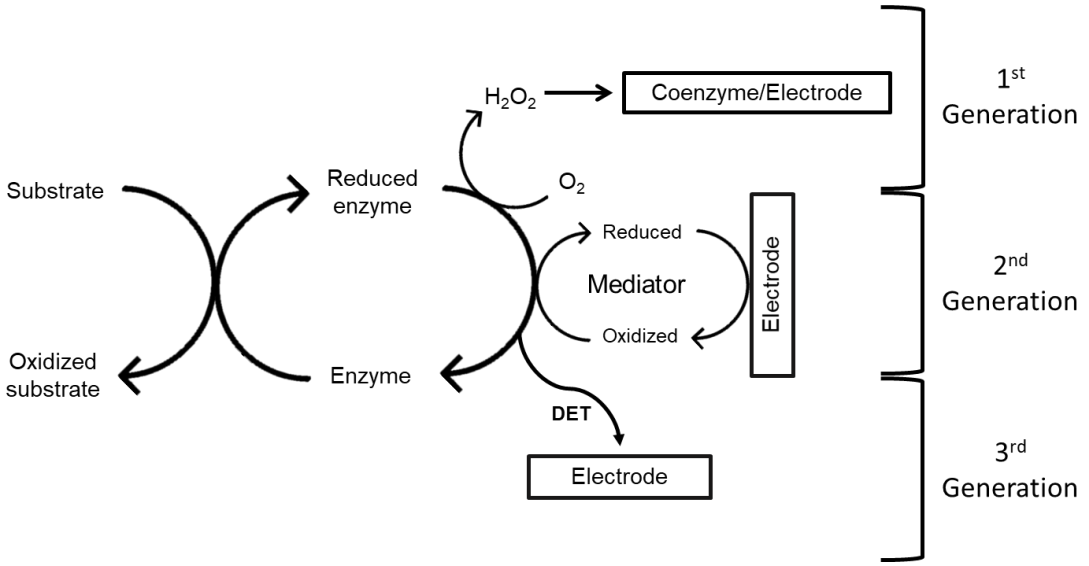

Figure 5 – Schematic representation of the three generations of amperometric enzyme biosensors for glucose. In first generation biosensors, coenzymes are used as a bridge between the main reaction and the electrode; O_2 consumption is usually monitored. In second generation biosensors, mediators are used to directly transfer electrons (DET – direct electron transfer) between the enzyme and the electrode. In third generation, biosensors the electrons are directly transferred to the electrode.

Glucose biosensors are the most important and commercially successful biosensors to date, for example for determination of glycemic levels in blood, the most popular being based on immobilized GOx [28]. This enzyme oxidizes β-glucose to gluconolactone, while reducing oxygen to hydrogen peroxide that can be detected by peroxidases in first-generation biosensors [28]. Studies with GOx/carbon nanotubes (CNT) based biosensors indicate that GOx is unable to transfer its electrons directly to the electrode, preventing third-generation biosensors to be designed with this enzyme. It was proposed that this happens because the redox center is shielded by an electronically nonconductive glycoprotein coat [29] and DET has already been proven on a biosensor with de-glycosylated Gox [27], [30]. Despite its noted commercial success in biosensors use, GOx has anomeric selectivity, being only able to oxidize β-D-glucose, hampering its affinity for D-glucose. In this regard, P2Ox shows higher electron transfer and glucose transport rates and does not have anomeric selectivity towards glucose [27]. P2Ox is also able to oxidize a considerable number of other monosaccharides [31]. In fact, this enzyme was already applied for multianalyte determination in a first-generation biosensor composed of co-immobilized P2Ox and horseradish peroxidase in carbon paste, used in combination with liquid chromatography for carbohydrates detection [32]. Other P2Ox-based biosensors for glucose detection, mainly for food and beverage processes monitoring, have been reported to date [33], [34], [35] with various types of electrodes being explored, such as glassy carbon, gold, and carbon nanotubes [27], [34], [35].

P2Ox biosensors can be used for quantitative measurement of 1,5-anhydro-D-glucitol, an analogue of D-glucose found in human serum and human cerebrospinal fluid [36]. This analyte is secreted in the urine and in normal conditions it is reabsorbed. However, when serum glucose levels are high, its reabsorption is suppressed due to competitive tubular reabsorption (between

1,5-AG and glucose) and ends up being excreted in the urine. Thus, its serum levels decrease can be measured and can act as a marker of hyperglycemia [37], [38].

Only fungal P2Ox have been studied and proposed as biological components of biosensors. To date, no third-generation P2Ox-based biosensors were created due to the fact that DET cannot be achieved in fungal P2Oxs since the active site does not have an easy access, being in the core of the tetrameric protein [5]. *As*P2Ox might bring new possibilities to P2Ox biosensors, since its monomeric state increases the chance of successful DET between the enzyme and the electrode, supporting the importance of advancing the knowledge of *As*P2Ox. However, *As*P2Ox does not have the desired catalytic efficiency and stability, and enzyme engineering could be used to improve the biocatalyst for its application in biosensors.

1.5 Protein engineering – from directed evolution to rational evolution

Protein engineering takes advantage of the sequence to function relation of enzymes and there are a few methodologies available that try to find the best sequence to perform a specific and desired function. Engineered proteins may have improved solubility, stability, activity, and selectivity, to better fit their applications [39]. Despite this sequence to function relation of proteins, there are clear knowledge and experimental limitations for its application. If we think of a small 100 amino acid protein and consider that each sequence position has 20 possible outcomes (20 natural amino acids), there are 20^{100} possible sequences. Adding the fact that most proteins have more than 100 amino acids, there is no *in vivo* approach to date that can cover such a vast sequence space. For this reason, *de novo* design of a protein is still not a valid option, and protein engineering uses already known proteins and improves or changes their functions. Usually, protein engineering utilizes rational design or directed evolution to create variants of a starting wild-type protein by mutagenesis [40], [41].

1.5.1 Directed evolution

Directed evolution uses evolutionary approaches to mimic natural evolution and create a library of mutants of the target protein – Figure 6. One of the main limitations of directed evolution is the requirement of sensitive, efficient, and reliable high-throughput screening methods to select the improved variants, since the library comprises a considerable number of mutants [42]. Insertion of mutations can be accomplished by random mutagenesis and/or recombination. Random mutagenesis uses chemical or physical agents to create point mutations. *In vitro* is commonly preferred to *in vivo* (*e.g.* mutator strains [43]) random mutagenesis, and typically error-prone PCR (epPCR) can be applied, where genetic diversification is introduced into the gene of interest using reaction conditions that reduce the fidelity of Taq DNA polymerase (error rate of 1 to 20 nucleotides per 1kb per PCR reaction) [44], [45].

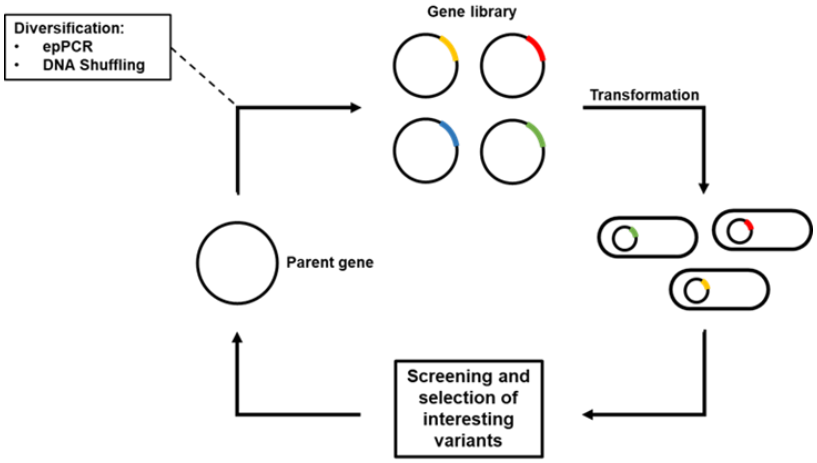

Figure 6 – Schematic representation of directed evolution. epPCR: error-prone polymerase chain reaction. Different colours in the genes/plasmids represent random mutations obtained by diversification methods (e.g. epPCR).

Recombination techniques such as homologous and non-homologous recombination can be used to combine already existing beneficial mutations and enrich the library of mutants. DNA shuffling is a widely used method for homologous recombination where gene variants are randomly cleaved into fragments that prime to each other in order to be reassembled in a PCR reaction [41], [45].

1.5.2 Rational design

Rational methods for protein engineering only focus on the replacement of amino acid positions of potential interest. Therefore it requires previous structural, mechanistic and functional knowledge of the target protein [40], [46]. Precisely, rational approaches are limited by the difficulty of predicting which amino acid replacement will lead to a required change in function [41]. In rational design, mutations are predetermined by educated guessing or computational modeling and then introduced using site-directed mutagenesis. Two main techniques (Figure 7) are used for this latter purpose: overlap extension method, where two out of four primers, used in two different PCR reactions, contain the mutant codon, and whole plasmid single-round PCR, which uses selective digestion of the methylated parental template and consequent transformation of mutated nicked dsDNA (circular) into competent cells that repair the nicks [41]. After these mutation methods, confirmation of the mutation by sequencing means is necessary and the variants can then be biochemically characterized.

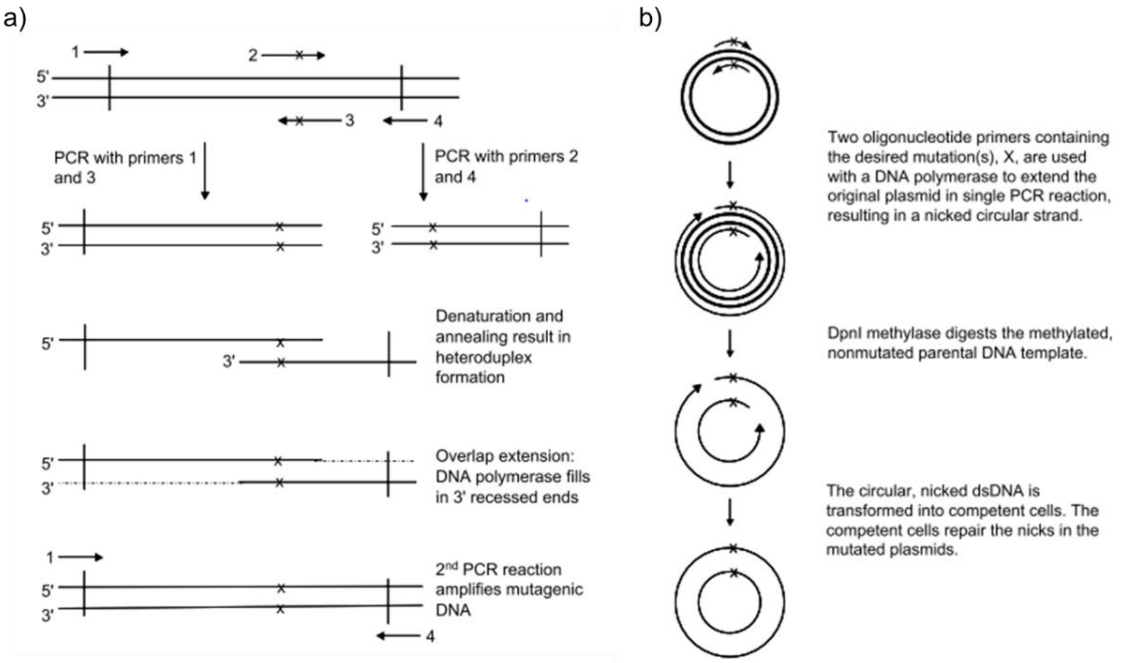

Figure 7 – a) Overlap extension PCR method and b) whole plasmid, single-round PCR. (→) represents a primer and x represents a mutagenic codon. Image taken from [41].

1.5.3 Other approaches and future perspectives

Although each method has its advantages and drawbacks, a combination of both might be the most effective route to enzyme engineering. For example, directed evolution can give structural and functional insight for later application of rational design [46] or it can uncover additional beneficial mutations in already rationally engineered enzymes [47].

Semi-rational approaches such as localized random mutagenesis (site saturation mutagenesis) have also shown great potential for enzyme engineering, where specific regions thought to be important for function are randomly mutated. Semi rational approaches also need some structural or biochemical knowledge. Simultaneous saturation of target residues can provide combinations of epistatic mutations that are for example, beneficial only in combination [41], [47]. Saturation mutagenesis was applied to residues in the active site of TmP2Ox for improved catalytic activity using D-glucose and D-galactose. The variants obtained following this semi-rational approach showed promising properties for applications in biofuel cells and production of D-tagatose, a non-glycemic and low caloric sweetener [48] [49].

Computational methods show an increasing value in biochemistry and are most probably the future of protein engineering. Machine learning guided directed evolution can be used to predict sequence to function relationships, being able to utilize unimproved sequences to predict new variants to screen. Linear regressions and classification and regression trees are some of the mostly used machine-learning models used for “directed evolution” [50]. Computational modeling is also a widely used tool to predict point mutations for rational design [41]. Single-point mutations, however, have little contribution to enzyme stability and new software for improved stability has been developed and successfully tested in the past recent year. Structure and sequence based

algorithms such as FireProt [51], PROSS [52] and FRESCO [53], combine molecular dynamics (energy based approach) with evolutionary approaches to reach multiple-point mutations. Despite their potential, all these engineering tools require homologous sequences and the 3D structure of the target enzyme, increasing the importance of the main goal of this work, the determination of *As*P2Ox's crystal structure.

1.6 General concepts of macromolecular X-ray crystallography

The knowledge of the three-dimensional structure of an enzyme is crucial to predict beneficial mutations for the enzyme catalytic activity and stability using rational design and computational tools. In order to determine structural information, a wide variety of techniques such as NMR spectroscopy, electron microscopy, fiber diffraction, small angle X-ray scattering (SAXS) and X-ray crystallography [54] can be used. X-ray crystallography is the most widely and efficient way of obtaining three-dimensional structures. For enzymes, these models not only give us insight into the overall structure of the protein, but also of the active site, catalytic residues and other important molecular details that allow us to understand biological processes at the most basic levels: how the enzyme interacts with the substrate and inhibitors, how enzymes catalyze reactions, and might help developing new drugs, for example.

Since the first protein structure (myoglobin) was published in 1958 [55], X-ray crystallography has improved considerably. New crystallization techniques, as well as new X-ray sources, and modern computer software have allowed an important acceleration and ease of structure determination [56]. In 1971, the Protein Data Bank (PDB) was created to function as an archive of protein and proteins-ligands structures and the number of structures available now exceeds 180 000, manly obtained by X-ray crystallography (87.6%) [57].

1.6.1 Macromolecule crystallization

Once we have the purified protein, the first step in X-ray crystallography is the production of high-quality crystal suitable for diffraction studies. The crystal is an orderly three-dimensional array of molecules held together by noncovalent interactions. The unit cell is the smallest and simplest volume element that represents the whole crystal (Figure 8). The unit cell reflects the symmetry and structure of the entire crystal that is built up by repetitive translation of the unit cell along its main axes. The lengths (*a,b,c*) of the principal axes of the unit cell and the angles (*α,β,γ*) between them are called as cell parameters. There are 14 Bravais lattices belonging to seven possible crystal systems. When combining the 32 point groups (all possible unique combinations of crystallographic symmetry elements) with the 14 Bravais lattices we find up to 230 space groups. The asymmetric unit of a space group is that part of the crystallographic unit cell which can be used to generate the entire unit cell by symmetry of the space group [58].

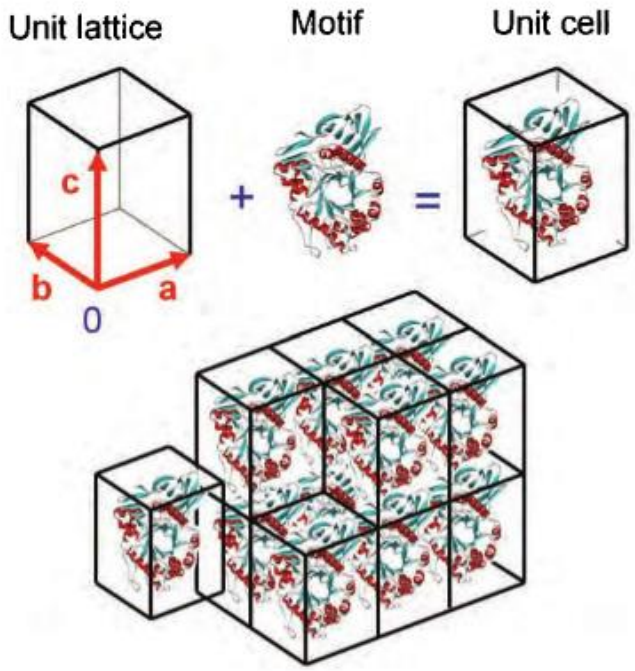

Figure 8 – Unit lattice, motif and unit cell of a crystal and translation of the same unit cell to build the crystal. Image obtained from [59].

Since there is no obvious correlation between crystallization conditions and protein structure, production of good quality crystals can be the most time-consuming step of this technique. Screening methods include sparse-matrix screens, relying on conditions that previously led to good quality crystals or based on properties of precipitants and additives. Systematic screenings are also part of X-ray crystallography screening methods, a more rational screening technique that explores information on the protein's properties. Crystallization screenings test a vast diversity of conditions, changing variables like pH, ionic strength, temperature and concentration of protein and precipitant [60], [61].

Crystals may start to form when the concentration of protein in solution passes the solubility limit, reaching the nucleation zone of the supersaturated state. Once nucleation occurs, the protein concentration drops and the crystals start growing as they enter in a metastable zone [62]. Thus, crystallization occurs in two phases: nucleation and growth. A schematic illustration of a protein crystallization phase diagram can be found in Figure 9.

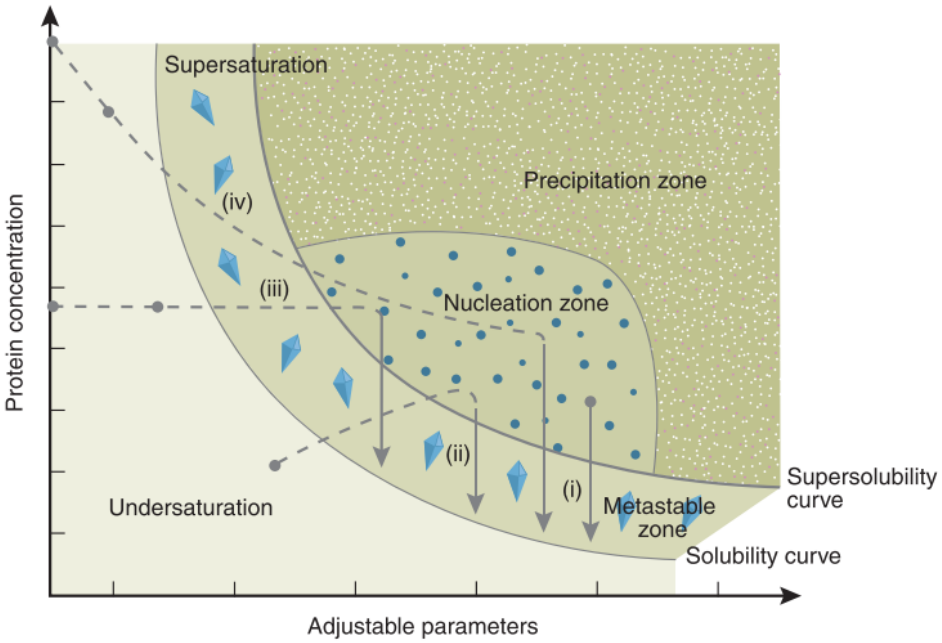

Figure 9 – Schematic representation of a phase diagram and routes, assuming the adjustable parameter is precipitant concentration, for different crystallization methods: (i) microbatch, (ii) vapor diffusion, (iii) dialysis and (iv) free interface diffusion. Grey dots represent the starting point (starting conditions) of each method. Adjustable parameters: precipitant and additives concentration, pH, ionic strength and temperature. Image taken from [61] (adapted from [63]).

Several approaches have been developed to combine factors that promote crystallization. The most widely used crystallization method is vapor diffusion (Figure 10), where a protein solution in a hanging or sitting drop equilibrates against a reservoir containing the reservoir solution. The concentration of precipitant in the reservoir will be higher than the one in the drop, forcing water equilibration by vapor diffusion. Due to this equilibration, the protein concentration in the drop gradually reaches supersaturation and consequently crystal formation and growth [62].

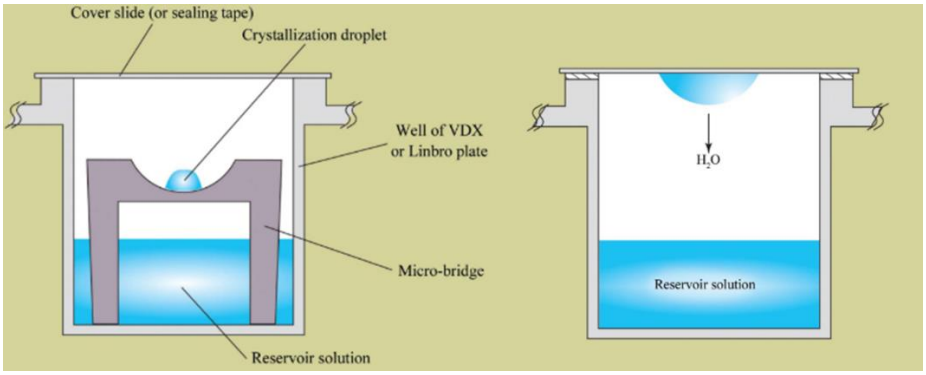

Figure 10 – Representation of the sitting drop (on the left) and hanging drop (on the right) diffusion crystallization methods. In both cases, a drop of protein solution mixed with the reservoir solution equilibrates with the reservoir. Since the drop has a lower concentration of precipitating agents, water equilibration through the vapor phase will occur, increasing the concentration of the protein, ideally until the nucleation zone is reached. Image adapted from [64].

Micro batch, dialysis, and free interface diffusion are also commonly used methods of crystallization (Figure 9) [61],[65], [66] . Microfluidics have been increasingly used for high throughput screening due to its minimal protein and precipitant consumption [67]. Seeding allows the separation of the nucleation and growth processes, a technique where crystals are transferred from nucleation conditions to growth conditions [61]. Robots are also a time-saving tool that allows high throughput screening with minimal amounts (nanoliter volumes) of protein. [56]. They can reproducibly set thousands of conditions in a few hours, using lower volumes of protein than previously necessary. After these high throughput screenings, scale-up of the hit crystallization condition (now using microliter volumes of protein) is of extreme importance and consists of optimization of parameters such as protein concentration, protein/precipitant ratio, temperature, pH and ionic strength.

1.6.2 X-ray diffraction experiment

Once crystals suitable for diffraction experiments are obtained, it is possible to get the diffraction pattern by making an x-ray beam cross the crystal. X-ray beams can easily damage or destroy the crystals and so a cryo-cooling of the crystals prior to (in liquid nitrogen) and during any diffraction experiment (with a nitrogen gas stream at 100 K) is of extreme importance. For that, the crystal must be transferred to a cryoprotective solution (e.g., by supplementing the mother liquor with 20% glycerol) to prevent damage from the cryo-cooling itself, such as ice formation within the crystal. X-rays can be generated when an accelerated electron beam in vacuum hits a target made of a metal such as copper (Cu) or molybdenum (Mo), each one producing X-rays with different characteristic wavelengths. In the 1970s, synchrotron radiation became a widely used tool for protein crystal diffraction. These facilities made it possible to choose the wavelength of the X-rays, a much-needed characteristic to exploit anomalous scattering (see below). Very recently, X-ray free electron lasers (XFEL) were introduced, where electrons accelerated in a linear device emit extremely intense and very short pulses of electromagnetic radiation [56].

X-ray crystallography takes advantage of the X-ray diffraction patterns produced by crystals, seen in Figure 11. According to Lawrence and William Bragg, diffraction can be viewed as reflections from families of parallel lattice planes with Miller indices h, k and l, where the incident and reflected rays make the same angle θ with the reflecting lattice planes. However, the θ angles have a strict relation with the wavelength of the incident beam and the interplanar spacing d_{hkl}, leading to selective conditions for reflection. Bragg's law defines that relation: $2\, d_{hkl} \sin\theta = n\lambda$, where n is the diffraction order (integer), meaning that the difference in optical paths between reflected rays from consecutive planes (hkl) must be equal to nλ (i.e., be in phase) for diffraction to happen. Thus, each diffracted beam (spot in Figure 11) corresponds to a set of reflecting planes with a specific interplanar distance (a spot further away from the center of the diffraction pattern corresponds to a tighter interplanar spacing). X-rays interact with the electrons of the molecules in the crystal, and the intensity of each diffracted beam depends on their distribution in the unit cell.

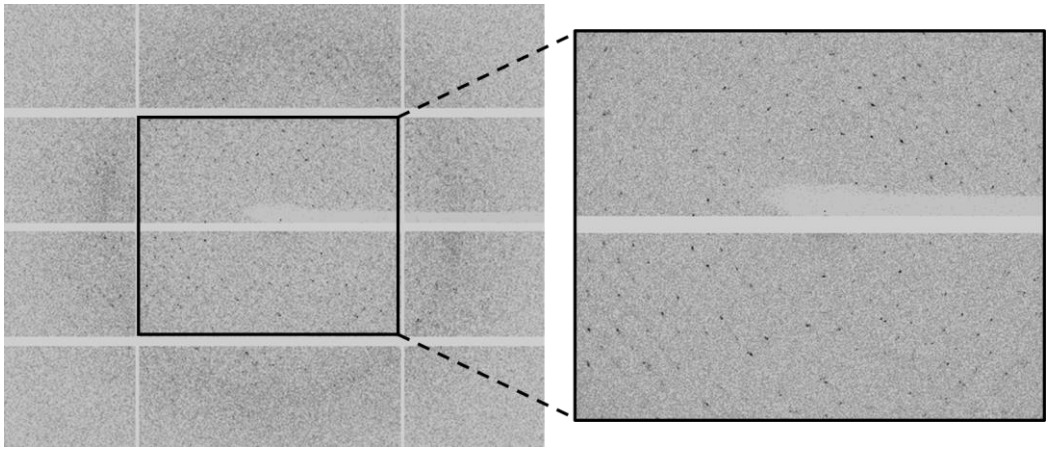

Figure 11 – Diffraction pattern image obtain by diffraction of an *As*P2Ox wild-type crystal in ESRF (European Synchrotron radiation facility). The crystal had around 100 µm x 100 µm x 100 µm and the crystallization condition used was: 0.1 M TRIS-HCl pH=8.5 and 2 M ammonium sulfate at 20°C. A cryoprotectant solution (20% glycerol) was used to transport and diffract the crystal under cryogenic conditions.

1.6.3 3D structure determination

Each diffraction spot (Figure 11) corresponds to a set of diffracted waves, characterized by an amplitude and a phase [54], [56], [59]. Every diffracted wave can be represented by a vector named structure factor (**F**) and includes the scattering contributions from all the n atoms in the unit cells to that specific reflection (direction of the diffracted wave).

$$F_{(hkl)} = \sum_{j=1}^{n} f_j . \mathrm{e}^{2\pi \mathrm{i}[\mathrm{hx+ky+lz}]}$$

where hkl is a reflection for which every j atom contributes a partial wave of atomic scattering factor f_j for the complex structure factor $F_{(hkl)}$. The atomic scattering factor is independent of the position of the atom in the unit cell, only depending on the type of atom and the direction of scattering.

Although each reflection is characterized by the amplitude |F| and phase angle α of its structure factor, only amplitudes (proportional to the square root of the intensity of the diffracted beam) can be measured experimentally [54], [59]. This constitutes the phase problem.

Phases are crucial for the construction of structural information, and the contribution of the phase information is predominant relative to structure factor amplitudes in the calculation of the electron density map. The inverse Fourier transform of the structure factors is the electron density:

$$\rho(xyz) = \frac{1}{V} \sum_{h,k,l=-\infty}^{+\infty} F(hkl) e^{-2\pi i(hx+ky+lz-\alpha(\mathrm{hkl}))}$$

where ρ is the electronic density at coordinates “xyz”, *F* is the structure factor amplitude for the “hkl” plane in the reciprocal space, α is the corresponding phase angle and V represents the unit cell volume. Thus, for macromolecular structure determination it is crucial to obtain accurate phases and there are several methods by which the phases can be determined, requiring different diffraction experiments.

Molecular replacement is applied when a homologous model is available and uses the Patterson function (a Fourier transform with reflection intensities as coefficients and all phase angles set to zero) where the peaks correspond to interatomic vectors. One can use these interatomic vectors to determine the orientation and position of the homologous model in the unit cell of the target crystal, using rotational and translational parameters. Therefore, it is possible calculate initial phases and consequently calculate an electron density map using those phases and the measured amplitudes.

The single and multiple isomorphous replacement methods (SIR and MIR) use native and heavy-atom derivatized crystals by soaking or co-crystallizing native crystals in solutions containing heavy-atom compounds. A heavy atom (e.g., Pt or Hg) is an atom with an atomic number Z much higher than that of the elements normally present in proteins (H, C, N, O, S). The heavy atoms create small intensity changes that can be used to deduce their positions using direct or Patterson methods [56], [68] and in turn use this information to obtain initial estimates of the phases for the structure factor amplitudes of the native crystal.

While the isomorphous replacement method works best with at least two different heavy-atom derivatives (i.e., atoms bound to different protein sites) the anomalous scattering only needs one crystal. This technique uses crystals with anomalously scattering elements like selenium in the form of selenomethionine. When crystals containing these elements are irradiated by specific X-ray wavelengths at or slightly above an elemental absorption edge, their “normal” X-ray scattering acquires components that are independent of the scattering angle θ. These “anomalous differences” can be used to locate the anomalous scattering atoms in the crystal structure and from that information, phases for the structure factor amplitudes of the crystal can be estimated. Synchrotron radiation can be used to measure diffraction data at these specific wavelengths. The Single-wavelength anomalous dispersion method (SAD) relies on measurements at a single wavelength while the Multiple-wavelength anomalous dispersion method (MAD) collects data at several wavelengths, usually at the absorption peak, at the point of inflection of the absorption curve and at a wavelength far away from the absorption edge [56], [68], [69].

Before model building, it is extremely important to assess diffraction data quality, for which statistical evaluation parameters like resolution, R_{merge}, completeness and signal-to-noise ratio (I/σ) are normally used. To produce an accurate model, crystallographic refinement and model rebuilding is carried out in an iterative manner to reach the best possible agreement between the model and the observed experimental data. An important consideration when refining protein crystal structures is that the number of observations (that primarily depends on the data resolution) is seldom enough to allow a fully unrestrained refinement. Instead, the known amino acid stereochemistry (bond lengths, bond angles, planar groups and chiral centers) is included in

the refinement and treated as additional “observations”, thus introducing restraints in the model, by limiting the variations of the structural parameters and leading to a structural model with physical and chemical significance.

The initial electron density map is used to obtain an approximate structural model that needs to be refined (i.e., its structural parameters need to be optimized), by alternating rounds of automated and manual optimization. Software like PHENIX [70] allows a wide variety of refinement options for automated optimization. Between each refinement round and to decide whether the model is of enough quality to be considered final, other statistical evaluation methods are used. Reciprocal space refinement aims at optimizing the agreement between the calculated structure factors (F_{calc}) and the experimental ones (F_{obs}) which is typically followed by monitoring the R-factor:

$$R = \frac{\sum_{hkl}|Fobs(hkl)| - |Fcalc(hkl)|}{\sum_{hkl}|Fobs(hkl)|}$$

To avoid over-refinement (i.e., using too many parameters) , an R_{free} is similarly calculated, using a small test set of reflection data that is omitted from the refinement process. Other quality assessment statistics like R_{sleep}, root-mean-square deviations (r.m.s.d.s) of the model stereochemistry w.r.t. dictionary values, Ramachandran plots and thermal motion parameters (or B-factors) are also used to determine the structure quality [54], [56]. Many validation procedures are already integrated in PHENIX as well as in model-building software like COOT [71].

1.7 Context of the present project

This master thesis arises as a complementary work of an ongoing project that aims to optimize *As*P2Ox for 3rd generation biosensors through protein engineering, in specific rational and semi-rational design. Since this type of enzyme engineering requires the protein structure, the main goa of the present work wa the determination of the crystallographic structures of wild-type *As*P2Ox, both with and without substrate (D-glucose). On top of that, and as a result of a detailed structural analysis, we were able to propose target residues that could be mutagenized, to create and test new variants.

When I started my work with the Met Lab group, *As*P2Ox wild-type crystals had already been obtained in optimal conditions. I was able to reproduce those conditions to get co-crystalized and soaked crystals with glucose and structures with and without the substrate were obtained for the first time for a bacterial P2Ox. However, seeding had to be used to obtain new crystals when crystals were not growing in the previously obtained conditions. A structural analysis was performed, and some important aspects of the model were identified, one of them being the difference in FAD binding between fungal P2Oxs and *As*P2Ox. To understand in a more detailed

manner the FAD attachment, mutants were designed using rational (site-directed mutagenesis) and semi-rational approaches (saturation mutagenesis).

The work performed resulted in the first bacterial P2Ox structure available and provided a better understanding of the main structural features of *As*P2Ox, namely the FAD non-covalent binding, the active site and its accessibility, and where this enzyme differs from its fungal counterparts.

2. Material and Methods

2.1 *As*P2Ox crystallization

*As*P2Ox wild-type was produced and purified according to protocols previously established in the MET Lab [24]. The initial, previously determined, crystallization conditions for wild-type *As*P2Ox consisted of 2 M ammonium sulfate and 0.1 M Tris-HCl, pH 8.5. To improve the crystallization conditions, a microliter scale optimization was performed using the hanging drop vapor-diffusion method in XRL 24-well crystallization plates (Molecular Dimensions, Newmarket, UK). In these experiments several parameters were tested, namely ammonium sulfate concentration (0.5 - 2.0 M) and the pH of Tris-HCl buffer (7.0 - 8.5). Additionally, several ratios of protein:reservoir solution were tested, 1:2, 1:1 and 2:1, and each drop was equilibrated against 500 µl of reservoir solution. All plates were inspected regularly, every day for the first week, and weekly or monthly after that. *As*P2Ox crystals with typical dimensions of 100 µm x 100 µm x 100 µm appeared within one week at 20°C in hanging drops obtained by mixing 1 µL of protein solution (18 mg/mL) with 1 µL or 2 µL of reservoir solution. The best crystals appeared in 2 M ammonium sulfate and 0.1 M Tris-HCl pH 8.5 using a protein:reservoir solution ratio of 1:2. To minimize the risk of crystal damage under X-rays, the *As*P2Ox wild-type crystals were cryo-protected by a quick soak in a solution of their reservoir solution supplemented with 20% (v/v) glycerol prior to flash-cooling in liquid nitrogen.

2.2 Crystal seeding

Crystal seeding was performed since *As*P2Ox crystals were no longer appearing in the previously described conditions. Preparation of the seed stock was carried out by adaptation of a Hampton Research protocol [72]. A crystal from a previous crystallization was transferred to a 2 µL drop of crystallization solution with a loop. This loop was used to crush the crystal before adding more 8 µL of crystallization solution to the drop. A total volume of 10 µL, with the crushed crystal, was transferred to an Eppendorf containing a Seed Bead. This process was repeated for a total of five crystals (final volume of 50 µL of crystallization solution with seeds). The Eppendorf was vortexed for 3 min, stopping every 30 seconds to cool the tube on ice. The dilutions 1:100, 1:1000 and 1:10000 were prepared from the final seed stock and used to test different seeding conditions. Hanging drop vapor-diffusion method in XRL 24-well crystallization plates (Molecular Dimensions, Newmarket, UK) was used. Each drop had a total volume of 3 µL (protein:reservoir solution ratio of 1:2), with 2 different combinations: 1 µL of protein, 1 µL of seed solution and 1 µL of reservoir solution; or 1 µL of protein, 0.5 µL of seed solution and 1.5 µL of reservoir solution. The seed stock and dilutions 1:1000 and 1:10000 were tested for both combinations of volumes. Different batches of protein were also tested. The reservoir solution used was the same used for the crystal growth, 2 M ammonium sulfate and 0.1 M Tris-HCl, pH 8.5. The best seeding condition was replicated in a different plate – 1 µL of protein, 1 µL of seed solution (1:1000) and 1 µL of reservoir solution – where crystals appeared within 1 week.

2.3 *As*P2Ox-Glucose complex crystallization

To obtain the crystal structure of *As*P2Ox:D-glucose complex, co-crystallization and soaking experiments using a wide range of glucose concentrations were performed.

The co-crystallization experiment was carried out by mixing 1 µL of protein solution and 2 µL of crystallization solution (2 M ammonium sulfate, 0.1 M Tris-HCl, pH 8.5), with different concentrations of glucose (1.6 mM, 3.3 mM, and 4.9 mM: 5, 10 and 15 times the concentration of protein, respectively), and using the hanging drop vapor-diffusion method in XRL 24-well crystallization plates (Molecular Dimensions, Newmarket, UK). Crystals appeared after one week with dimensions of 100 µm x 100 µm x 100 µm.

For the soaking procedures, *As*P2Ox crystals were transferred into a 2 µL crystallization drop, containing the reservoir solution and different glucose concentrations (300 mM, 600 mM, 800 mM, and 2 M). Different soaking times were tested for each glucose concentration – 5 min, 10 min, 30 min, 60 min, 2 h, and 24 h. The crystals were harvested with a cryo-loop and cryo-protected as described above. As mentioned before, crystals of wild-type *As*P2Ox were no longer being obtained from the scale-up experiments and thus crystals were obtained through seeding. These latter crystals were also transferred into 2 µL crystallization drops, containing the reservoir solution and different glucose concentrations (10 mM, 100 mM, 500 mM, 1 M and 2 M). For these crystals, quick soaking conditions were used. The crystals were soaked only for a couple of seconds and immediately frozen in liquid nitrogen, after cryo-protection.

2.4 Data collection and processing

Diffraction data was collected at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) and ALBA (Barcelona, Spain). At ESRF, a total of 5 datasets out of 7 *As*P2Ox crystals were measured. Moreover, 4 datasets from *As*P2Ox crystals co-crystallized with glucose were collected. At ALBA, a total of 12 glucose-soaked crystals were tested and 10 datasets were obtained. Diffraction spots were indexed, integrated, scaled and the final amplitudes calculated using XDS [73]. Additionally, the raw data were also indexed and integrated with XDS, space group determined with POINTLESS [74] and intensities scaled and merged using AIMLESS [75][76], all within the autoPROC [77] data-processing pipeline.

2.5 *As*P2Ox and *As*P2Ox-Glucose complex structure determination and refinement

The crystal structure of wild-type *As*P2Ox was determined using a 2.01 Å resolution dataset collected at the ESRF. The phase problem was solved by molecular replacement with MORDA [78], using the coordinates of *Pc*P2Ox (PDB 4MIF) [13] and ATP phosphoribosyltransferase regulatory subunit structure (PDB 3OD1) [79] as search models. The residue sequence was fitted into the initial electronic density map using PHENIX.AUTOBUILD [70], allowing completion of the structure through additional cycles of manual building and refinement. Iterative refinement cycles were performed with PHENIX.REFINE [80] followed by manual model building with COOT [71]. A total of 1.57% of the reflections were randomly excluded for cross-validation with R_{free}. The TLSMD server (http://skuld.bmsc.washington.edu/~tlsmd) [81] was used to define polypeptide chain regions to be treated as rigid bodies for translation, libration and screw motions (TLS) refinement of anisotropic atomic displacement parameters (ADPs). The protein structure was examined against σ_A-weighted electron density maps leading to improvement of the model. The *As*P2Ox crystal structure was inspected using COOT, and some regions of the structure were not visible in the electron density maps. Therefore, a BUSTER protocol [82], was employed to search for missing atoms.

The *As*P2Ox-Glucose complex crystal structure was determined using a 2.35 Å resolution dataset collected at the ESRF. The phases for this structure were obtained from the wild-type *As*P2Ox structure using a rigid-body refinement with PHENIX.REFINE [80]. For the refinement process, 1.36% of the reflections were randomly excluded for cross-validation with R_{free}. The TLSMD web server (http://skuld.bmsc.washington.edu/~tlsmd) [81] was once again used to define polypeptide chain regions as rigid bodies for TLS refinement of anisotropic ADPs. Iterative refinement cycles were performed with PHENIX.REFINE [80] followed by manual model building with COOT [71].

For both models, refinement included optimization of atomic positions and ADPs, refinement of TLS parameters, and automatic water solvent completion. Water solvent molecules were initially located automatically in difference Fourier maps, at peaks with hydrogen bonding distances within 2.45-3.40 Å. Different types of solvent molecules were also identified by matching their shapes against those of electron density peaks in σ_A difference maps, by comparing their refined ADPs with those of neighbouring atoms and by analysing their putative hydrogen bonding distances. The refinement was performed until convergence of R_{work} and R_{free} values. The validation of the stereochemistry of the polypeptide chains was performed with MOLPROBITY [83]. All figures were prepared with PyMOL [84].

2.6 Construction of site-directed mutagenesis mutants

Four single site variants were constructed based on rational design – N120V, V54K, S70Q and G96Q. All variants were, previously to any experimental work, modelled with SwissModell [85]: All variants were initially modelled with a second mutation – H127A – to avoid structural bias. In a second modelling step, His127 was re-introduced, and the first model (with the double mutation) was used as the template.

Single substitutions in the *As*P2Ox gene were created using the QuickChange Site-Directed Mutagenesis protocol (Stratagene). Primers (Table 2) were designed using SnapGene software (from Insightful Science; available at snapgene.com) to amplify the plasmid pSM-1, a pET-15b (Novagen) carrying the gene coding for wild-type *As*P2Ox to allow the change of the desired codon to mutate. Vector pSM-1 grants resistance to ampicillin through β-lactamase production, has a T7 promoter and can be induced by Isopropyl β-D-1-thiogalactopyranoside (IPTG). Moreover, the gene was cloned fused to a His-tag coding sequence (6 histidine residues), allowing to purify the enzyme using affinity chromatography.

Table 2 – Forward (FW) and reverse (RV) primers designed with SnapGene to perform a touchdown PCR for site-directed mutagenesis.

Variant	Primer	Sequence
N120V	*As*P2Ox-FW-N120V	5' CCATGTCCAGC**GTC**GTGGGCGGGATGG 3'
	*As*P2Ox-RV-N120V	5' CCATCCCGCCCAC**GAC**GCTGGACATGG 3'
V54K	*As*P2Ox-FW-V54K	5' CAATCCGCCCGGCGCGCAC**AAA**AAGAACATCGAGGAC 3'
	*As*P2Ox-RV-V54K	5' GTCCTCGATGTTCTT**TTT**GTGCGCGCCGGGCGGATTG 3'
S70Q	*As*P2Ox-FW-S70Q	5' CCAGCGGGCG**CAG**GAGGGTCCCGGTG 3'
	*As*P2Ox-RV-S70Q	5' CACCGGGACCCTC**CTG**CGCCCGCTGG 3'
G96Q	*As*P2Ox-FW-G96Q	5' CGTGCGCGCCCT**CAA**ACTTACCTGCTGCAG 3'
	*As*P2Ox-RV-G96Q	5' CTGCAGCAGGTAAGT**TTG**AGGGCGCGCAC 3'

PCRs were performed in a thermal cycler (MyCycler Thermal Cycler, Biorad). For each of the four variants, a touchdown PCR [86], was performed in reactional mixtures containing 100 ng of DNA template, 1 µM of each primer (FW and RV), 200 µM of dNTPs, NZYproof polymerase buffer and 1.3 U of NZYproof polymerase (NZYTech). The PCR program included an initial denaturation of 3 min at 95°C followed by 15 cycles of 30 sec at 95°C, 45 sec at 75-60°C (1°C decrease per cycle) and 8 min at 72°C. Then a second round of 15 cycles were performed: 30 sec at 95°C, 45 sec at 65°C and 8 min at 72°C; a final elongation step was performed at the end for 10 min at 72°C. The PCR product was analyzed in a 1% agarose gel and digested for 6 h with DpnI (ThermoFisher) to eliminate the DNA template. Then the PCR products were dialysed for 20 min with a 0.025 µm MCE membrane MF-Millipore for removal of unwanted salts used in the PCR mix.

Transformation was performed by mixing 5 µL of dialysed PCR product with 125 µL of electrocompetent DH5α cells, previously thawed on ice. The cell suspension was placed in the

electroporation cuvette (0.2 cm gap) and a pulse was applied in an Xcell ShockPod chamber (Gene Pulser XcellTM, Biorad), using the set conditions of C = 25 µF, PC = 200 Ω and V = 2.5 kV. 1 mL of LB media was added immediately after the pulse, and the cell suspension was incubated for 1 h at 37°C, 180 rpm. After incubation, the cells were spread in Luria Bertani (LB) agar plates supplemented with 100 µg/ml of ampicillin, in an appropriate dilution to obtain single colonies. The plates were incubated at 37°C overnight.

In the day after, single colonies were picked for 50 mL falcon tubes containing 10 ml of LB media supplemented with 100 µg/mL of ampicillin and the growth proceeded overnight, at 37°C at 180 rpm in an INNOVA shaker. The resulting cells were used to extract plasmid using the GeneJet kit (Thermofisher) and following the manufacturer protocol. The obtained plasmids were sent to sequencing (Stabvida) to confirm the presence of the desired mutation.

2.7 Construction of site-saturation library and screening protocol

Saturation mutagenesis was performed for the residue N120 using 4 different genetic backgrounds: wild-type and the variants V54K, S70Q and G96Q, generating 4 distinct libraries. The primers used were *As*P2Ox_FW_N120Sat: 5' GGCCATGTCCAGC**NNS**GTGGGCGG GATGGC 3' and *As*P2Ox_RV_N120Sat: 5' GCCATCCCGCCCAC**SNN**GCTGGACATGGCC 3' (where N represents any possible nucleotide and S represents G and C nucleotides). The PCR conditions were, in a total volume of 50 µL: 100 ng of DNA template (WT, V54K, S70Q or G96Q), 0.25 µM of each primer (*As*P2Ox_FW_N120Sat and *As*P2Ox_RV_N120Sat), 500 µM of dNTPs, NZYproof polymerase buffer and 1.3 U of NZYproof polymerase (NZYTech). The PCR program included an initial denaturation of 3 min at 95°C, followed by 30 cycles of 1min at 95°C, 1.5 min at 72°C, 8 min at 72°C and a final elongation step of 10 min at 72°C. The PCR products were digested with DpnI restriction enzyme (ThermoFisher) for 6 h to eliminate any DNA template. After digestion the products were purified and concentrated using the GFX PCR DNA and Gel Band Purification Kits (Merck).

Electrocompetent *Escherichia coli* KRX (Promega) cells were transformed with 5 µL of purified PCR products in an Xcell ShockPod chamber (Gene Pulser XcellTM, Biorad), using the following set conditions: C = 25 µF, PC = 200 Ω and V = 2.5 kV. The cells were resuspended with 1 mL of LB media and incubated for one hour at 37°C prior to the spreading of the cells in LB agar plates containing 100 µg/ml of ampicillin.

A total of 180 colonies were obtained for the saturation of N120 using the wild type as template, 64 for the V54K+N120Sat library, 166 for the S70Q+N120Sat library and 79 for the G96Q+N120Sat library. The colonies were picked to 500 µL of LB media supplemented with 100 µg/ml of ampicillin in 96-deep-well plates and incubated overnight at 37°C, 750 rpm (Heidolph Titramax 1000 plate shaker) for 16 h.

In the following day, 100 µL of the growth were transferred to 900 µL of fresh LB media supplemented with 100 µg/ml of ampicillin in a new 96-deep-well plate. The growth proceeded at

37°C, 750 rpm for 3 h. Then, the gene expression was induced with 12 mM rhamnose in each well and growth continued at room temperature under 750 rpm for 16 h. Cells were collected by centrifugation and used to screen *As*P2Ox's activity.

For the screening processes, the cell pellets were disrupted with 3 cycles of freeze and thaw (a few seconds in liquid nitrogen followed by 10 min at room temperature) and resuspended with 200 µL of lysis buffer containing (per mL): 2 mg of lysozyme, 2 µL of protease inhibitors (1.75 mM antipain and 2.5 mM leupeptin), 1 µL of 1 U/ml DNAse I and 5 µL of 1 M $MgCl_2$, in 20 mM Tris-HCl, pH 7.6. The cell resuspension was incubated for 30 min, 750 rpm at room temperature and then centrifuged for 30 min at 4000 rpm to separate the cell crude extracts from the cell debris.

Enzymatic activity was measured using a reaction coupled assay with horseradish peroxidase (HRP) in which HRP uses hydrogen peroxide produced by *As*P2Ox and two other substrates, 4-aminophenazone (AAP) and 3,5-dichloro-2-hydroxybenzene-sulfonic acid (DCHBS), to produce a pink chromogen, the concentration of which is followed at 515 nm. This enzymatic assay had been previously tested and optimized by André Taborda [24]. The activity of the variants was tested transferring 50 µL of crude extracts to a reactional mixture (200 µL final volume) containing: 0.5 M Glucose, 0.1 mM of 4-aminophenazone (AAP), 1 mM of 3,5-dichloro-2-hydroxybenzene-sulfonic acid (DCHBS), 10 U/mL HRP in 100 mM sodium phosphate pH 7.5. Activity was monitored using a Synergy2 microplate reader (BioTek) at 515 nm for 10 min; additionally, an endpoint activity was measured after 2 h of incubation at 37°C, 750 rpm. Relative activity was calculated by comparing the activity of each well with the activity of wild-type *As*P2Ox. The colour intensity after 2 h of incubation was also used to help picking wells with active variants.

2.8 Larger scale production and purification of *As*P2Ox variants

For biochemical and biophysical analysis of the constructed variants, 2.5 L scale production were performed using *E. coli* Rosetta™(DE3) pLysS (carrying resistance for ampicillin and chloramphenicol) as routine expression strain. The *E. coli* strain was transformed with the desired plasmid as described before and the resulting colonies were picked to perform pre-inocula in 50 mL of LB supplemented with 100 µg/mL of ampicillin and 20 µg/mL of chloramphenicol in a 250 mL Erlenmeyer. The cultures were left overnight at 37°C under 180 rpm (INNOVA Shakers). The pre-inocula were used to inoculate 2.5 L of LB media supplemented with 100 µg/mL of ampicillin and 20 µg/mL of chloramphenicol in Corning® 5 L Baffled PETG Erlenmeyer flasks. The cultures were incubated at 37°C, 100 rpm, and gene expression was induced with 100 µM IPTG when OD_{600nm} reached a value of around 0.8. Growth continued at 25°C, 100 rpm overnight for additional 16-17 h. Cells were then collected by centrifugation at 4,420 g (Avanti J26 centrifuge, JA-10 rotor) for 15 min at 4°C and stored at -20°C.

Prior to purification, cell pellets were resuspended in 20 mM Tris-HCl, pH 7.6, and supplemented (per mL) with 2 µL of protease inhibitor (1.75 mM antipain and 2.5 mM leupeptin), 1 µL of 1 U/mL DNAse I and 5 µL of 1 M $MgCl_2$. French press (Thermo IEC) operating at 1000

psi for 3 cycles was used to disrupt the cells. Cell debris were removed (pellet) after centrifugation at 39,200 g (Avanti J26 centrifuge, JA25.50 rotor) for 2h, 4°C.

An ÄKTA purifier system (GE Healthcare) was used for protein purification. The crude extracts were loaded onto a 5 mL His-Trap HP column (GE Healthcare), previously equilibrated with 20 mM Tris-HCl buffer at pH 7.6, 10 mM imidazole and 0.2 M NaCl. A first washing step was performed with that same buffer and the protein elution was achieved by applying a gradient 0 – 100% of 1 M imidazole in 20 min (1 mL/min of flow rate). The fractions were collected and analysed by SDS-PAGE to check the purity of the preparations. Fractions with purified enzymes were pooled and then a PD-10 desalting column (GE Healthcare) was applied to remove imidazole and change the buffer to 20 mM Tris-HCl, at pH 7.6, supplemented with 0.2 M NaCl.

Quantification of total protein was performed by the Bradford assay and the FAD content of enzyme preparations was quantified by measuring the absorption of the protein at 450 nm in a Nicolet Evolution 300 spectrophotometer (Thermo Industries, Waltham). The molarity of FAD in protein preparations was determined by using ε_{450}=11,300 M^{-1} cm^{-1}. For all analytical procedures such as steady-state kinetics the stock concentration of protein considered was the value determined with Abs_{450nm} which corresponds to the FAD-loaded fraction (holo-enzyme) of the enzyme preparation.

2.9 FAD interaction with *As*P2Ox

To assess FAD binding, spectra of protein preparations were acquired in a Nicolet Evolution 300 UV-Vis spectrophotometer (Thermo Electron Corporation) before and after a denaturation step. 500 µL of purified protein preparations were pipetted to a quartz cuvette with a 1 cm path and a full spectrum from 200 nm to 600 nm was acquired. The proteins were denatured with 10 % Trichloroacetic-acid for 10 min at 100°C, centrifuged at 5000 rpm in a top bench centrifuge and the supernatant was used to acquire a new UV-Vis spectra The comparison of the spectra before and after denaturation allowed to conclude whether the FAD was covalently or non-covalently bonded to the enzyme.

2.10 Release of FAD – thermal stability

To test the effect of the introduced mutations in the FAD cofactor stability, a fluorescence assay was performed by monitoring FAD release during a thermal unfolding assay with a temperature gradient from 20°C to 60°C, with increments of 1°C per minute. This assay takes advantage of the fact that the fluorescence of FAD is quenched by the protein. When the protein starts unfolding, the FAD gets more and more exposed, and its fluorescence intensity increases. A pure enzyme preparation at 0.1 mg/mL in 20 mM Tris-HCl pH 7.6 was excited at 450 nm and FAD emission was recorded at 520 nm (maximum of emission) in a Cary Elipse spectrofluorometer.

It was possible to calculate the equilibrium constant between the holoenzyme and the FAD exposed in solution at each point, using the fluorescence emission at each temperature; $\Delta G°$

(Gibbs free energy at a constant pressure) can then be calculated using $\Delta G° = - RTlnK_{eq}$ and a plot of ΔG° vs T (K) can be obtained, excluding all temperature points corresponding to folded and unfolded baselines. A linear regression of those points provides the values of ΔH° and ΔS° in $\Delta G° = \Delta H° – T\Delta S°$, where H and S represent the enthalpy and entropy, respectively. Knowing that $\Delta G° = 0$ when $K_{eq} = 1$, and since T_m corresponds to the temperature value where the folded and unfolded populations are equal ($K_{eq} = 1$), $T_m = \Delta H°/ \Delta S°$. An adjustment of the experimental points can be performed using the following expression: fraction unfolded = $\exp(-\Delta G°/RT)/(1+\exp(-\Delta G°/RT))$, where ΔG° values are theoretical and calculated using the ΔH° and ΔS° values previously acquired. During this work T_m was considered as the temperature at which half of the enzyme population has lost the FAD cofactor.

2.11 Apparent steady-state kinetics

To test how the catalytic parameters were affected by the introduction of the mutations, enzymatic activity of purified *As*P2Ox variants was monitored with a Synergy2 microplate reader (BioTek) using a coupled assay with HRP that allows to indirectly follow the release of hydrogen peroxide through the formation of a pink chromogen at 515 nm (ε_{515} = 26 000 M^{-1} cm^{-1}), as described before. All enzymatic assays were performed using at least three independent assays. The reaction mixtures of the coupled assays contained 0.1 mM AAP, 1 mM DCHBS, 10 U/mL HRP and different concentrations (0 – 1.25 M) of D-Glucose. The reaction buffer was 100 mM sodium phosphate, pH 7.5. The reaction was started with the addition of the enzyme and all steady state kinetics measurements were performed at 37°C. Apparent steady-state kinetic parameters (k_{cat} and K_m) were determined by fitting data directly into the Michaelis-Menten equation using Origin-Lab software.

3. Results and Discussion

3.1 Protein crystallization

Yellow *As*P2Ox crystals, with dimensions of 100 µm x 100 µm x 100 µm, appeared within one week at 20°C in the crystallization condition 2 M ammonium sulfate in 0.1 M Tris-HCl, pH 8.5, (Figure 12a). Crystals resulting from co-crystallization and soaking with glucose displayed a similar morphology (Figure 12b).

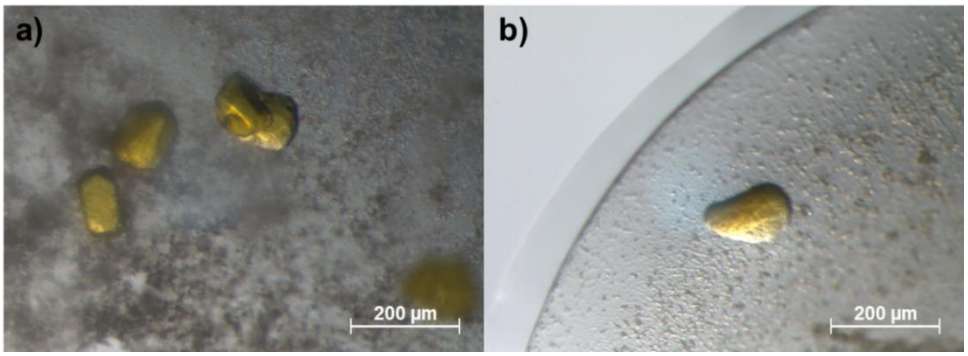

Figure 12 – *As*P2Ox crystals obtained by hanging drop in the crystallization condition 2 M ammonium sulfate in 0.1 M Tris-HCl pH 8.5. The drops contained a) 1 µL of protein solution (18 mg/mL) and 2 µL of crystallization solution; b) 1 µL of protein solution (18 mg/mL) and 2 µL of crystallization solution with 1.6. mM Glucose.

As mentioned in the "Materials and Methods" section, and due to inability to obtain additional crystals using the vapor diffusion conditions previously used, a seeding protocol was followed to obtain new crystals from formerly obtained *As*P2Ox crystals. Different ratios of protein, reservoir and seed solution were tested, as well as different seed concentrations and protein batches. A schematic representation of all seeding conditions tested, and its results can be found in Table 3. The best seeding condition contained 1 µL of protein solution,1 µL of reservoir protein and 1 µL of seed solution 1:1000.

Table 3 – Schematic representation of the different seeding conditions tested. Different dilutions from the seed stock were tested (1:1, 1:1000, 1:10000). In orange are represented the conditions where no crystals were found. In light blue are represented the conditions where at least one crystal grew, while in light green we can find the conditions where 3 or more crystals were found. The best condition is represented in a darker green. Batch 1 was an older batch while batch 2 had been recently purified; both had around 18 mg/mL of protein. Reservoir solution: 2 M ammonium sulfate in 0.1 M Tris-HCl, pH 8.5.

	Batch 1			**Batch 2**		
1 µL protein 1.5 µL reservoir sol. 0.5 µL seeds	**1 : 1**	**1 : 1000**	**1 : 10000**	**1 : 1**	**1 : 1000**	**1 : 10000**
1 µL protein 1 µL reservoir sol. 1 µL seeds	**1 : 1**	**1 : 1000**	**1 : 10000**	**1 : 1**	**1 : 1000**	**1 : 10000**

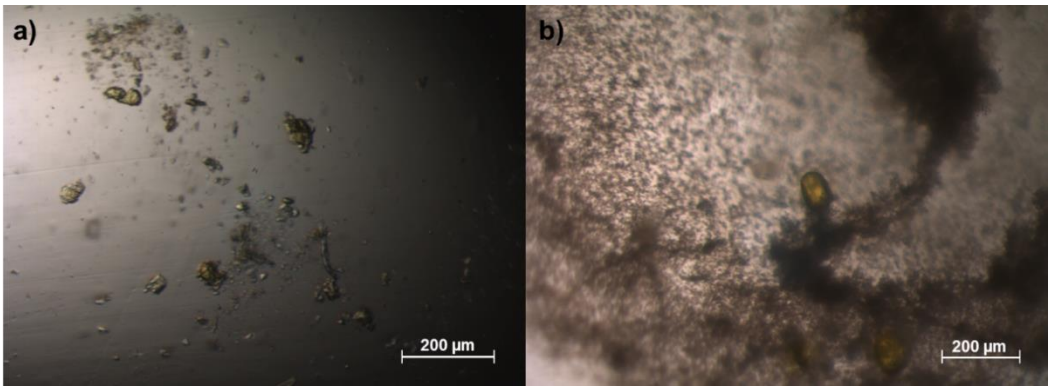

Figure 13 – *As*P2Ox crystal seeding. a) Crystal seeds obtained from the seeding procedure using previously obtained crystals (Figure 12.a). **b)** Crystals obtained by hanging drop in the crystallization condition 2 M ammonium sulfate in 0.1 M Tris-HCl, pH 8.5; 1 µL of protein solution (18 mg/mL), 1 µL of crystallization solution and 1 µL of seeds (1:1000).

A total of 7 crystals from *As*P2Ox wild type were tested at the ESRF and 5 datasets were obtained from those crystals. A total of 5 *As*P2Ox-Glucose crystals resulted from co-crystallization experiments (3 x 1.6 mM and 2 x 3.3 mM of glucose) were also tested at the ESRF, providing a total of 4 datasets. Since no *As*P2Ox-Glucose complex structure was obtained using the first co-crystalized crystals, 12 *As*P2Ox-Glucose crystals obtained by soaking with 0.3 and 0.6 M glucose at different incubation times were tested at ALBA, from which 10 datasets were collected. Additionally, 5 other *As*P2Ox crystals soaked with 0.01, 0.1, 0.5, 0.8, 1 and 2 M glucose for different time periods were tested at the ESRF, and 5 datasets were obtained. Finally, 10 more crystals obtained from seeding experiments and fast soaking conditions (soaking for a couple of seconds, in contrast to minutes or hours as previously performed) in 0.01, 0.1, 0.5 and 1 M of glucose were sent to the ESRF, but bad quality diffraction spots were found, probably due to internal disorder of the crystals, and no datasets were acquired.

From the 24 initial datasets obtained at the synchrotrons, some had poor processing statistics (mainly low completeness and resolution), as a result of many frames with missing diffraction spots due to unsuccessful centring of the crystal and/or bad quality crystals. A total of 15 datasets were fully processed, and their processing statistics can be found in table 4, in the next section.

3.2 Data processing

A total of 10 and 5 datasets were collected at ESRF and ALBA, respectively, with good quality diffraction statistics. These datasets were processed using XDS and autoPROC and the latter one allowed to reach the highest possible resolution values for almost every single dataset. Table 4 describes a summary of the data processing statistics.

For the *As*P2Ox crystals, dataset 2 (Table 4 in blue) had the highest resolution (2.01 Å). This dataset was used for structure determination of wild-type *As*P2Ox, as will be described later. Relatively to the datasets from co-crystalized and soaked *As*P2Ox-Glucose crystals, two of them can be highlighted for their high resolution, dataset 4 and 6 with resolutions of 2.10 Å and 1.91 Å, respectively. However, it was dataset 12 (Figure 4 in green) that allowed determination of

*As*P2Ox-Glucose complex crystal structure showing the substrate in its active site with a 2.35 Å resolution. This crystal resulted from 30 min of soaking in crystallization solution with 2 M of glucose.

Table 4 – Data processing statistics from *As*P2Ox crystals using XDS and autoPROC. Highest resolution and total resolution range statistics are presented. Only datasets that were able to be processed by at least one of the two processing sofware are presented. NP (not processed) represent datasets that could not be processed. The processed datasets used to determine the structures with and without glucose are highlighted in green and blue, respectively.

	Dataset	Glucose		Resolution (Å)	Completeness (%)	I/σ	R-meas (%)	CC(1/2)
XDS	1 (ESRF)	**0 mM**	Highest res	**4.03 – 3.97**	99.8	1.0	106.0	46.1
			Total	37.30 – 3.97	95.6	0.7	122.6	57.8
	2 (ESRF)	**0 mM**	Highest res	**2.63 – 2.62**	79.0	1.7	81.4	85.0
			Total	39.68 – 2.62	96.5	6.0	32.9	97.1
	3 (ESRF)	**0 mM**	Highest res	**3.85 – 3.77**	99.9	0.7	126.7	27.2
			Total	58.04 – 3.77	72.7	0.8	114.3	45.3
	4 (ESRF)	**Co-crys. 1.6 mM**	Highest res	**3.12 – 3.10**	98.8	1.0	97.7	46.3
			Total	56.23 – 3.10	92.7	2.0	77.0	75.1
	5 (ESRF)	**Co-crys. 3.3 mM**	Highest res	**2.78 – 2.76**	99.9	1.5	134.7	79.6
			Total	56.22 – 2.76	98.0	7.7	23.6	99.4
	6 (ALBA)	**10min soak. 300 mM**	Highest res	**2.35 – 2.34**	98.9	1.5	59.5	84.8
			Total	44.66 – 2.34	83.0	5.6	17.7	98.7
	7 (ALBA)	**30min soak. 300 mM**	Highest res	**2.62 – 2.61**	84.4	1.1	45.7	87.1
			Total	44.50 – 2.61	95.5	6.1	17.0	98.4
	8 (ALBA)	**10min soak. 600 mM**	Highest res	**3.04 – 3.03**	67.8	2.7	24.1	94.2
			Total	75.80 – 3.03	80.3	7.9	15.1	98.2
	9 (ALBA)	**30min soak. 600 mM**	Highest res	**2.52 – 2.51**	90.9	1.5	51.9	78.6
			Total	55.40 – 2.51	88.0	4.8	28.4	95.4
	10 (ALBA)	**1h soak. 600 mM**	Highest res	**2.83 – 2.82**	54.9	2.2	30.4	89.2
			Total	55.46 – 2.82	74.7	6.9	16.2	97.9
	11 (ESRF)	**5min soak. 2 M**	Highest res	**2.33 – 2.32**	71.8	0.7	111.2	53.8
			Total	71.32 – 2.32	84.3	3.0	35.0	90.8
	12 (ESRF)	**30min soak. 2 M**	Highest res	NP	NP	NP	NP	NP
			Total	NP	NP	NP	NP	NP
	13 (ESRF)	**5min soak. 800 mM**	Highest res	**5.18 – 4.99**	63.0	6.9	16.4	96.0
			Total	75.28 – 4.99	62.9	7.7	13.3	97.0
	14 (ESRF)	**1h soak. 2 M**	Highest res	NP	NP	NP	NP	NP
			Total	NP	NP	NP	NP	NP
	15 (ESRF)	**1h soak. 800 mM**	Highest res	**2.92 – 2.90**	46.4	3.6	30.2	94.4
			Total	74.38 – 2.90	80.3	8.8	12.3	98.9
autoPROC	1 (ESRF)	**0 mM**	Highest res	NP	NP	NP	NP	NP
			Total	NP	NP	NP	NP	NP
	2 (ESRF)	**0 mM**	Highest res	**2.02 – 2.01**	99.6	1.0	188.6	65.3
			Total	39.40 – 2.01	99.8	9.6	17.9	99.6
	3 (ESRF)	**0 mM**	Highest res	**3.76 – 3.73**	99.9	1.2	105.6	45.8
			Total	54.08 – 3.73	84.6	1.6	92.3	66.9
	4 (ESRF)	**Co-crys. 1.6 mM**	Highest res	**2.11 – 2.10**	92.2	0.9	196.9	39.4
			Total	52.76 – 2.10	98.7	5.0	37.4	98.7
	5 (ESRF)	**Co-crys. 3.3 mM**	Highest res	**2.61 – 2.60**	90.8	0.7	242.2	55.0
			Total	55.80 - 2.60	98.2	7.2	23.4	99.1
	6 (ALBA)	**10min soak. 300 mM**	Highest res	**1.92 – 1.91**	96.3	0.8	144.8	55.0
			Total	75.76 – 1.91	99.3	6.5	15.8	99.5
	7 (ALBA)	**30min soak. 300 mM**	Highest res	**2.52 – 2.50**	91.6	1.2	59.6	87.7
			Total	**75.37 – 2.50**	97.8	6.1	17.2	98.7
	8 (ALBA)	**10min soak. 600 mM**	Highest res	**3.26 – 3.24**	43.0	0.8	117.1	26.7
			Total	55.49 – 3.24	71.9	8.3	36.2	89.6
	9 (ALBA)	**30min soak. 600 mM**	Highest res	**2.63 – 2.62**	52.8	1.4	50.7	78.4
			Total	55.65 – 2.62	84.9	4.8	31.7	93.2
	10 (ALBA)	**1h soak. 600 mM**	Highest res	**2.62 – 2.61**	88.5	2.9	40.9	92.7
			Total	55.55 – 2.61	96.6	8.1	99.0	99.0
	11 (ESRF)	**5min soak. 2 M**	Highest res	**2.12 – 2.10**	93.8	1.1	85.4	79.9
			Total	75.24 – 2.10	98.6	5.2	19.9	98.8
	12 (ESRF)	**30min soak. 2 M**	Highest res	**2.37 – 2.35**	81.4	0.6	155.7	44.1
			Total	75.10 – 2.35	96.8	4.1	28.5	97.3
	13 (ESRF)	**5min soak. 800 mM**	Highest res	**3.21 – 3.20**	84.4	1.2	176.7	31.5
			Total	75.04 – 3.20	81.2	4.1	61.5	93.4
	14 (ESRF)	**1h soak. 2 M**	Highest res	**2.39 – 2.38**	88.1	0.8	129.0	63.9
			Total	74.55 – 2.36	96.9	3.9	30.6	96.8
	15 (ESRF)	**1h soak. 800 mM**	Highest res	**2.21 – 2.19**	71.2	1.4	107.8	71.5
			Total	74.19 – 2.19	77.2	6.1	16.8	99.2

3.2.1 *As*P2Ox Wild-Type data quality

*As*P2Ox crystalized in $C222_1$ space group with unit cell dimensions a = 78.45 Å, b = 78.80 Å, c = 151.31 Å and α = β = γ = 90° (dataset 2; Table 4). All other datasets had the same space group and similar cell dimensions. Dataset 2 was processed up to 2.01 Å resolution and a total of 361063 reflections were measured, with 31573 unique reflections after scaling and merging. R_{meas} values of 17.9% for the total resolution range and 188.6% for the highest resolution shell were obtained. Overall, the statistics obtained show a dataset with relatively good quality. High completeness was achieved at the highest resolution range (99.6%), granting coverage of most unique reflections. To sum up, the diffraction data used for structure determination of *As*P2Ox was processed up to 2.01 Å of resolution, with high completeness and good overall statistics, which allowed using this dataset to solve the *As*P2Ox wild-type crystal structure by means of molecular replacement for the phase problem, as will be described later. More detailed data collection and processing data statistics can be found in Table 5.

Table 5 – Data collection and processing statistics from dataset 2, processed with autoPROC. The values in parenthesis are for the highest resolution shell. $R_{sym} = \Sigma_{hkl} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$, $R_{meas} = \Sigma_{hkl} [N/(N(hkl) -1)]^{1/2} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma hkl \Sigma_i I_i(hkl)$ are indicators of the agreement between symmetry related observations and $R_{pim} = \Sigma_{hkl} [1/(N(hkl) -1)]^{1/2} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$ is an indicator of the precision of the final merged and averaged data set., where N(hkl) is the data multiplicity, $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity of multiple observations from symmetry-related reflections [87], [88].

Beamline	ESRF ID23-2
source/wavelength (Å)	0.8731
Space group	$C222_1$
Unit cell dimensions: a, b, c (Å)	78.45, 78.80, 151.31
Resolution range (Å)	39.39 - 2.01 (2.11 - 2.01)
No. Measured reflections	361063 (51690)
No. Unique reflections	31572 (5031)
Completeness (%)	99.8 (99.6)
Mosaicity (°)	0.43
Multiplicity	11.43 (10.27)
$CC^{1/2}$	99.6 (65.3)
R_{meas} (%)	17.9 (188.6)
R_{sym} (%)	15.2 (81.3)
R_{pim} (%)	4.6 (26.4)
$\langle I/\sigma \rangle$	9.57 (0.99)
V_M ($Å^3$ Da^{-1})	1.83
Estimated solvente content (%)	32.67
Wilson B-factor ($Å^2$)	42.02

3.2.2 *As*P2Ox-Glucose complex data quality

The *As*P2Ox-Glucose complex also crystalized in space group $C222_1$ and had identical unit cell dimensions a = 78.27 Å, b = 78.50 Å, c = 151.23 Å and α = β = γ = 90° (dataset 12; Table 4) to the wild type *As*P2Ox crystal structure. Dataset 12 was processed up to 2.35 Å resolution, with a total of 94778 measured reflections and 18899 unique reflections after scaling and merging. R_{meas} values of 28.4% for total resolution range and 154.9% for the highest resolution shell were obtained. A relatively high completeness was achieved at the total resolution range (96.7%) but only 80.8% was reached for the highest resolution range (Table 6). As referred in material and methods, the crystal structure of *As*P2Ox complexed with its substrate, glucose, was obtained using the crystal structure of wild-type *As*P2Ox as template for rigid-body refinement with PHENIX.REFINE.

Table 6 – Data collection and processing statistics from dataset 12, processed with autoPROC. The values in parenthesis are for the highest resolution shell. $R_{sym} = \Sigma_{hkl} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$, $R_{meas} = \Sigma_{hkl} [N/(N(hkl) -1)]^{1/2} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma hkl \Sigma_i I_i(hkl)$ are indicators of the agreement between symmetry related observations and $R_{pim} = \Sigma_{hkl} [1/(N(hkl) -1)]^{1/2} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$ is an indicator of the precision of the final merged and averaged data set., where N(hkl) is the data multiplicity, $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity of multiple observations from symmetry-related reflections [87], [88].

Beamline	ESRF ID30A-1
source/wavelength (Å)	0.9654
Space group	$C222_1$
Unit cell dimensions : a, b, c (Å)	78.27, 78.50, 150.23
Resolutivos range (Å)	75.11 – 2.35 (2.36 – 2.35)
No. Measured reflections	94766 (9314)
No. Unique reflections	18894 (2511)
Completeness (%)	96.2 (80.8)
Mosaicity (°)	0.43
Multiplicity	5.01 (3.71)
$CC^{1/2}$	97.3 (42.4)
R_{meas} (%)	28.4 (154.9)
R_{sym} (%)	18.83 (185.94)
R_{pim} (%)	14.66 (47.35)
$\langle I/\sigma \rangle$	4.11 (0.64)
V_M ($Å^3$ Da^{-1})	1.8
Estimated solvent content (%)	31.76
Wilson B-factor ($Å^2$)	50.88

3.3 Structure determination and refinement

3.3.1 *As*P2Ox Wild-Type

MORDA [78] was used to determine the initial calculated phases by molecular replacement. A preliminary *As*P2Ox model was built with 96% of probability using two different homologues models: *Pc*P2Ox (PDB 4MIF) [13] with 27.6% identity (438 residues), and ATP phosphoribosyltransferase regulatory subunit (PDB 3OD1) [79], with 29.3% identity (41 residues). The structure was solved in the $C222_1$ space group and the asymmetric unit (a.u.) contained one polypeptide chain. Initial R_{work} and R_{free} values of 41.5% and 48.1%, respectively, for the structural model were obtained. Phenix's Autobuild [70] tool was used to enhance the quality of the electron density maps, improving the R_{work} and R_{free} values to 29.7% and 34.3%, respectively. However, only 356 out of 519 residues were present in this initial model, i.e., a considerable number of residues were not visible, probably due to a higher flexibility of those regions. PHENIX [70] and BUSTER [82] were used for the structure refinement, with manual adjustments between each cycle using COOT [71]. A schematic representation of the refinement process can be found in Figure 14.

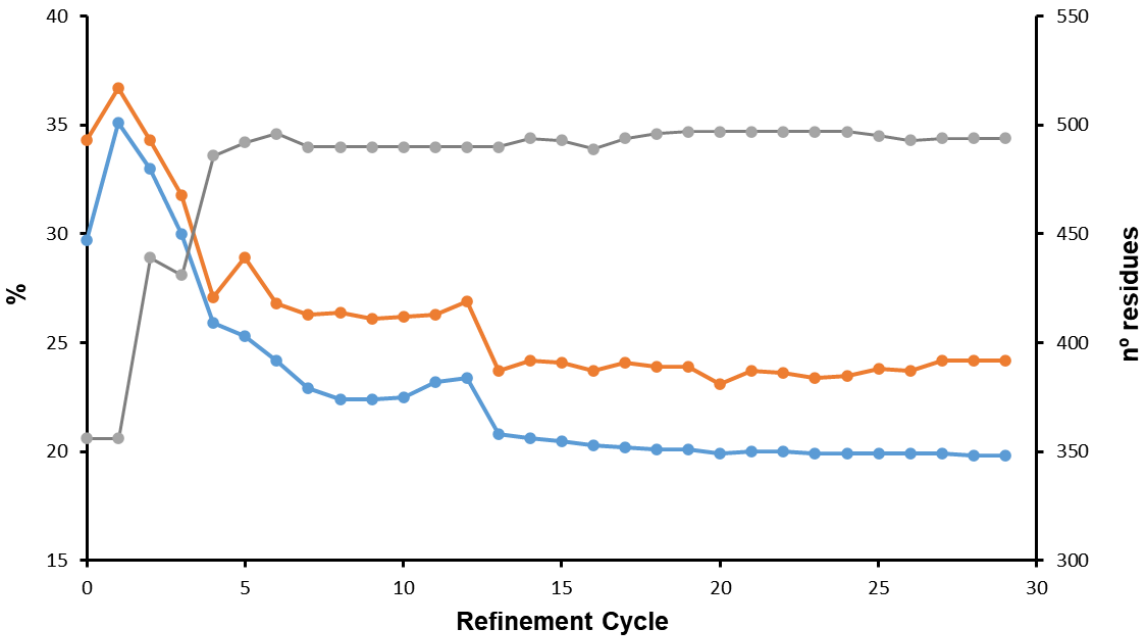

Figure 14 – Schematic representation of the refinement process. In grey, the number of residues per cycle of refinement is represented, while in blue and orange, R_{work} and R_{free} are represented (in %), respectively. Model refinement was done using Phenix between cycles 1 and 13. Buster was implemented from there until the end (cycle 29). TLSMD was used with both Phenix and Buster since cycle 4.

After 4 cycles of refinement with Phenix, a model with 486 residues was achieved with R_{work} and R_{free} of 25.9% and 27.1%, respectively. However, this model contained two regions without electron density, 201-206 and 306-320. TLS groups were introduced in the refinement, defining the anisotropic displacement of the atoms. These TLS groups were modeled individually: a total of 17 rigid bodies were created for *As*P2Ox. Since in the 13th refinement cycle neither the R values nor the electronic density maps of the non-visible regions were improving, the refinement proceeded with BUSTER, which uses a modified maximum-likelihood approach to improve the electron density maps in parts of the structure with missing atoms, such as the model gaps [82]. A final model with 494 residues (5-201, 205-308, 319-511) out of a total of 519 was obtained after 29 cycles of refinement, with R_{work} and R_{free} values of 19.8% and 24.4%, respectively. A total of 232 water molecules were automatically included in the model during structure refinement and were manually inspected. Three sulfate ions and one Tris molecule were modelled. MOLPROBITY was used to assess the model stereochemistry [83] and the final refinement statistics are summarized in Table 7. No Ramachandran outliers were found, with 97.95% residues in favored regions. Four poor rotamers (1.03%) can be found, with 94.83% of favoured rotamers. The statistics presented in table 7 indicate that it was possible to obtain a reliable model of good quality.

Table 7 – Refinement statistics of *As*P2Ox and model stereochemistry analysis.

Resolution range (Å)	39.40 - 2.01 (2.02 - 2.01)	**Poor rotamers**	4 (1.03%)
Rwork (%)	19.8	**Favored rotamers**	367 (94.83%)
Rfree (%)	24.4	**Ramachandran outliers**	0
No. Non-hydrogen atoms	4009	**Ramachandran favored**	478 (97.95%)
No. Protein atoms	3815	**Cβ deviation > 0.25 Å**	0
No. Residues	494	**Rama distribution Z-score**	-0.96 ± 0.33
No. Buffer species:		**Bad angles**	4 (0.10%)
SO_4^{2-}	3	**Bad bonds**	8 (0.15%)
TRIS	1	**Cis proline**	1/38 (2.63%)
No. water molecules	232		
Protein average B value (Å^2)	52.2		
rmsd for bond lengths (Å)	0.003		
rmsd for bond angles (°)	0.542		

3.3.2 *As*P2Ox-Glucose complex

For the structure determination of *As*P2Ox complexed with glucose, a total of 10 datasets were used for rigid-body refinement. Most of these datasets did allow determination of an initial *As*P2Ox structure, but without glucose. However, dataset 12 allowed to obtain a crystal structure of the *As*P2Ox-glucose complex. Like *As*P2Ox structure, the *As*P2Ox-glucose complex structure belongs to the $C222_1$ space group and the a.u. contains a single polypeptide chain. Initial R_{work} and R_{free} values were 48.48% and 52.01%, respectively. A total of 24 cycles of refinement were performed with PHENIX.REFINE. The R_{work} and R_{free} converged to 24.5% and 26.2%, respectively. Indeed, some regions of the enzyme were not visible in the electron density maps, such as the substrate loop which is suggested as being responsible for the entrance and accommodation of the substrate [17], as will be addressed in the "Structural Analysis" section. Due to this lack of electronic density, only 458 residues were included in the model (5-76, 90-200, 207-307, 321-344, 360-509), versus 494/519 of the *As*P2Ox wild-type structure (table 8).

Table 8 – Non-visible residue regions of AsP2Ox and AsP2Ox-Glucose in their respective electronic density maps.

***As*P2Ox**	***As*P2Ox-Glucose**
	77-89
202-204	201-206
309-318	308-320
	345-359

Two sulfate ions and one Tris molecule were modelled. The final model stereochemistry was assessed with MOLPROBITY [83] and refinement statistics are summarized in Table 9. Two Ramachandran outliers were found, with 93.53% residues in favored regions, which is a good value for the average quality of the dataset. Only 2 poor rotamers (0.55%) were found, with an excellent value of 98.34% of favoured rotamer. The average B-factor value of the *As*P2Ox-Glucose complex structure is higher than the *As*P2Ox model (64.89 Å vs 52.19 Å), showing its higher overall flexibility. The overall statistics of the *As*P2Ox-Glucose complex structure are not as good as of that without substrate but allowed a first analysis of the influence of glucose on *As*P2Ox.

Table 9 - Refinement statistics of *As*P2Ox-Glucose complex and model stereochemistry analysis.

Resolution range (Å)	75.10 - 2.35 (2.37 – 2.35)	**Poor rotamers**	2 (0.55%)
Rwork (%)	24.5	**Favored rotamers**	355 (98.34%)
Rfree (%)	26.2	**Ramachandran outliers**	2
No. Non-hydrogen atoms	3571	**Ramachandran favored**	419 (93.53%)
No. Protein atoms	3459	**Cβ deviation > 0.25 Å**	0
No. Residues	458	**Rama distribution Z-score**	-2.96 ± 0.33
No. Buffer atoms:		**Bad angles**	0
SO_4^{2-}	2	**Bad bonds**	0
TRIS	1	**Cis proline**	1/33 (3.03%)
No. Water molecules	29		
No. Glucose molecules	1		
Protein average B value (Å^2)	64.9		
rmsd for bond lengths (Å)	0.002		
rmsd for bond angles (°)	0.465		

3.4 Structural analysis

3.4.1 *As*P2Ox overall structure

The present work describes the first bacterial pyranose 2-oxidase crystal structure, *As*P2Ox. *As*P2Ox wild-type and *As*P2Ox-glucose complex structures diffracted up to 2.01 Å and 2.35 Å resolution, respectively. Both models crystallized in the $C222_1$ space group, with similar unit cell dimensions: a ≈ 78 Å, b ≈ 79 Å, c ≈ 151 Å (α = β = γ = 90°). The two models share the same architecture, with a root mean square deviation (r.m.s.d.) of 0.397 Å between Cα atoms, calculated with PDBeFold web server (Protein structure comparison service (http://www.ebi.ac.uk/msd-srv/ssm, authored by E. Krissinel and K. Henrick) [89]. This value shows that both structures are very similar, despite showing some small differences.

The crystallographic structure of the *As*P2Ox monomer can be divided into two major domains: the substrate binding domain (46-206; 290-443) and the FAD-binding domain with a Rossmann fold (5-45; 210-289; 444-511), containing the FAD cofactor. These domains can be visualized in Figures 16 (structural alignment) and 16 in the next section. While *As*P2Ox wild-type structure has 494 visible residues, the complex model only has 458 out of a total of 519 residues.

*As*P2Ox-Glucose complex structure has four non-visible regions (77-89, 201-206, 308-320 and 345-359) while *As*P2Ox wild-type only has two of those (202-204 and 309-318). This means that the major difference between both models is the absence of 77-89 and 345-359 residues in *As*P2Ox-Glucose structure. Both of these regions can be visualized in Figure 15. The lack of electronic density in certain regions of the protein could be due to the presence of glucose. The fact that the two non-visible regions (345-359 and 77-89) in *As*P2Ox-Glucose are surrounding the active site cavity supports this hypothesis. The occupancy of glucose in the AsP2Ox-substrate complex crystal structure is of 80%? and a change in conformation could occur when glucose is present to allow its entrance and accommodation in the active site. In fact, it was reported that

fungal P2Oxs have a dynamic loop gating the active site and the corresponding residues in *As*P2Ox (^{345}DASPVP^{350}) are included in one of the missing regions (345-359). If we take a closer look at the zoomed image in Figure 15, we can find this loop (in the wild-type model) overlapped with the glucose in the complex structure, emphasizing its dynamic function in the bacterial enzyme. Thus, the *As*P2Ox-glucose complex structure provides insights towards the binding mode of glucose in the active site which is of extreme importance and interest for the catalytic process.

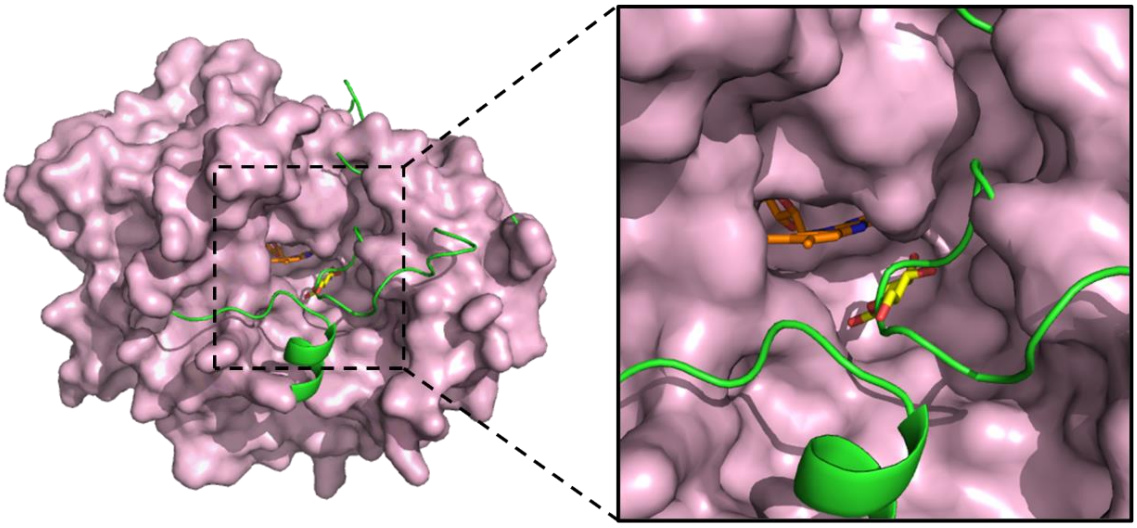

Figure 15 – Surface representation of *As*P2Ox-Glucose complex model in light purple and *As*P2Ox model shown in green as cartoon. The figure (overall and zoomed view) shows the missing regions of the enzyme-substrate complex model. Carbon atoms are represented in orange in FAD and in yellow in glucose.

3.4.2 Overall structure comparison of *As*P2Ox with fungal pyranoses

The DALI [90] web server (http://ekhidna2.biocenter.helsinki.fi/dali/) was used to confirm the structural similarity between *As*P2Ox, the available fungal P2Ox models and other members of the GMC superfamily. However, using PDBeFold web server (Protein structure comparison service – http://www.ebi.ac.uk/msd-srv/ssm - authored by E. Krissinel and K. Henrick) [89] for r.m.s.d. calculation between *As*P2Ox and every fungal model available (Table 10), one might find that *As*P2Ox is some kind of an outlier. Fungal P2Oxs have r.m.s.d. values from 0.26 Å up to 1.01 Å between them, while the values between *As*P2Ox and all fungal P2Oxs are around 1.80 Å. In terms of r.m.s.d. values, *As*P2Ox has significant structural differences when compared to the fungal models, while the latter are more similar among each other.

Table 10 **– Values of rmsd (Å)** between *As*P2Ox, *Pc*P2Ox, *Pe*P2Ox and *Tm*P2Ox, acquired using the PDBeFOLD web server.

	PDB | Structure	1	2	3	4
1	***As*P2Ox**	–	1.878	1.765	1.775
2	**4MIF| *Pc*P2Ox**	1.878	–	0.987	1.013
3	**1TZL| *Pe*P2Ox**	1.765	0.987	–	0.260
4	**1TT0| *Tm*P2Ox**	1.775	1.013	0.260	–

Fungal pyranoses with known crystal structures were reported as homotetramers [5], while *As*P2Ox assembles as a monomer in solution, as found by size-exclusion chromatography [4], and the crystals contain a single monomer in the asymmetric unit (a.u.). The dimensions of the *As*P2Ox monomer are 48.5 Å x 55.4 Å x 71.9 Å, relatively smaller than the homotetramer subunits from the fungal structures (around 50 Å x 55 Å x 90 Å).

A structural alignment (Figure 16) was performed using Modeller to identify the structural differences between *As*P2Ox and its fungal counterparts [91]. *As*P2Ox structure shows the typical substrate binding domain and FAD-binding domain with a Rossmann fold, present in P2Oxs belonging to the GMC superfamily. These domains can be easily visualized in the structural alignment in Figure 16, and in Figure 17, where the substrate domain is represented light orange, and the FAD-binding domain is represented in dark orange. When compared with the fungal P2Oxs (620-623 residues), *As*P2Ox has a considerably shorter residue sequence, 519 residues. The lower number of residues gives rise to several structural differences, the most significant being the absence of the arm and head domains reported in all three P2Ox fungal structures. This major difference can be easily verified in the structural alignment (Figure 16), where the fungal arm domain and the fungal head domain are highlighted by a green and a blue box, respectively. These two missing fungal domains in *As*P2Ox can be seen in Figure 17, where the arm and head domains are highlighted by dashed boxes in the fungal model (*Tm*P2Ox).

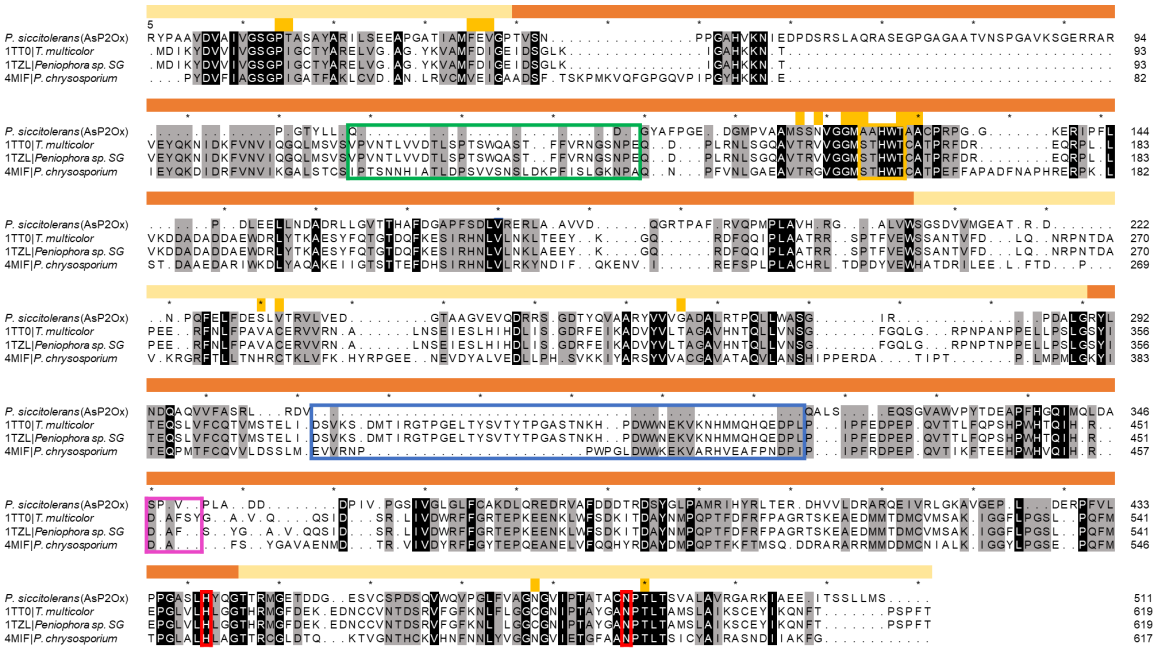

Figure 16 – Structural alignment between *As*P2Ox, *Tm*P2Ox (PDB 1TT0), *Pe*P2Ox (1TZL) and *Pc*P2Ox (PDB 4MIF). Each 10th position is marked by an asterisk (*). Strictly conserved residues are represented in a black background, while mostly conserved residues are represented in a light gray background. The green and blue boxes indicate the fungal arm and the head domain, respectively. The FAD binding motif is highlighted by a yellow box, while the catalytic residues are outlined by a red box. The pink box represents the dynamic substrate loop gating the active site. All residues that interact with the FAD cofactor in the *As*P2Ox structure have a yellow marker above them. The FAD-binding and substrate domains are represented above the sequences in dark and light orange, respectively. A larger version of the structural alignment can be found annexed in the supplementary material.

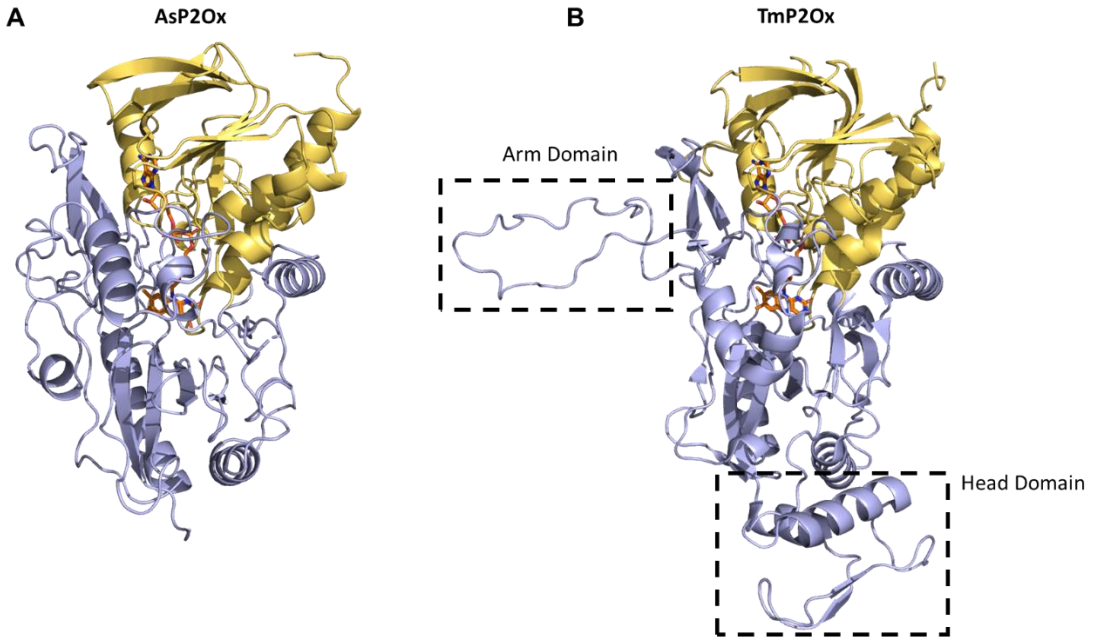

Figure 17 – (A) *As*P2Ox and (B) *Tm*P2Ox (PDB 1TT0) structures as cartoons. The arm and head domains of *Tm*P2Ox are highlighted with a box. The FAD binding domain is represented in yellow, and the substrate binding domain is shown in blue. FAD is shown in stick representation, with carbon atoms coloured in orange.

The arm domain (residues 116 to 143 in *Tm*P2Ox) present in fungal P2Oxs, was proposed to represent a structural determinant of oligomerization, and thus, a factor behind the homotetrameric nature of fungal P2Oxs, by promoting inter-subunit contacts [12] (Figure 18). In accordance to this, the oligomerization loops of *Pc*P2Ox (T51-V72) [13] and *Tm*P2Ox (I80-A109) [12] are partially or totally missing in *As*P2Ox. Therefore, the monomeric form reported for *As*P2Ox [5] is in agreement with the structural findings, as the *As*P2Ox model does not have the oligomerization domains found in its fungal counterparts. The exact role of the head domain (residues 374 to 424 in *Tm*P2Ox) is yet to be determined but it was proposed to be a mediator in the association to polysaccharides from the cell-wall, for immobilization of the enzyme at the fungal periplasmic space, a role that is not expected to occur in the gram-positive bacteria that harbors *As*P2Ox, which has a different cell morphology.

Although most differences between *As*P2Ox and its fungal counterparts are deletion of domains, in *As*P2Ox, region E58-P95 is an insertion not present in any of the fungal models. This insertion is located at the surface of the protein, close to the substrate loop and surrounding part of the fairly exposed substrate cavity, which may hypothetically indicate a catalytic role.

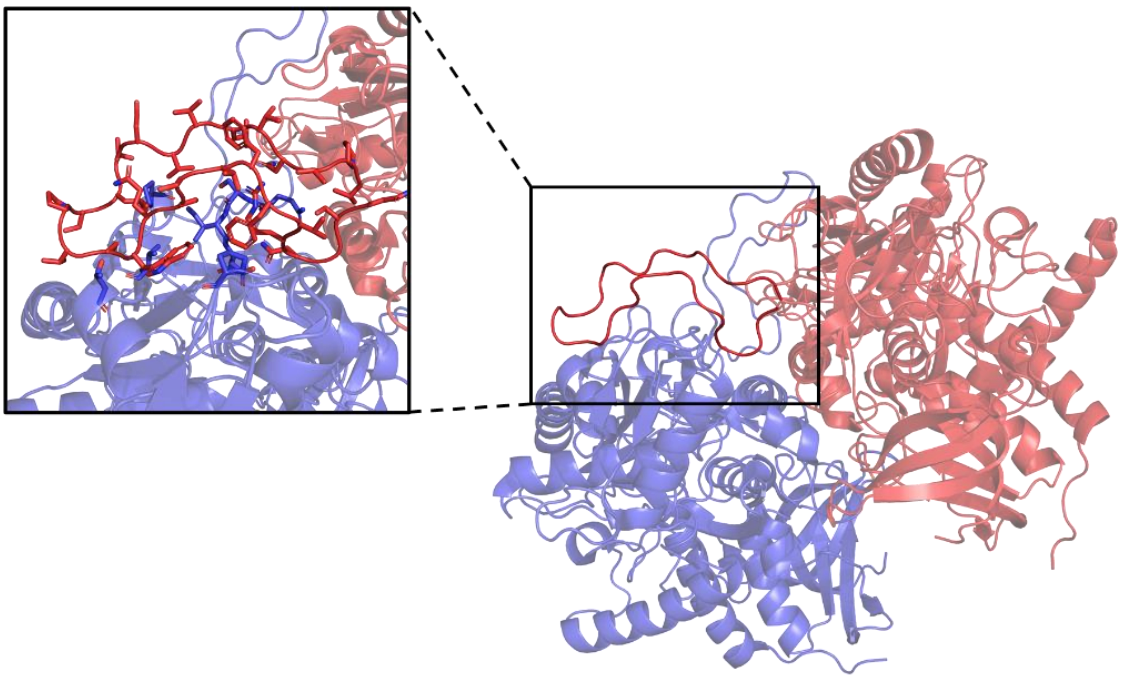

Figure 18 – Representation of the inter-subunit interaction from fungal *Tm*P2Ox (PDB 1TT0). The oligomerization arm of one tetramer subunit (in red) can be found interacting in close proximity with residues of a different subunit (in blue).

3.4.3 Substrate loop and substrate binding site

Structural studies on fungal P2Oxs in the presence and absence of substrate, showed a conserved substrate loop ^{452}DAFSYG^{457} (*Tm*P2Ox numbering, PDB 1TT0) gating the substrate binding site. Due to its flexibility, this loop can present itself in two main positions, open and closed. The loop substrate in the *Tm*P2Ox structure (PDB 1TT0) with acetate (simulating the presence of O_2 in the oxidative half-reaction) has a closed conformation. In models with and without reduced substrate, this loop shifts to an open conformation allowing substrate binding near the isoalloxazine moiety of FAD [12]. In a first step of the catalytic process, the loop is thought to help accommodating the substrate in an open conformation, for further reaction with catalytic residues, i.e. through interactions between e.g. glucose and residues D452, as well as Y456 [92]. Interestingly and in contrast to its fungal counterparts, *As*P2Ox has a different motif, ^{345}DASPVP^{350}, as compared to fungal P2Oxs.

By representing *As*P2Ox as solvent accessible surface and colouring it according to each residue B-factor value (Figure 19), it is possible to observe the B-factor distribution in different regions of *As*P2Ox.

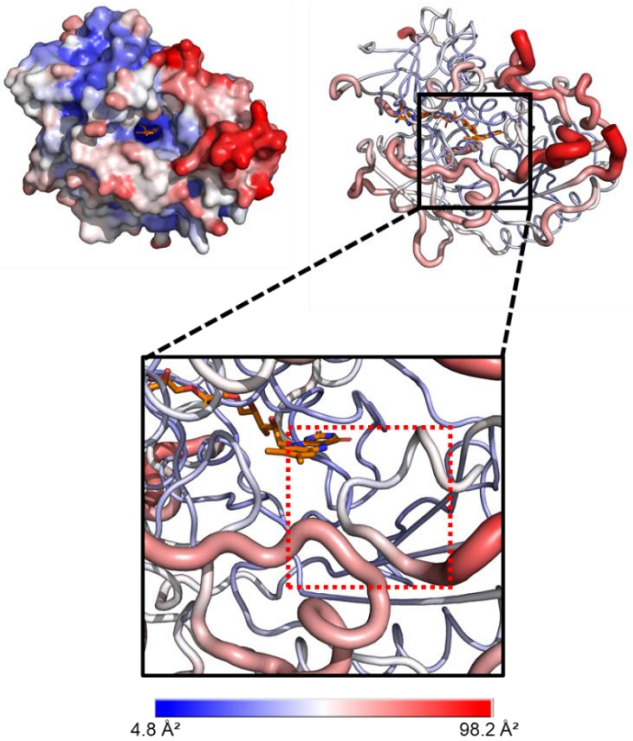

Figure 19 – Left upper image: Representation of B-factor values from *As*P2Ox mapped on the molecular surface. Right upper image: Representation of B-factor values from *As*P2Ox as a cartoon tube. Red (and thicker) regions represent higher B-factor values (high flexibility), and blue (and thinner) regions represent lower B-factor values. FAD is represented as sticks with its carbon atoms shown in orange. The lower image gives us a closer look at the cavity and dynamic loop region, where the loop is highlighted by a red dashed box.

*As*P2Ox has a higher global mobility when compared to its fungal counterparts (Figure 20 and Table 11), most probably due to its monomeric crystal packing. The loop region in fungal models is more flexible when compared to the rest of the enzyme, reinforcing the loop dynamic function in fungal P2Ox. The substrate loop of *As*P2Ox has similar B-factors when compared to the rest of the protein as well as when comparing to the substrate loop of fungal models PeP2Ox (1TZL) and PcP2Ox (4MIF) (Table 11), suggesting a dynamic function of this region also in this bacterial P2Ox. The lower B-factor values of *Tm*P2Ox are likely justified by the fact that the substrate loop finds itself in a closed conformation due to the presence of acetate, as mentioned before. A Graphical representation of the B-factors variation along the residue sequence can be found annexed in the supplementary material.

Table 11 – B-factor values for the substrate loop and overall enzyme of *As*P2Ox and available fungal models. A Graphical representation of the B-factors variation along the residue sequence can be found annexed in the supplementary material.

Protein (PDB)	Substrate Loop	B-factor (Å^2)	
		Global	Substrate Loop
***As*P2Ox**	^{345}DASPVP^{350}	52.2	48.7 – 57.2
***Tm*P2Ox (1TT0)**	^{452}DAFSYG^{457}	11.7	9.5 – 31.9
***Pe*P2Ox (1TZL)**	^{452}DAFSYG^{457}	17.8	41.9 – 48.9
***Pc*P2Ox (4MIF)**	^{458}DAFSYG^{464}	30.2	54.5 – 71.4

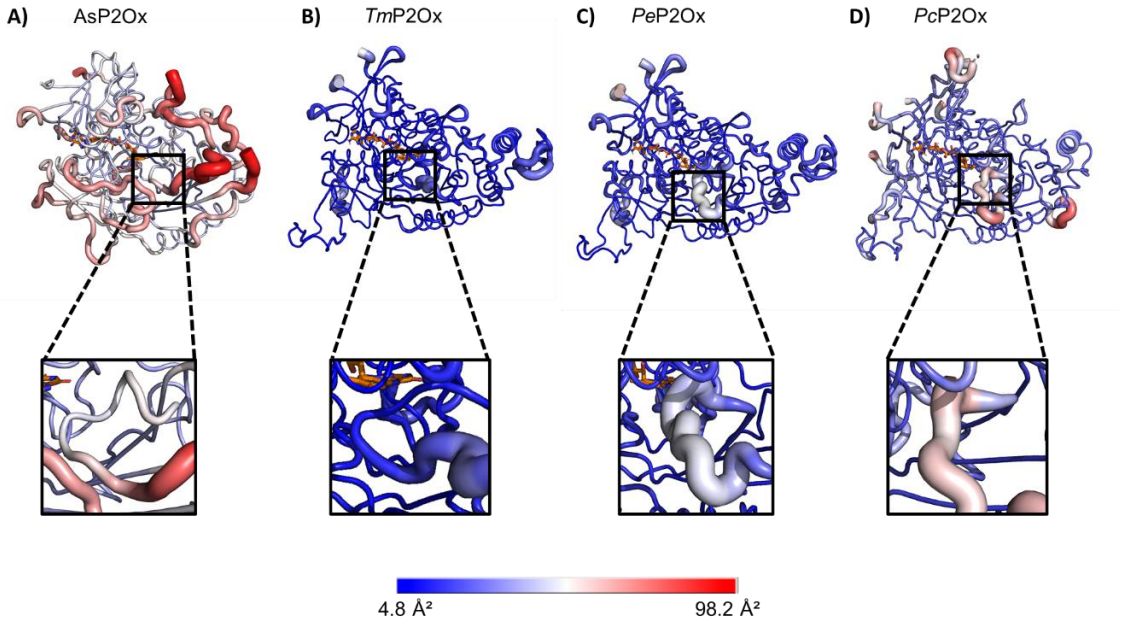

Figure 20 – Representation of B-factor values from *As*P2Ox, *Tm*P2Ox, *Pe*P2Ox and *Pc*P2Ox. Red and thicker regions represent higher B-factor values (high flexibility), while blue thin regions represent lower B-factor values. FAD is represented as sticks with its carbon atoms shown in orange. In the zoomed images, the dynamic substrate loop region is highlighted.

In the structure of the *As*P2Ox-Glucose complex, this substrate loop is not visible in the electronic density map (345-359), as mentioned before, neither is the loop region 77-89, close to the active site entrance and to the substrate loop itself. This is most probably due to the presence of glucose that could have led to a higher flexibility of regions surrounding the active site, hampering a clear electronic density for these loops. The *As*P2Ox-glucose structure has a higher overall B-factor, 62.9 $Å^2$ (vs 52.2 $Å^2$ for *As*P2Ox). The lack of electronic density for the substrate loop suggests a high flexibility and reinforces its dynamic function. Moreover, glucose is occupying part of the substrate loop region (in the same structural position as P348) of the model without glucose, possibly indicating that when glucose is present in the active site, the loop must change its conformation (Figure 21). The present results suggest that a conformational change of the loop region is required for the optimal catalytic function of the enzyme.

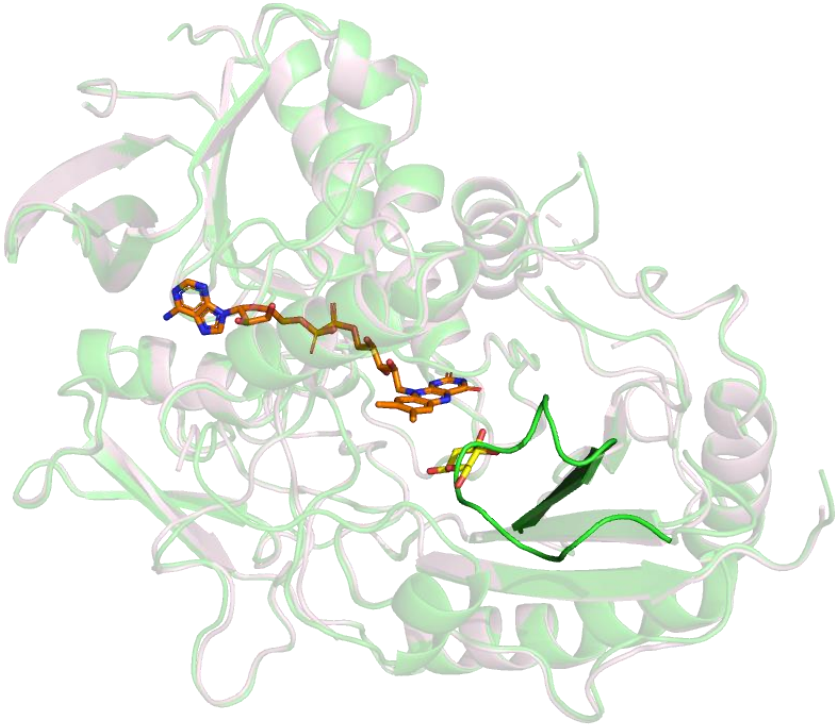

Figure 21 –Highlight of the substrate loop in *As*P2Ox. The glucose (shown as sticks with yellow carbon atoms) from the *As*P2Ox-glucose complex model (in light purple) occupies the loop substrate region of the *As*P2Ox (in green). FAD is also represented as sticks, with its carbon atoms in orange colour.

3.4.4 The *As*P2Ox active site

The residues H440 and N484 close to FAD's isoalloxazine group, are conserved among the studied pyranose 2-oxidase enzymes (Figure 22), including the newly characterized and focus of this work, *As*P2Ox. To confirm the catalytic role of H440, the variant H440A was constructed and characterized in the MET Lab (A. Taborda) and no activity was observed, suggesting that H440 is fundamental for the enzymatic function of *As*P2Ox. No studies have been done yet on N484, but like in other GMC family enzymes, such as cholesterol oxidase, it is possible that this residue by interacting with the flavin through a hydrogen bond is responsible for maintaining an optimal electrostatic potential of the FAD cofactor for the oxidation of the substrate [93]. This latter interaction can only be seen in the *As*P2Ox-glucose complex model, suggesting that this residue might undergo a small conformational change upon entrance of the substrate in the active site, in order to promote the reductive half-reaction. However, clearly more studies are needed to clarify the role of N484. In the *As*P2Ox-glucose complex model, glucose is stabilized by hydrogen bonds between its O1 and R94^{NH2}, as well as between its O3 and T129^{OH} (Figure 22). T129 can also be seen interacting with the flavin N5 through a hydrogen bond. Other interactions might exist between D-glucose and the substrate loop in the *As*P2Ox-Glucose model, since it is not visible, and it is thought to help in the substrate accommodation in the active site. The catalytic base H440^{NE2} is located 4.1 Å away from glucose O2, indicating that the substrate is not located in the same structural position as in fungal enzymes (interacting with the catalytic histidine and asparagine residues). It is possible that D-glucose is not in its final oxidation position, since a fine-tuning of the glucose and enzyme interactions must happen prior to oxidation [16].

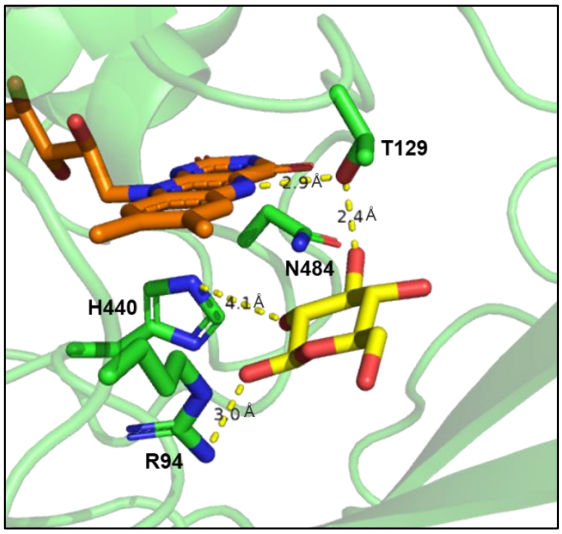

Figure 22 – The active site in the crystal structure of the *As*P2Ox-glucose complex. Glucose and FAD are displayed as sticks, with carbon atoms coloured yellow and orange, respectively. The catalytic residues (H440 and N484), as well as the residues interacting with glucose (R94 and T129) are also represented as sticks with green carbon atoms. The interatomic distances are shown as yellow dashed lines.

3.4.5 FAD binding domain

The FAD cofactor establishes contacts with several residues, displayed in Table 12 and Figure 23, and although some residues are not conserved, the only notable difference among *As*P2Ox and fungal P2Oxs is related with the absence of a covalent bond between H127 and the FAD in *As*P2Ox.

Table 12 – Interactions between FAD atoms and *As*P2Ox residues.

FAD atom	Residue/atom	Distance (Å)
N1	Thr486/OG1	3.12
O2	Thr486/OG1	2.81
	Thr486/N	3.24
	Ala131/O	3.47
N3	Ala131/O	3.01
O4	Ala131/N	2.82
	Ala130/N	3.38
	Thr129/OG1	3.39
N5	Thr129/N	3.11
	Thr129/OG1	3.18
O3'	Thr486/OG1	2.74
	Asn474/ND2	2.90
HO3'	Asn474/ND2	2.52
O4'	Ala125/N	3.40
O1P	Thr20/OG1	3.16
O2P	Thr20/OG1	2.89
	Thr20/N	3.02
	Pro19/N	3.31
O1A	Met124/N	2.94
O4B	Gly268/O	3.46
O3B	Glu41/OE1	2.79
	Glu41/OE2	3.32
	Asn120/O	2.95
HO3A	Gly123/N	2.75
	Asn120/O	2.88
	Glu41/OE1	2.91
O2B	Glu41/OE2	2.60
	Ser118/3.20	3.20
N1A	Val234/N	3.04
N3A	Val42/N	3.16
N6A	Val234/O	2.95
H61A	Val234/O	1.96

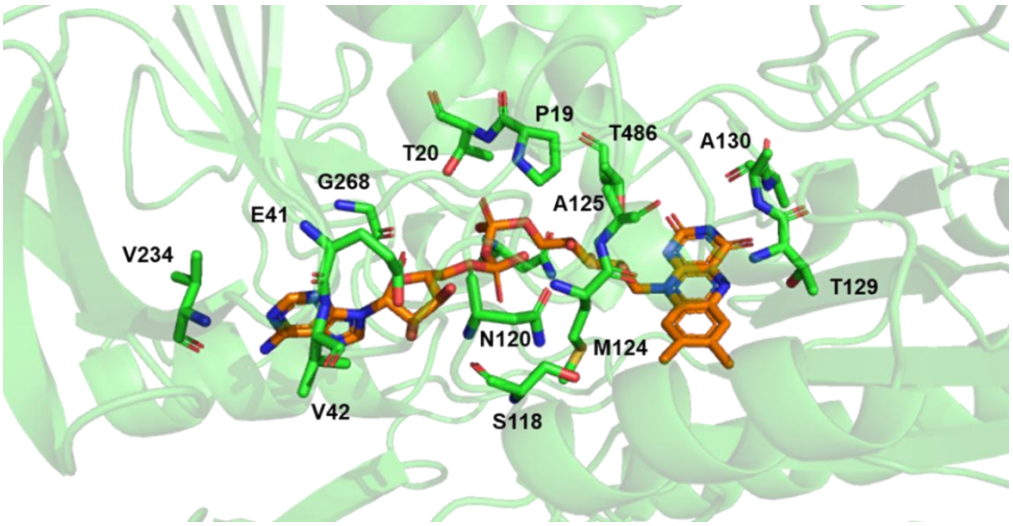

Figure 23 – Representation of the FAD binding motif. FAD is shown as sticks with carbon atoms coloured orange. Nitrogen and oxygen atoms are represented in dark blue and red, respectively. The carbon atoms of protein residues displayed as sticks are shown in green.

In fungal P2Ox structures, FAD is covalently bonded to a histidine residue (H167 in *T. mulicolor*) present in a conserved binding motif: ^{165}STHW^{168} [12]. This motif is highlighted by a yellow box in the structural alignment in a previous section (Figure 16). *As*P2Ox has a ^{125}AAHW^{128} flavinylation motif near the FAD cofactor. Despite being closer than expected, the histidine's $N^{\varepsilon 2}$ that binds FAD in other P2Oxs is 6.5 Å from the FAD 8-methyl group in *As*P2Ox (H127), confirming the experimental results that show a non-covalent binding of FAD. Despite the residue difference between *As*P2Ox and the fungal P2Ox's FAD binding motif (AAHW vs STHW, respectively), the non-covalent bond might not be directly related with these differences. Moreover, residue N120 is interacting through a hydrogen bond with H127 (Figure 24) while in fungal P2Oxs, this asparagine is replaced by hydrophobic residues (Figure 16 – structural alignment), namely a valine in *Tm*P2Ox (PDB 1TT0: V160) and *Pe*P2Ox (PDB 1TZL: V160), and a glycine in the P2Ox from *Pc*P2Ox (PDB 4MIF: G151). This hydrogen bond in *As*P2Ox might be preventing H127 proximity to the FAD, hampering the formation of a covalent bond.

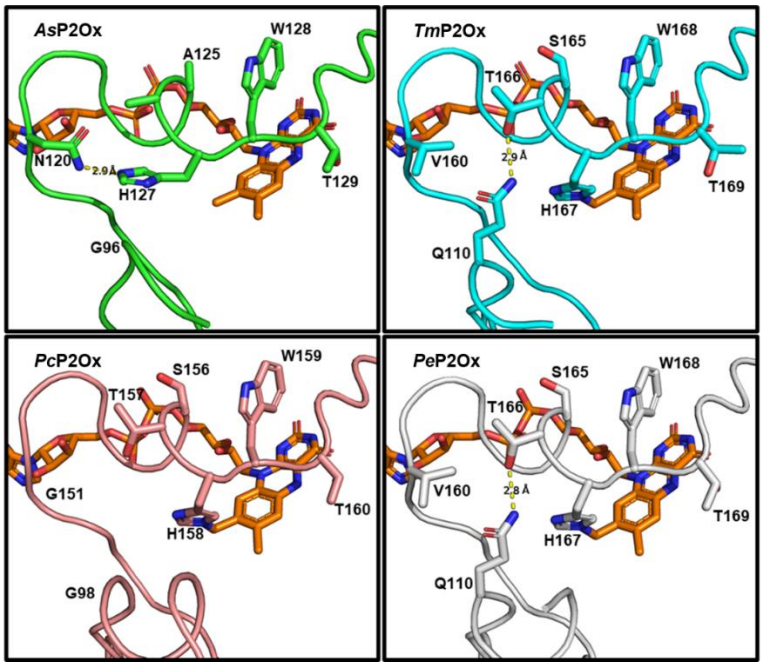

Figure 24 – Representation of the FAD binding motif and the residues that interact with this region, from *As*P2Ox in green, *Tm*P2Ox in light blue (PDB 1TT0), *Pc*P2Ox in pink (PDB 4MIF) and *Pe*P2Ox in gray (PDB 1TZL). Every relevant residue is identified with the respective code letter and residue number. FAD carbon atoms are represented in orange in all four models. Nitrogen and oxygen atoms are represented in dark blue and red, respectively.

Additionally, *Tm*P2Ox and *Pe*P2Ox have a glutamine residue (Q110) in the same structural position as H127 in *As*P2Ox (Figure 24). This glutamine residue position might be stabilized by T166 since they are close enough to form a hydrogen bond (2.9 Å in *Tm*P2Ox and 2.8 Å in *Pe*P2Ox). Instead of Q110, G96 can be found in *As*P2Ox, which is a small residue without side chain that probably has no impact on the proximity between H127 and the FAD. To further study the FAD binding, *As*P2Ox variants were designed by rational and semi-rational approaches.

3.5 Rational and semi-rational design to promote the covalent linkage of FAD

3.5.1 Rational design and computational “validation”

As previously mentioned, the FAD cofactor in *As*P2Ox is not covalently bonded, in contrast to all the fungal P2Oxs with available structural models, what could be hampering its stability. It was observed that in *As*P2Ox, the $N^{\varepsilon 2}$ of the histidine (H127) that covalently binds to FAD in the fungal enzymes, has a different conformation (Figure 24) probably due to a hydrogen bond with N120, preventing its proximity with the FAD, and consequently hampering the formation of a covalent bond. Additionally, it was observed that in fungal structures, residues with long side-chains such as lysine and glutamine were found in the vicinity of the histidine which could also force the histidine to adopt the right conformation in favor of a covalent linkage.

Engineering by rational design was followed to test if it was possible to force the covalent bond between the FAD and *As*P2Ox, since a covalently bonded cofactor could possibly increase the stability of the enzyme in biosensor applications. The variant N120V was proposed to break the hydrogen bond stabilizing the H127 in the conformation that prevents the covalent linkage with FAD. Moreover, three other variants V54K, S70Q and G96Q were proposed to introduce spatial constraints that could force the histidine to rotate to a conformation more likely to be covalently linked, as seen in the fungal models. Before the experimental construction of the proposed mutants, all variants were simulated using Swissmodel [85] to computationally validate the hypothesis. A comparison of the models obtained with Swissmodel and *As*P2Ox wild-type model can be found in Figure 25. When modelled with Swissmodel, no conformational change was seen on H127 – Figure 25 (A) – of N120V. This variant was produced and studied either way. The other three variants showed promising results when modelled with Swissmodel, with all H127 from these variants being close to the FAD – Figure 25 (B-D).

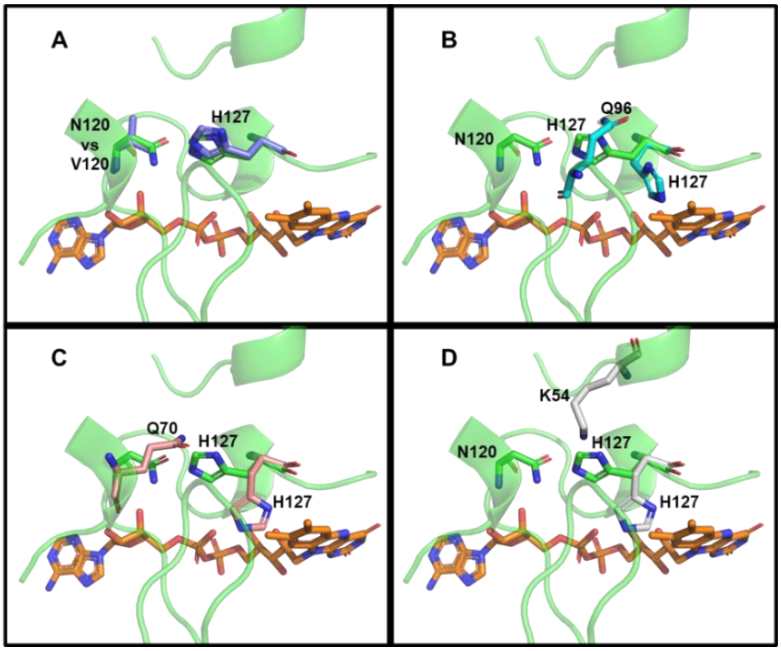

Figure 25 – Visual representation and comparation of the FAD binding region from *As*P2Ox wild-type model obtained by X-ray crystallography (control in green) and models of the N120V, G96Q, S70Q and V54K variants obtained with SwissModel. A) H127 and V120 from N120V are represented in purple. B) H127 and Q96 from G96Q are represented in blue. C) H127 and Q70 from S70Q are represented in pink. D) H127 and K54 from V54K are represented in white. *As*P2Ox wild-type H127 and N120 are represented in green in A, B, C and D, while FAD is represented in orange.

3.5.2 Construction and analysis of rational designed mutants to incorporate FAD covalently in *As*P2Ox

After the computational simulations, the proposed mutants were constructed, and the purified protein preparations analyzed. For all the four variants: N120V, V54K, S70Q (Figure 26) and G96Q (expressed and purified by a lab partner) the protein preparations lacked the typical yellow colour that FAD provides when present, suggesting that the flavin group had not been incorporated. The UV-Vis spectra confirmed that the typical FAD bands were not present (370 nm and 450 nm). This observation suggests that changes induced by the mutations, in that region of the protein, interfered with FAD incorporation. It can be hypothesized that if the mutations rotate the H127, as suggested by the computational simulations, maybe this rotation hampers the FAD binding instead of linking it covalently. This hypothesis could be addressed by determining the crystal structures of the variants, which was not performed in the scope of this thesis work.

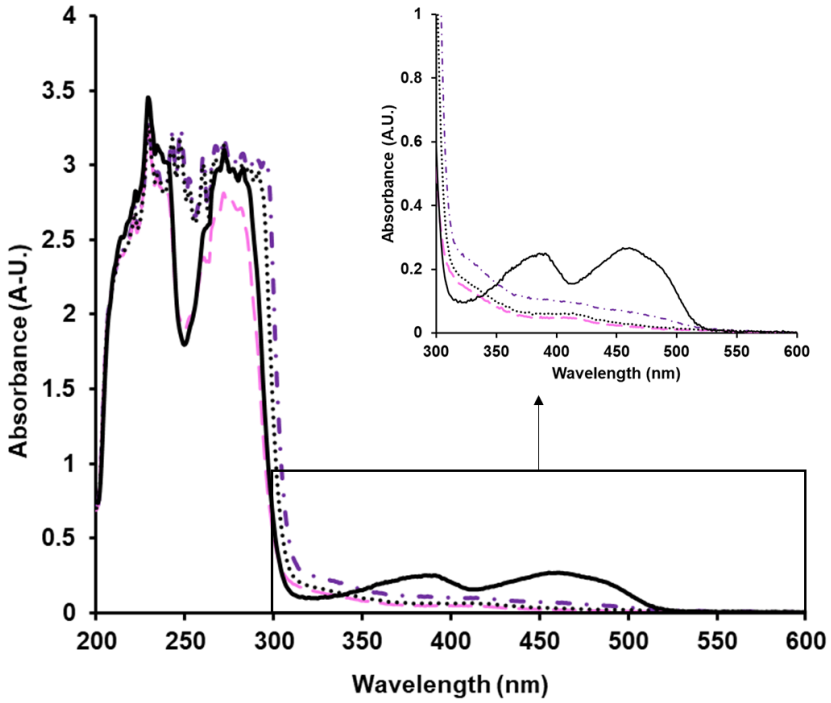

***Figure 26* – Graphical representation of the UV-Vis absorption spectra** of *As*P2Ox wild-type (——— in black), N120V (– · – · – in purple), V54K (····· in black) and S70Q (– – – – in pink). Spectra acquired using a Nicolet Evolution 300 UV Vis spectrophotometer (Thermo Electron Corporation).

3.5.3 Saturation of residue N120 to incorporate FAD covalently in *As*P2Ox

To continue exploring the binding of FAD, a semi-rational design was performed targeting the residue N120. Since N120V afforded a change in residue size and polarity at the FAD binding site, it was though that it could prevent the accommodation of FAD, consequently prevent its incorporation. For this reason, site saturation mutagenesis was applied to test all possible residues in this position. Besides using the wild-type as template, the previously tested variants – V54K, S70Q and G96Q – were also tested to check for possible synergistic effects (eliminate the hydrogen bond plus spatial constraints) in the double mutants.

The variants were produced and screened in a high-throughput mode using 96-well plates. The relative activity of the variants as compared to wild-type *As*P2Ox was calculated, and a threshold of 0.2 was chosen to select variants. The number of screened variants as well as the number of mutants with a relative activity higher than 0.2, for each one of the four libraries constructed, is presented in Table 13.

Table 13 – Number of screened colonies and variants displaying relative activity higher than 0.2 for the wild-type, G96Q, V54K and S70Q libraries.

DNA template	Number of screened colonies (96-well plates)	Number of variants displaying activity
Wild-type	180	**40** (~22 % of the library)
G96Q	79	**0**
V54K	64	**0**
S70Q	166	**0**

The library created using wild-type as background was the only one having variants with relative activity higher than 0.2 (around 22 % of the screened variants in the library) (Table 13). The absence of variants with activity in the double mutant libraries indicates that G96Q, V54K and S70Q mutations hinder the correct FAD incorporation, as previously observed when the single mutations were performed. These were all variants where a spatial constraint was applied to force the covalent bond of the FAD by breaking the hydrogen bond between H127 and N120 and are probably too extreme changes for a sensitive region of the protein.

From the 489 screened variants only 40 were left, around 22% of the wild-type library and 4% of the total number of screened variants. All 40 variants had their DNA extracted and sequenced. The results showed that the majority were the wild-type enzyme (N120), 2 carried out the mutation N120C (codon change from AAC to TGC) and 1 carried the mutation N120Y (codon change from AAC to TAC). N120C displayed a relative activity of 0.5, while N120Y showed a relative activity of 0.2 when compared with the wild-type. N120C transitioned to a less polar residue with a smaller side chain, while N120Y maintained the polarity of the residue, but had a considerable increase

in side-chain size. Interestingly, both variants still offer the possibility for a hydrogen bond with H127, the elimination of which was the objective of the semi-rational design.

N120C and N120Y where then grown in 1-L scale and the enzyme was produced and purified. To confirm that the pure protein preparation of the variants incorporated the cofactor, an absorbance spectrum was performed (Figure 27). N120C and N120Y showed yellow colour and the typical FAD absorbance band at 450 nm, as shown in Figure 27. To further study N120C and N120Y, both mutants were characterized in a more detailed manner for their FAD attachment, FAD thermal stability and activity.

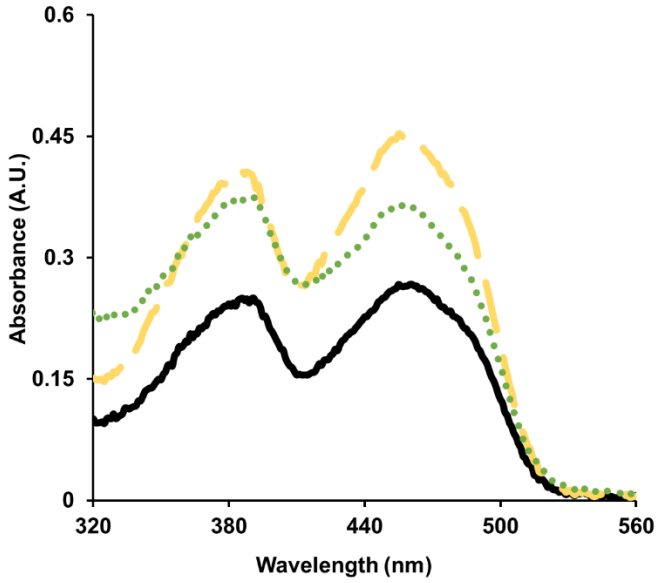

Figure 27 – Graphical representation of the UV-Vis absorption spectra of *As*P2Ox wild-type (——— in black), N120C (**– – – –** in yellow), and N120Y **(- - - - -** in green). Spectra acquired using a Nicolet Evolution 300 UV Vis spectrophotometer (Thermo Electron Corporation).

3.5.4 FAD attachment

To study the FAD attachment and observe the type of cofactor bond, covalent or non-covalent, a protein denaturation treatment with trichloroacetic acid at high temperature, was performed. If covalently bonded, the FAD would precipitate with the protein, while if not covalently bonded, the cofactor would be released upon protein denaturation and stay in solution that would remain yellow.

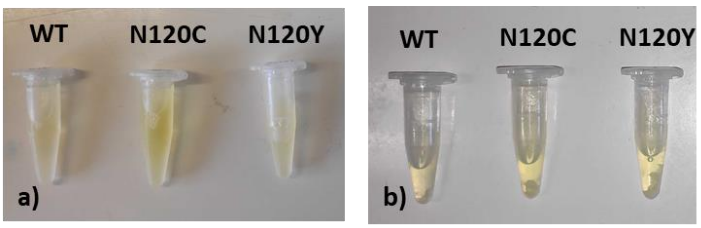

Figure 28 – *As*P2Ox wild-type, N120C and N120Y purified solutions a) prior to and b) after the denaturation treatment.

The images in Figure 28 unequivocally show that the solution maintained the yellow colour after the denaturation treatment, while the pellet was white, suggesting that FAD was not covalently incorporated in the protein. To confirm the latter observation, UV-Vis spectra were acquired for the wild-type, N120C and N120Y, prior to and after the denaturation treatment (Figures 29). The slight difference between the spectra might be due to the acidic state of the solution after the treatment. However, both spectra showed the typical FAD absorbance band at 450 nm. The non-covalent bond of FAD in these variants could be explained by the fact that both mutated residues can also form hydrogen bonds with H127, hindering the covalent bond between the histidine and the FAD, just like in the wild-type.

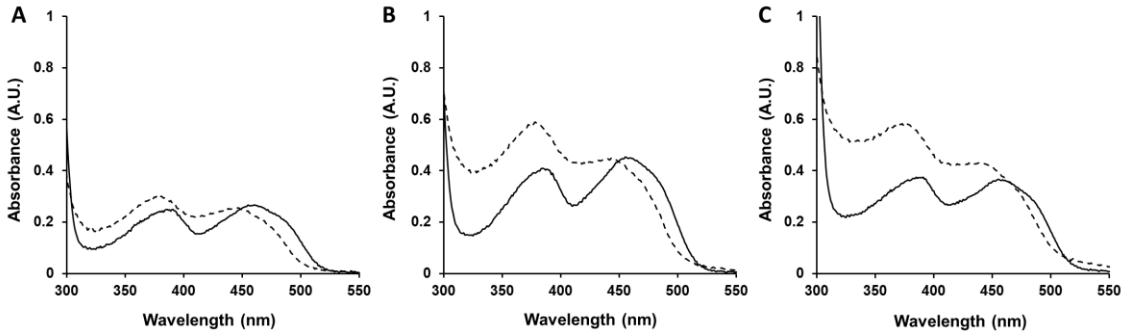

Figure 29 – Graphical representation of the UV-Vis absorption spectra of *As*P2Ox wild-type (A), N120C (B) and N120Y (C), prior (———) and after (- - - - -) the denaturation treatment. Spectra acquired using a Nicolet Evolution 300 UV Vis spectrophotometer (Thermo Electron Corporation).

Despite not being covalently bonded, it was observed (Figure 28) that N120C has a much more intense yellow colour when compared to N120Y variant and even the wild-type. This stronger colour could be the result of a higher concentration of protein or a consequence of a higher percentage of protein with FAD incorporated. To clarify this, FAD/protein ratios were determined (Table 14). The results show a higher incorporation of FAD in N120C, with 43.9% of the total protein having FAD incorporated, while the wild-type and N120Y only have 27.6% and 13.8% of its protein solution FAD-loaded, respectively. This shows that despite not having its FAD covalently bonded, N120C has a higher FAD incorporation. To test if the incorporation of FAD and its stability where correlated, the FAD thermal stability was studied by fluorescence.

Table 14 – Ratios of FAD cofactor per protein molecule in percentage, for the wind-type, N120C and N120Y.

	WT	N120C	N120Y
FAD/Protein (%)	27.6	43.9	13.8

3.5.5 Thermal stability

The FAD thermal stability was studied using a spectrofluorometer to follow intrinsic fluorescence of FAD during a temperature gradient (20-60°C). The fit to the experimental points showed a good fitting. The obtained data is summarized in the Figure 30.

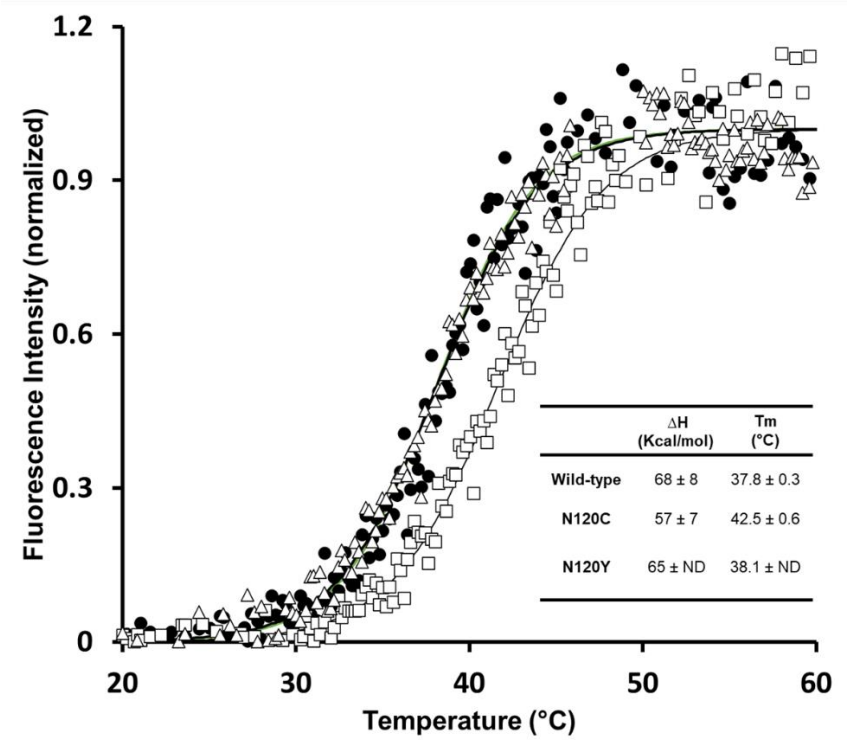

Figure 30 – Plot of the normalized fluorescence intensity with increasing temperature and values of ΔH° (Kcal/mol) and T_m (°C) for wild-type *As*P2Ox and N120C and N120Y variants. Experimental points are represented as filled circles for the wild-type, open triangles to N120Y and open squares to N120C. Solid lines represent the fit to the experimental points, performed using fraction unfolded = exp(-ΔG°/RT)/(1+exp(-ΔG°/RT)), where ΔG° values are theoretical and calculated using the ΔH° and ΔS° values previously acquired. N120Y does not have an associated error since only one experimental procedure was done for this variant.

N120C value showed a small increase (+5°C) of T_m as compared to the wild-type and N120Y. This shift can also be seen in the fit to the experimental points. However, despite N120Y having a lower incorporation of FAD when compared to the wild-type, they both have very similar values of T_m (around 38°C), meaning that, with the data obtained, it isn't possible to directly correlate the FAD incorporation with its stability.

3.5.6 N120C and N120Y enzymatic activity

The variants N120C and N120Y were also tested for their activity towards the oxidation of D-glucose, to investigate the effect of these mutations in catalysis. The obtained data is represented in Figure 31 and the apparent catalytic parameters are summarized in Table 15. Both variants showed an increase of around 2-fold in the value of K_m when compared to the wild-type, and N120C has a slightly higher k_{cat} value. Analysing the values of k_{cat}/K_m, it is possible to conclude that no catalytic improvements were obtained.

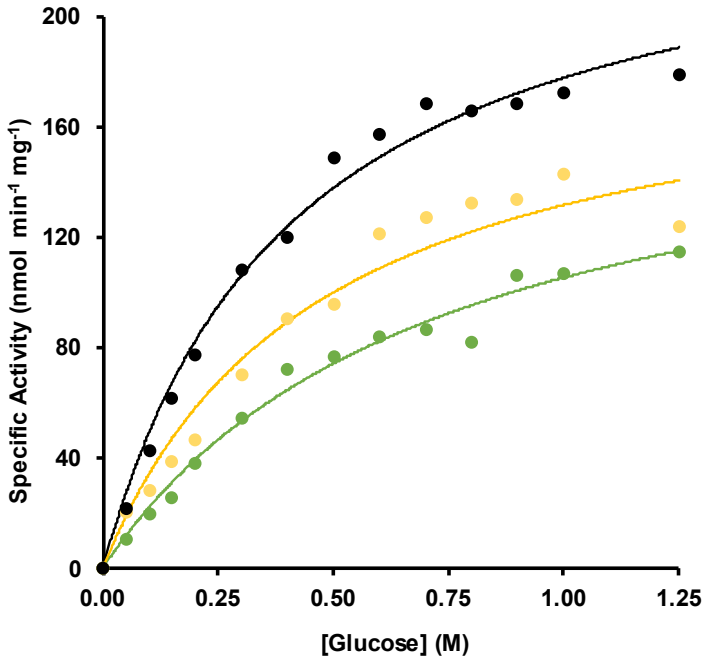

Figure 31 – Apparent steady-state kinetic curves for *As*P2Ox (black), N120C (yellow) and N120Y (green). Assays were performed by mixing AAP 4 mM, DCHBS 40 mM and 20 µL of protein (stock ~1 mg/mL) 10 U/mL HRP in 100 mM sodium phosphate pH 7.5, at different concentrations of glucose. Activity was monitored using a Synergy2 microplate reader (BioTek), pre-heated at 40°C for optimum activity, for 10 min at 515 nm (ε = 26,000 M^{-1} cm^{-1}). Experimental points were fitted into the Michaelis-Menten equation using the Origin-Lab software.

Table 15 – Values of K_m, k_{cat} and k_{cat}/K_m for *As*P2Ox wild-type, N120C and N120Y.

Variant	K_m (M)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
Wild-type	0.37 ± 0.13	0.19 ± 0.05	0.49 ± 0.10
N120C	0.88 ± 0.19	0.29 ± 0.05	0.34 ± 0.07
N120Y	0.70 ± 0.07	0.15 ± 0.03	0.22 ± 0.02

Based on these results it is possible to propose that this position shows a high sensitivity to mutations and changing the asparagine to other residues (cysteine and tyrosine seem to be the only exceptions) easily destabilizes its vicinity, hinders the FAD incorporation and consequently the enzymatic function. Furthermore, only mutations leading to the preservation of the hydrogen bond between residue 120 and H127 had the FAD incorporated, suggesting that the hydrogen bond is crucial for incorporation of the flavin group.

4. Conclusions

In this work the first bacterial P2Ox crystal structure was characterized, to a resolution of 2.01 Å, as well as of the enzyme with D-Glucose in its active site, at 2.35 Å resolution. The protein crystallized in space group $C222_1$ as a monomer, forming yellow crystals with dimensions around 100 µm x 100 µm x 100 µm. The crystals obtained allowed seeding experiments, providing apparently good crystals but with internal disorder. The *As*P2Ox crystal structure was refined to R_{work} and R_{free} values of 19.8% and 24.4%, respectively. The crystal structure of the *As*P2Ox-Glucose complex is of lesser quality, with the R_{work} and R_{free} values of 24.5% and 26.2%. The structures were compared with available crystal structures of fungal P2Oxs and two major P2Ox domains were confirmed in *As*P2Ox: the substrate binding domain and the FAD-binding domain. The lack of oligomerization domains in *As*P2Ox supports the observation of a monomeric form in solution and in the crystal structures. A fungal head domain – responsible for association with polysaccharides from the cell-wall for immobilization at the fungal periplasmatic space – is also missing in *As*P2Ox.

Previous studies indicated that in contrast to its fungal counterparts, the FAD cofactor is not covalently bonded in *As*P2Ox. The structural analysis of the FAD binding motif helped understanding the reasons behind this weaker attachment of FAD to the enzyme. A hydrogen bond between H127 and N120 seems to be preventing the histidine to covalently bind the FAD. Variants were designed using rational and semi-rational design to try to eliminate the hydrogen bond and force the covalent bond of FAD. Only two variants (N120C and N120Y), also capable of generating the hydrogen bond, incorporated the cofactor and generated catalytically active mutants, suggesting that the 120th position is very sensitive to mutations and that its interaction with H127 is crucial for FAD incorporation. Both mutants were obtained with saturation mutagenesis, and none had its FAD covalently bound. N120C showed a higher thermal stability than the wild-type (T_m = 43°C vs 38°C), responsible for its higher ratio of FAD incorporation (FAD/protein = 44% vs 28%). N120Y didn't show any improvements on its FAD stability, with a T_m of 38°C and very poor FAD incorporation (13.8%). Both variants had slight lower catalytic efficiencies than the wild-type (k_{cat}/K_m = 0.49 $M^{-1} s^{-1}$), with values of 0.34 and 0.22 $M^{-1} s^{-1}$ for N120C and N120Y, respectively.

A substrate loop found in fungal P2Ox was thought to gate and help accommodating the substrate in the active site. The corresponding loop in *As*P2Ox seems to have similar functions. The non-visibility of the substrate loop in the electron density maps of the *As*P2Ox-Glucose crystal structure, hampers any further conclusion on the role of the loop in the bacterial P2Ox. Despite this constraint, the active site seems exposed and optimal for direct electron transfer in third generation biosensors, something that cannot be achieved with the homotetrameric fungal P2Oxs.

The work performed allowed a better understanding of the main structural characteristics of *As*P2Ox and enabled the proposal of different hypotheses explaining the low catalytic values of the enzyme, opening perspectives to different approaches of enzyme engineering, not only

through experimental but also computationally assisted rational design. The dynamic loop of the complex structure could be generated by computational modelling, that might help understanding its role in the accessibility of the active site and accommodation of the substrate. New variants could be designed using computational prediction methods, since the rational design applied in the work didn't provide catalytically improved variants. Finally, determining the crystal structure of previously obtained variants – directed evolution was used by other lab members –could be of interest to look for structural alterations responsible for increased catalytic efficiency and/or stability.

5. References

[1] D. R. Cavener, “GMC oxidoreductases. A newly defined family of homologous proteins with diverse catalytic activities,” *J. Mol. Biol.*, vol. 223, no. 3, pp. 811–814, 1992.

[2] L. Sützl, G. Foley, E. M. J. Gillam, M. Bodén, and D. Haltrich, “The GMC superfamily of oxidoreductases revisited: Analysis and evolution of fungal GMC oxidoreductases,” *Biotechnol. Biofuels*, vol. 12, no. 1, pp. 1–18, 2019.

[3] D. Monti, G. Ottolina, G. Carrea, and S. Riva, “Redox reactions catalyzed by isolated enzymes,” *Chem. Rev.*, vol. 111, no. 7, pp. 4111–4140, 2011.

[4] S. Mendes *et al.*, “Characterization of a bacterial pyranose 2-oxidase from Arthrobacter siccitolerans,” *J. Mol. Catal. B Enzym.*, vol. 133, pp. S34–S43, 2016.

[5] A. T. Abrera, L. Sützl, and D. Haltrich, “Pyranose oxidase: A versatile sugar oxidoreductase for bioelectrochemical applications,” *Bioelectrochemistry*, vol. 132, p. 107409, 2020.

[6] F. Giffhorn, “Fungal pyranose oxidases: Occurrence, properties and biotechnical applications in carbohydrate chemistry,” *Appl. Microbiol. Biotechnol.*, vol. 54, no. 6, pp. 727–740, 2000.

[7] G. Daniel, J. Volc, E. Kubatova, and T. Nilsson, “Ultrastructural and immunocytochemical studies on the H2O2-producing enzyme pyranose oxidase in Phanerochaete chrysosporium grown under liquid culture conditions,” *Appl. Environ. Microbiol.*, vol. 58, no. 11, pp. 3667–3676, 1992.

[8] P. L. Herzog *et al.*, “Versatile oxidase and dehydrogenase activities of bacterial pyranose 2-oxidase facilitate redox cycling with manganese peroxidase in vitro,” *Appl. Environ. Microbiol.*, vol. 85, no. 13, 2019.

[9] W. Bao, Y. Fukushima, K. A. Jensen, M. A. Moen, and K. E. Hammel, “Oxidative degradation of non-phenolic lignin during lipid peroxidation by fungal manganese peroxidase,” *FEBS Lett.*, vol. 354, no. 3, pp. 297–300, 1994.

[10] J. Volc, E. Kubfitovfi, P. Sedmera, G. Daniel, and J. Gabriel, “Pyranose oxidase and pyranosone dehydratase: enzymes responsible for conversion of D-glucose to cortalcerone by the basidiomycete Phanerochaete chrysosporium,” pp. 297–301, 1991.

[11] M. Bannwarth, S. Bastian, D. Heckmann-Pohl, F. Giffhorn, and G. E. Schulz, “Crystal structure of pyranose 2-oxidase from the white-rot fungus Peniophora sp.,” *Biochemistry*, vol. 43, no. 37, pp. 11683–11690, 2004.

[12] B. Martin Hallberg, C. Leitner, D. Haltrich, and C. Divne, “Crystal structure of the 270 kDa homotetrameric lignin-degrading enzyme pyranose 2-oxidase,” *J. Mol. Biol.*, vol. 341, no. 3, pp. 781–796, 2004.

[13] N. Hassan *et al.*, “Crystal structures of Phanerochaete chrysosporium pyranose 2-oxidase suggest that the N-terminus acts as a propeptide that assists in homotetramer assembly,” *FEBS Open Bio*, vol. 3, pp. 496–504, 2013.

[14] T. Wongnate and P. Chaiyen, “The substrate oxidation mechanism of pyranose 2-oxidase and other related enzymes in the glucose-methanol-choline superfamily,” *FEBS J.*, vol. 280, no. 13, pp. 3009–3027, 2013.

[15] M. Prongjit, J. Sucharitakul, T. Wongnate, D. Haltrich, and P. Chaiyen, “Kinetic mechanism of pyranose 2-oxidase from Trametes multicolor,” *Biochemistry*, vol. 48, no. 19, pp. 4170–4180, 2009.

[16] T. Wongnate, P. Surawatanawong, L. Chuaboon, N. Lawan, and P. Chaiyen, “The Mechanism of Sugar C−H Bond Oxidation by a Flavoprotein Oxidase Occurs by a Hydride Transfer Before Proton Abstraction,” *Chem. - A Eur. J.*, vol. 25, no. 17, pp. 4460–4471, 2019.

[17] M. Kujawa *et al.*, “Structural basis for substrate binding and regioselective oxidation of

monosaccharides at C3 by pyranose 2-oxidase," *J. Biol. Chem.*, vol. 281, no. 46, pp. 35104–35115, 2006.

[18] J. Sucharitakul, M. Prongjit, D. Haltrich, and P. Chaiyen, "Detection of a C4a-hydroperoxyflavin intermediate in the reaction of a flavoprotein oxidase," *Biochemistry*, vol. 47, no. 33, pp. 8485–8490, 2008.

[19] W. Pitsawong *et al.*, "A conserved active-site threonine is important for both sugar and flavin oxidations of pyranose 2-oxidase," *J. Biol. Chem.*, vol. 285, no. 13, pp. 9697–9705, 2010.

[20] M. J. Han, H. T. Choi, and H. G. Song, "Purification and Characterization of Pyranose Oxidase from the White Rot Fungus Trametes multicolor," *J. Microbiol.*, vol. 43, no. 6, pp. 555–560, 2005.

[21] M. Bannwarth, D. Heckmann-Pohl, S. Bastian, F. Giffhorn, and G. E. Schulz, "Reaction geometry and thermostable variant of pyranose 2-oxidase from the white-rot fungus Peniophora sp.," *Biochemistry*, vol. 45, no. 21, pp. 6587–6596, 2006.

[22] K. Zhang *et al.*, "Identification and characterization of a novel bacterial pyranose 2-oxidase from the lignocellulolytic bacterium Pantoea ananatis Sd-1," *Biotechnol. Lett.*, vol. 40, no. 5, pp. 871–880, 2018.

[23] M. Manzanera, J. I. Vílchez, and C. Calvo, "Genome Sequence of Arthrobacter siccitolerans 4J27, a Xeroprotectant- Producing Desiccation-Tolerant Microorganism," vol. 2, no. 3, pp. 14–15, 2014.

[24] A. F. T. Taborda and A. F. Teixeira Taborda, "Development of novel bacterial enzymes for medical diagnosis through directed evolution," p. 66, 2019.

[25] H. H. Nguyen, S. H. Lee, U. J. Lee, C. D. Fermin, and M. Kim, "Immobilized enzymes in biosensor applications," *Materials (Basel).*, vol. 12, no. 1, pp. 1–34, 2019.

[26] C. R. Ispas, G. Crivat, and S. Andreescu, "Review: Recent Developments in Enzyme-Based Biosensors for Biomedical Analysis," *Anal. Lett.*, vol. 45, no. 2–3, pp. 168–186, 2012.

[27] K. Y. Kwon *et al.*, "Electrochemical activity studies of glucose oxidase (GOx)-based and pyranose oxidase (POx)-based electrodes in mesoporous carbon: Toward biosensor and biofuel cell applications," *Electroanalysis*, vol. 26, no. 10, pp. 2075–2079, 2014.

[28] A. A. Saei, J. E. N. Dolatabadi, P. Najafi-Marandi, A. Abhari, and M. de la Guardia, "Electrochemical biosensors for glucose based on metal nanoparticles," *TrAC - Trends Anal. Chem.*, vol. 42, no. x, pp. 216–227, 2013.

[29] M. Wooten, S. Karra, M. Zhang, and W. Gorski, "On the direct electron transfer, sensing, and enzyme activity in the glucose oxidase/carbon nanotubes system," *Anal. Chem.*, vol. 86, no. 1, pp. 752–757, 2014.

[30] S. Ferri, K. Kojima, and K. Sode, "Review of glucose oxidases and glucose dehydrogenases: A bird's eye view of glucose sensing enzymes," *J. Diabetes Sci. Technol.*, vol. 5, no. 5, pp. 1068–1076, 2011.

[31] S. Timur, Y. Yigzaw, and L. Gorton, "Electrical wiring of pyranose oxidase with osmium redox polymers," *Sensors Actuators, B Chem.*, vol. 113, no. 2, pp. 684–691, 2006.

[32] H. Lidén, J. Volc, G. Marko-Varga, and L. Gorton, "Pyranose Oxidase Modified Carbon Paste Electrodes for Monosaccharide Determination," *Electroanalysis*, vol. 10, no. 4, pp. 223–230, 1998.

[33] C. Ozdemir, F. Yeni, D. Odaci, and S. Timur, "Electrochemical glucose biosensing by pyranose oxidase immobilized in gold nanoparticle-polyaniline/AgCl/gelatin nanocomposite matrix," *Food Chem.*, vol. 119, no. 1, pp. 380–385, 2010.

[34] T. C. Gokoglan, S. Soylemez, M. Kesik, S. Toksabay, and L. Toppare, "Selenium containing conducting polymer based pyranose oxidase biosensor for glucose detection," *Food Chem.*, vol. 172, pp. 219–224, 2015.

[35] D. Odaci, A. Telefoncu, and S. Timur, “Pyranose oxidase biosensor based on carbon nanotube (CNT)-modified carbon paste electrodes,” *Sensors Actuators, B Chem.*, vol. 132, no. 1, pp. 159–165, 2008.

[36] F. Jing *et al.*, “A novel fully enzymatic method for determining glucose and 1,5-anhydro-d-glucitol in serum of one cuvette,” *Appl. Biochem. Biotechnol.*, vol. 150, no. 3, pp. 327–335, 2008.

[37] W. Nowatzke, M. J. Sarno, N. C. Birch, D. F. Stickle, T. Eden, and T. G. Cole, “Evaluation of an assay for serum 1,5-anhydroglucitol (GlycoMark™) and determination of reference intervals on the Hitachi 917 analyzer,” *Clin. Chim. Acta*, vol. 350, no. 1–2, pp. 201–209, 2004.

[38] K. M. Dungan, “1,5-Anhydroglucitol (GlycoMark™) as a marker of short-term glycemic control and glycemic excursions,” *Expert Rev. Mol. Diagn.*, vol. 8, no. 1, pp. 9–19, 2008.

[39] I. P. Petrounia and F. H. Arnold, “Designed evolution of enzymatic properties,” *Curr. Opin. Biotechnol.*, vol. 11, no. 4, pp. 325–330, 2000.

[40] A. Nixon and S. Firestine, “Rational and ‘Irrational’ Design of Proteins and Their Use in Biotechnology,” *IUBMB Life*, vol. 49, no. 3, pp. 181–187, 2000.

[41] N. M. Antikainen and S. F. Martin, “Altering protein specificity: Techniques and applications,” *Bioorganic Med. Chem.*, vol. 13, no. 8, pp. 2701–2716, 2005.

[42] S. Brakmann, “Discovery of superior enzymes by directed molecular evolution,” *ChemBioChem*, vol. 2, no. 12, pp. 865–871, 2001.

[43] A. Stefan, A. Radeghieri, A. Gonzalez Vara y Rodriguez, and A. Hochkoeppler, “Directed evolution of β-galactosidase from Escherichia coli by mutator strains defective in the 3′→5′ exonuclease activity of DNA polymerase III,” *FEBS Lett.*, vol. 493, no. 2–3, pp. 139–143, 2001.

[44] E. O. McCullum, B. A. R. Williams, J. Zhang, and J. C. Chaput, “Random mutagenesis by error-prone PCR,” *Methods Mol. Biol.*, vol. 634, no. 3, pp. 103–109, 2010.

[45] M. S. Packer and D. R. Liu, “Methods for the directed evolution of proteins,” *Nat. Rev. Genet.*, vol. 16, no. 7, pp. 379–394, 2015.

[46] R. Chen, “Enzyme engineering: Rational redesign versus directed evolution,” *Trends Biotechnol.*, vol. 19, no. 1, pp. 13–14, 2001.

[47] J. D. Bloom, M. M. Meyer, P. Meinhold, C. R. Otey, D. MacMillan, and F. H. Arnold, “Evolving strategies for enzyme engineering,” *Curr. Opin. Struct. Biol.*, vol. 15, no. 4, pp. 447–452, 2005.

[48] O. Spadiut *et al.*, “Engineering of pyranose 2-oxidase: Improvement for biofuel cell and food applications through semi-rational protein design,” *J. Biotechnol.*, vol. 139, no. 3, pp. 250–257, 2009.

[49] C. Salaheddin *et al.*, “Probing active-site residues of pyranose 2-oxidase from Trametes multicolor by semi-rational protein design,” *Biotechnol. J.*, vol. 4, no. 4, pp. 535–543, 2009.

[50] K. K. Yang, Z. Wu, and F. H. Arnold, “Machine-learning-guided directed evolution for protein engineering,” *Nat. Methods*, vol. 16, no. 8, pp. 687–694, 2019.

[51] M. Musil *et al.*, “FireProt: web server for automated design of thermostable proteins,” *Nucleic Acids Res.*, vol. 45, no. W1, pp. W393–W399, 2017.

[52] A. Goldenzweig *et al.*, “Automated Structure- and Sequence-Based Design of Proteins for High Bacterial Expression and Stability,” *Mol. Cell*, vol. 63, no. 2, pp. 337–346, 2016.

[53] H. J. Wijma, R. J. Floor, P. A. Jekel, D. Baker, S. J. Marrink, and D. B. Janssen, “Computationally designed libraries for rapid enzyme stabilization,” *Protein Eng. Des. Sel.*, vol. 27, no. 2, pp. 49–58, 2014.

[54] A. Wlodawer, W. Minor, Z. Dauter, and M. Jaskolski, “Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures,” *FEBS J.*, vol. 275, no. 1, pp. 1–21, 2008.

[55] J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, and D. C. Phillips, “A three-dimensional model of the myoglobin molecule obtained by x-ray analysis,” *Nature*, vol. 181, no. 4610, pp. 662–666, 1958.

[56] A. Wlodawer, W. Minor, Z. Dauter, and M. Jaskolski, “Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination,” *Febs J.*, vol. 280, no. 22, pp. 5705–5736, 2013.

[57] H. M. Berman, “Creating a community resource for protein science,” *Protein Sci.*, vol. 21, no. 11, pp. 1587–1596, 2012.

[58] G. Rhodes, *Crystallography made crystal clear*, Third edit. Academic Press, 2006.

[59] Bernhard rupp, *Biomolecular Crystallography: Principles, Practive, and Application to Structural Biology*. .

[60] J. Jancarik and S. H. Kim, “Sparse matrix sampling: a screening method for crystallization of proteins,” *J. Appl. Crystallogr.*, vol. 24, no. pt 4, pp. 409–411, 1991.

[61] N. E. Chayen and E. Saridakis, “Protein crystallization: From purified protein to diffraction-quality crystal,” *Nat. Methods*, vol. 5, no. 2, pp. 147–153, 2008.

[62] P. C. Weber, “Overview of protein crystallization methods,” *Methods Enzymol.*, vol. 276, no. 1988, pp. 13–22, 1997.

[63] N. E. Chayen, “Turning protein crystallisation from an art into a science,” *Curr. Opin. Struct. Biol.*, vol. 14, no. 5, pp. 577–583, 2004.

[64] A. McPherson and J. A. Gavira, “Introduction to protein crystallization,” *Acta Crystallogr. Sect. FStructural Biol. Commun.*, vol. 70, no. 1, pp. 2–20, 2014.

[65] D. H. Thomas, A. Rob, and D. W. Rice, “A novel dialysis procedure for the crystallization of proteins,” *Protein Eng. Des. Sel.*, vol. 2, no. 6, pp. 489–491, 1989.

[66] F. R. Salemme, “A free interface diffusion technique for the crystallization of proteins for X-ray crystallography,” *Arch. Biochem. Biophys.*, vol. 151, no. 2, pp. 533–539, 1972.

[67] J. Leng and J. B. Salmon, “Microfluidic crystallization,” *Lab Chip*, vol. 9, no. 1, pp. 24–34, 2009.

[68] G. Taylor, “The phase problem,” pp. 1881–1890, 2003.

[69] W. A. Hendrickson, “Anomalous diffraction in crystallographic phase evaluation,” *Q. Rev. Biophys.*, vol. 47, no. 1, pp. 49–93, 2014.

[70] P. D. Adams *et al.*, “PHENIX: A comprehensive Python-based system for macromolecular structure solution,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 66, no. 2, pp. 213–221, 2010.

[71] P. Emsley and K. Cowtan, “Coot: Model-building tools for molecular graphics,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 60, no. 12 I, pp. 2126–2132, 2004.

[72] Hampton Research, “Crystal Growth 101: Seeding,” *Hampt. Res.*, pp. 1–8, 2020.

[73] W. Kabsch, “XDS,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, pp. 125–132, 2010.

[74] P. R. Evans, “An introduction to data reduction: Space-group determination, scaling and intensity statistics,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 67, no. 4, pp. 282–292, 2011.

[75] P. Evans, “Scaling and assessment of data quality,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 62, no. 1, pp. 72–82, 2006.

[76] P. R. Evans and G. N. Murshudov, “How good are my data and what is the resolution?,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 69, no. 7, pp. 1204–1214, 2013.

[77] C. Vonrhein *et al.*, “Data processing and analysis with the autoPROC toolbox,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 67, no. 4, pp. 293–302, 2011.

[78] A. Vagin, “MoRDa, an automatic molecular replacement pipeline,” *29th Eur. Crystallogr. Meet.*, p. 31, 2015.

[79] A. Tan, K., Bigelow, L., Hamilton, J., Bearden, J., Joachimiak, “The crystal structure of a ATP phosphoribosyltransferase regulatory subunit/histidyl-tRNA synthetase from Bacillus halodurans C.” [Online]. Available: https://www.rcsb.org/structure/3OD1. [Accessed: 02-Apr-2021].

[80] P. V. Afonine *et al.*, “Towards automated crystallographic structure refinement with phenix.refine,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 68, no. 4, pp. 352–367, 2012.

[81] J. Painter and E. A. Merritt, “Optimal description of a protein structure in terms of multiple groups undergoing TLS motion,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 62, no. 4, pp. 439–450, 2006.

[82] E. Blanc, P. Roversi, C. Vonrhein, C. Flensburg, S. M. Lea, and G. Bricogne, “Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 60, no. 12 I, pp. 2210–2221, 2004.

[83] C. J. Williams *et al.*, “MolProbity: More and better reference data for improved all-atom structure validation,” *Protein Sci.*, vol. 27, no. 1, pp. 293–315, 2018.

[84] W. DeLano and S. Bromberg, “PyMOL User’s Guide (Original),” *DeLano Sci. LLC*, pp. 1–66, 2004.

[85] T. Schwede, J. Kopp, N. Guex, and M. C. Peitsch, “SWISS-MODEL: An automated protein homology-modeling server,” *Nucleic Acids Res.*, vol. 31, no. 13, pp. 3381–3385, 2003.

[86] D. J. Korbie and J. S. Mattick, “Touchdown PCR for increased specificity and sensitivity in PCR amplification,” *Nat. Protoc.*, vol. 3, no. 9, pp. 1452–1456, 2008.

[87] K. Diederichs and P. A. Karplus, “Improved R-factors for diffraction data analysis in macromolecular crystallography,” *Nat. Struct. Biol.*, vol. 4, no. 4, pp. 269–275, 1997.

[88] M. S. Weiss, “Global indicators of X-ray data quality research papers,” *Appl. Crystallogr.*, pp. 130–135, 2000.

[89] E. Krissinel and K. Henrick, “Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 60, no. 12 I, pp. 2256–2268, 2004.

[90] L. Holm, “DALI and the persistence of protein shape,” *Protein Sci.*, vol. 29, no. 1, pp. 128–140, 2020.

[91] B. Webb and A. Sali, “Comparative protein structure modeling using MODELLER,” *Curr. Protoc. Bioinforma.*, vol. 2016, no. June, pp. 5.6.1-5.6.37, 2016.

[92] T. C. Tan, D. Haltrich, and C. Divne, “Regioselective control of β-d-glucose oxidation by pyranose 2-oxidase is intimately coupled to conformational degeneracy,” *J. Mol. Biol.*, vol. 409, no. 4, pp. 588–600, 2011.

[93] T. Wongnate and P. Chaiyen, “The substrate oxidation mechanism of pyranose 2-oxidase and other related enzymes in the glucose-methanol-choline superfamily,” *FEBS J.*, vol. 280, no. 13, pp. 3009–3027, 2013.

Supplementary Material

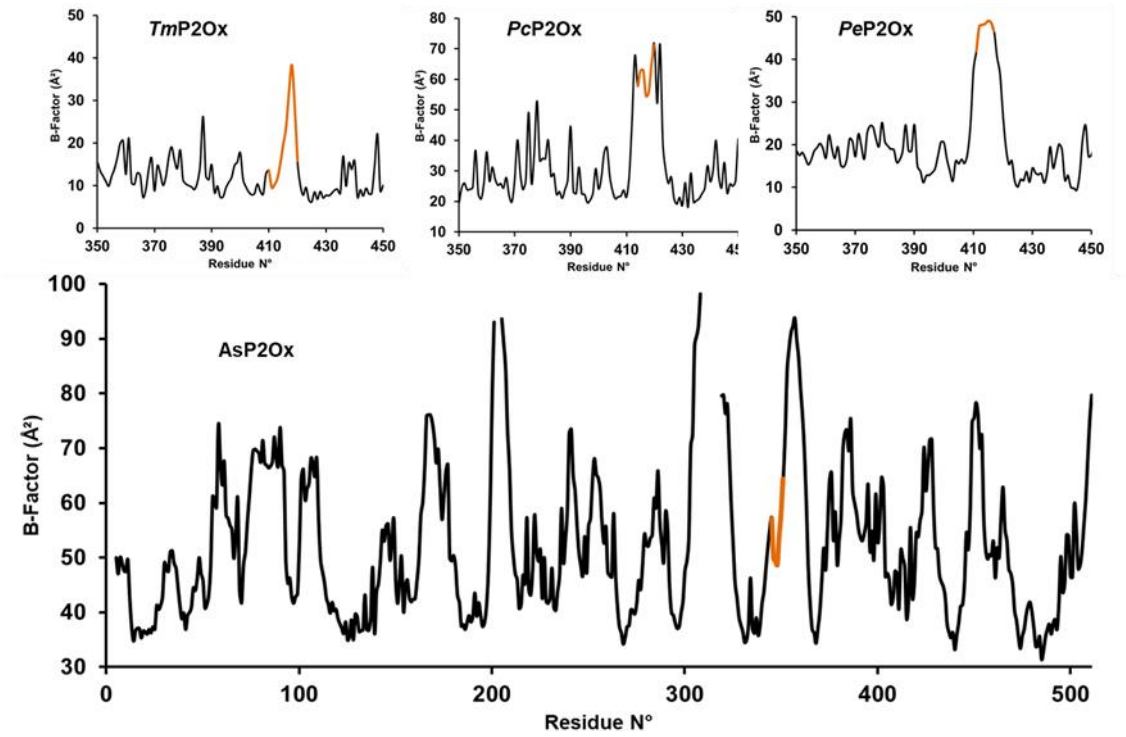

Figure 32 – Graphical representation of the B-factors variation along the residue sequence. At the top, B-factor variation from three fungal P2Ox with structural models available are represented: *Tm*P2Ox (PDB 1TT0), *Pc*P2Ox (PDB 4MIF) and *Pe*P2Ox (PDB 1TZL). At the bottom is represented the variation of temperature factors of *As*P2Ox. The sequence of residues belonging to the dynamic loops is represented in orange.

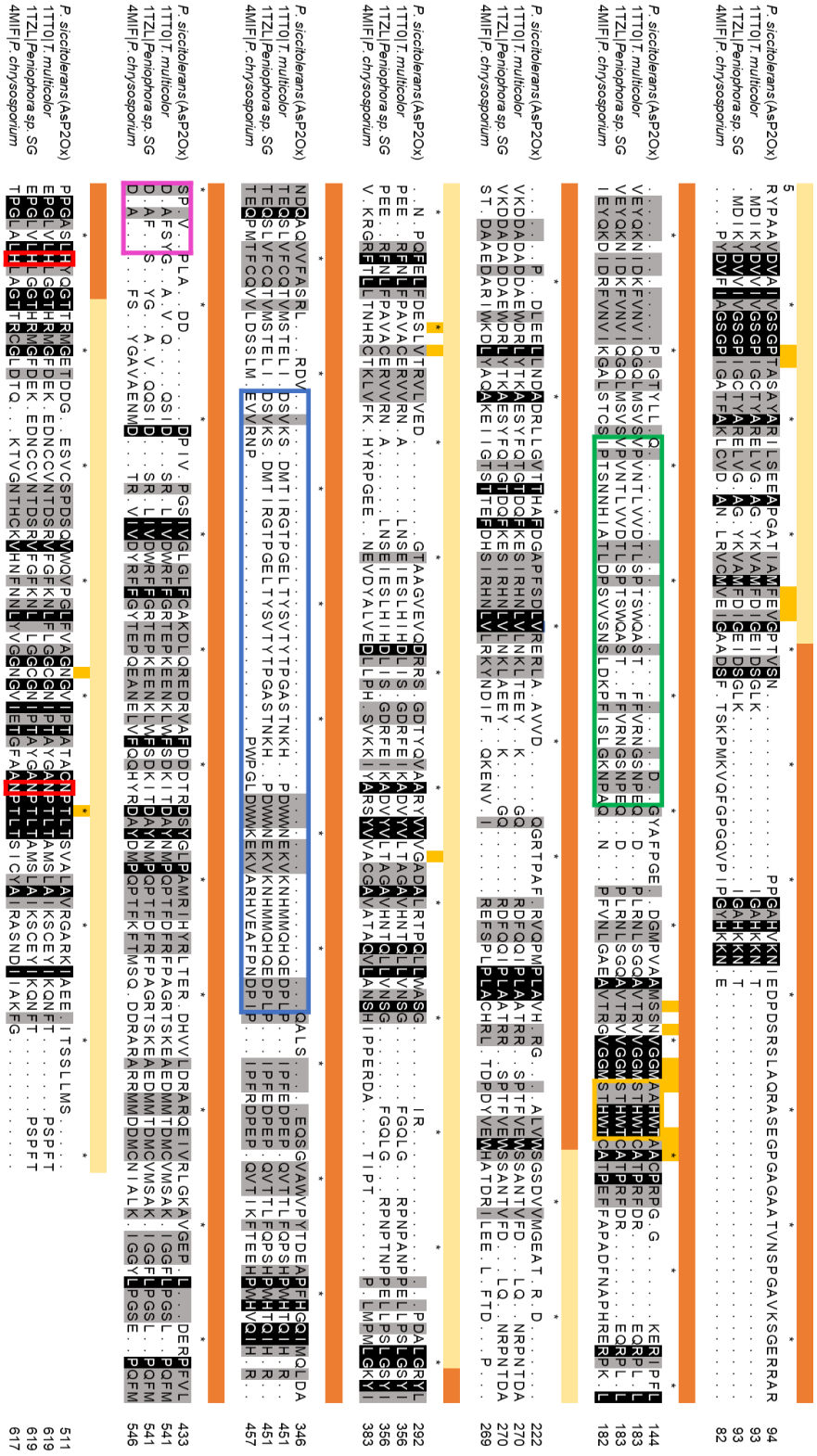

Figure 33 – Structural alignment between *As*P2Ox, *Tm*P2Ox (PDB 1TT0), *Pe*P2Ox (1TZL) and *Pc*P2Ox (PDB 4MIF). Each 10th position is marked by an asterisk (*). Strictly conserved residues are represented in a black background, while mostly conserved residues are represented in a light gray background. The green and blue boxes represent the fungal arm and the head domain, respectively. FAD binding domain is delineated by a yellow box, while the catalytic residues are outlined by a red box. The pink box represents the dynamic substrate loop gating the active site. All residues that interact with the FAD cofactor in *As*P2Ox model have a yellow marker above them. The FAD-binding and substrate domains are represented above the sequences in dark and light orange, respectively.