Expression, Purification and Characterization of Enzymes Involved in the Activation of Aromatic Ring Compounds in *Aspergillus niger*

Justin Raiche-Moyyen

A Thesis
In the Department
of
Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements
For the Degree of
Master of Science (Chemistry) at
Concordia University
Montreal, Quebec, Canada

April 2017

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Justin Raiche-Moyyen

Entitled: Expression, Purification and Characterization of Enzymes Involved in the Activation of Aromatic Ring Compounds in *Aspergillus niger*

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Chemistry)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examination committee:

_____ Chair

Dr. Brandon Findlay

_____ Examiner

Dr. Christine DeWolf

_____ Examiner

Dr. Peter Pawelek

_____ Supervisor

Dr. Justin Powlowski

Approved by _____

Chair of Department or Graduate Program Director

_____

Dean of Faculty

Date _____

Abstract

Expression, Purification and Characterization of Enzymes
Involved in the Activation of Aromatic Ring Compounds in *Aspergillus niger*

Justin Raiche-Moyyen

In microbial degradation of hydroxy-aromatic compounds, preparation of the ring for cleavage is a crucial step. Two enzymes types involved in this are hydroxylases and decarboxylases. Hydroxylases introduce an oxygen atom into the aromatic ring while decarboxylases remove a carboxyl group. Bacteria and fungi are known to elaborate these types of enzymes but fungal forms have so-far received relatively little attention.

A. niger is a fungus which has been shown previously to grow on salicylate, benzoate, and several other aromatics[1]. The *A. niger* genome has many genes annotated as coding for aromatic hydroxylases and some decarboxylases: fifteen such genes were chosen for recombinant expression in *E. coli*. Of these, four corresponding hydroxylases were successfully expressed, purified, and characterized according to substrate specificity, cofactor requirement and quaternary structure. These proteins, all FAD-containing, comprise a true salicylate hydroxylase, two salicylate hydroxylase-like enzymes, and a hydroxylase active on *p*-cresol, resorcinol, and hydroquinone. While the first enzyme is likely to be involved in metabolism of salicylate, the functions of the two salicylate hydroxylase-like enzymes are unclear, as is the function of the *p*-cresol active enzyme. A fifth enzyme, 2,3-dihydroxybenzoate decarboxylase, was also successfully expressed in *E. coli* and purified. Although its properties were similar to those previously reported for the enzyme isolated from *A. niger* [2,3,4], the pH optimum appeared to be somewhat higher. In addition, it was found that this enzyme could catalyze the reverse reaction, carboxylation of catechol, and that after treatment with a mixture of chelators Ca^{2+} appeared to stimulate the activity

Acknowledgements

First and foremost, I would like to thank my thesis supervisor Dr. Justin Powlowski for his constant guidance, suggestions, and endless support throughout this Master's thesis. His careful attention to detail has taught me to be more critical of everything I do both in the lab and life. His exceptionally detailed editing has contributed a great amount to the successfulness of conferences, graduate seminar, and most importantly the completion of this thesis document.

I would like to thank my committee members Dr. Christine DeWolf and Dr. Peter Pawelek for always listening and for their constructive feedback during my committee meetings. I would also like to thank Dr. Peter Pawelek for being the professor who initially sparked my interest in research and gave me the opportunity to experience it through a project in his lab during my undergrad studies.

Many thanks to my lab mates Patrick Semana, Felipe Venegas, Farnaz Olyaei, Dr. Elizabeth Cadieux, and Dr. Lena Sahlman, who have all been very helpful whenever I had questions and have made the lab experience particularly enjoyable. Additionally, I thank the Center for Structural and Functional Genomics (CSFG), Joyce, Pawelek, Turnbull, Finlay, and Wilds labs for all the equipment and reagents I borrowed during my studies. I thank Dr. Matthew Leibovitch (Joyce lab) who was often around when my lab members weren't and offered me assistance and advice on several occasions. He especially helped me in properly understanding how to interpret DNA sequencing data.

A big thanks to Annie Bellemare and her team at the CSFG who prepared the *Aspergillus niger* cDNA which made this project possible. I also thank Alain Tessier for running my ICP-MS samples and providing me with the data and answering any questions I may have.

Finally, I would also like to thank my family and friends for their support throughout my graduate studies, particularly my mom for always listening to me whenever things weren't going well and keeping my mindset positive.

Table of Contents

LIST OF FIGURES....x

LIST OF TABLES....xiii

LIST OF ABBREVIATIONS....xv

1 INTRODUCTION....1

1.1.1 Bacterial degradation strategies for benzoate derivatives....2

1.1.2 Fungal degradation strategies for benzoate derivatives....4

1.1.3 *A. niger* genes annotated to be involved in aromatic ring degradation....5

1.1.4 Properties of monooxygenase (hydroxylase) enzymes for hydroxyaromatics....6

1.1.5 Aromatic monooxygenases from fungi....8

1.1.6 (Di)-Hydroxybenzoate decarboxylases....10

1.1.7 Strategies for producing and characterization of fungal enzymes....12

1.2 Goals of this work....14

2 MATERIALS AND METHODS....15

2.1 Gene Selection, Amplification, Purification, and Analysis....15

2.1.1 Bioinformatic analysis of the *Aspergillus niger* genome to identify potential aromatic monooxygenases....15

2.1.2 PCR primers....15

2.1.3 PCR reactions....16

2.1.4 Analysis and purification of PCR products....17

2.1.5 Plasmid construction....18

2.1.6 Transformation of plasmids into *E. coli* DH5α....18

2.1.7 Growth of recombinant *E. coli* cells and plasmid extraction/purification....18

2.1.8 Plasmid analysis via restriction enzyme digestion....18

2.1.9 Plasmid analysis via PCR....19

2.1.10 Recombinant plasmid analysis via DNA sequencing....19

2.2 Protein Production, Analysis and Purification....20

2.2.1 Transformation of plasmids into *E. coli* expression strain....20

2.2.2 Growth of recombinant *E. coli* cells for protein expression.....20
2.2.3 Harvesting and cell lysis of recombinant *E. coli* cells.....20
2.2.4 Analysis of proteins via SDS-Electrophoresis (SDS-PAGE).....21
2.2.5 General strategy for purification of enzymes.....21
2.2.6 Procedure for DEAE (anion exchange) column chromatography.....21
2.2.7 Concentration of pooled fractions via ammonium sulfate precipitation.....22
2.2.8 Procedure for Phenyl-Sepharose (hydrophobic interaction) column chromatography.....22
2.2.9 Procedure for Sephacryl S-300 (size exclusion) column chromatography.....22
2.2.10 Second DEAE chromatography step.....22
2.2.11 TCA precipitation of enzymes in preparation for protein quantification.....23
2.2.12 Estimating protein concentrations.....24
2.3 Enzyme Characterization Experiments.....24
2.3.1 Identifying bound flavins.....24
2.3.2 Estimating number of FAD binding sites in purified proteins.....24
2.3.3 Preparation of substrates.....25
2.3.4 Enzymatic assays for hydroxylases and decarboxylases.....25
2.3.5 Assay for catechol production.....26
2.3.6 Sample preparation for product analysis by TLC.....26
2.3.7 TLC product analysis.....26
2.3.8 Oxygen monitoring experiments.....27
2.3.9 Substrate binding via monitoring FAD fluorescence, determining K_d.....27
2.3.10 Molecular weight and oligomeric state determination of proteins using gel filtration.....28
2.3.11 pH-activity profile of 2,3-dihydroxybenzoate decarboxylase.....29
2.3.12 Metal chelation of *A. niger* 2,3-dihydroxybenzoate decarboxylase.....29
2.3.13 Metal detection of enzymes via ICP-MS.....29
2.4 Bioinformatic and Computational Studies.....30

2.4.1 Construction of homology models.....30

2.4.2 Sequence alignments and construction of phylogenetic trees.....30

3 RESULTS.....31

3.1 Selection of *A. niger* Aromatic Ring Hydroxylase Genes.....31

3.2 Salicylate Hydroxylases.....33

3.2.1 PCR amplification to purification.....33

3.2.1.1 PCR amplification of *Aspergillus niger* genes annotated as salicylate hydroxylases.....34

3.2.1.2 Plasmid screening for correct inserts.....35

3.2.1.3 Heterologous expression of salicylate hydroxylases in *E. coli*.....37

3.2.1.4 Purification of three putative *A. niger* salicylate hydroxylase.....41

3.2.2 Salicylate Hydroxylases: Characterization.....45

3.2.2.1 UV-visible absorbance characteristics and identification of the accompanying flavin.....45

3.2.2.2 Number of FAD binding sites.....46

3.2.2.3 Oligomeric states of *A. niger* 767, 577, and 534 proteins.....47

3.2.2.4 NADH/NADPH specificity.....47

3.2.2.5 Aromatic substrate specificities of putative salicylate hydroxylases.....48

3.2.2.6 Measuring efficiency of substrate hydroxylation.....50

3.2.2.7 Product studies.....51

3.2.2.8 Substrate binding studies.....53

3.3 Phenol Hydroxylases.....55

3.3.1 PCR amplification to purification.....55

3.3.1.1 PCR amplification of phenol hydroxylase target sequences.....56

3.3.1.2 Cloning, plasmid screening. and DNA sequencing.....58

3.3.1.3 Protein expression results for putative phenol hydroxylases.....60

3.3.1.4 Purification of *A. niger* 999.....61

3.3.2 Phenol Hydroxylase: Characterization.....62

3.3.2.1 UV-VIS spectral properties of *A. niger* 999.....62

3.3.2.2 Oligomeric state of *A. niger* 999.....63

3.3.2.3 Nucleotide cofactor preference of *A. niger* 999.....64

3.3.2.4 *A. niger* 999 substrate specificity.....64

3.3.2.5 Measuring hydroxylation efficiency of putative phenol hydroxylase, *A. niger* 999.....65

3.3.2.6 *A. niger* 999 product identification with selected aromatic-ring substrates.....65

3.4 2,3-dihydroxybenzoate decarboxylase (*A. niger* 716).....67

3.4.1.1 PCR Amplification of *A. niger* 716.....67

3.4.1.2 2,3-dihydroxybenzoate decarboxylase plasmid insert verification.....68

3.4.1.3 *A. niger* 716 protein production.....69

3.4.1.4 *A. niger* 2,3-dihydroxybenzoate decarboxylase purification.....70

3.4.2.1 pH profile of *A. niger* 716.....72

3.4.2.2 *A. niger* 716 substrate specificity.....73

3.4.2.3 Oxygen sensitivity and reversibility of *A. niger* 2,3-dihydroxybenzoate decarboxylase.....73

3.4.2.4 Oligomeric state of *A. niger* 2,3-dihydroxybenzoate decarboxylase.....74

3.4.2.5 Effects of metal ions on *A. niger* 2,3-dihydroxybenzoate decarboxylase activity.75

3.5 Bioinformatic Studies.....78

3.5.1 Comparison of putative salicylate hydroxylase sequences.....78

3.5.1.1 Comparison of putative phenol hydroxylase sequences.....80

3.5.1.2 Comparison of aromatic ring decarboxylases.....82

3.5.2 *A. niger* salicylate hydroxylase active site motifs.....83

3.5.2.1 Accessibility of salicylate to salicylate hydroxylase active sites.....84

3.5.2.2 Potential catalytic serine residues.....85

3.5.3 Bioinformatic evidence suggests a potential salicylate 5-hydroxylase in *A.* niger...86

3.5.4 Phylogenetic analysis of *A. niger* target sequences.....87

4 DISCUSSION.....89

5 CONCLUSION.....101
REFERENCES.....103
APPENDIX.....112

LIST OF FIGURES

Figure 1.1: Structures of aromatic compounds mentioned in the text.....2

Figure 1.2: Ring cleavage pathways of microorganisms.....3

Figure 1.3: Crystal structure of salicylate hydroxylase from *pseudomonas* putida.....7

Figure 1.4: Salicylate 1-hydroxylase mechanism.....7

Figure 1.5: Diversity of products formed by salicylate hydroxylase from various microorganisms.....9

Figure 1.6: 2,3-Dihydroxybenzoate decarboxylase mechanism.....10

Figure 3.1: PCR Amplification of *A.niger* 767 (A), 577 (B), 534 (C), 860 (D), 655 (E), and 684 (F) run on 0.8% Agarose Gels.....35

Figure 3.2: Example of plasmid screening using restriction digestion.....36

Figure 3.3: Screening plasmid inserts for constructs of *A. niger* 767 (A), 577 (B), and 534 (C)..37

Figure 3.4: Soluble and insoluble fractions from BL21 (DE3) cells harbouring expression plasmids for *A. niger* 767 (A), 534 (B), 577 (C), 655 (D), and 684 (E) grown under various conditions.....39

Figure 3.5: Purification of *A. niger* 767 as monitored after electrophoresis on a 10 % SDS-polyacrylamide gel.....41

Figure 3.6: Purification of *A. niger* 577 as monitored after electrophoresis on a 10 % SDS-polyacrylamide gel.....43

Figure 3.7: Purification of *A. niger 534* as monitored on a 10 % SDS-polyacrylamide gel.....44

Figure 3.8: UV-VIS spectra of purified putative salicylate hydroxylases and released flavin for *A. niger* 767 (A), 577 (B), and 534 (C) in 50 mM Tris-HCl pH 7.6 at 25 ^{o}C.....45

Figure 3.9: Identification of putative *A. niger* salicylate hydroxylase flavins by TLC.....46

Figure 3.10: Example of oxygen uncoupling assay using *A. niger* 767 with salicylate.....50

Figure 3.11: Identification of *A. niger* 577 reaction products by TLC in 50:50:1 (ethyl acetate: chloroform: acetic acid) developing solvent where salicylate, catechol, and the salicylate reaction product have been spotted in lanes 1, 2, and 3 respectively.....52

Figure 3.12: Determination of salicylate binding to *A. niger* 767 (A) and *A. niger* 577 (B) K_d values by monitoring FAD florescence changes at λ_{ex} =450nm and λ_{em}= 525nm.....53

Figure 3.13: K_d determination for 4-methoxysalicylic acid binding to *A. niger* 767 (5 μM) by measuring changes in FAD fluorescence at λ_{ex} =450 nm and λ_{em}= 525 nm with varying additions of 4-methoxysalicylic acid.....54

Figure 3.14: PCR amplification of *A. niger* putative phenol hydroxylase genes.....57

Figure 3.15: Sequenced *A. niger* 146 plasmid construct with introns highlighted in red.....58

Figure 3.16: Sequenced *A. niger* 951 plasmid construct where introns are highlighted in red.....59

Figure 3.17: Phenol hydroxylase expression trials (induction at 16 ^{o}C).....60

Figure 3.18: Purification of *A. niger* 999 as monitored on a 10 % SDS-polyacrylamide gel.....61

Figure 3.19: UV-VIS Absorbance spectrum of purified *A. niger* 999 and released flavin in 50 mM Tris-HCl pH 7.6 buffer from 250-550 nm (A), 300-550 nm (B).....62

Figure 3.20: Identification of *A. niger* putative phenol hydroxylase flavin by TLC.....63

Figure 3.21: *A. niger* 999 product formation.....66

Figure 3.22: PCR amplification of *A.niger* 716 run on a 0.8% agarose gel.....67

Figure 3.23: *A. niger* 716 restriction digest map using *Kpn*I. pLATE11 derived DNA is shown in black while that from the insert fragment is shown in blue.....68

Figure 3.24: Restriction digest of recombinant *A. niger* 716 plasmids with *Kpn*I cleavage, followed by electrophoresis on a 1.0% agarose gel.....69

Figure 3.25: A. *niger* 716 protein expression in *E. coli* BL21 (DE3) at 16 ^{o}C as monitored by SDS-PAGE.....70

Figure 3.26: *A. niger* 716 protein purification - fractions analyzed by SDS-PAGE.....71

Figure 3.27: *A. niger* 716 activity as a function of pH over wide (A), and narrow (B) ranges.....72

Figure 3.28: Salicylate hydroxylase sequence alignments.....78

Figure 3.29: *A. niger* target phenol hydroxylase alignments.....80

Figure 3.30: Alignment of *A. niger* 2,3-dihydroxybenzoate decarboxylase with selected characterized aromatic-ring decarboxylases.....82

Figure 3.31: Alignment of the predicted active site structures of the *A. niger* putative salicylate hydroxylases.....84

Figure 3.32: Potential substrate entrances to salicylate hydroxylase active sites.....85

Figure 3.33: Potential catalytic serine residue in *A. niger* 577.....86

Figure 3.34: Screenshot of NRRL3_10244 and NRRL3_10245 (*A. niger* 1188051) in Gbrowse.....87

Figure 3.35: Phylogenetic tree of *A. niger* target and literature hydroxylase sequences.....88

LIST OF TABLES

Table 2.1: Primer Sequences used for Touchdown PCR....16

Table 2.2: PCR Reaction Components....17

Table 2.3: Touchdown PCR Program....17

Table 2.4: Plasmid Analysis PCR Reaction Components....19

Table 2.5: Plasmid Analysis PCR Program....19

Table 2.6: Purification steps used for each enzyme preparation....23

Table 3.1: List of all protein sequences chosen for this study....32

Table 3.2: Predicted superfamily domains of target proteins....33

Table 3.3: Expected DNA and protein sizes of putative salicylate hydroxylases....33

Table 3.4: Summary of steps achieved with *A. niger* putative salicylate hydroxylase....34

Table 3.5: Purification summary for *A. niger* 767....42

Table 3.6: Purification summary for *A. niger* 577....43

Table 3.7: Purification summary for *A. niger* 534....44

Table 3.8 FAD to protein ratio of putative salicylate hydroxylases....47

Table 3.9: Estimated molecular weights & oligomeric states for putative *A. niger* salicylate hydroxylases....47

Table 3.10: NADH/NADPH specificity of purified salicylate hydroxylases tested using salicylate as a substrate by monitoring the rate of NADH/NADPH consumption at 340 nm....48

Table 3.11: Substrate specificities of putative salicylate hydroxylases....49

Table 3.12: Oxygen uncoupling of salicylate hydroxylases....51

Table 3.13: Formation of catechol or catechol-like products via 3 hydroxylase enzymes....52

Table 3.14: Expected protein and gene sizes of putative phenol hydroxylases....55

Table 3.15: Summary of steps achieved for each putative phenol hydroxylase target....56

Table 3.16: Purification summary for *A. niger* 999....62

Table 3.17: Relative rates of *A. niger* 999 with NADH and NADPH....64

Table 3.18: *A. niger* 999 aromatic substrate specificity in the presence of NADPH....64

Table 3.19: *A. niger* 999 oxygen uncoupling with the most active phenolic substrates....65

Table 3.20: Purification summary for *A. niger* 716.....71

Table 3.21: *A. niger* 716 activity on different aromatic substrates.....73

Table 3.22: 2,3-Dihydroxybenzoate decarboxylase activity in the presence and absence of DTT.....74

Table 3.23: Reversibility of *A. niger* 716 catalyzed decarboxylation of 2,3-dihydroxybenzoate.....74

Table 3.24: Effects of metal ion additions to *A. niger* 716 prior to activity assay.....76

Table 3.25: ICP-MS metal analysis of *A. niger* 716.....76

Table 3.26: ICP-MS metal analysis of *A. niger* 716 samples after chelation with EDTA and 8-hydroxyquinoline-5-sulfanilic acid.....77

Table 3.27: Effects of $MgCl_2$ and $CaCl_2$ supplementation on *A. niger* 716 activity after chelation.....77

LIST OF ABBREVIATIONS

ATP, adenosine triphosphate

BCA, bicinchoninic acid

CFSG, Centre for Structural and Functional Genomics

DEAE, diethylaminoethyl

DNA, deoxyribonucleic acid

DTT, dithiothreitol

EDTA, ethylenediaminetetraacetic acid

FAD, flavin adenine dinucleotide

FI, fluorescence intensity

HPLC, high-performance liquid chromatography

ICP-MS, inductively coupled plasma mass spectrometry

IPTG, isopropyl β-D-1-thiogalactopyranoside

JGI, Joint Genome Institute

LIC, ligation independent cloning

NADH, nicotinamide adenine dinucleotide

PAGE, polyacrylamide gel electrophoresis

PCR, polymerase chain reaction

SDS, sodium dodecyl sulfate

TCA, trichloroacetic acid

TLC, thin-layer chromatography

Tris, tris(hydroxymethyl)aminomethane

UV-VIS, ultraviolet-visible

1. INTRODUCTION

Aromatic compounds constitute a large percentage of the Earth's biomass, mainly as the building blocks of lignin, the second most abundant plant biopolymer. Although aromatic compounds in central metabolism are mainly limited to a few amino acids and their derivatives, plants produce large quantities of secondary metabolites containing diverse aromatic rings. Aromatics are also found in large quantities in petroleum, and are produced and manufactured for industrial uses. In general, aromatic compounds are toxic to mammals, which have a relatively limited ability to degrade them. Instead mammals attempt to detoxify aromatic compounds and excrete them.

By contrast, many microorganisms have evolved to degrade aromatic compounds available in nature by channelling them into central metabolic processes, such as glycolysis and the TCA cycle, to capture energy in the form of ATP. Some have even been reported to break down compounds which have been synthesized: an example is pentachlorophenol degradation by *Sphingomonas chlorophenolica*[5]. A few examples of naturally-occurring compounds commonly broken down by both bacterial and fungal species are monocylic aromatics such as benzoate, salicylate, and phenol[6,7]. (Fig.1.1). Although these compounds are naturally-occurring, industrial use can dramatically increase their presence in the environment. Chloro- and nitro-substituted versions of these three simple aromatics can also be degraded by some bacterial species: for example, 4-chlorophenol or 2-chloro-4-nitrophenol are broken down by *Burkholderia sp. RKJ 800*[8]. Polycyclic aromatic compounds such as biphenyl, chlorobiphenyls, and naphthalene (Fig. 1.1), have also been reported to be degraded by bacteria[9]. Studying the degradation of aromatic compounds by microorganisms is not only of interest to learn about the natural catabolic diversity of microbes, but to potentially harness this catabolic power in bioremediation processes to decontaminate soil or water from toxic aromatic compounds.

Although microbial metabolism of aromatics can occur aerobically or anaerobically, this thesis only deals with aerobic catabolism.

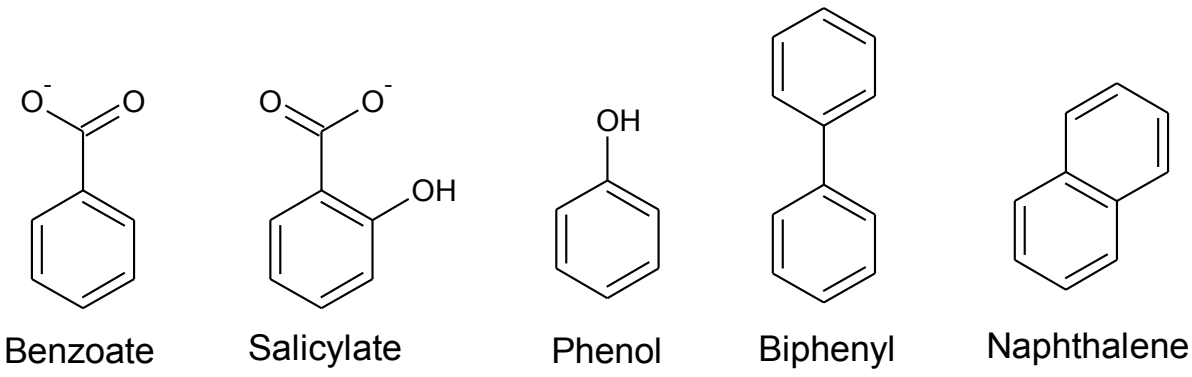


Figure 1.1: Structures of aromatic compounds mentioned in the text

1.1.1 Bacterial degradation strategies for benzoate derivatives

Aerobic bacterial aromatic degradation pathways for benzoates have been intensively studied and characterized. In common with other bacterial aromatic catabolic pathways, the key steps are: conversion to dihydroxylated ring-cleavage intermediate; oxygenolytic cleavage of the ring; and conversion of the ring-cleavage products into central metabolites. In the case of benzoates, the main ring-cleavage intermediates are catechol, protocatechuate, hydroxyquinol, hydroxyquinone, and gentisate, which can all be cleaved by non-heme iron-containing dioxygenase enzymes[10] (Fig. 1.2). Catechol and protocatechuate can be catabolized via either *meta* or *ortho* cleavage pathways, which involve dioxygenases that cleave between the hydroxyl groups of the ring cleavage substrate, or adjacent to them, respectively (Fig. 1.2). Via the *ortho* cleavage pathway, catechol, protocatechuate, and hydroxyquinol can be broken down to succinyl-CoA and acetyl-CoA to be used in the TCA cycle via the β-ketoadipate pathway.[1] By contrast, in the *meta*-cleavage pathways catechol, protocatechuate, and gentisate are all converted to pyruvate: in addition, catechol, protocatechuate and gentisate also produce acetyl-CoA, oxaloacetate and fumarate, respectively. Finally, hydroxyquinone is converted into the central metabolites, succinyl-CoA and acetyl-CoA, via a *meta*-cleavage pathway to be used in the citric acid cycle[10].

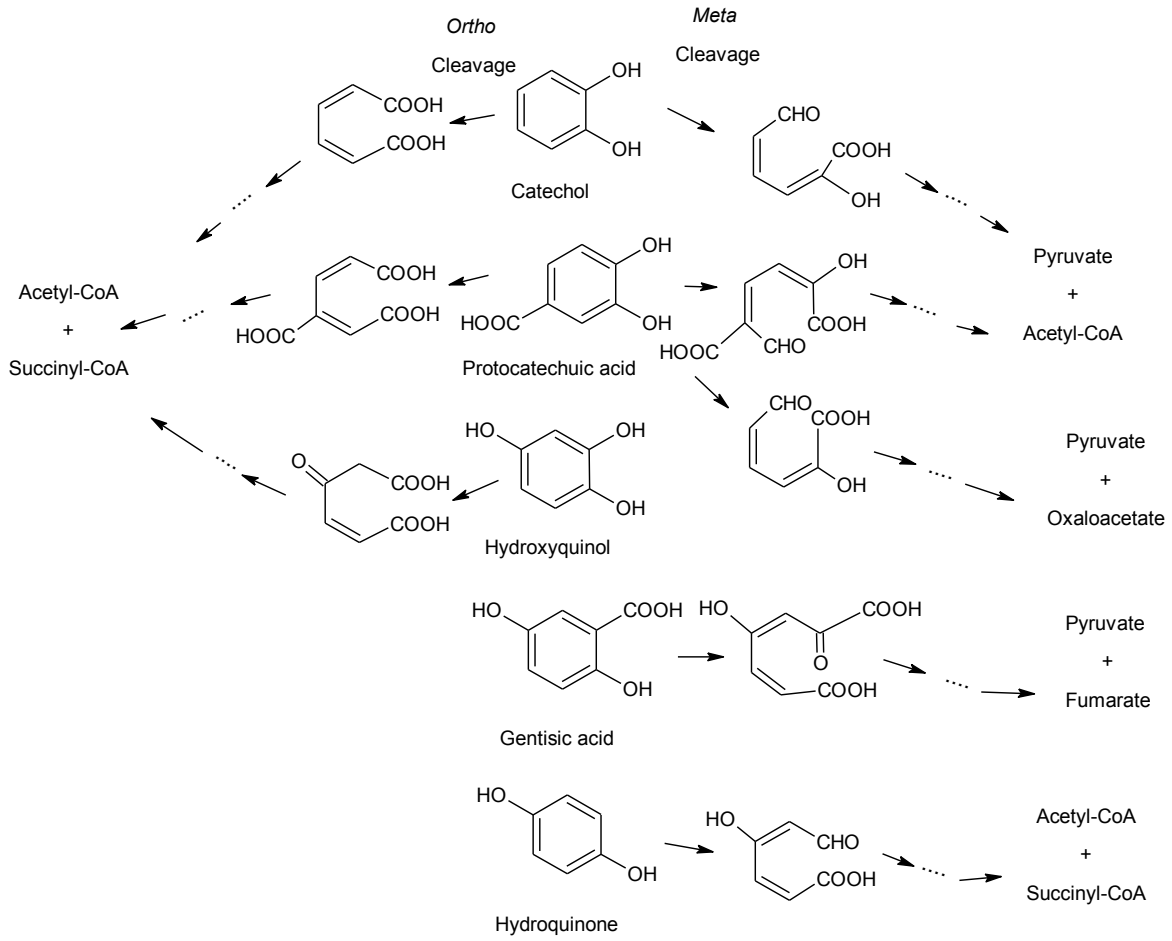


Figure 1.2: Ring cleavage pathways of microorganisms (Adapted from Vaillancourt et al, 2006[10])

The ring cleavage substrates described above are produced from benzoates by combinations of monooxygenase (hydroxylase) catalyzed reactions. For example, both catechol and gentisic acid may be formed from salicylate (2-hydroxybenzoate), via salicylate hydroxylase. Salicylate 1-hydroxylase[11] produces catechol from salicylate via a decarboxylation-hydroxylation reaction, while salicylate 5-hydroxylase[12] converts salicylate to gentisate with only a hydroxylation step. Salicylate 1-hydroxylase has been identified in *Pseudomonas putida PpG7*[11], while salicylate 5-hydroxylase has been isolated from *Ralstonia sp*[12]. Catechol can also be produced via hydroxylation of phenol by phenol hydroxylase which has been isolated from, for example, *Pseudomonas putida CF600*[13]. Protocatechuic acid is formed from *p*-hydroxybenzoate by *p*-hydroxybenzoate hydroxylase in *Pseududomonas*

fluorecens[14]. The formation of both hydroxyquinol and hydroquinone have been observed in *Candida parapsilosis* CBS604. Hydroxyquinol can be formed by the hydroxylation of either hydroquinone or resorcinol by hydroquinone hydroxylase [15], while benzoate is decarboxylated and hydroxylated to produce hydroquinone by as-yet unidentified enzymes[16].

1.1.2 Fungal degradation strategies for benzoate derivatives

The enzymes involved in bacterial aerobic aromatic degradation have been well studied and characterized, but relatively little characterization of benzoate-degrading enzymes from fungi has taken place. This is despite the fact that various authors have reported growth of a range of fungi and yeast on benzoates and their derivatives[1]. One of the exceptions is the soil yeast *Trichosporon cutaneum*, which exhibits considerable diversity in the range of single-ring aromatic compounds it can metabolize, including phenol, salicylate, benzoate, 2-3-dihydroxybenzoate, protocatechuate, and gentisate. *T. cutaneum* possesses catechol 1,2-dioxygenase and hydroxyquinone 1,2-dioxygenases, while it lacks protocatechuate, 2,3-dihydroxybenzoate or gentisate dioxygenases that are commonly found in bacteria [17]. 2,3-Dihydroxybenzoate is first decarboxylated via a 2,3-dihydroxybenzoic acid decarboxylase to form catechol, which is cleaved by catechol 1,2-dioxygenase. Additional hydroxyl groups are added to protocatechuate and gentisate, concomitantly with decarboxylation to form hydroxyquinol which is cleaved by hydroxyquinol 1,2-dioxygenase dioxygenase. For benzoate degradation, three hydroxylation reactions are catalyzed by *T. cutaneum* to form 4-hydroxybenzoate, protocatechuate and, finally, hydroxyquinol which can then be cleaved by hydroxyquinol 1,2-dioxygenase. Neither this organism or any other eukaryote has been reported to use a pathway that involves *meta*-ring cleavage (see Fig. 1.2), which appears to be limited to bacteria.[17]

A filamentous fungus that has been reported to grow on benzoate and some of its derivatives is the industrial workhorse and model fungus, *Aspergillus niger*[1]. *A. niger* has been shown to grow on many benzoate derivatives, including *p*-hydroxybenzoate, *p*-methoxybenzoate, protocatechuate, and catechol, as well as lignin building blocks such as vanillate, ferulate, and caffeate[1]. *A. niger* growth on these substrates indicates there must be aromatic degrading enzymes in the *A. niger* proteome but to date few have been isolated and characterized. Those that have, and which have roles in the preparation of aromatic compounds for ring cleavage,

include a *p*-hydroxybenzoate hydroxylase[18], a 2,3-dihydroxybenzoate decarboxylase[2] and an anthranilate hydroxylase[19] which are, respectively, responsible for benzoate conversion to *p*-hydroxybenzoate, 2,3-dihydroxybenzoate decarboxylation to catechol, and anthranilate conversion to 2,3-dihydroxybenzoate. However, in the genome of *A. niger* the genes encoding these enzymes are a small fraction of the genes annotated with functions related to the metabolism of benzoates.

1.1.3 *A. niger* genes annotated to be involved in aromatic ring degradation

The genome of *Aspergillus niger* was amongst the first published for a filamentous fungus[20]. Since then, many hundreds more have been sequenced. Currently, the Joint Genome Institute's 1000 genome project aims to obtain a better understanding of fungal diversity by sequencing two genomes of every fungal family.[21] Within all these genomes, some genes have previously been characterized while many others are annotated only by bioinformatic comparison with known genes.

In the *A. niger* genome many genes are annotated with functions suggesting that they are involved in aromatic degradation pathways. There are two publically-available *Aspergillus niger* genomes, *A. niger* ATCC 1015v4.0[22], prepared by the JGI (Joint Genome Institute), and the *A. niger* NRRL3 genome prepared during the Genozymes[23] project. The NRRL3[23] *A. niger* genome has been manually curated and has 65 genes annotated as hydroxylase enzymes. Of these 65 genes, almost all have been annotated as either salicylate hydroxylase, phenol hydroxylase, or aromatic ring hydroxylase enzymes. Although, as described above, *A. niger* has been shown to grow on benzoate and salicylate, no experimental evidence that *A. niger* grows on phenol has been reported. However, *Aspergillus fumigatus*[24] and *Aspergillus awarmori*[25] have both been shown to grow on phenol, so it is possible that *A. niger* also has this capability.

Upon examination, these annotated *A. niger* hydroxylase gene sequences are often only between 20% - 35% identical with characterized bacterial counterparts. Considering that aromatic hydroxylase enzymes encoded by fungal genes have rarely been expressed and characterized, and the sequence similarity between the fungal and bacterial versions is often low, there could potentially be some major differences in the properties of the enzymes they encode that reflect different lifestyles of bacteria and fungi. Bacteria are prokaryotic single-celled

organisms while fungi are eukaryotic multi-cellular organisms. Bacteria and fungi can live in the same environments, but they can also live under quite different conditions. For example *A. niger* has been reported to grow at temperatures ranging from 6-47 ^{o}C and at pH values from 1.4-9.8[26]. *A. niger* is a useful model fungus for studying aromatic degradation because in addition to having a fully sequenced and annotated genome, it is known to grow on aromatic substrates as sole carbon sources, and it has also been used in many other areas of research and in industrial applications.

1.1.4 Properties of monooxygenase (hydroxylase) enzymes for hydroxyaromatics

This thesis will mainly focus on fungal enzymes responsible for preparing hydroxyaromatics for subsequent ring-cleavage. Referred to as "monooxygenases" or "hydroxylases", these enzymes are responsible for hydroxylating aromatic rings with one (or sometimes more) OH group by inserting one atom of molecular oxygen into the ring and reducing the other to water (reviewed in Huijers et al, 2014[27]). The electron-donating OH group makes it easier for the ring to be broken open in the next metabolic step by ring-cleaving dioxygenase enzymes. In order for hydroxylation to occur, in addition to the aromatic substrate most of these monooxygenase require FAD and NAD(P)H cofactors. NAD(P)H is a dissociable cofactor responsible for reducing the enzyme-bound FAD to its $FADH_2$ state, which can react with O_2. The structures of FAD-containing aromatic hydroxylases generally are mixed α-helical and β-sheet-containing proteins with a substrate binding site in addition to the FAD and NAD(P)H binding sites. FAD and NADH are dinucleotide cofactors whose binding sites are located in the Rossmann fold domain which is a classical β-α-β motif[28]. As an example, the structure of a bacterial salicylate hydroxylase from *Pseudomonas putida* is displayed below: the Rossmann fold domain is indicated in red, FAD in yellow, and salicylate in grey[28].

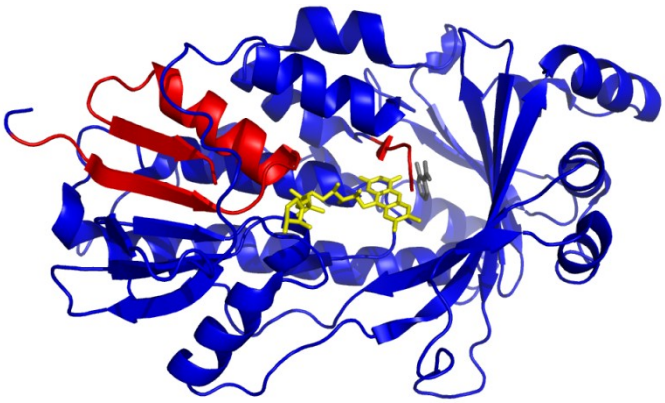


Figure 1.3: Crystal structure of salicylate hydroxylase from *Pseudomonas putida* (PDBID: 5evy)[28] generated using PyMOL.

O_2 activation by $FADH_2$ to hydroxylate aromatic substrates is tightly coupled if all of the oxygen consumed is used for hydroxylation. However, there can be hydrogen peroxide by-product formed with these enzymes when oxygen is not completely used for hydroxylation of the aromatic ring. This is referred to as "oxygen uncoupling"[29]. In order to understand this phenomenon, the mechanism of these enzymes must be considered (Fig. 1.4). The first step is the reduction of FAD to $FADH_2$ using the electrons donated from NAD(P)H as the reducing power. The second step consists of O_2 reacting with the reduced flavin to form the C4a-flavinhydroperoxide. Lastly, this C4a-flavinhydroperoxide intermediate transfers its terminal oxygen to the aromatic ring substrate, and the other atom remains with the flavin as a C4a-flavin hydroxide which eliminates water to regenerate FAD (reviewed in Ballou et al, 2005)[30]. Oxygen uncoupling occurs when FAD is reduced and the C4a-hydroperoxide is formed, either when the substrate is not present or if it is not positioned properly in the active site, and decays to release hydrogen peroxide.

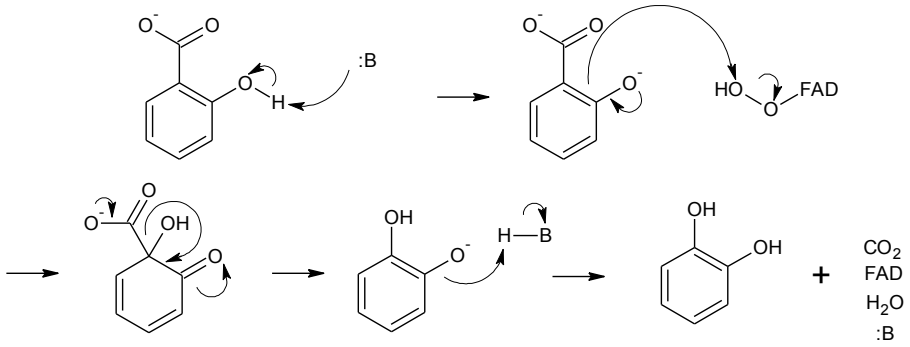


Figure 1.4: Salicylate 1-hydroxylase mechanism. The symbol (:B) is a catalytic base.

Usually, oxygen uncoupling with no substrate is low since flavin reduction is quite slow in the absence of aromatic substrate bound to the enzyme[30]. However it is possible for some flavoprotein hydroxylases to recognize a pseudo-substrate, which binds similarly to the true substrate, stimulates FAD reduction, but is not hydroxylated at all, resulting in complete diversion of oxygen to H_2O_2. This was first reported for a salicylate hydroxylase isolated from P*seudomonas putida* by White-Stevens et al. (1972), where benzoate is a non-substrate effector that causes formation of H_2O_2 while it's true substrate salicylate is fully coupled. They also observed partial hydroxylation of some substrates such as 3-methylsalicylate, 2,4-dihydroxybenzoate, and 2,5-dihydroxybenzote.[29] Partial hydroxylation occurs when a non-optimal but catalytically-competent substrate is bound and decay of the C4a-flavin hydroperoxide competes with hydroxylation. In addition, it has been shown that in other salicylate hydroxylases the order in which substrate and NADH bind can vary yet still result in hydroxylated product (catechol)[31].

1.1.5 Aromatic monooxygenases from fungi

In the literature, there are many bacterial aromatic ring hydroxylases, like the salicylate hydroxylase described above, that have been identified and studied, but far fewer from fungi have been reported. Previous studies have shown that fungal monooxygenases are active on benzoate and many of its derivatives such as 4-hydroxybenzoate, 3-hydroxybenzoate, phenol, and salicylate (reviewed in Wright et al[7]). In many cases, crude extracts of these fungi were used to assay the activities of these enzymes, but the specific enzymes were not purified or isolated. For example, some fungi have been reported to be able to degrade benzoate. Benzoate hydroxylase activity, converting benzoate to 4-hydroxybenzoate, has been reported for *Aspergillus niger, Rhodotorula graminis and Trichosporon cutaneum*[7]. 4-Hydroxybenzoate 3-monooxygenase activity has been observed in extracts of *A. niger, Penicillium spinulosum, Schizopyllum commune, and Penicillium* sp*., Aureobasidium pullulans, Neurosporacrassa, Rhodotorula mucilaginosa, Trichoderma viride* which converts 4-hydroxybenzoate to protocatechuate (3,4- dihydroxybenzoate).[7] In *A. niger* it was also reported that 3-hydroxybenzoate could be converted to protocatechuate via a 3-hydroxybenzoate-4-monooxygenase. This appears to be the only fungus reported to be capable of carrying out this reaction to date[32].

Salicylate hydroxylase is an enzyme of interest since it is involved in transforming salicylate, a very important bio-molecule especially in plants and fungi. In fact, salicylate is involved in the plant immune response and it is theorized that the salicylate hydroxylase expressed in some symbiotic fungi such as *Epichole festucae* is responsible for drawing out the host plant defence by breaking down salicylate.[33]

Salicylate hydroxylases are probably the most well studied flavoprotein hydroxylase enzymes from fungi, and enzymes from different species produce a variety of products (Fig. 1.5). Some have been observed to decarboxylate and hydroxylate salicylate to form catechol as is the case in *Trichosporon cutaneum*[17], *Rhodotorula graminis*[34], *and Epichloe festucae*[33]. However, others have also been observed to simply hydroxylate with no decarboxylation. This can involve: salicylate conversion to 2,4-dihydoxybenzoate, as observed in *A. niger, A. nidulans, and Trichosporon* sp.*;* salicylate conversion to 2,5-dihydroxybenzoate observed in *Trichosporon* sp.*;* or conversion of salicylate into 2,3-dihydroxybenzoic acid as in *A. niger, A. nidulans, and Trichoderma lignorum*.[7]

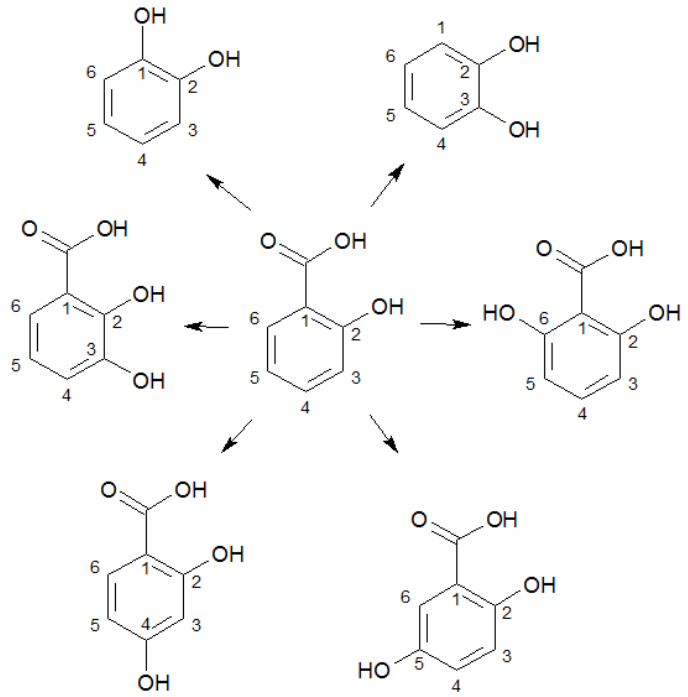


Figure 1.5: Diversity of products formed by salicylate hydroxylase from various microorganisms

Some fungi have been reported to degrade phenol[24,25]. Phenol hydroxylase has been identified and isolated from *Trichosporon cutaneum*, and shown to transform phenol into catechol[35]. Phenol is a common industrial by-product, and is highly toxic and can cause many undesired effects in humans when consumed[36]. At high levels phenol can even be lethal and phenol has also been classified as a potential mutagen and carcinogen[37]. Its mineralization by microbes is potentially of interest for bioremediation purposes.

1.1.6 (Di)-Hydroxybenzoate decarboxylases

Monooxygenases are the main type of enzyme responsible for converting aromatic rings into ring cleavage substrates, but another class of enzymes which can also perform this task is the aromatic ring decarboxylases. These enzymes remove carboxyl groups on the aromatic ring without the use of oxygen, releasing CO_2[38]. For example, 2,3-dihydroxybenzoate decarboxylation to catechol is catalyzed by 2,3-dihydroxybenzoate decarboxylase (Fig. 1.6). The proposed mechanism involves initiating the reaction by first forming a quinone via deprotonation of the OH group nearest to the carboxyl group. Due to the resonance destabilization of the intermediate, decarboxylation is facilitated and subsequently reprotonated to form catechol (Fig. 1.6)[3].

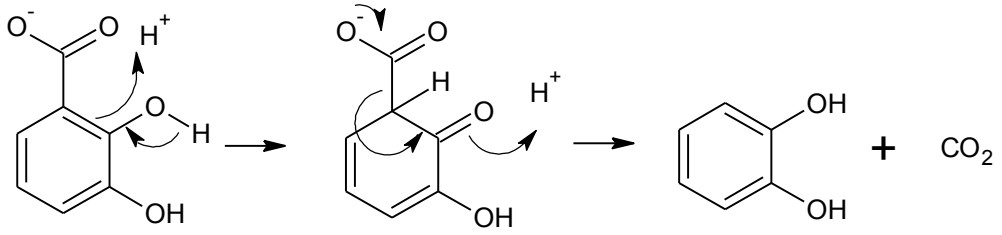


Figure 1.6: 2,3-Dihydroxybenzoate decarboxylase mechanism

Usually such aromatic decarboxylases are metal dependent and active on compounds which already have hydroxyl groups on the aromatic ring in addition to the carboxyl group. However, there have been several reports of aromatic decarboxylases not requiring metals, especially amongst those from fungi. 2,6-Dihydroxybenzoate decarboxylases from *Rhizobium sp*[39], and from *Polaromonas sp*[40], are examples of bacterial aromatic decarboxylases requiring Zn^{2+} and Mn^{2+} respectively. On the other hand, salicylate decarboxylase and 2,3-dihydroxybenzoate decarboxylase from *Trichosporon moniliiforme*[41], and *Aspergillus oryzae*[38],

respectively, are fungal decarboxylases that have been reported not to require metal ions. 4-Hydroxybenzoate decarboxylase and 3,4-dihydroxybenzoate decarboxylase from *Enterobacter cloacae P240*[42], *Clostridium hydroxybenzoicum*[43], respectively, are examples of bacterial decarboxylases reported not to require metal ions for activity. Thus, the metal ion requirement, and the identity of the metal ion itself, appear to vary from decarboxylase to decarboxylase.

A few common elements within this class of enzymes is that they appear to usually be tetramers, and have a common TIM-barrel folding motif comprising 8 alternating α,β strands, regardless of whether the enzyme has been reported to require a metal ion or not[39]. These decarboxylases usually exhibit very high specificity for one substrate: such is the case for the 2,3-dihydroxybenzoate decarboxylase from *A. niger*[2] which only decarboxylates 2,3-dihydroxybenzoate. However, there are also examples of enzymes displaying more substrate diversity such as the 2,6-dihydroxybenzote decarboxylase from *Rhizobium* sp.[44], which decarboxylates both 2,6-dihydroxybenzoate and 2,3-dihydroxybenzoate.

Many aromatic decarboxylases also have been reported to also catalyze the reverse reaction (carboxylation), as has been reported for enzymes such as salicylic acid decarboxylase from *Trichosporon monilliforme*[41] and 2,6-dihydroxybenzoate decarboxylase from *Rhizobium sp*[45]. For most of these enzymes, reversibility is affected by oxygen, and usually they require reducing agents such as DTT in assays to observe carboxylation in vitro [41].

Up to this point all the enzymes discussed have been treated as enzymes involved in aromatic compound degradation. Although this has been shown for a large number of these enzymes, they may not always be involved in aromatic compound degradation. For example, these types of enzymes could be responsible for forming intermediates in biosynthetic pathways. One example is 2,3-dihydroxybenzoate, which could be produced by salicylate 3-hydroxylase from salicylate. 2,3-Dihydroxybenzoic acid could be further processed into catechol and fed into the β-ketoadipate pathway for degradation, but several molecules of 2,3-dihydroxybenzoate could also be combined into siderophores such as enterobactin, anguibactin, and vibriobactin which are iron chelators produced by some bacteria[46].

1.1.7 Strategies for producing and characterization of fungal enzymes

Degradative enzymes may be obtained in two main ways: either extract them from the native organism, or express them in recombinant form. In the case of the first method, the organism must be grown under appropriate conditions in order to induce the expression of the enzymes of interest. For expression of aromatic degradative enzymes, aromatic compounds that can be used directly as sole carbon and energy sources are often used as the growth substrate to induce enzymes that are responsible for the breakdown of the carbon source, or responsible for their conversion to carbon dioxide and water. Another option is to use a more complex aromatic-containing compound, such as lignin, which can be broken down into benzoate or benzoate-like compounds by certain microorganisms.

For recombinant expression, the fungal gene is inserted into a plasmid, which is transformed into an expression host, such as *E. coli,* that expresses the enzyme encoded in the insert. There are several challenges to expressing fungal proteins in bacteria. First, mRNA from the fungus, grown under appropriate conditions to induce enzyme expression, must be extracted and then reverse transcribed into cDNA in order to be utilized in bacteria. The reason for this is that fungi are eukaryotes that contain introns in their genes. Higher organisms such as fungi have the ability to splice these non-coding DNA fragments out, but bacteria do not. cDNA must therefore be used to have a DNA template that does not contain introns. If intron-containing DNA is used in bacterial protein expression it will result in an abnormal protein containing unnecessary amino acid regions, premature stop codons, or frame shifts, and this will usually result in inactive enzyme. It can also be difficult to predict the correct start codon, and differences in codon usage can effect expression levels[47].

Although there are challenges to expressing fungal enzymes in *E. coli,* it has several advantages. *E. coli* expression plasmids and hosts are well-developed and easy to manipulate. *E. coli* has one of the shortest doubling time of any known organism, only about 20 min.[48] Fungi on the other hand have a much longer doubling time on the scale of hours: *Aspergillus niger* has a doubling time of about 3.6 h.[49] Fungi can take days to express enough protein to work with, while *E. coli* can produce these desired levels in a matter of hours. *E. coli* is a biohazard level 1 organism, whereas, *A. niger* is biohazard level 2. *A. niger* in general is fairly harmless but in immuno-compromised individuals, such as those allergic to its spores, it can

cause harm: in addition some strains can produce toxins. Expressing *A. niger* enzymes in *E. coli* eliminates the risk of exposure to *A. niger* spores and toxins.

A possible disadvantage of expressing fungal proteins in bacteria is that the protein produced may not be fully native since bacteria do not have as much machinery for protein post-translational modifications, if any are required. This becomes an issue in attempting to express enzymes which are normally secreted by fungi, since bacteria do not have the same capabilities to cleave signal peptides or carry out other important modifications such as glycosylation or phosphorylation. However, the enzymes that are targeted in this thesis are intracellular enzymes which in general require less post-translational processing than extracellular enzymes.

1.2 Goals of this work

Fungal genome sequencing efforts have resulted in published genomes that are replete with genes annotated as being involved in degradation of aromatic compounds, although almost no corresponding fungal enzymes have been characterized. The primary goal of this master's project was to identify, express and characterize enzymes, from the model fungus *Aspergillus niger,* that potentially encode enzymes involved in activation of aromatic ring compounds for ring cleavage. Most of the enzyme targets are of the monooxygenase class of enzymes, which hydroxylate aromatic rings to form di-hydroxylated single aromatic ring compounds. These enzymes were expressed in *E. coli* by first choosing genes that are annotated as hydroxylase enzymes in the *A. niger* genome. The workflow of the project consisted of: cloning genes into the pLATE11 *E. coli* plasmid; expressing them in *E. coli*; purifying the expressed proteins; and then characterizing them to see if they perform their bioinformatically assigned functions, most of which are based on weak sequence similarity to bacterial counterparts. Another goal was characterization of these enzymes to identify if there are any differences in their characteristics compared to those from bacteria. In studying these enzymes, we expected to find some properties or structure-function relationships of the fungal enzymes that differ from those of bacteria.

The long term goal of this project is to eventually combine this work with research on other enzymes associated with the degradation of aromatic ring compounds, and map out the aromatic degradation pathways in *A. niger* for comparison to those already known in bacteria and fungi. In addition, although the goal of this project is to understand these enzymes better, they may also have potential use in industries such as bioremediation, food, and pharmaceutical industries.

2. MATERIALS AND METHODS

2.1 Gene Selection, Amplification, Purification, and Analysis

2.1.1 Bioinformatic analysis of the A*spergillus niger* genome to identify potential aromatic monooxygenases

Genes were identified by using text such as "monooxygenase", "hydroxylase", and "aromatic-ring hydroxylase" to search the *Aspergillus niger* ATCC 1015 v4.0[22] and NRRL3[23] genomes prepared by the Joint Genome Institute (JGI) and Concordia Fungal Genomics Project respectively. Potential genes of interest were then Blasted against the Uniprot/Swissprot database, and those hits having at least 30% sequence identity with the query sequence were investigated further.

2.1.2 PCR primers

PCR primers were designed for use with the Thermo Scientific aLICicator LIC cloning kit following the manufacturer's recommendations[50]. All forward primers began with the nucleotides 5'-AGAAGGAGATATAACT-3', followed by the coding sequence of interest. For reverse primers 5'-GGAGATGGGAAGTCA-3' was followed by the coding sequence of interest. All primers were designed to have a Tm of 60 ^{o}C or greater, as well as a GC content of 40-60%[50]. Primers were synthesised and prepared by Integrated DNA Technologies (IDT).

Primer	Sequence (5'→3')
1185767 Forward	AGAAGGAGATATAACTATGACCGTGTCCAACAATGACAAACCTCTCCG AATCGC
1185767 Reverse	GGAGATGGGAAGTCATCAAGAACTAGCATGAAAAATGTCCAACGCATC TTGCAGGTGC
1127577 Forward	AGAAGGAGATATAACTATGTCCCCCTCCAAGAAAGACTTCCACGTCGC
1127577 Reverse	GGAGATGGGAAGTCATCACAACGAAGCCTTCTCAATCTTCCTCCCCAAC AC
1157716 Forward	AGAAGGAGATATAACTATGCTTGGTAAAATCGCTCTCGAAGAAGCCTTC GCG
1157716 Reverse	GGAGATGGGAAGTCACTAAGCCTTAACCTTCGCATCATAATCCCTAAAC GCCCC
1210999 Forward	AGAAGGAGATATAACTATGCCGGTCTTTACAGAATATGCATCCAAATCG CGGG
1210999 Reverse	GGAGATGGGAAGTCACTACTTCGTTCGGATACAGCGACGCAGGTACTCT CC
1205534 Forward	AGAAGGAGATATAACTATGGCGCAGCCAGCAGCGAGCATATC
1205534 Reverse	GGAGATGGGAAGTCATCATAATCGTGACTGAGACATCTGTACCAGGGC TATAAACCACCG

Table 2.1: Primer Sequences used for Touchdown PCR

2.1.3 PCR reactions

PCR reactions (50 µl) were prepared as indicated in Table 2.2 below. Genes were amplified using Touchdown PCR, using the conditions shown in Table 2.3 below [51]. Template DNA used was cDNA prepared by the Centre for Structural and Functional Genomics (CFSG) from *Aspergillus niger* N402 grown on 2 % alfalfa and barley. The dNTP mixture was purchased from Roche, and Phusion High-Fidelity DNA Polymerase was purchased from New

England Biolabs (NEB). PCR reactions were performed using the Applied Biosystems Verti 96 Well Thermal Cycler.

Component	Volume (µL)
Nuclease free water	24.0
5X Phusion HF or GC Buffer	10.0
1 mM dNTP's	10.0
10 µM Forward Primer	2.5
10 µM Reverse Primer	2.5
Aspergillus niger cDNA	0.5
Phusion DNA polymerase	0.5

Table 2.2: PCR Reaction Components

Phase	Cycles	Step	Temperature (^{o}C)	Time (s)
Initial	1	Initial Denaturation	98	180
1	10	Denaturation	98	30
		Annealing	77-1^{o}C each cycle	45
		Elongation	72	120
2	25	Denaturation	98	30
		Annealing	67	45
		Elongation	72	120
Final Extension	1	Final Extension	72	600
Hold	1	Hold	4	∞

Table 2.3: Touchdown PCR Program

2.1.4 Analysis and purification of PCR products

PCR products were prepared in 6X loading dye (ThermoScientific) and subjected to electrophoresis on 0.8-1.0 % agarose gels, containing ethidium bromide at 100V for 45-60 min. Fermentas Generuler 1kb ladder standards were run simultaneously with all PCR products. Gels were visualized under UV light using a UV illuminator. Bands of interest were excised using a surgical blade and purified using the Roche High Pure PCR Product Purification Kit [52].

2.1.5 Plasmid construction

Cloning of purified PCR products into the pLATE 11 vector was performed using the Thermo Scientific aLICator Ligation Independent Cloning Expression System[4], as specified by the manufacturer. 1 Unit of T4 DNA Polymerase was used to generate the 5' and 3' overhangs on PCR products in a 10 µl reaction volume, and reactions were quenched with EDTA after 5 min. The annealing reaction was then performed by adding the pLATE11 vector [50]. The resulting plasmid was then transformed into *E. coli* DH5α.

2.1.6 Transformation of plasmids into *E. coli* DH5α

For transformation, the entire reaction mixture (10 µl) from the previous step was mixed with DH5α competent cells (50 µl) and incubated on ice for 30 min. Cells then underwent 42^{o}C heat-shock treatment for 45 sec. They were then incubated on ice for 2 min followed by growth in LB at 37^{o}C with 225 rpm shaking for 1 h. Aliquots were then plated on LB + 100 µg/mL carbenicillin plates, and grown overnight (12-16 h) at 37 ^{o}C (Modified procedure from *Current Protocols in Molecular Biology*, 2003) [53].

2.1.7 Growth of recombinant *E. coli* cells and plasmid extraction/purification

A single colony from the transformation step was inoculated into 5 mL of LB + 100 µg/mL carbenicillin medium and grown overnight (12-16 h) at 37 ^{o}C and 225 rpm. Recombinant plasmids were extracted and purified using the Biobasic EZ-10 Spin Column Plasmid DNA MiniPreps Kit (BioBasic).[54]

2.1.8 Plasmid analysis via restriction enzyme digestion

Restriction digest analyses were performed using the appropriate NEB, Fermentas, or Promega restriction enzymes to cleave plasmids, followed by analysis of the resulting fragments by agarose gel electrophoresis. Restriction digests were performed according to the protocols provided by the restriction enzyme manufacturer. In all cases, this involved using 1-10 Units of restriction enzyme in the appropriate buffer to cleave 50-150 ng of plasmid DNA. Expected fragment sizes were determined by inputting plasmid and gene sequence into the NEBcutter V2.0 server[55], and compared with the estimated sizes of the fragments observed after agarose gel electrophoresis.

2.1.9 Plasmid analysis via PCR

Fragment insertion into the cloning vector was also checked by PCR amplification of the gene from plasmids using the pLATE11 sequencing primers provided with the Thermo Scientific aLICator Ligation Independent Cloning Expression kit. PCR reactions were set up as indicated in Table 2.4 and the PCR program utilized is presented in Table 2.5. The resulting PCR products were then observed via agarose gel electrophoresis using the same procedure as described above.

Component	Volume (µL)
10X Thermopol Reaction Buffer	2.0
1mM DNTP's	4.0
10 µM Forward LIC Sequencing Primer	0.4
10 µM Reverse LIC Sequencing Primer	0.4
Plasmid DNA	0.5
Taq DNA polymerase	0.1
Nuclease free water	12.6

Table 2.4: Plasmid Analysis PCR Reaction Components (20 µl)

Step	Temperature (^{o}C)	Time (s)	Cycles
Initial denaturation	98	180	1
Denaturation	98	30	25
Annealing	58	45	
Elongation	72	120	
Final Elongation	72	600	1
Hold	4	∞	1

Table 2.5: Plasmid Analysis PCR Program

2.1.10 Recombinant plasmid analysis via DNA sequencing

Plasmids were sequenced at the McGill University and Genome Quebec Innovation Centre using the pLATE11 forward and reverse sequencing primers provided with the cloning kit. Trimmed sequencing results representing the (reliable portions) of the forward and reverse complementary sequences were Blasted against each other to find the overlap region, which was deleted. The remaining sequences were then used to generate the full sequence, which was compared to the genomic sequences. In cases where there appeared to be missing residues at

either ends of the sequence the untrimmed sequences were used. In cases where differences were observed the chromatogram was manually interpreted to decide whether the difference was real or not.

2.2 Protein Production, Analysis and Purification

2.2.1 Transformation of plasmids into *E. coli* expression strain

Plasmids (1 µL, of 50-150 ng/µL, depending on the plasmid) were transformed into *E. coli* BL21 DE3[56] or Rosetta DE3[57] competent cells (50 µL). After addition of plasmid, cells were placed on ice for 30 min. Cells then underwent a 42 ^{o}C heat shock treatment for 45 s, followed by growth in LB at 37 ^{o}C with 225 rpm shaking for 1 h. Cells were then plated on LB + 100 µg/mL carbenicillin plates for *E. coli* BL21 DE3, and LB + 100 µg/mL carbenicillin + 30 µg/mL chloramphenicol plates for Rosetta DE3. Cells were then grown overnight (12-16 h) at 37 ^{o}C [56].

2.2.2 Growth of recombinant *E. coli* cells for protein expression

Single colonies from fresh transformation plates were inoculated into 5 mL of LB + 100 µg/mL carbenicillin (+ 30 µg/mL chloramphenicol for Rosetta DE3) medium and grown overnight (12-16 h) at 37 ^{o}C and 225 rpm shaking. Then 1 mL of overnight culture was added for every 100 mL of LB + 100 µg/mL ampicillin (+ 30 µg/mL chloramphenicol for Rosetta DE3) medium, and the culture was grown to an OD_{600} of 0.6. Protein expression was induced by the addition of 0.5 mM IPTG followed by continued incubation for 24 h at 16^{o}C.[56]

2.2.3 Harvesting and cell lysis of recombinant *E. coli* cells

Cells were pelleted by centrifugation at 8 500 x g at 4 ^{o}C for 10 min, and the resulting supernatant was removed. Cells were washed by resuspending with 50 mM Tris-HCl pH 7.6 buffer, followed by pelleting at 8 500 x g for 10 min: supernatant was removed. Cell pellets were then stored at -20 ^{o}C or used immediately. For lysis, cell pellets were resuspended in 50 mM Tris-HCl pH 7.6, at a 2 mL buffer per gram ratio, and sonicated using a Biologics Model 300 V/T Ultrasonic homogenizer at maximum power for 10 x 10 s, taking 15 s pauses in between each 10 s of sonication. Cell lysate was centrifuged at 40 000 x g for 10 min and supernatant ("crude extract") was retained. Centrifugation was repeated two more times in order to remove

as much cell debris as possible. Cells were kept on ice or at 4 ^{o}C at all times during the lysis procedure.

2.2.4 Analysis of proteins via SDS-Electrophoresis (SDS-PAGE)

Proteins were analyzed via SDS-PAGE on a Bio Rad electrophoresis system using the method of Laemmli[58]. Samples were electrophoresed on 10 % acrylamide gels for 45 min at 200 V in an electrophoresis tank with SDS running buffer (25 mM Tris, 192 mM glycine, 3.5 mM SDS). Samples were prepared with 2X loading buffer (125 mM Tris-HCl pH 6.8, 20% glycerol, 4% SDS, 3% DTT or 2 % 2-mercaptoethanol, 0.001% bromophenol blue)

2.2.5 General strategy for purification of enzymes

Flavin adenine dinucleotide (FAD) from Sigma was added in at least 2-fold excess (200 μM - 800 μM FAD) to recombinant flavoprotein-containing crude extracts before purifications, unless noted otherwise. Enzymes were purified from crude extracts using a series of chromatographic columns. The following columns were used, in order: DEAE Sepharose Fast Flow, Phenyl-Sepharose, Sephacryl S-300 gel filtration, and a second DEAE column if necessary (Table 2.6). In between the DEAE and Phenyl-Sepharose steps there was an ammonium sulfate precipitation step. During all purification steps, enzyme preparations were kept at 0-4 ^{o}C and in 50 mM Tris-HCl pH 7.6 buffer, unless otherwise stated.

2.2.6 Procedure for DEAE (anion exchange) column chromatography

The DEAE column (2.6 cm x 25 cm) was equilibrated with 50 mM Tris-HCl pH 7.6 buffer (or 100 mM Tris-HCl pH 7.6 in the case of *A. niger* 767). Crude extract was loaded onto the column at a flowrate of 2 mL/min. The column was washed with the same equilibration buffer at a flow rate of 2 mL/min. Proteins were eluted with a 0-500 mM NaCl gradient over 500 mL of equilibration buffer at a flow rate of 2 mL/min. Eluate was collected in 10 mL fractions. Proteins were located in fractions by measuring OD_{280} as well as OD_{450} for flavoproteins, and performing activity assays (see below), as well as analysis on 10 % SDS-polyacrylamide gels. Fractions containing the purest protein of interest were combined and carried on to the next step.

2.2.7 Concentration of pooled fractions via ammonium sulfate precipitation

Ammonium sulfate was added to the pooled DEAE fractions up to 70 % saturation (or 60 % in the case of *A. niger* 999 and *A. niger* 534), stirred for 30 min, and then the precipitated protein was collected by centrifugation at 8 500 x g for 10 min. Following removal of supernatant, protein precipitates were dissolved in 50 mM Tris-HCl pH 7.6 buffer, containing 20 % ammonium sulfate, and centrifuged to remove any remaining particulate material. The resulting supernatant then was ready to be loaded onto the Phenyl-Sepharose column.

2.2.8 Procedure for Phenyl-Sepharose (hydrophobic interaction) column chromatography

The Phenyl-Sepharose column (2.6 cm x 14 cm) was pre-equilibrated with 50 mM Tris-HCl pH 7.6 buffer containing 20 % ammonium sulfate, at a flow rate of 1 mL/min. Samples from the ammonium sulfate precipitation step were loaded onto the column, which then was washed with buffer containing 20 % ammonium sulfate at a flow rate of 1 mL/min and eluted with a decreasing ammonium sulfate gradient (to 0 % ammonium sulfate) over 450 mL. Proteins were located in fractions by measuring OD_{280}, OD_{450} (for flavoproteins), and performing activity assays (see below), as well as analysis on 10 % SDS-polyacrylamide gels. The purest fractions were combined and concentrated for the next step using a Millipore ultrafiltration membrane with a 30 000 Da cut-off.

2.2.9 Procedure for Sephacryl S-300 (size exclusion) column chromatography

The Sephacryl S-300 column (2.6 cm x 78 cm) was equilibrated with 50 mM Tris-HCl, pH 7.6 buffer containing 100 mM NaCl. The sample from the previous step was loaded at a flowrate of 1 mL/min, and proteins were eluted with the same buffer (also at a flowrate of 1 mL/min). Eluted fractions were collected and the enzyme was located as described for the previous steps. The purest fractions were combined and concentrated and then stored at -20 ^{o}C (short term) or -80 ^{o}C (long term). Some aliquots were stored in 10 % glycerol for increased stability.

2.2.10 Second DEAE chromatography step

In some cases (Table 2.6), a second DEAE chromatography step was added after the Sephacryl S-300 step. The DEAE column (1.6 cm x 15 cm) was equilibrated with 50 mM Tris-

HCl pH 7.6 buffer. Crude extract was loaded onto the column at a flowrate of 1 mL/min. The column was washed with the same equilibration buffer at a flow rate of 1 mL/min. Proteins were eluted with a 0-500 mM NaCl gradient over 300 mL of equilibration buffer at a flow rate of 2 mL/min. Eluate was collected in 5 mL fractions. Proteins were located in fractions by measuring OD_{280}, OD_{450} for flavoproteins, and performing activity assays (see below), as well as analysis on 10 % SDS-polyacrylamide gels. Fractions containing the purest protein of interest were combined and stored at -20 ^{o}C for short term storage and -80 ^{o}C for long term storage.

A. niger Protein ID	Purification Step				
	DEAE 1	Ammonium Sulfate Precipitation	Phenyl-Sepharose	Sephacryl S-300	DEAE 2
767	X	X	X		
577	X	X	X	X	X
716	X	X	X	X	
999	X	X	X	X	
534	X	X	X		

Table 2.6: Purification steps used for each enzyme preparation. The protein ID is based on the JGI ID number.

2.2.11 TCA precipitation of enzymes in preparation for protein quantification

In preparation for quantification, proteins were precipitated with trichloroacetic acid[59]. For this, 50 µL samples were prepared, with appropriate dilutions. Then, 1 mL of distilled H_2O and 100 µL of 0.15% sodium deoxycholate was added to each sample and mixtures were incubated for 15 min at room temperature. After incubation, 100 µL of trichloroacetic acid (72 %) was added to each sample, followed by mixing, and centrifugation at 13 000 rpm for 15 min at room temperature. Supernatants were removed and 50 µL of 5 % SDS in 0.1 M NaOH was added to each tube followed by careful vortexing.

2.2.12 Estimating protein concentrations

Samples (50 μL) and standards of appropriate dilutions were prepared in 1.5 mL microtubes. Standards (bovine serum albumin, BSA) were prepared according to the instructions provided for the Pierce BCA Protein Assay Kit. Samples and standards were treated via TCA precipitation as described above to remove any interfering compounds. Then, 1 mL of BCA working reagent (50 parts Pierce BCA Reagent A + 1 part of Pierce BCA Reagent B) was added to each tube and the contents were mixed. Reaction mixtures were incubated at room temperature for 2 h, and then their absorbances were measured at 562 nm. A standard curve was prepared by plotting the corrected A_{562} values vs. BSA standard concentrations. Concentrations of unknowns were then determined using the standard curve and applying any dilution factors necessary.

2.3 Enzyme Characterization Experiments

2.3.1 Identifying bound flavins

Protein bound flavin samples were prepared by boiling *A. niger* 767, 577, 534, and 999 to release any non-covalently bound flavin and centrifuging out the protein precipitate. These flavin sample supernatants for the 4 enzymes were spotted (10-20 times) on a Whatman silica-coated fluorescent TLC plate together with genuine FAD and FMN samples. The chromatogram was developed using a solvent of n-butanol: acetic acid: H_20 (4:1:5)[60]. Since FAD and FMN can be detected by their yellow color, no visualization procedure was required. However spots could also be visualized under UV light (254 nm).

2.3.2 Estimating number of FAD binding sites in purified proteins

The number of FAD binding sites was determined using the following equation:

$$\#\text{ of FAD Active Sites} = \frac{mol\ of\ FAD}{mol\ of\ Protein}$$

The concentration of enzyme was determined using the BCA protein assay described above. The concentration of FAD was estimated by first releasing FAD from the protein via denaturation at 100 ^{o}C for 5 min, then centrifuging the sample to remove precipitated protein, and finally

measuring the OD_{450} of the resulting supernatant. The concentration of FAD was then calculated using the extinction coefficient of 11 300 $M^{-1}cm^{-1}$.[31]

2.3.3 Preparation of substrates

All substrates were the purest available commercially and prepared in autoclaved distilled water or milli-Q, water to a concentration of 10 mM. If substrate was not fully soluble in water, a few microliters of 1 M NaOH was added until the substrate dissolved. NADH or NADPH were prepared in 2 mM Tris base and the concentration of NAD(P)H was verified by measuring the absorbance at 340 nm ($\varepsilon = 6\ 220\ M^{-1}cm^{-1}$) [12].

2.3.4 Enzymatic assays for hydroxylases and decarboxylases

Enzymatic activities of hydroxylases were monitored at 340 nm to follow NAD(P)H consumption. An assay mixture (1 mL) consisted of 200 µM NAD(P)H, and 200 µM aromatic substrate, in 50 mM Tris-HCl pH 7.6 buffer at 25°C. The amount of enzymes added to the assay mixtures were as follows for each *A. niger* preparation: 6.59 µg 767, 7.37 µg 577, 34.9 µg 534, or 15.0 µg 999). All reaction components except enzyme were mixed initially, and absorbance was monitored for a few seconds until a stable baseline was established: then enzyme was added and absorbance changes were recorded. Control assays without substrate were carried out to be able to subtract the consumption of NAD(P)H in the absence of aromatic substrate.

The enzymatic assay for the *A. niger* 2,3-dihydroxybenzote decarboxylase (*A. niger* 716) was carried out by measuring the rate of 2,3-dihydroxybenzoic acid consumption at 305 nm[38]. The assay (1 mL) contained 300 µM 2,3-dihydroxybenzoic acid (2,3-DHBA) and 16.27 µg of enzyme in 20 mM Bis-Tris pH 6.2 buffer at 25 °C. The buffer and substrate were combined and the absorbance of the mixture was monitored for a few seconds to ensure a stable baseline was achieved, then enzyme was added and the reaction was followed at 305 nm where the extinction coefficient of 2,3-dihydroxybenzoic acid at 305 nm is 3340 $M^{-1}cm^{-1}$ [38].

The reverse reaction of the decarboxylase was monitored by measuring the rate of 2,3-DHBA production at 305 nm when catechol was used as a substrate[38]. The assay (1 mL) consisted of 300 µM catechol, 0.25-1.00 M $KHCO_3$ (potassium bicarbonate) in 100 mM phosphate buffer pH 6.2 at 25 °C. The absorbance of the assay mixture was then measured for a

few seconds in order to ensure a stable baseline was achieved, and then finally 81.4 μg of *A. niger* 716 was added and the reaction was monitored at 305 nm until completion.

2.3.5 Assay for catechol production

Catechol production was observed by coupling reactions to catechol 2,3-dioxygenase[61,62]. In this assay, absorbance was measured at 400 nm to monitor 2-hydroxymuconate-6-semialdehyde formation. Assays were carried out in 50 mM Tris-HCl pH 7.6 buffer (1 mL) containing 200 μM NAD(P)H, monooxygenase enzyme (for each *A. niger* enzyme preparation 16.5 μg 767, 18.4 μg 577, 34.9 μg 534, or 15.0 μg 999) and *Pseudomonas CF600* catechol 2,3-dioxygenase (2 Units) at 25 ^{o}C. After a baseline absorbance was established, salicylate or phenol (200 μM) was added, and activity was measured from the change in absorbance at 400 nm.

In addition, catechol 2,3-dioxygenase was observed to be active on 4-methylcatechol, 3-methylcatechol, and hydroxyquinol, so the enzyme assay described above was used when there was a possibility of any of these compounds being produced, with any of the four hydroxylase enzymes investigated.

2.3.6 Sample preparation for product analysis by TLC

Standards of substrates and expected products were prepared in ethyl acetate at a 10 mM concentration. Reaction mixtures to generate products contained 200 μM aromatic substrate, 600 μM NAD(P)H, enzyme (13.2 μg 767 or 14.7 μg 577) in 50 mM Tris-HCl pH 7.6 buffer (1 mL), and were run for at least 10 min at 25 ^{o}C. A few drops of 6 M HCl were then added to the reaction mixtures to protonate any products and precipitate enzyme. Reaction mixtures were then centrifuged in a microfuge at maximum speed for 5 min, and the supernatant was removed and decanted into glass test tubes. Supernatants were extracted two times with 2 mL of ethyl acetate. Combined ethyl acetate extracts were then taken to dryness by evaporation under nitrogen gas, and residues were dissolved in 100 μL of ethyl acetate prior to TLC analysis.

2.3.7 TLC product analysis

Mobile phase consisting of chloroform: ethyl acetate: acetic acid (50:50:1)[63] was poured into a TLC chamber lined with a pierce of filter paper . TLC plates used were Whatman silica

coated TLC plates with fluorescent indicator. Samples were spotted (10-20 times) 1 cm from the bottom of the TLC plate using finely-drawn capillary tubes. TLC plates then were placed into the chamber and run until the solvent front reached a few millimeters from the top. The plate was then visualized under UV light (254 nm) and spots were marked using a pencil. R_f values (or retention factors) were calculated for each spot using the formula[64]:

$$R_f = \frac{distance\ of\ compound\ from\ origin}{distance\ of\ solvent\ front\ from\ origin}$$

2.3.8 Oxygen monitoring experiments

Reaction mixtures (1 mL) contained 50 mM Tris-HCl pH 7.6 buffer with 200 μM NADH, 300 μM aromatic substrate, and enzyme (for each *A. niger* protein 65.9 μg 767, 36.9 μg 577, 69.8 μg 534, or 150 μg 999). Reaction conditions were modified from the conditions used by White-Stevens et al[29]. Oxygen concentrations in these assays were measured using a Hansatech oxygen electrode (K140) Oxygraph + oxygen monitoring system.[65] The buffer was initially added to the chamber, and the oxygen concentration was monitored as NADH and aromatic substrate were added sequentially using a syringe. Hydroxylase enzyme was then added, and the decrease in oxygen concentration was observed. Once the oxygen concentration reached a plateau, 200 Units of catalase (Sigma-Aldrich) was added and the reaction was followed until oxygen production ceased. Calibration of the instrument was achieved by following the calibration procedure outlined in the instrument manual using air-saturated water and water with excess sodium dithionite (to completely consume O_2) added. All reactions and calibration were performed at 25 ^{o}C at ambient pressure. Atmospheric pressure was obtained from a weather website: (http://www.theweathernetwork.com/ca/weather/quebec/montreal), and was always around 101.3 kPa.

2.3.9 Substrate binding via monitoring FAD fluorescence, determining K_d

Determination of the K_d values for binding of aromatic substrate to *A. niger* hydroxylases were determined by measuring changes in FAD fluorescence emission at 525 nm, with an excitation wavelength of 450 nm, upon the addition of substrate. Successive substrate additions were made into 1 mL 50 mM Tris-HCl pH 7.6 buffer containing enzyme (5 μM) at ambient temperature. All measurements were conducted using a Varian Cary Eclipse fluorescence

spectrometer. Recorded fluorescence spectra were an average of 10 scans recorded at 600 nm/min with slit sizes of 5 nm and 10 nm for excitation and emission slit sizes respectively. The results were then plotted as $\Delta FI/\Delta FI_{max}$ versus the concentration of substrate, where ΔFI is defined as the change in fluorescence from the initial fluorescence and ΔFI_{max} is defined as the largest fluorescence change (at the highest substrate concentration). Florescence data were corrected for dilution by multiplying measured fluorescence emission by final volume/initial starting volume. K_d was then determined using the OriginLab software and fitting the data to the Hill1 equation (shown below):

$$y = Start + (End - Start)\frac{x^n}{k^n + x^n}$$

where "start" is the x value at the lowest substrate concentration and "end" is the x value of the highest substrate concentration, x is the concentration of substrate, k is the K_d, and n is the Hill coefficient.[15]

2.3.10 Molecular weight and oligomeric state determination of proteins using gel filtration

Gel filtration experiments were performed using the Superdex 200 10/300 GL column connected to a BioRad BioLogic DuoFlow HPLC. The column was equilibrated with 50 mM Tris-HCl, 100 mM NaCl pH 7.6 buffer, and eluted at a flow rate of 0.65 mL/min. All protein elution volumes were determined using a UV detector that measured OD_{280} as samples eluted from the column. The void volume was determined by measuring the elution volume of blue dextran (500 µL, 2 mg/mL). Elution volumes were also determined for Bio-Rad gel filtration standards (500 µL), consisting of thyroglobulin, γ-globulin, ovalbumin, myoglobin, and vitamin B12.[66] A standard curve of K_{av} vs $\log(M_w)$ was then plotted using elution volumes of the molecular weight standards. K_{av} was calculated using the equation[17]:

$K_{av}=\frac{V_e-V_0}{V_t-V_0}$, where V_e= elution volume, V_t= elution volume of smallest standard, V_0= void volume.

Purified *A. niger* proteins (0.5-1 mg, 500 µL) were then loaded onto the column and their elution volumes were measured. The K_{av} values of these unknowns were calculated, and their molecular weights were estimated based on the standard curve. In addition the oligomeric state of the

enzyme was determined by dividing the observed molecular weight by the expected mass of a single monomer based on the amino acid sequence.

2.3.11 pH-activity profile of 2,3-dihydroxybenzoate decarboxylase

The standard assay described above was used with buffers of varying pH. All buffers consisted of 20 mM Bis-Tris and were pH adjusted with HCl to pH values within a range of 5.8-7.2.

2.3.12 Metal chelation of *A. niger* 2,3-Dihydroxybenzoate decarboxylase

A protein sample (0.9 mL of 0.2 mM *A. niger* 2,3-dihydroxybenzoate decarboxylase) was inserted into a ThermoFisher Scientific Slide-A-Lyzer dialysis cassette (3 mL limit, 20 kDa cut-off), and dialyzed against 500 mL chelation buffer (50 mM Tris-HCl, 50 mM imidazole, 50 mM 8-hydroxyquinoline-5-sulfonic acid and 50 mM EDTA pH 7.4) at 4 ^{o}C with constant stirring. Dialysis was continued for 2 h followed by transfer to fresh dialysis buffer, and additional dialysis overnight. The following day the cassette was transferred to 500 mL 50 mM Tris-HCl buffer pH 7.6 (to remove chelator) and dialyzed for 2 h. The dialysis buffer was changed one final time, and dialysis was continued overnight to remove any remaining chelator (procedure adapted from Atkin et al, 1973[67]). Additionally, 50 mM Tris-HCl buffer pH 7.6 (control) also underwent dialysis simultaneously using the same methodology. The sample was removed from the dialysis cassette using a syringe and stored at -20 ^{o}C. Loading and removal of the sample from Slide-A-Lyzer dialysis cassette was performed following the manufacturer's directions[68].

2.3.13 Metal detection of enzymes via ICP-MS

All materials used in preparation of samples for ICP-MS were soaked in 10 % low trace metal nitric acid and rinsed with Milli-Q water. 50 mM Tris-HCl pH 7.6 and nitric acid solutions were prepared in Milli-Q water as well. For ICP-MS analysis samples containing 0.1 mg of protein were prepared: samples were then digested in 10 % HNO_3 at 50^{o}C for 2 days[69]. Digested samples were centrifuged in a microcentrifuge at max speed for a few minutes to remove all of the digested protein. Supernatants were transferred to a nitric acid treated centrifuge tube and brought to the Centre for Biological Applications of Mass Spectrometry.

Buffer controls were prepared in the same manner. The samples were run by Alain Tessier on the LC-ICP-MS (Agilent 7500ce). Additionally, standards of known metal concentrations were prepared and measured to be able to prepare a standard curve which would then be used to measure metal content of the unknowns. Metals investigated were Mg, Ca, Mn, Fe, Co, Cu, and Zn.

2.4 Bioinformatic and Computational Studies

2.4.1 Construction of homology models

Protein models were generated using the I-TASSER Protein Structure and Function prediction server (default parameters) by submitting the primary amino acid sequence.[70] I-TASSER models provide only the coordinates of the apo-enzyme. For a model with bound co-factor and/or substrate, these molecules were added manually. This was performed by finding a similar homologue with the molecule of interest, aligning the model to the homologue in PyMOL and then copying the new coordinates of the co-factor or substrate from the aligned template to the model being constructed.[71]

2.4.2 Sequence alignments and construction of phylogenetic trees

Sequence alignments and phylogenetic tree construction was performed by submitting query sequences to the Clustal Omega server[72,73]. In all cases, protein sequences of interest were aligned to determine similarity in specific regions. In many cases, other characterized enzymes (retrieved from UniProt) of the same class were included in the alignment. In addition the sequences of enzymes of interest were included in phylogenetic trees to help visualize the similarity between specific enzymes, and help to rationalize possible enzyme characteristics, such as the identity of their substrates.

3. RESULTS

3.1 Selection of *A .niger* Aromatic Ring Hydroxylase Genes

The *Aspergillus niger* genome contains many genes annotated as aromatic ring hydroxylases. Upon searching the two publically-available *A. niger* genomes (ATCC 1015v4.0[22] and the NRRL3[23]) using text searches such as "monooxygenase", "hydroxylase", or "aromatic-ring hydroxylase" more than 60 hits were found (data not shown). Further inspection of these hits showed that some have very low sequence identity with known aromatic hydroxylase enzymes. Most of these hits to "known" aromatic hydroxylases were for salicylate hydroxylase or phenol hydroxylase. For the salicylate hydroxylases, genes encoding the top five *A. niger* sequences most similar to *P. putida* (bacterial) salicylate hydroxylase (Uniprot ID: Q23262[11]) and a less similar enzyme were selected. In the case of the phenol hydroxylase annotated sequences, genes encoding proteins most similar to *Trichosporon cutaneum* phenol hydroxylase (Uniprot ID: P15245[74]) were initially chosen, although as will be shown later, most of these were not successfully amplified by PCR. Additional genes annotated as phenol hydroxylase were thus chosen based on the relatively high expression levels observed in the cDNA that was used as template (unpublished data).

The list in Table 3.1 represents all 14 predicted hydroxylases chosen for the experimental work in this thesis, and summarizes the best Blast hit from the Uniprot/Swissprot databases for each of the predicted sequences. Interestingly, the Blast results indicate that very few of the *A. niger* hydroxylase sequences showed identity above 35 % to their bacterial counterparts, with the highest identity only being 38 %.

Blast searches also reveal superfamily domains, summarized in Table 3.2. All the protein targets contain a NADB_Rossmann superfamily domain which identifies the Rossmann fold, the site of dinucleotide cofactor binding which is observed in flavoproteins utilizing NAD(P)H/NAD(P)[75]. The protein sequences annotated as phenol hydroxylase have an additional predicted thioredoxin_like superfamily domain which is responsible for dimer contacts[76].

For the remainder of this thesis protein ID's will be abbreviated to the last 3 digits of the JGI protein ID. For example, *A. niger* 1185767 is abbreviated to *A. niger* 767.

JGI Protein ID	NRRL3 Annotation	Top Blast Hit	Organism	Query Cover (%)	E value	Identity (%)
1127577	Salicylate Hydroxylase	Salicylate 1-monooxygenase	*Pseudomonas putida*	91	2E-61	36
1205534	Salicylate Hydroxylase	Salicylate 1-monooxygenase	*Pseudomonas putida*	86	2E-52	33
1223684	Salicylate Hydroxylase	3-hydroxybenzoate 6-hydroxylase	*Polaromonas naphthalenivorans CJ2*	82	4E-41	31
1114655	Salicylate Hydroxylase	Salicylate 1-monooxygenase	*Pseudomonas putida*	90	3E-38	33
1183860	Salicylate Hydroxylase	Salicylate 1-monooxygenase	*Pseudomonas putida*	91	6E-66	36
1206422	Phenol Hydroxylase	Phenol 2-monooxygenase	*Trichosporon cutaneum*	98	5E-121	35
1210999	Phenol Hydroxylase	Phenol 2-monooxygenase	*Trichosporon cutaneum*	91	5E-90	33
52617	Phenol Hydroxylase	Phenol 2-monooxygenase	*Trichosporon cutaneum*	77	7E-46	28
1125300	Phenol Hydroxylase	Phenol 2-monooxygenase	*Trichosporon cutaneum*	94	3E-86	32
1125146	Phenol Hydroxylase	Phenol 2-monooxygenase	*Trichosporon cutaneum*	96	2E-81	30
1159718	Phenol Hydroxylase	3-hydroxybenzoate 4-hydroxylase	*Comamonas testosteroni*	97	1E-39	25
1227951	Phenol Hydroxylase	Phenol 2-monooxygenase	*Trichosporon cutaneum*	99	5E-110	34
1085687	Phenol Hydroxylase	Phenol 2-monooxygenase	*Trichosporon cutaneum*	97	7E-103	33
1185767	Aromatic-ring hydroxylase-like protein	Salicylate 1-monooxygenase	*Pseudomonas putida*	89	4E-75	38

*Blast results calculated using protein sequences available on the Uniprot/Swissprot database in September 2014

Table 3.1: List of all protein sequences chosen for this study

JGI Protein ID	Superfamily/Domains	
1127577	NADB_Rossmann superfamily	
1205534	NADB_Rossmann superfamily	
1223684	NADB_Rossmann superfamily	
1114655	NADB_Rossmann superfamily	
1183860	NADB_Rossmann superfamily	
1206422	NADB_Rossmann superfamily	Thioredoxin_like superfamily
1210999	NADB_Rossmann superfamily	Thioredoxin_like superfamily
52617	NADB_Rossmann superfamily	Thioredoxin_like superfamily
1125300	NADB_Rossmann superfamily	Thioredoxin_like superfamily
1125146	NADB_Rossmann superfamily	Thioredoxin_like superfamily
1159718	NADB_Rossmann superfamily	Thioredoxin_like superfamily
1227951	NADB_Rossmann superfamily	Thioredoxin_like superfamily
1085687	NADB_Rossmann superfamily	Thioredoxin_like superfamily
1185767	NADB_Rossmann superfamily	

Table 3.2: Predicted superfamily domains of target proteins

3.2 Salicylate Hydroxylases

3.2.1 PCR amplification to purification

Predicted salicylate hydroxylase gene sizes are listed in Table 3.3 below, together with protein molecular weight estimated using the Protparam[77] tool which predicts protein characteristics by inputting the amino acid sequence.

A. niger JGI ID	Expected gene size (base pairs)	Expected protein molecular weight (kDa)
1127**577**	1311	48.4
1185**767**	1314	47.9
1205**534**	1317	48.2
1223**684**	1365	47.2
1114**655**	1332	48.4
1183**860**	1275	47.0

Table 3.3: Expected DNA and protein sizes of putative salicylate hydroxylases

Obtaining protein to characterize enzymes requires successful PCR amplification from template DNA, cloning, transformation, expression, and purification. Any one of these steps can

fail, for various reasons. Shown in Table 3.4 is the success of each step for the putative salicylate hydroxylase sequences of interest. The symbol (✓) in each cell indicates that the step was successful, while the symbol (×) cell indicates that the specific step was not successful.

	Steps Achieved				
***A. niger* JGI ID**	**PCR Amplification**	**Cloning**	**Expression**	**Purification**	**Characterization**
1127**577**	✓	✓	✓	✓	✓
1185**767**	✓	✓	✓	✓	✓
1205**534**	✓	✓	✓	✓	✓
1223**684**	✓	✓	✓	×	
1114**655**	✓	✓	✓	×	
1183**860**	✓	✓	×		

Table 3.4: Summary of steps achieved with *A. niger* putative salicylate hydroxylase

Three enzymes, *A. niger* 767, 577, and 534, were all successfully expressed and characterized. *A. niger* 684 was successfully expressed but not characterized any further since no hydroxylase activity was observed. *A. niger* 655 was successfully expressed but did not express in the soluble fraction in any of the conditions tested. Finally, *A. niger* 860 did not express in either the soluble or insoluble fractions in any of the conditions attempted. These results will be presented in detail below.

3.2.1.1 PCR amplification of *Aspergillus niger* genes annotated as salicylate hydroxylases

Putative salicylate hydroxylase genes listed in Table 3.3 were subjected to amplification by PCR using cDNA derived from *A. niger* grown on 2% alfalfa and barley, as described in *Materials and Methods.* Amplified fragments all appeared to be the expected sizes (Fig. 3.1). The resulting DNA fragments were cloned into the pLATE 11 vector and the presence of inserted DNA fragments was screened via restriction digestion and/or PCR amplification (See examples below in Fig. 3.2 and Fig 3.3). To verify that the correct DNA insert was cloned into the pLATE11 plasmid, the recombinant plasmids were sequenced at the McGill University Genome Quebec Innovation Centre. All putative salicylate hydroxylase sequences in Table 3.3 were found to be identical to the expected JGI gene sequence.

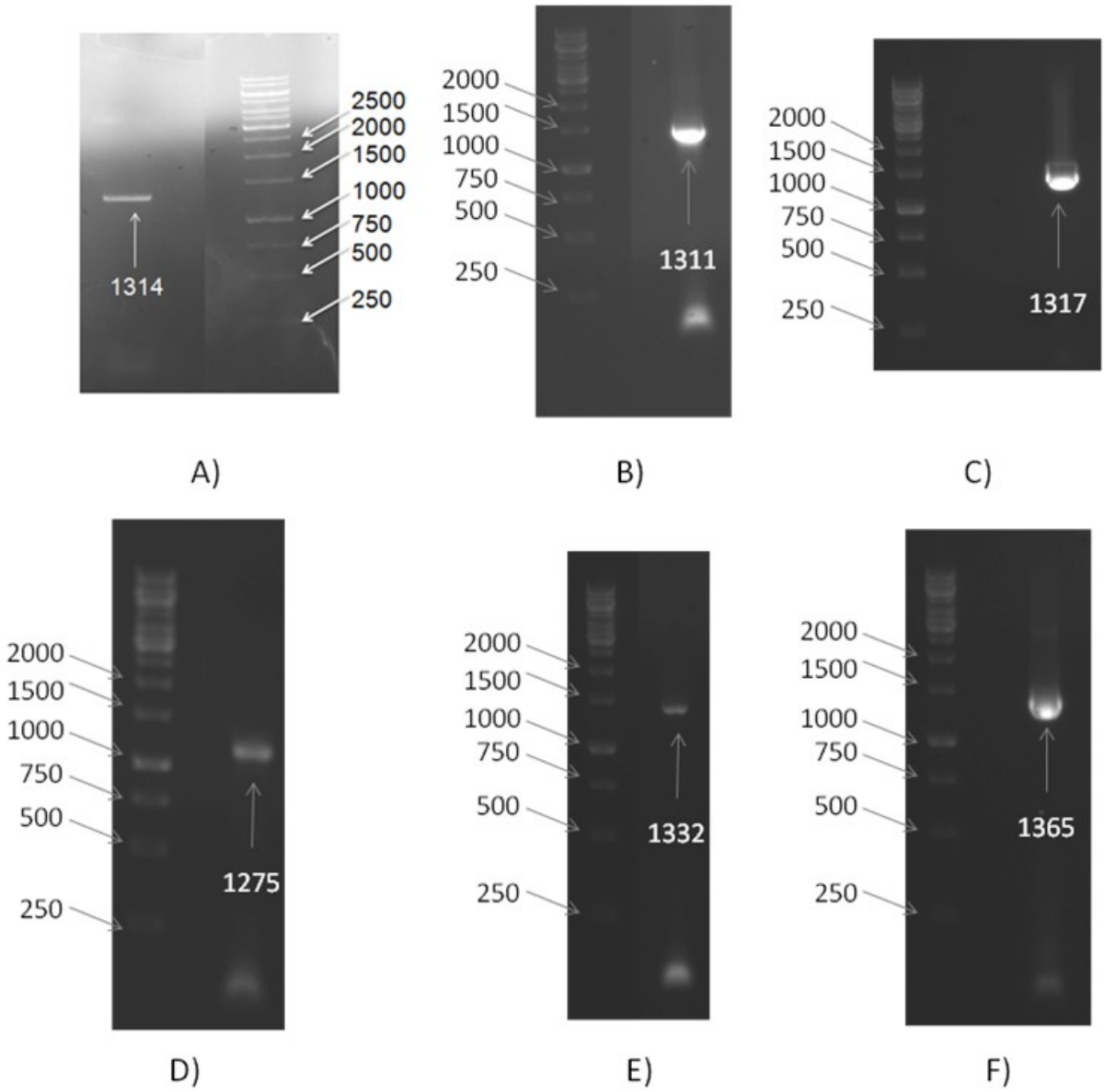


Figure 3.1: PCR Amplification of *A. niger* 767 (A), 577 (B), 534 (C), 860 (D), 655 (E), and 684 (F) run on 0.8% Agarose Gels. Numbers pointing to bands represent base pair sizes, where numbers pointing to PCR fragments represent *theoretical* base pair sizes.

3.2.1.2 Plasmid screening for correct inserts

Shown in Fig 3.2 is an example of a restriction digest where a plasmid construct has two *NdeI* cut sites, one located in the core pLATE11 plasmid and the other in the gene insert, thus,

cleavage of a plasmid with the correct insert should result in two bands. The expected correctly sized bands appear to be present at around 2201 bp and 3565 bp, indicating the insert is the expected size, and in the correct orientation.

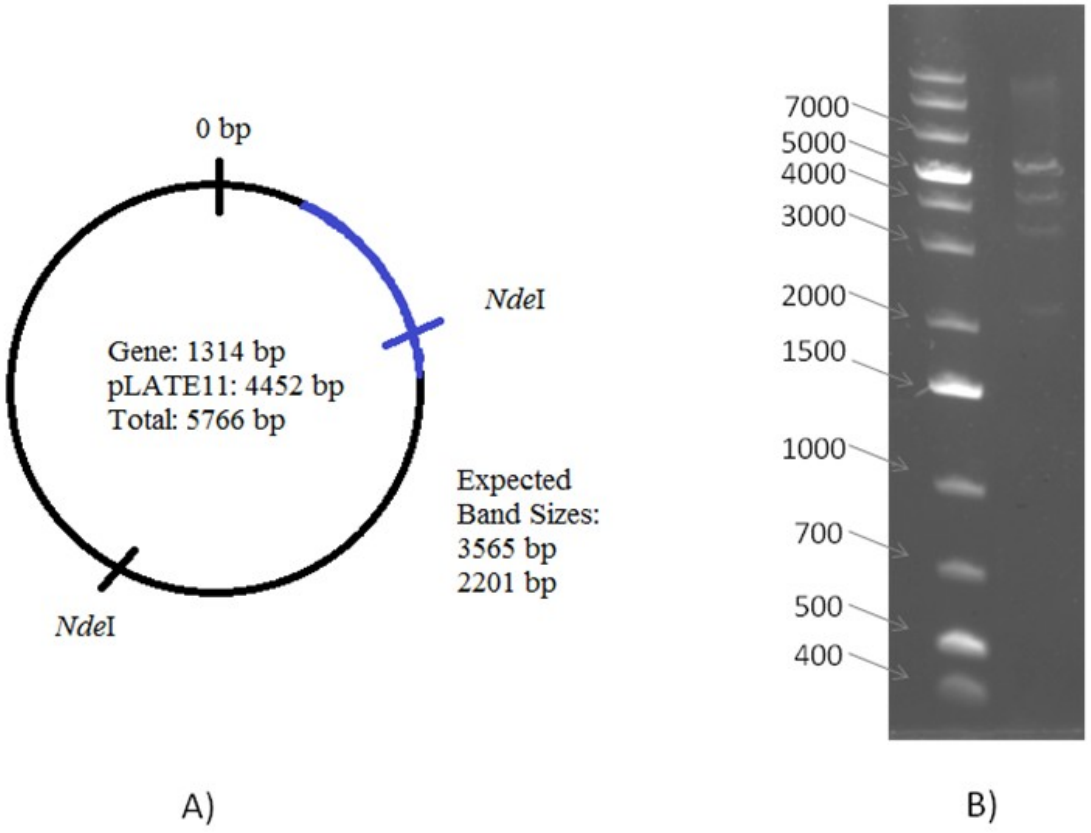


Figure 3.2: Example of plasmid screening using restriction digestion (A). Blue regions represent gene insert while black portion represents pLATE11. Cut sites are indicated in black by the *Nde*I label. *Nde*I restriction enzyme digest of *A. niger* 767 plasmid construct run on a 0.8% agarose gel (B). Similar tests were done with other plasmid constructs (data not shown).

A second method to screen plasmid DNA from single colonies of transformed cells involved using purified plasmids as PCR templates and amplifying inserts using the LIC sequencing primers. The sizes of the inserted fragments were then estimated on agarose gels, with some examples shown in Fig 3.3. Fragments amplified using this method were 178 bp larger than the size of the cloned gene since the LIC primers bind short distances upstream and downstream of the start and stop codons.

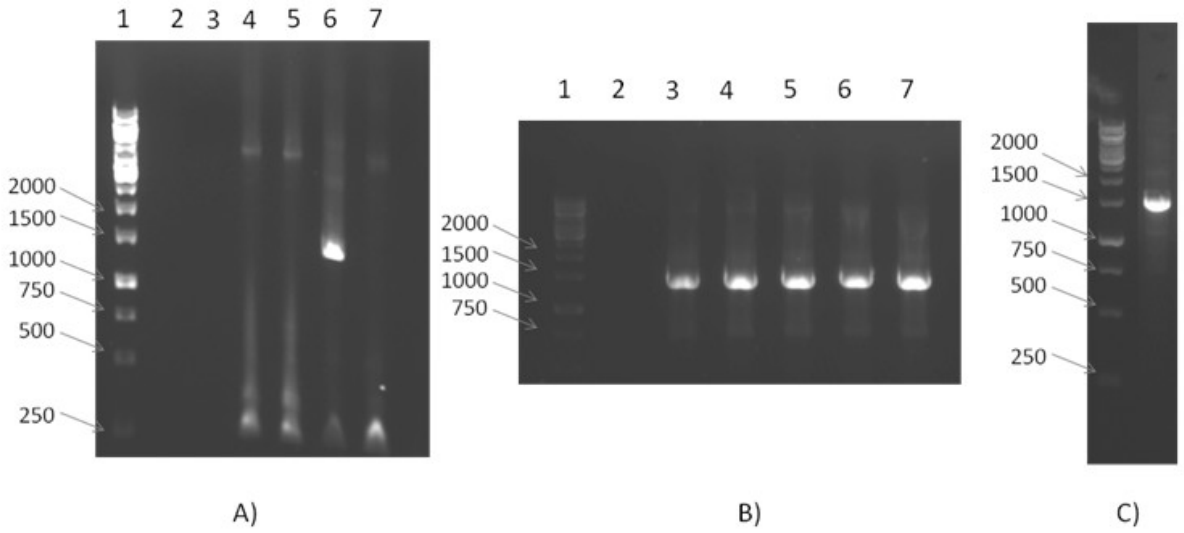


Figure 3.3: Screening plasmid inserts for constructs of *A. niger* 767 (A), 577 (B), and 534 (C). DNA fragments generated via PCR using plasmid as a template and amplifying insert fragment.

In Fig 3.3 the fragments visualized appear to be of the expected size taking into account these considerations (expected sizes of 1492 bp, 1489 bp and 1495 bp for *A. niger* 767, 577, and 534 respectively). As this method is used to screen for the correct insert, plasmids from all colonies do not always contain the correct insert. Thus, in Fig 3.3A only one colony (lane 6) contained a correctly-sized insert, while in Fig 3.3B plasmids from all colonies contained a correctly-sized insert.

Finally, the plasmids confirmed by the methods described above were then subjected to DNA sequencing to insure that only proteins with the correct intended sequence would be produced in subsequent steps.

3.2.1.3 Heterologous expression of salicylate hydroxylases in *E. coli*

Once the plasmid sequences were verified, protein expression was attempted by transforming the plasmid into *E. coli BL21* (DE3), growing the cells, inducting expression by the addition of IPTG, extracting the proteins, and running the samples on SDS polyacrylamide gels

with the appropriate controls. Both the soluble and pellet fractions were analyzed, in case the protein expressed in inclusion bodies. Induced samples were usually compared with samples from cells prior to addition of IPTG to see that the protein band was not simply a host protein. Different expression conditions and hosts were attempted when initial attempts to produce soluble protein proved to be unsuccessful.

As summarized in Figure 3.4, expression was successful for 5 of the 6 selected putative salicylate hydroxylases: SDS-polyacrylamide gel results are shown in Fig. 3.4. This figure is shown on the next page.

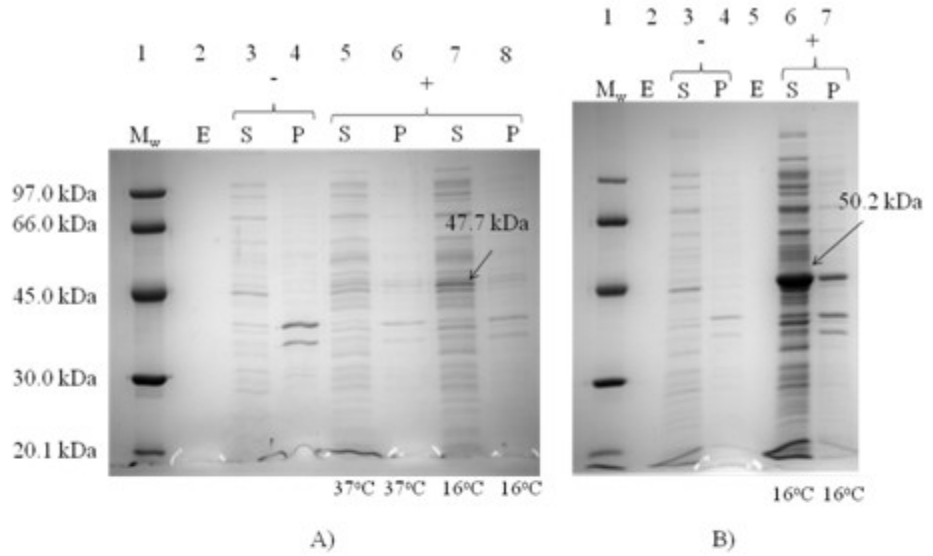

1 2 3 4 5 6 7 8
- +
Mw E S P S P S P
97.0 kDa
66.0 kDa
45.0 kDa
30.0 kDa
20.1 kDa
47.7 kDa
37°C 37°C 16°C 16°C
A)
1 2 3 4 5 6 7
- +
Mw E S P E S P
50.2 kDa
16°C 16°C
B)
1 2 3 4 5 6
+
Mw E S S P P
97.0 kDa
66.0 kDa
45.0 kDa
30.0 kDa
20.1 kDa
48.8 kDa
16°C 37°C 16°C 37°C
C)
1 2 3 4 5
- + - +
Mw S S P P
97.0 kDa
66.0 kDa
45.0 kDa
30.0 kDa
20.1 kDa
48.1 kDa
16°C 16°C
D)
1 2 3 4 5 6
- + - +
Mw E S S P P
97.0 kDa
66.0 kDa
45.0 kDa
30.0 kDa
20.1 kDa
48.9 kDa
16°C 16°C
E)

Figure 3.4: Soluble and insoluble fractions from BL21 (DE3) cells harbouring expression plasmids for *A. niger* 767 (A), 534 (B), 577 (C), 655 (D), and 684 (E) grown under various conditions. Samples were run on 10 % SDS-polyacrylamide gels. Lanes labelled M_w, E, S, and P represent molecular weight markers, empty lane, supernatant (soluble cell fraction), and pellet (insoluble cell fraction) respectively. Lanes with a negative sign (-) indicate samples from uninduced cells while a positive sign (+) indicates, that the cells have been induced. Lanes containing protein from induced cells have the induction temperature designated at the bottom of the lane. Finally, expressed target protein bands are indicated by an arrow pointing to them, with their estimated molecular weights listed alongside.

A. niger 767, 534, and 577 were all well expressed in soluble form (Fig. 3.4A-C). The *A. niger* 767 and 577 proteins were only expressed in soluble form when induction was done at 16 ^{o}C, thus all further expression trials were done at 16^{o}C. On the other hand, *A. niger* 655 protein was only expressed in insoluble form, even at 16 ^{o}C (Fig. 3.4D), and was not pursued further. The *A. niger* 684 target appeared to be produced in soluble form in small amounts (Fig. 3.4E), but crude extracts containing the protein showed no detectable activity on salicylate or any of the other related compounds tested using the NADH consumption assay described in *Materials and Methods*. The target *A. niger* 860 also did not show any expression under the conditions used for the above sequences (data not shown). Testing expression of these two targets in *E. coli* Shuffle Express [78] or Rosetta (DE3)[57] did not result in any expressed target protein even though the plasmid construct was deemed to be correct via DNA sequencing (data not shown). So these two targets were dropped from additional work.

3.2.1.4 Purification of three putative *A. niger* salicylate hydroxylases

The three successfully expressed and soluble putative salicylate hydroxylases, *A. niger* 767, 577, and 534, were each purified using combinations of column chromatography steps as outlined in *Materials and Methods.*

A. niger 767 (47.9 kDa) was purified to near homogeneity via a 3-step purification procedure consisting of the following steps: DEAE (ion-exchange chromatography), ammonium sulfate precipitation, and phenyl-Sepharose (hydrophobic interaction chromatography). The purity after each step was visualized after running samples on a 10 % SDS-polyacrylamide electrophoresis gel (Fig. 3.5). A summary of the overall purification is presented in Table 3.5. After the final step it is clear that nearly all contaminating *E. coli* proteins were removed (Fig 3.5, lane 6), and purification resulted in a specific activity increase of nearly 4 fold (Table 3.5).

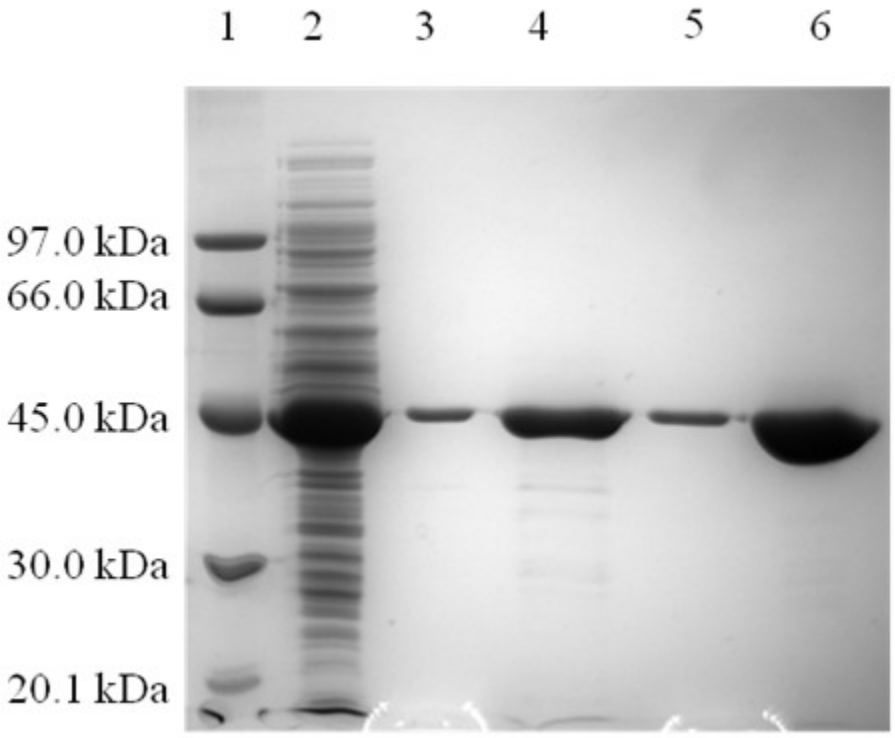


Figure 3.5: Purification of *A. niger* 767 as monitored after electrophoresis on a 10 % SDS-polyacrylamide gel. Protein low molecular weight standards (lane 1); crude extract (lane 2); pooled DEAE fractions (lane 3); ammonium sulfate concentrated DEAE fractions (lane 4). pooled phenyl-Sepharose fractions (lane 5); concentrated purified *A. niger* 767 (lane 6).

Purification Step	Volume (ml)	Concentration (mg/ml)	Activity (U/mL)	Total Protein (mg)	Total Activity (*Units)	Specific Activity (Units/mg)	Fold Purification	Activity Yield (%)
Crude Extract	26.0	16.9	23.9	439	621	1.41	-	100
DEAE pool	325	0.43	1.25	140	406	2.90	2.05	65.3
Ammonium sulfate precipitation	13.0	7.50	22.5	97.5	292	3.00	2.12	47.1
Phenyl-Sepharose Pool	45.0	0.84	4.13	37.8	186	4.91	3.47	29.9
Pure enzyme	4.50	6.59	33.8	29.7	152	5.13	3.63	24.5

*One unit (U) is defined as the amount of enzyme required to oxidize one μmol/min of NADH using salicylate as a substrate in 50 mM Tris-HCl pH 7.6 at 25 ^{o}C.

Table 3.5: Purification summary for *A. niger* 767

A. niger 577 (48.4 kDa) was purified in a similar fashion to *A. niger* 767. It was a 4-step purification in which the crude extract was chromatographically separated on a DEAE column, and continued with chromatography on a phenyl-Sepharose column, Sephacryl S-300 column (gel filtration), and finally a second DEAE column. Before being loaded on the phenyl-Sepharose column the protein was concentrated by ammonium sulfate precipitation (not shown). The purity of the protein preparation was monitored after each step (Fig. 3.6): Table 3.6 summarizes the purification. Note that activity and protein assays were not performed for the phenyl-Sepharose combined fractions because no aliquot of the pooled fractions was kept for analysis. The specific activity of the enzyme increased nearly 4 fold after the purification, and a nearly homogenous final preparation was achieved, although some minor contaminants are still visible (Fig. 3.6, lane 7). The second DEAE step does not significantly change the specific activity or improve the purity of the final preparation, and lowered the yield (although 25 % of the enzyme from the previous step was held back and is a contributor to the lower yield).

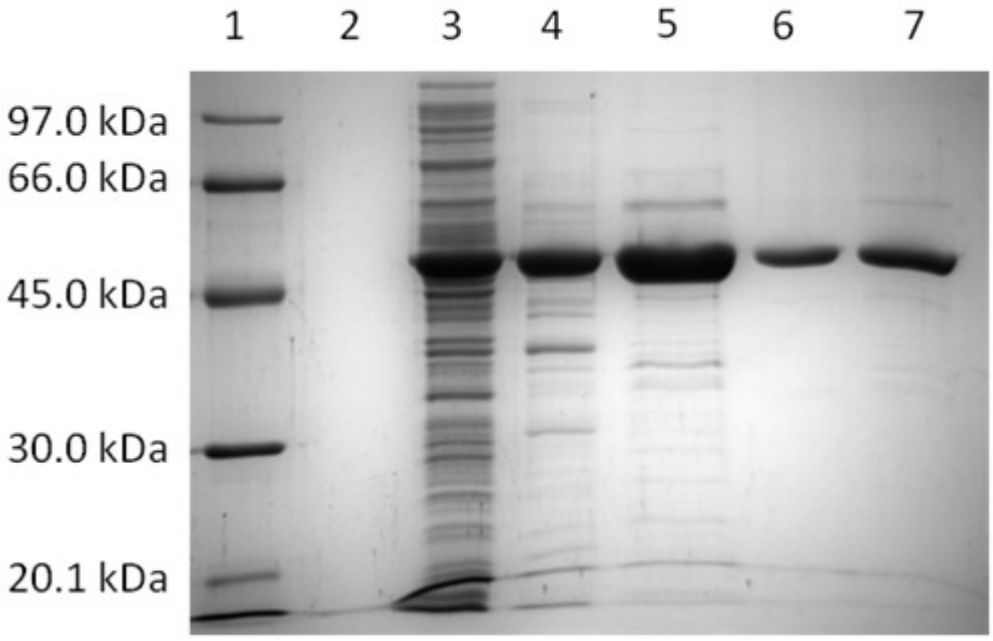


Figure 3.6: Purification of *A. niger* 577 as monitored after electrophoresis on a 10 % SDS-polyacrylamide gel. Protein low molecular weight markers (lane 1); empty well (lane 2); crude extract (lane 3); pooled DEAE fractions (lane 4); pooled phenyl-Sepharose fractions (lane 5); pooled Sephacryl S-300 fractions (lane 6); pooled second DEAE fractions (lane 7).

Purification Step	**Volume (ml)**	**Concentration (mg/ml)**	**Activity (U/mL)**	**Total Protein (mg)**	**Total Activity (*Units)**	**Specific Activity (Units/mg)**	**Fold Purification**	**Activity Yield (%)**
Crude Extract	60.0	12.5	85.5	751	5.13E+3	6.83	-	100
DEAE	112	1.94	43.3	217	4.85E+3	22.3	3.27	94.5
phenyl Sepharose	282	NA	NA	NA	NA	NA	NA	NA
S-300 (concentrated)	5.00	7.37	188	36.9	941	25.5	3.74	18.3
DEAE 2 (concentrated)	2.50	4.77	121	11.9	302	25.3	3.70	5.88

*One unit (U) is defined as the amount of enzyme required to oxidize one μmol/min of NADH using salicylate as a substrate in 50 mM Tris-HCl pH 7.6 at 25 ^{o}C.

Table 3.6: Purification summary for *A. niger* 577

A. niger 534 (48.2 kDa) was successfully purified via a 3-step purification procedure consisting of a DEAE column, ammonium sulfate precipitation, and finally a phenyl-Sepharose column. Fig. 3.7 and Table 3.7 summarizes the purification procedure and show that after purification the protein approaches near homogeneity resulting in approximately 3-fold increase in specific activity.

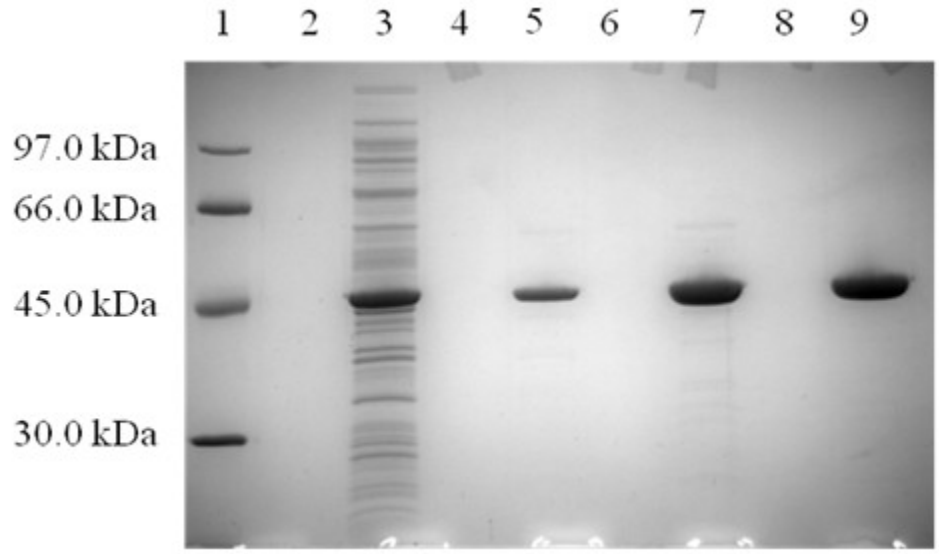


Figure 3.7: Purification of *A. niger 534* as monitored on a 10 % SDS-polyacrylamide gel. Protein low molecular weight markers (lane 1); crude extract (lane 3); pooled DEAE fractions (lane 5); ammonium sulfate concentrated DEAE fractions (lane 7); pooled phenyl-Sepharose fractions (lane 9)

Purification Step	Volume (ml)	Concentration (mg/ml)	Activity (U/mL)	Total Protein (mg)	Total Activity (*Units)	Specific Activity (Units/mg)	Fold Purification	Activity Yield (%)
Crude	32.0	21.1	3.01	676	96.3	0.143	1.00	100
DEAE + ASP	7.50	15.2	6.55	114	49.1	0.430	3.02	51.0
phenyl-Sepharose	4.00	14.0	6.06	55.8	24.2	0.434	3.05	25.2

*One unit (U) is defined as the amount of enzyme required to oxidize one μmol/min of NADH using salicylate as a substrate in 50 mM Tris-HCl pH 7.6 at 25^{o}C.

Table 3.7: Purification summary for *A. niger* 534

3.2.2: Salicylate Hydroxylases: Characterization

3.2.2.1 UV-visible absorbance characteristics and identification of the accompanying flavin

All known salicylate hydroxylases are flavoproteins which contain FAD (flavin adenine dinucleotide). FAD is a molecule that includes a chromophore with distinctive peaks at 375 nm and 450 nm, giving flavoproteins their distinct yellow color. UV-visible absorbance spectra characteristic of flavoproteins were observed for all 3 of the purified proteins (Fig 3.8). The peak at 280 nm is predominantly due to the protein's aromatic amino acids, but also flavin, while the other two peaks at 375 nm and 450 nm are due to bound flavin. The flavin peaks were identified as FAD, rather than FMN by TLC (Fig 3.9).

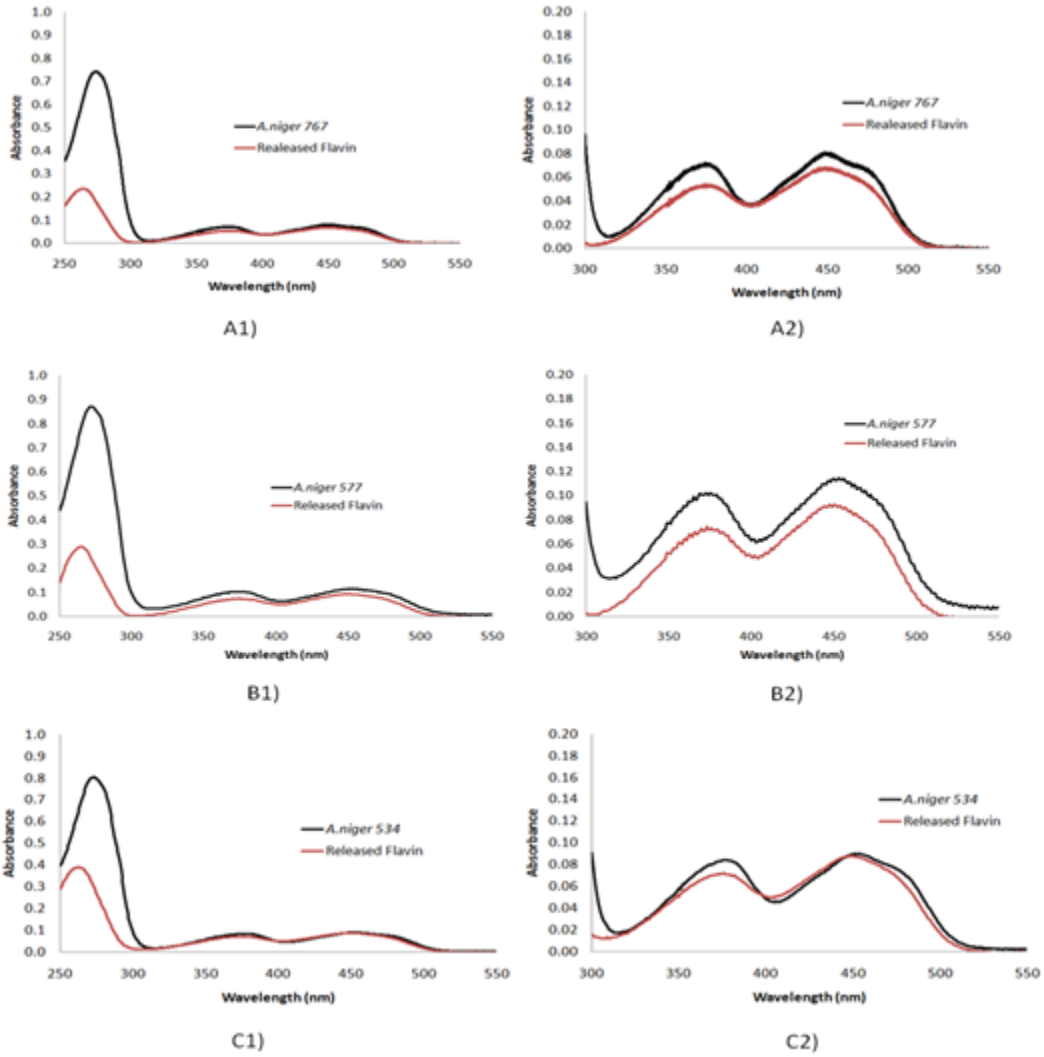


Figure 3.8: UV-VIS spectra of purified putative salicylate hydroxylases and released flavin for *A. niger* 767 (A), 577 (B), and 534 (C) in 50 mM Tris-HCl pH 7.6 at 25 ^{o}C. Each graph shows a full spectrum, and a focused flavin spectrum. Released flavin is the flavin absorbance of the supernatant after thermal denaturation and removal of the protein.

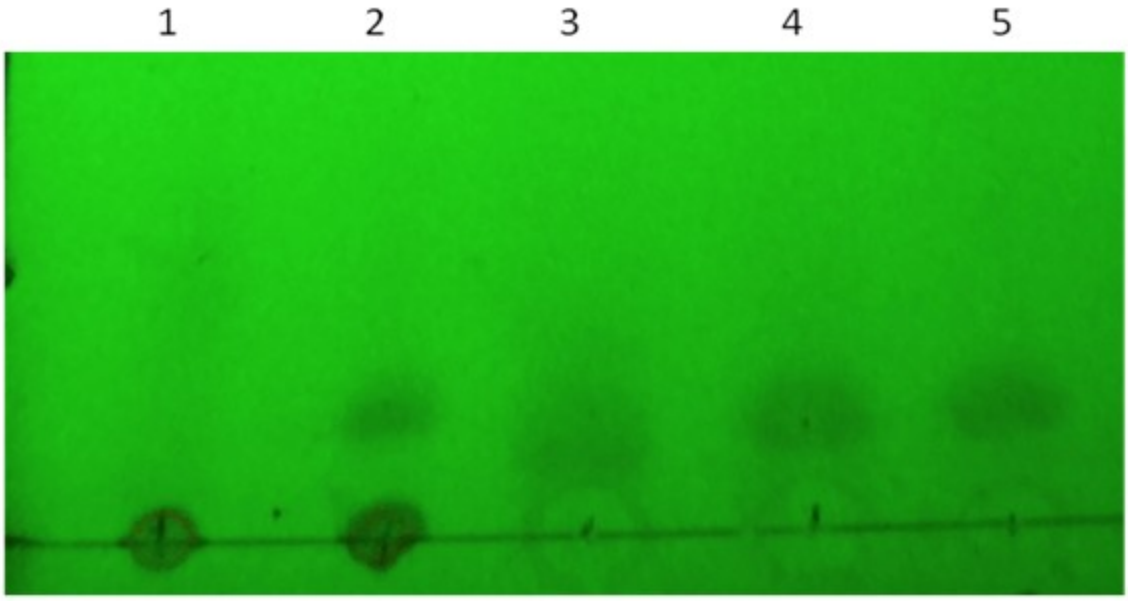


Figure 3.9: Identification of putative *A. niger* salicylate hydroxylase flavins by TLC. FMN (lane 1); FAD (lane 2); *A. niger* 767 flavin (lane 3); *A. niger* 577 flavin (lane 4); and *A. niger* 534 flavin (lane 5). Samples were developed on a Whatman silica-coated TLC plate using a developing solvent of n-butanol: acetic acid: H_20 (4:1:5)[60] and visualized under UV light.

3.2.2.2 Number of FAD binding sites

The next step was to determine the stoichiometry of FAD binding for each protein using the data shown in the previous section. The concentration of FAD was determined by using the FAD extinction coefficient of 11, 300 $M^{-1}cm^{-1}$ [31] at 450 nm after first releasing the FAD molecules from the protein by heating in a boiling water bath. Once the FAD concentration was determined it was compared to that of the protein. In all cases, there appeared to be one mole of FAD per mole of protein, indicating that all three of these putative salicylate hydroxylase binds a single FAD (Table 3.8).

A. niger protein ID	FAD : Protein Ratio
767	1.05
577	1.04
534	1.10

Table 3.8 FAD to protein ratio of putative salicylate hydroxylases

3.2.2.3 Oligomeric states of *A. niger* 767, 577, and 534 proteins

The oligomeric states of the purified proteins were determined by estimating the molecular weight of the proteins using gel filtration chromatography on a Superdex 200 10/300 GL column, and comparing to the predicted subunit molecular weights. The column was first calibrated by running protein standards (Bio-Rad) consisting of thyroglobulin, γ-globulin, ovalbumin, myoglobin, and vitamin B12. Elution times of each standard were measured and the K_{av} values calculated as described in *Materials and Methods* and plotted against the Log M_w. The K_{av} of the purified proteins were determined in the same manner as the standards and the standard curve was used to estimate the corresponding molecular weights. It was estimated that *A. niger* 767, 577, and 534 had molecular weights of 36.5 kDa, 80.1 kDa, and 77.7 kDa corresponding to a monomer, dimer, and dimer respectively (Table 3.9). Elution profiles of all three enzymes are included in the *appendix*.

A. niger Protein ID	Theoretical Molecular Weight (kDa)	Estimated Molecular Weight (kDa)	Oligomeric State
767	47.9	36.5	Monomer
577	48.4	80.1	Dimer
534	48.2	77.7	Dimer

Table 3.9: Estimated molecular weights & oligomeric states for putative *A. niger* salicylate hydroxylases.

3.2.2.4 NADH/NADPH specificity

All 3 of the purified enzymes preferred utilizing NADH compared to NADPH (Table 3.10), consistent with what is commonly reported for known bacterial salicylate hydroxylases.

Between the 3 enzymes, *A. niger* 767 seems to have a much higher preference for NADH to NADPH since only 13 % of the rate of consumption of NADH was achieved with NADPH compared to *A. niger* 577 and 534, where the relative values were 68.1 % and 49.2 % respectively.

***A. niger* JGI ID**	**Rate of NADH consumption (µmol/min/mg)**	**Rate of NADPH consumption (µmol/min/mg)**	**NADPH consumption relative to NADH (%)**
767	4.40	0.579	13.2
577	25.4	17.3	68.1
534	0.382	0.188	49.2

Table 3.10: NADH/NADPH specificity of purified salicylate hydroxylases tested using salicylate as a substrate by monitoring the rate of NADH/NADPH consumption at 340 nm. For all measured values NAD(P)H background consumption was subtracted, and all rate measurements were performed in duplicates with the average values represented. The following amounts of protein was used for each of the *A. niger* enzymes: 6.59 µg 767, 7.37 µg 577, and 34.9 µg 534.

3.2.2.5 Aromatic substrate specificities of putative salicylate hydroxylases

Once it was determined that the enzymes were more active with NADH as a co-substrate compared to NADPH, NADH was used to measure the enzyme activities on various aromatic substrates (Table 3.11). *A. niger* 767 was most active on 6-methylsalicylate with significant activity on 4-methoxysalicylate, salicylate, 4-methylsalicylate, and 5-methylsalicylate. *A. niger* 577 was most active on salicylate, 4-methylsalicylate, and 5-methylsalicylate while *A. niger* 534 was most active on salicylate. Comparison of the 3 purified hydroxylase enzymes reveals that their activities differ greatly on their most active substrates. *A. niger* 577 is the most active of the three enzymes (39.1 µmol/min/mg on 4-methylsalicylate), while 767 (13.91 µmol/min/mg on 6-methylsalicylate) has significantly lower activity, and 534 (0.40 µmol/min/mg on salicylate) has extremely low activity.

Substrate	Rate of NADH Consumption (μmol/min/mg)			*Relative Activity (%)		
	Enzyme			Enzyme		
	767	577	534	767	577	534
Benzoate	0.08	0.757	NA	1.0	2.4	NA
Salicylate	7.68	31.69	0.40	100.0	100.0	100.0
3-Hydroxybenzoate	0.02	0.02	0.00	0.3	0.1	-0.7
4-Hydroxybenzoate	0.02	0.03	0.05	0.3	0.1	12.4
2,3-Dihydroxybenzoic Acid	1.48	19.74	0.18	19.3	62.3	45.3
2,4-Dihydroxybenzoic acid	4.29	18.16	0.40	55.9	57.3	99.9
**2,5-Dihydroxybenzoic acid	0.42	44.16	0.13	5.5	139.3	32.3
2,6-Dihydroxybenzoic acid	0.69	15.14	0.37	9.0	47.8	92.6
3,4-Dihydroxybenzoic acid	0.00	0.02	0.09	0.0	0.1	22.8
3-Chlorosalicylic acid	0.41	3.15	0.06	5.4	10.0	15.7
4-Chlorosalicylic acid	0.42	10.91	0.34	5.5	34.4	85.2
5-Chlorosalicylic acid	0.26	3.89	0.37	3.4	12.3	91.6
3-Methylsalicylic acid	1.48	6.80	0.10	19.3	21.4	24.0
4-Methylsalicylic acid	7.23	39.08	0.40	94.1	123.3	99.1
5-Methylsalicylic acid	8.06	34.41	0.26	104.9	108.6	64.5
6-Methylsalicylic acid	13.91	5.95	NA	181.2	18.8	NA
4-Methyoxysalicylic acid	11.37	21.82	0.24	148.1	68.8	61.2
Anthranilic Acid	0.38	0.51	0.02	4.9	1.6	3.8
3-Hydroxyanthalic acid	0.45	NA	NA	5.8	NA	NA
2-Hydroxyphenylacetic acid	0.08	0.07	-	1.1	0.2	-
O-anisic acid	0.26	0.27	-	3.4	0.9	-
6-hydroxypyridine 3-carboxylic acid	0.21	0.03	0.03	2.7	0.1	7.6
Phenol	0.01	0.25	NA	0.1	0.8	NA

*Relative to the activity of the enzyme with salicylate as a substrate
** Measured at 370 nm (ε=2600 $M^{-1}cm^{-1}$)[79] instead of 340 nm (ε=6200 $M^{-1}cm^{-1}$)
Rates denoted as (-) were not tested with the specific enzyme.
Rates denoted as (NA) represent negative values

Table 3.11: Substrate specificities of putative salicylate hydroxylases. All above measurements represent duplicate averages in which background NADH consumption was subtracted. All replicate measurements had a maximum difference of 5 %. Detailed reaction conditions are described in *Materials and Methods*.

3.2.2.6 Measuring efficiency of substrate hydroxylation

After the 3 enzymes' most active substrates according to NADH consumption (Table 3.11) were determined, a similar assay assessing the amount of oxygen uncoupled (proportion of oxygen used to produce H_2O_2 instead of hydroxylated product) was performed. In this assay, reaction components were mixed in a sealed oxygen electrode chamber to measure the concentration of oxygen over time. The putative salicylate hydroxylase was added and the decrease in oxygen concentration was observed until it plateaued, indicating that all the NADH was consumed. At this point, catalase was added, and resulted in an increase in oxygen concentration due to the conversion of H_2O_2 to O_2 (Figure 3.10). The oxygen uncoupling was then determined by comparing the oxygen consumed compared to the oxygen returned by catalase, where enzymes with the lowest amount of oxygen uncoupled were more efficient at hydroxylating their substrate. Below in Table 3.12 is the oxygen uncoupling for the three purified enzymes with their most active substrates. *A. niger* 767 is highly uncoupled with all substrates except 4-methoxysalicylate (43.7 %), whereas *A. niger* 577 has low uncoupling with most substrates tested including salicylate. *A. niger* 534 is highly uncoupled with all the aromatic substrates tested.

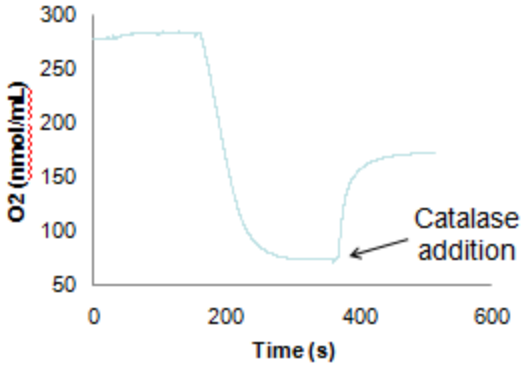


Figure 3.10: Example of oxygen uncoupling assay using *A. niger* 767 with salicylate. Reaction conditions were as described in *Materials and Methods*

	Oxygen uncoupled (%)		
Substrate	***A. niger* 767**	***A. niger* 577**	***A. niger* 534**
Salicylate	85.4	4.7	98.6
3-methylsalicylate	91.2	58.1	-
4-methylsalicylate	77.3	2.5	76.6
5-methylsalicylate	90.3	13.4	91.1
6-methylsalicylate	84.4	40.1	-
4-methoxysalicylate	43.7	47.1	95.4
2,3-dihydroxybenzoic acid	101	10.2	88.7
2,4-dihydroxybenzoic acid	93.0	28.2	88.4
2,5-dihydroxybenzoic acid	89.8	14.2	-
2,6-dihydroxybenzoic acid	87.8	0.93	71.6
3-chlorosalicylic acid	95.6	24.9	-
4-chlorosalicylic acid	93.7	3.74	86.3
5-chlorosalicylic acid	95.8	5.17	86.2

Table 3.12: Oxygen uncoupling of salicylate hydroxylases. Oxygen uncoupling measured in an oxygen electrode at 25 ^{o}C in 50 mM Tris-HCl pH 7.6 buffer. All measurements were carried out in either duplicates or triplicates with averages represented. All replicate differences were within 2 %. Measurements denoted as (-) were not tested with the specific enzyme.

3.2.2.7: Product studies

Once the optimal hydroxylase enzyme substrate was determined by considering both enzyme activity and oxygen uncoupling, the products of the reactions were examined. This was performed by using a coupled assay in which the hydroxylation product catechol (or a methylcatechol) was cleaved by catechol 2,3-dioxygenase (gift of Dr. Elisabeth Cadieux). These cleaved products absorb at 400 nm and extinction coefficients are known. *A. niger* 767 and 577 both formed catechol and 4-methylcatechol using salicylate and 5-methylsalicylate as substrates, respectively. Additionally, both enzymes produced 4-methylcatechol using 4-methylsalicylate as a substrate. *A. niger* 577 appeared to produce 3-methylcatechol with 3-methylsalicylate and 6-methylsalicylate as substrates, while *A. niger* 767 did not. *A. niger* 534 apparently did not produce any catechol (or catechol-like) product with any of the substrates tested. Additionally, some enzyme reaction products were detected using TLC such as the example of *A. niger* 577 conversion of salicylate to catechol (Fig 3.11).

	A. niger 767		*A. niger* 577		*A. niger* 534	
Substrate	**Oxygen uncoupling (%)**	***Observable product formation**	**Oxygen uncoupling (%)**	***Observable product formation**	**Oxygen uncoupling (%)**	***Observable product formation**
Salicylate	85.4	Yes	4.7	Yes	98.6	No
3-methylsalicylate	91.2	No	58.1	Yes	NA	NA
4-methylsalicylate	77.3	Yes	2.5	Yes	NA	NA
5-methylsalicylate	90.3	Yes	13.4	Yes	91.1	No
6-methylsalicylate	84.4	No	40.1	Yes	NA	NA

* observable product refers to the formation of catechol or methylcatechol which is produced by the hydroxylase enzyme and cleaved by catechol 2,3-dioxygenase.

Table 3.13: Formation of catechol or catechol-like products via 3 hydroxylase enzymes. Assays performed as described in *Materials and Methods*. Measurements denoted as NA were not tested with the specific enzyme.

Figure 3.11: Identification of *A. niger* 577 reaction products by TLC in 50:50:1 (ethyl acetate: chloroform: acetic acid) developing solvent where salicylate, catechol, and the salicylate reaction product have been spotted in lanes 1, 2, and 3 respectively.

3.2.2.8: Substrate binding studies

Tightness of substrate binding to the *A. niger* hydroxylases was determined by measuring the change in fluorescence associated with the enzyme-bound FAD upon the addition of aromatic substrates. Excitation was carried out at 450 nm which resulted in emission at 525 nm**.** The data were fit to the Hill1 equation in OriginLab as described in *Materials and Methods*. In comparing binding of salicylate to *A. niger* 767 and 577 it was found that *A. niger* 577 binds salicylate nearly 4 times more tightly than *A. niger* 767, with apparent K_d values of 480 µM and 95 µM respectively. Additionally, the n values (Hill coefficients) for the fitted curves in Figure 3.12 was about 1. On the other hand, *A. niger* 767 binds 4-methoxysalicylic acid nearly 4 times tighter than salicylate with apparent Kd values of 98 µM for 4-methoxysalicylic acid and 480 µM for salicylate (Fig 3.13). (Some raw fluorescence data have been included in the appendices)

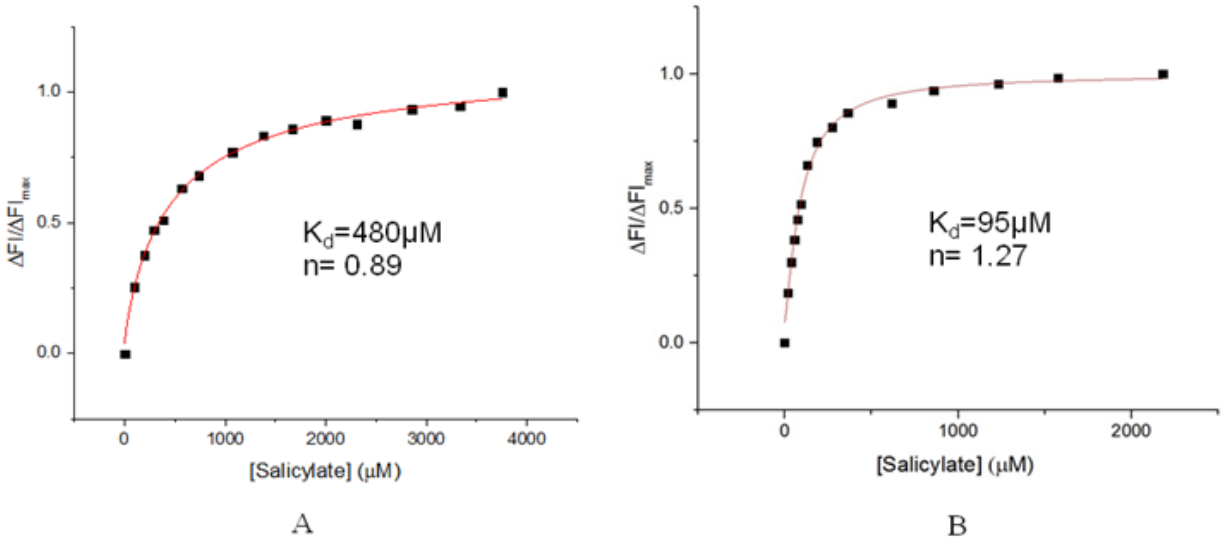


Figure 3.12: Determination of salicylate binding to *A. niger* 767 (A) and *A. niger* 577 (B) K_d values by monitoring FAD florescence changes at λ_{ex} =450 nm and λ_{em}= 525 nm. Successive amounts of salicylate were titrated into a fluorescence cuvette containing 5 µM *A. niger* 767 or 577, and the change in fluorescence was measured. Note: reported K_d values are apparent K_d.

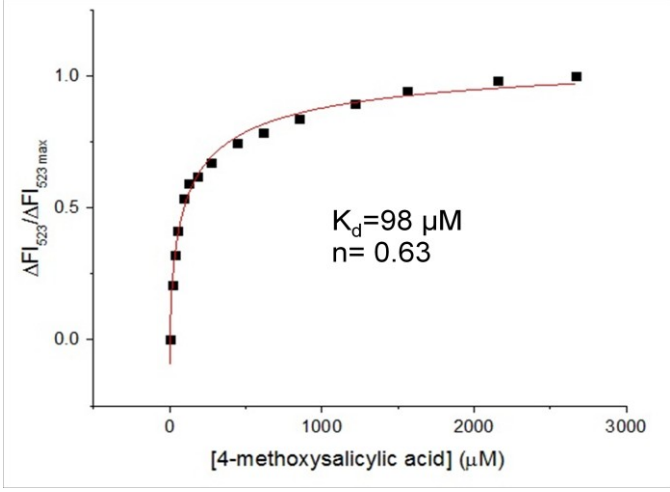


Figure 3.13: K_d determination for 4-methoxysalicylic acid binding to *A. niger* 767 (5 µM) by measuring changes in FAD fluorescence at λ_{ex} =450 nm and λ_{em}= 525 nm with varying additions of 4-methoxysalicylic acid. Note: reported K_d value is apparent K_d.

3.3 Phenol Hydroxylases

3.3.1 PCR amplification to purification

Predicted phenol hydroxylase gene sizes and calculated protein sizes are listed in Table 3.14: protein molecular weight is estimated based on the amino acid sequence submitted to the ProtParam server. Table 3.15 summarizes the success of each of the experimental steps for these targets starting from PCR amplification and continuing through to characterization. Expressing these target proteins proved very challenging, and was successful only with *A. niger* 999. *A. niger* 687 could not be amplified by PCR while *A. niger* 422 did not produce enough PCR product for cloning. *A. niger* 300, 146, 718, and 951 were successfully cloned but transformants never expressed protein under the various conditions attempted. On the other hand, protein expression was observed for *A. niger* 617 but it was present in inclusion bodies. Targets where production of soluble protein was not possible were not pursued any further.

***A. niger* JGI ID**	**Expected gene size (base pairs)**	**Expected protein molecular weight (kDa)**
1206**422**	1809	67.0
1210**999**	1977	73.2
52**617**	1701	63.0
1125**300**	1911	71.0
1125**146**	1857	69.4
1159**718**	1734	64.1
1227**951**	1887	71.4
1085**687**	1812	67.4

Table 3.14: Expected protein and gene sizes of putative phenol hydroxylases

	Steps Achieved				
***A. niger* JGI ID**	**PCR Amplification**	**Cloning**	**Expression**	**Purification**	**Characterization**
1206**422**	✓	×			
1210**999**	✓	✓	✓	✓	✓
52**617**	✓	✓	✓		
1125**300**	✓	✓	×		
1125**146**	✓	✓	×		
1159**718**	✓	✓	×		
1227**951**	✓	✓	×		
1085**687**	×				

Table 3.15: Summary of steps achieved for each putative phenol hydroxylase target . Check mark (✓) indicates successful step, (×) indicates unsuccessful step, and a blank box indicates untested step.

3.3.1.1 PCR amplification of phenol hydroxylase target sequences

DNA fragments were amplified by PCR using the same set of conditions used for salicylate hydroxylase gene amplification (see section 3.2.1.1). Figure 3.14 shows the PCR amplification results for *A. niger* 422, 999, 300, 146, 617, 718, and 951: all amplified fragments appear to be the correct size.

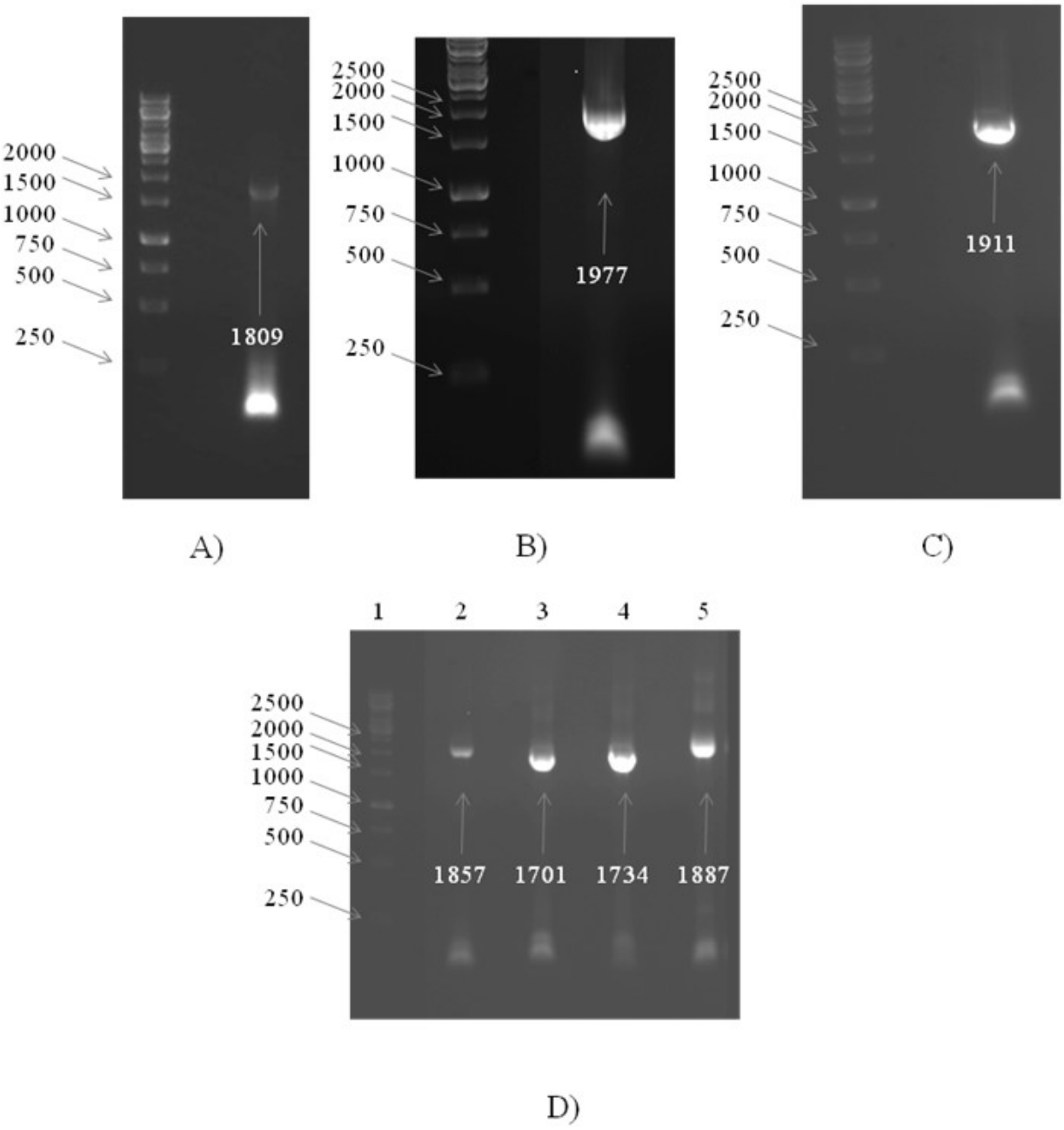


Figure 3.14: PCR amplification of *A. niger* putative phenol hydroxylase genes. In all 4 gels DNA size markers are present in lanes on the far left. The agarose gels show samples from the PCR amplification of *A. niger* 422 (A), 999 (B), 300 (C) 146 (D lane 2), 617 (D lane 3), 718 (D lane 4), and 951 (D lane 5). Arrows pointing to size marker bands represent base pair sizes. Additionally, arrows pointing to amplified PCR products indicate the expected size of the DNA fragment.

3.3.1.2 Cloning, plasmid screening, and DNA sequencing

Amplified phenol hydroxylase target DNA sequences were cloned into the pLATE11 vector, generating plasmids whose inserts were screened using restriction enzyme digestion and PCR experiments similar to those described in section 3.2.1.2 for salicylate hydroxylases (data not shown). Then, all plasmids were sequenced and confirmed to contain the expected JGI target sequences, with the exception of *A. niger* 146 and 951 which contained introns as shown in Figures 3.15 and 3.16 respectively. These were subsequently discarded.

ATGACTATCGCACCGTATCCAGTTGGCGCCTTCAAATCGCAGAGAACGGATGTTCTCATTGTAGGCGCG
GGCCCTGCTGGATGCATGGCCGCCGCAACTTTGCAAAGATACGGCATTGATTTTTGCCTAATTGATAAAA
GACCCACGAGAACACAAACTGGGCATGCAAGTGGTGCGTCTATGAGATGCCCTTCTGTTGATCTGACGA
GAAAGCTGATCGTTTCGCAAAGCATTCCAACCTCGAACCCAAGAGATCCTGCAGACGATGAATCTACTTC
ATGATCTTGACAAGAGAGGACATCGCCTTACCGAAACCTCCTTTTGGATGCGTGATAGCACAGGCGCTC
TAATCAGCAACTTCACGGGCGCTGAGGTGGTTCATGGTGAGTCCTTAATCATGAGGAGAGGTGATTATG
TGACACCCGACTTAGCTGACTTTGTAAAAAGCCACACCTTACCAGTACCTCTTTAACACCGACCAAGGTA
TGACGGAGGACGTATTTGAACAGTATCTGAACACCAAAGGGCAAAAGATACAGCGGTTCATGGAATTG
GTTCATTATGAGCACGACTTAGATCCCGAATGGCCACTCACAGCATACATCAAAAACAATGCTAGCGGG
GCCATTGAAGCGTGGCAGACGAAGTATATCCTGGGTACTGACGGGGCGCGAAGCGCTACACGGCGGG
CTACCGGCGTGCAATCATCCTCTCAGGGTGGTGAAGACGTTTGGGCTGTGGCGGATGTTTATGTTGATA
CGAACTTTCCCGATTATAGAAGACGTTGCGCGATCCGAACCCCTGATGGTGGTTGCATGTTAATTCCTCG
CAAGGATGAAGGACTCCGGATTTTTCTCCAGGTAGATGAGAAAAGCCAAGAGCATCTCGATGAAAACG
GAGCTTCTGGGCAAGATGCCCTCACTGGAAACAGCGCTTTCAAACTCACCCAGACCGTCCAGTCGCACA
TCAACAAGGTCATTCATCCTTACAAGATGAATATTACTGATATTGTTTGGATCAGCCAGTACCGTGTTGCC
CAGCGAGTTGTTCACCACTTCAGTGATCCGACCAAGCGAGTTTTCCTTCTTGGAGATGCATGCCATACAC
ATAGTCCTAAGGCGGGACAAGGGATGAATGTCAGCATTTCAGACGCATACAACCTCACATGGAAGCTG
GCACTTGTCATGAAGGGGGTCGCAAAGGCATCGTTGCTCGAAACATACGAACAAGAGAGACTTTGGGT
TGCGCAACAGCTTATTGAATTTGATGCCTTGTTTGCCCGTCAATTCGGACAGAAAGATAAGCTTGACAGT
CAGAATCTCCGCGAAACATGGGAAATGGGACACGGCTTCACAAGCGGGTGTGGATATGAGTACCCTGC
AAACTTGCTTGTTAACCCCGATGTCCGGACATCAATCAACAATCAAGCAGTAGAACCGCTCACGCCGGG
CAAACGTCTTTTACCAATCGATCTGATCCGACACATCGACGGCAATCATGTCCGTATGCTGGACATCATG
CCATCCAATGGTCGATTCCATTTGTTCATTTTTGCGGGCAATGACCTATCTTCGCCGACCCTTCAAAAGCT
CGGAAATACCTTGGATTCACCTCACTCACCAATGAGCCTCTTCAATCTTCTGCCGCTGGAACTCATGGAA
CGATTCCGTCATGAGGACATTACAACGGCATCAGTTCCGTTCACCAACAAGGCATATGTTCTTGATCTTTT
CTTGATTCATTCACAAAATCACCTGGATGTTCAACTGGGGAATATCCCTTCTCCATTCTCCACTAAATGGC
CCATGAACGTATACTCTGATTTTGATGGTGCTGCACAAAATCAGCTTGGTGTCCCTGAAGACTCTGGCGC

CTTGGTGGTGGTGAGACCAGATGGTTATATTGGACTTGTCACTGGCTTGGATAATGTTGAGGACGTGAC
GTCCTATTTCGATGGCTTCATGCACCGGCGCATCTAG

Figure 3.15: Sequenced *A. niger* 146 plasmid construct with introns highlighted in red

ATGACTGTCTCAATCGAATCTACAACCGATGTGCAAGTCTTCCCCGTCCGTCCCTCAGCTGCTACTACTGA
CCCAGATGCAGTCTCATCATTGGCGCGGGCCCATCAGGTCTGACTATTGCTTATTGGATGGCGCAGTATG
GCATAAACGCTCGCATCATCGACAAGCGTAGCACGAAACTCTTTCGAGGTCATGCAGATGGCCTGAGAG
GCCGAACGCTCGAACTCATCGATAGTATGGGGTTTGTCCACCGGATCTTGCAAGAAGGGTACCATGGAG
CAGGTCTTAGGTATTGGGTAGGTCGTCTGATGACCTTTAAGTCGGACCGGTAATTAATGATGTACACTA
GAGACCAGACGGAAAGGGGGGTTTGAAGCGGGCCGGAGAGCTGCGAAGCACAATCTGGAATTCCTCG
CCCTATCGCATGGCCTTATTGTCCCAGGGAACGATTCTTTCTCGATGTTATCCGGGATGCTTCCAGTATCG
ACGTAGAAAGAGGTGTCATAGGAGAGTCACTGAGCTATGATGAGAGCCTTGAGAACGATCCCCAGGCG
TACCCCATCACGATCAAGTTGAGGACTCTAAGTGAGGAAGAAGCAACTCCAATCCCCAACCCGGGTATT
ACCCCCTACAAACCAGCCCGGGATGACCTGGAGGATCTGGTTCCAAAACGCAAGCACGATGTTGGGAC
GATCGAGACGGTTCGGGCCAAGTATCTCATTGGCTGCGATGGGGCGCACAGCTGGACAAGGAAACAGT
TGAATATCCCCTTCGAGGGATCCCACACAGATCATATATGGTAGCCTACCATACCGTGCTTTGATAGTCC
GTAGCTAACATGCATGCTCTAGGGGTGTTATGGATGTGGTCCCAATCACGGATTTCCGTAAGTTCAAAAC
CTGGGATGGAAAGAAAGACCTCTAATGTCATACCACAGCGGACATCCGCTGGCCATCTTCGATTGACAG
TCCGCACGGGCATCTGCTTGTCATCACGAGAGAAGGCAATACAGTGCGCTTTTATGTCCCTTTGCAAGAA
GAAGATGAGATAACGCCCTCATTCGATCGGTCGGCGGTGACTTTGGACATGCTCTGCGAAAGAGTGAA
AAGCATCATGTCTCCCTTCTATTTCGACTTCAAAGTCTGCCACTGGTGGGCTGCATATCAGGTATACACAA
ACTCTCGCTTTTTCGGTAGGAAGCTTACTGATGATCGGTTAATGCAAATTATAGGTCGGTCGGAGACTCG
CACCCGCCGTTAATAAAGGACGTATGTATCTCGTCGGTGATGCCGTCCATACCCACAGCCCCAAGATTGG
GCTGGGAATGAACATGAGCATTCAGGACGGCTTTAACCTCGGTTGGAAACTTGCGCTGGCCGTGAAGG
AGATGGCGTCTCCTTCCATTCTCGATACCTACAAGTCCGAGCGTCATCAGCTCGCCAAGATGCTCCTCGA
GTTCGACCAGAAATGGGCTGCGTTTTTCCTGAAGCAGAAGAAAAAGCAGCAGCAGCTTGGACTTGAGG
CTCCTCCGGAAGCAAACCCAGAAGACATCCAGGCCATGCAGGATGTTTTCAGCGAAAATGAGCTGTTCG
CGGAGGGCCATGTATCATTCTACAAGGCAAGCCCCATAGTACACAAAGGCAGTACAACAGTGGCCCGAC
ACTTGACAGCAGGAGAAAGGTTTCCCGTGGCACTCATTCGCAAGCAAGCAGATGGTCAGCCATGGTGG
ACTTCACGGCTCCTGAAGAGCGATGGTCGTTTTCGGATTTTGCTTCTGGTCGGAGACTGCCGACTGGAA
GATCAGAAAAGCCGCATATTGGCTCTCAATAATGGGCTATTGGAGCTTCAGAAACGCTTCACGCCCCCC
AAGCAAAAGCTCGACTCTGTCATCGAAGTCAAGGCAGTCCACAGTGCTCCGGTGGATGACGTTGAACTG
AGTGACTTCCCCCCGATTTTGCGCCAGTTCGACGACACCACTGGCTGGGACTACACAAAGGTCTGGAGT
GACAGTGAGTGCTTCTGGGACGAACAGTGCAAAGGCAACGCATATGAGATTTGGGGGGTGGACAGGT
CAAGGGGTGCCATGGTTGCTTTACGACCCGATCAATATATCGGATGGATTGGAGAGCTCGACGACGTG
GCCGGCGTGACCACATACTTTGAAGGTTTCCTTCGCGAGCCTCGGCCGGAAGCGGTGCTGTAG

Figure 3.16: Sequenced *A. niger* 951 plasmid construct where introns are highlighted in red

3.3.1.3 Protein expression results for putative phenol hydroxylases

Protein expression in different *E. coli* strains, at 16 ^{o}C was attempted for all successful phenol hydroxylase plasmid constructs (*A. niger* 999, 617, 300, and 718). The only two proteins that were successfully expressed, as indicated by the appearance of the appropriately-sized bands on SDS-PAGE gels, are shown in Figure 3.17. *A. niger* 617 was expressed in the pellet fraction in all *E. coli* strains tested (BL21 DE3, Rosetta DE3, and Shuffle) while *A. niger* 999 was expressed in the soluble cell fraction in BL21 (DE3). Higher levels of *A. niger* 999 expression were observed in Rosetta (DE3) *E. coli* as opposed to *E. coli* BL21 (DE3).

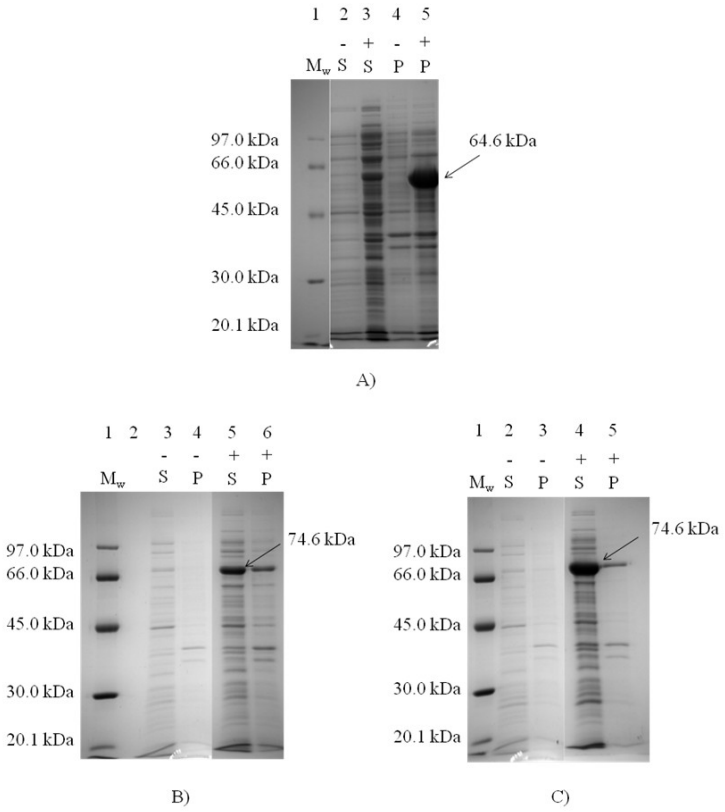


Figure 3.17: Phenol hydroxylase expression trials (induction at 16 ^{o}C). 10% SDS-polyacrylamide gels of: *A. niger* 617 in *E. coli* BL21 (DE3) (A); *A. niger* 999 in *E. coli* BL21 (DE3) (B); and *A. niger* 999 in *E. coli* Rosetta (DE3) (C). Lanes labelled M_w, S, and P represent molecular weight markers, supernatant (soluble cell fraction), and pellet (insoluble cell fraction) respectively. Lanes with a negative sign (-) indicate uninduced cells while a positive sign (+) indicates that the cells were induced with 0.5 mM IPTG (at 16 ^{o}C). Arrows indicate estimated protein sizes.

3.3.1.4 Purification of *A. niger* 999

A. niger 999 (73.2 kDa) was the only expressed and soluble phenol hydroxylase target obtained. It was purified via a 4-step procedure as described in *Materials and Methods*. Samples from each step of the purification were run on a 10 % SDS-polyacrylamide gel, from which it is clear that a significant improvement in purity was obtained, compared to the crude extract (Fig 3.18). However, by comparing specific activities at each step of the purification, little change is observed (Table 3.16).

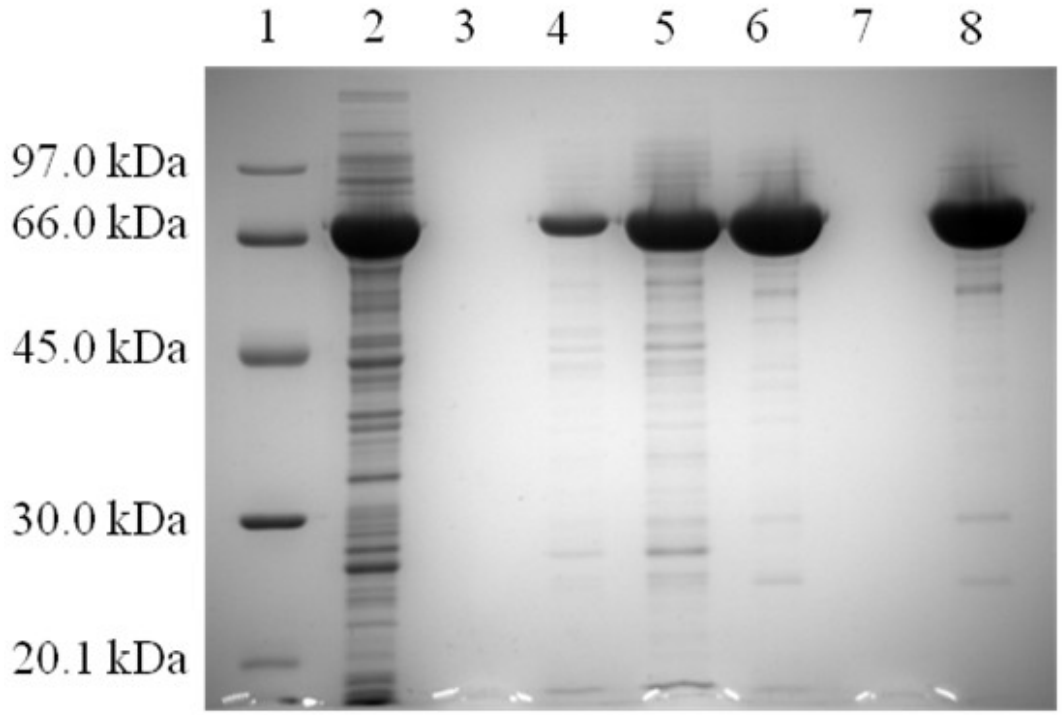


Figure 3.18: Purification of *A. niger* 999 as monitored on a 10 % SDS-polyacrylamide gel. Protein low molecular weight markers (lane 1); crude extract (lane 2); empty (lane 3); DEAE pool (lane 4); ammonium sulfate concentrated DEAE pooled fractions (lane 5); phenyl-Sepharose pooled fractions (lane 6); empty (lane 7); and Sephacryl S-300 pooled fractions (lane 8).

Purification Step	Volume (ml)	Concentration (mg/ml)	Activity (U/mL)	Total Protein (mg)	Total Activity (*Units)	Specific Activity (Units/mg)	Fold Purification	Activity Yield (%)
Crude	45.0	28.2	27.8	1.27E+3	1.25E+3	0.987	1.00	100
DEAE + Ammonium sulfate precipitation	31.0	18.8	18.7	582	579	0.995	1.01	46.3
phenyl-Sepharose	9.50	26.1	20.4	248	194	0.780	0.79	15.5
Sephacryl S-300	7.50	29.9	32.4	224	243	1.08	1.10	19.4

*One unit (U) is defined as the amount of enzyme required to oxidize one µmol/min of NADPH in the presence of phenol in 50 mM Tris-HCl pH 7.6 at 25 ^{o}C.

Table 3.16: Purification summary for *A. niger* 999

3.3.2 Phenol Hydroxylase: Characterization

3.3.2.1: UV-VIS spectral properties of *A. niger* 999

The absorbance of purified *A. niger* 999 was measured over a range of 250 nm to 500 nm, and the resulting spectrum is shown in Fig 3.19. The characteristic peaks at 280 nm, 375 nm, and 450 nm for amino acid and flavin absorbances previously described for the putative salicylate hydroxylases (section 3.2.2.2) are also present in this enzyme. The flavin moiety was determined to be FAD by TLC (Fig 3. 20) while the FAD to protein ratio was determined to be 0.62 after FAD addition, using the same methodology described in section 3.2.2.3. As with all the other hydroxylase enzymes investigated addition of FAD to the crude extract resulted in increased activity.

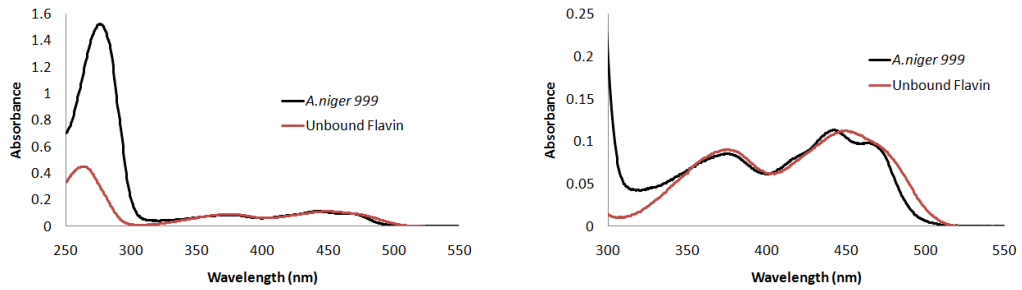


Figure 3.19: UV-VIS Absorbance spectrum of purified *A. niger* 999 and released flavin in 50 mM Tris-HCl pH 7.6 buffer from 250-550 nm (A), 300-550 nm (B).

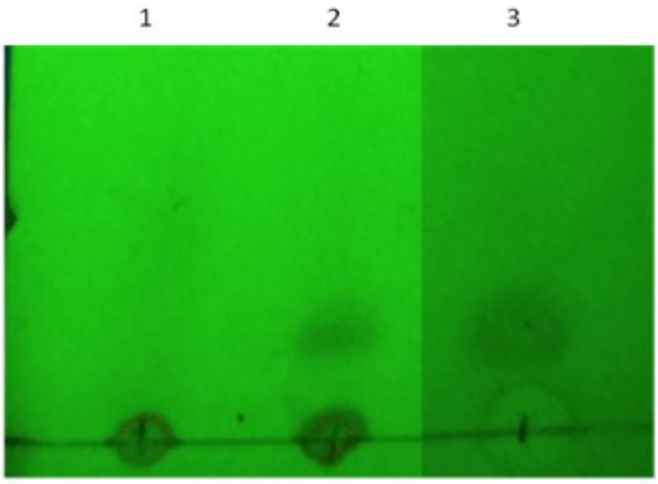


Figure 3.20: Identification of *A. niger* putative phenol hydroxylase flavin by TLC. FMN (lane 1); FAD (lane 2); and *A. niger* 999 flavin (lane 3). Samples were developed on a Whatman silica-coated TLC plate using a developing solvent of n-butanol: acetic acid: H_20 (4:1:5)[60], and visualized under UV light.

3.3.2.2 Oligomeric state of *A. niger* 999

The estimation of *A. niger* 999 molecular weight and oligomeric state was carried out using data from chromatography on a calibrated Superdex 200 10/300 GL gel filtration column, as described for the putative salicylate hydroxylases in section 3.2.2.4. The estimated molecular weight of 132.9 kDa suggests that *A. niger* 999 is dimeric based on the predicted monomeric subunit molecular weight of 73.2 kDa. The elution profile has been included in the *Appendix*.

3.3.2.3 Nucleotide cofactor preference of *A. niger* 999

NADPH was found to be a preferred co-substrate compared to NADH in the presence of phenol, as shown in Table 3.17.

A. niger JGI ID	Rate of NADH consumption (µmol/min/mg)	Rate of NADPH consumption (µmol/min/mg)	NADPH / NADH
999	0.19	1.09	5.7

Table 3.17: Relative rates of *A. niger* 999 with NADH and NADPH. Assays were done using phenol as a substrate by measuring the rate of NAD(P)H consumption at 340 nm. Measurements were performed in duplicate with averages represented: background NAD(P)H consumption has been subtracted.

3.3.2.4 *A. niger 999* substrate specificity

The enzyme activity was measured with phenol and some of its derivatives, as well as other aromatic ring compounds. *A. niger* 999's greatest activity was observed on hydroquinone, phenol, resorcinol, and p-cresol, which all exhibited similar specific activities, 4.18, 3.54, 3.38, and 3.11 (µmol/min/mg) respectively.

Substrate	Rate of NADPH Consumption (µmol/min/mg)	*Relative Activity (%)
Phenol	3.54	100
o-cresol	0.09	2.5
m-cresol	0.91	25.8
p-cresol	3.11	87.9
resorcinol	3.38	95.4
phloroglucinol	0.14	3.9
4-nitrophenol	0.15	4.1
hydroquinone	4.18	118.1
catechol	0.98	27.6

* Relative to the activity of *A. niger* 999 with phenol as a substrate

Table 3.18: *A. niger* 999 aromatic substrate specificity in the presence of NADPH. Rate measurements were performed at 340 nm, in duplicate: background NADPH consumption was subtracted and replicate measurements differed by a maximum of 5 %.

3.3.2.5 Measuring hydroxylation efficiency of putative phenol hydroxylase, *A. niger* 999

Oxygen uncoupling in *A. niger 999* with the four most active aromatic substrates was measured using an oxygen electrode, as described in *Materials and Methods*. Activity in the presence of, phenol was the most uncoupled, with 32.1 % of the O_2 consumed diverted to hydrogen peroxide formation. On the other hand, *p*-cresol, resorcinol, and hydroquinone exhibited significantly lower oxygen uncoupling, with values of 5.49 %, 4.12 %, and 5.34 %.

Substrate	Oxygen Uncoupled (%)
Phenol	32.1
p-Cresol	5.49
Resorcinol	4.12
Hydroquinone	5.34

Table 3.19: *A. niger* 999 oxygen uncoupling with the most active phenolic substrates. All measurements were carried out in duplicate or triplicate, with averages shown. All replicate differences were within 5%.

3.3.2.6 *A. niger* 999 product identification with selected aromatic ring-substrates

Product formation was examined using the most active aromatic substrates, phenol, *p*-cresol, resorcinol and hydroquinone. Product identification proved possible by coupling catechol 2,3-dioxygenase to *A. niger* 999 in these assays, as described in *Materials and Methods*. This enzyme was tested to cleave catechol, 4-methylcatechol, and hydroquinone which results in a yellow product (2-hydroxymuconate semialdehyde or a derivative depending on the initial substrate) which absorbs at 400 nm (data not shown). Additionally, catechol 2,3-dioxygenase does not appear to cleave pyrogallol or phloroglucinol (when measuring absorbance at 400 nm, data not shown). All 4 of these substrates for *A. niger* 999, when turned over in the presence of catechol 2,3-dioxygenase, resulted in the appearance of an intense yellow color absorbing at 400 nm (data not shown). This is consistent with conversion of phenol to catechol, p-cresol to 4-methylcatechol, resorcinol to hydroxyquinol, and hydroquinone to hydroxyquinol (Figure 3.21).

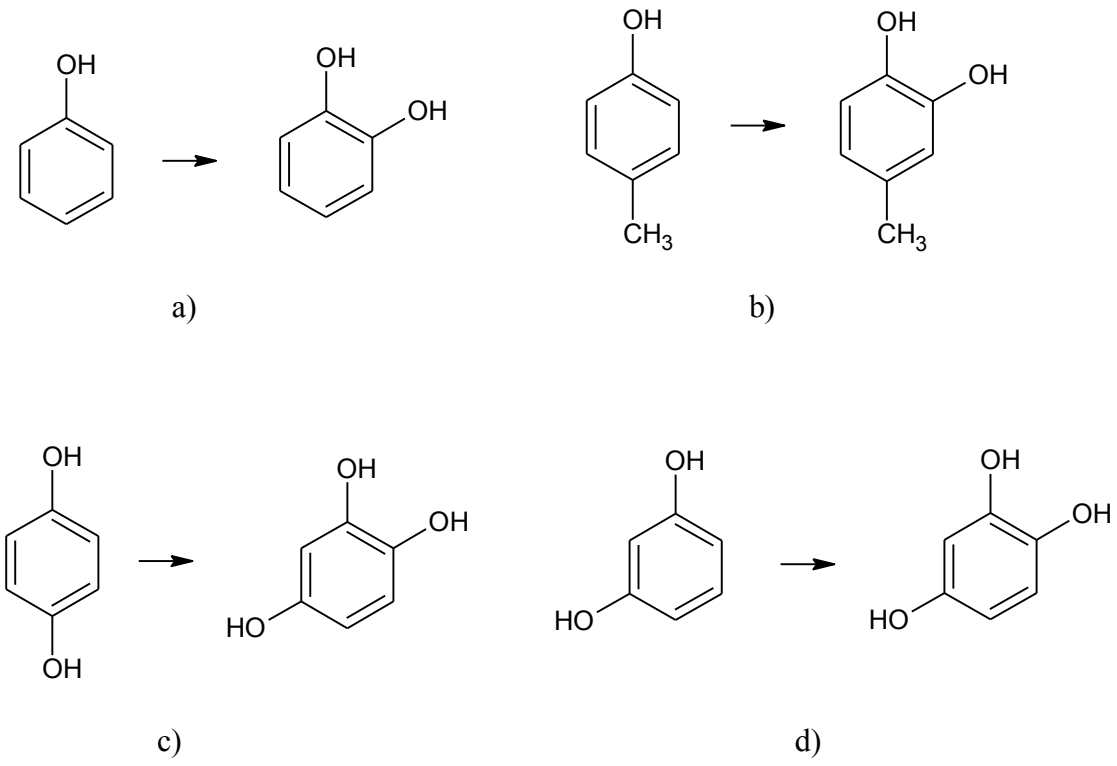


Figure 3.21: *A. niger* 999 product formation. Phenol to catechol (a), *p*-cresol to 4-methylcatechol (b), hydroquinone to hydroxyquinol (c), and resorcinol to hydroxyquinol (d).

3.4 2,3-dihydroxybenzoate decarboxylase (*A. niger* 716)

In addition to the aromatic ring-hydroxylases, other enzymes are involved in preparation of some aromatic compounds for ring cleavage dioxygenases. One such group of enzymes are the aromatic ring decarboxylases. Here, I will present results obtained for the *A. niger* 2,3-dihydroxybenzoate decarboxylase, which has previously been studied by Rao et al. (1967)[2] and Kamath et al (1987)[4]. In their studies the enzyme was isolated from its native organism *A. niger*, while here the enzyme was cloned and expressed in *E. coli*.

3.4.1.1 PCR Amplification of *A. niger* 716

PCR amplification was carried out using the same procedures as described for salicylate and phenol hydroxylase target genes (section 3.2.1.1 and 3.3.1.1). This enzyme is predicted to have a DNA gene size of 1029 bp and a corresponding protein molecular weight of 39.2 kDa. The amplified DNA fragment appears to be of the correct size (Fig 3.22).

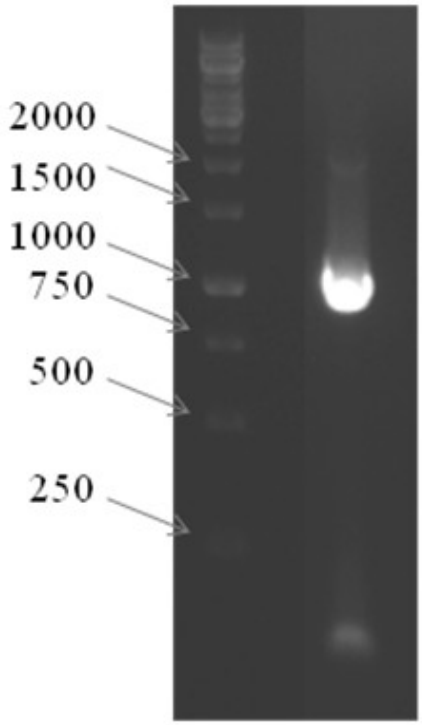


Figure 3.22: PCR amplification of *A.niger* 716 run on a 0.8 % agarose gel. Numbers pointing to bands represent base pair sizes.

3.4.1.2 2,3-dihydroxybenzoate decarboxylase plasmid insert verification

After LIC cloning of the successfully-amplified DNA target into the pLATE11 plasmid, as described in *Materials and Methods,* the recombinant plasmid was examined by restriction digest analysis and DNA sequencing to determine whether it contained the correct DNA insert. For restriction digest analysis, the recombinant plasmid was cleaved using *Kpn*I, for which there is one cut site in the gene and another in the pLATE11 plasmid (Fig 3.23). *Kpn*I cleavage of the plasmid resulted in bands of 315 bp and 5166 bp, which were observed on an agarose DNA gel (Fig. 3.24). DNA sequencing further confirmed that the target sequence (*A. niger* 716) was indeed cloned correctly into the recombinant plasmid used for expression.

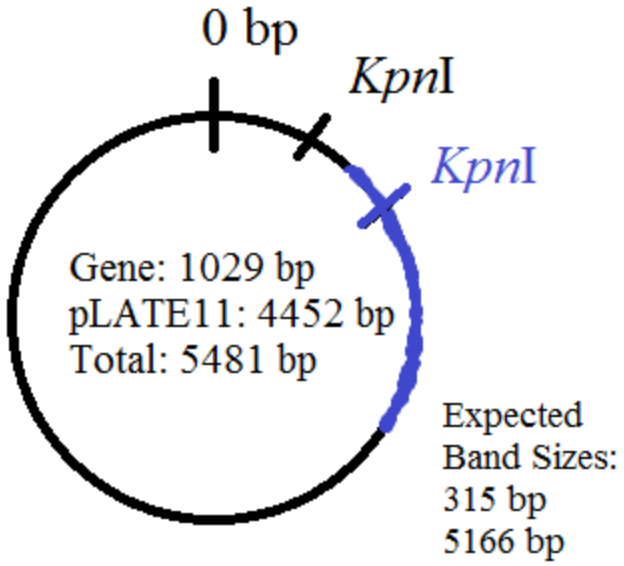


Figure 3.23: *A. niger* 716 restriction digest map using *Kpn*I. pLATE11 derived DNA is shown in black while that from the insert fragment is shown in blue.

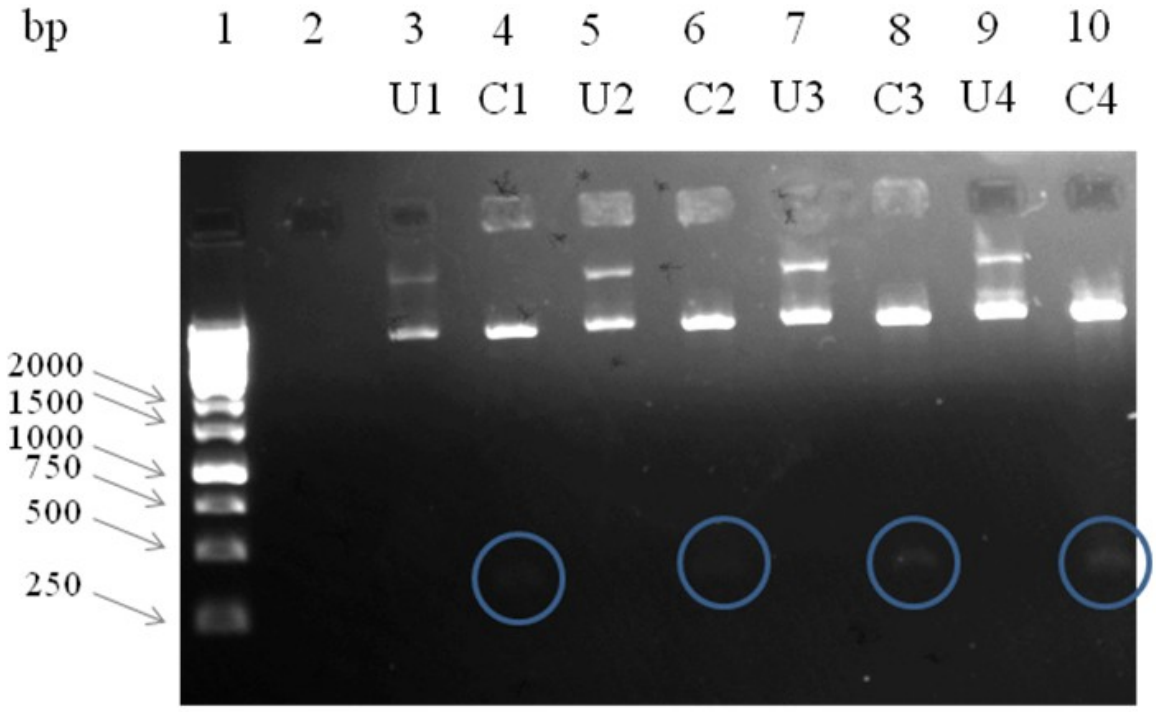


Figure 3.24: Restriction digest of recombinant *A. niger* 716 plasmids with *Kpn*I cleavage, followed by electrophoresis on a 1.0% agarose gel. Base pair sizes are indicated with arrows, in lane 1. Lanes 3-10 represent samples of 4 recombinant plasmids each isolated from independent single colonies (numbered 1-4), where U indicates uncut plasmid while C indicates *Kpn*I treated plasmids. Transparent circles highlight bands which are barely visible, representing the smaller band (≈ 315 bp) .

3.4.1.3 *A. niger* 716 protein production

The recombinant plasmid was introduced into *E. coli* BL21 (DE3), and protein expression was monitored using SDS-PAGE. As shown in Fig 3.25, a protein of the expected size (39.2 kDa observed, vs. 39.5 kDa theoretical size) was successfully expressed in the soluble fraction at 16 ^{o}C.

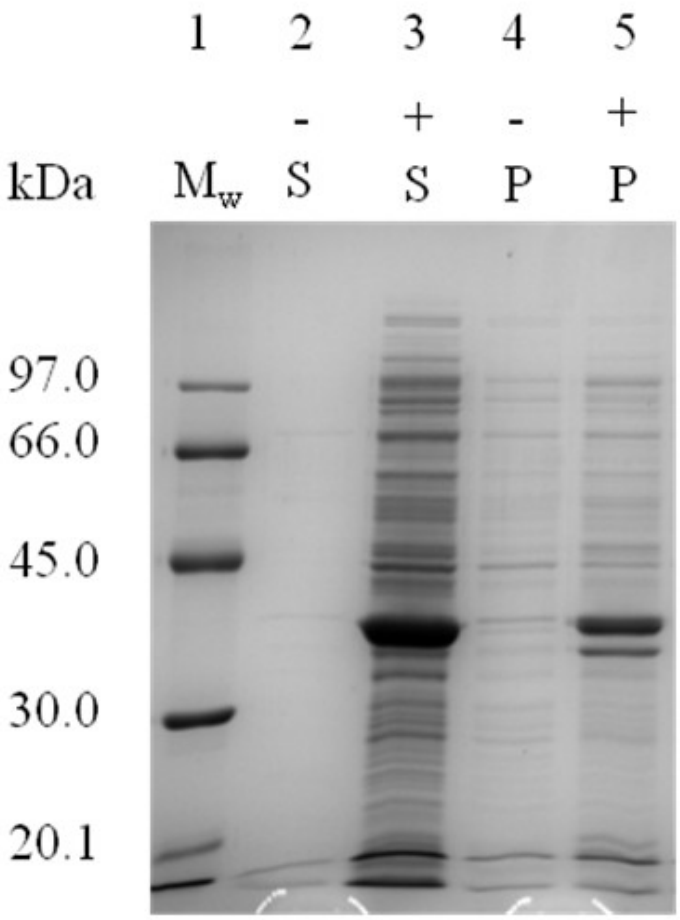


Figure 3.25: A. *niger* 716 protein expression in *E. coli* BL21 (DE3) at 16 ^{o}C as monitored by SDS-PAGE. Lanes labelled Mw, S, and P indicate lanes that have been loaded with protein molecular weights, lysed cell supernatant and pellet, respectively. (-) and (+) signs represent samples from uninduced and induced (0.5 mM IPTG) cell cultures, respectively.

3.4.1.4 *A. niger* 2,3-dihydroxybenzoate decarboxylase purification

A. niger 716 was partially purified via a 3-step purification procedure consisting of ion-exchange chromatography, hydrophobic interaction chromatography, and size-exclusion chromatography. Samples from each purification step were run on a 10 % SDS-polyacrylamide gel (Fig 3.26), which shows that some contaminants were removed and the purity of the final preparation was improved compared to the initial crude extract. However, little change in the specific activity was observed after each step (Table 3.20).

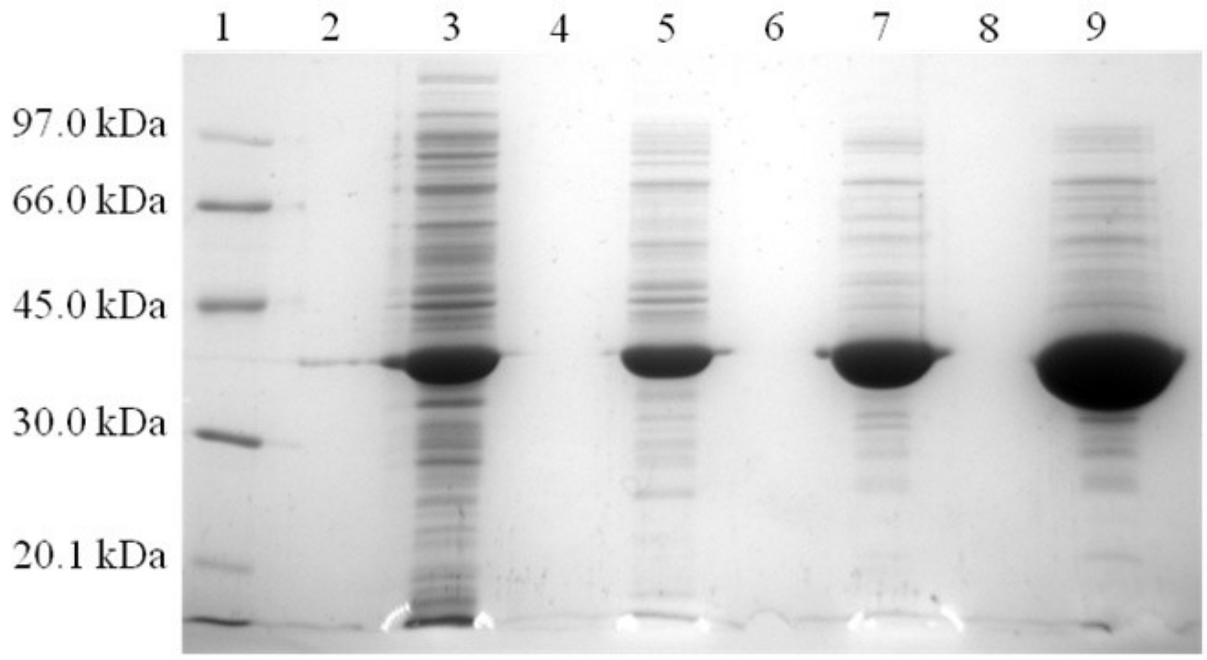


Figure 3.26: *A. niger* 716 protein purification - fractions analyzed by SDS-PAGE. Samples are: protein low molecular weight markers (lane 1); crude extract (lane 3); pooled DEAE fractions (lane 5); pooled phenyl-Sepharose fractions (lane 7); and pooled Sephacryl S-300 fractions (lane 9).

Purification Step	Concentration (mg/ml)	Volume (ml)	Total Protein (mg)	Activity (U/mL)	Total Activity (U)	Specific Activity (U/mg)	Fold Purification	Activity Yield (%)
Crude	31.1	60.0	1.87E+3	91.7	5.50E+3	2.94	1.00	100
DEAE + ammonium sulfate precipitation	16.6	50.0	831	64.1	3.20E+3	3.85	1.31	58.2
PS	30.4	11.5	350	77.7	893	2.55	0.87	16.2
S-300 (pure)	81.4	2.50	203	248	619	3.04	1.03	11.3

Table 3.20: Purification summary for *A. niger* 716.

3.4.2.1 pH-activity profile of *A. niger* 716

The enzymatic activity of *A. niger* 716 at 25 ^{o}C was scanned across various pH values to determine the pH optimum of the enzyme using 2,3-dihydroxybenzoate (2,3-DHB) as a substrate. Some overlapping pH values were tested to determine whether the buffering system influenced the enzymatic activity. The pH optimum was found to be around 6.0 while surveying the wide range of pH values shown in Fig 3.27A, while in Fig 3.27B the use of a much narrower pH range with more data points established a pH optimum of 6.2.

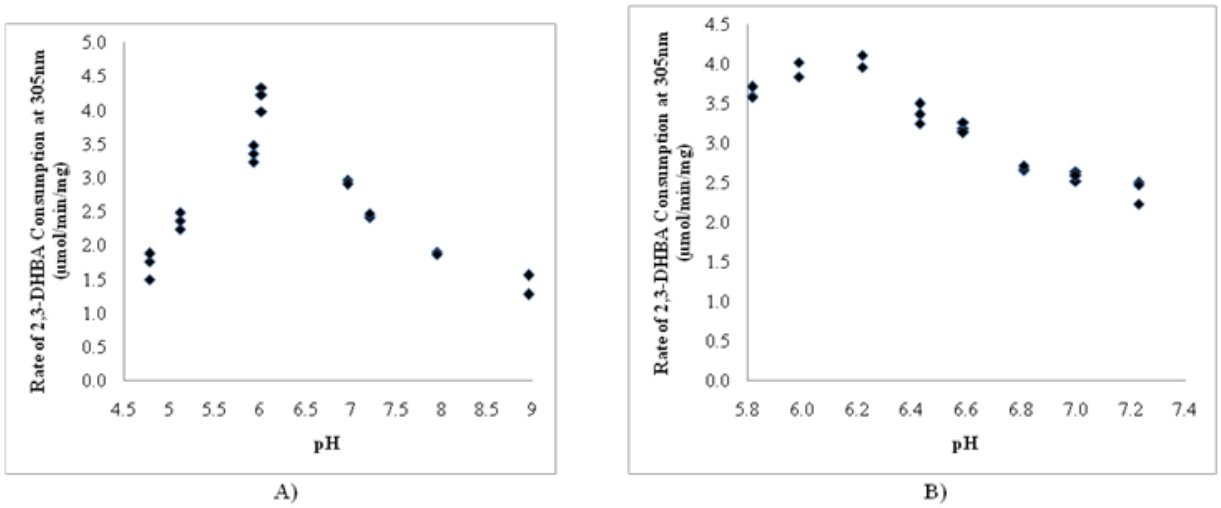


Figure 3.27: *A. niger* 716 activity as a function of pH over wide (A), and narrow (B) ranges. Buffers were 20 mM sodium acetate (pH 4.8-5.9), 20 mM PIPES (pH 6.0-7.2), 20 mM Tris (pH 6.9-9) for panel A and 20 mM Bis-tris was used for all buffers in panel B. Data were obtained in triplicate for all pH values.

3.4.2.2 *A. niger* 716 substrate specificity

As was shown previously using enzyme purified from the native *A. niger* host[4], the recombinant enzyme was specific for 2,3-dihydroxybenzoate decarboxylation: all other tested substrates showed activities less than 2 % (Table 3.21).

Substrate	Substrate consumption (µmol/min/mg)	Relative activity* (%)
2,3-dihydroxybenzoic acid	3.72	100.0
2,4-dihydroxybenzoic acid	0.01	0.4
2,5-dihydroxybenzoic acid	-	-
2,6-dihydroxybenzoic acid	0.02	0.5
3,4-dihydroxybenzoic acid	0.06	1.7
Salicylate	-	-
3-hydroxybenzoic acid	0.00	0.0
Anthranilic acid	-	-

* relative to 2,3-dihydroxybenzoic acid activity
- no activity

Table 3.21: *A. niger* 716 activity on different aromatic substrates. Reported activities are averages of duplicate measurements with a maximum difference of 5 %.

3.4.2.3: Oxygen sensitivity and reversibility of *A. niger* 2,3-dihydroxybenzoate decarboxylase

The activity of the enzyme was tested to see if reducing agents affected its activity since many aromatic-ring decarboxylases, such as 3,4-dihydroxybenzoate decarboxylase from *Clostridium hydroxybenzoicum* JW/Z-1T, are sensitive to O_2[41]. Table 3.22 shows that the activity was unaffected by the presence of 1 mM DTT in the assay. Furthermore, pre-incubation of the enzyme with 10 mM DTT for 10 min did not have a significant effect on the activity in comparison with a control sample diluted appropriately with buffer.

The enzymatic carboxylation of catechol to 2,3-dihydroxybenzoate was also examined in the presence and absence of reducing agent since many aromatic-ring decarboxylases are also sensitive to oxygen for the reverse reaction[41]. Carboxylation occurred when the enzyme was added to catechol and CO_2 supplemented by potassium bicarbonate, as shown in Table 3.23. Similarly to the decarboxylation reaction, the carboxylation reaction is oxygen insensitive since similar rates are observed with and without DTT.

Condition	Relative Activity* (%)
No DTT	100.0
Assay with DTT (1 mM)	101.6
Incubated with DTT (10 mM)	83.5
Dilution Control	80.8

* relative activity compared to activity in the absence of DTT

Table 3.22: 2,3-Dihydroxybenzoate decarboxylase activity in the presence and absence of DTT. When enzyme was pre-incubated with 10 mM DTT, addition of a 1 in 10 volume resulted in the presence of 1 mM DTT in the assay. The dilution control was when the same volume of enzyme was added with no DTT to account for enzyme dilution effects. Measurements above are averages of duplicates, which vary by a maximum of 5 %.

Assay with	Rate of 2,3-Dihydroxybenzoate formation (µmol/min/mg)	
	No DTT	With DTT
No $KHCO_3$	0.0258	0.0099
No Enzyme	0.0092	0.0040
0.25 M $KHCO_3$	0.1170	0.1284

Table 3.23: Reversibility of *A. niger* 716 catalyzed decarboxylation of 2,3-dihydroxybenzoate decarboxylase in the presence and absence of DTT. Reactions were performed in 100 mM phosphate buffer pH 6.8 with 300 µM catechol at 25 ^{o}C Values represent duplicate averages which had a maximum difference of 2%.

3.4.2.4 Oligomeric state of *A. niger* 2,3-dihydroxybenzoate decarboxylase

Protein molecular weight and oligomeric state were determined for the *A. niger* 2,3-dihydroxybenzoate decarboxylase in the same manner as the putative salicylate and phenol hydroxylases on the calibrated Superdex 200 10/300 GL gel filtration column (section 3.2.2.4 and 3.3.2.2). The eluted protein was estimated to have a molecular weight of 122.4 kDa, which corresponds to a trimer based on the 39.2 kDa predicted monomeric molecular weight. The elution profile has been included in the *Appendix*.

3.4.2.5 Effects of metals ions on *A. niger* 2,3-dihydroxybenzoate decarboxylase activity

Since similar decarboxylase enzymes, such as 2,6-dihydroxybenzoate decarboxylase from *Rhizobium sp*[39], have been reported to be metal-ion dependant, the effects of metal ions on *A. niger* 2,3-dihydroxybenzoate decarboxylase activity was examined. Prior to assaying the enzyme it was incubated in various metal ions. None of the metal ions resulted in increased activity, however, zinc, copper, and cobalt had inhibitory effects on the enzyme (Table 3.24).

The amount of each individual metal ion bound to the purified enzyme was determined by ICP-MS both before and after chelator treatment. Samples were prepared for analysis as described in "Materials and Methods", including appropriate buffer controls. A standard curve was prepared for each metal (see Appendix). Of all the metals examined only magnesium and calcium were detected in the purified enzyme, albeit in sub-stoichiometric amounts (Table 3.25). Enzyme samples were also treated with a mixture of two chelators, EDTA and 8-hydroxyquinoline-5-sulfanilic acid, prior to metal analysis, in an effort to remove any metal ions present (procedure adapted from Atkin et al, 1973[67]). After chelation treatment, magnesium was present (Table 3.26) at similar sub-stoichiometric levels as prior to chelation. Calcium levels appeared to be below 0 ppb, a result of the fact that buffer control had a higher calcium concentration than the sample. The reason for this has not been determined.

The enzyme samples treated with chelators were examined for activity in the presence of various metal ions. As shown in Table 3.27 only a small possible difference in activity was observed after chelation supplementing the enzyme with Ca^{2+} revealed that there is a increase in the enzymatic activity after incubation with Ca^{2+}, and a decrease upon Mg^{2+} addition (Table 3.28).

Metal	*Relative activity (%)
None	100
Zn^{2+}	3
Mg^{2+}	91
Mn^{2+}	92
Ca^{2+}	99
Co^{2+}	69
Cu^{2+}	73

*relative activity represents activity compared to that with no metal

Table 3.24: Effects of metal ion additions to *A. niger* 716 prior to activity assay. *A. niger* 716 was incubated with 5 mM metal ion for 10 min before being assayed. Reported values represent averages of duplicates with a maximum difference of 5 %.

Metal	*Theoretical [metal] (ppb)	Experimental [metal] (ppb)	% (Theoretical/ Experimental)
Mg	310	117.6 ± 0.5	37.9
Ca	510	84.8 ± 12.3	16.6
Mn	700	14.2 ± 0.1	2.0
Fe	715	18.9 ± 0.2	2.6
Co	750	0.1 ± 0.001	0.0
Ni	752	3.6 ± 0.1	0.5
Cu	810	1.4 ± 0.01	0.2
Zn	835	20.1 ± 0.3	2.4

* Theoretical [metal] is based on a 1:1 metal to enzyme ratio

Table 3.25: ICP-MS metal analysis of *A. niger* 716. 0.5 mg of purified protein was subjected to nitric acid digestion and the resulting supernatant analyzed by ICP-MS. A buffer control was prepared and measured in the same manner as the protein sample, and subtracted.

Metal	*Theoretical [metal] (ppb)	Experimental [metal] (ppb)	% (Theoretical/ Experimental)
Mg	310	93.8 ± 0.4	30.3
Ca	510	-140.0 ± 4.7	-27.5
Mn	700	14.2 ± 0.1	2.0
Fe	715	8.2 ± 0.4	1.1
Co	750	0.1 ± 0.002	0.0
Ni	752	1.7 ± 0.05	0.2
Cu	810	-3.3 ± 0.015	-0.4
Zn	835	5.5 ± 0.1	0.7

* Expected [metal] is based on a 1:1 metal to enzyme ratio

Table 3.26: ICP-MS metal analysis of *A. niger* 716 samples after chelation with EDTA and 8-hydroxyquinoline-5-sulfanilic acid. 0.5 mg of protein was subjected to nitric acid digestion and the resulting supernatant analyzed by ICP-MS. A buffer control was prepared and measured in the same manner as the protein sample and its metal content was subtracted.

A. niger 716 Sample	Rate of 2,3-dihydroxybenzoate consumption (µmol/min/mg)	*Relative activity (%)
No chelator	4.42	100
Chelator treated	4.13	93.3
$CaCl_2$ supplemented	7.02	159
$MgCl_2$ supplemented	3.11	70.3

*relative activity compares the activity without metal addition and before chelation

Table 3.27: Effects of $MgCl_2$ and $CaCl_2$ supplementation on *A. niger* 716 activity after chelation. Metal supplemented samples containing *A. niger* 716 were incubated in buffer containing 5 mM $MgCl_2$ or $CaCl_2$ for 5 min before being assayed. Assays were performed in duplicate, with a maximum difference of 5%, with averages shown. The activity of chelated sample is corrected for enzyme dilution that occurred during dialysis.

3.5 Bioinformatic Studies

3.5.1 Comparison of putative salicylate hydroxylase sequences

Figure 3.28 presents the amino acid sequence alignments of all target *A. niger* salicylate hydroxylases. Alignments were generated using the ClustalOmega sequence alignment tool[72], with important residues that have been manually highlighted. The characterized *E. festucae* and *P. putida* salicylate hydroxylases sequences have been included to serve as a model for comparing the putative *A. niger* salicylate hydroxylases. In general, residues are well conserved in all regions including the FAD fingerprint (yellow highlight) and FAD and NADH binding site regions (pink), which is consistent with the data shown in the preceding sections that these enzymes are all flavoproteins that use NAD(P)H. The substrate active site motif (grey highlight) is reasonably well conserved as well. Differences in the sequences are apparent in the region with the potential catalytic serine residue (green).

```
E.fest              ----------MATKKDHVFAIIGGGIAGLTLAVALHH----RGLRIKIFEQAGQMQEIGA     46
P.putida(5EVY)      ------------SKSPLRVAVIGGGIAGTALALGLSKS---SHVNVKLFETAPAFGEIGA     45
P.putida23262       ---------MKNNKLGLRIGIVGGGISGVALALELCRY---SHIQVQLFEAAPAFGEVGA     48
JGI577              ---------MSPSKKDFHVAIIGGGIAGVTLAIALHH----RNIPVTIYEQAHEFGEVGA     47
JGI767              -------MTVSNNDKPLRIAIVGGGIAGVTLLLALLKHTSRQNIAPHLYEATPEFSEIGA     53
JGI534              MAQPAASISGQAGSDTMHIAIVGAGIGGLALAIGLTN----RGVSYTLYESAAQFSAVGA     56
JGI655              -----------MSTDSIEVAIIGAGITGITLALGLLS----RGIPVRVYERARDFHEIGA     45
JGI860              ----------MGPPRPFTVTIIGGGIGGLVLAVGLLR----RQVPVQIYEAAAAFAEIGL     46
JGI684              ---------MEQAKTPLQVIVVGAGIGGMAAALTLGL----RGHHVTILESAPKLMEVGA     47
                                    . ::*.** * .  : *           : * :  :  :*

E.fest              GVAFTPNALQAMRVCHPAIYEAYERVRT--RNL--WPCKQKVWFDFYDTRNEDG---D-A     98
P.putida(5EVY)      GVGFGVNAVEAIQ--RLGIGELYKSVADSTPA-----PWQDIWFEWRHAHDASL------     92
P.putida23262       GVGFGPNAVRAIV--GLGLGEAYLQVADRTSE-----PWEDVWFEWRRGSDASY------     95
JGI577              GVGFSPNAVQAMKYCHEGIHNAFEKVCT--RNL--WPTKQKVWFDYLDGYNHKESPQN-G    102
JGI767              GLAFGSNSIRAMELIDSGLAAAYNQHATFQPDT--GALDKKLWSSYRMGMDGRGPPSNSL    111
JGI534              GIGLGPNALHAMELIDTRFRQLYNQISSGNLT----PNKEHIAMEVLYAEEGFGAKRG-W    111
JGI655              GIGFTPNAEWAMKVVDPRIQAAFKRVAT--PN-------ASDWFQWVDGFNESGTDPRET     96
JGI860              GLGIGPAAHRAMPLIDPEIRAIYDALVTTHADSPGYERYKQTWFELVWATGDH-------     99
JGI684              GLQVSPNMLTLFDRWGVSPLIHAQDVALEHIH-------V---RRWEDGSMLGT------     91
                    *: .       :

E.fest              AAKPAFTI-SNDVGQNGVLRAHFLDELIKLVPRE--AAHVGKRLASYEEGGPD-GRLRLR    154
P.putida(5EVY)      ----VGATVAPGIGQSSIHRADFIDMLEKRLPAG--IASLGKHVVDYTENAEG---VTLN    143
P.putida23262       ----LGATIAPGVGQSSVHRADFIDALVTHLPEG--IAQFGKRATQVEQQGGE---VQVL    146
JGI577              RQNIEFTI-SNSLGQNGVHRAHFLDELIKLVPKE--ISRFNKRLEDIHERISD-GKLIMK    158
JGI767              KAGDYLHEVHAPVARSSVHRARFLEAMTALLPPEGGFVSFGKRLVSIEEGADG---ATVH    168
JGI534              QPAPFG---AEFYTRTSAHRRDLLEIMTSLIPQE--RVQFNKRVKHINQTGEN-SKIIIT    165
JGI655              EEQLLFKIYLGERGFEGCHRADFLGELARLLPEG--VVTFQKALDTVEPAADNSLGQLLR    154
JGI860              EGETLMDLKALPSGQTTLRRADFLAALVALVPPE--IVHFGKRLNQLAETADG---VRLQ    154
JGI684              -----LPVNKTYGQQTVIHRADLHNALIEKALA---LSNVELRVNSTVTGV-QFDHAAVQ    142
                                  *  :   :           .                :

E.fest              FADGSEDEADVILACDGIKSRVRQLLFG-AHHPCALPSYTHRYAYRALVPMEEAVDA---    210
P.putida(5EVY)      FADGSTYTADVAIAADGIKSSMRNTLLRAAGHDAVHPQFTGTSAYRGLVETSALREAYQA    203
P.putida23262       FTDGTEYRCDLLIGADGIKSALRSHVLEGQGLAPQVPRFSGTCAYRGMVDSLHLREAYRA    206
JGI577              FADGSEDVADLVIGCDGIKSQVRQIIVG-ADHPSAKPSYTHKYAYRGLVPMDKAVEA---    214
JGI767              FADGTSVVADAVLGCDGIKSRVRQILLE-GE-EAANAVFTGKYAYRGLIPMDEAVAA---    223
JGI534              FEDGSIAEADAVIGCDGVKGPTRQAVLGTLYPDKVQPIFSGKYVYRSIVPMEDAQRL---    222
JGI655              FQDGTTATAHAVIGCDGIRSRVRQILLG-EDHPTASAHYSHKYAARGLIPMDRAREA---    210
JGI860              FEDGTTAMADVVVGCDGIKSKVKESMLP-EESAVKQPRYSGMYGYRAVFDMDEMVAA---    210
JGI684              LANGTIVKGDIVIAADGIKSALRDHLLE----EASVAIPTGDAAYRIMLPRSALENDPEL    198
                    : :*:    .  :..**::.  :. :.           :     * :.
```

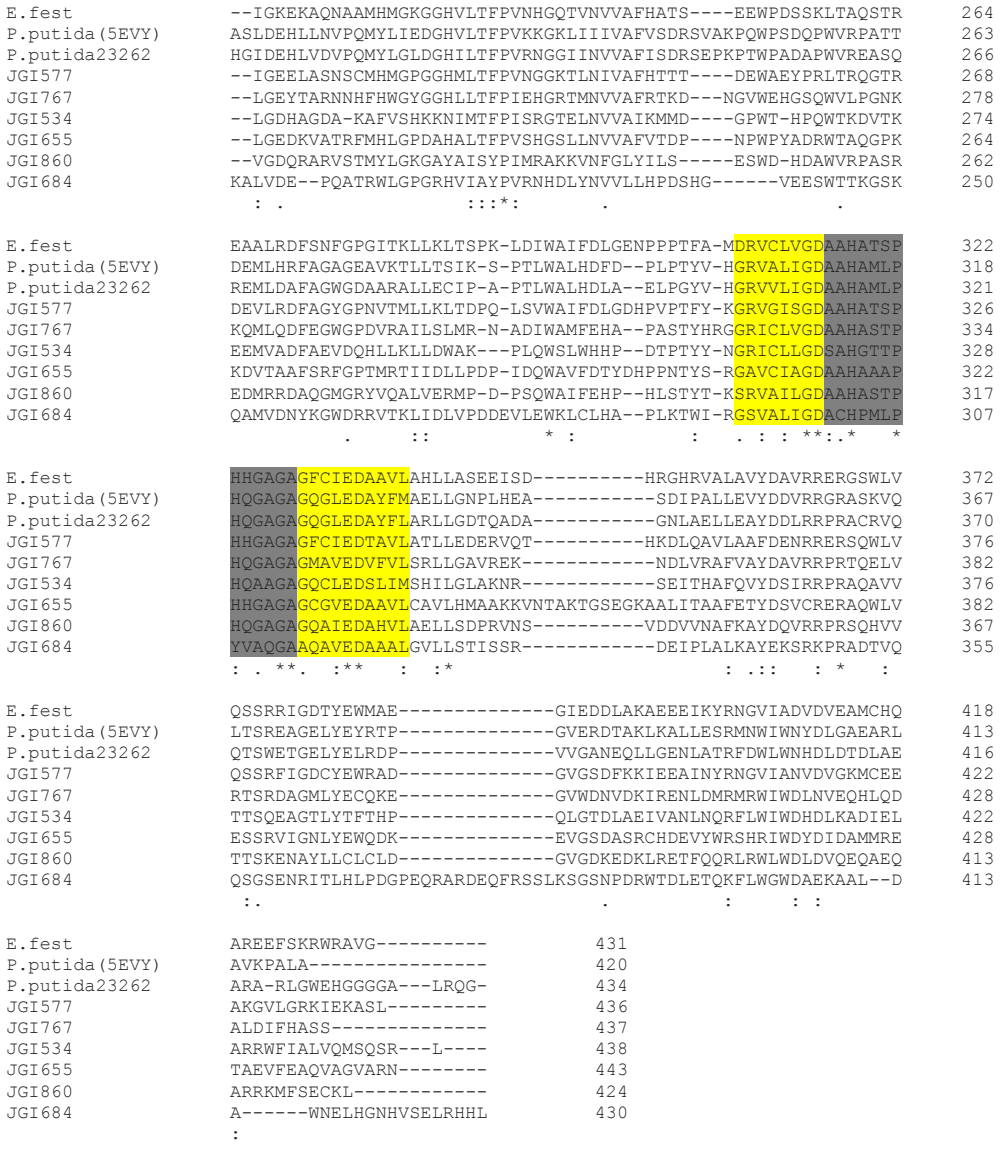


Figure 3.28: Salicylate hydroxylase sequence alignments. Alignments were generated using ClustalOmega[72]. An asterix (*) represents conserved, a colon (:) strongly conserved, and a period (.) weakly conserved residues. *A. niger* target sequences from this study and *E. festucae* and *P. putida* salicylate hydroxylase sequences have been included[33,80]. Important residues identified in salicylate hydroxylases are highlighted; FAD fingerprint regions (yellow); NADH and FAD binding sites (pink); salicylate binding and active site (gray); potential catalytic serine (green).

3.5.1.1 Comparison of putative phenol hydroxylase sequences

The sequence alignment of the putative phenol hydroxylases is presented in Fig 3.29. The *T. cutaneum* sequence has been included since it is a known eukaryotic phenol hydroxylase that is induced by the growth of *T. cutanueum* on phenol, and is highly active and tightly coupled in turning over phenol compared with similar compounds[74,81]. The alignment of key catalytic residues (Asp 54 and Tyr 289) shows that in most of the sequences these residues are well conserved, but in some cases these residues are different and may lead to differences in enzymatic properties. Additionally, residues that are part of the FAD fingerprint and Rossmann fold are all well conserved, as would be expected for NAD(P)H-dependent flavoproteins.

```
PhenolHydroxylase_Tcutaneum      ----------------------------------MTKYSESYCDVLIVGAGPAGLMAARVL   27
A_niger999                       MPVFTEYASKSRDLRVLPSFSLPVPRLTPPFEPTVNDGVERYEVVIVGAGPSGLMLELLL   60
A_niger687                       ---------------------------------MAPNSSPSQTDLLIIGAGPAGLMAACWA   28
A_niger951                       ----------------------------------MTVSIESTTDVLIIGAGPSGLTIAYWM   27
A_niger422                       -------------------------------------MTQKVDVLICGSGSAGLCAATWL   23
A_niger146                       --------------------------MTIAPYPVGAFKSQRTDVLIVGAGPAGCMAAATL   34
A_niger300                       ----------------------------------MAPSQQEKYDIVIVGAGPVGIVLSLCM   27
A_niger617                       -----------------------------------MTEIKKTDVFISGGGPTGLLTAYAL   25
A_niger718                       ----------------------------------MPSPETYDVTIIGGGPVGLLLAYQL   25
                                                                        .  :: * *.*  *

PhenolHydroxylase_Tcutaneum      SEYVRQKPDLKVRIIDKRS---TKVYNGQADGLQCRTLESLKNLGLADKILS-EANDMST   83
A_niger999                       ARYGLS--DESLLCIDSKP---STLKSGQADGVQPRTLEVFKTLGISDEIEN-EACQMWQ   114
A_niger687                       SQY----TNMTTRIIDKHP---SRTATGHADGIQSRTLEILESFGLVDPIVR-KGVQHAE   80
A_niger951                       AQY-----GINARIIDKRS---TKLFRGHADGLRGRTLELIDSMGFVHRILQ-EGYHGAG   78
A_niger422                       ARY-----GIKCKVLERRD---GPMTMGQADGVQCRTVEIFESFGIGEELLR-ESYHVVE   74
A_niger146                       QRY-----GIDFCLIDKRP---TRTQTGHASAFQPRTQEILQTMNLLHDLDK-RGHRLTE   85
A_niger300                       SRW-----GYKVKHIDNRP---VPTATGRADGIQPRSTEILRNLGLKRQIMAFKPAKVYD   79
A_niger617                       ARQ-----GVDTLFVEQHDK-AQQAMYGRATTLHPRTLELLDQLSILDDLNQ-IGYVARD   78
A_niger718                       SRF-----ELSVCVVEKQDKNSPEGRYGRATTLFPRTLELLDQLDLVHTMLQ-QGFACRS   79
                                  .           :: :         *:*  .  *: * :  :.:   :

PhenolHydroxylase_Tcutaneum      IALYNPDENGH-IRRTDRIPDTL-PGISRYHQV-VLHQGRIERHILDSIAEISDTRIKVE   140
A_niger999                       FAFWNTSSDPSKIIERKSVVPEV-IPPARFRYEATIHQGRIERIMETDLLRYSKRGVQRG   173
A_niger687                       MCFWGYDKNTG-GITRQKMDPAQ-TDKSRYNMV-LLNQGMIEQIILDYLELEREKRVEVE   137
A_niger951                       LRYWRPDGKGG-LKRAGELRSTI-WNSSPYRMA-LF--------------------IDVE   115
A_niger422                       VVFWADDSGSGVIRRTGKTADTQ-PGLSHQPHV-ILNQARINGLLIELMKKYNGQEIDYG   132
A_niger146                       TSFWMRDSTGALISNFTGA--EV-VHATPYQYLFNTDQGMTEDVFEQYLNTK-GQKI---   138
A_niger300                       VAFWDPLPGGQDIHRTGSWPSCPRFIDTRYPFTTLVHQGKIERVFLDEIQKA-GTTVERP   138
A_niger617                       SVTYRD---GKRVSSRGWEIMYQHMRGTFLDYCLNIRQKFSEEVIRDAYVSLRGEPYIGW   135
A_niger718                       SVTYKD---GVRQFPGKVWTFMENIQGTVFDFALVLRQMYTEGILRKRLDKEKV-TYHGS   135
                                    :                         :

PhenolHydroxylase_Tcutaneum      RPLIPEKMEIDSSKAEDPEAYPVTMTLRYMSDHESTPLQFGHKTENSLFHSNLQTQEEED   200
A_niger999                       TKLVHV------QLNEDDKEFPVLAEIE--------------------------------   195
A_niger687                       YAKQAKTLEF------LKDEYPIVVGVKRVG----------------------L-------   163
A_niger951                       RGVIGESLSYDESLENDPQAYPITIKLRTLSEEEATPIPNPGITPYKPARDDL-------   168
A_niger422                       YTVQDVRVD-ESQLSAGPEAYPVKVTAE--------------------------------   159
A_niger146                       ----QRFMELVHYEHDLDPEWPLTAYI--------------------------------   161
A_niger300                       WTIVGF------KNDGLDETYPVEVQLK--------------------------------   160
A_niger617                       -RLDGFT-----VEEACDDDYKVVSHVT--------------------------------   157
A_niger718                       MECVAFE-----IG-LDGSEYPVT--VH--------------------------------   155
                                                    : :

PhenolHydroxylase_Tcutaneum      ANYRLPEGKEAGEIETVHCKYVIGCDGGHSWVRRTLGFEMIG-EQTDYIWGVLDAVPASN   259
A_niger999                       ---------TRGVRRLIRTKHLVGADGAHSMVRQCMGLRLEG-QSLDHIWGVIDLVVDTN   245
A_niger687                       ----DGANGHCDQVETIRARYVIGCDGARSWVREQLQVPMRV-CSGDTTWGVVDIAPITD   218
A_niger951                       EDLVPKRKHDVGTIETVRAKYLIGCDGAHSWTRKQLNIPFEG-SHTDHIWGVMDVVPITD   227
A_niger422                       ---------KDGKVESFEAKYALACDGAHSTVRKALGYNMIG-DSTDAVWGVMDMVPRTD   209
A_niger146                       ------KNNASGAIEAWQTKYILGTDGARSATRRATGVQSSS-QGGEDVWAVADVYVDTN   214
A_niger300                       -------SIDTNVIETVRTKYLFSGEGARSFIRQQLGIQIQYKDPISYVWGVMDGVVRTN   213
A_niger617                       -------EVSSGRSLTVRSKFIVGADGSHSLVRRLAGIPFEG-DHTNFRWVRIDGRFKTD   209
A_niger718                       -------CSGPGGMMTAKSKYLVGADGGHSLVRRYANIPFDG-DSSEDQWIRIDGIVETN   207
                                            .     . :. .. :*.:*  *.           .  *   *    ::
```

```
PhenolHydroxylase_Tcutaneum      FPDI-RSPCAIHSAESGSIMIIPREN-----NLVRF-YVQLQARAEKG-GRV--------   303
A_niger999                       FPDI-RRRCAIHSP-AGSVMVIPRERIVTGEYLTRL-YVQVPGMANPDSDIAVGATPVSE   302
A_niger687                       FPDI-RQSCAIHAGDRGSIMTAPREN-----RLVRF-YIYPKGDGRLSVEGK--------   263
A_niger951                       FPDI-RWPSSIDSP-HGHLLVITREG-----NTVRF-YVPLQEEDEIT-PSF--------   270
A_niger422                       FPDI-RKKSSIRSK-AGNILVIPREGDS--YNLTRF-YIELPTGT---------------   249
A_niger146                       FPDY-RRRCAIRTP-DGGCMLIPRKD-----EGLRI-FLQVDEKSQEHLDEN-GA---SG   262
A_niger300                       FPDI-ETKCTIHSD-AGSIMVIPRED-----NMVRL-YVQIASSDDPDFD----------   255
A_niger617                       MPEADIGFASIESKSHGNVLWVQLDH-----GVKRIGFAMTPEMLAR-------------   251
A_niger718                       MPIN-RAYGAIETTTHGNVLWAPLDH-----GATRIGVAYTPEIAAK-------------   248
                                 :*       :* :   *  :    .        *: :

PhenolHydroxylase_Tcutaneum      ------DRTKFT-PEVVIANAKKIFHPYTFDV---QQLDWFTAYHIGQRVTEKFS-KD--   350
A_niger999                       LDDARARRQKVT-KEGILKQASDAFKPYYLRPKSDESIDWWAAYQIGQRVSPEFTVKDSA   361
A_niger687                       ------DRSEVS-LDEIVGAMAKVMKPYTLTY---KHCDWWSIYKIGQRLVETYR-PH--   310
A_niger951                       ------DRSAVT-LDMLCERVKSIMSPFYFDF---KVCHWWAAYQVGRRLAPAVN--K--   316
A_niger422                       ------KAKDVR-LEDLQNAAKQIFSPYQIDF---VETVWWSAYAIGQRHADFFH-KD--   296
A_niger146                       QD-ALTGNSAFKLTQTVQSHINKVIHPYKMNI---TDIVWISQYRVAQRVVHHFSDPT--   316
A_niger300                       -------PRKTATAEDVQATARKILQPYWVEW---DRVEWYSVYPIGQGISEKYT-LD--   302
A_niger617                       ------YGDKLT-LEDAKVEAVKSMEPFSLEF---ETIDWWTLYGINQRVAECFYN-N--   298
A_niger718                       ------YPEGVT-EEVAVNEAIACLRPFNLKF---KEVHWWTLYKIGQRMARTFATHN--   296
                                               :         : *: .         * : * : :

PhenolHydroxylase_Tcutaneum      --ERVFIAGDACHTHSPKAGQGMNTSMMDTYNLGWKLGLVLTGRAKR--------DILKT   400
A_niger999                       GVPRVFIVGDACHTHSPKAGQGMNVSMMDSYNLAWKLAYHIHGLSPASAEPGQLHPLLDT   421
A_niger687                       --DRIFLAGDAAHTHSPKAGQGLNVSMQDTYNLTWKLCSVITGEASP--------SILDT   360
A_niger951                       --GRMYLVGDAVHTHSPKIGLGMNMSIQDGFNLGWKLALAVKEMASP--------SILDT   366
A_niger422                       --YRVFLAGDACHTHSPKAGQGMNVSLQDGYNMGWKLASVLKGLAHP--------SLLET   346
A_niger146                       --KRVFLLGDACHTHSPKAGQGMNVSISDAYNLTWKLALVMKGVAKA--------SLLET   366
A_niger300                       --ERVFMGGDACHTHSPKAGQGMNTAFHDALNMAWKLHAVESGLAKR--------SILST   352
A_niger617                       --DCVLLAGDACHTHSSGTAQGMNTGVHDAVNLAWKLGGIVKGWYND--------DVLRS   348
A_niger718                       --NRVFICGDAAHTHSSGAAQGLNTGIHDAVNLAWKLALEVHGLSHP--------EVLNT   346
                                    : : *** ****   . *:* .. *  *: ***                  :* :

PhenolHydroxylase_Tcutaneum      YEEERHAFAQALIDFDHQFSRLFSGRPAKDVADEM--------GVSMDVFKEAFVKGNEF   452
A_niger999                       YHLERHANAQQLIDFDRKFSTIFSGQLDTEE-----------SGMSHAEFVATFNTGNGF   470
A_niger687                       YELERRPVGEELLKLDTELVQAYEQDTSVTE------------G--VERVR---GQYIDF   403
A_niger951                       YKSERHQLAKMLLEFDQKWAAFFLKQKKKQQQLGL-EAPPEANPEDIQAMQDVFSENELF   425
A_niger422                       YVLEREKVAIDLINFDRYFAKLFSSQGSAT----------------PKEFQEGFIKSGKY   390
A_niger146                       YEQERLWVAQQLIEFDALFARQFGQKDKLD----------------SQNLRETWEMGHGF   410
A_niger300                       YETERKDIAETLLSFDNKYAALFSKRRPTAGEVGEASHTTAAAGGEEDEFVKTFKSSCEF   412
A_niger617                       YDTERRPVAQHLIDLDKDFSATISGTVPDRYKD-L-------SLSSNELFAKVFDDTIDF   400
A_niger718                       YTTERQSAVQRLLNYDRDISLLMTHKWPVWYDGDR-------SADLNVLLGEIFQDAAQF   399
                                 *  **      *:. *                                .        :

PhenolHydroxylase_Tcutaneum      ASGTAINYDENLVTDKK---SSK------QELAKNCVVGTRFKSQPVVRHSEGLWMHFGD   503
A_niger999                       TSGCGVEYGESMIVDRTASDGQDGPIRGTDYLSGILRPGRRLLNVKLRRHADGWHRDIQD   530
A_niger687                       MTGVGVTYGPNQLVAGK---KAD------TAVANNIKIGMRLSSHPVVGQADGISMQLSS   454
A_niger951                       AEGHVSFYKASPIVHK-----GS------TTVARHLTAGERFPVALIRKQADGQPWWTSR   474
A_niger422                       TAGLTAKYDESPITSDT---GVS------NELAKNVVVGMRLPSAQVVRFCDSKPLQLAA   441
A_niger146                       TSGCGYEYPANLLVNPDVRTSIN------NQAVEPLTPGKRLLPIDLIRHIDGNHVRMLD   464
A_niger300                       TSGYGVAYKPNVFTWDATHPAQSPL---FDVPGVRLTPGRAFTPTTVTRLADSNHVHLEQ   469
A_niger617                       CIGLGIHYNESTINKT--------------SLTTAISAGWRAPDALVYAPGSRLPVRLFH   446
A_niger718                       NTGLGISYEANVINQPL---EP------STEVAVGVQPGSRAPDTELTMPGTFQSVRMHQ   450
                                   *    *  . .                          *       :

PhenolHydroxylase_Tcutaneum      RLVTDGRFRIIVFAGKATDATQMSRIKKFSAYLDSENSVISLYTPKV-------------   550
A_niger999                       DLASTGRFRVLTLTSDDLLNAQSASA----------RSLQDVSALLG-------------   567
A_niger687                       RLTSNGAWRLLVFPGDLREPRNQSRLSRFTNQFNEQPHLPCLHKDQS-------------   501
A_niger951                       LLKSDGRFRILLLVGDCRLEDQKSRILALNNGLLE---LQKRFTPPK-------------   518
A_niger422                       SLKSDCRWRVIAFSGDLTVSQNMAALQAIGGYLGSDDSPLQRFRLRS-------------   488
A_niger146                       IMPSNGRFHLFIFAGNDLSS---PTLQKLGNTLDSPHSPMSLFNLLPLELMERFRHEDIT   521
A_niger300                       EIPANGAFRIFIFAGKQDKTS--KAITDLAANLEKERSFLSVYRRADIADVSFFEN----   523
A_niger617                       LMPNVGAWSILLLAGH-PEKTRQSLI-RAVEMVKQLSES----------------------   483
A_niger718                       VLRNRCQFLAVVFTGGDIETAKLDLL-PLREYLDSHPEL----------------------   488
                                  :     :  . : .

PhenolHydroxylase_Tcutaneum      ------SDRNSRIDVITIHSCHRDDIEMHDFPAPA----LHP-KWQYDFIYADCDSW--H   597
A_niger999                       ------TFPAGVIEQIVVHPQLPNDFEWDNLPACVKEHAEMR-------FY--------N   606
A_niger687                       ------NRCSPLLETIVVHSSPRTSVNLLDLPEIFHPFDETW-GWDYSRVFADDGSY--G   552
A_niger951                       ------QKLDSVIEVKAVHSAPVDDVELSDFPPILRQFDDTT-GWDYTKVWSDSECF-WD   570
A_niger422                       ------GEPDSLIEPILVAHGERHGFELEQIPQSFYPFTGKNHIRDLHKIFFDDESY--N   540
A_niger146                       TASVPFTNKAYVLDLFLIHSQNHLDVQLGNIPSPFS---TK---WPMN-VYSD-------   567
A_niger300                       --H---LPHSKLFSICLVYAGEKNQIDVDSIPKILRDYHHHL--------YADNIPDVRV   570
A_niger617                       ------LPREMIRSLTVVAQSPAGGDQRFSIPTIGR-------------LYYD-------   517
A_niger718                       ------STHPAIAWLTVCGSAGCSPYEVLGMTGFGD-------------TYFD-------   522
                                                        :   :                :
```

```
PhenolHydroxylase_Tcutaneum      H-PHPKSYQAWGVDETKGAVVVVRPDGYTSLVTDLE----GTAEIDRYFSGILVEPK-EK   651
A_niger999                       GSELEDAYGIYGVDPHQGAMVVVRPDGYVGIVVALG----EVDRVGEYLRRCIRTK----   658
A_niger687                       E-ACGQAYERYGIRDETGCLVLCRPDQHVAWLGSLV----EVEALDQYFGQFVPGS----   603
A_niger951                       EQCKGNAYEIWGVDRSRGAMVALRPDQYIGWIGELD----DVAGVTTYFEGFLREPR-PE   625
A_niger422                       R-GHGHLYEHLGISPDKGAIVIVRPDQYVSAVFSLD----EYRRIGQFFEGFMIPQERPE   595
A_niger146                       --FDGAAQNQLGVPEDSGALVVVRPDGYIGLVTGLD----NVEDVTSYFDGFMHRRI---   618
A_niger300                       PQATYAAHEKLGFDVEKGGVVVTRPDSHVACTIQLSEGSGTVDALNAFFGSFATKPLGQD   630
A_niger617                       --PEGSAHAAYTVSDSNGAVVVVRPDGILGYATQLD----EASNVIKYFAGFITR-----   566
A_niger718                       --ARGIAHIAYKLEPRKGRLVVIRPDGLIAFTCTLD----GEAIRQHFFRILKSQPEKTC   576
                                           .    * :*  ***   .    *           ::

PhenolHydroxylase_Tcutaneum      SGAQTEADWTKSTA      665
A_niger999                       --------------      658
A_niger687                       --------------      603
A_niger951                       AVL-----------      628
A_niger422                       SLTGSKL-------      602
A_niger146                       --------------      618
A_niger300                       SQ-QSRL-------      636
A_niger617                       --------------      566
A_niger718                       S-------------      577
```

Figure 3.29: *A. niger* target phenol hydroxylase alignments. Alignments were generated using ClustalOmega[72] where an asterix (*) represents conserved, a colon (:) strongly conserved, and a period (.) weakly conserved residues. *T. cutaneum* phenol hydroxylase has been included as a reference sequence where its two known catalytic residues have been highlighted in gray[76]. In addition, Rossman fold fingerprint regions have been indicated in yellow.

3.5.1.2 Comparison of aromatic ring decarboxylases

The sequence alignment in Fig 3.30 aligns the sequence of the *A. niger* 2,3-dihydroxybenzoate decarboxylase with several other characterized aromatic ring decarboxylases from other organisms (both bacterial and fungal). The alignments reveal that the *A. niger* sequence contains three of the four proposed metal binding residues[39], consisting of Glu8, His164, and Asp287 while the fourth residue, His10, is not conserved. These four proposed metal binding residues are highlighted in gray (Fig 3.30).

```
2,3_DHB_decarboxylase_A_niger            MLGKIALEEAFALPRFEEKTRWWASLFSVD--PETHVKEITDINKLRIEHADKYGVGYQI      58
2,6_DHB_decarboxylase_Rhizobium          MQGKVALEEHFAIPETLQDSA----GFVPGDYWKELQHRLLDIQDTRLKLMDAHGIETMI      56
Salicylate_decarboxylase_T_monoliiforme  MRGKVSLEEAFELPKFAAQTKEKAELYIAPNNRDRYFEEILNPCGNRLELSNKHGIGYTI      60
2,3_DHB_decarboxylase_A_oryzae           MLGKIALEEAFALPRFEEKTRWWASLFSTD--AETHVKEITDINKIRIEHADKHGVGYQI      58
DHB_decarboxylase_Polarmonas_sp          MNGKIALEEHFATEETLMDSA----GFVPDKDWPELRSRLLDIQDRRVRLMDEHGIETMI      56
                                         * **::*** *   .   .:     :          .: :    *:.  : :*:   *

2,3_DHB_decarboxylase_A_niger            LSYTAPGVQDIWDPVEAQALAVEINDYIAEQIRDKPDRFGAFATLSMHNPQEAASELRRC      118
2,6_DHB_decarboxylase_Rhizobium          LSLNAPAVQAIPDRRKAIEIARRANDVLAEECAKRPDRFLAFAALPLQDPDAATEELQRC      116
Salicylate_decarboxylase_T_monoliiforme  YSIYSPGPQGWTERAECEEYARECNDYISGEIANHKDRMGAFAALSMHDPKQASEELTRC      120
2,3_DHB_decarboxylase_A_oryzae           LSYTAPGVQDIWDPVEAQALAVEINDYIAEQVRVNPDRFGAFATLSMHNPKEAADELRRC      118
DHB_decarboxylase_Polarmonas_sp          LSLNAPAVQAIADSTRANETARRANDFLAEQVAKQPTRFRGFAALPMQDPELAARELERC      116
                                          *  :*. *   :  ..   * . ** :: :   .  *: .**:* :::*. *: ** **

2,3_DHB_decarboxylase_A_niger            VQTYGFKGALVNDTQRAGPDGDDMIFYDNASWDIFWQTCTELDVPLYLHPRNPTGTIYEK      178
2,6_DHB_decarboxylase_Rhizobium          VNDLGFVGALVNGFSQEG-DGQTPLYYDLPQYRPFWGEVEKLDVPFYLHPRNPLPQD-SR      174
Salicylate_decarboxylase_T_monoliiforme  VKELGFLGALVNDVQHAGPEGETHIFYDQPEWDIFWQTCVDLDVPFYLHPEPPFGSYLRN      180
2,3_DHB_decarboxylase_A_oryzae           VEKYGFKGALVNDTQRAGPDGDDMIFYDNADWDIFWQTCTELDVPFYMHPRNPTGTIYEK      178
DHB_decarboxylase_Polarmonas_sp          VKELGFVGALVNGFSQDN-RSAVPLYYDMAQYWPFWETVQALDVPFYLHPRNPLPSD-AR      174
                                         *:  ** *****. .: .  .   ::**  .:  **     ****:*:**. *      .

2,3_DHB_decarboxylase_A_niger            LWADRKWLVGPPLSFAQGVSLHVLGMVTNGVFDRHPNLQLIMGHLGEHVPFDMWRINHWF      238
2,6_DHB_decarboxylase_Rhizobium          IYDGHPWLLGPTWAFAQETAVHALRLMASGLFDEHPRLNIILGHMGEGLPYMMWRIDHRN      234
Salicylate_decarboxylase_T_monoliiforme  QYEGRKYLIGPPVSFANGVSLHVLGMIVNGVFDRFPKLKVILGHLGEHIPGDFWRIEHWF      240
2,3_DHB_decarboxylase_A_oryzae           LWADRKWLVGPPLSFAHGVSLHVLGMVTNGVFDRHPKLQIIMGHLGEHVPFDMWRINHWF      238
DHB_decarboxylase_Polarmonas_sp          IYDGHAWLLGPTWAFGQETAVHALRLMGSGLFDKYPALKIILGHMGEGLPYSMWRIDHRN      234
```

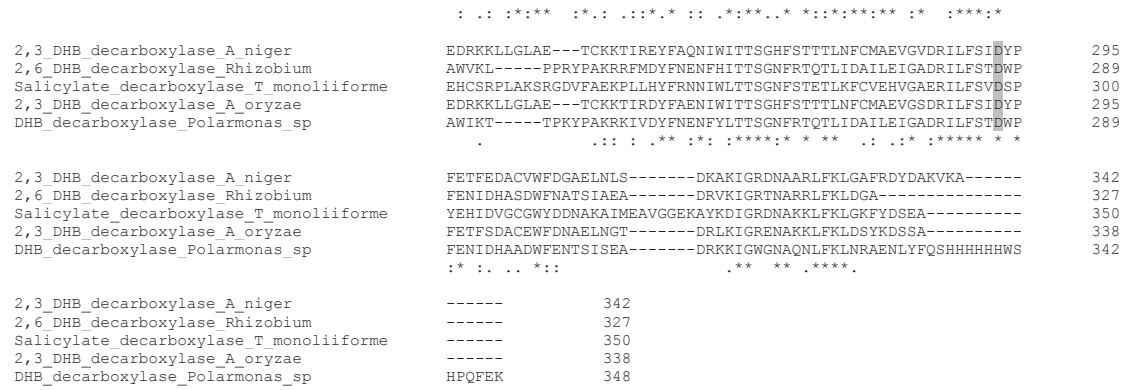


Figure 3.30: Alignment of *A. niger* 2,3-dihydroxybenzoate decarboxylase with selected characterized aromatic-ring decarboxylases. Alignments were generated using ClustalOmega[72] where an asterix (*) represents conserved, a colon (:) strongly conserved, and a period (.) weakly conserved residues. *A. niger* 716 has been aligned against 2,6-dihydroxybenzoate decarboxylase from *Rhizobium* sp. Strain MT5-10005[39], salicylate decarboxylase from *Trichosporon monoliiforme*[41], 2,3-dihydroxybenzoate decarboxylase from *Aspergillus oryzae*[38], and dihydroxybenzoate decarboxylase from *Polarmonas sp*[40]. Residues potentially involved in metal ion binding have been highlighted in gray.

3.5.2 *A. niger* salicylate hydroxylase active site motifs

Salicylate hydroxylases have been proposed to have two domains involved in the salicylate hydroxylase active site consisting of AAH(A/S)(M/L)(L/V)PH(Q/H)G(Q/A)GA and AGV[82]. The *A. niger* putative salicylate hydroxylases described above, *A. niger* 577, 767, and 534 have been modeled, and features at the predicted active sites have been aligned using PyMOL and are depicted in Figure 3.31. This alignment shows there are no major differences in the protein folds for these regions.

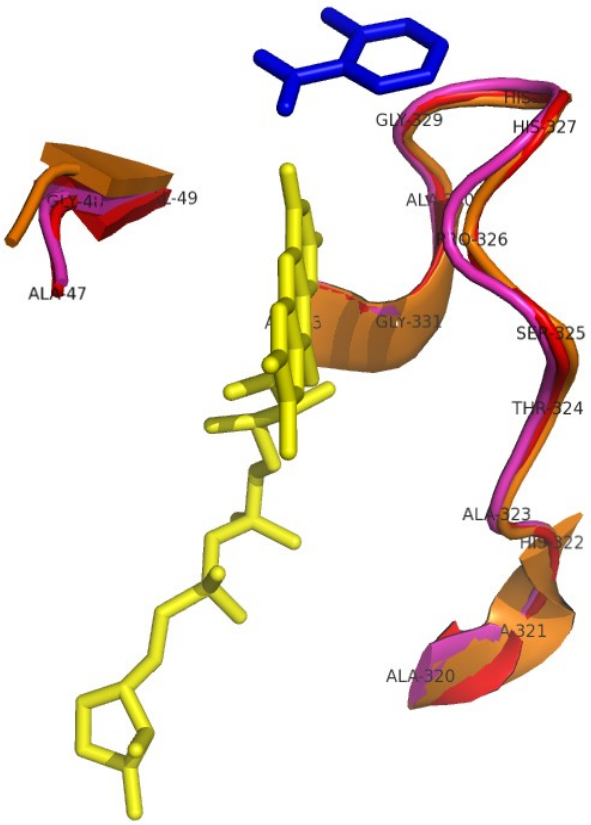


Figure 3.31: Alignment of the predicted active site structures of the *A. niger* putative salicylate hydroxylases. Models were predicted using I-TASSER and the overlaid image was generated using PyMOL; *A. niger* 577 (red); *A. niger* 767 (orange); *A. niger* 534 (fuchsia); FAD (yellow); salicylate (blue). Residues have been labelled according to *A. niger* 577.

3.5.2.1 Accessibility of salicylate to salicylate hydroxylase active sites

Potential entrances for substrate (salicylate) to the hydroxylase active sites are shown in Figure 3.32. Panel A, B, C, and D represent salicylate (blue) in the *P. putida*, *A. niger* 577, *A. niger* 767, and *A. niger* 534 active sites respectively. In *P. putida* and *A. niger* 577 salicylate has easy access through an appropriately sized entrance to reach the active site while in the case of *A. niger* 767 and 534 the entrance seems to be much smaller and substrate may not be able to reach the active site efficiently through this opening.

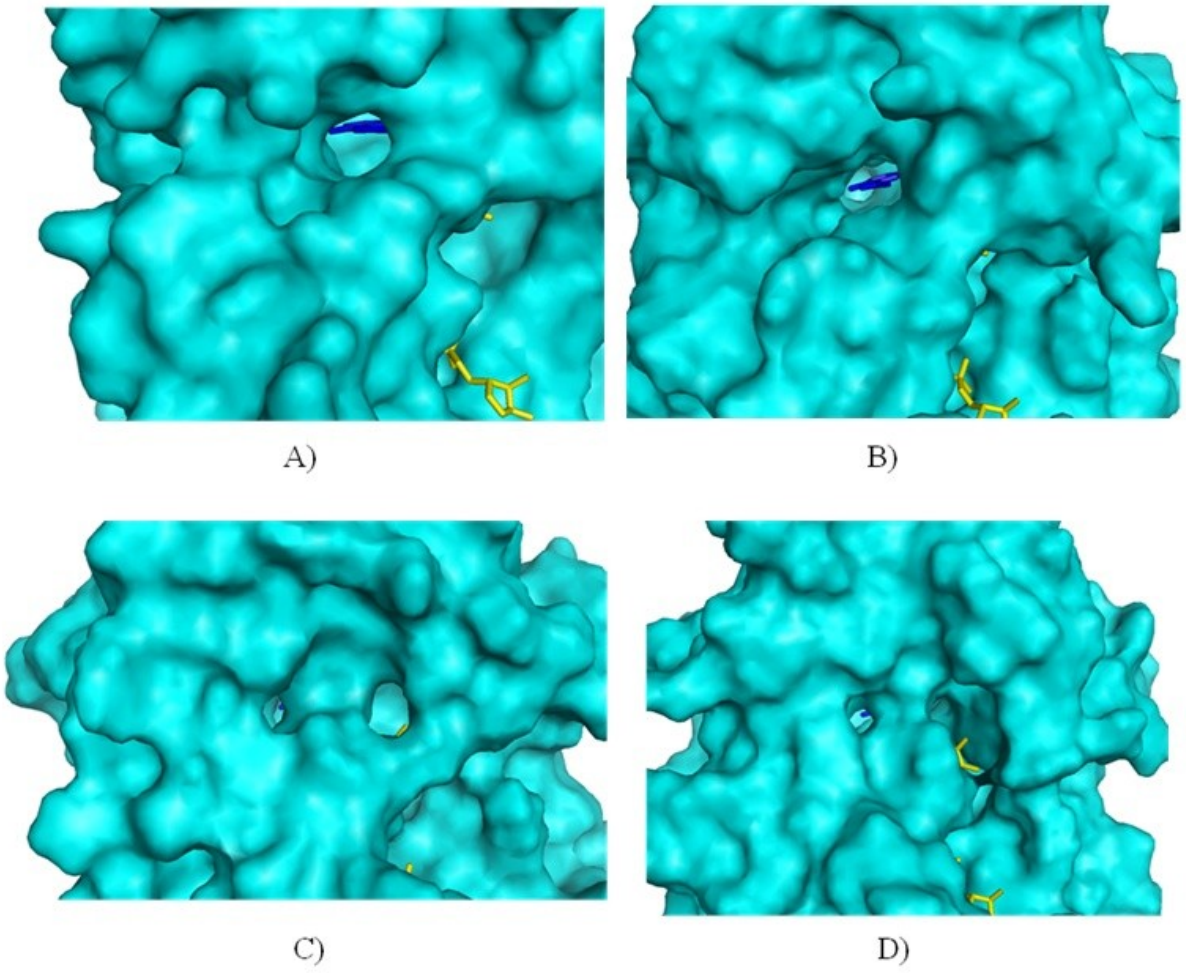


Figure 3.32: Potential substrate entrances to salicylate hydroxylase active sites. Openings for substrate are shown for *P. putida* salicylate hydroxylase (A), *A. niger* 577 (B), *A. niger* 767 (C), and, *A. niger* 534 (D) in the presence of salicylate (blue) and FAD (yellow). Enzyme models were produced using I-TASSER and visualized using PyMOL.

3.5.2.2 Potential catalytic serine residues

Serine (S49) has been proposed as a catalytic residue involved in the deprotonation of the salicylate hydroxyl group in *P. putida* salicylate hydroxylase[28]. This serine is located 3.11 Å away from the salicylate hydroxyl group shown in panel A. It is also conserved in *A. niger* 577 (B) located 3.48 Å away, but not in *A. niger* 767, and *A. niger* 534.

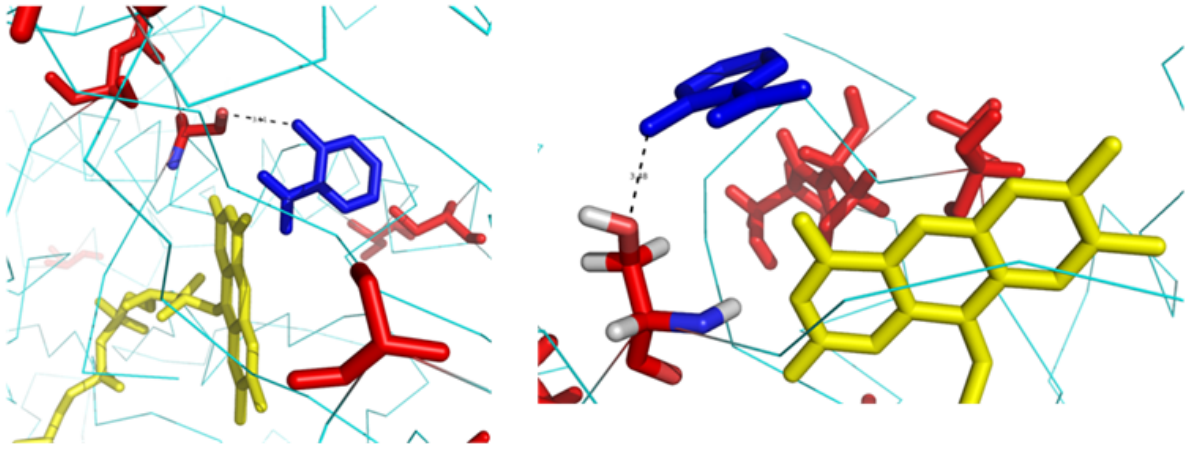


Figure 3.33: Potential catalytic serine residue in *A. niger* 577. Serine residues are indicated in red, with the distance between the nearest serine to the salicylate hydroxyl group measured in Å. FAD and salicylate are represented in yellow and blue respectively. *P. putida* and *A. niger* 577 are shown in panel A and B respectively where the *A. niger* 577 protein model was generated using I-TASSER and visualised using PyMOL. *P. putida* salicylate hydroxylase is a published structure (PDB ID: 5EVY).

3.5.3 Bioinformatic evidence suggests a potential salicylate 5-hydroxylase in *A. niger*

Fig 3.34 shows the location of 2 genes in the NRRL3 Gbrowse app located directly next to each other[23]. On the left is a putative gentisate 1,2-dioxygenase and on the right an aromatic-ring hydroxylase-like protein. NRRL3_10244 has been expressed in our lab (Alain Patrick Semana, personal communication) and shows gentisate 1,2-dioxygenase activity so NRRL3_10245 (*A. niger* 1188051) is possibly a salicylate 5-hydroxylase since genes coding for enzymes of the same metabolic pathway are often located next to each other in the genome, especially in prokaryotes[83].

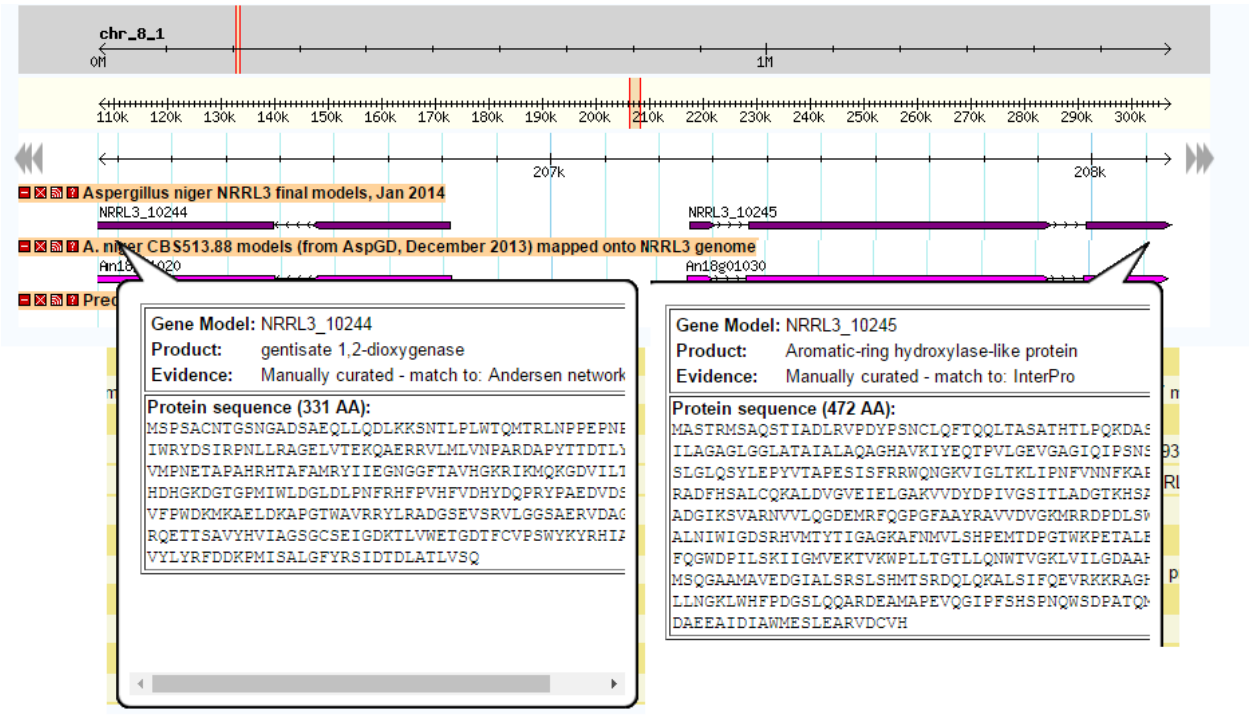


Figure 3.34: Screenshot of NRRL3_10244 and NRRL3_10245 (*A. niger* 1188051) in Gbrowse

3.5.4 Phylogenetic analysis of *A. niger* target sequences

A phylogenetic tree was constructed using alignments of all the target putative salicylate and phenol hydroxylase sequences used in this study (Figure 3.35). Additionally, a few known salicylate and phenol hydroxylase sequences from the literature were added, including P*seudomonas putida* salicylate hydroxylase[80], *Epichloë festucae* salicylate hydroxylase[33], *Pseudomonas alcaligenes* 3-hydroxybenzoate 6-hydroxylase[84], and *Trichosporon cutaneum* phenol hydroxylase[85]. The phylogenetic tree shows two main clusters, one with the annotated salicylate hydroxylase sequences (top half) and the second with those annotated as phenol hydroxylases (bottom half). However, with the addition of the sequences from the literature the tree reveals that the *A. niger* sequences belong in several different branches potentially relating to the differences in substrate specificity and other enzymatic properties. Three main branches are observed. One consists of sequences that resemble most closely salicylate hydroxylase or 3-hydroxybenzoate 6-hydroxylase. A second branch comprises two sequences (*A. niger* 999 and 146) that are distinct from either salicylate hydroxylase or phenol hydroxylase. A third branch consisting of sequences most similar to *T. cutaneum* phenol hydroxylase. Greater analysis of

these main branches and smaller branches within the main branches will be covered in the *Discussion*.

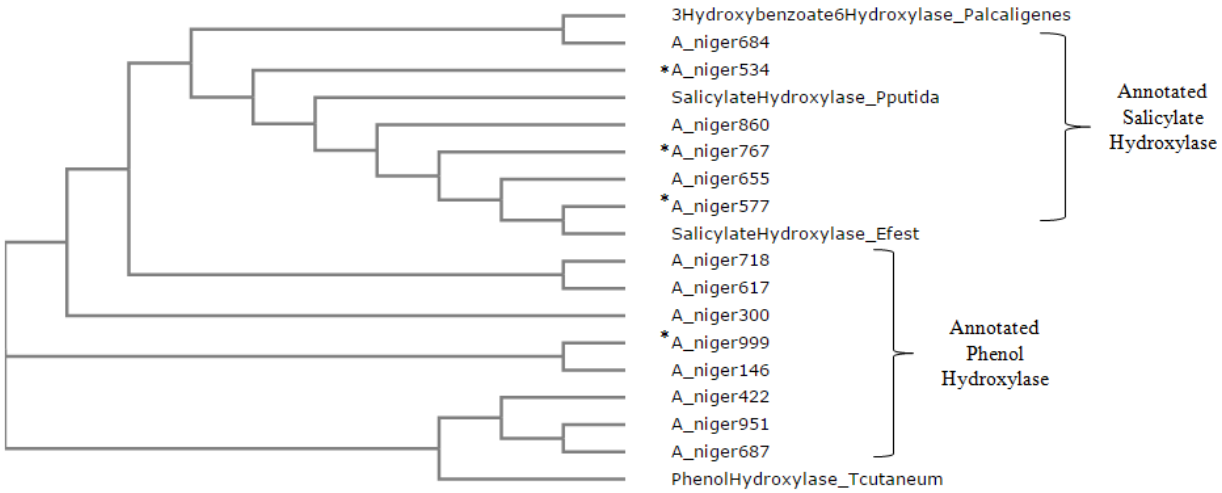


Figure 3.35: Phylogenetic tree of *A. niger* target and literature hydroxylase sequences. Sequences with an asterix (*) indicate the sequences of enzymes that were successfully expressed and characterized. The phylogenetic tree was constructed using the ClustalOmega sequence alignment tool[72].

4. DISCUSSION

Aromatic ring hydroxylases and other enzymes that prepare aromatic compounds for ring cleavage have been quite well characterized in bacteria, but much less work has been carried out in fungi. With advances in DNA sequencing and increased fungal genome availability, many fungal genes have been annotated as coding for enzymes that may be involved in aromatic degradation. However, most of these annotations are based on weak sequence conservation with bacterial counterparts. One of the main goals of this project was to express predicted aromatic ring hydroxylases from the model fungus, *Aspergillus niger*, characterize them, and compare them with their bacterial counterparts. The experimental evidence thus obtained would allow a better understanding of fungal aromatic ring activating enzymes and the genes encoding them.

In addition to *A. niger* being fully sequenced and having a genome annotated with a large number of supposed aromatic ring-hydroxylase sequences, there are other reasons why it was chosen for this study. *A. niger* is one of the most well-studied fungi, and it has been shown to grow on various aromatic compounds, including salicylate and benzoate, which can each be used as single carbon energy sources for growth. *A. niger* also has many industrial applications.

During the course of this project five *A. niger* enzymes were successfully expressed, purified, and characterized. These included three putative salicylate hydroxylases (*A. niger* 767, 577, and 534), a putative phenol hydroxylase (*A. niger* 999), and a 2,3-dihydroxybenzoate decarboxylase (*A. niger* 716). The 2,3-dihydroxybenzoate decarboxylase had been previously isolated from *A. niger* but was cloned and expressed in *E. coli* for this project.

Aromatic hydroxylase target sequences in the *A. niger* genome were chosen based on their annotations and Blasted against the Swissprot/ Uniprot database to search for similarity against already characterized enzymes. Fourteen sequences were chosen (Table 3.1): most have a query cover >80 % and a sequence identity between 30-38 % with other published aromatic ring-hydroxylating enzyme sequences. This may seem like a low sequence conservation for proteins that are predicted to carry out similar function, however, since these sequences come from different organisms, including prokaryotes, it may not be too surprising. In addition, substrate specificity may differ, contributing to the low sequence similarity.

Selected genes were amplified by PCR using *A. niger* cDNA and cloned into the pLATE11 vector using the LIC cloning method. Of the 15 (14 aromatic-ring hydroxylases and 1 2,3-dihydroxbenzoate decarboxylase) gene targets only 11 were successfully cloned. *A. niger* 687 was not amplified by PCR, and *A. niger* 422 was not amplified in sufficient quantity for cloning. This may have been due to low cDNA levels of these target genes in the *A. niger* template DNA. The cDNA used was extracted from *A. niger* grown on alfalfa and barely, plant material that contains lignin, and aromatic secondary metabolites that could up-regulate expression of aromatic degrading enzymes. *A. niger* 146 and 951 were cloned but DNA sequencing revealed they contained introns (Fig 3.15 and 3.16) suggesting that they were probably amplified by contaminating genomic DNA (gDNA).

Of these eleven successfully-cloned genes, only five were expressed in soluble form in *E. coli* BL21(DE3), and were active. These include: the putative salicylate hydroxylases, *A. niger* 577, 767, and 534 (Fig 3.4); the putative phenol hydroxylase *A. niger* 999 (Fig 3.17); and *A. niger* 2,3-dihydroxybenzoate decarboxylase (*A. niger* 716) (Fig 3.25). In the case of *A. niger* 999, Rosetta (DE3) was used instead of *E. coli* BL21 (DE3) for expression since greater expression levels were detected in Rosetta (Fig 3.17). Rosetta (DE3) is an *E. coli* variant which contains an additional chloramphenicol-resistant plasmid which provides tRNA genes for codons which are rarely used in *E. coli*[86]. For the other target genes, *A. niger* 655 and 617 were expressed but insoluble, while no expression was observed for *A. niger* 860, 300, and 718. Additionally, *A. niger* 684 was soluble and inactive. Using Rosetta (DE3) and Shuffle *E. coli* strains instead of BL21 (DE3) for these targets did not improve the expression.

Once the putative hydroxylases were expressed, enzyme activities in the presence of salicylate/phenol and NAD(P)H were measured in the crude extracts using the assays described in the *Materials and Methods*. For all four enzymes, activity was observed in crude extracts, but increased (150-500 %) in the presence of FAD (data not shown). Therefore, FAD was supplemented in at least 2-fold excess (200 μM - 800 μM FAD) in all 4 enzyme crude extracts to ensure that all the enzymes had FAD cofactor bound prior to purification.

The 5 protein purifications were carried out with various degrees of success. In general the putative salicylate hydroxylases, *A. niger* 767, 577, and 534, were quite well purified, with between 3 and 4 fold increases in purity (Fig 3.5, 3.6, and 3.7). On the other hand, the putative

phenol hydroxylase, *A. niger* 999, and the 2,3-dihydroxybenzoate decarboxylase, *A. niger* 716, showed minimal increases in specific activity during purification, although the polyacrylamide gels summarizing the purification steps showed that some contaminants were removed (Fig 3.18 and 3.26).

By UV-VIS spectroscopy the three putative salicylate hydroxylases (*A. niger* 577, 767, and 534) were confirmed to contain flavin based on the two characteristic flavin peaks at 375 and 450 nm. In all three cases there was little change in the absorbance of the bound and free flavin (Fig 3.8). This bound flavin was identified as FAD by TLC (Fig 3.9). *A. niger* 767, 577, and 534 had FAD/protein ratios of 1.05, 1.04, and 1.10 respectively indicative of a single bound FAD. Most reported salicylate hydroxylases including P*seudomonas putida*[87] salicylate hydroxylase, and the only other characterized fungal salicylate hydroxylase, isolated from *Trichosporon cutaneum*[88], were also reported to contain one single FAD/polypeptide.

In addition to FAD, aromatic-ring hydroxylases require NADH or NADPH as electron donors to reduce the FAD prior to reaction with oxygen. Many aromatic ring hydroxylases show preference for either NADH or NADPH. Like most other salicylate hydroxylases[87,88] reported in the literature, the putative salicylate hydroxylases expressed in this project appear to have a preference for NADH over NADPH (Table 3.10).

Previously-purified salicylate hydroxylases have been reported to be monomers in some organisms and dimers in others. For example, the *P. putida* PpG7 (bacterial) salicylate hydroxylase is a monomer[11], while the *T. cutaneum* (eukaryotic) salicylate hydroxylase has been reported to be a dimeric protein consisting of two non-identical subunits of 39.6 kDa and 38.2 kDa, giving it an apparent molecular weight of 77.8 kDa[88]. The oligomeric states of the three putative *A. niger* salicylate hydroxylases were investigated. *A. niger* 577 and 534 were found to be dimers, while *A.niger* 767 behaved like a monomer. The dimeric enzymes are likely homodimers since only one gene was expressed to produce each protein, although post-translational modification has not been ruled out.

Substrate specificities of the three putative salicylate hydroxylases were examined by measuring NADH consumption rates, coupling of oxygen consumption to product formation, and product detection (Table 3.11, 3.12, and 3.13). On the basis of these studies *A. niger* 577 is

the only enzyme of the three that appears to be a "true" salicylate hydroxylase. It is very active with salicylate as a substrate, and oxygen consumption is tightly coupled to product formation with only 4.7 % diverted to H_2O_2 formation (Table 3.12). Additional tightly-coupled (<15% H_2O_2 formed) aromatic substrates were 4- or 5- methylsalicylate, and 2,3-, 2,5-, or 2,6-dihydroxybenzoate (Table 3.12). Hydroxylated products were detected for salicylate, 4-methylsalicylate, and 5-methylsalicylate, yielding catechol, 4-methylcatechol, and 4-methylcatchol (Table 3.13 and Fig 3.11).

The substrate specificity of *A. niger* 577 resembles that of other salicylate hydroxylases reported in the literature. *T. cutaneum* salicylate hydroxylase is also active on salicylate, 2,3-dihydroxybenzoate, and 2,6-dihydroxybenzoate, resulting in formation of catechol for salicylate and pyrogallol for the two dihydroxybenzoates[17,88]. However, the coupling of oxygen activation for these substrates by the *T. cutaneum* enzyme were not reported. For the bacterial salicylate hydroxylase isolated by White-Stevens et al[29] virtually no oxygen uncoupling was observed with salicylate as a substrate (as observed with *A. niger* 577). Similarly, the specific activity of *A. niger* 577 and oxygen uncoupling did not vary much with 2,3-, 2,4-,and 2,5-dihydroxybenzoates compared to the bacterial enzyme. However, the oxygen uncoupling with *A. niger* 577 was much lower than the enzyme isolated by White-Stevens et al[29] with 2,6- dihydroxybenzoate (1% H_2O_2 formation compared to 28 % H_2O_2 formation). Other substrates could not be compared since no data have been reported.

According to the data compiled on BRENDA for salicylate 1-hydroxylase (EC 1.14.13.1), salicylate hydroxylase specific activity ranges from as low as 0.58 μmol/min/mg to as high as 37 μmol/min/mg[89]. Most of these data are for bacterial salicylate hydroxylases but for the two other known fungal salicylate hydroxylases, from *T. cutaneum* and *Epichloë festucae* the specific activities are, respectively 37 μmol/min/mg[88] and 0.71 μmol/min/mg[33]. The specific activity of *A. niger* 577 is 31.69 μmol/min/mg, comparable to that of the *T. cutaneum* enzyme.

A. niger 767 is, on the other hand, a much poorer salicylate hydroxylase than *A. niger* 577, having much lower activity with salicylate (7.68 μmol/min/mg) and with 85 % of oxygen consumed diverted to H_2O_2 (Table 3.11 and 3.12). In fact, this enzyme exhibited oxygen uncoupling between 85-100% with nearly all the substrates tested, with the exception of 4-methoxysalicylate (Table 3.12). With 4-methoxysalicylate, oxygen uncoupling was significantly

lower (only 43.7% diversion to H_2O_2) and the specific activity (11.4 μmol/min/mg) was higher than that observed with salicylate (7.68 μmol/min/mg). Furthermore, the binding of 4-methoxysalicylate to *A. niger* 767 is about 4-fold tighter than with salicylate, consistent with the assay data that suggested 4-methoxysalicylate is a better substrate for *A. niger* 767 than salicylate. Unfortunately the product formed was never identified. *A. niger* 767 showed the highest NADH consumption rate with 6-methylsalicyate, however this was accompanied by a high degree of oxygen uncoupling, with no 3-methylcatechol formation. It is interesting to note that *A. niger* has been shown to express a 6-methylsalicylate synthase, but no catabolic enzyme has yet been identified to break down 6-methylsalicylate[90].

A. niger 534 was shown to be an NADH-dependent flavin-containing hydroxylase, but the aromatic substrate has not been identified: nearly all the oxygen was diverted to H_2O_2 with all substrates tested (Table 3.12). Even with substrates where oxygen consumption was not completely uncoupled, no products were detected. With salicylate, this enzyme had an NADH consumption rate of only 0.4 μmol/min/mg and an oxygen uncoupling of 98.6 % with no catechol formation. The aromatic substrates tested act as non-substrate effectors that convert the enzyme to its catalytic form and only produce hydrogen peroxide with no hydroxylated product[29].

The similarities between the three hydroxylase sequences are quite low. Using *A. niger* 577 as the reference sequence, *A. niger* 767 and 534 display sequence identity of only 39 % (59 % positives) and 33 % (53 % positives), respectively. However, functionally important residues are mostly conserved in the FAD fingerprints, active site, substrate binding and NADH binding site regions (Fig 3.28). The salicylate hydroxylase substrate active site is proposed to consist of 2 highly conserved sequence motifs. These include the 13 residue motif, AAH(A/S)(M/L)(L/V)PH(Q/H)G(Q/A)GA, which is located near the C-terminal end, and the short residue AGV sequence located near the N-terminal end of the protein[82]. The first region also partly overlaps with FAD fingerprint 2 (Fig 3.28). These conserved regions were determined using only bacterial sequences, and in *E. festucae* and the *A. niger* target sequences the 5th and 6th residues consisting of M/L and L/V are rarely conserved, which may be a general occurrence in fungi.

In the first conserved motif, *A. niger* 577 and 767 have 11 out of 13 residues conserved (the 5th and 6th residues not conserved), while *A. niger* 534 has only 8 out of 13 residues conserved. The three residues that differ in *A. niger* 534 from the other two sequences are a serine substituted for alanine (residue # 1), a glycine for alanine or serine (residue # 4), and an alanine instead of a glycine (residue # 10). The second motif (AGV) was conserved in *A. niger* 577 while in *A. niger* 767 and 534 the sequence is AGI. Protein models were generated for all 3 enzymes using the I-TASSER[70] modelling server and visualized in PYMOL: the protein folds in these two regions are very similar (Fig 3.31).

Upon further inspection of the enzyme models, it is apparent that in *A. niger* 767 and 534 there is a much smaller opening to the active site compared with that of *A. niger* 577 and the 5EVY salicylate hydroxylase crystal structure the models were based on (Fig 3.32). This may affect substrate access to the active site and therefore be a factor in substrate specificity.

Recently, with the publication of the first salicylate hydroxylase crystal structure, from *Pseudomonas putida* (PDB ID: 5EVY), a serine residue has been proposed to be involved in salicylate hydroxylase catalysis. In the *P. putida* enzyme this is S49, which deprotonates the hydroxyl group of salicylate located only 3.11 Å away.[28] It is interesting to note that in the three *A. niger* proteins studied here, this residue is only conserved in *A. niger* 577, the "true" salicylate hydroxylase (3.48 Å away) (Fig 3.33). Therefore this residue could be important in determining optimal activity on salicylate.

However, *E. festucae* salicylate hydroxylase does not possess the catalytic serine found in *A. niger* 577 and *P. putida* salicylate hydroxylases. This enzyme still catalyzes the conversion of salicylate to catechol albeit at a much lower rate of 0.71 μmol/min/mg compared to 31.69 μmol/min/mg for *A. niger* 577. It would be interesting to know the oxygen uncoupling of the *E. festucae* salicylate hydroxylase (thus far unreported) since it is possible that the serine residue may play a role in the extent of coupling.

A number of other putative *A. niger* salicylate hydroxylase target sequences weren't successfully expressed. These sequences are included in a phylogenetic tree including model salicylate hydroxylases sequences from *E. festucae* and *P. putida,* and the expressed *A. niger* enzymes. Unfortunately the sequence for *T. cutaneum* salicylate hydroxylase is unavailable.

The phylogenetic tree (Fig 3.35) shows that *A. niger* 577 is most similar to the *E. festucae* salicylate hydroxylase, the only previously-characterized fungal salicylate hydroxylase sequence in this tree. Of the other two putative salicylate hydroxylases, *A. niger* 767 is the closest sequence to *A. niger* 577 while *A. niger* 534 is more distant: the two latter sequences seem to specify enzymes that cannot function as efficient salicylate hydroxylases. *A. niger* 684 appears to be more closely related to a 3-hydroxybenzoate 6-hydroxylase from *P. alcaligenes* than to any salicylate hydroxylases. However, crude extracts of *E. coli* cells that expressed this enzyme did not show any activity on 3-hydroxybenzoate or salicylate even with the addition of FAD (data not shown). Additionally, the insoluble protein *A. niger* 655 is the next most similar target to *A. niger* 577 and is thus potentially a salicylate hydroxylase as well. Another unexpressed enzyme, *A. niger* 860, is unlikely to be a salicylate hydroxylase since it is more similar to *A. niger* 767 than *A. niger* 577

The only successfully expressed and purified putative phenol hydroxylase in this study was *A. niger* 999. The most extensively studied fungal phenol hydroxylase is that from *Trichosporon cutaneum*, originally isolated by Neujahr et al (1973)[35]. In addition, a few other eukaryotic fungal enzymes capable of hydroxylating phenol to catechol are hydroquinone hydroxylase isolated from *Candida parapsilosis* CBS604[15] and resorcinol hydroxylase from *Gloeophyllum trabeum* (Zhuoxuan Yang, MSc Thesis, Concordia University[91].) Unlike these eukaryotic enzymes, which are comprised of a single polypeptide, most bacterial phenol hydroxylases are multicomponent[13]. The results obtained with *A. niger* 999 therefore will be compared with these three fungal enzymes.

A. niger 999 contains FAD, which was identified by TLC (Fig 3.20). The FAD absorbance peak was found at 443 nm, with a shoulder at 462 nm, comparable spectral features were reported for *T. cutaneum* phenol hydroxylase[35], hydroquinone hydroxylase[15] and resorcinol hydroxylase[91]. *A. niger* 999 was also determined to be associated with 1 FAD per polypeptide, which was also reported for *T. cutaneum* phenol hydroxylase[35].

Single-component phenol hydroxylases such as these are usually reported to be dimers: they have a thioredoxin-like fold domain occurring at the C-terminal end of the protein which is involved in dimerization[76]. *A. niger* 999 was also found to be a dimer, as size exclusion

chromatography indicated a molecular weight of 132.9 kDa (with an expected monomeric weight of 73.2 kDa.)

Different types of aromatic ring-hydroxylating enzymes seem to vary in their specificity for NADH vs. NADPH. Salicylate hydroxylase type enzymes usually prefer NADH while phenol hydroxylases such as the *T. cutaneum* enzyme usually prefer NADPH[35]. *A. niger* 999 also prefers NADPH, with a rate nearly 6-fold faster observed with NADPH compared to NADH (Table 3.17).

A. niger 999 was active on a variety of aromatic substrates with the greatest activities observed with phenol (3.54 µmol/min/mg), *p*-cresol (3.11 µmol/min/mg), resorcinol (3.38 µmol/min/mg), and hydroquinone (4.18 µmol/min/mg) (Table 3.18). Amongst these substrates, the greatest oxygen uncoupling is observed with phenol, where 32 % of the oxygen consumed was diverted to H_2O_2, while the other three substrates exhibited uncoupling of about 5 % (Table 3.19). For all 4 of these substrates, product formation was detected using the catechol 2,3-dioxygenase coupled assay. The products formed for phenol, *p*-cresol, resorcinol, and hydroquinone appeared to be catechol, 4-methylcatechol, hydroxyquinol, and hydroxyquinol respectively (Fig 3.21).

These data indicate that this enzyme prefers *p*-cresol, resorcinol, and hydroquinone to phenol, but since the rates of NADPH consumption and oxygen uncoupling are so similar for all three substrates it is difficult to distinguish further between them on the basis of the data collected.

Comparison with the phenol hydroxylase from *T. cutaneum* reveals some significant differences. *T. cutaneum* phenol hydroxylase has a specific activity of 8.3 µmol/min/mg[35] with phenol as a substrate and an oxygen uncoupling of only 10 %[81] compared to *A. niger* 999s' specific activity of 3.54 µmol/min/mg and 32 % oxygen uncoupling. The *T. cutaneum* enzyme also has a broad substrate specificity but it exhibits much higher degrees of oxygen uncoupling with *p*-cresol and resorcinol than *A. niger* 999. *T. cutaneum* phenol hydroxylase showed 40 % uncoupled oxygen with resorcinol as a substrate and 20 % with *p*-cresol, compared to about 5% oxygen diverted to H_2O_2 production for both substrates with *A. niger* 999 [81].

Compared to the resorcinol hydroxylase from *Gloeophyllum trabeum* which has a specific activity of 1.99 μmol/min/mg with phenol and 0.83 μmol/min/mg with resorcinol, *A. niger* 999 appears to be more active with specific activities of 3.54 μmol/min/mg (phenol) and 3.38 μmol/min/mg (resorcinol). The *G. trabeum* enzyme is nearly completely oxygen uncoupled with phenol and 25 % of the oxygen is diverted to hydrogen peroxide in the presence of resorcinol[91]. Thus, *A. niger* 999 acts as a superior resorcinol hydroxylase since it displays 4 fold higher activity and far less oxygen uncoupling (only 5 %).

Another characterized fungal aromatic-ring hydroxylase is hydroquinone hydroxylase isolated from *Candida parapsilosis* CBS604, which is highly active with many aromatic-ring substrates and shows virtually no oxygen uncoupling (unspecified by the authors). This enzyme had a specific activity of 2.07 μmol/min/mg with phenol and 4.26 μmol/min/mg with hydroquinone[15]. This is comparable with *A. niger* 999, for which specific activities were 3.54 μmol/min/mg with phenol (32 % uncoupled) and 4.18 μmol/min/mg with hydroquinone (5 % uncoupled). The exact proportion of oxygen uncoupled from hydroxylation is not reported for either phenol or hydroquinone for the *C. parapsilosis* enzyme and would be needed to fully compare the efficiency of hydroxylation between *A. niger* 999 and *C. parapsilosis* hydroquinone hydroxylase.

In summary, based on the data obtained, *A. niger* 999 is more effective at hydroxylating hydroquinone, *p*-cresol, and resorcinol than phenol, so having it annotated as a "phenol hydroxylase" is misleading.

In addition to *A. niger* 999, there were 7 other putative phenol hydroxylase genes from *A. niger* cloned for expression, without success in obtaining the corresponding enzymes. However, their potential enzymatic activities are suggested by analysis of sequence alignments and a phylogenetic tree (Fig 3.29 and 3.35). In this analysis, the inducible *T. cutaneum* "phenol hydroxylase" *A. niger* 999, was used as a reference sequence to identify which *A. niger* sequences are most likely to be phenol hydroxylases. Sequence comparison shows that the protein characterized in this study, *A. niger* 999, is actually one of the more distantly related sequences to *T. cutaneum* phenol hydroxylase, with a sequence identity of only 34 % (48 % positives). This low sequence identity, combined with its phylogenetic distance from the *T. cutaneum* enzyme (Fig 3.35) may account for the differing substrate specificity and catalytic

activities. However, despite the low sequence identity, proposed catalytic residues Tyr 289 (Tyr 279 in *A. niger* 999) and Asp 54 (Asp 86 in *A. niger* 999) are well conserved as are the sequences in the FAD and NADPH binding domains[76].

A. niger 422, 951, and 687 are the sequences most likely to be phenol-specific hydroxylases since they share the same phylogenetic tree branch as *T. cutaneum* phenol, hydroxylase, and possess both conserved catalytic residues highlighted in gray (Fig 3.29). *A. niger* 146 shares the same branch as *A. niger* 999 but phenylalanine is substituted for tyrosine and serine is substituted for aspartate at the two catalytic positions, which would likely affect catalytic activity and/or specificity. *A. niger* 617 and 718 are the most distantly related sequences on the phylogenetic tree and appear to be closer to salicylate hydroxylase than phenol hydroxylase (Fig 3.35). In fact, *A. niger* 617 is missing both catalytic residues (phenylalanine instead of tyrosine and threonine instead of aspartate) and *A. niger* 718 has isoleucine substituted for aspartate.

One additional target sequence was *A. niger* 300, which is distantly related to phenol hydroxylase, as well as salicylate hydroxylase (Fig 3.35). This enzyme may be a different type of aromatic-ring hydroxylating enzyme. Recent studies in *A. nidulans* have identified enzymes which are promisingly involved in the degradation pathway of salicylate and benzoate. One of these is AN10952, which on the basis of proteomic studies, is believed to be a 4-hydroxybenzoate 3-monooxygenase (an enzyme that converts 4-hydroxybenzoate to protocatechuate)[16]. The sequence alignment between this enzyme and *A. niger* 300 is 82.6 % over a sequence coverage of 98.6 %, suggesting that *A. niger* 300 may be a 4-hydroxybenzoate 3-monooxygenase. This is the first sequence identified in this study that is potentially involved in the formation of protocatechuate, rather than catechol.

Apart from the aromatic-ring hydroxylase sequences already discussed there was one other target sequence, *A. niger* 1188051 (NRRL3_10245), which is annotated as an aromatic-ring hydroxylating enzyme. Its closest blast hit is 6-hydroxynicotinate 3-monooxygenase, but it could potentially be a salicylate 5-hydroxylase which converts salicylate into gentisate because the gene coding for this enzyme is directly next to an annotated gentisate 1,2-dioxygenase (Fig 3.34). In fact, Alain Patrick Semana in our laboratory has managed to express, and confirm that the enzyme encoded by this gene has gentisate 1,2-dioxygenase activity. To date, salicylate 5-

hydroxylase is an enzyme that has only been found in bacteria[12]. Unfortunately, PCR amplification attempts of this enzyme's gene sequence have been unsuccessful under all conditions attempted.

Another enzyme involved in hydroxy-aromatic metabolism, *A. niger* 716, is an aromatic-ring decarboxylase for 2,3-dihydroxybenzoate. This enzyme had been previously isolated from *A. niger* and characterised[2,4], but in this work it was expressed in *E. coli* and purified. Like the salicylate and phenol hydroxylases already discussed it forms the same product, catechol. However, instead of hydroxylating the aromatic ring, the 2,3-dihydroxybenzoate decarboxylase removes the carboxyl group of 2,3-dihydroxybenzoate which is released as CO_2.

This enzyme has been expressed and characterized from *A. niger* by two research groups in the past, firstly by Subba Rao et al (1967)[2], and later by Kamath et al (1987)[4]. Most of the findings with the *E. coli*-produced enzyme are as those with the enzyme expressed by the native organism, although there are a few differences as well. For one, the pH optimum was previously reported to be 5.2[2], but in this study it was found to be 6.2 (Figure 3.27). Another difference was that the enzyme appears to be a trimer with a molecular weight of 122.4 kDa and an expected monomeric weight of 39.2 kDa, instead of a tetramer as previously reported. However, the shape of a protein also influences the gel filtration elution profile so additional confirmation of the oligomeric structure (for example using analytical ultracentrifugation) would be desirable. The closely related 2,3-dihydroxybenzoate decarboxylase from *Aspergillus oryzae* had an estimated molecular weight of 130 kDa (by gel filtration) for a 38.0 kDa monomer which would appear to be a trimer, but was shown to be a tetramer after correction using its $S_{20,w}$ value[38]. The 2,3-dihydroxybenzoate decarboxylase from *T. cutaneum* was shown to be a dimer[92] and shares almost identical properties with the two *Aspergillus* enzymes.

One common element of aromatic ring decarboxylases is that they share a TIM-barrel motif which may be involved in binding of cofactors such as metal ions[93]. However, even though all enzymes of this class are predicted to adopt this fold, activity has often been reported to be unaffected by metal ion addition to assays[39]. This is the case for the eukaryotic aromatic ring-decarboxylases such as the 2,3-dihydroxybenzoate decarboxylases from *A. oryzae*[38] and *T. cutaneum*[92] as well as salicylate decarboxylase from *Trichosporon moniliiforme*[41]. Previous studies have reported that *A. niger* 2,3-dihydroxybenzoate decarboxylase also appears not to

require metal ions for activity[2, 3, 4]. The data reported in this thesis confirm that *A. niger* 2,3-dihydroxybenzoate decarboxylase activity is not increased by the addition of various metal ions (Table 3.24). It was determined by ICP-MS that sub-equimolar amounts of Mg^{2+} and Ca^{2+} were present in the purified protein (Table 3.25) After extensive dialysis in chelator-containing buffer it was observed that the activity was enhanced in the presence of Ca^{2+} (but not other metal ions) by nearly 60 % (Table 3.27). This is the first evidence that the activity may be metal ion dependent. However, the results of ICP-MS measurements for Ca^{2+} on the dialysis/chelator treated samples were inconclusive since the buffer control sample showed a higher calcium content than the protein sample.

Few crystal structures of aromatic ring decarboxylases have been solved but one of them is a 2,6-dihydroxybenzoate decarboxylase from *Rhizobium* sp. Strain MTP-10005 (PDB ID: 2DVT)[39]. The protein was crystallized in the presence of 2,6-dihydroxybenzoate (substrate) and Zn^{2+} (metal co-factor). In this crystallographic study, potential metal binding residues have been identified; Glu8, His10, His164, and Asp287. These four residues seem to be well conserved amongst bacterial aromatic-ring decarboxylases, however, in fungal sequences alanine is substituted for His10 (Fig 3.30). It is possible that this would weaken metal binding to the enzyme.

Many of the reactions catalyzed by other aromatic ring decarboxylase enzymes are reversible and oxygen sensitive [41]. *A. niger* 2,3-dihydroxybenzoate decarboxylase was also shown to catalyze the reverse reaction, carboxylation of catechol to 2,3-dihydroxybenzoic acid (Table 3.23). The presence of reducing agents such as DTT had no effect on the enzyme activity in either the forward or reverse reactions, indicating that this enzyme is oxygen insensitive (Table 3.22 and 3.23). Neither of these results had been previously reported for this enzyme.

5. CONCLUSIONS

The goal of the work described in this thesis was to identify and characterize new enzymes from the fungus, *Aspergillus niger*, that are active on the aromatic-rings of simple benzene derivatives. Although bacterial versions of these enzymes have been isolated and studied in great deal, very few fungal variants have been examined. Nonetheless, large numbers of genes from fungal genomes have been annotated to be these enzymes via bioinformatic methods, with little to no experimental verification. *Aspergillus niger* is one of these fungi, and since it has been shown to grow on various aromatic ring compounds, it was chosen as the source for a representative group of target genes.

Fifteen target genes were chosen for study. However only a third of the selected enzyme targets were successfully cloned, expressed, and characterized. These five characterized enzymes, along with additional bioinformatic analysis of the other ten targets, provide experimental insight into several types of enzymes *A. niger* has to degrade aromatic compounds. Three of these five characterized enzymes were shown to be a: salicylate hydroxylase (*A. niger* 577); a hydroxylase with poor activity on salicylate (*A. niger* 767); and a hydroxylase enzyme similar to salicylate hydroxylases but with no identified aromatic substrate (*A. niger* 534). The other two were: an aromatic-ring hydroxylase annotated as phenol hydroxylase which is more efficient at hydroxylating hydroquinone, *p*-cresol, and resorcinol than phenol (*A. niger* 999); and a 2,3-dihydroxybenzoate decarboxylase (*A. niger* 716). In addition, bioinformatic studies suggest that *A. niger* may have genes encoding phenol hydroxylases, a 4-hydroxybenzoate 3-monooxygenase, and a salicylate 5-hydroxylase.

One of the significant findings is that *A. niger* has three different enzymes capable of producing catechol from 3 different substrates. Salicylate, phenol, and 2,3-dihydroxybenzoate are converted to catechol by *A. niger* 577, *A. niger* 999, and *A. niger* 716. There are several microorganisms that can form catechol from one or two of these three starting compounds, but apart from *T. cutaneum*, no other organism (bacterial or fungal) appears to have been identified that can express enzymes to convert all three of these substrates to catechol[35,88,92]. Also of note is that the salicylate hydroxylases and 2,3-dihydroxybenzoate decarboxylase from *A. niger* and *T. cutaneum* have similar activities and enzymatic characteristics. A few putative phenol hydroxylases from *A. niger* (*A. niger* 687, 951, 422) that were not successfully expressed in this

study are likely to be similar to the well characterized *T. cutaneum* phenol hydroxylase[35] based on phylogenetic analysis and sequence alignment. *A. niger* 999, which was successfully expressed, acts as a phenol hydroxylase but with lower activity and more oxygen diverted to H_2O_2 than its *T. cutaneum* counterpart. In contrast, *A. niger* 999 seems to be able to process hydroquinone, p-cresol, and resorcinol than the known *T. cutaneum* enzymes.

A. niger has the potential to be a fungus capable of degrading a wider range of benzene derivative compounds then most, and in our lab enzymes capable of cleaving catechol, protocatechuate, hydroxyquinol, and gentisate have all been confirmed. This thesis has identified enzymes capable of producing catechol and hydroxyquinol and has proposed the existence of enzymes capable of producing protocatechuate and gentisate, i.e. a 4-hydroxybenzoate 3-monooxygenase (*A. niger* 300), and a salicylate 5-hydroxylase (*A. niger* 051) respectively.

REFERENCES

1. Cain, R.B.; Bilton, R.F.; Darrah, J. A. The metabolism of aromatic acids by micro-organisms. *Biochem. J.* **108,** 797–828 (1968).

2. Subba Rao, P., Moore, K. & Towers, G. *O*-Pyrocatechuic Acid Carboxy-Lase from Aspergillus niger. *Arch. Biochem. Biophys.* **122,** 466–473 (1967).

3. Santha, R., Savithri, H. S., Rao, N. A. & Vaidyanathan, C. S. 2,3-Dihydroxybenzoic acid decarboxylase from Aspergillus niger: A Noved Decarboxylase. *Eur. J. Biochem.* **110,** 104–110 (1995).

4. Kamath, A., Dasgupta, D. & Vaidyanathan, C. . Enzyme-catalyzed Non-oxidative Decarboxylation of Aromatic Acids: I.Purification and Spectroscopic Properties of 2,3-Dihydroxybenzoic acid Decarboxylase from *Aspergillus niger*. *Biochem. Biophys. Res. Commun.* **145,** 586–595 (1987).

5. Xu, L., Resing, K., Lawson, S. L., Babbitt, P. C. & Copley, S. D. Evidence that pcpA encodes 2,6-dichlorohydroquinone dioxygenase, the ring cleavage enzyme required for pentachlorophenol degradation in *Sphingomonas chlorophenolica* strain ATCC 39723. *Biochemistry* **38,** 7659–7669 (1999).

6. Harwood, C. S. & Parales, R. E. The β-Ketoadipate Pathway and the Biology of Self-Identity. *Annu. Rev. Microbiol* **50,** 553–590 (1996).

7. Wright, J. D. Fungal related degradation compounds of benzoic acid and. *World J. Microbiol. Biotechnol.* **9,** 9–16 (1993).

8. Arora, P. K., Srivastava, A. & Singh, V. P. Bacterial degradation of chlorophenols and their derivatives. *J. Hazard. Mater.* **266,** 42–59 (2014).

9. Barriault, D., Durand, J., Maaroufi, H., Eltis, L. D. & Sylvestre, M. Degradation of polychlorinated biphenyl metabolites by naphthalene-catabolizing enzymes. *Appl. Environ. Microbiol.* **64,** 4637–4642 (1998).

10. Vaillancourt, F. H., Bolin, J. T. & Eltis, L. D. The Ins and Outs of Ring-Cleaving

Dioxygenases. *Crit. Rev. Biochem. Mol. Biol.* **41,** 241–267 (2006).

11. You, I.-S., Murray, R. I., Jollie, D. & Gunsalus, I. C. Purification and characterization of salicylate hydroxylase from *Pseudomonas putida* PpG7. *Biochem. Biophys. Res. Commun.* **169,** 1049–1054 (1990).

12. Zhou, N. Y., Al-Dulayymi, J., Baird, M. S. & Williams, P. A. Salicylate 5-hydroxylase from *Ralstonia* sp. strain U2: A monooxygenase with close relationships to and shared electron transport proteins with naphthalene dioxygenase. *J. Bacteriol.* **184,** 1547–1555 (2002).

13. Powlowski, J. & Shingler, V. In vitro analysis of polypeptide requirements of multicomponent phenol hydroxylase from *Pseudomonas* sp. strain CF600. *J. Bacteriol.* **172,** 6834–6840 (1990).

14. Howell, L., Spector, T. & Massey, V. Purification from and Properties of p-Hydroxybenzoate Hydroxylase *Pseudomonas fuorescens*. *J. Biol. Chem.* **247,** 4340–4350 (1972).

15. Eppink, M. H. M., Cammaart, E., Van Wassenaar, D., Middelhoven, W. J. & Van Berkel, W. J. H. Purification and properties of hydroquinone hydroxylase, a FAD-dependent monooxygenase involved in the catabolism of 4-hydroxybenzoate in *Candida parapsilosis* CBS604. *Eur. J. Biochem.* **267,** 6832–6840 (2000).

16. Martins, T. M. *et al.* The old 3-oxoadipate pathway revisited: New insights in the catabolism of aromatics in the saprophytic fungus *Aspergillus nidulans*. *Fungal Genet. Biol.* **74,** 32–44 (2015).

17. Anderson, J. J. & Dagley, S. Catabolism of aromatic acids in *Trichosporon cutaneum*. *J. Bacteriol.* **141,** 534–543 (1980).

18. Faber, B. W., van Gorcom, R. F. & Duine, J. A. Purification and characterization of benzoate-para-hydroxylase, a cytochrome P450 (CYP53A1), from *Aspergillus niger*. *Arch. Biochem. Biophys.* **394,** 245–254 (2001).

19. Subramanian, V. & Vaidya Anathan, C. S. Anthranilate hydroxylase from *Aspergillus*

niger: New type of NADPH-linked nonheme iron monooxygenase. *J. Bacteriol.* **160,** 651–655 (1984).

20. Jones, M. G. The first filamentous fungal genome sequences: *Aspergillus* leads the way for essential everyday resources or dusty museum specimens? *Microbiology* **153,** 1–6 (2007).

21. Tsang, A. Fungal genomics. *Brief. Funct. Genomics* **13,** 421–423 (2014).

22. Andersen, M. R. *et al.* Comparative genomics of citric-acid-producing *Aspergillus niger* ATCC 1015 versus enzyme-producing CBS 513.88. *Genome Res.* **21,** 885–897 (2011).

23. Genozymes Project Public Genomes. *Aspergillus niger* NRRL3. Available at: http://genome.fungalgenomics.ca/new_gene_model_pages/species_search_page.php?predname=Aspni_NRRL3. (Accessed: 18th April 2016)

24. Gerginova, M. *et al.* Biodegradation of phenol by Antarctic strains of *Aspergillus fumigatus*. *Zeitschrift fur Naturforsch. - Sect. C J. Biosci.* **68 C,** 384–393 (2013).

25. Stoilova, I. *et al.* Biodegradation of high amounts of phenol, catechol, 2,4-dichlorophenol and 2,6-dimethoxyphenol by Aspergillus *awamori cells*. *Enzyme Microb. Technol.* **39,** 1036–1041 (2006).

26. Schuster, E., Dunn-Coleman, N., Frisvad, J. & Van Dijck, P. On the safety of *Aspergillus niger* - A review. *Appl. Microbiol. Biotechnol.* **59,** 426–435 (2002).

27. Huijbers, M. M. E., Montersino, S., Westphal, A. H., Tischler, D. & Van Berkel, W. J. H. Flavin dependent monooxygenases. *Arch. Biochem. Biophys.* **544,** 2–17 (2014).

28. Uemura, T. *et al.* The catalytic mechanism of decarboxylative hydroxylation of salicylate hydroxylase revealed by crystal structure analysis at 2.5 Å resolution. *Biochem. Biophys. Res. Commun.* **469,** 158–163 (2016).

29. White-Stevens, R. & Kamin, H. Studies of a Flavoprotein , Salicylate Hydroxylase I Preparation, Properties, and the Uncoupling of Oxygen Reduction from Hydroxylation. *J. Biol. Chem.* **247,** 2358–2371 (1972).

30. Ballou, D. P., Entsch, B. & Cole, L. J. Dynamics involved in catalysis by single-component and two-component flavin-dependent aromatic hydroxylases. *Biochem. Biophys. Res. Commun.* **338,** 590–598 (2005).

31. Katagiri, M., Takemori, S., Suzuki, K. & Yasuda, H. Mechanism of the Salicylate Reaction. *J. Biol. Chem.* **241,** 5675–5678 (1966).

32. Haribabu, B. Microbial degradation of substituted benzoic acids. *J. Indian Inst. Sci* **65,** 69–107 (1984).

33. Ambrose, K. V *et al.* Functional characterization of salicylate hydroxylase from the fungal endophyte *Epichloë festucae*. *Sci. Rep.* **5,** 10939 (2015).

34. Durham, D. O. N. R., Mcnamee, C. G. & Stewart, D. B. Dissimilation of Aromatic Compounds in *Rhodotorula graminis* : Biochemical Characterization of Pleiotropically Negative Mutants. *J. Bacteriol.* **160,** 771–777 (1984).

35. Neujahr, H. Y. & Gaal, A. Phenol hydroxylase from yeast. Purification and Properties of the Enzyme from *Trichosporon cutaneum*. *Eur. J. Biochem.* **35,** 386–400 (1973).

36. McCall, I., Betanzos, A. & Weber, D. Effects of phenol on barrier function of a human intestinal epithelial cell line correlate with altered tight junction protein localization. *Toxicol. Appl. ...* **241,** 61–70 (2009).

37. Bruce, R. M., Santodonato, J. & Neal, M. W. Summary review of the health effects associated with phenol. *Toxicol. Ind. Health* **3,** 535–68 (1987).

38. Santha, R., Rao, N., Vaidyanathan, C. Identification of the active-site peptide of 2,3-dihydroxybenzoic acid decarboxylase from *Aspergillus oryzae*. *Biochim. Biophys. Acta* **1293,** 191–200 (1996).

39. Goto, M. *et al.* Crystal structures of nonoxidative zinc-dependent 2,6-dihydroxybenzoate (γ-resorcylate) decarboxylase from *Rhizobium* sp. strain MTP-10005. *J. Biol. Chem.* **281,** 34365–34373 (2006).

40. Patskovsky, Y. *et al.* Crystal Structure of Dihydroxybenzoate Decarboxylase from

polaromonas Sp With Bound Manganese and 2-Nitroresorcinol. *To be Publ.* doi:10.2210/pdb4qro/pdb

41. Kirimura, K., Gunji, H., Wakayama, R., Hattori, T. & Ishii, Y. Enzymatic Kolbe-Schmitt reaction to form salicylic acid from phenol: Enzymatic characterization and gene identification of a novel enzyme, *Trichosporon moniliiforme* salicylic acid decarboxylase. *Biochem. Biophys. Res. Commun.* **394,** 279–284 (2010).

42. Matsui, T., Yoshida, T., Hayashi, T. & Nagasawa, T. Purification , characterization , and gene cloning of 4-hydroxybenzoate decarboxylase of *Enterobacter cloacae* P240. *Arch Microbiol* **2910,** 21–29 (2006).

43. Wiegel, J. Purification and Characterization of an Oxygen-Sensitive , Reversible 3 , 4-Dihydroxybenzoate Decarboxylase from *Clostridium hydroxybenzoicum*. *J. Bacteriol.* **178,** 3539–3543 (1996).

44. Yoshida, M., Fukuhara, N. & Oikawa, T. Thermophilic, reversible γ-resorcylate decarboxylase from *Rhizobium* sp. strain MTP-10005: purification, molecular characterization, and expression. *J. Bacteriol.* **186,** 6855–63 (2004).

45. Yoshida, T., Hayakawa, Y., Matsui, T. & Nagasawa, T. Purification and characterization of 2,6-dihydroxybenzoate decarboxylase reversibly catalyzing nonoxidative decarboxylation. *Arch. Microbiol.* **181,** 391–397 (2004).

46. Marín, M., Plumeier, I. & Pieper, D. H. Degradation of 2,3-dihydroxybenzoate by a novel meta-cleavage pathway. *J. Bacteriol.* **194,** 3851–3860 (2012).

47. Griffiths, A. J., Gelbart, W. M., Miller, J. H. & Lewontin, R. C. *Cloning a Specific Gene*. (W. H. Freeman, 1999).

48. Sezonov, G., Joseleau-Petit, D. & D'Ari, R. *Escherichia coli* physiology in Luria-Bertani broth. *J. Bacteriol.* **189,** 8746–8749 (2007).

49. Mesojednik, S. & Legiša, M. Posttranslational modification of 6-phosphofructo-1-kinase in *Aspergillus niger*. *Appl. Environ. Microbiol.* **71,** 1425–1432 (2005).

50. Aslanidis, C. *et al.* Thermo Scientific aLICator Ligation Independent Cloning and Expression System. 1–27 (2013). doi:10.1016/j.cell.2013.08.021

51. Korbie, D. J. & Mattick, J. S. Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat. Protoc.* **3,** 13–15 (2008).

52. Roche. Product Purification Kit. 1–18 (2010).

53. Ausubel, F. M. *et al. Current Protocols in Molecular Biology*. *John Wiley & Sons Inc* **1,** (2003).

54. Bio Basic Inc. EZ-10 Spin Column Handbook. 20 (2011).

55. Vincze, T., Posfai, J. & Roberts, R. J. NEBcutter: A program to cleave DNA with restriction enzymes. *Nucleic Acids Res.* **31,** 3688–3691 (2003).

56. New England Biolabs. BL21(DE3) Competent E.coli.

57. EMD Millipore. Rosetta^{TM}(DE3) Competent Cells - Novagen | 70954. Available at: https://www.emdmillipore.com/CA/en/product/Rosetta^{TM}(DE3)-Competent-Cells---Novagen,EMD_BIO-70954#overview. (Accessed: 17th February 2017)

58. Laemmli, U. K. (1970): Cleavage of Structural Proteins during Assembly of Head of Bacteriophage T4. *Nature* **227,** (1970).

59. Brown, R. E., Jarvis, K. L. & Hyland, K. J. Protein measurement using bicinchoninic acid: elimination of interfering substances. *Anal. Biochem.* **180,** 136–139 (1989).

60. Yagi, K. & Ohishi, N. Separating Determination of Flavins Thin Layer Chromatography1. *J. Vitaminol. (Kyoto).* **17,** 49–51 (1971).

61. Shingler, V., Powlowski, J. & Marklund, U. Nucleotide sequence and functional analysis of the complete phenol/3,4-dimethylphenol catabolic pathway of *Pseudomonas* sp. strain CF600. *J. Bacteriol.* **174,** 711–724 (1992).

62. Jouanneau, Y., Micoud, J. & Meyer, C. Purification and characterization of a three-component salicylate 1-hydroxylase from *Sphingomonas* sp. strain CHY-1. *Appl. Environ.*

Microbiol. **73,** 7515–7521 (2007).

63. Sharma, O. P., Bhat, T. K. & Singh, B. Thin-layer chromatography of gallic acid, methyl gallate, pyrogallol, phloroglucinol, catechol, resorcinol, hydroquinone, catechin, epicatechin, cinnamic acid, *p*-coumaric acid, ferulic acid and tannic acid. *J. Chromatogr. A* **822,** 167–171 (1998).

64. Joseph C. Touchstone. *Practice of Thin Layer Chromatography*. (John Wiley & Sons, 1983).

65. *Operations Manual Setup, Installation & Maintenance: Oxygraph System*. **44,** (Hansatech Instruments Ltd, 2006).

66. Bio-Rad laboratories. *Gel Filtration Standard (#151-1901)*.

67. Atkin, C. L., Thelander, L. & Reichard, P. Iron Radical in Ribonucleotide Reductase. *J. Biol. Chem.* **248,** 7464–7472 (1973).

68. Thermo Scientific. Slide-A-Lyzer ^{TM} Dialysis Cassettes. (2014).

69. Cadieux, E. Characterisation of phenol hydroxylase and its auxiliary proteins, DmpM and DmpK, from Pseudomonas sp. Strain CF600. (Concordia University, 2002).

70. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9,** 40 (2008).

71. *PyMol Molecular Graphics System*. (Schrödinger, LLC, 2006).

72. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7,** 539 (2011).

73. Li, W. *et al.* The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res.* **43,** W580-4 (2015).

74. Kälin, M. *et al.* Phenol hydroxylase from Trichosporon cutaneum: gene cloning, sequence analysis, and functional expression in Escherichia coli. *J. Bacteriol.* **174,** 7112–7120 (1992).

75. Kleiger, G. & Eisenberg, D. GXXXG and GXXXA motifs stabilize FAD and NAD(P)-binding rossmann folds through C-H...O hydrogen bonds and van der Waals interactions. *J. Mol. Biol.* **323,** 69–76 (2002).

76. Enroth, C., Neujahr, H., Schneider, G. & Lindqvist, Y. The crystal structure of phenol hydroxylase in complex with FAD and phenol provides evidence for a concerted conformational change in the enzyme and its cofactor during catalysis. *Structure* **6,** 605–17 (1998).

77. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., B. A. Protein Identification and Analysis Tools on the ExPASy Server. *Proteomics Protoc. Handb.* 571–607 (2005). doi:10.1385/1592598900

78. Bessette, P. H., Aslund, F., Beckwith, J. & Georgiou, G. Efficient folding of proteins with multiple disulfide bonds in the Escherichia coli cytoplasm. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 13703–13708 (1999).

79. Marquez, L. a & Dunford, H. B. Transient and steady-state kinetics of the oxidation of scopoletin by horseradish peroxidase compounds I, II and III in the presence of NADH. *Eur. J. Biochem.* **233,** 364–71 (1995).

80. You, I. S., Ghosal, D. & Gunsalus, I. C. Nucleotide sequence analysis of the *Pseudomonas putida* PpG7 salicylate hydroxylase gene (nahG) and its 3'-flanking region. *Biochemistry* **30,** 1635–41 (1991).

81. Neujahr, H. Y. & Kjellen, K. G. Phenol Hydroxylase from Yeast. Reaction with Phenol Derivatives. *J. Biol. Chem.* **253,** 8835–8841 (1978).

82. Bosch, R., Moore, E. R. B., García-Valdés, E. & Pieper, D. H. NahW, a novel, inducible salicylate hydroxylase involved in mineralization of naphthalene by *Pseudomonas stutzeri* AN10. *J. Bacteriol.* **181,** 2315–2322 (1999).

83. Lee, J. M. & Sonnhammer, E. L. L. Genomic gene clustering analysis of pathways in eukaryotes. *Genome Res.* **13,** 875–882 (2003).

84. Gao, X., Tan, C. L., Yeo, C. C. & Poh, C. L. Molecular and biochemical characterization

of the xlnD-encoded 3-hydroxybenzoate 6-hydroxylase involved in the degradation of 2,5-xylenol via the gentisate pathway in *Pseudomonas alcaligenes* NCIMB 9867. *J. Bacteriol.* **187,** 7696–7702 (2005).

85. Kalin, M. *et al.* Phenol Hydroxylase from *Trichosporon cutaneum*: Gene Cloning, Sequence Analysis, and Functional Expression in *Eschenichia coli*. *J. Bacteriol.* **174,** (7112-7120) (1992).

86. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Front. Microbiol.* **5,** 1–17 (2014).

87. Yamamoto, S., Katagiri, M., Maeno, H. & Hayaishi, O. Salicylate Hydroxylase, a Monooxygenase Requiring Flavin Adenine Dinucleotide. *J. Biol. Chem.* **240,** 3408–3413 (1965).

88. Sze, I. S. Y. & Dagley, S. Properties of salicylate hydroxylase and hydroxyquinol 1,2-dioxygenase purified from *Trichosporon cutaneum*. *J. Bacteriol.* **159,** 353–359 (1984).

89. BRENDA - Information on EC 1.14.13.1 - salicylate 1-monooxygenase. Available at: http://www.brenda-enzymes.org/enzyme.php?ecno=1.14.13.1#SPECIFIC ACTIVITY [μmol/min/mg].

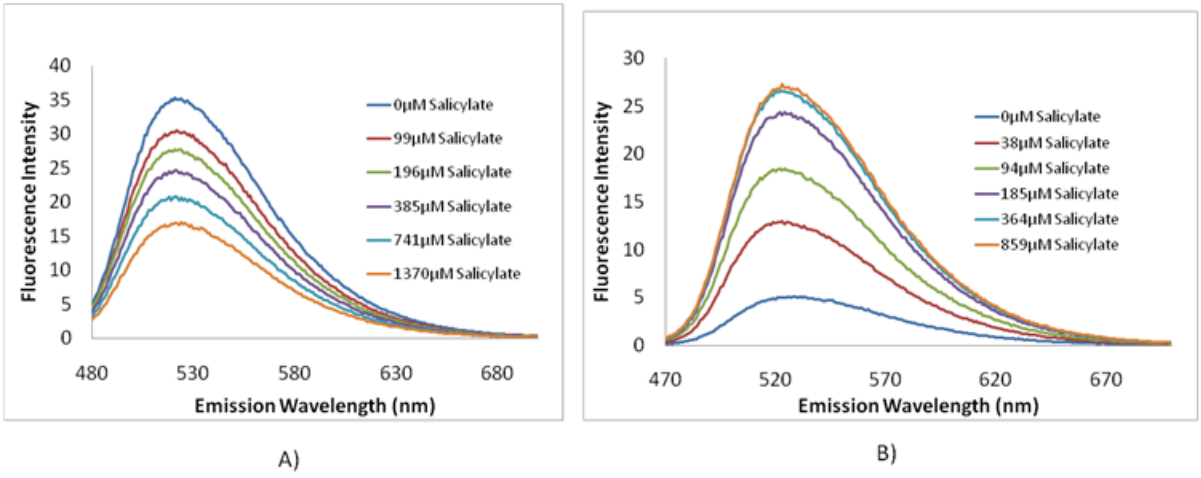
90. Holm, D. K. *et al.* Molecular and chemical characterization of the biosynthesis of the 6-MSA-derived meroterpenoid yanuthone D in *Aspergillus niger*. *Chem. Biol.* **21,** 519–529 (2014).

91. Yang, Z. Overexpression, purification and characterization of diverse oxygenases from fungi. (Concordia University, 2010).

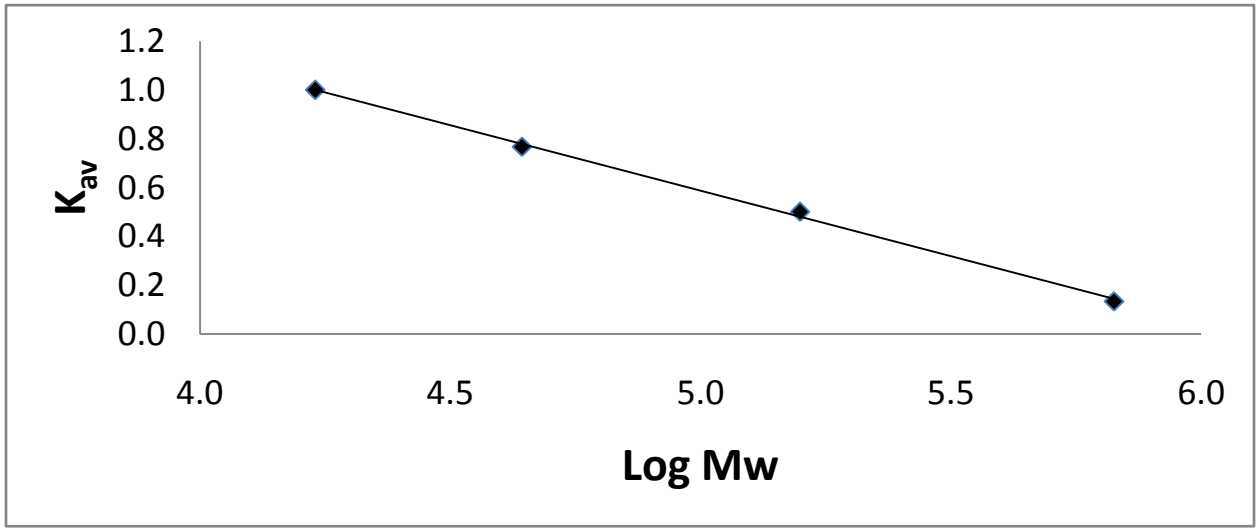
92. Anderson, J. J. & Dagley, S. Catabolism of tryptophan, anthranilate, and 2,3-dihydroxybenzoate in *Trichosporon cutaneum*. *J. Bacteriol.* **146,** 291–297 (1981).

93. Goldman, A. D., Beatty, J. T. & Landweber, L. F. The TIM Barrel Architecture Facilitated the Early Evolution of Protein-Mediated Metabolism. *J. Mol. Evol.* **82,** 17–26 (2016).

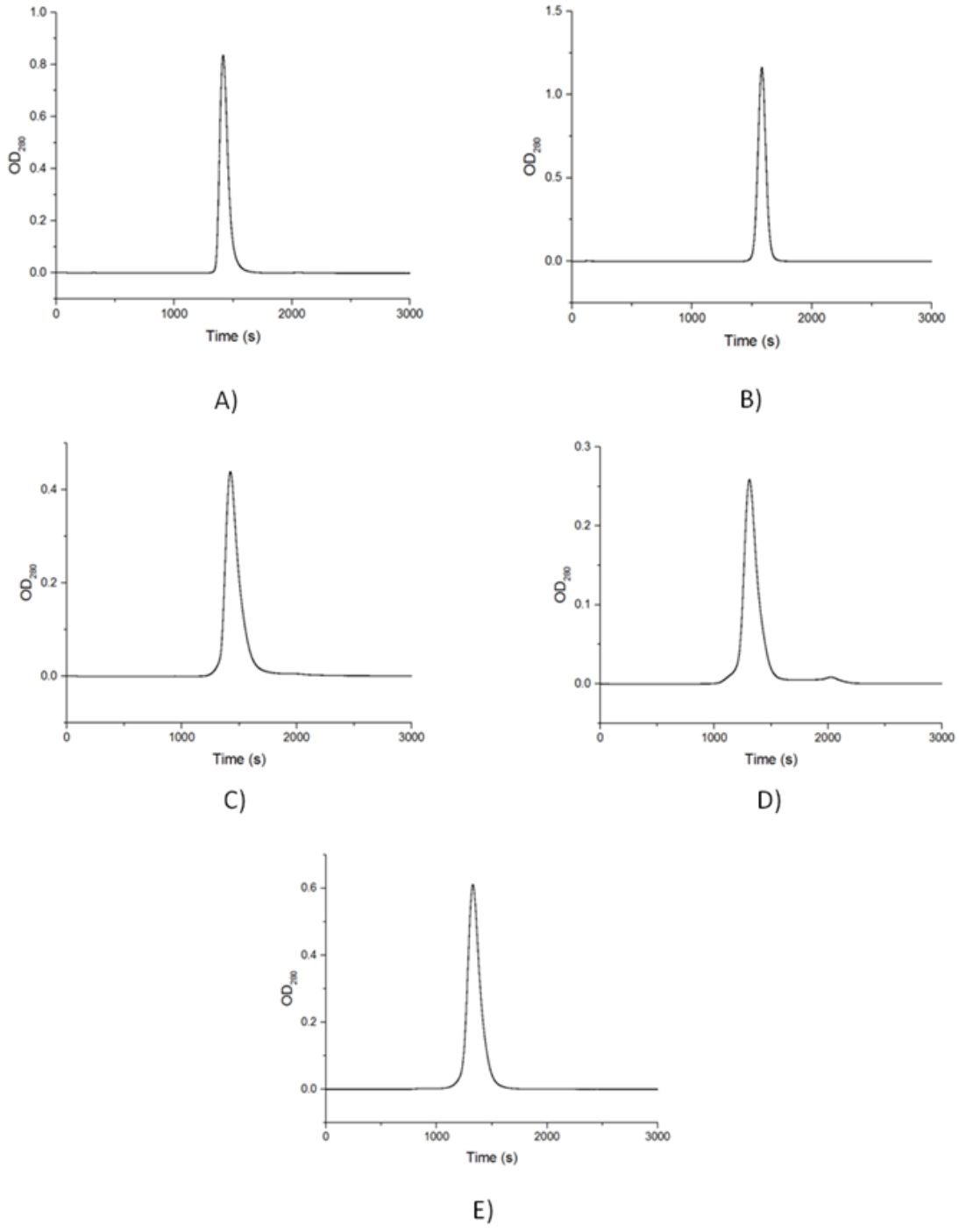
APPENDIX



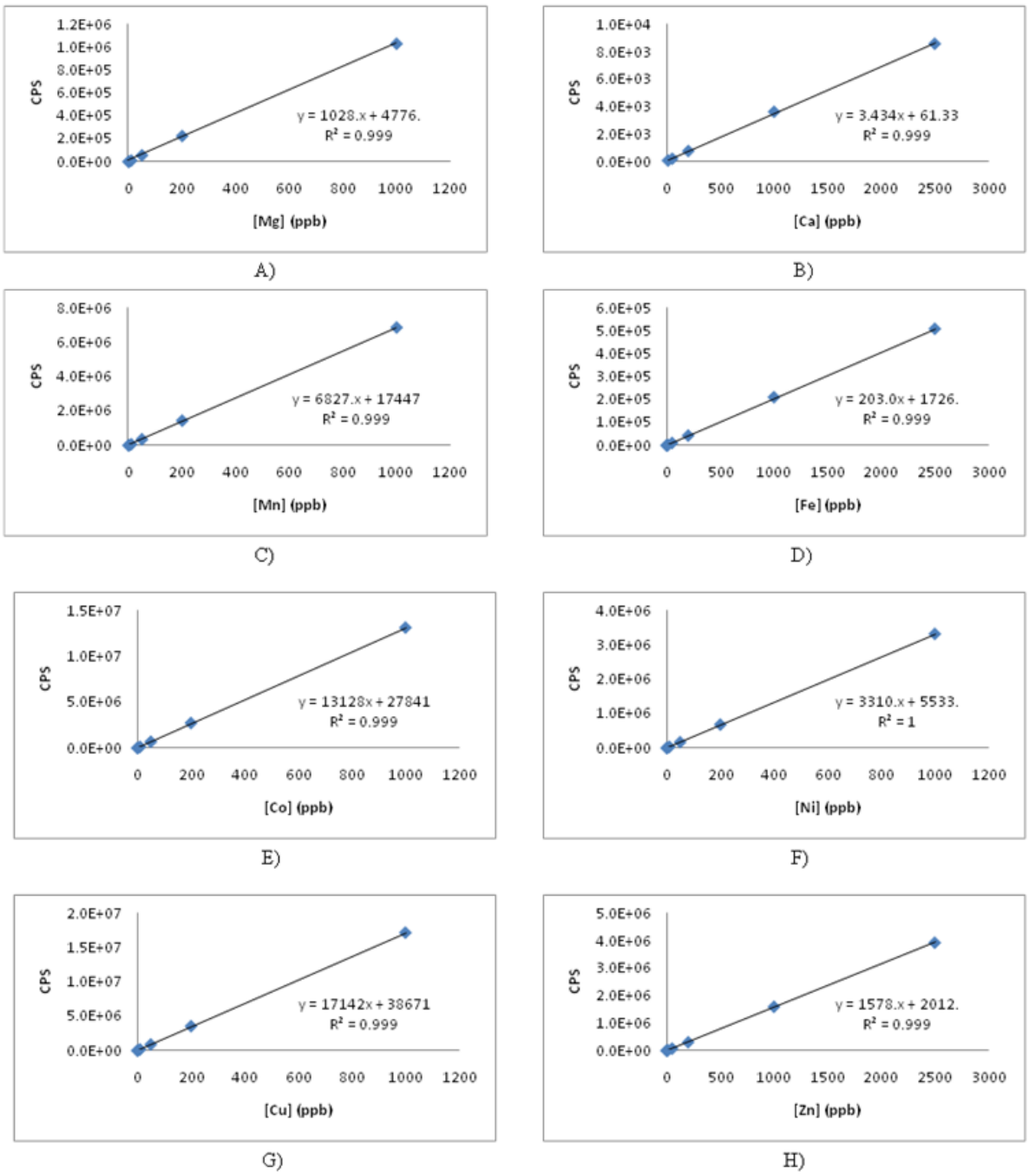
Appendix 1: Effects of salicylate binding to *A. niger* 767 (A) and *A. niger* 577 (B) on the respective protein fluorescence spectra where λ_{ex}= 450 nm. Successive additions of salicylate were made to 50 mM Tris-HCl pH 7.6 buffer containing enzyme (5 µM) at ambient temperature. For clarity, not all salicylate concentrations used are shown.



Appendix 2: Protein standards calibration proteins run on a Superdex 200 10/300 GL. Protein standards used are thyroglobulin (670 kDa), γ-globulin (158kDa), ovalbumin (44kDa) and, myoglobin (17 kDa). The equation of the fitted line was -0.5368x + 3.2710 with an R^2 value of 0.9984.



Appendix 3: Gel filtration profiles of *A. niger* purified enzymes; 767 (A); 577 (B); 534 (C); 999 (D); and 716 (E). Samples were loaded on a Superdex 200 10/300 GL column connected to a BioRad BioLogic DuoFlow HPLC. Proteins were eluted using 50 mM Tris-HCl, 100 mM NaCl pH 7.6 buffer, with a flow rate of 0.65 mL/min at room temperature.



Appendix 4: Metal calibration standards for ICP-MS. Sample above are: Mg (A); Ca (B); Mn (C); Fe (D); Co (E); Ni (F); Cu (G); and Zn (H). Standards were run by Alain Tessier on the LC-ICP-MS (Agilent 7500ce)