Characteristics of Phenylacrylic Acid Decarboxylase (PAD1) and Ferulic Acid Decarboxylase (FDC) from *Saccharomyces cerevisiae*

Thesis by: David Boyer
Thesis Advisor: E. Neil G. Marsh

In partial fulfillment of the degree
Honors Bachelor of Science
(Biochemistry)
2014

Department of Chemistry
University of Michigan
Ann Arbor, MI

Table of Contents

Abstract

The use of enzymes in organic synthesis offers many advantages over traditional catalysts including high catalytic turnover, mild reaction conditions, and specificity. The increasing interest in environmentally friendly production of fuels and chemicals has spurred research and development of effective and versatile biocatalysts either to be used in organic synthesis or in living systems. Due to the potential for decarboxylases to produce optically pure products under mild reaction conditions, we have sought to further investigate phenylacrylic acid decarboxylase 1 (PAD1) and ferulic acid decarboxylase (FDC) from *Saccharomyces cerevisiae.* These enzymes have been shown to be responsible for the decarboxylation of antimicrobial phenylacrylic acids; however, their specific roles in these reactions remain unknown. We have over-expressed and purified PAD1 and FDC from *E. coli* in order to elucidate the function of each enzyme. Traditional biochemical, kinetic, and mechanistic techniques were employed to characterize these two enzymes. We have shown that PAD1 is a flavo-protein that utilizes FMN as a cofactor, but does not function as a decarboxylase. Instead, PAD1 catalyzes the formation of a novel, diffusible cofactor required for decarboxylase activity of FDC. Co-expression of PAD1 and FDC results in FDC with high cofactor occupancy and ability to decarboxylate a variety of phenylacrylic acids with a k_{cat} ranging from 1.4 – 4.6 s^{-1}. *In silico* experiments were also performed to supplement *in vitro* mechanistic studies. We find that a revised mechanism involving resonance contributions from the *para* substituents on the benzyl ring and the presence of a counter-ion (Mg^{2+}) were necessary for the construction of a reasonable reaction coordinate diagram for ferulic acid. The decarboxylation of ferulic acid was found to likely proceed under nucleophilic, rather than cationic, conditions.

Chapter 1:

Introduction and Background

Introduction

Traditional chemical synthesis and energies production has presented numerous problems in our society such as environmental stress, economic concerns, and reliance on quickly depleting fossil fuel reserves. One of the most prominent methods to solve these problems is the *in vivo* production of non-natural metabolites in microorganisms via a combination of metabolic engineering, systems biology, and synthetic biology. This process often requires elements of bioprospecting – identifying enzymes capable of producing metabolites of interest from both genomic and proteomic databases – and the introduction of heterologous metabolic pathways in a desired host organism (e.g., *Escherichia coli*). This method is beneficial as it opens a vast resource of enzymatic chemistry capable of producing a wide variety of valuable metabolites.

The use of natural catalysts (i.e., enzymes) for chemical synthesis is not new and has been a trend in organic chemistry for some time. Biocatalysts have been used for more than one hundred years, employed as either whole cells, cell organelles, or isolated enzymes[1,2]. They present many advantages over traditional chemical synthesis. First, enzymes are very efficient catalysts; typically enzymes increase the rates of chemical reactions which they catalyze by 10^8-10^{12}, which can be far above the rates that chemical catalysts are capable of achieving[3]. Enzymes are also environmentally friendly because, unlike heavy metals, they are composed of amino acids and are completely biodegradable. They are also able to act under mild conditions, typically at biological pH, and a temperature range of 20 – 40 °C. This permits the minimization

of undesired side-reactions (e.g., decomposition, isomerization, racemization) and reduction of the energy burden of traditional synthesis. Enzymes can also catalyze almost every type of organic reaction, and even some that are not found in the organic synthesis repertoire such as the selective functionalization of nonactivated groups in organic molecules[4].

The repertoire of enzymatic chemistry is only expected to increase due to the large amount of genomic data that is becoming available. A recent publication described over one *million* novel proteins, known only as open reading frames encoded in genomic DNA[5]. These proteins will be the design parts of many future biological systems, enabling sustainable solutions to many of society's problems. However, enzymatic pathways that often look promising from preliminary genomic data are poorly understood and, when transferred into a host organism, do not perform to optimal levels or even fail to function as desired[6]. Therefore, detailed biochemical studies of potentially useful enzymatic reactions are highly important when attempting the construction of new, nonnative metabolic pathways.

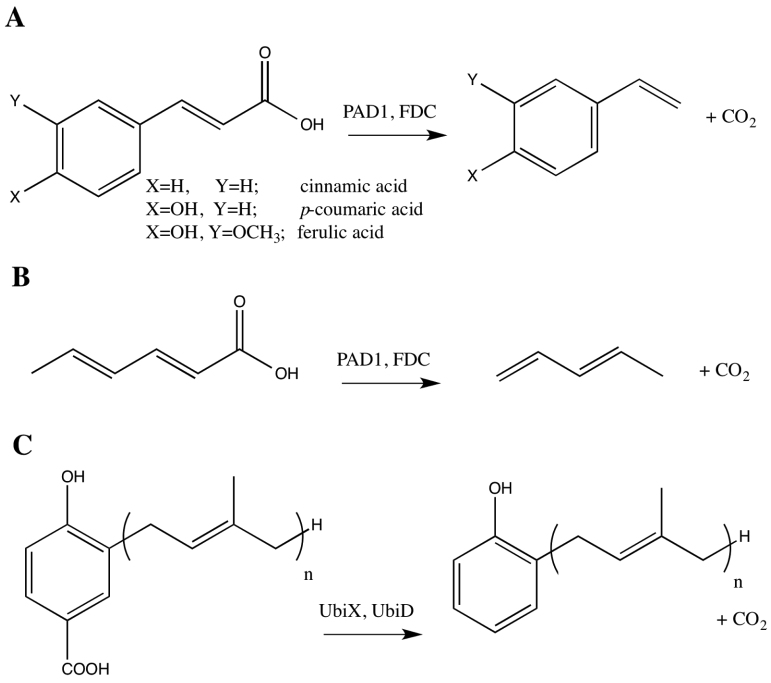
Enzyme-catalyzed decarboxylations are one such class of reactions that are becoming increasingly important in both chemical synthesis and fermentation processes. However, decarboxylation reactions are characterized by high-energy barriers because the transition state involves a buildup of negative charge on the α-carbon. Nature has therefore evolved decarboxylases to overcome this high-energy barrier through the use of cofactors such as pyridoxal phosphate and thiamin pyrophosphate that serve as electron sinks, and also Lewis acidic metal ions[7,8,9].

Decarboxylases are of interest for their potential use as industrial catalysts in the synthesis of optically pure intermediates for organic synthesis under mild reaction conditions[10].

They have also been used in microbial fermentation processes to produce short-chain alcohols and in the conversion of overexpressed amino acids to chemical feedstocks such as styrene and acylamide[11,12,13]. The investigation of decarboxylases is especially relevant to the search for effective yet sustainable methods for chemical synthesis as many valuable chemicals can be generated from the phenylpropanoid pathway in higher plants[14]. Using plants (i.e., biomass) as a source of cheap chemical feedstocks to produce commodity chemicals is becoming increasingly desired. Therefore, the extraction of certain chemicals from plant cell cultures[15] or from the waste products of biomass used in bioenergy processes is an important goal in sustainable chemistry[16]. Recent research has demonstrated that the production of various phenylpropanoids can also be achieved in engineered microorganisms, making the microbial production of a variety of phenylpropanoids a realistic possibility[17]. Combining phenylpropanoid producing strains with strains containing phenylacrylic acid decarboxylation activity could yield many valuable chemicals, including biofuels, monomers, and industrial flavorings. A recent example of particular interest is the use of ferulic acid decarboxylase (FDC) from *Saccharomyces cerevisiae* in conjunction with phenylalanine ammonia lyase (PAL) in an engineered *E. coli* strain that converts endogenously produced phenylalanine into styrene[18].

FDC is found in various strains of yeast that detoxify antimicrobial compounds commonly used as food preservatives and flavorings, such as sorbic acid and cinnamic acid (phenylacrylic acid), by decarboxylation to yield CO_2 and volatile products pentadiene and styrene[19,20] (Scheme 1 A, B). Initial studies into the nature of decarboxylation of phenylacrylic acids in *S. cerevisiae* have begun to elucidate the basis for these biosynthetic reactions. Originally, a genetic study characterized phenylacrylic acid decarboxylase 1 (PAD1) as

conferring cinnamic acid resistance to *S. cerevisiae*[21]. PAD1 was cloned into a cinnamic acid-sensitive mutant strain of yeast, restoring cinnamic acid resistance and phenylacrylic acid decarboxylation. However, the product of the *pad1* gene was not shown to possess decarboxylation activity, despite its name. Importantly, Stratford et al. identified the gene product of *pad1* from *S. cerevisiae* as being responsible for decarboxylation of sorbic acid, ferulic acid, coumaric acid, and cinnamic acid into their corresponding products, labeling PAD1 as a unique decarboxylase capable of accepting aromatic and aliphatic substrates. This, among other examples in the literature, confuse the role of PAD1. The contents of this thesis will help clarify the role that PAD1 plays in decarboxylating phenylacrylic acids. In a separate study, the FDC gene was introduced into sake yeast, which lacks the ability to decarboxylate ferulic acid, allowing the yeast to decarboxylate ferulic acid[22]. More recently, the nature of the decarboxylation of phenylacrylic acids has been more fully investigated. It has been shown that FDC needs the gene product of the *pad1* gene from *S. cerevisiae* for decarboxylation activity[20]. However, FDC was not shown to contain decarboxylation activity unambiguously. It was not shown until recently that both PAD1 and FDC are required for phenylacrylic acid decarboxylation activity in yeast[20]. The ability of *S. cerevisiae* to decarboxylate various phenylacrylic acids was investigated using *fdc, pad1,* and *fdc/pad1* knockout strains. It was found that ferulic acid, *p*-coumaric acid, and *trans*-cinnamic acid could not be decarboxylated into their respective products without both the *pad1* and *fdc* genes present. However, the nature of the interaction between PAD1 and FDC has not been evaluated biochemically and the work described herein identifies how these proteins work cooperatively to achieve decarboxylation activity.



Scheme 1: (A and B) Reactions shown to have been catalyzed by PAD1, FDC in yeast and recombinant *E. coli*. (C) Intermediate step in ubiquinone synthesis catalyzed by UbiX, UbiD in bacteria.

Interestingly, PAD1 and FDC are thought to be homologs of UbiX/UbiD, two enzymes involved in an intermediate step in the biosynthesis of ubiquinone in a wide range of bacteria. They catalyze the decarboxylation of the aromatic intermediate 4-hydroxy-3-octaprenylbenzoic acid to 2-octaprenylphenol[23,24] (Scheme 1C). It is important to note that previous literature confused the roles of UbiD/UbiX as isofunctional[24]. However, the work described herein clarifies the interaction between these two enzymes. Crystal structures of related decarboxylases and sequence comparisons suggest that UbiX and PAD1, and by extension UbiD and FDC, may

be flavo-proteins. However, the function of the flavin prosthetic group in these enzymes remains unknown.

The nature of this study is to further characterize FDC/PAD1 by elucidating their structural properties, biochemical properties, and how their interaction brings about the decarboxylation of a range of phenylacrylic acids. Preliminary studies into the mechanism of decarboxylation by FDC are also presented. As a result of this work, conclusions can also be made about the role of UbiX/UbiD in bacterial ubiquinone biosynthesis.

Background

It is important to note here that three key experiments illuminating the interaction between PAD1 and FDC were performed prior to my work on this project. First, after initial expression and purification of FDC in *E. coli* BL21 by standard methods, it was noted that after dialysis of FDC against buffer (MW cutoff = 3500 Da), decarboxylation of *trans*-cinnamic acid into styrene was nearly undetectable (Figure 1). However, upon addition of BL21 cell lysate, decarboxylase activity was fully restored. This suggested that a low-molecular weight cofactor was needed for FDC decarboxylation activity. In order to confirm this hypothesis, BL21 cell lysate was similarly dialyzed and upon addition to dialyzed FDC, activity of FDC was almost undetectable.

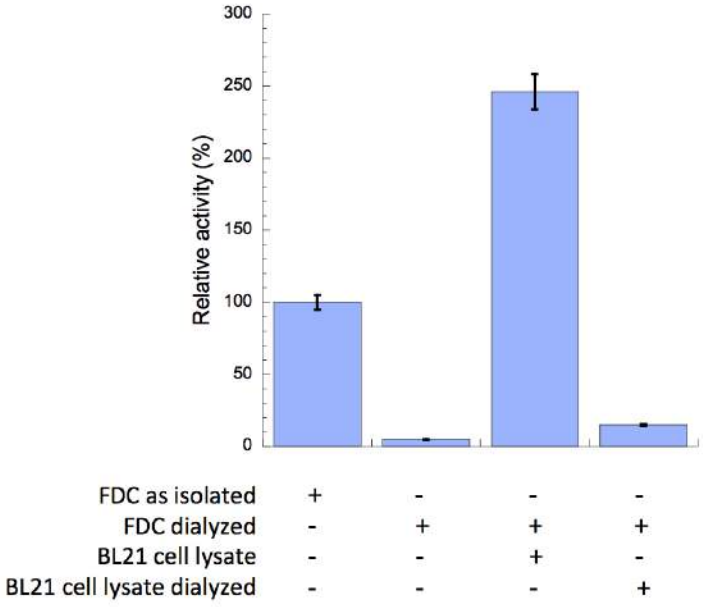


Figure 1: Relative activities of FDC as isolated from *E. coli* BL21, dialyzed FDC, FDC in BL21 cell lysate, dialyzed FDC in dialyzed BL21. (Adapted from Lin, et al. ACS Chem. Bio. 2014)

Due to PAD1's sequence similarity to UbiX, it was hypothesized that PAD1 might make a cofactor that is used by FDC. It was thought that UbiX might substitute for PAD1 in cells expressing FDC to give decarboxylation activity. Therefore, an *E. coli* BL21 Δ*ubiX* strain was created in order to investigate the possible isofunctionality of UbiX and PAD1. This strain resulted in FDC with no activity and is hence named apo-FDC. Apo-FDC was reacted in a solution containing Δ*ubiX* cell lysate, and no activity was discerned (Figure 2). Alternatively, when tPAD1 (truncated PAD1 without mitochondrial localization sequence) was added to apo-FDC, a small amount of decarboxylation activity was restored. This suggests that tPAD1 retains a small amount of the cofactor after its expression and purification and is able to release it to FDC. Lastly, when tPAD1 and Δ*ubiX* cell lysate were added to apo-FDC, decarboxylation activity was greatly increased.

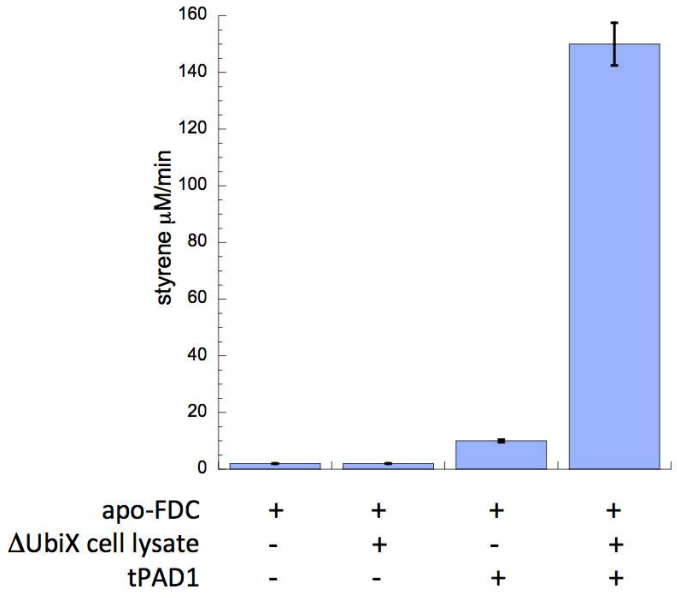


Figure 2: Specific activities of apo-FDC, apo-FDC in Δ*ubiX* cell lysate, apo-FDC with tPAD1 purified from *E. coli* BL21 (DE3), apo-FDC in Δ*ubiX* cell lysate with tPAD1. (Adapted from Lin, et al. ACS Chem. Bio. 2014).

In order to definitively confirm that PAD1/UbiX creates a small cofactor necessary for FDC activity, and that PAD1 itself does not participate directly in the decarboxylation reaction, the following experiment was performed (Figure 3). A dialysis chamber was set up with FDC and 6.7 mM of *trans*-cinnamic acid on one side while the other side had either BL21 cell lysate, Δ*ubiX* cell lysate, or Δ*ubiX* cell lysate with tPAD1. Production of styrene from *trans*-cinnamic acid was monitored over time. The results indicate that although BL21 cell lysate and tPAD1 were able to restore some decarboxylation activity, Δ*ubiX* cell lysate with tPAD1 allowed restoration of maximal activity. This suggested that PAD1 makes a diffusible low-molecular weight cofactor from endogenous elements of *E. coli* that is necessary for FDC to decarboxylate phenylacrylic acids. It also indicates that UbiX makes the same cofactor necessary for activity of FDC. In tandem with the information that UbiD and UbiX are necessary for the decarboxylation of 4-hydroxy-3-octaprenylbenzoic acid to 2-octaprenylphenol in ubiquinone biosynthesis[22], UbiD most likely makes the same cofactor as PAD1 that is required by UbiX to perform decarboxylation of the aromatic 4-hydroxy-3-octaprenylbenzoic acid.

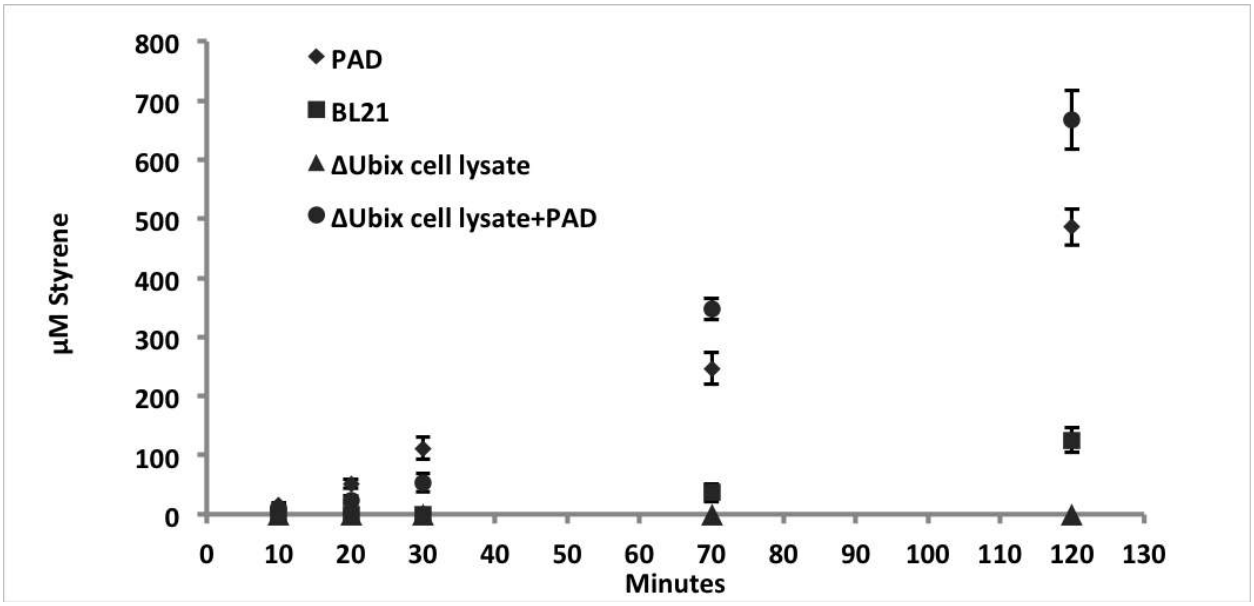


Figure 4: Direct interaction of tPAD1 and FDC is not required for activation of FDC. FDC in buffer containing 6.7 mM cinnamic acid was separated by a 3500 Da cutoff dialysis membrane from either BL21 cell lysate, Δ*ubiX* cell lysate, or Δ*ubiX* cell lysate with tPAD1. The production of styrene was monitored over time. (Adapted from Lin, et al. ACS Chem. Bio. 2014)

These experiments demonstrate that UbiX/tPAD1 make a low-molecular weight, diffusible cofactor that is necessary for FDC to perform decarboxylation. The direct interaction of tPAD1 and FDC is not required for transfer of this prosthetic group. It is also found that UbiX can substitute for tPAD1, however, it is not sufficient in creating highly active FDC, presumably due to a low occupancy. These experiments were critical in forming the basis for the following work in characterizing both FDC, tPAD1, and the cofactor necessary for activity of FDC.

Chapter 2:

Results and Discussion

Results and Discussion

Properties of recombinantly expressed and purified FDC

As mentioned in the background, FDC requires a low-molecular weight, diffusible cofactor for activity. Therefore, an expression vector containing both the *tpad1* gene (*tpad1* has no 6-his tag and was not purified from this construct) and a 6-his tagged *fdc* gene under the control of a single T7 promoter was constructed and transformed into *E. coli* BL21 to create a highly active FDC.

The specific activity of FDC expressed as a single gene on the expression vector in decarboxylating *trans*-cinnamic acid to styrene was determined by a GC-MS assay (see Materials and Methods section) to be ~31 μM styrene.min^{-1}.μM^{-1} enzyme. However, when FDC was co-expressed in a plasmid also containing the *tpad1* gene , activity was shown to increase to 247 μM styrene.min^{-1}.μM^{-1}.enzyme, a roughly 8-fold increase.

NATIVE-PAGE analysis of FDC performed by Fengming Lin indicated that the protein migrated with an apparent molecular weight of ~ 480 kDa, suggesting that the protein may adopt an octameric form in solution (Figure 5A). SDS-PAGE analysis indicates that monomeric FDC has a weight of ~58 kDa (Figure 4B) . This is in agreement with LC-ESI-MS results performed by Fengming Lin that demonstrate a monomeric weight of FDC being 57898.15 Da (Figure 5B). Both results are in excellent agreement with the predicted mass of 57897.8 Da. The activity of

FDC was found by Fengming Lin to be optimal between pH 7.0 - 8.0, declining sharply at pH below 7 (Figure 5C).

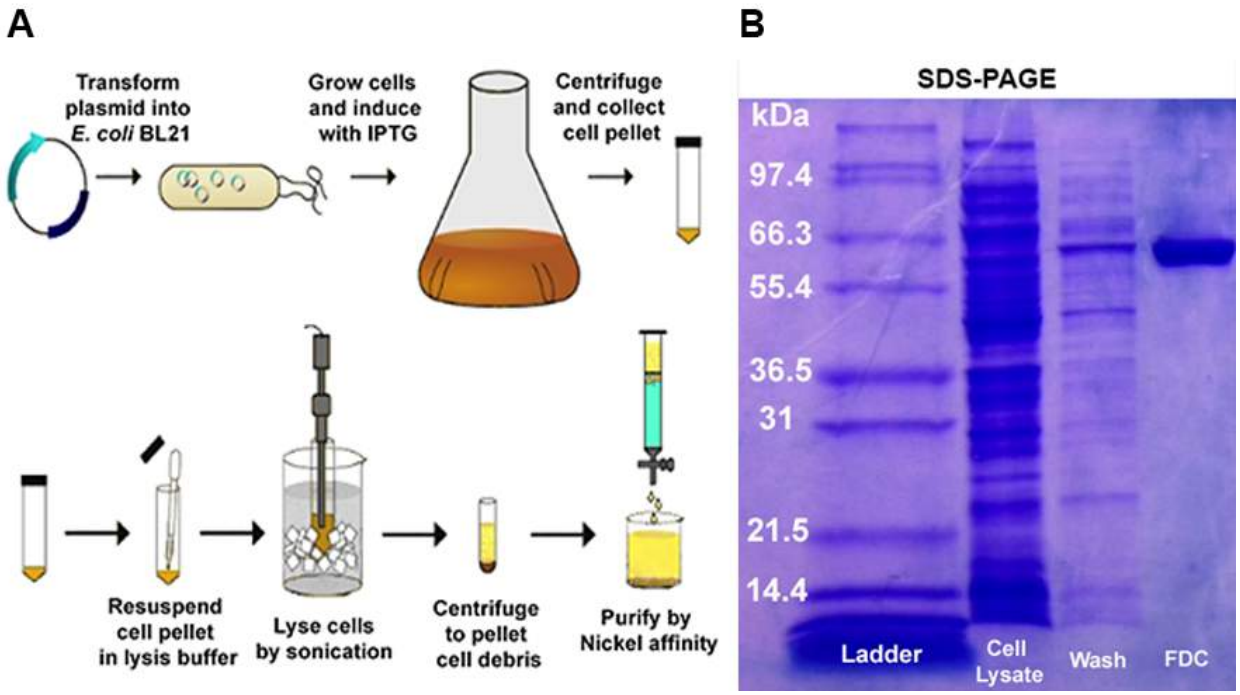


Figure 4: In order to further investigate the characteristics and mechanism of FDC, 6-his tagged FDC was successfully overexpressed and purified using standard methods (A) Typical purification scheme (B) SDS-PAGE analysis indicative of typical purity of eluted FDC to be high.

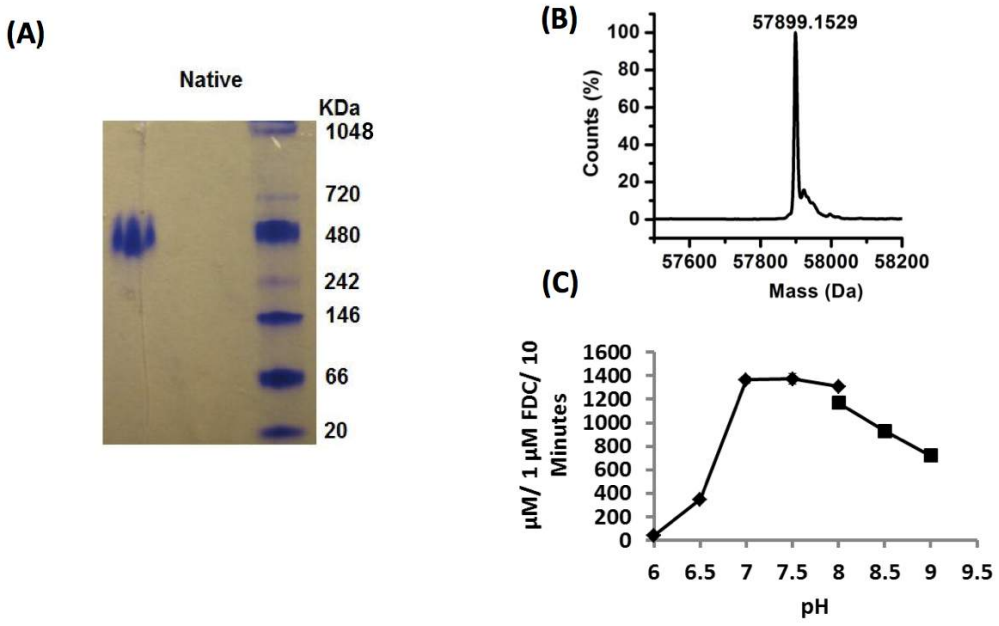


Figure 5: (A) NATIVE-PAGE analysis of purified FDC. (B) ESI-MS of purified FDC. (C) pH dependence of cinnamic acid decarboxylase activity.

Characterization of tPAD1

Characterization of tPAD1 was carried out via NATIVE-PAGE and SDS-PAGE. The genetic characterization of PAD1 indicated that PAD1 consists of 242 amino acid peptide with a MW of 26743.7 Da[21]. However, in order to properly purify the protein, a 58 residue leader sequence believed to encode a mitochondrial targeting sequence was removed, this protein is hence referred to as tPAD. This leads to a predicted molecular weight of 21806.4 Da. SDS-PAGE analysis corroborates the predicted molecular weight of tPAD1 as being ~ 22 kDa (Figure 6A). Liquid-chromatography electrospray ionization mass spectrometry (LC-ESI-MS) performed by Fengming Lin confirmed the monomeric weight of tPAD1 as being 21808.6 Da (Figure 6B), in excellent agreement with the predicted molecular weight of 21806.4 Da. NATIVE-PAGE analysis indicated that tPAD1 migrated with a molecular weight of ~ 300 kDa, suggesting that tPAD1 adopts a dodecameric structure in solution (Figure 6A). Gel-filtration chromatography performed by Fengming Lin also supports a molecular weight of the native protein as ~ 300 kDa, suggesting it is a dodecameric protein in solution. A crystallographic study indicated that a homolog of PAD1 with 54% sequence identity, FMN-dependent UbiX-like decarboxylase Pad1 from Escherichia coli O157:H7, forms a dodacameric structure with individual trimers forming the basic units. Our results indicate that tPAD1 may not only share sequence homology to this protein, but also adopt similar high order quaternary structures.

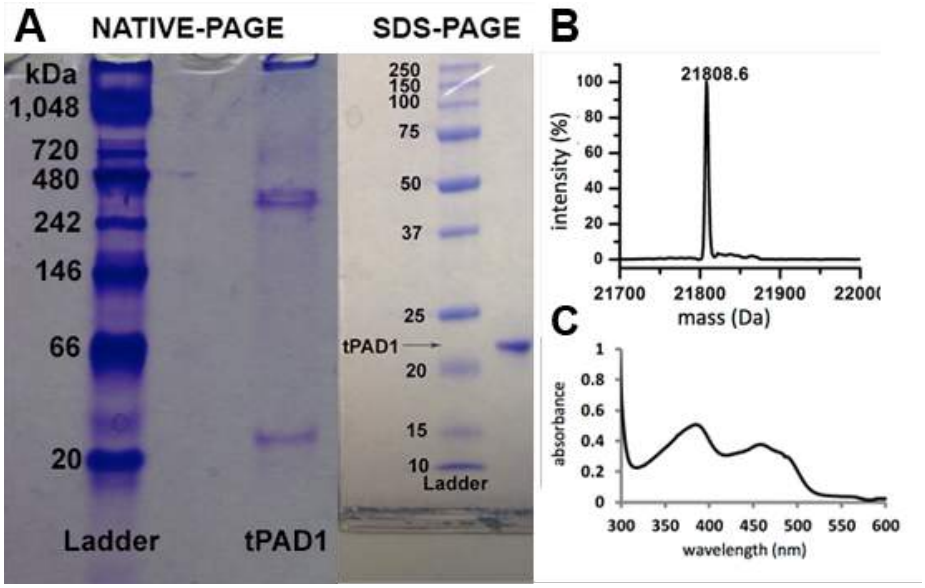


Figure 6: (A) NATIVE-PAGE and SDS-PAGE analysis of purified tPAD1. (B) ESI-MS analysis of purified tPAD1. (C) UV/vis spectrum of holo-tPAD1

Identification of FMN as cofactor for tPAD1

Purified tPAD1 shows homology to enzymes that utilize FMN as a cofactor[27]. Purified tPAD1 was also visibly yellow and the UV-vis spectrum exhibits absorbance peaks at 384 nm and 458 nm (Figure 6C), indicating the presence of oxidized FMN. The presence of FMN as a cofactor was confirmed first using reverse-phase high performance liquid chromatography (HPLC).

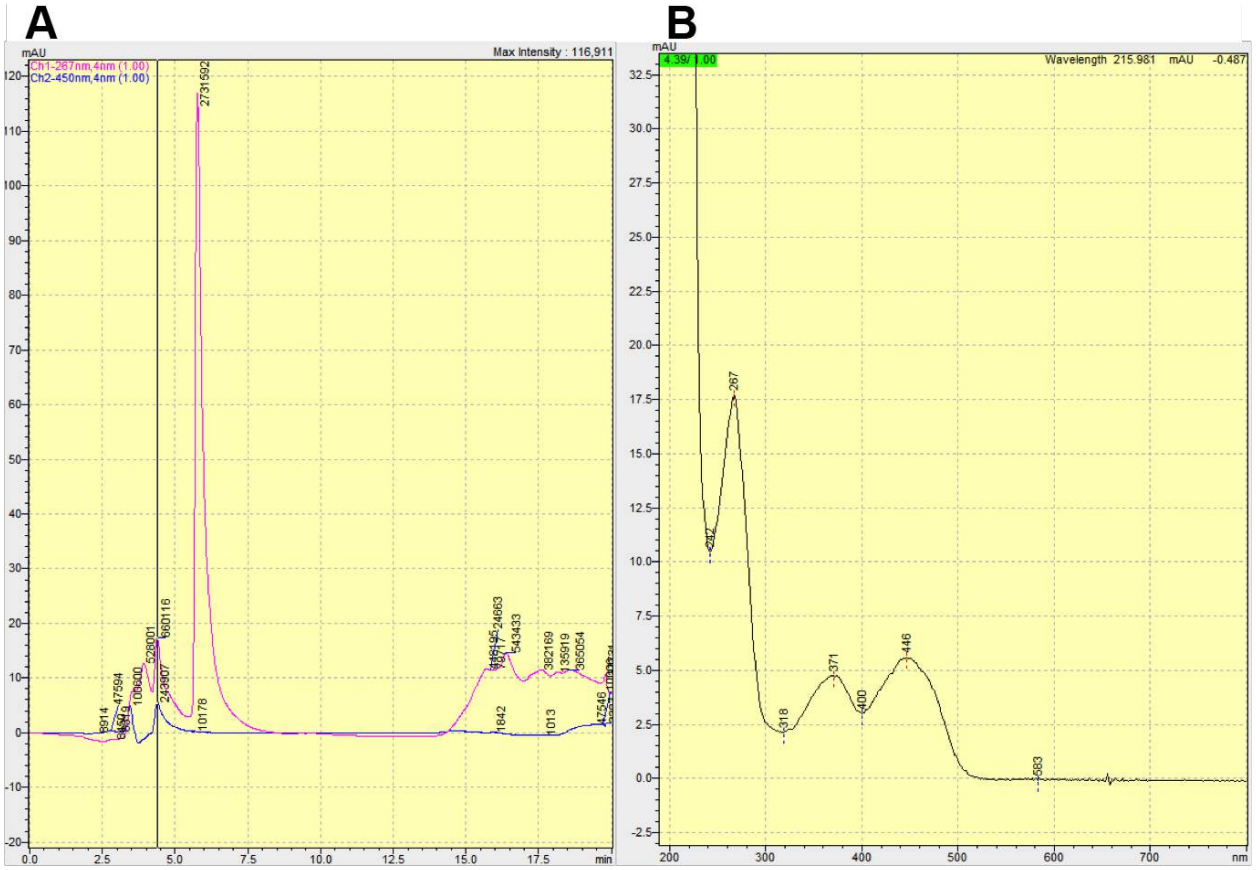


Figure 7: (A) HPLC chromatogram of isolated cofactor solution from tPAD1. (B) UV/vis spectrum of putative cofactor at retention time = 4.3 min.

The enzyme was thermally denatured at 100 ℃ for 10 min. and the resulting solution was filtered using an Amicon 3K spin filter to isolate the cofactor. The isolated cofactor was run on HPLC (see Methods and Materials for conditions). FAD and FMN standards were also run on the same conditions for comparison. The retention times of the cofactor was determined to be 4.3 min. (Figure 7A) while the retention times of FMN and FAD were found to be 4.3 min. and 3 min., respectively. The UV spectrum of the cofactor exhibited maxima at 371 and 446 nm, indicating the presence of a oxidized flavin (Figure 7B).

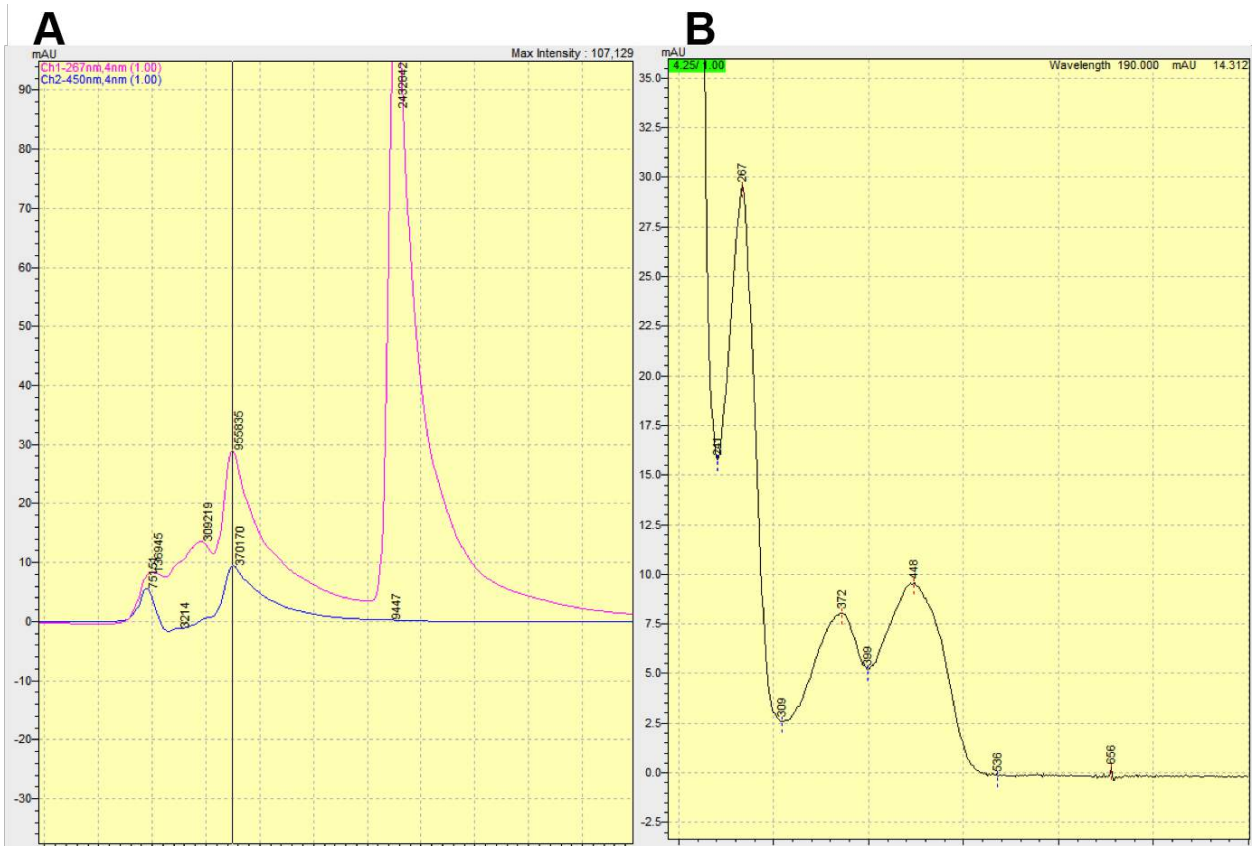


Figure 8: (A) HPLC chromatogram of isolated cofactor solution from purified tPAD1 supplemented with 5 µM FMN . (B) UV/vis spectrum of peak at retention time = 4.3 min.

When the isolated cofactor was spiked with 5 µM FMN, the peak corresponding to the cofactor grew in size proportional to the amount of FMN added (Figure 8A) while also maintaining the spectral properties of FMN (Figure 8B). This result strongly suggested that tPAD1 uses FMN as a cofactor. When supplemented with FAD, the peak corresponding to 4.3 min. did not grow in size and a separate peak at 3 min. appeared with spectral properties of FAD (data not shown).

LC-ESI-MS analysis was used to further confirm the identity of the cofactor used by tPAD1 as being FMN. Purified tPAD1 was denatured and the cofactor isolated as previously described. The isolated cofactor was then analyzed by LC-ESI-MS to determine its molecular weight.

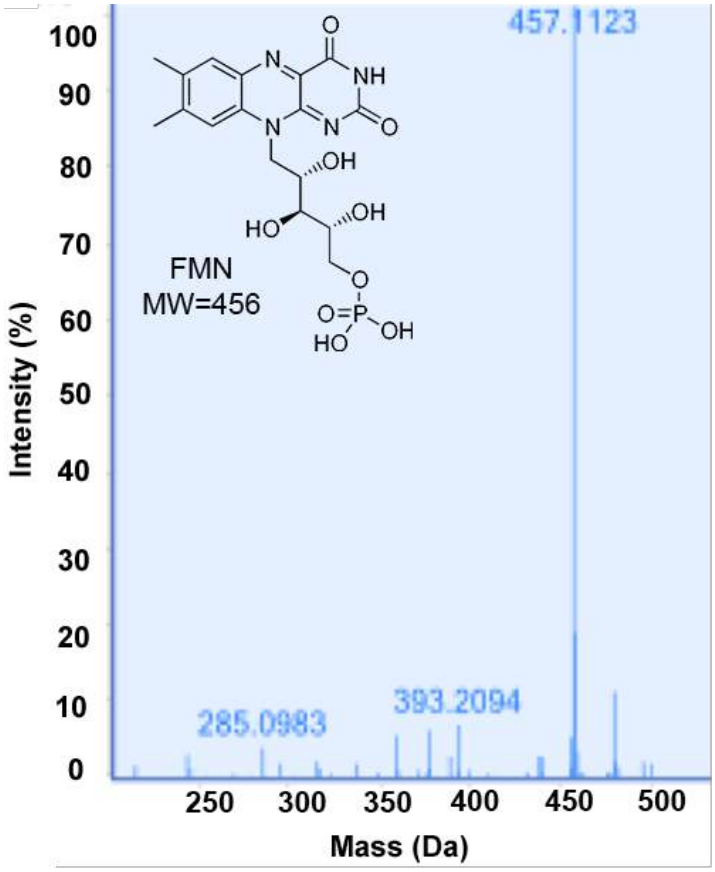


Figure 9: LC-ESI-MS chromatogram of tPAD1 cofactor confirmed to be FMN.

Chromatographic analysis identified a molecule with (m+1)/z of 457.1123 Da matching an authentic standard of FMN, M_r = 456 (Figure 10). These results confirm that tPAD1 utilizes FMN as a cofactor, although it is unclear the exact role that FMN plays in the function of tPAD1. The UV spectrum of purified tPAD1 (Figure 7C) had maxima at 384 nm and 458 nm. However, the UV spectrum of tPAD1 also exhibited characteristics of a second chromophore, due to the fact that the absorbance at 384 nm was stronger than at 458 nm, whereas the opposite is true for oxidized FMN. This indicates that tPAD1 may contain another chromophore, although no other peaks were detected on LC-ESI-MS.

Isolation and characterization of cofactor from FDC

In order to study the properties of the novel flavin-cofactor, it was necessary to find an efficient method of enzyme denaturation to obtain highly pure cofactor. Several methods were explored including thermal denaturation, chaotropic agents such as SDS, and organic solvents such as acetonitrile. The fastest and most efficient method was determined to be precipitation with 1:1 solution of methanol and dichloromethane. In typical isolations, the cofactor was released from holo-FDC by precipitation with 1:1 methanol:dichloromethane and removed by centrifugation (Figure 10). The organic phase was removed under a stream of N_2 and the aqueous phase was typically filtered using a 3K Amicon spin filter to remove any remaining protein.

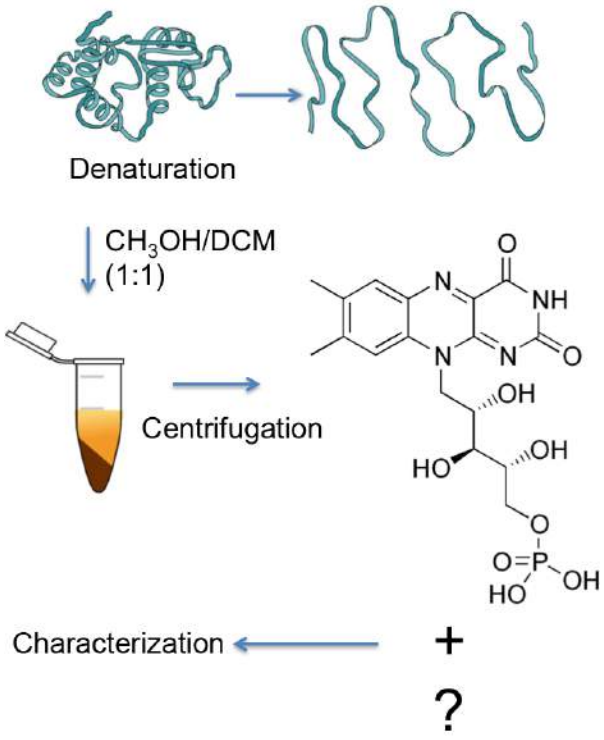


Figure 10: Cofactor isolation scheme.

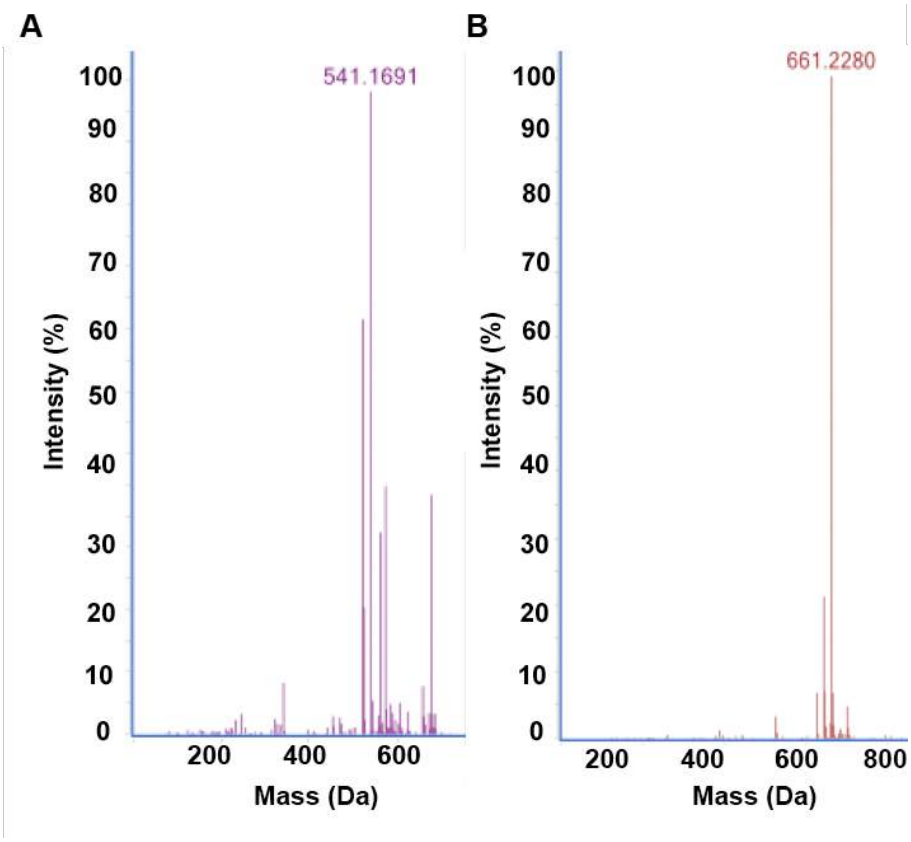


Figure 11: (A) LC-ESI-MS of possible breakdown product (B) MW of putative cofactor

The isolated cofactor was analyzed by LC-ESI-MS. The resulting chromatogram consistently indicated the presence of several molecules (Figure 11). Most prominently, a molecule with (m+1)/z of 661.2280 Da was identified on the resulting chromatogram (Figure 11B). Another consistent peak was a molecule with (m+1)/z of 541.1691 (Figure 11A). This peak is most likely a breakdown product of the peak at ~661 Da, which represents the true cofactor. These data, in conjunction with the knowledge that PAD1 uses FMN as a prosthetic group, suggest that PAD1 makes a covalent modification of FMN to produce the cofactor necessary for activity of FDC.

Fluorescence of Isolated Cofactor

The cofactor was isolated from purified FDC as previously described and analyzed using fluorescence spectroscopy. The isolated cofactor was excited at 345 nm and the emission profile was observed from 450-700 nm (Figure 12A). A 50 μM FMN standard was also run under the same conditions. The resulting redshift in maximum absorbance suggests that the cofactor is likely to be a FMN-derivative in the reduced state.

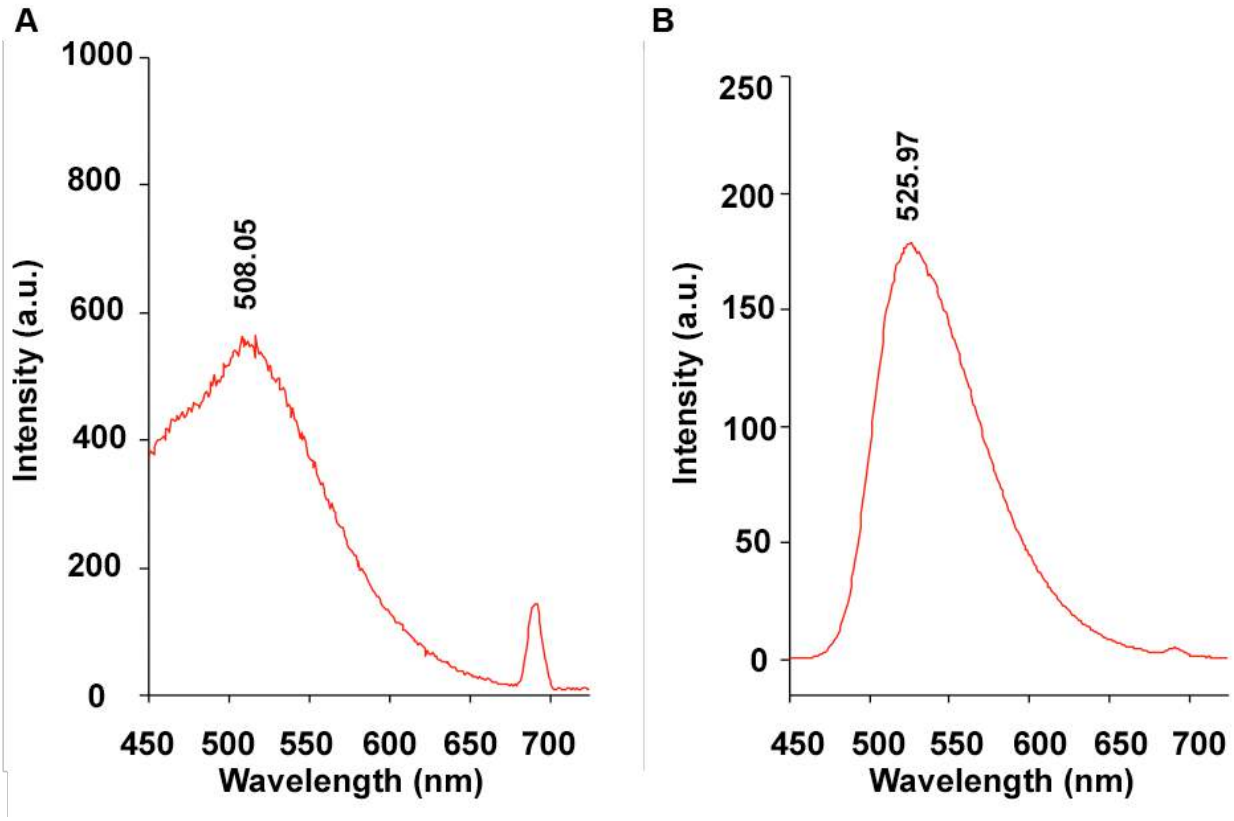


Figure 12: (A) Fluorescence emission spectrum of isolated cofactor. (B) Fluorescence emission spectrum of 50 μM FMN standard.

Enzyme Kinetics Studies

FDC was characterized using a UV/vis spectroscopic assay to determine its kinetic properties. The decarboxylation of *trans*-cinnamic acid to give styrene was used as an initial model, although other substrates (e.g., *p*-coumaric acid, ferulic acid) have been subsequently examined. A summary of the observed kinetic parameters is found in Table 1. The disappearance of *trans*-cinnamic acid at 304 nm was used to follow the progress of the reaction. A range of substrate concentrations was used to fit reaction velocities to a Michaelis-Menten curve (Figure 13A) and Lineweaver-Burke plot (Figure 13B) in order to determine kinetic parameters. The initial results indicate an observed k_{cat} of 2.5 s^{-1}, K_m of 232 μM, V_{max} of 147.1 μM/min. and k_{cat}/K_m of 1.08×10^5 $M^{-1}s^{-1}$.

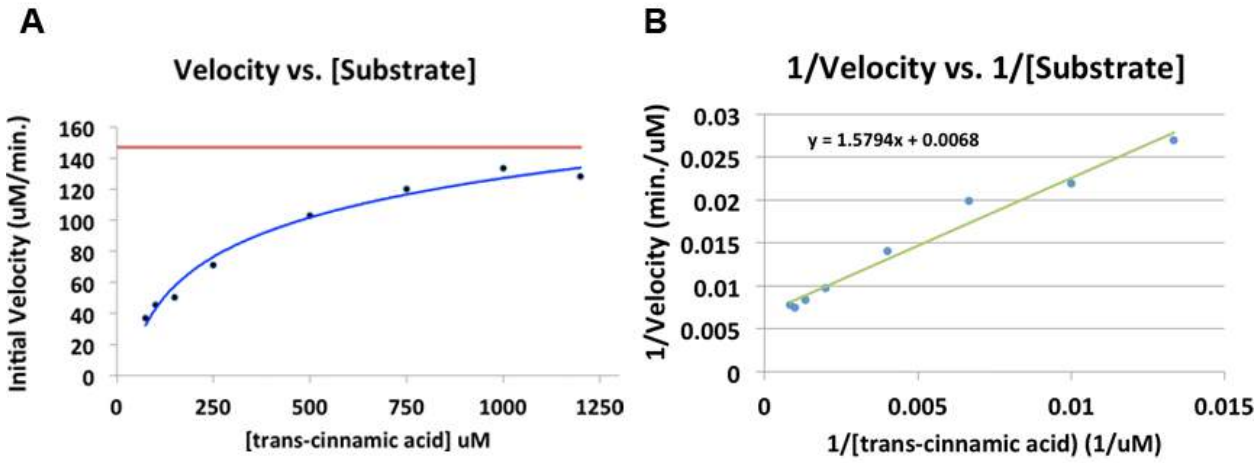


Figure 13: (A) Michaelis-Menten curve produced by monitoring initial reaction velocities as a function of varying cinnamic acid concentration. (B) Corresponding Lineweaver Burke plot of 1/Initial Velocity vs. 1/[substrate].

Substrate	k_{cat} **(s^{-1})**	K_m **(mM)**	k_{cat}/K_m **(s^{-1} M^{-1})**
cinnamic acid	4.6 ± 0.2	180 ± 24	25500 ± 3500
ferulic acid	3.8 ± 0.3	180 ± 41	21000 ± 4500
p-coumaric acid	1.5 ± 0.1	110 ± 26	13600 ± 880

Table 1: Observed kinetic parameters of FDC decarboxylating a variety of phenylacrylic acids.

Mechanistic Studies

Although PAD1 and FDC have been characterized for the first time in terms of their basic structural and biochemical properties, the reactions which they catalyze have yet to be understood at a mechanistic level. Efforts to understand the mechanism by which PAD1 creates the novel FMN-derived cofactor have been hindered by a low yield and lack of precedence in flavo-protein chemistry. Attempts to reconstitute an *in vitro* system by which PAD1 can create the cofactor necessary to activate apo-FDC have thus far produced no results. The addition of many commonly found and commercially available metabolites such as thiamine pyrophosphate, pyridoxal phosphate, nucleotides, NADH, NADPH, FMN and FAD and various metal ions were unsuccessful in allowing PAD1 to activate apo-FDC.

However, the mechanism by which FDC catalyzes the decarboxylation of various phenylacrylic acids has met with more success. Related enzymes have been shown to decarboxylate phenylacrylic acids and their respective mechanisms have been suggested as the result of structural data, most notably the decarboxylation of ferulic acid by ferulic acid decarboxylase from *Enterobacter*[24] (named FADase, not to be confused with FDC from *S. cerevisiae*). However, these related enzymes have not been shown to need the same cofactor as FDC and may proceed through significantly different intermediates.

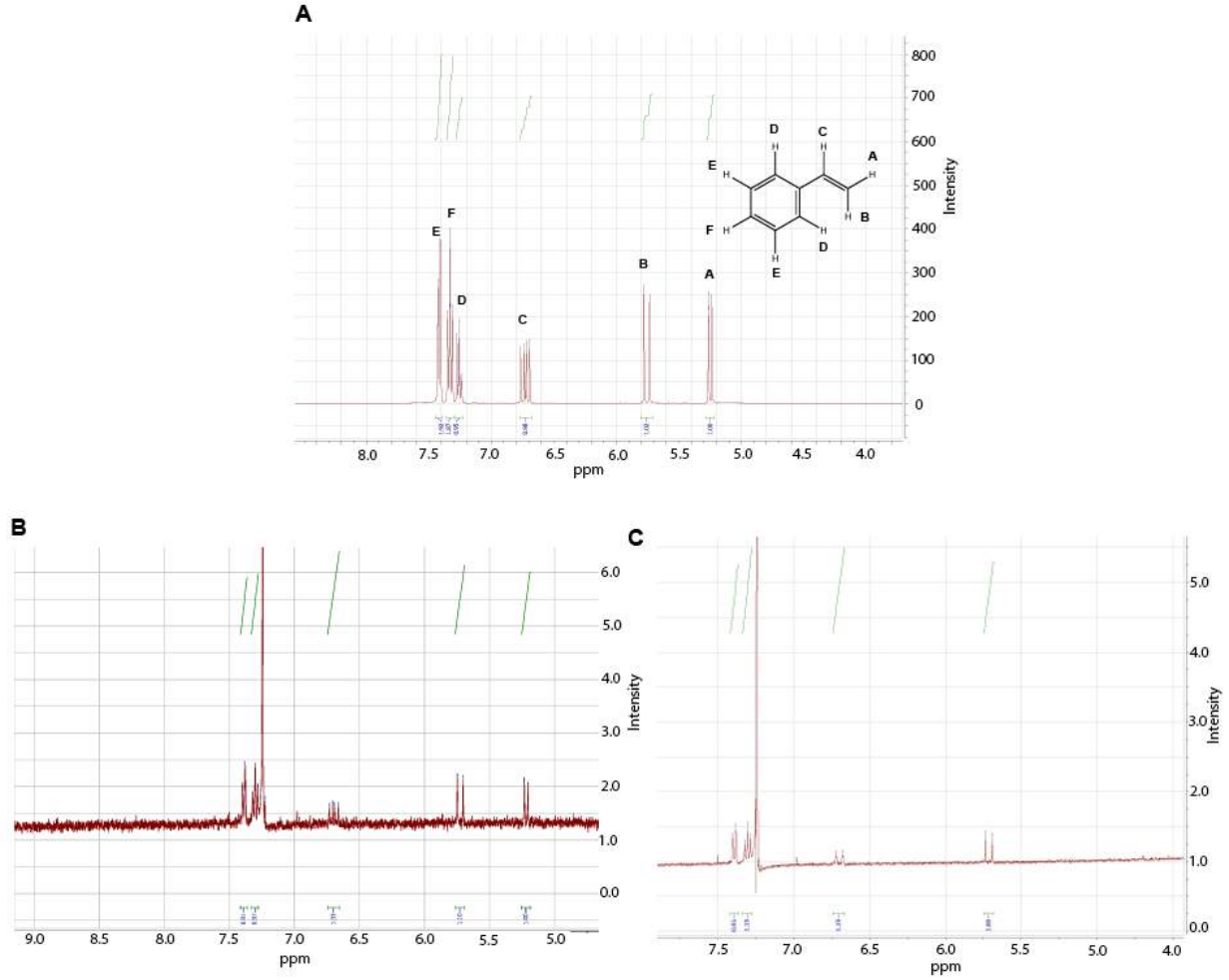
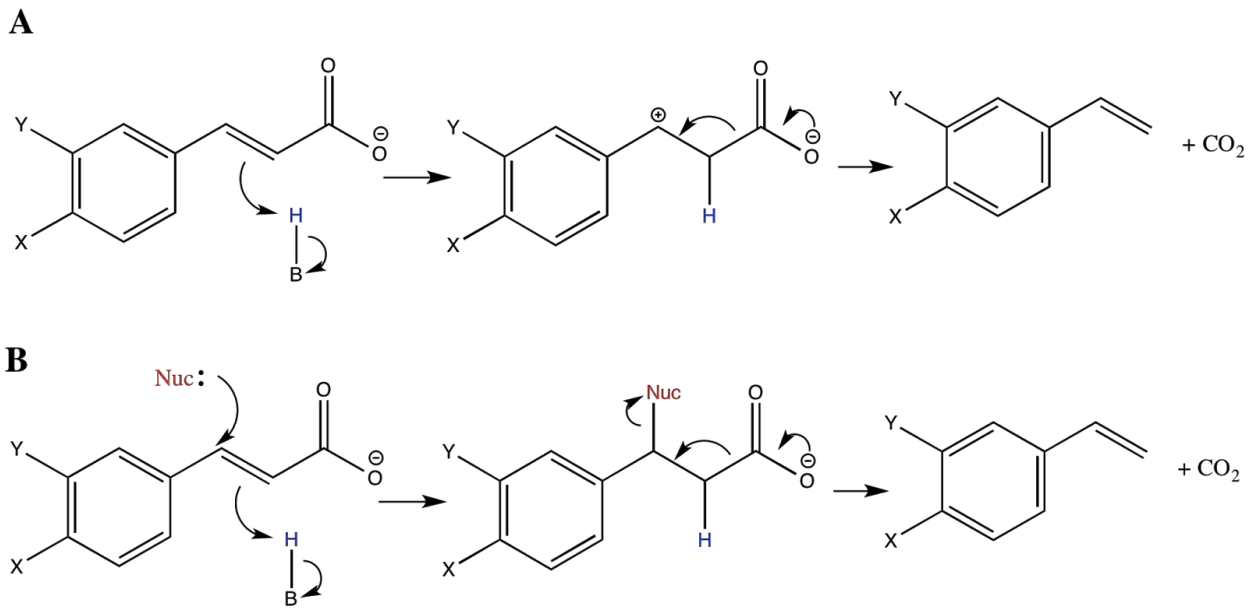


Figure (14): ^{1}H-NMR isotopic exchange experiment demonstrates that the *trans*-terminal hydrogen of styrene originates from the solvent. (A) Styrene standard in phosphate buffer (B) Styrene produced from cinnamic acid in non-deuterated phosphate buffer. (C) Styrene produced from cinnamic acid in deuterated HEPES buffer.

Further studies in our lab have refined the initial kinetic parameters of cinnamic acid decarboxylation as well as examined a wider range of substrates including *p*-coumaric acid and ferulic acid (Table 1). Preliminary NMR data has also demonstrated that the *trans*-terminal

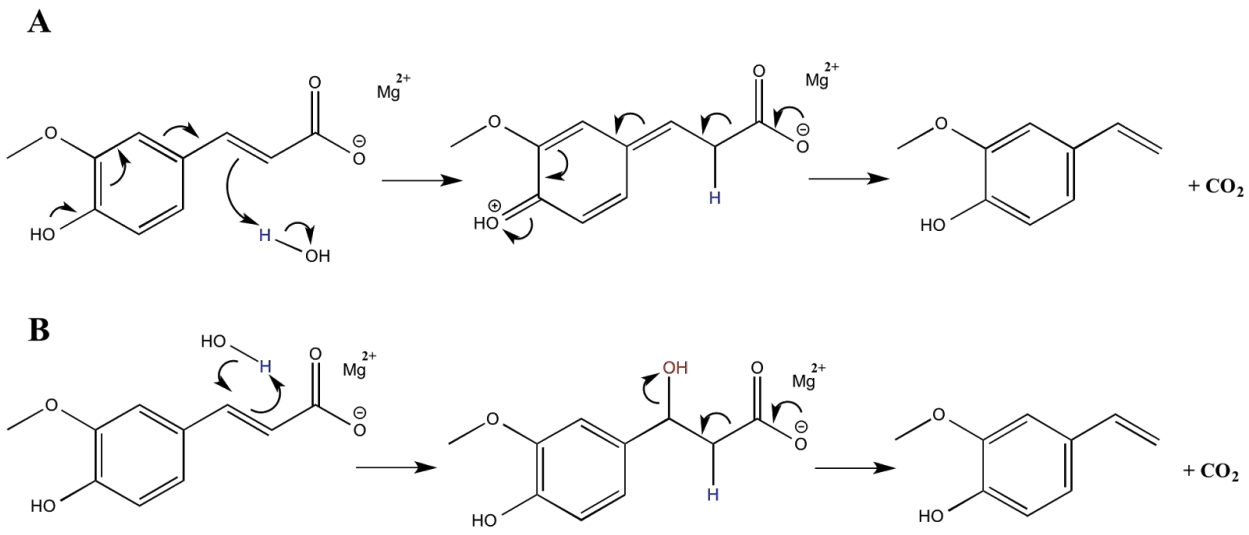
hydrogen of styrene originates from the solvent stereospecifically upon decarboxylation, suggesting that a protonation of the terminal double bond is a likely first step (Figure 14). Taking into account these data along with a review of enzyme-catalyzed decarboxylations, we believe two mechanisms are likely to occur, each occurring through a distinct intermediate. The two possible mechanisms are i) nucleophilic attack on the β-carbon in conjunction with α-carbon protonation followed by subsequent decarboxylation and leaving of the nucleophile or ii) α-carbon protonation leading to a carbocation intermediate which then undergoes spontaneous decarboxylation (Scheme 2).



Scheme 2: Possible mechanisms of polyacrylic acid decarboxylation by FDC

In order to inform mechanistic studies, *in silico* methods were utilized to identify which reaction pathway is more likely. We set out to use the computational software Spartan 14 to determine the energies of the various chemical species along the possible reaction pathways. Spartan 14 allows quantum calculations using different levels of theory (e.g., semi-empirical and *ab initio*) and basis sets (e.g., AM1, PM3, B3LYP) that account for the electronic structure of atomic and molecular species to varying degrees of accuracy.

Although much progress has been made in quantum chemistry in the last two decades due largely to the development of faster and more powerful computers, predicting chemical reactions remains a non-trivial matter[26]. We have found that to be the case here as well. With no structural information regarding the active site of FDC, the possible reaction space seems almost infinite. We have attempted to probe the mechanisms presented in Scheme 2 using both semi-empirical and Density Functional Theory (DFT) calculations. Our initial method of action was to use Spartan 14 to find equilibrium and transition state geometries with the semi-empirical AM1 level of theory and then determine subsequent energy profiles with DFT. However, finding reasonable transition states, especially for the initial protonation step, proved difficult. Due to the large amount of trial and error, limited computing time, and little progress, a revised mechanism was constructed that seemed to give reasonable results and insight into the likelihood of the two possible mechanisms. In particular, both the carbocation and nucleophilic mechanism of the decarboxylation of ferulic acid were examined using the revised mechanism (Scheme 3). Energies of starting materials, intermediates, transition states, and final products were examined using semi-empirical PM3 level of theory. This basis set was chosen in order to account for the added orbitals of magnesium while still being computationally feasible.



Scheme 3: Revised mechanism of decarboxylation of ferulic acid by FDC

The revised mechanism proved to give more reasonable energies as predicted by the PM3 level of theory in Spartan 14 and can help us make some basic assumptions about the nucleophilic vs. carbocation mechanism. First, it was found to be necessary to add a counter-ion to the carboxyl group in order for both mechanisms to proceed. Otherwise, our results suggested that the intermediate would revert back to the starting materials, or an intramolecular proton transfer would result in the carboxyl group protonated. Therefore, Mg^{2+} was added in simulations involving the revised mechanism. Mg^{2+} was chosen due to the fact that preliminary data suggests that FDC uses an FMN-modified cofactor, which often coordinates with a Mg^{2+} ion for stabilization of the negatively charged phosphate group. It should be noted however that no other counter ion was selected, and it would be plausible that a positively charged amino acid side chain (e.g., arginine) could fulfill this role. Second, the resonance contribution from the electron-donating *para*-hydroxyl group was found to be necessary for protonation of the double bond in the first step. Perhaps suggesting that the *para*-hydroxyl group may first get

deprotonated by an amino acid side chain in order to donate its electrons into the pi-conjugated system, as suggested in previous studies[25].

It should be noted that the first step in the revised nucleophilic mechanism was chosen to be concerted, having water act as a both a proton source and a nucleophile. It is much more likely that there involves more chemical species to perform this reaction (i.e., a separate proton source and nucleophile); however, it order to keep energy calculations commensurate between both both reactions, we chose to use water in both cases. Also, the calculations were performed in gas phase, and this would presumably differ greatly from the conditions of the enzyme active site. However, the gas phase calculations serve as starting point for future studies.

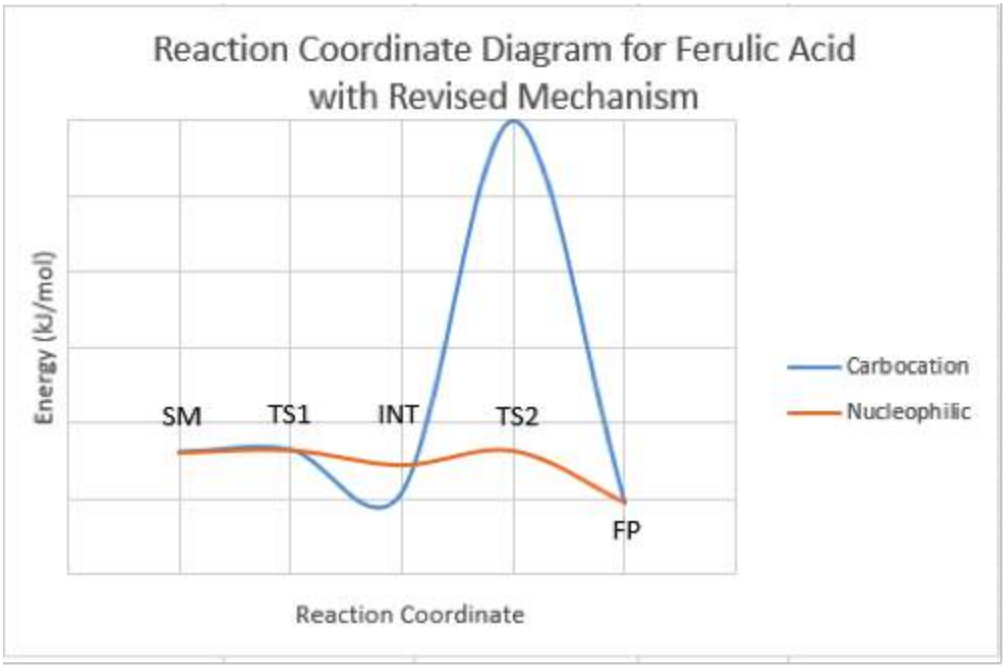


Figure 15: Reaction coordinate diagram for revised mechanism of ferulic acid decarboxylation. Reaction progress monitored at SM (starting materials), TS1 (first transition state, INT (intermediate), TS2 (second transition state), and FP (final products).

It is clear from Figure 15 that the carbocation mechanism includes a lower energy intermediate, but much higher transition state between the intermediate and the final product. Contrary to previous assumptions, the intermediates of both mechanisms have very similar energies. It was thought that the carbocation intermediate would be higher in energy due to the presence of an open-shell carbon. However, this was not found to be the case, and it is proposed that the resonance stabilization from the *para*-hydroxyl contributes to the stability of the intermediate. Indeed, when examining the electron density maps, it was surprisingly found that the most positive charge was located on the α-carbon of the carboxyl group. Presumably, this signifies the readiness of the intermediate to lose the carboxyl group in order to reform the double bond between the exocyclic carbon atoms. The energy of the final products indicate that this reaction is overall exothermic, most likely due to the creation of a stable carbon dioxide species. The nucleophilic mechanism is much more iso-energetic relative to the carbocation mechanism and may represent a more likely mechanism, under these conditions. This was expected due to the fact that no open-shell atoms exist in any state of the reaction. Both reaction mechanisms are exothermic under these conditions.

Chapter 3:

Conclusions and Future Directions

Conclusions

FDC and PAD1 from *S. cerevisae* were successfully recombinantly expressed and purified in *E. coli*. The results described characterize biochemically the enzymes PAD1 and FDC from *S. cerevisiae* for the first time. Importantly, it was found conclusively that PAD1 does not function as a decarboxylase, despite its name. It was found that PAD1 utilizes a FMN cofactor, most likely adopts a high order dodecameric structure in solution, and synthesizes a novel cofactor necessary for activity of FDC. This cofactor is most likely a covalently modified FMN in the reduced state. Its molecular weight was determined to be ~661 Da by LC-ESI-MS. It has been shown that this cofactor is released to FDC and that direct interaction of PAD1 and FDC are not needed for decarboxylation activity. FDC was shown to have optimal activity between pH 7.0 and 8.0, declining sharply at pH less than 7. FDC most likely adopts an oligomeric structure in solution. FDC was shown to decarboxylate a variety of phenylacrylic acids with k_{cat} ranging from ~1.5 - 4.6 s^{-1}. When FDC was expressed without tPAD1 present, the activity was low, but still present. When tPAD1 was co-expressed with FDC, the activity was dramatically increased by roughly 8-fold in decarboxylation of cinnamic acid. These data suggest that UbiX can substitute for tPAD1 in activating FDC, suggesting that the same cofactor is required by UbiD for decarboxylation of 3-hydroxy-4-octaprenylbenzoic acid by UbiD.

NMR and kinetic experiments continue to elucidate the mechanism of the enzyme catalyzed decarboxylation. These studies indicate that there are two likely decarboxylation

routes: one with a carbocation intermediate and the other occurring under nucleophilic (basic) conditions. In order to supplement *in vitro* experiments with computational studies, *in silico* methods were used to elucidate the likelihood of each mechanistic route. We find that a revised mechanism involving resonance contributions from the *para* substituents on the benzyl ring and the presence of a counter-ion (Mg^{2+}) were necessary for the construction of a reasonable reaction coordinate diagram for ferulic acid. Under these conditions, it is believed that although both the carbocation and nucleophilic reaction pathways are found to be exothermic, the nucleophilic mechanism is more plausible due to lower energy transition states.

Future Directions

Our current work is focused on further elucidating the mechanism whereby FDC decarboxylates phenylacrylic acids. Studies using various substituted phenylacrylic acids to determine new kinetic parameters are being performed to allow the construction of a Hammett correlation plot. These experiments should help support one of the two possible mechanistic routes outlined above. In order to increase the activity of FDC, a synthetic operon is being constructed with the *pad1* and *fdc* genes both under the control of two individual T7 promoters. This should allow increased levels of PAD1 in *E. coli* expressing this dual construct. Correspondingly, increased production of PAD1 should allow for a higher cofactor occupancy of FDC, which should lead to higher observable k_{cat}. Crystallization of FDC leading to X-ray structures would greatly facilitate not only mechanistic studies, but also protein engineering efforts through the use of active site mutagenesis and perhaps even structure-guided recombination.

Further computational studies could be performed at higher levels of theory (e.g., DFT) to obtain more accurate geometries and energies. A great variety of proton sources, nucleophiles, and counter-ions could be examined to determine the robustness of the revised mechanism obtained from preliminary calculations. An X-ray structure of FDC would also greatly facilitate enzyme active site modeling such as docking studies for virtual screening of substrates. In addition, an X-ray structure would allow amino acid residues to be used in future computational studies in place of water and magnesium to act as the proton source/counter-ion. The X-ray structure of either PAD1 or FDC with the flavin-derived cofactor could also help understand what role this prosthetic group plays in the reaction, if any.

Methods and Materials

Protein Expression and Purification

Protein expression was performed by standard methods. *E. coli* cultures were grown at 32 ℃ in LB medium supplemented with 50 μg/mL kanamycin to an OD_{600} of 0.6 ~ 0.8 and gene expression induced with 0.2 mM isopropyl-D-thiogalactopyranoside (IPTG) overnight before harvesting the cells by centrifugation.

In a typical purification of FDC or PAD1, 7 g (damp weight) of cells were resuspended in 50 mL binding buffer (20 mM phosphate buffer, pH 8.0, 500 mM NaCl, 20 mM imidazole, 5% glycerol, 1 mM THP). 0.5 mg/mL of lysozyme, 1 protease inhibitor tablet (Roche), 1 μL of DNAse (Novagen) and 1 mM tris(hydroxylpropyl) phosphine (THP) were added to the cell suspension and incubated on ice for 1 h with gentle shaking. The cells were lysed by sonication at maximum power using 2 s pulses separated by 8 s to prevent overheating for a total time of 30 min. The supernatant was separated from cell debris by centrifugation at 15000 *g* at 4 °C for 30 min. Protein purification was performed using an ÄKTAexplorer™ chromatography system. The supernatant was loaded onto a HisTrap column (GE Healthcare) and the column was washed with buffer (20 mM phosphate buffer, pH 8.0, containing 500 mM NaCl, 60 mM imidazole, 5 % glycerol,) at a flow rate of 1.0 mL/min. The enzyme was eluted from the column with 500 mM imidazole in the same buffer. The resulting fractions were analyzed by SDS-PAGE on a 15 % gel. The fractions containing the pure target protein were pooled and passed through PD-10 column (GE Healthcare) for desalting into PBS buffer plus 5% glycerol.

GC-MS Enzyme Assay

Routine assays of FDC activity were carried out in PBS buffer, pH 8.0, with a total volume of 500 µL, using cinnamic acid, 6.7 mM, as the substrate; the concentration of FDC was typically 1 µM. Assays utilizing PAD1 typically included this enzyme at 10 µM. The assay mixtures were incubated at 32 °C for varying lengths of time, after which the reaction was quenched and the reaction products were extracted by addition of 500 µL ethyl acetate; samples were vortexed for 15 min and then centrifuged in a microfuge at 12000 rpm for 20 min to separate the organic and aqueous phases. The amount of styrene produced was determined by GC-MS.

GC-MS analysis was performed using a Shimadzu QP- 2010S GC-MS instrument equipped with a quadrupole mass detector and a DB-5 column (Restek, 30m x 0.25 mm x 0.25 µm). The flow rate of the helium carrier gas is constant at 1 mL/min and the inlet temperature was maintained at 200 °C. The interface temperature was maintained at 250 °C. 10 µL injections of the ethyl acetate extracted solution were made in splitless mode. Oven temperature was held initially at 40 °C for 3.5 min, gradually increased to 90 °C at 14 °C/min, and then gradually increased from 90 °C to 315 °C at 20 °C/min and finally maintained at 315 °C for 1 min. Chromatographic data were analyzed by GC-MS PostRun analysis software. Enzymatic conversion of cinnamic acid to styrene was quantified using a calibration plot of styrene standards of known concentration.

Spectroscopic enzyme assay

For high activity preparations of FDC, i.e. enzyme prepared from *E. coli* strains co-expressing tPAD1, the decarboxylation of substrates could be followed by the decrease in absorbance of the u.v.-active substrates. Typical assays were performed in 100 mM potassium phosphate buffer,

pH 7.0, at room temperature and contained 1 μM FDC. The substrate concentrations ranged between 0.05 mM and 1.5 mM; at higher concentrations the absorbance of the substrate became too high. Decarboxylase activity was followed by monitoring the decrease in absorbance at 304 nm for cinnamic acid; 334 nm for para-coumeric acid; and 344 nm for ferulic acid. Velocity data were fitted to the Michaelis-Menten equation using Origin™ software.

LC-MS analysis of proteins and small molecules

The molecular weights of proteins and small molecules were determined using an Agilent 6520 LC - accurate-mass Q-TOF MS system. The protein was passed through a desalting column and acidified with 0.1% formic acid. 5 ml of the sample was injected onto a Poroshell 300SB-C8 column equilibrated with 0.1% formic acid and 5% acetonitrile. Proteins were eluted for 5 min with 95% water: 5% acetonitrile followed by an increasing gradient of acetonitrile to 100% over 7 min at a flow rate of 0.5 mL/min. Eluting proteins were detected at 280 nm. Mass data were obtained using intact protein mode and analyzed using Agilent MassHunter Qualitative Analysis software. The raw data was deconvoluted with respect to maximum entropy.

For small molecule analysis proteins were precipitated with an equal volume of 1:1 dichloromethane:methanol in glass tubes and removed by centrifugation at 3000 *g* for 10 min. The organic solvents was removed under a stream of nitrogen. 20 ml of the sample then was injected onto a Zorbax 300SB-C18 column equilibrated with 0.1% formic acid and 5% acetonitrile, followed by a 2 min washing with 95% water: 5% acetonitrile. Then the gradient of acetonitrile was increased to 100% over 10 min at a flow rate of 0.4 mL/min. Mass data were

obtained using small molecular mode and analyzed using Agilent MassHunter Qualitative Analysis software.

Reverse Phase HPLC

For HPLC analysis of isolated cofactor, purified tPAD1 was thermally denatured for 10 min. at 100 °C and filtered using an Amicon Ultra 0.5 mL 3K spin column. 50 µL injections were made and eluted for 10 min. at a constant flow rate of 1.0 mL/min in isocratic mode with 75% 5 mM NH_4(OAc) pH 6.5 and 25% methanol. Eluents were monitored at 260 nm and 450 nm. HPLC analysis was performed using a Shimadzu LC-20AT Liquid Chromatograph with a SPD-M20A Diode Array Detector equipped with a Vydac 201SP C_{18} Reversed Phase column (90 Å, 5 µm, 4.6 mm i.d. x 250 mm).

Fluorescence Characterization

Cofactor was isolated as previously described and analyzed on a Varian Cory Eclipse Fluorescence Spectrophotometer. Isolated cofactor was excited at 345 nm using a "high" photomultiplier effect (PMT) and emission was observed from 450 nm to 700 nm. A 50 µM FMN standard was excited at 345 nm using a "medium" PMT and emission was observed from 450 nm to 700 nm.

Isotopic NMR experiments

3 mM of cinnamic acid was reacted with 1 μM holo-FDC in deuterated HEPES (100% D_20, 100 mM HEPES, 100 mM KCl, pH 6.8) for 30 min. at 25 degrees Celsius. Styrene was extracted by addition of 500 μL deuterated chloroform. Samples were gently inverted for 2 min. and then centrifuged at 12000 rpm for 20 min. to separate the organic and aqueous phases. The solution was dried using Molecular Sieves from Fischer Scientific for 5 min. Resulting precipitate was analyzed using a Variance 400 MHz ^{1}H-NMR, and compared to a control experiment using non-deuterated HEPES buffer.

Computational Studies

Spartan 14 was used to predict equilibrium/transition state geometries and their corresponding energies. Semi-empirical PM3 level of theory was used to calculate vibrational modes, charge densities, orbitals, and energies. The models were calculated in gas phase (*in vacuo*). To insure a realistic transition state geometry was achieved, curved arrows were input before performing the calculations. After the calculations had completed, the IR spectra were examined to confirm the existence of an imaginary vibration frequency representative of the transition state. The frequency was visualized and compared to expected vibrational motion of the transition state. Energies were found in the output file of each semi-empirical geometry calculation and were given as heat of formations (kJ/mol). These energies were plotted against the each chemical species to give a reaction coordinate diagram.

References

[1] Neidelman, S.B. The archeology of Enzymology. In: Abramowicz D (ed) Biocatalysis, Van Nostrand Reinhold, New York, 1-24 (1990).

[2] Roberts SM, Turner NJ, Willetts AJ, Turner MK. Introduction to Biocatalysis using enzymes and Microorganisms, Cambridge University Press, Cambridge. 1-33 (1995).

[3] Menger, F. Enzyme reactivity from an organic perspective. *Accounts of Chemical Research* **26,** 206212 (1993).

[4] Faber, Kurt. *Biotransformations in Organic Chemistry.* 5th Ed. Springer.

[5] Pedroso, I. *et al.* AlterORF: a database of alternate open reading frames. *Nucleic acids research* **36,** 517–8 (2008).

[6] Li, L.-L. L., McCorkle, S. R., Monchy, S., Taghavi, S. & van der Lelie, D. Bioprospecting metagenomes: glycosyl hydrolases for converting biomass. *Biotechnology for biofuels* **2,** 10 (2009).

[7] Cleland, W.W. (1999) Mechanisms of enzymatic oxidative decarboxylation. *Accounts of Chemical Research* **32**, 862-868.

[8] Belcher J, McLean KJ, Matthews S, Woodward LS, Fisher K, et al. Structure and Biochemical Properties of the Alkene Producing Cytochrome P450 OleTJE (CYP152L1) from the Jeotgalicoccus sp 8456 Bacterium. *Journal of Biological Chemistry* **289**, 6535-6550 (2014).

[9] Jordan F., Patel H. Catalysis in Enzymatic Decarboxylations: Comparison of Selected Cofactor-Dependent and Cofactor-Independent Examples. *ACS Catalysis* **3**, 1601-1617 (2013).

[10] Kourist, R., Guterl, J., Miyamoto, K. & Sieber, V. Enzymatic Decarboxylation—An Emerging Reaction for Chemicals Production from Renewable Resources. *ChemCatChem* **6,** 689–701 (2014).

[11] Jung, D.-H. H. *et al.* Bioconversion of p-coumaric acid to p-hydroxystyrene using phenolic acid decarboxylase from B. amyloliquefaciens in biphasic reaction system. *Applied Microbiology and Biotechnology* **97,** 1501–11 (2013).

[12] Straathof, A. J. J. Transformation of biomass into commodity chemicals using enzymes or cells. *Chemical Reviews* **114,** 1871–908 (2014).

[13] Shin, J. H., Kim, H. U., Kim, D. I. & Lee, S. Y. Production of bulk chemicals via novel metabolic pathways in microorganisms. *Biotechnology Advances* **31**, 925–35 (2013).

[14] Winkel-Shirley, B. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiology* **126**, 485–93 (2001).

[15] Rao, R. S. & Ravishankar, GA. Plant cell cultures: chemical factories of secondary metabolites. *Biotechnology Advances* **20**, 101–153 (2002).

[16] Yaman, S. Pyrolysis of biomass to produce fuels and chemical feedstocks. *Energy Conversion and Management* **45**, 651-671 (2004).

[17] Kang, S.-Y. Y. *et al.* Artificial biosynthesis of phenylpropanoic acids in a tyrosine overproducing Escherichia coli strain. *Microbial Cell Factories* **11,** 153 (2012).

[18] McKenna, R. & Nielsen, D. Styrene biosynthesis from glucose by engineered E. coli. *Metabolic Engineering* **13,** 544–54 (2011).

[19] Stratford, M., Plumridge, A. & Archer, D. B. Decarboxylation of sorbic acid by spoilage yeasts is associated with the PAD1 gene. *Applied and Environmental Microbiology* **73**, 6534–42 (2007).

[20] Mukai, N., Masaki, K., Fujii, T., Kawamukai, M. & Iefuji, H. PAD1 and FDC1 are essential for the decarboxylation of phenylacrylic acids in Saccharomyces cerevisiae. *Journal of Bioscience and Bioengineering* **109,** 564–9 (2010).

[21] Clausen, M., Lamb, C., Megnet, R. & Doerner, P. PAD1 encodes phenylacrylic acid decarboxylase which confers resistance to cinnamic acid in Saccharomyces cerevisiae. *Gene* **142,** 107–12 (1994).

[22] Ago, S., Kawasaki, H., and Kikuchi, Y., Abstr. 50th Annu. Meet. Soc. Biothechnol., Jpn., p. 24, 1998.

[23] Gulmezian, M., Hyman, K., Marbois, B., Clarke, C. & Javor, G. The role of UbiX in Escherichia coli coenzyme Q biosynthesis. *Archives of Biochemistry and Biophysics* **467**, 144–53 (2007).

[24] Zhang, H. & Javor, G. T. Regulation of the isofunctional genes ubiD and ubiX of the ubiquinone biosynthetic pathway of Escherichia coli. *FEMS microbiology letters* **223**, 67–72 (2003).

[25] Gu, W. *et al.* Structural Basis of Enzymatic Activity for the Ferulic Acid Decarboxylase (FADase) from Enterobacter sp. Px6-4. *PLoS ONE* **6,** (2011).

[26] Deglmann, P. et al. Application of Quantum Calculations in the Chemical Industry—An Overview. *J. Quantum Chem.* (2014).

[27] Rangarajan, E. *et al.* Crystal structure of a dodecameric FMN-dependent UbiX-like decarboxylase (Pad1) from Escherichia coli O157: H7. *Protein Science* **13**, 3006–16 (2004).