UAlg
UNIVERSIDADE DO ALGARVE

Universidade do Algarve

Bacterial Peroxide Forming Enzymes

Joaquim Paulo Curre Madeira

Dissertação para obtenção do Grau de Mestre em Engenharia Biológica

Trabalho efetuado sob a orientação da Professora Doutora Lígia O. Martins e coorientado pelo Professor Doutor Eduardo P. Melo

2015

Bacterial Peroxide Forming Enzymes

Declaração de autoria de trabalho

Declaro ser o autor deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

Bacterial Peroxide Forming Enzymes

Joaquim Paulo Curre Madeira

2015

The work presented in this thesis was conducted at the Instituto de Tecnologia Química e Biológica, in the Microbial and Enzymatic Technology (MET) Lab under the supervision of Professor Lígia Oliveira Martins and co-supervision of Dr. Vânia Brissos.

AGRADECIMENTOS

Quero aqui espressar o meu agradecimento a todos os que me ajudaram na elabocação desta tese.

À Professora Lígia Martins, por me ter orientado e por ter demostrado, sempre, total disponibilidade. A sua forma de me acolher no laboratório ajudou imenso no meu crescimento como pessoa e investigador, pois incitou-me a tentar, lutar e alcançar os objetivos por mim próprio. Fico-lhe também com profunda admiração pela sua dedicação ao trabalho e pela sua energia e perseverança contagiante.

À Dra. Vânia Brissos, por todo o acompanhamento ao longo do trabalho, pelas técnicas que me ensinou, pelo olhar crítico e por puxar por mim. A sua dedicação e preocupação pelo trabalho foram também uma inspiração.

Ao Professor Eduardo Melo, por ter aceite ser meu orientador, pela disponibilidade demonstrada e pela oportunidade de vir estagiar para o ITQB.

Aos que estão ou passaram pelo MET lab durante a minha estadia: À Sónia Mendes, por estar sempre disposta a ajudar mesmo quando tinha muito trabalho pela frente. Ao Diogo, pelo companheirismo e pelos bons momentos. À Lúcia, pela boa disposição e espírito que aligeiraram as longas horas de trabalho. Ao Marcin e Kamil, dois polacos cinco estrelas. E também ao Marcelo, Bruna, Antonieta e Mara, que até hoje não sei o que toma para rir tanto.

Aos meus colegas de curso. Flávio, que me influenciou a ter um pensamento mais critico; Juaci, que me mostrou que não importa as bases, o esforço é recompensado; Margarida; Raquel; Ana; João Charneca; João Estêvão; Laura; Diogo e Fábio a todos agradeço pelas saídas, pelos jantares, pelas muitas horas de estudo mas sobretudo pela amizade e companheirismo.

Ao Carlos, ao que parece sempre entreguei a tese antes de acabares a mota.

À minha família, principalmente Mãe, Pai e Irmã pelo apoio e por me proporcionarem a oportunidade de tirar um curso superior.

À minha namorada e amiga Ana, obrigado pela amizade, pelo apoio incondicional e por me aturar durante todo este período.

ABSTRACT

Lignin, after cellulose, is the most abundant organic polymer on Earth and has vital functions as a constituent of plant cell walls including structural resistance and protection against pathogens and hydrolysis. Notwithstanding lignin degradation by microbes represents a key-step in the completion of the carbon cycle in land ecosystems. The most well-known microorganisms involved in lignin degradation are fungi using oxidoreductases such as laccases, high redox peroxidases and peroxide forming enzymes. Bacterial lignin degradation is much less characterized and might allow the development of new cost-affordable bioprocesses for lignin depolymerisation, considering the relative easiness of bacterial enzyme production and engineering systems.

This work focuses on *in silico* screening, cloning, expression and characterisation of bacterial lignin-degrading auxiliary enzymes, more specifically, H_2O_2 forming enzymes such as aryl-alcohol oxidase, pyranose-2-oxidases, glyoxal oxidases or galactose oxidases with the intent to advance the scope of our understanding of lignin degradation in bacteria.

Here we report the successful cloning of the first bacterial pyranose-2-oxidase and also the heterologous cloning, production and preliminary biochemical characterization of a bacterial galactose oxidase, both from *Arthrobacter siccitolerans*. The recombinant pyranose oxidase revealed activity towards D-glucose and the typical ultra-violet/visible spectrum of a flavoenzyme. The recombinant galactose oxidase showed an optimal temperature between 25-35°C and an optimal pH of 8.0. The *A. siccitolerans* galactose oxidase is able to oxidize the typical substrates such as D-galactose, D-raffinose, lactose and glycerol, but also exhibited activity towards L-arabinose which has not been reported for other galactose oxidases.

It is expected that this work will open new perspectives for the understanding of the structure-function relationships of bacterial hydrogen peroxide producing enzymes and optimization of multienzymatic systems, along with bacterial dye-decolorizing peroxidases and bacterial laccases, recently identified in the laboratory, for the set-up of lignin depolymerisation and valorisation.

Key words: Lignin depolymerisation; bacterial lignin-degrading auxiliary enzymes; aryl-alcohol oxidase; pyranose-2-oxidase; glyoxal oxidase; galactose oxidase

RESUMO

A biomassa vegetal é composta principalmente por celulose, hemicelulose e lenhina. A lenhina é, a seguir à celulose, o polímero orgânico mais abundante na Terra e desempenha funções fundamentais como constituinte da parede das células de plantas, na sua rigidez, impermeabilidade, bem como na proteção contra agentes invasores e as suas eventuais enzimas degradativas. A lenhina é extremamente resistente tanto à degradação física como biológica, devido à sua complexa estrutura polimérica. No entanto, a degradação microbiana da lenhina representa um processo fundamental para a conclusão do ciclo de carbono em ecossistemas terrestres. Presentemente a maior parte da lenhina que advém de processos de produção de pasta de papel, por exemplo, não é isolada nem tem qualquer aplicação tecnológica mas apenas queimada para produção de energia. Com crescente interesse no uso de biomassa linhocelulósica renovável, para a produção de compostos químicos, materiais e combustíveis, é reconhecido o enorme potencial da lenhina como um importante recurso para a produção de produtos de valor acrescentado como compostos químicos aromáticos e polímeros. Por outro lado, o maior desafio para a utilização industrial plena da biomassa linhocelulósica é a eficiente separação da lenhina dos polissacáridos, a fim de permitir a sua total utilização. Tem-se assim investido no desenvolvimento de variados processos, físico-químicos e biológicos, que permitam a eficiente separação da lenhina dos restantes componentes da biomassa vegetal. Considera-se que a biotecnologia baseada em organismos que degradam a lenhina e nas suas enzimas, podem contribuir para uma utilização mais eficiente e ambientalmente saudável da biomassa linhocelulósica como matéria-prima para futuras biorrefinarias.

Os microrganismos mais ativos, identificados até hoje, envolvidos na degradação da lenhina são fungos, mais especificamente os denominados fungos da “podridão branca”. Este nome deriva do facto de estes microrganismos acederem e degradarem a lenhina sem removerem os polissacáridos, resultando num material final branco. O mecanismo de atuação destes microrganismos envolve um sistema multienzimático com enzimas oxidativas; lacases, peroxidases com elevado potencial redox e enzimas formadoras de peróxido. Apesar dos vários estudos de degradação da lenhina por fungos, reportados na literatura, não existe até à data nenhum processo biocatalítico comercial para a despolimerização da lenhina. Tal pode dever-se, em parte, aos problemas práticos de expressão de proteínas fúngicas e à sua manipulação genética e também devido ao facto destes não serem estáveis sob condições ambientais extremas, não sendo de menosprezar complexidade da lenhina como um substrato. A este respeito faz sentido recordar a grande adaptabilidade ambiental e versatilidade bioquímica

exibida por bactérias. Apesar dos processos de degradação da lenhina por bactérias estarem muito menos caracterizados em comparação com os dos fungos, o seu estudo pode ser importante para o desenvolvimento de processos biotecnológicos que visem a despolimerização da lenhina. A recente identificação de lacases e peroxidases com elevado potencial redox em bactérias, apontam para que estas também possam possuir processos de despolimerização da lenhina semelhantes aos mecanismos fúngicos descritos. No que diz respeito às enzimas "auxiliares" de degradação da lenhina, mais especificamente às enzimas formadoras de peróxido de hidrogénio, existem até à data apenas quatro enzimas identificadas em fungos ("aryl-alcohol" oxidase; piranose-2-oxidase; glioxal oxidase e galactose oxidase) e em bactérias apenas uma enzima com semelhanças com a galactose oxidase fúngica. A enzima "aryl-alcohol" oxidase catalisa a desidrogenação oxidante de uma ampla gama de álcoois primários insaturados, com a produção de peróxido de hidrogénio. A enzima piranose oxidase catalisa a oxidação de várias aldopiranoses na posição C-2 para obter as correspondentes 2-cetoaldoses. A enzima glioxal oxidase catalisa a oxidação de aldeídos a ácidos carboxílicos, e a enzima galactose oxidase, da mesma família, denominada de "radical copper oxidases", catalisa a oxidação estéreoespecífica de D-isómeros de uma gama álcoois primários, incluindo D-galactose e polissacáridos com D-galactose na sua extremidade não redutora, levando à produção do aldeído correspondente. Durante as reações de oxidação descritas os eletrões "extraídos" aos diferentes substratos são transferidos para o oxigénio molecular, resultando na formação de peróxido de hidrogénio.

Neste trabalho efetuámos um rastreio *in silico* das enzimas formadoras de peróxido acima descritas em genomas bacterianos que estão disponíveis. Selecionámos sete possíveis genes codificantes para as enzimas glioxal oxidase e galactose oxidase e um possível gene codificante para a enzima piranose oxidase. Não identificámos nenhum gene codificante para a enzima "aryl-alcohol" oxidase. De seguida foram feitas várias tentativas de clonagem e expressão heteróloga dos genes selecionados em várias estirpes de *Escherichia coli*. Como resultado desse esforço reportamos a primeira clonagem de uma piranose oxidase bacteriana e também a clonagem e caracterização primária de uma galactose oxidase bacteriana, ambas as clonadas da bactéria *Arthrobacter siccitolerans*. A enzima recombinante piranose oxidase revelou atividade para com o substrato D-glucose e também as assinaturas espectroscópicas de uma flavoproteína.

A enzima recombinante galactose oxidase foi recuperada a partir de corpos de inclusão através de um protocolo de desenrolamento e enrolamento. Verificou-se no entanto uma forte propensão para agregação. Para efetuar a sua caraterização primária foi utilizada uma

preparação de proteína parcialmente purificada. Esta enzima exibiu uma temperatura ótima de reação de entre 25 a 35°C e um pH ótimo de 8,0. Verificámos a sua capacidade para oxidar D-galactose, para o qual apresenta um K_m de 30 ± 3 mM, assim como outros substratos típicos, D-raffinose, lactose e glycerol, exibindo também atividade para o açúcar L-arabinose que, até à data não tinha sido reportado para nenhuma outra galactose oxidase. A atividade detetada para os açúcares D-raffinose e lactose indica que pode oxidar derivados de galactose com substituições no carbono-1, como também foi demonstrado em outras galactose oxidases isoladas a partir de fungos. A enzima não demonstrou atividade para com o substrato D-glucose, indicando que a configuração do grupo hidroxilo no carbono-4 é importante para a atividade da enzima e que esta é específica para o substrato galactose e derivados, o que está em boa concordância com os relatórios de outras galactose oxidases previamente estudadas. O elevado valor de K_m para o substrato D-galactose também se verificou para os outros substratos estudados. Este facto pode ser uma consequência da ampla especificidade da enzima, resultando num centro ativo capaz de ligar uma variedade de substratos diferentes Por fim, estudos de estabilidade revelaram que a enzima tem um tempo de meia vida aos 40°C de 1 h. Espera-se que este trabalho possa contribuir para o futuro estudo e elucidação de características estruturais e funcionais de enzimas oxidativas envolvidas em sistemas multienzimáticos de degradação da lenhina com interesse fundamental e aplicado.

Palavras-chave: Despolimerização da lenhina; enzimas “auxiliares” de degradação da lenhina; “aryl-alcohol” oxidase; piranose-2-oxidase; glioxal oxidase; galactose oxidase

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

a.a. – Amino acid
AAO – Aryl-alcohol oxidases
ABTS – 2,2'-azino bis (3- ethylbenzthiazoline-6-sulfonic acid)
APPL – Acid-precipitable polymeric lignin
APS – Ammonium persulfate
BCA – trichloroacetic acid/bicinchoninic acid
BLAST – Basic Local Alignment Search Tool
BSA – Bovine Serum Albumin
DMSO – Dimethyl sulfoxide
dNTP – 2'-deoxynucleotide 5'-triphosphate
DyP – Dye-decolorizing peroxidase
EDTA – Ethylenediaminetetraacetic acid
FAD – Flavin adenine dinucleotide
FOLy – Fungal oxidative lignin enzymes
Gal-GalNAc – D-galactose-β-(1-3)-N-acetyl-D-galactosamine
GalOx – Galactose oxidase
GdnHCl – Guanidine hydrochloride
GlyOx – Glyoxal oxidase
GMC – Glucose-methanol-choline
GSH - Reduced glutathione
GSSG - Oxidized glutathione
HRP – Horseradish peroxidase
IPTG – Isopropyl β-D-1-thiogalactopyranoside
kDa – kilodalton
Lac – Laccase
LB – Luria-Bertani
LDA – Lignin-degrading auxiliary enzymes
LEM – Lignin-modifying enzymes
LiP – Lignin peroxidase
LLCT – Ligand to ligand charge transfer
LMCT – Ligand to metal charge transfer
MnP – Manganese peroxidase
NAD(H) – Nicotinamide adenine dinucleotide (reduced)

NGS – Next generation sequencing

OD_{600nm} – Optical density at 600 nm

ORF – Open Reading Frame

PCR – Polymerase chain reaction

POx – Pyranose oxidase

RCO – Radical copper oxidase

SDS-PAGE – Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SOB – Super Optimal Broth

TEMED – N, N, N', N'-tetramethylethylenediamine

Tris – Tris (hidroximetil) aminometano

UV – Ultra-violet

Vis – Visible

VP – Versatile peroxidase

WT – Wild-type

1. INTRODUCTION

1.1. Lignin as a natural Renewable Resource

Plant biomass is composed by a number of chemical components, mainly polymeric, such as cellulose, hemicellulose and lignin as well as smaller molecules, such as lipids (fatty acids, terpenoids) and phenolic compounds (simple phenolics, flavonoids) , among other thousands of different "extractives" (Hinterstoisser *et al*., 2000). In 1838, Anselme Payen reported on the intricate composition of wood and referred to a carbon-rich substance as the "encrusting material" which embedded cellulose in wood. Later Schulze in 1865 defined this encrusting material as "lignin" (Sjöström, 1993). The word lignin derives from the Latin word lignum, which means wood.

Lignin is the second most abundant organic polymer on Earth, after cellulose, comprising around 15-25% of the total carbon fixed by photosynthesis in land ecosystems. The ability to synthesize lignin has been essential in the evolutionary adaptation of plants from aquatic to terrestrial environments. Lignin is a vital constituent of the cell wall of vascular plants and is mainly present in the middle lamella providing structural integrity and stiffness and strength of the stem. Its main purpose is to waterproof the cell wall, enabling transport of water and solutes through the vascular system. Lignin has also an additional important function in the plant's natural defence against pathogens by impeding penetration of destructive enzymes through the cell wall (Sarkanen and Ludwig 1971). Notwithstanding lignin degrading microbes are perceived to have evolved simultaneously with the colonization of vascular plants to land ecosystems, around 400 million years ago (Taylor and Osborne, 1996). Degradation of lignin by microorganisms represents a key step in the completion of the carbon cycle on Earth, since the removal of the lignin barrier enables the subsequent microbial degradation of plant wall polysaccharides (Martínez *et al*, 2005).

Lignin structure is formed by an amorphous optically inactive phenylpropanoid polymer and its complexity makes it highly recalcitrant towards both chemical and biological degradation (Hammel and Cullen, 2008). Lignin is synthesized without defined repetitive units, through the arrangement of three main p-hydroxycinnamyl alcohols or monolignols (coniferyl, p-coumaryl and sinapyl alcohol) and their acylated forms (Ralph *et al*., 2004; Martínez *et al*., 2008).

Even though the building blocks of lignin are phenolic compounds, the resulting polymer is basically non-phenolic. In fact it has been proposed that in the last step of lignin biosynthesis, plant peroxidases (and possibly also laccases) oxidize monolignols to their phenolic radicals (Higuchi 1997). As a result of the prevalence of certain radical resonant forms, as well as the different stability of the coupling products, ether linkages between the phenolic position (C4) and the side-chain (or aromatic ring) β-carbon of the p-hydroxyphenylpropenoid precursors (Fig. 1.2., substructure A, B and D) are strongly predominant in the polymer, resulting in its non-phenolic nature (Ruiz-Dueñas and Martínez, 2009).

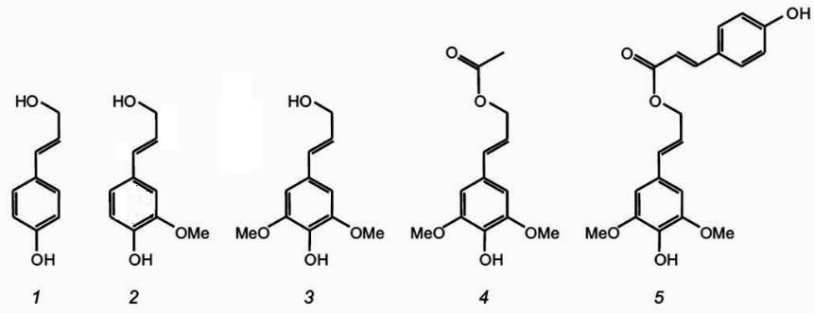


Figure 1.1. Lignin precursors or monolignols. Classical: p-coumaryl (1), coniferyl (2), sinapyl (3). Acylated: derived from sinapyl alcohol γ-esterified with acetic (4) and p-coumaric acid (5) (adapted from Ruiz-Dueñas and Martínez, (2009)).

There are presently an estimated 50 million tonnes of lignin available world-wide derived in pulping processes from pulp and paper industries however, the majority of this material is not isolated or used for technological purposes but burned onsite to provide steam for heat and power production (Gosselink *et al*, 2004). With the escalating interest in the use of renewable lignocellulosic plant biomass and residues for production of chemicals, materials, and biofuels, the potential of lignin was recognized as a source for the production of value-added compounds and polymers. Moreover, the major challenge for the industrial utilization of plant biomass is the efficient separation of lignin from the plant polysaccharides in order to enable the full use of the respective components. A good example is the bioethanol production from lignocellulosic biomass, where the lignin barrier limits the accessibility of microorganism to fermentable sugars (derived from cellulose and hemicelluloses). In the manufacture of cellulose pulp, disruption of this barrier is presently mainly achieved through traditional chemico-physical methods, including high-temperature and strong chemical reagents (Salvachúa *et al*, 2011), that leads to the breaking down of the plant wall and wood fibres separation (Sixta, 2006).

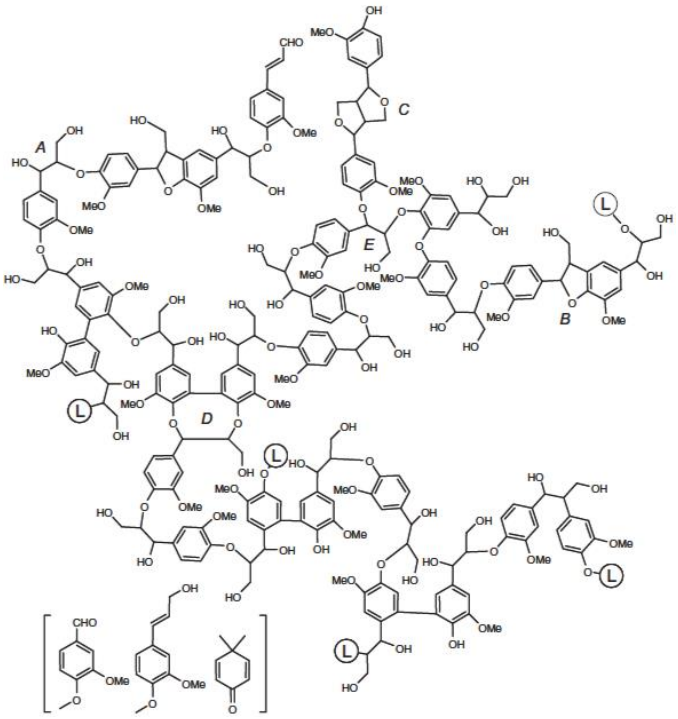


Figure 1.2. Lignin structure. Substructures: β-O-4′ (A); phenylcoumaran (B); pinoresinol (C) and dibenzodioxocin (D). L-containing circles indicate linkages to additional lignin chains. Brackets indicate other minor structures, such as vanillin, coniferyl alcohol and dimethylcyclohexadienone-type units, and the latter in new spirodienone substructures (adapted from Ruiz-Dueñas and Martínez, (2009)).

In the future, biorefinery processes are expected to lead to the extraction of high-value chemicals present in the biomass, such as high-value nutraceuticals, fragrances, food-related products, flavouring agents, and other fine chemicals, while plant polysaccharides and lignin are processed into feedstocks for bio-derivate materials, bulk chemicals, and fuels (Ragauskas *et al*. 2006; Ruiz-Dueñas and Martínez, 2009). Biotechnology based on lignin-degrading organisms and their enzymes can contribute to a more efficient and environmentally sound use of lignocelluloses as raw materials for the future biorefineries.

1.2. Biodegradation of lignin

Even knowing that the recalcitrance of lignin have evolved naturally, selected microorganisms have developed enzymatic mechanisms to its depolymerisation. The most active microorganisms involved in lignin degradation identified so far are fungi. In that light, studies have focussed primarily on the breakdown of lignin by white-rot fungi which are able to access and degrade lignin without removing wood polysaccharides (Hatakka, 1994; Sanchez 2009). A different group of basidiomycetes called brown-rot fungi act trough a different

mechanism than white-rot fungi and instead degrade wood polysaccharides, without a previous total breaking of lignin, resulting in a final brown material consisting of modified lignin (Martínez *et al*., 2011; Yelle *et al*., 2011). White-rot fungi have been described as the most efficient lignin degraders in nature and are thought to play a key role in carbon recycling on Earth exposing the plants polyssacharides for microbial degradation. While these latter are degraded mostly via hydrolytic enzymatic mechanisms, lignin degradation uses a unique mechanism involving oxidative enzymes, through a mechanism termed as "enzymatic combustion" (Kirk and Farrel, 1987).

Different combinations of extracellular oxidative lignin-modifying enzymes (LEMs) are used by white-rot fungi to generate highly reactive nonspecific free radicals, which initiate lignin depolymerisation by cleavage of carbon-carbon and ether inter-unit bonds. There are two main classes of LEMs: multicopper oxidases with laccase activity (Lac) and high-redox peroxidases. Laccases are a part of the large multicopper oxidase family of enzymes that catalyse the four-electron reduction of oxygen to water (at the T2-T3 trinuclear Cu centres) by the sequential one-electron uptake from a suitable reducing substrate (at the T1 mononuclear copper centre) (Solomon *et al*. 1996; Stoj and Kosman 2005). Most of the known laccases have fungal (e.g. white-rot fungi) or plant origins. However, many laccases have been isolated and also identified from bacteria in the last decade (Giardina *et al*. 2010, Martins *et al*. 2015). Ligninolytic peroxidases, lignin peroxidase (LiP), manganese peroxidase (MnP) and versatile peroxidase (VP) harbouring the highest redox potential among peroxidases, are haem peroxidases catalysing the H_2O_2-dependent oxidation of a variety of lignin related substrates (Ruiz-Dueñas and Martínez, 2009, Martinez *et al* 2009).

Additional oxidative enzymes named lignin-degrading auxiliary (LDA) enzymes act in concert with the LEMs facilitating lignin degradation. Most enzymes classified as LDAs are H_2O_2 generators. These include radical copper oxidases (glyoxal oxidase (GlyOx), galactose oxidase (GalOx)), aryl-alcohol oxidases (AAO) and pyranose-2-oxidase (POx) (Martinez *et al*. 2005). A relatively recent study has classified the enzymes potentially involved in lignin catabolism into sequence-based families and integrated them into a newly developed database, designated fungal oxidative lignin enzymes (FOLy) (Levasseur *et al*, 2008).

Table 1.1. Enzymes involved in lignin degradation (adapted from El-Enshasy, *et al*, (2013)).

Enzyme	EC No.	Systematic Name	Cofactor	Catalysed Reaction	Comments
Lignin oxidases and peroxidases					
Lignin peroxidase	1.11.1.14	Diarylpropane: oxygen, hydrogen-peroxide oxireductase	Heme	1,2-bis(3,4-dimethoxyphenyl) propane-1,3-diol + H_2O_2 ↔ 3,4-dimethoxybenzaldehyde + 1-(3,4-dimethoxyphenyl) ethane-1,2-diol + H_2O	Oxidative cleavage of C-C and C-O-C bonds.
Manganese peroxidase	1.11.1.13	Mn(II): hydrogen-peroxide oxidoreductase	Heme	2 Mn^{2+} + 2 H^+ + H_2O_2 ↔ 2 Mn^{3+} + 2 H_2O	Oxidative degradation of lignin
Laccase	1.10.3.2	Benzenediol: oxygen oxidoreductase	Copper	4 benzenediol + O_2 ↔ 4 benzosemiquinone + 2 H_2O	The formed semiquinone can further react either enzymatically or spontaneously
Versatile peroxidase	1.11.1.16	Reactive black-5: hydrogen-peroxide oxidoreductase	Heme	Reactive Black 5 (or donor) + H_2O_2 ↔ oxidized Reactive Black 5 (or oxidized donor) + 2 H_2O	Combines substrate specificities of both LiP and MnP; can oxidize phenols and hydroquinones as well
Lignin-degrading auxiliary enzymes					
Glyoxal oxidase	1.1.3.-	-	Copper	RCHO + O_2 ↔ RCOOH + H_2O	Characterized by its radical-copper complex acting as a two-electron redox active site

Enzyme	EC No.	Systematic Name	Cofactor	Catalysed Reaction	Comments
Galactose oxidase	1.1.3.9	D-galactose: oxygen 6-oxidoreductase	Copper	D-galactose + O_2 ↔ D-galacto-hexodialdose + H_2O_2	Contains a free radical-coupled copper active site (a characteristic shared with glyoxal oxidase)
Pyranose oxidase	1.1.3.10	Pyranose: oxygen 2-oxidoreductase	FAD	D-glucose + O_2 ↔ 2-dehydro-D-glucose + H_2O_2	Can oxidize D-xylose, L-sorbose and D-glucono-1,5-lactone
Vanillyl-alcohol oxidase	1.1.3.38	Vanillyl alcohol: oxygen oxidoreductase	FAD	Vanillyl alcohol + O_2 ↔ vanillin + H_2O_2	4-hydroxybenzyl alcohols and 4-hydroxybenzylamines are good substrates
Cellobiose dehydrogenase	1.1.99.18	Cellobiose: acceptor 1-oxidoreductase	FAD Heme	Cellobiose + Aceptor ↔ Cellobiono-1,5-lactone + Reduced acceptor	Acceptors can be either two-electro oxidants; e.g., redox dyes benzoquinones, and molecular oxygen; or one-electron oxidants, e.g. semiquinone species, iron(II) complexes, and cytochrome c
Glucose oxidase	1.1.3.4	Beta-D-glucose: oxygen 1-oxidoreductase	FAD	Beta-D-glucose + O_2 ↔ D-glucono-1,5-lactone + H_2O_2	Used with peroxidases coupled reactions
p-Benzoquinone reductase	1.6.5.6	NADPH: *p*-benzoquinone oxidoreductase	NADPH	NADPH + H^+ + : *p*-benzoquinone ↔ $NADP^+$ + hydroquinone	Accepts quinones or similar compounds as oxidants

Apparently, there are two major fungal strategies based on H_2O_2 to overcome the lignin barrier. A strategy followed by white-rot fungi and a strategy followed by the brown-rot fungi. *Phanerochaete chrysosporium* and *Pleurotus eryngii* are two of the most investigated white-rot fungi that evolved different extracellular enzymatic machineries to degrade lignin. *P. chrysosporium* secretes LiP, MnP (Hammel and Cullen, 2008) and GlyOx (Kersten, 1990), whereas *P. eryngii* secretes MnP, VP (Ruiz-Dueñas *et al*., 2009), together with Lac (Muñoz *et al*., 1997) and AAO (Guillén *et al*., 1990). In these strategies H_2O_2 formed by the LDAs "fuels" the lignolytic peroxidases as H_2O_2 is the oxidizing substrate of peroxidases. Nonetheless these enzymes, due to their size, cannot penetrate the compact cell-wall in sound wood, therefore as mentioned above, small chemical oxidizers, aromatic radicals, metal cations and activated oxygen species including H_2O_2 are proposed to act as the precursor of highly reactive OH· and are most likely involved in the initial stages of fungal decay of wood (Evans *et al*., 1994).

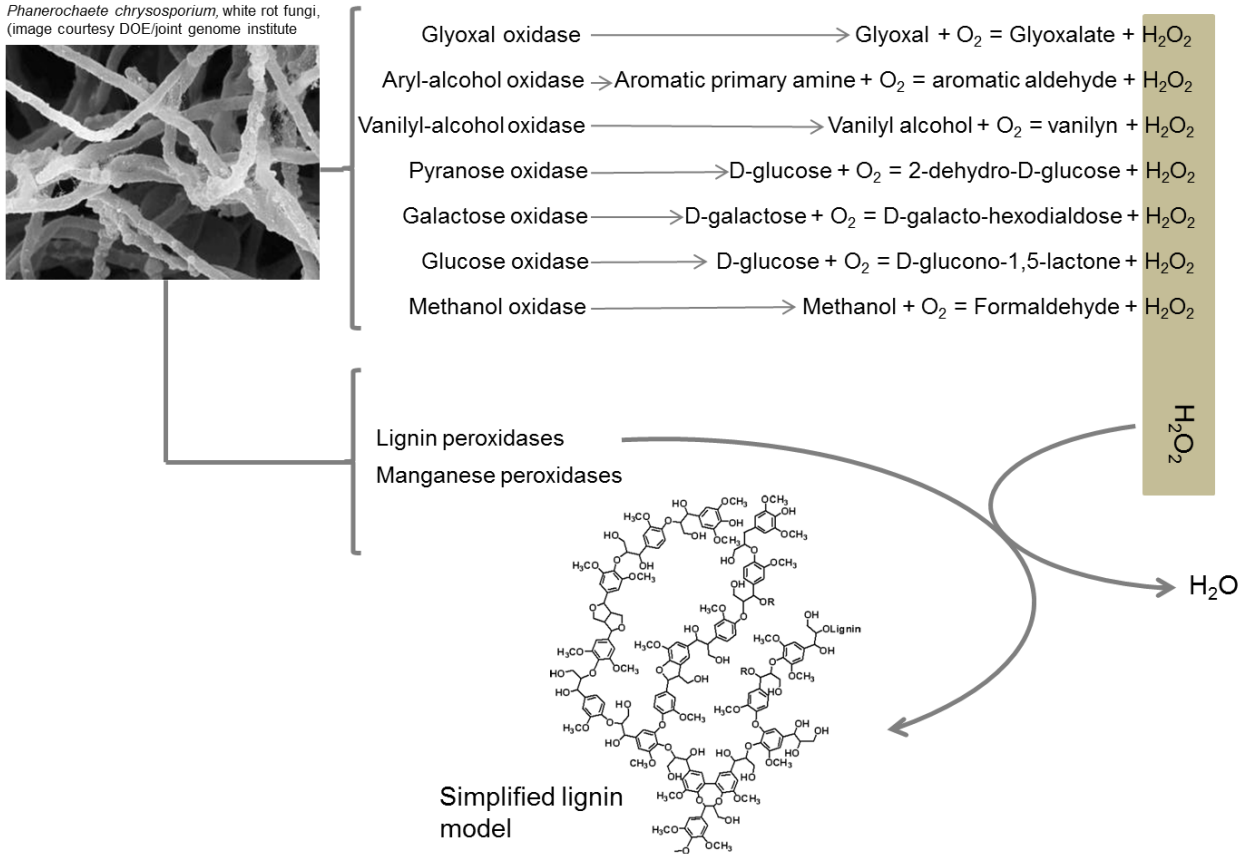


Figure 1.3. Extracellular enzymes involved in lignin degradation in *Phanerochaete chrysosporium*. The fungus is able to excrete a large array of oxidases in the extracellular medium. These enzymes generate hydrogen peroxide, which is required for peroxidase activity to cleave lignin linkages (adapted from Morel *et al*. (2009)).

The strategy adopted by brown-rot fungi is still poorly understood. As mentioned before, these fungi are able to access and degrade wood polysaccharides without removing the major part of the lignin present. However it seems now that the partial lignin degradation achieved by these microorganism relates to the combination of extracellular Fe^{2+} (oxidized to Fe^{3+} by hydroquinones), H_2O_2 (formed by alcohol oxidases) and the highly reactive hydroxyl free radical generated by the Fenton reaction ($Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + H_2O + OH\cdot$) which are small enough to attack the wood cell-wall and cleave lignin (Hammel *et al*. 2002).

Fungal lignin degradation has been studied since the mid-1980s, however, there is yet no available commercial biocatalytic process for lignin depolymerisation. This may relate to the practical challenges of fungal protein expression and genetic manipulation, e.g. fungi are not stable under extreme environmental and substrate conditions (Hatakka, 1994), and also due to the complexity of lignin as a substrate. Moreover, the identity and properties of the full systems involved in lignin metabolism (enzymes and mediators) remains to be fully elucidated and therefore assessment and exploitation of new biological diversity is expected to allow to further understanding the biocatalytic mechanisms as well as to identify new or alternative molecular approaches involved in lignin metabolism in the environment.

Bacteria are much less well characterized in what concerns lignin metabolism but deserve to be studied for their lignolytic potential considering their immense environmental adaptability and biochemical versatility. Bacteria may interact differently with the lignin polymer as compared to fungi and use the lignin itself as a substrate i.e. as an aromatic carbon source, considering the high number of aromatic compounds metabolic pathways recognized in bacteria, rather than carry out its rapid degradation to access the sugar-based cellulose and hemicellulose substrates as it is believed to happen with many fungal systems (Bianchetti *et al*., 2013).

In this context researchers are keened to advance the scope of our understanding of lignin degradation by bacteria. Interestingly, bacteria seem to exhibit different reactivity towards lignin as compared to fungi, based on differences in the resulting product profile. Various genera of bacteria have been reported to be involved in the metabolism of lignin and showed to release $^{14}C\text{-}CO_2$ from labelled lignin indicating the existence of total mineralization of the polymer (Zimmermann, 1990). Other studies have shown that bacteria also have the capacity to catabolize lignin non-phenolic related-compounds (Pellinen *et al*., 1984).

The bacterial strains identified so far exhibiting activity for lignin degradation fall into three main classes: actinomycetes, α-proteobacteria, and γ-proteobacteria (Bugg *et al*., 2010). Actinomycetes (commonly found in soil) were the first to be reported to react with lignin to both solubilize and produce a high molecular weight acid-precipitable polymeric lignin (APPL) (Kirby, 2005). The first molecular information on formation of bacterial APPL reported led to the isolation of a secreted heme peroxidase from *Streptomyces viridosporus* T7A, indicating that this bacteria, like white-rot fungi, may possess a set of oxidative enzymes involved in lignin metabolism and degradation. Lignolytic activity of *S. viridosporus* T7A was performed and found to be highly dependent on the presence of hydrogen peroxide (Ahmad *et al*, 2010). The degradation of lignin was also acknowledged in well-known and characterized bacterial aromatic degraders, such as *Pseudomonas putida* mt-2, a γ-proteobacteria, and *Rhodococcus jostii* RHA1, a actinobacteria, which is concordant, to the fact that lignin is the ultimate source for much of the aromatic material present in soil (Ahmad *et al*, 2010). The ability to react with lignin was associated in these strains with the presence of dye-decolorizing peroxidases (DyPs), as well as lacases. DyPs belong to a new family of heme peroxidases showing a different overall fold as compared with classic peroxidases and contain a distal aspartate residue in place of a histidine close to the haem moiety (Sugano *et al*., 2000). Additionally, they have often been found responsible for the oxidation of dyes and other aromatic xenobiotics (Sugano, 2009). These enzymes have additionally the potential to be used in biological decolourization processes, which has a high environmental and economic impact since around 800,000 t of synthetic dyes are produced annually of which 10-15% are released into effluents (Santos *et al*., 2014).

Advances in DNA sequencing with next generation sequencing (NGS) methodologies have revolutionized de novo bacterial genome sequencing and as a result the oxidative systems of lignin-reactive bacteria are currently being identified for further characterization. Upon screening of several different soil bacteria and ranking both APPL formation and extracellular peroxidase activity, *Amycolatopsis* sp. 75iv2 was identified as a highly active bacterial strain (Brown *et al*., 2011; Antai and Crawford, 1981). Using the Illumina platform, a rapid *de novo* genome sequencing and assembling of *Amycolatopsis* sp. 75iv2 revealed the presence of several potential lignin oxidative systems, such as secreted laccases, peroxidases, and peroxide generating enzymes (Brown *et al*., 2011; Davis *et al*., 2012). Another strategy for lignin degradation was observed in *Sphingobium* sp. SYK-6 a α-proteobacteria where production of

unique and specific enzymes, such as β-etherases, O-demethylases, and ring fission dioxygenases, was correlated with lignin catabolism. *Sphingobium* sp. SYK-6 is able to grow on several low-molecular-weight lignin compounds as a sole source of carbon and energy (Sonoki *et al*., 2002). An intricate lignin metabolic pathway was proposed, were cleavage of β-aryl ether, the most abundant lignin linkage (approximately 50%) occurs and seems to represent the most important step in the degradative scheme (Reiter *et al*., 2013). Overall, bacteria seem to share similar mechanisms as those exhibited by fungi and also additional strategies, mostly unknown, to process lignin, given the recognized span of bacterial enzymes and pathways that metabolize, functionalize or degrade aromatic compounds.

Certainly the exploitation of microbial biodiversity in new environments will help to extend our scope of understanding the biological lignin metabolisms. Recent screenings in unusual environments where active biomass degradation occurs allowed the identification of many new bacterial species involved in lignin degradation (Taylor *et al*., 2012). The study of these strains is expected to reveal the presence of new mechanisms involved in lignin biological reactivity. Additionally, the increased availability of microbial (meta)genomes from sources where biomass is rapidly degraded, such as wood-feeding insects (Warnecke *et al*., 2007; Shi *et al*., 2010); the rumen of cows (Hess *et al*., 2011); or active soils (Daniel, 2005) will allow to assess a greater diversity at the genetic level. A focus in anaerobic or microaerobic environments (Kato, 1998; Kajikawa, 2000) may be especially interesting as canonical lignin degradation involves peroxide- or oxygen-dependent enzymes. Indeed, the insufficient knowledge of the suite of genes involved in lignin modification creates roadblocks in functional annotation of (meta)genomes by sequence homology alone (Warnecke *et al*., 2007), however physiological (Geib *et al*., 2008) and structural studies (Mansfield *et al*., 2012) may help in the identification of new chemical transformations on lignin as a result of microbial metabolism. Comparative genomic studies of microbes isolated from environments such as rainforest soil (Huang *et al*., 2013) or forest areas (Eastwood *et al*., 2013) are also expected to contribute not only to improve our general understanding of the function of various members within microbial consortia (Brown and Chang, 2014) but also the identification of the bacterial consortium that are most effective in lignin degradation (Wang *et al*., 2013).

1.3. Lignin-Degrading Auxiliary Enzymes- H_2O_2 forming enzymes

A variety of auxiliary enzymes have been characterized in different fungi that have a supporting role in lignin degradation, by providing H_2O_2 required for the ligninolytic peroxidases. There are increasing evidences that bacterial auxiliary enzymes also play a pivotal role in the degradation of lignin, although to date only a galactose like radical copper oxidase has been identified from bacterial sources, *Streptomyces coelicolor* A3(2) (Whittaker and Whittaker, 2006).

1.3.1. Aryl-alcohol oxidases (AAO)

(aromatic primary alcohol + O_2 $\rightleftharpoons$ aromatic aldehyde + H_2O_2)

AAOs are extracellular flavin adenine dinucleotide (FAD)-containing enzymes first identified in *Tetrames versicolor* but later in other white-rot fungi belonging to *Pleurotus* and *Bjerkendera* genera (Gutiérrez *et al*., 1994; Muheim *et al*., 1990, Varela *et al*., 2000b). In these fungal species, H_2O_2 generation is coupled to the redox cycling of aromatic fungal metabolites such as *p*-anisaldehyde (Guillén *et al*., 1992; Marzullo *et al*., 1995). AAOs use other substrates besides *p*-anisyl such as benzyl, cinnamyl, naphthyl, and polyunsaturated primary alcohols (Guillén *et al*., 1992). These enzymes are structurally related to the glucose-methanol-choline (GMC) family of oxidoreductases, such as glucose oxidase, pyranose-2-oxidase and cellobiose dehydrogenase. Structural studies performed with *P. eryngii* AAO allowed the identification of two conserved residues, His502 and His546, involved in catalysis. A cyclic system of H_2O_2 production involving the AAO and two dehydrogenases was proposed where by the action of AAOs on a number of aromatic alcohols, acids and aldehydes derived from fungal metabolism and lignin degradation, H_2O_2 is produced extracellularly (Varela *et al*., 1999). Subsequently intracellular dehydrogenases recycle the oxidized products and redox equilibrium is reached (Varela *et al*., 1999).

1.3.2. Pyranose 2 oxidase

(D-glucose + O_2 ⇌ 2-dehydro-D-glucose + H_2O_2)

Pyranose 2-oxidase (POx) is a FAD-dependent member of the GMC oxidoreductase family (Hallberg *et al*, 2004). The enzyme is preferentially located in the hyphal periplasmic space and associated with membranous components of lignocellulolytic fungi (Daniel *et al*., 1992) POx transcript patterns are similar to LiP and GlyOx and thus it is believed to play a role in lignocellulose degradation by providing the essential co-substrate hydrogen peroxide for lignin and manganese peroxidases (Daniel *et al*., 1994; de Koker *et al*., 2004). The enzyme was first isolated from *Polyporus obtusus* (Ruelius *et al*., 1968), and later from several other white or brown-rot basidiomycetes including *P. chrysosporium* (Pisanelli *et al*., 2009).

The reaction catalysed by POx follows a "bi bi ping pong" type mechanism typically observed in flavoprotein oxidoreductases (Ghisla and Massey, 1989; Prongjit *et al*., 2009). The reaction scheme can be divided into two half reactions: the first reductive half reaction, where an aldopyranose substrate, preferentially D-glucose (Leitner *et al*., 2001), reduces the FAD cofactor to $FADH_2$ and 2-dehydroaldose (2-ketoaldose) as result of the sugar oxidation at position C-2 (reaction 1) (Freimund *et al*., 1998) and a second oxidative half reaction, where the cofactor is re-oxidized by reducing molecular oxygen to H_2O_2 (reaction 2) (Ruelius *et al*., 1968). Alternatively electron acceptors, including quinones, radicals or chelated metal ions can also be used by pyranose oxidase instead of oxygen (reaction 3) (Leitner *et al*., 2001).

$$\mathrm{FAD} + \mathrm{aldopyranose} \rightarrow \mathrm{FADH_2} + 2 - \mathrm{keto} - \mathrm{aldopyranose}\ (1)$$

$$\mathrm{FADH_2} + \mathrm{O_2} \rightarrow \mathrm{FAD} + \mathrm{H_2O_2}\ (2)$$

$$\mathrm{FADH_2} + \mathrm{benzoquinone} \rightarrow \mathrm{FAD} + \mathrm{hydroquinone}\ (3)$$

The natural substrates of POx are D-glucose, D-galactose, and D-xylose, highly abundant in lignocellulose which are respectively oxidized to 2-keto-D-glucose (D-arabino-hexos-2-ulose, 2-dehydro-D-glucose), 2-keto-D-galactose (D-lyxo-hexos-2-ulose, 2-dehydro-D-galactose), and 2-keto-D-xylose (D-threopentos-2-ulose, 2-dehydro-D-xylose) (Leitner *et al*., 2001). In addition, POx is also able to use other carbohydrates, including L-sorbose, D-glucono-1,5-lactone, and D-allose as substrates (Freimund *et al*., 1998). In some fungi, the oxidized sugar can be further metabolized into, for instance, the antibiotic cortalcerone (Giffhorn, 2000).

POx's, have received increased attention considering their high potential for i) biotransformation of carbohydrates *e.g.* the oxidized 2-keto sugars obtained from D-glucose and D-galactose that can be reduced at position C-1 to yield high levels of virtually pure D-fructose or D-tagatose, both ketoses with a number of known or proposed applications in food industry; ii) clinical analytics *e.g.* for determination of 1,5-anhydro-Dglucitol, an important marker for glycemic control in diabetes patients (Giffhorn, 2000).

1.3.3. Radical copper oxidases (RCO)

The radical copper oxidases constitute a family of enzymes that includes the glyoxal oxidase (reaction 4) and galactose oxidase (reaction 5) which are enzymes that are defined by a distinctive metalloradical complex in their active sites. This radical-Cu motif performs a two-electron oxidation of simple alcohols and aldehydes with concomitant reduction of O_2 to H_2O_2, accordingly to the following reactions:

$$\text{GlyOx: } RCHO + H_2O + O_2 \rightarrow RCO_2H + H_2O_2 \quad (4)$$

$$\text{GalOx: } RCH_2OH + O_2 \rightarrow RCHO + H_2O_2 \quad (5)$$

Current evidences suggest that RCOs have evolved to serve as factories for H_2O_2 production in extracellular spaces (Whittaker JW and Whittaker, 1998).

1.3.3.1. Glyoxal Oxidases

Glyoxal oxidase (GlyOx) was first reported in *P. chrysosporium*, as a secreted enzyme providing hydrogen peroxide, which is required as an oxidant in the peroxidative reactions of lignin degradation (Kersten, 1990; Kirk and Farrell, 1987). The enzyme has a broad specificity in the oxidation of simple aldehydes, hydroxycarbonyl, and dicarbonyl compounds to the corresponding carboxylic acids. GlyOx is an acidic monomeric glycoprotein with a molecular mass of 57 kDa. The active site of the enzyme is nearly identical to that of galactose oxidase, although comparison of the nucleotide sequence for GlyOx and GalOx shows no obvious homology in primary structure (McPherson *et al.*, 1992), which shows less than 20% amino acid sequence similarity based on standard sequence alignment algorithms (Whittaker *et al.*, 1996). The presence of two distinct one-electron acceptors, a Cu^{2+} metal center and an internal

Cys-Tyr radical forming a metallo-radical complex, lead to the classification of GlyOx as an RCO (Whittaker *et al*., 1996, 1999). Biological evidence have been found that *P. chrysosporium* specifically secretes simple dicarbonyls (glyoxal and methylglyoxal) to drive GlyOx reaction (Kersten and Kirk, 1987) and that further metabolism of glyoxylic acid, leads to the formation of oxalic acid, which was identified as a cofactor for manganese peroxidase turnover (Kuan and Tien, 1990). Some of the products that appear upon fragmentation of lignin, like glycoaldehyde, can also act as GlyOx substrates. GlyOx has a regulatory mechanism responsive to peroxidases, peroxidase substrates and peroxidase products that are suggestive of a close physiological connection with *P. chrysosporium* peroxidases (Kersten, 1990). Under nitrogen and carbon starvation the GlyOx gene is upregulated (Kersten, 1990) and the purified enzyme is catalytically inactive but can be activated by lignin peroxidase and nonphenolic peroxidase substrates or by treatment with strong oxidant (Whittaker *et al*., 1996). Phenolic compounds impair the GlyOx activation by lignin peroxidase (Kurek and Kresten, 1995).

1.3.3.2. Galactose oxidase

1.3.3.2.1. General properties

Galactose oxidase (GalOx) is a glycosylated protein that was first reported in the culture medium of the fungus *Polyporus circinatus* (Cooper *et al*., 1959) and that was re-classified as *Dactylium dendroides* (Nobles and Madhosingh, 1963) and later as *Fusarium graminearum* (Ogel *et al*., 1994). GalOx was proposed to catalyse the production of H_2O_2, as its primary biological relevant product, with a putative role in either antibiotic defence or in the synergistically action with peroxidases in lignin degradation (Whittaker *et al*., 1998). The broad substrate specificity of GalOx appears to support these functions being a compromise to achieve high rates of turnover to generate H_2O_2 from a range of available substrates (Whittaker *et al*., 1998).

Initially GalOx was suggested to be a metalloenzyme with zinc and flavin mononucleotide as prosthetic groups but later, copper was established as the sole exogenous cofactor (Avigad *et al*., 1962). Mature galactose oxidase (GalOx, oxidoreductase: galactose 6-oxidase, EC 1.1.3.9) from *F. graminearum*, the most studied GalOx, contains 639 amino acids and one cupric ion, and has a molecular weight of 68 kDa. Before fungal cells secrete GalOx, a peptide containing 41 amino acids is removed. This peptide contains a signal sequence necessary for

the transport of GalOx out of cells and a precursor sequence that was proposed to facilitate the folding of GalOx (McPherson *et al*., 1993).

The native GalOx from *F. graminearum* exhibits unusual resistance to denaturation and reduction. At 6 M of urea the enzyme is still active and the titration of its thiol groups with 5,5'-dithio-bis(2-nitrobenzoic acid) has proved to be very difficult. These results indicated an inherent conformational stability with inaccessible cysteine groups (Kosman *et al*., 1974). *F. graminearum* GalOx has an unusually high isoelectric point (pI = 12) that accounts for its high affinity to glass and macromolecules and its inability to be removed without a diluted acid wash (10% HNO_3) (Kosman *et al*., 1974).

GalOx oxidises the 6-hydroxyl group of D-galactose into the corresponding aldehyde with concomitant reduction of oxygen to hydrogen peroxide (Fig 1.4.). The enzyme has a surprisingly low specificity for primary alcohols but is regioselective and sterioselective. For example D-galactose is an enzyme substrate, but not L-galactose or D-glucose, where secondary alcohol groups are not oxidised (Amaral *et al*., 1963).

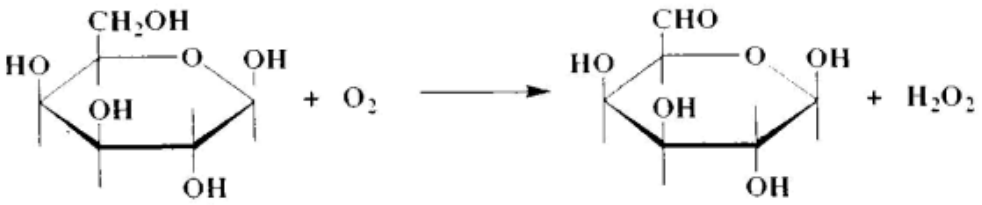


Figure 1.4. Oxidation of D-galactose to aldehyde and hydrogen peroxide catalysed by GalOx (adapted from Halcrow *et al*. (2000)).

When hydrogen peroxide accumulates during catalysis, GalOx is irreversibly inactivated. However, when no substrates are present, GalOx is stable in the presence of hydrogen peroxide (Hamilton *et al*., 1978). At higher concentrations, oxygen was observed to retard the enzyme inactivation by hydrogen peroxide. The rate of GalOx catalysis is significantly increased by some oxidizing reagents, such as ferricyanide and EDTA complex of Mn(III) (Cleveland *et al*., 1975; Hamilton *et al*., 1977). GalOx is activated also by peroxidases, which are commonly used to assay GalOx activities. It was proposed that activation of GalOx by peroxidase was due to the oxidation of GalOx by the high redox potential intermediate radical produced in the peroxidase reaction (Hamilton *et al*., 1978).

1.3.3.2.2. Structure of GalOx

The crystal structure of GalOx was solved in 1991 at the University of Leeds (Ito *et al*., 1991). The crystallographic structure of GalOx at pH 4.5 reveals a three-domain structure where β-sheets are predominant (Fig 1.5.) and the presence of a novel thioether bond in the active site which had not been previously observed (Ito *et al*., 1991; Ito *et al*., 1994). Domain I (residues 1 to 155) is located at the N-terminus of the protein and consists of 8 β-strands in a β-sandwich structure with a jelly-roll motif with a 5-stranded antiparallel β-sheet facing a 3-stranded antiparallel β-sheet. The function of domain I is presently unclear, since it does not contain catalytic residues and it is not expected to be directly involved in the catalysis, however it may be involved in the binding of polysaccharides present in cell walls, thus anchoring the enzyme in a favourable position for catalysis (Ito *et al*., 1994) or in facilitating the correct folding of GalOx (McPherson *et al*., 1993; Mahmoud *et al*., 2000). Domain II (residues 156-532) has a pseudo 7-fold symmetry called a β-propeller (Murzin, 1992), consisting in 1 α-helix and 28 β-strands, arranged in 7 antiparallel β-sheets, each containing 4 β-strands (Ito *et al*., 1991, Ito *et al*., 1994). The 7 β-sheets are thought to contribute for a higher stability of the enzyme due to its tight packing. It has also been confirmed that each 'blade' of the β-propeller is encoded by a *kelch* sequence motif (Ito *et al*., 1994). Cupric ion is close to the axis of the β-propeller at the surface exposed active site and where the catalytic residues are found. Domain III (residues 533-639) consists of 7 β-strands surrounding a hydrophobic core on the opposite side of the cupric ion of domain II (Ito *et al*., 1994). An antiparallel β ribbon, formed by two β-strands that are significantly longer than the others, pierces the middle of Domain II along the pseudo 7-fold axis to provide a ligand (His 581) for the copper ion and therefore, Domain III is thought to stabilise the β-propeller of Domain II (Ito *et al*., 1994).

The active site of GalOx (Fig 1.6.) contains a Cu(II) centre where the copper is coordinated by five ligands: Tyr272, His496, Tyr495, His581 and an exogenous ligand such as water or acetate ion through a square pyramidal coordination (Ito *et al*., 1991). The position of the acetate molecule in the crystal structure is proposed to mimic the binding site of the alcohol substrate during catalysis (Whittaker, 2005). The tight packing of Domains II and III gives significant physical rigidity to the site, which has been suggested to result in the high specificity for copper as the metal cofactor (Ito *et al*., 1994).

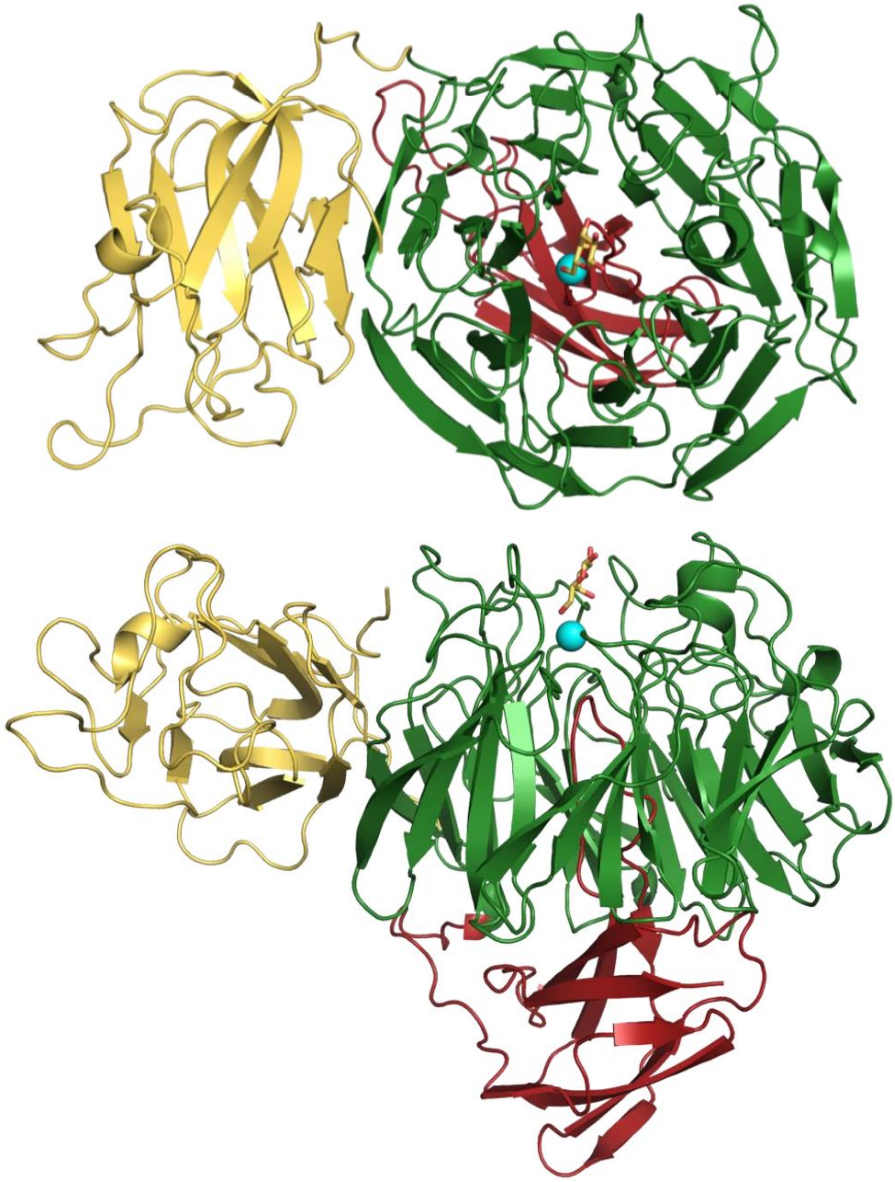


Figure 1.5. X-ray crystal structure of galactose oxidase. Two orthogonal views are shown, parallel (top), and perpendicular (bottom) to the axis of the β-propeller. Domain 1 is shown in yellow, domain 2 is shown in green and domain 3 is shown in brown. Copper is shown as a cyan sphere and the D-galactose molecule in yellow is modelled into the active site according to Wachter and Branchaud (1996). Based on PDB ID 1GOG and rendered with PyMOL (adapted from Whittaker (2003)).

The crystal structure of GalOx showed also that one of the copper ligands, Tyr272, in the active site, was covalently bonded at $C_{\varepsilon 1}$ to the sulphur of Cys228 to give a thioether bond (Ito *et al*., 1991) (Fig 1.7.). This finding is consistent with distinct and independent evidences by protein chemistry (Ito *et al*., 1994; McPherson *et al*., 1993; Kosman *et al*., 1974), EPR data (Whittaker and Whittaker, 1990; Babcock *et al*., 1992) and resonance Raman (Whittaker *et al*., 1989). This arrangement results in an increased rigidity of the active site; indeed, the higher stability of the thioether bond generates a loop between Cys228 and Tyr272 resulting in a

molecular mass difference from 68 to 65 kDa as assessed by SDS-PAGE analysis (McPherson *et al*., 1993). Additionally, the formation of the thioether bond significantly diminishes the redox potential of Tyr272 from around 1000 (free tyrosine) to 400 mV making it more accessible for catalysis. This is the lowest redox potential of Tyr found in nature (Hamilton *et al*., 1978; Borman *et al*., 1998). Upon oxidation of Tyr272 a protein radical is formed and is stabilised by delocalisation across the thioether bond (Babcock *et al*., 1992). Within the active site Trp290 residue is thought to contribute to radical stabilisation; Trp290 is parallel to the plane of Tyr272 ring and staked over the thioether bond such that the six-membered ring of the Trp is located above the Cys sulphur atom (Ito *et al*., 1991). The other face of the Trp indole ring is exposed to solvent, thus Trp290 protects the stabilised radical site from solvent (Ito *et al*., 1994). The structure, formed by the bond Tyr-Cys and Trp, supports the radical reaction mechanism (Whittaker and Whittaker, 1988).

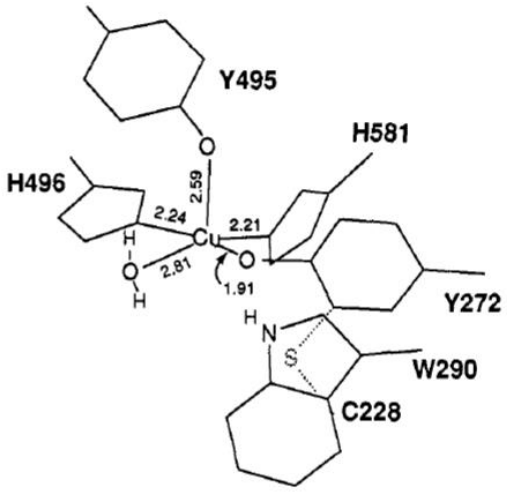


Figure 1.6. Structure of the active site of galactose oxidase (adapted from Whittaker and Whittaker (1998)).

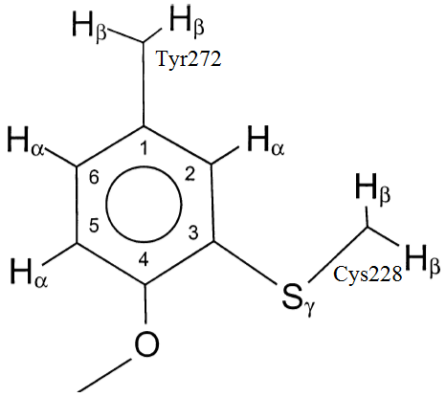


Figure 1.7. Structure of the tyrosine-cysteine cross-link site (adapted from Whittaker (2003)).

1.3.3.2.3. Reaction Mechanism

GalOx and its reaction mechanism have raised tremendous interest and a number of hypothesis have been formulated on how the mononuclear cupric ion active site of GalOx is able to catalyse the reduction of dioxygen and the oxidation of the alcohol group that are both two-electron processes. Generally, redox cofactors such as NAD(H) or quinones are involved in this kind of catalysed reactions (Whittaker, 2005). One of the first theories suggested a mechanism involving the conversion between Cu(III) and Cu(I) (Hamilton *et al*., 1978), but later the theory was disproved (Clark *et al*., 1990) and another mechanism was proposed, where a pyrroloquinoline quinone cofactor was implied (van der Meer *et al*., 1989). In the same period, Whittaker and co-workers were the first to prepare stable and homogeneous forms of GalOx and to propose that oxidation of the enzyme resulted in a protein-based radical involved in catalysis where the copper ion suffered a number of oxidation states (Whittaker and Whittaker, 1988). Crystallization of GalOx with the identification of the thioether bond (Ito *et al*., 1991) and site-directed mutagenesis studies (Baron *et al*., 1994) further supported the radical mechanism proposed.

There are three distinct redox states in which GalOx can exist and two of them are stable: an oxidized Cu(II)-Tyr• state (active); a semireduced Cu(II)-Tyr state (inactive); and a reduced state Cu(I)-Tyr which only occurs during catalysis or upon addition of substrate under anaerobic conditions (Fig 1.8.).

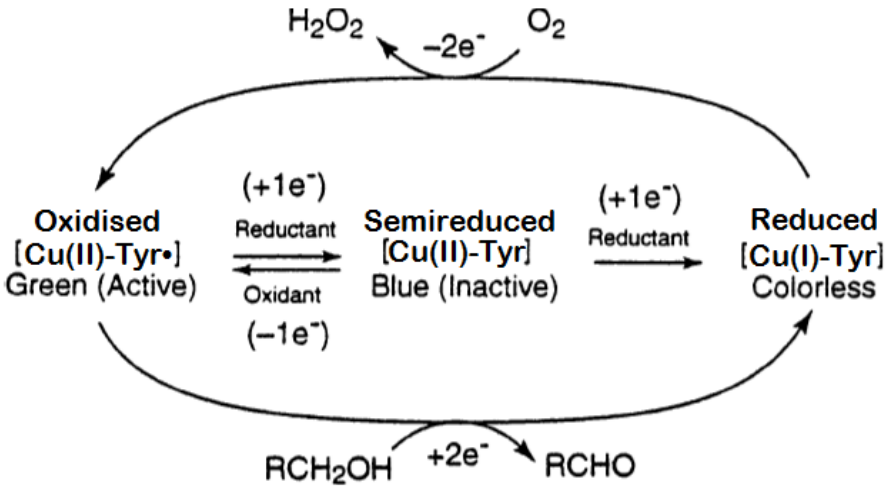


Figure 1.8. Schematic redox interconversion of Galactose Oxidase species (adapted from Whittaker (2003)).

Native GalOx from *F. graminearum* was isolated as a mixture of the stable states, which can be interconverted by inorganic oxidation and reduction (generally by adding ferricyanide and ferrocyanide, respectively) (Whittaker and Whittaker, 1988). As mentioned for GlyOx, peroxidases have also been shown to oxidise GalOx to the active form, (Cleveland *et al*., 1975, Tressel and Kosman, 1980; Baron *et al*., 1994). Each of the three states of the enzyme show characteristic absorbance spectral features (Fig 1.9.). The semireduced state shows a band at 450 nm, while the oxidised state exhibits strong bands both at 450 nm and 800 nm. The band at 450 nm has been assigned to a ligand to metal charge transfer (LMCT) between Tyr495 and Cu(II), combined with a $\pi - \pi^*$ transition of the Tyr272 radical (only present in the oxidised state). The broad band at 800 nm has been proposed to derive from a ligand to ligand charge transfer (LLCT) spanning both Tyr ligands (Tyr272-Tyr495) combined with smaller contributions from Cu(II) *d-d* and intraradical absorption (Whittaker and Whittaker, 1998). A charge transfer between Y272 and W290 is also proposed to contribute to the absorption at 800 nm (Baron *et al*., 1994).

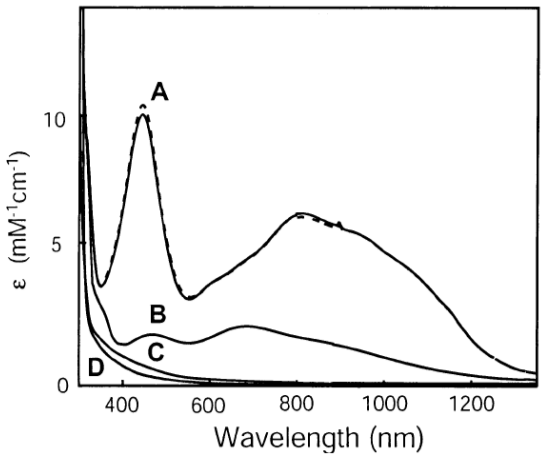


Figure 1.9. Optical absorption spectra for galactose oxidase in accessible redox states. (A) Redox-activated enzyme (oxidised) (pH 5.6, ---; pH 7.3,—). (B) One-electron reduced inactive enzyme (semireduced). (C) Substrate-reduced anaerobic enzyme (reduced). (D) Metal-free apoprotein (adapted from Whittaker (2003)).

It is currently accepted that oxidation of substrates by GalOx occurs by a "bi bi ping pong" mechanism where GalOx binds and oxidises the alcohol, releases the corresponding aldehyde, and consecutively O_2 binds, is reduced, and H_2O_2 is released (Fig 1.10.).

In the first step of the reductive half reaction, substrate binds to copper by displacing the equatorial solvent (Ito *et al*., 1994, Wachter and Branchaud, 1996), with abstraction of a proton from the acidified C-6 hydroxyl by the copper axial ligand, Tyr 495, thus acting as the

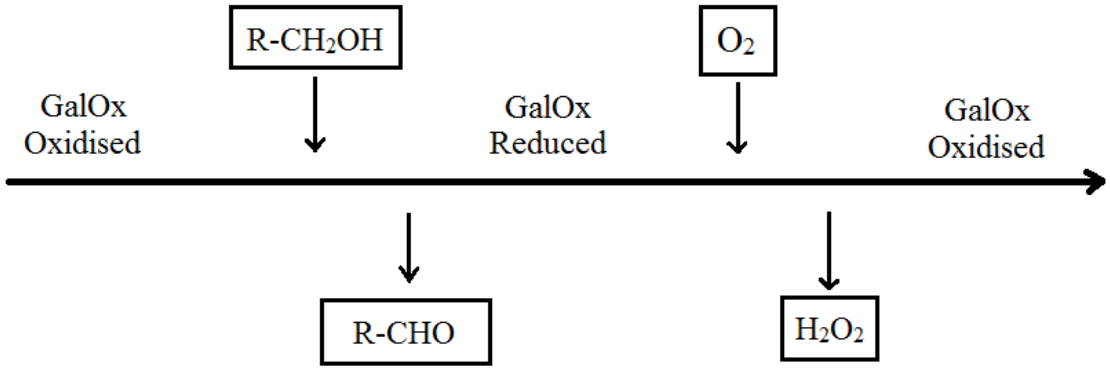


Figure 1.10. Schematic representation of the bi-bi ping-pong mechanism proposed for GalOx.

catalytic base (Fig 1.11., (i)). This step is supported by an inhibitor azide binding experiment (Whittaker and Whittaker, 1993) and a site-directed mutagenesis study, in which the mutation of Tyr495 to Phe removes the ability of the residue to abstract protons (Reynolds *et al*., 1995). The *pro*–S hydrogen atom is then abstracted from the C-6 methylene group of the substrate intermediate by the radical present at Tyr272, leading to a partially oxidized radical substrate (Fig 1.11., (ii)). Kinetic isotopes effects backed up this hydrogen atom transfer (Whittaker *et al*., 1998), and also the stereoselectivity of the transfer, which was due to constraints imposed by the active site structure on the orientation of substrate binding (Wachter and Branchaud, 1996, Minasian *et al*., 2004). In the reductive half reaction, Cu(II) is reduced to Cu(I) by the transfer of a single electron to the partially oxidised radical substrate, oxidizing it, and consequently the aldehyde product is released (Fig 1.11., (iii)) (Wachter and Branchaud, 1996).

The second oxidative half reaction is much less well characterized, probably due to the particularly fast reaction between Cu(I) with O_2. O_2 binds to Cu(I) and transfers an electron yielding a metal-bound superoxide (O_2^-) and reforms the Cu(II) center (Fig 1.12., (iv)) (Himo *et al*., 2000). A hydrogen atom from the Tyr272 is then abstracted by O_2^- to form hydroperoxide (HO_2^-) and reform the Tyr272 radical simultaneously (Fig 1.12., (v)). Finally a proton is transferred from Tyr495 and hydrogen peroxide is formed and is released from the active site (Fig 1.12., (vi)), with renovation of the hydrogen bond between Tyr 495 and Cu(II) (Humphreys *et al*., 2009). The cycle is closed and GalOx is restored to its native (fully reoxidised) state.

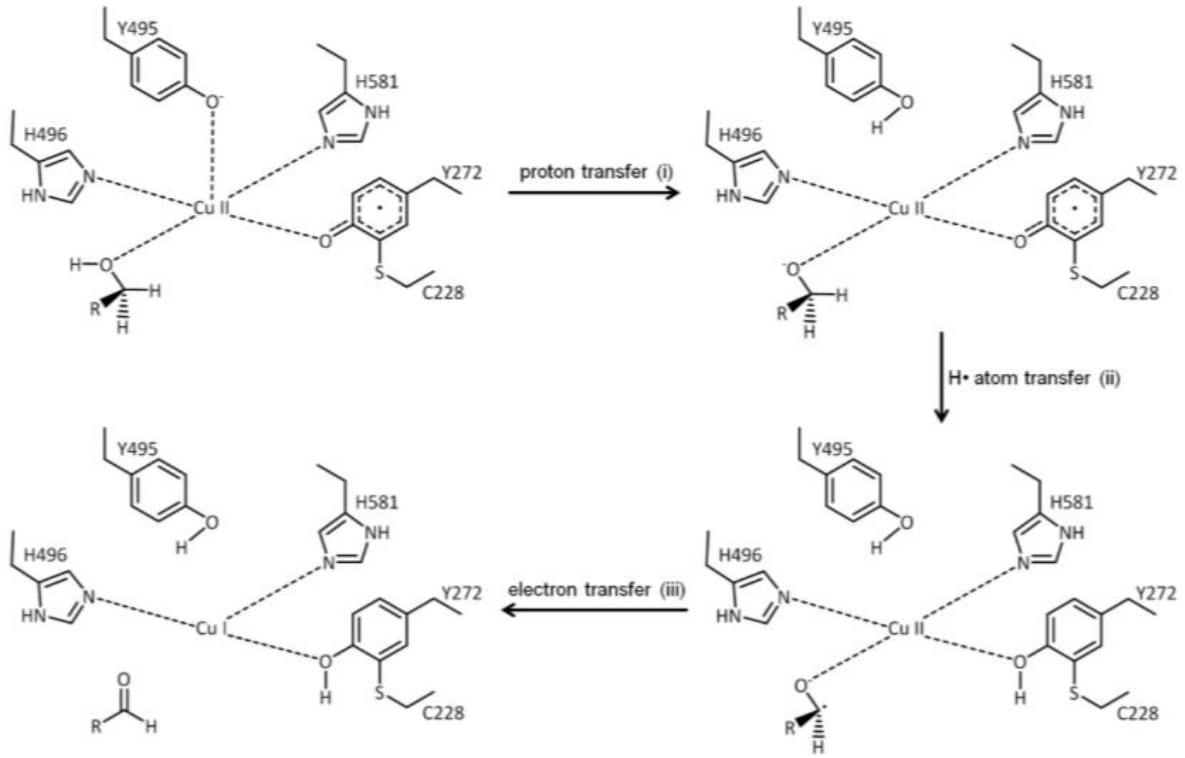


Figure 1.11. The reductive half reaction of GalOx (adapted from Messerschmidt *et al*. (2001)).

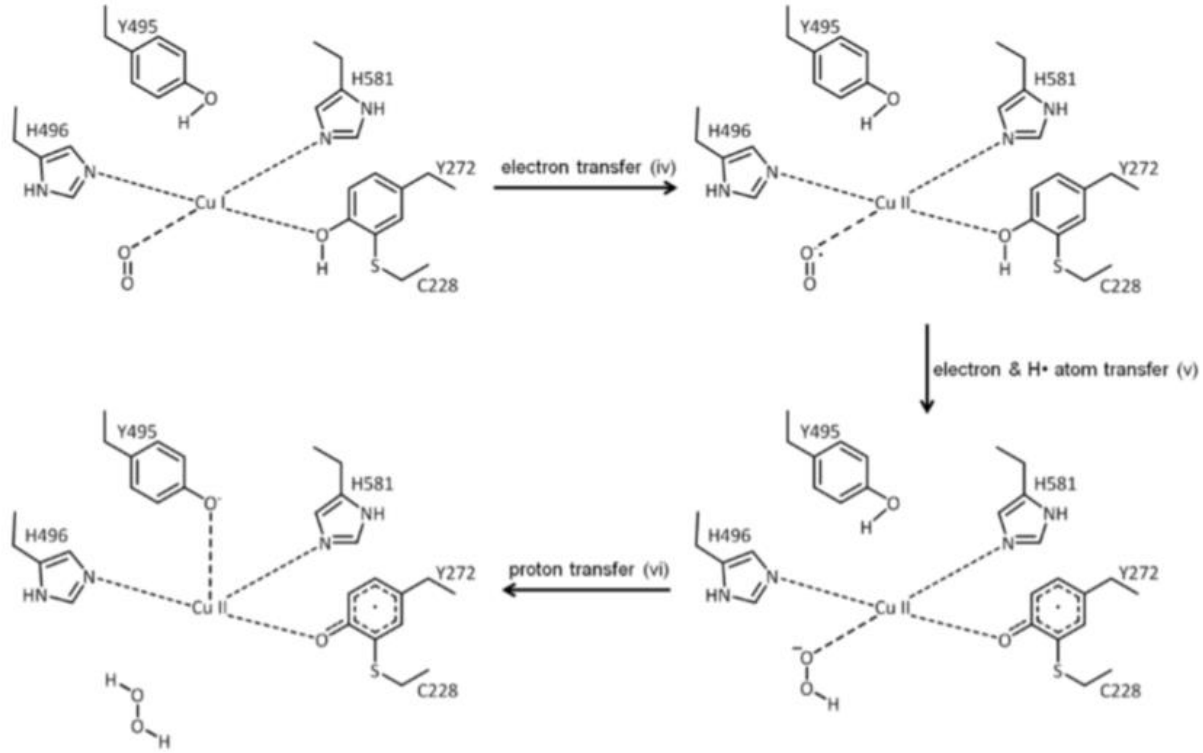


Figure 1.12. The oxidative half reaction of GalOx (adapted from Messerschmidt *et al*. (2001)).

1.3.3.2.4. Substrate Specificity and Binding

A GalOx from *F. graminearum* displays relatively broad substrate specificity to a range of primary alcohols. The native substrate of GalOx is D-galactose, having a K_m of 67 mM and k_{cat} of 2990 s^{-1} (Baron *et al*., 1994). GalOx activity is higher towards polysaccharides containing D-galactose on the non-reducing end, rather than D-galactose alone (Avigad, 1985, Schlegel *et al*., 1968). This seems to be primarily due to an increased K_m of the enzyme for the monosaccharide D-galactose; for example the K_m for guar gum, a polysaccharide composed by linked mannose with side branches of galactose units, is 1000 × lower than for galactose (Avigad, 1985). The enzyme shows lower K_m values also for liposomes with galactose residues on the surface (Ohno and Kitano, 1998) and glycoproteins with terminal or sub-terminal D-galactose units (Goudsmit *et al*., 1984). The model for the substrate binding site proposed for GalOx is consistent with oxidation of oligo and polysaccharides, in which the substrate binds to copper at an equatorial position and the long polysaccharide chain is exposed to the solvent (Ito *et al*., 1994). GalOx also oxidises various aliphatic alcohols, like glycerol, but with higher K_m values than D-galactose (Klibanov *et al*., 1982). Investigations of the water accessibility in the surface of GalOx active site using X-ray studies revealed a pocket located at the copper site that is structurally complementary to the D-galactose chair form (Ito *et al*., 1994). This study has also showed, that R330 and Q406 residues form hydrogen bonds with O4 and O3 of D-galactose and that the backbone of D-galactose has hydrophobic interactions with the side chains of F464 and F194, which are residues located at the pocket wall. In addition, D-glucose, a C-4-epimer of D-galactose not used by GalOx as a substrate, seems to break the hydrogen bond with R330 and cause serious steric clash with Y495 due to the presence of a O-4 hydroxyl instead of a O-6 (D-galactose) (Fig 1.13.). This study was corroborated by molecular modelling research (Wachter and Branchaud, 1996). Modelling studies indicated the most favourable conformation for D-galactose to bind in the active site was through seven hydrogen bonds. The study revealed also that the good fit of D-galactose for the hydrophobic residues F464 and F194, may be important for positioning the substrate for hydrogen bonding interactions in the active site, due to the more hydrophobic nature of the C-H groups rather than the side containing C-OH groups. Concerning D-glucose, as well as the steric clash with Tyr495, it is proposed that D-glucose binding is disfavoured as a result of its equatorial hydrophilic O-4 hydroxyl that would be in close proximity to the mentioned phenylalanine hydrophobic residues if bound in the same conformation as D-galactose in native GalOx (Wachter and Branchaud, 1996). Present

in a mixture even high concentrations of D-glucose do not interfere with the D-galactose oxidation reaction (Wachter and Branchaud, 1996).

An important feature of GalOx in what concerns substrate specificity is the strict regioselectivity *e.g.* only the C-6 hydroxyl of D-galactose is oxidised (Whittaker and Whittaker, 1998), making it a valuable catalyst in organic synthesis with high potential for industrial applications.

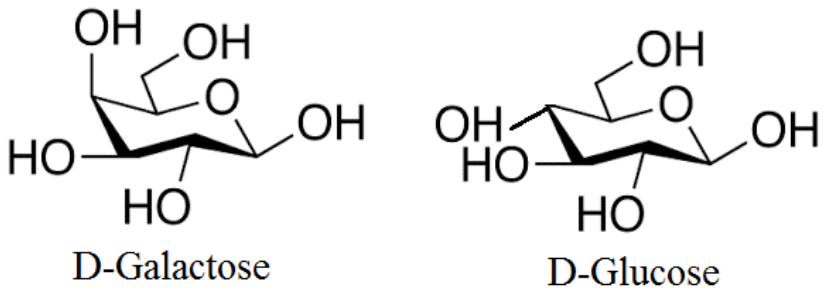


Figure 1.13. Representation of the sugars D-galactose and D-glucose.

GalOx like a number of other enzymes have been the target of several mutagenesis studies aiming to improve its activity or altering the substrate specificity (Sun *et al*., 2001; Deacon and McPherson, 2011).

1.3.3.2.5. Functional Expression of GalOx

Production of native GalOx by *F. graminearum* was shown to result in a very time-consuming and inefficient process. In order to allow, biochemical characterization and protein engineering studies, the expression of GalOx was carried out in a recombinant expression system, to efficiently generate high yields of functional enzyme. The first optimized expression system, in the filamentous fungus *Aspergillus nidulans*, resulted in protein yields of 50 mg per liter of culture, 10-fold to those obtained using the native microrganims (Baron *et al*., 1994). The spectral properties and kinetic parameters of the expressed GalOx are identical to the native enzyme. All crystal structures of GalOx that derived from mutational studies were performed in this expression system (Baron *et al*., 1994; Firbank *et al*., 2001). The methylotrophic yeast *Pichia pastoris* is the other system that has been used and represent, thus far, the best generating heterologous system to generate functional GalOx (Whittaker and Whittaker, 2000), with yields

as high as 500 mg per liter culture. Since the industrial use of enzymes is primarily limited by the associated costs, GalOx produced in in this system is used for numerous applications.

E. coli, the most widely used expression system considering its fast growth rate and well established genetic manipulation was also utilized for GalOx functional expression. It was demonstrated that active *F. graminearum* GalOx in its mature form is produced by *E. coli* cells (Lis and Kuramitsu, 1997); native GalOx is translated as a precursor protein with an N-terminal pro region, including a secretion signal peptide that is cleaved during maturation of the enzyme within the eukaryotic host (Rogers *et al*., 2000).

The omission of the pro sequence, followed by directed evolution, enabled the generation of a variant produced in *E.coli* with 18-fold increased expression levels when compared to the wild type (WT), and also with improved activity and stability towards D- galactose (Sun *et al*., 2001). A more recent report led to a mutant with even greater production levels through the combination of previous mutations with new ones (Deacon and McPherson, 2011).

1.3.3.2.6. Applications of GalOx

GalOx's unique features make them ideal candidates for industrial applications, and have raised great interests in various fields. There are numerous examples of the use of GalOx in biotechnological applications other than the application of GalOx as being part of a multienzymatic system for lignin processing. One of the advantages is the excellent selectivity of GalOx towards D-galactose and its derivatives. This feature has been used to develop analytic tools to detect and quantify D-galactose and derivatives in various mixtures. An important example is the use of GalOx to monitor the levels of D-galactose in patients that have galactosemia (Holton, 1990). There are various other examples where it is necessary to monitor the levels of D-galactose and other GalOx substrates, as in the food industry. Two interesting examples are found in the processing of sugar beet (Rorem and Lewis, 1962) and in the dairy industry (Adanyi *et al*., 1999, Mannino *et al*., 1999). GalOx based biological sensors, as we previous mentioned, have suffered significant development to enable accurate detection of D-galactose and lately a variety of methods for immobilisation of GalOx within a biosensor have been reported (Taylor *et al*., 1977; Dicks *et al*., 1986; Vrbova *et al*., 1992; Lourenço *et al*., 2003; Sharma *et al*., 2006; Sharma *et al*., 2007; Cevik *et al*., 2010; Kim *et al*., 2012; Charmantray *et al*., 2013; Kanyong *et al*., 2013).

GalOx is also used for modification of oligo and polysaccharides, often followed by further modifications using other enzymes, providing the access to a wide range of derivatives (Yalpani and Hall, 1982); manufacture of sweeteners and flavorants for the food industry (Xu *et al*., 2000); as well as paper strength additives for the paper industry, with the oxidation of the natural plant-based polymer guar (Delagrave *et al*., 2001), and also derivatives of galactoxyloglucan (from blend of galactose, xylose and glucan) with potential uses in abrasion resistant clothing or diagnostic strips (Parikka *et al*., 2012).

We can also point out the generation of chemicals and polymers by GalOx as another interesting application such as in the synthesis of sugar-containing polyamines (sugar-based polymers) used in hydrogels, adsorbents and bio recognition agents (Liu and Dordick, 1999); of L-glucose, L-galactose and L-xylose (Yadav *et al*., 2002); of unnatural sugars (Root *et al*., 1985); and of 5-C-(hydroxymethyl) hexoses, a potential group of artificial sweeteners (Mazur and Hiler, 1997), which are usually difficult to synthesize through chemical methods. The catalysed generation of the previous compounds using GalOx removes the need for protection of hydroxyl groups resulting in a more straightforward and efficient process.

Medical research is also an area where GalOx has been a valuable tool. Different types of cells have been characterised and quantified through labelling of D-galactose-terminating glycans on cell surfaces, such as in brain tissue (Singh and Kanfer, 1980) or in the characterization of changes in cell surface glycan expression, as occurred upon virus infection of fibroblasts (Critchley, 1974) or human erythrocyte aging (Gattegno *et al*., 1981). This method consists in the oxidation of terminal galactosyl residues by GalOx where the aldehyde product is detected by fluorescent (Rannes *et al*., 2011) or chemiluminescent methods (Han *et al*., 2012), instead of previously performed reduction of the aldehyde product by tritiated borohydride (Morell *et al*., 1966). An additional use for cell surface glycans labelling is relevant in cancer diagnosis as the disaccharide D-galactose-β-(1-3)-N-acetyl-D-galactosamine (Gal-GalNAc) is expressed on the surface of colon cells in the early stages of carcinogenesis. After oxidation by GalOx, the aldehyde groups are detected using the Schiff's reagent which produces a magenta coloration (Shamsuddin and Elsayed, 1988). This assay has been a patient's cost-effective screening test (Said *et al*., 1999). The expression of this marker in carcinomatous breast, ovary, pancreas, stomach, and endometrium has expanded the use of this simple screening method for many types of cancers (Shamsuddin *et al*., 1995).

Research into lymphocytes and associated disorders such as leukaemia has been possible (Berzins *et al*., 1983, Zafar *et al*., 1981) using treatments with neuraminidase and also GalOx to induce lymphocyte blastogenesis and generation of growth factors (Novogrodsky and Katchalski, 1973). GalOx has the potential to be applied in other fields and the modification of the enzyme in order to introduce or improve its activity towards other substrates rather than D-galactose is much sought. For example enhancing the activity towards glycerol to which the enzyme already shows some activity can lead to the production of high value products with a range of potential uses, such as, lactic acid which can be used to synthetize petroleum alternatives, food additives, textile and leather treatments (Paster *et al*., 2003).

1.4. Context of the project

The research focus of Microbial and Enzyme Technology Lab at NOVA ITQB has been to explore microorganisms and enzymes with potential for applications in the areas of environmental and industrial biotechnology. The set-up of biocatalytic processes for best performance depends on the cyclic development and combination of multidisciplinary research strategies at the biocatalyst level: i) selection from *in vitro* or *in silico* screening, ii) biochemical characterization for providing structural, catalytic and stability fingerprints, iii) engineering to improve their performance and robustness and iv) application for the synthesis, degradation or modification of compounds of interest. The research activities are in the interface of protein science and protein technology and have focused in the isolation, characterization and engineering of useful bacterial oxidoreductases (laccases, azoreductases, peroxidases) for the degradation and/or valorisation of xeno or natural aromatic compounds, including synthetic dyes and lignin-related compounds, with impact in the bioremediation and biorefinery fields.

Pioneering studies were performed in the structural and functional characterisation of bacterial laccases (Martins *et al*., 2015). Important insights were given on the role of synthetic and phenolic redox mediators that enhanced the substrate range of these biocatalysts (Rosado *et al*., 2012). A toolbox of experimental approaches was mounted at ITQB to explore these biocatalysts not only for the detoxification of industrial synthetic dyes (Mendes *et al*., 2015). Efficient expression systems to produce large quantities of these enzymes and tools to genetically manipulate them were developed. Thorough multidisciplinary investigations of wild type and engineered variants revealed key functional aspects of these enzymes, such as the solvent accessibility of the catalytic centres, the electrostatic interactions modulating the redox

potential and the molecular mechanisms behind their thermodynamic stability (Durao *et al*., 2006; Durao *et al*., 2008; Brissos *et al*., 2012; Chen *et al*., 2010). The studies of laccase-like enzymes from hyperthermophilic microorganisms provided the first evidence that these enzymes possess notable metal oxidase activity and an extreme intrinsic thermostability, which is worthwhile exploring for engineering of enzymes with improved chemical robustness (Fernandes *et al*., 2007; Fernandes *et al*., 2010). We have recently focused on bacterial dye-decolourising peroxidases (DyP-type), novel enzymes that have primary sequence, structural and apparently mechanistic features unrelated to those of other known peroxidases (Santos *et al*., 2104; Sezer *et al*., 2013; Mendes *et al*., 2015). They successfully degrade a wide range of dyes, including anthraquinone-based and azo dyes, and non-phenolic lignin units in the absence of redox mediators. It was demonstrated that DyPs from different sources reveal distinct stability, substrate specificity, catalytic efficiency and redox properties, as a basis for exploring their potential to replace the high-redox fungal lignin and versatile peroxidases in biotechnological applications.

In this study the aim was to further expand the repertoire of oxidoreductase enzymes involved in the biotransformation of biotechnological important substrates. Biocatalysis offers an environmentally friendly tool for lignin degradation, holding additionally the key for its successful valorisation. Our proposal is based on the central hypothesis that degradation of the complex, irregular and mostly insoluble lignin polymeric structure can only be satisfactorily addressed through the application of multiple enzymatic or chemo-enzymatic systems. In nature, lignin depolymerization is catalysed via poorly understood mechanisms but recent genome mining suggest that laccases and high-redox peroxidases act synergistically with an array of enzymes such as radical-copper oxidases, glutathione-S-transferases, aldo-keto reductases, esterases and others that likely have a concerted action on both (hemi)cellulose and lignin. Therefore this study focusing at identifying and characterizing novel bacterial peroxide forming enzymes represents a step forward on the understanding and application of the interplay of bacterial ligninolytic characterized (e.g. CotA laccase and Dyp-type peroxidases) and auxiliary enzymes for the set-up of multi-enzymatic reactions using lignin models and lignin preparations leading to the production of new added-value chemical compounds in a biorefinery context.

2. MATERIALS E METHODS

2.1. Bacterial strains, plasmids and media

The *Escherichia coli* strains DH5α (Novagen) and KRX (Promega) were used for routine propagation and amplification of plasmid (Table 2.1). The *E. coli* strains BL21 (DE3, Novagen), BL21 star (DE3, Novagen), Tuner (DE3, Novagen) and Rosetta (DE3, pLysS, Novagen) were used for recombinant protein expression (Table 2.1). The DE3 designation means the strains contain the λ DE3 lysogen which carries the gene for T7 RNA polymerase under control of the lacUV5 promoter. Isopropyl β-D-1-thiogalactopyranoside (IPTG) is required to induce expression of the T7 RNA polymerase (Casali and Preston, 2003).

The pET-21a(+) vector (Novagen) (appendix A) was the plasmid used to clone the selected genes. This vector carry an N-terminal T7•Tag sequence plus an optional C-terminal His•Tag sequence and an ampicillin resistance marker.

Luria-Bertani medium (LB) was used for the growth of *E. coli* strains. LB medium contains: tryptone (10 g/L), yeast extract (5 g/L) and NaCl (10 g/L). Super Optimal Growth medium (SOB) was used for the growth of competent cells. SOB medium contains: tryptone (20 g/L), yeast extract (5 g/L) and NaCl (0.584 g/L) and KCl (0.186 g/L). All culture media solutions were autoclaved and stored at room temperature until use.

2.2. PCR amplification and cloning

A few putative glyoxal oxidase and/or galactose oxidase and pyranose oxidase genes were identified through BLAST search (please see Figs 3.1. and 3.2.) in the genomes of several bacteria including *Myxococcus stipitatus, Gloeobacter violaceus, Streptomyces clavuligerus, Stigmatella aurantiaca* and *Arthrobacter siccitolerans* (Table 2.2). The DNA of these microorganisms was obtained respectively from Lotte Sogaard-Andersen (Max-Planck Institute, Marburg, Germany), Satoshi Tabata (Kazusa DNA Research Institute, Japan), Paloma Liras (Univ Leon, Spain), Rolf Mueller (Helmholtz Institute for Pharmaceutical Research Saarland, Germany) and Maximino Manzanera (Univ Granada, Spain).

Primers were designed, based on the genomic DNA sequence of GOase and POx from several bacteria in GenBank (Table 2.2.) and custom synthesized (StabVida). The primers have *Nde*I and *Eco*RI (for the genes 1 to 6) or *Hind*III and *Xho*I (for the genes 7 and 8) restriction enzyme sites on the flanking ends of the ORF's (Table 2.3.).

Table 2.1. *Escherichia coli* strains used in this study.

Strain	Genotype	Key Feature(s)
DH5α	*deo*R *end*A1 *gyr*A96 *hsd*R17 Δ(*lac*)U169 *rec*A1 *rel*A1 *sup*E44 *thi*-1 (Δ80*lacZ* ΔM15)	Allows replication of large plasmids. Improves the yield and quality of isolated plasmidic DNA and provides increased stability of inserts
KRX	[F´ *tra*D36 Δ*omp*P *pro*A+B+ *lac*Iq Δ(*lac*Z)M15] Δ*omp*T *end*A1 *rec*A1 *gyr*A96 (Nalr) *thi*-1 *hsd*R17 (rK^- mK^+) e14– (McrA–) *rel*A1 *sup*E44 Δ(*lac-pro*AB) Δ(*rha*BAD)	It is an *E. coli* K12 derivative that has salient features associated with cloning plus engineered attributes to optimize controlled protein expression
BL21 (DE3)	F− *omp*T *gal dcm lon hsd*SB(rB-mB-) λ(DE3[*lac*I *lac*UV5-T7 gene 1 *ind*1 *sam*7 *nin*5])	It is the most widely used host background for protein expression and has the advantage of being deficient in the *lon* (8) and *omp*T proteases
BL21star (DE3)	F- *omp*T *gal dcm lon hsd*SB (rB-mB-) *rne*131 λ (DE3)	In addition to the λ DE3 lysogen the BL21star strain also contains the *rne*131 mutation to enhance the expression capabilities
Tuner (DE3)	*gal hsd*SB *lac*Y1 *omp*T λ (DE3)	It is a *lac*ZY deletion mutant of BL21, which enable adjustable levels of protein expression throughout all cells in a culture. The *lac* permease (*lac*Y) mutation allows uniform entry of IPTG into all cells in the population
Rosetta (DE3)	*gal hsd*SB *lac*Y1 *omp*T (pRARE araW argU glyT ileX leuW proL metT thrT tyrU thrU Camr) λ(DE3)	It is a BL21 derivative designed to enhance the expression of eukaryotic proteins that contain codons rarely used in *E. coli*. The original Rosetta strains supply tRNAs for the codons AUA, AGG, AGA, CUA, CCC, and GGA on a compatible chloramphenicol-resistant plasmid, pRARE

Table 2.2. Information on the genes targeted in the present study.

Nº	Protein	Strain	Base pairs	a.a.	MW (kDa)	PI	Accession number
1	hGalOxMs1	*Myxococcus stipitatus* DSM 14675	2778	925	101.1	6.2	AGC46318.1
2	hGalOxMs2	*Myxococcus stipitatus* DSM 14675	1419	472	51. 3	9.6	AGC42194.1
3	GalOxGv	*Gloeobacter violaceus* PCC 7421	2286	761	81.9	5.6	NP_927197.1
4	GalOxSt	*Streptomyces clavuligerus* ATCC 27064	2391	796	86.1	6.0	WP_003962715.1
5	hGalOxSa	*Stigmatella aurantiaca* DW4/3-1	2424	807	86.5	5.2	WP_002617931.1
6	GalOxSa	*Stigmatella aurantiaca* DW4/3-1	2568	855	92.1	6.5	ADO74264.1
7	GalOxAs	*Arthrobacter siccitolerans*	2283	760	78.9	5.3	CCQ45078.1
8	POxAs	*Arthrobacter siccitolerans*	1560	519	55.2	4.8	CCQ48064.1

Table 2.3. List of primers used to amplify the targeted genes in the present study.

Nº	Gene	Name	Primer
1	*hgalOxMs1*	FWD	5'-GAACCATATGGCCCACCCAGGGGGTGCCTAC-3'
		REV	5'-GCGAATTCTACTTCGCCTCGAGGTTCAGCTCGAGGTC-3'
2	*hgalOxMs2*	FWD	5'-CCAGCATATGACACCCGACCAGGTGGG-3'
		REV	5'-CCGGGAATTCAGGGAGGGGGGATGCGCAG-3'
3	*galOxGv*	FWD	5'-CTTAGTCCGAAGAACATATGCGTAAATCTC-3'
		REV	5'-CGCTAGGGAATTCCCTGTGCCAGTG-3'
4	*galOxSt*	FWD	5'-CTGGCTGTCGCGCATATGACCGCGGCACTGCTCGTCAC-3'
		REV	5'-CGGGAATTCAGCCGCGCCGCAGATCCCCCTTTC-3'
5	*hgalOxSa*	FWD	5'-GTAGGGAGGAATCATATGGTGGAAGGACAATCCAGAAGTGCGTGG-3'
		REV	5'-GTGGTCAAGAATTCATCAGGGCGCCGGGACCTGAACCCCGACATC-3'
6	*galOxSa*	FWD	5'-G GAGCAAACATATGCACTGCACTGTGCGGATCTGGATGGC-3'
		REV	5'-GAATCGAATTCGGCAGCCTTCTCAAGGGCCAGCAACCCAG-3'
7	*galOxAs*	FWD	5'-GCGCGAAGCTTAGATGGCCACCGCAAGCGACGAGGAAAC-3'
		REV	5'-AAGCCCCTCGAGTCACGAGATCTGGATCGTCGTCGCGAC-3'
8	*pOxAs*	FWD	5'-CAGGAACTGGAAGCTTGGCGATGAGCGGTCACCGGTATC-3'
		REV	5'-GCGCTACCGGCTCGAGTTCATTATTTAGACAGTCTATTG-3'

FWD- Forward Primer; REV- Reverse Primer. Restriction enzyme sites are underlined.

The polymerase chain reactions (PCR) were performed in a thermal cycler (*MyCycler*^{TM} Thermal Cycler, Biorad) in a final volume of 50 µL containing 20 nmol of each primer, the respective genomic DNA, 0.2 mM of 2'-deoxynucleotide 5'-triphosphate (dNTP's) (NZYTech), NZYProof polymerase buffer and 2.5 U of NZYProof polymerase (NZYTech). The mixtures were submitted to temperature cycles as described in table 2.4.

The PCR products were visualized by electrophoresis using 1% agarose gels (unless otherwise stated) in a 45 min-run under an electrical field of 100 V and purified using the kit GFX PCR DNA purification (GE Healthcare).

Table 2.4. Conditions used in the performed PCR reactions.

Gene	Denaturation	30 cycles	Final amplification
hgalOxMs-1	2 min at 95°C	95°C : 1 min 66°C: 1 min 72°C: 3 min	10 min at 72°C
hgalOxMs-2	2 min at 95°C	95°C : 1 min 64°C: 1 min 72°C: 1.5 min	10 min at 72°C
galOxGv	2 min at 95°C	95°C : 1 min 53°C: 1 min 72°C: 3 min	10 min at 72°C
galOxSt	2 min at 95°C	95°C : 1 min 62°C: 1 min 72°C: 3 min	10 min at 72°C
hgalOxSa	2 min at 95°C	95°C : 1 min 63°C: 1 min 72°C: 3 min	10 min at 72°C
galOxSa	2 min at 95°C	95°C : 1 min 62°C: 1 min 72°C: 3 min	10 min at 72°C
galOxAs	2 min at 95°C	95°C : 1 min 68°C: 1 min 72°C: 3 min	10 min at 72°C
pOxAs	2 min at 95°C	95°C : 1 min 62°C: 1 min 72°C: 2 min	10 min at 72°C

The final PCR products of the genes were digested simultaneously with 10 U of *Nde*I/*Eco*RI (for the genes 1 to 6) or 10 U of *Hind*III/*Xho*I (for the genes 7 and 8) (Thermo Scientific Fermentas) restriction enzymes for 2 h at 37°C. Products were purified using the kit GFX PCR DNA purification (GE Healthcare).

The cloning vector pET21a (+) (Novagen) was digested using identical conditions: with 10 U *Nde*I/*Eco*RI or 10 U of *Hind*III/*Xho*I (Thermo Scientific Fermentas) restriction enzymes for 2 h at 37°C for the cloning genes 1 to 6 and 7 to 8, respectively. The digested plasmid was dephosphorylated, i.e. the 5' phosphate is removed to prevent self-ligation (Sambrook *et al*, 1989), by adding 1 U of alkaline phosphatase (FastAP, Fermentas). Products were purified using the kit GFX PCR DNA purification (GE Healthcare).

The PRC products were cloned into plasmid pET21a (+) using 1 U of T4 DNA ligase (Thermo Scientific Fermentas) using several ratios of insert to vector. Mixtures were incubated overnight at room temperature (RT). Ligations were then submitted to a temperature of 65°C for 15 min to inactivate the ligase followed by dialysis on an MF^{TM} membrane (Milipore) during 30 min at RT.

2.3. Transformation and expression in *E. coli* cells

2.3.1. Transformation using heatshock

Chemocompetent cells were prepared as follows: a LA plate was streaked with a frozen stock of the appropriate *E. coli* strain and incubated overnight at 37°C. A single colony was picked and used to inoculate 20 mL of SOB medium which was then incubated overnight at 37°C with shaking at 150 rpm. Growth at a scale of 150 mL in SOB medium (in 1 L Erlenmeyers) was started with an optical density at 600 nm (OD_{600nm}) of 0.05 and incubated at 37°C and150 rpm. When OD_{600nm} reached ~ 0.6, cells were transferred to cold centrifuge vessels and spun down at 5000 × g for 10 min 4°C. The supernatant was discarded and the cell pellet washed with 75 mL of a sterile cold 100 mM $MgCl_2$ solution. A second centrifugation and cell wash with 75 mL of a sterile cold 100 mM $CaCl_2$ solution was performed as described above, followed by incubation on ice for 2 h. A third centrifugation at the previous conditions was performed and the cell pellet re-suspended in 1 mL of 100 mM $CaCl_2$ and 0.3 mL of 50% (v/v) glycerol. Aliquots of 100 µL were flash-frozen in liquid nitrogen and stored at - 80°C.

The host competent *E. coli* cells were transformed as follows: 10-50 ng of recombinant plasmid DNA were added to 100 µL of *E. coli* competent cells and the mixture was gently mixed and placed for 30 min on ice followed by a heat pulse at 42°C for 90 s and again placed on ice for 5 min. LB medium (1 mL) was added to the reaction mixture, which was incubated at 37°C for 1 h with shaking at 150 rpm. The cells were then centrifuged at ~ 2000 × g for 5 min, the supernatant was removed and the pellet was suspended in ~ 100 µL. This volume was

plated on LA plates supplemented with appropriate antibiotic and incubated at 37°C for about 16 h.

2.3.2. Transformation using electroporation

The electrocompetent cells were prepared by streaking a LA plate with a stock of the appropriate *E. coli* strain and incubated overnight at 37°C. A single colony was picked and used to inoculate 20 mL of SOB medium and incubated overnight at 37°C with shaking at 150 rpm. Growth at a scale of 500 mL in SOB medium was started OD_{600nm}= 0.05 and incubated at 37°C, with 150 rpm. When OD_{600nm} reached ~ 0.8, cells were transferred to cold centrifuge vessels and spun down at ~ 3000 × g for 15 min 4°C. The supernatant was discarded and the cell pellet washed with 500 mL of a sterile cold 10% glycerol solution. The described procedure was performed one more time. A final centrifugation was performed, and the cells re-suspended in ~ 1 mL of the glycerol solution (50% v/v). Aliquots of 100 µL were flash-frozen in liquid nitrogen and stored at - 80°C.

The host *E. coli* strains were transformed as follows: 10 - 50 ng of recombinant plasmid DNA were added to 100 µL of *E. coli* competent cells. The mixtures were then carefully transferred to a chilled, sterile electroporation cuvette with a 0.2 cm gap (BioRad). The reactions were electroporated in a GenePulser XCell^{TM} Electroporator (BioRad) (using the set C = 25 µF, PC = 200 Ω, V = 2.5 kV) and 1mL of LB medium was immediately added. The cells were incubated at 37°C for 1 h with shaking at 150 rpm. The cultures were plated on LA plates supplemented with appropriate antibiotic and incubated at 37°C for about 16 h. Dilutions of the culture were made (10^{-1} to 10^{-3}) in order to have single colonies.

2.3.3. Analysis of positive clones by colony PCR and extraction of plasmidic DNA

In order to confirm the success of cloning, single colonies were picked and swirled gently in 20 µL of MiliQ H_2O. The cells were then heated at 99°C for 5 min followed by centrifugation at ~ 12,000 g for 1 min to pellet down cell debris leaving DNA in suspension. One µL aliquot of these suspensions were added to a PCR reaction mixture, total volume of 50 µL, containing 20 nmol of each primer, 0.2 mM of dNTP's (NZYTech), 1.5 mM of $MgCl_2$ buffer and 0.5 U of Taq polymerase (Fermentas). PCR was carried out according to the parameters described in section 2.1.2. The PCR products were then analysed by agarose gel electrophoresis to confirm

the presence of a band corresponding to the amplified fragment showing the expected molecular mass. The positive colonies were picked again and inoculated in 10 mL LB medium containing the appropriate antibiotic and incubated at 37°C for about 16 h. The plasmidic DNA were isolated using the GeneJET^{TM} Plasmid Miniprep Kit (Fermentas). DNA elution was performed with 30 μL of MiliQ H_2O. The DNA sequences were confirmed by Sanger DNA sequencing (Stab Vida or GATC).

2.4. Heterologous expression of targeted genes

2.4.1. Cell growth of recombinant strains

Plasmids containing the genes of interest were transformed into several competent *E. coli* strains: BL21 (DE3), BL21star (DE3), Tuner (DE3) and Rosetta (DE3) strains (Table 2.1), in order to verify the most promising strain for expressing the gene and producing the heterologous enzyme. One colony of each recombinant strain was used to perform a pre-inoculum in 10 mL of LB supplemented with ampicillin (100 μg/mL) at 37°C overnight with agitation. An aliquot of the pre-inoculum was used to initiate the growth with an OD_{600}= 0.05 in 100 mL of LB medium. The growth conditions were 37°C with shaking at 150 rpm and growth was followed up to an OD_{600} of ~ 0.6. At that point 100 μM IPTG and 250 μM $CuCl_2$ (in the case of GalOx overexpression) were added to the culture medium, and the temperature was lowered to 25°C. For *galOx* overexpression the incubation was continued for an additional period of 6 h and then agitation was stopped and incubation under anaerobic conditions was carried for another 12 h; this procedure was performed in order to improve the incorporation of copper into the heterologous enzyme (Durao *et al*., 2008). In the case of *pOx* overexpression the incubation was carried for 18 h with shaking.

2.4.2. Cell disruption

Cells were harvested by centrifugation (~ 7000 × g, 10 min, 4°C), supernatants were discarded and the cellular pellet was suspended in 20 mM Tris-HCl buffer (pH 7.6) containing 5 mM $MgCl_2$, 10 μg/mL DNase I and 5 μg/mL of a mixture of protease inhibitors, antipain and leupain. The cell suspensions were disrupted in a French press (Thermo IEC) operating at 900/1000 psi in a process repeated 5 times, allowing the complete cellular rupture. The lysates

were then centrifuged (~ 25000 × g, 3 h, 4°C) to separate the non-soluble fraction (the pellet) from the soluble fraction (the supernatant).

2.4.3. Determination of protein concentration

Protein concentration was determined using the Bradford assay (Bradford 1976) with bovine serum albumin (BSA) as standard.

A calibration curve was performed with known concentrations of BSA: 0 - 0.6 mg/mL. 20 μL of sample was added to 1 mL of Bradford reagent (100 mg of Comassie Blue G dissolved in 45% of ethanol, 100 mL orthophosphoric acid and 750 mL of distilled water, for 1 L of solution), mixed and 5 min later, the absorbance was measured at 595 nm in a spectrophotometer Novaspec III (Amersham Biosciences). The measurements were performed in triplicate.

2.4.4. SDS-PAGE analysis

Protein production was examined with SDS-PAGE. Stock solutions used for SDS-PAGE gel were prepared as follows:

For the running gel: 12.5% (v/v) acrylamide, Lower Tris (375 mM Tris-HCl, 0.1% SDS, pH 8.8), 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulfate (APS), 0.06% (v/v) N, N, N', N'-tetramethylethylenediamine (TEMED).

For the stacking gel: 5% (v/v) acrylamide, Upper Tris (125 mM Tris-HCl, 0.1% SDS, pH 6.8), 0.1% (w/v) SDS, 0.1% (w/v) APS, 0.1% (v/v) TEMED.

The electrophoresis was performed at 220 V during 45 min in the Electrophoresis Buffer (192 mM glycine, 25 mM Tris, 0.1% SDS, pH 8.3). Before the application of the protein samples these were mixed with 2 × Sample Preparation Solution (2% SDS, 6.255 mM Tris-HCl pH 6.8, 5% 2-mercaptoethanol, 0.5 mM DTT, 5% glycerol, and 0.025% bromophenol blue) and denatured at 99°C for 20 min.

Gels were stained with Coomassie® blue staining solution (50% methanol, 10% acetic acid, 0.05% (v/v) Coomassie brilliant blue R-250). Slow agitation was provided for homogeneous staining. Destaining of the gels was performed by using a solution of 10% methanol and 10% acetic acid.

2.4.5. Enzyme activity assays

The activity of GalOx and POx was carried out using a horseradish peroxidase (HRP) coupled assay (Baron *et al*, 1994) in which the putative hydrogen peroxide generated by the GalOx and POx reactions is used as substrate by HRP to oxidise two molecules of 2,2'-Azino-bis(3-ethylbenzthiazoline-6-sulfonic acid (ABTS) generating a coloured (green) product that could be easily monitored at 420 nm (ε = 36,000 $M^{-1}cm^{-1}$). Reactions were performed in a Synergy2 microplate reader (BioTek, Vermont, USA). One Unit of GalOx activity was defined as the amount of enzyme that is necessary for the consumption of 2 μmol of ABTS per min, which equals the consumption of 1 μmol of O_2 per min.

Activity tests in a range of different conditions were performed to assess the activity of the GalOx and POx. Buffers: 100 mM acetate pH 4 and 5, 100 mM phosphate pH 6 and 7. Substrates: D-galactose (10-1000 mM), D-glucose (10-1000 mM), glycerol (150-1500 mM) and methylglyoxal (1-100 mM). Copper (Cu(II)) was added at different concentrations (0.1, 0.25, 0.5 mM) as well as the oxidants Ferricianide (0.1; 0.25 and 1 mM), Iridate (20 μM) and $MgCl_2$ (1mM). The incubation of cell extracts or protein preparations (for ~ 30 min) with Cu^{2+} (5-fold protein concentration) and with Ferricianide (5 fold protein concentration), followed by dialyses, was also tested. The concentration of ABTS used was 1 mM and HRP was added at different amounts: 1, 5, 10 and 20 U. Other standard controls were also performed: Positive control of HRP (addition of ABTS, H_2O_2 and HRP, to verify the activity of HRP; negative control on GalOx reducing substrate (addition of ABTS, cell extracts and HRP, to verify that no reaction occurs when no substrate was added); negative control on cell extracts (addition of GalOx reducing substrate, ABTS and HRP, to verify that no reaction occurs when no cell extracts were added) and negative control on HRP (addition of GalOx reducing substrate, ABTS and cell extracts, to verify that no coloured product is formed without addition of HRP).

2.4.6. Enzyme production and purification

For enzyme purification cell cultures at larger scales were performed (1 and 2 L of culture medium in 5 L Erlenmeyers). One single colony was used to perform a pre-inoculum in 100 mL of LB supplemented with ampicillin (100 μg/mL) at 37°C, overnight, at 150 rpm. An aliquot of the pre-inoculum was used to initiate growth at an OD_{600} = 0.05 in 1 L of LB medium. The growth conditions were 37°C with shaking at 150 rpm and growth was followed up to an OD_{600} of 0.6. At that point 100 μM IPTG and 250 μM $CuCl_2$ (only for GalOx) were added to the

culture medium, and the temperature was lowered to 25°C. For GalOx incubation was continued for an additional period of 6 h, at that point agitation was stopped and incubation carried for another 12 h under anaerobic conditions. In the case of POx production the incubation was carried for an additional period of 18 h with shaking.

Recombinant enzymes were purified from crude extracts using ÄKTA purifier system (GE HealthCare, BioSciences, Uppsala, Sweden) using the SP-Sepharose, Q-Sepharose, Hi-Load 16/60 Superdex 75 prep-grade and Hi-Load 16/60 Superdex 200 prep-grade columns. All steps described below were performed at room temperature and the elution was always monitored at 280 nm for total protein content. Crude extracts were filtered with a 0.2 μm membrane to remove large particles in suspension prior to loading into the columns.

2.4.6.1. *M. stipitatus* hypothetical Galactose oxidase (hGalOxMs2)

This recombinant enzyme was only partially purified. Crude extracts were loaded on to a SP-Sepharose column (anionic exchange chromatography column) equilibrated with 20 mM Tris-HCl, pH 7.6, and eluted with a NaCl gradient from 0 to 0.5 M in 5 column volumes and from 0.5 to 1 M in 2 column volumes. A SDS-PAGE gel was performed and the fractions containing the desired protein were pooled together and concentrated using an amicon (Ultra-15 Centrifugal Filter Units) with a cut off of 30kDa and stored at - 20°C.

2.4.6.2. *G. violaceus* Galactose oxidase (GalOxGv)

This recombinant enzyme was only partially purified. Crude extracts were loaded on to a Q-Sepharose column (cationic exchange chromatography column) equilibrated with 20 mM Tris-HCl, pH 7.6, and eluted with a NaCl gradient from 0 to 0.5 M in 5 column volumes and from 0.5 to 1 M in 2 column volumes. A SDS-PAGE gel was performed and fractions containing the desired protein were pooled together and concentrated using an amicon (Ultra-15 Centrifugal Filter Units)with a cut off of 30kDa and stored at - 20°C.

2.4.6.3. *A. siccitolerans* Galactose oxidase (GalOxAs)

Crude extracts were loaded onto a Q-Sepharose column equilibrated with 20 mM Tris-HCl buffer, pH 7.6, and eluted with a NaCl gradient from 0 to 0.5 M in 5 column volumes and from 0.5 to 1 M in 2 column volumes. Other conditions were also tested,: Q-Sepharose column

equilibrated with different buffers (10 mM Tris-HCl buffer, pH 7.6; 5 mM Tris-HCl buffer, pH 7.6; 20 mM Tris-HCl buffer, pH 8.6; 5 mM Tris-HCl buffer, pH 9) using the previous described NaCl gradient; an SP-Sepharose column equilibrated with 20 mM Tris-HCl buffer, pH 7.6, was also used applying the the previous mentioned NaCl gradient. Identification of the recombinant enzyme using SDS-PAGE and activity tests were performed to assign the fractions to be pooled together. The pooled active fractions were concentrated to 1 mL and then applied on to a size-exclusion Hi-Load 16/60 Superdex 200 prep-grade column (GE HealthCare, BioSciences) previously equilibrated with 20 mM Tris-HCl buffer, pH 7.6 with 0.2 M NaCl.

2.4.6.3.1. Inclusion body refolding of GalOxAs

Soluble protein was recovered from inclusion bodies with the method described by Bollag *et al*, (1996). Inclusion bodies were washed with 20 mM Tris-HCl buffer, pH 7.6, containing 0.5% Triton X-100, and spun down by centrifugation (25,000 × g, 15 min, 4°C). Inclusion bodies were unfolded in 20 mM Tris-HCl buffer, pH 7.6, containing 8 M of urea, 2 mM of reduced glutathione (GSH) and 0.2 mM of oxidized glutathione (GSSG) in a volume so that the final protein concentration was below 2.5 mg/mL, for 1 h at RT. Refolding was performed by slowly adding 9 mL for each mL of urea-protein solution of 20 mM Tris-HCl buffer, pH 7.6, containing 2 mM of GSH and 0.2 mM of GSSG and 50 μM of CuCl2 for 2-4 h at RT. Dialysis in a Diaflow (Millipore Amicon stirred ultrafiltration filtration cell filter model 8200 200) with a 30 kDa membrane was performed for removing the urea using 20 mM Tris-HCl buffer, pH 7.6, 0.4 M NaCl to prevent further protein aggregation.

2.4.6.4. *A. siccitolerans* Pyranose oxidase (POxAs)

Crude extracts were loaded onto a Q-Sepharose column equilibrated with 20 mM Tris-HCl buffer, pH 7.6, and eluted with a NaCl gradient from 0 to 0.5 M in 5 column volumes and from 0.5 to 1 M in 2 column volumes. Other conditions were also tested, such as using a Q-Sepharose column initialy equilibrated with 10 mM of Tris-HCl buffer, pH 7.6 and a NaCl gradient using 1 M NaCl solution: 5 column volumes between 0-50%, 2 column volumes between 50-100%; initial equilibration with 20 mM of Tris-HCl buffer, pH 7.6 and a gradient using 1 M NaCl: 2 column volumes 0-30%, 5 column volumes 30-60% and 2 column volumes 60-100%. Identification of the recombinant protein using SDS-PAGE and activity tests was performed to confirm the fractions to be pooled. The pooled active fractions were concentrated to 1 mL and

then applied on to a size-exclusion Hi-Load 16/60 Superdex 75 prep-grade column (GE HealthCare, BioSciences) equilibrated with 20 mM Tris-HCl buffer, pH 7.6 with 0.2 M NaCl.

2.5. Enzyme characterization of *A. siccitolerans* Galactose oxidase (GalOxAs) and Pyranose oxidase (PoxAs)

2.5.1. Spectroscopic analysis

The UV-visible absorption spectra of enzymes were recorded on a Nicolet Evolution 300 spectrophotometer from Thermo Industries (Madison, USA) at room temperature. GalOxAs protein samples in 20 mM Tris-HCl buffer, pH 7.6, 0.4 M NaCl were chemically oxidized by incubation in the presence of 10-100 mM of ferricyanide or 100 mM of iridate at RT or on ice for 5-30 min. The oxidants were subsequently removed by diafiltration using Amicon®Ultra centrifugal filters (30K).

2.5.2. Kinetic analysis

The enzymatic activity of GalOxAs was monitored using either a Nicolet Evolution 300 spectrophotometer (Thermo Industries, Madison, USA) or a Synergy2 microplate reader (BioTek, Vermont, USA). All enzymatic assays were performed at least in triplicate. A pH-activity profile was determined using Britton-Robinson buffer (100 mM phosphoric acid, 100 mM boric acid, 100 mM acetic acid mixed with NaOH to the desired pH in the range of 2-11). Temperature profile was achieved by measuring the activity in a range of temperatures between 20-60°C in 100 mM phosphate buffer, pH 8.0. Activity measurements were performed, unless otherwise stated, by using a standard assay (250 mM D-galactose; 1 mM ABTS; 15 U HRP and adequate quantity of enzyme at 25°C in 100 mM phosphate buffer, pH 8.0). Apparent steady-state kinetic constants (k_{cat} and K_m) were determined for GalOxAs at 25°C in 100 mM phosphate buffer, pH 8.0 for different sugar substrates: D-galactose (1-200 mM), D-raffinose, α-lactose (1-250 mM), L-arabinose (10-500 mM) and glycerol (10-2500 mM). Kinetic data were fitted directly using to the Michaelis-Menten equation (Origin-Lab software, Northampton, MA, USA). POxAs activity measurements were performed using the following assay: 250 mM of D-glucose; 1 mM ABTS; 15 U HRP and adequate quantity of enzyme at 25°C in 100 mM phosphate buffer, pH 7.0.

2.5.3. Enzyme stability

The enzymes were incubated at 40°C in 20 mM Tris-HCl buffer, pH 7.6, and at fixed time intervals, samples were withdrawn and tested for activity following the ABTS oxidation at 25°C. The assays were performed on a Nicolet Evolution 300 spectrophotometer from Thermo Industries (Madison, USA).

2.5.4. Other methods

The GalOxAs and POxAs proteins were identified by MALDI-TOF/TOF after tryptic digestion using a 4800plus MALDI-TOF/TOF (AB Sciex) mass spectrometer, at Instituto de Tecnologia Química e Biológica. The molecular mass of GalOxAs was determined on a gel filtration Superose 12 HR10/30 column (Amersham Biosciences) equilibrated with 20 mM Tris-HCl buffer, pH 7.6 containing 0.2 M NaCl. Ribonuclease (13.7 kDa), chymotrypsinogen A (25 kDa), ovalbumin (43 kDa), albumin (67 kDa) and aldolase (158 kDa) were used as standards. For GalOxAs the copper content was determined through the trichloroacetic acid/bicinchoninic acid (BCA) method (Brenner and Harris, 1995).

3. RESULTS AND DISCUSSION

3.1. Bioinformatic analysis

As a starting point for the bioinformatics analysis, we performed a Basic Local Alignment Search Tool (BLAST) using all bacterial genomes available (appendix B) for identifying bacterial peroxide forming enzymes potentially involved in lignin degradation as auxiliary enzymes of high-redox peroxidases. The fungal auxiliary peroxide forming enzymes used as template have the crystal structure solved, with exception of glyoxal oxidase for which no structure is available, and were the following: glyoxal oxidase from *Phanerochaete chrysosporium* (AAA87595); galactose oxidase from *Fusarium graminearum* (P0CS93; PDB:1GOF); pyranose oxidase from *Tricholoma matsutake* (Q8J2V8; PDB:1TT0) and *Trametes multicolour* (AAP40332; PDB:2IGO); aryl-alcohol oxidase from *Pleurotus eryngii* (O94219; PDB:3FIM).

3.1.1. Glyoxal/Galactose Oxidase

As mentioned in the Introduction section both glyoxal and galactose oxidases contain a catalytic motif based in an unusual tyrosyl-cysteine thioether cross-linked redox cofactor forming a stable amino acid side chain radical (Whittaker and Whittaker, 1988). The active site of GlyOx is spectroscopically similar to that of GalOx, except for the tryptophan overlying the Tyr-Cys site in GalOx that is replaced by a histidine residue in GlyOx (Whittaker *et al*., 1996). The structure of *F. graminearum* galactose oxidase revealed the basic geometric details of the enzyme active site: a β-propeller (or *kelch*) architecture formed from seven four-stranded antiparallel β-sheets, with three of the modules contributing to essential metal ligands (Ito *et al*, 1991).

In this study the BLAST search was performed and compared for both GalOx and GlyOx (Fig 3.1.) based in matching motifs, predominantly the *kelch* motifs, and active site residues.

The BLAST search tool allowed identifying a large number of bacterial genes (appendix B) with similarity to GalOx and GlyOx. However, at the protein sequence level, the similarity seems to be higher with GalOx than GlyOx (appendix B). This seems to indicate the fact that bacteria only have GalOx enzymes and not GlyOx, but without a crystal structure of GlyOx no serious assumptions can be made at this respect.

Two genes were identified showing the typical motifs of galactose and glyoxal oxidases in the genome of *M. stipitatus* DSM 14675 (Genbank accession No. NC_020126): 1) MYSTI_05030, an Open Reading Frame (ORF) encoding a 925 amino acid (a.a.) polypeptide having 23% and 30% identity to the fungal GlyOx and GalOx, respectively, renamed hGalOxMs-1 and 2) MYSTI_00845 an ORF encoding a 472 a.a. having 24% and 30% identity to the fungal GlyOx and GalOx, respectively, renamed hGalOxMs-2.

In the genome of *G. violaceus* PCC 7421 (Genbank accession No. NC_005125) the gene gll4251 (an ORF encoding a 761 a.a. polypeptide) was identified having 25% and 28% identity to the fungal GlyOx and GalOx, respectively, renamed GalOxGv.

In the genome of *S. clavuligerus* ATCC 27064 (Genbank accession No. NZ_CM000913) the gene SCLAV_5541 (an ORF encoding a 796 a.a. polypeptide) was identified having 23% and 27% identity to the fungal GlyOx and GalOx, respectively, renamed GalOxSt.

In the genome of *S. aurantiaca* DW 4/3-1 (Genbank accession No. NC_014623) two genes were identified 1) STAUR_4893 (an ORF encoding a 807 a.a. polypeptide) having 24% and 29% identity to the fungal GlyOx and GalOx, respectively, renamed hGalOxSa and 2) STAUR_6507 (an ORF encoding a 855 a.a. polypeptide) was identified having 18% and 39% identity to the fungal GlyOx and GalOx, respectively, renamed GalOxSa.

In the genome of *A. siccitolerans* (Genbank accession No. PRJEB1189) the gene ARTSIC4J27_1011 was identified as an ORF encoding a 760 a.a. polypeptide having 22% and 47% identity to the fungal GlyOx and GalOx, respectively, renamed GalOxAs.

We also identified a putative galactose oxidase (SSFG_04533) in the genome of *S. viridosporus* T7A, a bacterium reported to react with lignin to both solubilize and produce a high molecular weight metabolite APPL (Kirby, 2005) as described in section 1.2.

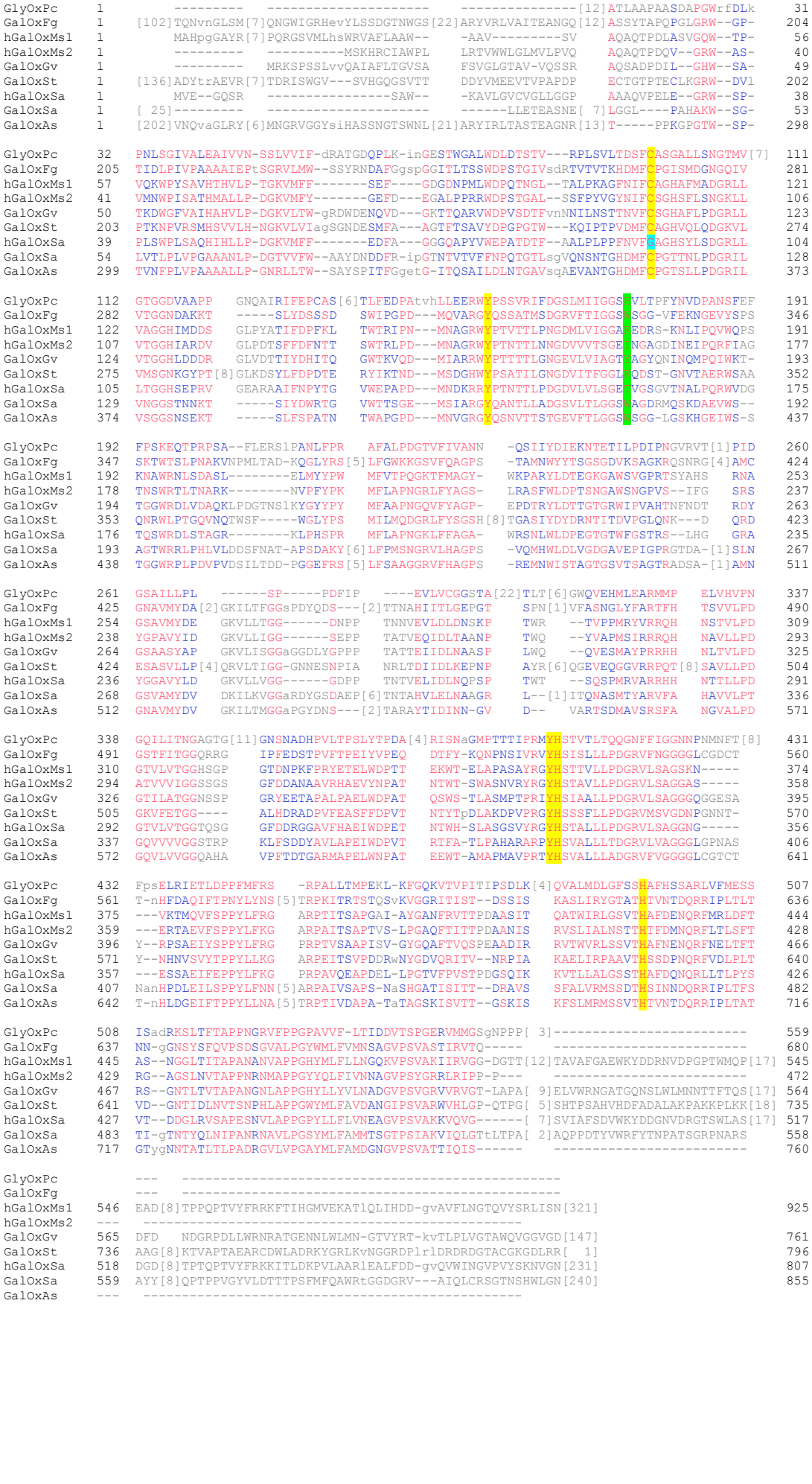


Figure 3.1. Amino acid multiple sequence alignment of GalOx generated by Constraint-based Multiple Alignment Tool. GlyOxPc: glyoxal oxidase of *P. chrysosporium*; GalOxFg: galactose oxidase of *F.graminearum*; hGalOxMs1: hypothetical galactose oxidase 1 of *M. stipitatus* DSM 14675; hGalOxMs2: hypothetical galactose oxidase 2 of *M. stipitatus* DSM 14675; GalOxGv: galactose oxidase of *G. violaceus* PCC 7421; GalOxSt: galactose oxidase of *S. clavuligerus* ATCC 27064; hGalOxSa: hypothetical galactose oxidase *S. aurantiaca* DW 4/3-1; GalOxSa: galactose oxidase *S. aurantiaca* DW 4/3-1; GalOxAs: galactose oxidase *A. siccitolerans*. Multiple sequence alignment columns with no gaps are colored in blue or red. The red color indicates highly conserved columns and blue indicates less conserved ones. Active site residues are highlighted in yellow for matching residues, light blue for non-matching residues and in green for the only active site residue that differs between GlyOx and GalOx.

The sequence alignment (Fig 3.1.) reveals that the active site residues that are identical in GalOxFg and GlyoxPc (Cys, Tyr, Tyr, His, His, marked in yellow) are present in all the analysed sequences, with exception of hGalOxSa that shows a glycine instead of the putatively conserved cysteine (marked in light blue). Additionally, we can verify that the sequences that have the higher percentage of identity to GalOxFg, namely GalOxSa (39%) and GlaOxAs (49%) also have conserved tryptophans (marked in green) and are thus more likely to share GalOx instead of GlyOx features. In the remaining 5 sequences neither a tryptophan nor a histidine are observed at that position. However, since it is still unclear which are the exact structural features behind the catalytic differences between GalOx and GlyOx and the exact details of the active site of GlyOx we cannot rule out the existence of other residues key different from Trp and His and/or additional structural determinants of specificity.

3.1.2. Pyranose Oxidase

The crystal structure of POx of *T. multicolour* ligated to the poor substrate 2-deoxy-2-fluoro-D-glucose (Kujawa *et al*., 2006) allowed the identification of the conserved active-site residues that most likely take part in the substrate(s) binding. These residues are in POx of *T. multicolour*, (in brackets the POx of *T. matsutake*): Gln448 (Gln392), Asp452 (absent, corresponds to a Tyr396), Arg472 (Arg418), His548 (His498), and Asn593 (Asn532). In addition, the residues His548 (His489), Asn593 (Asn532), Thr169 (160) and His167 (His158) were suggested to be putatively important for catalysis and covalent binding of the prosthetic group (Kujawa *et al*., 2006).

The BLAST alignment was performed with all the sequences of fungal POx crystal structures available and in the genome annotation for *A. siccitolerans* (Genbank accession No. PRJEB1189) the gene ARTSIC4J27_4061 was identified as an ORF encoding a 519 a.a. polypeptide having 32% identity, renamed POxAs (appendix A) (Fig 3.2.). In Fig. 3.2 the sequence is displayed along *T. matsutake* Pox since this is one showing higher similarity to POxAs.

The BLAST search tool allowed us to identify other bacterial genes with similarity to fungal POx. For example a putative POx (WP_020420830) from *Amycolatopsis* sp. 75iv2 (as referred in the section 1.2. a bacteria previously reported to be potentially involved in lignin degradation (Brown *et al*., 2011; Davis *et al*., 2012)) suggesting once again that these peroxide forming enzymes are involved in lignin degradation.

```
POxTm    1    MPIRLSKEKINDLLQRSQGDLTSSQDEIVHYTDVFIAGSGPIACTYARHIIDNTSTTKVYMAEIGSQDNPVIGAHHRNSI  80
POxAs    1    M----------------------SGHRYPAAVDVAIVGSGPTASAYARILSEEAPGATIAMFEVGPTVSNPPGAHVKNI-  57

POxTm    81   KFQKDTDKFVNIINGALQPISISP---SDTYQPTLAVAAWAPPIDPAEGQLVIMGHNPNQEAGLNLPGSAVTRTVGGMAT  157
POxAs    58   ---EDPDS-----RSLAQRASEGPGAGAATVNSPGAVKSGERRARPGT-YLLQDGYAFPGEDGM--PVAAMSSNVGGMAA  126

POxTm    158  HWTCACPTPHDEERVNNPVDKQEFDALLERAKTLLNVHSDQYDDSIRQIVVKETLQQTLDASRG----VTTLPLGVERRT  233
POxAs    127  HWTAACPRPGGKERIPFLPDLEE---LLNDADRLLGVTTHAFDGAPFSDLVRERLAAVVDQGRTPAFRVQPMPLAVHRRQ  203

POxTm    234  DNPIYVTWTGADTVLGDVPK-SPRFVLVTETRVTKFIVSETNPTQVVAALLRNLNTSNDELVVAQSFVIACGAVCTPQIL  312
POxAs    204  DGAL--VWSGSDVVMGEATRDNPQFELFDESLVTRVLVEDGTAAGVEVQDRRSGDTYQ---VAARYVVVGADALRTPQLL  278

POxTm    313  WNSNIRPHALGRYLSEQSMTFCQIVLKRSIVDSIATDPRFAAKVEAHKKKHPDDVLPIPFHEPEPQVMIPYTSDFPWHVQ  392
POxAs    279  WASGIRPDALGRYLNDQA----QVVFASRLRDVQPEDAPAAANGA--------------LSEQSGVAWVPYTDEAPFHGQ  340

POxTm    393  VHRYAFGDVGPKAD-----PRVVVDLRFFGKSDIVEENRVTFGPNPKLRDWEAGVTDTYGMPQPTFHVKRTNADGDRDQR  467
POxAs    341  IMQLDASPV-PLADDDPIVPGSIVGLGLFCAKDLQREDRVAFDDDTR---------DSYGLPAMRIHYRLTERDHVVLDR  410

POxTm    468  MMNDMTNVANILGGYLPGSYPQFMAPGLAQHITGTTRIG-TDDQTSVADPTSKVHNFDNLWVGGNGCIPDATACNPTRTS  546
POxAs    411  ARQEIVRLGKAVGEPLD-ERPFVLPPGASLHYQGTTRMGETDDGESVCSPDSQVWQVPGLFVAGNGVIPTATACNPTLTS  489

POxTm    547  VAYALKGA----EAVVSYLGVS--------                                                   564
POxAs    490  VALAVRGARKIAEEITSSLLMSESDNRLSK                                                   519
```

Figure 3.2. Amino acid multiple sequence alignment of POx generated by Constraint-based Multiple Alignment Tool. POxTm: pyranose oxidase of *T. matsutake*; POxAs: pyranose oxidase of *A. siccitolerans*. Active site residues are highlighted in yellow for matching residues and light blue for non-matching residues.

The alignment reveals (Fig 3.2.) that sequences from *A. siccitolerans* and f *T. matsutake* POx's are highly conserved and that all residues responsible for the catalysis and covalent binding of the prosthetic group match between POxTm and POxAs (His540, Asn574, Thr129 and His127).

Additionally, most of the active-site residues identified in the crystal structure of POx of *T. multicolour* (AAP40332; PDB:2IGO) that directly take part in substrate binding are conserved in POxAs (Gln340, His440 and Asn474, marked in yellow).

The active site residue Asp345 (marked in yellow) is not conserved in *T.matsutake* POx, corresponding to an Ala597 (marked in light blue), but it is conserved in *T. multicolour* POx. Gly366 (marked in light blue) is the only POxAs residue that is not conserved neither in *T. matsutake* POx nor in *T. multicolour* POx corresponding to an Arg in these enzymes.

3.1.3. Aryl-alcohol oxidase

No ORF was identified on the available genomes for aryl-alcohol oxidase.

3.2. Cloning of the selected genes

3.2.1. *M. stipitatus* Galactose oxidase-1 (hGalOxMs1)

A range of annealing temperatures between 56 and 70°C were tested for the PCR amplification of the gene coding for hGalOxMs-1 (2778 bp) from *M. stipitatus* based on the theoretical melting temperature (Tm). However, the persistent appearance of nonspecific bands (Fig 3.3.) in the agarose gel lead to a number of troubleshooting procedures. The following causes/procedures were considered: (i) non appropriate concentration of DNA template: several concentrations of DNA template were tested varying from 10 to 100 ng; (ii) a denaturation time too short and/or temperature too low: a number of increments in time (5 s) and temperature (0.5°C) were tested; (iii) an extension time too short: increments of 1 min were added to the extension time; (iv) cycle number might be too high: reduction of the number of cycles in decrements of 2 cycles was performed; and finally (v) nonspecific binding of primers on DNA template/ addition of 10% dimethyl sulfoxide (DMSO) in the reaction mixtures was considered since DMSO inhibits secondary structure formations in DNA templates or primers.

However in spite of all the attempts to solve the problem the nonspecific bands persisted in the agarose gels after the PCR amplification steps. Therefore, a different strategy was followed to clone *hGalOxMs1*: the band showing the desired molecular mass was excised from the gel, purified and a new PCR amplification was performed using as DNA template the excised DNA. Nonetheless, the amplification resulted once again in the appearance of nonspecific bands and we have concluded that this should be caused by the presence of (i) hairpins or other DNA conformations in the region that is ought to be cloned or close to it, or else, (ii) the primers designed were not appropriate to amplify the target sequence. A new set of primers were designed but unfortunately their use resulted in the absence of amplification. Considering that after several attempts using different approaches, the PCR amplification problem could not be successfully solved, we ended up not able to clone this gene.

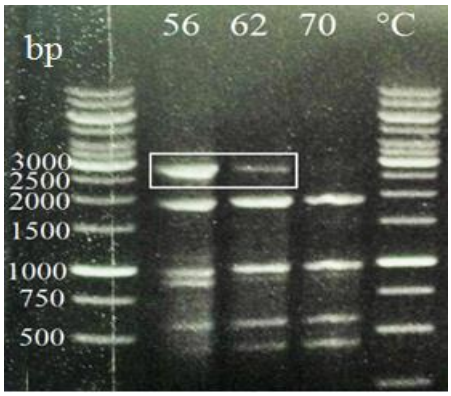


Figure 3.3. Ethidium bromide-stained agarose gel showing amplification of the gene *hgalOxMs1* (2778 bp) in PCR using different annealing temperatures.

3.2.2. *M. stipitatus* hypothetical Galactose oxidase-2 (hGalOxMs2)

An empirical search for the annealing temperature (60 - 66°C) was performed based on the theoretical melting temperature and hGalOxMs2 gene amplification was observed for temperatures between 60 - 63°C. Modification of the PCR protocol was performed due to the appearance of unspecific bands on the agarose gel (Fig 3.4. A) and several attempts performed as described above in 3.2.1 in order to try to solve the problem. One hypothesis was that the designed primers lacked appropriate specificity to the target DNA. Therefore, longer primers were designed and tested and the PCR amplification using the new forward and reverse primers (Table 2.3.) successfully amplified the predicted 1419 bp gene fragment (Fig 3.4. B) using the conditions described in the methods section 2.2. (Table 2.4.). The gene was successfully inserted in to the pET21a (+) expression vector. The cloning of the gene was confirmed with DNA sequencing analysis (appendix C).

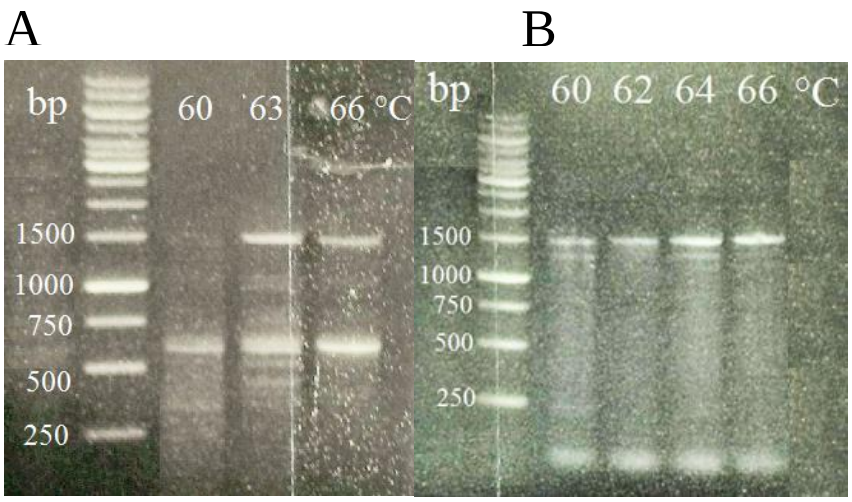


Figure 3.4. Ethidium bromide-stained agarose gel showing amplification of genes *hgalOxMs-2* (1419 bp) using a different set of primers (A) shorter and (B) longer primers (Table 2.3).

3.2.3. *G. violaceus* Galactose oxidase (GalOxGv)

Based on the theoretical Tm of the gene coding for GalOxGv, an empirical search for annealing temperatures was performed (53 - 60°C) and all the temperatures used led to gene amplification. The PCR amplification resulted in the accumulation of a gene fragment with the predicted 2286 bp (Fig 3.5.) using the conditions mentioned in the method section 2.2. (Tables 2.3 and 2.4.) The gene was successfully inserted in to the pET21a (+) expression vector. The cloned gene was confirmed with DNA sequencing (appendix C).

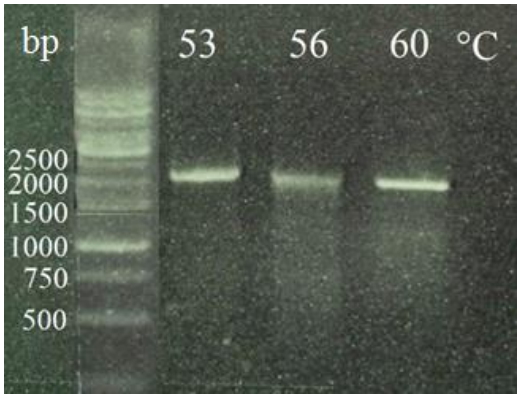


Figure 3.5. Ethidium bromide-stained agarose gel showing amplification of the gene *galOxGv* (2286 bp) in PCR using different annealing temperatures.

3.2.4. *S. clavuligerus* Galactose oxidase (GalOxSt)

The *galOxSt* gene was successfully amplified by PCR using annealing temperatures between 64 – 66°C with the predicted 2391 bp gene size (Fig 3.6.) using the designed forward and reverse primers (Table 2.3.) in the conditions described in the Material and Methods section 2.2. (Table 2.4.). Numerous attempts always unsuccessful were made in order to insert the gene fragment into the pET21a (+) cloning vector. Several problems were identified; a low number of transformants were observed and all colonies tested showed the cloning vector without insert. Presumably this was caused by inefficient ligations that lead to an increased susceptibility to recombination i.e. self ligation of the vector or to the inefficient action of the restriction enzyme(s), which did not cleave completely the amplified fragment. Therefore a number of actions were taken: (i) individual digestion of the insert and of the amplified fragment; (ii) variation of the molar ratio of insert to vector from 1:1 to 1:10; (iii) different temperatures (4 to 25°C for a more efficient ligation), and time of incubation (8 to 16 h) were tested and, finally, (iv) two different *E. coli* strains were used KRX and DH5α, as host strains. In spite of all the efforts described it was not possible to clone this gene in *E. coli*.

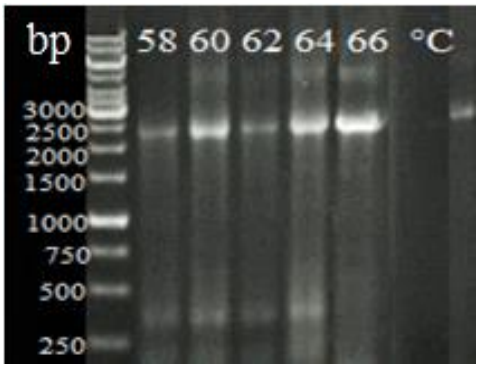


Figure 3.6. Ethidium bromide-stained agarose gel showing amplification of the gene *galOxSt* (2391 bp) in PCR using different annealing temperatures.

3.2.5. *S. aurantiaca* hypothetical Galactose oxidase (hGalOxSa)

Based on the theoretical melting temperature of the gene coding for hGalOxSa, an empirical search for annealing temperatures was performed (60 - 70°C) and a gene amplification with the predicted 2424 bp was observed using the annealing temperatures between 60 - 66°C (Fig 3.7.) with the designed forward and reverse primers (Tables 2.3.) under the conditions described in the Material and Methods section 2.2. (Table 2.4.). This gene was cloned successfully into the pET21a (+) expression vector. The cloning of the gene was confirmed with DNA sequencing analysis (appendix C).

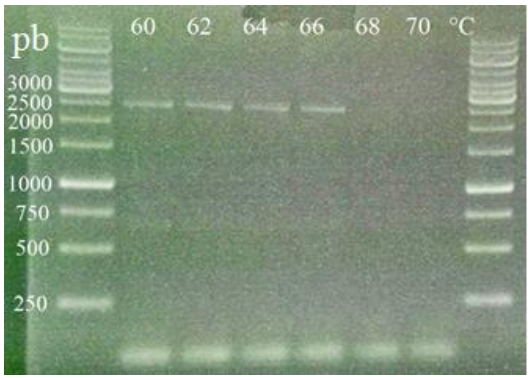


Figure 3.7. Ethidium bromide-stained agarose gel showing amplification of the gene *hgalOxSa* (2424 bp) in PCR using different annealing temperatures.

3.2.6. *S. aurantiaca* Galactose oxidase (GalOxSa)

The *galOxSa* gene was successfully amplified by PCR using annealing temperatures between 60 – 66°C, with the predicted 2568 bp gene size (Fig 3.8.) using the designed forward and reverse primers (Table 2.3.) in the conditions described in the Material and Methods section 2.2. (Table 2.4.). However, all the attempts to clone this gene in the expression vectors available were unsuccessful. The troubleshooting procedures described for the cloning of the previous *galOxSt* gene were also followed, as the problems found were similar, ending without success.

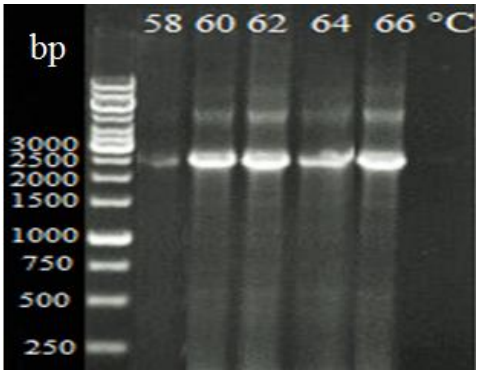


Figure 3.8. Ethidium bromide-stained agarose gel showing amplification of the gene *galOxSa* (2568 bp) in PCR using different annealing temperatures.

3.2.7. *A. siccitolerans* Galactose oxidase (GalOxAs)

Based on the theoretical Tm of the gene coding for GalOxAs an empirical search for annealing temperatures (62 - 70°C) was performed and gene amplification was observed for temperatures between 66 – 70 °C. PCR amplification of the predicted 2283 bp gene fragment (Fig 3.9.) was carried out using the forward and reverse primers (Table 2.3.) using the conditions mentioned in the Material and Methods section 2.2. (Table 2.4.). The gene was successfully inserted in to the pET21a (+) expression vector. The cloning of the gene was confirmed using DNA sequencing analysis (appendix C).

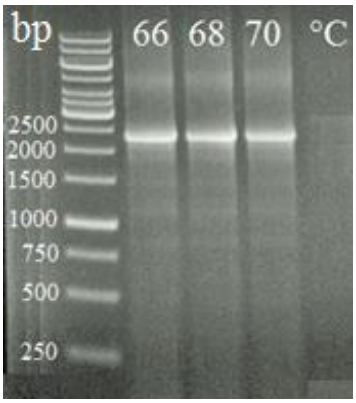


Figure 3.9. Ethidium bromide-stained agarose gel showing amplification of the gene *galOxAs* (2283 bp) in PCR using different annealing temperatures.

3.2.8. *A. siccitolerans* Pyranose oxidase (POxAs)

The *pOxAs* gene was successfully amplified by PCR using annealing temperatures between 60 – 68°C (Fig 3.10.). This led to the appearance of the predicted 1560 bp gene fragment in the agarose gel (Table 2.4.) The gene was successfully inserted in to the pET21a (+) expression vector and the presence of the cloned gene was confirmed with DNA sequencing analysis (appendix C).

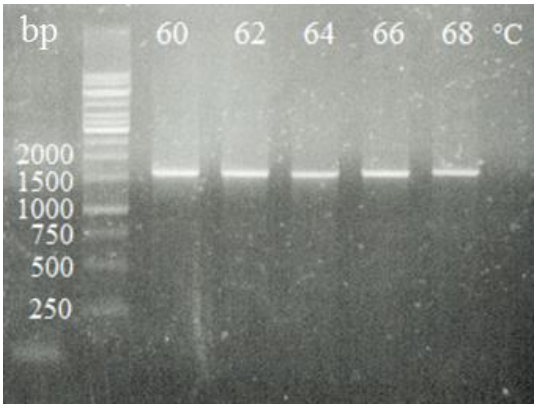


Figure 3.10. Ethidium bromide-stained agarose gel showing amplification of the gene *pOxAs* (1560 bp) in PCR using different annealing temperatures (60, 62, 64, 66 and 68°C).

3.3. Overexpression of selected genes in *E. coli* cells

3.3.1. *M. stipitatus* hypothetical Galactose oxidase-2 (hGalOxMs2)

Four different protein expression systems were used *E. coli* BL21 (DE3), BL21 Star (DE3), Tuner (DE3) and Rosetta (DE3), to examine the levels of expression of *hgalOxMs-2*. We can point out that *E. coli* BL21 Star was the only strain that revealed expression of the recombinant protein but only in the insoluble fraction. SDS-PAGE analysis of crude extracts from the recombinant cells revealed that upon addition of IPTG, a major band of around 51 kDa, absent in extracts prepared from non-induced cells, appeared in the insoluble fraction of cell lysates, most probably in the form of inclusion bodies (Fig 3.11.).

However, low expression levels were observed. A DNA sequencing analysis of the gene was performed in order to assess the presence of a signal peptide in its sequence, which many times results in lower levels of heterologous recombinant proteins in *E. coli* (this analysis using SignalP 4.1 was additionally performed for all the cloned genes). The analysis pointed to the presence of a signal peptide with 30 bp long. NZYtech cloning services cloned the gene without the signal peptide. More tests to assess the levels of expression were performed using the four different protein expression systems mentioned above but no differences were observed: the expression occurred in *E. coli* BL21star at similar levels as the full-length gene and in the insoluble fraction.

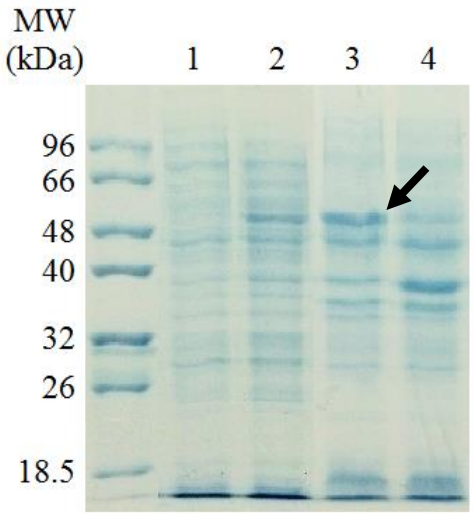


Figure 3.11. SDS-PAGE analysis of hGalOxMs2 overproduction in *E. coli* BL21 star. Lane 1: soluble fraction of an IPTG induced crude extract; Lane 2: soluble fraction of a non-induced crude extract; Lane 3: insoluble fraction of an IPTG induced crude extract; Lane 4: insoluble fraction of a non-induced crude extract. The arrow indicates the band corresponding to the recombinant protein (51 kDa).

Enzymatic activity assays were performed on both soluble and insoluble fractions, of the IPTG induced crude extracts using the HRP coupled assay in a range of different conditions as described in the methods section 2.4.5. Despite of the variety of conditions tested no activity was detected both in soluble and insoluble fractions. One step of purification was performed to rule out the possibility of activity inhibition due to contaminants present in crude extracts using an SP-Sepharose column, using the soluble fraction of *E. coli* BL21star crude extracts (Fig 3.12.).

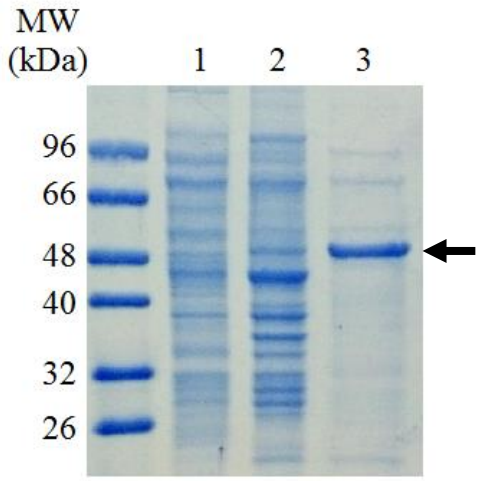


Figure 3.12. SDS-PAGE analysis of hGalOxMs-2 overproduction in *E. coli* BL21 Star. Lane 1: soluble fraction of an IPTG induced crude extract; Lane 2: insoluble fraction of an IPTG induced crude extract; Lane 3: sample after the SP-sepharose chromatography. The arrow indicates the size of the recombinant protein (51 kDa).

This purification step resulted in a semi purified enzyme solution with the expected size of 51 kDa. HRP coupled activity assays were performed using this semi-purified fraction using a large range of conditions (described in the methods section 2.4.5) but no activity was detected.

The band that corresponded to protein of interest was excised and identified by MALDI-TOF/TOF-MS after tryptic digestion. The result from this analysis revealed with a score of (1100) and a coverage of (57) that the protein was Trigger factor of *E. coli* (appendix D). No further experimental work was performed with this recombinant protein.

3.3.2. *G. violaceus* Galactose oxidase (GalOxGv)

Four different protein expression systems (*E. coli* BL21 (DE3), BL21 Star (DE3), Tuner (DE3) and Rosetta (DE3)) were used in parallel to examine the levels of expression of *galOxGv*. SDS-PAGE analysis of crude extracts from *E. coli* BL21 Star and Tuner cells revealed that the addition of IPTG to the culture media resulted in the accumulation of an additional band of around 32 kDa, absent in extracts prepared from non-induced cells (Fig 3.13.). This result

turned out to be unexpected, since the predicted protein should have 82 kDa. In *E. coli* BL21star this band was observed only on the insoluble fraction while in *E. coli* Tuner, *galOxGv* expression was verified in the soluble and insoluble fractions.

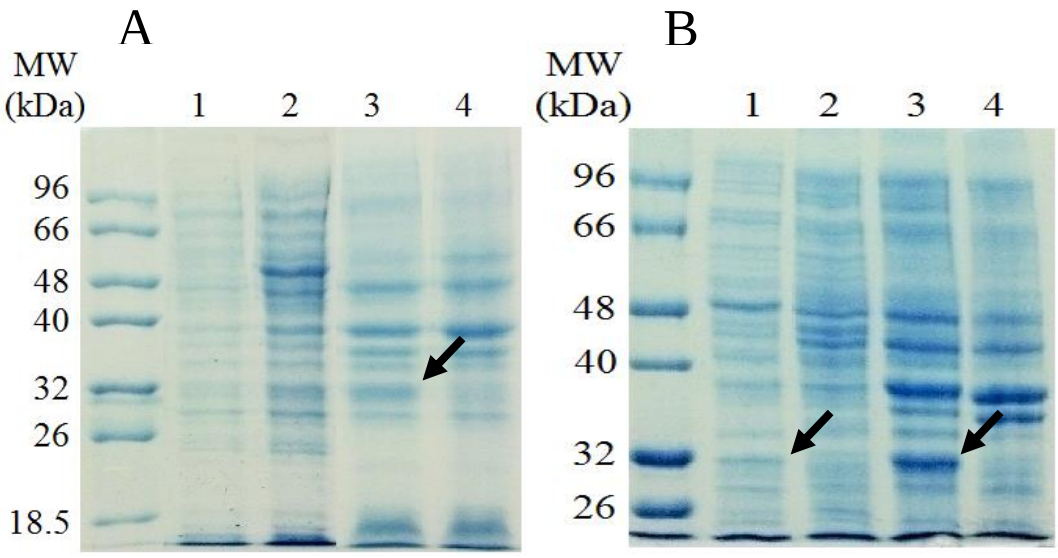


Figure 3.13. (A) SDS-PAGE analysis of GalOxGv overproduction in *E. coli* BL21 Star. Lane 1: soluble fraction of an IPTG induced crude extract; Lane 2: soluble fraction of a non-induced crude extract; Lane 3: insoluble fraction of an IPTG induced crude extract; Lane 4: insoluble fraction of a non-induced crude extract. The arrow indicates the band corresponding to the recombinant protein. (B) SDS-PAGE analysis of GalOxGv overproduction in *E. coli* Tuner. Lane 1: soluble fraction of an IPTG induced crude extract; Lane 2: soluble fraction of a non-induced crude extract; Lane 3: insoluble fraction of an IPTG induced crude extract; Lane 4: insoluble fraction of a non-induced crude extract. The arrow indicates the band corresponding to the recombinant protein (32 kDa).

It was hypothesized that the expected protein with 82 kDa was not totally denatured and migrate to a molecular mass corresponding to 32 kDa in the SDS-PAGE. In order to verify this possibility, thermal incubation at 99°C up to 60 min and chemical incubation with 10-50 mM of guanidine hydrochloride (GdnHCl) were performed prior to the SDS-PAGE analysis (Fig 3.14.).

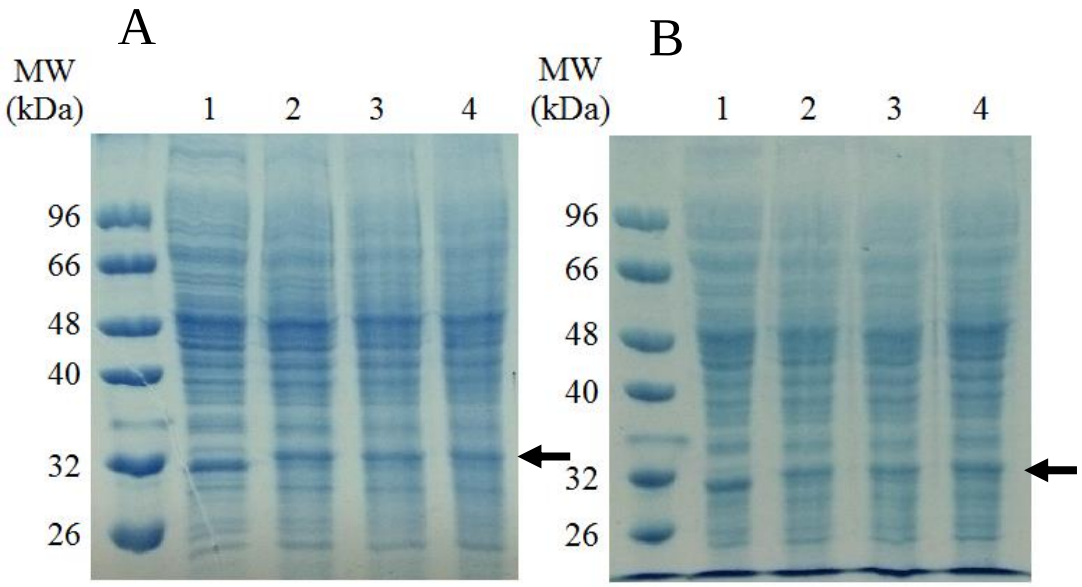


Figure 3.14. (A) SDS-PAGE analysis of GalOxGv overproduction in *E. coli* Tuner after thermal denaturation at 99°C. Lane 1: no denaturation; Lane 2: 20 min of denaturation; Lane 3: 40 min of denaturation; Lane 4: 60 min of denaturation. (B) SDS-PAGE analysis of GalOxGv overproduction in *E. coli* Tuner after chemical denaturation. Lane 1: no denaturation; Lane 2: 10 mM of GdnHCl plus 20 min at 99°C; Lane 3: 25 mM of GdnHCl plus 20 min at 99°C; Lane 4: 50 mM of GdnHCl plus 20 min at 99°C. The arrows indicates the size of the recombinant protein (32 kDa).

The results of the different denaturation conditions showed no significant alterations on the migration of the recombinant protein in the SDS-PAGE that showed the non-predicted molecular mass of 32 kDa. Further tests were conducted due to the fact that even though the size does not match to the theoretical previsions, this was the only additional band visible upon addition of IPTG and the analysis of the DNA sequence confirmed the presence of the cloned gene in the expression vector (appendix C).

Activity assays were performed on both soluble and insoluble fractions, of the IPTG induced crude extracts, to verify the GalOx enzymatic activity using the HRP coupled assay in a range of different conditions as described in the methods section 2.4.5. Despite the variety of conditions tested no activity was observed. In order to rule out the possibility that some other enzymes or components present in the crude extract could interfere with the activity of the recombinant enzyme, one primary purification step, using a Q-Sepharose column, was performed using the soluble crude extract from *E. coli* Tuner strain (Fig 3.15.), that showed an increased expression in the soluble fraction as compared to *E. coli* BL21 Star.

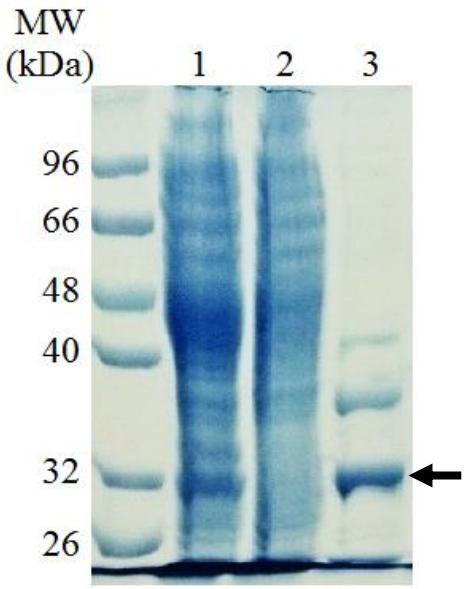


Figure 3.15. SDS-PAGE analysis of GalOxGv overproduction in *E. coli* Tuner. Lane 1: soluble fraction of an IPTG induced crude extract; Lane 2: soluble fraction of a non-induced crude extract; Lane 3: sample after the Q-sepharose chromatography. The arrow indicates the size of the recombinant protein (32 kDa).

As a result it was possible to separate many of the proteins and contaminants present on the crude extract, yielding a semi purified enzyme. Activity assays using the HRP coupled assay were performed as described in the methods section 2.4.5. but in spite of a variety of different conditions tested, no enzymatic activity was observed.

Additionally, the protein of interest was excised from the SDS-PAGE gel and identified it by MALDI-TOF/TOF-MS after tryptic digestion. The result from this analysis revealed with a score of (644) and a coverage of (66%) that the protein was (β-Lactamase of *E. coli*) (appendix D).

In the light of this result no further experimental work was conducted with this recombinant protein.

3.3.3. *S. aurantiaca* hypothetical Galactose oxidase (hGalOxSa)

Four different protein expression systems (*E. coli* BL21 (DE3), BL21 Star (DE3), Tuner (DE3) and Rosetta (DE3)) were used in parallel to examine the levels of expression of *hgalOxSa*. *E. coli* BL21 Star was the only strain that showed expression of the recombinant protein but only in the insoluble fraction. SDS-PAGE analysis of crude extracts of the recombinant cells revealed that upon addition of IPTG, a major band of around 86.5 kDa, absent in extracts prepared from non-induced cells, appeared in the insoluble fraction of cell lysates, most probably in the form of inclusion bodies (Fig 3.16.).

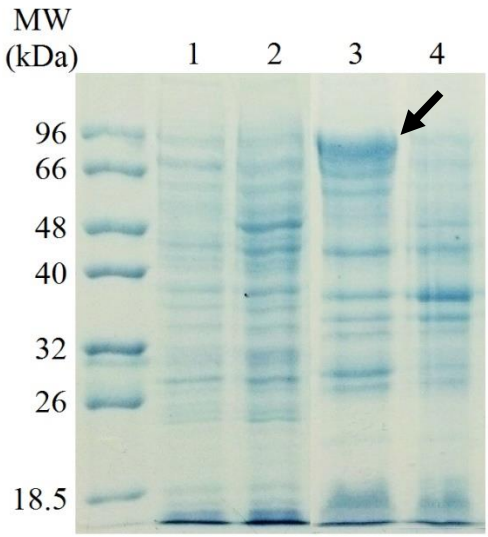


Figure 3.16. SDS-PAGE analysis of hGalOxSa overproduction in *E. coli* BL21 Star. Lane 1: soluble fraction of an IPTG induced crude extract; Lane 2: soluble fraction of a non-induced crude extract; Lane 3: insoluble fraction of an IPTG induced crude extract; Lane 4: insoluble fraction of a non-induced crude extract. The arrow indicates the band corresponding to the recombinant protein (86.5 kDa).

Activity assays were performed on both soluble and insoluble fractions, of the IPTG induced crude extracts, using the HRP coupled assay. A range of different conditions to detect activity of the GalOx were tested as described in the Material and Methods section 2.4.5. Despite the variety of conditions tested, no enzymatic activity was observed in both fractions. As this recombinant protein was the one that showed the lowest sequence similarity with the fungal GalOxFg (Fig 3.1.) no further attempts were made to overexpress this enzyme in the soluble fraction.

3.3.4. *A. siccitolerans* Galactose oxidase (GalOxAs)

The expression of recombinant GalOxAs, using four different protein expression systems, *E. coli* BL21, BL21 Star, Tuner and Rosetta, were performed to examine the heterologous levels of expression. *E. coli* BL21 Star and Rosetta were the strains that displayed observable expression of the recombinant protein. SDS-PAGE analysis of crude extracts from the recombinant cells revealed the presence, upon addition of IPTG, of a major band of around 78 kDa, absent in extracts prepared from non-induced cells (Fig 3.17.). The protein is mostly present in the insoluble fraction of cell lysates of *E. coli* BL21 Star and Rosetta strains, most likely in the form of inclusion bodies. However, the levels of expression in the Rosetta strain are higher than in the BL21star strain. Moreover, since it was possible to observe the protein in the soluble fraction of *E. coli Rosetta*, as opposed to the BL21 Star, this strain was used for further overexpression and purification of the recombinant protein.

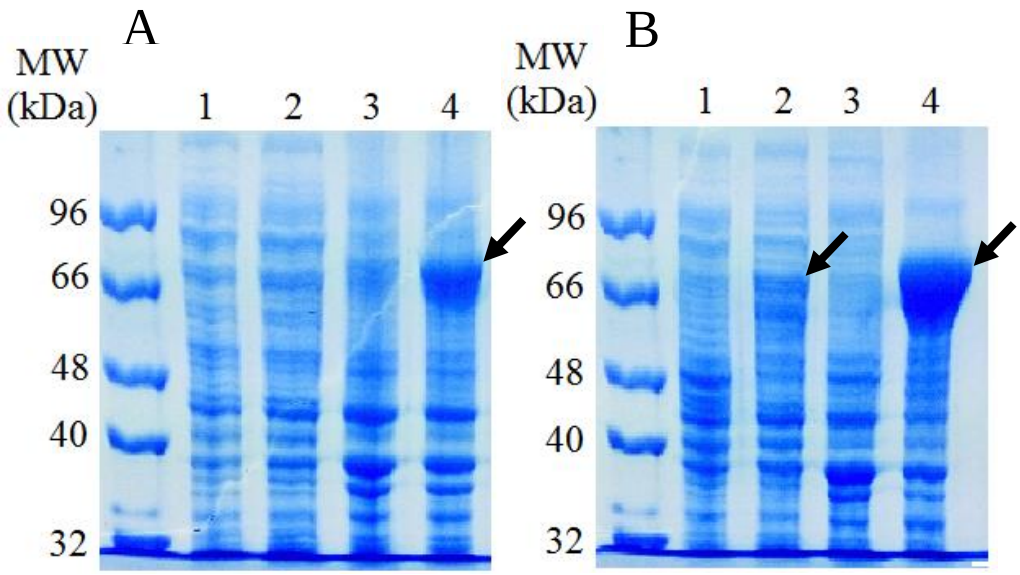


Figure 3.17. (A) SDS-PAGE analysis of GalOxAs overproduction in *E. coli* BL21 star. Lane 1: soluble fraction of a non-induced crude extract; Lane 2: soluble fraction of an IPTG induced crude extract; Lane 3: insoluble fraction of a non-induced crude extract; Lane 4: insoluble fraction of an IPTG induced crude extract. The arrow indicates the band corresponding to the recombinant protein (78 kDa). (B) SDS-PAGE analysis of GalOxAs overproduction in *E. coli* Rosetta. Lane 1: soluble fraction of a non-induced crude extract; Lane 2: soluble fraction of an IPTG induced crude extract; Lane 3: insoluble fraction of a non-induced crude extract; Lane 4: insoluble fraction of an IPTG induced crude extract. The arrow indicates the band corresponding to the recombinant protein (78 kDa).

Enzymatic activity assays were performed using both soluble and insoluble fractions of IPTG-induced crude extracts, using the HRP coupled assay. Activity was detected using D-galactose and glycerol as substrates using both the soluble and insoluble fraction of both strains, however, a higher activity was always measured using *E. coli* Rosetta. No activity was observed

for methylglyoxal or D-glucose. Attempts to produce the protein in higher amounts in the soluble fraction were unsuccessful (increasing the IPTG concentration from 0.1 mM to 1 mM; lowering temperature (to 25°C) after induction; different periods of incubation 12 to 24 h). Several purification procedures were implemented using the soluble fraction of *E. coli* Rosetta cell lysates. Q-Sepharose column was used with different sets of parameters: equilibrated with 5-20 mM Tris-HCl buffer at pH 7.6, 8.6 or 9. A SP-Sepharose column (equilibrated with 20 mM Tris-HCl buffer, pH 7.6) was also performed, since the theoretical pI (5.3) could be different from the actual pI of the protein. However, it was observed that in spite of all the conditions tested the quantity of protein that bind to the column was very small. As the attempts to purify the protein on the soluble fraction failed and the major part of the produced protein was on the insoluble fraction, a protocol of protein unfolding and refolding was followed as described in the 2.4.6.3.1. Material and Methods section. The protein was successfully recovered and the enzyme preparation exhibited an activity for D-galactose of 14 nmol/min.mL. After dialysis the protein appeared to be partially purified as observed in SDP-PAGE (Fig 3.18.) (Table 3.1.). A Q-Sepharose column (equilibrated with 20 mM Tris-HCl buffer, pH 7.6) was performed to further purify the recombinant protein but once again the majority of the protein did not show to bind to the column. Moreover, the eluted protein shows no enzymatic activity and thus we hypothesized that the protein precipitated inside the chromatographic column. Other procedures confirmed the propensity of the recombinant protein to aggregate; molecular mass determination by size-exclusion chromatography indicated that the protein was aggregated, appearing with a molecular mass in between ~ 300 and ~ 1000 kDa. As it was not possible to further purify this enzyme, the partially purified protein obtained after dialysis was used to perform a preliminary biochemical characterization of the protein.

Table 3.1. Purification parameters of GalOxAs. CEs: soluble fraction of GalOxAs crude extract; CEi: insoluble fraction of GalOAs crude extract; Re: GalOxAs refolded and after dialysis.

	Protein (mg/ml)	**Total volume (ml)**	**Total Protein (mg)**	**Activity (nmol/ min.ml)**	**Total Actitvity (nmol.min)**	**Specific Activity (nmol/ min.mg)**
CEs	9 ± 0.3	30	271	0.15 ± 0.02	22 ± 3	0.1 ± 0.01
CEi	6 ± 0.3	15	91	1.4 ± 0.3	106 ± 22	1.2 ± 0.2
Re	0.7 ± 0.02	125	84	14.1 ± 0.3	8823 ± 162	105 ± 2

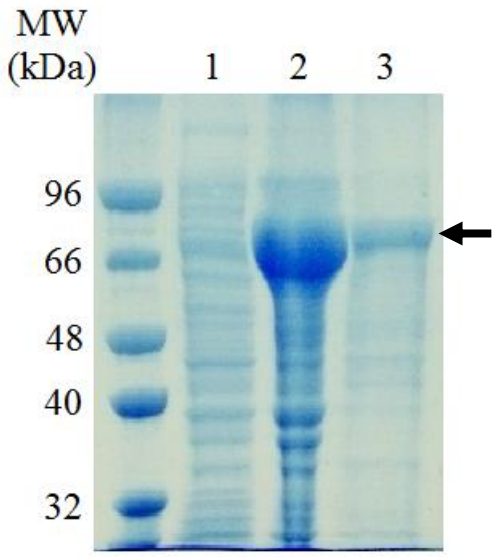


Figure 3.18. SDS-PAGE analysis of GalOxAs overproduction and purification in *E. coli* Rosetta. Lane 1: soluble fraction of an IPTG induced crude extract; Lane 2: insoluble fraction of an IPTG induced crude extract; Lane 3: GalOxAs refolded and after dialysis. The arrow indicates the size of the recombinant protein (78 kDa).

The specific activity measured in the soluble and insoluble crude extracts was 0.1 mU/mg and 1.2 mU/mg, respectively. After the dialysis step, the specific activity for GalOxAs was 105 mU/mg, an increase of ~ 10,000 fold and of ~ 100 fold in relation to the initial activity, in the soluble and insoluble crude extracts, respectively.

The protein was positively identified by MALDI-TOF/TOF-MS after tryptic digestion with an identification score of 2900 with sequence coverage of 66% (appendix D).

3.3.4.1. Spectroscopic properties

The enzyme solution showed a pale bluish colour, which led us to believe that the recovered enzyme existed mostly in the semireduced state (Whittaker, 2003). The UV-visible spectrum (200 nm to 1100 nm) performed with this sample did not evidence any distinguishable bands. In order to generate fully oxidized protein, the state that leads to the characteristic visible bands of galactose oxidase, samples of the protein were incubated in the presence of ferricyanide (5-30 min, 10-100 mM of $K_3Fe(CN)_6$) and also iridate (5-30 min, 100 mM). Following incubation the samples were diafiltrated with Amicon®Ultra centrifugal filters (30K) to remove the oxidizing agents. However and despite all the attempts a UV-visible spectrum with the characteristic spectral features of oxidized enzyme (bands at 450 nm and at 800 nm) was never observed. We conclude that these results are most likely related to the fact that the enzyme is in a semi-aggregated state.

3.3.4.2. Optimal pH and temperature

Temperature effect on GalOxAs activity was examined at temperatures ranging from 20 to 60°C using the HRP coupled assay with D-Galactose as substrate. GalOxAs presents a defined bell-shape activity profile with optimal temperature between 25-35°C (Fig 3.19 A) in agreement with the optimum growth temperature of *A. siccitolerans* at 30°C (SantaCruz-Calvo *et al*, 2013). It is also important to emphasize that between 20 and 40°C the enzyme shows more than 90% of the maximum activity. Recent reports from two fungal GalOx, from *Fusarium acuminatum* (GalOxFa) (Alberton *et al*., 2007) and from *Fusarium oxysporum* (GalOxFo) (Paukner *et al*., 2014), also show a defined bell shaped activity profile with optimal temperatures of 30°C and 40°C, respectively. GalOx from *Fusarium sambucinum* (GalOxFs) show a more abrupt temperature shape profile with an optimum of 35°C (Paukner *et al*., 2015). The activity of GalOxAs for D-galactose showed a maximum at pH 8, and at pH 7 and 9 the activity is around 90% of the maximum (Fig 3.19. B). This optimum pH is similar to the values described for other (fungal) galactose oxidase. For GalOxFa the optimum pH is 8.0 (Alberton *et al*., 2007), for GalOxFo is pH 7.0 (Paukner *et al*., 2014) and for GalOxFs optimal pH range from 6.0 to 7.5 (Paukner *et al*., 2015).

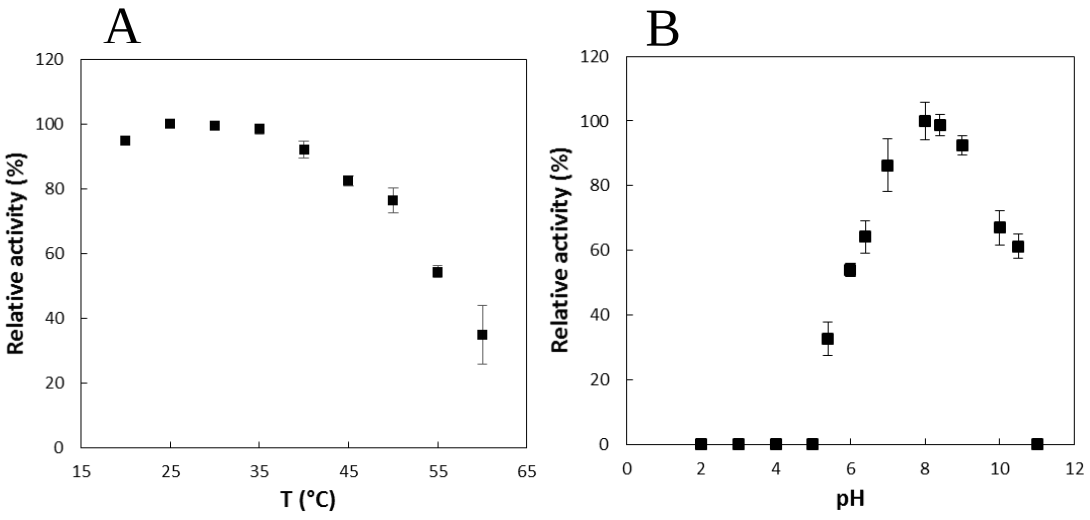


Figure 3.19. (A) Temperature profile of GalOxAs using 1 mM of ABTS, 15 U of HRP and 250 mM of D-galactose at pH 8. (B) pH profile of GalOxAs using 1 mM of ABTS, 15 U of HRP and 250 mM of D-galactose at 25°C.

3.3.4.3. Kinetic properties of GalOxAs

GalOx is characterized for having broad substrate specificity, which is one of the most important characteristics of the enzyme (Whittaker, 2002). The apparent steady-state kinetic parameters for different substrates were determined using oxygen (air) as electron acceptor.

The initial rates of substrate turnover were recorded using different substrate concentrations in the standard HRP-coupled assay at 25°C and pH 8.0, and the enzyme was shown to follow a Michaelis-Menten kinetics with all the tested substrates (Fig 3.20., Table 3.2.).

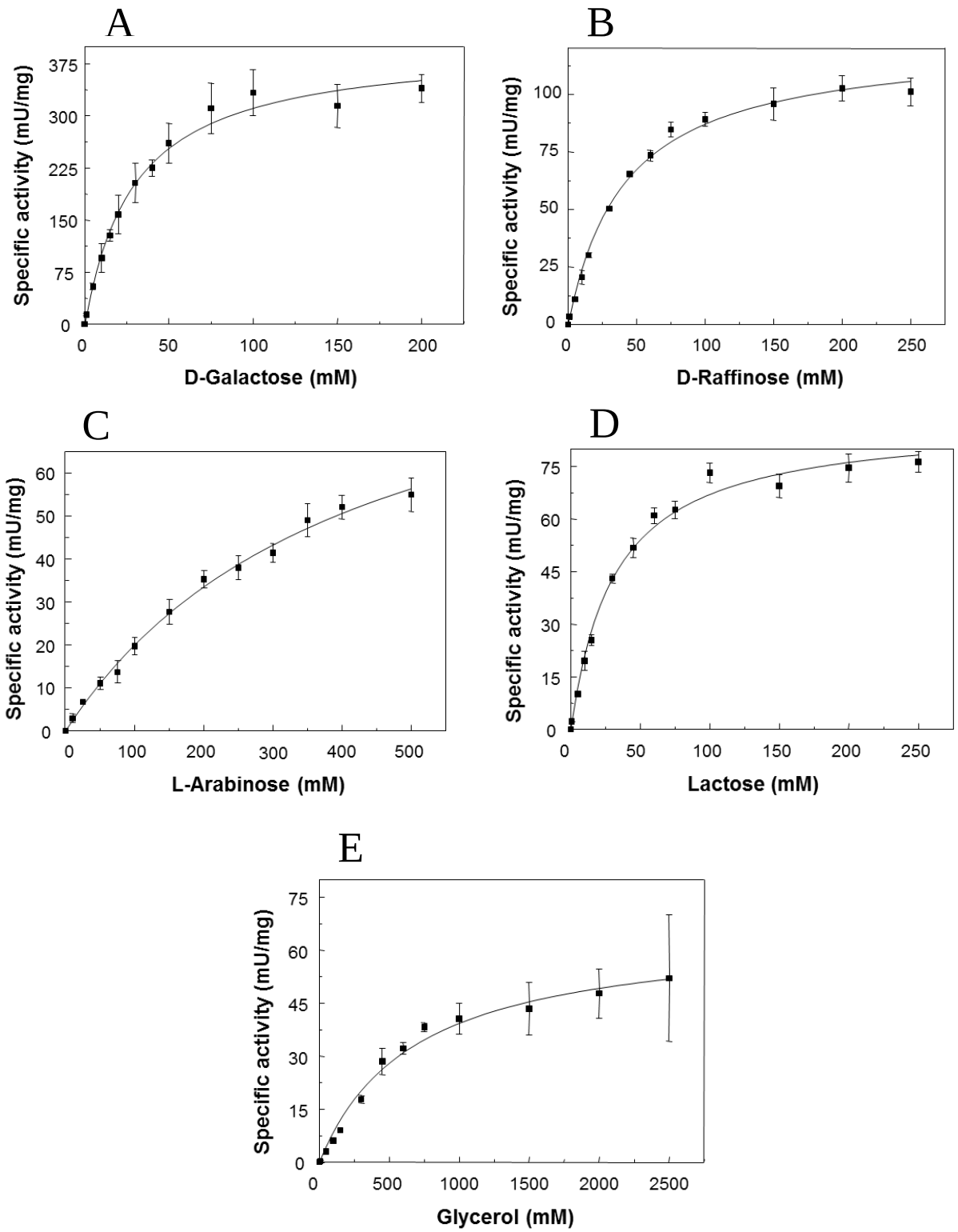


Figure 3.20. GalOxAs apparent saturation kinetics using 1 mM of ABTS and 15 U of HRP at 25°C and pH 8 for (A) D-galactose; (B) D-raffinose; (C) L-arabinose; (D) Lactose; (E) Glycerol.

Table 3.2. Steady-state apparent kinetic parameters of GalOxAs. Reactions were performed in the presence of 1 mM of ABTS and 15 U of HRP at 25°C and pH 8.

Substrate	V_{max} (nmol/min.mg)	$K_{m\ app}$ (mM)	$k_{cat\ app}$ (s^{-1})	k_{cat}/K_m ($mM^{-1}s^{-1}$)
D-galactose	403 ± 14	30 ± 3	5×10^{-1}	2×10^{-2}
D-raffinose	123 ± 3	41 ± 3	2×10^{-3}	4×10^{-5}
L-arabinose	103 ± 7	416 ± 46	1×10^{-3}	3×10^{-6}
Lactose	88 ± 3	32 ± 3	8×10^{-4}	3×10^{-5}
Glycerol	66 ± 3	676 ± 87	8×10^{-4}	3×10^{-6}

Galactose oxidase from *A. siccitolerans* exhibits a K_m of 30 ± 3 mM for D-galactose, a similar value as of *F. graminearum* GalOx, produced in *E. coli* (35 mM) (Choosri *et al*., 2010). GalOxFo also shows a similar K_m of 47 mM (Paukner *et al*., 2014), GalOxFa presents a lower K_m of 16.2 mM (Alberton *et al*., 2007) and GalOxFs a higher K_m of 61 mM (Paukner *et al*., 2015). Native *F. graminearum* GalOx presented a K_m of 67 mM (Baron *et al*., 1994). In terms of catalytic efficiency (k_{cat}/K_m), GaloxAs shows the highest efficiency for D-galactose (2×10^{-2} $mM^{-1}s^{-1}$) but this value is two orders of magnitude lower than others reported GalOx's (2.0 $mM^{-1}s^{-1}$ GalOxFo (Paukner *et al*., 2014); 0.89 $mM^{-1}s^{-1}$ GalOxFs (Paukner *et al*., 2015)), which is again most probably due to the fact that the enzyme is in a semi-aggregated state. The reported GalOxFg; GalOxFo and GalOxFs K_m for D-raffinose (20 – 25 mM) (Avigad *et al*., 1962; Paukner *et al*., 2014; Paukner *et al*., 2015) is usually lower than for D-galactose (47 – 175 mM) (Avigad *et al*., 1962; Paukner *et al*., 2014; Paukner *et al*., 2015), however GalOxAs shows a higher K_m for D-raffinose (41 ± 3 mM) than for D-galactose (30 ± 3 mM). The activity of the enzyme with raffinose and lactose, indicates that GalOxAs can oxidize galactose derivatives with substitutes at the carbon-1 site, as previously reported for all GalOx's (Alberton *et al*., 2007). In contrast, the lack of enzyme activity on glucose indicates that the stereo configuration of the hydroxyl group at the carbon-4 site is important for its
activity, and that the enzyme is specific to galactose and derivatives, which is in good agreement with reports for other characterized GalOx's. One interesting feature of GalOxAs is the K_m for lactose (32 ± 3 mM) which is 20-fold lower as compared to other reported GalOx's (683 mM for GalOxFs (Paukner *et al*., 2015; 685 mM for GalOxFo (Paukner *et al*., 2014)); this can be an advantage of this enzyme for specific applications when lactose is used. Another interesting

feature is the detected activity towards L-arabinose, which have not been reported in the other mentioned GalOx's. The aliphatic alcohol, glycerol, is also a substrate for this GalOx with much higher K_m than D-galactose in agreement to other reports (Klibanov *et al*., 1982; Kosman, 1984). The relatively high K_m values for different substrates for GalOx seem to be a consequence of the broad substrate specificity of the enzyme resulting in an active site capable of binding a range of different substrates, but therefore being relatively weak at binding any particular substrate (Wilkinson *et al*. 2004).

3.3.4.4. Enzyme stability

In this study, GalOxAs denaturates irreversibly according to a first-order process, which can be described by the classical Lumry-Eyring model applied to the majority of the enzymes (N ↔ U → D, where N, U and D, are the native, the reversible unfolded and the irreversible denaturated enzyme), pointing to a simple pathway of unfolding and deactivation (Fig 3.21.).

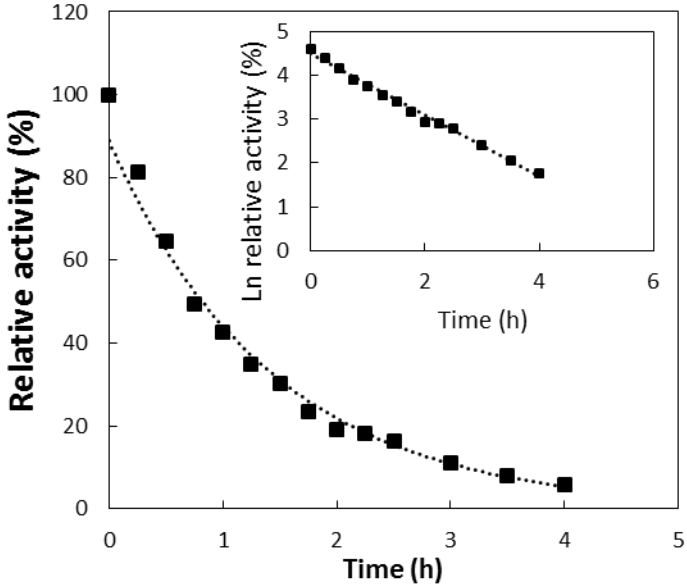


Figure 3.21. Kinetic stability of GalOxAs. The activity decay at 40°C was fitted accurately considering an exponential decay (dotted line shows the fit) with a haft-life time of 1 h. The inset clearly shows that the activity decay of GalOxAs can be fitted to a single first order process, as the logarithm of activity displays an inverse linear relationship with time.

GalOxAs presents a half-life time of 1 h at 40°C, 5 fold higher than the measured half-life of GalOxFs (11 min) (Paukner *et al*., 2015) but 9 fold lower than GalOxFo (9 h) ((Paukner *et al*., 2014). Nonetheless, the half-life time of GalOxAs is most probably limited as the enzyme is on a semi-aggregated state and complete solubilized enzyme is expected to present an diferent half-life time.

3.3.5. *A. siccitolerans* Pyranose oxidase (POxAs)

Expression of recombinant POxAs was performed to examine the efficiency and levels of expression *E. coli*, using four different protein expression systems, *E. coli* BL21, BL21 Star, Tuner and Rosetta. *E. coli* BL21 Star and Rosetta were the strains that showed expression of the recombinant protein under the conditions tested. Crude extracts from these recombinant cells revealed that the addition of IPTG to the culture medium resulted in the accumulation of an additional band around 55 kDa in *E. coli* BL21star (the expected size) and around 28 kDa in *E. coli* Rosetta, both absent in extracts prepared from non-induced cells as visualized by SDS-PAGE (Fig 3.22.).

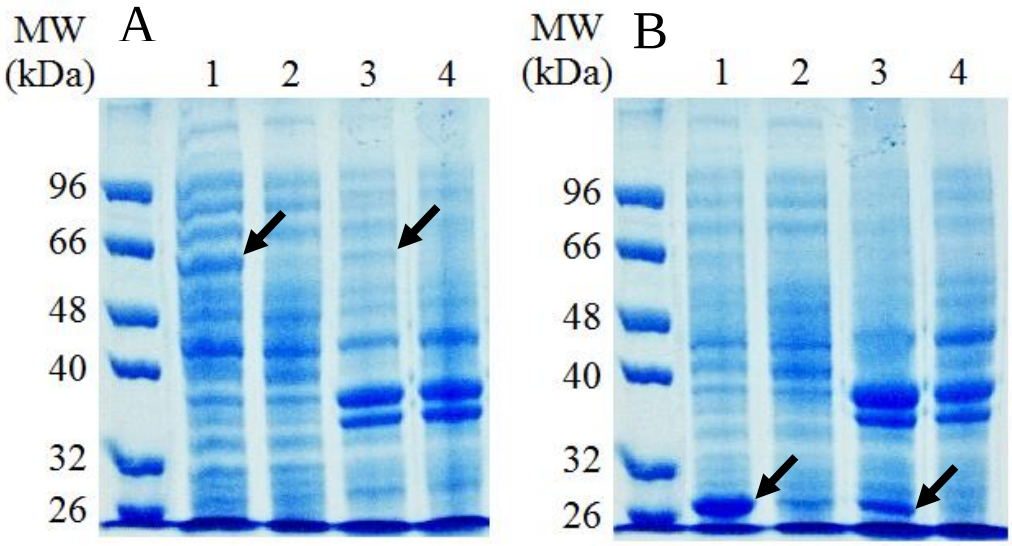


Figure 3.22. (A) SDS-PAGE analysis of POxAs overproduction in *E. coli* BL21 star. Lane 1: soluble fraction of an IPTG induced crude extract; Lane 2: soluble fraction of a non-induced crude extract; Lane 3: insoluble fraction of an IPTG induced crude extract; Lane 4: insoluble fraction of a non-induced crude extract. The arrow indicates the band corresponding to the recombinant protein (55 kDa). (B) SDS-PAGE analysis of POxAs overproduction in *E. coli* Rosetta. Lane 1: soluble fraction of an IPTG induced crude extract; Lane 2: soluble fraction of a non-induced crude extract; Lane 3: insoluble fraction of an IPTG induced crude extract; Lane 4: insoluble fraction of a non-induced crude extract. The arrow indicates the band corresponding to the recombinant protein (28 kDa).

Enzymatic activity assays were performed using both soluble and insoluble fractions, of the IPTG-induced crude extracts, using the HRP coupled assay. Activity was detected using D-glucose as substrate (which is the best substrate for the described enzymes in the literature (Leitner *et al*., 2001; Takakura and Kuwata, 2003; Pisanelli *et al*., 2009; Salaheddin *et al*., 2010)) on the soluble fraction of *E. coli* BL21 Star. Accordingly, this strain was used for expression of the recombinant protein and POxAs was partially purified from crude cell extracts supernatant through two chromatographic steps (Fig 3.23.).

However, in spite of all the different conditions tested, it was not possible to reach an acceptable level of purification for biochemical and kinetic characterization. We have

additionally observed a problem on the expression of the enzyme *i.e.* the levels of the recombinant protein are unstable, since its expression did not always occurr even if the same growth and induction protocol was followed (data not shown). Due to these problems it was not possible, in this work, to perform a proper characterization of this protein but nevertheless, the UV-visible absorption spectr from partially purified POxAs produced a typical flavoprotein spectrum with absorption maxima at 456 nm and 345 nm is shown in Fig 3.24., confirming the reports in the literature (Leitner *et al*., 2001; Takakura and Kuwata, 2003; Pisanelli *et al*., 2009; Salaheddin *et al*., 2010).

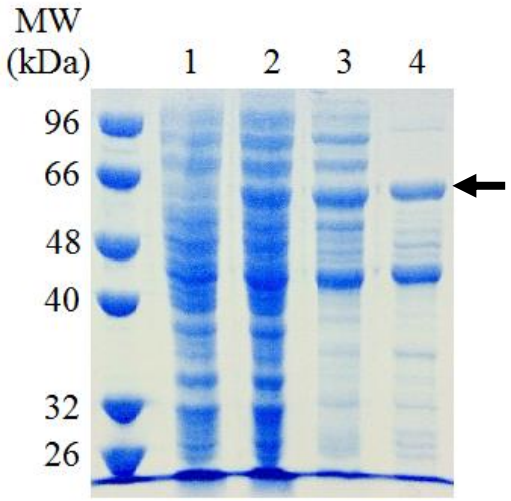


Figure 3.23. SDS-PAGE analysis of POxAs overproduction and purification in *E. coli* BL21 star. Lane 1: soluble fraction of a non-induced crude extract; Lane 2: soluble fraction of an IPTG induced crude extract; Lane 3: sample after Q-sepharose chromatography; Lane 4: sample after Superdex-200 chromatography. The arrow indicates the size of the recombinant protein (55 kDa).

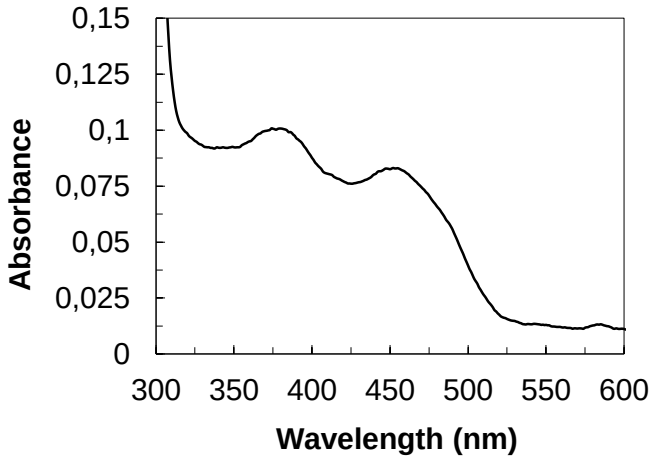


Figure 3.24. UV-visible spectrum of POxAs in 20 mM Tris-HCl buffer, pH 7.6 with 0.2 M of NaCl.

The protein was positively identified by MALDI-TOF/TOF-MS after tryptic digestion with an identification score of 1190 and a sequence coverage of 50% (appendix D).

4. CONCLUSIONS AND FUTURE PERSPECTIVES

Lignin-degrading auxiliary enzymes represent an important group in the multienzymatic system required for lignin depolymerisation. Thus far all the LDA's described enzymes are from fungal origins with exception of a small report of a galactose like RCO enzyme in *S. coelicolor* (Whittaker and Whittaker, 2006).

In this work we have identified a large number of putative glyoxal/galactose oxidases including 7 genes in the available genomes (appendix B). We have also identified several putative pyranose oxidase genes including one gene in the available genomes (appendix B). Considering the putative glyoxal/galactose and pyranose oxidase genes, we were able to clone 5 out of 8 genes (hGalOxMs2; GalOxGv; hGalOxSa; GalOxAs and POxAs). Several problems of expression were observed for the five genes cloned and only POxAs and GalOxAs were further studied. POxAs was identified in the soluble fraction of an IPTG-induced cell culture by SDS-PAGE and activity for D-glucose was detected as well as the flavin-dependent fingerprint in the UV-visible spectrum. However the levels of expression were erratic and the enzyme could not be further purified using standard chromatographic approaches. GalOxAs, expressed in *E. coli* Rosetta, shows activity in IPTG-induced cell extracts towards D-galactose and glycerol, was further studied after an unfolding and refolding protocol since the vast majority of the recombinant protein was located in the insoluble fraction of cell lysates, most probably in the form of inclusion bodies. However the final protein preparation was found to be in a semi-aggregated state as assessed by size-exclusion chromatography. Despite all attempts no further purification was achieved most probably related to the enzyme high propensity to aggregate. Several actions can be taken in order to try to solve these problems. Optimization of the unfolding/refolding protocol can be performed/attempted aiming at obtain a completely soluble enzyme. Considering that GalOxAs is a prone-to-aggregate protein several other measures can be adopted such as (i) developing procedures to tightly control the expression of the proteins using specialized promoters; (ii) attaching the protein to solubility-enhancing fusion proteins; (iii) developing specialized bacterial host strains; (iv) screening for specific growth and induction conditions; (v) considering practical methods to reduce aggregation during the purification steps; and (vi) screening for suitable buffer conditions for protein purification. A different line of action can be used through mutagenesis in order to increase the solubility of the protein.

The preliminary characterization of GalOxAs showed that both the optimal pH and temperature are very similar to recently reported GalOx enzymes from fungal sources. GalOxAs showed broad substrate specificity in agreement with the stereo and regio selectivity presented by other GalOx's, and also activity towards L-arabinose which has not been reported to date.

GalOxAs and POxAs have fundamental interest and are important enzymes for the set-up of bacterial multienzymatic systems involved in aromatic compounds or ligning degradation and valorization. Additionally they present advantages in relation to the fungal enzymes as higher yields of production are expected as well as the easiness in manipulating their properties by molecular biology techniques. It is expected that this work will open new perspectives for the understanding of the structure-function relationships of hydrogen peroxide producing enzymes in bacteria and for the use of bacterial oxidoreductive enzymes in a large range of biotechnological applications.

5. REFERENCES

Adanyi N, Szabo EE, and Varadi M (1999) Multi-enzyme biosensors with amperometric detection for determination of lactose in milk and dairy products. Eur. Food Res Technol, 209: 220-226

Ahmad M, Taylor CR, Pink D, Burton K, Eastwood D, Bending GR, Bugg TDH (2010) Development of novel assays for lignin degradation: comparative analysis of bacterial and fungal lignin degraders. Mol Biosystems, 6: 815-821

Alberton D, Oliveira LS, Peralta RM, and Barbosa-Tessmann IP (2007) Production, purification, and characterization of a novel galactose oxidase from *Fusarium acuminatum*. J. Basic Microbiol., 47: 203–212

Amaral D, Bernstein L, Morse D, and Horecker BL (1963) Galactose oxidase of *Polyporus circinatus*: a copper enzyme. J. Biol. Chem, 238: 2281-2284

Antai SP, Crawford DL (1981) Degradation of softwood, hardwood, and grass lignocelluloses by two *Streptomyces* strains. Appl Environ Microbiol, 42: 378-380

Avigad G (1985) Oxidation Rates Of Some Desialylated Glycoproteins By Galactose-Oxidase. Arch. Biochem. Biophys., 239: 531-537

Avigad G, Asensio C, Horecker BL, and Amaral D (1962) D-galactose oxidase of *Polyporus circinatus*. J. Biol. Chem, 237: 2736-2743

Badcock GT, Eldeeb MK, Sandusky PO, Whittaker MM, and Whittaker JW (1992) Electron-paramagnetic resonance and electron nuclear double-resonance spectroscopies of the radical site in galactose-oxidase and of thioether-substituted phenol model compounds. J. Am. Chem. Soc., 114: 3727-3734

Baron AJ, Stevens C, Wilmot C, Seneviratne KD, Blakeley V, Dooley DM, Phillips SEV, Knowles PF, and McPherson MJ (1994) Structure and mechanism of galactose oxidase − The free radical site. J. Biol. Chem, 269: 25095-25105

Berzins T, Wiken M, Hellstrom U, and Perlmann P (1983) The use of supernatants from neuraminidase and Galactose-Oxidase (nago)-treated lymphocytes as a source of Tcell growth-factor. J. Immunol. Meth., 63: 309-319

Bianchetti CM, Harmann CH, Takasuka TE, Hura GL, Dyer K, and Fox BG (2013) Fusion of dioxygenase and lignin-binding domains in a novel secreted enzyme from cellulolytic *Streptomyces* sp. SirexAA-E. J Biol Chem, 288: 18574-18587

Bollag D, Rozycki M, and Edelstein S, (1996) Protein Methods. John Wiley & Sons, New York.

Borman CD, Saysell CG, Wright C, and Sykes AG (1998) Mechanistic studies on the single copper tyrosyl radical containing enzyme galactose oxidase. Pure Appl. Chem, 70: 897-902

Bradford MM, (1976). A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding, Anal. Biochem, 72, 248-254.

Brenner AJ, and Harris ED, (1995) A quantitative test for copper using bicinchoninic acid, Anal Biochem, 226: 80-84

Brissos V, Chen Z, and Martins LO (2012) The kinetic role of carboxylate residues in the proximity of the trinuclear centre in the O2 reactivity of CotA-laccase. Dalton Trans., 41: 6247-6255

Brown ME and Chang MCY (2013) Exploring bacterial lignin degradation. Curr. Opin. Chem. Biol., 19: 1–7

Brown ME, Walker MC, Nakashige TG, Iavarone AT, and Chang MCY (2011) Discovery and characterization of heme enzymes from unsequenced bacteria: application to microbial lignin degradation. J Am Chem Soc, 133: 18006-18009

Bugg TDH, Ahmad M, Hardiman EM, and Singh R (2010) The emerging role for bacteria in lignin degradation and bio-product formation. Curr Opin in Biotechnol, 22: 394-400

Casali N, and Preston A, (2003). E. coli Plasmid Vectors Methods and Applications. Meth. Mol. Biol., 235, 24-48

Cevik E, Senel M, and Abasiyanik MF (2010. Construction of biosensor for determination of galactose with galactose oxidase immobilized on polymeric mediator contains ferrocene. Curr. Appl. Phys., 10: 1313-1316

Charmantray F, Touisni N, Hecquet L, and Mousty C (2013) Amperometric Biosensor Based on Galactose Oxidase Immobilized in Clay Matrix. Electroanalysis, 25: 630-635

Chen Z, Durão P, Silva CS, Pereira MM, Todorovic S, Hildebrandt P, Bento I, Lindley PF, and Martins LO (2010) The role of E498 in the dioxygen reactivity of CotA-laccase from *Bacillus subtilis*. Dalton Trans, 39: 2875-2882

Choosri W, Paukner R, Wuhrer P, Haltrich D, and Leitner C (2010) Enhanced production of recombinant galactose oxidase from *Fusarium graminearum* in E. coli. World J Microbiol Biotechnol, 27: 1349-1353

Clark K, Pennerhahn JE, Whittaker MM, and Whittaker JW (1990) Oxidation state assignments for galactose oxidase complexes from X-ray absorption spectroscopy evidence for Cu(II) in the active enzyme. J. Am. Chem. Soc, 112: 6433-6434

Cleveland L, Coffman RE, Coon P, and Davis L (1975) The investigation of the role of the copper in galactose oxidase. Biochemistry, 14: 1108-1115

Cooper JAD, Smith W, Bacila M, and Medina H (1959) Galactose oxidase from *Polyporus circinatus*. Fr. J. Biol. Chem. 234: 445-448

Critchley DR (1974) Cell-surface proteins of nil1 hamster fibroblasts labeled ay a Galactose oxidase, Tritiated Borohydride Method. Cell, 3: 121-125

Daniel G, Volc J, and Kubátová E (1994) Pyranose oxidase, a major source of H_2O_2 during wood degradation by *Phanerochaete chrysosporium, Trametes versicolor,* and *Oudemansiella mucida*. Appl. Environ. Microbiol, 60, 2524–2532

Daniel G, Volc J, Kubátová E, and Nilsson T (1992) Ultrastructural and immunocytochemical studies on the H_2O2-producing enzyme pyranose oxidase in *Phanerochaete chrysosporium* grown under liquid culture conditions. Appl Environ Microbiol, 58: 3667-3676

Daniel R (2005) The metagenomics of soil. Nat Rev Microbiol, 3: 470-478

Davis JR, Goodwin LA, Woyke T, Teshima H, Bruce D, Detter C, Tapia R, Han S, Han J, Pitluck S *et al*. (2012) Genome sequence of *Amycolatopsis* sp. strain ATCC 39116, a plant biomassdegrading actinomycete. J Bacteriol, 194: 2396-2397

de Koker TH, Mozuch MD, Cullen D, Gaskell J, and Kersten PJ (2004) Isolation and purification of pyranose 2-oxidase from *Phanerochaete chrysosporium* and characterization of gene structure and regulation. Appl Environ Microbiol, 70: 5794-5800

Delagrave S, Murphy DJ, Pruss JLR, Maffia AM, Marrs BL, Bylina EJ, Coleman WJ, Grek CL, Dilworth MR, Yang MM, and Youvan DC (2001) Application of a very high-throughput digital imaging screen to evolve the enzyme galactose oxidase. Protein Eng, 14: 261-267

Dicks JM, Aston WJ, Davis G, and Turner APF (1986) Mediated amperometric biosensors for D-galactose, glycolate gnd L-amino-acids based on a ferrocene modified carbon paste electrode. Anal. Chim. Acta, 182: 103-112

Durão P, Bento I, Fernandes AT, Melo EP, Lindley PF, and Martins LO (2006) Perturbations of the T1 copper site in the CotA-laccase from *Bacillus subtilis*: structural, biochemical, enzymatic and stability studies. J Biol Inorg Chem, 11: 514-526

Eastwood DC, Floudas D, Binder M, Majcherczyk A, Schneider P, Aerts A, Asiegbu FO, Baker SE, Barry K, Bendiksby M *et al*. (2011) The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. Science, 333: 762-765

El-Enshasy HA, Yang S-T, and Thongchul N (2013) Bioprocessing technologies in biorefinery for sustainable production of fuels, chemicals, and polymers. John Wiley & Sons, New Jersey, 163-192

Evans CS, Dutton MV, Guillén F, and Veness RG (1994) Enzymes and small molecular mass agents involved with lignocellulose degradation. FEMS Microbiol Rev, 13: 235–240

Fernandes AT, Damas J, Soares CM, Todorovic S, Huber R, Pogni R, and Martins LO (2010) The multicopper oxidase from the archaeon *Pyrobaculum aerophilum* shows nitrous oxide reductase activity. FEBS J, 277: 3176-3179

Fernandes AT, Soares CM, Pereira MM, Huber R, Melo EP, and Martins LO (2007) A robust metallo-oxidase from the hyperthermophilic bacterium *Aquifex aeolicus*. FEBS J, 274: 2683-94

Firbank SJ, Rogers MS, Wilmot CM, Dooley DM, Halcrow MA, Knowles PF, McPherson MJ, and Phillips SEV (2001) Crystal structure of the precursor of galactose oxidase: An unusual self-processing enzyme. Proc. Natl. Acad. Sci. U. S. A., 98: 12932-12937

Freimund S, Huwig A, Giffhorn F, and Kopper S (1998) Rare Keto-Aldoses from Enzymatic Oxidation: Substrates and Oxidation Products of Pyranose 2- Oxidase. Chem – Eur J, 4: 2442–2455

Galbe M and Zacchi G (2007) Pretreatment of lignocellulosic materials for efficient bioethanol production. Adv Biochem Engineer Biotechnol, 108: 41–65

Gattegno L, Perret G, Fabia F, Bladier D, and Cornillot P (1981) In vivo aging of human-erythrocytes and cell-surface labeling by D-galactose oxidase and sodium borotritide. Carboh. Res., 95: 283-289

Geib SM, Filley TR, Hatcher PG, Hoover K, Carlson JE, Jimenez- Gasco MM, Nakagawa-Izumi A, Sleighter RL, and Tien M (2008) Lignin degradation in wood-feeding insects. Proc Natl Acad Sci USA, 105: 12932-12937

Ghisla S and Massey V (1989) Mechanisms of flavoprotein-catalyzed reactions. Eur J Biochem, 181: 1-17

Giardina P, Faraco V, Pezzella C, Piscitelli A, Vanhulle S, and Sannia G (2010) Laccases: a never-ending story, Cell. Mol. Life Sci, 67: 369-385

Giffhorn F (2000) Fungal pyranose oxidases: occurrence, properties and biotechnical applications in carbohydrate chemistry. Appl. Microbiol. Biotechnol, 54: 727–740

Gosselink RJA, de Jong E, Guran B, and Abächerli A (2004) Co-ordination network for lignin—standardisation, production and applications adapted to market requirements (EUROLIGNIN), Ind. Crops Prod., 20: 121–12

Goudsmit EM, Matsuura F, and Blake DA (1984) Substrate specificity of D-galactose oxidase − Evidence for the oxidation of internally linked galactosyl residues of helix-pomatia galactogen. J. Biol. Chem, 259: 2875-2878

Guillén F, Martínez AT, and Martínez MJ (1992) Substrate specificity and properties of the aryl-alcohol oxidase from the ligninolytic fungus *Pleurotus eryngii*. Eur J Biochem, 209: 603–611

Gutiérrez A, Caramelo L, Prieto A, Martínez MJ, and Martínez AT (1994) Anisaldehyde production and aryl-alcohol oxidase and dehydrogenase activities in ligninolytic fungi from the genus *Pleurotus*. Appl Environ Microbiol, 60: 1783–1788

Halcrow M, Phillips S, and Knowles P (2000) Enzyme-catalyzed electron and radical transfer. In Subcellular Biochemistry vol 35, edited by Holzenburg and Scrutton. Academic / Plenum Publisher, New York, 184-197

Hallberg BM, Leitner C, Haltrich D, and Divne C (2004) Crystal structure of the 270 kDa homotetrameric lignin-degrading enzyme pyranose 2-oxidase. J Mol Biol, 341: 781-796

Hamilton GA, Adolf PK, Jersey JD, Dubois GC, Dyrkacz GR, and Libby RD (1978) Trivalent copper, superoxide, and galactose oxidase. J. Am. Chem. Soc, 100: 1899-1912

Hamilton GA, de Jersey J, and Adolf PK (1973) Galactose oxidase, in Oxidases and related redox systems. University Park Press Baltimore, 1: 103-124

Hammel KE and Cullen D (2008) Role of fungal peroxidases in biological ligninolysis. Curr Opin Plant Biol, 11: 349–355

Hammel KE, Kapich AN, Jensen KA Jr, and Ryan ZC (2002) Reactive oxygen species as agents of wood decay by fungi. Enzyme Microb Technol, 30: 445–453

Han E, Ding L, Qian RC, Bao L, and Ju HX (2012) Sensitive chemiluminescent imaging for chemoselective analysis of glycan expression on living cells using a multifunctional nanoprobe. Anal Chem., 84: 1452-1458

Hatakka A (1994) Lignin-modifying enzymes from selected white-rot fungi: production and role in lignin degradation. FEMS Microbiol. Rev, 13: 125-135

Hess M, Sczyrba A, Egan R, Kim T-W, Chokhawala H, Schroth G, Luo S, Clark DS, Chen F, Zhang T *et al*. (2011) Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. Science, 331: 463-467

Higuchi T (1997) Biochemistry and molecular biology of wood. Springer, London

Himo F, Eriksson LA, Maseras F, and Siegbahn PEM (2000) Catalytic mechanism of galactose oxidase: A theoretical study. J. Am. Chem. Soc, 122: 8031-8036

Hinterstoisser B, Stefke B, and Schwanninger M (2000) Wood: Raw material- material-Source of Energy for the future. Lignovisionen, 2: 29-36

Holton JB (1990) Galactose Disorders - An Overview. J.f Inherited Metabolic Disease, 13: 476-486

Huang X-F, Santhanam N, Badri DV, Hunter WJ, Manter DK, Decker SR, Vivanco JM, and Reardon KF (2013) Isolation and characterization of lignin-degrading bacteria from rainforest soils. Biotechnol Bioeng, 110: 1616-1626

Humphreys KJ, Mirica LM, Wang Y, and Klinman JP (2009) Galactose Oxidase as a Model for Reactivity at a Copper Superoxide Center. J. Am. Chem. Soc., 131: 4657-4663

Ito N, Phillips SEV, Stevens C, Ogel ZB, McPherson MJ, Keen JN, Yadav KDS, and Knowles PF (1991) Novel Thioether Bond Revealed by a 1.7-A Crystal structure of Galactose-Oxidase. Nature, 350: 87-90

Ito N, Phillips SEV, Yadav KDS, and Knowles PF (1994) Crystal structure of a free radical enzyme − Galactose oxidase. J. Mol. Biol, 238: 794-814

Kajikawa H, Kudo H, Kondo T, Jodai K, Honda Y, Kuwahara M, and Watanabe T (2000) Degradation of benzyl ether bonds of lignin by ruminal microbes. FEMS Microbiol Lett, 187: 15-20

Kanyong P, Pemberton RM, Jackson SK, and Hart JP (2013) Development of an amperometric screen-printed galactose biosensor for serum analysis. Anal. Biochem., 435: 114-119

Kato K, Kozaki S, and Sakuranaga M (1998) Degradation of lignin compounds by bacteria from termite guts. Biotechnol Lett, 20: 459-462

Kelly-Falcoz F, Greenberg H, and Horecker BL (1965) Galactose Oxidase - Studies on structure and role of disulfide linkages. J. Biol. Chem., 240: 2966-2970

Kersten PJ (1990) Glyoxal oxidase of Phanerochaete chrysosporium: its characterization and activation by lignin peroxidase. Proc Natl Acad Sci USA, 87: 2936–2940

Kersten PJ, and Kirk TK (1987) Involvement of a new enzyme, glyoxal oxidase, in extracellular H_2O_2 production by *Phanerochaete chrysosporium*. J. Bacteriol. , 169: 2195-2201

Kim MI, Shim J, Li T, Woo MA, Cho D, Lee J, and Park HG (2012) Colorimetric quantification of galactose using a nanostructured multi-catalyst system entrapping galactose oxidase and magnetic nanoparticles as peroxidase mimetics. Analyst, 137: 1137-1143

Kirby R (2005) Actinomycetes and lignin degradation. In Advances in Applied Microbiology, vol 58. Edited by Laskin AI, Bennett JW, Gadd GM, Sariaslani S. Waltham, MA: Academic Press: 125- 168

Kirk TK and Farrell RL (1987) Enzymatic combustion: the microbial degradation of lignin. Ann. Rev. Microbiol, 41: 465-505

Klibanov AM, Alberti BN, and Marletta MA (1982) Stereospecific oxidation of aliphatic alcohols catalyzed by galactose oxidase. Biochem. Biophys. Res. Commun. 108: 804-808

Kosman DJ (1984) Galactose Oxidase. In: LONTIE, R. (ed.) Copper Proteins and Copper Enzymes. CRC Press

Kosman DJ, Ettinger MJ, Weiner RE, and Massaro EJ (1974) The molecular properties of the copper enzyme galactose oxidase. Arch. Biochem. Biophys, 165: 456-467

Kuan I and Tien M (1993) Stimulation of Mn peroxidase activity: a possible role for oxalate in lignin biodegradation. Proc. Natl. Acad. Sci. USA, 90: 1242–1246

Kujawa M, Ebner H, Leitner C, Hallberg BM, Prongjit M, Sucharitakul J, Ludwig R, Rudsander U, Peterbauer C, Chaiyen P, *et al* (2006) Structural basis for substrate binding and regioselective oxidation of monosaccharides at C3 by pyranose 2-oxidase. J Biol Chem, 281: 35104-35115

Kurek B and Kersten PJ (1995) Physiological regulation of glyoxal oxidase from *Phanerochaete chrysosporium* by peroxidase systems. Enzyme Microb. Technol, 17: 751–756

Leitner C, Volc J, and Haltrich D (2001) Purification and characterization of pyranose oxidase from the white rot fungus *Trametes multicolor*. Appl Environ Microbiol, 67: 3636-3644

Lis M and Kuramitsu HK (1997) Galactose oxidase-glucan binding domain fusion proteins as targeting inhibitors of dental plaque bacteria. Antimicrob. Agents Chemother, 41: 999-1003

Liu XC and Dordick JS (1999) Sugar-containing polyamines prepared using galactose oxidase coupled with chemical reduction. J. Am. Chem. Soc., 121: 466-467

Lourenço RJM, Serralheiro MLM, and Rebelo MJF (2003) Development of a New Amperometric Biosensor for Lactose Determination. Portugaliae Electrochimica Acta 21: 171-177

Mahmoud KMAG, Spooner RK, Deacon S, and McPherson MJ (2000) Domain I of galactose oxidase is not required for catalytic activity. Biochem. Soc. Trans., 28: 72-72

Mannino S, Cosio MS, and Buratti S (1999) Simultaneous determination of glucose and galactose in dairy products by two parallel amperometric biosensors. Italian Journal of Food Science, 11: 57-65

Mansfield SD, Kim H, Lu F, and Ralph J (2012) Whole plant cell wall characterization using solution-state 2D NMR. Nat Protoc 2012, 7: 1579-1589

Martínez AT, Rencoret J, Marques G, Gutiérrez A, Ibarra D, Jiménez-Barbero J, and del Río JC (2008) Monolignol acylation and Lignin Structure in Some Nonwoody Plants: a 2D NMR Study. Phytochemistry, 16: 2831–2843

Martínez AT, Rencoret J, Nieto L, Jiménez-Barbero J, Gutiérrez A, and del Río JC (2011) Selective lignin and polysaccharide removal in natural fungal decay of wood as evidenced by in situ structural analyses. Environ Microbiol, 13: 96–107

Martínez AT, SperanzaM, Ruiz-Dueñas FJ, Ferreira P, Camarero S, Guillén F, Martínez MJ, Gutiérrez A, and del Río JC (2005) Biodegradation of lignocellulosics: microbiological, chemical and enzymatic aspects of fungal attack to lignin. Intern Microbiol, 8: 195–204

Martins LO, Durão P, Brissos V, and Lindley PF (2015) Laccases of prokaryotic origin – enzymes at the interface of protein science and protein technology. Cell Mol Life Sci., 72: 911-22

Marzullo L, Cannio R, Giardina P, Santini MT, Sannia G (1995) Veratryl alcohol oxidase from *Pleurotus ostreatus* participates in lignin biodegradation and prevents polymerization of laccase-oxidized substrates. J Biol Chem, 270: 3823–3827

Mazur AW and Hiler GD (1997) Chemoenzymic approaches to the preparation of 5-C-(hydroxymethyl)hexoses. J. Org. Chem, 62: 4471-4475

McPherson MJ, OGEL ZB, Stevens C, Yadav KDS, KEEN JN, and Knowles PF (1992) Galactose oxidase of *Dactylium dendroides* - Gene cloning and sequence analysis. J. Biol. Chem. 267: 8146-8152

McPherson MJ, Stevens C, Baron AJ, Ogel ZB, Seneviratne K, Wilmot C, Ito N, Brocklebank I, Phillips SEV, and Knowles PF (1993) Galactose oxidase: Molecular analysis and mutagenesis studies. Biochem. Soc. Trans, 21: 752-756

Mendes S, Gabriel A, Catarino T, Silveira C, Turner DL, Todorovic S, Martins LO (2015) An integrated view of redox and catalytic properties of B-type PpDyP from *Pseudomonas putida* MET94 and its distal variants Arch Biochem Biophys, 574: 99-107

Mendes S, Robalo MP, and Martins LO (2015) Bacterial Enzymes and Multi-Enzymatic Systems for Cleaning-up Dyes from the Environment in "Microbial Degradation of Synthetic Dyes in Waste Waters". Ed. S. N. Singh. Springer. pp. 27-55

Messerschmidt A, Huber R, Poulos T, and Wieghardt K (2001) Handbook of Metalloproteins, England, John Wiley & Sons, Ltd

Minasian SG, Whittaker MM, and Whittaker JW (2004) Stereoselective hydrogen abstraction by galactose oxidase. Biochemistry, 43: 13683-13693

Morel M, Ngadin AA, Jacquot J, and Gelhaye E (2009) Reactive oxygen species in *Phanerochaete chrysosporium*: Relationship between extracellular oxidative and intracellular antioxidant systems. Adv. Bot. Res., 52: 0065-2296

Morell AG, Van Den Hamer CJA, Scheinberg IH, and Ashwell G (1966) Physical and chemical studies on ceruloplasmin. J. Biol. Chem., 241: 3745-3749

Muheim A, Leisola MSA, and Schoemaker HE (1990) Aryl-alcohol oxidase and lignin peroxidase from the white-rot fungus *Bjerkandera adusta*. J Biotechnol, 13: 159–167

Murzin AG (1992) Structural principles for the propeller assembly of beta-sheets - The preference for 7-Fold symmetry. Proteins-Structure Function and Genetics, 14: 191-201

Nobles MK and Madhosingh C (1963) *Dactylium dendroides* (bull.) Fr. Misnamed as *Polyporus circinatus* Fr. Biochem. Biophys. Res. Commun. , 12: 146-147

Novagrodsky A and Katchalsky E (1973) Induction of lymphocyte transformation by sequential treatment with neuraminidase and galactose oxidase. Proc. Nat. Acad. Sci. USA, 70: 1824-1827

Ogel ZB, Brayford D, and McPherson MJ (1994) Cellulose-triggered sporulation in the galactose oxidase-producing fungus *Cladobotryum (Dactylium) dendroides* NRRL-2903 and its reidentification as a species of Fusarium. Mycol. Res., 98, 474-480

Ohno K and Kitano H (1998) Catalytic properties of galactose oxidase to liposome-forming amphiphiles which have many pendent galactose residues. Bioconjugate Chem, 9: 548-554

Parikka K, Leppanen AS, Xu CL, Pitkanen L, Eronen P, Osterberg M, Brumer H, Willfor S, and Tenkanen M (2012) Functional and Anionic Cellulose-Interacting Polymers by Selective Chemo-Enzymatic Carboxylation of Galactose-Containing Polysaccharides. Biomacromolecules, 13: 2418-2428

Paster M, Pellegrino JL, and Carole TM (2003) Industrial Bioproducts: Today and Tomorrow. Report prepared by Energetics, Inc., for the United States Department of Energy, Office of Energy Efficiency and Renewable Energy, Office of the Biomass Program, Washington, DC

Paukner R, Staudigl P, Choosri W, Haltrich D, Leitner C *et al*. (2015) Expression, purification, and characterization of galactose oxidase of *Fusarium sambucinum* in E. coli. Prot. Exp. Purif., 108: 73–79

Paukner R, Staudigl P, Choosri W, Sygmund C, Halada P, *et al*. (2014) Galactose Oxidase from *Fusarium oxysporum* - Expression in E. coli and P. pastoris and Biochemical Characterization. PLoS ONE 9(6): e100116.

Pellinen J, Vaisanen E, Salkinoja-Salonen M, and Brunow G (1984) Utilization of dimeric lignin model compounds by mixed bacterial cultures. Appl Microbiol Biotechnol, 20: 77-82

Pisanelli I, Kujawa M, Spadiut O, Kittl R, Halada P, Volc J, Mozuch MD, Kersten P, Haltrich D, and Peterbauer C (2009) Pyranose 2-oxidase from *Phanerochaete chrysosporium*–expression in E. coli and biochemical characterization. J Biotechnol, 142: 97-106

Prongjit M, Sucharitakul J, Wongnate T, Haltrich D, and Chaiyen P (2009) Kinetic mechanism of pyranose 2-oxidase from *Trametes multicolor*. Biochemistry, 48: 4170-4180

Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J, Eckert CA, Frederick WJ, Hallett JP, Leak DJ, Liotta CL, Mielenz JR, Murphy R, Templer R, and Tschaplinski T (2006) The path forward for biofuels and biomaterials. Science, 311: 484–489

Ralph J, Lundquist K, Brunow G, Lu F, Kim H, Schatz PF, Marita JM, Hatfield RD, Ralph SA, Christensen JH, and Boerjan W (2004). Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. Phytochemistry Reviews, 2: 29-60

Ramachandra M, Crawford DL, and Hertel G (1988) Characterization of an extracellular lignin peroxidase of the lignocellulolytic actinomycete *Streptomyces viridosporus*. Appl Environ Microbiol, 54: 3057-3063

Rannes JB, Ioannou A, Willies SC, Grogan G, Behrens C, Flitsch SL, and Turner NJ (2011) Glycoprotein labeling using engineered variants of Galactose Oxidase obtained by directed evolution. J. Am. Chem. Soc., 133: 8436-8439

Reiter J, Strittmatter H, Wiemann LO, Schiedera D, and Sieber V (2013) Enzymatic cleavage of lignin β-O-4 aryl ether bonds via net internal hydrogen transfer. Green Chem., 15: 1373–1381

Rogers MS, Baron AJ, McPherson MJ, Knowles PF, and Dooley DM (2000) Galactose oxidase pro-sequence cleavage and cofactor assembly are self-processing reactions. J. Am. Chem. Soc., 122: 990-991

Rorem E.S, and Lewis JC (1962) A test paper for detection of galactose and certain galactose-containing sugars. Anal. Biochem., 3: 230-235

Rosado T, Bernardo P, Koci K, Coelho AV, Robalo MP, and Martins LO (2012) Methyl syringate: an efficient phenolic mediator for bacterial and fungal laccases. Bioresour. Technol., 124: 371-378

Ruelius HW, Kerwin RM, and Janssen FW (1968) Carbohydrate oxidase, a novel enzyme from *Polyporus obtusus*. I. Isolation and purification. Biochim Biophys Acta, 167: 493-500

Ruiz-Dueñas FJ and Martínez AT (2009). Microbial Degradation of Lignin: How a Bulky Recalcitrant Polymer is Efficiently Recycled in Nature and How We Can Take Advantage of this. Microbial Biotechnol, 2: 164–177

Ruiz-Dueñas FJ, Morales M, García E, Miki Y, Martínez MJ, and Martínez AT (2009) Substrate oxidation sites in versatile peroxidase and other basidiomycete peroxidases. J Exp Bot, 60: 441–452

Salaheddin *et al*. (2010) Characterisation of recombinant pyranose oxidase from the cultivated mycorrhizal basidiomycete *Lyophyllum shimeji* (hon-shimeji). Microbial Cell Factories, 9: 57-69

Salvachúa D, Prieto A, Lopez-Abelairas M, Lú-Chau T, Martínez AT, and Martínez MJ (2011) Fungal pretreatment: an alternative in second-generation ethanol from wheat straw. Bioresource Technol, 102: 7500–7506

Sambrook J, Fritsch E, and Maniatus T (1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory Press.

Sanchez C (2009) Lignocellulosic residues: biodegradation and bioconversion by fungi. Biotechnol Adv, 27: 185-194

Santa Cruz-Calvo L, González-López J, and Manzanera M, (2013) *Arthrobacter siccitolerans* sp. nov., a highly desiccation-tolerant, xeroprotectant-producing strain isolated from dry soil. Intern Microbiol, 63: 4174–4180

Santos A, Mendes S, Brissos V, Martins LO (2014) New dye decolourising peroxidases from *Bacillus subtilis* and *Pseudomonas putida*: towards biotechnological applications. Appl. Microbial Biotechnol, 98: 2053-2065

Sarkanen KV and Ludwig CH (1971) Lignins: Occurrence, Formation, Structure, and Reactions. New York: Wiley-Intersci.

Schlegel RA, Gerbeck CM, and Montgomery R (1968) Substrate specificity of D-galactose oxidase. Carboh Res., 7: 193-199

Schoevaart R and Kieboom T (2004) Application of galactose oxidase in chemoenzymatic one-pot cascade reactions without intermediate recovery steps. Topics in Catalysis, 27: 3-9

Sezer M, Santos A, Kielb P, Pinto T, Martins LO, Todorovic S (2013) Distinct structural and redox properties of the heme active site in bacterial dye decolorizing peroxidase-type peroxidases from two subfamilies: resonance Raman and electrochemical study. Biochemistry, 52: 3074-3084

Shamssuddin AM and Elsayen AM (1988) A test for detection of colorectal-cancer. Human Pathology, 19: 7-10

Shamsuddin AM, Tyner GT, and Yang GY (1995) Common expression of the tumor-marker D-galactose-β-[1-]3]-N-acetyl-D-galactosamine by different adenocarcinomas - Evidence of field-effect phenomenon. Cancer Res. 55: 149-152

Sharma SK, Kumar A, Chaudhary R, Suman, Pundir CS, and Sehgal N (2007) Lactose biosensor based on lactase and galactose oxidase immobilized in polyvinyl formal. Artif Cells Blood Substit Immobil Biotechnol, 35: 421-30

Sharma SK, Suman, Pundir CS, Sehgal N, and Kumar A (2006) Galactose sensor based on galactose oxidase immobilized in polyvinyl formal. Sensors and Actuators B-Chemical, 119: 15-19

Shi W, Syrenne R, Sun J-Z, and Yuan JS (2010) Molecular approaches to study the insect gut symbiotic microbiota at the 'omics' age. Insect Sci, 17: 199-219

Singh H and Kanfer JN (1980) Quantitation of galactocerebrosides and sulfatides by Galactose-Oxidase and sodium borotritide. Anal. Biochem., 109: 27-31

Sixta H (2006) Handbook of Pulp. Weinheim, Germany: Wiley-VCH

Sjöström E (1993). Wood Chemistry: Fundamentals and Application. Academic Press: Orlando. 293

Solomon EI, Sundaram UM, and Machonkin TE (1996) Multicopper oxidases and oxygenases. Chem. Rev, 96: 2563–2605

Sonoki T, Iimura Y, Masai E, Kajita S, and Katayama Y (2002) Specific degradation of β-aryl ether linkage in synthetic lignin (dehydrogenative polymerizate) by bacterial enzymes of *Sphingomonas paucimobilis* SYK-6 produced in recombinant Escherichia coli. J Wood Sci, 48: 429-433

Stoj CS and Kosman DJ (2005) Copper proteins: Oxidases. In: King RB (ed) Encyclopedia of inorganic chemistry, Vol II, 2nd edn. John Wiley and Sons, New York, 1134-1159

Sugano Y (2009) DyP-type peroxidases comprise a novel heme peroxidase family. Cell Mol Life Sci, 66: 1387-1403

Sugano Y, Muramatsu R, Ichiyanagi A, Sato T, and Shoda M (2000) DyP, a unique dye-decolorizing peroxidase, represents a novel heme peroxidase family. J Biol Chem, 282: 36652-36658

Takakura Y and Kuwata S (2003) Purification, characterization, and molecular cloning of a pyranose oxidase from the fruit body of the basidiomycete, Tricholoma matsutake. Biosci Biotechnol Biochem, 67: 2598-2607

Taylor CR, Hardiman EM, Ahmad M, Sainsbury PD, Norris PR, and Bugg TDH (2012) Isolation of bacterial strains able to metabolize lignin from screening of environmental samples. J Appl Microbiol, 113: 521-530

Taylor PJ, Kmetec E, and Johnson JM (1977) Design, construction, and applications of a galactose selective electrode. Anal. Chem., 49: 789-794

Taylor TN and Osborne JM (1996) The importance of fungi in shaping the paleoecosystem. Rev Paleobot Palyn, 90: 249–262

Tressel P and Kosman DJ (1980) O,O-Dityrosine in native and horseradish peroxidase-activated galactose-oxidase. Biochem. Biophys. Res. Commun. , 92: 781-786

van der Meer RA, Jongejan JA and Duine JA (1989) Pyrroloquinoline quinone as cofactor in galactose oxidase. J. Biol. Chem, 264: 7792-7794

Varela E, Martínez AT, and Martínez MJ (1999) Molecular cloning of aryl-alcohol oxidase from *Pleurotus eryngii*, an enzyme involved in lignin degradation. Biochem J, 341: 113–117

Varela E, Martínez MJ, and Martínez AT (2000b) Aryl-alcohol oxidase protein sequence: a comparison with glucose oxidase and other FAD oxidoreductases. Biochim Biophys Acta, 1481: 202–208

Vrbova E, Peckova J, and Marek M (1992) Preparation and utilization of a biosensor based on galactose oxidase. Collect. Czech. Chem. Commun, 57: 2287-2294

Wachter RM and Branchaud BP (1996) Molecular modelling studies on oxidation of hexopyranoses by galactose oxidase - An active site topology apparently designed to catalyze radical reactions, either concerted or stepwise. J. Am. Chem. Soc, 118: 2782-2789

Wang Y, Liu Q, Yan L, Gao Y, Wang Y, and Wang W (2013) A novel lignin degradation bacterial consortium for efficient pulping. Bioresource Technology, 139: 113–119

Warnecke F, Luginbuhl P, Ivanova N, Ghassemian M, Richardson TH, Stege JT, Cayouette M, McHardy AC, Djordjevic G, Aboushadi N *et al*. (2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. Nature, 450: 560-565

Whittaker JW (2002) Galactose oxidase. Adv Protein Chem, 60: 1–49

Whittaker JW (2003) Free Radical Catalysis by Galactose Oxidase. Chem. Rev, 103: 2347-2363

Whittaker JW (2005) The radical chemistry of galactose oxidase. Arch. Biochem. Biophys. , 433: 227-239

Whittaker JW and Whittaker MM (1990) A tyrosine-derived free radical in apogalactose oxidase. J. Biol. Chem, 265: 9610-9613

Whittaker JW and Whittaker MM (1998) Radical copper oxidases, one electron at a time. Pure Appl. Chem, 70: 903-910

Whittaker MM and Whittaker JW (1988) The active site of galactose oxidase. J. Biol. Chem, 263: 6074-6080

Whittaker MM and Whittaker JW (1993) Ligand interactions with galactose oxidase mechanistic insights. Biophys. J, 64: 762-772

Whittaker MM and Whittaker JW (2006) *Streptomyces coelicolor* oxidase (SCO2837p): A new free radical metalloenzyme secreted by Streptomyces coelicolor A3(2). Arch Biochem and Biophysics, 452: 108–118

Whittaker MM, Ballou DP, and Whittaker JW (1998) Kinetic isotope effects as probes of the mechanism of galactose oxidase. Biochemistry, 37: 8426-8436

Whittaker MM, DeVito VL, Asher SA, and Whittaker JW (1989) Resonance Raman evidence for tyrosine involvement in the radical site of galactose oxidase. J. Biol. Chem., 264: 7104-7106

Whittaker MM, Kersten PJ, Cullen D, and Whittaker JW (1999) Identification of catalytic residues in glyoxal oxidase by targeted mutagenesis. J. Biol. Chem., 274: 36226-36232

Whittaker MM, Kersten PJ, Nakamura N, Sanders-Loehr J, Schweizer ES, and Whittaker JW (1996) Glyoxal oxidase from *Phanerochaete chrysosporium* is a new Radical-Copper Oxidase. J. Biol. Chem., 271: 681-687

Wilkinson D, Akumanyi N, Hurtado Guerrero R, Dawkes H, Knowles PF, *et al*. (2004) Structural and kinetic studies of a series of mutants of galactose oxidase identified by directed evolution. Prot Eng Design Selec, 17: 141–148

Xu F, Golightly EJ, Schneider P, Berka RM, Brown KM, Johnstone JA, Baker DH, Fuglsang CC, Brown SH, Svendsen A, and Klotz AV (2000) Expression and characterization of a recombinant *Fusarium* spp. galactose oxidase. Appl. Biochem. Biotechnol, 88: 23-32

Yadav KK, Vernwal SK, Afaq Z, and Yadav KDS (2002) Enzymatic preparation of L-glucose, L-galactose and L-xylose using galactose oxidase immobilised on crab-shell particles. J. Sci. Ind. Res, 61: 361-365

Yalpani M and Hall LD (1982) A high-yielding, specific method for the chemical derivatization of galactose-containing polysaccharides - oxidation with galactose-oxidase followed by reductive amination. J. Polymer Sc. Part A-Polymer Chemistry, 20: 3399-3420

Yelle DJ, Wei DS, Ralph J, and Hammel KE (2011) Multidimensional NMR analysis reveals truncated lignin structures in wood decayed by the brown rot basidiomycete. Environ Microbiol, 13: 1091–1100

Zafar MN, Pittman S, Cherchi M, and Catovsky D (1981) Stimulation of chronic lymphatic-leukemia cells by pokeweed mitogen after treatment with neuraminidase-Galactose Oxidase. Clin. Exp. Immunol. , 44: 124-128

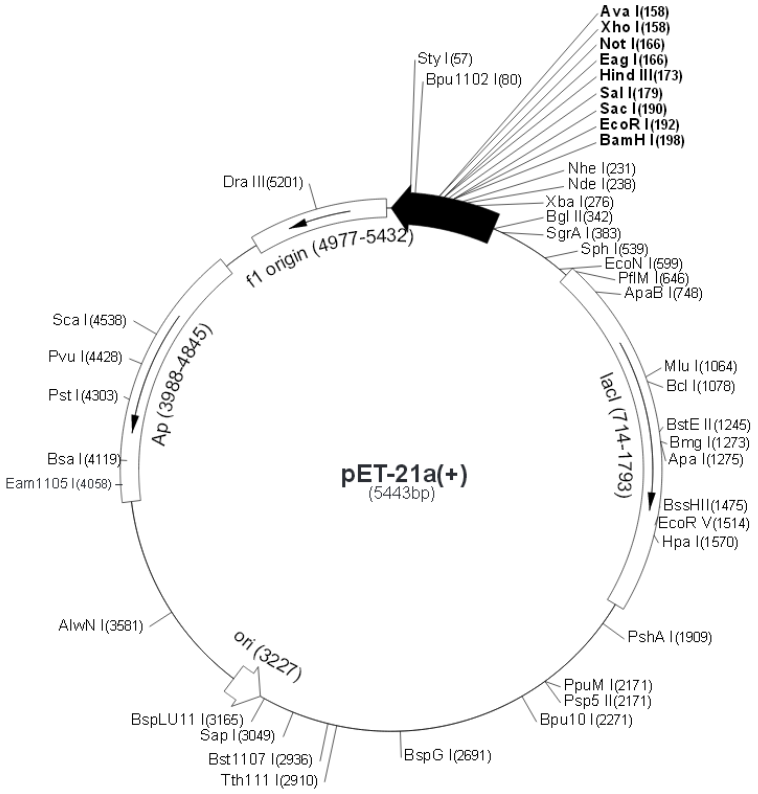
Zimmermann W (1990) Degradation of lignin by bacteria. J Biotechnol, 13: 119-130

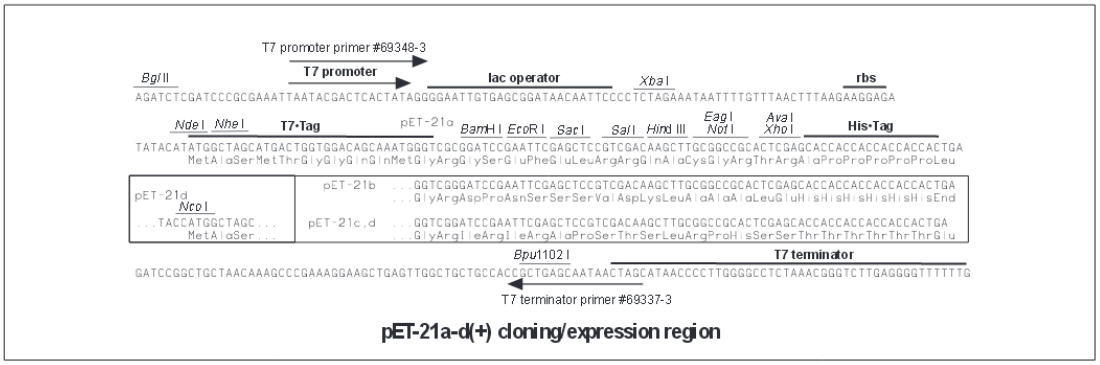
6. APPENDICES

Appendix A

pET-21a(+) sequence landmarks	
T7 promoter	311-327
T7 transcription start	310
T7•Tag coding sequence	207-239
Multiple cloning sites (*Bam*H I - *Xho* I)	158-203
His•Tag coding sequence	140-157
T7 terminator	26-72
lacI coding sequence	714-1793
pBR322 origin	3227
bla coding sequence	3988-4845
f1 origin	4977-5432

The maps for pET-21b(+), pET-21c(+) and pET-21d(+) are the same as pET-21a(+) (shown) with the following exceptions: pET-21b(+) is a 5442bp plasmid; subtract 1bp from each site beyond *Bam*H I at 198. pET-21c(+) is a 5441bp plasmid; subtract 2bp from each site beyond *Bam*H I at 198. pET-21d(+) is a 5440bp plasmid; the *Bam*H I site is in the same reading frame as in pET-21c(+). An *Nco* I site is substituted for the *Nde* I site with a net 1bp deletion at position 238 of pET-21c(+). As a result, *Nco* I cuts pET21d(+) at 234, and *Nhe* I cuts at 229. For the rest of the sites, subtract 3bp from each site beyond position 239 in pET-21a(+). *Nde* I does not cut pET-21d(+). Note also that *Sty* I is not unique in pET-21d(+).





Appendix B

Table S1- Details of the 90 bacterial sequences showing the highest identity to GlyOx from *P. chrysosporium.* Genes 1 to 7 indicated in the table 2.2. are highlighted in bold.

Description	Identity	Accession
kelch motif family protein [Burkholderia pseudomallei]	24%	AIP72504.1
galactose oxidase [Burkholderia pseudomallei]	24%	KKI76564.1
galactose oxidase [Gloeobacter violaceus]	**25%**	**NP_927197.1**
hypothetical protein [Solirubrobacter soli]	24%	WP_037498701.1
hypothetical protein [Pseudogulbenkiania ferrooxidans]	25%	WP_021477152.1
Galactose oxidase [Frankia symbiont of Datisca glomerata]	24%	AEH09003.1
kelch domain-containing protein [Acidobacteriaceae bacterium KBS 83]	25%	WP_035222260.1
galactose oxidase [Frankia sp. BMG5.1]	24%	WP_047224956.1
hypothetical protein [Microcystis aeruginosa]	23%	WP_002786378.1
galactose oxidase [Actinoplanes sp. SE50/110]	23%	WP_043515771.1
galactose oxidase [Kitasatospora mediocidica]	23%	WP_035797022.1
hypothetical protein [Streptomyces prunicolor]	25%	WP_019065236.1
hypothetical protein [Conexibacter woesei]	23%	WP_027007190.1
galactose oxidase [Streptacidiphilus jeojiense]	25%	WP_037598087.1
hypothetical protein [Stigmatella aurantiaca]	**24%**	**WP_002617931.1**
hypothetical protein [Arthrobacter sp. 9MFCol3.1]	24%	WP_026537250.1
hypothetical protein [Arthrobacter sp. 131MFCol6.1]	23%	WP_043466075.1
galactose oxidase [Streptomyces peruviensis]	24%	WP_030047441.1
galactose oxidase [Meiothermus cerbereus]	24%	WP_027876292.1
Galactose oxidase [Meiothermus silvanus DSM 9946]	24%	ADH63319.1
hypothetical protein [Meiothermus silvanus]	24%	WP_041652412.1
hypothetical protein [Cyanothece sp. PCC 7822]	22%	WP_013334263.1
hypothetical protein [Streptomyces collinus]	24%	WP_020941904.1
galactose oxidase [Streptomyces mirabilis]	24%	WP_037740247.1
hypothetical protein [Actinokineospora inagensis]	25%	WP_035304644.1
hypothetical protein [Methyloglobulus morosus]	22%	WP_023495756.1
galactose oxidase [Actinoplanes rectilineatus]	23%	WP_045746824.1
hypothetical protein, partial [Nitrosospira lacus]	28%	WP_040851844.1
hypothetical protein [Lechevalieria aerocolonigenes]	24%	WP_030469926.1
hypothetical protein DB31_5318 [Hyalangium minutum]	24%	KFE70276.1
hypothetical protein [Arthrobacter sp. L77]	24%	WP_043444640.1
hypothetical protein [Nostoc sp. PCC 7120]	24%	WP_010995753.1
hypothetical protein [Hyalangium minutum]	24%	WP_044186153.1

hypothetical protein [Pseudanabaena biceps]	22%	WP_040687471.1
galactose oxidase [Streptomyces sp. NRRL S-118]	23%	WP_031068564.1
hypothetical protein [Arthrobacter sp. SPG23]	24%	WP_043482214.1
galactose oxidase [Streptomyces niveus]	26%	WP_031229249.1
hypothetical protein [Acidobacteriaceae bacterium KBS 83]	21%	WP_035223630.1
hypothetical protein [Cystobacter violaceus]	25%	WP_043410852.1
hypothetical protein [Blastomonas sp. AAP53]	24%	WP_040370344.1
galactose oxidase [Streptomyces virginiae]	24%	WP_037942203.1
galactose oxidase [Streptomyces chartreusis]	23%	WP_029181573.1
hypothetical protein [Anabaena variabilis]	24%	WP_011320876.1
hypothetical protein ACPL_4146 [Actinoplanes sp. SE50/110]	23%	AEV85041.1
hypothetical protein [Limnoraphis robusta]	23%	WP_046277970.1
galactose oxidase [Streptomyces violaceoruber]	24%	WP_030941853.1
DUF1929 domain-containing protein [Cylindrospermum stagnale]	23%	WP_015210248.1
galactose oxidase [Streptomyces sp. NRRL F-2747]	24%	WP_037846722.1
hypothetical protein MYSTI_05030 [Myxococcus stipitatus DSM 14675]	**23%**	**AGC46318.1**
galactose oxidase [Streptomyces sp. RSD-27]	24%	WP_042811049.1
hypothetical protein [Streptomyces scopuliridis]	24%	WP_030355043.1
galactose oxidase [Streptomyces niveus]	22%	WP_037784611.1
galactose oxidase [Actinoplanes sp. SE50/110]	23%	WP_043514838.1
putative copper-dependent oxidase [Stigmatella aurantiaca]	22%	AAL25195.1
galactose oxidase [Streptomyces sp. NRRL S-475]	24%	WP_030842479.1
hypothetical protein [Stigmatella aurantiaca]	22%	WP_002617156.1
galactose oxidase [Streptomyces sp. WM6386]	24%	WP_046260831.1
galactose oxidase [Streptomyces tsukubaensis]	23%	WP_040914640.1
galactose oxidase [Streptomyces viridochromogenes]	24%	WP_048586170.1
galactose oxidase [Meiothermus ruber]	22%	WP_013012666.1
galactose oxidase [Streptomyces sp. NRRL WC-3744]	24%	WP_030989673.1
galactose oxidase [Streptomyces cellulosae]	24%	WP_030666701.1
galactose oxidase [Streptomyces regensis]	24%	KMS86259.1
galactose oxidase [Meiothermus ruber]	22%	WP_036199241.1
hypothetical protein MYSTI_00845 [Myxococcus stipitatus DSM 14675]	**24%**	**AGC42194.1**
galactose oxidase [Meiothermus ruber DSM 1279]	22%	AGK03600.1
galactose oxidase [Streptomyces sp. NRRL WC-3725]	24%	WP_031028837.1
conserved hypothetical protein [Raphidiopsis brookii D9]	23%	EFA71896.1
hypothetical protein [Actinoplanes sp. N902-109]	23%	WP_015621785.1
galactose oxidase [Streptomyces sp. 303MFCol5.2]	24%	WP_020128082.1
galactose oxidase [Streptomyces sp. NRRL F-4474]	24%	WP_037834645.1
galactose oxidase [Streptomyces nodosus]	24%	WP_043440481.1
hypothetical protein AFR_18160 [Actinoplanes friuliensis DSM 7358]	23%	AGZ41910.1
galactose oxidase [Streptomyces sp. NRRL WC-3795]	24%	WP_031018860.1

galactose oxidase [Streptomyces sp. MBT28]	24%	WP_046248203.1
hypothetical protein [Oscillatoria acuminata]	23%	WP_015151626.1
galactose oxidase [Streptomyces bicolor]	24%	WP_031483601.1
galactose oxidase [Streptomyces sp. TOR3209]	24%	WP_037909413.1
galactose oxidase [Streptomyces clavuligerus]	**23%**	**WP_003962715.1**
galactose oxidase [Actinoplanes friuliensis]	23%	WP_041842374.1
hypothetical protein [Nocardia vinacea]	25%	WP_040698768.1
galactose oxidase [Streptomyces sp. FxanaA7]	24%	WP_045563175.1
galactose oxidase [Streptomyces sp. NRRL WC-3742]	28%	WP_037973874.1
galactose oxidase [Streptomyces sp. PBH53]	23%	AKN75379.1
galactose oxidase [Chroococcidiopsis thermalis]	25%	WP_015153763.1
galactose oxidase [Streptomyces avermitilis]	23%	WP_037645234.1
hypothetical protein [Amycolatopsis methanolica]	23%	WP_038533215.1
galactose oxidase [Arthrobacter siccitolerans]	**22%**	**CCQ45078.1**
galactose oxidase [Stigmatella aurantiaca]	**18%**	**ADO74264.1**

Table S2 - Details of 130 bacterial sequences showing the highest identity to GalOx from *F. graminearum*. Genes 1 to 7 indicated in the table 2.2. are highlighted in bold.

Description	Identity	Accession
arabinogalactan endo-1 4-beta-galactosidase [Actinoplanes subtropicus]	46%	WP_030440774.1
arabinogalactan endo-1 4-beta-galactosidase [Actinokineospora enzanensis]	45%	WP_033400428.1
hypothetical protein [Actinoplanes globisporus]	48%	WP_020510009.1
Galactose oxidase precursor [Actinokineospora spheciospongiae]	47%	EWC62380.1
galactose oxidase [Arthrobacter siccitolerans]	**47%**	**CCQ45078.1**
arabinogalactan endo-1 4-beta-galactosidase [Saccharothrix syringae]	45%	WP_037330498.1
arabinogalactan endo-1 4-beta-galactosidase [Lechevalieria aerocolonigenes]	49%	WP_035910667.1
galactose oxidase [Burkholderiaceae bacterium 16]	47%	WP_045237679.1
galactose oxidase [Cupriavidus sp. SK-3]	46%	KDP85841.1
galactose oxidase [Variovorax paradoxus]	46%	WP_012747437.1
galactose oxidase [Burkholderia fungorum]	46%	WP_046566618.1
hypothetical protein [Variovorax paradoxus]	45%	WP_018906126.1
galactose oxidase [Cupriavidus sp. SK-3]	46%	KDP85828.1
galactose oxidase [Burkholderia glathei]	44%	KDR37639.1
galactose oxidase [Burkholderia mimosarum]	46%	WP_028209577.1
galactose oxidase [Methylibium sp. CF468]	44%	WP_047506310.1
galactose oxidase [Burkholderia sp. ABCPW 1]	45%	WP_038729779.1
hypothetical protein [Methylobacterium aquaticum]	45%	WP_048462414.1
galactose oxidase [Methylibium sp. T29-B]	42%	WP_036235640.1
galactose oxidase [Sphaerotilus natans]	42%	WP_037478273.1
galactose oxidase [Stigmatella aurantiaca]	**39%**	**ADO74264.1**
Galactose oxidase [Burkholderia sp. H160]	44%	EDZ97950.1
hypothetical protein [Methylopila sp. M107]	44%	WP_040577414.1
cytolethal distending toxin A/C family protein [Burkholderia pseudomallei]	44%	KGS99202.1
galactose oxidase [Burkholderia pseudomallei]	44%	CRY25856.1
cytolethal distending toxin A/C family protein [Burkholderia pseudomallei]	44%	KGD15792.1
Galactose oxidase-related protein [Burkholderia pseudomallei 1258a]	44%	EIF64344.1
hypothetical protein [Chromobacterium haemolyticum]	43%	WP_043590065.1
hypothetical protein [[Polyangium] brachysporum]	39%	WP_047193637.1

hypothetical protein X805_06830 [Sphaerotilus natans subsp. natans DSM 6575]	42%	KDB53705.1
galactose oxidase [Burkholderia caledonica]	43%	WP_027777999.1
galactose oxidase [Burkholderia sp. JPY251]	46%	WP_026228577.1
hypothetical protein [Burkholderia bryophila]	43%	WP_020069811.1
hypothetical protein [Leucothrix mucor]	43%	WP_040504286.1
hypothetical protein [Methylopila sp. M107]	44%	WP_040577415.1
hypothetical protein [Chromobacterium sp. C-61]	42%	WP_019101693.1
hypothetical protein [Gelidibacter mesophilus]	42%	WP_035479207.1
galactose oxidase [Methylobacterium extorquens]	41%	WP_012605219.1
hypothetical protein [Methylococcaceae bacterium 73a]	35%	WP_045225564.1
hypothetical protein [Colwellia psychrerythraea]	34%	WP_033082486.1
hypothetical protein [Marinobacter sp. MCTG268]	35%	WP_036205421.1
galactose oxidase [Hahella chejuensis]	32%	WP_011397221.1
hypothetical protein [Deinococcus misasensis]	31%	WP_034341839.1
hypothetical protein [Neosynechococcus sphagnicola]	31%	WP_036533638.1
hypothetical protein [Meiothermus silvanus]	32%	WP_041652412.1
Galactose oxidase [Meiothermus silvanus DSM 9946]	32%	ADH63319.1
hypothetical protein [Hahella ganghwensis]	31%	WP_020407857.1
hypothetical protein [Hymenobacter sp. MIMtkLc17]	32%	WP_046247228.1
kelch domain-containing protein [Acidobacteriaceae bacterium KBS 83]	34%	WP_035222260.1
hypothetical protein [Meiothermus chliarophilus]	31%	WP_036219226.1
galactose oxidase [Meiothermus ruber DSM 1279]	28%	AGK03600.1
galactose oxidase [Meiothermus rufus]	29%	WP_027882868.1
hypothetical protein [Hyalangium minutum]	30%	WP_044186153.1
galactose oxidase [Meiothermus cerbereus]	29%	WP_027876292.1
hypothetical protein [Meiothermus chliarophilus]	31%	WP_036219228.1
hypothetical protein [Deinococcus peraridilitoris]	29%	WP_041231486.1
hypothetical protein MYSTI_05030 [Myxococcus stipitatus DSM 14675]	**30%**	**AGC46318.1**
protein of unknown function (DUF1929) [Deinococcus peraridilitoris DSM 19664]	29%	AFZ67980.1
hypothetical protein [Solirubrobacter soli]	32%	WP_037498701.1
hypothetical protein MFUL124B02_04795 [Myxococcus fulvus 124B02]	30%	AKF84810.1
galactose oxidase [Actinoplanes sp. N902-109]	31%	WP_015621352.1
hypothetical protein [Myxococcus stipitatus]	30%	WP_044282935.1
hypothetical protein MYSTI_00845 [Myxococcus stipitatus DSM 14675]	**30%**	**AGC42194.1**
hypothetical protein [Actinokineospora inagensis]	33%	WP_035304644.1
kelch domain-containing protein [Myxococcus stipitatus]	31%	WP_044900611.1
galactose oxidase [Gloeobacter violaceus]	**28%**	**NP_927197.1**
Galactose oxidase precursor [Archangium gephyra]	28%	AKJ08173.1

hypothetical protein [Archangium gephyra]	28%	WP_047863243.1
hypothetical protein [Cystobacter violaceus]	28%	WP_043410852.1
kelch domain-containing protein [Edaphobacter aggregans]	30%	WP_035356820.1
galactose oxidase [Actinoplanes utahensis]	30%	WP_043532124.1
galactose oxidase [Burkholderia graminis]	45%	WP_044021995.1
hypothetical protein [Aquabacterium sp. NJ1]	32%	WP_035041416.1
hypothetical protein [Modestobacter marinus]	30%	WP_041796556.1
hypothetical protein [Actinokineospora enzanensis]	32%	WP_033400636.1
kelch repeat-containing protein [Nostoc punctiforme]	28%	WP_012409261.1
Putative galactose oxidase [Modestobacter marinus]	30%	CCH87947.1
Kelch domain protein [Myxococcus xanthus DK 1622]	28%	ABF88956.1
galactose oxidase [Actinoplanes globisporus]	29%	WP_040432629.1
hypothetical protein [Enterovibrio calviensis]	29%	WP_039735715.1
hypothetical protein [Conexibacter woesei]	28%	WP_027007190.1
galactose oxidase [Nitrosomonas sp. AL212]	29%	WP_013648874.1
Galactose oxidase precursor [Archangium gephyra]	29%	AKJ03338.1
galactose oxidase [Frankia sp. BMG5.1]	30%	WP_047224956.1
hypothetical protein [Leucothrix mucor]	29%	WP_022953373.1
hypothetical protein [Deinococcus apachensis]	29%	WP_040383602.1
hypothetical protein [Lysobacter antibioticus]	29%	WP_036149502.1
galactose oxidase [Methylosinus sp. LW4]	29%	WP_043332818.1
hypothetical protein [Acidobacteriaceae bacterium KBS 83]	28%	WP_035223630.1
galactose oxidase [Actinoplanes sp. SE50/110]	30%	WP_014694539.1
hypothetical protein [Sorangium cellulosum]	27%	WP_012237503.1
Galactose oxidase [Frankia symbiont of Datisca glomerata]	29%	AEH09003.1
hypothetical protein [Stigmatella aurantiaca]	**29%**	**WP_002617931.1**
hypothetical protein [Myxococcus xanthus]	29%	WP_020477777.1
hypothetical protein AAW51_2416 [[Polyangium] brachysporum]	30%	AKJ29107.1
hypothetical protein, partial [Actinoplanes subtropicus]	27%	WP_034214340.1
hypothetical protein [Archangium gephyra]	29%	WP_047862268.1
putative copper-dependent oxidase [Stigmatella aurantiaca]	28%	AAL25195.1
hypothetical protein [Aureimonas ureilytica]	28%	WP_019995593.1
hypothetical protein, partial [Polaromonas glacialis]	29%	WP_029527695.1
galactose oxidase [Beijerinckia indica]	28%	WP_012385679.1
galactose oxidase [Streptomyces sp. NRRL F-5008]	28%	WP_030783542.1
hypothetical protein [Methylibium sp. CF059]	29%	WP_047494934.1
galactose oxidase [Scytonema tolypothrichoides VB-61278]	28%	KIJ85569.1
galactose oxidase [Methylobacterium extorquens]	27%	WP_003602644.1
galactose oxidase [Streptomyces iakyrus]	28%	WP_033311290.1
hypothetical protein [Aquabacterium sp. NJ1]	29%	WP_035038762.1
putative Galactose oxidase [Methylobacterium extorquens AM1]	27%	ACS38886.1
galactose oxidase [Chroococcidiopsis thermalis]	27%	WP_015153763.1

galactose oxidase [Streptomyces bicolor]	28%	WP_031483601.1
hypothetical protein [Cystobacter violaceus]	28%	WP_043409563.1
galactose oxidase [Streptomyces durhamensis]	29%	WP_031163897.1
hypothetical protein [Methyloglobulus morosus]	25%	WP_023495756.1
hypothetical protein [Alteromonadaceae bacterium Bs12]	27%	WP_045861546.1
galactose oxidase [Streptomyces toyocaensis]	28%	WP_037932663.1
hypothetical protein M271_30460 [Streptomyces rapamycinicus NRRL 5491]	31%	AGP57527.1
galactose oxidase [Streptomyces roseochromogenus]	29%	WP_023547107.1
galactose oxidase [Scytonema millei]	27%	WP_039715878.1
hypothetical protein [[Polyangium] brachysporum]	28%	WP_047193573.1
galactose oxidase [Kitasatospora mediocidica]	28%	WP_035797022.1
galactose oxidase [Streptomyces mutabilis]	28%	WP_043383326.1
galactose oxidase [Streptacidiphilus carbonis]	27%	WP_042403307.1
hypothetical protein [Methylotenera sp. 1P/1]	27%	WP_033424972.1
galactose oxidase [Streptomyces clavuligerus]	**27%**	**WP_003962715.1**
hypothetical protein [Lautropia mirabilis]	26%	WP_040530453.1
hypothetical protein [Methylocapsa aurea]	27%	WP_036257144.1
hypothetical protein [Lysobacter daejeonensis]	28%	WP_036138908.1
hypothetical protein [Streptomyces wedmorensis]	48%	WP_033208135.1
hypothetical protein [Streptomyces collinus]	28%	WP_020941904.1

Table S3 - Details of 40 bacterial sequences showing the highest identity to POx from *T. matsutake.* Gene 8 indicated in the table 2.2. is highlighted in bold.

Description	Identity	Accession
pyranose oxidase [Kitasatospora azatica]	42%	WP_035850787.1
pyranose oxidase [Streptomyces aureofaciens]	43%	WP_046385855.1
pyranose oxidase [Streptomyces flavidovirens]	40%	WP_028814754.1
Pyranose oxidase [Streptomyces somaliensis]	40%	WP_010473650.1
pyranose oxidase [Streptomyces sp. NRRL S-920]	40%	WP_030777486.1
hypothetical protein [Streptomyces davawensis]	40%	WP_015655935.1
pyranose oxidase [Amycolatopsis japonica]	38%	WP_038519531.1
Pyranose oxidase [Amycolatopsis decaplanina DSM 44594]	38%	EME57748.1
hypothetical protein [Nocardiopsis halotolerans]	36%	WP_017571642.1
CetL [Actinomyces sp. Lu 9419]	43%	ACH85575.1
putative Pyranose oxidase precursor (Glucose 2-oxidase) [Frankia alni ACN14a]	35%	CAJ63283.1
pyranose oxidase [Frankia alni]	36%	WP_041940790.1
pyranose oxidase [Streptomyces cyaneogriseus]	37%	WP_044388762.1
choline dehydrogenase [Modestobacter marinus]	35%	WP_041795258.1
choline dehydrogenase [Geodermatophilaceae bacterium URHA0031]	35%	WP_026845614.1
Pyranose oxidase [Modestobacter marinus]	36%	CCH88189.1
choline dehydrogenase [Microbacterium trichothecenolyticum]	34%	WP_045298611.1
choline dehydrogenase [Arthrobacter sp. I3]	34%	WP_043434989.1
choline dehydrogenase [Geodermatophilaceae bacterium URHB0048]	34%	WP_029336749.1
pyranose oxidase [Microbacterium sp. B19]	33%	WP_022877589.1
choline dehydrogenase [Microbacterium testaceum]	33%	WP_013584811.1
choline dehydrogenase [Microbacterium sp. SUBG005]	33%	KEP75059.1
choline dehydrogenase [Arthrobacter phenanthrenivorans]	33%	WP_043453839.1
choline dehydrogenase [Microbacterium sp. ZOR0019]	33%	WP_047520202.1
choline dehydrogenase [Nocardioides insulae]	31%	WP_028659842.1
choline dehydrogenase [Microbacterium yannicii]	32%	WP_040569273.1
GMC oxidoreductase family protein [Arthrobacter siccitolerans]	**32%**	**CCQ48064.1**
Pyranose oxidase (glucose 2-oxidase) [Streptomyces himastatinicus]	30%	WP_009712611.1
choline dehydrogenase [Microbacterium mangrovi]	33%	WP_039399471.1
choline dehydrogenase [Microbacterium hydrocarbonoxydans]	32%	WP_045257742.1
hypothetical protein [Microbacterium hydrocarbonoxydans]	32%	WP_045257736.1
pyranose oxidase [Microbacterium sp. 292MF]	32%	WP_018170256.1
oxidoreductase [Arthrobacter globiformis]	32%	WP_003803563.1
pyranose oxidase [Arthrobacter phenanthrenivorans]	32%	WP_013599330.1

choline dehydrogenase [Marmoricola sp. URHB0036]	33%	WP_036226406.1
choline dehydrogenase [Pseudonocardia autotrophica]	32%	WP_037039929.1
hypothetical protein [Micromonospora sp. ATCC 39149]	30%	WP_036372947.1
pyranose oxidase precursor [Pluralibacter gergoviae]	31%	WP_043083901.1
Pyranose oxidase [Modestobacter marinus]	32%	CCH88167.1
choline dehydrogenase [Modestobacter marinus]	32%	WP_041795256.1

Appendix C

```
   1 ATGGTGGAAG GACAATCCAG AAGTGCGTGG AAGGCTGTAT TGGGGGTGTG
  51 CGTGGGGCTG CTGGGAGGGC CCGCGGCAGC GCAGGTTCCG GAGCTGGAGG
 101 GAAGGTGGTC CCCTCCGCTG TCGTGGCCGC TGTCCGCCCA GCACATTCAC
 151 CTGCTGCCGG ATGGCAAGGT GATGTTCTTC GAGGACTTCG CGGGGGGCGG
 201 GCAGGCGCCG TACGTCTGGG AGCCCGCCAC GGACACGTTC GCGGCGCTGC
 251 CTCTGCCGCC CTTCAATGTC TTCGGGGCCG GCCACTCGTA CCTTTCCGAT
 301 GGCCGCCTGC TGCTCACGGG CGGGCATTCG GAGCCGCGCG TGGGCGAGGC
 351 CCGCGCGGCC ATCTTCAATC CGTACACGGG CGTCTGGGAG CCCGCGCCGG
 401 ACATGAATGA CAAGCGCCGG TACCCCACCA ACACCACGTT GCCGGATGGG
 451 GATGTGCTGG TGCTCTCCGG AGAGACGGTG GGTTCGGGGG TGACCAATGC
 501 GCTGCCCCAG CGGTGGGTCG ATGGCACCCA ATCGTGGCGG GACCTGTCCA
 551 CCGCCGGGCG CAAGCTGCCG CACTCACCCC GCATGTTCCT GGCGCCCAAC
 601 GGCAAGCTCT TCTTCGCCGG CGCCTGGCGC TCCAACCTGT GGTTGGACCC
 651 CGAAGGCACG GGGACCTGGT TCGGCTCCAC GCGCAGCCTC CATGGGGGCA
 701 GGGCGTACGG GGGCGCCGTC TACCTCGATG GAAAGGTGCT CCTCGTGGGC
 751 GGCGGTGATC CTCCCACGAA CACCGTGGAG CTCATCGATC TGAATCAGCC
 801 CTCGCCCACG TGGACGAGCC AGAGCCCCAT GCGCGTGGCC CGGCGGCACC
 851 ACAACACGAC GTTGTTGCCC GATGGCACGG TGCTCGTGAC GGGAGGAACC
 901 CAGTCGGGTG GGTTCGATGA CCGCGGGGGG GCCGTCTTCC ATGCCGAGAT
 951 CTGGGATCCT GAGACCAACA CCTGGCATTC GCTGGCCAGT GGGAGCGTCT
1001 ACCGCGGCTA CCACTCCACG GCCCTCCTGC TGCCAGACGG GCGCGTCCTG
1051 AGCGCGGGAG GAAACGGGGA GTCCTCCGCG GAGATCTTCG AGCCGCCTTA
1101 CCTCTTCAAG GGGCCTCGTC CCGCCGTCCA GGAGGCCCCG GACGAGCTTT
1151 TGCCGGGCAC GGTCTTCCCG GTGAGCACCC CGGACGGGTC CCAGATCAAG
1201 AAGGTGACCT TGCTGGCCCT GGGGTCCTCG ACGCATGCCT TCGATCAGAA
1251 CCAGCGGCTC CTCACGCTGC CGTACTCCGT GACGGACGAT GGGCTGCGCG
1301 TGAGCGCTCC CGAGAGCAAC GTTCTGGCCC CTCCTGGGCC GTACCTGCTC
1351 TTCCTGGTGA ACGAGGCCGG GGTTCCCTCG GTGGCGAAGA AGGTGCAAGT
1401 CGGCAAGGTG CCCTCCCGGT TCGCGAGCGT CATCGCCTTC AGCGACGTGT
1451 GGAAGTACGA CGATGGCAAC GTGGACCGCG GCACGTCCTG GCTCGCCAGC
1501 GATTACGATG ACTCGGCGTG GAAGTCGGGC CCGGGCCAGT TCGGCCATGG
1551 CGATGGTGAT GAGGCCACGG TGCTGAACGC CACCACGCCC ACCCAGCCCA
1601 CGGTCTACTT CCGCAAGAAG ATCACCTTGG ACAAGCCCGT TCTCGCCGCC
1651 CGGCTGGAGG CGCTCTTCGA TGACGGTGTG CAGGTGTGGA TCAACGGTGT
1701 GCCCGTCTAT TCCAAGAACG TCGGCAACGG GCTCGACTTT GGCAAATACG
1751 CCTCGGGCTC CACGAGCAAC GAGTACCGCC GCGAGAGCCT CTCGCTGGAC
1801 AGTGCCCCCT TCGTCGTGGG GGAGAACCTC GTCACGGCGC TGGTGAAGCA
1851 GGTGGGGCCC CGCTCGGACG ATCTCTCCTT TGCCCTGGCC TTGCAGGTGC
1901 AGTTCGAGAA CTTCCAGGAG GAGCACCAGC GGGCCATCGC CTTCGGCGAT
1951 CGCTGGGAGT ACGACGACAG CGGCGTGGAT CCGGGCCCTT CCTGGATGGC
2001 CCCGGCGTTC GACAGCGCGT CCTGGAAGGT GGGGGTGGGC CAGTTCGGCC
2051 ACGGCGATGG GGATGAGGTC ACCGTGCTGG CCTCCCCCCG CCCCGCCCGG
2101 CCGTCCACCT ACTTCCGCAA GAAAATCCAC CTCTCGGGGC CGGTGACTTC
2151 CGCCTCCTTG GAGCTCCTCT TCGAGGACGG TGTGGCCGTC TTCGTGAACG
2201 GCACCCAGGT GTTCGCGCGC AACGTGGGGC GCCTGGCGCA TTCCAAGTAC
2251 GCCACCGGCT CACTGGAGAA TGCGCGGGAG ACGGTGGAGC TGGTGCTCCA
2301 GCCCAATCCC TTCGTGAAGG GGGAGAACAC CGTCGCGGTG GTGGTGAAGC
2351 AGGTGGGGGC CACCTCGCCG GATCTCTCCT TTGATCTCGC GCTGGATGTC
2401 GGGGTTCAGG TCCCGGCGCC CTGA
```

Figure S1 – *S. aurantiaca* DW4/3-1 hypothetical GalOx coding gene after DNA sequencing.

```
   1 ATGCGTAAAT CTCCTTCCAG CCTCGTGGTT CAAGCGATTG CATTTTTGAC
  51 GGGCGTTTCG GCTTTTTCTG TGGGCCTTGG CACTGCTGTC GTGCAAAGCA
 101 GTCGTGCCCA GTCTGCAGAT CCGGATATAC TCGGCCATTG GAGCGCGACC
 151 AAAGATTGGG GTTTTGTCGC CATCCACGCC CACGTACTGC CCGACGGCAA
 201 GGTTCTCACC TGGGGTCGCG ATTGGGACGA GAATCAAGTC GATGGCAAGA
 251 CCACACAAGC GCGCGTTTGG GATCCGGTAA GCGACACTTT CGTCAACAAC
 301 AACATCCTCA ACAGCACCAC CAACGTCTTT TGCTCCGGGC ATGCCTTTTT
 351 GCCCGACGGG CGGTTGCTGG TCACCGGCGG CCATTTAGAT GACGACCGCG
 401 GCCTGGTGGA CACGACCATT TACGACCACA TCACCCAGGG TTGGACCAAA
 451 GTCCAGGACA TGATCGCCCG GCGGTGGTAT CCGACGACCA CCACTTTGGG
 501 CAACGGCGAG GTGCTCGTGA TCGCCGGAAC CTACGCGGGT TACCAGAACA
 551 TCAATCAGAT GCCCCAGATC TGGAAGACGA CGGGGGGGTG GCGGGATCTT
 601 GTCGACGCCC AAAAGCTGCC GGACGGCACC AACTCGCTGA AGTACGGCTA
 651 CTACCCGTAT ATGTTTGCCG CTCCCAACGG CCAGGTCTTC TACGCCGGTC
 701 CGGAGCCGGA CACCCGCTAC CTCGACACCA CCGGTACCGG CCGCTGGATC
 751 CCGGTGGCGC ACACGAACTT TAACGATACG CGCGACTACG GTTCGGCGGC
 801 GAGTTACGCT CCCGGCAAAG TACTCATCTC GGGGGGGGCC GGCGGCGACC
 851 TGTACGGCCC GCCCCCCACC GCCACCACCG AGATCATCGA TCTGAACGCC
 901 GCCTCCCCGC TCTGGCAGCA GGTCGAATCG ATGGCTTACC CGCGCCGCCA
 951 CCACAACCTC ACGGTGCTGC CGGACGGCAC GATCCTCGCC ACCGGCGGCA
1001 ACAGTTCGCC CGGCAGGTAC GAGGAGACCG CCCCCGCTCT GCCCGCCGAA
1051 CTGTGGGACC CGGCCACCCA AAGCTGGAGC ACCCTCGCAA GCATGCCCAC
1101 TCCCCGCATT TATCACTCGA TAGCGGCGCT ATTGCCGGAC GGCCGGGTAC
1151 TCTCGGCAGG CGGCGGCCAG GGTGGCGAAT CGGCCTACCG CCCGAGTGCA
1201 GAAATTTATT CGCCGCCCTA TCTGTTTCGT GGGCCGCGCC CCACAGTGAG
1251 CGCCGCGCCG ATTTCTGTGG GCTACGGCCA GGCGTTTACC GTCCAGAGCC
1301 CCGAGGCGGC CGATATCCGG CGCGTCACCT GGGTGCGTCT CTCCTCGGTG
1351 ACCCACGCCT TCAATGAAAA TCAGCGCTTC AACGAATTGA CCTTCACCCG
1401 CTCGGGCAAC ACCCTCACCG TCACCGCCCC GGCAAACGGC AATCTTGCCC
1451 CCCCCGGGCA CTACCTGCTC TATGTCCTCA ACGCCGACGG CGTCCCCTCG
1501 GTGGGGCGGG TGGTACGTGT CGGCACCCTG GCGCCTGCCG GGGATTTTAG
1551 CGGCGACCGC AAATCCGAAC TGGTCTGGCG CAACGGCGCC ACCGGTCAAA
1601 ACAGCCTCTG GCTGATGAAC AACACGACCT TCACCCAGAG TGTGGCCCTG
1651 CCCGCTGTCT CCGATCCCAA CTGGCTCATC ACCGGCAGCG GCGATTTTGA
1701 CAACGACGGC CGGCCGGACC TGCTCTGGCG CAATCGGGCC ACCGGCGAAA
1751 ACAACCTCTG GCTGATGAAC GGCACCGTCT ACCGCACCAA AGTCACCCTG
1801 CCGCTGGTGG GCACCGCCTG GCAGGTGGGT GGGGTGGGCG ATTTCGACGG
1851 CGACGGGCGC CCGGACATTT TCTGGCGCAA TCCGAGCACC GGCCAGGTGA
1901 GCGCCTGGCT GATGAACGGC ACCCAGTACC TCGTCTCGCG GGCCCTACCG
1951 TCGGTAATCG ACCCCAACTG GCAACTGGTG GCAGCAGGAC AGTTCAACAA
2001 CGACGCCCAG GCGGACCTGC TCTGGCGGCA CAAGCAGACC GGGGCCAACA
2051 GCGTCTGGCT GATGAGTGGC ACGGCACTGG CGCAGTCGGT AGCCTTACCG
2101 GCGGTCTCGA CCGCCTGGCA GATTGGGGCG GCGGGCGATT ACAACGGCGA
2151 CGGCAAAGCG GACATCGTCT GGCGCGACAC GACCACCGGC GCCAACGCCG
2201 TCTGGCTGAT GAACGGCACG GCTTTTGTCC AGTCGGTGGG GCTGTCGGCG
2251 GTCGCGGCCG GCTGGCAACT GAGCGGACCG CGCTAG
```

Figure S2 - *G. violaceus* PCC 7421 GalOx coding gene after DNA sequencing.

```
   1 ATGTCGAAGC ACCGCTGCAT CGCCTGGCCC TTGTTACGCA CCGTCTGGTG
  51 GCTGGGGCTC ATGGTGCTTC CCGTCCAAGC CCAGGCCCAG ACACCCGACC
 101 AGGTGGGCCG CTGGGCCAGC GTCATGAACT GGCCCATCTC CGCCACGCAC
 151 ATGGCCCTGC TGCCCGACGG CAAGGTGATG TTCTACGGCG AGTTCGACGA
 201 AGGCGCCCTG CCCCCGCGAC GCTGGGACCC CTCCACCGGC GCCCTCTCGT
 251 CCTTCCCCTA CGTGGGCTAC AACATCTTCT GCTCGGGGCA CTCCTTCCTC
 301 TCCAACGGCA AGCTGCTCGT CACCGGAGGC CACATCGCCC GCGACGTGGG
 351 CCTGCCCGAC ACGAGCTTCT TCGACTTCAA CACCACCAGC TGGACGCGCC
 401 TGCCGGACAT GAACGCCGGG CGCTGGTACC CCACCAACAC CACGCTCAAC
 451 AACGGCGACG TGGTGGTGAC CTCGGGCGAA ATCAACGGGG CGGGCGACAT
 501 CAACGAGATT CCCCAGCGCT TCATCGCCGG CACCAACTCC TGGCGCACGC
 551 TCACCAACGC GCGCAAGAAC GTGCCCTTCT ATCCCAAGAT GTTCCTGGCC
 601 CCGAACGGCC GGCTCTTCTA CGCGGGCTCC CTGCGCGCCT CGTTCTGGCT
 651 GGACCCCACC AGCAACGGCG CGTGGAGCAA CGGCCCTGTC AGCATCTTCG
 701 GCAGCCGCAG CTACGGCCCC GCCGTGTACA TCGACGGCAA GGTGCTCCTC
 751 ATCGGTGGAA GCGAGCCGCC CACCGCCACC GTCGAGCAGA TCGACCTCAC
 801 CGCGGCGAAC CCCACCTGGC AGTACGTGGC ACCCATGAGC ATCCGCCGGC
 851 GTCAGCACAA CGCGGTGCTG CTGCCCGACG CCACCGTCGT CGTCATCGGC
 901 GGAAGCAGCG GCAGCGGCTT CGACGACGCC AACGCGGCCG TGCGCCATGC
 951 GGAGGTCTAC AACCCCGCCA CCAACACGTG GACGTCCTGG GCCAGCAACG
1001 TGCGCTACCG GGGCTATCAC TCCACCGCCG TGCTCCTCCC GGACGGGCGG
1051 GTGCTGTCCG CGGGCGGCGC CAGCGAGCGC ACCGCGGAGG TGTTCTCCCC
1101 GCCCTACCTC TTCAAGGGCG CGCGCCCCGC CATCACCTCC GCGCCCACCG
1151 TGTCCCTGCC GGGGGCCCAG TTCACCATCA CCACGCCCGA CGCGGCCAAC
1201 ATCAGCCGCG TGAGCCTCAT CGCCCTCAAC TCGACGACAC ACACCTTCGA
1251 CATGAACCAG CGGTTCCTCA CGCTGAGCTT CACGCGCGGC GCCGGCAGCC
1301 TCAACGTCAC CGCGCCGCCC AACCGGAACA TGGCGCCACC GGGTTACTAT
1351 CAACTCTTCA TCGTCAACAA CGCGGGCGTG CCCTCCTACG GGCGGCGGCT
1401 GCGCATCCCC CCTCCCTGA
```

Figure S3 – *M. stipitatus* DSM 14675 hypothetical GalOx 2 coding gene after DNA sequencing.

```
   1 ATGAGCGGTC ACCGGTATCC CGCCGCAGTT GACGTCGCCA TCGTCGGCAG
  51 CGGTCCCACG GCTTCGGCCT ATGCGCGGAT CCTCAGCGAG GAAGCCCCCG
 101 GTGCCACGAT CGCGATGTTC GAAGTGGGCC CGACTGTCAG CAATCCGCCC
 151 GGCGCGCACG TCAAGAACAT CGAGGACCCT GATAGCCGCA GCCTCGCCCA
 201 GCGGGCGTCG GAGGGTCCCG GTGCCGGTGC TGCAACAGTG AATTCGCCGG
 251 GCGCCGTCAA GAGCGGCGAA CGCCGTGCGC GCCCTGGAAC TTACCTGCTG
 301 CAGGACGGCT ACGCCTTCCC GGGCGAGGAC GGCATGCCCG TCGCGGCCAT
 351 GTCCAGCAAC GTGGGCGGGA TGGCCGCCCA CTGGACCGCC GCCTGCCCCC
 401 GCCCGGGCGG CAAGGAACGC ATCCCGTTCC TGCCGGACCT GGAAGAGCTC
 451 CTTAACGACG CCGACCGCCT TTTGGGCGTC ACCACGCACG CTTTTGATGG
 501 TGCCCCGTTC TCGGACCTGG TCCGTGAACG ACTCGCCGCG GTCGTGGATC
 551 AGGGCCGCAC GCCTGCCTTT CGGGTCCAGC CCATGCCCCT TGCTGTACAC
 601 CGGCGGCAGG ATGGCGCCCT CGTATGGTCC GGCTCCGACG TCGTCATGGG
 651 GGAGGCCACC CGCGATAATC CGCAGTTTGA ACTGTTTGAT GAATCGCTGG
 701 TGACCCGTGT GCTGGTGGAG GACGGCACTG CTGCCGGCGT CGAAGTCCAG
 751 GACCGCCGCA GCGGTGACAC TTATCAGGTG GCGGCCCGTT ACGTTGTGGT
 801 GGGGGCGGAC GCCCTGCGCA CACCGCAGCT GCTGTGGGCG TCCGGGATCC
 851 GGCCCGACGC CCTGGGCCGC TACCTGAACG ACCAGGCGCA GGTGGTGTTC
 901 GCGAGCAGGC TCCGCGACGT CCAGCCCGAG GACGCGCCGG CAGCGGCCAA
 951 TGGTGCCCTC AGTGAGCAGA GCGGAGTGGC CTGGGTTCCC TACACGGACG
1001 AGGCGCCCTT CCACGGCCAG ATCATGCAGC TCGATGCTTC CCCGGTTCCC
1051 CTGGCCGATG ATGATCCCAT CGTCCCGGGC TCCATCGTGG GGCTGGGCCT
1101 GTTCTGCGCG AAAGACCTGC AGCGTGAGGA CCGGGTGGCG TTCGACGACG
1151 ATACCCGCGA CTCCTACGGC CTGCCCGCCA TGCGCATCCA CTACCGGCTG
```

```
1201 ACCGAGCGGG ACCACGTGGT ACTGGACCGG GCCAGGCAGG AAATTGTCCG
1251 TCTGGGCAAG GCGGTGGGCG AACCGCTGGA CGAGCGGCCC TTCGTCCTGC
1301 CGCCGGGTGC GTCGCTGCAC TACCAGGGCA CCACGCGGAT GGGTGAGACG
1351 GACGACGGCG AGAGCGTCTG TTCGCCGGAC AGCCAGGTGT GGCAGGTCCC
1401 CGGCCTCTTT GTGGCCGGCA ACGGCGTTAT CCCCACCGCT ACAGCATGCA
1451 ATCCCACCCT GACGTCGGTG GCGCTCGCCG TGCGCGGCGC CCGGAAAATC
1501 GCTGAAGAAA TCACCAGCTC TTTACTTATG TCCGAATCAG ACAATAGACT
1551 GTCTAAATAA
```

Figure S4 – *A. siccitolerans* POx coding gene from DNA sequencing.

```
   1 ATGGCCACCG CAAGCGACGA GGAAACGGCA GGTGAAAACG GCCGCGCCGC
  51 GAACGTCCTC GACGGCAATG CCGCCACCAT CTGGCACAGC AGGTATTCAC
 101 CGGCACCGGC GGCACCACTG CCGCATACCC TCACCATCGA CATGGGTGTC
 151 TCAAACCAGG TCTCGGGGCT GAGCTACCTC CCCCGCACAG ACCACATGAA
 201 CGGCCGGGTG GGCGCCTACG CCATCCACGC CAGCAGTGAC GGAACCATCT
 251 GGAGCGTGGT CGCCAGCGGC ACCTGGGCGG ACAACGCGGA CGAGAAAACC
 301 GCGGCCTTCA CTTCAACCAC GGTGCGCTAC ATCCGGTTGA CAGCGAGCAC
 351 CGAGGCAGGC AACCGCGGAC CGTGGAGCAG CGCCGCGGAA ATCAACATCC
 401 TTGGGGTCCA GCCAAAGCCC GAGTCGGCTT CCGGTCCCCT GCCCCGGGAC
 451 GGCTGGCTGG CCTCCGCGAG CGACCAGGAG ACAGCCAACG AAAACGGCCG
 501 CGCCGCGAAC GTGCTCGACG GCGACGCCGC CACCCTTTGG CACAGCAGGT
 551 ATTCGCCCGC TCCCGCGGCC CCGCTGCCGC ATACCCTGAC CATCGACATG
 601 GGTGTGGTCA ACCAGGTCGC GGGGCTGCGC TACCTCCCCC GCTTCGACAA
 651 TATGAACGGT CGGGTGGGCG GGTACTCAAT CCACGCCAGC TCGAACGGAA
 701 CCTCCTGGAA CCTCCTCGCG CGCGGCACCT GGGCCGACAA CGCGGACGAG
 751 AAGACTGTCA CTTTTGCGGC AGCTTCAGCG CGTTACATCC GGCTGACGGC
 801 GAGCACGGAA GCTGGCAACC GGGGACCGTG GAGCAGCGCA GCTGAAATCA
 851 ACCTGCTTGG AACACCACCG AAGGGGCCCG GAACATGGAG TCCCACGGTC
 901 AACTTCCCCC TCGTCCCGGC CGCCGCGGCA CTGCTGCCCG GCAACAGGCT
 951 GCTCACGTGG TCCGCCTACT CGCCTATTAC TTTTGGCGGC GAGACCGGCA
1001 TCACGCAGTC AGCCATCCTG GACCTGAATA CCGGGGCTGT CAGCCAGGCG
1051 GAAGTGGCCA ACACCGGCCA CGACATGTTC TGCCCCGGCA CATCGCTGTT
1101 GCCGGACGGC CGGATCCTGG TCAGCGGCGG CAGCAACAGC GAAAAAACCT
1151 CTCTGTTCAG TCCTGCAACG AACACGTGGG CCCCGGGGCC GGACATGAAT
1201 GTTGGCCGCG GCTACCAGTC CAACGTGACC ACTTCCACGG GGGAAGTGTT
1251 CACCCTCGGC GGTTCATGGT CGGGCGGACT GGGCAGCAAG CACGGCGAGA
1301 TCTGGAGCAG CACCGGCGGC TGGCGGCCGT TGCCTGATGT TCCGGTGGAC
1351 AGCATCCTGA CCGATGATCC CGGCGGCGAG TTCCGTTCCG ACAACCACGC
1401 CTGGCTGTTT TCCGCGGCCG GCGGACGGGT GTTCCATGCC GGCCCCAGCC
1451 GTGAGATGAA CTGGATCTCC ACGGCCGGCA CCGGCAGCGT CACGTCGGCC
1501 GGCACCCGGG CGGACAGCGC AGACGCGATG AACGGCAACG CGGTGATGTA
1551 CGACGTCGGA AAGATCCTCA CCATGGGCGG GGCGCCCGGC TACGACAACT
1601 CCGATGCCAC GGCGCGGGCC TACACCATCG ACATCAATAA CGGGGTTGAC
1651 GTTGCCCGGA CCTCGGACAT GGCGGTCAGC CGGTCCTTCG CCAACGGCGT
1701 CGCGCTGCCG GACGGCCAGG TGCTGGTGGT GGGCGGCCAG GCCCATGCCG
1751 TCCCGTTCAC GGACACAGGC GCCAGGATGG CCCCCGAACT CTGGAATCCG
1801 GCTACCGAGG AGTGGACCGC GATGGCACCG ATGGCGGTCC CCCGGACGTA
1851 TCACAGCGTG GCATTGCTCC TTGCCGACGG CCGGGTGTTT GTGGGCGGCG
1901 GCGGACTGTG CGGCACGTGC ACCACCAACC ACCTGGACGG CGAGATCTTC
1951 ACGCCGCCCT ACCTGCTGAA CGCGGACGGC AGCGCACGGA CCCGGCCAAC
2001 GATCGTTGAC GCCCCCGCCA CCGCTACGGC CGGCAGCAAG ATCTCCGTGA
2051 CCACCGGATC GAAGATCAGC AAGTTTTCCC TCATGCGGAT GTCCTCGGTC
2101 ACCCACACGG TAAACACCGA CCAGCGGAGA ATTCCGCTGA CGGCAACGGG
2151 GACCTACGGC AACAACACCG CAACGCTGAC GTTGCCGGCA GACCGCGGCG
2201 TGCTGGTACC GGGCGCCTAC ATGCTGTTCG CGATGGACGG CAACGGGGTG
2251 CCCAGCGTCG CGACGACGAT CCAGATCTCG TGA
```

Figure S5 – *A. siccitolerans* GalOx coding gene from DNA sequencing.

Appendix D

UniMS

Mass Spectrometry Unit

UniMS- Mass Spectrometry Unit

Tel. 351 – 214469468/744; http://www.itqb.unl.pt/facilities/UniMS/

Data measured by the UniMS

Publications that use data measured by the Mass Spectrometry Unit (UniMS) should include this information in the appropriate section. A typical sentence: "Data provided/obtained by the Mass Spectrometry Unit (UniMS), ITQB/iBET, Oeiras, Portugal".

Methodology:

The protein bands, excised from a 1D-PAGE gel, were destained, reduced, alkylated and digested with trypsin (Promega, 6.7 ng/µl) overnight at 37 ^{o}C. The tryptic peptides were desalted and concentrated using POROS R2 (Applied Biosystems) and eluted directly onto the MALDI plate using 0.6 µl of 5 mg/ml CHCA (alpha-cyano-4-hydroxycinnamic acid, Sigma) in 50 % (v/v) acetonitrile and 5 % (v/v) formic acid.

The data was acquired in positive reflector MS and MS/MS modes using a 4800plus MALDI-TOF/TOF (AB Sciex) mass spectrometer and using 4000 Series Explorer Software v.3.5.3 (Applied Biosystems). External calibration was performed using Pepmix1 (Laser BioLabs).

The twenty fifth most intense precursor ions from the MS spectra were selected for MS/MS analysis.

The raw MS and MS/MS data were analyzed using Protein Pilot Software v. 4.5 (ABSciex) with the Mascot search engine (MOWSE algorithm). The search parameters were as follows: monoisotopic peptide mass values were considered, maximum precursor mass tolerance (MS) of 50 ppm and a maximum fragment mass tolerance (MS/MS) of 0.3 Da. The searches were performed against UniProtKB protein database with taxonomic restriction to Bacteria (24161833 sequences; 7271637236 residues). A maximum of two missed cleavage was allowed. Carboxyamidomethylation of cysteines, oxidation of methionines and N-Pyro Glu of the N-terminal Q were set as variable modifications.

Protein identification was only accepted when significant protein homology scores were obtained ($p<0.05$, protein scores greater than 86) and at least one peptide was fragmented with a significant individual ion score ($p<0.05$).

Mass spectrometry result of recombinant hGalOxMs2 and GalOxGv

ASSAY REPORT: UNIMS226/14

Protein sequence coverage:

Band 1- hGalOxMs2

PROTEIN VIEW

Match to: TIG_ECO45 Score: 1100 Expect: 3.3e-105

Trigger factor OS=Escherichia coli O45:K1 (strain S88 / ExPEC) GN=tig PE=3 SV=1

Nominal mass (Mr): 47836; Calculated pI value: 4.83

NCBI BLAST search of TIG_ECO45 against nr

Unformatted sequence string for pasting into other applications

Taxonomy: Escherichia coli S88

Variable modifications: Carbamidomethyl (C),Oxidation (M)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

Sequence Coverage: 55%

Matched peptides shown in Bold Red

```
  1 MQVSVETTQG LGRRVTITIA ADSIETAVKS ELVNVAKKVR IDGFRKGKVP
 51 MNIVAQRYGA SVRQDVLGDL MSRNFIDAII KEKINPAGAP TYVPGEYKLG
101 EDFTYSVEFE VYPEVELQGL EAIEVEKPIV EVTDADVDGM LDTLRKQQAT
151 WKEKDGAVEA EDRVTIDFTG SVDGEEFEGG KASDFVLAMG QGRMIPGFED
201 GIKGHKAGEE FTIDVTFPEE YHAENLKGKA AKFAINLKKV EERELPELTA
251 EFIKRFGVED GSVEGLRAEV RKNMERELKS AIRNRVKSQA IEGLVKANDI
301 DVPAALIDSE IDVLRRQAAQ RFGGNEKQAL ELPRELFEEQ AKRRVVVGLL
351 LGEVIRTNEL KADEERVKGL IEEMASAYED PKEVIEFYSK NKELMDNMRN
401 VALEEQAVEA VLAKAKVTEK ETTFNELMN
```

Band 2- GalOxGv

PROTEIN VIEW

Match to: BLAT_ECOLX Score: 644 Expect: 1.3e-059

Beta-lactamase TEM OS=Escherichia coli GN=bla PE=1 SV=1

Nominal mass (Mr): 31495; Calculated pI value: 5.69

NCBI BLAST search of BLAT_ECOLX against nr

Unformatted sequence string for pasting into other applications

Taxonomy: Escherichia coli

Variable modifications: Carbamidomethyl (C),Oxidation (M)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

Sequence Coverage: 66%

Matched peptides shown in Bold Red

```
  1 MSIQHFRVAL IPFFAAFCLP VFAHPETLVK VKDAEDQLGA RVGYIELDLN
 51 SGKILESFRP EERFPMMSTF KVLLCGAVLS RVDAGQEQLG RRIHYSQNDL
101 VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLLTTI GGPKELTAFL
151 HNMGDHVTRL DRWEPELNEA IPNDERDTTM PAAMATTLRK LLTGELLTLA
201 SRQQLIDWME ADKVAGPLLR SALPAGWFIA DKSGAGERGS RGIIAALGPD
251 GKPSRIVVIY TTGSQATMDE RNRQIAEIGA SLIKHW
```

Mass spectrometry result of recombinant GalOxAs and POxAs

ASSAY REPORT: UNIMS56/15

Protein sequence coverage:

Sample VB_1: GalOxAs

PROTEIN VIEW

Match to: A0A024GZ97_9MICC Score: 2900 Expect: 2.4e-283

Galactose oxidase OS=Arthrobacter siccitolerans GN=GAOA PE=4 SV=1

Nominal mass (Mr): 78814; Calculated pI value: 5.33

NCBI BLAST search of A0A024GZ97_9MICC against nr

Unformatted sequence string for pasting into other applications

Variable modifications:

Carbamidomethyl (C),Deamidated (NQ),Gln->pyro-Glu (N-term Q),Oxidation (M)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

Sequence Coverage: 66%

Matched peptides shown in **Bold Red**

```
  1 MATASDEETA GENGRAANVL DGNAATIWHS RYSPAPAAPL PHTLTIDMGV
 51 SNQVSGLSYL PRTDHMNGRV GAYAIHASSD GTIWSVVASG TWADNADEKT
101 AAFTSTTVRY IRLTASTEAG NRGPWSSAAE INILGVQPKP ESASGPLPRD
151 GWLASASDQE TANENGRAAN VLDGDAATLW HSRYSPAPAA PLPHTLTIDM
201 GVVNQVAGLR YLPRFDNMNG RVGGYSIHAS SNGTSWNLLA RGTWADNADE
251 KTVTFAAASA RYIRLTASTE AGNRGPWSSA AEINLLGTPP KGPGTWSPTV
301 NFPLVPAAAA LLPGNRLLTW SAYSPITFGG ETGITQSAIL DLNTGAVSQA
351 EVANTGHDMF CPGTSLLPDG RILVSGGSNS EKTSLFSPAT NTWAPGPDMN
401 VGRGYQSNVT TSTGEVFTLG GSWSGGLGSK HGEIWSSTGG WRPLPDVPVD
451 SILTDDPGGE FRSDNHAWLF SAAGGRVFHA GPSREMNWIS TAGTGSVTSA
501 GTRADSADAM NGNAVMYDVG KILTMGGAPG YDNSDATARA YTIDINNGVD
551 VARTSDMAVS RSFANGVALP DGQVLVVGGQ AHAVPFTDTG ARMAPELWNP
601 ATEEWTAMAP MAVPRTYHSV ALLLADGRVF VGGGGLCGTC TTNHLDGEIF
651 TPPYLLNADG SARTRPTIVD APATATAGSK ISVTTGSKIS KFSLMRMSSV
701 THTVNTDQRR IPLTATGTYG NNTATLTLPA DRGVLVPGAY MLFAMDGNGV
751 PSVATTIQIS
```

Sample VB_2: POxAs

PROTEIN VIEW

Match to: A0A024H8G7_9MICC Score: 1190 Expect: 2.4e-112

GMC oxidoreductase family protein OS=Arthrobacter siccitolerans GN=ARTSIC4J27_4061 PE=4 SV=1

Nominal mass (Mr): 55212; Calculated pI value: 4.82

NCBI BLAST search of A0A024H8G7_9MICC against nr

Unformatted sequence string for pasting into other applications

Variable modifications: Carbamidomethyl (C),Deamidated (NQ),Gln->pyro-Glu (N-term Q),Oxidation (M)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

Sequence Coverage: 50%

Matched peptides shown in **Bold Red**

```
  1 MSGHRYPAAV DVAIVGSGPT ASAYARILSE EAPGATIAMF EVGPTVSNPP
 51 GAHVKNIEDP DSRSLAQRAS EGPGAGAATV NSPGAVKSGE RRARPGTYLL
101 QDGYAFPGED GMPVAAMSSN VGGMAAHWTA ACPRPGGKER IPFLPDLEEL
151 LNDADRLLGV TTHAFDGAPF SDLVRERLAA VVDQGRTPAF RVQPMPLAVH
201 RRQDGALVWS GSDVVMGEAT RDNPQFELFD ESLVTRVLVE DGTAAGVEVQ
251 DRRSGDTYQV AARYVVVGAD ALRTPQLLWA SGIRPDALGR YLNDQAQVVF
301 ASRLRDVQPE DAPAAANGAL SEQSGVAWVP YTDEAPFHGQ IMQLDASPVP
351 LADDDPIVPG SIVGLGLFCA KDLQREDRVA FDDDTRDSYG LPAMRIHYRL
401 TERDHVVLDR ARQEIVRLGK AVGEPLDERP FVLPPGASLH YQGTTRMGET
451 DDGESVCSPD SQVWQVPGLF VAGNGVIPTA TACNPTLTSV ALAVRGARKI
501 AEEITSSLLM SESDNRLSK
```