Confirmation of Genes Involved in the Degradation of Protocatechuate in *Aspergillus niger* through Characterization of their Encoded Enzymes

Nicholas Chow

A Thesis

In The Department Of

Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements

for the Degree of

Master of Science (Biochemistry)

at Concordia University

Montreal, Québec, Canada

February 2024

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Nicholas Chow

Entitled: Confirmation of genes involved in the degradation of protocatechuate in *Aspergillus niger* through characterization of their encoded enzymes

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 Examining Committee:

_____Chair

Dr. Paul Joyce

_____Examiner

Dr. Paul Joyce

_____Examiner

Dr. David Kwan

_____Supervisor

Dr. Adrian Tsang

Approved by _____

Dr. Louis Cuccia Graduate Program Director

_____2024 _____

Dr. Pascale Sicotte Dean of Faculty

ABSTRACT

Confirmation of genes involved in the degradation of protocatechuate in *Aspergillus niger* through characterization of their encoded enzymes

Nicholas Chow

Lignocellulosic biomass is an abundant and renewable source of aromatic compounds and carbohydrates. Microbial lignin catabolism leads to the formation of seven central aromatic intermediates that have potential markets valuated in the billions. Protocatechuate is one of the most common central intermediates and is the precursor to chemical building blocks, including *cis,cis*-muconic acid, used for production of bioplastics, cosmetics, food preservatives and antioxidants. Understanding the full extent of the microbial aromatic catabolic pathways is necessary to design the most efficient strains to produce valuable chemicals from protocatechuate. Aromatic catabolic pathways have been mapped out in bacteria, however, the equivalent pathways in fungi have not been as well characterized. In this study, the candidate genes *NRRL3_01405*, *NRRL3_02586* and *NRRL3_01409* encoding protocatechuate-3,4-dioxygenase, 3-carboxy-*cis, cis*-muconate cyclase and β-carboxymuconolactone hydrolase/decarboxylase in the *Aspergillus niger* protocatechuate catabolic pathway were expressed in *Escherichia coli*. The target proteins were purified by column chromatography and characterized by enzyme activity assays.

The products of each enzyme were characterized by time-of-flight mass spectrometry, nuclear magnetic resonance spectroscopy, and ultraviolet-visible spectroscopy. NRRL3_00837 was determined to not participate in the *in vitro* catabolism of protocatechuate to β-ketoadipate despite having been previously reported to participate in the *in vivo* protocatechuate catabolism. Mapping out the central aromatic catabolic pathways in fungi may lead to the refinement of strains engineered for lignin valorization and bioremediation.

ACKNOWLEDGEMENTS

I would like to first thank the late Dr. Justin Powlowski and Dr. Adrian Tsang as my thesis supervisors. Dr. Powlowski for his guidance in every step in the research portion of my thesis. Dr. Adrian Tsang for his aid in the completion of the writing and defense. I would also like to thank my committee members, Dr. Paul Joyce and Dr. David Kwan for their comments and feedback.

I am grateful for the collaboration of Michael Sgro for his work on the transcriptomics and mutagenic portion of this project, Patrick Semana for the preparation of recombinant pLATE11-NRRL3_01405 and Farnaz Olyaei for the preparation of recombinant pLATE11-NRRL3_02586. I would also like to thank Dr. My Pham for her help on the isolation of metabolites and Marie-Claude Moissan for the preparation of *Aspergillus niger* crude extracts and cDNA.

I would like to mention the support of my lab mates and the consistent presence of Dr. Lena Sahlman and her management of the lab.

Finally, I would like to thank my family for their support throughout the years.

Table of contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACN	Acetonitrile
BCA	Bicinchoninic acid
BlastP	Protein basic local alignment search tool
BSA	Bovine serum albumin
CRISPR	Clustered regularly interspaced short palindromic repeats
DEAE	Diethylethanolamine
DMSO-d6	Deuterated dimethyl sulfoxide
FA	Formaldehyde
^{1}H NMR	Hydrogen nuclear magnetic resonance
IPTG	Isopropyl β-D-1 thiogalactopyranoside
LB	Luria broth
LC-MS	Liquid chromatography - tandem mass spectrometry
LIC	Ligation independent cloning
rpm	Rotations per minute
RT	Room temperature
SDS-PAGE	Sodium dodecyl sulfate – polyacrylamide gel electrophoresis
TCA	Trichloroacetic acid
ToF-MS ES+	Time of flight – Mass spectrometry electrospray positive mode
TPM	Transcripts per million
UV-Vis	Ultraviolet-Visible

1. Introduction

1.1 Aromatic carbon as part of the carbon cycle

The depletion of fossil fuel reserves and the increasing demand for sustainable energy have brought attention to alternative and renewable energy sources. Lignocellulosic biomass, also known as plant dry matter, is the most abundant sustainable raw material that is currently treated to be used as biofuel. Lignocellulosic biomass is divided into the carbohydrates cellulose and hemicellulose and the complex polymer lignin. Cellulose and hemicellulose have been converted to bioethanol and biobutanol using biological treatments (Ren et al., 2016; Liu et al., 2021; Radhika et al., 2021; Zhu and Pan, 2010) and to furan-based compounds using thermochemical treatments (Zhang et al., 2017). Lignin is a major by-product in the pulp and paper and biofuel sectors, however only 5% of lignin is used as a low-value combustible fuel due to its recalcitrant properties (Cao et al., 2018).

Lignin is composed of aromatic monolignols; coniferyl alcohol, *p*-coumaryl alcohol and sinapyl alcohol, making it an ideal target as a renewable source of aromatic compounds (Freudenberg, 1965, Humphreys et al., 2002). Lignin may be valorized for industrial applications as a natural and renewable source of building blocks for bioplastics, biofuels, cosmetics, food preservatives and antioxidants (Arif, 2015; Kaur and Chakraborty., 2013; McKenna et al., 2013; Tsuge et al., 2016; Wu et al., 2017). Lignin valorization and depolymerization have thus become a growing field of interest (Ragauskas et al., 2014; Xu et al., 2014). The largest obstacle to lignin valorization is the recalcitrant heterogenous polymeric structure of lignin. Studies using chemical, biological and thermal treatments of lignin have converted lignin into biofuels and chemicals as shown in Figure 1.1 (Ragauskas et al., 2014; Wu et al., 2017). For example, *cis, cis*-muconic acid and its derivatives, the degradative intermediates of protocatechuate, have an estimated market value of $22 billions (Figure 1). However, the depolymerization of lignin releases a heterogenous mixture of aromatic compounds, whereas the production of chemicals demands high purity and requires costly methods for isolating aromatic components (Schutyser et al., 2018; Liu et al., 2022). To tackle the issue of heterogeneity, lignin valorization using a biological approach may facilitate lignin conversion in commercial biorefineries due to the specificity of enzymatic reactions. Microbial degradation of lignin has been shown to be sustainable with biocatalytic refineries functioning under mild environmental conditions using

bacterial and fungal strains capable of degrading lignin (Xu et al., 2018; Ni and Tokuda, 2013; Xu et al., 2014). Lignin valorization using a biological approach still faces many challenges as the resulting chemical is of low concentration. The construction of a robust strain to degrade all components of lignin has been difficult. A more conservative approach is the construction of several bacterial and fungal strains to convert lignin into different value-added compounds (Singhvi and Kim, 2021). To capitalize on the development of synthetic biology and metabolic engineering of strains for lignin degradation, the catabolic pathways of bacterial and fungal species capable of breaking down lignin and its aromatic derivatives must first thoroughly be investigated.

One of the entryways into the synthesis of highly valued chemicals is from the conversion of protocatechuate-derived *cis, cis*-muconic acid to terephthalic acid, then to plastics, lubricants, nylon and resins (Figure 1.1). Intradiol cleavage of protocatechuate forms 3-carboxy-*cis, cis*-muconic acid, however protocatechuate can be decarboxylated to catechol, followed by aromatic ring fission to form *cis, cis*-muconic acid (Figure 1.2). Engineered metabolic pathways in *Escherichia coli*, *Corynebacterium glutamicum*, *Pseudomonas putida and Saccharomyces cerevisiae* have been designed to maximize *cis, cis*-muconic acid yield from lignin-derived aromatic compounds (Brückner et al., 2018; Choi et al., 2020; Lee et al., 2018)

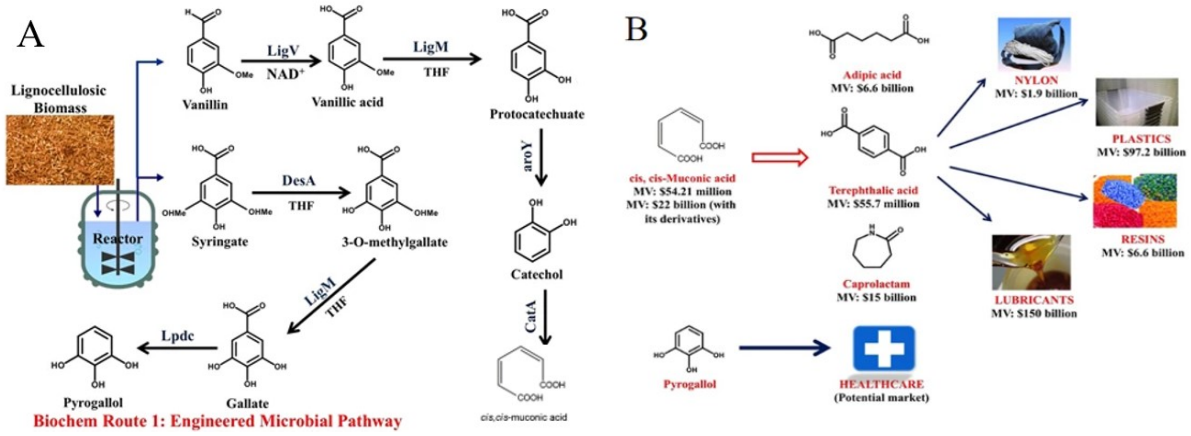


Figure 1.1 Lignin valorization by converting lignin to protocatechuate and its degradative intermediates. (A) Genetically engineered *Escherichia coli* and tobacco plant lignin catabolic route to convert lignin-derived vanillin to protocatechuate then to *cis*, *cis*-muconic acid and

syringate to pyrogallol; (B) Market values of protocatechuate-derived chemicals from *cis*, *cis*-muconic acid and pyrogallol. Modified from Wu et al., (2017).

Microbial degradation of lignin releases aromatic compounds that are toxic to most microorganisms even at low concentrations (Adeboye et al., 2014; Friedman et al., 2003, Guiraud et al., 1995; Lima et al, 2018). Microorganisms that have survived and developed enzymatic pathways to metabolize aromatic compounds as carbon sources have flourished in environments containing aromatic compounds. Lignin degradation has been observed to begin with basidiomycete white-rot and brown-rot fungi and the resulting aromatic metabolites are funnelled into central ring-cleavage pathways (Mäkaelä et al, 2015). Many studies in the past decades have focused on the bacterial and fungal enzymes that participate in aromatic catabolism (Buchan et al., 2000; Cain et al., 1968; Crawford et al., 1979; Dennis et al., 1973; Martins et al., 2015; Thatcher and Cain, 1975). The bacterial pathways have been extensively identified and characterized in comparison to their fungal counterparts (Doten et al., 1973; Elsemore and Ornstone, 1994; Gerischer et al., 1998; Hara et al., 2003; Harwood et al., 1994; Nogales et al., 2005; Ornston, 1966; Yamanashi et al., 2015).

Aromatic compounds are stable structures due to resonance energy within the aromatic ring. Bacteria and fungi have evolved aromatic ring-fission enzymes under aerobic conditions, making them an essential component of the carbon cycle. Aromatic ring cleavage is centralized through seven intermediates; catechol, protocatechuic acid, hydroxyquinol, hydroquinone, gentisic acid, gallic acid and pyrogallol. In this thesis, we focus on the metabolism of protocatechuate in the filamentous fungi *Aspergillus niger*.

Protocatechuate has been reported as an intermediate in the catabolism of the secondary metabolites of lignin: benzoic acid, vanillin, *p*-coumaric acid, ferulic acid and their derivatives (Lubbers et al, 2019) and is the precursor to *cis, cis*-muconic acid, a highly valued chemical (Figure 1.1). Ring cleavage of protocatechuate has been reported between the 2,3-, 3,4- and 4,5-carbon in bacteria and only between the 3,4-carbon in fungi (Figure 1.2). The most common pathway observed in both fungi and bacteria is the intradiol 3,4- carbon cleavage pathway.

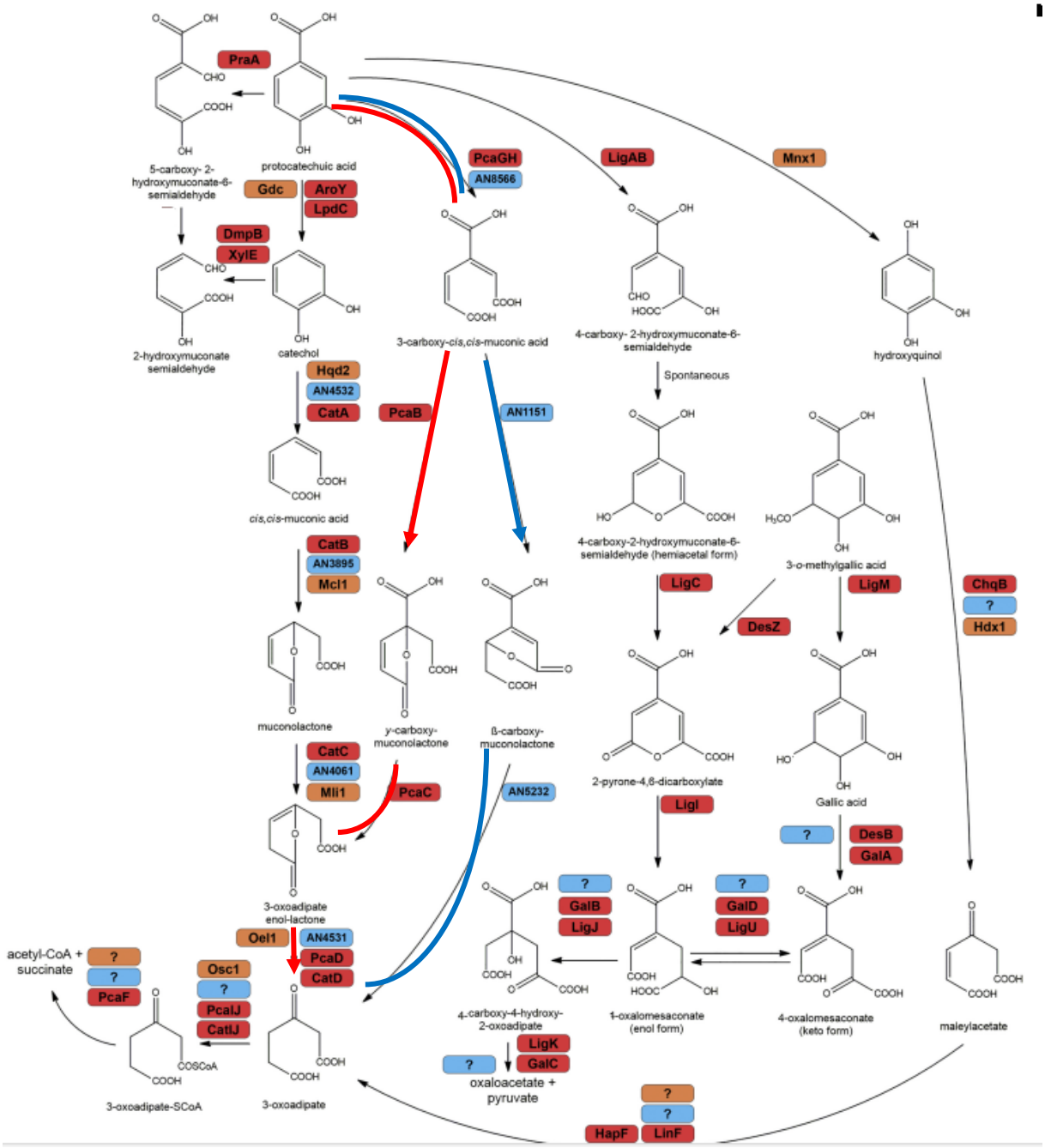


Figure 1.2 Catabolism of aromatic central intermediates in bacteria, fungi and yeast. The coloured arrows indicate the intradiol cleavage pathway of protocatechuate. Highlighted in blue is the fungal protocatechuate catabolic pathway, highlighted in red are the pathways observed in bacteria and highlighted in orange are the equivalent pathways in yeast. Modified from Lubbers et al., (2019).

1.2 Ring cleavage of protocatechuate in bacteria

Intradiol cleavage of protocatechuic acid is the most common ring cleavage step in bacteria. In *Pseudomonas putida*, this step is catalyzed by protocatechuate 3,4-dioxygenase to form 3-carboxy-*cis, cis*-muconic acid (Harwood et al., 1994). After isomerization to γ-carboxymuconolactone by 3-carboxy-*cis, cis*-muconate cycloisomerase, the γ-carboxymuconolactone is then decarboxylated by 4-carboxymuconolactone decarboxylase to form 3-oxoadipate enol-lactone (Buchan et al., 2000). The 3-oxoadipate enol-lactone is then converted to β-ketoadipate (also referred to as 3-oxoadipate) by 3-oxoadipate enol-lactonase. The genes encoding the protocatechuate catabolic enzymes in *P. putida* have been sequentially named as *pcaGH*, *pcaB*, *pcaC* and *pcaD* (Harwood et al, 1994; Buchan et al., 2000). Protocatechuate catabolic genes in bacteria have been observed to be clustered within operons. In *Acinetobacter*, the *pcaIJFBDKCHG* operon encodes six enzymes that catabolize protocatechuate to citric acid cycle intermediates and is directly upstream of the *qui* and *pob* genes that encode enzymes that catabolize quinate and *p*-hydroxybenzoate (Elsemore and Ornston, 1994). The regulatory protein PcaU has been identified to induce expression of the *pca* genes in the presence of protocatechuate (Gerischer et al., 1998)

The 2,3- and 4,5- extradiol ring cleavage pathways have been reported in *Paenibacillus* sp. (Crawford et al., 1979; Kasai et al., 2009) and *Sphingobium* sp. (Dennis et al., 1973; Masai et al., 2000; Hara et al., 2003; Nogales et al., 2005), respectively.

1.3 Ring cleavage of protocatechuate in fungi

Although intradiol ring cleavage of protocatechuate has been observed in both bacteria and fungi, the pathways split after the formation of 3-carboxy-*cis, cis* muconic acid. The proteomic analysis of the benzoic acid pathway was reported in *Aspergillus nidulans* (Martins et al., 2015). As protocatechuate was one of the intermediate metabolites in benzoic acid catabolism in *A. nidulans*, therefore the protocatechuate proteomic data were uncovered in this analysis. Ring cleavage of protocatechuate to form 3-carboxy-*cis, cis* muconic acid was reported to be catalyzed by AN8566, a protocatechuate 3.4-dioxygenase. In contrast to the bacterial pathway, 3-carboxy-*cis, cis* muconic acid is converted to β-carboxymuconolactone by AN1151, a carboxy-*cis, cis*-muconate cyclase. β-carboxymuconolactone is then converted to β-ketoadipate by AN5232, a β-

carboxymuconolactone hydrolase/decarboxylase. Individual deletion of *AN8566, AN1151,* or *AN5232* each resulted in accumulation of the precursor of their respective catalyzed reactions and complete growth inhibition on benzoic acid, indicating there are no additional catabolic pathways for benzoic acid or protocatechuate in *A. nidulans* (Martins et al., 2015).

1.4 Enzymology of protocatechuate catabolism in *Aspergillus niger*

Aspergillus niger was observed to proliferate using protocatechuate, benzoate, quinate, *p*-hydroxybenzoate, *p*-methoxybenzoate, vannilate, ferrulate, caffeate, *p*-coumarate or catechol as a carbon source. The enzymology of protocatechuate catabolism was studied by characterizing the enzymes and metabolites involved using ultraviolet-visible (UV-Vis) and infrared (IR) spectroscopies (Cain et al., 1968). The metabolites in the *A. niger* protocatechuate catabolic pathway were isolated and characterized as 3-carboxy-*cis, cis* muconic acid, β-carboxymuconolactone and β-ketoadipate (Cain et al., 1968).

The metabolites from protocatechuate catabolism in *A. niger* were compared to those observed in the bacterial pathways. γ-carboxymuconolactone is observed in bacteria following the lactonization of 3-carboxy-*cis, cis* muconic acid and has an absorption maximum at 230 nm with an extinction coefficient of 4100 (Ornston, 1966). Whereas the metabolite observed by Cain et al. (1968) had an absorption maximum at 217 nm with an extinction coefficient of 10300. The infrared spectroscopy of the *Aspergillus niger* lactonizing product was observed to have bands at 1740 cm^{-1} for a β-lactone and 1600 and 685 cm^{-1} for a *cis*-double bond which are nearly identical to the natural and synthetic β-carboxymuconolactone IR spectrum (MacDonald et al. 1954). The IR bands observed in γ-carboxymuconolactone wereat 1735 cm and 1635 cm^{-1} for an αβ-unsaturated γ-lactone and a C=C—C=O conjugated bond respectively. Based on these observations, the metabolite observed from *A. niger* protocatechuate cleavage was said to be β-carboxymuconolactone (Cain et al., 1968).

The crude extract cultivated on protocatechuate was reported as inactive on γ-carboxymuconolactone (Cain et al., 1968), supporting that the protocatechuate intradiol cleavage pathway is different from the bacterial pathway. The hydroxylation and decarboxylation of β-carboxymuconolactone was observed in a single step catalyzed by β-carboxymuconolactone

hydrolase/decarboxylase forming β-ketoadipate, whereas in bacteria, 4-carboxymuconolactone decarboxylase and 3-oxoadipate enol-lactonase catalyze separate decarboxylation and hydrolysis reactions.

Protocatechuate 3,4-dioxygenase has been annotated as NRRL3_01405 using transcriptomics, mutagenic and biochemical studies (Lubbers et al., 2019; Semana and Powlowski, 2019). 3-carboxy-*cis, cis*-muconate cyclase has been reported as an octamer with identical subunits of 24 kDa or 47 kDa in the presence or absence of a reducing agent (Thatcher and Cain, 1974) and a K_m of 57 µM in pH 6.0 at 30 °C (Thatcher and Cain, 1975). β-carboxymuconolactone hydrolase/decarboxylase was reported as a 54±5 kDa enzyme with a K_M of 70 µM in pH optimum of 8.6 in *Aspergillus niger* (Thatcher and Cain, 1970), and 26 kDa in *Aspergillus nidulans* (Martins et al., 2015). The degradation of protocatechuate to β-ketoadipate in fungi and bacteria is summarized in Figure 1.3.

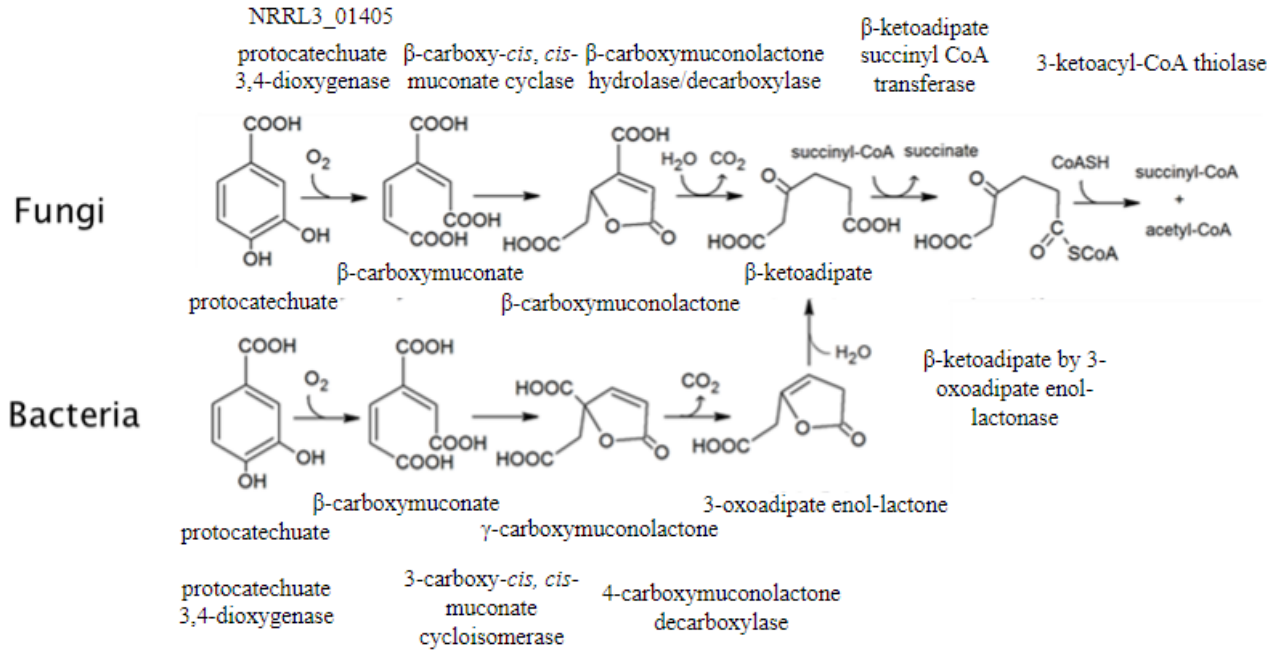


Figure 1.3 Bacterial and fungal protocatechuate intradiol cleavage pathways

1.5 *Aspergillus niger* as the model organism for protocatechuate catabolism

Aspergillus niger is a filamentous fungus that is ubiquitous in the environment (Baker, 2006). It is one of the most widely studied filamentous fungi with high-quality genome sequences publicly available, curated and annotated (The *Aspergillus niger* NRRL3 genome was sequenced and manually curated by researchers at the Centre for Structural and Functional Genomics at Concordia University). This genome resource is publicly hosted by the Joint Genome Institute, https://mycocosm.jgi.doe.gov/Aspni_NRRL3_1/Aspni_NRRL3_1.home.html). *Aspergillus niger* strains are widely known as cell factories for the production of citric acid (Cairns et al. 2018; Pel et al., 2007; Tong et al., 2019), proteases for detergents, food ingredients and additives including amylase, glucose oxidase, phospholipase, triacylglycerol lipase, xylanase among many other enzymes (Acourene and Ammouche, 2012; Archer, 1994; Cesário et al., 2021; Fasiku et al., 2023; Meng et al., 2014). *Aspergillus niger* also has potential in lignocellulose degradation as bleaching aids, and in catabolizing phenolic compounds in wastewater treatment (Duarte and Costaferreira, 1994). The CRISPR-Cas9 system has been applied to *Aspergillus niger* and has shown to be effective on this already genetically tractable organism (Song et al., 2018).

In terms of using *Aspergillus niger* as a model organism for protocatechuate catabolism, the organism was among the microbes that are capable of catabolizing lignocellulosic-derived monocyclic aromatic compounds as a sole carbon source (Cain *et al.*, 1968; Lubbers *et al*. 2019).

Strategies to develop lignin valorization requires the understanding of the catabolism of lignin-derived aromatic compounds. Bacterial pathways have been thoroughly studied. However, the studies on the fungal pathways have been relatively scarce and scattered throughout different species of fungi. Protocatechuate is one of the most common lignin-derived intermediates and is converted to *cis, cis* muconic acid, which is the precursor to products with multibillion dollar markets. Although the intradiol cleavage of protocatechuate is similar between the bacterial and fungal pathways, the paths split at the lactonization of 3-carboxy-*cis, cis*-muconic acid forming β-carboxymucolactone in fungi and γ -carboxymuconic acid in bacteria (Figure 1.3). There are still many challenges to bringing lignin valorization to industrial standards. Fungal species have been reported to degrade aromatic compounds; potentially eliminating limitations on the toxic nature of aromatic compounds on microbes in bioreactors. *Aspergillus niger* also has potential in bioremediation strategies to degrade chlorinated aromatic compounds in pesticides and industrial

waste (Sahasrabudhe and Modi, 1985, Shailubhai et al., 1983) and aromatic hydrocarbons from petroleum (Luo et al., 2019; Priyanka and Lens 2022). The information from enzyme characterization from fungal species may be the key to solve these underlying issues.

1.6 Goals of this study

The goal of this study is to characterize the enzymes encoded by the protocatechuate catabolic genes in *Aspergillus niger*. The most likely candidate genes in the *A. niger* protocatechuate catabolic pathway were expressed in *Escherichia coli* and purified by column chromatography. The purified enzymes were characterized by enzyme activity assays and SDS-PAGE and the metabolites produced by each of the purified enzymes were finally isolated and identified by Time of Flight-Mass spectrometry (ToF-MS), ^{1}H Nuclear Magnetic Resonance (^{1}H NMR) and Ultraviolet-Visible spectroscopy (UV-Vis).

2. Materials and Methods

2.1 *Aspergillus niger* and *Escherichia coli* strains

The gene sequences used in this study were based on the *Aspergillus niger* NRRL3 strain (https://gb.fungalgenomics.ca/fgb2/gbrowse/NRRL3_public). Table 1 lists the *A. niger* and *E. coli* strains used in this study.

Escherichia coli DH5α was used to propagate recombinant plasmids. *Escherichia coli* BL21 (DE3) and, Rosetta (DE3were used to induce recombinant gene expression and protein production. The λ prophage DE3 carries the T7 RNA polymerase gene under the control the *lac* UV5 promoter where IPTG (isopropyl ß-D-1-thiogalactopyranoside) is required to induce the expression of the T7 RNA polymerase. Rosetta (DE3) contains pRARE (Cam^R) that produces tRNAs with rare *E. coli* codons. Addition of chloramphenicol in the media is necessary to maintain the plasmids in Rosetta (DE3).

Table 1. *Aspergillus niger* and *Escherichia coli* strains used in this study

Species	Strain	Genotype
Aspergillus niger	NRRL3	Wild type
Escherichia coli	DH5α	F $^{-}$, Φ*80dlacZ*△*M15,* △ (*lacZYA-argF*)*U169, deoR, recA1, endA1, hsdR17* (rk^-,mk^+), *phoA, supE44,* λ^- , *thi-1, gyrA96, relA1*
Escherichia coli	BL21 (DE3)	F^- *omp*T *hsdS*$_B$ (r_B^-, m_B^-) *gal dcm* (DE3)
Escherichia coli	Rosetta (DE3)	F^- *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (DE3) pRARE (Cam^R)

2.2 Culture conditions

The solutions used for the cultures, identifications of proteins and protein purification are listed in Table 2. For plasmid propagation, single *E. coli* DH5α colonies were inoculated in 2 mL of LB media with 100 µg/mL ampicillin and incubated overnight at 37 °C, with shakeing at 225 rpm. For the screening of induced gene expression, single *E. coli* BL21(DE3) colonies were cultured in 1 mL of LB medium with 100 µg/mL of ampicillin overnight and inoculated in 10 mL of fresh LB media with ampicillin until an OD_{600} of 0.6-0.8 was reached. At that optical density, a final concentration of 400 µM of IPTG was added and the culture was incubated overnight at 17 °C, with shaking at 225 rpm. In the case of Rosetta (DE3), 30 µg/mL

chloramphenicol was included in the medium. For scaled up protein production, the procedure was identical to the above induced gene expression, but a single colony was inoculated into 10 mL of medium cultured overnight (as above) and was used as the starting culture for 1 L of medium which was processed as above.

Table 2. Media, buffers and chemicals used in this study

Solutions	Content
Luria Broth (LB)	10 g/L tryptone, 10 g/L NaCl and 5 g/L yeast extract, (with 1.5% agar for plates)
Tris-HCl buffer	100 mM Tris-HCl, pH 6 or 7.5
Selective Media	LB supplemented with 100 µg/mL of ampicillin (and 30 µg/mL chloramphenicol)
TBE Buffer (10X)	Tris 108 g/L, boric acid 55 g/L, 20 mM EDTA pH 8.0
SDS-PAGE acrylamide solution	300 g/L acrylamide, 8 g/L N-N[1]-methylenebisacrylamide
SDS-PAGE stacking buffer (4X)	60.5 g/L Tris-HCl, 4 g/L SDS, pH 6.8
SDS-PAGE resolving buffer (4X)	182 g/L Tris-HCl, 20 g/L SDS, pH 8.8
In-gel trypsin digestion, reduction	50 mM NH_4HCO_3, 10 mM DTT
In-gel trypsin digestion, alkylation	50 mM NH_4HCO_3, 50 mM iodoacetamide
In-gel trypsin digestion, washes	50 mM NH_4HCO_3. 25 mM NH_4HCO_3, 5 % ACN. 25 mM NH_4HCO_3, 50 % ACN. 100 % ACN
In-gel trypsin digestion, extraction	60 % ACN, 0.5 % formic acid
DEAE-sepharose, Sephacryl S-300, phenyl-sepharose elution buffers	50 mM Tris-HCl, 0 – 1M NaCl, pH 7.5
Sephadex S300 running buffer	50 mM Tris-HCl, 0.1M NaCl, pH 7.5
His-trap elution buffer	50 mM Tris-HCl, 0.1M NaCl, 5 – 400 mM imidazole, pH 7.5

2.3 Identifying candidate protocatechuate catabolic genes by sequence similarity

Genes from the *Aspergillus niger* NRRL3 genome were queried along with experimentally supported annotated gene sequences involved in protocatechuate catabolism in *Aspergillus nidulans, Pseudomonas putida* and *Acinetobacter baylyi* on the Uni/SwissProt database using BLASTP. Sequence identities greater than 30% were selected as potential candidate genes. Polymerase chain reaction primers were designed based on these sequences; the N-terminus of

NRRL3_00837 was modified to improve protein solubility. Additional plasmids were provided by Patrick Semana (pLATE11-*NRRL3_01405*) and Farnaz Olyaei (pLATE11-*NRRL3_02586*).

2.4 Construction of plasmids containing candidate genes

The selected genes in the *A. niger* genome were amplified using the standard 50 µL PCR reaction with Phusion high-fidelity DNA polymerase (New England Biolabs), as per the manufacturer's instructions. Table 3 lists the oligonucleotide primers (synthesized by Integrated DNA Technologies) used to amplify the open reading framess predicted in the NRRL3 genome. The cDNA was prepared from *A. niger* grown on 2% alfalfa and barley by Marie-Claude Moisan of the Centre for Structural and Functional Genomics. PCR products of the expected size were purified by agarose gel electrophoresis followed by gel extraction with the Roche DNA purification kit. The purified PCR products were was then annealed to the bacterial expression vector pLATE11 or pLATE52 (aLICator, Thermo Scientific), as described by the manufacturer.

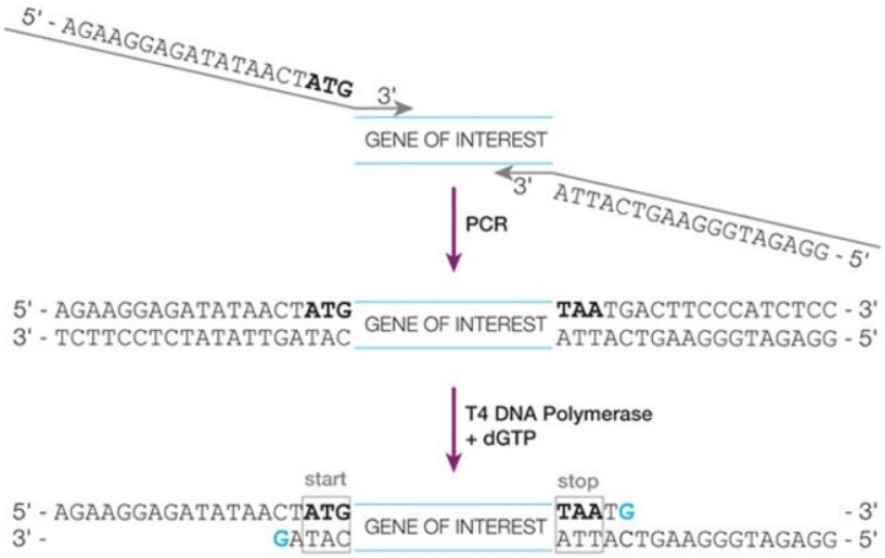
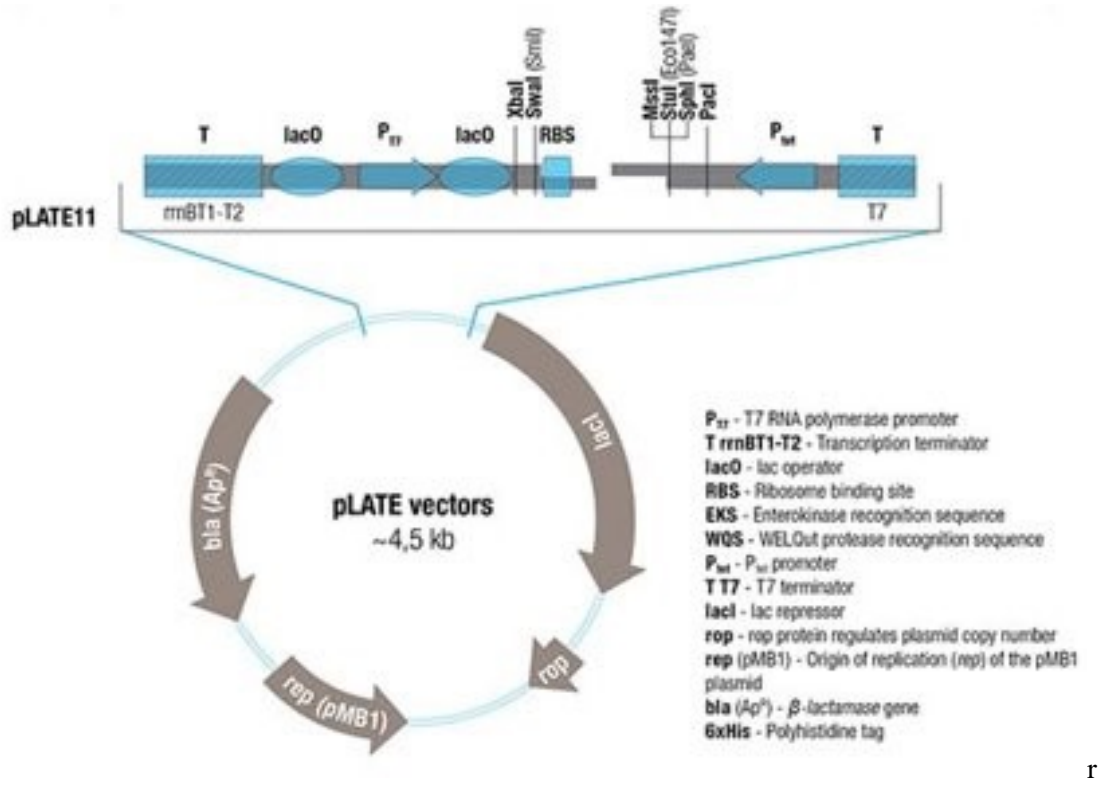


Figure 2.1 Sticky end formation to generate overhangs on vector and gene inserts using the aLICator cloning kit (Adapted from aLICator, ThermoFisher)



r

Figure 2.2 pLATE expression vectors controlled by bacteriophage T7 in the aLICator expression kit (Adapted from aLICator, ThermoFisher)

Table 3. Primers used in this study

Primer	Sequence (5'-3')	Description
NRRL3_01409 Fw	**AGAAGGAGATATAACT**ATGGCGACCCCCGG	pLATE11 - forward
NRRL3_01409 Rv	**GGAGATGGGAAGTCA**CTAGTGCAGGTACTCATTCACCA	pLATE - reverse
NRRL3_01886 Fw	**GGTTGGGAATTGCAA**ATGTCGCGGCCTGTTCCTG	pLATE52 - forward
NRRL3_01886 Rv	**GGAGATGGGAAGTCA**TTACAGCATAGGCTTCAGATCAGC	pLATE - reverse
NRRL3_01526 Fw	**GGTTGGGAATTGCAA**ATGGCCTCTCCAATCCC	pLATE52 - forward

NRRL3_01526 Rv	**GGAGATGGGAAGTCA**TTACTCCCGGA CGAACATGCC	pLATE - reverse
NRRL3_00837 ΔM1-K23 R25S Fw	**GGTTGGGAATTGCAA**ATGAGTCTCCC CTACGCC	pLATE52 - forward
NRRL3_00837 Rv	**GGAGATGGGAAGTCA**TCACTTCCTCT CAAGATCAGCC	pLATE - reverse

2.4.1 Propagation of recombinant plasmids in DH5α

Competent *E. coli* DH5α cells (50 µL) were incubated with recombinant plasmids (5 ng) on ice for 30 mins and treated by heat shock at 42 °C for 30 seconds. The treated sample was then plated onto LB medium with 100 µg/mL ampicillin and incubated overnight at 37 °C for selection. Plasmids were recovered from single colonies grown overnight at 37 ºC with shaking in 1.5 mL of LB and 100 µg/mL ampicillin overnight following the manufacturer's instructions on the EZ-10 spin 10 column kit (BioBasic).

2.4.2 Growth and cell lysis of *Escherichia coli* recombinant cells for protein production

For transformation of expression cassettes, 2 µg of the recombinant plasmid was added to 50 µL of competent *E. coli* cells and left on ice for 30 mins, treated by heat shock at 42°C for 30 seconds, then returned to ice for 15 mins. The cells were then plated onto LB medium with 100 µg/mL ampicillin (+ 30 µg/mL chloramphenicol for Rosetta) and incubated overnight at 37 °C.

A single colony was then transferred into 1 mL of LB supplemented with 100 µg/mL ampicillin (+ 30 µg/mL chloramphenicol for Rosetta and Lemo21) as a starter culture and grown overnight at 37 °C with shaking at 225 rpm. The overnight culture was then used to inoculate 10 mL of LB supplemented with 100 µg/mL ampicillin (+30 µg/mL chloramphenicol for Rosetta and Lemo21), which was grown at 37 °C with shaking at 225 rpm until an OD_{600} of 0.6-0.8 was reached. In the case of scaled-up protein production, 10 mL of starter culture and 1 L final volumes were prepared. Protein production was induced with a final concentration of 400 µM IPTG and the culture was incubated with shaking at 225 rpm overnight at 17°C. The cells were collected by centrifugation at 17 700 x*g* for 30 mins, then washed with 50 mM Tris-Cl pH 7.5,

followed by another round of centrifugation at 17 700 x*g*. The cell pellet was then suspended in 50 mM Tris-Cl pH 7.5 in a 2 mL buffer :1 g of wet cell ratio. The cells were lysed by sonication in 10 s intervals with 30 s pauses at 40% power for 6 minutes using the ultrasonic homogenizer (Model 300 V/T, by Biologics, Inc.). The lysed cells were then centrifuged at 17 700 g for 30 mins and the supernatant was retained as the crude extract and stored at -80°C.

2.5 Purification of recombinant proteins by column chromatography

The crude extracts were purified by column purification and salt precipitation. Crude extracts containing recombinant proteins with His-tags were loaded onto 5 mL HisTrap columns (Cytiva) equilibrated with 50 mM Tris-Cl pH 7.5, 0.1 M NaCl and 5 mM imidazole. Unbound proteins were removed by washing with 10 column volumes of 50 mM Tris-Cl pH 7.5, 0.1 M NaCl and 5 mM imidazole. The bound proteins were eluted stepwise using 5 mL solutions of 50 mM Tris-Cl pH 7.5, 0.1 M NaCl containing imidazole concentrations of 20 mM, 200 mM and 500 mM. Fractions of 1 mL were collected. Fraction purity was examined by SDS-PAGE, the fractions containing the recombinant proteins based on expected molecular size with lowest amount of contaminating proteins, were combined and concentrated using an Amicon concentrator with a 10 kDa cutoff membrane, then buffer exchanged with two 5-fold volumes of 50 mM Tris-Cl pH 7.5 containing 0.1 M NaCl. The column-purified protein was then splint into 1 mL aliquots and stored at –80 °C

Crude extract with recombinant proteins without a His-tag were passed (3 mL/min) through a Fast Flow DEAE-Sepharose column (2.6 cm x 27.5 cm) equilibrated with 50 mM Tris-Cl pH 7.5 collecting 12 mL fractions; the column was washed with 50 mM Tris-Cl pH 7.5 and 16 12 mL fractions were collected, followed by a 74-fraction gradient elution with 0 to 1 M NaCl in 50 mM Tris-Cl pH 7.5. Fractions with the highest enzyme activity or highest purity based on SDS-PAGE were pooled and concentrated using an Amicon concentrator with a 10 kDa cutoff membrane and stored at -80 °C.

Partially purified proteins from the DEAE-Sepharose column were loaded onto a Sephacryl S-300 column (2.6 cm x 76 cm), equilibrated with 50 mM Tris-Cl pH 7.5, and eluted with the same buffer at 3 mL/min over 90 15 mL fractions. Fractions with the highest enzyme activity or purity based on SDS-PAGE were pooled and concentrated using an Amicon concentrator with a 10 kDa

cutoff membrane and stored at -80 °C. The concentrated DEAE pool was slowly brought to 35% ammonium sulfate saturation. Precipitated proteins were removed by centrifugation for 30 min at 17, 200 x *g*. The supernatant was adjusted to 75% ammonium sulfate saturation, and the precipitated proteins were collected by centrifugation for 30 min at 17, 200 x g. The pellet was dissolved in 6 mL of ice-cold 20% ammonium sulfate with 50 mM Tris-Cl pH 7.5. This solution was then loaded onto a phenyl-Sepharose column (2.6 cm x 11.8 cm) equilibrated with 20% ammonium sulfate in 50 mM Tris-Cl pH 7.5 (3 mL/min) collecting 10 mL fractions; the column was washed with 2 column volumes of 50 mM Tris-Cl pH 7.5 followed by a 69-fraction gradient elution from 20% to 0% ammonium sulfate in 50 mM Tris-Cl pH 7.5. Fraction purity was examined by SDS-PAGE. The fractions with highest concentration of the recombinant protein were pooled, concentrated and buffer exchanged using an Amicon concentrator with a 10 kDa cutoff membrane then stored at –80 °C.

2.6 Protein concentration estimation

2.6.1 Trichloroacetic acid (TCA)-mediated protein precipitation.

In cases where it was necessary to remove imidazole and other interfering compounds for protein quantification, 25 µL of sample intended for the protein assay was incubated in 1 mL distilled H_2O with 100 µL of 0.15% sodium deoxycholate at room temperature for 15 mins. The proteins were precipitated with 100 µL of 72% trichloroacetic acid, vortexed, and centrifuged at 16 300 x *g* for 15 mins at 4°C. The supernatant was decanted, the pellet was resuspended in 0.5% SDS and 0.1M NaOH and immediately used for protein quantification (ThermoScientific).

2.6.2 Bicinchoninic acid (BCA) assay

Protein samples were prepared according to the Pierce BCA Protein Assay Kit manufacturer's protocols with slight modifications. In a ratio of 50:1 of Reagent A to Reagent B, 1 mL of the mixed reagent was added to 50 µL of protein sample. Samples were incubated at 30 °C for 30 minutes and the absorbance was measured at 562 nm. A standard curve of absorbance values based on varying amounts of Bovine Serum Albumin (BSA) was used to determine the unknown protein concentrations.

2.7 Enzymatic activity assays and substrate preparation

In a 125 mL Erlenmeyer flask, 4.625 mg of protocatechuate was dissolved in 10 mL of 50 mM Tris-HCl pH 7.5 containing 60 µg of NRRL3_01405 and 515 µg of NRRL3_02586 to produce β-carboxymuconolactone at RT on a rotary shaker. β-ketoadipate was produced under the same conditions with the addition of 260 µg of NRRL3_01409. An aliquot was withdrawn and a UV spectrum from 220-300 nm was taken every 30 min until no further spectral changes were observed.

Table 4. Enzyme assays used in this study

Enzyme	Wavelength (nm)	Buffer	Substrate
protocatechuate 3,4 dioxygenase	290	50 mM Tris-HCl pH 7.5	100 µM protocatechuate
3-carboxy-*cis, cis*-muconate cyclase	260	50 mM Tris-HCl pH 7.5	Product of 100 µM protocatechuate and NRRL3_01405
ß-carboxymuconolactone hydrolase/decarboxylase	230	50 mM Tris-HCl pH 7.5	Product of 100 µM protocatechuate, NRRL3_01405, NRRL3_02586
β-ketoadipate:succinyl CoA transferase	305	200 mM Tris-Cl pH 8.0, 25 mM $MgCl_2$,	150 µM succinyl-CoA and product of 3.5 mM protocatechuate, NRRL3_01405, NRRL3_02586 and NRRL3_01409

2.8 UV-Vis scanning kinetics assay

In a 3 mL cuvette, the conversion of protocatechuate to β-ketoadipate was observed following the sequential addition of NRRL3_01405, protocatechuate 3,4 dioxygenase (6.35 µg); NRRL3_02586, 3-carboxy-*cis, cis* muconate cyclase (0.19 µg); and NRRL3_01409, β-

carboxymuconolactone hydrolase/decarboxylase (0.38 µg) in 100 mM Tris-Cl pH 7.5 by scanning from 220 nm to 300 nm at a speed of 360 nm/min. The rate of formation and concentration of each enzymatic product were calculated following Beer's Law, where the observed extinction coefficient is the difference in extinction coefficients between the reactant and product at the absorption maxima of the reactant. Protocatechuate and 3-carboxy-*cis,* cis-muconate have ε 3890 and 1590 M^{-1} cm^{-1} at 290 nm respectively. 3-carboxy-*cis, cis*-muconate has ε 8150 M^{-1} cm^{-1} and β-carboxymuconolactone does not absorb significantly at 260 nm. β-carboxymuconolactone has ε 8500 M^{-1} cm^{-1} and β-ketoadipate does not absorb at 230 nm.

2.9 Organic solvent extraction of metabolites in the protocatechuate catabolic pathway.

Following the accumulation of about 30 µmoles of the enzyme product in 10 mL of Tris-Cl pH 7.5 buffer, 2 g of NaCl was added and dissolved, followed by acidification with 12 drops of 6 M HCl. The solution, 5 mL, was then transferred to 20-mL closed vials. An equal volume of ethyl acetate was added and vortexed over 10 minutes. Once the phases separated, the organic layer was transferred into a closed 50-mL vial. This extraction process was repeated twice and the organic lazer was then frozen at -80°C for 10 min. The extract was then thawed at room temperature for 10 min and the remaining aqueous layer was removed from the bottom of the vial. The solvent was evaporated with a stream of air and the residue was stored at -80 °C.

2.10 ^{1}H NMR.

Metabolite samples extracted by organic solvent were run on a Varian VNMRS-500 MHz equipped with 5 mm AutoX DB (Dual Broadband) probe^{1}H-^{19}F/X[^{15}N-^{31}P], with z-PFG and automatic tuning for all nuclei by the ProTune accessory. The system operates with VNMRJ 3.2 software under LINUX Red Hat 5. The data were processed using the Mnova V.14 software by Mestrelab Research.

2.11 In-gel tryptic digestion of recombinant polypeptides and analysis by mass spectrometry.

Liquid chromatography-tandem Mass Spectrometry (MS) analyses were performed (by Dr. Heng Jiang of the Centre for Biological Applications of Mass Spectrometry) on a Thermo EASY nLC II LC system coupled to a Thermo LTQ Orbitrap Velos mass spectrometer equipped with a nanospray ion source.. Purified protein samples were loaded onto a 12 % SDS-PAGE. Protein

bands of 1 cm x 0.5 cm were excised and were in-gel digested using trypsin (Trypsin Gold, Promega) in 200 µL of buffers listed in Table 2 for 16 hours at 30°C. A volume of 2 µL of each sample containing around 100 ng of tryptic peptides was injected onto a 10 cm × 100 µm column packed in-house with Michrom Magic C18 (5 µm particle diameter and 300 Å pore size). Peptides were eluted using a 35-min gradient at a flow rate of 400 nL/min with mobile phase A (96.9% water, 3% ACN and 0.1% FA) and B (97% ACN, 2.9% water and 0.1% FA). A full MS spectrum (*m/z* 400-1400) was acquired in the Orbitrap at a resolution of 60000 FWHM, and the ten most abundant multiple charged ions were selected for MS/MS sequencing in linear trap with the option of dynamic exclusion. Peptide fragmentation was performed using a collision induced dissociation at normalized collision energy of 35% with activation time of 10 ms. The resulting MS spectrum were then compared to the target protein sequence and the *Escherichia coli* and human databases.

3. Results

The enzymology of protocatechuate catabolism in the filamentous fungus *Aspergillus niger* has been reported by Cain et al. (1968). A recent study assigned genes encoding aromatic catabolic enzymes in *Aspergillus nidulans* based on proteomic analysis of proteins differentially upregulated in the presence of benzoate (Martins et al., 2015). With the genes and enzymes involved in the protocatechuate catabolic pathway identified in bacteria and fungi as reference, Sgro et al. (2023) used multiple approaches to identify candidate genes for protocatechuate catabolism in *A. niger*: orthologues were identified based on sequence similarity, comparative transcriptomics, and mutational studies. In this thesis, the most likely candidate genes identified by Sgro et al. (2023) were expressed in *Escherichia coli*, with the recombinant enzymes purified and biochemically characterized using activity assays. The chemical structure of the metabolites generated by the recombinant enzymes were identified by ^{1}H NMR and mass spectrometry.

3.1 Candidate protocatechuate catabolic pathway genes

Table 5. Sequence identity for potential protocatechuate catabolic genes in *Aspergillus niger*

Enzyme	Candidate genes	Query	Identity (%)
Protocatechuate 3,4 dioxygenase	***NRRL3_01405***	*AN8566*[1]	94.8
3-carboxy-*cis,* cis-muconate cyclase	***NRRL3_02586***	*AN1151*[1]	91.4
		CMLE_NEUCR[2]	67
β-carboxymuconolactonehydrolase/decarboxylase	*NRRL3_00837*	*AN5232*[1]	70.2
	NRRL3_01409	*ELH2_ACIAD*[3]	28
β-ketoadipate:succinyl-CoA transferase	***NRRL3_01886***	*AN10495*[1]	86.2
	NRRL3_01593		65.7
	NRRL3_11640		53.7
β-ketoadipyl-CoA thiolase	***NRRL3_01526***	*AN5698*	91
	NRRL3_07786		48.7
	NRRL3_11162		49.5

1. Queries from *Aspergillus nidulans* reported by Martins et al., 2015. 2. Query from *Neurospora crassa* reported by Mazur et al., 1994. 3. Query from *Acinetobacter baylyi* reported by Doten et al., 1987

Potential candidate genes in the protocatechuate metabolic pathway in *Aspergillus niger* had been determined using BlastP, querying the NRRL3 genome of *A. niger* and experimentally assigned monocyclic aromatic catabolic genes from *Aspergillus nidulans*, *Neurospora crassa*, and *Acinetobacter baylyi*. Comparative transcriptomics of *A. niger* identified *NRRL3_02586* as the only candidate to encode 3-carboxy-*cis, cis*-muconate cyclase; *NRRL3_00837, NRRL3_01409 and NRRL3_08340* as candidates to encode β-carboxymuconolactone hydrolase/decarboxylase; *NRRL3_01886, NRRL3_01526* as the candidates to encode β-ketoadipate:succinyl-CoA transferase and β-ketoadipyl-CoA thiolase respectively (Sgro et al., 2023). Transcriptome data on *Aspergillus niger* growth on protocatechuate compared to fructose suggests the emboldened genes in Table 5 are the most likely candidates to encode the respective enzyme. *NRRL3_01405* was reported to encode protocatechuate 3,4-dioxygenase (Lubbers et al., 2019; Semana and Powlowski, 2019). *Aspergillus niger* growth on protocatechuate indicated a 120-fold increase in *NRRL3_01405* transcript level compared to no carbon source (Lubbers et al., 2019). Growth of *A. niger* on protocatechuate was abolished when *NRRL3_01405* was knocked out (Lubbers et al., 2019). Biochemical assays of the enzyme encoded by *NRRL3_01405* supports the annotation as protocatechuate 3,4-dioxygenase (Semana and Powlowski, 2019).

Alignment results with BlastP for the second enzyme, 3-carboxy-*cis*-*cis* muconate cyclase indicate *NRRL3_02586* as the only candidate within the NRRL3 genome. NRRL3_02586 candidacy is supported by an identity of 91.4 % and 67 % with AN1151 in *A. nidulans* and CMLE_NEUCR in *N. crassa* respectively. AN1151 was observed to have increased protein and gene levels when grown in benzoate and growth was completely inhibited when *AN1151* was knocked out (Martins et al., 2015). CMLE_NEUCR as a purified protein was reported to have carboxy-cis-cis cyclase activity (Mazur et al., 1994). Sgro et al. (2023) reported a 60-fold increase in expression of *NRRL3_02586* when grown on protocatechuate as compared to fructose. The *NRRL3_02586*-encoded enzyme was reported to have 3-carboxy-*cis*-*cis* muconate cyclase activity (Sgro et al., 2023).

The *A. nidulans* gene *AN5232* was reported to encode β-carboxymuconolactone hydrolase/decarboxylase. The strain with Δ*AN5232* resulted in accumulation of protocatechuate

catabolic metabolites and was unable to grow on protocatechuate (Martins et al., 2015). The *A. niger* orthologue of AN5232 is NRRL3_00837, which displays a sequence identity of 70.2 %. Sgro et al. (2023) reported a 1038-fold upregulation in NRRL3_00837 when grown on protocatechuate and Δ*NRRL3_00837* resulted in complete growth inhibition on protocatechuate as a sole carbon source. Another candidate, NRRL3_1409 that shares no significant identity with AN5232 and an identity of 28% with ELH2_ACIAD, a β-ketoadipate enol-lactonase from *Acinetobacter baylyi*. Sgro et al. (2023) reported NRRL3_01409 with a 728-fold upregulation when grown on protocatechuate and observed growth inhibition of the *ΔNRRL3_01409* strain on protocatechuate. The conversion of 3-carboxymuconolactone to β-ketoadipate in *A. niger* has been reported to be catalyzed by an enzyme with hydrolase and decarboxylase activities with a molecular weight of 54(±5) kDa (Thatcher and Cain, 1970). Although NRRL3_00837 is a more likely candidate based on homology, NRRL3_00837 has a predicted molecular weight of 30 kDa, whereas the predicted molecular weight of NRRL3_01409 of 60 kDa is similar in molecular size to the enzyme characterized by Thatcher and Cain (1970).

The most likely candidates for the two downstream enzymes; succinyl-CoA:3-ketoacid coenzyme A transferase and 3-ketoacyl-CoA thiolase candidate are NRRL3_01886 and NRRL3_01526 based on protein sequence homology and transcriptomics data. NRRL3_01886 and NRRL3_01526 were reported to have an upregulation difference of 73-fold when grown on protocatechuate and fructose.

3.2. Diagnostic PCR to confirm the cloning of candidate genes

Protocatechuate catabolic candidate genes from Table 5 were amplified by PCR using cDNA synthesized from *A. niger* poly(A)+RNA grown on 2 % alfalfa and barley as described in the Material and Methods. pLATE11 and pLATE51 were the cloning and expression vectors (Figure 2.1) used in this thesis. pLATE51 includes a 6xHis-tag at the N-terminus of the target protein (Figure 2.2). A diagnostic PCR was performed to check the size of the amplified fragments from the resulting recombinant plasmids. Figure 3.1 shows the results of diagnostic PCR, confirming that *NRRL3_01405*, *NRRL3_02586* and *NRRL3_01409* were correctly cloned into pLATE11 and *NRRL3_01886*, *NRRL3_ 01526* and *NRRL3_00837* were correctly cloned into pLATE51.

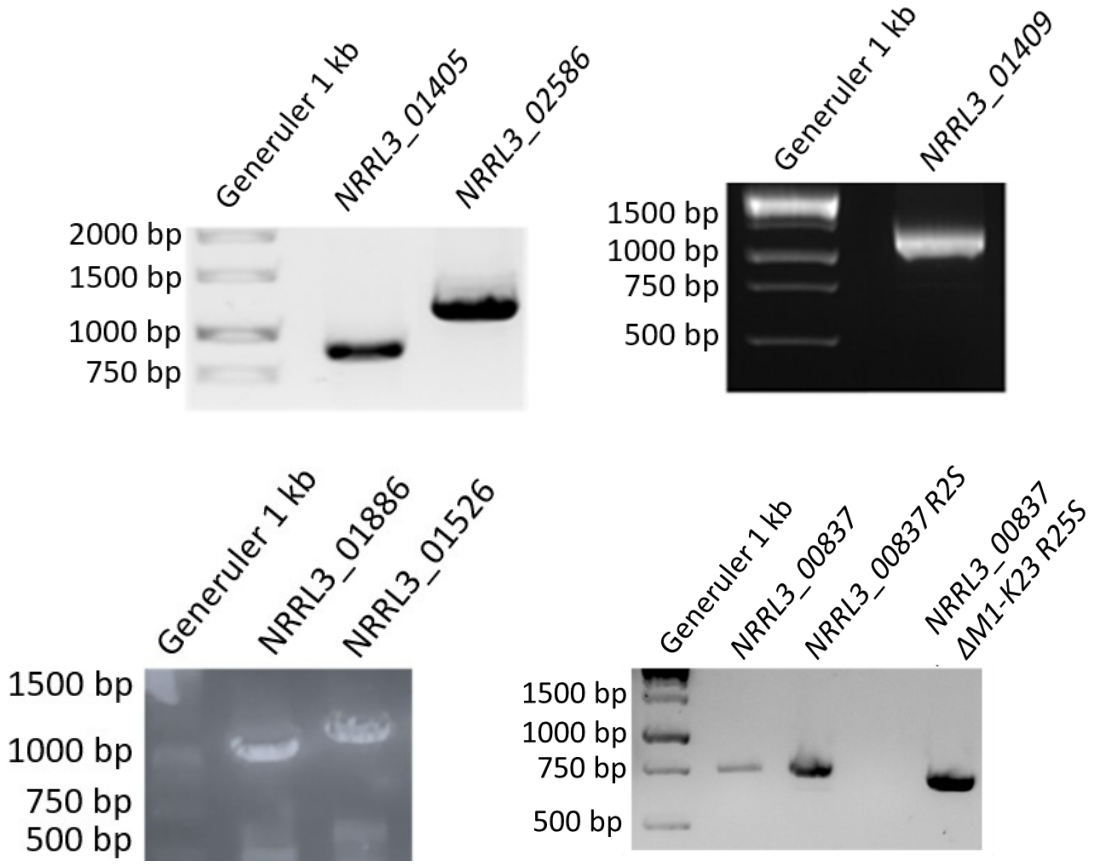


Figure 3.1 Agarose gels (0.8%) of PCR products from the amplification of *Aspergillus niger* genes. Primers used are listed under Table 3. Fragment lengths are 979 bp for *NRRL3_01405*, 1114 bp for *NRRL3_02586*, 1672 bp for *NRRL3_1409*, 1273 bp for *NRRL3_01526,* 1546 bp for *NRRL3_1886*, 661 bp for *NRRL3_00837, NRRL3_00837 R2S* and 592 bp for *NRRL3_00837 ΔM1-K23 R25S* including the 31 bp for the ligation independent cloning kit as described in Material and Methods.

3.2.2 Production of recombinant proteins – verification by LC-MS/MS mass spectrometry

The recombinant proteins were produced in *Escherichia coli* and their identities after purification were verified by LC-MS/MS (Figure S1-S5) as described in Materials and Methods. The soluble target proteins are indicated by arrows in Figure 3.2. Successful induced expression of the candidate genes was observed by the increased band intensity at the molecular weight of the target protein relative to the native *E. coli* proteins in the induced soluble lane compared to the uninduced soluble lane.

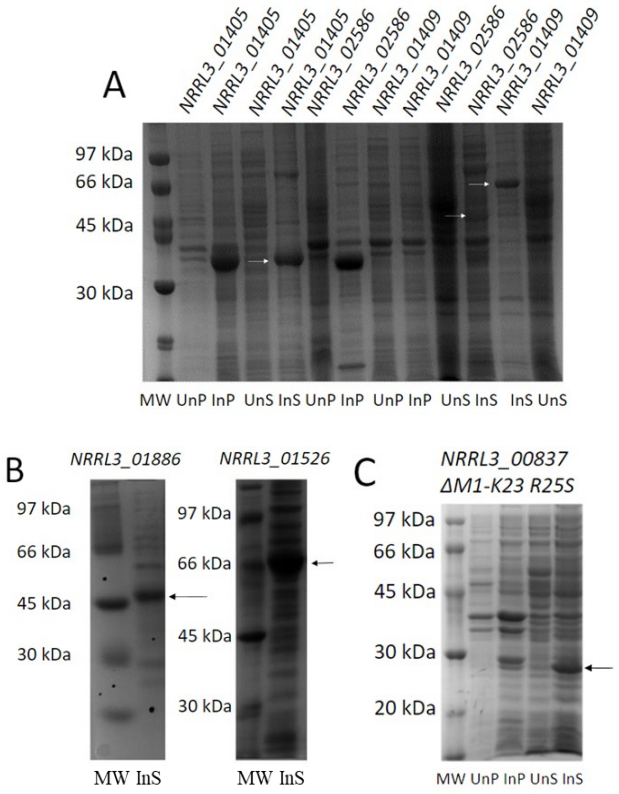


Figure 3.2 Protein production monitored on 12% SDS-PAGE gels (A) *NRRL3_01405*, *NRRL3_02586* and *NRRL3_01409* (B) *NRRL3_01886* and *NRRL3_01526* (C) *NRRL3_00837 ΔM-K23 R25S* expressed in *Escherichia coli*. The arrows point to the band representing the theoretical molecular weight of the protein of interest. "u" represents uninduced, "i" represents induced, "P" represents the pellet and "S" represents the supernatant.

3.2.3 Purification of recombinant proteins by column chromatography

The purification protocol for each protein is described in Table 6. In the case of NRRL3_01409, the purification process is shownin Table 7. The purification of NRRL3_01409 was achieved with a purification fold of 5.9, from a specific activity of 23.7 to 139.0 $\mu molmin^{-1}mol^{-1}$. The purified enzymes in this thesis are presented in Figure 3.5. Visible contamination of native *Escherichia coli* proteins is observed in the purified fractions. The soluble extract of untransformed *E. coli* was not observed to be active in any step in the conversion of protocatechuate to β-ketoadipate based on activity assays.

Table 6. Purification chart of host strain and column chromatography of purified proteins

Gene	Host strain	DEAE-Sephadex	Sepharose S-300	Phenyl Sepharose	His-trap
NRRL3_01405	BL21 (DE3)	1	-	-	-
NRRL3_02586	BL21 (DE3)	1	2	-	-
NRRL3_01409	Rosetta (DE3)	1	2	3	-
NRRL3_00837 ΔM1-K23 R25S	Rosetta (DE3)	-	-	-	1
NRRL3_01886	BL21 (DE3)	-	-	-	1
NRRL3_01526	BL21 (DE3)	-	-	-	1

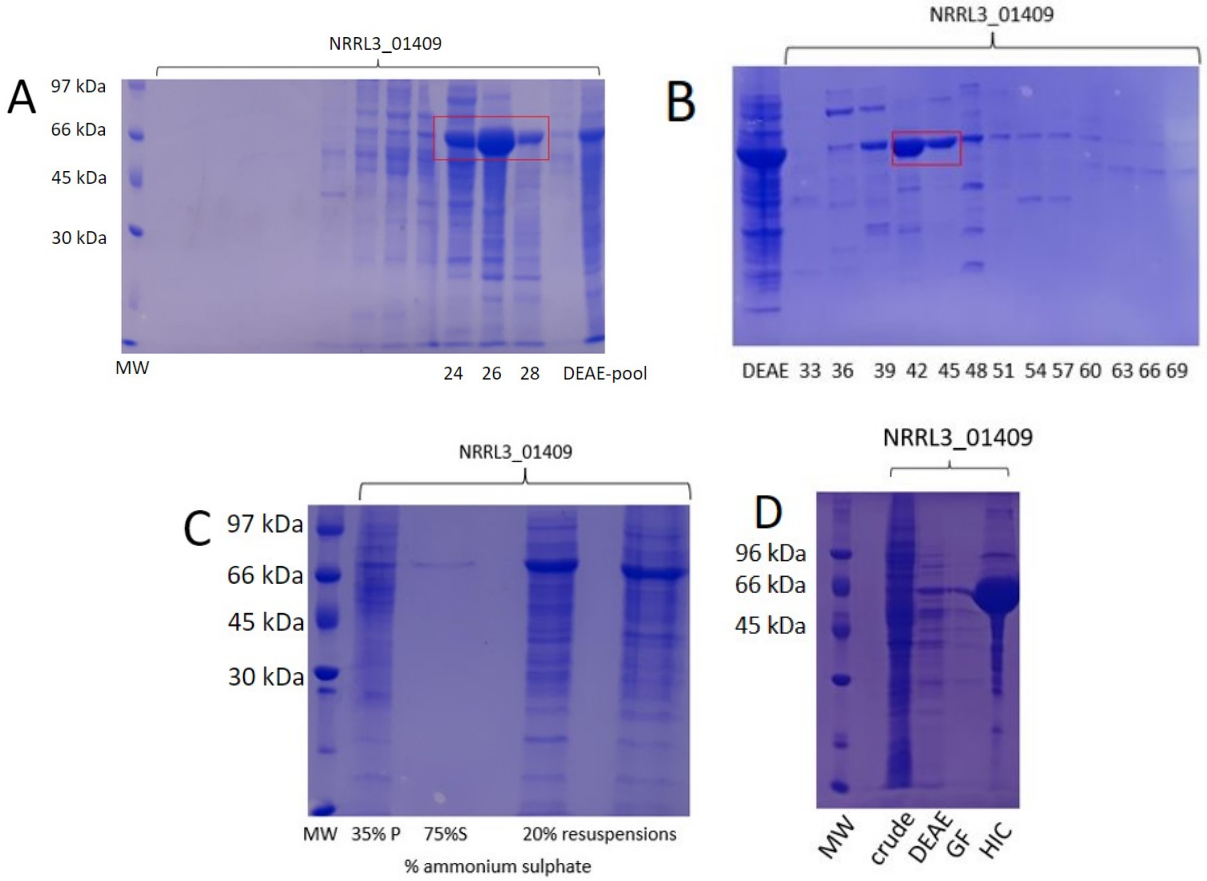


Figure 3.3 Purification of NRRL3_01409. (A) DEAE-Sephadex column (B) S-300 Sepharose column (C) salt precipitation (D) summary of purification; including Sephadex S300 (GF) and Phenyl-Sepharose (HIC) pooled fractions. The boxed bands highlight the predicted protein of interest.

Table 7. Purification table of NRRL3_01409

Purification Step	Total volume (mL)	Total Protein (mg)	Total Activity (units)	Specific Activity ($\mu molmin^{-1}mg^{-1}$)	Yield (%)	Purification fold
Crude extract	60.0	236.82	5620.6	23.73	100.0	1
Ionic exchange	8.0	176.10	4519.3	25.66	80.4	1.1
Size-exclusion	5.0	33.92	2391.1	70.70	42.5	3.0
Hydrophobic interaction	2.1	8.62	1198.2	139.0	21.3	5.9

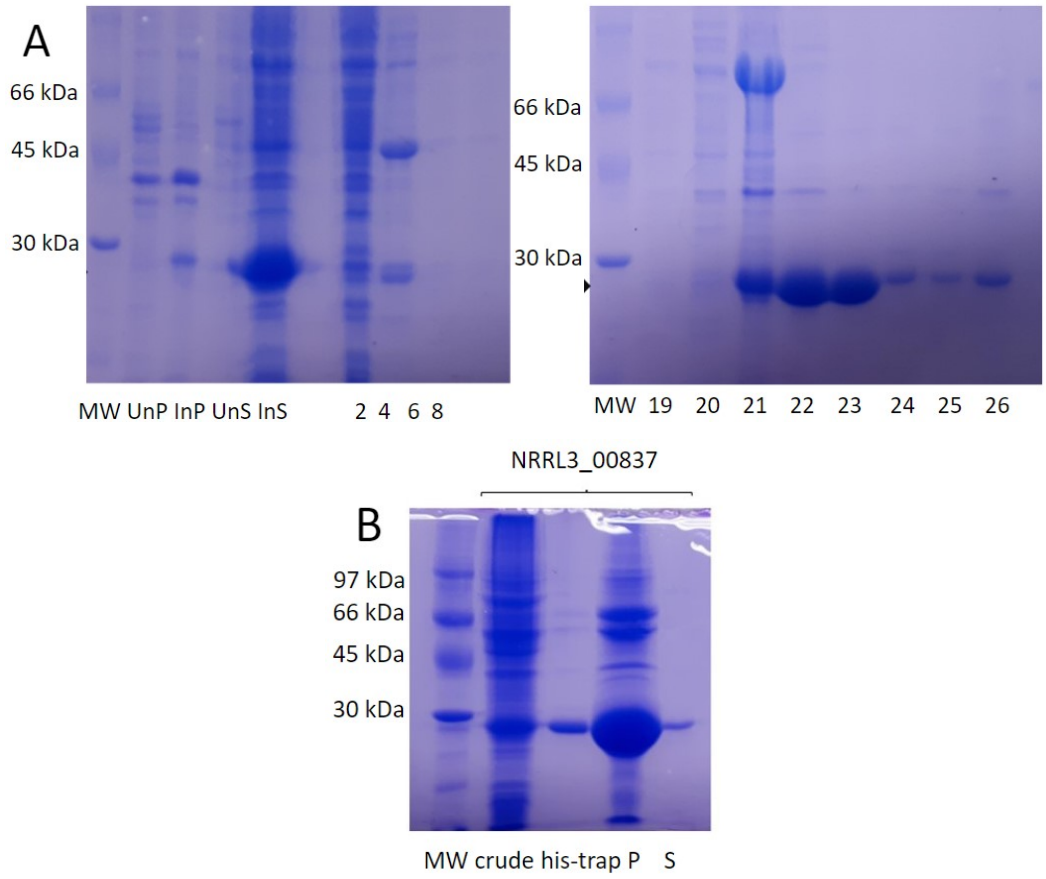


Figure 3.4 Purification of NRRL3_00837 ΔM1-K23 R25S crude extract (A) his-trap column (B) summary purification of NRRL3_00837 ΔM1-K23 R25S. NRRL3_00837 ΔM1-K23 R25S was purified in a 5 mL His-trap column as described in Materials and Methods followed by centrifugation after concentration. "P" represents the pellet and "S" represents the supernatant. The arrows point to the molecular weight of the target protein.

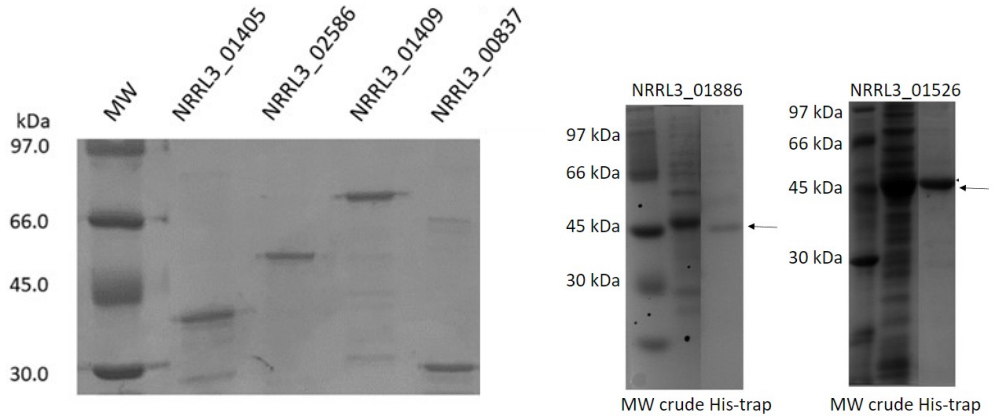


Figure 3.5 Purified recombinant proteins of the protocatechuate degradation pathway. Molecular weight marker (lane 1), protocatechuate 3,4, dioxygenase (lane 2), 3-carboxy-*cis, cis*-muconate cyclase (lane 3), ß-carboxymuconolactone decarboxylase/hydrolase (lane 4),NRRL3_00837 (lane 5) **(left)**. Purification of β-ketoadipyl-CoA thiolase (NRRL3_01886) and β-ketoadipate:succinyl-CoA transferase (NRRL3_01526) **(right)**

3.3 Biochemical characterization of protocatechuate catabolic enzymes

The activity of the candidate protocatechuate catabolic enzymes for the steps leading to the generation of β-ketoadipate were measured by enzyme activity assays calculated from difference in the extinction coefficients between protocatechuate, 3-carboxy-*cis*, *cis*-muconate, β-carboxymuconolactone and β-ketoadipate at 290, 260, and 230 nm, respectively. As observed in Figure 3.7A, after the addition of NRRL3_01405, absorption decreased at 290 nm and increased at 260 nm indicating conversion of protocatechuate to 3-carboxy-*cis, cis*-muconate. In Figure 3.7B, the addition of NRRL3_02586 led to a decrease in absorption at 260 nm and an increase at 230 nm, indicating a conversion of 3-carboxy-*cis, cis*-muconate to β-carboxymuconolactone. Finally, addition of NRRL3_01409 to the mixture led to a decrease in absorption at 230 nm indicating metabolism of β-carboxymuconolactone. Isosbestic points observed around 250 nm and 280 nm after the addition of NRRL3_01405 and at 235 nm after the addition of NRRL3_02586 and NRRL3_01409 were consistent with previous reports (Cain et al., 1968).

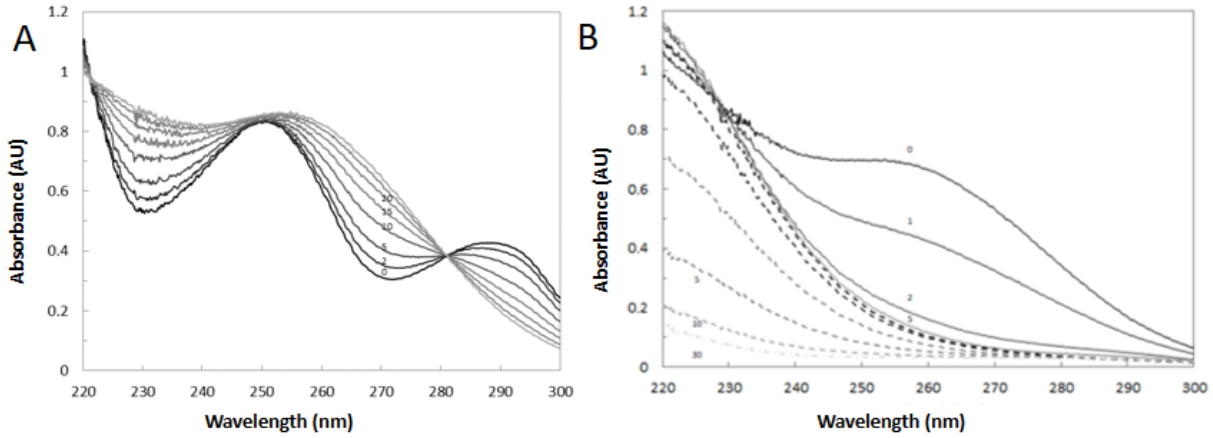


Figure 3.6 (A) Scanning kinetics of NRRL3_01405 protocatechuate 3,4 dioxygenase (6.35 µg) in 100 µM protocatechuate and 100 mM Tris-Cl pH 7.5 at a scan speed of 360 nm/min (left). **(B) Scanning kinetics of NRRL3_02586 3-carboxy-*cis*,*cis* muconate cyclase (0.19 µg) and NRRL3_01409 ß-carboxymuconolactone decarboxylase/hydrolase (0.38 µg)** in substrate as prepared in the previous figure at a scan speed of 360nm/min. Solid line; NRRL3_02586, dotted line; NRRL3_01409 (right).

The protocatechuate intradiol cleavage catabolic pathway differs between the bacteria and fungi in the formation of a γ-substituted and β-substituted metabolite, respectively. 3-carboxy-*cis*, *cis* muconate cyclase and ß-carboxymuconolactone decarboxylase/hydrolase are unique to the fungal pathway. The most likely candidate for ß-carboxymuconolactone decarboxylase/hydrolase activity was NRRL3_00837 based on homology with the *Aspergillus nidulans* AN5232. However, NRRL3_01409 was observed to have ß-carboxymuconolactone decarboxylase/hydrolase activity. The enzyme encoded by *NRRL3_1409* was characterized for ß-carboxymuconolactone decarboxylase/hydrolase kinetic parameters in triplicate as depicted in Figure 3.6 and for stability in Table 8. The determined K_m and k_{cat} for NRRL3_01409 are 187 µM and 0.336 ± 0.072 s^{-1} respectively.

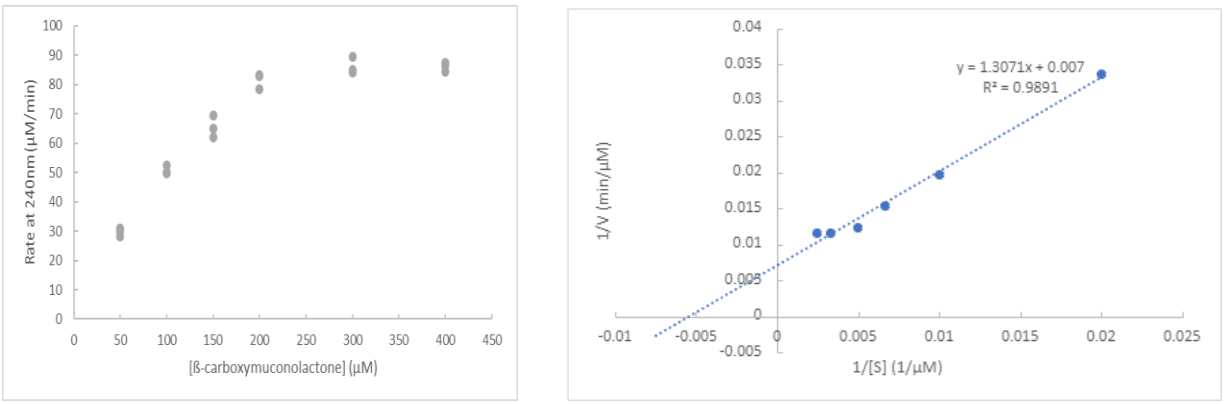


Figure 3.7 Michaelis-Menten (left), Lineweaver-Burke (right) plots of NRRL3_1409. Assay conditions; 100 mM Tris-Cl pH 7.5, 0.454 µg NRRL3_01409 and ß-carboxymuconolactone. ß-carboxymuconolactone was prepared by incubating 1 mM protocatechuate with NRRL3_01405 and NRRL3_02586.

3.4 Mass spectrometric spectra of protocatechuate metabolites

None of the intermediates of the protocatechuate pathway are commercially available. Therefore, production of intermediates in the protocatechuate catabolic pathway was necessary for enzymatic activity assays. The intermediates were prepared by incubating protocatechuate with the enzymes upstream of the enzyme to be assayed. In this study, the molecular weight and chemical structures of the intermediates at each step of the protocatechuate catabolic pathway were identified using mass spectrometry and ^{1}H NMR.

The intermediates analyzed by mass spectrometry were bound to sodium before being measured and were observed to be 22.98 m/z larger. The mass to charge ratios of β-carboxymuconolactone (Figure 3.8A) and β-ketoadipate (Figure 3.8B) were detected at 209.01 and 183.03mz/e respectively, supporting that *NRRL3_01405*, *NRRL3_02586* and *NRRL3_01409* encode the protocatechuate 3,4-dioxygenase, 3-carboxy-*cis*-*cis* muconate cyclase and ß-carboxymuconolactone decarboxylase/hydrolase respectively in the protocatechuate catabolic pathway. No significant change was observed after the addition of NRRL3_00837 to the enzyme cocktail producing β-ketoadipate (Figure 3.8C).

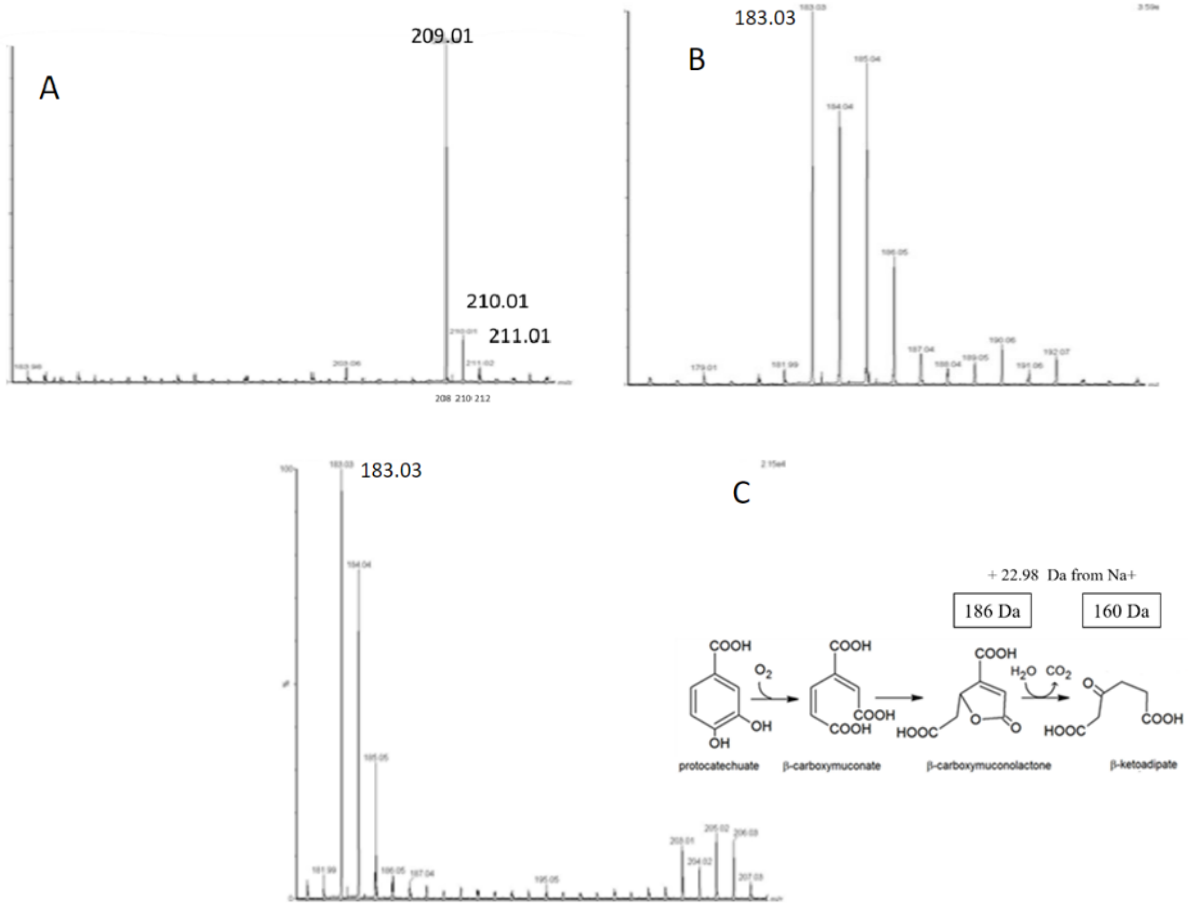


Figure 3.8 TOF MS ES+ spectra of the compounds produced from the incubation of protocatechuate with: **A** NRRL3_01405 + NRRL3_02586; **B** NRRL3_01405 + NRRL3_02586 + NRRL3_01409; **C** NRRL3_01405 + NRRL3_02586 + NRRL3_01409 + NRRL3_00837. M/z ratios of the compounds.

3.5 ^{1}H NMR spectra of metabolites

The ^{1}H-NMR spectrum in Figure 3.9A shows that chemical shifts and coupling constants of the mixture containing NRRL3_01405 and protocatechuate were consistent with the predicted spectra of 3-carboxy-*cis*,*cis*-muconic acid; δ 6.01(d, 1H, J = 11.5 Hz), 6.50 (d,1H, J = 2.0 Hz), 7.05 (dd, 1H, J = 2.0 Hz, 11.5 Hz). Additional peaks at 2.63, 3.06, 5.53 and 6.85 ppm were consistent with the presence of the next intermediate of the pathway, 3-carboxymuconolactone.

Although no lactonizing enzyme was present, it is likely that acid-catalyzed lactonization occurred when 3-carboxy-*cis, cis*-muconic acid was prepared for extraction in ethyl acetate, forming 3-carboxymuconolactone (MacDonald, 1954).

^{1}H-NMR analysis of NRRL3_01405 and NRRL3_02586 with protocatechuate in Figure 3.9B confirms the intermediate as β-carboxymuconolactone; δ 2.63 (dd, 1H, J = 7.9, 16.4 Hz), 3.06 (dd, 1H, J = 3.3, 16.4 Hz,), 5.53 (ddd, 1H, J = 2.1, 3.3, 7.9 Hz), 6.80 (d, 1H, J = 2.1 Hz).

^{1}H-NMR analysis following the addition of NRRL3_01409 to the mixture in Figure 3.9C was identical to that of predicted spectrum for β-ketoadipate: δ 2.38 (t, 2H, J = 6.7 Hz), 2.65 (t, 2H, J = 6.7 Hz), 3.48 (s, 2H).

In the upper panels of Figure 3.9, NRRL3_00837 was added to each mixture to observe any potential enzyme activity. However, the spectrum remained unchanged, indicating NRRL3_00837 does not play a protocatechuate catabolic role *in vitro*.

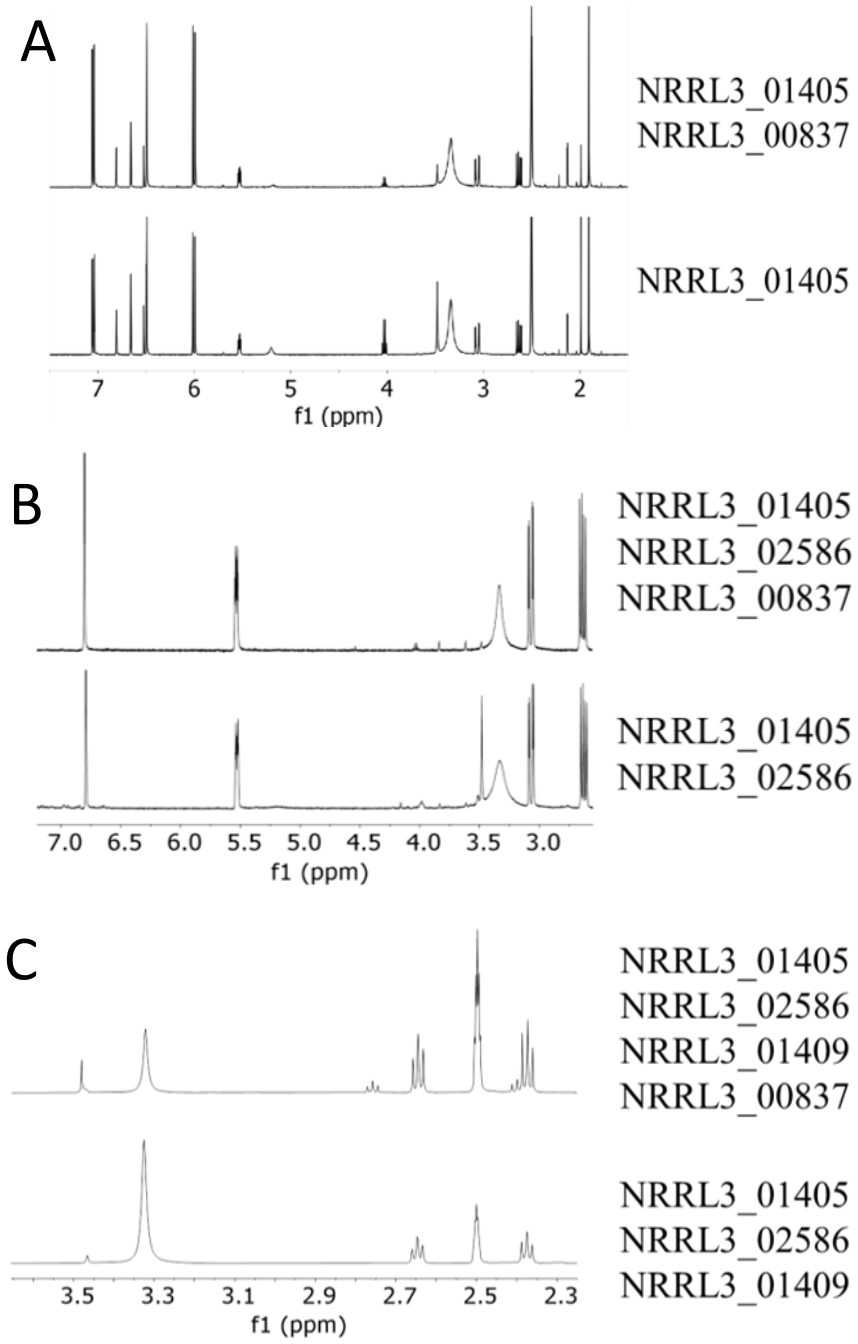


Figure 3.9 ^{1}H NMR (500 MHz, DMSO) spectra of the compounds produced from the incubation of protocatechuate with: **A** NRRL3_01405 (lower), NRRL3_1405+ NRRL3_00837 (upper); **B** NRRL3_1405 + NRRL3_02586 (lower), NRRL3_1405+ NRRL3_02586 + NRRL3_00837 (upper); **C** NRRL3_1405+ NRRL3_02586 + NRRL3_01409 (lower), NRRL3_1405+ NRRL3_02586 + NRRL3_01409 + NRRL3_00837 (upper)

3.6 Complete *in vitro* conversion of protocatechuate to β-ketoadipate does not require NRRL3_00837

Based on the biochemical evidence from characterizing the chemical structures of the intermediates and by enzyme activity assays, *NRRL3_01405*, *NRRL3_02586* and *NRRL3_01409* are assigned as the genes encoding protocatechuate 3,4, dioxygenase, 3-carboxy-*cis, cis*-muconate cyclase and β-carboxymuconolactone hydrolase/decarboxylase, respectively. These assignments are supported by mutational and transcriptomic studies of Sgro et al. (2023). The enzyme encoded by *NRRL3_00837* was reported to play a role in *in vivo* catabolism of protocatechuate as Δ*NRRL3_00837* strains are unable to grow using protocatechuate as sole carbon source (Sgro et al., 2023). After confirming that the chemical composition and chemical structure of the intermediates remains unchanged in the presence of NRRL3_00837, the enzymatic rates were measured in the presence and absence of NRRL3_00837. As observed in Table 8, no significant change in rates was observed at any step of the conversion of protocatechuate to β-ketoadipate, including the downstream conversion activity of β-ketoadipate:succinyl CoA transferase. No spectral changes occurred between 220-300 nm with the addition of NRRL3_00837 to protocatechuate and sequential addition of NRRL3_01405, NRRL3_02586 and NRRL3_01409 compared to an identical run in the absence of NRRL3_00837. Based on these assays, we conclude that NRRL3_00837 is not required for the *in vitro* conversion of protocatechuate to β-ketoadipate.

Table 8. Rates of protocatechuate catabolism in the presence and absence of NRRL3_00837. The assays were run in triplicate, the average rate and standard deviation are tabulated.

NRRL3_00837 (µg)	NRRL3_01405 Activity (µM/min)	NRRL3_02586 Activity (µM/min)	NRRL3_01409 Activity (µM/min)
0	78.3 ± 0.75	70.2 ± 5.32	43.5 ± 3.90
2.45 µg	82.2 ± 4.14	75.2 ± 5.30	46.1 ± 1.16
12.25 µg	83.2 ± 8.20	70.1 ± 2.90	42.1 ± 5.06

4. Discussion

The bacterial intradiol cleavage of protocatechuate has been extensively mapped out by biochemical and molecular biological studies whereas the annotations in the fungal pathways are sparsely supported by experimental evidence. The bacterial pathway involves four metabolites; protocatechuate, 3-carboxy-*cis, cis*-muconic acid, γ-carboxymuconolactone, 3-oxoadipate enol-lactone, and β-ketoadipate. Each step is catalyzed by a unique enzyme; protocatechuate 3,4 dioxygenase, 3-carboxy-*cis, cis*-muconate cycloisomerase, 4-carboxymuconolactone decarboxylase, and β-ketoadipate enol-lactone hydrolase respectively. Whereas the fungal pathway was described in *Aspergillus niger* by only three metabolites; protocatechuate, 3-carboxy-*cis, cis*-muconic acid, β-carboxymuconolactone, and β-ketoadipate catalyzed by protocatechuate 3,4-dioxygenase, 3-carboxy-*cis, cis*-muconate cyclase and β-carboxymuconolactone hydrolase/decarboxylase. (Cain et al., 1968). The overall difference between the bacterial and fungal pathways are the γ- and β-substituted carboxymuconolactone, respectively and the hydrolysis and decarboxylation of β-carboxymuconolactone to form β-ketoadipate by a single enzyme whereas γ-carboxymuconolactone undergoes hydrolysis and decarboxylation by two enzymes. The pathways downstream of β-ketoadipate are identical between bacteria and fungi; β-ketoadipate:succinyl CoA transferase and β-ketoadipyl-CoA thiolase form β-ketoadipyl-CoA, and acetyl-CoA/succinyl CoA respectively.

Experimental evidence for the annotation of the *Aspergillus nidulans* protocatechuate catabolic pathway has been reported in the form of transcriptomic and genetic mutation studies (Martins et al., 2015). The enzymes participating in the pathway; protocatechuate 3,4-dioxygenase, 3-carboxy-*cis, cis*-muconate cyclase, β-carboxymuconolactone hydrolase/decarboxylase, and β-ketoadipate:succinyl-CoA transferase have been as assigned to protein IDs AN8566, AN1151, AN5232 and AN10495 respectively (Martins et al., 2015). Based on the genomic and mutational analyses of Sgro et al. (2023) and the biochemical characterization in this thesis, we conclude that protocatechuate 3,4-dioxygenase, 3-carboxy-*cis, cis*-muconate cyclase, β-carboxymuconolactone hydrolase/decarboxylase, β-ketoadipate:succinyl-CoA transferase and β-ketoadipyl-CoA thiolase in *A. niger* are encoded by the genes *NRRL3_01405*, *NRRL3_02586*, *NRRL3_00837*, *NRRL3_01886* and *NRRL3_01526* (Table 9).

Table 9. *Aspergillus niger* protocatechuate catabolic pathway proteins derived from *Aspergillus nidulans* orthologues

Protocatechuate catabolic proteins	*Aspergillus nidulans*	*Aspergillus niger*	Sequence identity (%)
Protocatechuate 3,4-dioxygenase	AN8566*	NRRL3_01405	94.8
3-carboxy-*cis, cis*-muconate cyclase	AN1151*	NRRL3_02586	91.4
β-caboxymuconolactone hydrolase/decarboxylase	AN10520	NRRL3_01409	70.4
β-caboxymuconolactone hydrolase/decarboxylase	AN5232*	NRRL3_00837	70.2
β-ketoadipate:succinyl CoA transferase	AN10495*	NRRL3_1886	86.2
β-ketoadipyl-CoA thiolase	AN5698	NRRL3_01526	91

*Proteins assigned by Martins et al. (2015)

4.1 Product analysis of NRRL3_01405 confirms annotation as protocatechuate 3,4-dioxygenase

The purification and UV-Vis spectral analysis of the intradiol cleavage of protocatechuate to 3-carboxy-*cis, cis*-muconate by NRRL3_01405 was previously reported (Semana and Patrick, 2019), however the reaction product was not analyzed. The ^{1}H-NMR spectrum of the product of NRRL3_01405 in protocatechuate (Figure 3.9A) shows that chemical shifts and coupling constants of 3-carboxy-*cis, cis*-muconate are similar to reports by Yamanashi et al. (2015) with the exception of multiplicity observed in allylic coupling; δ 6.01(d, 1H, J = 11.5 Hz), 6.50 (d,1H, J = 2.0 Hz), 7.05 (dd, 1H, J = 2.0 Hz, 11.5 Hz) versus δ 6.01(dd, 1H, J = 0.9, 12.0 Hz), 6.50 (m, 1H), 6.77 (dd, 1H, J = 1.7, 12.0 Hz). Additional peaks at 2.63, 3.06, 5.53 and 6.85 ppm are consistent with the presence of β-carboxymuconolactone. Although no lactonizing enzyme was present, it is likely that acid-catalyzed lactonization occurred when the sample was prepared for extraction in ethyl acetate. Rapid acid-catalyzed isomerization and lactonization of

protocatechuate has been reported by MacDonald et al. (1954), acidification was not performed in the preparation done by Yamanashi et al. (2015).

4.1 NRRL3_02586 annotation confirmation as 3-carboxy-*cis, cis*-muconate cyclase

The conversion of 3-carboxy-*cis, cis*-muconic acid to β-carboxymuconolactone by NRRL3_02586 is supported by the isosbestic points at 230 nm and decrease in absorbance at 260 nm observed in UV-Vis spectral assays (Figure 3.6B) that have been previously reported in enzymology studies by Cain et al. (1968). The purified NRRL3_02586 migrated at approximately 50 kDa on 12 % SDS-PAGE compared to the 47 kDa 3-carboxy-*cis, cis*-muconate cyclase characterized by Thatcher and Cain (1974).

The product of protocatechuate incubated with NRRL3_01405 and NRRL3_02586 was isolated and analyzed by MS and ^{1}H-NMR. The molecular weight of β-carboxymuconolactone bound to sodium was confirmed as 209.1 Da by MS. ^{1}H-NMR analysis confirms the NRRL3_02586 product as β-carboxymuconolactone (Figure 3.9B) by comparison with the ^{1}H-NMR of muconolactone from Yamanashi et al. (2015). The additional peaks and coupling constants are consistent with the report by Yamanashi et al. (2015) on muconolactone due to the additional hydrogen at C_5: δ 2.63 (dd, 1H, J = 7.9, 16.4 Hz), 3.06 (dd, 1H, J = 3.3, 16.4 Hz,), 5.53 (ddd, 1H, J = 2.1, 3.3, 7.9 Hz), 6.80 (d, 1H, J = 2.1 Hz) versus δ 2.71 (dd, 1H, J = 8.2, 16.6 Hz), 2.95 (dd, 1H, J = 4.8, 16.6 Hz), 5.57 (dddd, 1H, J = 1.4, 1.9, 4.8, 8.2 Hz), 6.24 (dd, 1H, J = 1.9, 5.8 Hz), 7.81 (dd, 1H, J = 1.4, 5.8 Hz). The reported chemical shifts are also consistent with β-carboxymuconolactone values from Kondo et al. (2016): δ 2.67, 3.10, 5.55, 6.81 ppm.

The 3-carboxy-*cis, cis*-muconate lactonizing enzyme in *N. crassa* was observed to lactonize 3-carboxy-*cis, cis*-muconic acid but also *cis, cis*-muconic acid with significantly reduced activity (Mazur et al., 1994). The 3-carboxy-*cis, cis*-muconate lactonizing enzyme homologues in *A. nidulans* and *A. niger* have not been assayed for activity with *cis, cis*-muconic acid. In terms of protocatechuate valorization, *cis, cis*-muconic acid is a highly valued chemical, therefore repressing the expression of *AN1151* and *NRRL3_02586* may be necessary to maximize *cis, cis*-muconic acid production in filamentous fungi.

4.2 NRRL3_00837 plays no significant role in the *in vitro* protocatechuate degradation pathway of *Aspergillus niger*

NRRL3_00837 was the most likely candidate as β-carboxymuconolactone hydrolase/decarboxylase based on sequence identity with the orthologue AN5232 in *Aspergillus nidulans* (Table 9). Deletion of *NRRL3_00837* was reported to inhibit growth on protocatechuate *in vivo,* however, accumulation of β-carboxymuconolactone was not detected (Sgro et al., 2023) Transcriptomics studies illustrate a 50-fold increase in *NRRL3_00837* with a difference of 21 transcripts per million (TPM) to 1048 TPM when grown on fructose and protocatechuate respectively (Sgro et al., 2023), however accumulation of β-carboxymuconolactone was not detected (Sgro et al., 2023). Enzymology studies by Cain et al. (1968) reported β-carboxymuconolactone hydrolase/decarboxylase in *Aspergillus niger* as a 54±5 kDa protein whereas *NRRL3_00837* encodes a 30 kDa protein. In this thesis, biochemical studies with purified NRRL3_00837 indicated the enzyme has no activity in the protocatechuate catabolic pathway *in vitro*.

The protein encoded by *NRRL3_00837* was included in the reaction mixtures with protocatechuate, NRRL3_01405 (protocatechuate 3,4-dioxygenase), NRRL3_02586 (3-carboxy-*cis, cis*-muconate cyclase) and NRRL3_01409 (β-carboxymuconolactone hydrolase/decarboxylase) in UV-Vis enzyme activity assays and metabolite extraction for ^{1}H NMR and mass spectrometry. No significant UV-Vis spectral changes were detected in the presence of NRRL3_00837 compared to the omission of NRRL3_00837 (Figure 3.6), nor did it catalyze any one of these reaction steps alone. There was also no significant spectral difference in the mass spectrometry or ^{1}H NMR spectra in the presence of NRRL3_00837 (Figures 3.8 and Figure 3.9). Thus, NRRL3_00837 does not affect the chemistry of the protocatechuate to β-ketoadipate conversion *in vitro*. The possibility that the protein encoded by NRRL3_00837 could enhance the rates of one of the enzymes between protocatechuate and β-ketoadipate was tested by adding this protein at one of two different concentrations to single wavelength assays of each of the three enzymes. As shown in Table 8, none of these rates was significantly affected. The role of NRRL3_00837 in the degradation of protocatechuic acid *in vivo* remains unknown.

4.3 NRRL3_01409 annotation identified as β-carboxymuconolactone hydrolase/decarboxylase

The conversion of β-carboxymuconolactone to β-ketoadipate by NRRL3_01409 was assayed as a β-carboxymuconolactone hydrolase/decarboxylase and the spectral changes are identical to those reported by Cain et al. (1968). NRRL3_01409 is a 61 kDa protein and is consistent with the previously reported 54±5 kDa β-carboxymuconolactone hydrolase/decarboxylase in *Aspergillus niger* (Cain et al.,1968). The determined Km and K_{cat} for NRRL3_01409 are 187 μM and 0.336 ± 0.072 s^{-1} respectively at pH 7.5 and 25 °C compared to the K_m of 70 μM in optimal pH of 8.6 reported by Thatcher and Cain (1970). NRRL3_01409 shares no significant sequence identity to AN5232, a protein assigned as β-carboxymuconolactone hydrolase/decarboxylase in *Aspergillus nidulans* (Martins et al., 2015) and a 70.4% sequence identity with AN10520, however, there are no studies on AN10520. NRRL3_01409 does share a 28% identity with the homologue ELH2_ACIAD in *Acinetobacter baylyi*. *NRRL3_01409* knockout was reported to impair growth on protocatechuate, and transcriptome studies indicate a 70-increase increase *NRRL3_01409* transcripts with a difference of 10.5 TPM to 727 TPM when grown on fructose and protocatechuate respectively (Sgro et al., 2023).

The product of protocatechuate incubated with NRRL3_01405 (protocatechuate 3,4-dioxygenase), NRRL3_02586 (3-carboxy-*cis,cis*-muconate cyclase) and NRRL3_01409 was isolated and analyzed by MS and ^{1}H-NMR. The molecular weight of the product bound to sodium was confirmed as 183.03 Da by MS. ^{1}H-NMR analysis following the addition of NRRL3_01409 to the mixture in Figure 3.9C was identical to that of β-ketoadipate reported by Yamanashi et al. (2015): δ 2.38 (t, 2H, J = 6.7 Hz), 2.65 (t, 2H, J = 6.7 Hz), 3.48 (s, 2H) versus δ 2.55 (t, 2H, J = 6.5 Hz), 2.86 (t, 2H, J = 6.5 Hz), 3.39 (s, 2H). Hence, we conclude that *NRRL3_01409* encodes β-carboxymuconolactone hydrolase/decarboxylase.

4.4 Conclusions and Future Plans

In this thesis, recombinant NRRL3_01405, NRRL3_02586, NRRL3_01409 and NRRL3_00837 were purified, assayed for enzymatic activity in the protocatechuate catabolic pathway with the

metabolites analyzed by MS and ^{1}H NMR. NRRL3_01405 was confirmed as protocatechuate 3,4 dioxygenase with enzyme activity and molecular weight similar to the enzyme characterized by Cain et al. (1968) and Lubbers et al. (2019). NRRL3_02586 was confirmed as 3-carboxy-*cis, cis*-muconate cyclase with similar enzyme activity and molecular weight reported by Cain et al. (1968) and to produce β-carboxymuconolactone by metabolite analysis. NRRL3_01409 was assigned as β-carboxymuconolactone hydrolase/decarboxylase with similar enzyme activity and molecular weight reported by Cain et al, (1968) and to produce β-ketoadipate by metabolite analysis. *NRRL3_01886* and *NRRL3_01526* had been expressed and the encoded enzymes purified as potential β-ketoadipate:succinyl-CoA transferase and β-ketoadipyl-CoA thiolase candidates. NRRL3_00837 was investigated as the most likely β-carboxymuconolactone hydrolase/decarboxylase candidate, however, no significant activity was observed by NRRL3_00837 at any step in the *in vitro* conversion of protocatechuate to β-ketoadipate.

NRRL3_00837 was reported to participate in the *in vivo* catabolism of protocatechuate (Sgro et al., 2023), although the function of NRRL3_00837 remains unknown. Post-translational modifications have been reported in the bacteria operons of protocatechuate catabolism. In the presence of protocatechuate, PcaU has been annotated as the transcriptional activator of *pca* genes by binding to the *pcaU-pcaI* intergenic region. In this thesis, *NRRL3_00837* was expressed in *Escherichia coli* as bacterial host. Investigation of NRRL3_00837 as a potential regulatory protein may be worthwhile as NRRL3_00837 has been reported to be essential to *in vivo* protocatechuate catabolism and with no significant activity in the *in vitro* protocatechuate catabolism.

Characterization of AN10520 in *A. nidulans*, predicted to be a β-carboxymuconolactone decarboxylase/hydrolase based on sequence identity with NRRL3_01409 may identify a second enzyme with activity on β-carboxymuconolactone. Characterization of AN5698, predicted to be a β-ketoadipyl-CoA thiolase based on sequence identity with NRRL3_01526 may complete the annotation of the β-ketoadipate to citric acid cycle intermediates in *A. nidulans.*

5. References

Acourene, S., & Ammouche, A. (2012). Optimization of ethanol, citric acid, and α-amylase production from date wastes by strains of Saccharomyces cerevisiae, Aspergillus niger, and Candida guilliermondii. *Journal of Industrial Microbiology & Biotechnology*, *39*(5), 759–766. https://doi.org/10.1007/s10295-011-1070-0

Adeboye, P. T., Bettiga, M., & Olsson, L. (2014). The chemical nature of phenolic compounds determines their toxicity and induces distinct physiological responses in Saccharomyces cerevisiae in lignocellulose hydrolysates. *AMB Express*, *4*, 46. https://doi.org/10.1186/s13568-014-0046-7

Archer, D. B. (1994). Enzyme production by recombinant Aspergillus. *Bioprocess Technology*, *19*, 373–393.

Arif, T. (2015). Salicylic acid as a peeling agent: A comprehensive review. *Clinical, Cosmetic and Investigational Dermatology*, *8*, 455–461. https://doi.org/10.2147/CCID.S84765

Baker, S. E. (2006). Aspergillus niger genomics: Past, present and into the future. *Medical Mycology*, *44 Suppl 1*, S17-21. https://doi.org/10.1080/13693780600921037

Brückner, C., Oreb, M., Kunze, G., Boles, E., & Tripp, J. (2018). An expanded enzyme toolbox for production of cis, cis-muconic acid and other shikimate pathway derivatives in Saccharomyces cerevisiae. *FEMS Yeast Research*, *18*(2), foy017. https://doi.org/10.1093/femsyr/foy017

Buchan, A., Collier, L. S., Neidle, E. L., & Moran, M. A. (2000). Key aromatic-ring-cleaving enzyme, protocatechuate 3,4-dioxygenase, in the ecologically important marine Roseobacter lineage. *Applied and Environmental Microbiology*, *66*(11), 4662–4672. https://doi.org/10.1128/AEM.66.11.4662-4672.2000

Cain, R. B., Bilton, R. F., & Darrah, J. A. (1968). The metabolism of aromatic acids by micro-organisms. Metabolic pathways in the fungi. *Biochemical Journal*, *108*(5), 797–828. https://doi.org/10.1042/bj1080797

Cairns, T. C., Nai, C., & Meyer, V. (2018). How a fungus shapes biotechnology: 100 years of Aspergillus niger research. *Fungal Biology and Biotechnology*, *5*, 13. https://doi.org/10.1186/s40694-018-0054-5

Cao, L., Yu, I. K. M., Liu, Y., Ruan, X., Tsang, D. C. W., Hunt, A. J., Ok, Y. S., Song, H., & Zhang, S. (2018). Lignin valorization for the production of renewable chemicals: State-of-the-art review and future prospects. *Bioresource Technology*, *269*, 465–475. https://doi.org/10.1016/j.biortech.2018.08.065

Cesário, L. M., Pires, G. P., Pereira, R. F. S., Fantuzzi, E., da Silva Xavier, A., Cassini, S. T. A., & de Oliveira, J. P. (2021). Optimization of lipase production using fungal isolates from oily residues. *BMC Biotechnology*, *21*(1), 65. https://doi.org/10.1186/s12896-021-00724-4

Choi, S., Lee, H.-N., Park, E., Lee, S.-J., & Kim, E.-S. (2020). Recent Advances in Microbial Production of cis,cis-Muconic Acid. *Biomolecules*, *10*(9), 1238. https://doi.org/10.3390/biom10091238

Crawford, R. L., Bromley, J. W., & Perkins-Olson, P. E. (1979). Catabolism of protocatechuate by Bacillus macerans. *Applied and Environmental Microbiology*, *37*(3), 614–618. https://doi.org/10.1128/aem.37.3.614-618.1979

Dennis, D. A., Chapman, P. J., & Dagley, S. (1973). Degradation of protocatechuate in Pseudomonas testosteroni by a pathway involving oxidation of the product of meta-fission. *Journal of Bacteriology*, *113*(1), 521–523. https://doi.org/10.1128/jb.113.1.521-523.1973

Doten, R. C., Ngai, K. L., Mitchell, D. J., & Ornston, L. N. (1987). Cloning and genetic organization of the pca gene cluster from Acinetobacter calcoaceticus. *Journal of Bacteriology*, *169*(7), 3168–3174. https:/doi: 10.1128/jb.169.7.3168-3174.1987

Duarte, J. C., & Costa-Ferreira, M. (1994). Aspergilli and lignocellulosics: Enzymology and biotechnological applications. *FEMS Microbiology Reviews*, *13*(2–3), 377–386. https://doi.org/10.1111/j.1574-6976.1994.tb00057

Elsemore, D. A., & Ornston, L. N. (1994). The pca-pob supraoperonic cluster of Acinetobacter calcoaceticus contains quiA, the structural gene for quinate-shikimate dehydrogenase. *Journal of Bacteriology*, *176*(24), 7659–7666. https://doi.org/10.1128/jb.176.24.7659-7666.1994

Fasiku, S. A., Bello, M. A., & Odeniyi, O. A. (2023). Production of xylanase by Aspergillus niger GIO and Bacillus megaterium through solid-state fermentation. *Access Microbiology*, *5*(6), acmi000506.v5. https://doi.org/10.1099/acmi.0.000506.v5

Freudenberg, K. (1965). Lignin: Its Constitution and Formation from p-Hydroxycinnamyl Alcohols: Lignin is duplicated by dehydrogenation of these alcohols; intermediates explain formation and structure. *Science (New York, N.Y.)*, *148*(3670), 595–600. https://doi.org/10.1126/science.148.3670.595

Friedman, M., Henika, P. R., & Mandrell, R. E. (2003). Antibacterial activities of phenolic benzaldehydes and benzoic acids against Campylobacter jejuni, Escherichia coli, Listeria monocytogenes, and Salmonella enterica. *Journal of Food Protection*, *66*(10), 1811–1821. https://doi.org/10.4315/0362-028x-66.10.1811

Gerischer, U., Segura, A., & Ornston, L. N. (1998). PcaU, a Transcriptional Activator of Genes for Protocatechuate Utilization in Acinetobacter. *Journal of Bacteriology*, *180*(6), 1512–1524. doi: 10.1128/JB.180.6.1512-1524.1998.

Guiraud, P., Steiman, R., Seigle-Murandi, F., & Benoit-Guyod, J. L. (1995). Comparison of the toxicity of various lignin-related phenolic compounds toward selected fungi perfecti and fungi imperfecti. *Ecotoxicology and Environmental Safety*, *32*(1), 29–33. https://doi.org/10.1006/eesa.1995.1081

Hara, H., Masai, E., Miyauchi, K., Katayama, Y., & Fukuda, M. (2003). Characterization of the 4-carboxy-4-hydroxy-2-oxoadipate aldolase gene and operon structure of the protocatechuate 4,5-cleavage pathway genes in Sphingomonas paucimobilis SYK-6. *Journal of Bacteriology*, *185*(1), 41–50. https://doi.org/10.1128/JB.185.1.41-50.2003

Harwood, C. S., Nichols, N. N., Kim, M. K., Ditty, J. L., & Parales, R. E. (1994). Identification of the pcaRKF gene cluster from Pseudomonas putida: Involvement in chemotaxis, biodegradation, and transport of 4-hydroxybenzoate. *Journal of Bacteriology*, *176*(21), 6479–6488. https://doi.org/10.1128/jb.176.21.6479-6488.1994

Humphreys, J. M., & Chapple, C. (2002). Rewriting the lignin roadmap. *Current Opinion in Plant Biology*, *5*(3), 224–229. https://doi.org/10.1016/s1369-5266(02)00257-1

Kasai, D., Fujinami, T., Abe, T., Mase, K., Katayama, Y., Fukuda, M., & Masai, E. (2009). Uncovering the protocatechuate 2,3-cleavage pathway genes. *Journal of Bacteriology*, *191*(21), 6758–6768. https://doi.org/10.1128/JB.00840-09

Kaur, B., & Chakraborty, D. (2013). Biotechnological and molecular approaches for vanillin production: A review. *Applied Biochemistry and Biotechnology*, *169*(4), 1353–1372. https://doi.org/10.1007/s12010-012-0066-1

Kondo, S., Sugimura, K., Okamura, Y., & Mase, K. (2016). Stable Chiral Carboxymuconolactone Production from a Lignin-Related Aromatic Compound, Protocatechuic Acid. *Fermentation Technology*, *05*(03). https://doi.org/10.4172/2167-7972.1000135

Lee, H.-N., Shin, W.-S., Seo, S.-Y., Choi, S.-S., Song, J., Kim, J., Park, J.-H., Lee, D., Kim, S. Y., Lee, S. J., Chun, G.-T., & Kim, E.-S. (2018). Corynebacterium Cell Factory Design and Culture Process Optimization for Muconic Acid Biosynthesis. *Scientific Reports*, *8*(1), Article 1. https://doi.org/10.1038/s41598-018-36320-4

Lima, T. C., Ferreira, A. R., Silva, D. F., Lima, E. O., & de Sousa, D. P. (2018). Antifungal activity of cinnamic acid and benzoic acid esters against Candida albicans strains. *Natural Product Research*, *32*(5), 572–575. https://doi.org/10.1080/14786419.2017.1317776

Liu, Y., Cruz-Morales, P., Zargar, A., Belcher, M. S., Pang, B., Englund, E., Dan, Q., Yin, K., & Keasling, J. D. (2021). Biofuels for a sustainable future. *Cell*, *184*(6), 1636–1647. https://doi.org/10.1016/j.cell.2021.01.052

Liu, Z.-H., Li, B.-Z., Yuan, J. S., & Yuan, Y.-J. (2022). Creative biological lignin conversion routes toward lignin valorization. *Trends in Biotechnology*, *40*(12), 1550–1566. https://doi.org/10.1016/j.tibtech.2022.09.014

Lubbers, R. J. M., Dilokpimol, A., Visser, J., Mäkelä, M. R., Hildén, K. S., & de Vries, R. P. (2019). A comparison between the homocyclic aromatic metabolic pathways from plant-derived compounds by bacteria and fungi. *Biotechnology Advances*, *37*(7), 107396. https://doi.org/10.1016/j.biotechadv.2019.05.002

Lubbers, R. J. M., Dilokpimol, A., Peng, M., Visser, J., Mäkelä, M. R., Hildén, K. S., & de Vries, R. P. (2019). Discovery of Novel p-Hydroxybenzoate-m-hydroxylase, Protocatechuate 3,4 Ring-Cleavage Dioxygenase, and Hydroxyquinol 1,2 Ring-Cleavage Dioxygenase from the Filamentous Fungus Aspergillus niger. *ACS Sustainable Chemistry & Engineering*, *7*(23), 19081–19089. https://doi.org/10.1021/acssuschemeng.9b04918

Luo, M., Li, Z., Su, M., Lubbers, R. J. M., Dilokpimol, A., Peng, M., Visser, J., Mäkelä, M. R., Hildén, K. S., & de Vries, R. P. (2019). Discovery of Novel p-Hydroxybenzoate-m-hydroxylase, Protocatechuate 3,4 Ring-Cleavage Dioxygenase, and Hydroxyquinol 1,2 Ring-Cleavage Dioxygenase from the Filamentous Fungus Aspergillus niger. *ACS Sustainable Chemistry & Engineering*, *7*(23), 19081–19089. https://doi.org/10.1021/acssuschemeng.9b04918

MacDonald, D. L., Stanier, R. Y., & Ingraham, J. L. (1954). THE ENZYMATIC FORMATION OF β-CARBOXYMUCONIC ACID. *Journal of Biological Chemistry*, *210*(2), 809–820. https://doi.org/10.1016/S0021-9258(18)65408-0

Mäkelä, M. R., Marinović, M., Nousiainen, P., Liwanag, A. J. M., Benoit, I., Sipilä, J., Hatakka, A., de Vries, R. P., & Hildén, K. S. (2015). Aromatic metabolism of filamentous fungi in relation to the presence of aromatic compounds in plant biomass. *Advances in Applied Microbiology*, *91*, 63–137. https://doi.org/10.1016/bs.aambs.2014.12.001

Martins, T. M., Hartmann, D. O., Planchon, S., Martins, I., Renaut, J., & Silva Pereira, C. (2015). The old 3-oxoadipate pathway revisited: New insights in the catabolism of aromatics in the saprophytic fungus Aspergillus nidulans. *Fungal Genetics and Biology: FG & B*, *74*, 32–44. https://doi.org/10.1016/j.fgb.2014.11.002

Masai, E., Momose, K., Hara, H., Nishikawa, S., Katayama, Y., & Fukuda, M. (2000). Genetic and biochemical characterization of 4-carboxy-2-hydroxymuconate-6-semialdehyde dehydrogenase and its role in the protocatechuate 4,5-cleavage pathway in Sphingomonas paucimobilis SYK-6. *Journal of Bacteriology*, *182*(23), 6651–6658. https://doi.org/10.1128/JB.182.23.6651-6658.2000

Mazur, P., Henzel, W. J., Mattoo, S., & Kozarich, J. W. (1994). 3-Carboxy-cis, cis-muconate lactonizing enzyme from Neurospora crassa: An alternate cycloisomerase motif. *Journal of Bacteriology*, *176*(6), 1718–1728. doi: 10.1128/jb.176.6.1718-1728.1994

McKenna, R., Pugh, S., Thompson, B., & Nielsen, D. R. (2013). Microbial production of the aromatic building-blocks (S)-styrene oxide and (R)-1,2-phenylethanediol from renewable resources. *Biotechnology Journal*, *8*(12), 1465–1475. https://doi.org/10.1002/biot.201300035

Meng, Y., Zhao, M., Yang, M., Zhang, Q., Hao, J., & Meng, Y. (2014). Production and characterization of recombinant glucose oxidase from Aspergillus niger expressed in Pichia pastoris. *Letters in Applied Microbiology*, *58*(4), 393–400. https://doi.org/10.1111/lam.12202

Ni, J., & Tokuda, G. (2013). Lignocellulose-degrading enzymes from termites and their symbiotic microbiota. *Biotechnology Advances*, *31*(6), 838–850. https://doi.org/10.1016/j.biotechadv.2013.04.005

Nogales, J., Canales, A., Jiménez-Barbero, J., García, J. L., & Díaz, E. (2005). Molecular characterization of the gallate dioxygenase from Pseudomonas putida KT2440. The prototype of a new subgroup of extradiol dioxygenases. *The Journal of Biological Chemistry*, *280*(42), 35382–35390. https://doi.org/10.1074/jbc.M502585200

Ornston, L. N. (1966). The conversion of catechol and protocatechuate to beta-ketoadipate by Pseudomonas putida. II. Enzymes of the protocatechuate pathway. *The Journal of Biological Chemistry*, *241*(16), 3787–3794.

Pel, H. J., de Winde, J. H., Archer, D. B., Dyer, P. S., Hofmann, G., Schaap, P. J., Turner, G., de Vries, R. P., Albang, R., Albermann, K., Andersen, M. R., Bendtsen, J. D., Benen, J. A. E., van den Berg, M., Breestraat, S., Caddick, M. X., Contreras, R., Cornell, M., Coutinho, P. M., … Stam, H. (2007). Genome sequencing and analysis of the versatile cell factory Aspergillus niger CBS 513.88. *Nature Biotechnology*, *25*(2), 221–231. https://doi.org/10.1038/nbt1282

Priyanka, U., & Lens, P. N. L. (2022). Enhanced removal of hydrocarbons BTX by light-driven Aspergillus niger ZnS nanobiohybrids. *Enzyme and Microbial Technology*, *157*, 110020. https://doi.org/10.1016/j.enzmictec.2022.110020

Radhika, N. L., Sachdeva, S., & Kumar, M. (2021). Microbe assisted depolymerization of lignin rich waste and its conversion to gaseous biofuel. *Journal of Environmental Management*, *300*, 113684. https://doi.org/10.1016/j.jenvman.2021.113684

Ragauskas, A. J., Beckham, G. T., Biddy, M. J., Chandra, R., Chen, F., Davis, M. F., Davison, B. H., Dixon, R. A., Gilna, P., Keller, M., Langan, P., Naskar, A. K., Saddler, J. N., Tschaplinski, T. J., Tuskan, G. A., & Wyman, C. E. (2014). Lignin valorization: Improving lignin processing in the biorefinery. *Science (New York, N.Y.)*, *344*(6185), 1246843. https://doi.org/10.1126/science.1246843

Ren, N.-Q., Zhao, L., Chen, C., Guo, W.-Q., & Cao, G.-L. (2016). A review on bioconversion of lignocellulosic biomass to H2: Key challenges and new insights. *Bioresource Technology*, *215*, 92–99. https://doi.org/10.1016/j.biortech.2016.03.124

Sahasrabudhe, S. R., & Modi, V. V. (1985). Hydroxylation of benzoate and its chlorinated derivatives in Aspergillus niger. *Biochemistry International*, *10*(4), 525–529.

Schutyser, W., Renders, T., Van den Bosch, S., Koelewijn, S.-F., Beckham, G. T., & Sels, B. F. (2018). Chemicals from lignin: An interplay of lignocellulose fractionation, depolymerisation, and upgrading. *Chemical Society Reviews*, *47*(3), 852–908. https://doi.org/10.1039/c7cs00566k

Semana, P., & Powlowski, J. (2019). Four Aromatic Intradiol Ring Cleavage Dioxygenases from Aspergillus niger. *Applied and Environmental Microbiology*, *85*(23), e01786-19. https://doi.org/10.1128/AEM.01786-19

Sgro, M., Chow, N., Olyaei, F., Arentshorst, M., Geoffrion, N., Ram, A. F. J., Powlowski, J., & Tsang, A. (2023). Functional analysis of the protocatechuate branch of the β-ketoadipate pathway in Aspergillus niger. *The Journal of Biological Chemistry*, *299*(8), 105003. https://doi.org/10.1016/j.jbc.2023.105003

Shailubhai, K., Somayaji, R., Rao, N. N., & Modi, V. V. (1983). Metabolism of resorcinol and salicylate in Aspergillus niger. *Experientia*, *39*(1), 70–72. https://doi.org/10.1007/BF01960634

Singhvi, M., & Kim, B. S. (2021). Lignin valorization using biological approach. *Biotechnology and Applied Biochemistry*, *68*(3), 459–468. https://doi.org/10.1002/bab.1995

Song, L., Ouedraogo, J.-P., Kolbusz, M., Nguyen, T. T. M., & Tsang, A. (2018). Efficient genome editing using tRNA promoter-driven CRISPR/Cas9 gRNA in Aspergillus niger. *PloS One*, *13*(8), e0202868. https://doi.org/10.1371/journal.pone.0202868

Thatcher, D. R., & Cain, R. B. (1970). Metabolism of aromatic compounds by fungi: Conversion of beta-carboxymuconolactone into 3-oxoadipate in Aspergillus niger. *Biochemical Journal*, *120*(4), 28P-29P. https://doi.org/10.1042/bj1200028Pb

Thatcher, D. R., & Cain, R. B. (1974). Metabolism of Aromatic Compounds by Fungi. 1. Purification and Physical Properties of 3-Carboxy-cis-cis-muconate Cyclase from Aspergillus niger. *European Journal of Biochemistry*, *48*(2), 549–556. https://doi.org/10.1111/j.1432-1033.1974.tb03796.x

Thatcher, D. R., & Cain, R. B. (1974). Metabolism of aromatic compounds by fungi. 2. Subunit structure of the 3-carboxy-cis-cis-muconate cyclase of Aspergillus niger. *European Journal of Biochemistry*, *48*(2), 557–562. https://doi.org/10.1111/j.1432-1033.1974.tb03797.x

Thatcher, D. R., & Cain, R. B. (1975). Metabolism of Aromatic Compounds by Fungi: Kinetic Properties and Mechanism of 3-Carboxy- *cis, cis* -muconate Cyclase from *Aspergillus niger*. *European Journal of Biochemistry*, *56*(1), 193–204. https://doi.org/10.1111/j.1432-1033.1975.tb02222.x

Tong, Z., Zheng, X., Tong, Y., Shi, Y.-C., & Sun, J. (2019). Systems metabolic engineering for citric acid production by Aspergillus niger in the post-genomic era. *Microbial Cell Factories*, *18*, 28. https://doi.org/10.1186/s12934-019-1064-6

Tsuge, Y., Kawaguchi, H., Sasaki, K., & Kondo, A. (2016). Engineering cell factories for producing building block chemicals for bio-polymer synthesis. *Microbial Cell Factories*, *15*, 19. https://doi.org/10.1186/s12934-016-0411-0

Wu, W., Dutta, T., Varman, A. M., Eudes, A., Manalansan, B., Loqué, D., & Singh, S. (2017). Lignin Valorization: Two Hybrid Biochemical Routes for the Conversion of Polymeric Lignin into Value-added Chemicals. *Scientific Reports*, *7*(1), 8420. https://doi.org/10.1038/s41598-017-07895-1

Xu, C., Arancon, R. A. D., Labidi, J., & Luque, R. (2014). Lignin depolymerisation strategies: Towards valuable chemicals and fuels. *Chemical Society Reviews*, *43*(22), 7485–7500. https://doi.org/10.1039/c4cs00235k

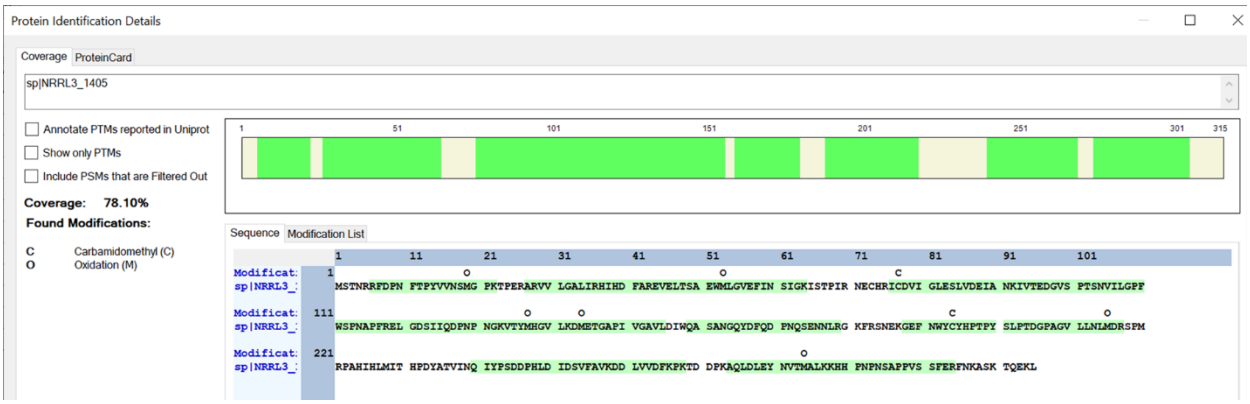
Xu, R., Zhang, K., Liu, P., Han, H., Zhao, S., Kakade, A., Khan, A., Du, D., & Li, X. (2018). Lignin depolymerization and utilization by bacteria. *Bioresource Technology*, *269*, 557–566. https://doi.org/10.1016/j.biortech.2018.08.118

Yamanashi, T., Kim, S.-Y., Hara, H., & Funa, N. (2015). In vitro reconstitution of the catabolic reactions catalyzed by PcaHG, PcaB, and PcaL: The protocatechuate branch of the β-ketoadipate pathway in Rhodococcus jostii RHA1. *Bioscience, Biotechnology, and Biochemistry*, *79*(5), 830–835. https://doi.org/10.1080/09168451.2014.993915

Zhang, L., Xi, G., Zhang, J., Yu, H., & Wang, X. (2017). Efficient catalytic system for the direct transformation of lignocellulosic biomass to furfural and 5-hydroxymethylfurfural. *Bioresource Technology*, *224*, 656–661. https://doi.org/10.1016/j.biortech.2016.11.097

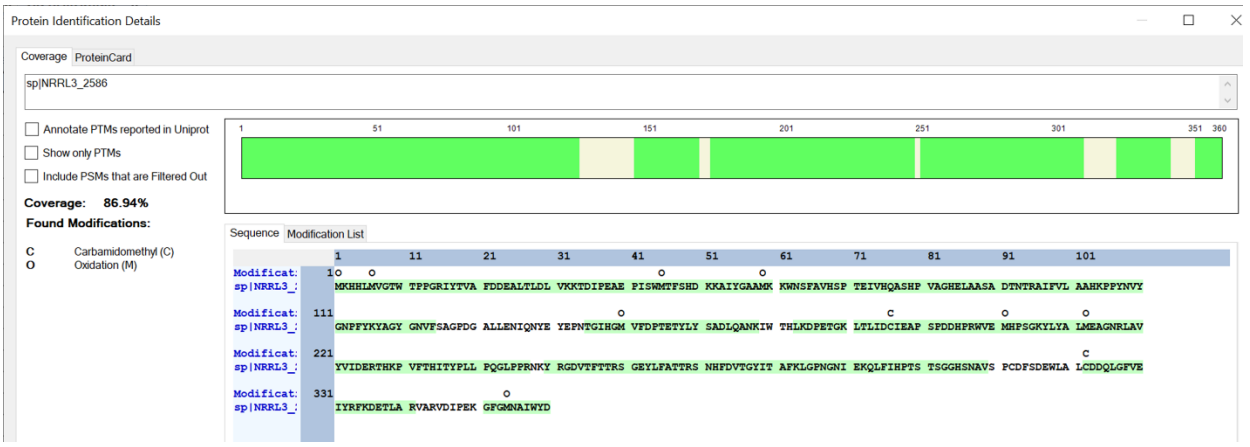
Zhu, J. Y., & Pan, X. J. (2010). Woody biomass pretreatment for cellulosic ethanol production: Technology and energy consumption evaluation. *Bioresource Technology*, *101*(13), 4992–5002. https://doi.org/10.1016/j.biortech.2009.11.007

6. Appendices



Description	Sum PEP Score	Coverage	# Peptides	# Unique Peptides	# AAs	MW [kDa]	calc. pI	Score Sequest HT	Abundance	Found in Sample:
sp|TRYP_PIG	106.705	46	32	32	231	24.4	7.18	316.59	1.04E+09	High
sp|NRRL3_1405	241.495	78	70	70	315	35.3	5.76	444.73	8.49E+08	High
sp|K2C1_HUMAN	117.019	54	37	33	643	65.8	8.12	200.58	43133338	High
sp|K1C9_HUMAN	84.039	48	31	30	623	62.1	5.3	130.68	27103355	High
sp|K1C10_HUMAN	57.029	38	22	16	593	59.5	5.21	87.55	16296218	High
sp|K22E_HUMAN	43.507	31	18	14	645	65.8	8	56.72	3051003	High
sp|K1C15_SHEEP	12.916	13	6	1	453	48.7	4.79	19.32	1086904	High
sp|NRRL3_837	9.719	18	4	4	209	22.6	5.08	8.25	939028.4	High
sp|KRHB4_HUMAN	3.209	2	2	1	600	64.9	7.74	2.13	703645.6	High
sp|ALBU_BOVIN	13.86	13	7	7	607	69.2	6.18	19.2	363418.8	High
sp|RS27A_HUMAN	1.2	6	1	1	156	18	9.64	0	80393.82	High
sp|BGAL_ECOLI	1.68	1	1	1	1024	116.4	5.5	2.51	50986	High
sp|CAH2_BOVIN	1.517	5	1	1	260	29.1	6.92	1.67	11745.01	High

Figure S1. Peptide mass identification following in-gel tryptic digestion of NRRL3_01405 expressed in *Escherichia coli* (BL21). (top) Peptide coverage map of NRRL3_01405 and observed modifications. **(bottom)** Alignment of mass spectrometry data with the SEQUEST search engine using the Thermo Proteome Discoverer software (v2.4) against the sequences of the candidate protocatechuate catabolic enzymes, the UniProt *Escherichia coli* proteome database (Uniprot UP000002032) and cRAP protein sequences (https://www.thegpm.org/crap/).



Checked	Protein FD	Description	Sum PEP Score	Coverage	# Peptides	# Unique Peptides	# AAs	MW [kDa]	calc. pI	Score Sequest HT	Abundance	Found in S
FALSE	High	sp|NRRL3_2586	410.14	87	110	110	360	40.4	6.2	860.33	1.02E+09	High
FALSE	High	sp|TRYP_PIG	122.259	51	42	42	231	24.4	7.18	342.25	1E+09	High
FALSE	High	sp|K2C1_HUMAN	59.679	37	25	20	643	65.8	8.12	93.08	10558673	High
FALSE	High	sp|K1C9_HUMAN	60.071	36	20	19	623	62.1	5.3	98.02	7432905	High
FALSE	High	sp|K1C10_HUMAN	39.071	29	17	13	593	59.5	5.21	48.44	5877817	High
FALSE	High	sp|K22E_HUMAN	48.73	30	20	15	645	65.8	8	62.22	1931666	High
FALSE	Medium	sp|PDGFB_HUMAN	0.839	3	1	1	241	27.3	9.16	0	1253146	High
FALSE	Medium	sp|TRFL_HUMAN	0.692	1	1	1	710	78.1	8.12	0	1123250	High
FALSE	High	sp|NRRL3_1409	15.428	18	9	9	546	60.9	5.54	18.98	1121266	High
FALSE	High	sp|K1H8_HUMAN	2.844	3	2	1	456	50.5	4.84	1.84	326501.7	High
FALSE	Medium	KKA1_ECOLX	0.734	4	1	1	271	30.9	5.39	0	315738	High

Figure S2. Peptide mass identification following in-gel tryptic digestion of NRRL3_02586 expressed in *Escherichia coli* (BL21). (top) Peptide coverage map of NRRL3_02586 and observed modifications. **(bottom)** Alignment of mass spectrometry data with the SEQUEST search engine using the Thermo Proteome Discoverer software (v2.4) against the sequences of the candidate protocatechuate catabolic enzymes, the UniProt *Escherichia coli* proteome database (Uniprot UP000002032) and cRAP protein sequences (https://www.thegpm.org/crap/).

Checked	Description	Sum PEP Score	Coverage [%]	# Peptides	# Unique Peptides	# AAs	MW [kDa]	calc. pI	Score Sequest HT	Abundance	Found in Sample
FALSE	sp|NRRL3_1409	663.625	92	174	174	546	60.9	5.54	1331.79	1.19E+10	High
FALSE	sp|TRYP_PIG	129.571	46	34	34	231	24.4	7.18	176.51	1.3E+09	High
FALSE	sp|K2C1_HUMAN	95.011	43	29	24	643	65.8	8.12	127.7	37746858	High
FALSE	sp|LEP_HUMAN	1.066	16	1	1	167	18.6	6.37	0	29099222	High
FALSE	sp|K1C9_HUMAN	102.906	54	28	27	623	62.1	5.3	132.96	21787299	High
FALSE	sp|K1C10_HUMAN	32.41	21	13	9	593	59.5	5.21	44.87	8527692	High
FALSE	sp|K22E_HUMAN	48.686	26	17	12	645	65.8	8	61.64	5054526	High
FALSE	sp|K1C15_SHEEP	12.889	12	6	3	453	48.7	4.79	17.12	3394983	High
FALSE	sp|TRFL_HUMAN	0.903	1	1	1	710	78.1	8.12	0	2386445	High
FALSE	sp|KRHB4_HUMAN	3.376	2	2	1	600	64.9	7.74	6.34	977727.4	High
FALSE	sp|NRRL3_1526	9.787	13	4	4	413	43.2	6.89	10.17	254234.5	High
FALSE	sp|ALBU_BOVIN	4.194	4	2	2	607	69.2	6.18	5.78	231182.1	High
FALSE	sp|CATA_HUMAN	1.009	5	1	1	526	59.6	7.39	0	181214.1	High
FALSE	sp|CAH2_BOVIN	7.254	8	2	2	260	29.1	6.92	6.73	149453.9	High
FALSE	sp|NRRL3_2586	6.177	11	2	2	360	40.4	6.2	5.5	87343.45	High
FALSE	sp|NRRL3_1405	4.765	10	2	2	315	35.3	5.76	2.59	74403.93	High
FALSE	sp|OVAL_CHICK	5.587	4	1	1	386	42.9	5.29	8.91	48595.17	High

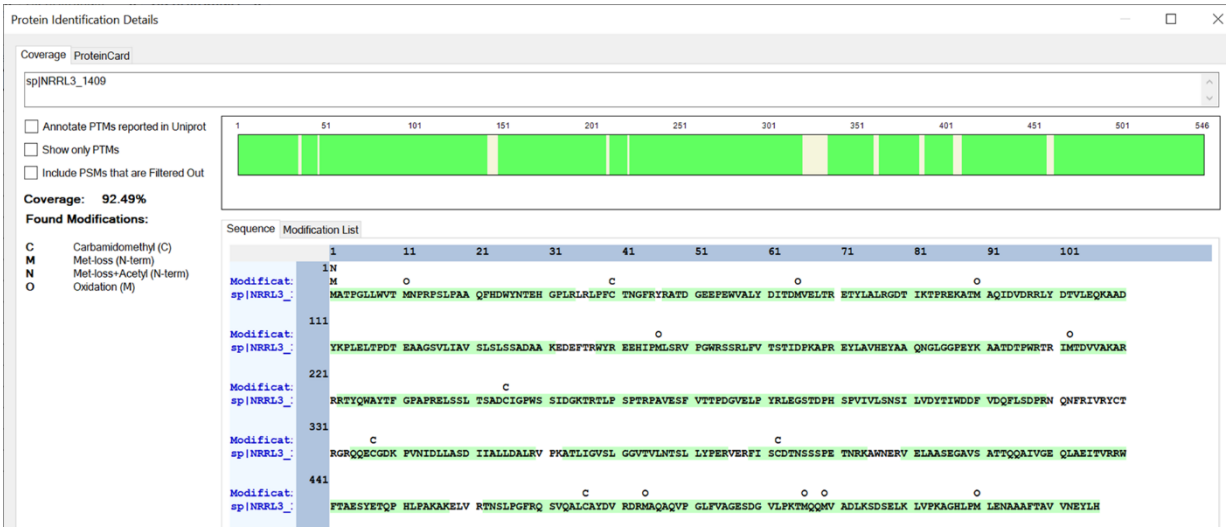
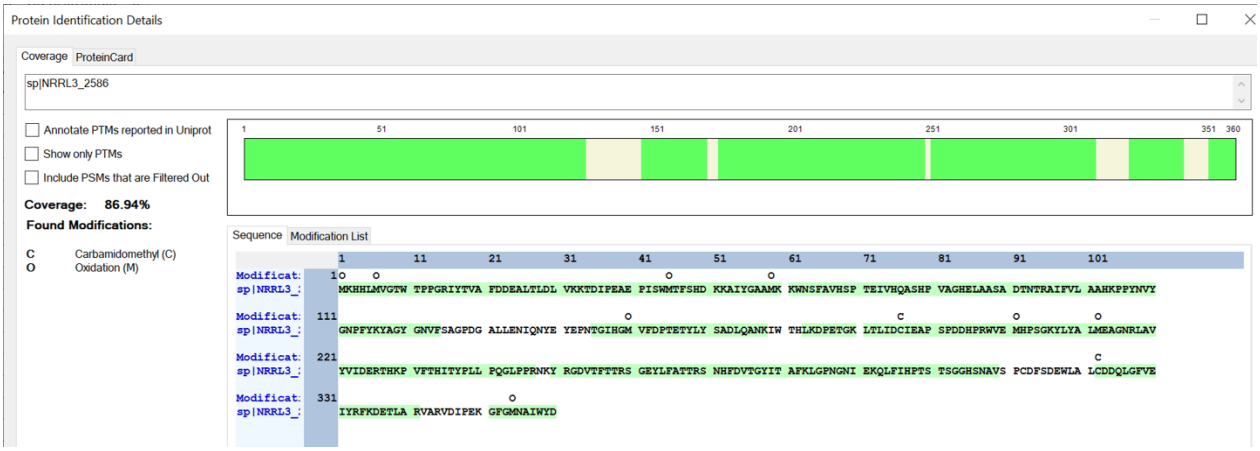
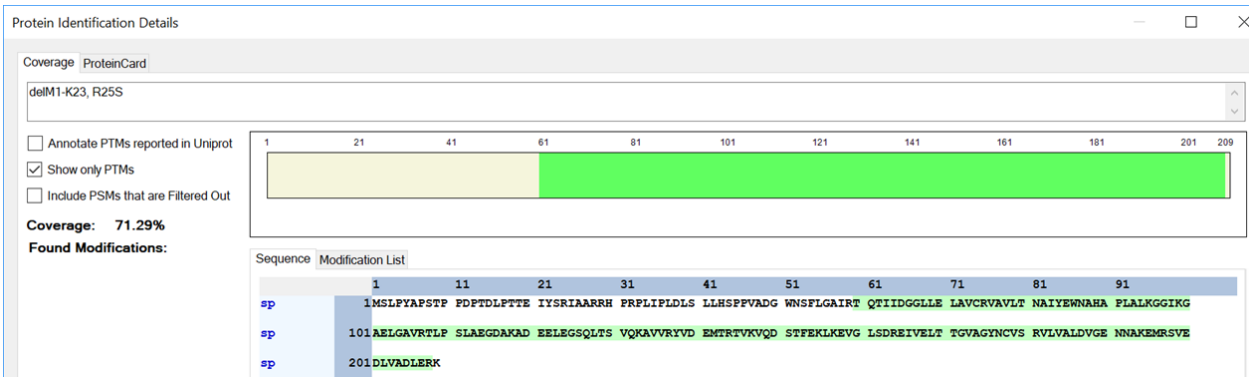


Figure S3. Peptide mass identification following in-gel tryptic digestion of NRRL3_01409 expressed in *Escherichia coli* (BL21). (top) Peptide coverage map of NRRL3_01409 and observed modifications. **(bottom)** Alignment of mass spectrometry data with the SEQUEST search engine using the Thermo Proteome Discoverer software (v2.4) against the sequences of the candidate protocatechuate catabolic enzymes, the UniProt *Escherichia coli* proteome database (Uniprot UP000002032) and cRAP protein sequences (https://www.thegpm.org/crap/).



Checked	Protein FDR Confidence	Description	Sum PEP Sc	Coverage	# Peptides	# Unique Peptides	# AAs	MW [kDa]	calc. pI	Score Sequest HT	Abundance	Found in Sample:
FALSE	High	sp|NRRL3_837	325.519	92	93	93	209	22.6	5.08	830.31	4.52E+09	High
FALSE	High	sp|TRYP_PIG	134.797	50	38	38	231	24.4	7.18	302.62	1.24E+09	High
FALSE	High	sp|K2C1_HUMAN	79.739	40	30	25	643	65.8	8.12	119.6	13295523	High
FALSE	High	sp|K1C9_HUMAN	76.981	42	25	24	623	62.1	5.3	114.89	9044662	High
FALSE	High	sp|NRRL3_1409	70.827	53	21	21	546	60.9	5.54	154.35	6126493	High
FALSE	High	sp|K1C10_HUMAN	48.995	34	19	15	593	59.5	5.21	63.97	6036245	High
FALSE	High	sp|K22E_HUMAN	43.205	28	16	11	645	65.8	8	70.5	1745418	High
FALSE	High	sp|ANXA5_HUMAN	1.426	3	1	1	319	35.8	5.05	0	1334628	High
FALSE	High	sp|BGAL_ECOLI	1.289	4	1	1	1024	116.4	5.5	0	768690.4	High
FALSE	High	sp|CO5_HUMAN	0.909	2	1	1	1676	188.2	6.52	0	454529.9	High
FALSE	High	sp|K2M2_SHEEP	0.904	2	1	1	491	53.6	5.57	0	195757.3	High
FALSE	High	sp|KRB2B_SHEEP	1.013	12	1	1	156	16	6.06	0	173315.9	High
FALSE	Medium	sp|TRFE_HUMAN	0.664	1	1	1	698	77	7.12	0	69189.3	High
FALSE	High	sp|NRRL3_1526	3.549	10	3	3	413	43.2	6.89	8.49	62184.58	High
FALSE	High	sp|NRRL3_1405	1.362	5	1	1	315	35.3	5.76	2.01	12595.22	High
FALSE	Medium	sp|ALBU_BOVIN	0.7	2	1	1	607	69.2	6.18	2.8	12434.84	High

Figure S4. Peptide mass identification following in-gel tryptic digestion of NRRL3_00837 expressed in *Escherichia coli* (BL21). (top) Peptide coverage map of NRRL3_00837 and observed modifications. **(bottom)** Alignment of mass spectrometry data with the SEQUEST search engine using the Thermo Proteome Discoverer software (v2.4) against the sequences of the candidate protocatechuate catabolic enzymes, the UniProt *Escherichia coli* proteome database (Uniprot UP000002032) and cRAP protein sequences (https://www.thegpm.org/crap/).



Description	Sum PEP Score	Coverage [%]	# Peptides	# Unique Peptides	# AAs	MW [kDa]	calc. pI	Score Sequest HT:	Abundance:	Found in Sample:
NRRL3_00837 ΔM1-K23, R25S	72.148	71	18	18	209	22.6	5.08	524.18	2.426E+09	High
sp|TRYP_PIG	33.41	35	8	8	231	24.4	7.18	142.83	2.06E+09	High
sp|K2C1_HUMAN	91.315	49	32	28	643	65.8	8.12	199.77	41162586	High
sp|K1C9_HUMAN	69.192	45	23	22	623	62.1	5.3	140.11	30707426	High
sp|K1C10_HUMAN	38.64	28	20	15	593	59.5	5.21	79.87	20455679	High
sp|K22E_HUMAN	74.699	49	28	24	645	65.8	8	155.02	13097877	High
peptidyl-prolyl cis-trans isomerase [(	7.487	39	4	4	196	20.8	5.05	19.96	8292050.1	High
GTP cyclohydrolase 1 [OS=Escherich	17.277	28	7	7	222	24.8	7.33	17.03	1852534.4	High
LPP repeat-containing protein [OS=E	8.386	33	2	2	78	8.3	9.25	10.64	582139.27	High
transcription termination factor Rho	2.685	5	2	2	419	47	7.25	4.84	357288.2	High
30S ribosomal protein S3 [OS=Escher	3.954	8	2	2	233	26	10.27	5.31	339486.33	High
transcriptional regulator, DeoR famil	3.143	10	2	2	243	27.4	7.9	2.7	233197.69	High
pyridoxine/pyridoxamine 5'-phospha	1.64	10	1	1	218	25.5	9.16	2.39	210748.48	High

Figure S5. Peptide mass identification following in-gel tryptic digestion of NRRL3_00837 ΔM1-K23, R25S expressed in *Escherichia coli* (BL21). (top) Peptide coverage map of NRRL3_00837 ΔM1-K23, R25S and observed modifications. **(bottom)** Alignment of mass spectrometry data with the SEQUEST search engine using the Thermo Proteome Discoverer software (v2.4) against the sequences of the candidate protocatechuate catabolic enzymes, the UniProt *Escherichia coli* proteome database (Uniprot UP000002032) and cRAP protein sequences (https://www.thegpm.org/crap/).