Overexpression, purification and characterization of diverse oxygenases from fungi

Zhuoxuan Yang

A Thesis

In

The Department

of

Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements

For the Degree of Master of Science at

Concordia University

Montreal, Quebec, Canada

August 2010

Your file Votre référence
ISBN: 978-0-494-71103-3
Our file Notre référence
ISBN: 978-0-494-71103-3

ABSTRACT

Overexpression, purification and characterization of diverse oxygenases from fungi

Zhuoxuan Yang

Bacteria and fungi are able to grow using diverse aromatic compounds as sole sources of energy and carbon *(1) (2)*. These abilities are of interest since many of these compounds are toxic to higher organisms, which cannot degrade them. While bacterial enzymes responsible for degradation of aromatics have been quite well studied, relatively little is known about their fungal counterparts. The aim of this thesis is to expand knowledge of aromatic compound degradation in fungi.

cDNA libraries from 15 diverse fungal species are available from the Concordia fungal genomics group and within this database there are many BLAST (Basic Local Alignment Search Tool) hits to genes encoding enzymes involved in aromatic degradation pathways. Using bioinformatics techniques, DNA sequences encoding putative mono- and dioxygenases that initiate aromatic compound degradation were identified. Genes encoding these enzymes were amplified and used to construct recombinant plasmids for overexpression of the corresponding proteins, which were then purified, characterized and compared to their bacterial counterparts.

A putative salicylate hydroxylase from *L. edodes*, the shiitake mushroom, and a putative phenol hydroxylase and a catechol-1,2-dioxygenase from *Gloeophyllum trabeum* were expressed in *E. coli*. The putative salicylate hydroxylase was purified in good yield using two chromatographic steps. The purified protein contained a 1:1 ratio of non-

covalently bound FAD (flavin adenine dinucleotide). Interestingly, salicylate was a non-substrate effector, stimulating production of hydrogen peroxide from NADPH and oxygen, rather than hydroxylated product formation. In contrast, 2-aminobenzoate was rapidly converted to 2,3-dihydroxbenzoate. Furthermore, spectroscopic probes of binding indicated that the modes of binding of salicylate and 2-aminobenzoate are quite different. Together, these results suggest that the enzyme is not a salicylate hydroxylase, and that the true substrate is an amino aromatic compound. The putative phenol hydroxylase from *G. trabeum* was also purified in good yield using two chromatographic steps and was found to contain 1 mol of FAD per mol of protein. Uncoupling assays and product analysis results suggest that the enzyme is a resorcinol-specific hydroxylase, while phenol is a non-substrate effector. The putative catechol-1,2-dioxygenase was purified using four chromatographic steps, however, no activity was observed under any condition.

In summary, potential functions were elucidated for two new fungal flavoprotein hydroxylases, whose activities were not accurately predicted using sequence comparisons. Additional work will be required to further examine the activities of the putative catechol 1,2-dioxygenase.

ACKNOWLEDGEMENTS

First and foremost, I thank my supervisor Dr. Justin Powlowski for his unending support, his constructive guidance, encouragement and advice he has provided throughout my time as his student.

My sincere gratitude goes out to Dr. Lena Sahlman, whom I owe special credit for giving me a start in research and for fostering my curiosity in science.

Appreciation goes to Dr. Ann M. English and Dr. Paul Joyce for useful suggestions and discussion.

Special thanks to Dr. Yu Lei for unconditional contributions to my research and constant support.

I also wish to thank to Dr. Mengwei Ye for her help with MS and HPLC.

A big thanks to my lab mates, Ekwe Enongene, Racha Cheikh-Ibrahim, Joseph Napoletano and Mayukh Choudhury, whose bright spirits and willingness to help each other out made working in the lab a pleasure.

Finally I want to thank my parents and my brother for their love and encouragement.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Aromatic compounds are common environmental contaminants in soil and ground water, because they are widely used as solvents and intermediates in chemical syntheses as well as in products such as pesticides and herbicides. Many of these synthetic compounds cause environmental pollution and human health problems as a result of their persistence, toxicity, or biotransformation into hazardous intermediates *(1)*. However, since microbes have evolved to degrade naturally-occurring aromatic compounds as sole sources of carbon and energy, they often have the capacity to fully or partially degrade man-made aromatics as well *(1) (3)*. Such activities have the potential to be used to remove aromatic pollutants from the environment, or to clean up chemical wastes. In contrast, the ability of humans and other higher organisms to degrade aromatic compounds other than amino acids is limited, and generally involves cytochrome P450-mediated oxygenation followed by excretion *(4)*.

In the past 20 years, studies have been performed on both aerobic and anaerobic treatment of aromatic pollutants by using pure or mixed cultures of microorganisms. Many microbial species that metabolize natural and synthetic organic compounds have been isolated, pathways have been elucidated *(3)* and their enzymes have been studied (e.g. *(5) (6) (7) (8)*). As a result of these types of studies, the metabolic strategies that are used by bacteria for degradation of aromatic compounds are quite well known.

Common features of microbial aromatic degradation pathways

The common aspect of all microbial aromatic mineralization pathways is the cleavage of the benzene ring. In anaerobic degradation pathways, the ring is first reduced and then opened *(9)*. In aerobic catabolic pathways, which will be discussed in more detail here, the ring is opened by the action of a ring-cleavage dioxygenase. As a prelude to ring cleavage, the ring is hydroxylated: the catabolism of phenol may be considered as an example (Figure 1). A hydroxyl group is added to phenol by an enzyme that uses NADH or NADPH to incorporate one atom of molecular oxygen into the product, while reducing the other atom to water. Enzymes that catalyze such reactions are called *monooxygenases (10)*. In the second step, the catechol ring is cleaved to an unsaturated aliphatic acid by the addition of both atoms of O_2 into the substrate. Enzymes that catalyze such reactions are called *dioxygenases* and they are metalloenzymes which require iron or manganese for activity *(11)(12)*.

As is shown in Figure 1, ring cleavage of catechol can occur at either of two positions: between the dihydroxylated carbons, or adjacent to them. Addition of oxygen between the dihydroxylated carbons is called *ortho* or intradiol cleavage, and is catalyzed by Fe^{3+} containing dioxygenases which produce *cis*, *cis*-muconic acid. Cleavage adjacent to the dihydroxylated carbons is catalyzed by *meta*-cleavage (or extradiol) dioxygenases, which contain Fe^{2+} and produce 2-hydroxymuconic semialdehyde.

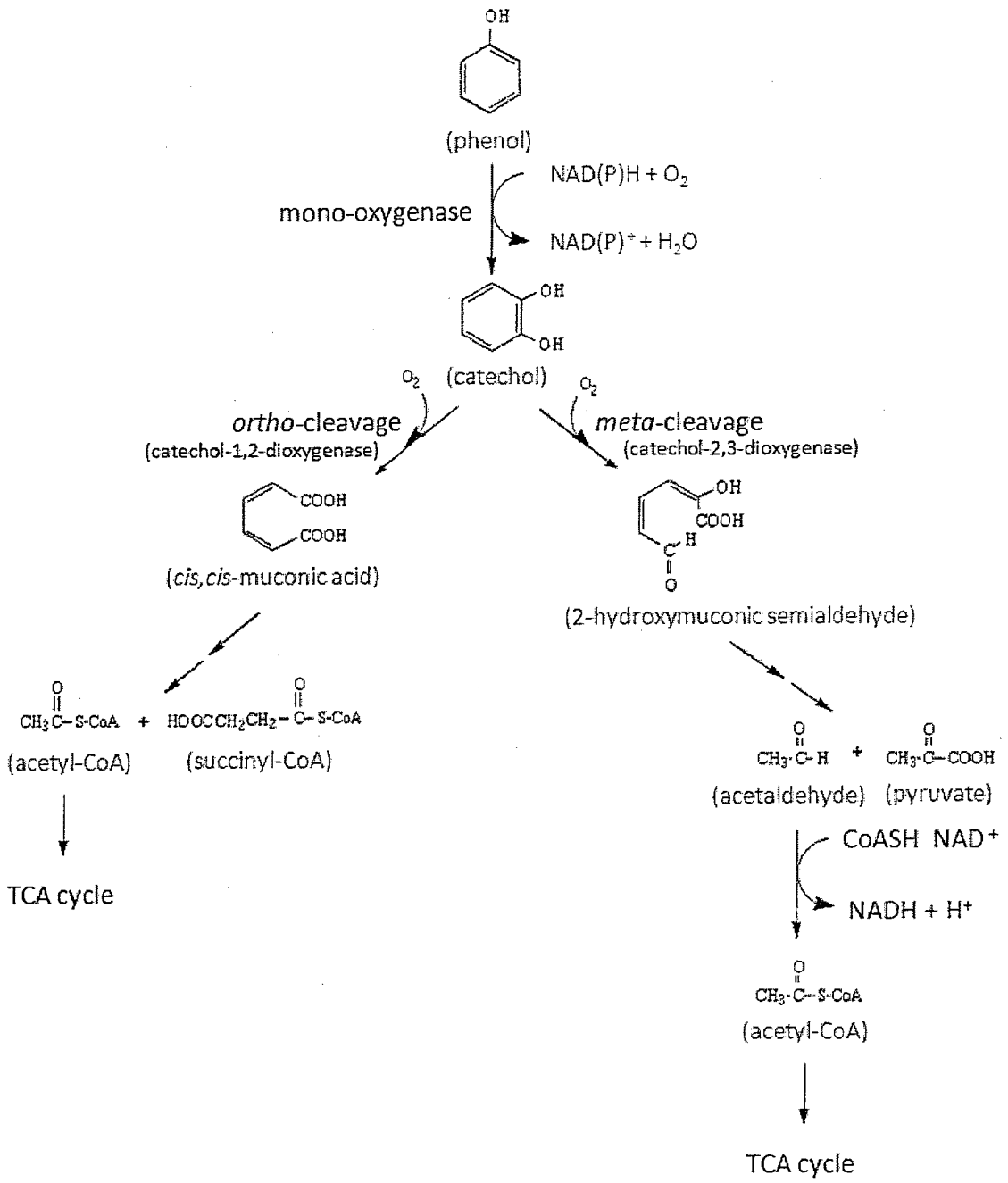


Figure 1: Biodegradation pathways of phenol

An example of an *ortho*-cleavage enzyme is catechol 1, 2-dioxygenase from *Acinetobacter app.*, while catechol 2, 3-dioxygenase from *Pseudomonas sp.* strain CF 600 is an example of a *meta*-cleavage enzyme *(13) (14).* While different species of

bacteria use *ortho*- or *meta*-cleavage pathways, thus far fungi able to completely degrade aromatic compounds have only been found to do so via *ortho*-cleavage pathways, and no fungal *meta*-cleavage enzyme has been reported.

Oxygenases involved in microbial aromatic degradation

In this thesis, the focus is on putative mono- and di-oxygenases involved in phenol and 2-hydroxybenzoate (salicylate) metabolism in two species of fungi. Although mammalian monooxygenases are, like cytochrome P-450, predominantly membrane-bound, heme-containing enzymes, microbial hydroxylases are commonly soluble proteins that contain flavin or non-heme iron prosthetic groups for the activation of O_2 *(30)*. Electron-donating co-substrates such as reduced nicotinamide adenine dinucleotide or tetrahydropterine cofactors are always required for monooxygenase-catalyzed hydroxylation reactions (see step 1, Figure 1), as well as for ring-hydroxylating non-heme iron dioxygenases such as biphenyl dioxygenase. Ring cleavage dioxygenases do not require additional electrons for the reactions that they catalyze.

Fungal aromatic degradation-growth on aromatic compounds

Although pathways and enzymes involved in bacterial degradation of aromatic compounds are the most extensively studied, similar pathways and enzymes have also been reported in fungi that can grow at the expense of aromatic compounds *(15)*. Fungi can thrive in challenging environments that are unfavourable for bacterial growth, and thus their degradative abilities may complement those of bacteria. Furthermore, it is important to know what the degradative capabilities of fungi and bacteria are as aromatic

compounds in the environment are degraded by microbial communites, rather than pure cultures.

Studies of oxygenases involved in benzoate, hydroxybenzoate and phenol degradation by fungi have been reported. Benzoate-4-hydroxylases have been purified from *Aspergillus niger (16)* and *Rhodotorula graminis (17)* and shown to require NADPH for activity. The *R. graminarus* enzyme was stimulated by FAD while the *A. niger* enzyme was not. Premkumar *et al.* reported that 3-hydroxybenzoic acid can be hydroxylated to 3, 4-dihydroxybenzoic acid by 3-hydroxybenzoate-4-monooxygenase from *Aspergillus niger*, the enzyme was found to be a flavoprotein requiring NAD(P)H for activity *(18)*. Salicylate (2-hydroxybenzoate) is commonly converted to 2, 3-dihydroxybenzoate in fungi such as *Aspergillus nidulans*, *A. niger*, and *Trichodermalignorum (19)*. However, Haribabu *et al.* noted that salicylate can be converted to 2, 4-dihydroxybenzoate in *Aspergillus niger* and to 2, 5-dihydroxybenzoate in *Trichosporon* sp *(19)*. Finally, Gaal & Neujahr *(21)* reported a NADPH-dependent flavoprotein, phenol hydroxylase from *Trichosporon cutaneum*, that could be expressed in the presence of phenol. The structure of this enzyme, determined using X-way crystallography, has been published *(20)*.

Some *ortho*-cleavage dioxygenases from fungi including enzymes catalyzing the cleavage of catechol 3, 4-dihydroxybenzoate and 1, 3, 4-trihydroxybenzene have also been studied,. The 3, 4-dihydroxybenzoate-3, 4-dioxygenase from *Rhodotorula mucilaginosa (17)* was reported to be able cleave the ring of 3, 4-dihydroxybenzoate to form 3-carboxy-*cis, cis*-muconate. This enzyme was found to be a non-heme, iron-containing protein. Varga and Neujahr purified a catechol-1, 2-dioxygenase from

Trichosporon cutaneum (21), which catalyses the ring-cleavage of catechol to form *cis,cis* muconate. Catechol ring cleavage activity has also been found in *Aspergillis niger*, *Penicillium spinulosum* and *Schizophyllum commune*. Karasevich et al. *(22)* reported the ring cleavage of 1, 3, 4-trihydroxybenzene by 1, 3, 4-trihydroxybenzene 3, 4-dioxygenase by *Candida tropicalis* grown on 4-hydroxybenzoate. Gaal and Neujahr *(23)* reported the growth of *Trichosporon cutaneum* on 1, 3-dihydroxybenzene, with the hydroxylation of the substrate to 1, 3, 4-trihydroxybenzene, which was subsequently cleaved by an oxygenase. The ring cleavage of 1, 3, 4-trihydroxybenzene was also reported in *Sporotrichum pulverulentum (22)*, *Rhodotorula rubra* and *Penicillium citrinum (24)*.

Some fungal enzymes catalyzing reactions subsequent to *ortho*-cleavage of the aromatic ring have also been described. 3-Carboxy-*cis, cis*-muconate lactonizing enzyme (CMLE) from *Neurospora crassa* catalyzes the reversible γ-lactonization of 3-carboxy-cis, cis-muconate by a syn-1, 2 addition-elimination reaction *(25)*. An X-ray structure of this enzyme has been determined *(26)*. The enzyme functions in the β-ketoadipate pathway, which is used in *N.crassa* and other fungi, yeast and bacteria for degradation of catechol and protocatachuate in aerobic bacteria, fungi, and yeasts *(27)*.

Although other enzyme activities of the β-ketoadipate pathway in fungi have been described, there is little information in the published literature on their characterization. Aside from the *T.cutaneum* phenol hydroxylase (EC #: 1.14.13.7) and *T.cutaneum* and *N. crassa* lactonizing enzyme (EC #: 5.5.1.5) sequences, which are directly linked to enzyme characterization, there is very little sequence information published for aromatic degaradation.

Fungal degradation-oxidation of aromatic compounds

In addition to enzymes involved in degrading aromatic compounds that are used as growth substrates, some fungi are able to oxidize aromatic compounds by virtue of secreted peroxidases and laccases associated with lignin degradation *(29)*. Lignin is a nonrepeating heteropolymer of aromatic acids found in plant cell walls. Although it is generally resistant to microbial attack, some white rot fungi and a small number of species of bacteria can metabolize lignin by cleavage of linkages between aromatic rings, and by ring cleavage to form CO_2. In addition to lignin, the ligninolytic systems of some white rot species have been shown to oxidize a very large variety of compounds. For instance, *Phanerochaete chrysosporium* can mineralize a variety of persistent environmental pollutants such as indane, benzo(a)-pyrene, DDT (1, 1-bis(4-chlorophenyl)-2,2,2-trichlorethane) and 2,3,7,8- TCDD (2,3,7,8-terachlorodibenzo-p-dioxin) *(28)*. It has been suggested that these lignin-degrading white rot fungi could be used for applications in environmental bioremediation such as the cleanup of toxic organic chemicals in soils and waters *(29)*.

Studies also have shown that some species of filamentous fungi hydroxylate complex polyaromatic compounds using monooxygenase enzymes belonging to the cytochrome P450 superfamily. *Cunninghamella elegans, C. bainieri, Mortierella isabellina,* and *Beauveria bassiana* have all been shown to hydroxylate one or more polyaromatic hydrocarbons (PAHs) *(30)*. Bezalel *et al.* suggested that both lignolytic and cytochrome P450 enzymes are involved in the degradation of PAHs *in vivo*. In this scenario, the cytochrome P450 enzymes perform the initial modification of the aromatic groups followed by degradation of the PAH derivative by the lignolytic system *(30)*.

Available fungal cDNA libraries

There are cDNA libraries of over 70000 genes from 15 diverse fungal species available from the Concordia fungal genomics project. The 15 species of fungi include *Lentinula edodes, Aspergillus niger* and *Gloeophyllum trabeum.* These cDNA sequences are available at the following website: https://fungalgenomics.concordia.ca. Sequences used in this thesis are from two species, *Lentinula edodes and Gloeophyllum trabeum.*

Lentinula edodes is commonly referred to as the shiitake mushroom. Initially, mycelia burrow through the dead tissue of various hardwoods such as oak, beech, or chestnut gradually utilizing the wood components to form fungal tissue and eventually produce basidiospores. Basidiospores are then dispersed by the wind, and eventually colonize another host *(31).* Due to the pleasant flavor and its nutritional as well as medicinal value *(32) (33), Lentinula edodes* has become the most popular edible mushroom. Consumption is thought to bestow anti-mutagenic, antibiotic and immunomodulatory effects *(34) (35). L. edodes* is also capable of degrading the major wood biopolymers, including cellulose, hemicellulose and lignin *(36)* and neutralizing environmentally persistent pesticide contaminants such as chlorophenols and dioxins *(37) (38) (39).*

Gloeophyllum trabeum can be found growing on the surface of dead trees, in particular on hardwoods in temperate North American forests *(40).* It is a brown-rot fungus, meaning that it degrades the cellulosic and hemicellulosic components of wood, leaving the lignin biopolymers oxidized but intact. The resulting darkening and deterioration of the wood substrate *(41),* gives rise to the term "brown rot". In order to

degrade wood lignocellulose, *Gloeophyllum trabeum* uses Fenton chemistry to produce extracellular hydroxyl radicals *(42)*. An extracellular low molecular weight peptide *(43)* and two intracelluar quinone reductases *(44)* are the most important components of this biodegradation system which can be used in bioremediating polyethylene oxide *(45)* and TNT. *Gloeophyllum trabeum* can also express various cellulose- and hemicellulose-degrading enzymes, such as endoglucanses *(46) (47)*, β-glucosidases *(46)*, xylanases *(47)*, and hemicellulases. When this fungus is grown in media with high carbon and low nitrogen content, several of these hydrolytic enzymes can reach their maximal expression level *(46)*. *Gloeophyllum trabeum* also can cause the mineralization of benzaldehyde catalysed by intracellular enzymes *(48)* and chlorophenols by manganese-independent peroxidase *(49)*.

Flavoprotein monooxygenases

Flavoprotein monooxygenases, such as the phenol hydroxylase shown in Figure 1, use NAD(P)H as the hydride donor to catalyze the hydroxylation of aromatic compounds in the presence of molecular oxygen. They are classified into two groups: single-component flavoprotein monooxygenases and two-component flavin-dependent monooxygenases *(50)*. Single component flavoprotein monooxygenases contain tightly-bound FAD, which, after reduction by NAD(P)H, activates O_2 for insertion into the aromatic substrate. In a two-component flavin-dependent monooxygenase system, a flavin reductase produces reduced flavin, which is delivered to the monooxygenase component to react with O_2 and hydroxylate the substrate *(50)*. These two component enzymes will not be considered further since they are outside the scope of this thesis.

Single component flavoprotein hydroxylases such as salicylate and phenol hydroxylases are considered in more detail below. Generally speaking, flavoprotein monooxygenase-catalyzed reactions involve three chemical steps *(50)*. First, reduction of the flavin by the hydride donor, NAD(P)H. Second, the formation of the oxygenating reagent, C4a-flavin hydroperoxide from the reaction of the reduced flavin and O_2. Finally, binding, orienting, and activating the substrate for its oxygenation by the C4a-hydroperoxide, and release of the second atom of oxygen as water from flavin C4a-hydroxide. For most of the single-component flavoprotein aromatic hydroxylases that have been characterized, reduction of the flavin is quite ineffective in the absence of substrate, because it is often a critical control point for catalysis by flavoprotein monooxygenases *(50)*. Thus, this feature can be used to prevent the wasteful use of NAD(P)H which would form reactive oxygen species such as H_2O_2 and waste reducing equivalents.

NADPH reduces these flavoprotein monooxygenases and the oxidized product NADP remains tightly bound. The reduced enzyme-bound flavin then reacts with oxygen to form a C4a- (hydro) peroxyflavin that is quite stable in the absence of substrates. Meanwhile, the bound NADP stabilizes the intermediate. As soon as substrate is present, reaction with the C4a-(hydro) peroxyflavin occurs to yield the oxygenated product *(50)*.

Salicylate hydroxylase

Salicylic acid is a phenolic phytohormone and is found in plants with roles in growth and development, photosynthesis, transpiration, ion uptake and transport *(51)*.

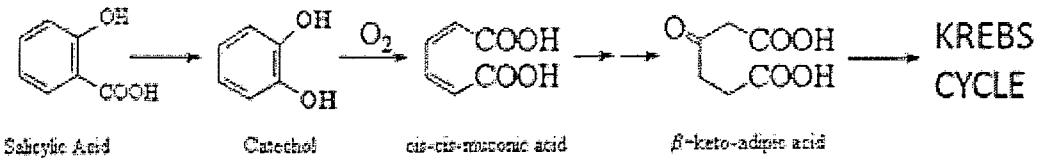


Figure 2: Biodegradation of salicylate

Several soil bacteria are able to utilize salicylic acid (salicylate) as the sole source of carbon and energy for growth. Microbial catabolism of salicylate is initiated by salicylate hydroxylase to form catechol which can be further oxidized by the *ortho* (intradiol) cleavage pathway (Figure 2) or *meta* (extradiol) cleavage pathway *(52)*. Bacterial salicylate hydroxylase is a flavin adenine dinuleotide (FAD) containing protein, which catalyzes the decarboxylative hydroxylation of salicylate to form catechol with stoichiometric consumption of molecular oxygen and NAD(P)H *(53)(54)*.

Salicylate hydroxylases have been isolated from many different bacterial sources and characterized. Sze and Dagley reported the purification of a salicylate hydroxylase from *T. cutaneum* grown with 4-hydroxybenzoate *(55)*. The enzyme contained FAD and was monomeric. Salicylate hydroxylase studied by Shozo *et al. (5)* was isolated from a strain of *Pseudomonas putida* and is a monomer with one FAD per 57,000 molecular weight. Salicylate hydroxylase studied by White-Stevens et.al *(7)*, was also isolated from a soil bacterium and contained two moles of FAD and two subunits per 91000 molecular weight.

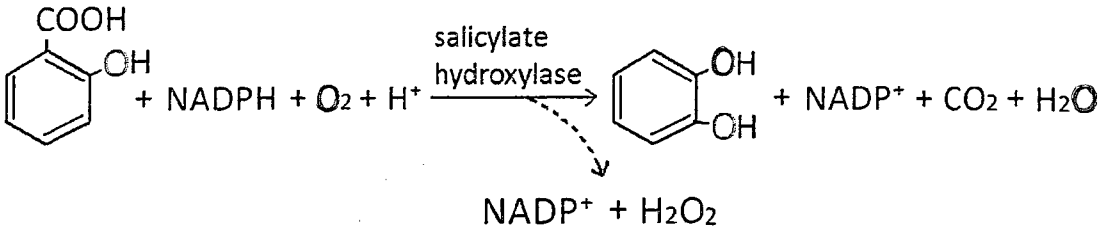


Figure 3: Reaction catabolised by salicylate hydroxylase

As shown in Figure 3, in the presence of NADPH one atom of molecular oxygen (which is indicated in red) is incorporated per molecule of the product, catechol, while the other atom is incorporated into water. Thus salicylate hydroxylase is a monooxygenase *(56)*. In the presence of the substrate salicylate, the enzyme catalyzes the aerobic oxidation of NAH(P)H resulting in the hydroxylation of salicylate. In the absence of salicylate, NAD(P)H oxidation can be also catalyzed by the enzyme to produce hydrogen peroxide. In the absence of oxygen, it promotes the reduction of FAD by NAD(P)H in the presence of salicylate *(54)*.

The mechanism of salicylate hydroxylase is shown in Figure 4. The hydroxylation of salicylate starts with the formation of a substrate-enzyme complex which contains apoenzyme, FAD and salicylate with the ratio of 1:1:1. Then NAD(P)H is oxidized with the formation of enzyme-substrate-bound $FADH_2$. Subsequently this complex reacts with molecular oxygen to form the product catechol, CO_2, and water *(53) (57)*. In the absence of salicylate, the holoenzyme is also stoichiometrically reduced by NAD(P)H to form enzyme-bound $FADH_2$, then this compound reacts with molecular oxygen stimulating the formation of hydrogen peroxide and enzyme bond FAD. Enzyme-bound $FADH_2$ can also react in the presence of salicylate to form the product catechol.

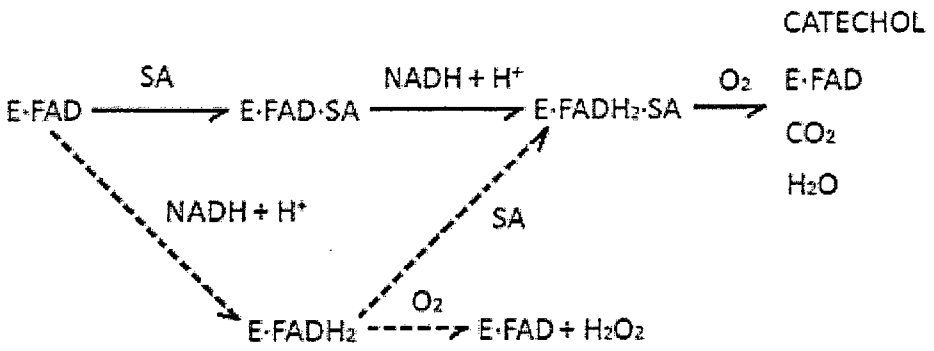


Figure 4: Mechanism of salicylate hydroxylase reaction

E and SA denote protein moiety of the enzyme and salicylate, respectively. Figure was taken from reference *(58)*.

The phenomenon of uncoupling has been well studied in salicylate hydroxylase. For the enzyme characterized by White Stevens *et al. (7)*, salicylate served as the perfect substrate, with no uncoupling of oxygen to form hydrogen peroxide. When salicylate was substituted by benzoate, oxygen was reduced in the same fashion as when salicylate was present, but H_2O_2 was formed stoichiometrically with NAD(P)H oxidized and benzoate unchanged. Both salicylate and benzoate thus facilitated NAD(P)H binding. Benzoate can competitively bind at the salicylate binding site, inhibiting salicylate binding and permitting NAD(P)H binding and oxidation. Thus benzoate is a "pseudosubstrate" that utilizes molecular oxygen to produce H_2O_2 without hydroxylation of the aromatic substrate, referred to as "uncoupling of oxygen activation from hydroxylation" *(7)*. Other salicylate analogues such as orcinol *(59)* and *p*-hydroxybenzoate *(60)* were also tested as substrates. Some of these compounds were hydroxylated, but some of the oxygen consumption was uncoupled, yielding hydrogen peroxide. These compounds were bound to the enzyme with a K_m higher than the

inducer, salicylate, and also facilitated NAD(P)H binding, but less effectively than did salicylate. Thus the potential for uncoupling of oxygen activation from hydroxylation to peroxide formation appears to be a general phenomenon for flavin hydroxylases *(7)*.

Phenol hydroxylase

Phenol and its derivatives are widely distributed environmental pollutants, since they are used in many industrial processes. A number of microorganisms have been found to degrade phenol and metabolic pathways of degradation have been delineated (e.g. *(6) (61) (8)*.)

Phenol hydroxylase is a monooxygenase that catalyzes the first step in phenol catabolism. As shown in Figure 1, in phenol degradation by aerobic bacteria, the aromatic ring is initially hydroxylated by a phenol hydroxylase at a position *ortho* to the pre-existing hydroxyl group to form catechol. This enzyme uses NAD(P)H and O_2 as co-substrates. The second step in phenol degradation in either *ortho* or *meta* cleavage pathway, is catalyzed by catechol 1, 2-dioxygenase or catechol 2, 3-dioxygenase, respectively. After *ortho* cleavage of the ring, enzymes convert the product, *cis, cis*-muconic acid, to succinyl-CoA and acetyl-CoA, whereas after *meta*-cleavage, the product, 2-hydroxymuconic semialdehyde, is cleaved to pyruvate and acetaldehyde, and acetaldehyde is subsequently converted to acetyl-CoA.

Phenol hydroxylase is composed of either one or multiple components. The phenol hydroxylase from the soil yeast *Trichosporon cutaneum (6)* and the bacterium *Pseudomonas pickettii* PKO1 *(62)* are single component flavoproteins, where the enzyme from *Bacillus thermoglucosidasius* A7 studied by Ulrike Kirchner et al. *(61)* is a two-

component enzyme, and the phenol hydroxylase from *Pseudomonas* CF600 studied by Shingler et al.*(63)*, is a multicomponent protein. The two component phenol hydroxylase from *Bacillus thermoglucosidasius* A7 consists of two homodimeric proteins: an oxygenase and a flavin reductase. Multicomponent phenol hydroxylases contain three components *(14) (63) (64)*: one is an oxygenase that binds substrate and oxygen and is responsible for hydroxylation, while another is a reductase that binds NAD(P)H and is responsible for electron transfer from NAD(P)H to the oxygenase. The former is generally an oligomeric protein, while the latter a monomeric iron-sulfur flavoprotein. An activator protein is the third component *(65)*.

Most relevant to this thesis is the phenol hydroxylase studied by Neujahr *et al*, which was purified from the soil yeast *Trichosporon cutaneum* grown on phenol or resorcinol as a major carbon source *(6)(66) (67)*. This enzyme contains one mole of FAD per 148,000 Da subunit. As shown in Figure 5, in the presence of NADPH, one atom of molecular oxygen (which is indicated in red) is incorporated into the product, catechol, while the other atom is incorporated into water. Thus phenol hydroxylase is a monooxygenase. As has been described for salicylate hydroxylase, some aromatic compounds are non-substrate effectors and stimulate phenol hydroxylase to form hydrogen peroxide from NADPH and oxygen. For example, resorcinol is a partially uncoupled non-substrate effector for the phenol hydroxylase studied by Neujahr *et al. (66)*

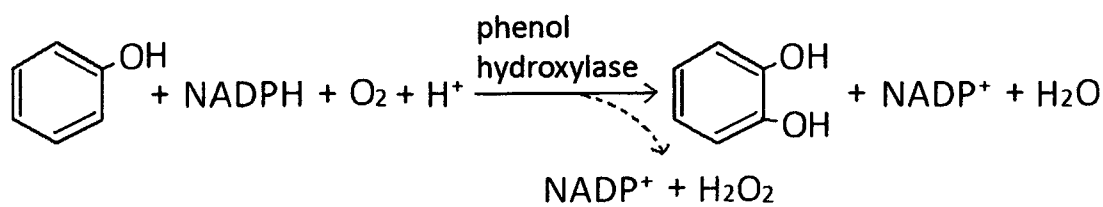


Figure 5 : Reaction catalyzed by phenol hydroxylase

Catechol-1, 2-dioxygenase

The second step in the *ortho*-cleavage pathway of phenol is catalyzed by catechol 1, 2-dioxygenase. Catechol 1, 2-dioxygenases have been purified from diverse microbial sources and are nonheme, Fe^{3+}-containing enzymes. (e.g. *(21)(68)(69)*). Also, a number of crystal structures of catechol 1,2-dioxygenases that has been solved *(70)*. The product of catechol 1, 2-dioxygenase is *cis, cis*-muconic acid, in which the aromatic ring is split between the two carbon atoms bearing hydroxyl groups with the incorporation of two atoms of molecular oxygen into the substrate *(69) (71)* (Figure 6).

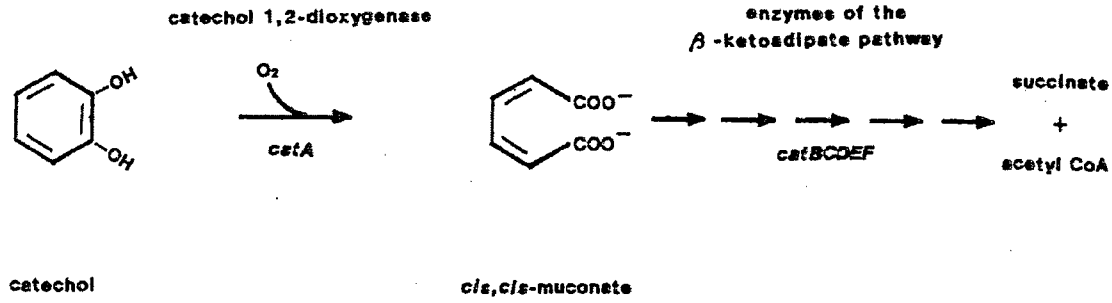


Figure 6: Reaction catalyzed by catechol-1, 2-dioxygenase

Figure was modified from reference *(69)*.

Kojima et al *(72)* purified the catechol 1, 2-dioxygenase from a cell-free extract of benzoate-grown *Pseudomonas arvilla C-1* and reported the spectral properties associated with its catalytic activity. The purified enzyme preparation of catechol-1, 2-dioxygenase from *A. calcoaceticus* had a red color with broad absorption between 390 and 650 nm. The enzyme contained 2 atoms of iron per 80000 Da of enzyme protein *(69)*. Varga and Neujahr reported purification a catechol 1, 2-dioxygenase with a molecular weight of 109 kDa from the soil yeast, *Trichosporon cutaneum (21)*. Both di- and trivalent iron are

present in the enzyme. Trivalent rather than divalent iron is the active agent in the oxygenation reaction. The *T. cutaneum* enzyme had a pH optimum similar to its bacterial counterparts as well as similar resistance to oxidizing agents and sensitivity to reducing agents. However, the specific activity of this enzyme was higher than that of the enzymes from other sources *(21)*. Significant differences are also found in the absorption spectra of the purified enzymes. There is a single peak in the ultraviolet region at 280 nm, a shoulder at 330 nm, and a rather pronounced peak at 560 nm. While the ultraviolet region shows a great resemblance to the spectra of catechol 1,2-dioxygenases isolated from other microorganisms the visible region is rather different. Instead of the broad maximum between 400-600 nm, which is usually ascribed to enzyme-bound ferric iron, the enzyme purified from *T. cutaneum* has a relatively sharp maximum at 560 nm. This type of spectrum gives association to a charge transfer complex between ferric iron and catechol, instead. Similar to other catechol 1,2-dioxygenases isolated from other microorganisms, the characteristic red colour of the active enzyme disappeared upon inactivation (by iron chelating agents and reducing agents) *(21)*.

Hydroxyquinol-1, 2-dioxygenase

Hydroxyquinol 1,2-dioxygenase (1,2-HQD), catalyzes the ring cleavage of hydroxyquinol (1,2,4-trihydroxybenzene), which is an intermediate in the degradation of various aromatic compounds (including some polychloro- and nitro- aromatic pollutants) to 3-hydroxy-cis,cis-muconates *(73)*. Hydroxyquinol 1, 2-dioxygenase is a member of the aromatic dioxygenase family, a family of mononuclear non-heme intradiol-cleaving enzymes *(73)*.

Sze and Dagley reported the purification of a hydroxyquinol 1, 2-dioxygenase from *T. cutaneum (55)* grown with 4-hydroxybenzoate. It was shown to be a red ferric iron-containing enzyme, and it was specific for hydroxyquinol. Catechol and pyrogallol were oxidized at less than 1% of the rate for hydroxyquinol. The enzyme was reported to be a dimer with two subunits having molecular weights of 39600 and 38200.

Latus *et al. (74)* studied a hydroxyquinol 1, 2-dioxygenase purified from the soil bacterium *Azotobacter* sp. strain GP1 grown with 2, 4, 6-trichlorophenol as the sole source of carbon. The enzyme was highly specific for 6-chlorohydroxyquinol (6-chloro-1, 2, 4-trihydroxybenzene) and hydroxyquinol (1,2,4-trihydroxybenzene), and was found to perform *ortho* cleavage of the hydroxyquinol compounds, yielding chloromaleylacetate and maleylacetate, respectively. The addition of Fe^{2+} ions significantly activated enzyme activity. The protein was determined to be a dimer with a molecular weight of 58000, and the protein can be inhibited by metal-chelating agents.

The goal of my project

cDNA libraries of over 70,000 genes from 15 diverse fungal species are available from the Concordia fungal genomics project and within this database, https://fungalgenomics.concordia.ca, there are many BLAST hits to enzymes involved in aromatic degradation pathways. Studies of degradation of aromatic compounds by fungi has to a large extent focused on oxidation of aromatics by lignin and other peroxidases, but as discussed above fungi do have the ability to degrade aromatic compounds as sole sources of carbon and energy, much as bacteria do. In addition to the species and growth substrates described in the previous sections, other species of fungi are also known to

have the ability to degrade benzoate and hydroxybenzoates using *ortho* -cleavage pathways channeled through 3,4-dihydroxybenzoate *(75)*. Many of the studies of fungal aromatic degradation pathways and their enzymes were done prior to the ready availability of gene sequences and gene sequencing technology, so there is very little information about them that relates structure to function. By contrast, much is known about structure-function relationships in bacterial enzymes involved in aromatic degradation and, in fact, this information appears to have been applied extensively to the annotation of fungal genes. The availability of this cDNA library therefore presents an opportunity to obtain experimental support for structure-function relationships in fungal enzymes involved in aromatic degradation pathways.

In this project fungal genes encoding proteins with predicted similarity to bacterial monooxygenases and dioxygenases involved in aromatic degradation pathways were selected for further study. The goals of this project were to amplify these genes, construct recombinant plasmids, overexpress and purifiy the corresponding proteins, and characterize them. It was expected that this information would expand our knowledge of fungal enzymes involved in aromatic degradation pathways, and the genes that encode them.

MATERIALS AND METHODS

Materials

All chemicals were reagent grade or of the highest available purity. Oligonucleotide primers for PCR were synthesized by BioCorp Inc. (Montreal, Quebec, Canada). *Pfu* DNA polymerase for PCR was purchased from Fermentas. Recombinant plasmids were constructed using the pETBlue-1 vector *(76)*. For expression of recombinant proteins, *E. coli* (DE3) *pLac- I* cells were used *(76)*. Restriction enzymes were purchased from Fermentas, Promega, Roche or MBI. NADPH and FAD were purchased from Roche Science and Boehringer Mannheim GmbH, and made up freshly in 2mM Tris·Cl buffer and distilled water, respectively. The concentration of NADPH and FAD stock solutions were calculated using $\varepsilon_{340} = 6220\ M^{-1}\ cm^{-1}$ and $\varepsilon_{450} = 11300 M^{-1}\ cm^{-1}$.

Target sequences

*Ledo*1986, *Gtra*1516 and *Gtra*1270 represent the fungal genes used in this study and were obtained from the Concordia fungal genomics group (https://fungalgenomics.concordia.ca). The following sequences were translated into amino acid sequences and used for bioinformatics studies.

ATTAGCACTTTCTCTCATACATTGGACCTCTACTTTTTCGAAATGGTATCCTCAACAAACCCCAACTTCAAACAACT
CCGCGTGGCAGTAATTGGTGGAGGTCTTGGAGGTCTCTCTGCTGCCGTCGCTTTGCGTCGTGCTGGTCATCTTGTTG
AAATTTACGAACGTCGCGACTTCAACGTAGAAGTCGGTGCCTCCATCTCTTGTGCCGCCAATGGAACCCAGTGGCTG
CGAGAATGGGAAGTAGATATTCCCGATATGAAACCTGTTATCCTTATGAAGCTAGTAATGCGCGACTGGGAGACTGG
AAGGATTCTGAACCAATACAACTTGGATAAATATGAGGAGGAATGGGGAAATGTCTACAACATGTTACATAGACAAG
ACATGCATGCTACTTTGCTCAAGACCGCGACTTCTCCAGAAGGCAAAGGAACGCCATGTATAGTGAAAATCGACCAC
ATCTGCGAGACAGTCGATTCAGAAGCGGGAACCGTCACGTTCAAAAATGGTGTTACAGTCAAAGCGGACATTATCAT
TGGTGCAGATGGTATCCGGTCTGTTGTTCGAGGCCAAATCGGGGTCGTACCGGACATGAAGTCGGCTCCGCAGACGT
GCTATCGCCTTAACGTCAAAAAGTCAGTTGTTGACGAACTAGGGCTTGTGAAGTACTCTTATGAGCCAGCTATTCAA
TTCTGGGGTGGTCTTCAAGGGAAAAATGGTCGATCCAAATATTACAAGATTGTGATGTCGCCGTGTTCTGATGGAGA
GATCGTTTCCTTCTATTGCTTTATGCCAACTGAATTGACAAAGCACCACGAAGAGGGCTTTACCTTCGCGGAAGTGC
CAGTATCAGACGTTATTGCCGGACGGTACGACGAACTAGACCCAGACTGCGTGAACCTGTTGAAAAATTCGGTCGAT
CGGATGCCTTGGCGTCTGTACGTTCACCAACCGTACGACCGGTGGTTCGCGGGCAACACTTGCATTCTTGGTGATGC
TGCTCATCCCATGATGCCCCACCAGTCTCAAGGCGCGTGTCAAGCTATCGAAGATGCCGCCGCTCTAGGTATCATCT
TCTCCGACAAGTACAACTTTACCACAGATGTTACCGCCGGATTGGCTATGTATCAGACAATTCGAAAGCCCAGGGCA
ACCAGAGTACAGTCTGCATCTGCGCGCGCAACAGAGAATTTGAACGAGAGGATCGGTTTTACTAGCTTGACACCTCA
TGACGCGTCCTTAGCTGCTGCTGAGGGAAAATTGACCATTAATGAGATGAACAGCTATAAAATGCACAATCATATCG
CGACCGAGGTGGACAAGTTAGGAGGTGGAATGGGCCCTACTATTGCAGTGATGTGAACGTAAATACAGGACGTCTAG
AGGCAAAGATGGTTGGTTTTGACTTTTCTGGGTGCTGGTGATGATTTATTACGATATTCCGCGCTGATTGTATTTTC
TTGTTCCCACTCCTTTCTCCGAGTTCTTTGGTCGGATTATATGTAATAAAAGTGAAGCACGGTCCCGTGACACCATT
TCTCTCAAAAAAAAAAAAAAAA

Figure 7: DNA sequence of *Ledo*1986 (obtained from https://fungalgenomics.concordia.ca)

The predicted start codon is indicated in red and the highlighted sequences were used as primers for PCR amplification

GTTGAACTCTCAGCCCGCCCAACATGCCTGTACCCGCCATGAAGGAATCCGACGTGGATGTGCTCGTTATCGGTGCG
GGACCTGCGGGCTTGATGTGTGCGCAGGGGCTGGCCAGGGCAGGTGTAAATGTCCGGATTGTGGACAAGAGACCAGG
CAAGGTCGCGGCGGGACAGGCTGATGGCATCCAACCACGCACAATTGAAGTGCTACAGAGCTATGGACTCGCCGAGC
GACTATTGAAGGAAGGCAACCAAATGCACATGGCAGCATTCTATGACCCCAGCCCCGAAGGAGGAATCCACCGTACA
GGACGAACACCGGACATCAATGCCCCTACGGCCCGATTCCCGTTCGAGGTCACCCTCCATCAGGGCGCGATCGAGTC
CATATTCTTAGACTGCCTCCGCTCGATGGGGCACGAAGTCGAGCGTCCAGTGGTCCCGGAATCGCTTGAGATCTCAG
ATAACCGAGACGCCTTGAAGGATCCGCAAGCGCGCGCTGTAAAGGTTGTACTGAAGCACGTGGACGCACCCGAGGGG
AAAGATACCGAAGTCGTTCACGCGAAATACGTTGTCGGCGCAGACGGCGCCCATTCATGGGTGAGGAAACAGCTCGG
CTTTGCCATGGAGGGCGAACAAACAGATTATGTTTGGGGTGTGGTCGACATGATACCCGACACAGACTTCCCGGACA
TCCGCAACCGGTGCGCGATCCATTCGAACAACGGCTCATGCATGGTTATACCGCGCGAAGGGGATGTTGTGAGGTTG
TATATCCAGCTTACGGACGAGGACGTCCGAGACGTGACGACGGGTCGCGTTGACACGCAGAAGTGCAGCGCTGAAAA
ATTGCTGGAGGTTGCAAAGAAGTCTTTCCATCCCTATCGCATCAAGGCAAAGGGCGATATACTGTGGTGGACGATAT
ACATCATTGGTCAGCGCGTGGCTTCAAAGTTCTCAGCGCACGAACGCGTATTCATCGCTGGTGATGCCTGTCACACG
CATTCTCCCAAAGCCGGACAGGGGATGAACGCCAGCATGAACGATACCCACAACCTGATCTGGAAGTTGACGCATGT
GCTTCGAGGATGGGCGGATATTTCGCTGCTAAAGACTTACGAGTTGGAGCGCCGAAAGTACGCTCAGGACCTGATCG
CGTTCGACAAAGAGTTCGCCTCACTGTTTTCCAAGAAGCCGAAGAGCGAAGAGAACCCCGACGGCGTCACACAAGAG
GAATTCGTAGAGGCCTTCCGTACGTTCGGTGGCTTTACGAGCGGCATCGGGATCCACTACGCTCCGTCCACGATCGT
CGACGCAGCCCACCAGGCCTCCGCCTCCAAGCTCATCATCGGCCAGCGTGTGCTCCCGCAGACCGTCATCCGCGCAG
CCGACGCGCGCCCGTACGAGCTGCAGGACCTCCTTCCCGCCGATACCCGCTTCAAGCTCCTCGTCTTCACGGGCGAC
ATCGGCGCGCCGGAGCAGAAGCGGAAAGTCGACGAACTTGCGAAGCACCTGGAGCGGAAGGAGAGCTTCTTGGGCAG
GTTTGGGGAGCAGAAGAATGATGTGTTCGATGTGTTTGCTGTCTGTCTCTCGAGAAAAGAGGACGTGGTGTATACTG
ACGTGCCCGAGGTATTCCGGCCCCACTGGTCCAAGGTTCTCCTCGACGACGTCGATACCACGGGCAAGGTCGGCGCG
GGCGTGTACGAGAAGTTTGGTATTGGCTCCGAGGGCGCTATTGTCGTAGTCCGGCCGGACGGGTATGTCGGGATTGT
TGCGCCTTTGGACGACGTTTCCGTTCTGGAGAGTTACTTTGCTGGGTTTTTCGCGAAGTCATGACCTGTTGCCCTTG
TGATATCTGAGTCGAGATCTTTTGCAGTTTGTGTTTAATAGCAGGTC

Figure 8: DNA sequence of *Gtra*1516 (obtained from https://fungalgenomics.concordia.ca)

The predicted start codon is indicated in red and the highlighted sequences were used as primers for PCR amplification

CGACGACAAGAAGCCCCTTCAAGCCAAGGCAGTGTCGAACGGCCTTCCGCCTATGCCTAACCTTGACTTGCCGTACC
CGGACAGACCAGAGCTCATTACGGAGAACCTTCTGAAGTTGACCAACTTGATTACCGATGAACGAAAGAAGTACATC
TTCAAGAATTTAATTACGCATATACACCAATTCATCAATGAGACAAGTATCACCACGGACGAATGGATGAACACCAT
CCAATTTCTGACCCGGACAGGTCAAATCTGCACGCCCATTCGTCAAGAGTTCATCCTGCTCTCCGACGTCCTCGGCA
TCTCCGCGCTCGTAGACGCACTTAACAACCCCCCAGTCAGCGGCGGGACGGAGAGCAGCGTGCTCGGGCCTTTCTTC
ACTGAGGACGCTCCCGACGTCAACAACGGCGATTCTATTGCGTCCGAGGGCAAGGGGCAGTACATGTACGTCGAGGG
GCGTGTAATCGATACGCACGGGAAGCCAGTCCCGAACGCTCTCATCGAAACGTGGGAGACAGACGAGTATGGCTTCT
ACGACACGCAATATGCGGACAGAAGCAAACCCGACTGCAGGGGCCGCTTGCGGACGGACAAGGACGGTAGATACGGC
TACCGCGCTGTCGTTCCAGTAGCGTACCCCATCCCCGGTGACGGCCCCGTCGGAGACTTACTGCTCATGCTCAACCG
ACACAATATGCGCCCGAACCATCTGCATATGATGATCGAGGCACCGGGCTACCAGAAACTCACGACCGCGTTCTACC
CCGAGGGCGATGAGTGGCTTGCGAGCGACGCTGTTTTCGGCGTCAAGAAATCCCTGGTCGTCACATTGAAGGATGTG
GACAATGAGCAGGAGGCGCGCAAGCGTGGGTTCCCCAAGGGGAGCCACTTCAAGTTGCTTGAGCATGATCTCGTGTT
GGTTCCTGAGGCTGAGTCCAAGGCGGCGCGGGAACAGTACGCGAGGGAACATGCTGTTAACAGGAGTAACGAGATTC
AGGCATGAAGTATGATTTGCGTCTTGTATGATGGGCGGACGTGCCAGTGCAGGCTGTATTATGATAGTTCAACCTGT
CAGGAATGTATTTCTATTGCACATCCAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAC

Figure 9: DNA sequence of *Gtra*1270

Inserted start codon is indicated in red and the highlighted sequences were used as primers for PCR amplification (*Gtra*1270B). Underlined sequences are the primers used for *Gtra*1270A

Recombinant plasmid construction and restriction digestion

PCR was performed using standard protocols *(77)*, and Table 1 shows the parameters used for amplification of each of the genes under study.

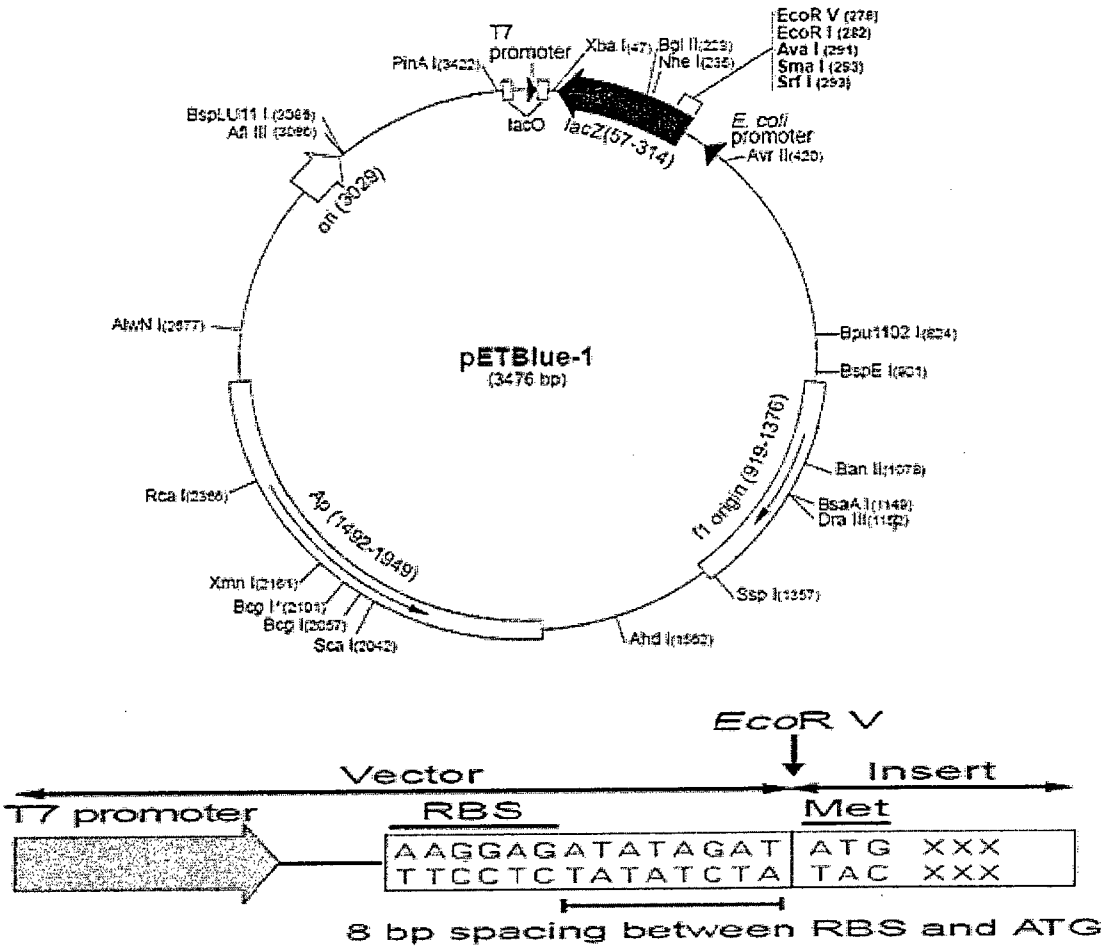
Step	Temperature (°C)	Time (min)	Number of cycles
Initial Denaturation	94	2	1
Denaturation	94	0.5	25
Annealing	41 (for *Ledo*1986) 51 (for *Gtra*1516) 37 (for *Gtra*1270A) 40 (for *Gtra*1270B)	0.5	
Extention	72	2.8 (for *Ledo*1986) 3.7 (for *Gtra*1516) 1.7 (for *Gtra*1270A) 2 (for *Gtra*1270B)	
Final Extention	72	7	1

Table 1: PCR settings for *Ledo*1986, *Gtra*1516, *Gtra*1270A and *Gtra*1270B

Table 2 shows the primers used for amplification of the target genes. Amplified fragments were subsequently ligated into the *Eco*RV site of the expression vector pETBlue-1 (Figure 10) which was supplied by the manufacturer in the open form. Thus, the salicylate hydroxylase, phenol hydroxylase or catechol dioxygenase (hydroxyquinol dioxygenase) coding region was inserted downstream from the T7 promoter (as shown in Figure 10). Recombinant plasmids then were used for transformation of the *E. coli* (DE3) *pLac- I* strain. Transformants were plated on LB agar plates containing 50 μg/ml carbenicillin and grown overnight at 37 °C; selected individual transformants were grown in liquid LB medium with the same concentration of carbenicillin. Plasmid DNA was purified using a Wizard Miniprep Kit (Promega) and screened for correct insert orientation by restriction enzyme digestion. Recombinant plasmid DNA was sent to Bio S&T (Montreal, Quebec, Canada) for sequencing of the insert region. pETBlueUP primer, TCACGACGTTGTAAAACGAC and pETBlueDOWN primer, CAATTTAACGATTGCGTCAG were used for DNA sequencing.

Target Gene	**Primer Sequence**	**Target Gene Size**	**Protein Size**
***Ledo*1986**	5'- ATGGTATCCTCAACAAACC-3' 5'- TCATCACCAGCACCCAG-3'	1388 bp	48600 Da
***Gtra*1516**	5'-ATGCCTGTACCCGCCATGAAGG-3' 5'-CAAGGGCAACAGGTCATGACTTCG-3'	1895 bp	66600 Da
***Gtra*1270A**	5'- ATGAACACCATCCAATTTCTG-3' 5'- CATAATACAGCCTGCACTGG-3'	844 bp	29100 Da
***Gtra*1270B**	5'- ATGCCTAACCTTGACTTG-3' 5'- CCATCATACAAGACGCAA-3'	983 bp	35900 Da

Table 2: Primers used for construction of target genes



TB258 pETBlue-1 Vector Map, Novagen

Figure 10: Expression vector pETBlue-1 map

Target gene is inserted into *Eco*RV restriction site in the multiple cloning site.

Agarose gel electrophoresis

A 0.7 % agarose gel was used to separate and analyze DNA fragments. The GeneRuler^{TM} 1 kb DNA Ladder Plus was purchased from Fermentas. DNA was visualised in the gel by addition of ethidium bromide.

Bacterial growth and protein expression

E.coli strain *(DE3) pLacI* was employed as the expression host, which carries a chromosomal copy of T7 RNA polymerase under the control of the *lacUV5* promoter and *lac* repressor from the compatible pLacI plasmid to ensure repression unless induced *(76)*. The recombinant plasmids with the genes encoding *Ledo*1986, *Gtra*1516 and *Gtra*1270 under the control of the T7 promotor were used for transformation. Transformation was carried out as described *(78)*. Subsequently a single colony was picked, suspended in 300 µl of LB and the cells were then evenly plated on six LB agar plates containing 50 µg/ml carbenicillin, 34 µg/ml chloramphenicol and 1 % glucose. The plates were then incubated overnight at 37 °C. Cells were scraped from the plates, resuspended in 50 ml of LB, and inoculated into 6 L of LB containing 100 µg/ml ampicillin. For protein expression the cell cultures were grown at 37 °C with shaking (225 rpm) until OD_{600} reached 0.8~1.0, and then IPTG was added to a final concentration of 0.5 mM. The induced cultures of *Ledo*1986, *Gtra*1516 and *Gtra*1270 were grown under optimum induction conditions, which were determined experimentally as described in *Results*. Then the cells were harvested by centrifugation and washed once with 50 mM Tris·Cl buffer, pH 7.5, and the cell pellets were weighed and stored at -80°C until use.

Optimization of expression condition

In order to determine conditions allowing the highest expression levels of soluble proteins, small scale (50 ml) IPTG-induced cultures expressing *Ledo*1986, *Gtra*1516 and *Gtra*1270 were grown for an additional 1, 2 or 3 h at 37 °C, or 3, 6 h and overnight at 20 °C. Then the cells were harvested by centrifugation, resuspended in 50 mM Tris·Cl

buffer, pH 7.5 (2 ml/ g cells), and subjected to sonication using a Branson Sonifier 250 under 50% full power, output of 4, for 10 bursts of 6 s each. After centrifugation (62000 x *g*), the expression level of each protein in the supernatant at different induction condition was examined by SDS-PAGE.

Purification of *Ledo*1986

Preparation of crude extract

All procedures were carried out at 0-4 °C. Frozen cell paste was thawed and the wet bacterial cells (24-30 g) were suspended in cold Tris·Cl buffer (pH 7.5 at room temperature, pH 7.9 at 4°C) (2 ml/ g cells) with DNase (1 mg/ml cell suspension) and RNase (0.25 mg/ml cell suspension). Batches (30-40 ml) of re-suspended cells on ice were sonicated using a Branson Sonifier 250 at 50% full power, output of 4, for 15 bursts of 10 s each. The resulting slurry was centrifuged using a Backman ultracentrifuge at 62000 x *g* for 1 h in order to separate the cell-free extract from cell debris. The yellowish brown supernatant was carefully decanted from the pellet and used as the crude extract.

Fast-Flow DEAE-Sepharose chromatography

Crude extract was applied to the column (20 cm x 3 cm) equilibrated with 50 mM Tris·Cl buffer, pH 7.5, at a rate of 6 ml/min. Once the sample was loaded, the column was washed (6 ml/min) with about 800 ml 50 mM Tris·Cl buffer pH 7.5, and 200 ml of the same buffer containing 50 mM NaCl until no protein was detected in the eluate fractions. A linear gradient (800 ml) from 50 mM to 150 mM NaCl in 50 mM Tris·Cl buffer was applied at a flow rate of 2 ml/min and the Absorbance of the fractions was measured at wavelengths of 280 nm (for protein) and 450 nm (for flavin). The enzyme

activity in each fraction was also measured. The fractions containing peak enzyme activity were combined.

HP-Phenyl-Sepharose chromatography

The pooled fractions from the Fast-Flow DEAE-Sepharose column were concentrated to 2-5 ml using a YM-30 ultrafiltration membrane and brought to 20% saturation (10.7 g / 100 ml) with ammonium sulfate. The sample was then centrifuged at 15600 x *g* for 15 min to remove precipitate, and the supernatant was then loaded on a High Performance Phenyl-Sepharose chromatography column (15 cm x 3 cm) equilibrated with 50 mM Tris·Cl buffer, pH 7.5 containing 20% saturated ammonium sulfate. A linear gradient (800 ml) from 20% - 0% saturation of ammonium sulfate in 50 mM Tris·Cl buffer, pH 7.5, was applied at a flow rate of 2 ml/min. The Absorbance at wavelengths of 280 nm and 450 nm as well as enzyme activity, were measured for the collected fractions. The fractions containing peak enzyme activity were combined.

Purification of *Gtra*1516

Preparation of crude extract

The procedure was the same as described for *Ledo*1986.

Fast-Flow DEAE-Sepharose chromatography

The procedure was the same as described for *Ledo*1986. The combined fractions were then applied to an octyl-Sepharose column.

Octyl-Sepharose chromatography

The combined fractions from the Fast-Flow DEAE-Sepharose column were concentrated to 2-5 ml using a YM-30 ultrafiltration membrane and brought to 20% saturation (27.1 g/ 100 ml) with ammonium sulfate. The sample was centrifuged at 15600 x *g* for 15 min to remove precipitate, and the supernatant was then loaded on an octyl-Sepharose chromatography column (10 cm x 2 cm) equilibrated with 50 mM Tris·Cl buffer, pH 7.5, containing 20% saturated ammonium sulfate. A linear gradient (600 ml) from 20% - 0% saturated ammonium sulfate in 50 mM Tris·Cl buffer, pH 7.5, was applied at a flow rate of 2 ml/min. An additional 300 ml of 50 mM Tris·Cl buffer, pH 7.5, and then 300 ml of 5 mM Tris·Cl buffer, pH 7.5, was used to elute remaining proteins. The Absorbance at wavelengths of 280 and 450 nm, as well as enzyme activity, were measured for the collected fractions. The fractions containing peak enzyme activity were combined.

Purification of *Gtra*1270B

Preparation of crude extract

The procedure was the same as described for *Ledo*1986.

Fast-Flow DEAE-Sepharose chromatography

The procedure was the same as described for *Ledo*1986.

High Performance Phenyl-Sepharose chromatography

The procedure was the same as described for *Gtra*1516.

Octyl-Sepharose chromatography

The procedure was same as described for *Gtra*1516. This was followed by gel filtration chromatography.

Gel Filtration chromatography

The combined fractions from the octyl-Sepharose column were concentrated to 2-5 ml using a YM-30 ultrafiltration membrane. The sample was centrifuged at 62000 x *g* for 15 min to remove precipitate, and the supernatant was then loaded on an S300 gel-filtration chromatography column (40 cm x 2 cm) equilibrated with 50 mM Tris·Cl buffer, pH 7.5. Then 240 ml of 50 mM Tris·Cl buffer, pH 7.5, was used to elute the proteins at a flow rate of 0.5 ml/min.

Since the purified *Gtra*1270B protein was not active, SDS-PAGE was used to locate the fractions containing the target protein though-out the purification steps.

Enzyme activity assays

All assays were carried out using a Cary Bio50 spectrophotometer at room temperature, with 50 mM Tris·Cl, pH 7.5 as the assay buffer, unless indicated otherwise.

For bacterial salicylate hydroxylase, the consumption of NADH is proportional to the formation of product *(5)*. Thus, enzyme activity for *Ledo*1986 was estimated by the consumption of NADPH (the enzyme was not active with NADH). Each assay was routinely carried out in 50 mM Tris·Cl buffer, pH 7.5, containing 0.2 μmol NADPH, 0.15 μmol salicylate and 2.8 μM enzyme in a total volume of 1 ml. The activity indicated is the oxidation of NADPH in the presence of salicylate (i.e. with background NADPH

consumption subtracted). One unit of activity is defined as the amount of enzyme required to oxidize one μmol of NADPH per min at room temperature, and the specific activity is defined as enzyme units per mg of protein.

The activity assay for *Gtra*1516, a predicted phenol hydroxylase, was as described for salicylate hydroxylase, but using phenol as substrate instead of salicylate. Assays were carried out in 50 mM Tris·Cl buffer, pH 7.5, containing 0.2 μmol NADPH, 0.15 μmol phenol and 0.8 μM enzyme in a total volume of 1 ml.

Enzyme activity of *Gtra*1270 (A and B), a predicted catechol-1, 2-dioxygenase, was determined spectrophotometrically by measuring the increase in Absorbance at 260 nm *(79)*. The assay was carried out in 50 mM Tris·Cl buffer, pH 7.5, containing 0.15 μmol of catechol and 0.95 μM enzyme in a total volume of 1 ml at room temperature. One unit of enzyme activity is defined as the amount of enzyme that produces 1 μmol of *cis,cis*-muconic acid per min under the standard assay conditions. Enzyme activity of *Gtra*1270 (A and B) was also determined spectrophotometrically as for hydroxyquinol-1, 2-dioxygenase by measuring the increase in absorbance at 245 nm *(74)*. The assay was carried out in 50 mM Tris·Cl buffer, pH 7.5, containing 0.15 μmol of hyroxyquinol and 0.95 μM enzyme in a total volume of 1 ml. One unit of enzyme activity is defined as the amount of enzyme that produces 1 μmol of maleylacetate per min under the standard assay conditions. It is the maleylacetate that absorbs at 245 nm.

Protein analysis by SDS-PAGE and BCA assay

SDS-PAGE (12 % acrylamide separating in gel) was used to examine the purity of protein preparations, and estimate subunit molecular weights. The BCA assay (Pierce) was used to estimate protein concentration, as described by the manufacturer. To remove interfering substances, TCA precipitation was used in conjunction with the BCA assay *(80)*. BSA supplied with the assay kit was used as the standard to construct a calibration curve.

Mass spectrometry

The purified protein was digested by trypsin either in solution or in gel depending on the purity. Subsequently, the peptides were analyzed by MALDI or ESI-Q-TOF mass spectrometry *(81)*. Protein mass measurement was done using a Waters CapLC system coupled to an ESI-Q-TOF 2 mass spectrometer. The CapLC system was equilibrated with a mixture of 90% solvent A (97% H_2O: 3% acetonitrile: 0.1% formic acid) and 10% solvent B (97% acetonitrile : 3% H_2O : 0.1% formic acid). The samples were loaded and desalted in the Symmetry 300 C18 trap column (0.35 mm × 5mm) with the equilibration solvent. Then the proteins were eluted into the mass spectrometer with a gradient of 10% to 90% solvent B over 30 min.

Thin layer chromatography

Thin layer chromatography was used to identify flavins and the products from enzymatic reactions. Both n-butanol-acetic acid-water (12: 3: 5) and $Na_2HPO_4{\cdot}4H_2O$ (5% in water) *(82)* were used as solvent systems to separate samples on 0.25 mm silica gel with fluorescent indicator (Kodak). The quenching of fluorescence of FAD standard,

FMN standard, enzymatic reaction products and the corresponding standards were detected under UV light.

High performance liquid chromatography

HPLC was performed on an Agilent 1100 system equipped with a variable-wavelength UV-vis detector set at 254 nm. A C_{18} reversed phase column (4.6 mm x 150 mm) equilibrated with a solvent mixture containing 2% acetic acid (60%) and methanol (40%) *(83)* was used to separated substrate and product at a flow rate of 1 ml/min. Samples were treated with ZipTip (Millipore) and then injected to the column.

Oxygen consumption assay

For *Ledo*1986, reactions were carried out in 50 mM Tris·Cl buffer, pH 7.5, containing 300 μM NADPH, 2.8 μM purified enzyme and 100 μM aromatic compounds in a total volume of 1 ml at 25 °C. For *Gtra*1516, reactions were carried out in 50 mM Tris·Cl buffer, pH 7.5, containing 300 μM NADPH, 8.3 μM purified enzyme and 100 μM aromatic compounds in a total volume of 1 ml at 25 °C. Oxygen consumption was measured using an oxygen electrode from Hansatech with a voltage of 1 V, and the oxygen trace. In order to set 1 V as 100 % oxygen content on the chart and 0 V as 0%, The voltage was set as 1 V at 100 %, and then sodium dithionite was added to the reaction chamber containing 1 ml of buffer. For example, when the voltage shows 75 mV and the oxygen content is indicated as 7.5 % on the chart, the machine is well set.

Fluorescence spectroscopy

An Aminco Bowman series 2 fluorimeter was used for fluorescence measurements. The emission spectra were scanned from 500 nm to 700 nm following excitation at 280 nm and 530 nm for *Ledo*1986 and *Gtra*1516, respectively. Bandwidth and voltage were set as 4 and 750V, respectively. Samples were placed in 500 µl cuvette and data were collected at a scan speed of 1 nm/s. Buffer was 50 mM Tris·Cl, pH 7.5.

Substrate binding study

For *Ledo*1986, 50 mM Tris·Cl buffer, pH7.5 containing 0.56 µM enzyme was placed in 500 µl cuvette and data were collected by titrating the enzyme with gradually increasing amounts of salicylate or anthranilate. For *Gtra*1516, 5 µM enzyme in 50 mM Tris·Cl buffer, pH7.5 was placed in 500 µl cuvette and titrated with gradually increasing amounts of phenol or resorcinol. K_d values were analyzed by a computerized, weighted nonlinear regression method with Grafit v4.0 software (Erithacus).

Enzyme kinetics

The kinetic constants, V_{max} and K_M, for anthranilate and resorcinol were estimated for *Ledo*1986 and *Gtra*1516, respectively. For *Ledo*1986, the reactions were initiated by the addition of different amounts of anthranilate to 50 mM Tris·Cl buffer, pH 7.5, containing 300 µM NADPH and 0.56 µM of the purified enzyme at room temperature. The reactions were also carried out in 50 mM Tris·Cl buffer, pH 7.5, containing 160 µM anthranilate and 0.56 µM of enzyme with the addition of different amounts of NADPH. Initial rates corresponding to different concentrations of anthranilate or NADPH

(corrected for the rates in the absence of aromatic substrate) were converted from A_{370}/min to μmole/min using an extinction coefficient of 2660 $cm^{-1}M^{-1}$, since anthranilate does not absorb at this wavelength *(83)*. For *Gtra*1516, the reactions were initiated by the addition of different amounts of resorcinol in 50 mM Tris·Cl buffer, pH 7.5, containing 300 μM NADPH and 0.48 μM of the purified enzyme at room temperature. The reactions were also carried out in 50 mM Tris·Cl buffer, pH 7.5, containing 120 μM resorcinol and 0.48 μM of enzyme with the addition of different amounts of NADPH. The Michaelis-Menten equation with Grafit v4.0 software (Erithacus) was used to initial rates at different substrate concentrations, and to estimate the kinetic parameters.

Iron quantitation for *Gtra*1270

An aliquot (~100 μl) of desalted sample from the iron-complexation experiment was transferred to a 1.5 ml microcentrifuge tube and 30 % (w/v) trichloroacetic acid was added to obtain a final concentration of 5 % TCA (w/v). After centrifugation for 5 min at 16000 x g, an aliquot of the supernatant (50 μl) was transferred to a new tube. Saturated ammonium acetate, 0.12 M ascorbic acid, 0.25 M Ferrozine and MilliQ water were added to a final volume of 1.0 ml. After 30 min, samples were centrifuged for 1 min at 16000 x g and the absorbance at 562 nm of the supernatant was measured. A standard curve was constructed using ferrous ammonium sulfate hexahydrate *(84)*.

Assay of *Gtra*1270 for catechol or hydroxyquinol dioxygenase activity

When catechol was used as a substrate, enzyme activity was estimated spectrophotometrically by measuring the increase in absorbance at 260 nm *(79)*. When hydroxyquinol was used as a substrate, enzyme activity was determined

spectrophotometrically by measuring the increase in absorbance at 245 nm *(74)*. Assays were carried out with the as-isolated enzyme, as well as enzyme which was incubated with a ten-fold excess of Fe^{2+} under aerobic conditions. Purified enzyme was also pre-incubated with Fe^{2+} under a stream of nitrogen. Reactions were carried out in a sealed cuvette with solutions prepared anaerobically. Enzyme assays were conducted in the presence of 0.53 μM enzyme in 50 mM air-saturated Tris·Cl buffer, pH 7.5, containing 0.15 μmoles of substrate in a total volume of 1 ml.

RESULTS

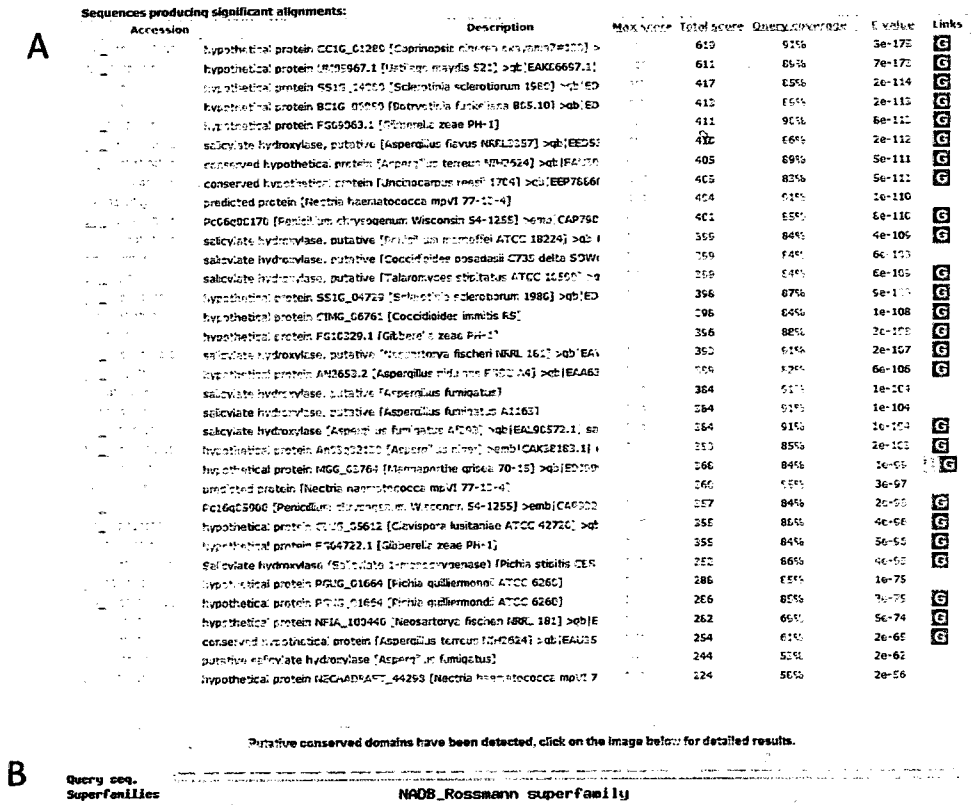
Bioinformatics studies

The cDNA libraries of over 70,000 genes from 15 diverse fungal species are available from the Concordia fungal genomics project. By searching with the term "oxygenase" in this database, https://fungalgenomics.concordia.ca, 426 sequences were identified as potential oxygenases, some of which may play roles in aromatic degradation pathways. For example, using text searching for "phenol hydroxylase" two *Aspergillus niger* sequences and two *Amorphotheca resinae* sequences were found.

The sequences in the database had been annotated when the database was first constructed. BLAST was used to compare the sequences of interest against the current set of sequences in Genbank. BLAST is short for Basic Local Alignment Search Tool. It is a set of similarity search programs designed to explore all of the available sequence databases regardless of whether the query is protein or DNA *(85)*.

BLAST analysis of the *Ledo*1986 sequence from *Lentinula edodes* against the GenBank database generated a large number of sequences producing significant alignment, as shown in Figure 11. This Figure is only a portion of the sequence list, which arrays from maximum similarity to minimum similarity with the sequence of *Ledo*1986. A putative conserved domain of the *NADB Rossman superfamily* was detected within these sequences. The Rossman-fold NAD(P)H/ $NAD(P)^+$ binding (NADB) domain is found in a large number of dehydrogenases involved in various metabolic pathways. As indicated by underlining in the alignment list, many of the

sequences having high similarity with the query were hypothetical or putative salicylate hydroxylases. Thus, it is hypothesized that *Ledo*1986 from *Lentinula edodes* may be a salicylate hydroxylase. In the bioinformatics analysis (Figure 12), the starting methionine of *Ledo*1986 locates upstream from the beginning of the alignment.



http://blast.ncbi.nlm.nih.gov/Blast.cgi
Data collected on March 1, 2010

Figure 11: BLAST analysis of the sequence for *Ledo*1986 from *Lentinula edodes*

A is a list of sequences producing significant alignment with the query; B: putative conserved domain was detected

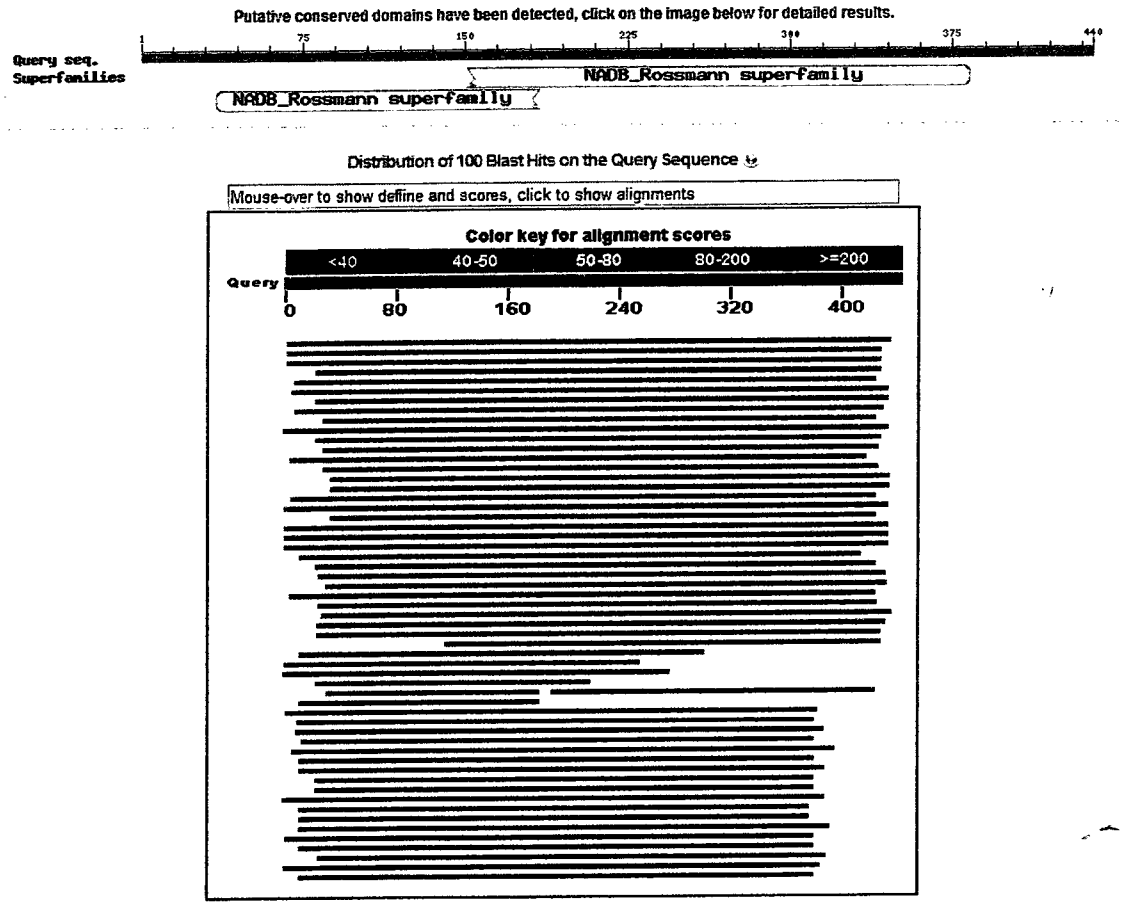


Figure 12: BLAST of *Ledo*1986

The sequence of *Ledo*1986 included the full sequence of the predicted conserved domain.

BLAST analysis of the sequence of *Gtra*1516 from *Gloeophyllum trabeum* against the GenBank database showed a large number of sequences producing significant alignment, as shown in Figure 13. This Figure is only a portion of the sequence list, which arrays from maximum similarity to minimum similarity, with the sequence of *Gtra*1516. Putative conserved domains of the NADB Rossman superfamily and PHOX-C family were detected within these sequences. The PHOX-C family is referred to as "FAD-dependent phenol hydroxylase (PHOX) family, C-terminal TRX-fold domain".

Many of the hits are to genes that are annotated as "putative", "hypothetical", or come from genome sequencing efforts. In addition, as underlined in the BLAST alignment list, *Gtra*1516 from *Gloeophyllum trabeum* showed strong sequence similarity with phenol hydroxylase from the aerobic topsoil yeast, *Trichosporon cutaneum,* whose structure has been solved *(66)*. The E value of its alignment with *Ledo*1986 is 3e-129. Thus, it is hypothesized that *Gtra*1516 from *Gloeophyllum trabeum* is a putative phenol hydroxylase. In the bioinformatics analysis of *Gtra*1516, the first start codon was selected as the first Methionine which was upstream from the beginning of the alignments (Figure 14).

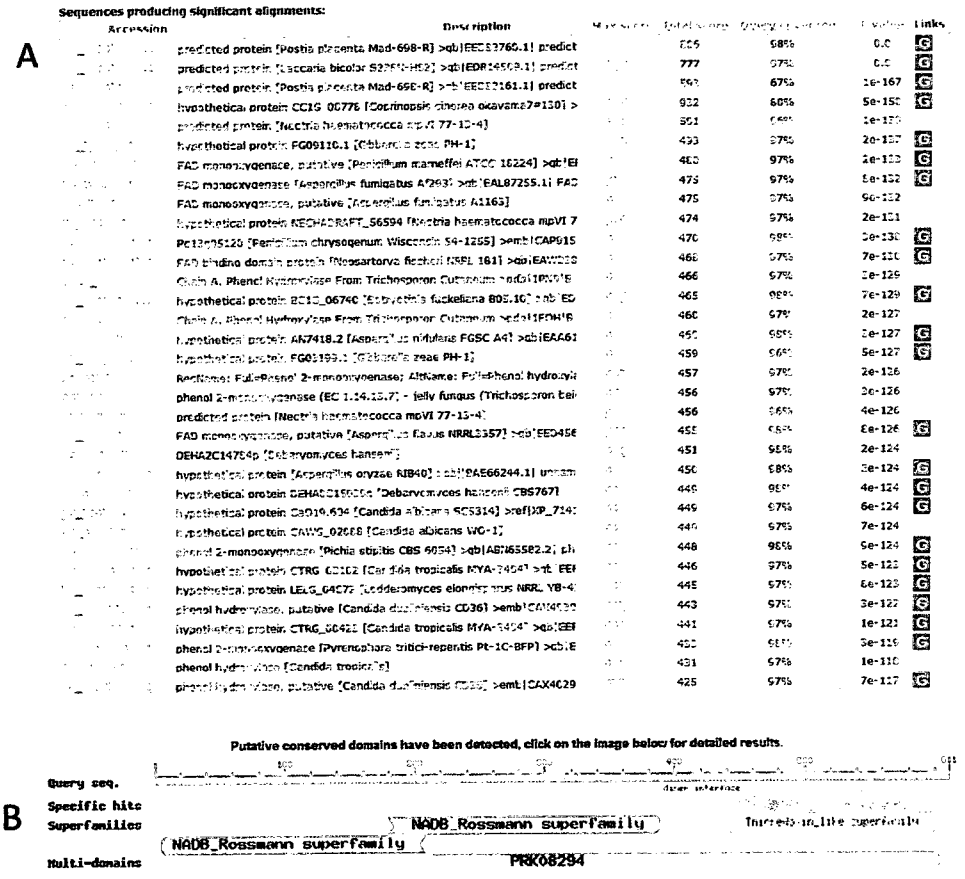


Figure 13: BLAST analysis of the sequence for *Gtra*1516 from *Gloeophyllum trabeum*

A is a list of sequences producing significant alignment with the query; B: putative conserved domains were detected

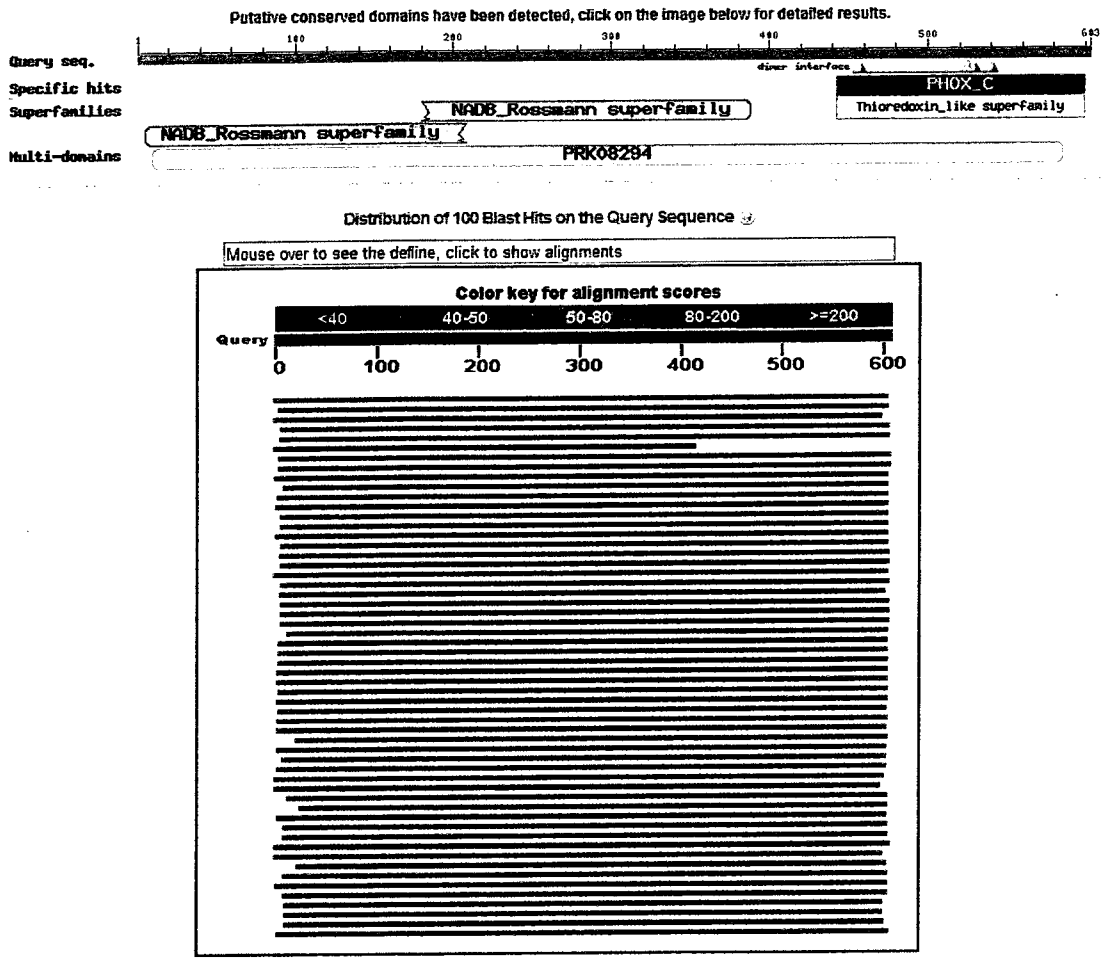


Figure 14: BLAST of *Gtra*1516

The sequence of *Ledo*1986 included the full sequence of the predicted conserved domain.

BLAST analysis of the sequence of *Gtra*1270 from *Gloeophyllum trabeum* against the GenBank database indicated a large number of sequences producing significant alignment, as shown in Figure 15. Putative conserved domains of intradiol_dioxygenase superfamily and catechol proteo were detected within these sequences. Most of these sequences are hits to genome sequences where the enzyme activity has never actually been characterized.

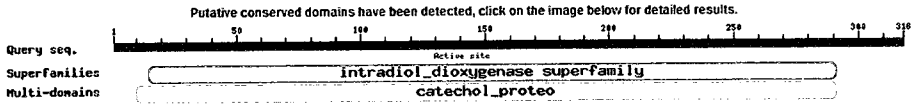
As underlined in the alignment list, BLAST searches for *Gtra*1270 from *Gloeophyllum trabeum* showed similarity with catechol dioxygenase and hydroxyquinol 1, 2-dioxygenase. Thus, it is hypothesized that *Gtra*1270 from *Gloeophyllum trabeum* is a putative catechol dioxygenase or hydroxyquinol-1, 2-dioxygenase. At the fungal genomics website, the first ATG (highlighted in green) was predicted for the sequence (Figure 16), and thus the primers (highlighted in green) were used to generate the PCR product that was named *Gtra*1270A. However, as shown in Figure 17, the start codon was downstream from the beginning of most of the aligned genes which means this sequence was missing the N-terminal part of the predicted conserved domain. In order to include the full conserved domain in the target gene, a pair of new primers (highlighted in yellow, Figure 16) was designed with an insertion of an ATG in the forward primer to extend the N-terminal of Gtra1270A, and the corresponding gene was named *Gtra*1270B. BLAST of *Gtra*1270B (Figure 18) shows that this target gene included the full alignment sequence of the predicted conserved domain.

A

Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Links
XP_001839720.1	hypothetical protein CC1G_09575 [Coprinopsis cinerea okayama7=	336	336	91%	3e-90	G
EEU35456.1	hypothetical protein NECHADRAFT_51181 [Nectria haematococca m	294	294	92%	9e-78	
XP_381523.1	hypothetical protein FG11347.1 [Gibberella zeae PH-1]	284	284	92%	9e-75	G
XP_001249756.1	hypothetical protein CIMG_03197 [Coccidioides immitis RS] >gb|EE	283	283	92%	2e-74	G
XP_002566535.1	Pc16g12860 [Penicillium chrysogenum Wisconsin 54-1255] >emb|C	283	283	91%	3e-74	G
CBI56604.1	unnamed protein product [Sordaria macrospora]	282	282	92%	4e-74	
EEU33965.1	hypothetical protein NECHADRAFT_54469 [Nectria haematococca m	281	281	91%	8e-74	
XP_658368.1	hypothetical protein AN0764.2 [Aspergillus nidulans FGSC A4] >gb|	281	281	94%	1e-73	G
XP_002146603.1	catechol dioxygenase, putative [Penicillium marneffei ATCC 18224]	280	280	91%	1e-73	G
XP_002478299.1	catechol dioxygenase, putative [Talaromyces stipitatus ATCC 1050(	279	279	91%	3e-73	G
XP_964363.2	hypothetical protein NCU01979 [Neurospora crassa OR74A] >emb|(	279	279	91%	4e-73	UG
XP_002377677.1	catechol dioxygenase, putative [Aspergillus flavus NRRL3357] >gb|I	279	279	91%	4e-73	G
XP_002562281.1	conserved hypothetical protein [Uncinocarpus reesii 1704] >gb|EEP	276	276	92%	2e-72	G
XP_001210715.1	conserved hypothetical protein [Aspergillus terreus NIH2624] >gb|E	276	276	92%	2e-72	G
XP_001942183.1	chlorocatechol 1,2-dioxygenase [Pyrenophora tritici-repentis Pt-1C-	276	276	92%	2e-72	G
XP_682632.1	hypothetical protein AN9363.2 [Aspergillus nidulans FGSC A4] >gb|	274	274	92%	1e-71	G
XP_001551553.1	hypothetical protein BC1G_09927 [Botryotinia fuckeliana B05.10] >	273	273	91%	2e-71	G
XP_001389637.1	hypothetical protein An01g12310 [Aspergillus niger] >emb|CAK441	273	273	92%	2e-71	G
XP_002547425.1	conserved hypothetical protein [Candida tropicalis MYA-3404] >gb|I	272	272	94%	4e-71	G
EEQ33768.1	hydroxyquinol 1,2-dioxygenase [Microsporum canis CBS 113480]	272	272	90%	4e-71	
EEU38445.1	hypothetical protein NECHADRAFT_55328 [Nectria haematococca m	271	271	97%	7e-71	
XP_001216586.1	conserved hypothetical protein [Aspergillus terreus NIH2624] >gb|E	271	271	92%	9e-71	G
EEU36617.1	hypothetical protein NECHADRAFT_94105 [Nectria haematococca m	270	270	92%	1e-70	
XP_001592996.1	hypothetical protein SS1G_05424 [Sclerotinia sclerotiorum 1980] >(	270	270	91%	2e-70	G
XP_001264329.1	catechol dioxygenase, putative [Neosartorya fischeri NRRL 181] >gl	269	269	92%	3e-70	G
XP_752797.1	catechol dioxygenase [Aspergillus fumigatus Af293] >gb|EAL90759.	268	268	92%	5e-70	G
XP_001902990.1	unnamed protein product [Podospora anserina] >emb|CAP60762.1|	268	268	92%	7e-70	G
XP_714371.1	hypothetical protein CaO19.2282 [Candida albicans SC5314] >ref|X	267	267	89%	1e-69	G
XP_001263832.1	catechol dioxygenase, putative [Aspergillus clavatus NRRL 1] >gb|E	267	267	92%	1e-69	G
EEQ47255.1	hypothetical protein CAWG_05821 [Candida albicans WO-1]	266	266	89%	3e-69	
XP_002762698.1	hydroxyquinol 1,2-dioxygenase [Paracoccidioides brasiliensis Pb01]	265	266	92%	3e-69	
EEH50700.1	hydroxyquinol 1,2-dioxygenase [Paracoccidioides brasiliensis Pb18]	264	264	92%	1e-68	
EFY15693.1	hydroxyquinol 1,2-dioxygenase [Verticillium albo-atrum VaMs.102]	263	263	92%	2e-68	
XP_001941989.1	hydroxyquinol 1,2-dioxygenase [Azospirillum sp. B510] >dbj|BAI74	262	262	88%	4e-68	
CAE32140.1	dioxygenase, putative [Aspergillus fumigatus]	261	261	92%	6e-68	
XP_002428630.1	6-chlorohydroxyquinol-1,2-dioxygenase, putative [Candida dubliniie	260	260	89%	1e-67	G
XP_001525983.1	hypothetical protein LELG_02461 [Lodderomyces elongisporus NRRL	259	259	92%	3e-67	G
XP_366543.1	hypothetical protein MGG_02619 [Magnaporthe grisea 70-15] >ref|X	259	259	94%	3e-67	G

B



http://blast.ncbi.nlm.nih.gov/Blast.cgi

Data collected on March 1, 2010

Figure 15: BLAST analysis of the sequence for *Gtra*1270 from *Gloeophyllum trabeum* (Data collected on May 7, 2010)

A is a list of sequences producing significant alignment with the query; B: putative conserved domain was detected

CGACGACAAGAAGCCCCTTCAAGCCAAGGCAGTGTCGAACGGCCTTCCGCCT(ATG)CC
TAACCTTGACTTGCCGTACCCGGACAGACCAGAGCTCATTACGGAGAACCTTCTGAAGT
TGACCAACTTGATTACCGATGAACGAAAGAAGTACATCTTCAAGAATTTAATTACGCAT
ATACACCAATTCATCAATGAGACAAGTATCACCACGGACGAATGG[redacted]
[redacted]ACCCGGACAGGTCAAATCTGCACGCCCATTCGTCAAGAGTTCATCCTGCTCT
CCGACGTCCTCGGCATCTCCGCGCTCGTAGACGCACTTAACAACCCCCCAGTCAGCGGC
GGGACGGAGAGCAGCGTGCTCGGGCCTTTCTTCACTGAGGACGCTCCCGACGTCAACAA
CGGCGATTCTATTGCGTCCGAGGGCAAGGGGCAGTACATGTACGTCGAGGGGCGTGTAA
TCGATACGCACGGGAAGCCAGTCCCGAACGCTCTCATCGAAACGTGGGAGACAGACGAG
TATGGCTTCTACGACACGCAATATGCGGACAGAAGCAAACCCGACTGCAGGGGCCGCTT
GCGGACGGACAAGGACGGTAGATACGGCTACCGCGCTGTCGTTCCAGTAGCGTACCCCA
TCCCCGGTGACGGCCCCGTCGGAGACTTACTGCTCATGCTCAACCGACACAATATGCGC
CCGAACCATCTGCATATGATGATCGAGGCACCGGGCTACCAGAAACTCACGACCGCGTT
CTACCCCGAGGGCGATGAGTGGCTTGCGAGCGACGCTGTTTTCGGCGTCAAGAAATCCC
TGGTCGTCACATTGAAGGATGTGGACAATGAGCAGGAGGCGCGCAAGCGTGGGTTCCCC
AAGGGGAGCCACTTCAAGTTGCTTGAGCATGATCTCGTGTTGGTTCCTGAGGCTGAGTC
CAAGGCGGCGCGGGAACAGTACGCGAGGGAACATGCTGTTAACAGGAGTAACGAGATTC
AGGCATGAAGTATGATTTGCGTCTTGTATGATGGGCGGACGTG[redacted]
[redacted]ATAGTTCAACCTGTCAGGAATGTATTTCTATTGCACATCCAAAAAAAAAAAAAAG
AAAAAAAAAAAAAAAAAAAC

Figure 16: DNA sequence of Gtra1270

Start codon is indicated in red, stop codon is in the box marked with red line.

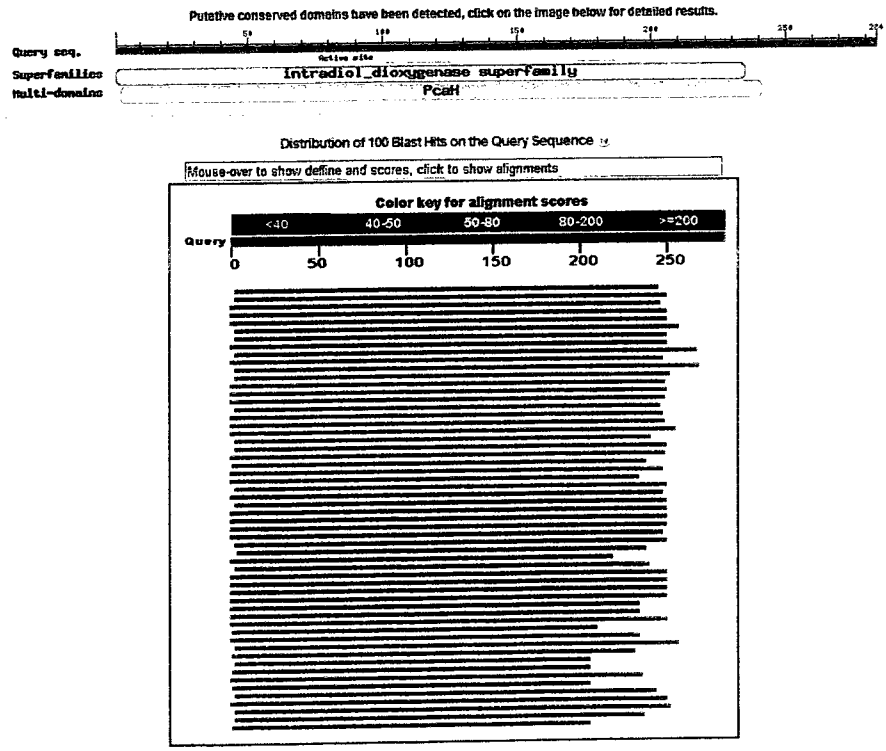


Figure 17: BLAST of *Gtra*1270A

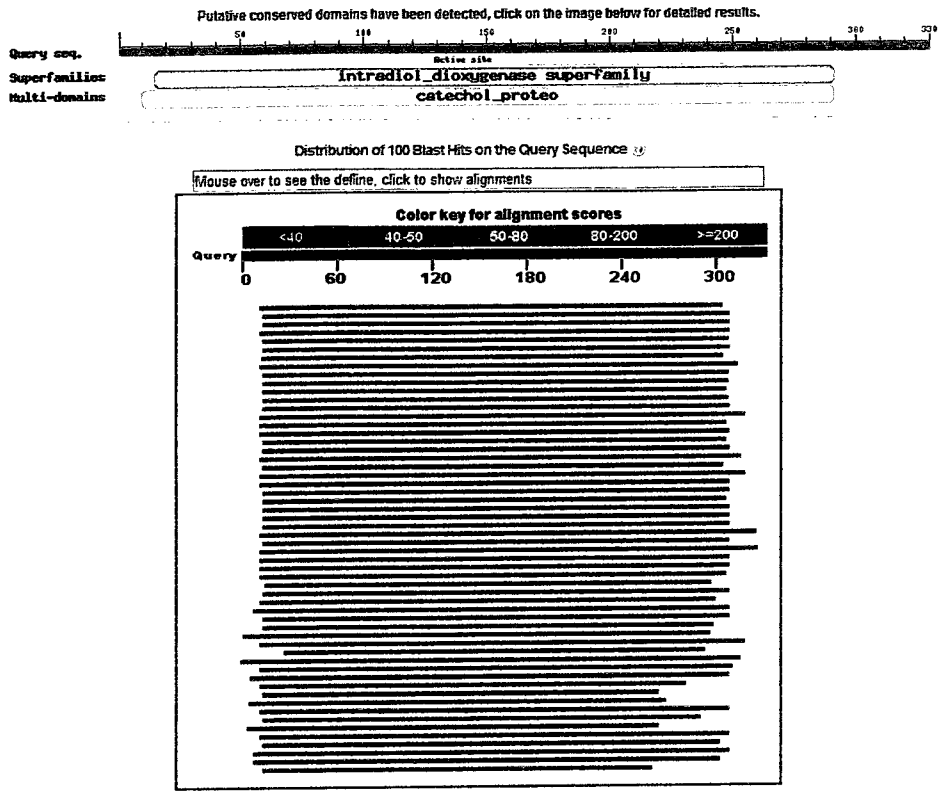


Figure 18: BLAST of *Gtra*1270B

PCR amplification of *Ledo*1986, *Gtra*1516 and *Gtra*1270

The expected target lengths for *Ledo*1986, *Gtra*1516, *Gtra*1270A and *Gtra*1270B are 1388 bp, 1825 bp, 844 bp and 983 bp, respectively. Fragments of the correct size were detected for each gene amplified (Figure 19). PCR products were purified and ligated into the blunt-ended pETBlue-1 vector.

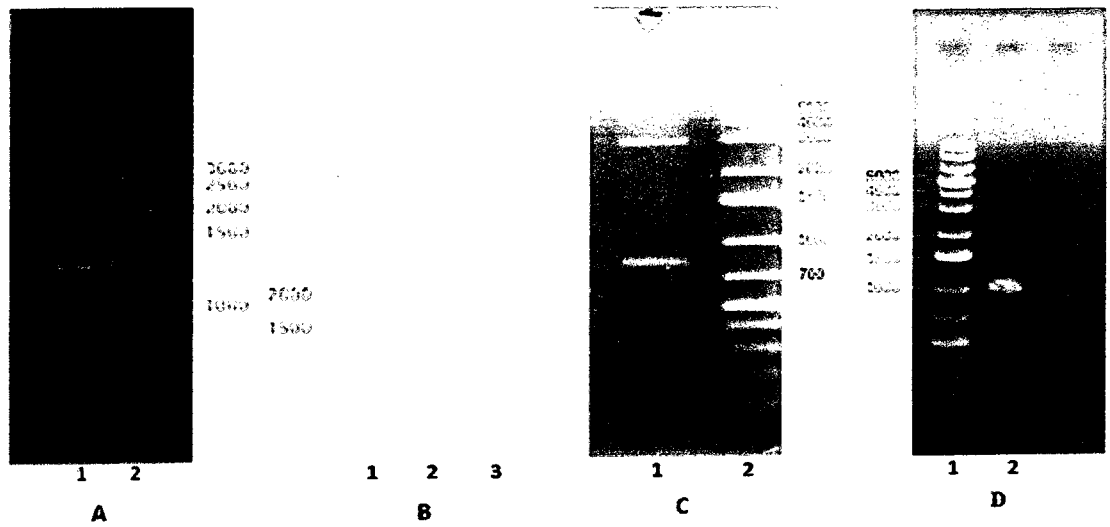


Figure 19: Agarose gel (0.7 %) of PCR products from *Ledo*1986, *Gtra*1516, *Gtra*1270A and *Gtra*1270B

Panel A: PCR amplification of *Ledo*1986. Lane 1, PCR product of *Ledo*1986, lane 2, DNA ladder. Panel B: PCR amplification of *Gtra*1516. Lane 1, DNA ladder, lane 2 and lane 3, PCR product of *Gtra*1516. Panel C: PCR amplification of *Gtra*1270A. Lane 1, PCR product of *Gtra*1270A, lane 2, DNA ladder. Panel D: PCR amplification of *Gtra*1270B. Lane 1, DNA ladder, lane 2, PCR product of *Gtra*1270B.

Restriction enzyme digestion

Plasmids were purified and checked for inserts in the correct orientation using restriction enzyme digestion. Recombinant plasmids from individual transformants harbouring *Ledo*1986 were digested using *Sph*I and *Eco*RI. Plasmid with the insert in the correct orientation should release a fragment of 1040 bp. As shown in Figure 20, among the eight plasmids tested, three showed the expected fragment (lanes 4, 5, and 7). Recombinant plasmids harbouring *Gtra*1516 were screened by *Eco*RI-digestion: a fragment of 619 bp will be excised if insert is in the correct orientation. As shown in Figure 21, among the eight recombinant plasmids screened, five showed the expected fragment for plasmid with insert in the correct orientation (lanes 1, 2, 3, 5 and 8).

Recombinant plasmid harbouring *Gtra*1270A were screened by digestion with *Xba*I and *Cla*I: for plasmids with inserts in the correct orientation, release of fragment with the length of 650 bp was expected. As shown in Figure 22, among the ten recombinant plasmids screened, five showed the expected fragment (lanes 1, 2, 6, 7, and 9). Recombinant plasmids harbouring *Gtra*1270B were also screened by *Xba*I and *Cla*I digestion, a fragment with the length of 789 bp was expected. As indicated in Figure 23, lanes 1, 3, 6, 7, and 9 showed the expected fragments.

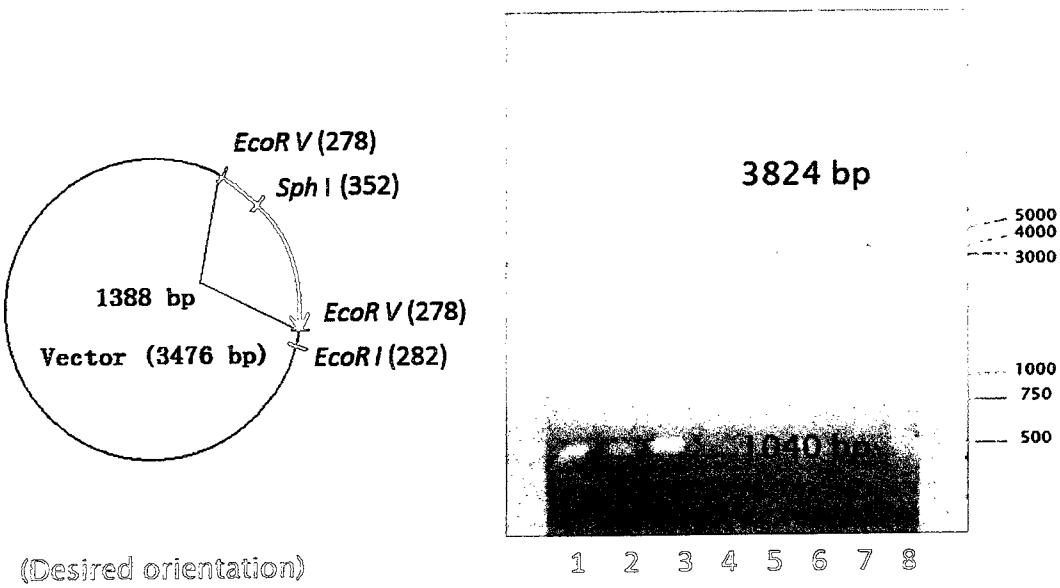


Figure 20: Agarose gel (0.7%) demonstrating restriction enzyme digestion of *Ledo*1986 with *Sph* I and *EcoR* I

(Left) Recombinant plasmid construct with the target gene inserted in the correct orientation into the *Eco*RV site. The length of the insert is 1388 bp and the length of the vector is 3476 bp. There is a single *Sph*I site at position 352 on the insert, and a single *Eco*RI site at position 282 on the vector. (Right) Agarose gel showing the fragments obtained from digestion of recombinant plasmids by *Sph* I and *Eco*RI. Lanes 1 to 8 are eight different plasmids purified from randomly picked colonies after transformation.

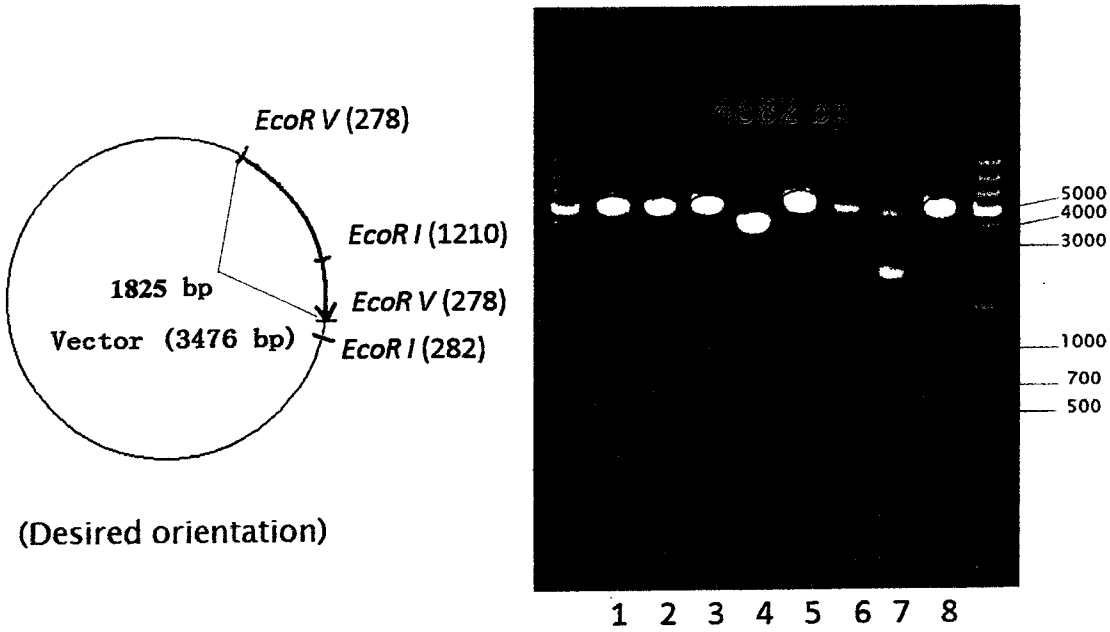


Figure 21: Agarose gel (0.7%) demonstrating *Eco*RI restriction enzyme digestion of plasmid with *Gtra*1516

(Left) Recombinant plasmid construct with the target gene inserted in the correct orientation into the *Eco*RV site. The length of the insert is 1825 bp and the length of the vector is 3476 bp. There is a single *Eco*RI site at position 1210 on the insert, and a single *Eco*RI site at position 282 on the vector. (Right) Agarose gel showing the fragments obtained from digestion of recombinant plasmids by *Eco*RI. Lanes 1 to 8 are eight different plasmids purified from randomly picked colonies after transformation.

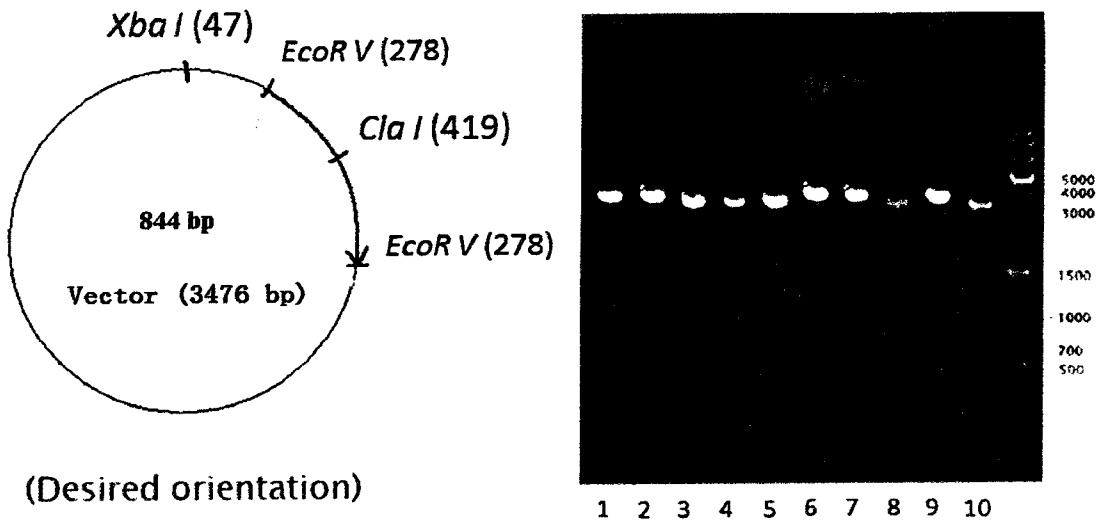


Figure 22: Agarose gel (0.7%) demonstrating restriction digestion of plasmids harbouring *Gtra1270*A with *Xba* I and *Cla* I

(Left) Recombinant plasmid construct with the target gene inserted in the correct orientation into the *Eco*RV site. The length of the insert is 844 bp and the length of the vector is 3476 bp. There is a *Cla* I site at position 419 on the insert, and a single *Xba* I site at position 47 on the vector. (Right) Agarose gel showing the fragments obtained from restriction digestion of *Gtra*1270 by *Xba* I and *Cla* I. Lanes 1 to 10 are ten different plasmids purified from randomly picked colonies after transformation.

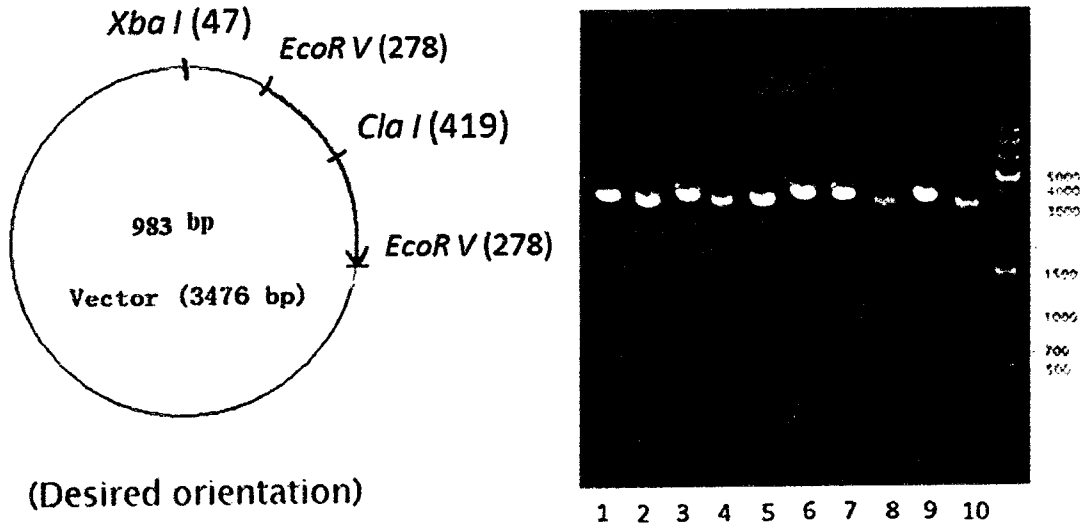


Figure 23: Agarose gel (0.7%) demonstrating restriction digestion of plasmids harbouring *Gtra1270*B with *Xba* I and *Cla* I

(Left) Recombinant plasmid construct with the target gene inserted in the correct orientation into the *Eco*RV site. The length of the insert is 983 bp and the length of the vector is 3476 bp. There is a *Cla* I site at position 419 on the insert, and a single *Xba* I site at position 47 on the vector. (Right) Agarose gel showing the fragments obtained from restriction digestion of *Gtra*1270 by *Xba* I and *Cla* I. Lanes 1 to 9 are nine different plasmids purified from randomly picked colonies after transformation.

DNA sequencing

DNA sequencing was performed for all of the recombinant plasmid using primers based on the sequence of the pETBlue-1 vector. No difference was found between the sequencing data and the known sequence.

Characterization results for *Ledo*1986

Ledo1986 is a putative salicylate hydroxylase, which catalyzes the hydroxylation of salicylate in the presence of oxygen and NAD(P)H. The protein was purified using a spectrophotometric assay for salicylate hydroxylase, as described in *Materials and Methods*.

Optimization of expression conditions for *Ledo*1986

Small scale cultures of *E. coli* (DE3) *lacI* pETBlue-1 *Ledo*1986 were grown for different lengths of time, and at various temperatures, after induction with IPTG to determine the optimal conditions for expression of soluble protein. As shown in Figure 24, *Ledo*1986, which has a predicted molecular mass of 48.6 kDa, appears to be expressed at the highest level in a soluble form after inducing at 18 °C for 16 h (Figure lane 7).

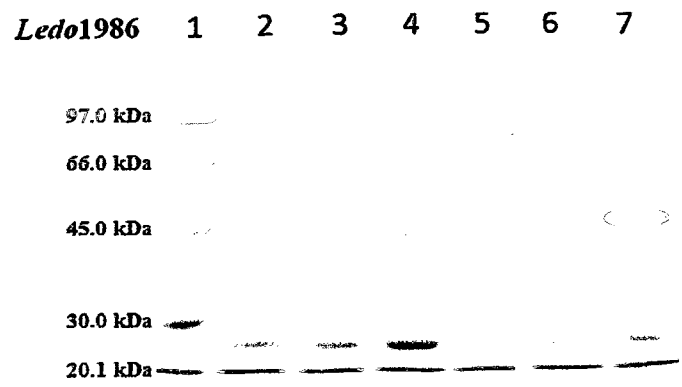


Figure 24: SDS-PAGE gel (12%) demonstrating the optimum expression conditions for *Ledo*1986

Low molecular weight standards (lane 1). Crude extracts from cells induced at 37 °C for 1 h (lane 2), for 2 h (lane 3), or 3 h (lane 4); 20 °C for 3 h (lane 5), for 6 h (lane 6), or overnight at 18 °C (lane 7).

Purification of *Ledo*1986

*Ledo*1986 was successfully overexpressed in *E.coli* and purified in good yield using two chromatographic steps as described in "Materials and Methods". Briefly, after inducing by IPTG, cells were harvested by centrifugation, disrupted by sonication, and subjected to ultracentrifugation to obtain crude extract.

Fast-flow DEAE-Sepharose chromatography and High Performance phenyl-Sepharose chromatography were applied in succession. In the Fast-Flow DEAE-Sepharose chromatography step, the peak of enzyme activity was eluted at approximately 0.1 M sodium chloride. The peak salicylate hydroxylase activity coincided with the appearance of a protein peak and fractions also showed absorbance at 450 nm: these fractions were pooled and concentrated by ultracentrifugation prior to the next step. In the High Performance phenyl-Sepharose chromatography step, the peak of the enzyme activity was eluted at approximately 13 % saturation with ammonium sulphate. The peak activity coincided with absorbance at 280 nm and 450 nm: fractions with peak specific activity and similar A_{450}/A_{280} ratios were pooled, concentrated by ultrafiltration, and stored at - 80°C.

A typical purification procedure for *Ledo*1986 is summarized in

Table 3. The purified enzyme preparation with a specific activity of 0.28 U/mg was obtained by 5-fold purification starting from the crude extract. Figure 25 demonstrates the protein purity observed after each purification step. In early preparations, an added step of gel-filtration chromatography was included, but this procedure did not give appreciable additional purification, and subsequent preparations omitted this step.

Table 3: Purification summary for *Ledo*1986

Purification step	**Vol (ml)**	**Total protein (mg)**	**Total activity * (U)**	**Specific activity (U/mg)**	**Purification (fold)**	**Yield (%)**
Crude extract	80	1290	71	0.055	1	100
DEAE-Sepharose chromatography	12	270	69	0.254	4.6	97
Phenyl-Sepharose chromatography	11	230	65	0.28	5.1	92

*The activity indicated is the oxidation of NADPH in the presence of salicylate. One unit of activity (U) is defined as the amount of enzyme required to oxidize one µmol of NADPH per min.

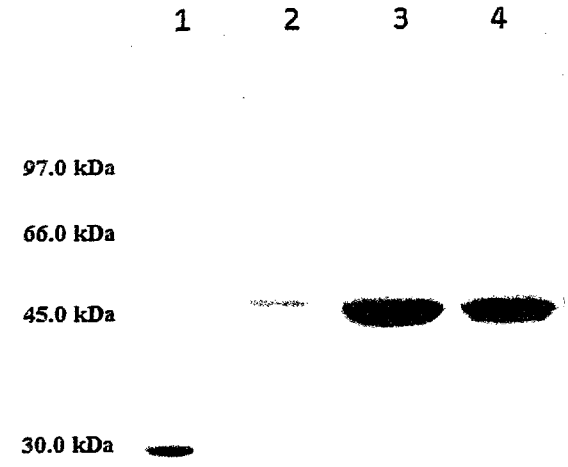


Figure 25: SDS-PAGE gel (12%) demonstrating the purity of recombinant *Ledo*1986 preparation after different purification steps

Lane 1: Low molecular weight standards. Lane 2: crude extract. Lane 3: pooled fractions from DEAE column. Lane 4: pooled fractions from Phenyl-Sepharose column.

Mass spectrometry

In solution Trypsin digestion and MALDI-TOF MS was conducted for the purified *Ledo*1986 protein. Figure 26 shows the spectrum for the peptides. Fragments observed in the spectrum are underlined in Figure 27, and cover 20% of the sequence.

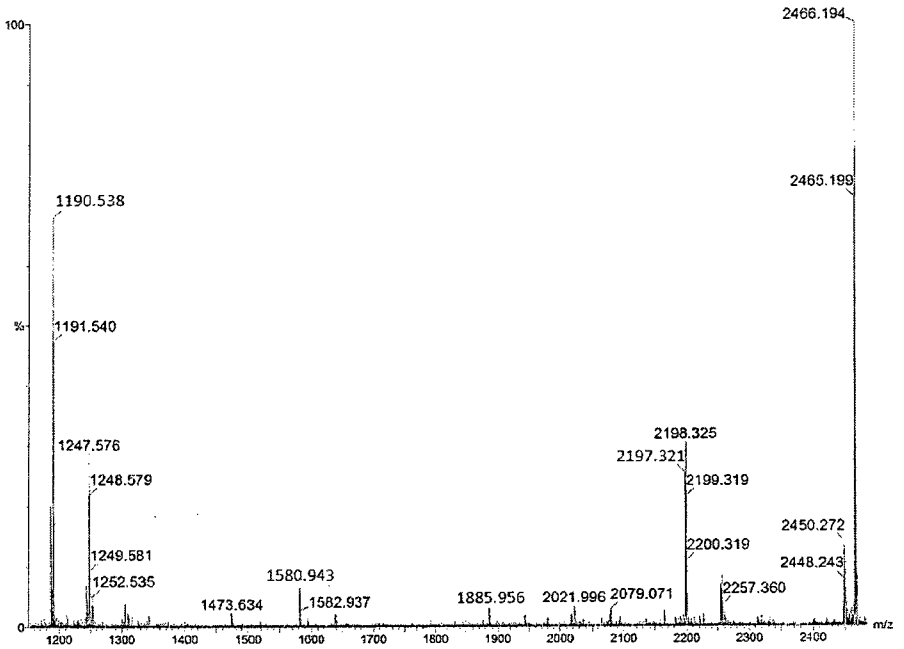


Figure 26: MALDI-TOF mass fingerprint for the tryptic digest of *Ledo*1986

MVSSTNPNFKQLRVAVIGGGLGGLSAAVALRRAGHLVEIYERRDFNVEVGASISC
AANGTQWLREWEVDIPDMKPVILMKLVMRDWETGRILNQYNLDKYEEEWGNVYNM
LHRQDMHATLLKTATSPEGKGTPCIVKIDHICETVDSEAGTVTFKNGVTVKADII
IGADGIRSVVRGQIGVVPDMKSAPQTCYRLNVKKSVVDELGLVKYSYEPAIQFWG
GLQGKNGRSKYYKIVMSPCSDGEIVSFYCFMPTELTKHHEEGFTFAEVPVSDVIA
GRYDELDPDCVNLLKNSVDRMPWRLYVHQPYDRWFAGNTCILGDAAHPMMPHQSQ
GACQAIEDAAALGIIFSDKYNFTTDVTAGLAMYQTIRKPRATRVQSASARATENL
NERIGFTSLTPHDASLAAAEGKLTINEMNSYKMHNHIATEVDKLGGGMGPTIAVM

Figure 27: Identification of the mass finger print for tryptic digests of *Ledo*1986

Absorption spectrum of purified *Ledo*1986

The UV-visible absorption spectrum of purified *Ledo*1986 was recorded using a Cary Bio50 spectrophotometer. The visible spectrum of this enzyme is typical of flavin-containing proteins *(7)*, exhibiting maxima at around 370 nm and 450 nm and minima at 320 nm and 400 nm, respectively (Figure 28). The peak at 450 nm showed a small shouder at 480 nm.

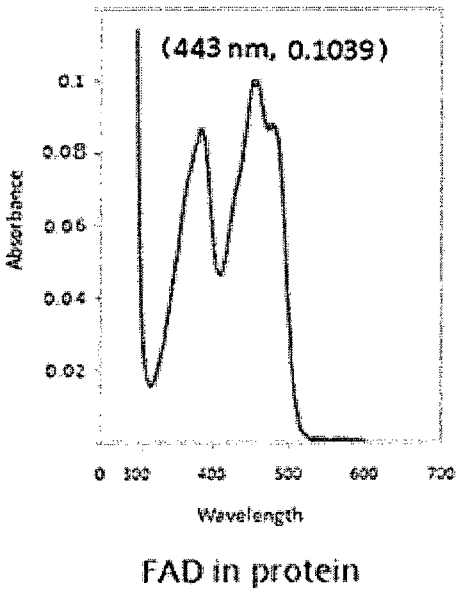


Figure 28: Absorption pectrum of purified *Ledo*1986

The spectrum was acquired in 50 mM Tris·Cl buffer, pH, 7.5 with 1 cm light path cuvette in the Cary Bio50 spectrophotometer. The protein concentration in the solution was 12.2 μM. A peak was observed at 443 nm with an absorbance of 0.1039.

Identification of the protein bound flavin in *Ledo*1986

The enzyme bound flavin was dissociated from the enzyme by heating at 100 °C for 5 min, followed by centrifugation at 15,600 x *g* for 10 min to remove denatured

protein. The supernatant from the boiled solution of the enzyme was bright yellow in color and was subjected to chromatography on a silica gel thin layer chromatography plate. Solutions of FAD, FMN and enzyme supernatant were spotted on the plate which was then developed with n-butanol-acetic acid-water (12: 3: 5) or $Na_2HPO_4 \cdot 4H_2O$ (5% in water) as solvents *(82)*. In both solvent systems the sample chromatographed as a single yellow spot, and the Rf values of each sample was calculated: in the first developing solvent, $Rf_{FAD} = 0.15$, $Rf_{sample} = 0.15$, $Rf_{FMN} = 0.26$, in the second developing solvent, $Rf_{FAD} = 0.63$, $Rf_{sample} = 0.63$, $Rf_{FMN} = 0.76$. The Rf value of boiled enzyme supernatant was the same as that of pure FAD, but quite different from that of FMN. Thus, the protein bound flavin of Ledo1986 was identified as FAD not FMN.

Quantitation of protein bound FAD in *Ledo*1986

The amount of FAD bound to the purified protein, and its extinction coefficient, were estimated by absorbance measurements. The spectrum of diluted native protein was recorded and showed an absorbance of 0.1039 at 450 nm. Then the FAD present in this sample was dissociated from the protein by heating at 100 °C for 5 min. The boiled supernatant showed an absorbance of 0.13 at 450 nm (Figure 29). The extinction coefficient of free FAD in the boiled supernatant is 11,300, the same as that of free FAD *(5)*. Protein concentration was determined by the BCA assay, and the molecular weight was determined by gel filtration (the subunit molecular weight was use in the calculation). The amount of FAD can be determined in the boiled sample, using the absorbance and extinction coefficient. The concentration of protein in the sample can be also calculated. The extinction coefficient at 450 can be obtained by taking the absorbance of the enzyme at 450 nm divided by the absorbance of the boiled sample at 450 nm, and multiplying by

11300. The extinction coefficient of the FAD on the protein was determined to be 9035 $M^{-1}cm^{-1}$. The number of molecules of FAD bound to 1 mole of enzyme was calculated to be approximately 1. Thus, *Ledo*1986 contains 1 mol FAD per 48600 Da subunit.

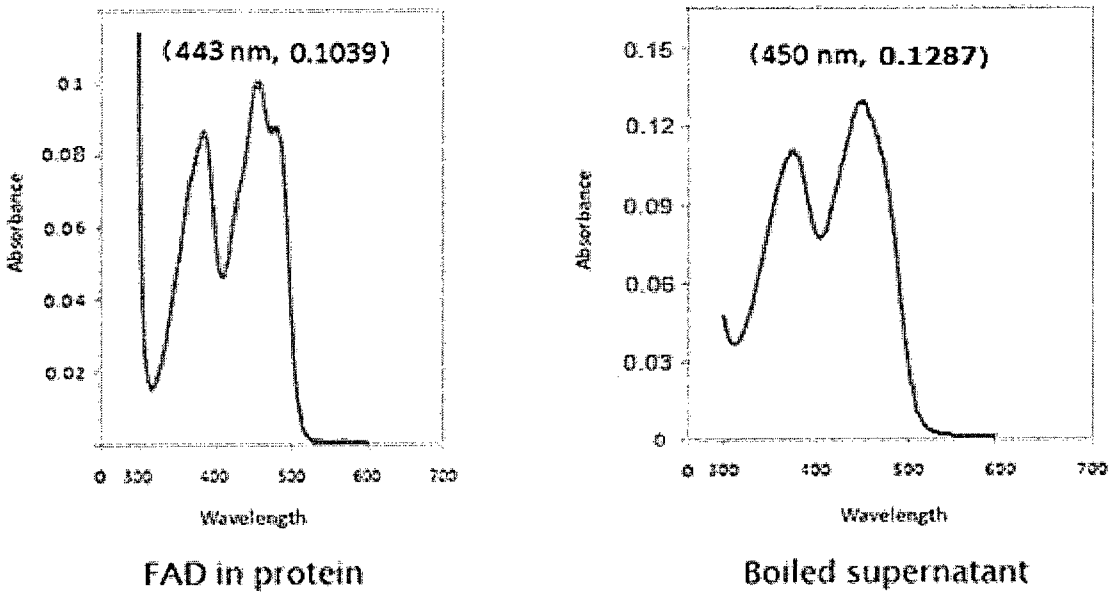


Figure 29: Absorption spectra of native protein and boiled supernatant of *Ledo*1986

The protein was diluted in 50 mM Tris·Cl buffer, pH, 7.5, with a concentration of 12.2 μM. At left is the spectrum of native protein and at right is the spectrum of boiled supernatant.

Estimation of molecular weight of *Ledo*1986 by gel filtration chromatography

The molecular weight of purified Ledo1986 was estimated by gel filtration on a Superdex S-200 column (Pharmacia) (40 cm x 2 cm) equilibrated with 0.1 M NaCl in 50 mM Tris·Cl buffer, pH 7.5. Molecular weight standards were loaded onto the column separately and a standard curve was obtained by plotting K_{av} vs. log MW (Figure 30). A sample of purified protein (100 μl) was chromatographed separately on the column. From the K_{av} value of the purified protein, log MW was calculated. The molecular

weight of Ledo1986 was estimated to be 49000 Da, close to the calculated molecular weight from the amino acid sequence (48600 Da). Thus *Ledo*1986 is a monomer.

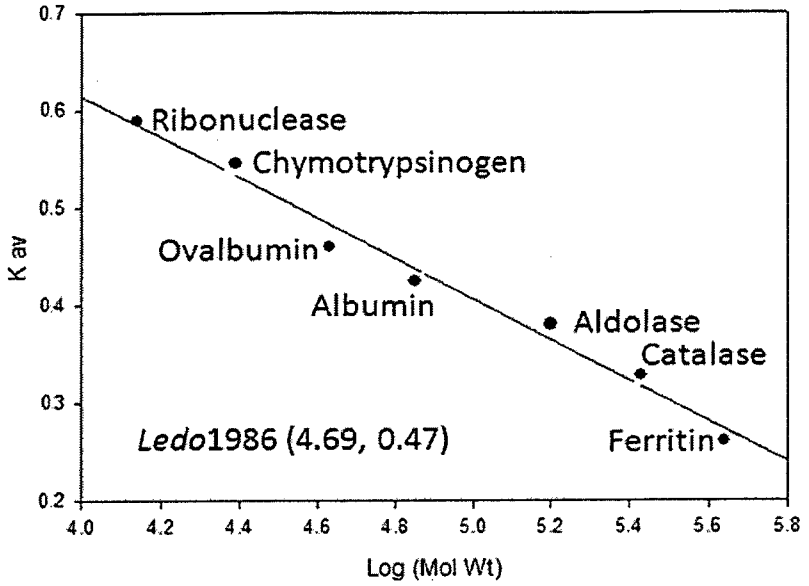


Figure 30: Estimation of the molecular weight of Ledo1986 by gel filtration chromatography

The standard protein used were Ribonuclease (13700 Da), Chymotrypsinogen (25000 Da), Ovalbumin (43000 Da), Albumin (67000 Da), Aldolase (158000 Da), Catalase (232000 Da), Ferritin (440000 Da). The equation obtained from K_{av} vs. log MW was y = - 0.21 x + 1.45. The position data for *Ledo*1986 was (4.69, 0.47).

Substrate specificity of *Ledo*1986

The substrate specificity of the enzyme was examined for a number of salicylate analogues by measuring consumption of NADPH in the presence of these compounds (Table 4). NADPH consumption in the presence of anthranilate was 310% of that with salicylate. Other analogues were less reactive. In the absence of aromatic substrate, the NADPH oxidase activity of the purified enzyme was almost negligible under standard

assay conditions. The non-enzymatic activity was subtracted from all of the values shown in Table 4.

Table 4: NADPH oxidation by *Ledo*1986 in the presence of various aromatic compounds

Aromatic compound	**Relative activity (%)**	**Standard error**
Salicylate	100	0.002
Anthranilic acid	309	0.0047
4-Chlorosalicylic acid	138	0.0067
2,4-Dhihydroxybenzoic acid	71	0.0008
Benzoate	61	0.0006
2,6-Dihydroxybenzoic acid	36	0.0004
2,5-Dihidroxybenzoic acid	27	0.002
2,3-Dihydroxybenzoic acid	24	0.0004
5-Chlorosalicylic acid	23	0.0018
3-Chlorosalicylic acid	18	0.0012
o-Anisic acid	9	0.0005
2-Hydroxyphenylacetic acid	5.6	0.0006
3,4-Dihydroxybenzoic acid	5.4	0.0005
m-Hydroxybenzoic acid	4.2	0.0011
4-Aminosalicylate	1.9	0.0007

The enzyme activity was determined in 50 mM Tris·Cl buffer, pH, 7.5 containing 0.2 µmol NADPH, 0.15 µmol of the respective substrate and 60 µg of the purified enzyme in a total volume of 1 ml, at room temperature. The oxidation of NADPH was measured at 340 nm. The enzyme activity with salicylate (0.28 U/mg) was taken as 100%. The values shown in the Table are averages of triplicates.

Product identification by TLC and HPLC

In order to identify products produced from salicylate, anthranilate or 4-chlorosalicylate (the most active compounds shown in Table 4), reactions were carried out in 50 mM Tris·Cl buffer, pH 7.5, containing 600 μM NADPH, 150 μM substrate and 0.3 mg enzyme in a total volume of 1 ml. Reaction mixtures were incubated when the reaction finished and then acidified with HCl. The precipitate formed was removed by centrifugation and the supernatant solution was extracted three times with 2 ml portions of ethyl acetate. The extracts were combined and dried over $MgSO_4$. Ethyl acetate extracts were then decanted into a new container, and solvent was evaporated under a stream of nitrogen. The residue was dissolved in ethyl acetate and analyzed by TLC, then the rest of the solvent was evaporated and the residue was dissolved in distilled water and analyzed by HPLC. For TLC, anthranilate, salicylate, catechol and 2, 3-dihydroxybenzoate standards were spotted on the plate together with extracts from reaction mixtures. No product was detected when either salicylate or 4-chlorosalicylate were used as substrates. In contrast, using anthranilate as a substrate, the Rf value of the reaction product was 0.49, almost the same as that of pure 2, 3-dihydroxybenzoate. In HPLC chromatography (Figure 31), 2, 3-dihydroxybenzoate was also detected using anthranilate as a substrate, but no product was detected for the salicylate or 4-chlorosalicylate reactions (data not shown).

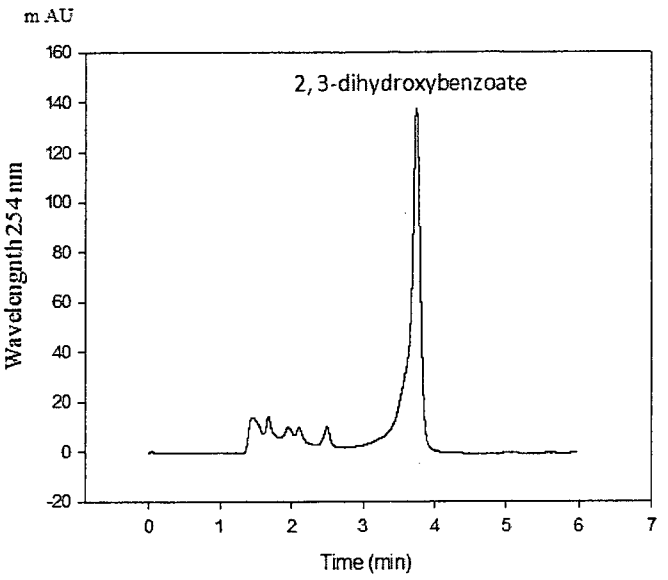


Figure 31: HPLC of reaction mixture of *Ledo*1986 incubated with NADPH and anthranilate

The major peak eluted in the same position as 2, 3-dihydroxybenzoate.

Oxygen consumption assays for *Ledo*1986

Since hydroxylases consume oxygen, the measurement of oxygen uptake is another way to monitor enzyme activity. Oxygen consumption by *Ledo*1986 in the presence of NADPH and aromatic compounds was measured using an oxygen electrode, as described in "Material and Methods". Reactions were performed at limiting levels of aromatic compounds, to permit observation of the stoichiometry between oxygen consumption and hydrogen peroxide formation in the presence of anthranilate, salicylate or 4-chlorosalicylate. The addition of catalase at the end of the reaction allow detection of any hydrogen peroxide produced *(7)*. The addition of catalase to the salicylate- or 4-chlorosalicylate-mediated oxygen consumption assay caused a release of one-half of the oxygen consumed (Figure 32). By contrast, catalase had a smaller effect upon oxygen

uptake in the anthranilate containing reaction mixture. Stoichiometries were estimated and shown in Figure 3.

The data in Table 5 indicate that hydrogen peroxide is formed in the salicylate or 4-chlorosalicylate mediated reaction, with a 1:1 stoichiometry observed between NADPH and H_2O_2. No product was observed from either salicylate or 4-chlorosalicylate using thin-layer chromatography (data not shown). Together these data indicate that oxygen was completely diverted to hydrogen peroxide production in the presence of these compounds. Thus, salicylate and 4-chlorosalicylate appear to be "non-substrate effectors" *(66)* which stimulate NADPH consumption (i.e. act as effectors) but where oxygen activation is not coupled to hydroxylation. Much less H_2O_2 was formed in the anthranilate reaction, from which 2, 3-dihydroxybenzoate was detected using TLC and HPLC, indicating anthranilate is a true substrate as well as an effector of NADPH oxidation.

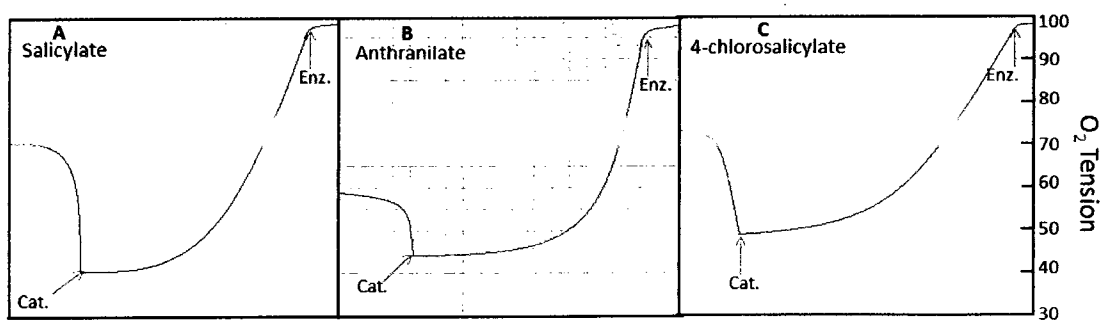


Figure 32: Consumption of oxygen by reaction mixtures for *Ledo*1986

O_2 consumption was measured with a Gilson model KM oxygraph equipped with a Clark oxygen electrode. The O_2 electrode traces are recorded backwards. The reaction chamber contained 600 μM NADPH, 2.8 μM of the purified protein and 150 μM of substrate in 50 mM Tris·Cl buffer, pH 7.5 with a total volume of 1 ml. Reactions were initiated by addition of salicylate (A), anthranilate (B) or 4-chlorosalicylate (C). Then 200 units of catalase (X) was added as shown by the arrows after reactions were almost complete.

Table 5: Comparison of the relative activities of *Ledo*1986 and the ratio of hydrogen peroxide produced to oxygen consumed for aromatic compounds.

(The values shown in the Table are averages of triplicates)

Substrate	**Relative activity**	**H_2O_2 / O_2**
Salicylate	100	1/1
Anthranilate	309	~0.3/1
4-chlorosalicylate	138	1/1

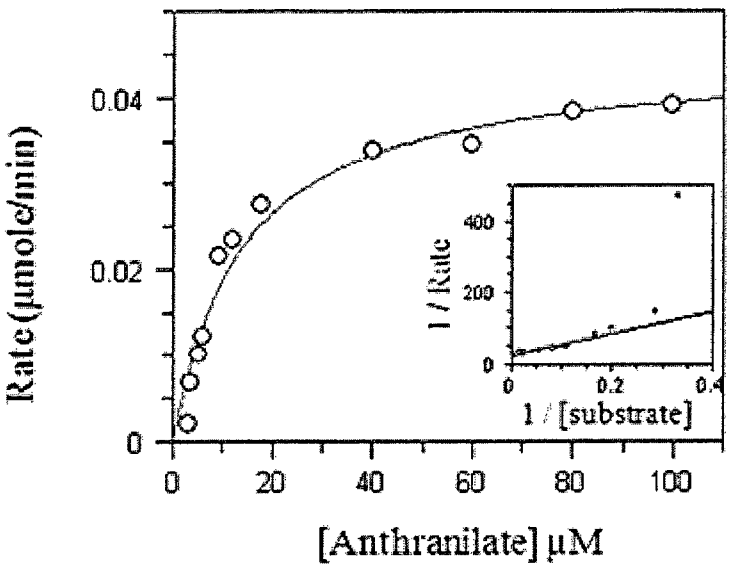
Binding of aromatic compounds to *Ledo*1986

As shown in (Appendix I), binding of anthranilate caused extensive spectral changes, and a K_d value of approximately 25 μM was obtained, which indicates that the binding between *Ledo*1986 and anthranilate is relatively tight.

Compared with anthranilate binding, salicylate binding caused a much less extensive spectral change and considerably weaker binding than anthranilate (Appendix II).

Steady state kinetics of *Ledo*1986 with anthranilate as substrate

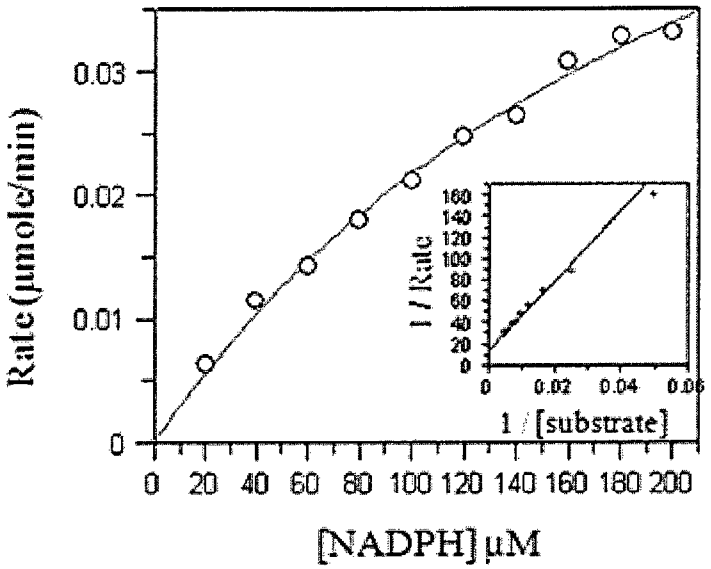
Kinetics constants, V_{max} and K_M when anthranilate was used as substrate were estimated for Ledo1986. Initial rates corresponding to different concentrations of anthranilate or NADPH (corrected for the rates in the absence of anthranialte) were obtained. Then data were fitted using the Michaelis-Menten equation with Grafit v4.0 software (Erithacus). With the software, Lineweaver-Burk plots were constructed from non-linear regression fitting of the data. Apparent V_{max} and K_m values of 0.05 μmoles/ min and 13.8 μM for anthranilate binding (Figure 33) and 0.08 μmoles/ min and 255 μM for NADPH binding were obtained from the fitted parameters (Figure 34).



Parameter	Value	Std. Error
Vmax (µmoles/min)	0.05	0.0027
Km (µM)	13.8	2.5322

Figure 33: Steady state kinetics of *Ledo*1986

Reactions were initiated by the addition of different amounts of anthranilate to 50 mM Tris·Cl buffer, pH 7.5, containing 300 µM NADPH and 0.56 µM enzyme at room temperature.



Parameter	Value	Std. Error
Vmax (µmoles/min)	0.08	0.0069
Km (µM)	255	35.4471

Figure 34: Steady state kinetics of *Ledo*1986

Reactions were initiated by the addition of different amounts of NADPH to 50 mM Tris·Cl buffer, pH 7.5, containing 140 µM anthranilate (saturated) and 0.56 µM enzyme at room temperature.

Characterization results for *Gtra*1516

*Gtra*1516 is a putative phenol hydroxylase, an enzyme that catalyzes the hydroxylation of phenol in the presence of oxygen and NADPH. Phenol hydroxylase purified from the soil yeast, *Trichosporon cutaneum*, is an FAD-containing protein *(67)*.

Optimization of expression conditions for *Gtra*1516

Small scale cultures of *Gtra*1516 were induced with IPTG and growth was continued at different temperatures for various lengths of time to determine optimal

conditions for expression of soluble protein. Figure 35 shows the expression level of the proteins under the conditions tested. *Gtra*1516, which is predicted to have a molecular weight of 66.6 kDa, was expressed at the highest level in a soluble form after induction for 3 h at 37 °C

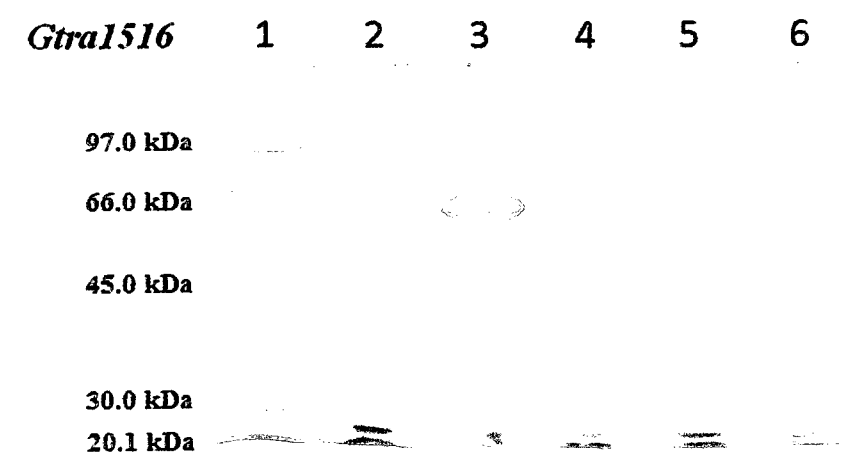


Figure 35: SDS-PAGE gel (12%) demonstrating the optimum expression conditions for *Gtra*1516

Low molecular weight standards (lane 1). Crude extracts from cells induced at 37 °C for 1 h (lane 2), or 3 h (lane 3); 20 °C for 3 h (lane 4), for 6 h (lane 5), or overnight (lane 7).

Purification of *Gtra*1516

*Gtra*1516 was overexpressed in *E.coli*(DE3)*lacI* cells and purified in good yield using two chromatographic steps as described in "Materials and Methods". However, the initial purification procedure used had to be modified to allow the isolation of the most active preparation of enzyme.

The first time this protein was purified, cells were harvested after inducing with IPTG, suspended in 50 mM Tris-HCl buffer, and disrupted by sonication. The crude

extract of soluble proteins represented the supernatant after ultracentrifugation, and this was subjected to ion-exchange chromatography followed by phenyl-Sepharose chromatography. After the first chromatographic column, the activity fractions were yellow in colour, and the enzyme appeared to lose a considerable amount of yellow colour during the second chromatography step. In order to establish the identity of the yellow compound, the partially-purified extracts were applied to TLC to identify the flavin. Also, when FAD was added to the assay the activity was stimulated.

Having established that the yellow-colour cofactor was FAD, a modified purification method using an FAD-containing buffer was used. Buffers used for crude extract preparation and during purification contained 10 μM FAD, 10 mM DTT, and 0.1 mM EDTA, as was the case for the successful purification of phenol hydroxylase from *T.cutaneum (6)*. The crude extract was subjected to ion-exchange chromatography followed by octyl-Sepharose chromatography, and fractions were monitored for phenol hydroxylase activity, as described in "Materials and Methods". In the Fast-flow DEAE-Sepharose chromatography step, the peak of phenol hydroxylase activity eluted at approximately 0.09 M sodium chloride in purification buffer. The peak activity coincided with peaks of protein absorbance at 280 nm and of flavin absorbance at 450 nm absorbance. In this purification step, the spectrophotometer was blanked with purification buffer which contained FAD, so the absorbance detected at 450 nm was due to the protein bound flavin. The peak fractions were pooled and concentrated prior to the next step. However, in the octyl-Sepharose chromatography step, the enzyme bound FAD gradually dissociated from the protein during the purification procedures, since the

enzyme appeared to lose a considerable amount of yellow colour as well as activity. So the collected fractions were assayed with addition of FAD.

The peak of enzyme activity from the octyl-Sepharose column was confirmed by SDS-PAGE to contain a protein with the expected molecular weight of 66 kDa (result not shown). Since the protein lost some flavin and the activity assay was stimulated by the addition of FAD, the collected fractions exhibiting peak enzyme activity were pooled, FAD (to 1 mM) was added, the protein was concentrated using an Ultrafree-CL centrifugal filter unit (Millipore), and buffer-exchanged (50 mM Tris·Cl buffer, pH 7.5) until no FAD was detected in the filtrate. Then the protein was aliquoted and stored at -80 °C.

A typical purification procedure is summarized in Table 6. The purified enzyme preparation had a specific activity of 2 U/mg after 20-fold purification from the crude extract. Figure 36 demonstrates the protein purity at each purification step: as can be observed, an addition S300 gel-filtration chromatography did not improve the purity (Figure 36, lane 7). The SDS-PAGE results indicate that the purified protein had a subunit molecular weight consistent with the 66600 Da predicted from the *Gtra*1516 amino acid sequence.

Table 6: Purification summary for *Gtra*1516

Purification step	**Vol (ml)**	**Protein (mg)**	**Total activity * (U)**	**Specific activity (U/mg)**	**Purification (fold)**	**Yield (%)**
Crude extract	63	1100	110	0.1	1	100
Fast-flow DEAE-Sepharose chromatography	5.1	108	105	0.97	9.6	95
Octyl-sepharose chromatography	1.93	20.7	41.4	1.99	19.9	38

*The activity indicated is the oxidation of NADPH in the presence of phenol. One unit of activity (U) is defined as the amount of enzyme required to oxidize one μmol of NADPH per min under the standard assay conditions ("Materials and Methods"). The protein collected after octyl-Sepharose chromatography was combined, had FAD added, ultrafiltered and assayed for its activity.

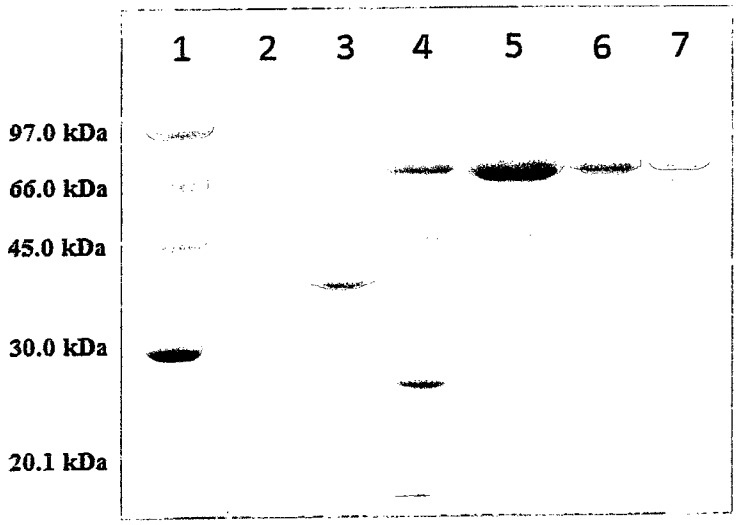


Figure 36: SDS-PAGE gel (12%) demonstrating the purity of recombinant Gtra1516 preparation after different purification steps

Low molecular weight standards (lane 1). Uninduced crude extract (lane 2). Pellet resuspension (lane 3). Crude extract (lane 4). Pooled fractions from Fast-flow DEAE column (lane 5). Pooled fractions from octyl-Sepharose column (lane 6). Pooled fractions from S300 gel-filtration column (lane 7).

Mass spectrometry

In-solution Trypsin digestion and ESI-Q-TOF MS was conducted on the purified *Gtra*1516 protein (Figure 37). Fragments observed in the spectrum are underlined in Figure 38, and cover 19% of the sequence.

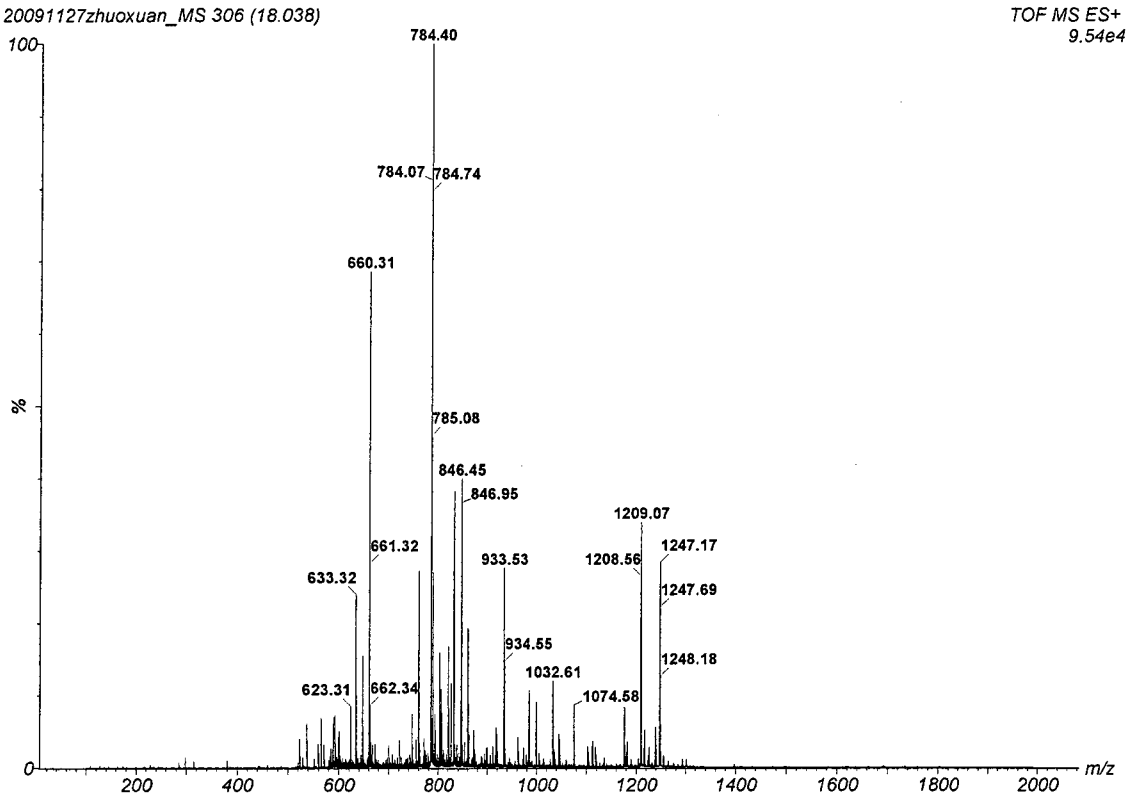


Figure 37: ESI-Q-TOF spectrum for the peptides of *Gtra*1516

MPVPAMKESDVDVLVIGAGPAGLMCAQGLARAGVNVRIVDKRPGKVAAGQADGIQPRTIEVLQSYGL
AERLLKEGNQMHMAAFYDPSPEGGIHRTGRTPDINAPTARFPFEVTLHQGAIESIFLDCLRSMGHEV
ERPVVPESLEISDNRDALKDPQARAVKVVLKHVDAPEGKDTEVVHAKYVVGADGAHSWVRKQLGFAM
EGEQTDYVWGVVDMIPDTDFPDIRNRCAIHSNNGSCMVIPREGDVVRLYIQLTDEDVRDVTTGRVDT
QKCSAEKLLEVAKKSFHPYRIKAKGDILWWTIYIIGQRVASKFSAHERVFIAGDACHTHSPKAGQGM
NASMNDTHNLIWKLTHVLRGWADISLLKTYELERRKYAQDLIAFDKEFASLFSKKPKSEENPDGVTQ
EEFVEAFRTFGGFTSGIGIHYAPSTIVDAAHQASASKLIIGQRVLPQTVIRAADARPYELQDLLPAD
TRFKLLVFTGDIGAPEQKRKVDELAKHLERKESFLGRFGEQKNDVFDVFAVCLSRKEDVVYTDVPEV
FRPHWSKVLLDDVDTTGKVGAGVYEKFGIGSEGAIVVVRPDGYVGIVAPLDDVSVLESYFAGFFAKS

Figure 38: Identification of the mass fingerprint for the tryptic digest of *Gtra*1516

Identification of the protein bound flavin in *Gtra*1516

The enzyme-bound flavin was dissociated from the enzyme by heating the protein at 100 °C for 5 min, then the denatured protein was removed by centrifugation at 62000 x *g* for 10 min. The flavin in the resulting yellow supernatant was identified by thin layer chromatography on silica gel. Samples of FAD, FMN and enzyme supernatant were spotted on the plate, which was developed with n-butanol-acetic acid-water (12: 3: 5) or $Na_2HPO_4 \cdot 4H_2O$ (5% in water) as solvent. In both solvent systems the sample chromotographed as a single yellow spot. The Rf value of each spot was calculated in the first developing solvent, $Rf_{FAD} = 0.16$, $Rf_{sample} = 0.16$, $Rf_{FMN} = 0.24$, and in the second developing solvent, $Rf_{FAD} = 0.627$, $Rf_{sample} = 0.627$, $Rf_{FMN} = 0.75$. The Rf values of the yellow compound from boiled enzyme supernatant were the same as those of pure FAD, but quite different from FMN. Thus the protein-bound flavin of Gtra1516 was identified as FAD, not FMN.

Absorption spectrum of purified *Gtra*1516

The UV-vis absorption spectrum of purified *Gtra*1516 was recorded using a Cary Bio50 spectrophotometer. The visible spectrum of this enzyme is typical of flavin-containing proteins *(7)*, exhibiting maxima at approximately 370 nm and 443 nm and minima at 320 nm and 400 nm, respectively, as well as a shoulder at 480 nm (Figure 39).

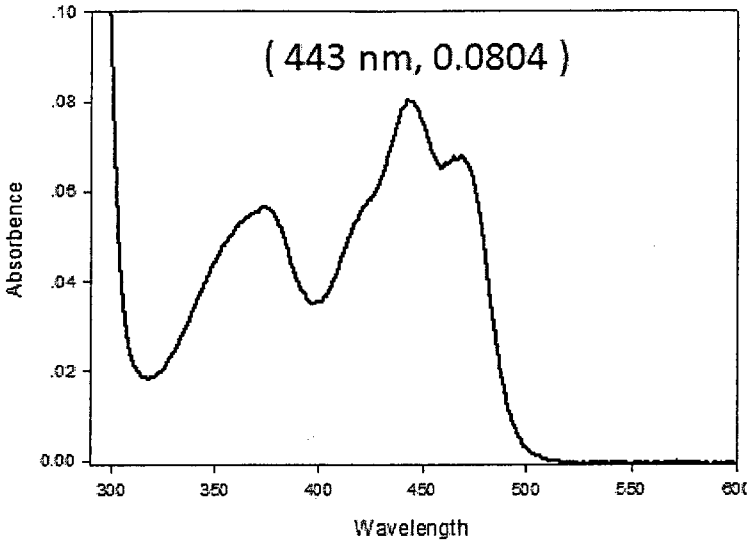


Figure 39: Absorption pectrum of purified *Gtra*1516

The enzyme (8 μM) was in 50 mM Tris·Cl buffer, pH, 7.5, in a 1 cm light path cuvette in the Cary Bio50 spectrophotometer.

Quantitation of protein-bound FAD in *Gtra*1516

The amount of FAD bound to the purified protein was estimated by spectral measurements. The spectrum of diluted native protein was recorded and showed an absorbance of 0.0804 at 443 nm (Figure 40). Then the FAD presented in the diluted protein was dissociated from the protein by heating at 100 °C for 5 min. The spectrum of the boiled supernatant showed an absorbance of 0.078 at 450 nm. The extinction coefficient of the FAD in the boiled supernatant is 11300, the same as that of free FAD *(5)*. Protein concentration was estimated using the BCA assay and the subunit molecular weight of 66600 Da was used to calculate molar concentrations. The amount of FAD can be determined in the boiled sample, using the absorbance and extinction coefficient. The concentration of protein in the sample also can be calculated. The extinction coefficient at 450 can be obtained by taking the absorbance of the enzyme at 450 nm dividing by the

absorbance of the boiled sample at 450 nm, and multiplying by 11700. The mol FAD bound per mol enzyme was calculated to be approximately 1. Thus, *Gtra*1516 contains 1 FAD per-subunit.

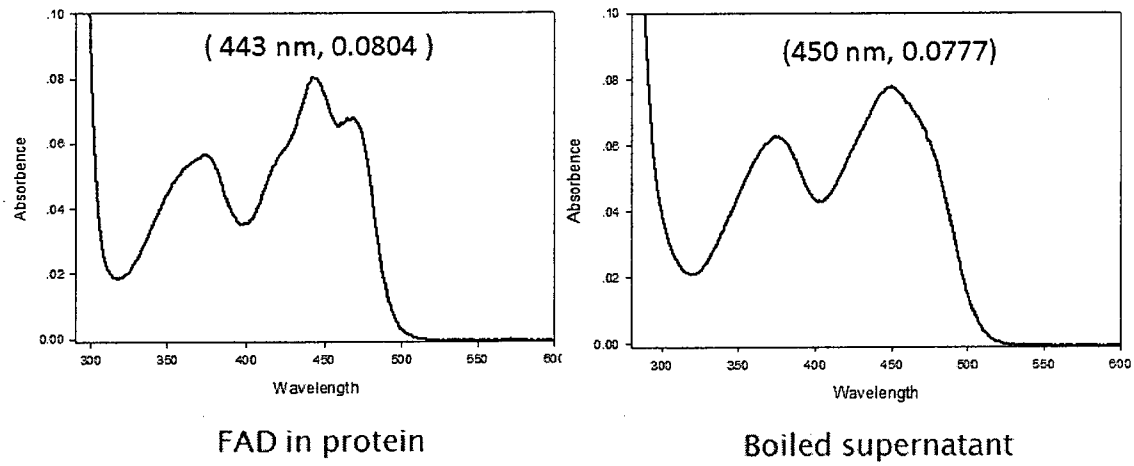


Figure 40: Absorption spectrum of native protein and boiled supernatant of *Gtra*1516

The protein (8 μM) was diluted in 50 mM Tris·Cl buffer, pH, 7.5. To the left is the spectrum of native protein, and to the right is the spectrum of the boiled supernatant.

Estimation of molecular weight of *Gtra*1516 by gel filtration chromatography

The molecular weight of purified Gtra1516 was estimated by gel filtration on a Superdex S-200 column (Pharmacia, 40 cm x 2 cm) equilibrated with 0.1 M NaCl in 50 mM Tris·Cl buffer, pH 7.5. Molecular weight standards were loaded onto the column separately and a standard curve was obtained by plotting K_{av} vs. log MW. For 100 μl of the purified protein was loaded on the column, the K_{av} value of the purified protein, log MW was calculated using the standard curve. The molecular weight of Gtra1516 was estimated to be 138000 Da which indicates that this protein is a dimer (Figure 41).

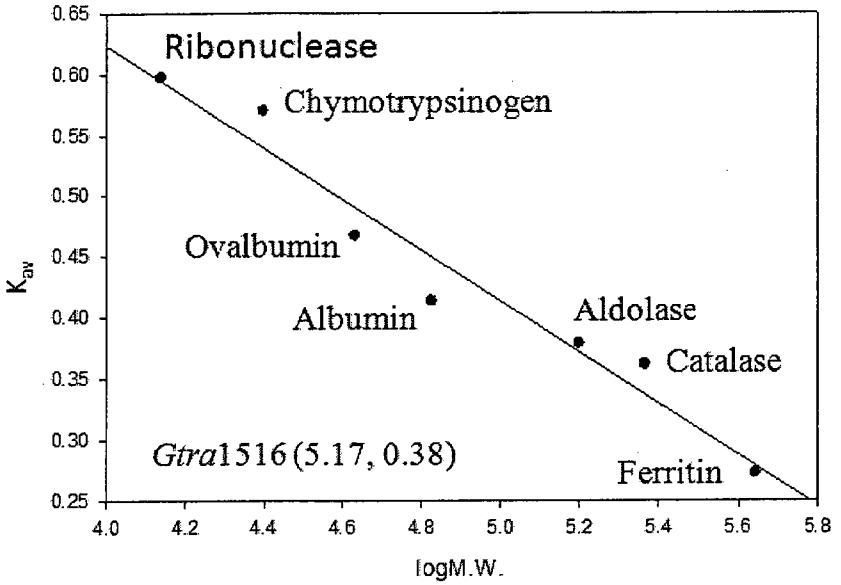


Figure 41: Estimation of the molecular weight of Gtra1516 by gel filtration

The standard protein used were Ribonulease (13,700 Da), Chymotrypsinogen (25000 Da), Ovalbumin (43000 Da), Albumin (67000 Da), Aldolase (158000 Da), Catalase (232000 Da), Ferritin (440000 Da). The equation obtain from K_{av} vs. log MW was y = - 0.2093 x + 1.4601. The position data for *Gtra*1516 was (5.17, 0.38).

Substrate specificity of *Gtra*1516

The activity of the enzyme in stimulating NADPH oxidation was determined for a number of aromatic compounds using the assay conditions outlined in "Materials and Methods" (Table 7). In the absence of aromatic substrate, NADPH oxidase activity of the purified enzyme was almost negligible under the standard assay conditions. The highest rate of NADPH oxidation was observed using phenol as the substrate, followed by resorcinol and p-methylphenol. The anzyme was not active with NADH.

Substrate	Relative activity (%)
Phenol	100
Resorcinol	41.9
p-Methylphenol	21.2
Catechol	14.7
m-Methylphenol	13.9
phloroglucinol	6.6
o-Methylphenol	0.7
Benzentriol	0

Table 7: Substrate specificity of purified Gtra1516

The enzyme activity was determined in 50 mM Tris·Cl buffer, pH, 7.5 containing 0.2 µmol NADPH, 0.15 µmol of the respective substrate and 53.7 µg of the purified enzyme in a total volume of 1 ml. The oxidation of NADPH was determined at 340 nm. The enzyme activity with phenol (2 U/mg) was taken as 100%. The values shown in the Table are averages of triplicates.

Product identification for *Gtra*1516 by TLC and HPLC

In order to identify products produced from the aromatic compounds listed in Table 7, reactions were carried out in 50 mM Tris·Cl buffer, pH 7.5, containing 600 µM NADPH, 150 µM of the aromatic compound and 0.5 mg enzyme in a total volume of 1 ml. Using phenol, *p*-methylphenol or resorcinol, the reaction mixture was incubated for 10 min and then acidified with HCl. The precipitate formed was removed by centrifugation and the supernatant solution was extracted three times with 2 ml portions of ethyl acetate. The extracts were combined and dried over $MgSO_4$. The supernatant was decanted and the ethyl acetate was evaporated under a stream of nitrogen. The residue was dissolved in ethyl acetate and analyzed by TLC. Phenol, resorcinol, *p*-

methylphenol, catechol, salicylate and hydroxyquinol were spotted on TLC plates as standards. Using phenol or *p*-methylphenol as a substrate, no product was observed (data not shown). Using resorcinol as a substrate, a spot was observed and the Rf value of the reaction product was 0.82, almost the same as that of pure hydroxyquinol (Rf = 0.824). The reactions were carried out again and applied to HPLC. In HPLC (Figure 42), using resorcinol as a substrate, hydroxyquinol was detected, while using phenol as a substrate, very little catechol was present.

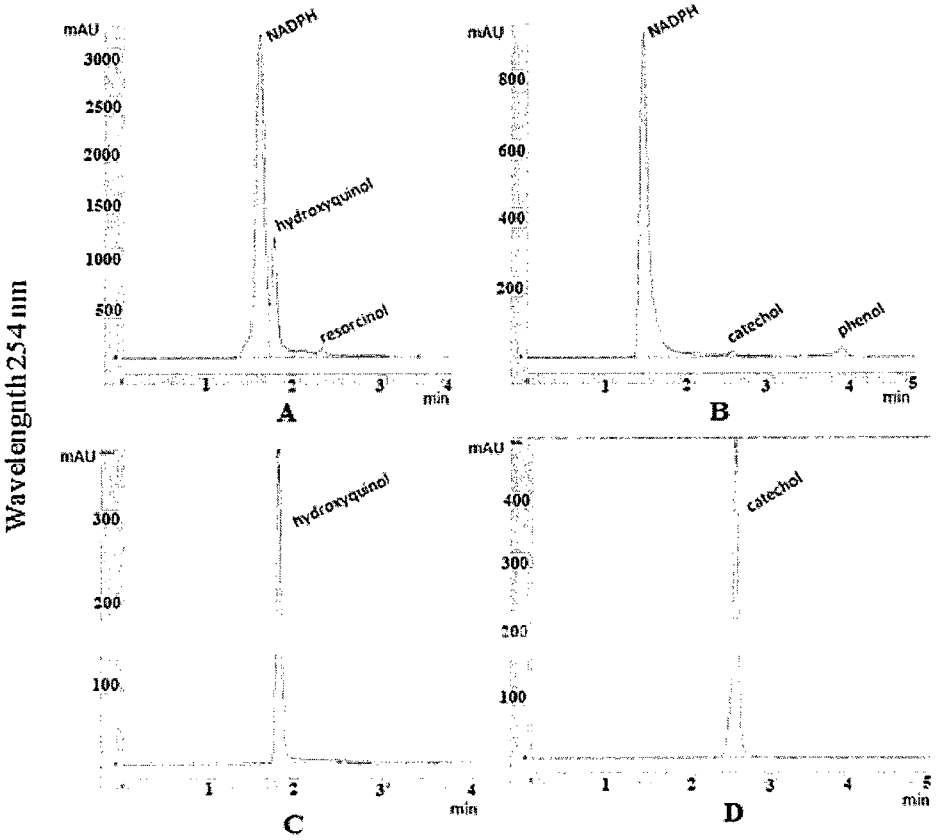


Figure 42: Product identification of *Gtra1516* by HPLC

The reaction mixtures were subjected to HPLC without extracting with ethyl acetate. (A) is the reaction mixture using resorcinol as the substrate. (B) is the reaction mixture using phenol as the substrate. (C) is hydroxyquinol standard (the same concentration as resorcinol added in A). (D) is catechol standard (the same concentration as phenol added in D). Using resorcinol as a substrate, the product was identified as hydroxyquinol. Using phenol as the substrate, the formation of catechol was negligible.

Oxygen consumption assays for *Gtra*1516

Oxygen consumption in the presence of different aromatic substrates was measured using an oxygen electrode from Hansatech. Reactions were run with limiting aromatic compound, to permit observation of the stoichiometry between oxygen consumption and hydrogen peroxide production in the presence of phenol, *p*-methylphenol or resorcinol. Catalase was added after oxygen consumption ceased, to determine whether hydrogen peroxide was produced. The addition of catalase to the reaction mixture containing phenol resulted in the return of almost one-half of the oxygen consumed (Figure 43). However, catalase had little effect upon oxygen "return" in the reaction mixture containing resorcinol. The data in Table 8 indicate that hydrogen peroxide is formed in the phenol and *p*-methylphenol mediated reaction, with a 1:1 stoichiometry observed between NADPH and H_2O_2. No product was observed using *p*-methylphenol as a substrate. Very little product was observed in HPLC analysis of the reaction mixtures containing phenol (Figure 42), also indicating that oxygen was mostly diverted to hydrogen peroxide production. Thus, phenol appears to be mainly a non-substrate effector. Much less H_2O_2 was formed in the resorcinol mediated reaction, from which hydroxyquinol was detected using TLC and HPLC. Resorcinol thus appears to be a much better substrate for the enzyme, with relatively little uncoupling.

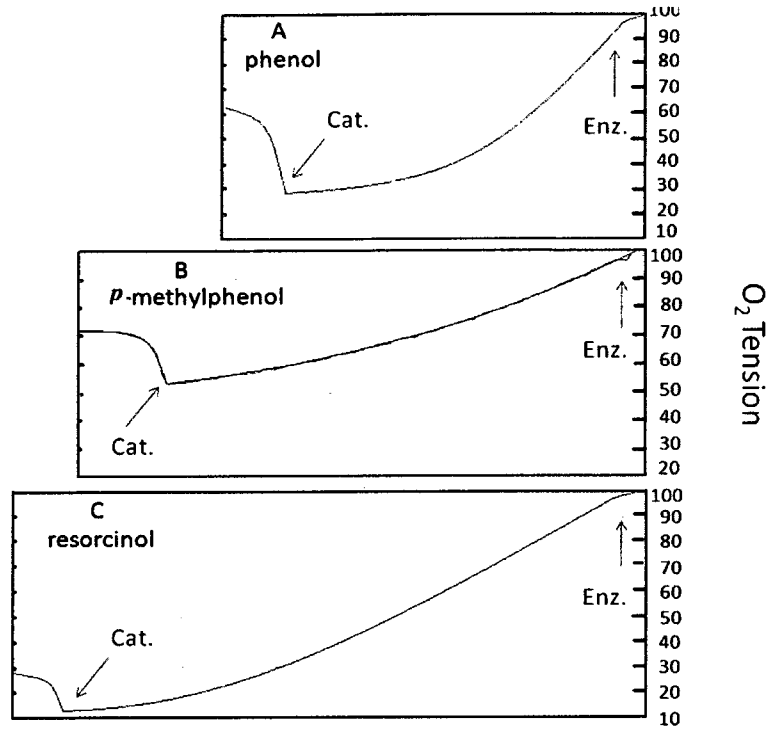


Figure 43: Consumption of O_2 by reaction mixture of *Gtra*1516

O_2 consumption was measured with a Gilson model KM oxygraph equipped with a Clark oxygen electrode. The O_2 electrode traces are recorded backwards. The reaction chamber was hooked up to a 20 °C water bath and contained 300 μM NADPH, 0.8 μM of the purified protein and 100 μM of substrate in 50 mM Tris·Cl buffer, pH 7.5 with a total volume of 1 ml. Reactions were initiated by addition of phenol (A), 4-methylphenol (B) or resorcinol (C). Then 200 units of catalase (X) were added as shown by the arrows after reactions were almost complete.

Table 8: Comparison of the relative activity and the ratio of hydrogen peroxide produced to oxygen consumed for different substrates of *Gtra1516*

Substrate	**Relative enzyme activity**	**H_2O_2 / O_2**
phenol	100	1/1
4-methylphenol	21.2	1/1
resorcinol	41.9	~0.25/1

The data showed in the Table are averages of replicates.

Substrate binding study for *Gtra*1516

The visible and fluorescence spectra of enzyme-bound FAD were used to monitor aromatic substrate binding to the enzyme. As shown in Figure 44 and Figure **45**, tighter binding of resorcinol relative to phenol was observed: for phenol, no sign of saturation was observed, even at 4 mM phenol while for resorcinol the data were fitted with a Kd of 1.6 mM.

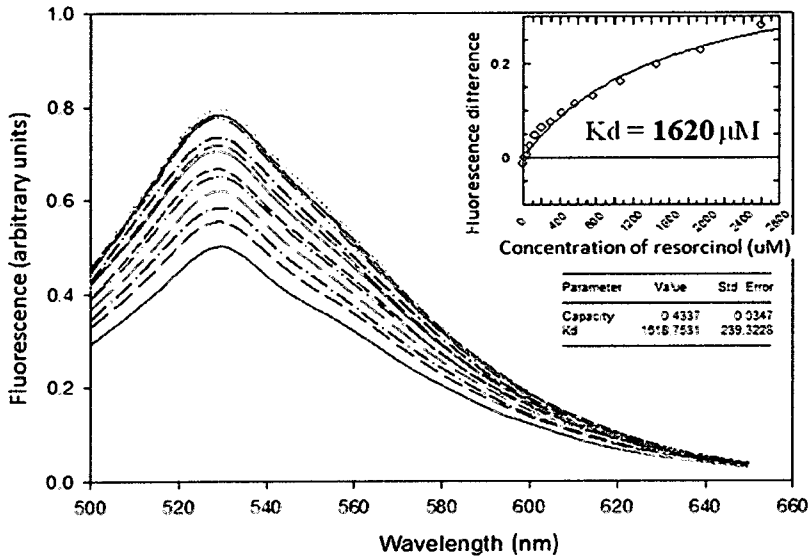


Figure 44: Resorcinol binding to *Gtra*1516

Binding study was carried out using Fluorescence spectroscopy. Samples were placed in 500 µl cuvette and data were collected by titrating 5 µM enzyme with gradually increasing amount of resorcinol. The inset shows the ΔAbsorbance vs. total resorcinol concentration. K_d value were analyzed by a computerized, weighted nonlinear regression method. (Fluorimeter setting is described in "Materials and Methods")

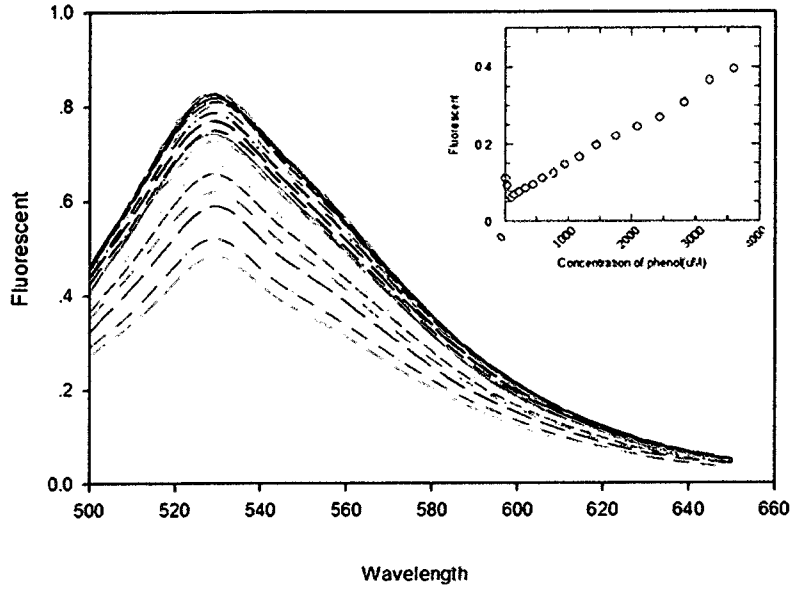


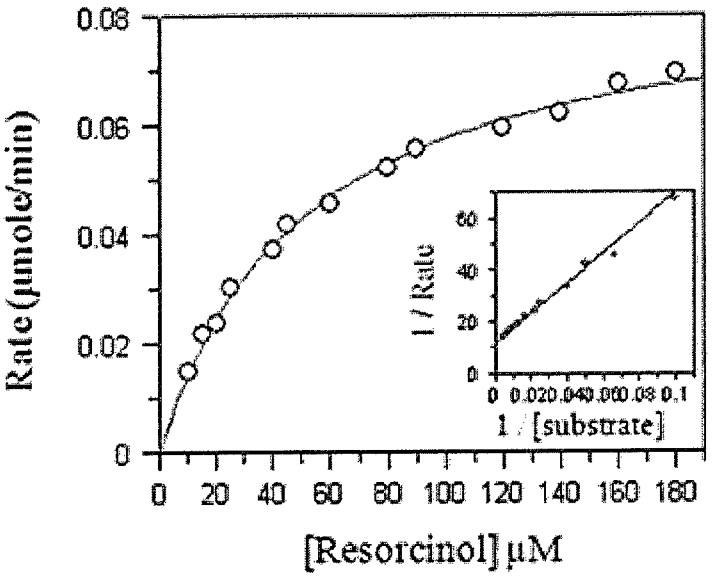
Figure 45: Phenol binding to *Gtra*1516

Binding study was carried out using fluorescence spectroscopy. Samples were placed in 500 µl cuvette and data were collected by titrating 5 µM enzyme with gradually increased amount of phenol. The inset shows the Δabsorbance vs. total phenol concentration. K_d value were analyzed by a computerized, weighted nonlinear regression method. (Fluorimeter setting is described in the "Materials and Methods")

Resorcinol hydroxylase steady state kinetics

The V_{max} and K_M values for resorcinol were estimated for *Gtra*1516. Initial rates observed with different concentrations of resorcinol were corrected for the rates in the absence of aromatic substrate and A_{340}/min values were converted to µmole/min using an extinction coefficient of 6220 $cm^{-1}M^{-1}$. Data were then fitted with the Michaelis-Menten equation using Grafit v4.0 software (Erithacus). With Grafit, Lineweaver-Burk plots were constructed from non-linear regression fitting of the data. The apparent V_{max} and K_m obtained from the fitted parameters were 0.09 µmoles/min and 50.6 µM for resorcinol

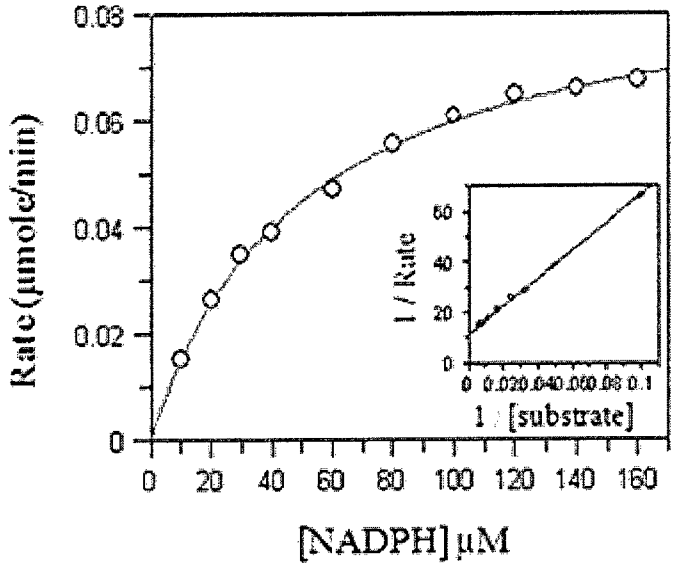
binding (Figure 46) and 0.09 μmoles/ min and 49.7 μM for NADPH binding (Figure 47), respectively.



Parameter	Value	Std. Error
Vmax (μmoles/min)	0.09	0.0027
Km (μM)	50.6	2.5322

Figure 46: Steady state kinetics of *Gtra*1516

Reactions were initiated by the addition of different amounts of resorcinol to 50 mM Tris·Cl buffer, pH 7.5, containing 160 μM NADPH and 0.48 μM enzyme at room temperature.



Parameter	Value	Std. Error
Vmax (µmoles/min)	0.09	0.0069
Km (µM)	49.7	35.4471

Figure 47: Steady state kinetics of *Gtra*1516

Reactions were initiated by the addition of different amounts of NADPH to 50 mM Tris·Cl buffer, pH 7.5, containing 160 µM resorcinol (saturated) and 0.56 µM enzyme at room temperature.

Characterization results for *Gtra*1270

*Gtra*1270 is a putative catechol-1, 2-dioxygenase or hydroxyquinol dioxygenase, catalyzing the ring cleavage of phenol and hydroxyquinol to form *cis,cis*-muconic acid or maleylacetate, respectively.

Optimization of expression conditions for *Gtra*1270A

Small-scale induced cultures of *E. coli* transformed with expression plasmid harbouring *Gtra*1270 were grown under different conditions to maximize expression of

this protein. Figure 48 shows the expression level of the proteins under different conditions. *Gtra1270* had the highest expression level after induction for 3 h at 37°C. However, instead of being expressed in a soluble form as were *Ledo*1986 and *Gtra*1516, Gtra1270 was mainly expressed in an insoluble form Figure 48.

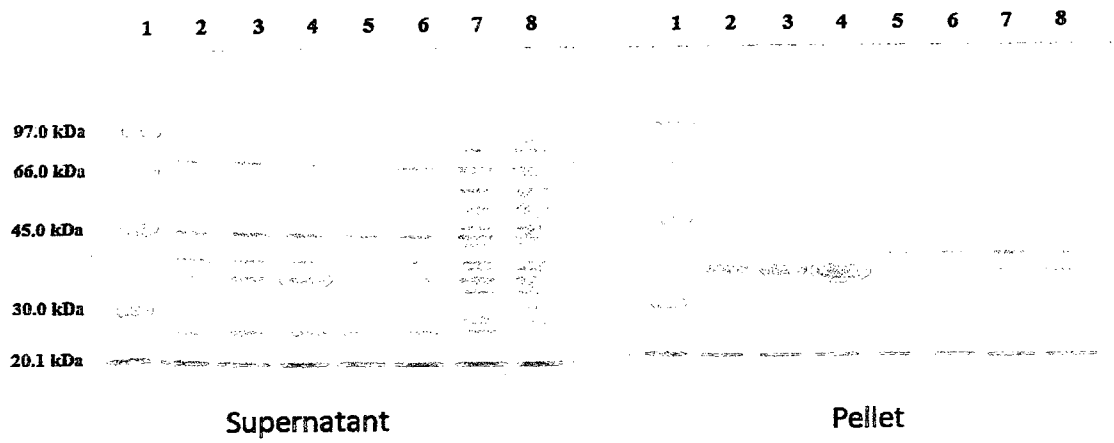


Figure 48: SDS-PAGE gel (12%) demonstrating the optimum expression condition for *Gtra*1270A

(Left) Expression levels in the crude extracts obtained under different induction conditions. Low molecular weight standards (lane 1). Crude extracts from cell induced at 37 °C for 1, 2, and 3 h (lanes 2, 3 and 4). Crude extracts from cells induced at 20 °C for 1, 3, 6 h and overnight (lanes 5, 6, 7 and 8). (Right) Expression levels of protein in the pellet obtained from different induction condition. Low molecular weight standards (lane 1). Pellet resuspension from cells induced at 37 °C for 1, 2, and 3 h (lanes 2, 3 and 4). Pellet resuspension from cells induced at 20 °C for 1, 3, 6 h and overnight (lane 5, 6, 7 and 8).

ESI-Q-TOF for *Gtra*1270A

The identity of the major band identified as *Gtra*1270A in the gels shown in Figure 48 was confirmed by mass spectrometry. Trypsin in-gel digestion followed by ESI-Q-TOF MS was conducted for the circled protein bands in Figure 48 (lane 4,

supernatant, and lane 4, pellet). Figure 49 shows the spectrum for digestion of the band of protein observed in the pellet fraction. Peptide fragments observed in the spectrum are underlined in Figure 50, and cover 16% of the sequence. None of the peptide masses from the digest of the band of protein observed in the supernatant matched the peptide masses expected from Gtra1270A (ex. Figure 51).

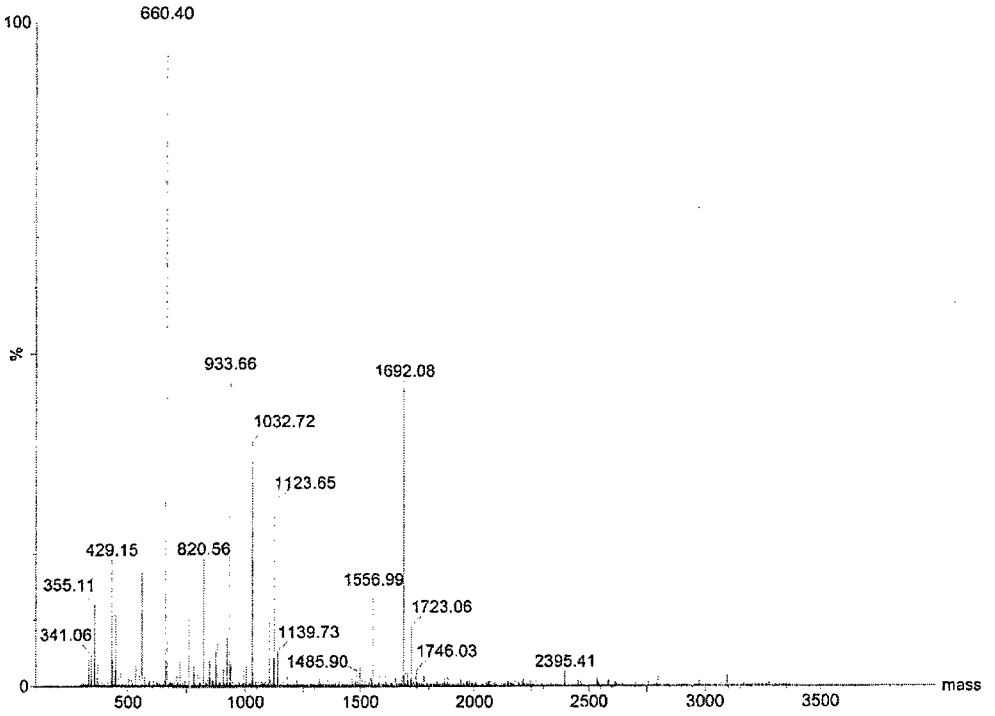


Figure 49: ESI-Q-TOF spectrum for the peptides of *Gtra*1270 (Figure 48, pellet, lane 4)

MNTIQFLTRTGQICTPIRQEFILLSDVLGISALV
DALNNPPVSGGTESSVLGPFFTEDAPDVNNGDSI
ASEGKGQYMYVEGRVIDTHGKPVPNALIETWETD
EYGFYDTQYADRSKPDCRGRLRTDKDGRYGYRAV
VPVAYPIPGDGPVGDLLLMLNRHNMRPNHLHMMI
EAPGYQKLTTAFYPEGDEWLASDAVFGVKKSLVV
TLKDVDNEQEARKRGFPKGSHFKLLEHDLVLVPE
AESKAARERQYARERHAVNRSNEIQA

Figure 50: Identification of the mass fingerprint for the tryptic digest of *Gtra*1270

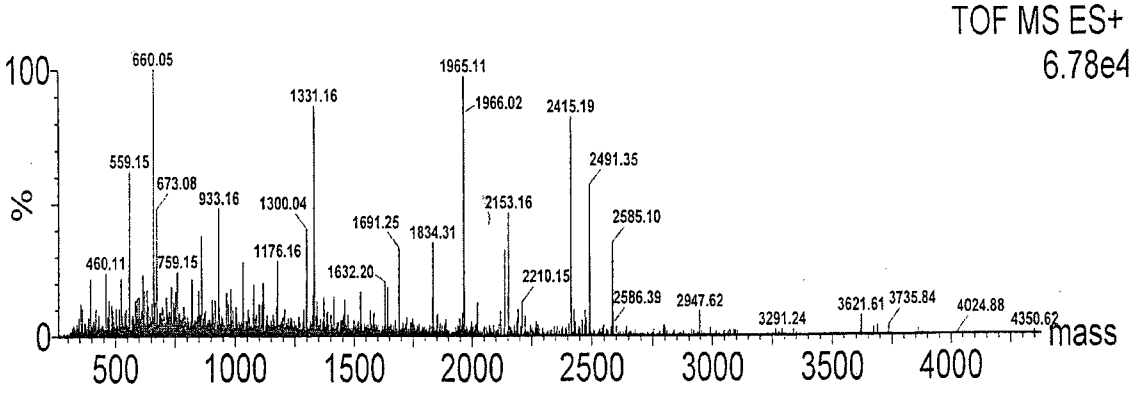


Figure 51: ESI-Q-TOF spectrum for the peptides of *Gtra*1270 (Figure 48, supernatant, lane 4)

Characterization of *Gtra*1270

Since Gtra1270A is expressed in an insoluble form, it is possible that this may be due to the missing N-terminal part of the predicted conserved domain suggested by the sequence alignments (Figure 17). With an N -terminal extension, a BLAST analysing of *Gtra*1270B (Figure 18) shows that this target gene includes the full alignment sequence with the predicted conserved domain.

Optimization of expression conditions for *Gtra*1270B

Small scale induced cultures of *E. coli* cells transformed with expression plasmid harbouring *Gtra*1270B were grown for different lengths of time at various temperatures to determine the best conditions for protein expression (Figure 52). Compared to the pellet, appreciable amounts of expressed protein were detected in the crude extract. Expression of a protein with a molecular mass close to that of *Gtra1270*B was highest after inducing for 3 h at 37 °C (Figure 52, lane 4, supernatant). Thus, by extending the N-

terminal sequence of the protein, the protein appeared to be successfully expressed in a soluble form.

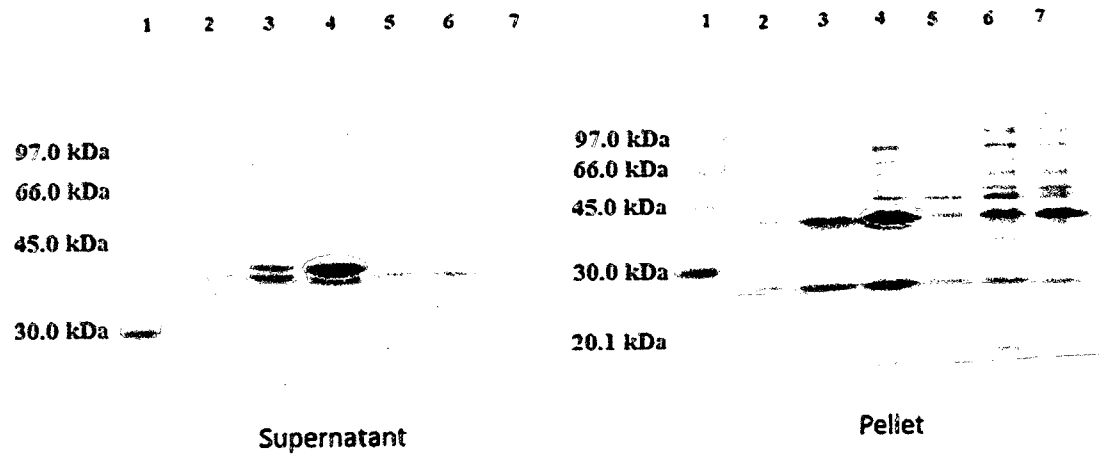


Figure 52: SDS-PAGE gel (12%) demonstrating the optimum expression condition for *Gtra*1270

(Left) Expression levels in the crude extracts from *E. coli* transformants harbouring Gtra1270 obtained under different induction condition. Low molecular weight standards (lane 1). Crude extract from cells induced at 37 °C for 1, 2, and 3 h (lane 2, 3 and 4). Crude extract from cells induced at 20 °C for 3, 6 h and overnight (lane 5, 6 and 7). (Right) Expression levels of protein in the pellet obtained under different induction conditions. Low molecular weight standards (lane 1). Pellet resuspension from cells induced at 37 °C for 1, 2, and 3 h (lane 2, 3 and 4). Pellet resuspension from cells induced at 20 °C for 3, 6 h and overnight (lane 5, 6 and 7).

ESI-Q-TOF for *Gtra*1270B

The identity of the putative Gtra1270B band was confirmed by mass spectrometry. Trypsin in-gel digestion followed by ESI-Q-TOF MS was conducted for the protein bands circled in Figure 52 (lane 4, supernatant and lane 4, pellet). Figure 53 shows the spectrum obtained for the digest of the protein band observed in the supernatant. Peptide

fragments observed in the spectrum and they are underlined in Figure 54, and cover 29% of the sequence.

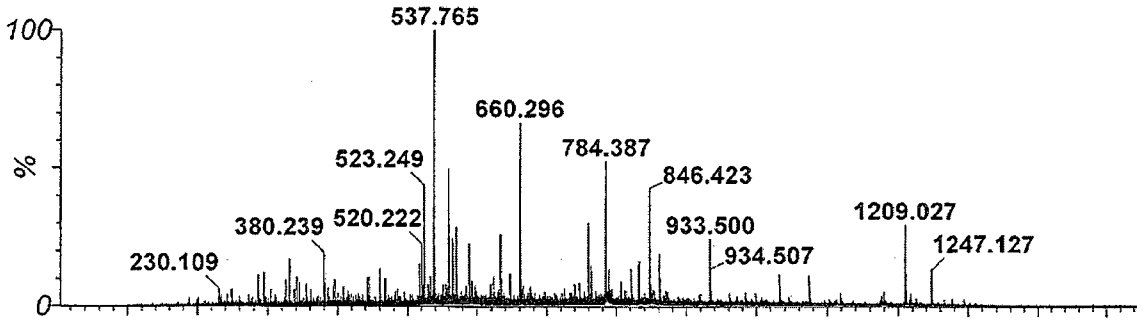


Figure 53: ESI-Q-TOF spectrum for the peptides of *Gtra*1270B

M P N L D L P Y P D R P E L I T E N L L K L T N L I T D E R K K Y I F K N L I
T H I H Q F I N E T S I T T D E W M N T I Q F L T R T G Q I C T P I R Q E F I
L L S D V L G I S A L V D A L N N P P V S G G T E S S V L G P F F T E D A P D
V N N G D S I A S E G K G Q Y M Y V E G R V I D T H G K P V P N A L I E T W E
T D E Y G F Y D T Q Y A D R S K P D C R G R L R T D K D G R Y G Y R A V V P V
A Y P I P G D G P V G D L L L M L N R H N M R P N H L H M M I E A P G Y Q K L
T T A F Y P E G D E W L A S D A V F G V K K S L V V T L K D V D N E Q E A R K
R G F P K G S H F K L L E H D L V L V P E A E S K A A R E Q Y A R E H A V N R
S N E I Q A

Figure 54: Identification of the mass fingerprint for the tryptic digest of *Gtra*1270B

Purification of *Gtra*1270B

*Gtra*1270B was successfully overexpressed in *E.coli* (DE3)*lacI* cells and purified from 6 L of culture using four chromatographic steps as described in "Materials and Methods". After preparation of crude extract, Fast-flow DEAE-Sepharose

chromatography, High Performance phenyl-Sepharose chromatography, octyl-Sepharose chromatography and gel-filtration (S-300) chromatography were applied in succession to purify the target protein. Since the crude extract was not active in oxygenase assays with hydroxyquinol or catechol, through all chromatography steps, the target protein was located by SDS-PAGE. The peak fractions were pooled and concentrated prior to the next step. Figure 55 demonstrates the protein purity observed after each purification step. The subunit molecular weight of *Gtra*1270B was calculated to be 35900 Da from the amino acid sequence.

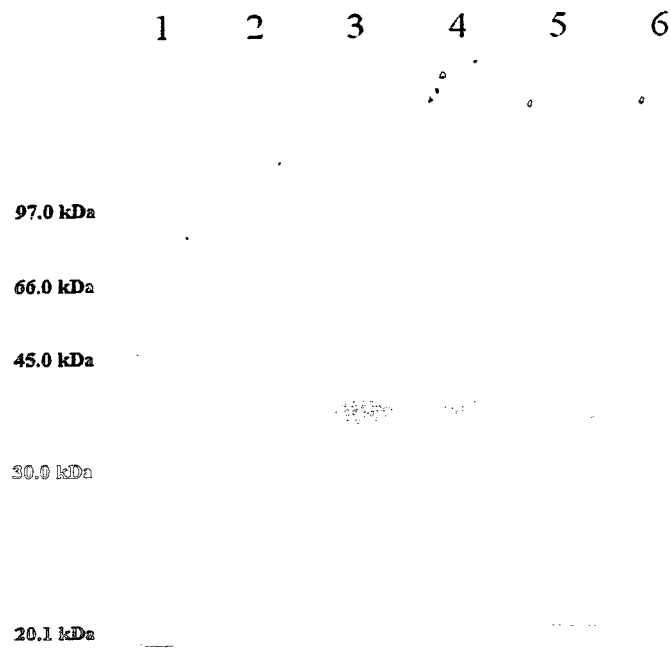


Figure 55: SDS-PAGE gel (12%) demonstrating the purity of recombinant *Gtra*1270B after different purification steps

Low molecular weight standards (lane 1). Crude extract (lane 2); Pooled fractions from Fast-Flow DEAE column (lane 3); Pooled fractions from phenyl-Sepharose column (lane 4); Pooled fractions from octyl-Sepharose column (lane 5); Pooled fractions from S-300 gel-filtration column (lane 6).

Estimation of molecular weight of *Gtra*1270B by gel filtration chromatography

The molecular weight of purified *Gtra*1270B was estimated by gel filtration on a Superdex S-200 (Pharmacia, 40 cm x 2 cm) column equilibrated with 0.1 M NaCl in 50 mM Tris·Cl buffer, pH 7.5. Molecular weight standards were loaded onto the column separately and a standard curve was obtained by plotting K_{av} vs. log MW (Figure 56). Purified protein (100 µl) was loaded on the column, and eluted in two peaks (Figure 57). From the K_{av} value of the purified protein, log MW was calculated using the standard curve. The molecular weights of the two peaks of *Gtra*1270B were estimated to be 143300 Da and 215400 Da which indicated that this protein appeared to aggregate.

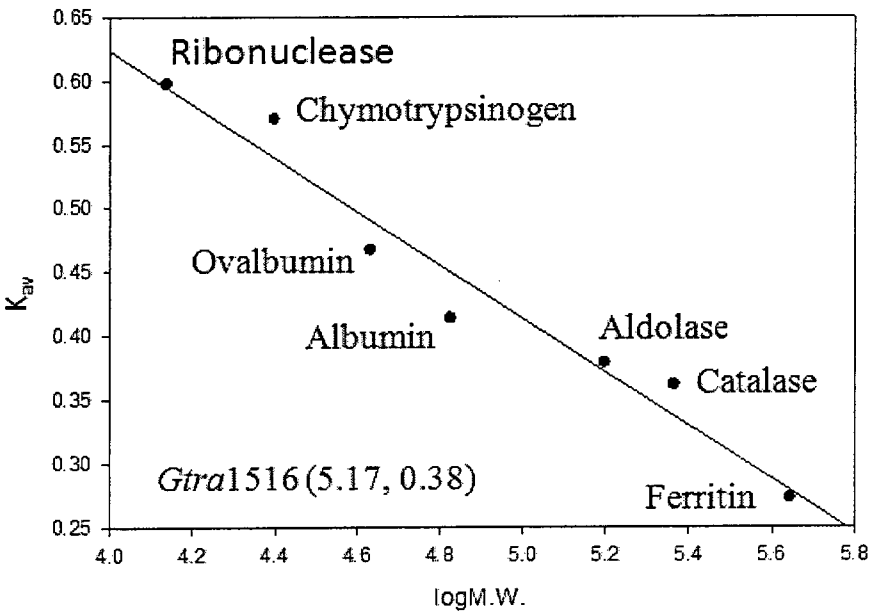


Figure 56: Estimation of the molecular weight of *Gtra*1270B by gel filtration chromatography

The standard protein used were Ribonulease (13700 Da), Chymotrypsinogen (25000 Da), Ovalbumin (43000 Da), Albumin (67000 Da), Aldolase (158000 Da), Catalase (232000 Da), Ferritin (440000 Da). The equation obtain from K_{av} vs. log MW was y = - 0.2093 x + 1.4601.

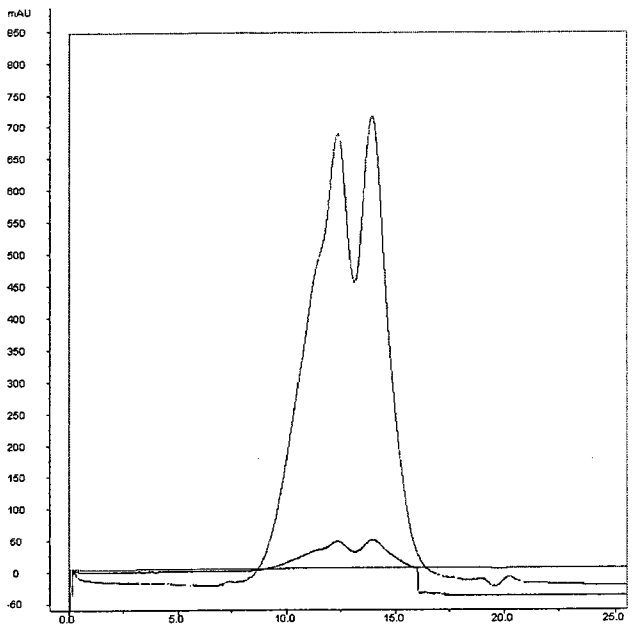


Figure 57: Gel filtration traces with Gtra1270

Attempted activation of *Gtra*1270B by addition of ion under aerobic and anaerobic condition

The result of iron quantitation (Materials and Methods) indicated that there was no iron in the purified protein, so the activity assay was carried out with the reconstitution of iron. The activity of Gtra1270B was measured spectrophotometrically at either 260 or 245 nm using catechol or hydroxyquinol as the substrate. Reactions were carried out under aerobic or anaerobic conditions in the presence of Fe^{2+}, as described in "Materials and Methods", however no activity was observed.

DISCUSSION

Fungi are able to degrade a wide array of different molecules that higher organisms are incapable of metabolizing. Much of the degradative power of fungi is associated with secreted hydrolases that attack plant-based macromolecules such as cellulose, starch, and hemicelluloses. Fungi also produce a wide array of extracellular proteases and peptidases, many of which have been exploited for industrial purposes *(86)*. Secreted oxidative enzymes, such as peroxidases and laccases, play important roles in the degradation of lignin, a highly cross-linked network of aromatic compounds that is very resistant to hydrolysis *(29)*. Free radicals and active oxygen species generated by these enzymes attack the linkages in lignin and break it down into simpler subunits. Oxidative enzymes, such as cellobiose dehydrogenase, also play a role in the degradation of polysaccharides.

Considerable attention has been paid to the role of extracellular fungal oxidative enzymes, such as lignin peroxidase and laccase, in dye decolourization and in the degradation and mineralization of toxic aromatic compounds such as chlorinated aromatics *(86)*. Less well-studied are fungal intracellular pathways for the use of aromatic compounds as sole sources of carbon and energy. Decades ago, pathways were described in fungi such as *Aspergillus niger, Neurospora crassa and Rhodotorula sp.* for the degradation of lignin-derived aromatics and hydroxybenzoates *(75)*. Although differing in some details, in general the chemistry of these pathways is quite similar to pathways used by bacteria for growth on these compounds. Thus, the aromatic ring is prepared for ring fission by insertion of hydroxyl groups, such that it is then cleaved by

dioxygenases to produce a ring-fission product that is converted in a series of reactions to intermediates of central metabolism.

Although the genetics and molecular biology of bacterial aromatic degradation pathways have been studied extensively, very few corresponding studies have been published about fungal pathways. Nevertheless, in the past few years there has been an explosion of information in the form of sequences of fungal genomes. The Joint Genome Institute (http://www.jgi.doe.gov/) and the Broad Institute (http://www.broadinstitute.org/) are two prominent centers for fungal genomics that run sequencing programs and are repositories of publicly-available fungal genome sequences. There are 207 and 160 fungal genomes, respectively available in the two website. The annotation of these genomes relies to a large extent on sequence comparisons with known genes. In the case of genes encoding enzymes for aromatic degradation pathways, most of the annotation is based on sequence comparisons with bacterial genes. In some cases, the enzymes are similar enough that the interpretation is straightforward. However, in other cases where fungi use a somewhat different strategy than bacteria, the lack of fungal sequence information makes it difficult to predict gene function accurately. For example, Cain et al. *(75)* studied a protocatechuate dioxygenase from *Aspergillus niger*, however, protocatechuate dioxygenase can not be found by searching the JGI sequence of the genome of *Aspergillus niger*.

In order to better understand fungal aromatic degradation pathways, especially the information that is available from fungal genome sequencing programs, it is clear that

more structure-function information is needed. In this study, I expressed genes that appear to encode enzymes from aromatic degradation pathways in fungi for which cDNA was available from the fungal genomics program at Concordia University. Two new flavoprotein hydroxylases were successfully expressed and functional information was obtained. A putative ring cleavage dioxygenase was also expressed, but the enzyme lacked a metal cofactor and was not active in any of the activity assays that were tested.

In the past a few decades a variety of aromatic hydroxylases including many flavoproteins have been studied. Mono- or dioxygenases which catalyze the catabolism of aromatic compounds by introducing hydroxyl groups or cleaving the benzene ring either at the *ortho* or *para* position. Flavoprotein monooxygenases use NAD(P)H as the hydride donor to catalyze the hydroxylation of aromatic compounds in the presence of molecular oxygen. In addition to the flavin binding site, flavoprotein monooxygenases-catalyzed reactions seem to have a common mechanism. Generally, three steps are involved: first, reduction of the flavin by the hydride donor, NAD(P)H, second, the formation of the oxygenating reagent, C4a-flavin hydroperoxide, from the reaction of the reduced flavin and O_2, and binding, orienting, and activating the substrate for its oxygenation by the C4a- hydroperoxide, with release of the second atom of oxygen as water from flavin C4a-hydroxide. Dioxygenases such as catechol-1,2- dioxygenase and hydroxyquinol-1,2-dioxygenase incorporate both atoms of molecular oxygen into the products of the reaction.

*Ledo*1986 and *Gtra*1516 were successfully overexpressed in *E.coli* and purified to homogeneity so that they could be characterized with respect to function. Bioinformatics

analysis suggested that *Ledo*1986 was a putative salicylate hydroxylase which catalyzes the hydroxylation of salicylate to catechol. The specific activity of the purified enzyme was 0.28 units/mg of protein and the yield was 97%. This is considerably lower than the salicylate hydroxylases from *Pseudomonas putida (87)*, which had activity of 10.6 units/mg and 37 units/mg, respectively. *Ledo*1986 was determined to be a monomer with a mass of 48600 Da by SDS polyacrylamide gel electrophoresis and Superdex S-200 chromatography. Similar to the enzymes from *Pseudomonas putida (87)* and *Pseudomonad (5)(54)*, *Ledo*1986 contains one molecule of FAD per subunit as a prosthetic group. Salicylate hydroxylase from *Psedomonas (57)* is also a monomer with one FAD per 57 kDa subunit, while the enzyme from *Pseudomonas putida (87)* was found to be a 91 kDa dimer, also with 1 FAD per subunit.

As with bacterial salicylate hydroxylases, *Ledo*1986 uses a reduced nicotinamide nucleotide cofactor as the external electron donor. The enzyme from *Pseudomonads* uses NADH as an electron donor whereas salicylate hydroxylase from *Pseudomonas putida* uses NADPH. The *L.edodes* enzyme uses NADPH as a hydrogen donor, and was practically inactive with NADH. This is probably due to the sequence feature of the NADPH binding site, where the steric hindrance on binding of the extra phosphate group of NADPH is weaker than that of the NADH-specific enzyme *(88)*. Under the standard assay conditions in the absence of anthranilate the oxidation of NADPH was practically negligible. The oxidation of NADPH could be observed in the presence of some of the derivatives of anthranilate as shown in Table 4.

The substrate specificity studies yield some of the most interesting results for this enzyme. Table 4 shows that a variety of aromatic compounds stimulated NADPH consumption. The two salicylate analogues, anthranilate and 4-chlorosalicylate, exhibited reaction rates 309 and 138% of that of salicylate, respectively. Other compounds were less active. Thus, it appears that *Ledo*1986 was more active with anthranilate than other aromatic compounds, including salicylate. When anthranilate was used as a substrate, the reaction product, 2, 3-dihydroxybenzoate, was identified by TLC and HPLC (Figure 31). On the other hand, although *Ledo*1986 was active in NADPH oxidation in the presence of salicylate or 4-chlorosalicylate, no product was observed. Taken together, these results indicate that anthranilate may be a true substrate for *Ledo*1986, while salicylate and 4-chlorosalicylate are "pseudosubstrates" *(7)*. This may explain why *Ledo*1986 catalyzes salicylate oxidation at a much lower rate compared to some other salicylate hydroxylases that have been studied (see above).

In order to further distinguish substrates from psedosubstrates, oxygen consumption assays for *Ledo*1986 were carried out. Since hydroxylase catalyzed reactions consume oxygen, the addition of catalase at the end of the reaction allows detection of any hydrogen peroxide produced in the presence of a non-substrate effector *(7)*. One mole of H_2O_2 forms 0.5 mole of O_2 in the presence of catalase, thus, the ratio of hydrogen peroxide produced to oxygen consumed for salicylate and 4-chlorosalicylate in Table 5 indicates that in both cases all the oxygen was used to produce H_2O_2. In the cases of salicylate or 4-chlorosalicylate bind at the active site of the enzyme, permitting NADPH binding and oxidation. At the same time, the FAD moiety is reduced rapidly by

NADPH. The reduced species reacts with O_2 to form the flavin 4a-hydroperoxide, which then decays to oxidized flavin and H_2O_2 leaving salicylate unchanged *(89)*. Thus, salicylate or 4-chlorosalicylate is a "pseudosubstrate" (non-substrate effectors), but cannot be hydroxylated, so the oxygen utilized decomposes to H_2O_2, and oxygen reduction is considered as "uncoupled" from hydroxylation. The salicylate or 4-chlorosalicylate effect is termed "uncoupling of oxygen activation from hydroxylation" *(7)*.

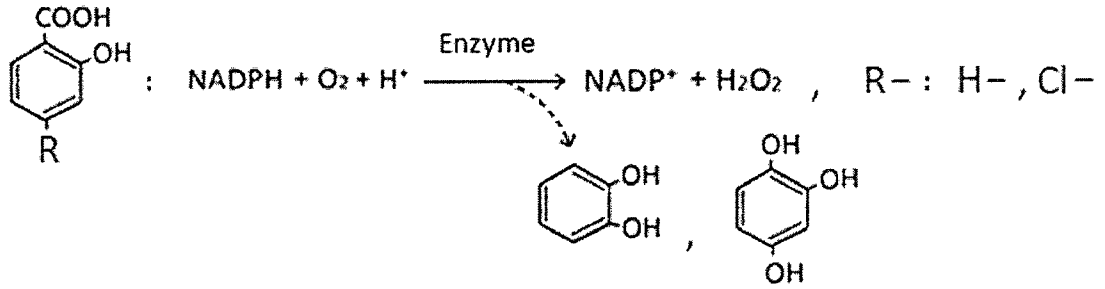


Figure 58: Non-substrate effectors of *Ledo*1986

R- represents the substituent group on the benzene ring. When R- is substituted with H- the substrate is salicylate, when R- is substituted with Cl- the substrate is 4-chlorosalicylate. O_2 is completely incorporated into H_2O_2 in the presence of NADPH. Substrate is unchanged and no aromatic product is formed

The ratio of hydrogen peroxide produced to oxygen consumed (Table 5) indicates that anthranilate is a substrate for *Ledo*1986, since the product, 2, 3-dihydroxybenzoate, was produced and only some hydrogen peroxide was detected. This means that most of the molecular oxygen was used to hydroxylate the aromatic substrate anthranilate, and only a small part of the oxygen was used to produce hydrogen peroxide (Figure 59). Only one other anthranilate hydroxylase, from *T. cutaneum (83, 83)*, has been described and it, like *Ledo*1986 used NADPH as a reducing agent. The enzyme from *T.cutaneum*

catalyzes the double hydroxylation of anthranilate by incorporating one atom of oxygen from O_2 and one from H_2O to form the product, 2,3-dihydroxybenzoate, from anthranilate (Figure 59). Although the *T.cutaneum* enzyme has been studied extensively, its sequence is not known, and this is why it did not turn up in the BLAST search using the sequence of *Ledo*1986.

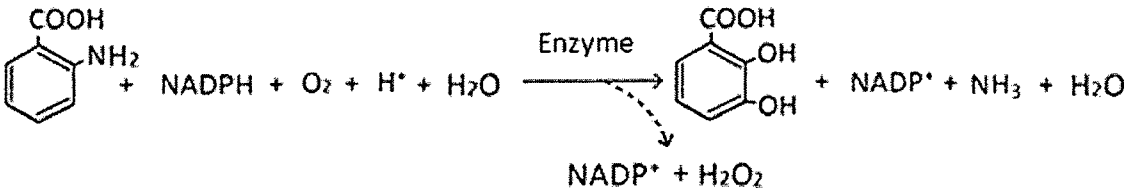


Figure 59: Reaction catalyzed by *Ledo*1986

In order to compare the binding mode of anthranilate and salicylate to the enzyme, spectroscopic probes of binding were employed. The results indicated that the modes of binding of salicylate and anthranilate to the enzyme must be quite different. As shown in Appendix I, binding of anthranilate to *Ledo*1986 causes extensive spectral change, and a Kd value was obtained which indicates that the binding between *Ledo*1986 and anthranilate is relatively tight. Compared with anthranilate binding, the spectral changes observed for salicylate binding to *Ledo*1986 were much less extensive and the interaction was weaker than with anthranilate (Appendix II). This is consistent with the activity results in that it suggests anthranilate is a better substrate.

Talking all of these results together, the enzyme thus appears to be an anthranilate hydroxylase rather than a salicylate hydroxylase. The specific activity of the purified

enzyme with anthranilate was 0.87 units/mg, compared with the anthranilate hydroxylase from *Trichosporon cutaneum (83)* which had a specific activity of 4.4 units/mg. The specific activities of the enzymes are thus similar. Although anthranilate hydroxylases from fungi are known (6), no sequences are identified as such in Genbank. This may explain why the closest BLAST hits were salicylate hydroxylases.

Bioinformatics analysis indicated that *Gtra*1516 is a putative phenol hydroxylase which catalyzes the hydroxylation of phenol to catechol. The sequence is quite similar to phenol hydroxylase from *Trichosporon cutaneum (6).* The specific activity of the purified enzyme was 1.99 units/mg of protein and the yield was 38%, compared with the enzyme from *T. cutaneum (6)*, which had a specific activity of 5.5 units/mg. The specific activity of the enzymes are thus similar. *Gtra*1516 was determined to be a dimer with a mass of 66600 Da by SDS polyacrylamide gel electrophoresis and Superdex S-200 chromatography. Similar to the *T. cutaneum (6)(90)* enzyme, *Gtra*1516 contains 1 molecule of FAD per subunit as a prosthetic group. Phenol hydroxylase from *Trichosporon cutaneum* was determined to contain 1 FAD per 148000 molecular weight of protein.

Similar to salicylate hydroxylase *(5)* and anthranilate hydroxylase *(61)*, phenol hydroxylase is also a flavoprotein monooxygenase, and catalyzes the hydroxylation of phenol by incorporating one atom of molecular oxygen into the substrate, and the other atom into water in the presence of an electron donor. In the case of phenol hydroxylase there is a strict requirement for NADPH as a hydrogen donor *(6).*

Table 7 shows that a number of aromatic compounds-stimulated NADPH consumption. Phenol showed the highest activity, and another analogue, resorcinol exhibited a reaction rate 42% of that with phenol. Other compounds were less reactive. Thus it appears that *Gtra*1516 was more active with phenol than other aromatic compounds. Although *Gtra*1516 was active with NADPH and phenol, the amount of catechol formed was almost negligible (Figure 42). On the other hand, when resorcinol was substituted for phenol, the reaction product hydroxyquinol was identified by HPLC and TLC. These results indicate that resorcinol appears to be a much better substrate for the *Gtra*1516. In contrast, resorcinol is a poor non-substrate effector of the hydroxylase from *Trichosporon cutaneum* but an excellent growth substrate for the living cell *(66)*.

In order to further distinguish substrates from psedosubstrates, oxygen consumption assays for *Gtra*1516 were carried out. The ratio of hydrogen peroxide produced to oxygen consumed is another way to test enzyme activity and substrate specificity since the ratio of 1/1 indicates that all oxygen is used to form hydrogen peroxide and no product should be produced *(7)*. Table 8 indicates that when phenol was used as a substrate, almost all the oxygen was used to incorporate into H_2O_2 with the stoichiometric formation of H_2O_2 and oxidation of NADPH. Thus phenol is a "non-substrate effectors" for *Gtra*1516, stimulating formation of hydrogen peroxide from NADPH and oxygen (Figure 60).

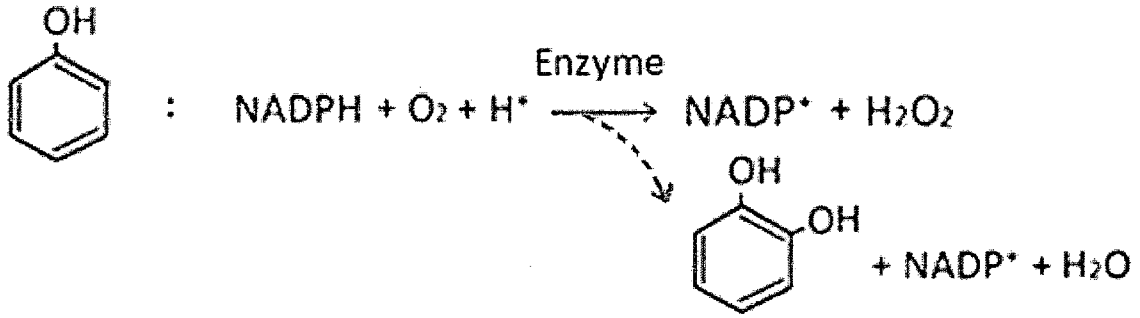


Figure 60: Non-substrate effectors of *Gtra*1516

Most of the O_2 is incorporated into H_2O_2 in the presence of NADPH. Very little oxygen is incorporated into phenol to produce catechol, and this reaction is negligible.

The ratio of hydrogen peroxide produced to oxygen consumed (Table 8) indicates that resorcinol appears to be a much better substrate for *Gtra*1516, since the product hydroxyquinol was produced and only some hydrogen peroxide was detected. This means that most of the molecular oxygen was used to hydroxylate the aromatic substrate, resorcinol, and only a small part of the oxygen was used to produce hydrogen peroxide. Similar to the resorcinol hydroxylase studied by Yoshiyuki *et al (91)*, *Gtra*1516 used NADPH as a reducing agent, catalyzing the hydroxylation of resorcinol by incorporating one atom of oxygen from O_2 and one from H_2O to form the product hydroxyquinol from the substrate (Figure 61).

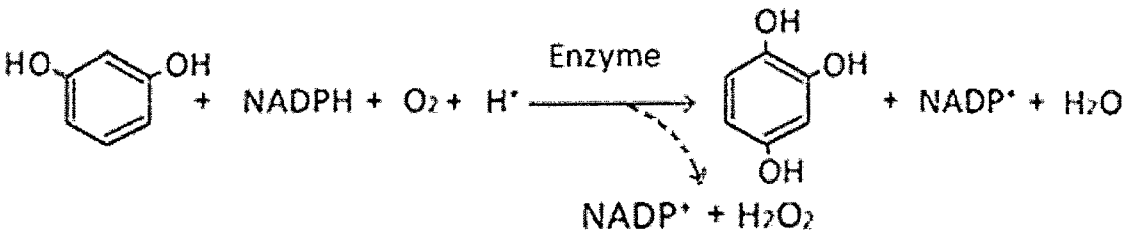


Figure 61: Reaction catalyzed by *Gtra*1516

In order to compare the binding modes of phenol and resorcinol to the enzyme, spectroscopic probes of binding were employed. As shown in Figure 44 and Figure **45**, and the *Kd* for the interaction of phenol with *Gtra*1516 is too low to measure accurately, while the ΔA vs. total concentration of phenol could not be fit. Tighter binding of resorcinol relative to phenol was observed. However, the Kd values both appeared to be greater than 1 mM.

Combined with the previous observations, the enzyme thus appears to be a resorcinol hydroxylase rather than a phenol hydroxylase. Since resorcinol showed relative activity with *Gtra*1516 equal to 42% that of phenol, the specific activity of this enzyme with resorcinol is 0.83units/mg which indicates that this enzyme is relative less active.

BLAST analysis of the sequence of *Gtra*1270 against the GenBank database detected putative conserved domains of *intradiol_dioxygenase superfamily* and *catechol proteo*. Most of the sequences that showing similarity with *Gtra*1270 are hits to genome sequences where the enzyme activity has never actually been demonstrated. As underlined in the alignment list in Figure 15, BLAST searches for *Gtra*1270 from *Gloeophyllum trabeum* showed similarity with catechol dioxygenase and hydroxyquinol 1, 2-dioxygenase. Thus, *Gtra*1270 was predicted to be a fungal catechol-1,2-dioxygenase which catalyzes the ring cleavage of catechol at the *ortho* position and leads to Krebs cycle intermediates.

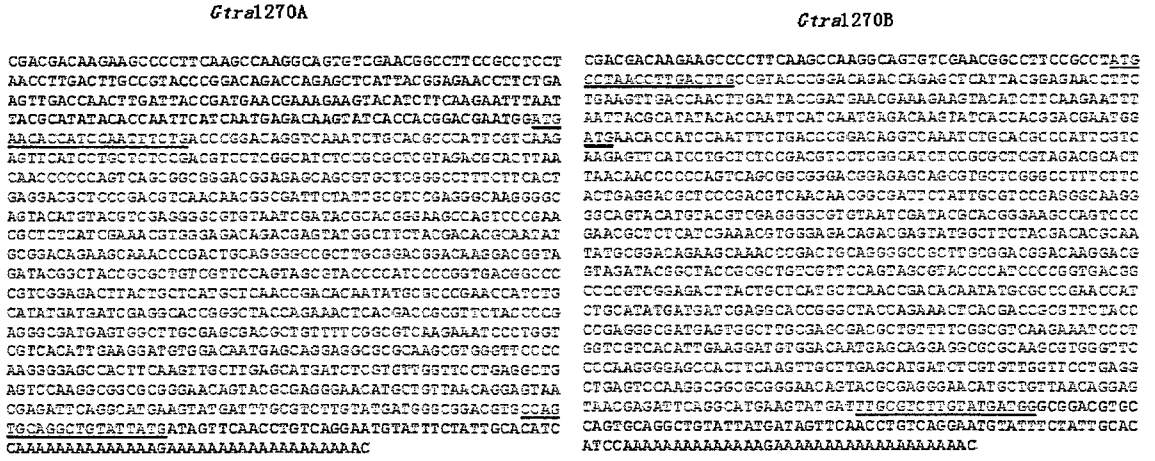


Figure 62 : Primers and target genes of *Gtra*1270A and *Gtra*1270B

At first, a pair of primers was used to amplify the target gene which starts from the predicted first start codon of the template, and the amplified gene was named *Gtra*1270 A (Figure 62, left figure). When expressed in *E.coli*, but the protein was produced in an insoluble form Figure 48, and was not active. Thus the sequence of *Gtra*1270 A was re-examined. From the BLAST result shown in Figure 17, I found that most of the alignments with other Genbank sequences start upstream of the first amino acid of *Gtra*1270 A, which means for *Gtra*1270A some part of the conserved domain shared by all sequences listed may have been left out. Thus a pair of new primers was designed in order to extend the N terminus of *Gtra*1270A. As shown in Figure 16, *Gtra*1270B was extended for 56 codons past the N terminus of *Gtra*1270A, with a new start codon inserted at the beginning. A BLAST search using the *Gtra*1270B sequence showed that this target gene included the full alignment sequence of the listed genes (Figure 18). *Gtra*1270B was then successfully overexpressed, isolated and purified from *E.coli*. Thus, with an extension on the N terminus, *Gtra*1270B was able to be expressed in a soluble form. However, gel filtration results indicated that the purified enzyme (which has a 29

kDa subunit molecular weight) appeared to aggregate, the protein was not active in the assay with a variety of dihydroxylated substrates.

Catechol 1,2-dioxygenase from *Acinetobacter calcoaceticus (69, 79)* is determined to be a nonheme, trivalent, iron-containing enzyme, which catalyzes the cleavage of the aromatic ring of catechol to *cis,cis*-muconate with the incorporation of two atoms of molecular oxygen into the substrate. This enzyme has a red color with broad absorption between 390 and 650 nm. Catechol 1,2-dioxygenase from *Trichosporon cutaneum (21)* was also determined to be an iron containing enzyme showing red color, and the inactive form was colourless. All of the intradiol catechol dioxygenases studied to date contain high-spin Fe (III) in a rhombic environment and exhibit a distinct red-brown colour. The colour is a result of an electronic transition in the visible region, typically centred around 430 - 450 nm. The wavelength and intensity are indicative of a charge-transfer interaction from the co-ordination of tyrosinate side chains to the Fe (III) *(92)*. On the basis of a variety of spectroscopic techniques, the active sites of these enzymes are proposed to consist of a high-spin ferric center coordinated to two tyrosine, two histidine, and a water. The reaction sequence involves initial substrate binding to the ferric center generating an iron(III)- catecholate complex and subsequent attack by dioxygen on the enzyme-substrate complex. Sze and Dagley reported the purification of a hydroxyquinol 1, 2-dioxygenase from *T. cutaneum (55)* grown with 4-hydroxybenzoate. It was shown to be a red ferric iron-containing enzyme, and it was specific for hydroxyquinol.

Purified *Gtra*1270 was colourless, thus an iron assay was carried out to quantify the iron content of *Gtra*1270B. However, no iron was observed and this may be the reason why the enzyme was inactive. The intradiol catechol dioxygenses have Fe (III) at

the active site, whereas the extradiol dioxygenases have Fe (II). In order to activate the enzyme, the protein was reconstituted with Fe^{3+} aerobically or anaerobically, however no activity was observed.

CONCLUSIONS

Significant advances have been made in recent years in the elucidation of the genetic and biochemical bases of aerobic mineralization of aromatic compounds, and detailed knowledge is available of different catabolic pathways and enzymology. Much of these data relate to bacterial pathways, and relatively little has been published on fungal aromatic catabolic enzymes. In this project new information about the properties of two new fungal monooxygenases was obtained, and a putative dioxygenase was successfully expressed though it failed to show activity.

Bioinformatics analysis suggested that *Ledo*1986 is a salicylate hydroxylase. It was successfully overexpressed in *E. coli*, purified in good yield, and found to contain FAD as do many bacterial hydroxylases. Although *Ledo*1986 was active with NADPH and salicylate, no product was observed. Thus, salicylate is a non-substrate effector, stimulating formation of hydrogen peroxide from NADPH and oxygen, as has been reported for a number of other flavoprotein hydroxylases. In contrast, *Ledo*1986 was active with anthranilate, producing 2,3-dihydroxbenzoate as a product. Spectroscopic probes of binding indicated that the modes of binding of salicylate and anthranilate are quite different. Combined with the previous observation, the enzyme thus appears to be an anthranilate hydroxylase rather than a salicylate hydroxylase . Although anthranilate hydroxylases from fungi are known, no sequences are identified as such in Genbank. This may explain why the closest BLAST hits were salicylate hydroxylases.

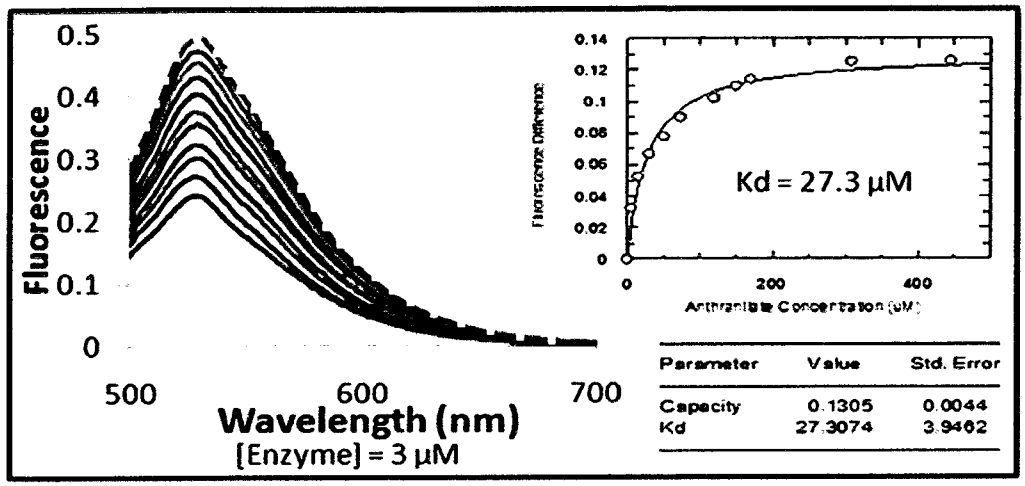
Bioinformatics analysis suggested that *Gtra*1516 is a phenol hydroxylase. It was successfully overexpressed in *E. coli* and purified in good yield. As with *Ledo*1986, the purified protein contained a non-covalently bound FAD but, unlike *Ledo*1986, the FAD was readily lost during purification. For this reason, FAD was included in the purification buffer. *Gtra*1516 was active with NADPH and phenol, producing very little catechol as a product. However, the uncoupling assay suggested that phenol is mainly a non-substrate effector, stimulating formation of hydrogen peroxide from NADPH and oxygen. *Gtra*1516 was also active with resorcinol, producing hydroxyquinol and some hydrogen peroxide. Combined with the previous observation, the enzyme thus appears to be a resorcinol hydroxylase rather than a phenol hydroxylase.

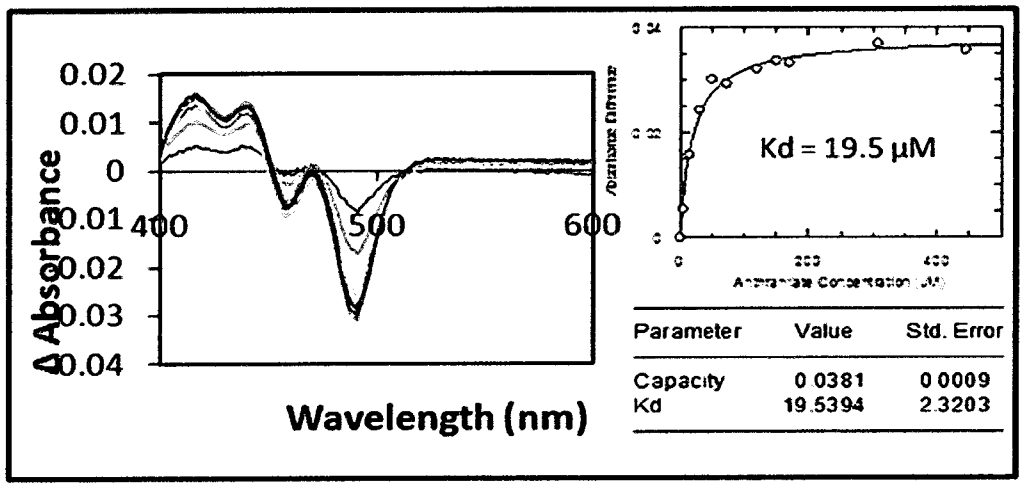
Bioinformatics analysis suggested that Gtra1270 is a catecho-1,2-dioxygenase. The protein was successfully overexpressed in *E.coli*, after extension of the N-terminus of the originally-predicted gene sequence, and was purified using four chromotographic steps. Purified enzyme appeared to aggregate, and was not active either as isolated or after the addition of iron.

APPENDIX

Ligand binding was examined by titrating the enzyme with aliquots of the ligand and measuring the resulting perturbations of the enzyme-bound flavin absorbance or fluorescence spectrum. Data were analyzed by a computerized, weighted nonlinear regression method. The equation in the method used is shown below, where E_T is the total enzyme concentration, L_T is the total ligand concentration, and ΔA, $\Delta Amax$ and K_d are the fitted variables *(93)*.

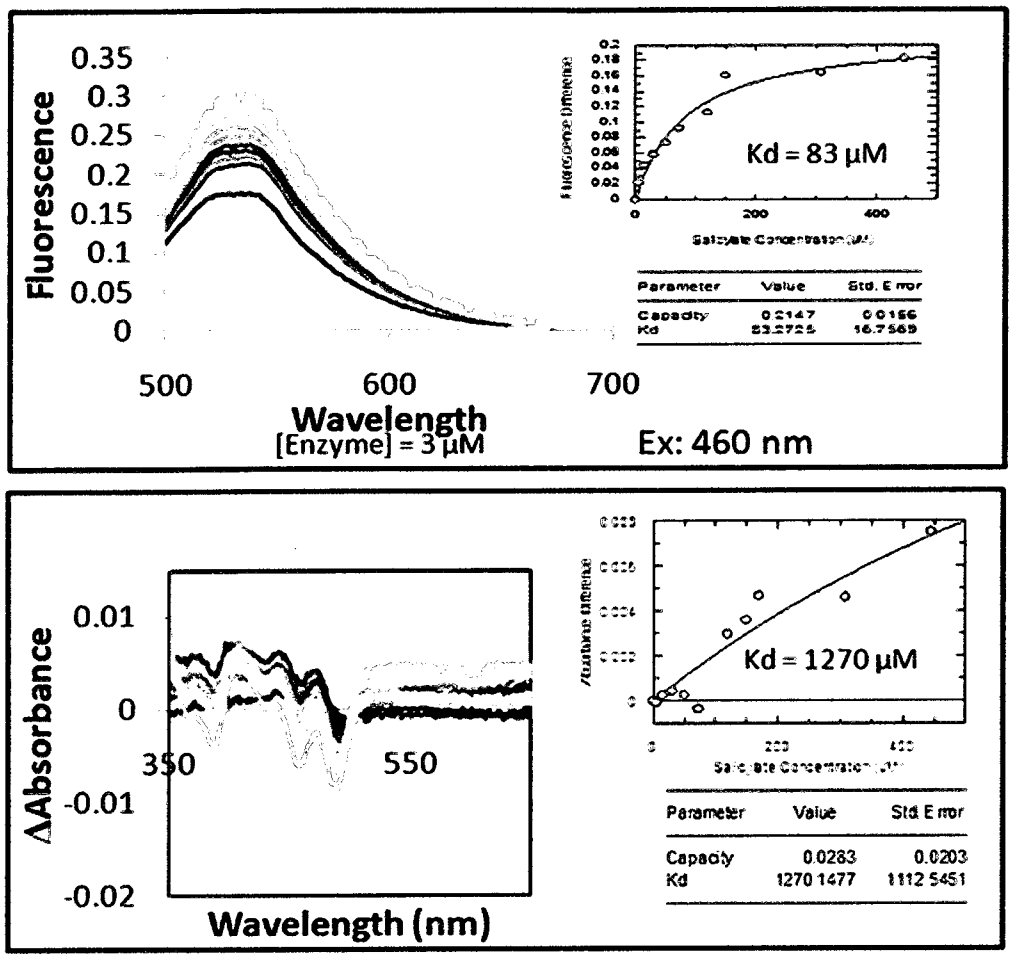
$$\frac{\Delta A}{\Delta A_{max}} = \frac{(L_T + E_T + K_d) - \sqrt{(L_T + E_T + K_d)^2 - 4E_T L_T}}{2E_T}$$





Appendix I: Titration of *Ledo*1986 with anthranilate A) changes in fluorescence emission spectrum and B) absorbance difference spectrum

This work was done by Joseph Napoletano. The experiment for A) was carried out using fluorescence spectroscopy. Samples were placed in 500 µl cuvette and data were collected by titrating 3 µM enzyme with gradually increasing amount of anthranilate. The experiments for B) monitored the visible spectrum of enzyme-bound FAD. Enzyme (3 µM) in a 1 ml cuvette was titrated with gradually increasing amounts of anthranilate. The inset shows the ΔAbsorbance vs. total anthranilate concentration.



Appendix II: Titration of *Ledo*1986 with Salicylate A) changes in fluorescence emission spectrum and B) absorbance difference spectrum

This work was done by Joseph Napoletano. The experiment for A) was carried out using fluorescence spectroscopy. Samples were placed in 500 µl cuvette and data were collected by titrating 3 µM enzyme with gradually increasing amount of salicylate. The experiments for B) monitored the visible spectrum of enzyme-bound FAD. Enzyme (3 µM) in a 1 ml cuvette was titrated with gradually increasing amounts of salicylate. The inset shows the ΔAbsorbance vs. total salicylate concentration

REFERENCES

1. W. C. Evans. *Journal of General Microbiology*. 1963;32:177-184.

2. C. Enroth, H. Neujahr, G. Schneider and Y. Lindqvist. *Structure*. 1998;6:605-617.

3. J. Gao, L. B. Ellis, and L. P. Wackett. *Nucleic Acids Research*. January, 2010;38:488–491.

4. I. Gut, V. Nedelcheva, P. Soucek, P. Stopka, and B. Tichavska. *Environmental Health Perspectives*. 1996;104:1211-1218.

5. Y. Shozo , K. Masayuki, M. Hiroo, and H. Osamu, *The Journal of Biological Chemistry*. 1965;240:3408-3413.

6. H. Y. Neujahr and A. Gaal. *European Journal of Biochemistry*. 1973;35:386-400.

7. R. H. White-Stevensi, H. Kamin. *The Journal of Biologocal Chemistry*. 1972;247:2358-2370.

8. H. Zouari, S. Moukha, M. Labat, and S. Sayadi. *Applied Biochemistry and Biotechnology*. 2002;102–103:261-276.

9. C. S. Harwood, G. Burchhardt, H. Herrmann and G. Fuchs. *FEMS Microbiology Reviews*. 1999;22:439-458.

10. D NM. Hayaishi. *Science*. 1969;164:389-396.

11. Y. R. Boldt, M. J. Sadowsky, L. B.M. Ellis, L. Que, JR. and L. P. Wackett. *Journal of Biotechnology*. 1995;177:1225-1232.

12. JB. Broderick. *Essays Biochem*. 1999;34:173-189.

13. P. A. Williams and J. R. Sayers. *Biodegradation*. 1994;5:195-217.

14. J. Powlowski and V. Shingler. *Biodegradation*. 1994;5:219-236.

15. W. J. Middelhoven. *Antonie van Leeuwenhoek*. 1993;63:125-144.

16. C. C. Reddy and C. S. Vaidyanathan. *Archives of Biochemistry and Biophysics*. December, 1976;177:488-498.

17. C. G. McNamee and D. R. Durham. *Biochemical and Biophysical Research Communications*. 1985;129:485-492.

18. R. Premkumar, P. V. Subba Rao, N. S. Sreeleela, and C. S. Vaidyanathan. *Biochemistry and Cell Biology*. 1969;47:825–827.

19. B. Haribabu, A. V. Kamath and C.S. Vaidyanathan. *FEMS Microbiology Letters*. 1984;21:197-200.

20. C. Enroth. *Acta Crystallographica Section D*. 2003;59:1597-1602.

21. M. Varga and H. Y. Neujahr. *European Journal of Biochemistry*. 1970;12:427.

22. J. D. Wright. *Springer Netherlands*. 1993;9:1573-0972.

23. A. Gaal and H. Y. Neujahr. *Journal of Bacteriology*. 1979;137:13-21.

24. J. D. Wright and C. Ratledge. *Applied Microbiology and Biotechnology*. 1991;35:0175-7598.

25. P. Mazur, W. J. Henzel, S. Mattoo and J. W. Kozarich. *Journal of Bacteriology*. 1994;176:1718-1728.

26. T. Kajander, M. C. Merckel, A. Thompson and A. M. Deacon. *Structure*. 2002;10:483-492.

27. C. S. Harwood and R. E. Parales. *Annual Review of Microbiology*. 1996;50:553-590.

28. A. Zeddel, A. Majcherczyk and A. Httermann. *Toxicological & Environmental Chemistry*. 1993;40:255 - 266.

29. A. Leonowicz, A. Matuszewska, J. Luterek. *Fungal Genetics and Biology*. 1999;27:175-185.

30. H. M. van den Brink, R. F. M. van Gorcom. *Fungal Genetics and Biology*. 1998;23:1-17.

31. D. Royse. Cultivation of shiitake on natural and synthetic logs *Penn State University, Department of Plant Pathology, University Park, PA*. 2001.

32. J. S. Birmingham. *Advances in Applied Microbiology*. 1993;39:153-184.

33. T. Mizuno. *Food Reviews International*. 1995;11:111-128.

34. N. Hatvani. *International Journal of Antimicrobial Agents*. 2001;17:71-74.

35. H. Mea. *International Journal of Antimicrobial agents*. Feb.1999;11:151-157.

36. C. C. Lee, D. W.S. Wong, and G. H. Robertson. *FEMS Microbiology Letters*. 2001;205:355-360.

37. N. Hatvani and I. Mécs. *Enzyme and Microbioal Technology*. 2002;30:381-386.

38. R. S. Makkar, A. Tsuneda, K. Tokuyasu, and Y. Mori. *FEMS Microbiology Letters*. 2001;200:175-179.

39. A. C. Grabski, H. J. Grimek,and R. R. Burgess. *Biotechnol Bioeng*. 1998;60:204-215.

40. L.O. Overholts. *The polyporaceae of the U.S. Alaska and Canada University of Michigan Press*. 1967;Ed. 2:110-111.

41. F.Green, and T.L.Highly, *Trends in Plant Pathology*. 1997;1:1-17.

42. K. A.Jensen, Z. C. Ryan, A.V. Wymelenberg, D. Cullen, and K. E. Hammel. *Applied and Environmental Mocrobiology*. 2002;68:2699-2703.

43. W. Wang and P.J. Gao. *Journal of Biotechnology*. 2003;101:119-130.

44. R. Cohen, M. R. Suzuki, and K. E. Hammel. *Applied and Environmental Mocrobiology*. 2004;70:324 - 331.

45. Z. Kerem, W. Bao, and K. E. Hammel. *PNAS*. 1998;95:10373-10377.

46. E. Varela, T. Mester and M. Tien. *Archives of Microbiology*. 2003;180:251-256.

47. G. Gea. *Journal of Biotechnology*. 1998;65:209-215.

48. F. Kamada, S. Abe, N. Hiratsuka, H. Wariishi, and H. Tanaka. *Microbiology*. 2002;148:1939 - 1946.

49. K. Fahr, H. G. Wetzstein, R. Grey, D. Schlosser. *FEMS Microbiology Letters*. 1999;175:127-132.

50. D. P. Ballou, B. Entsch, L. J. Cole. *Biochemical and Biophysical Research Communications*. 2005;338:590–598.

51. S. Hayat AA. *Springer*. 2007.

52. J. Lee, K. R. Min, Y. C. Kim, C. K. Kim, J. Y. Lim, H. Yoon, K.H. Min, K. Lee, and Y. Kim. *Biochemical and Biophysical Research Communication*. 1995;211:382-388.

53. S. Takemori, M. Nakamura, K. Suzuki and M. Katagiri. *FEBS Letters*. Feb. 1970;6 (4).

54. K.Masayuki, M.Hiroo, Y.Shozo, and H.Osamu. *The Journal of Biological Chemistry*. 1965;240:3414-3417.

55. S.Y. Isaac Sze and D. Stanley *Journal of Bacteriology*. 1984;159:353-359.

56. O.Hayaishi. *Sixth International Congress of Biochemistry*. 1964;33:31-43.

57. K. Masaytjki, T. Shigeki, S. Kenzi, and Y. Hiroshi. *The Journal of Biological Chemistry*. 1966;241:5675-5677.

58. K. Suzuki, M. Mizuguchi, K. Ohnishi, and E. Itagaki. *Biochimica et Biophysica Acta*. 1996;1275:154-156.

59. Y. Ohta, I. Higgins and D. W. Ribbons. *Applied and Environmental Mocrobiology*. 2006;72:7238-7245.

60. L. G., and V. Massey. *Biochemical and Biophysical Research Communications*. 1970;40:887-893.

61. U. Kirchner, A. H. Westphal, R. Muller, and W. J. H. van Berkel. *The Journal of Biological Chemistry*. 2003;278:47545–47553.

62. J. J. Kukor and R. H. Olsen. *Journal of Bacteriology*. 1990;172:4584-4630.

63. V. Shingler, C. H. Franklin, M. Tsuda, D. Holroyd and M. Bagdasarian. *Journal of General Microbiology*. 1989;135:1083-1092.

64. S. Ehrt, F. Schirmer, and W. Hillen. *Molecular Microbiology*. 1995;18:13-20.

65. E. Pessione, S. Divari, E. Griva, M. Cavaletto, G. L. Rossi, G. Gilardi and C. Giunta. *European Journal of Biochemistry*. 1999;265:549-555.

66. H. Y. Neujahr, K. G. Kjellen. *The Journal of Biologocal Chemistry*. 1978;253:8835-8841.

67. H.Y. Neujahr, G. Andras. *European Journal of Biochemistry*. 1975;58:351 -357.

68. K. Aoki, T. Konohana, R. Shinke and H. Nishira. *Agricultural Biology and Chemistry*. 1984;48:2087-2095.

69. R. N. Patel, C. T. Hou, A. Felix, and M. O. Lillard. *Journal of Bacteriology.* 1976;127:536-544.

70. M.W. Vetting and D.H. Ohlendorf. *Structure Fold Des.* 2000;8:429-440.

71. I. Matera, M. Ferraroni, M. Kolomytseva, and L. Golovleva. *Journal of Structural Biology.* 2010;170:548-564.

72. Y. Kojima, H. Fujisawa, A. Nakazawa, T. Nakazawa, and F. Kanetsuna. *The Journal of Biological Chemistry.* 1967;242:3270- 3278.

73. A. Marchler-Bauer et al. *Nucleic Acids Research.* 2009;37(D):205-210.

74. M. Latus, H. Seitz, J. Eberspacher, and F. Lingens. *Applied and Environmental Mocrobiology.* 1995;61:2453–2460.

75. R. B. Cain, R. F. Bilton, and J. A. Darrah. *Biochemistry Journal.* 1968;108:797–828.

76. User Protocol TB249 Rev. A 0204. pETBlue^{TM} system manual. *Novagen.* :1-32.

77. M. A. Frederick, B. Roger, E. K. Robert, D. M. David, A. S.J ohn. Current protocols in molecular biology, volumne 1.

78. User Protocol TB248 Rev. C 1108, Technical Bulletins,Novagen, 1-9. User protocol TB248 rev. C 1108, technical bulletins, Novagen, 1-9. .

79. R. N. Patel, C.T. Hou, A. Felix. *Journal of Bacteriology.* 1976;127:536-544.

80. R. E. Brown, K. L. Jarvis, K. J. Hyland. *Analytical Biochemistry.* 1989;180:136-139.

81. Y. Lei, P. D. Pawelek, and J. Powlowski. *Biochemistry.* 2008;47:6870-6882.

82. J. B. Powlowski. *Ph D Thesis.* 1983;University of Minnesota.

83. J. B. Powlowski, S. Dagley, V. Massey, and D. Ballou. *The Journal of Biological Chemistry.* 1987;262:69-74.

84. P. M. David. *Journal of Biological Chemistry.* 1991;266:10058-10061.

85. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. *Journal of Molecular Biology.* 1990;215:403–410.

86. X. Francesc, P. Boldu, R. Summerbell and G. Sybren de Hoog. *FEMS Microbiology Reviews.* 2006;30:109–130.

87. R. H. White-Stevens and H. Kamin. *The Journal of Biological Chemistry.* 1972;247:2371-2381.

88. R. Feeney, A.R. Clarke and J. J. Holbrook. *Biochemical and Biophysical Research Communications.* 1990;166:667-672.

89. J. Powlowski, V. Massey, and D. P. Ballou. *The Journal of Biological Chemistry.* 1989;264:6606-5612.

90. T. Sejlitz, H. Y. Neujahr. *European Journal of Biochemistry.* 1987;170:343 - 349.

91. Y. Ohta and D. W. Ribbons. *European Journal of Biochemistry.* 1976;61:259-269.

92. J. B. Broderick. *Assay in Biochemistry.* 1999;34:173-189.

93. G. H. Einarsdottir, M. T. Stankovich, J. Powlowski, D. P. Ballou and V. Masse. *Biochemistry.* 1989;28:4161–4168.