

**Active Site Studies of Arogenate Dehydrogenase & the Purification of its Substrate**

**Natalie Rachel**

**A Thesis**

**In the Department**

**of**

**Chemistry and Biochemistry**

**Presented in Partial Fulfillment of the Requirements**

**For Degree of Master of Science at**

**Concordia University**

**Montréal, Québec, Canada**

**December 2011**

**© Natalie Rachel, 2011**

**CONCORDIA UNIVERSITY**

**School of Graduate Studies**

This is to certify that the thesis prepared

By: Natalie Rachel

Entitled: Active Site Studies of Arogenate Dehydrogenase & the  
Purification of its Substrate

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:

\_\_\_\_\_  
Dr. Gilles Peslherbe

\_\_\_\_\_  
Dr. Paul Joyce

\_\_\_\_\_  
Dr. Peter Pawelek

\_\_\_\_\_  
Dr. Joanne Turnbull

Approved by \_\_\_\_\_  
Chair of Department or Graduate Program Director

---

Date

---

Dean of Faculty

## Abstract

### Active Site Studies of Arogenate Dehydrogenase and the Purification of its Substrate

Natalie Rachel

Arogenate dehydrogenase (AD) from the cyanobacterium *Synechocystis sp.*

PCC6803 is one enzyme within the family of TyrA proteins dedicated to the biosynthesis of L-tyrosine (L-Tyr) and catalyzes the NADP<sup>+</sup>-dependant oxidative decarboxylation of L- arogenate to L-Tyr. To circumvent some of the difficulties in obtaining L-arogenate using current methodology, we developed an improved methodology to synthesize L-arogenate enzymatically using the recently annotated prephenate aminotransferase from *Petunia hybrida*. L-arogenate was then purified from major contaminants by high performance liquid chromatography. Guided by sequence alignments and the crystal structure of AD, residues were targeted for the first mutagenesis experiments on this enzyme; the variant proteins were characterized by kinetic, biophysical and computational methods. We identified that His112, a highly conserved residue, is the catalytic hydrogen bond acceptor and is critical for enzyme function, whereas electrostatic interactions afforded by Asp171 appear critical for the overall stability of the enzyme. Both Arg213 and Arg217, a highly conserved residue, appear to play a role in enzyme function, which contradicts the conclusion inferred from the crystal structure. His170, was also identified to be an important active site residue, and likely maintains His112 in a catalytically competent conformation via hydrogen bonding. Ser178, His179, T117 and Gly221 were all also found to be important for enzyme function. Even though the exact roles of these residues require further clarification, our studies have provided valuable insight into how arogenate dehydrogenase functions.

## Acknowledgements

I'll start by extending my gratitude to my supervisor, Dr. Joanne Turnbull, for her guidance, patience, and advice. She respected me, provided unwavering encouragement, and gave me an absurd amount of freedom with my project, for which I am very thankful.

I would also like to thank my committee members, Dr. Paul Joyce and Dr. Peter Pawelek, for their advice, suggestions, and support with my research. Dr. Judith Kornblatt's invaluable knowledge, advice and ideas were also vastly appreciated. I would like to extend my thanks to Dr. Guillaume Lamoreux for helping me with molecular docking and homology modelling experiments.

Infinite thanks go to Dr. Marc Ouellet, Dorival Martins, and Avid Hassanpour, for helping me set up and run analytical experiments, as well as offering their expertise in subjects like mass spectrometry and liquid chromatography. In particular, Marc: extra thanks for being such a wealth of information, consistently helpful, and for having more ideas than there are Mormons in Utah.

Much thanks goes to my excellent friend (and Turnbull lab emeritus) Guillaume Beaudoin. Even though he had to leave Montréal to go do his PhD in some big shot lab in Calgary, the distance did not keep him from being a great person to bounce hair-brained ideas around with, offering advice, and being extremely thoughtful when it came to trouble shooting. I'm very grateful for that.

Next is my other excellent friend, Lauren Narcross, my local molecular biology guru. Oh Lauren. Thank you so much for reviewing my primers designs, the endless brainstorming, and being such a plethora of information for molecular biology theorycrafting and troubleshooting.

I would also like to thank my friend and future colleague, Maria Gundersen, for also being highly excellent, helpful, supportive, a great friend, and a consistent source of (in)sanity. I also cannot overlook France Desrochers, who has been an excellent friend and roomie for two-and-a-half years and counting. Her support was very precious even when I thought I didn't need it. Additional thanks go to previous and current Turnbull lab members, Natascha Hotz and Hayline Kim for their support and advice, as well as Francis McManus for his help.

Thank you, everyone. This is all your fault.

*“Let your life sing.”*

*Charity Tillemann-Dick*

## Table of Contents

<b>List of Figures.....</b>	<b>IX</b>
<b>List of Tables.....</b>	<b>XI</b>
<b>List of Abbreviations.....</b>	<b>XII</b>
<b>Chapter 1: Introduction.....</b>	<b>1</b>
1.1 Aromatic Amino Acids Biosynthesis.....	1
1.2 The TyrA Protein Family.....	4
1.3 The Active Sites of the Bifunctional CM-PD from <i>E. coli</i> .....	6
1.4 Proposed Mechanism of Prephenate and Arogenate Dehydrogenases.....	7
1.5 Crystallographic Studies on Different TyrA Protein.....	10
1.6 Prephenate Dehydrogenase from <i>Aquifex aeolicus</i> .....	12
1.7 Contrasting L-Tyrosine Feedback Inhibition in <i>E. coli</i> and <i>H. influenzae</i> CM-PD and <i>A. aeolicus</i> PD.....	17
1.8 Arogenate Dehydrogenase from <i>Synechocystis sp.</i> PCC 6083.....	18
1.9 Synthesis and Purification of Arogenate.....	22
1.10 Research Objectives.....	24
<b>Chapter 2: Materials and Methods.....</b>	<b>28</b>
2.1 Materials.....	28
2.2 Strains and Plasmids.....	39
2.3 Bacterial Growth Media.....	30
2.4 Site-directed mutagenesis.....	30
2.5 Transformation.....	33
2.6 Agarose Gel Electrophoresis.....	33
2.7 Expression and Purification of AD Proteins.....	34
2.8 Expression and Purification of Prephenate Aminotransferase.....	35
2.9 L-arogenate Synthesis.....	36
2.10 High Pressure Liquid Chromatography and Liquid Chromatography-Mass Spectrometry.....	37

2.11 Determination of Protein Concentration.....	37
2.12 Polyacrylamide Gel Electrophoresis.....	38
2.13 ESI-ToF Mass Spectrometry.....	39
2.14 Determination of Enzyme Activities and Kinetic Parameters.....	40
2.15 Far-UV Circular Dichroism Spectroscopy.....	41
2.16 Near-UV Circular Dichroism Spectroscopy.....	42
2.17 Fluorescence Emission Spectroscopy.....	43
2.18 Molecular Modeling and Docking.....	44
<b>Chapter 3: Results.....</b>	<b>46</b>
3.1 Site-Directed Mutagenesis.....	46
3.2 Expression and Purification of WT <i>Synechocystis</i> AD and Variant Proteins.....	46
3.3 Electrospray Ionization Mass Spectrometry.....	49
3.4 Circular Dichroism Spectroscopy.....	51
3.4.1 Far-UV Circular Dichroism.....	51
3.4.2 Near-UV Circular Dichroism.....	53
3.5 Fluorescence Emission Spectroscopy.....	54
3.6 Kinetic Studies of Wild-Type AD and Variants.....	56
3.7 Molecular Docking of L-rogenate in the Active Site of AD from <i>Synechocystis</i> .....	58
3.8 Enzymatic synthesis of L-rogenate.....	64
3.8.1 Expression and Purification of Prephenate Aminotransferase from <i>Petunia Hybridia</i> .....	64
3.8.2 Coupled enzymatic PAT Assay for L-rogenate Synthesis.....	66
3.9 Purification of L-rogenate.....	67
3.9.1 Analytical HPLC and LC-MS.....	67
3.9.2 Preparative HPLC.....	72
<b>Chapter 4: Discussion.....</b>	<b>74</b>
<b>Chapter 5: Summary and Future Work.....</b>	<b>94</b>
<b>References.....</b>	<b>98</b>

## List of Figures

Figure 1: The pathway for aromatic amino acid bio synthesis.....	2
Figure 2: Biosynthesis of L-tyrosine and L-phenylalanine by the HPP/PP pathway and the L-rogenate Pathway.....	3
Figure 3: Proposed mechanism for the prephenate dehydrogenase reaction.....	8
Figure 4: Multiple sequence alignment of AD and PD domains of TyrA proteins.....	9
Figure 5: Selected active site residues of <i>A. aeolicus</i> Δ19PD complexed with NAD <sup>+</sup> at pH 3.2 with prephenate modeled in the active site.....	11
Figure 6: Selected active site residues of <i>A. aeolicus</i> Δ19PD complexed with NAD <sup>+</sup> and HPP or L-Tyr at pH 7.8.....	15
Figure 7: Crystal structure of <i>Synechocystis sp.</i> AD complexed with NADP <sup>+</sup> , pH 8.0.....	16
Figure 8: Selected active sitre residues of <i>Synechocystis sp.</i> AD complexed with NADP <sup>+</sup> at pH 8.0 with L-rogenate placed in the active site.....	21
Figure 9: The pathways of tyrosine and phenylalanine biosynthesis in <i>N. crassa</i> , <i>S. typhimurium</i> , and <i>S. typhimurium</i> tyrA19 mutant.....	23
Figure 10: SDS-PAGE analysis of wild-type AD purification.....	48
Figure 11: SDS-PAGE analysis of purified AD variants.....	49
Figure 12: Deconvoluted ESI-MS spectrum of wild-type AD enzyme.....	50
Figure 13: Far-UV spectra of wild-type and variant ADs.....	52
Figure 14: Variable temperature far-UV spectra of wild-type and select variant PD enzymes.....	52
Figure 15: Near-UV CD spectra of wild-type AD.....	54
Figure 16: Fluorescence emission spectra for selected AD enzymes.....	55
Figure 17: Fluorescence intensity of AD upon binding of NADP <sup>+</sup> .....	56
Figure 18: Alignment of the backbone atoms of <i>Synechocystis</i> AD and <i>A. aeolicus</i> Δ19PD crystal structures.....	62
Figure 19: Alignment of the backbone atoms of the selected <i>Synechocystis</i> AD model and its Δ19PD template.....	62
Figure 20: Active site of <i>Synechocystis sp.</i> AD with L-rogenate modeled into the active site of the NADP <sup>+</sup> -bound crystal structure and docked into the homology model.....	63
Figure 21: Comparison of arginine positions in the crystal structure of homology model of <i>Synechocystis sp.</i> AD.....	63
Figure 22: SDS-PAGE analysis of the purification of PAT from <i>Petunia hybrida</i> .....	65
Figure 23: Coupled PAT assay for L-rogenate synthesis .....	66
Figure 24: Time-dependant consumption of NADH by MDH-catalyzed reaction.....	67
Figure 25: Analytical HPLC chromatograms of standards and L-rogenate reaction mixture.....	69
Figure 26: Chromatogram of acidified experimental L-rogenate mixture and spectra of peaks corresponding to the masses of phenylalanine and NADH.....	71

Figure 27: Preparative HPLC chromatogram of L-arenogenate reaction mixture.....	73
Figure 28: Possible interactions of S178 and H179 with L-arenogenate.....	76
Figure 29: Possible interactions of H112, H170, and D171 in the NADP <sup>+</sup> -bound crystal structure of AD .....	79
Figure 30: Comparison of the location of the hypothesized hydrophobic pocket in the active site of <i>Synechocystis</i> AD .....	81
Figure 31: Potential interactions formed by G221 after being converted to glutamic acid in both the modeled and docked structures .....	86

## List of Tables

Table 1: Primer sequences for site-directed mutagenesis of <i>Synechocystis</i> AD.....	31
Table 2: Temperature cycling parameters for PCR using <i>Phusion® HF DNA Polymerase</i> .....	32
Table 3: Purification table for wild-type AD.....	48
Table 4: Summary of molecular masses of wild-type and variant AD proteins.....	50
Table 5: Specific activities for wild-type AD and variants.....	58
Table 6: Summary of RMSD and DOPE score values of AD homology models.....	60

## List of Abbreviations

Å	Angstrom
AD	Arogenate dehydrogenase
ADT	Arogenate dehydratase
Ala (A)	Alanine
Arg (R)	Arginine
Asn (N)	Asparagine
Asp (D)	Aspartic acid
bp	Base pair
CD	Circular dichroism
CM	Chorismate mutase
CM-PD	Chorismate mutase prephenate dehydrogenase
Da	Dalton
DNA	Deoxyribonucleic acid
EDTA	Ethylenediamine tetra-acetic acid
ESI-Q-ToF	Electrospray ionization Quadrupole Time-of-Flight
FA	Formic acid
Gln (Q)	Glutamine
Glu (E)	Glutamic acid
Gly (G)	Glycine
H-bond	Hydrogen bond
His (H)	Histidine
HPLC	High performance liquid chromatography
HPP	4-hydroxyphenyl pyruvate
HPpropionate	4-hydroxyphenyl propionate
IPTG	Isopropyl-β-D-thiogalactopyranoside
kcat	Turnover number
Kd	Dissociation constant
Ki	Inhibition constant
Km	Michaelis constant
LB	Luria-Bertani broth
LC-MS	Liquid chromatography mass spectrometry
Leu (L)	Leucine
Lys (K)	Lysine
MDH	Malate dehydrogenase
MS	Mass spectrometry
MW	Molecular weight
NAD(P) <sup>+</sup> / NAD(P)H	Oxidized/reduced nicotinamide adenine dinucleotide (phosphate)
Ni-NTA	Nickel-nitrilotriacetic acid
PAT	Prephenate aminotransferase
PCR	Polymerase chain reaction
Phe (F)	Phenylalanine
PD	Prephenate dehydrogenase
PDT	Prephenate dehydratase

PLP	Pyridoxal-5-phosphate
PMSF	Phenylmethylsulfonyl fluoride
PP	Phenylpyruvate
Pre	Prephenate
Pro (P)	Proline
SDS-PAGE	Sodium dodecyl sulfate- polyacrylamide gel electrophoresis
Ser (S)	Serine
Thr (T)	Threonine
Tris	Tris(hydroxymethyl)aminomethane
Trp (W)	Tryptophan
Tyr (Y)	Tyrosine
U	Units
UV	Ultraviolet
Val (V)	Valine
WT	Wild-type

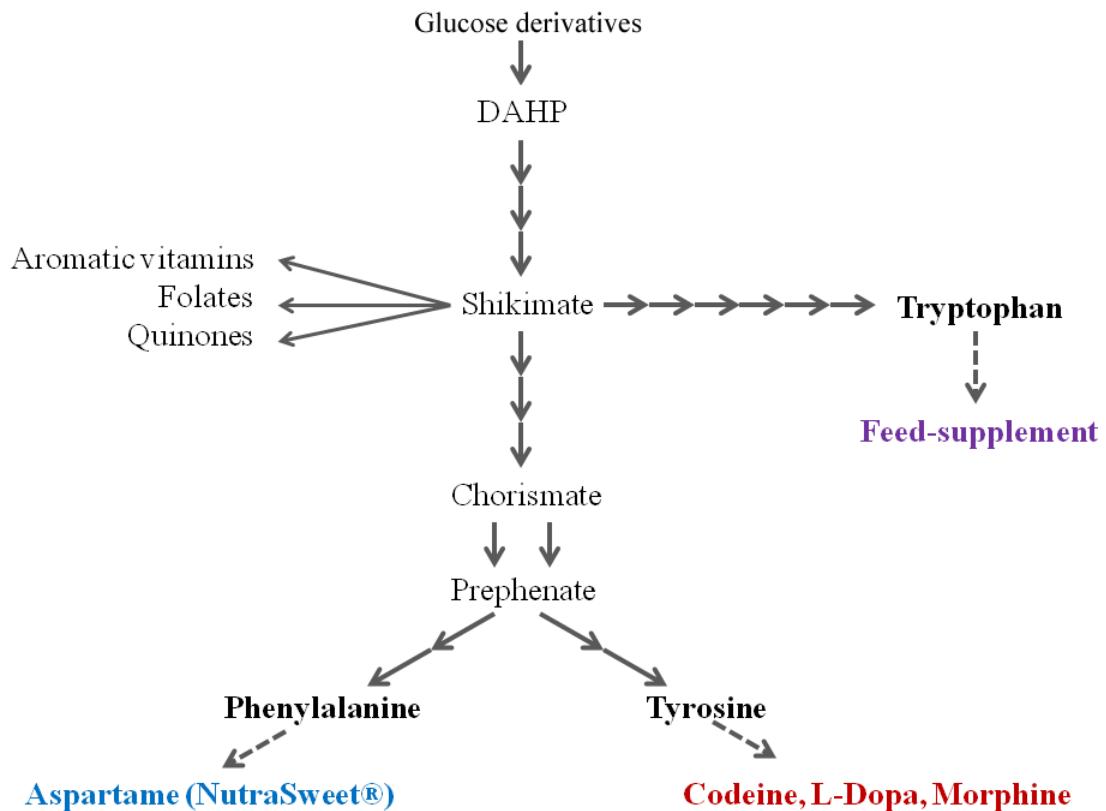
## **Chapter 1: Introduction**

### **1.1 Aromatic Amino Acids Biosynthesis**

The aromatic amino acids L-tyrosine (L-Tyr), L-tryptophan (L-Trp) and L-phenylalanine (L-Phe) are critical for the growth and survival of all living organisms. They are essential components for protein synthesis, but they also serve as precursors for a number of important aromatic metabolites such as flavonoids (1), quinones (2,3), cyanogenic glycosides (4) and alkaloids (5,6). Unlike mammals, who require these essential amino acids from their diet, archae- and eubacteria, fungi and plants have the ability to synthesize them. The metabolic pathways involved in this synthesis are the “shikimate” and the “phenylpyruvate/4-hydroxyphenylpyruvate (PP/HPP)” or “arogenate” pathways. Considering the absence of these metabolic pathways in mammals, these enzymes are targets for the design of inhibitors that can act as herbicides, fungicides or antimicrobial agents (7,8). Furthermore, many of these enzymes are of interest to the food, pharmaceutical, and agricultural industries as their products can serve as precursors for the synthesis of commercially valuable compounds such as drugs and biodegradable polymers (9, 10, 11, 12).

The shikimate pathway (Figure 1) links carbohydrate metabolism to the biosynthesis of aromatic compounds (13,14). In this pathway, glucose and other six-carbon sugars are converted through a series of seven enzymatic steps to chorismate, which then serves as a branch point for the biosynthesis of several intermediates. The first reaction of the pathway is the condensation of erythrose-4-phosphate and phosphoenolpyruvate, to yield the seven-carbon compound 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP). The reaction is catalyzed by 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase

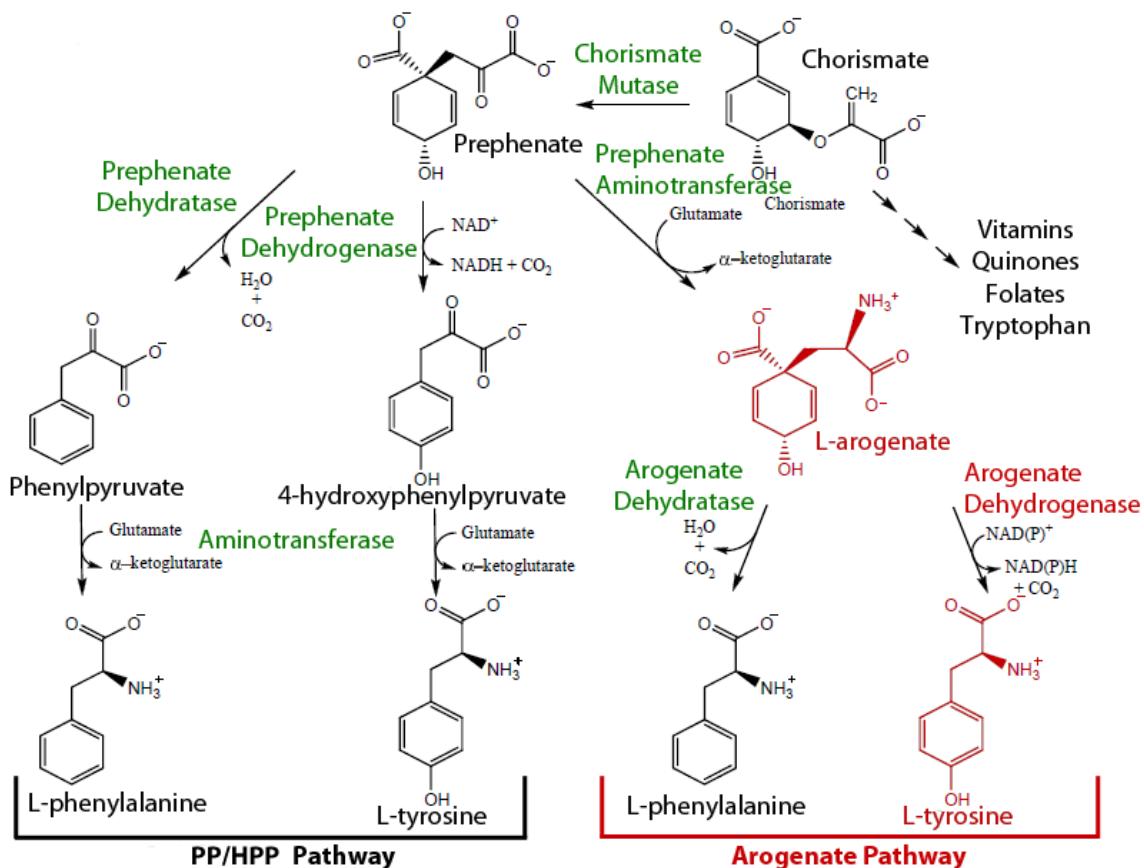
(DAHP synthase), whose isozymic forms are sensitive to regulation by each of the aromatic end products. In the steps that follow, DAHP is cyclised to yield shikimate, which is then converted to chorismate.



**Figure 1: The pathway for aromatic amino acid biosynthesis.** The shikimate pathway consisting of the first seven enzyme-catalyzed steps ending in the production of chorismate; DAHP denotes 3-deoxy-D-arabino-heptulosonate.

The remaining biosynthetic steps leading to L-Tyr and L-Phe can occur by two possible routes, shown in Figure 2. First, chorismate undergoes a Claisen rearrangement to the common intermediate, prephenate. This reaction is catalyzed by chorismate mutase (CM). Prephenate is then either oxidatively decarboxylated by prephenate dehydrogenase (PD) in the presence of  $\text{NAD}^+$  to form *p*-hydroxyphenylpyruvate (HPP) and carbon dioxide, or dehydrated and decarboxylated by prephenate dehydratase (PDT) to form

phenylpyruvate (PP). HPP or PP are then transaminated to form L-Tyr or L-Phe, respectively. This branch of the pathway is known as the “HPP/PP pathway”. In the “arogenate pathway”, prephenate is transaminated by prephenate aminotransferase (PAT) using either L-glutamate or L-aspartate as amino donors, yielding L-arogenate. L-arogenate is then either oxidatively decarboxylated by arogenate dehydrogenase (AD) in the presence of NADP<sup>+</sup> to form L-Tyr and carbon dioxide, or dehydrated by arogenate dehydratase (ADT) to form L-Phe. Both amino acid end products function as feed-back inhibitors of the CM-catalyzed reaction; this inhibition is also observed with PD, PDT, and some AD and ADT enzymes. The focus of this study is on the reaction catalyzed by AD, shown in bold and red (Figure 2).



**Figure 2: Biosynthesis of L-tyrosine and L-phenylalanine by the phenylpyruvate/4-hydroxyphenylpyruvate (PP/HPP) pathway and the L-arogenate pathway.**

## 1.2 The TyrA Protein Family

The TyrA protein family consists of enzymes that catalyze a cofactor-dependent irreversible decarboxylation step in a terminal step in L-Tyr biosynthesis. These homologous dehydrogenases are categorized according to their substrate specificity (15, 16). The two cyclohexadienyl substrates that are accepted by TyrA proteins are either prephenate or L-arogenate. Prephenate and L-arogenate are structurally very similar; the pyruvyl side chain in prephenate is replaced by an alanyl group in L-arogenate. Despite this similarity, most TyrA proteins are highly exclusive to either prephenate or L-arogenate. TyrA<sub>p</sub> proteins exclusively decarboxylate prephenate to yield HPP; TyrA<sub>a</sub> proteins exclusively decarboxylate L-arogenate to yield L-Tyr, and TyrA<sub>c</sub> proteins are promiscuous for either prephenate or L-arogenate. In addition to their selected cyclohexadienyl substrate, these enzymes also require the presence of either NAD<sup>+</sup> or NADP<sup>+</sup> as a nucleotide cofactor (15). Generally, TyrA<sub>p</sub> dehydrogenases prefer NAD<sup>+</sup> as a co-substrate, while TyrA<sub>a</sub> dehydrogenases prefer NADP<sup>+</sup> (17, 18). The different TyrA categories are also generally linked to specific organisms, as deduced through extensive phylogenetic analysis and sequence alignments (reviewed by Jensen and colleagues (16, 27)): bacteria and yeast are hypothesized to possess TyrA<sub>p</sub> or TyrA<sub>c</sub> activity, while plants and a few bacteria (such as cyanobacteria) possess TyrA<sub>a</sub> activity.

As TyrA proteins are always responsible for the terminal decarboxylation step in L-Tyr biosynthesis, they all share a core catalytic domain of approximately 30 kDa, are catalytically active as dimers, and maintain the same fundamental reaction group chemistry. However, a number of TyrA proteins are multi-functional due to the presence of fused domains. Some examples of bifunctional TyrA proteins are as follows: PDs from

*E. coli* and *Haemeophilus influenza*, which are coupled with CM activity associated with the N-terminal region of the protein (19); PDs from *Pseudomonas fluorescens* and *Acinetobacter calcoaceticus* are bifunctional with phosphoshikimate carboxyvinyltransferases activity at the proteins' C-terminal portion (20). Trifunctional PDs also exist, as seen in *Nanoarchaeum equitans* and *Archaeoglobus fulgidus* (21); these enzymes possess CM, PD, and PDT activities in separate domains. This work is primarily interested with the monofunctional TyrA proteins, examples of which include: PDs from *Aquifex aeolicus* (22) and *Bacillus subtilis* (23), ADs from *Nicotiana silvestris* (24) and *Synechocystis sp.* (15), and the cyclohexadienyl dehydrogenases from *Pseudomonas stutzeri* (23) and *Zymomonas mobilis* (26).

Despite the existence of TyrA proteins in many different bacteria, fungi, and plants (15, 27), relatively few have been purified and characterized. The most well characterized TyrA protein is the bifunctional CM-PD from *E. coli*; the kinetic and biophysical properties of both wild-type and variant forms have been reported (19, 28, 29, 30), although there is no crystal structure available for the bifunctional enzyme. Another well-characterized TyrA<sub>p</sub> protein is PD from the hyperthermophilic bacterium *Aquifex aeolicus* (22). The crystal structure was determined both in the presence and absence of product and product analogs, and is the only monofunctional TyrA<sub>p</sub> structurally well characterized to date (31). TyrA<sub>p</sub> proteins from *Streptococcus mutans* (32) and *Mycobacterium tuberculosis* (33) have also been studied, while TyrA<sub>c</sub> from *Z. mobilis* and *Pseudomonas stutzeri* are two of the only TyrA<sub>c</sub> proteins to have been characterized (23, 26, 27). From the TyrA<sub>a</sub> protein class, isozymic forms from

*Arabidopsis thaliana* have been reported (56) and the crystal structure and some kinetic data is available for AD from *Synechocystis sp* (15, 48).

### **1.3 The Active Sites of the Bifunctional CM-PD from *E.coli***

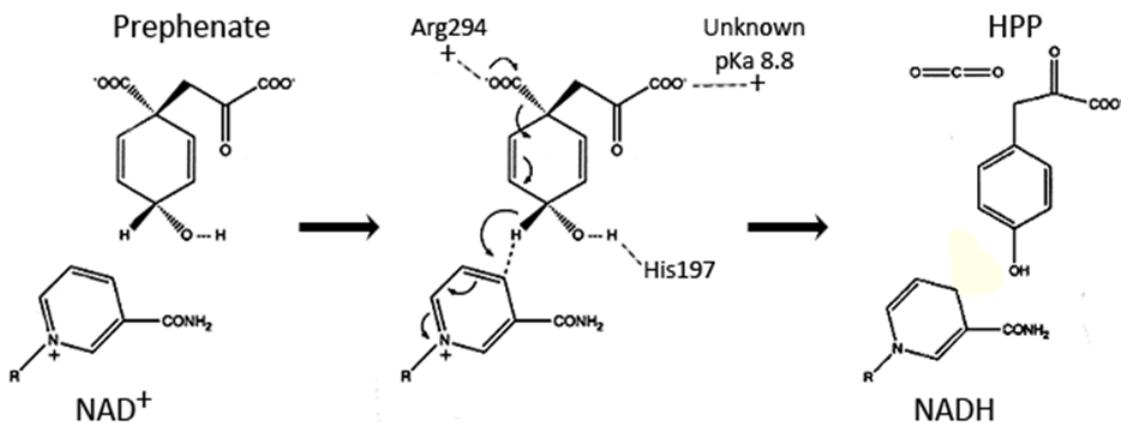
CM-PD from *E. coli* is a homodimer with a molecular weight of ~ 42 kDa per monomer. In the absence of a crystal structure, residues that compose each domain were determined by aligning the primary amino acid sequence of the enzyme with that of *E. coli* CM-PDT. It was found that the first 109 amino acids of the sequence likely encode the CM domain of the enzyme, while the remaining 277 residues constitute the PD domain. Initial kinetic and protein chemistry studies on CM-PD supported the hypothesis that both CM and PD activities occur at a common active site. (34, 35) More recent studies, however, have suggested that the enzyme maintains two separate active sites for each reaction; both the CM and PD activities were independently inhibited by substrate analogs that mimic the transition states of the CM and PD reactions (28, 35). This hypothesis was further enforced by mutagenesis studies whereby K37A and H197N variants were reported to selectively abolish CM and PD activities, respectively (36, 37). Studies also support the idea that the two active sites may be structurally and functionally interdependent. NAD<sup>+</sup>, the cofactor for the PD-catalyzed reaction, activates the mutase reaction (19). Efforts to genetically separate CM-PD's two domains have lead to proteins that are inactive or unstable (81). Several site-specific replacements in the dehydrogenase domain also affect mutase activity; for example, the K178R, H189N, R286A variants were found to have reduced CM and PD activities (37).

Mutagenesis studies on *E. coli* CM-PD have also helped to identify residues that are critical for chorismate and prephenate binding (37), although most mutagenesis work on CM is reported for the independently expressed mutase domain from *E. coli* CM-PDT, guided by a crystal structure of the “mini-mutase” bound with a transition-state analog (82). The residues involved in prephenate binding, however, are of particular interest in the present study. For example, K178R and R294Q variants from CM-PD were reported to have increased the  $K_m$  for prephenate 36- and 120-fold, respectively (37). Additionally, both cationic residues are conserved amongst all TyrA proteins (Figure 4).

#### **1.4 Proposed Mechanism of the Prephenate and Arogenate Dehydrogenases**

The proposed catalytic mechanism for the prephenate dehydrogenase reaction is shown in Figure 3. In PDs, prephenate is oxidatively decarboxylated in a  $\text{NAD}^+$ -dependant reaction, to form HPP and carbon dioxide. Similarly, in ADs, arogenate is decarboxylated in a  $\text{NADP}^+$ -dependant reaction, to form L-Tyr and carbon dioxide. The formation of the aromatic products is the driving force of this irreversible reaction. There are two chemical processes that must occur in this reaction: first, the ring carboxyl group must leave to form carbon dioxide, and second, the subsequent hydride transfer from the C-4 of prephenate to the C-4 of the nicotinamide ring of the cofactor.  $^{13}\text{C}$  and deuterium kinetic isotope studies on *E. coli* CM-PD, which monitor the cleavage of the C-C bond in the decarboxylation step, revealed that for this enzyme-catalyzed reaction, these two steps occurred concomitantly (38).

In order to help identify the residues involved in the PD-catalyzed reaction of *E. coli* CM- PD, the  $\text{pK}_a$  values of these residues were determined by analyzing pH rate



**Figure 3: Proposed mechanism for the prephenate dehydrogenase reaction.** Prephenate and NAD<sup>+</sup> bind to different cavities at the active site of PD domain. His197, a deprotonated group with pK<sub>a</sub> value of 6.5 is believed to assist hydride transfer from 4-hydroxyl group of prephenate to NAD<sup>+</sup> and concomitant decarboxylation by polarizing prephenate's 4-hydroxyl group. This and the formation of aromatic product, HPP, provide the driving force to lower the activation energy barrier. Arg294 was proposed to interact with the ring carboxylate of prephenate bound to the enzyme-NAD<sup>+</sup> complex. In addition, another protonated group (identity unknown) with pK<sub>a</sub> value of ~8.8 is believed to interact with the pyruvyl side chain of prephenate to lock prephenate in the active site. This mechanism was proposed by Christendat and Turnbull (36,37).

profiles (29, 36, 38). The variation of log V/E<sub>t</sub> as a function of pH identified one ionisable group (pK<sub>a</sub> 6.8) which must be deprotonated for maximum activity and is involved in catalysis and/or product release (38). However, the plot of (V/E<sub>t</sub>)K<sub>prephenate</sub> vs. pH indicated the presence of another residue, with a pK<sub>a</sub> value of ~8.8, involved in the binding of prephenate to the enzyme-NAD<sup>+</sup> complex. The identity of this residue has yet to be determined. The catalytic residue, however, was identified by chemical modification and site-directed studies as H197 (25, 30, 36). The PD activity of the H197N variant was dramatically reduced, while CM activity was unaffected (36). Interestingly, and the K<sub>m</sub> for prephenate and NAD<sup>+</sup> for this variant were found to be similar to

<i>E_coli</i>	PP--LPKDCILVDSLAVKNGPLQAMLVAHDGP--VLGLHPMPGPDGSGLAKQVVVWCDG	217
<i>H_influenzae</i>	KPY-LTENMLLADLTSVKREPLAKMLEIHTGA--VLGLHPMPGADIASMAKQVVVRCDG	220
<i>S_pneumoniae</i>	ANLDLKEGIISDAGSTKAAIVEVAEEYLAGKSIRFVGAGHPMAGSHKTGAASADVNLFEN	142
<i>B_subtilis</i>	AHSGIEHELLITDVGSTKQKVVVDYADQVLPQR-YQFVGGHPMAGSHKGVAAAKEFLFEN	144
<i>Z_mobilis</i>	AP-ALKDVIICDTGSVKVSVIKTLQDNLPNH-IIIVPSHPLAGTENNGPDAGFAELFQD	141
<i>A_aeolicus</i>	SY-ILSEDATVTIDQGSVKGKIVYDLENILGKR--FVGGHPPIAGTEKSGVEYSLDNLYEG	167
<i>H_pylori</i>	TP--IKKSATIIDLGGAQAKILHNIPKSIRQN--FIAAHPMCTEFYGPKAQVKGLYEN	133
<i>Synechocystis</i>	IP-HLSPTAIVTDVASVTKTAIAEPASQLWSGF---IGGHPMAGTAQGIDGAEENLFVN	132
<i>S_cerevisiae</i>	GP-SSKVGTIVGGQTSCKLPEIEAFEKYLPKD-CDIITVHS LHGPKVTEGQPLVIINHR	150
	: . . * : * . : * . : *	
<i>E_coli</i>	RK-----PEAYQWFLEQIQVW GARLHRISAVEHDQNMFIQALRHFA TFAYGLHL	267
<i>H_influenzae</i>	RF-----PERYEWLLEQIQIWGAKIYQTNATEHDHNMTYI QALRHFA STFANGLHL	270
<i>S_pneumoniae</i>	AYYIFTPSSLTTPETLSEMKDLLSGLHARFIEIDAKEHDQVTSQISHFPHILASG----L	198
<i>B_subtilis</i>	AFYILTPGQKTDKQAVEQLKNLLKGTNAHFVEMSPPEEHDGVTSVISHFPHIVAAS----L	200
<i>Z_mobilis</i>	HPVILTPDAHTPAQAIAYIADYWE EIGGRINLMSAEHHHDHV LALTSHLPHVIA YQ---LI	198
<i>A_aeolicus</i>	KKVILTPKKTDKKRLKLKVKR VWE DVGGVVEYMSPELHDYVFGVVS HLPHAVAFA----L	223
<i>H_pylori</i>	ALVILCDLEDSTGTQV ELAKEIF LGIKARLIK MKSNEHDTHVAYISHLP HVLSYA----	188
<i>Synechocystis</i>	APYVLTPTEYTDPEQLACLRSVLEPLGVKIYLCTPADHQAVAWISHLPVMVSAA--LIQ	190
<i>S_cerevisiae</i>	SQY-----PESFEFVN S VMA CLSKQVYLT YEEHD KITADT QAV THA A FLS MGS AW	201
	*: . . : : . . : :	
<i>E_coli</i>	AEENVQLEQL--LALSSPIYRLELAMVGR LF AQDPQLYADIIMSS-ERNLALIKRYYKRF	324
<i>H_influenzae</i>	SKQP IN LANL--LALSSPIYRLELAMIGR LFAQDAE LYADIIMDK-PENLAVIETI LKQTY	327
<i>S_pneumoniae</i>	MEQTAVYAQE--HEMARRFAAGGFRDMTRIAESEP GMWT SILL SNRETILERIADFKERL	256
<i>B_subtilis</i>	VHQTHHSEN L--YPLVKRFAAGGFRDTRIA ESSSPAMWRDILLHNKDKILD RFDE WIREI	258
<i>Z_mobilis</i>	GMVSGYEKK S--RTPIMRYSAGSF RDAT RVAASE PRLWQD IMLENAPALLPVLDHFIADL	256
<i>A_aeolicus</i>	VDTL IHMSTP--EV DLF KYP GGGFKDTRIAKSDPIMWRDIFLENKENVMKAIEGFEKSL	281
<i>H_pylori</i>	LANS VLKQND--PEMILS LAGGGFRDMS RLSKSSPLMWKDFKQNRDNV LEAIEKCEK I	246
<i>Synechocystis</i>	ACAGEK DGD I--LKLAQNLASSGF DTSVGGCNP ELS T MATYNQR ALLKSLQDYRQHL	248
<i>S_cerevisiae</i>	AKIKIY PWT LGV NKWYGGLEN V KV NISL RIYSNKWHVYAGLAI TNPSAHQQ ILQY AT SAT	261
	*: . . : : . . : :	
<i>E_coli</i>	GEAI ELL E QGD KQAFIDS FRK VEH WFG-----	351
<i>H_influenzae</i>	EEALSFF ENND RQGF IDA FH KVR DWFG-----	354
<i>S_pneumoniae</i>	EAIGQAI SRG DEEQIWNFFFNQARA QRQAME-----IH KRGGV DSSYDLYVDV	303
<i>B_subtilis</i>	DKIR TYVEQEDAENLFRYF KTA KDYRD GLP-----LRQKGAIPAFYDLYVDV	305
<i>Z_mobilis</i>	KKLRTA IASQD GDY LLH FKES-----	278
<i>A_aeolicus</i>	NHLKELIV REA EEE LVE YLKEVK-----	304
<i>H_pylori</i>	AQAKAWIEN NDYES LAEW MAQAN-----	269
<i>Synechocystis</i>	DQLITL ISN QWPEL HRT LQQT N-----	271
<i>S_cerevisiae</i>	ELFSLMIDNKEQELTD RLLKAKQFVFGKHTG LLLDDTILEK YSL SKSSIS DSN NCK PVP	321
	:	

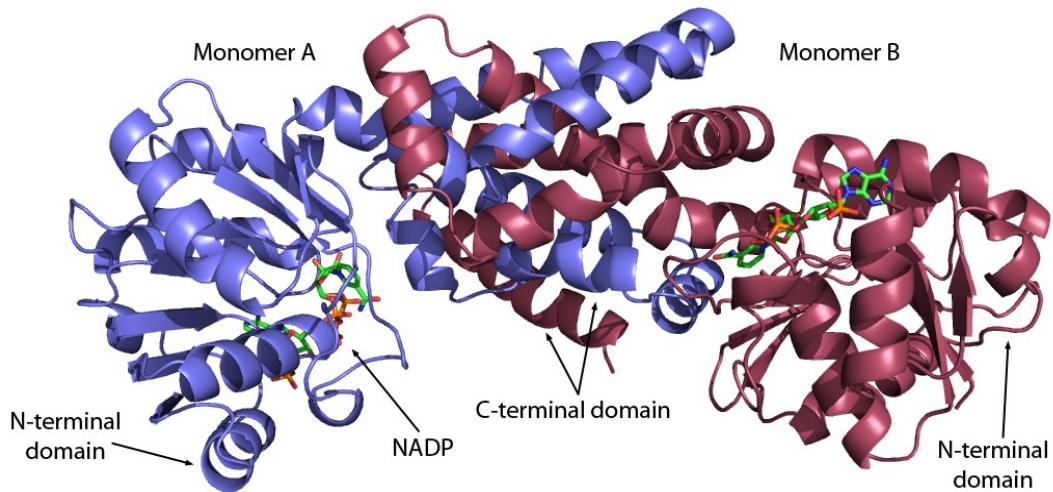
**Figure 4: Multiple sequence alignment of AD and PD domains of TyrA proteins.** The sequences of the following species are listed: Bifunctional CM-PDs from *Escherichia coli* and *Haemophilus influenzae*, *Streptococcus pneumoniae*, *Bacillus subtilis*, *Zymomonas mobilis*, *Aquifex aeolicus*, *Helicobacter pylori*, *Synechocystis* sp., and *Saccharomyces cerevisiae*. Residues conserved to identity are highlighted in yellow, and include the catalytic H-bond acceptor (H112 in *Synechocystis* sp.) and the cationic arginine interacting with prephenate's carboxylate group in PD enzymes (R217 in *Synechocystis* sp.) Residues studied in this thesis are in green. The multiple sequence alignment was performed using ClustalW2.

the values reported for wild-type enzyme, suggesting that H197 is involved solely in catalysis. H197 is also a highly conserved residue in the TyrA proteins (Figure 4). Another conserved residue, R294, was found to be critical for the binding of prephenate to the enzyme-NAD<sup>+</sup>-complex, as the R294 to glutamine replacement increased the K<sub>m</sub> for prephenate by 120-fold without affecting catalysis (37). Further studies with inhibitory substrate analog studies were carried out in order to probe R294's interaction with prephenate. A series of HPP analogues, all missing the ring carboxyl group, were found to have similar dissociation constants for wild-type enzyme and the R294Q variant (37), suggesting that an electrostatic interaction exists between the arginine and the ring carboxylate. This finding contradicts the hypothesis of Hermes et al. (38), in which they proposed that the ring carboxyl group might reside in a hydrophobic pocket in order to promote decarboxylation concomitant hydride transfer.

## 1.5 Crystallographic Studies on Different TyrA Proteins

It is only recently that crystallographic data has become available to guide mutagenesis studies on TyrA proteins. In 2006, Christendat, Turnbull and colleagues (31) reported the first crystal structure of a TyrA protein; an N-terminally deleted variant of *A. aeolicus* PD ( $\Delta$ 19PD) was co-crystallized with NAD<sup>+</sup> at pH 3.2, yielding crystals that diffracted to 1.9 Å resolution (22, 40). It was not until 2009, however, that the same authors reported the structures of  $\Delta$ 19PD in complex with NAD<sup>+</sup> plus either L-Tyr (at 2.21 Å), HPP (at 2.15 Å) or hydroxyphenyl propionate (at 2.25 Å) (31). Also in 2006, and pertinent to the present study, Legrand et al. reported the first structure of AD from *Synechocystis sp* in complex with NADP<sup>+</sup>, at a resolution of 1.55 Å (48). Its diagram is

shown in Figure 5. Since then, two other PDs have been crystallized but in the absence of any ligands and their structures reported – that of PD from *Streptococcus thermophilus* at 2.46 Å resolution in 2008, and PD from *Corynebacterium glutamicum* at 2.60 Å resolution in 2010. The most recent crystal structure is that of a variant of *H. influenzae* CM-PD ( $\Delta$ 80 CM-PD) in which the first 80 residues of the CM domain had been genetically deleted and the protein recombinantly expressed in *E. coli*. The structure of  $\Delta$ 80 CM-PD in complex with NAD<sup>+</sup> and L-Tyr was reported at a resolution of 2.0 Å (47).



**Figure 5: Crystal structure of *Synechocystis* sp AD complexed with NADP<sup>+</sup>, pH 8.0.** The two monomers are coloured in blue and burgundy. Each monomer can be divided into a N-terminal dinucleotide-binding domain and a C-terminal dimerization domain, which have been labelled. NADP<sup>+</sup> is represented as green sticks. This figure was created using PyMOL (44).

Despite having moderately low sequence identity, the crystallized TyrA proteins display remarkably similar overall structural folds. The sequence identity between *Synechocystis* sp AD and *A. aeolicus* PD is only 29%, but the monomers of the NADP<sup>+</sup>-bound AD and NAD<sup>+</sup>-tyrosine-bound Δ19PD have a backbone RMSD alignment of 1.3 Å. As such, the structures of *A. aeolicus* Δ19PD have been utilized to facilitate modelling and ligand docking studies in this thesis. The level of identity of AD with the other PDs

and the PD domain of bifunctional CM-PDs are all very similar: *Synechocystis sp* AD shares 26% identity with *H. influenzae* CM-PD, 23% with *E. coli* CM-PD, 28% identity with *C. glutamicum* PD, and 31% with *S. thermophilus* PD. The two enzymes that share the highest sequence identity, at 57%, are the CM-PD enzymes from *E. coli* and *H. influenzae*. The structure of *H. influenzae* Δ80 CM-PD is currently being used in the Turnbull lab to guide further mutagenesis studies on CM-PDs from *E. coli* and *H. influenzae*.

A representative structure of AD from *Synechocystis sp* in complex with NADP<sup>+</sup> is shown in Figure 5. The enzyme is dimeric and comprised of identical monomers containing a C-terminal dimerization domain, and an N-terminal nucleotide binding domain. Each monomer contains one active site, which is located at the interface of the two domains and encompasses amino acids from both monomers. The overall geometry of the active site between the five crystallized enzymes is reasonably well conserved, particularly with respect to several residues that have been identified to be important for binding and catalysis in both *E. coli* CM-PD and *A. aeolicus* PD.

## 1.6 Prephenate Dehydrogenase from *Aquifex aeolicus*

*A. aeolicus* is a microaerophilic, hydrogen-oxidizing, obligate chemolithautotroph organism, originally isolated from thermal vents found in Yellowstone National Park and can thrive at temperatures up to 95 °C. Its genome was sequenced (39) revealing the identity of a putative *tyrA* gene, which was predicted to encode a protein of 311 residues. The protein was recombinantly expressed in *E. coli* and both the full-length and crystallisable Δ19PD forms have been kinetically and biophysically characterized (22),

confirming that they are thermophilic PDs. The enzyme has an absolute requirement for NAD<sup>+</sup> as a cofactor and prefers prephenate as the cyclohexadienyl substrate to L-  
arogenate by a factor of 241-fold (22). Both the full-length and Δ19PD forms are  
sensitive to feedback inhibition by L-Tyr.

To gain further insight into the role of active site residues in substrate binding and catalysis, Bonvin (22), Hou (43) and Hotz (42) completed the first site-directed mutagenesis studies on Δ19PD, guided by sequence alignments and the available crystal structures of the enzyme bound with NAD<sup>+</sup> and in complex with NAD<sup>+</sup> and HPP or L-Tyr (Figure 7)). Given the structural similarities between prephenate, L-arogenate and L-Tyr, it was hoped that such studies might also shed light on the factors that determine substrate specificity in TyrA proteins. The residues that are summarized here are of particular interest in this study, as they are all conserved in *Synechocystis* AD, and subjected to investigation by mutagenesis.

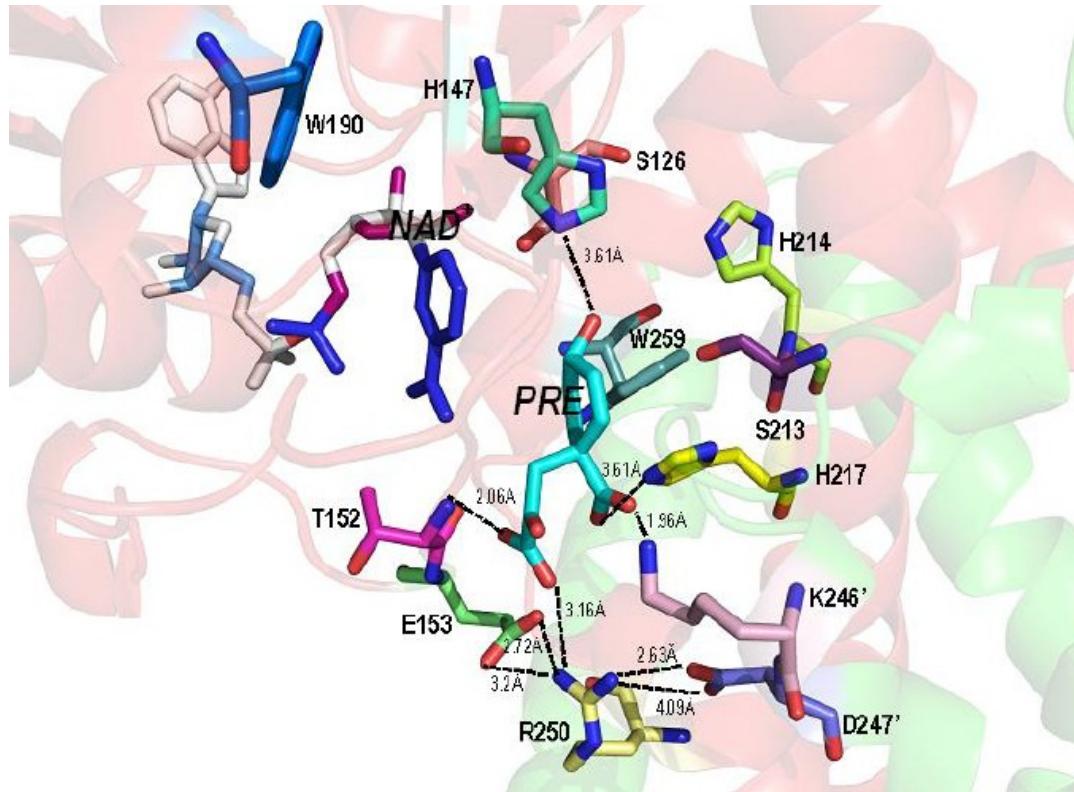
The results from the mutagenesis studies revealed that a number of residues were important for catalysis, particularly H147 and H205. **H147** was identified as the catalytic hydrogen bond acceptor in the reaction (H197 in *E. coli* CM-PD) (31); the H147N variant was found to be inactive, but maintains an apparent affinity for prephenate similar to that of the wild-type enzyme. The structures of liganded Δ19PD also revealed that **H205** is within hydrogen bonding distance of H147. Accordingly, the H205L variant was 772-fold less active than wild-type enzyme, suggesting that this replacement perturbs the correct orientation of H147 within the active site.

Other residues were identified to play a role primarily in prephenate binding without affecting catalysis or the binding of NAD<sup>+</sup> (R250, K246', T152). Two cationic residues,

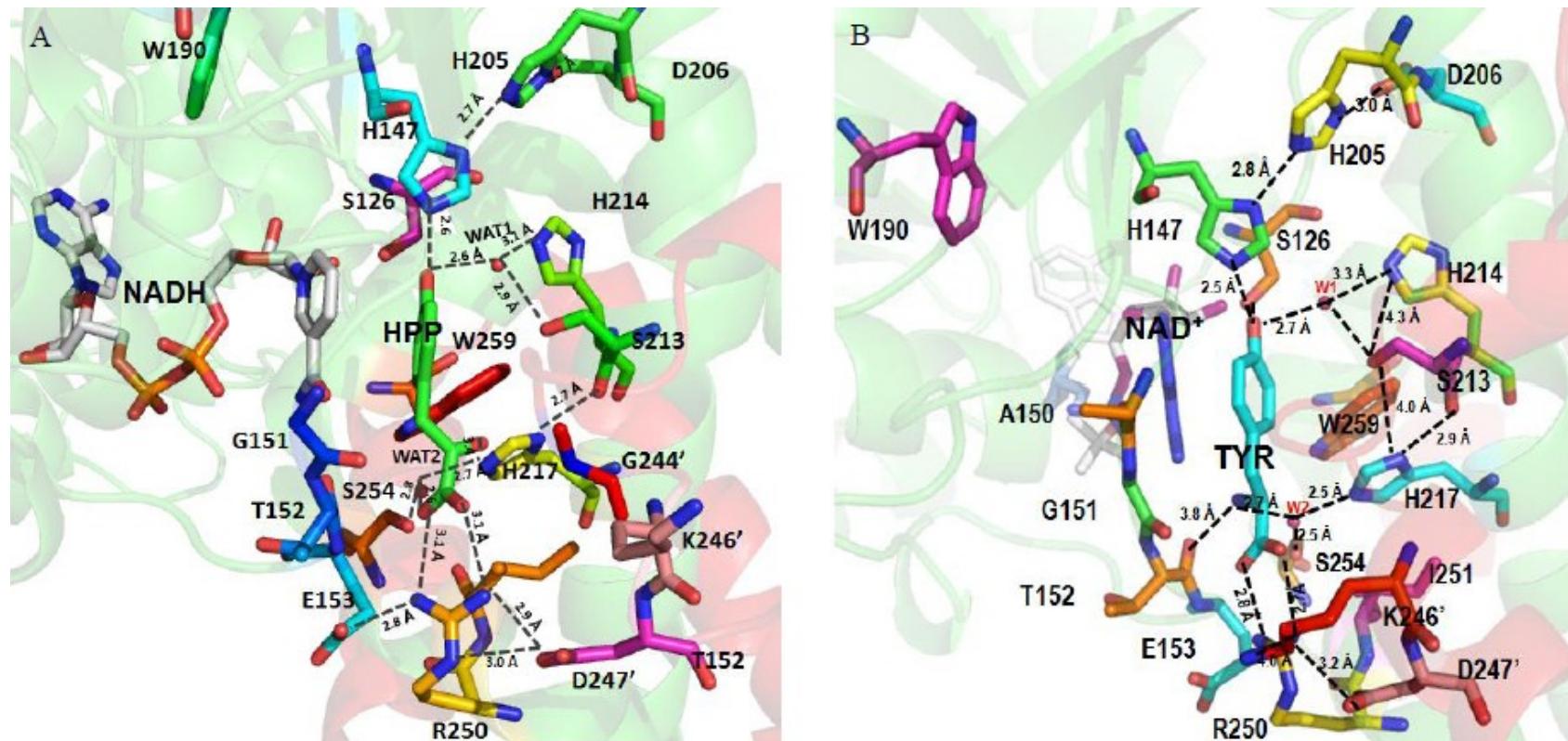
**R250** and **K246'**, when replaced with alanine individually, resulted only in a 10-fold and 20-fold increase in  $K_m$  for prephenate, respectively (42, 43). This was especially surprising for R250A since replacement of the homologous residue in *E. coli* CM-PD (R294Q) resulted in over a 100-fold increase in  $K_m$  for prephenate (37). Although, not critical for prephenate binding individually, presumably R250 and K246' might act in concert, as the  $K_m$  for prephenate of the double variant was increased by 221-fold (42). Docking prephenate into the active site of the NAD<sup>+</sup>-bound crystal structure (42) revealed that these two residues likely form electrostatic interactions with prephenate: K246 with the ring carboxyl group, and R250 with the chain carboxyl group (Figure 6). **T152** resides on a flexible loop in the active site, and interacts with NAD<sup>+</sup> in the crystal structure of PD. Docking experiments with prephenate revealed a potential hydrogen bond with the backbone amine of T152 (42). Additionally in the L-Tyr and NAD<sup>+</sup>-bound crystal structure, the amine of L-Tyr was found to have an interaction with the backbone carbonyl group of T152 (Figure 7).

Many TyrA proteins, including AD from *Synchocystis sp*, maintains a conserved **SH pair** in the active site. In *A. aeolicus* PD, **S213** is thought to be important in the binding of prephenate by helping to position a conserved water molecule, WAT1, near the C-4 hydroxyl group of the substrate or product HPP (Figure 7). WAT1 is coordinated through a network of hydrogen bonds involving S213, H214, H217, and the substrate's hydroxyl group. S213's role in binding was confirmed when the S213A variant displayed a 21-fold increase in  $K_m$  for prephenate, without affecting NAD<sup>+</sup> binding. Mutagenesis studies also revealed that of the SH pair, S213 is the more important than **H214** for this interaction (42). Finally, **D206**, a residue conserved in all TyrA proteins, was found to be critical for

enzyme function. The D206A variant displayed no detectable activity as well as a significantly reduced yield when expressed and purified, relative to the wild-type enzyme and other variants (42).



**Figure 6: Selected active site residues of *A. aeolicus* Δ19PD complexed with NAD<sup>+</sup> at pH 3.2 with prephenate modeled in the active site.** The primed residues denote those groups associated with the adjacent monomer (43). Figure and docking by Hotz (42).



**Figure 7:** Selected active site residues of *A. aeolicus* Δ19PD complexed with NAD<sup>+</sup> and (A) HPP or (B) L-Tyr at pH 7.8. Pictures created using PyMOL (44) using the coordinates derived from reference (31). The primed residues denote those groups associated with the adjacent monomer. Figure by Hotz (42).

## **1.7 Contrasting L-Tyrosine Feedback Inhibition in *E. coli* and *H. influenzae* CM-PD and *A. aeolicus* PD**

Many TyrA proteins are at a branch point in the pathways for the biosynthesis of L-Phe and L-Tyr; the ability to regulate enzyme activities is necessary for shunting metabolites to the appropriate aromatic amino acid end product. Thus, certain TyrA enzymes are highly regulated by L-Tyr. Such is the case for both the dehydrogenase and mutase activities of *E. coli* CM-PD, as well as *A. aeolicus* PD, which are feedback inhibited by L-Tyr. With both enzymes, this inhibition is dependent on the presence of NAD<sup>+</sup>. Double reciprocal plots of velocity at varying prephenate concentrations and fixed, increasing concentrations of L-Tyr are notably concave upward, suggestive of cooperative interactions between subunits in the binding of L-Tyr (18, 28, 35, 41).

Several models for *E. coli* CM-PD have been proposed to explain these kinetic findings. The results of analytical ultracentrifugation experiments with *E. coli* CM-PD suggest that the enzyme exists in active dimer and inactive tetrameric forms and that NAD<sup>+</sup>/L-Tyr shifts the equilibrium in favour of the tetramer (41). In contrast, Christopherson and colleagues (18), proposed that the concave upward kinetics (observed at low enzyme concentrations) could be due to a tertiary structural changes such that the binding of L-Tyr to one subunit promotes the binding of a second tyrosine molecule to the other subunit (18, 45). In both of these models, L-Tyr was presumed to bind at the active site. Turnbull and colleagues (29) proposed an alternate mechanism based on the results of fitting initial velocity data to several models and by analyzing the patterns of inhibition in the presence of both HPP and L-Tyr. They hypothesized that tertiary structural changes that occur upon L-Tyr binding occurs by L-Tyr's interaction at an allosteric site, since L-

arogenate (which also possesses a side chain amino group) is a poor substrate for the PD reaction; the L-Tyr-enzyme-prephenate complex remains active although markedly less so than the enzyme-prephenate complex. Interestingly, neither *A. aeolicus* PD or *H. influenzae* Δ80 CM-PD undergo tetramerization in the presence of L-Tyr, thus favouring a model of tertiary structural changes (22, 49). Additionally, crystal structures of these two enzymes in complex with NAD<sup>+</sup> plus L-Tyr show the end product bound at the active site (31,40, 47), which argues against the presence of an allosteric site for L-Tyr to bind.

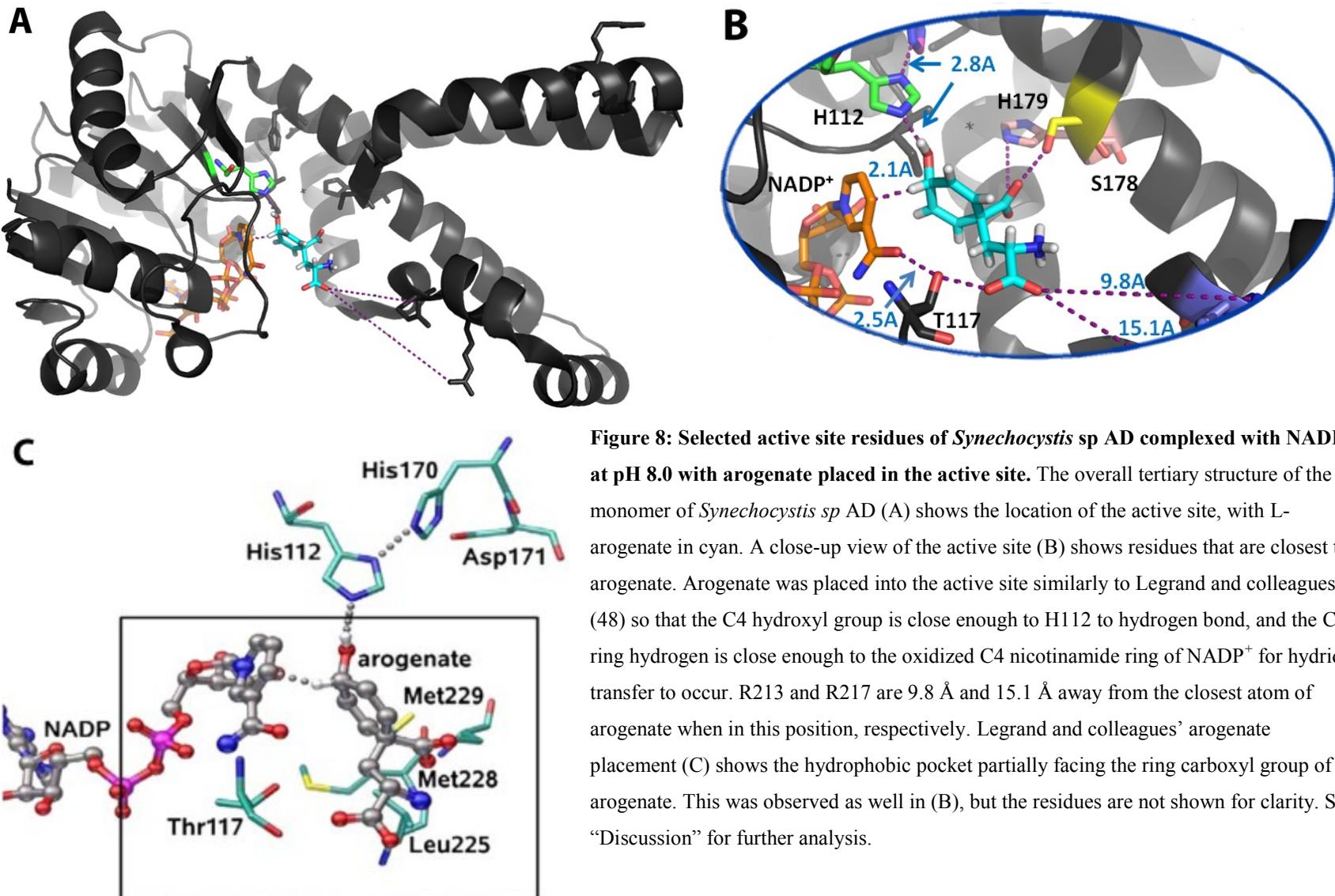
### **1.8 Arogenate Dehydrogenase from *Synechocystis* sp. PCC 6083**

*Synechocystis* sp. PCC6803 is a freshwater cyanobacterium, and is one of the most highly studied in its class. It is an autotroph when in the presence of sunlight and a heterotroph when sunlight is absent. The arogenate (formerly called pretyrosine) pathway was first discovered by Stenmark and colleagues (57) in cyanobacteria. As no PD activity could be detected in this organism, they proposed that L-Tyr biosynthesis must proceed via an alternate intermediate. It was not until 2004 that a putative *tyrA* gene encoding a monofunctional AD was identified from the sequenced genome of *Synechocystis* sp., recombinantly expressed in *E. coli* and purified (15). The enzyme consists of 279 amino acids with a predicted molecular weight of 30.2 kDa per monomer. The molecular weight of the native wild-type enzyme has been reported in the range of 57 – 65 kDa, confirming that AD is a dimer in solution (15). The NADP<sup>+</sup>-bound crystal structure was determined in 2006 (48), and it currently remains the only AD for which there is a crystal structure.

Bonner and colleagues (15) and Legrand and colleagues (48) have conducted kinetic studies on recombinant AD from *Synechocystis sp.* Both groups have shown that this enzyme displays an absolute specificity for NADP<sup>+</sup> and L-arogenate (15, 48), although other results are not in agreement. Legrand et al. reported that the enzyme exhibited a rapid equilibrium random kinetic mechanism, mirroring the findings for the *E. coli* PD-catalyzed reaction (83). In contrast, the finding of Bonner et al. from dead-end inhibition studies in the presence of a nucleotide analog 2',5'-ADP indicated that the enzyme followed a preferred ordered kinetic mechanism with L-arogenate combining with the free enzyme prior to the co-factor (15). Furthermore, Bonner et al showed that L-Tyr was a competitive inhibitor with respect to L-arogenate, while Legrand et al reported that the same enzyme was insensitive to inhibition by the end product of the reaction. The latter group proposed that this discrepancy could be due to amino acid replacements (H264Q, R265H, L266R, and Q268L) present in the C-terminal domain in the enzyme studied by Bonner and colleagues. This discrepancy is under investigation in the present study.

Legrand and colleagues also proposed that the active site residues H179, G219, G221, G226, M228, and Tyr232 might help define the enzyme's strict preference for L-arogenate, although there have been no mutagenesis studies reported to test this hypothesis. These residues were selected after positioning L-arogenate in the active site of the NADP<sup>+</sup>-bound structure (Figure 8C). This was done by placing the C4 hydroxyl group of L-arogenate on top of a water molecule located 2.5 Å from H112, presumed to be the reaction's catalytic hydrogen bond acceptor (H197 in *E. coli* CM-PD and H147 in *A. aeolicus* Δ19PD). L-arogenate's C-4 hydrogen atom was close enough (2.6 Å) to the C-4 atom of the oxidized nicotinamide so that hydride transfer could occur. Of the

residues identified by Legrand and colleagues, only H179 is highly conserved. G221 was speculated to be particularly important for determining L-arogenate selectively; the small side chain was argued to be important for accommodating the larger, charged nitrogen atom of L-arogenate's side chain (48). Legrand also speculated that a group of aliphatic side-chains (V182, V218, L225, M228, and M229) comprises a hydrophobic pocket which was predicted by Hermes and colleagues (38) to promote a concerted reaction mechanism. The reaction's dienone intermediate becomes so short-lived that C-C cleavage occurs in the same transition state as hydride transfer (38). In Legrand's model (Figure 8), the side chains of these groups are pointing towards the ring carboxyl group of arogenate, supporting this theory. R217 in AD from *Synechocystis sp* (R294 and R250 in *E. coli* CM-PD and *A. aeolicus* Δ19PD, respectively) is located on an α-helix in a very polar, solvent-exposed environment and too far away from the active site to interact with L-arogenate. This prompted Legrand and colleagues to suggest that this cationic residue was not important for substrate binding in TyrA proteins (48), despite kinetic results on variants of PDs from *E. coli* and *A. aeolicus* to indicate the contrary (37, 42).



Interestingly, this arginine is absent in plant ADs, such as that from *Arabidopsis thaliana*. Another residue in the *Synechocystis sp* AD, R213 (equivalent K246<sup>7</sup> in *A. aeolicus* Δ19PD), is located on the same helix as R217, but even more distant from the active site and closer to the C-terminal dimerization domain.

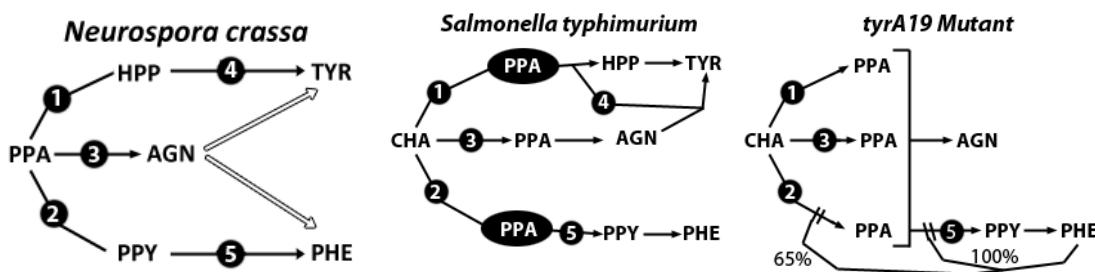
No mutagenesis studies have been performed on any AD TyrA proteins to date. Accordingly, the above mentioned residues in *Synechocystis sp* AD and a number of others have been examined in this thesis, and their importance in the enzyme's catalytic mechanism will be addressed.

## 1.9 Synthesis and Purification of Arogenate

The substrate for AD, L-arogenate, is not commercially available, hence a number of different protocols have been developed in order to obtain the compound. Jensen and colleagues (50) first reported the isolation of arogenate from an auxotroph of *N. crassa* that carries multiple metabolic blocks in the aromatic amino acid biosynthetic pathway (Figure 8, left hand panel), thus favouring the accumulation of L-arogenate. L-arogenate was enriched from the accumulation media by barium precipitation and Dowex-chloride-21K-anion exchange chromatography. Using the same strain, Zamir and colleagues (15) modified the procedure to include two more chromatography steps (AG1-X8-Cl anion exchange and Sephadex G-10 size exclusion) and reported a purity of 95% as assessed by <sup>1</sup>H-NMR. Bonner and colleagues (52) continued to use the protocol of Zamir and colleagues, but replaced the *N. crassa* strain with a metabolically blocked mutant of *Salmonella typhimurium* (Figure 8, right panel) in an attempt to increase the yield of L-arogenate. They reported as much as 30% of the total cyclohexadienyl accumulation in

the culture supernatant can be L-rogenate from the *tyrA19* mutant, although the isolation procedure is very lengthy.

Zamir and Jensen (53) were also the first to use HPLC in order to quantify L-rogenate and to follow its enzymatic conversion by AD. Prior to HPLC analysis, reaction mixtures were derivatized with o-phthalaldehyde (OPA) – a reagent that reacts covalently with primary amino groups of amino acids (such as L-rogenate, L-Tyr) to yield compounds that can be detected fluorometrically at 455 nm when  $\lambda_{ex} = 360$  nm. This method remains a useful way to verify enzyme activity, particularly for enzymes whose substrates or products do not exhibit a unique spectrophotometric or fluorometric



**Figure 9: The pathways of tyrosine and phenylalanine biosynthesis in *N. crassa*, *S. typhimurium*, and *S. typhimurium* *tyrA19* mutant.** In *N. crassa* (left), prephenate dehydrogenase (1) and prephenate dehydratase (2) transform prephenate (PPA) into HPP and phenylpyruvate (PPY), respectively. Prephenate can also be transaminated to arogenate (AGN) by prephenate aminotransferase (3). The mutant used for arogenate accumulation possesses deficiencies in enzymes (1) and (2), favouring arogenate production. The hollow arrows represent enzymatic reactions that exist in other organisms, but are not present in *N. crassa*. In wild type *S. typhimurium* (center), chorismate (CHA) becomes prephenate (PPA) by enzymes (1), (2), or (3). These enzymes correspond to monofunctional chorismate mutase enzymes, which are along the pathway for the synthesis of L-Tyr, L-Phe and L-Tyr (through L-rogenate), respectively. Prephenate will be directed to the synthesis of tyrosine through enzyme activity (4), prephenate dehydrogenase or to phenylalanine through enzyme activity (5), prephenate dehydratase. In the *tyrA19* mutant (right), cyclohexadienyl dehydrogenase activity (4) has been removed and (2) and (5) (chorismate mutase and prephenate dehydratase, respectively) are sensitive to phenylalanine feedback inhibition. As such, arogenate production is favoured.

signal. The activity of prephenate aminotransferase, which transaminates prephenate to L-arogenate, was also assayed in this manner (54). Due to the covalent modification by OPA, however, the resulting derivatized arogenate is not suitable for use in kinetic studies.

In 1983, Danishefsky and colleagues (55) reported the total synthesis of L-arogenate from pyroglutamate. The identity of the final product was confirmed by<sup>1</sup>H NMR analysis but the poor yield and multi-step synthesis rendered this protocol impractical for sustaining moderate quantities of L-arogenate for enzymatic and biophysical studies on AD.

Most recently, Rippert and colleagues (56) reported the enzymatic synthesis of L-arogenate enzymatically by the transamination of prephenate via prephenate aminotransferase (PAT). For this purpose, these authors obtained a crude preparation of PAT by harvesting and grinding the leaves of *Nicotiana tabacum*. Fortunately, the genes encoding PAT from *Petunina hybrid* and *Arabidopsis thaliana* have very recently been reported (54). As will be outlined in this thesis, purified recombinant PAT can now be used to synthesize L-arogenate from prephenate.

## 1.10 Research Objectives

The overall goals of this study were two-fold: 1) to commence the first mutagenesis study on AD, targeting those residues which may play a role in substrate binding, catalysis and end product inhibition; 2) to develop a procedure in order to obtain L-arogenate for use in enzymatic and biophysical assays.

In our first goal, both the crystal structure of *Synechocystis sp* AD and multiple amino acid sequence alignments revealed a number of residues in the active site that could potentially play a role in arogenate specificity, substrate binding and/or catalysis. As previously discussed, **H112** is expected to be the catalytic hydrogen bond acceptor in the reaction; as there are no mutagenesis studies reported for this TyrA protein, this residue was changed to alanine to note the effects on enzyme activity. **G221** was speculated by Legrand and colleagues (48) to be critical for arogenate selectivity. Thus, G221 was replaced with serine and glutamine, the equivalent residues in the *A. aeolicus* PD and *E. coli* CM-PD, respectively, to note any effects on substrate preference. Both **S178 and H179** are conserved in many TyrA proteins, although they are replaced by glutamine and alanine in other members of the family. Furthermore, the equivalent residues in *A. aeolicus* PD were shown to play a role in prephenate binding (42). Accordingly, both single and double variants, S178A, H179A, and S178Q/H179A were constructed to probe the effects on AD activity. **M228** was speculated to be important for AD catalysis, as it was hypothesized to help define a hydrophobic pocket within the active site which might be critical for product release (38, 48). Accordingly, this residue was replaced with the hydrophilic amino acid aspartic acid that might disrupt this hydrophobic pocket, and perturb enzyme activity. **G226** was another residue speculated by Legrand and colleagues to play a role in arogenate specificity. Furthermore, equivalent residues in *E. coli* CM-PD and *A. aeolicus* PD, Y303 and W259, respectively were found to play an important role in L-Tyr binding and inhibition. As such, G226 was replaced with tyrosine to note the effect on enzyme activity, particularly in the presence of the end product. Despite the fact that the crystal structure of AD in complex with NADP<sup>+</sup> revealed that **R213 and R217**

are too far away from the active site to interact directly with L-aro-ogenate, these two residues were replaced with alanine to probe their importance in enzyme function; the equivalent residues in other TyrA proteins were found to be important in binding prephenate (37, 42). The highly conserved **D171** and **H170** were targeted for mutagenesis and replaced with alanine and asparagines, respectively; in particular, D171's homolog in *A. aeolicus* Δ19PD was found to be critical for enzyme function and stability (42). Lastly, the importance of C-terminal residues **H264, R265** and **L266** in conferring end product inhibition were probed by characterizing the single variant H264Q as well as triple variants derived combinatorially with R265H and L266R. These studies targeted three of the four amino acid replacements which might have lead to a discrepancy in the results reported by Bonner et al(15) and Legrand et al (48) with regard to L-Tyr sensitivity.

In total, fifteen *Synechocystis sp* AD variants were generated by site-directed mutagenesis, expressed in *E. coli* and purified to homogeneity using Reactive Red120 Agarose affinity chromatography. Their specific activities were determined spectrophotometrically and the values compared to the wild-type enzyme. Circular dichroism (CD) was used to confirm that secondary structure of the variants was not disrupted, and to determine the thermal stabilities of selected enzymes. The effects of amino acid substitutions on tertiary structure of AD were assessed using fluorescence emission spectroscopy. Additionally, the binding of NADP<sup>+</sup> to the wild-type enzyme was also investigated by recording changes in fluorescence emission. Where applicable, molecular modelling was used to help explain the effects produced by the amino acid replacements.

Addressing our second goal, we chose to obtain L-arogenate using a protocol adapted from Rippert and colleagues (56). This method involves the enzymatic conversion of prephenate to L-arogenate via prephenate aminotransferase, which is then linked to a coupled enzyme reaction that can be monitored spectrophotometrically. Prephenate was readily available in our laboratory and we had recently obtained access to the clone encoding PAT from *Petunia hybrida*. Next, purification of L-arogenate using HPLC was developed on the analytical scale and adapted to the preparative scale. Identities of the compounds detected by HPLC analysis were verified using LC-MS.

## Chapter 2: Materials and Methods

### 2.1 Materials

Chorismate (free acid form) was isolated and purified from *Klebsiella pneumoniae* (60). Prephenate (barium salt) was prepared enzymatically from chorismate as previously described (61). Crude L-arogenate (purity ~4%) used in specific activity measurements was a kind gift from Carol Bonner. L-arogenate was prepared enzymatically from prephenate as described in section 2.9. NAD<sup>+</sup> (grade I) was obtained from Roche while NADP<sup>+</sup> (grade I) was obtained from Sigma Aldrich. Stock solutions of substrates were prepared in the appropriate buffers and stored at -20°C in small aliquots. Their concentrations were determined using published extinction coefficients (NAD<sup>+</sup>/NADP<sup>+</sup>) or enzymatic end-point analysis (arogenate, prephenate and chorismate) (62, 63). L-Tyr and L-Phe were purchased from ICN Biochemicals Inc. Ampicillin (sodium salt), kanamycin sulphate, chloramphenicol, IPTG, phenyl-methyl-sulfonyl fluoride (PMSF, prepared to 0.5M stock solutions in methanol and stored at -20°C) and dithiothreitol were obtained from BioShop. Oligonucleotides of standard purity were ordered from BioCorp (Montréal, QC). GeneJET<sup>TM</sup> Plasmid Miniprep Kit was purchased from Fermentas Life Sciences, while PureLink<sup>TM</sup> Quick Gel Extraction Kit was purchased from Invitrogen. Restriction enzymes *Dpn*I, *Nde*I, *Bam*HI, *Xho*I, *Nco*I, (all at 10 U/μL), and the deoxy-NTP (dNTP) mixture (5 mM of each dNTP, stored at -20°C in small aliquots) were purchased from MBI Fermentas. *Phusion® High-Fidelity* DNA Polymerase (2.0 U/μL) was purchased from New England Biolabs Inc. Benzonase Nuclease was obtained from Novagen. Complete<sup>TM</sup>, Mini, EDTA-free protease inhibitor cocktail tablets were purchased from Roche. Amersham Biosciences supplied thrombin

protease (purified from bovine plasma, 500 U resuspended in 500 µL of phosphate buffered saline) and NAP<sup>TM</sup>-5 size exclusion buffer exchange columns pre-packed with DNA grade Sephadex<sup>TM</sup> G-25, while Ni-NTA Superflow<sup>TM</sup> chromatography resin was purchased from Qiagen. Reactive Red Agarose 120 affinity chromatography resin was supplied by Sigma. Dialysis membrane (MW cut-off 12-14 kDa) from Spectrapor was washed according to manufacturer's instructions. All other chemical reagents and solvents were purchased commercially and were of the highest quality available.

## 2.2 Strains and Plasmids

The *E. coli* strain XL10-Gold ultracompetent (Stratagene) Tetr<sup>r</sup>Δ(*mcrA*)I83 Δ(*marCB-hsdSMR-mrr*)I73 *endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F'proAB *lacI*<sup>q</sup>ZΔM15 Tn10 (Tet<sup>r</sup>) Amy Cam<sup>r</sup>] was used for plasmid production while BL21(DE3) (Stratagene) [F<sup>-</sup> dcm<sup>+</sup> Hte ompT hsdS(rB<sup>-</sup> mB<sup>-</sup>) gal λ (DE3) endA Tet<sup>r</sup>] was used for protein expression. The cells were either purchased commercially ready to use or rendered competent using calcium chloride (20). Recombinant wild-type AD plasmid was previously constructed by Dr. P. Hanic-Joyce by cloning *Synechocystis* sp. PCC 6803 *tyrA<sub>a</sub>* encoding amino acid residues 1-279 of AD into the *Xba*I and *Nde*I restriction sites of *E. coli* expression vector pET24a (Novagen), yielding pET24a-*tyrA<sub>a</sub>*. The construct, pET28a-PAT, which contains the truncated *Petunia hybrida PhPPA-AT* gene cloned into the *Nde*I and *Bam*HI sites, encodes amino acid residues 59-479 of PAT and was an extremely kind gift from Dr. Natalia Doudareva (Purdue).

### **2.3. Bacterial Growth Media**

Luria-Bertani (LB) medium was prepared with 1% tryptone, 0.5% yeast extract and 1% NaCl in distilled H<sub>2</sub>O (pH was adjusted to 7.5) while LB agar was prepared with 1.5% agar in LB medium. Both media were sterilized by autoclaving at 121°C. Stock solutions of ampicillin (100 mg/mL), and kanamycin (10 mg/mL) were prepared in MilliQ-H<sub>2</sub>O, and chloramphenicol (30 mg/mL) in ethanol, and were filter sterilized by passage through a 0.45 µm syringe (VWR) and stored at -20°C.

### **2.4. Site-Directed Mutagenesis**

Oligonucleotides used for the mutagenesis reactions (see Table 1) were resuspended in MilliQ-H<sub>2</sub>O and their concentrations were determined by measuring using an Implen Nanophotometer located in the Center for Structural and Functional Genomics (Concordia University). The plasmid pET24a-*TyrA<sub>a</sub>* and was used as template plasmid for site-directed mutagenesis. The plasmid was isolated from a 5 mL culture of *E. coli* XL-10-Gold ultracompetent cells harbouring the plasmid using the GeneJET™ Plasmid Miniprep Kit (Fermentas Life Sciences), and its concentration was determined using the Nanophotometer.

Site-directed mutagenesis was conducted according to the instructions supplied in the QuickChange™ XL Site-Directed Mutagenesis Kit. The reaction mixtures were prepared with 15 ng of double stranded (ds) DNA template, 125 ng of each oligonucleotide primer (forward and reverse), 0.5 µL of *Phusion*® High-Fidelity DNA polymerase (2 U/µL), 10 µL of 5 x *Phusion*® HF buffer, 1 µL of a 5 mM dNTP solution.

Reaction mixtures for the R217A and S178A also contained 10% DMSO due to the high G/C content of their primers. DNA polymerase was added just prior to the first denaturation cycle. PCR amplification was carried out using a GeneAmp PCR system 9700 (Applied Biosystems). The temperature cycle parameters are listed in Table 2. Briefly, double-stranded plasmid DNA was denatured at 98°C then oligonucleotide primers were annealed to the plasmid at 60 °C. Synthesis and extension of the new DNA strand was catalyzed by the DNA polymerase at 72°C. Reaction mixtures were stored at 4°C in the apparatus until processed further.

**Table 1: Primer sequences for site-directed mutagenesis of *Synechocystis* AD**

AD Variants	Oligonucleotide Primer Sequence
SH178QA	5' CCAAGCAGTAGCCTGGATT <u>CAGG</u> CTTACCTGTAA <u>TGGT</u> GAGTGC 3'
G221Q	5' GATACCAGTCGGTGGGAGGG <u>CAA</u> AACCCGGAGTTGGGCACCATG 3'
H264Q	5' CAATGGCCTGA <u>ACTCCA</u> <u>AC</u> GTCTTTACAACAAAC 3'
HRL264QHR	5' GTAACCAACAATGGCCTGA <u>ACTCCA</u> <u>AC</u> <u>ATCG</u> TTACAACAAACCAACGGCG 3'
M228D	5' GCAACCCGGAGTTGGGCACC <u>GAC</u> ATGCCACCTATAACCAACG 3'
G226Y	5' GAGGGGGCAACCCGGAGTT <u>GTA</u> ACCATGATGCCACCTATAAC 3'
G221S	5' GATACCAGTCGGTGGGAGGC <u>A</u> GCAACCCGGAGTTGGGCACCATG 3'
T117P	5' GGTCACCCC <u>ATGGCCGGCC</u> CAGCGGCCAGGGCATCG 3'
R217A	5' GGTTTCGGGATACCAGT <u>GC</u> GGTGGAGGC <u>GG</u> CAACCCGG 3'
R213A	5' GGCCAGTT <u>CGGG</u> TTT <u>GC</u> GGATACCAGTCGGTGGG 3'
S178A	5' CCAAGCAGTAGCCTGGATT <u>GCC</u> CATTACCTGTAA <u>TGGT</u> G 3'
D171A	5' GCACTCCCGCAGACC <u>AT</u> GGCAGTAGCCTGG 3'
H112A	5' GGGTTCATTGGTGGT <u>GCC</u> CCATGGCCGGCACAG 3'
H170N	5' CACTCCCGCAGAC <u>A</u> ATGACCAAGCAG 3'
H179A	5' CAGTAGCCTGGATT <u>CCG</u> CTTACCTGTAA <u>TGGT</u> G 3'

Primers were designed considering length, GC content and location, and melting temperature using Primer3 Output program (<http://frodo.wi.mit.edu/primer3/input.htm>) (64)). Bold and underlined letters represent bases that differed from the wild-type sequence. Complementary primers are not shown.

**Table 2: Temperature cycling parameters for PCR using *Phusion® HF DNA Polymerase***

Number of Cycles	Temperature (°C)		Time (s)
1	Denaturation	98	120
18	Denaturation	98	60
	Annealing	60	60
	Extension	72	300
1	Extension	72	600
	Cooling	4	∞

Upon completion of the temperature cycling, methylated and hemi-methylated parental DNA were digested by incubation of the reaction mixtures with *DpnI* restriction endonuclease (10 U) at 37°C for 3 h. The presence of amplified DNA was confirmed by electrophoresis on a 0.7% agarose gel (see section 2.6) and transformed into *E. coli* strain XL10-Gold ultracompetent cells (see section 2.5). All cells were then plated on LB/agar containing kanamycin (50 µg/mL). After overnight incubation at 37°C, single colonies were selected and inoculated separately into 10 mL of liquid LB medium containing kanamycin (50 µg/mL) and grown overnight at 37°C with shaking at 225 rpm. The plasmid DNA was extracted and purified using the GeneJET™ Plasmid Miniprep Kit (Fermentas). The protocol was modified by eluting the DNA using MilliQ-H<sub>2</sub>O instead of the buffer provided in the kit. The concentration and purity of the plasmid DNA were determined using an Implen Nanophotometer.

Aliquots of purified plasmid DNA were sent for sequencing of the *tyrA<sub>a</sub>* gene (McGill University and Génome Québec Innovation Centre). The resulting sequences were aligned with that of WT *tyrA<sub>a</sub>*, using the BLAST tool on NCBI (<http://www.ncbi.nlm.nih.gov/blast>) to ensure that the desired mutation was achieved and that no other mutations had been introduced.

## 2.5 Transformation

*DpnI*-treated PCR products were transformed using a rapid transformation protocol described by Sambrook and Russell (65). Briefly, purified *DpnI*-treated PCR products (10 µL) were added to 50 µL of XL10-Gold ultracompetent cells, mixed and incubated on ice for 15 min. The cells were then incubated at 42°C for 45 s (heat-shock) and subsequently kept on ice for 5 min. One mL of LB broth (section 2.3) was added to the reaction, and the cells were incubated for 1 h at 37°C, with shaking at 225 rpm. The transformed cells were centrifuged at 5000 rpm for 10 min (Eppendorf Centrifuge 5415 C), and excess LB broth was removed. The transformed cells were then plated on LB/agar containing 50 µg/mL kanamycin and grown for 16 h at 37°C.

Transformation for AD protein expression was performed using the same method as described above. Briefly, ~ 200 ng AD plasmid DNA was mixed with 50 µL of BL21(DE3) competent cells, placed on ice for 15 min followed by heat-shock for 45 s at 42°C and then incubated on ice for 5 min. The entire mixture was plated on LB/agar containing 50 µg/mL kanamycin and incubated for 16 h at 37°C.

## 2.6 Agarose Gel Electrophoresis

Agarose gel electrophoresis (0.7% agarose) was used to verify the amplification of plasmid DNA and PCR products. Agarose (0.35 g) was dissolved in 50 mL TBE buffer (45 mM Tris-borate/1 mM EDTA, pH ~ 8.3) and heated for 2 min in a microwave. After the agarose solution had cooled to 55°C, ethidium bromide (0.5 µg/mL) was added. DNA samples and molecular weight markers (Gene Ruler 1kbp DNA ladder, Fermentas)

were mixed with 6 x loading dye (Fermentas). Electrophoresis was performed at 100V until the loading dye migrated to the middle of the gel. DNA was visualized using a FluoroChem FC2 Imaging Illuminator.

## 2.7 Expression and Purification of AD Proteins

AD and variants were purified by the method of Legrand and colleagues (48) with minor modifications. Single colonies were selected from freshly plated transformants of *E. coli* BL21(DE3) cells containing plamids of wild-type or mutant *TyrA<sub>a</sub>*, and a 10 mL culture (LB medium containing 50 µg/mL of kanamyacin) was grown overnight at 37°C with shaking (225 rpm). Five mL of this culture were then diluted into each of 2 x 1.5 L of the same medium. These cultures were grown at 37°C with shaking until they reached an OD<sub>600</sub> of 0.6-0.8 (after about 3 h). Protein expression was then induced by adding 0.4 mM IPTG and the cultures grown for an additional 3 h at 37°C. The cultures were then centrifuged at 5,000 rpm (Beckman centrifuge J2-HS, JA-10 rotor) at 4°C for 15 min, and the supernatant discarded. The pellet was stored at -80°C overnight, and then placed at -20°C until use. All steps of the purification were performed at 4°C.

The pellet was thawed and then resuspended into buffer A (50 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 7.5) using 5 mL/g of wet cell pellet. Buffer A was supplemented with Complete™ protease inhibitor cocktail tablet (one tablet/50 mL of resuspension buffer). The mixture was homogenized using a Dounce Homogenizer (at least 5 up-and-down strokes), and the cell walls were disrupted by two passages through a French pressure cell with a setting of 1000 psi. The cell lysate was then centrifuged to clarity using a Beckman centrifuge for 1 h at 20,000 rpm (JA-20 rotor, Beckman) at 4°C.

A 50 µL aliquot of the supernatant was removed and stored for SDS-PAGE analysis. The cell-free extract was dialysed in 4 L of buffer A overnight at 4°C.

The dialysed extract was then applied to Reactive Red Agarose 120 resin (25 mL bed volume) in a 2.6 x 5 cm glass column conditioned previously with ice-cold buffer A at 1 mL/min. After the flow-through was collected, the resin was washed at a flow rate of 3 mL/min with approximately 150 mL of buffer A. The protein was eluted with a 100 mL linear gradient of buffer A containing 0-2 M NaCl. Fractions (about 5 mL) were examined at 280 nm to estimate protein concentration and the flow-through, wash, and elution contents were examined by SDS-PAGE. Fractions containing AD (electrophoretic band at 30 kDa) were then pooled and dialysed in 4 L of buffer A overnight at 4°C. Aliquots of protein (1 mL) were stored in buffer A containing 50% glycerol at -20°C.

## 2.8 Expression and Purification of Prephenate Aminotransferase

PAT protein were expressed, and cell-free extracts harbouring recombinant proteins were prepared as described for AD and its variants (section 2.7) with the exception that the cells were resuspended in buffer B (50 mM Tris-HCl, 100 mM NaCl, 10% glycerol) prior to lysis. The cell-free extract was applied to a 10 ml Ni-NTA Superflow column (binding capacity: 5-10 mg of His-tagged protein per mL resin), previously equilibrated with ice-cold buffer B containing 5 mM imidazole using a flow-rate of 1 mL/min. After the flow-through was collected, the column was washed at a flow-rate of 3 mL/min with 300 mL of buffer A supplemented with 30 mM imidazole. Bound His-tagged protein was eluted with buffer A containing 300 mM imidazole and 1.5 mL fractions supplemented with 1 mM EDTA were collected during elution. Protein

elution was monitored quantitatively by adding 10 µL of every third fraction collected to 990 µL Bio-Rad Bradford dye and recording the absorbance at 595 nm. Those fractions with the highest protein content and were pooled and dialyzed (Spectrapore, 12kDa cut-off) overnight at 4°C against buffer B containing 2.5 mM CaCl<sub>2</sub>. Protein was concentrated, if necessary, to 2-10 mg/mL (Amicon Ultra-15, MW cut-off 10 kDa) and stored at -20°C in buffer B supplemented with 50% glycerol. Each step of the purification procedure was monitored by SDS-PAGE (12% acrylamide). Protein concentration was determined using the Biorad protein assay (see section 2.12).

## 2.9 Arogenate Synthesis

L-argenate was synthesized enzymatically using a coupled assay adapted from Rippert et al. (56) with purified PAT (section 2.8) and malate dehydrogenase (MDH, EC 1.1.1.37, Sigma). In the presence of pyridoxal-5-phosphate, PAT produces arogenate from prephenate and an amino donor, which can be either L-aspartate or L-glutamate. The reaction catalyzed by PAT was enzymatically coupled to MDH since neither the appearance of L-argenate nor the disappearance of prephenate or L-aspartate can be easily monitored spectrophotometrically or fluorometrically. Reactions were carried out in 20 mM sodium phosphate buffer (pH 8.0) at 37°C in the presence of 5.4 mM barium prephenate, 10 mM L-aspartate, 16.2 mM NADH, 50 µM PLP, 10 µg of PAT, and 44 units of MDH in a final volume of 1 mL. The reaction was monitored by removing 5 µL aliquots of the reaction mixture at numerous time intervals, diluting it into 995 µL of 50 mM NaOH, and recording the absorbance at 340 nm. Once the absorbance had decreased to a value corresponding to the conversion of prephenate to L-argenate ( $\Delta A_{340} = 0.22$ ),

the reaction was stopped by flash-freezing in liquid nitrogen. The reaction vials were kept at -80°C until further use.

## **2.10 High Pressure Liquid Chromatography and Liquid Chromatography-Mass Spectrometry**

High pressure liquid chromatography (HPLC) was performed both on the analytical and preparative scales at ambient temperature. For analytical analysis, samples prepared in section 2.9 were centrifuged for 10 min at 13,000 x g in order to remove any precipitated protein debris. An Agilent HPLC system was used equipped with a C<sub>18</sub> column (5 um, 150 mm x 4.6 mm). A 15 min linear gradient from 1 to 30 % acetonitrile in 20 mM sodium phosphate buffer (pH 7.6). Arogenate, NAD<sup>+</sup>, and NADH peaks were monitored at 210 nm and quantified based on a standard calibration curve generated with individual standards, when available. Peak identities were further confirmed by liquid chromatography-mass spectrometry (LC-MS) using a Quattro-LC mass spectrometer (Waters micromass) equipped with an Agilent HPLC system. The analytical HPLC method was adapted to a preparative scale using a XBridge C<sub>18</sub> column (5 um, 100 x 10 mm, Waters) and a Agilent HPLC system with fraction collector, kindly provided by Dr. Pat Forgione (further experimental details are described in Results).

## **2.11 Determination of Protein Concentration**

Concentrations of protein purified to near homogeneity (>95% homogenous) were calculated by using the extinction coefficient of 25,672 M<sup>-1</sup>cm<sup>-1</sup> for all AD proteins (with the exception of G226Y) and measuring the OD<sub>280</sub>. Extinction coefficients were

determined experimentally using the Edelhoch method (66) in duplicate. Identical dilutions of protein were made into 30 mM MOPS, 1 mM DTT buffer, pH 7.0, and into the same buffer containing 6 M GdnHCl. UV spectra were measured between 250-350 nm with buffer alone and with protein. The extinction coefficient of the protein in 6 M GdnHCl was calculated using determined extinction coefficients for tryptophan, tyrosine and cystine in 6 M GdnHCl:

$$(6MG) = (\#Trp) \varepsilon_{\lambda}(Trp, 6MG) + (\#Tyr) \varepsilon_{\lambda}(Tyr, 6MG) + (\#cystine) \varepsilon_{\lambda}(cystine, 6MG)$$

$$(6MG) = (3Trp)5,685(Trp, 6MG) + (7Tyr)1,285(Tyr, 6MG) + (0cystine) 125(cystine, 6MG)$$

From this, the protein concentration in 6 M GdnHCl is calculated, which is equal to the concentration of protein in the buffer. With this, the extinction coefficient of the protein is calculated using:

$$\varepsilon_{280}(\text{buffer}) = A_{280}(\text{buffer})/C(\text{buffer})$$

The  $\varepsilon_{280}$  that was calculated for wild-type AD and the G226Y variant were found to be  $25,672 \text{ M}^{-1}\text{cm}^{-1}$  and  $28,084 \text{ M}^{-1}\text{cm}^{-1}$ , respectively. When compared to the theoretical values for the wild-type and G226Y variant ( $26,930 \text{ M}^{-1}\text{cm}^{-1}$  and  $28,670 \text{ M}^{-1}\text{cm}^{-1}$ ) calculated using the ExPasy ProtParam server, the differences between the experimental and theoretical coefficients were found to be 4.6% and 2.1%, respectively.

Protein concentration of crude extracts and prephenate aminotransferase was calculated using the Bio-Rad protein assay kit (Bio-Rad Laboratory) with bovine serum albumin (BSA, Sigma) as a standard. (85) BSA was dissolved in 10 mM Tris-HCl, pH 7.4, filtered using a  $0.2 \mu\text{m}$  syringe and its concentration was determined by  $\text{OD}_{280}$  readings using an extinction coefficient of  $0.667 \text{ mL/mg/cm}$ . (86)

## **2.12 Polyacrylamide Gel Electrophoresis**

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was used to assess the purity and estimate the molecular weights of proteins under denaturing conditions. SDS-PAGE was performed using a 4% acrylamide stacking gel (pH 6.8) and 12% acrylamide resolving gel (pH 8.3) as reported by Sambrook *et al.* (65). The 30% acrylamide stock solution was prepared by mixing 29% (w/v) acrylamide and 1% (w/v) N, N'-methylene-bis-acrylamide in distilled H<sub>2</sub>O. Resolving gel buffer (1.5 M Tris-HCl, pH 8.3) and stacking gel buffer (0.5 M Tris-HCl, pH 6.8) were prepared in distilled water and the pH was adjusted using HCl. The acrylamide and Tris buffer solutions were stored at 4°C. Ten percent SDS (w/v) in distilled water was stored at room temperature while 10% (w/v) ammonium persulfate prepared in distilled water was stored at -20°C. Electrophoresis buffer was composed of 25 mM Tris-base, 250 mM glycine and 0.1% SDS, pH 8.3. Protein samples and protein molecular weight marker (Fermentas) were diluted 1:1 (v/v) into 2 x SDS gel-loading buffer (1.5 M Tris-HCl, 4% SDS, 20% glycerol (v/v), 0.002% Bromophenol Blue, pH 6.8) and denatured in boiling water for 5 min before loading onto the gel. Electrophoresis was conducted at 120 V and terminated when the Bromophenol Blue tracking dye migrated off the resolving gel. Proteins were visualized by first washing the gel in warm distilled water three times for 2 min, followed by staining in a solution of 0.1% (w/v) Brilliant Blue-G250 (BioShop) and 35 mM HCl in distilled water, until bands appeared (66). Contrast could be enhanced by placing the gel in distilled water to remove excess staining solution.

## 2.13 ESI-ToF Mass Spectrometry

The molecular weights of wild-type and variant AD proteins were confirmed using ESI-ToF-MS. Samples were prepared using a protocol developed by J. Manioudakis for *E. coli* CM-PD (84). Briefly, to 100 µg of protein was added 300 µL methanol, 100 µL chloroform and 200 µL MilliQ water. The mixture was vortexed for 5 s after each solvent addition. Next, the sample was centrifuged (Eppendorf Centrifuge 5415 C) at 16 000 x g for 3 min. The two visible liquid phases were now separated by a thin layer of precipitate containing the protein. After carefully removing all of the liquid, 300 µL methanol was added to the precipitated protein, and the sample was vortexed for 5 s and then centrifuged (16 000 x g) for 2 min. This wash step repeated once more. After the second wash step the methanol was decanted and the protein pellet was dried at room temperature in the fume hood. The dried pellet was resuspended in 300 µL of 30% methanol/0.2% formic acid (v/v) and then the sample was centrifuged for 5 min to remove all insoluble material. The samples were applied to a Waters Micromass Q-ToF 2 triple-quadrupole mass spectrometer by direct injection at a flow rate of 1 µL/min. Samples were analyzed in the positive-ion mode within an m/z range of 700-2000 using Mass Lynx 4.0 software (Waters Micromass). Instrument parameters were as follows: source block temperature: 80 °C; capillary voltage: 3.5 kV; cone voltage: 35 V; ToF: 9.1 kV; MC: 1.8 kV; resolution: 8000. Calibration of the instrument was performed with myoglobin (Sigma).

## 2.14 Determination of Enzyme Activities and Kinetic Parameters

The oxidative decarboxylation of L-aro-*g*enate in the presence of NADP<sup>+</sup> was monitored at 340 nm as described previously (48). The reactions (total volume 500 µL) were monitored continuously by using a Varian Cary 50 spectrophotometer equipped with a thermostated cuvette holder.

Standard activity assays for AD proteins were measured at 25°C in a reaction buffer containing 50 mM Tris-HCl, pH 7.5. Buffer was incubated in a 1 mL quartz cuvette of 1 cm path length at 25°C for 2 min, while NADP<sup>+</sup> and 500 µM crude L-aro-*g*enate (purity ~4% as determined by enzyme end-point assay) was added. The reaction was initiated by the addition of enzyme. Components were mixed by inversion of the cuvette. All substrates and enzyme were kept on ice prior to their addition.

The OD<sub>340</sub> was recorded continuously for 1 min and reaction rates were calculated from the linear portion of the progress curves using the software supplied with the spectrophotometer. Exact concentrations of substrates are listed in figure legends in the Results section. Values of specific activities were calculated using the following equation:

$$\text{Specific Activity } (\mu\text{mol/min}/\mu\text{g}) = \text{Activity} / (\text{total amount of AD in the cuvette, } \mu\text{g})$$

$$\text{Activity } (\mu\text{mol/min}) = \Delta C \text{ (mol/min)} \times \text{Total reaction volume (L)} / (1 \times 10^{-6})$$

$$\Delta C(\text{mol/min}) = \Delta A / (\epsilon_{NADH} \times L)$$

## 2.15 Far-UV Circular Dichroism Spectroscopy

Far-UV CD spectra of wild-type AD and variants were obtained using a Jasco-815 spectropolarimeter equipped with a Pelletier heating/cooling temperature control

system. Spectra were recorded at 20°C in a 0.2 cm path-length rectangular cell (600 µL) from 260 to 200 nm with the following parameters: 20 nm/min scan rate, 0.2 nm resolution, 0.25 sec response time, 1 nm bandwidth and a sensitivity of 100 mdeg. For each spectrum, five accumulations were averaged and the absorbance contribution of the buffer was subtracted. Protein samples were prepared fresh from a -20°C stock solution of enzyme stored in 50 mM Tris-HCl and 50% glycerol (v/v). Proteins were first exchanged into a buffer containing 50 mM KH<sub>2</sub>PO<sub>4</sub> / K<sub>2</sub>HPO<sub>4</sub>, pH 7.5 (buffer P) by overnight dialysis. Protein concentrations were then adjusted to ~ 5 µM monomer in the same buffer. Buffers were filtered (Millipore, 0.22µm) and their pH corrected.

For variable temperature experiments, changes in ellipticity at 222 nm (1 nm bandwidth) were recorded from 25°C to 95°C by using instrument's software controlled temperature ramping program. The following parameters were used: ΔT of 20°C/h, a 0.2°C step resolution, and a 0.25 s response time. The protein samples were prepared according as described above. Spectral scans from 200-260 nm were recorded for each sample at the beginning and the end of each the variable temperature experiments.

## **2.16 Near-UV Circular Dichroism Spectroscopy**

Near-UV CD spectra of wild-type AD and variants were obtained using a Jasco-715 spectropolarimeter. Spectra were recorded at 20°C in a 1 cm path-length rectangular cell (3 mL) from 320 to 250 nm with the following parameters: 20 nm/min scan rate, 0.2 nm resolution, 0.25 sec response time, 1 nm bandwidth and a sensitivity of 100 mdeg. For each spectrum, ten accumulations were averaged and the absorbance contribution of the buffer was subtracted. Protein samples were prepared in the same manner as

described for far-UV CD (section 2.15). To determine any changes in aromatic CD signal in the presence of ligands, stock solutions of NADP<sup>+</sup> (29.6 mM) and crude L-arogenate (44 mM) were prepared in 10 mM Tris buffer, pH 7.5, and titrated directly into the reaction cuvette.

## 2.17 Fluorescence Emission Spectroscopy

Fluorescence spectra for wild-type AD and each protein variant were recorded at 20°C using a Varian Cary Eclipse spectrofluorimeter equipped with Scan Software version 1.1. Protein samples were prepared as described in 2.13 using a monomer concentration of 5 µM. Excitation wavelengths of 280 for tryptophan and tyrosine emission and 295 nm for tryptophan emission were used. Fluorescence emission was scanned from 300 to 400 nm in 2 nm increments at a fast scan speed (1200 nm/min) with the PMT voltage set at 700 V. Both excitation and emission slit widths were set at 5 nm. The average of 20 scans was used to obtain the final spectrum which was then processed by Savitzky-Golay smoothing with a filter size of 5. A Varian 400 µL fluorescence micro cell (1 cm x 1cm) was used throughout all experiments. Data were exported in ASCII (.csv) format and spectra were constructed in Windows Word Excel 2010. All spectra were corrected for buffer contributions, changes in volume, and corrected for the inner filter effect using the following equation:

$$F_{\text{corr}} = F_{\text{obs}} \text{ antilog } [(A_{\text{ex}} + A_{\text{em}}) / 2]$$

Where  $F_{\text{obs}}$  and  $F_{\text{corr}}$  represent the observed fluorescence intensities and those corrected for the inner filter effect respectively. Absorbance readings (A) were determined at both the excitation (ex) and emission (em) wavelengths. Experiments involving titration of the

wild-type AD with ligands were set up as described in 2.16. The apparent dissociation constant was determined by fitting the change in fluorescence intensity at a selected wavelength to the Michaelis-Menthen equation using Grafit Software version 5.0 (Erathicus Software).

## 2.18 Molecular Modeling and Docking

Molecular modelling was used to dock the substrate prephenate into a homology model for *Synechocystis sp* AD. First, the homology model of AD was generated using *A. aeolicus* Δ19PD as the template. The PDB file “3GGG” (*A. aeolicus* Δ19PD bound with NAD<sup>+</sup> and L-Tyr) was prepared by removing all atoms and heteroatoms attributed to chains A and C and D using WordPad. Therefore only chain B was used for computing the model but additionally all heteroatoms forming tyrosine and all water molecules in chain B were removed. Next, a PIR file containing sequence alignment information of AD and Δ19PD was prepared using Notepad. MODELLER 9v8 software was used to generate 10 homology models using the default script, model-default.py. After evaluating the DOPE scores of the 10 models and aligning each onto the backbone of AD, it was determined that all of the models were reproducible. L-arojenate was then docked into the homology model using Autodock Vina (75). First, a PDBQT file was prepared by uploading the homology model into AutoDock Tools (Molecular Graphics Laboratory), removing all non-polar hydrogen atoms and saving the file as PDBQT. Next the search space was defined in AutoDockTools. The search space restricts where the movable atoms, including those in the flexible side chains, should lie. The spacing of the grid points was set to 1.000 Å and 35 grid points in all three directions were used. Therefore

the search space was cubic with a volume of in which each site measured 35 Å. Using important residues as a guideline, the center for x was chosen to be at X, for y at X and for z at X. The PDB file of L-argenate was then converted into a PDBQT file by loading the PDB file into AutoDockTools, defining all bonds as rotatable bonds and saving the file as PDBQT file. Docking was performed using AutoDock Vina, the defined centers for x, y, and z and an exhaustiveness of 20. (Molecular Graphics Laboratory) and the resulting structures were visualized using PyMol (Schroedinger) (68).

## Chapter 3: Results

### 3.1 Site-Directed Mutagenesis

The *tyrA* gene from *Synechocystis sp.* had been previously cloned into the pET24b expression vector by Dr. P. Hanic-Joyce between *NdeI* and *XhoI* restriction sites, maintaining its stop codon at the end of the gene so that no His-tag is incorporated into the sequence. The recombinant plasmid is approximately 6.1 kbp including a 5310 bp pET24b vector and an 840 bp insert (data not shown). The variant proteins were obtained using site-directed mutagenesis according to Stratagene's QuickChange<sup>TM</sup> protocol. Template plasmid was digested by incubation with *DpnI* so that only the mutagenized PCR products remained for transformation. PCR products were transformed into XL-10 Gold *E. coli* cells that are able to repair the nicked mutagenic strand and plasmid preparations were performed on selected colonies. DNA sequencing of the inserts confirmed that the mutant plasmids carried only the desired base substitutions.

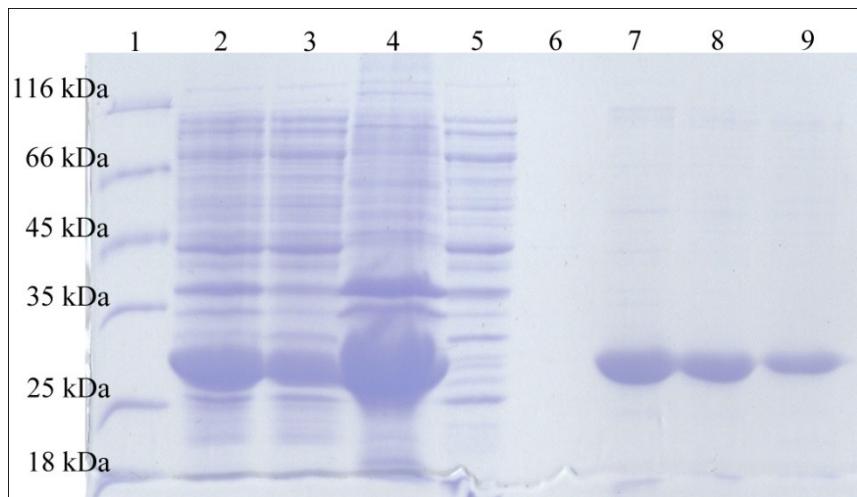
### 3.2 Expression and Purification of WT *Synechocystis* AD and Variant Proteins

The following fifteen AD proteins were expressed and purified: Wild-type enzyme and the variants G221Q/S, M228D, G226Y, T117P, R213A, R217A, D171A, H112A, H170N, S178A, H179A, S178Q/H179A (double substitution), H264Q, H264Q/R265H/L266R (triple substitution).

The purification procedure adapted was based on the protocol originally developed for wild-type AD as outlined by Legrand and colleagues (48). Wild-type or mutant AD plasmids were transformed into BL21(DE3) cells which were then grown in LB medium supplemented with kanamycin, and protein expression was induced by the

addition of IPTG. Cells were disrupted by high pressure in the presence of protease inhibitors and the recombinant proteins were purified using Reactive Red 120 Agarose affinity chromatography at 4°C. The Reactive Red 120 resin is specifically designed for the purification of NADP<sup>+</sup>-dependant dehydrogenases by interacting with the cofactor binding pocket (70). The pooled purified protein fractions were dialyzed overnight at 4°C, and AD was stored at -20°C in buffer containing 50% glycerol. Purified protein lost activity if frozen at -80°C in 15% glycerol and thawed.

Protein purification was monitored by measuring the A<sub>280</sub> of eluted fractions and denaturing gel electrophoresis. A detailed analysis of each step of purification for the wild-type enzyme is shown in Figure 10. Over-expression of the recombinant AD enzyme is clearly seen in the cell lysate and the cell-free extract (lanes 2 and 3). However, some protein remained in the particulate fraction as noted in the analyzed cell debris which has been resuspended in buffer containing 2% SDS (lane 4). After applying the cell-free extract to the Reactive Red 120 Agarose column, no significant band corresponding to the monomer molecular weight of AD (~30 kDa) was observed in the flow-through (lane 5), indicating that all of the AD protein had bound to the column. After washing with ~15 column volumes of buffer (lane 6), a salt gradient was applied and the protein was eluted usually in fractions 3-8. The purification appears to yield homogenous protein as judged by Coomassie blue staining.



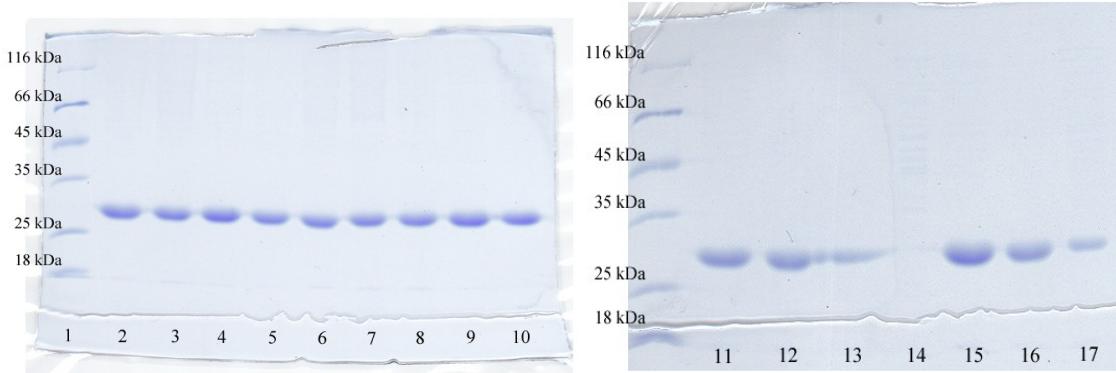
**Figure 10: SDS-PAGE analysis of wild-type AD purification.** Lane 1: Protein molecular weight marker, Lane 2: Cell lysate (1/2 dilution), Lane 3: Cell-free extract (1/2 dilution), Lane 4: Cell debris in buffer containing 2% SDS, Lane 5: Red 120 Agarose flow-through, Lane 6: Red 120 Agarose wash, Lane 7-9: Pooled protein after Red 120 Agarose. Analysis performed on a 12% acrylamide gel.

**Table 3: Purification table for wild-type AD**

	Total protein <sup>a</sup> (mg)	Total activity (Units)	Specific activity (Units/mg)	Activity yield (%)	Purification fold
<b>Cell-free extract</b>	2166	1673	0.77	100	1.0
<b>Flow-through</b>	1570	97	0.06	5.8	0.1
<b>Pooled</b>	153	1105	7.2	66.1	9.4

Activity assays performed at 25°C at 2 mM NADP<sup>+</sup> and 200 µM crude arogenate. <sup>a</sup>From 3L of cell culture, determined by Bio-Rad protein assay.

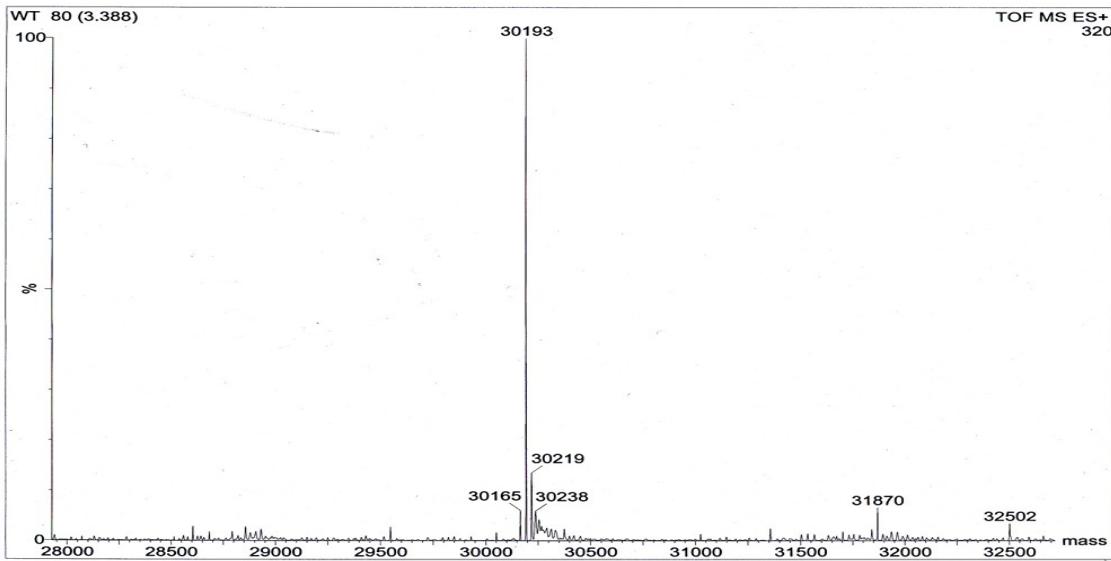
All AD variant proteins were purified using the same methodology as the wild-type enzyme. Almost all AD proteins were purified to homogeneity (Figure 11) and behaved similarly during the purification process. The yields of purified protein typically varied between 40 – 70 mg/L of cell culture, with the exception was D171A, which yielded 3mg of protein per litre of culture.



**Figure 11: SDS-PAGE analysis of purified AD variants.** Lane 1: Protein molecular weight marker, Lane 2: Wild-type AD, Lane 3: S178Q/H179A, Lane 4: G221Q, Lane 5: H264Q, Lane 6: H264Q/R265H/L266R, Lane 7: M228D, Lane 8: G226Y, Lane 9: G221S, Lane 10: T117P, Lane 11: R217A, Lane 12: R213A, Lane 13: S178A, Lane 14: D171A, Lane 15: H112A, Lane 16: H179A, Lane 17: H170N. Each of the protein samples were purified with Agarose Red 120 affinity chromatography prior to being analyzed on 12% acrylamide gels. Approximately 6-10 µg of protein was applied in each lane.

### 3.3 Electrospray Ionization Mass Spectrometry

The masses of the purified AD proteins were determined by ESI-Q-ToF MS to confirm that the correct amino acid substitutions were present (Figure 12). Overall, the theoretical and observed mass values of each recombinant protein were in good agreement (Table 4). The exception was D171A, whose deconvoluted spectrum did not reveal major peaks between 26500 and 32300 (data not shown). This result supports the data obtained from SDS-PAGE analysis, which shows that the enzyme is likely not stable and prone to degradation.



**Figure 12: Deconvoluted ESI-MS spectrum of wild-type AD enzyme.** The peak shows the molecular weight of the monomeric form of the protein.

**Table 4: Summary of molecular masses of wild-type and variant AD proteins.**

AD Variants	Expected Mass <sup>*</sup> (Da)	Observed Mass (Da)	Difference (Da)	AD Variants	Expected Mass (Da)	Observed Mass (Da)	Difference (Da)
WT	30197.7	30193.0	4.7	R217A	30112.6	30110.4	1.8
S178Q/H179A	30172.7	30170.3	2.4	R213A	30112.6	30109.5	2.9
G221Q	30268.7	30266.7	2.0	S178A	30181.7	30180.3	1.4
H264Q	30188.7	30186.5	2.2	D171A	30153.7	ND	ND
H264Q/R265H /L266R	30212.7	30209.4	3.3	H112A	30131.7	30127.6	4.1
M228D	30181.7	30180.8	0.9	H179A	30131.7	30128.2	3.5
G226Y	30303.7	30301.7	2.0	H170N	30174.7	30171.9	2.8
T117P	30193.7	30192.1	1.6	G221S	30227.7	30226.0	1.7

\* Expected mass based on the amino acid sequence was calculated using EXPASY-PeptideMass program which gives the average molecular weight of the deprotonated side chain. A mass difference of +/- 5 Da was considered to be acceptable. Values differ from sample to sample based on the particular instrument conditions and ambient temperature fluctuations on the day of analysis. ND denotes not detected.

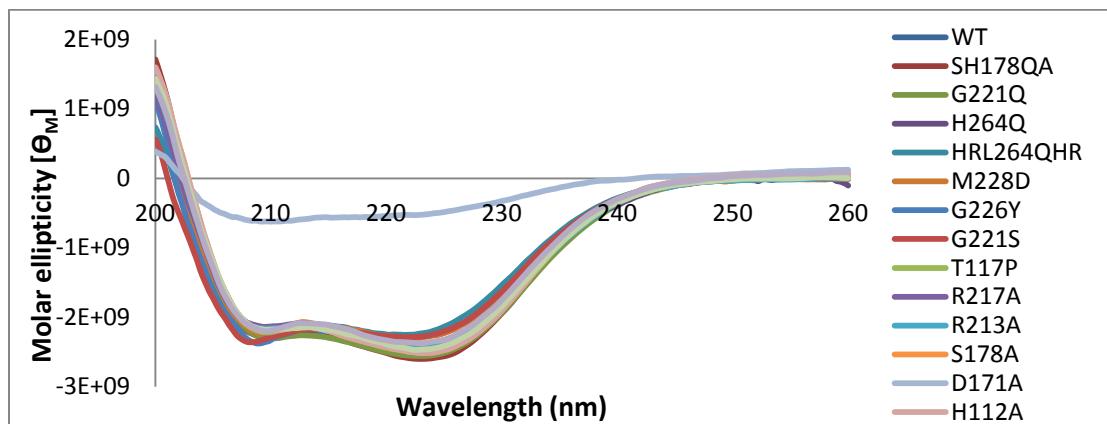
### **3.4 Circular Dichroism Spectroscopy**

#### **3.4.1 Far-UV Circular Dichroism**

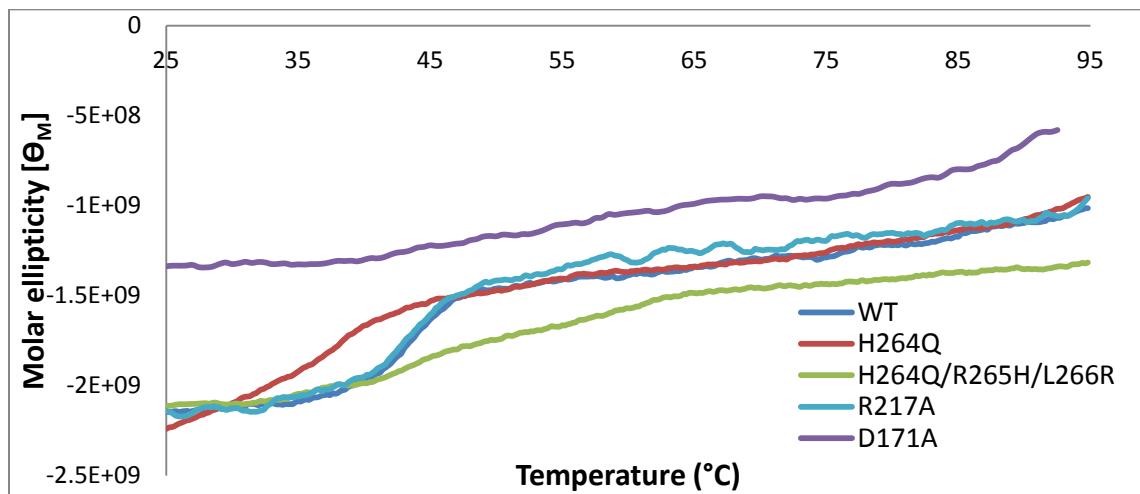
Far-UV circular dichroism (CD) spectroscopy was used to determine if any of the amino acid substitutions within the AD variants had altered the global secondary structure of the protein. CD spectroscopy uses circularly polarized light, which is composed of both left-handed and right-handed components. Light in the far-UV region can be absorbed by the amide and carbonyl groups in a protein's peptide backbone. Depending on the conformation adopted by the protein backbone, the amount of left- and right-handed circularly polarized light that is absorbed will vary (71), giving distinctive absorption bands for different global secondary structure arrangements. CD spectra were acquired between 200 – 260 nm at 20°C. Almost all AD proteins showed pronounced absorption minimal at 208 and 222 nm (Figure 13), which are characteristic of proteins with a high  $\alpha$ -helical content. The spectra of the variant proteins and wild-type enzyme are nearly super-imposable, which indicates that AD's global secondary structures have not been perturbed by the amino acid replacements. The exception is D171A, whose signal is considerably less intense than that of the wild-type protein and other variants, presumably because there is much less than 6  $\mu$ M in solution.

The thermal stability of the wild-type AD and selected variants (Figure 14) was also probed by monitoring the change in ellipticity at 222 nm while increasing the temperature under conditions listed in section 2.15. The  $T_m$  of wild-type AD was determined to be 42.8°C by taking the 1<sup>st</sup> derivative of the spectrum, with one other variant having a  $T_m$  within 0.5°C of this value. The variants that displayed markedly altered  $T_{mS}$  (determined by the 1<sup>st</sup> derivative of the melting curves) and unfolding

patterns were H264Q ( $36.1^{\circ}\text{C}$ ), H264Q/R265H/L266R ( $\sim 50^{\circ}\text{C}$ ), and D171A. None of the enzymes completely lose their  $\alpha$ -helical content, even when heated to  $95^{\circ}\text{C}$ .



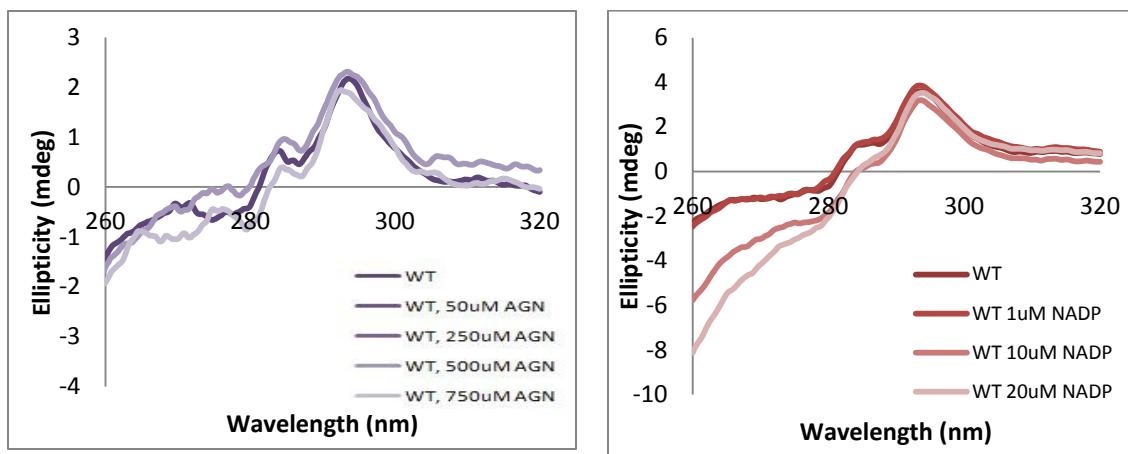
**Figure 13: Far-UV CD spectra of wild-type and variant AD enzymes.** Variant proteins ( $\sim 6 \mu\text{M}$  monomer) were exchanged into 50 mM potassium phosphate, pH 7.5. Spectra were recorded at  $20^{\circ}\text{C}$  in a 0.2 cm path-length rectangular cell with a scan rate of 50 nm/min. Each curve represents the average of 5 accumulations.



**Figure 14: Variable temperature far-UV spectra of wild-type and select variant AD enzymes.** Each protein ( $\sim 6 \mu\text{M}$ ) was prepared in the same manner as described in figure 13. Spectra were recorded at 222 nm in a 0.2 cm path-length rectangular cell from  $25^{\circ}\text{C}$  to  $95^{\circ}$ , using a ramping speed of  $30^{\circ}\text{C}/\text{h}$ .

### **3.4.2 Near-UV Circular Dichroism**

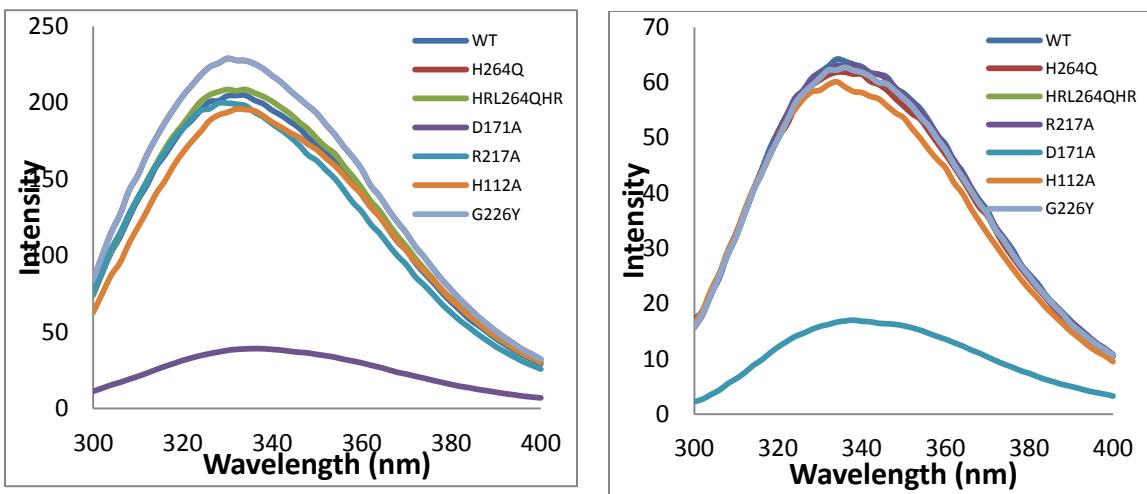
Near-UV CD spectroscopy was used to obtain information concerning the global tertiary structure of wild-type AD enzyme, and if it was perturbed by interaction with either of its substrates, NADP<sup>+</sup> or L-argenate (Figure 15). This absorption is a magnitude lower than that of peptide bonds, and generally occurs over the 260 – 320 nm wavelength range. The majority of this absorption is due to  $\pi - \pi^*$  transitions within the aromatic chromophores, which corresponds to tryptophan, tyrosine, and phenylalanine residues. Furthermore, a signal occurs only if the chromophores are held rigidly in an asymmetric environment (72). CD spectra were acquired between 250 – 320 nm at 20°C. A significant peak was observed at 295 nm and a smaller shoulder at 285 nm, in keeping with the contribution of Trp or Tyr residues to the aromatic CD signal. No significant change in ellipticity was observed with the wild-type AD when titrated with high concentrations of L-argenate. When titrated with NADP<sup>+</sup>, a change in ellipticity intensity was observed from 290 nm and below, increasing as the wavelength decreased. NADP<sup>+</sup> begins to absorb at ~290 nm, however, exhibiting an absorbance maximum at 259 nm. As such, the change in ellipticity is likely due to a change in the environment of NADP<sup>+</sup> upon binding to AD. Buffer containing 20  $\mu$ M of NADP<sup>+</sup> or 750  $\mu$ M L-argenate displayed no notable ellipticity signal (data not shown).



**Figure 15: Near-UV CD spectra of wild-type AD.** Spectra on the left show AD (49  $\mu$ M) titrated with crude arogenate, and spectra on the right show AD titrated with NADP<sup>+</sup>. Spectra were recorded at 20°C in a 0.2 cm path-length rectangular cell with a scan rate of 20 nm/min. Each curve represents the average of 7 accumulations.

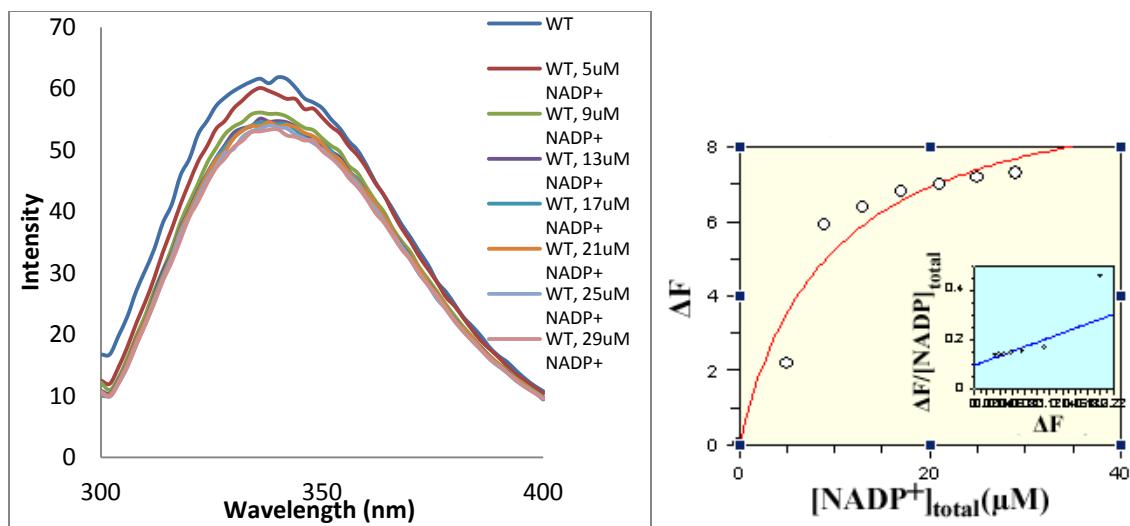
### 3.5 Fluorescence Emission Spectroscopy

Fluorescence emission spectroscopy was used to determine if the amino acid replacements had altered the environment of the aromatic chromophores, tryptophan and tyrosine, in AD. The fluorescence emission spectra of most variants were comparable to the wild-type enzyme, with the exceptions of G226Y and D171A (Figure 16). G226Y is observed to have an increase in fluorescence intensity at an excitation wavelength of 280 nm, but not 295 nm. This is consistent the introduction of an additional tyrosine residue into the protein's primary sequence. The intensity of D171A at both excitation wavelengths is much lower than that wild-type protein, indicating that the amount of soluble protein present is quite low.



**Figure 16: Fluorescence emission spectra for selected AD enzymes.** Left: Fluorescence emission spectra for variants at 20°C. Proteins (~6 μM monomer) were buffer exchanged into 50 mM potassium phosphate buffer, pH 7.5 and placed in a Varian 400 μL fluorescence micro cell (1.0 cm path-length). Excitation wavelength was fixed at 280 nm. The fluorescence emission was scanned from 300 to 400 nm using excitation and emission slit widths set at 5 nm. Spectra were corrected for buffer contribution. Right: Fluorescence emission spectra for selected proteins at an excitation wavelength of 295 nm. Spectra were recorded as described for the left panel.

The fluorescence spectra of wild-type AD enzyme were also recorded in the presence of increasing amounts of NADP<sup>+</sup> (Figure 17). Stepwise changes in fluorescence emission intensity were observed, and an apparent K<sub>d</sub> of 9.5 ± 3.9 μM was calculated. While the decreases in fluorescence intensity were relatively small, it was apparent by 17 μM NADP<sup>+</sup> no further signal were observed indicating that the AD is saturated with NADP<sup>+</sup>. AD contains three tryptophan residues per monomer, but none are present in the putative L-argenate binding site or the cofactor binding site. This decrease in intensity is likely attributed to minor quenching effects of cofactor binding. This observation was not seen upon titration with crude L-argenate (data not shown).



**Figure 17: Changes in fluorescence intensity of AD upon binding of NADP<sup>+</sup>.** Left: The changes in intrinsic tryptophan fluorescence of wild-type AD (5  $\mu$ M monomer) was observed by excitation at 295 nm and measuring the emission from 300 to 400 nm in 50 mM potassium phosphate buffer, pH 7.5. NADP<sup>+</sup> was varied from a concentration of 0 to 29  $\mu$ M. All spectra are corrected for buffer contribution, dilution effects upon titration and inner filter effects. Right: Fluorescence quenching is determined by plotting change in fluorescence at 335 nm as a function of total [NADP<sup>+</sup>]. An apparent  $K_d$  for NADP<sup>+</sup> was calculated to be  $9.5 \pm 3.9 \mu\text{M}$ .

### 3.6 Kinetic Studies of Wild-Type AD and Variants

In order to investigate the importance of selected active site residues of AD for enzyme function, kinetic assays were performed on the wild-type and variant proteins. For these studies, a crude preparation of L-arogenate (a kind gift from Carol Bonner) was used; its purity was estimated to be ~4%, determined by end point assay. Attempts to construct substrate saturation curves with L-arogenate or NADP<sup>+</sup> as the variable substrate were unsuccessful, precluding the determination of the Michaelis constant ( $K_m$ ) and the turnover number ( $k_{cat}$ ). Although the  $K_m$  for L-arogenate of AD from *Synechocystis sp* has been previously determined to be 93  $\mu\text{M}$  (48) and 331  $\mu\text{M}$  (15); a concentration of up to 3 mM L-arogenate in our assay did not yield zero order kinetic plot. We speculated that the crude preparation of L-arogenate may be contaminated with high salt content

and/or with product analogs that would act as inhibitors of AD, therefore hindering enzyme function. Accordingly, specific activities of each variant were obtained (Table 5) and compared relative to that of the wild-type enzyme. Specific activity is described as the micromoles of substrate converted per minute per microgram of total protein.

The most common substitution employed in this collection of variants was alanine. Its small, non-polar properties eliminate hydrogen bonds or ionic interactions. Glutamine was used to introduce a longer, polar side chain group, while aspartic acid's polar, charged side chain was used to replace a hydrophobic residue in the case of one variant. Backbone flexibility was investigated by introducing a proline. Another replacement for histidine was asparagine; the amide group preserves electrostatic interactions and hydrogen bonds associated with the imidazole ring. A number of amino acid replacements were made corresponding to residues within *A. aeolicus* PD or *E. coli* CM-PD in order to probe substrate specificity.

The variants that displayed the greatest decrease in specific activity (380 to 3200-fold) were S178A, S178Q/H179A, G221Q, T117P, and R217A. G221S showed a moderate loss in activity (36-fold), while others showed activities comparable to the wild-type enzyme. H112A and D171A variants did not display any detectable activity. The activities for the variant proteins will be described further in the “Discussion”, by relating the substitutions to crystal structures, sequence alignments, literature and molecular docking.

Activities of each of the variants were also determined under the following conditions: in the presence of L-tyrosine, replacing arogenate with prephenate, and replacing NADP<sup>+</sup> with NAD<sup>+</sup>. Activity measurements in the presence of prephenate were

performed with NAD<sup>+</sup> and NADP<sup>+</sup>. No turnover was observed for the wild-type AD or any of the variants when prephenate and/or NAD<sup>+</sup> were used as substrates. Additionally, there was no significant decrease in specific activities (5% or less) when variants were assayed in the presence of 500 μM L-Tyr, indicating that the enzymes are not sensitive to feedback inhibition.

**Table 5: Specific activities for wild-type AD and variants.**

Protein	Specific Activity (umol/min/μg)	Fold-decrease in Activity to WT AD
<b>WT</b>	647 ± 55	1
<b>S178A</b>	0.40 ± 0.03	1617
<b>H179A</b>	143 ± 19	4.5
<b>S178Q/H179A</b>	0.20 ± 0.01	3205
<b>M228D</b>	612 ± 73	1.1
<b>G226Y</b>	524 ± 80	1.2
<b>H170N</b>	2.5 ± 0.3	258
<b>H112A</b>	N/A <sup>a</sup>	N/A <sup>a</sup>
<b>D171A</b>	N/A <sup>a</sup>	N/A <sup>a</sup>
<b>G221Q</b>	1.7 ± 0.1	380
<b>G221S</b>	18 ± 2	36
<b>T117P</b>	1.6 ± 0.4	404
<b>R213A</b>	178 ± 12	3.6
<b>R217A</b>	1.1 ± 0.08	588
<b>H264Q</b>	499 ± 23	1.3
<b>H264Q/R265H/L266R</b>	437 ± 50	1.5

Reactions were performed at pH 7.5 and 25°C. Values are the average of duplicate reactions. Substrate concentrations were fixed at 500 μM arogenate and 1 mM NADP<sup>+</sup> in each assay. <sup>a</sup> No detectable activity when adding 60 μg to reaction mixture.

### 3.7 Molecular Docking of Arogenate in the Active Site of AD from *Synechocystis*

The crystal structure of *Synechocystis sp.* AD was solved at pH 8.0 in complex with NADP<sup>+</sup> (PDB ID: 2F1K), and is currently the only AD enzyme for which the crystal structure is available. Unfortunately, there are no structures of AD bound with L-

arogenate. In order to gain further insight into the interactions of L-arogenate with *Synechocystis sp.* AD, attempts were made to dock the cyclohexadienyl substrate into the active site of AD using AutoDock Vina. Two criteria were established to evaluate the validity of the docking conformations: first, the C-4 ring hydroxyl group of L-arogenate must be close enough to the catalytic residue, H112, to form a hydrogen bond. Second, the C-4 hydrogen atom of L-arogenate must be close enough to the C-4 carbon of NADP<sup>+</sup> to facilitate hydride transfer. However, none of the conformations that were generated met either of these criteria. As such, an alternative strategy was adopted; when the backbone atoms of the NAD<sup>+</sup>-L-Tyr-enzyme and the NAD<sup>+</sup>-enzyme crystal structures of *A. aeolicus* Δ19PD were aligned, it was found that the structures possessed a root-mean square deviation (RMSD) of 0.6 Å. This small difference was visible in the active site, however, showed that the L-Tyr-complexed structure was in a slightly more closed conformation. Accordingly, it was hypothesized that this more “closed” form of the active site could be the preferable conformation for L-arogenate to dock. Homology models were generated using the crystal structure of *A. aeolicus* Δ19PD in complex with L-Tyr and NAD<sup>+</sup> (PDB ID: 3GGG) with MODELLER 9v7 software. Ten homology models were generated, and their discrete optimized protein energy (DOPE) scores and RMSDs between the model and both the AD and Δ19PD crystal structures were evaluated (Table 6). The DOPE score is a statistical potential that takes into account non-interacting atoms within a calculated area based on the structure of the template (74). A lower, more negative score indicates that the generated model is of higher quality. All ten models showed very similar DOPE scores and RMSDs to both crystal structures. Model nine, with the lowest RMSD, was selected for docking with AutoDock Vina. L-arogenate

was successfully docked into the active site of the homology model of AD; the two previously outlined criteria were met (Figure 20). For comparison purposes, L-arogenate was also modeled into the active site of the NADP<sup>+</sup>-bound crystal structure.

Nine different L-arogenate conformations were generated simultaneously by the docking simulation. All of the conformations had binding affinities between -6.1 kcal/mole and -4.8 kcal/mole. These differences are very small and cannot exclusively be used to evaluate the conformations.

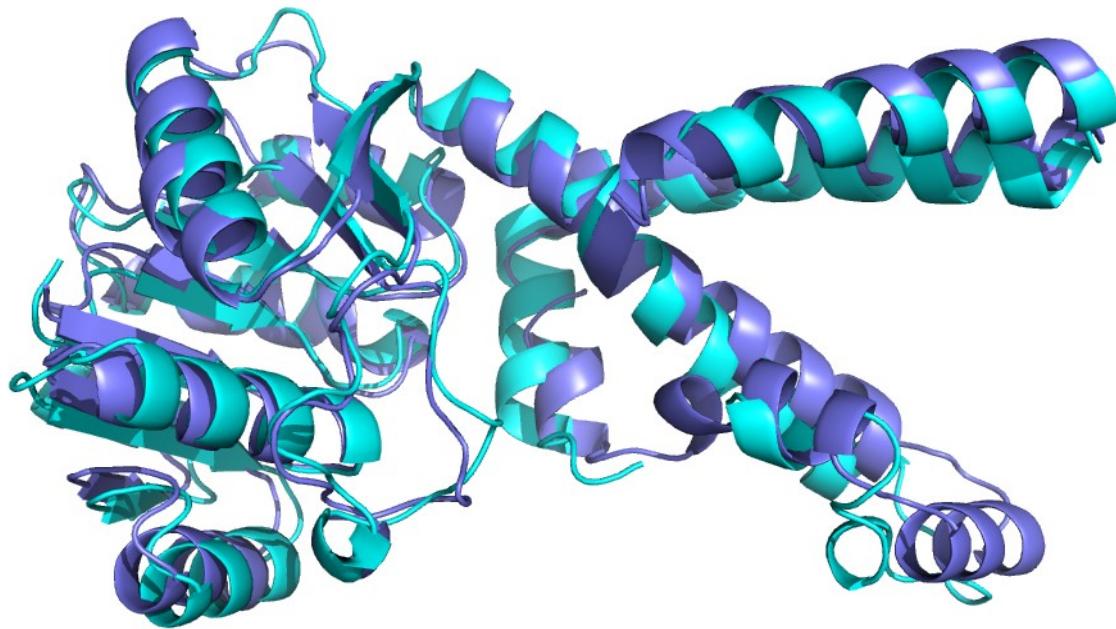
**Table 6: Summary of RMSD and DOPE score values of AD homology models**

Model Number	DOPE Score	RMSD from AD (Å)	RMSD from Tyr-bound 19PD (Å)
1	-26304	2.502	0.324
2	-26135	2.492	0.373
3	-26253	2.731	0.335
4	-26401	2.773	0.347
5	-26094	2.810	0.338
6	-26645	2.199	0.386
7	-26362	2.991	0.345
8	-26267	2.662	0.326
9	-26709	2.883	0.309
10	-26154	2.229	0.394

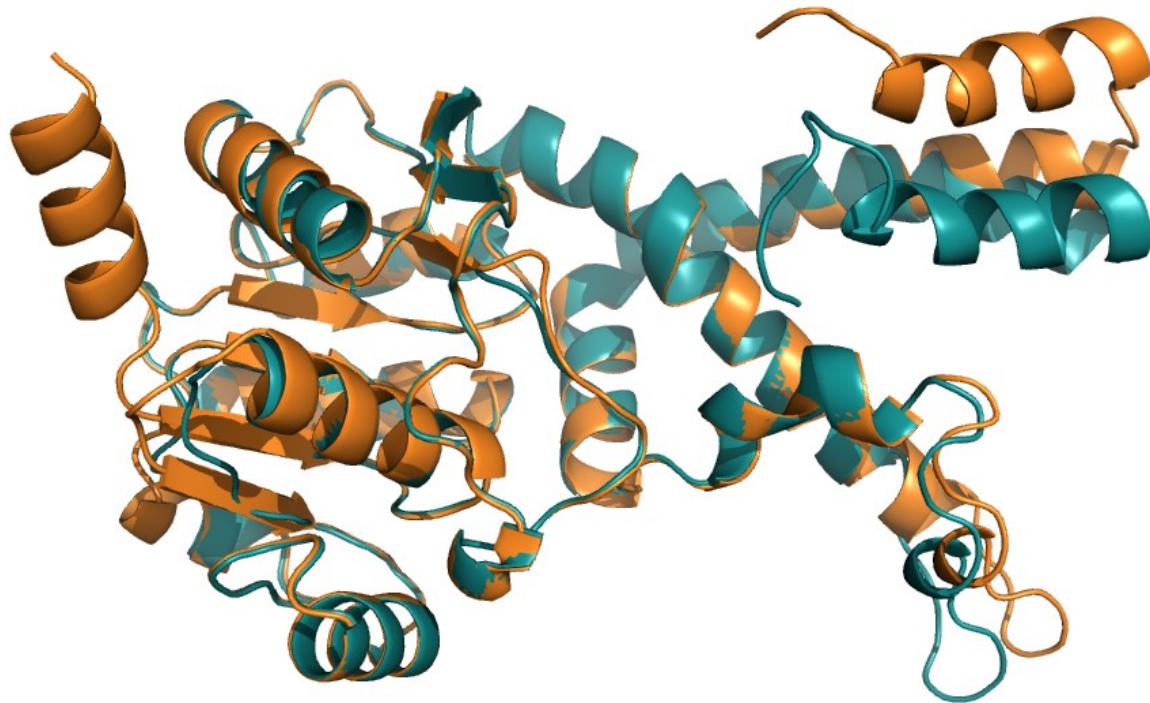
These binding affinities also suggest that it is possible for L-arogenate to exist in more than one of the proposed conformations, or that it can oscillate between a number of conformations while in solution.

A number of similarities as well as differences were observed between the residues interacting with L-arogenate when placed into the active site of NADP<sup>+</sup>-bound crystal structure of AD and docked in the homology model. When L-arogenate was placed into the active site of the crystal structure, the C4 ring hydroxyl of L-arogenate is

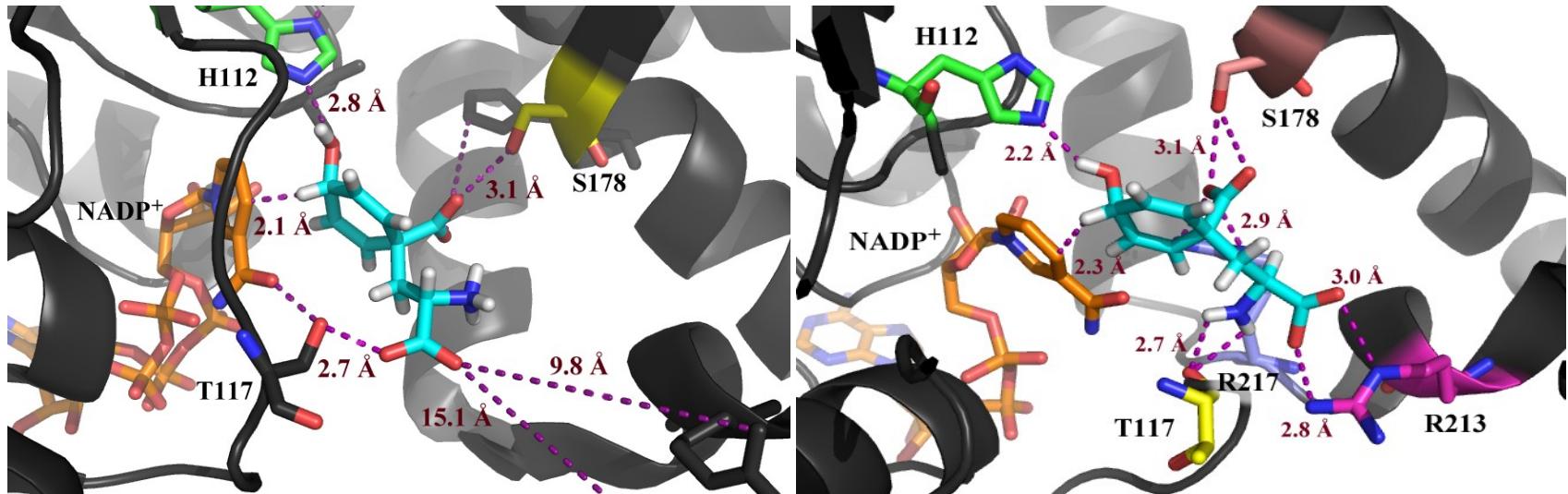
2.8 Å from H112, and the C-4 hydrogen of arogenate is 2.1 Å from the C-4 atom of NADP<sup>+</sup>. There are only two other residues that are close enough to potentially interact with L-rogenate directly. The first is S178, which maintains a distance of 3.4 Å away from the ring carboxyl group of L-rogenate. The second is the side chain of T117, which is 2.6 Å away from the side chain carboxyl group of L-rogenate. In the docked model, these residues are observed to interact with L-rogenate as well. In this structure, the ring hydroxyl of L-rogenate is 2.2 Å from H112, and the C4 ring hydrogen of L-rogenate is 2.3 Å from the C4 atom of NADP<sup>+</sup>. Additionally, S178 is 3.1 Å away from the ring carbonyl group, and the backbone carbonyl of T117 is interacting with the amine group of L-rogenate, at a distance of 2.7 Å. The most striking observation is that R213 and R217 are both interacting directly with L-rogenate in this conformation. R217 is in the center of the active site, located 2.9 Å or 3.0 Å away from the ring carboxylate of L-rogenate. Additionally, R213 is 2.8 Å or 3.0 Å away from the side chain carboxyl group of L-rogenate. In the NADP<sup>+</sup>-bound crystal structure, these two arginine residues are too far away from L-rogenate (Figure 21) to interact with this substrate. As such, the enzyme would have to undergo a considerable tertiary rearrangement upon the binding of L-rogenate in order for these residues to interact with the cyclohexadienyl substrate. Mutagenesis data supports the idea that both of these arginines play a role in enzyme function, with R217 being much more important than R213.



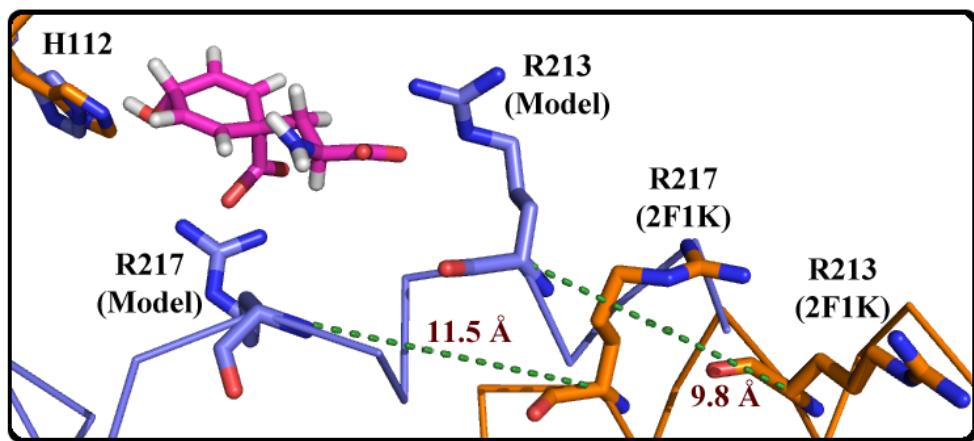
**Figure 18: Alignment of the backbone atoms of monomers of *Synechocystis* AD and *A. aeolicus* Δ19PD crystal cofactor bound structures.** AD (PDB ID: 2F1K) is colored in blue while Δ19PD (PDB ID: 2G5C) is colored in purple. The RMSD of the alignment was 1.3 Å. Figure created using PyMOL (44).



**Figure 19: Alignment of the backbone atoms of the *Synechocystis* AD model nine and its Δ19PD template.** The AD homology model (teal) and its template, Δ19PD in complex with NAD<sup>+</sup> and L-Tyr (orange, PDB ID: 3GGG) were found to have an RMSD of 0.309 Å. Figure created using PyMOL (44).



**Figure 20: Active site of *Synechocystis* sp. AD with L-aro-ogenate modeled into the active site of the NADP<sup>+</sup>-bound crystal structure and docked into the homology model.** Left: L-aro-ogenate (cyan) modeled into the active site of the experimentally determined NADP<sup>+</sup>-bound crystal structure of AD. The closest atoms on R217 and R213 are 9.8 Å and 15.1 Å away from the closest atom of L-aro-ogenate, respectively. Shown are possible interactions of L-aro-ogenate with key residues of the active site. Right: L-aro-ogenate docked into the active site of the AD homology model that had been generated using the NAD<sup>+</sup>-L-Tyr-bound crystal structure of *A. aeolicus* Δ19PD as a template. Possible interactions of L-aro-ogenate and NADP<sup>+</sup> with key residues are also shown. Images were generated in PyMOL (44), docking files were prepared using AutoDock tools and AutoDock Vina (75) as described in Materials and Methods.



**Figure 21: Comparison of arginine positions in the crystal structure and homology model of *Synechocystis* sp. AD.** The backbone atoms of the crystal structure (labelled 2F1K) were aligned with those of the homology model of AD. The distances between the α-carbons were measured to determine their differences for R213 and R217 to be present in the active site. Image generated in PyMOL (44).

### **3.8 Enzymatic Synthesis of L-arogenate**

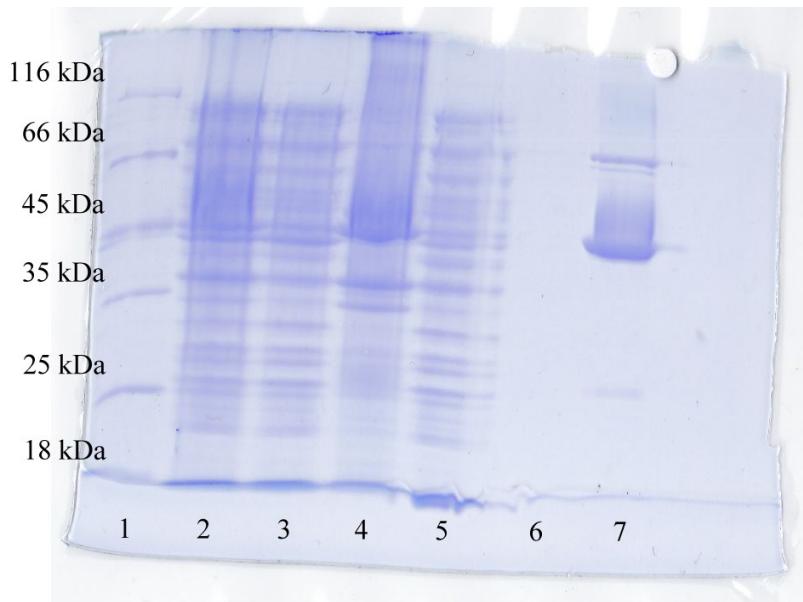
#### **3.8.1 Expression and Purification of Prephenate Aminotransferase from *Petunia Hybrida***

Preliminary studies were attempted to obtain L-arogenate from the accumulation media of *S. typhimurium* tyrA19 mutant which carried metabolic blocks in the L-arogenate pathway (Figure 9). We also followed the procedure reported by Rippert et al (56, section 3.8.2) to synthesize L-arogenate enzymatically from prephenate on a small scale using the crude preparations of prephenate aminotransferase (PAT) from the extracts of tobacco leaves. In our hands, both procedures yielded low amounts of impure L-arogenate and were very time consuming (data not shown). We chose to develop the enzymatic route further, using a recombinant form of PAT and a HPLC separation methodology to remove the nucleotide cofactors from the synthesis mixture. This will be described below.

Prephenate aminotransferase (PAT) was expressed and purified using a protocol based on the report by Maeda and colleagues (54). The expression vector, pET24b, containing the gene encoding the PAT was transformed into BL21(DE3) cells, and then the cells were grown in LB medium supplemented with 50 µg/µL of kanamycin. Protein expression was induced by the addition of IPTG. Cells were lysed by high pressure while in the presence of protease inhibitors, and the recombinant protein was purified using Ni-NTA affinity chromatography at 4°C. The purified protein was dialyzed overnight at 4°C.

The purity of the protein was verified at each step of the purification by denaturing acrylamide gel electrophoresis (Figure 22). Surprisingly, the presence of the over-expressed protein is not clearly visible in the cell lysate and the cell-free extract

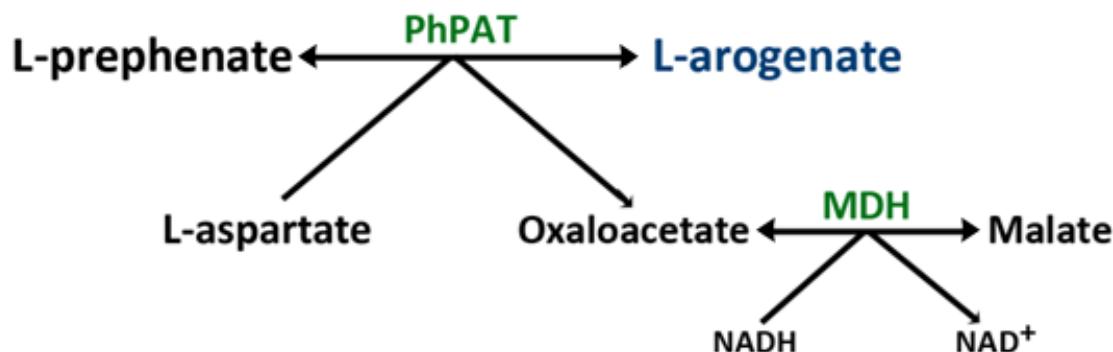
(Lanes 2 and 3), although, some of the enzyme appeared to be associated with the particulate fraction (Lane 4). The lysate was applied to the Ni-NTA column and the flow through was collected. No relevant band at the weight corresponding to PAT can be observed in the flow through (Lane 5), suggesting that the majority of the recombinant hexa his-tagged enzyme bound to the column. After washing the column extensively with buffer (~15 column volumes, Lane 6), the protein was eluted with buffer containing a high concentration (300 mM) of imidazole. The molecular mass of His-tagged PAT is calculated to be 47.6 kDa, and is in agreement with the major band observed by SDS-PAGE analysis (Lane 7). The purification protocol also yielded a higher molecular weight species (>66.2 kDa) as well as a protein smear just above the band for the recombinant PAT enzyme which might represent degradation products of the high-molecular weight contaminant (Lane 7).



**Figure 22: SDS-PAGE analysis of the purification of PAT from *Petunia hybrida*.** Lane 1: Protein molecular weight marker, Lane 2: Cell lysate, Lane 3: Cell-free extract, Lane 4: Cell pellet dissolved in 2% SDS, Lane 5: Ni-NTA flow-through, Lane 6: Wash with 30 mM imidazole wash, Lane 7: pooled protein eluate with 300 mM imidazole wash. Analysis is performed on a 12% acrylamide gel.

### 3.8.2 Coupled enzymatic Assay for L-rogenate Synthesis

Partially purified PAT can be used to enzymatically synthesize L-rogenate. PAT transaminates prephenate to L-rogenate in the presence of PLP and an amino donor, which can be either L-aspartic acid or L-glutamic acid (54). However, unlike the rearrangement of chorismate to prephenate by CM or the decarboxylation of prephenate to HPP by PD, the reaction catalyzed by PAT is reversible. Furthermore, none of the substrates or products of the reaction possess any distinctive spectrophotometric or fluorometric properties,

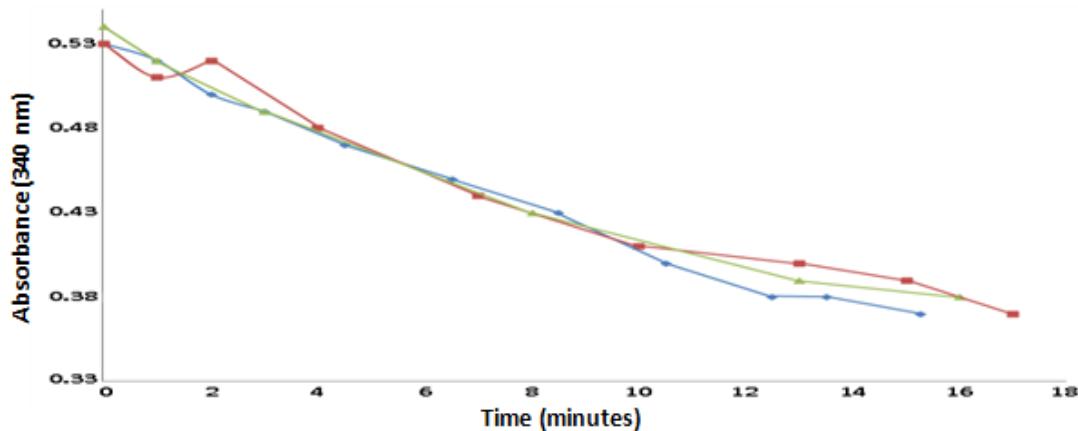


**Figure 23: Coupled PAT assay for L-rogenate synthesis.** Prephenate is reversibly transaminated to L-rogenate by PhPAT in the presence of L-aspartate or L-glutamate. When L-aspartate is the amino donor, both oxaloacetate and L-rogenate are produced. The oxaloacetate can then be reversibly converted to malate in the presence of NADH and malate dehydrogenase (MDH).

making the enzymatic reaction difficult to monitor continuously. As such, the reaction catalyzed by PAT was coupled to the reaction catalyzed by malate dehydrogenase (MDH), as described by Rippert and colleagues (56). The coupled assay is summarized in Figure 23, which shows that since NADH is a component of the assay, the progress of the reaction can now be monitored by observing the decrease in absorbance at 340 nm.

Furthermore, the production of L-rogenate is favoured by the conversion of oxaloacetate by MDH.

The coupled enzymatic reactions were performed at 37°C, the optimum temperature for PAT-catalyzed reaction (54). Until all the prephenate was consumed ( $\Delta A_{340}$  of 0.33, 15-20 minutes), the reaction was stopped by snap-freezing in liquid nitrogen; aggregated proteins can later be separated by centrifuging of the thawed solution.



**Figure 24: Time-dependant consumption of NADH by MDH-catalyzed reaction.** NADH disappearance is monitored spectrophotometrically at 340 nm as oxaloacetate is converted to malate. Oxaloacetate is a product of the transamination of prephenate to L-rogenate by PAT. Barium prephenate (5.4 mM) was converted the presence of 16.2 mM NADH (3x [prephenate]), 10 mM aspartate, 50  $\mu$ M PLP, 12-17  $\mu$ g PAT, and ~44 units MDH, in a volume of 1 mL. The  $A_{340}$  was monitored by making a 1:200 dilution into 50 mM H<sub>2</sub>O. At this dilution, the concentration of NADH is 81  $\mu$ M, which corresponds to an absorbance of 0.55. The expected  $A_{340}$ , if 5.4 mM prephenate is totally consumed, would decrease to 0.33, corresponding to a [NADH] of 54  $\mu$ M after being diluted. N = 3

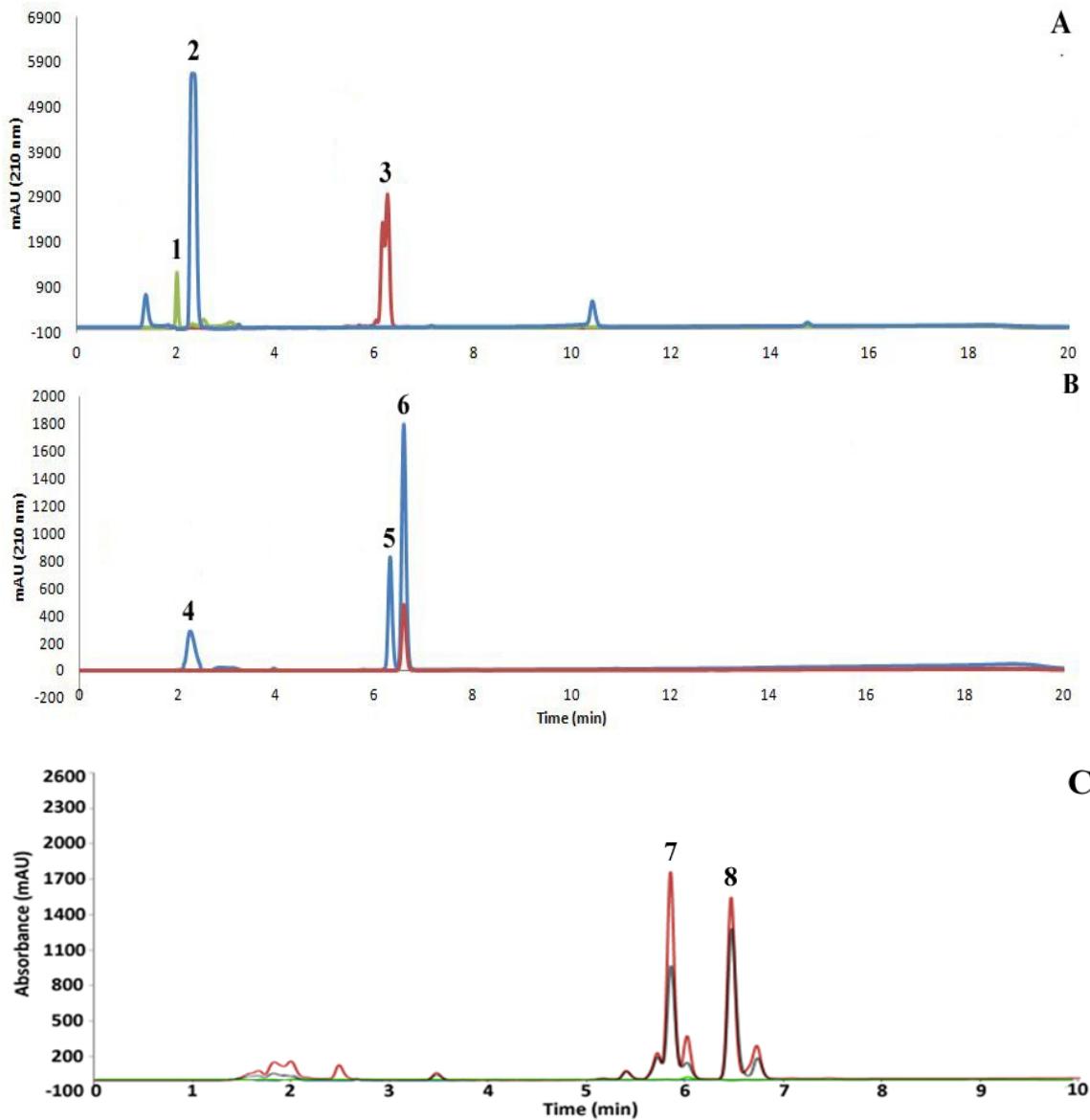
### 3.9 Purification of Arogenate

#### 3.9.1 Analytical HPLC and LC-MS

With the enzymatic conversion of prephenate to L-rogenate by the coupled assay, a number of other compounds remain in solution: NADH, NAD<sup>+</sup>, aspartate,

malate, prephenate, oxaloacetate and PLP. Of these compounds, NADH and NAD<sup>+</sup> are present in the highest concentrations and are the most problematic. The activity of AD is measured by following the time-dependent production of NADPH at 340 nm; contaminating NADH interferes with this measurement, and NAD<sup>+</sup> in solution will degrade into compounds that also absorb at 340 nm, and this process is hastened by repeated freeze-thaw cycles (69). Therefore, it is essential that L-rogenate is separated from NADH and NAD<sup>+</sup>. HPLC is a standard method to separate small compounds, and has been used previously to separate arogenate from other compounds (54, 58). However, in these experiments, arogenate and any other compounds containing a primary amine were covalently derivatized with OPA for fluorometric detection at 336 nm. Since the OPA-derivatized arogenate cannot be used for kinetic studies, the development of a HPLC method to separate unmodified L-rogenate from NAD<sup>+</sup> and NADH was of clear interest. A separation protocol was developed first on the analytical scale, and then scaled up to the preparative level.

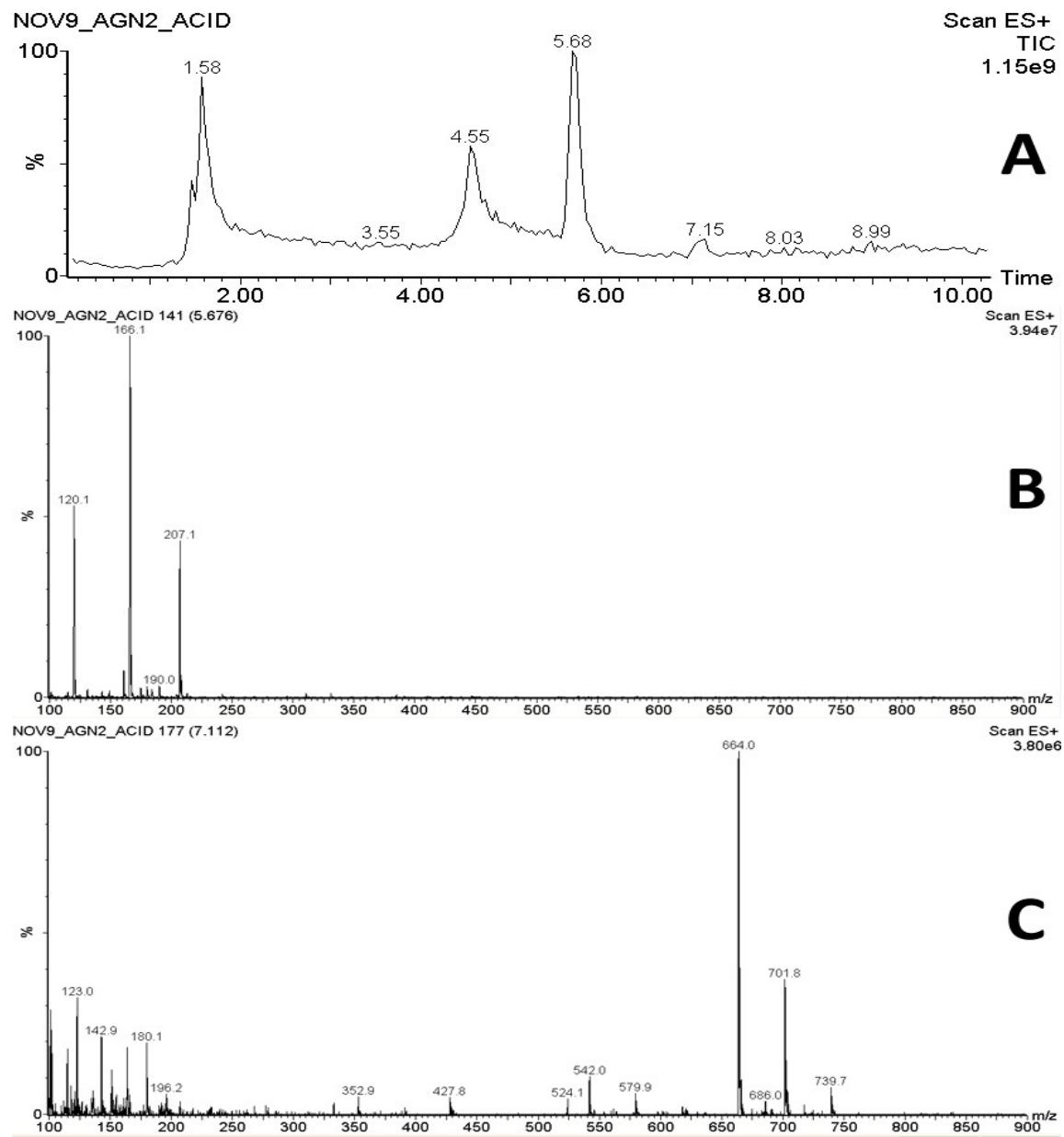
Preliminary analytical HPLC separations were performed by using the protocol reported by Maeda and colleagues (54). Standard solutions of NAD, NADH, and prephenate (the latter was used as a standard for L-rogenate) were chromatographed separately on a C<sub>18</sub> column (5 μm, 80 Å, 150 x 3 mm) using a 15-min linear gradient of 15-65% methanol in sodium phosphate buffer. Regrettably, all compounds eluted at approximately 2.1 min, and further attempts to resolve the compounds using different gradient and isocratic runs yielded similar results (data not shown). However, when the organic solvent was replaced with a 1-30% linear gradient of acetonitrile, NADH eluted at 6.2 min, approximately 4 min after prephenate (Figure 25A, peak 3). Prephenate eluted



**Figure 25: Analytical HPLC chromatograms of standards and arogenate reaction mixture.** A: Approximately 2.2 µg of prephenate (peak 1) and 6.6 µg of NADH (peak 3) were individually injected onto a C<sub>18</sub> column and resolved with a 15-minute linear gradient of 1-30% acetonitrile and 20 mM sodium phosphate (pH 7.6). Peaks were detected at 210 nm. Prephenate was also acidified to produce phenylpyruvate by the addition of concentrated HCl and 2.2 µg was injected as well (peak 2). B: A 100 µL aliquot of a 1:400 dilution of the experimental arogenate reaction was injected onto the same column and compounds separated using the same conditions as the standards. Detection was performed at both 210 nm (blue) and 340 nm (red). C: A small amount of concentrated HCl was added to the sample and incubated for 20 min at room temperature. The reaction mixture after acidification was monitored at 210 nm (red), 259 nm (dark purple), and 340 nm (green). A 100 µL aliquot of a 1:400 dilution of the acidified experimental arogenate reaction was injected onto the same column and run at the same condition as the standards.

at 2.0 min (Figure 25A, peak 1) and phenylpyruvate (produced by the acidification of prephenate) eluted at 2.3 min (Figure 25A, peak 2). Next, HPLC separation was performed on the experimental arogenate reaction mixture under the conditions used to resolve the standards (Figure 25B). As shown in Figure 25A, two intense peaks at 6.3 (Figure 25B, peak 5) min and 6.5 min (Figure 25B, peak 6) were observed likely corresponding to NAD<sup>+</sup> and NADH, respectively. The peak at 6.5 min also displayed an absorbance at 340 nm, thus confirming the presence NADH. A small peak at 2.2 min (Figure 25B, peak 4) was observed, likely corresponding to the elution of arogenate, aspartic acid, and malate. In order to confirm the presence of L-arogenate in the reaction mixture, the sample was also acidified with HCl. The chromatogram of the acidified mixture is shown in Figure 25C. After the sample is acidified, two peaks are observed at 5.8 and 6.4 min (Figure 25C, peaks 7 and 8). The injection was also monitored at 259 nm, and these two peaks absorb at this wavelength as well. The absorbance at 340 nm drops to a nearly undetectable level at 6.0 minutes after acidification, suggesting that NADH has decomposed under the acidic conditions.

LC-MS was used to help identify the compounds from the acidified mixture (Figure 26). The presence of phenylalanine confirms the acidification of L-arogenate, for which a m/z peak was observed (Peak of 5.68 min, m/z 166.1, Figure 26B). The m/z peak corresponding to the mass of NADH was observed at 7.1 minutes (m/z 664.0, Figure 26C), which is notably higher than previously observed. As the LC-MS was performed with MilliQ water rather than phosphate buffer, it is possible that this change in the mobile phase has altered the retention times of the compounds. Its intensity has also been markedly reduced, supporting the hypothesis that the cofactor has undergone acidic

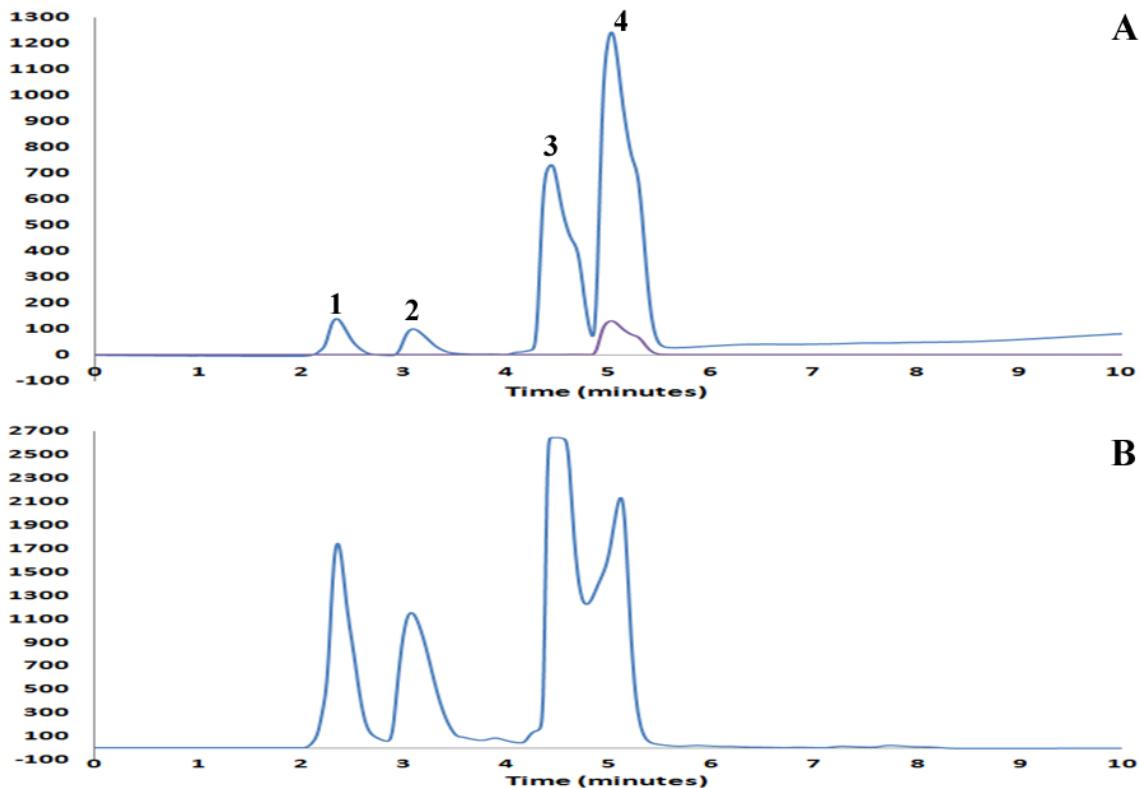


**Figure 26: Chromatogram of acidified experimental arogenate mixture and spectra of peaks corresponding to the masses of phenylalanine and NADH.** A: The chromatogram of ions detected by the MS after being subjected to HPLC on the same C<sub>18</sub> column used for previous separations. 100 μL of 1:400 diluted experimental acidified arogenate mixture was injected onto a C<sub>18</sub> column with a 15-minute linear gradient of 1-30% MilliQ H<sub>2</sub>O and acetonitrile. Major peaks are observed at 1.58, 4.55, and 5.68 minutes. B: The mass spectrum corresponding to the 5.68 minute peak of the chromatogram. The major peak has a m/z of 166.1, which corresponds to the molecular weight of phenylalanine. C: The mass spectrum corresponding to the 7.11 minute peak of the chromatogram. The major peak m/z peak of 664.0 corresponds to the molecular weight of NADH.

decomposition. The presence of peaks at 1.58 and 4.55 minutes (Figure 26A) further supports this view, as the peaks appear to correspond to the m/z peaks of degradation products. The peaks at 1.58 and 4.55 minutes had major m/z peaks of 277.1 and 143.1, respectively (data not shown).

### **3.9.2 Preparative HPLC**

A diluted aliquot of the reaction mixture yielding L-arogenate was resolved by preparative HPLC using a XBridge C<sub>18</sub> column (5 µM, 100 x 10 mm, Waters), kindly provided by Dr. Pat Forgione. Conditions were identical to those on the analytical scale experiments, with the exception that the flow rate was increased from 1.0 mL/min to 4.0 mL/min. A small amount diluted sample mixture was first applied to observe the retention times of the compounds in the mixture on the preparative column (Figure 27). A similar pattern is observed, although the peaks are broader and the NAD<sup>+</sup> and NADH peaks are observed at 4.5 and 5.1 min, respectively. Additionally, two small peaks are observed at elution times of 2.4 min and 3.1 min. When a larger amount of the reaction mixture is chromatographed separation is maintained between these two peaks (Figure 27). Fractions corresponding to these two peaks were collected for future MS analysis.



**Figure 27: Preparative HPLC chromatogram of arogenate reaction mixture.** A: An aliquot of 20  $\mu$ L of a 1:2 diluted experimental L-argenate mixture was injected onto a C<sub>18</sub> column with a 15-min linear gradient of 1-30% 20 mM sodium phosphate (pH 7.6) and acetonitrile and detected at 210 nm (blue) as well as 340 nm (purple). B: An aliquot of 100  $\mu$ L of the undiluted experimental L-argenate mixture was injected onto the same column. Detection was performed at 210 nm.

## **Chapter 4: Discussion**

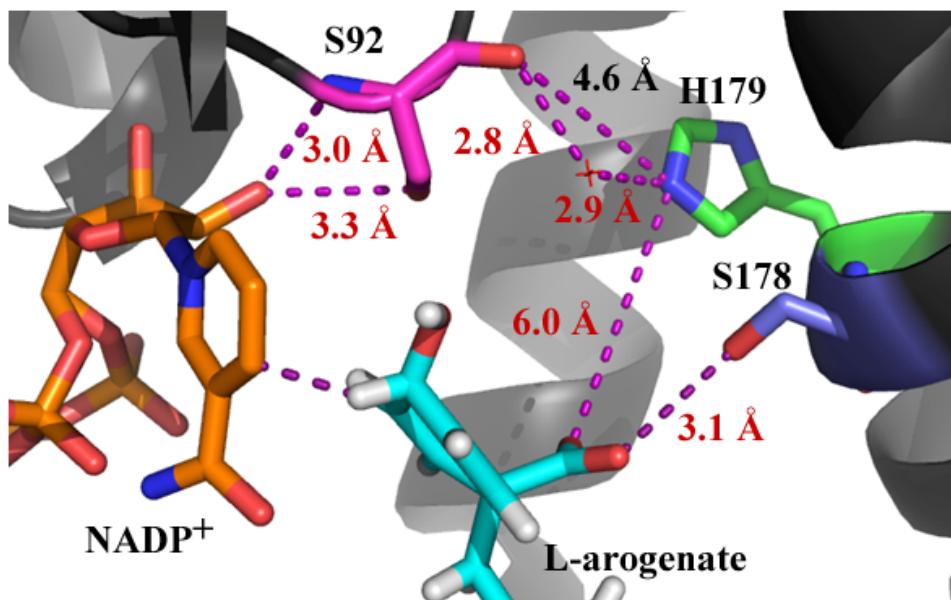
The purpose of this study was to identify which amino acid residues which are important for the function of the monofunctional arogenate dehydrogenase from the cyanobacterium *Synechocystis sp.* Additionally, we attempted to develop a novel method to synthesize and purify arogenate using analytical techniques. This work presents the first mutagenesis studies on *Synechocystis sp.* AD. In total, fifteen different variants were constructed and studied. Amino acid substitutions were selected based on the crystal structure of AD, docking simulations to a homology model of AD, and sequence alignments with other TyrA proteins. Some residues selected are highly conserved; these residues include H112, H170, D171, and R217. The majority of variants harboured single amino acid substitutions, along with a double substituted variant (S178Q/H179A) as well as a triple variant (H264Q/R265H/L266R).

It was found that all variants, with the exception of D171A, were purified to homogeneity using Reactive Red 120 agarose affinity chromatography and in good yields. The specific activity values were found to be reproducible. The results have been interpreted while referring to the crystal structure of AD, as well as the crystal structures and previous mutagenesis studies on other TyrA proteins.

### **Serine 178 and Histidine 179**

S178 and H179 are located in the middle of the active site of AD from *Synechocystis sp.* Their proximity to L-rogenate after being both placed and docked in the active site suggests that these residues may play a role in substrate binding and/or catalysis, with S178 predicted to be of greater importance as its side chain is within

hydrogen-bonding distance from L-rogenate (Figure 28). Both S178 and H179 are conserved in several organisms (Figure 4), but are respectively a glutamine and alanine in other TyrA proteins, including CM-PD from *E. coli* and *H. Influenzae*. When this glutamine residue in *E. coli* CM-PD was replaced with an alanine residue, a decrease in  $k_{cat}$ , without a marked change in  $K_m$  for prephenate, was observed. In the monofunctional PD from *A. Aeolicus*, the equivalent residues for S178 and H179 are, S213 and H214, respectively. Both residues are involved in a hydrogen bonding network with S126, H147 (the catalytic residue in PD), and WAT1, a highly conserved water molecule and were found to be important for prephenate binding, only showing modest decreases in  $k_{cat}$ . (42) Of these residues, S213 was found to be more important than H214 in positioning WAT1 with the C4-hydroxyl group of the ligand. However, in AD, no evidence for such a hydrogen bonding network is observed in the NADP<sup>+</sup>-bound crystal structure. S178 was converted to an alanine, to probe the possibility of any hydrogen bonds involving the side chain hydroxyl group of the residue. When L-rogenate was modeled into the NADP<sup>+</sup>-bound crystal structure and docked into the homology model, the hydroxyl of S178 is proposed to form a hydrogen bond with the ring carboxyl group of L-rogenate (Figure 28). The substitution to alanine resulted in a ~1600-fold drop in activity relative to the wild type enzyme, which indicates that it is essential in the catalytic mechanism of AD. Although S178's interaction with the ring carboxyl group of L-rogenate may help lock it in the active site, S178 might also play a catalytic role.



**Figure 28: Possible interactions of S178 and H179 with L-arogenate.** L-arogenate was placed into the active site of the  $\text{NADP}^+$ -bound crystal structure of *Synechocystis* AD and the distances of S178 and H179 from different atoms of L-arogenate were measured in PyMOL. Figure generated in PyMOL (44).

A hydrogen bond formed between the side chain of S178 and the carbonyl group of the L-arogenate's ring carboxyl moiety, might help to polarize the carbonyl oxygen thus promoting decarboxylation of the substrate;  $\text{CO}_2$  is one of the products of the reaction. The kinetic data greatly supports the hypothesis that S178 interacts directly with L-arogenate and could be involved in binding and/or catalysis.

H179 was also individually converted to an alanine residue, and this variant only revealed a modest ~4-fold reduction in enzyme activity. As reflected in both the modeled and docked AD structures with L-arogenate, H179 appears to be too far away to interact directly with L-arogenate (Figure 28). Since H179 appears too far away from L-arogenate, it is possible that the residue may play an indirect role in  $\text{NADP}^+$  binding. One of the residues that binds to  $\text{NADP}^+$  is S92, which is located on a flexible loop in the active site. A water molecule bridges H179's imidazole nitrogen and S92's side chain

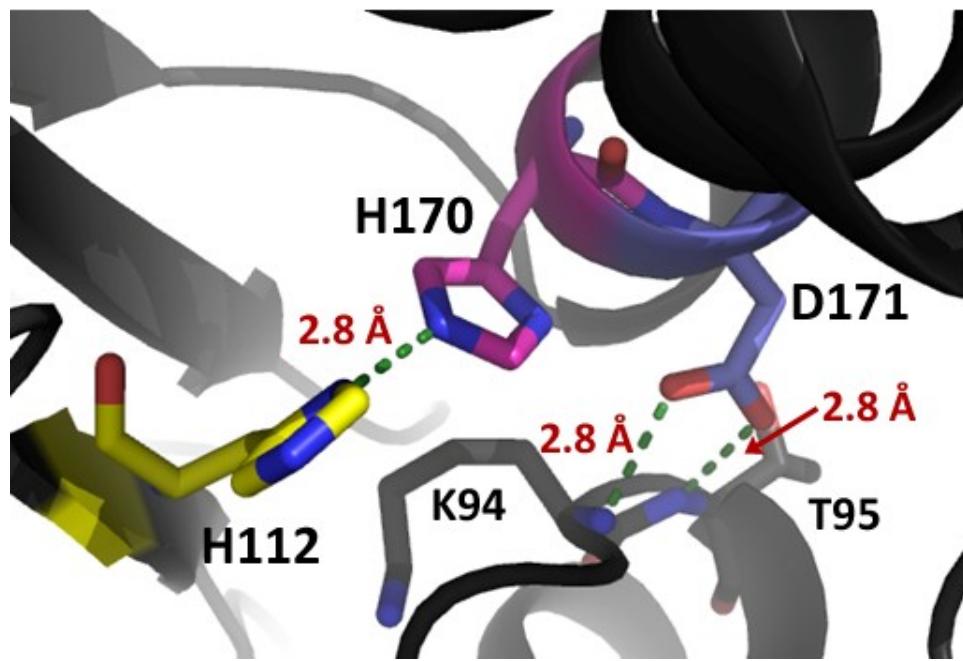
hydroxyl group in the NADP<sup>+</sup>-bound crystal structure of AD. Additionally, the backbone amide and carbonyl groups of S92 appear to coordinate the ribose group of NADP<sup>+</sup>. Thus, this extensive hydrogen bonding network appears important to help coordinate NADP<sup>+</sup> into the active site.

S178 was also converted to a glutamine residue in combination with H179 to an alanine in the double variant. In this case, the length and flexibility of the side chain has increased, although glutamine retains the potential for hydrogen bonding interactions with L-rogenate and other enzymatic residues. If the S178Q substitution affected enzyme activity to the same degree as the S178A single variant, a ~6400-fold decrease in enzyme activity would be expected. However, a ~3200-fold reduction in enzyme activity was observed, which suggests that the S178Q substitution is not as detrimental to the architecture of the active site. It is possible that glutamine's polar side chain can rescue some loss of enzyme function incurred by the ability to retain hydrogen bonding interactions with L-rogenate and/or the water molecule. (Figure 28) The longer side chain length of glutamine compared to serine appears to be less favourable, due to increased flexibility and/or the introduction of steric clashes. An additional reason to study the S178Q/H179A double variant was to probe any changes in substrate specificity, since the glutamine-alanine pair is only observed in TyrA proteins with a high preference for prephenate. However, no activity for the double variant was detected in the assays with prephenate and either NADP<sup>+</sup> or NAD<sup>+</sup> as substrates.

## **Histidine 112 and 170**

H112 is another key residue located in the core of the active site of *Synechocystis* sp AD. It is highly conserved among all TyrA proteins, and has been identified as the catalytic hydrogen bond acceptor in a number of these proteins, including *E. coli* and *H. influenzae* CM-PD, as well as *A. aeolicus* Δ19PD (42). H112A is the only AD variant that lacks any detectable enzyme activity while maintaining global secondary and tertiary structural arrangements comparable to that of the wild-type enzyme as determined by spectroscopic measurements (Figure 13, 16). The conversion to alanine removes the H112's ability to hydrogen bond with the C-4 hydroxyl group of L-argenate. This hydrogen bond is hypothesized to be essential for the polarization of the ring hydroxyl group in order to facilitate electron transfer and hydride transfer. (48) As such, it is highly likely that H112 is the key catalytic hydrogen bond acceptor for the AD-catalyzed reaction. It remains to be determined if this histidine group is solely involved in catalysis and not binding the cyclohexadienyl substrate, as has been proposed for TyrA from *A. aeolicus* (43) and *E. coli* CM-PD (37). Determination of a  $K_m$  or  $K_d$  for the interaction of L-argenate with H112A is necessary.

H170 is another highly conserved residue in TyrA proteins. It is located directly above H112 in the NADP<sup>+</sup>-bound crystal structure of AD as well as the AD homology model, and appears to form an interaction with H112, as the imidazole nitrogen atoms of the two residues are 2.8 Å away from each other (Figure 29). The equivalent residue in *A. aeolicus* Δ19PD, H205 when changed to glutamine or leucine resulted in a 3-fold and 500-fold decrease in  $k_{cat}/K_m$ , respectively (42). It was hypothesized that H205 (possibly a protonated ring) hydrogen bonds to H112 so that the deprotonated H112 can remain in a



**Figure 29: Possible interactions of H112, H170, and D171 in the NADP<sup>+</sup>-bound crystal structure of AD.** Figure generated in PyMOL (44).

catalytically competent conformation. As the positions of the histidine groups in the AD enzyme are almost identical to that of *A. aeolicus* PD, it is likely that H170 plays the same role. H170 was converted to asparagine, and hence, the potential to form this hydrogen bond with H112 might still be retained. The 250-fold decrease in specific activity of the H170N variant confirmed that H170 and H112 do interact but argued that an appropriate hydrogen bond was not retained with the glutamine substitution to help direct H112 for catalysis. It would be informative to measure and compare the activity for a H170L and H170Q variants, which would verify whether the presence of a hydrogen bond with H112 is critical for proper enzyme function.

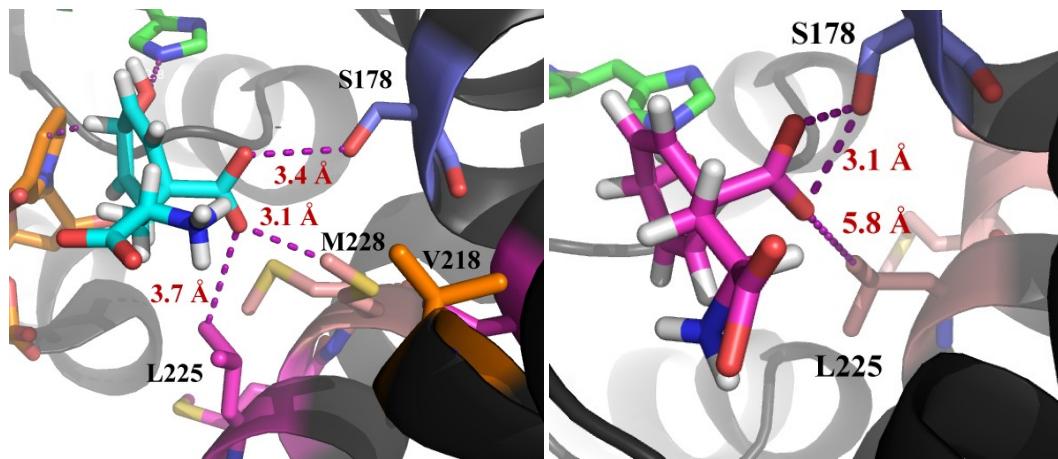
### **Aspartic Acid 171**

D171 is a highly conserved neighbour of H170. When D171 in *Synechocystis* AD was converted to alanine, the variant found to be was easily degraded, poorly expressed, and yielded a non competitive pattern of thermal unfolding as detected by far-UV CD (Figure 16). Interestingly, the compromised structure and function appear to be due to the depletion of different interactions, mainly this aspartic acid in AD. These observations mirrored the findings reported for the alanine variant of the equivalent residue in Δ19PD, D206. (42) In the NAD<sup>+</sup>-bound crystal structures *A. aeolicus* Δ19PD, the negative carboxyl moiety form a hydrogen bond with Nδ1 of H205, which in turn coordinate H147. The triad of H147, H205, and D171 had been hypothesized to play a role in substrate binding and catalysis. (42) In contrast, in the NADP<sup>+</sup>-bound crystal structure of AD, as well as the homology model of AD, D171's side chain points away from the active site and instead interacts with the backbone amide of K94, the side chain and backbone of T95, the backbones of A168 and A175, and the side chain of N233. As such, it appears that D171 creates a number of interactions that helps maintain the structure of AD in a functional conformation (Figure 29).

### **Methionine 228**

M228 is one of several residues found forming a hydrophobic pocket in the bottom of the active site. This residue is not conserved in other TyrA proteins, and is often found as an aspartic acid, which is the residue observed in the CM-PD enzymes from *E. coli* and *H. influenzae*, as well as the PD from *A. aeolicus*. It was hypothesized by Legrand and colleagues that this residue played a role in substrate specificity, and that

it and the other residues forming the hydrophobic pocket (L225, M229, V182, V218) are necessary for ensuring that a concerted reaction mechanism occurs (48).



**Figure 30: Comparison of the location of the hypothesized hydrophobic pocket in the active site of *Synechocystis* AD.** Left: L-arogenate is modeled into the active site of the NADP<sup>+</sup>-bound crystal structure of AD. Of the residues hypothesized to compose the pocket, M228 and L225 are 3.1 Å and 3.7 Å, respectively, away from one of the oxygen atoms of the ring carboxyl group of L-arogenate. Right: L-arogenate is docked into the active site of the homology model of AD. None of the residues are in proximity to the ring carboxyl group of L-arogenate, the closest being L225, at a distance of 5.8 Å. Figures generated in PyMOL (44).

This was originally proposed by Hermes and colleagues (38); a hydrophobic pocket facing the ring carboxyl group of arogenate could markedly shorten the lifetime of the dienone intermediate so that C-C cleavage occurs in the same transition state as hydride transfer. Legrand and colleague's support for this hypothesis was based on their structural model of placing L-arogenate into the active site by replacing WAT43 with the ring hydroxyl group of arogenate and ensuring that the C-4 hydrogen atom was close enough to the C-4 carbon of NADP<sup>+</sup> (Figure 8). When L-arogenate was modeled into the NADP<sup>+</sup>-bound crystal structure as well as the homology model, the ring carboxyl group was in proximity of this hydrophobic pocket, but was also interacting with the side chain

of S178 (Figure 30). This is in contrast to the docked homology model, in which the hydrophobic pocket is not observed within the proximity of L-arogenate. The M228D variant was characterized to see if the introduction of a polar, charged amino acid would disrupt the hydrophobic nature of this pocket, and affect enzyme activity. However, the specific activity of M228D was found to be comparable to that of the wild-type enzyme (Table 5), suggesting that if this hydrophobic pocket is required for a concerted reaction mechanism more residues within this pocket would need to be replaced. It is noteworthy that M228 is the closest residue to L-arogenate in the NADP<sup>+</sup>-bound crystal structure of AD (Figure 30). Considering the results of the modeling and docking studies combined with the structure and function importance of S178 presented in this thesis, it seems unlikely that this hydrophobic pocket is critical for catalysis to occur efficiently in AD.

### **Arginine 213 and 217**

In the NADP<sup>+</sup>-bound crystal structure of *Synechocystis sp* AD, R213 and R217 are considerable distances away from the active site (Figure 8), with R213 located in close proximity to the C-terminal dimerization domain of the enzyme. Legrand and colleagues (48) were the first to report this after determining the crystal structure, and as such, dismissed the possibility for either residue playing a role in L-arogenate binding or catalysis, further pointing out that there is no R217 equivalent in plant ADs. The highly conserved nature of R217 made it an attractive target for mutagenesis studies in other TyrA proteins, such as *E. coli* CM-PD (R294), *H. influenzae* CM-PD (R297), and *A. aeolicus* PD (R250). In all three of these enzymes, this arginine was found to play an important role in prephenate binding (37, 42, 43, 49). The *E. coli* CM-PD variant, R294Q, yielded a 130-fold increase in the K<sub>m</sub> for prephenate (37); the *A. aeolicus* PD

variant, R250A, was observed to have a moderate increase in  $K_m$  of 13-fold (43). The  $K_m$  for prephenate in the *H. influenzae* CM-PD variant, R297Q, was the most severely affected of the three, showing a ~750-fold increase (49). The other arginine, R213, appears to be an alanine residue in the primary sequence of the CM-PD enzyme from both organisms, and is a lysine residue, K246, in *A. aeolicus* PD. Interestingly, the PD variant, K246A, was found to have a  $K_m$  for prephenate that was 20-fold higher, suggesting that K246 is more important than R250 in prephenate binding.

Although the NADP<sup>+</sup>-bound crystal structure of AD shows, R213 and R217 are too far away to interact with L-aroogenate when this substrate was modeled into the active site (Figure 8), the AD variants, R217A and R213A, were constructed to determine any correlation with the previous functional studies on other TyrA protein variants. Surprisingly, R217A and R213A were observed to have a ~580-fold and ~4-fold drop in activity, respectively. As this result contradicts the information obtained from the available crystal structure of AD, molecular modeling and docking techniques were employed in order to attempt to rationalize how these residues, particularly R217, could play a role in maintaining enzyme activity.

We had carried out numerous docking simulations with the NADP<sup>+</sup>-bound crystal structure of AD, under different conditions, which included: introducing flexible residues; using different docking software; and increasing the exhaustiveness of the search to high levels. As previously mentioned, two criteria were used to judge the credibility of the docked conformations: the C-4 ring hydroxyl group of L-aroogenate must be close enough to H112 so that they can hydrogen bond, and the C-4 ring hydrogen of L-aroogenate must be close enough to the C-4 carbon of the nicotinamide ring of NADP<sup>+</sup>.

so that hydride transfer can occur. However, no conformations were found that matched these two criteria. It had been previously observed that aligned backbone atoms of the NAD<sup>+</sup>-bound crystal structure of *A. aeolicus* Δ19PD with that of the NAD<sup>+</sup>-enzyme-L-Tyr structure of the same enzyme possessed an RMSD of 0.6 Å. This difference was visible in the active site, where the amino acid residues in the L-Tyr-bound enzyme had “closed in” around this ligand. Thus, it is reasonable to hypothesize that this closed conformation may occur upon L-argenate binding to AD, and that generating a structure of AD in this conformation may yield informative docking results. The crystal structure of Δ19PD in complex with NAD<sup>+</sup> and L-Tyr (PDB ID: 3GGG) was then used as a template for homology modeling, and a homology model of AD was obtained that possessed a backbone RMSD of 0.3 Å with the template structure. Upon the first docking attempt with the homology model, a conformation satisfying the previously mentioned criteria was found. Upon closer inspection of the docked model, it was observed that both arginine residues are interacting directly with L-argenate (Figure 20). R217 is now forming an interaction with the ring carboxyl group of L-argenate, while R213 is interacting with both oxygen atoms of the side chain carboxyl group. These results suggest that the two residues play a role in binding L-argenate, and that by converting the arginine residues to alanines should result in a drop in enzyme activity. While a reduction in specific activity observed, it cannot be confirmed whether the residues are involved in L-argenate binding until values for the K<sub>m</sub> for L-argenate are determined.

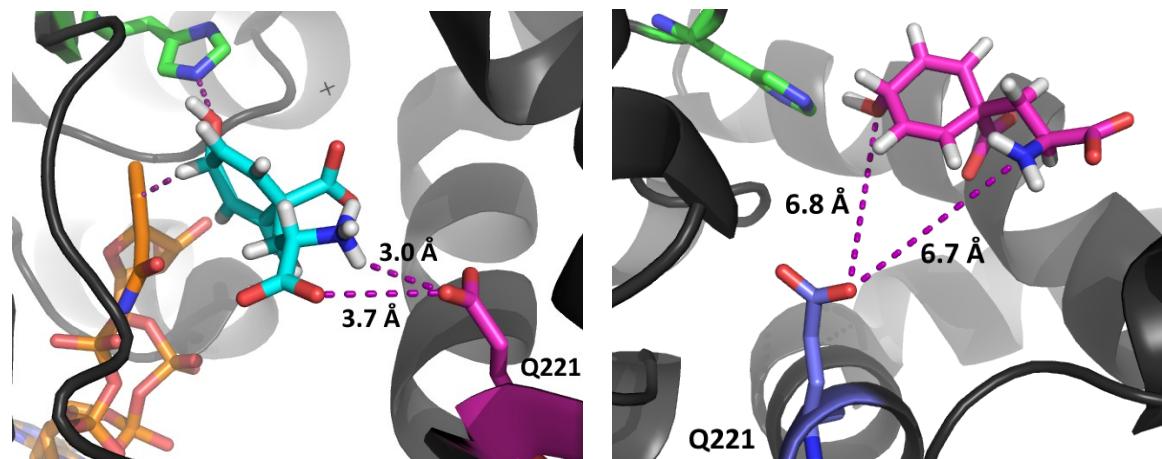
In order for R213 and R217 to interact directly with argenate, AD would have to undergo a very large tertiary structural rearrangement relative to its NADP<sup>+</sup>-bound form; indeed, the α-carbons of R213 and R217 shift 9.8 Å and 11.5 Å, respectively (Figure 21).

However, large tertiary structural rearrangements have been reported previously with enzymes that bind NADP<sup>+</sup>/NADPH. The crystal structures of the apo and NADPH-bound form of 2,5-diketro-D-gluconic acid reductase A (2,5-DKGR A) from *Cornyebacterium* were found to have markedly different tertiary structures, with regions of the active site undergoing coordinated conformational changes of up to 8 Å (77, 78). This rearrangement was hypothesized to occur so that upon NADPH binding, the enzyme would be in a catalytically competent form of 2,5-DKGR A. In the present study, near-UV circular dichroism was performed first to see if such a structural change could be detected upon titration with arogenate or NADP<sup>+</sup>, but no such change in the signal relative to that of the apo form of AD was observed (Figure 15). Additionally when the complimentary technique of fluorescence emission spectroscopy was also used, no change in the emission intensity was observed at an excitation wavelength of 295 nm upon titration with L-rogenate. Small but significant decreases that mimicked saturation of a receptor by a ligand were observed upon titration with NADP<sup>+</sup> (Figure 17). The closest tryptophan residue (3 per monomer) is located 11 Å away from the NADP<sup>+</sup> binding site in the cofactor liganded crystal structure. Consistent with NADP<sup>+</sup> quenching the fluorescence intensity upon binding and/or perturbing the environment of the tryptophan residues participating in the emission spectra, although we could not determine a K<sub>m</sub> for either substrate from the kinetic studies in this thesis, the change in fluorescence intensity yielded a K<sub>d</sub> value of 9.5 ± 3.9 μM (Figure 17) for the interaction of NADP<sup>+</sup> with the enzyme. This value is similar to the K<sub>m</sub> of NADP<sup>+</sup> determined by Legrand et al. (48) of 9 μM and is therefore in keeping with the hypothesis that AD

follows a rapid equilibrium random kinetic mechanism. In this mechanism,  $K_m$  is a good approximation of  $K_d$ .

### Glycine 221

G221 is located at the bottom of the active site, and had been previously hypothesized by Legrand and colleagues (48) for being a key residue in determining the AD's strict substrate preference for L-arogenate. They argued that G221's small chain length is more accommodating for L-arogenate's more bulky amine group than the equivalent glutamine residues found in the *E. coli* and *H. influenza* CM-PD enzymes. Additionally, this residue is a serine in a number of other PD enzymes, including that from *A. aeolicus*, whose corresponding residue is S254. In the crystal structure with



**Figure 31: Potential interactions formed by G221 after being converted to glutamic acid in both the modeled and docked structures.** G221 was converted to a glutamic acid residue using the “mutagenesis” feature in PyMOL. In the crystal structure with arogenate modeled into the active site (left), the new Q221 residue is close enough to potentially interact with L-arogenates’ side chain. However, in the docked homology model (right), Q221 is moved directly into the active site, but remains distant from L-arogenate. Figures generated using PyMOL (44).

L-arogenate modeled into the active site as well as the L-arogenate-docked homology model, the amine group of L-arogenate is the atom closest to G221, is still 9.7 Å and 11.5 Å away, respectively. Thus, there appears to be ample space for L-arogenate to enter the active site. In the L-arogenate-docked homology model of AD, when G221 is changed to glutamine, the residue is still too far away to form an interaction with L-arogenate (Figure 31). However, in the crystal structure with L-arogenate modeled into the active site, when G221 is changed to glutamine, the side chain resides 3.3 Å away from L-arogenate's amine group (Figure 31). The AD variants, G221S and G221Q were characterized to determine whether these replacements play any role in substrate specificity. Surprisingly, G221Q and G221S were observed to have a ~380-fold and ~36-fold decrease in specific activity, respectively. The reduction in activity with the G221S variant is difficult to explain because its side chain doesn't move much closer to L-arogenate in either model (Figure 31). It is possible that the introduction of residues capable of hydrogen bonding results in an interaction with the amine group; the glutamine in Figure 31 is a distance of 3.3 Å away from the amine. This interaction could hinder enzyme activity by slowing down product release, affecting catalysis. The hydroxyl side chain of serine is further away, making the formation of this interaction more difficult, which would explain why the G221S variant is 10 times more active than G221Q. Since this interaction is not observed in the arogenate docked homology model, this would also suggest that either the position of G221 in this model is incorrect, or that the enzyme returns to a similar conformation as the NADP<sup>+</sup>-bound form after the reaction has occurred, and manages to interact with tyrosine before it leaves the active site. It

would be worthwhile to investigate this further by converting G221 to a leucine; if a drop in activity remained, then this would suggest that this is due to a steric clash.

### **Threonine 117**

T117 is located on a flexible loop, and composes a portion of the active site. T117 appears to be interacting with L-aro-*genate* in both the modeled NADP<sup>+</sup>-bound crystal structure and the docked homology model (Figure 20), although with different atoms. In the crystal structure, the backbone carbonyl and side chain hydroxyl groups are close enough to the side chain carboxyl group of L-aro-*genate* to interact. In the homology model, the backbone carbonyl and amine groups of T117 are the groups close enough to the docked L-aro-*genate* to interact with its side chain amine group. While the side chain hydroxyl group of T117 also participates in NADP<sup>+</sup> binding by forming a hydrogen bond with the carbonyl group of the nicotinamide moiety. In order to investigate these potential backbone interactions, T117 was replaced by the cyclic imino acid proline to see if the imposed rigidity of the substitution on the loop region would affect the function of AD. A ~400-fold drop in specific activity (Table 5) was observed relative to wild-type AD. This suggests that increasing the rigidity of this active site loop affects enzyme activity negatively. In support of this view, when the equivalent residue in *A. aeolicus* Δ19PD (T152) was changed to a proline residue, the efficiency constant for the reaction catalyzed by this variant was reduced by a factor of 100, with equal but opposite effects on  $k_{cat}$  (decrease) and  $K_m$  (increase) for prephenate; when NAD<sup>+</sup> was the variable substrate, a 10-fold decrease in  $k_{cat}$  was observed, with a significant increase to the  $K_m$  (42). Interestingly the carbonyl group of T152 was shown to interact with the amine of L-Tyr. (Figure 7) Thus, it is possible that T117 in AD could be involved in catalysis and the

binding of both substrates, L-argenate and/or NADP<sup>+</sup>. Converting T117 to a glycine residue to increase flexibility in the loop could be worthwhile to further investigate the importance of flexibility with this loop.

### **Histidine 264, Arginine 265, and Leucine 266**

H264, R265, and L266 are all located in on  $\alpha$ -helix  $\alpha$ 11 that composes part of the C-terminal dimerization domain of AD. These residues were targeted for mutagenesis based on the controversy regarding the effect of L-tyrosine on AD activity. Bonner and colleagues (15) had reported previously that AD was sensitive to end-product inhibition, and determined a  $K_i$  of 70  $\mu$ M by kinetic inhibition studies. Legrand and colleagues (48) observed no L-tyrosine inhibition with concentrations up to 500  $\mu$ M, in agreement with what was observed in this study. The “wild-type” enzyme which Bonner and colleagues characterized was reported by Legrand and colleagues to contain four amino acid changes: H264Q, R265H, R266L, and Q268L. In an effort to try to explain this discrepancy, three of these four amino acid substitutions were made on two different variants; one variant contained a single substitution, H264Q, while the second variant contained H264Q, R265H, and R266L. The structure of the NADP<sup>+</sup>-bound AD shows that of the four residues targeted, the side chain of H264 interacts with the side chain of E224 on the adjacent monomer. While E224 is not part of the active site, it is located on helix  $\alpha$ 6 that is part of the active site and this changes the position of this helix, which could in turn affect AD’s sensitivity to L-Tyr. This hypothesis was not supported by the kinetic results; it was observed that the wild-type enzyme along with H264Q and H264Q/R265H/L266R were L-tyrosine insensitive up to a L-tyrosine concentration of 1 mM. Thus, the discrepancy in the two published reports remains a mystery. Interestingly,

although the variants were effective catalysts, the C-terminal region affected the enzyme's thermal induced unfolding properties; H264Q had a  $T_m$  value 6°C lower than the wild-type enzyme and the triple variant unfolded was considerably less cooperative.

### **Arogenate Synthesis and Purification**

Another purpose of this study was to develop a reliable method to synthesize and purify L-arogenat for use in kinetic and biophysical studies. Synthesis of L-arogenate was performed enzymatically using PAT from *Petunia hybrida*. This assay was done similarly to Rippert and colleagues (56), whom performed the assay on a smaller scale and used a crude PAT extract obtained from tobacco leaves.

The expression of the PAT was performed in the same manner as the wild-type and variant AD enzymes, although the yield of the purification of PAT, ~6 mg/3L of culture, was much lower than the AD enzymes. Expression of the PAT was previously performed in *E. coli* Rosetta<sup>TM</sup> cells (54), which are BL21 derivatives designed to enhance the expression of eukaryotic proteins that contain codons rarely used in *E. coli*; the cells contain a compatible plasmid that express tRNAs for AGG, AGA, AUA, CUA, CCC, and GGA codon. The absence of this plasmid is likely the reason for the low expression level seen in *E. coli* BL21(DE3) cells. However, for purposes of this study, 6 mg of partially purified PAT was more than sufficient. The enzyme was purified by Ni-NTA affinity chromatography. The pooled PAT fractions contained a high molecular weight contaminant, approximately 75 kDa in size. However, the PAT was functional in the presence of this contaminant, and so further purification steps were not pursued.

The conversion of prephenate to L-rogenate enzymatically by PAT could be readily predicted and followed spectrophotometrically by oxaloacetate produced by the PAT-catalyzed transamination reaction in an NADH-dependant reaction with malate dehydrogenase (Figure 23, 24). As NADH used in this coupled enzymatic synthesis interferes with the NADP-dependant reaction catalyzed by AD, we successfully developed an HPLC method using an acetonitrile gradient in phosphate buffer to potentially separate the NAD<sup>+</sup> and NADH from the other components of the reaction mixture (L-aspartate, oxaloacetate, malate, and L-rogenate). This is performed under neutral pH conditions that would note decarboxylate L-rogenate.

In efforts to follow L-rogenate production by acidification of the synthesis reaction mixture, it was observed that NAD<sup>+</sup> and NADH likely decomposed under the acidification conditions (Figure 25C). Any absorbance at 340nm due to reduced cofactor concentration was eliminated. Some of these peaks correspond to phenylalanine, as well as breakdown products of NAD<sup>+</sup> and NADH are aromatic, as both peaks were detected at 259 nm (Figure 25C). Figures 25A and 25B show that NAD<sup>+</sup> and NADH elute from analytical HPLC at 6 – 6.5 min and much later than other reaction components. Further analysis by LC-MS of the acidified reaction mixture confirmed the presence of breakdown products of the nucleotide cofactors and additionally revealed that the reaction mixture did contain L-rogenate.

LC-MS was attempted on the experimental L-rogenate sample prior to acidification; however, despite detection being observed on the LC chromatogram, detection was not observed on the MS chromatogram (data not shown). Normally, the mobile phase will contain formic acid to improve the detection and resolution observed

on the MS. When the neutral experimental arogenate sample is run on the LC-MS, neither acetonitrile or MilliQ H<sub>2</sub>O buffer contains formic acid, otherwise arogenate will decarboxylate. After acidification, this is no longer a concern, and the sample is run with each component of the mobile phase containing 0.2% formic acid. This suggests that the presence of acid in the mobile phase is critical for proper ionization to occur.

The MS chromatogram reveals four distinctive peaks, observed at: 1.58 minutes, 4.55 minutes, 5.68 minutes, and 7.15 minutes (Figure 26A). The 5.68 minute peak was found to have a major m/z peak of 166.1, which corresponds to the exact mass of phenylalanine. This confirms the presence of phenylalanine in the mixture, derived from L-rogenate, which had been present prior to acid treatment. The 7.15 minute peak was found to have a major m/z peak of 664.0, which closely corresponds to the exact mass of NAD<sup>+</sup>, which is found to be 664.1 g/mol. There are two characteristics about this peak that differ from the sample prior to acidification: the first of which is the peak is much less intense, and the retention time has increased by approximately 1 minute. The low intensity of the peak corresponding to NAD<sup>+</sup> supports the HPLC chromatogram, which indicates that the nucleotide has degraded under the acidic conditions. The change in retention time is likely due to the change in mobile phase; MilliQ H<sub>2</sub>O is used in place of phosphate buffer for LC-MS experiments. It's possible that in the presence of phosphate buffer, the negatively charged phosphate ion will be attracted to the positive charge of the NAD<sup>+</sup> molecule, and could make the molecule more hydrophilic than usual. The remaining peaks, observed at 1.58 and 4.55 minutes, were found to have major m/z peaks of 277.1 and 143.1, respectively. These signals are more difficult to explain, but appear to be degradation products of NAD<sup>+</sup>/NADH after acidification. After NAD<sup>+</sup>/NADH, the

compound with the largest molecular weight is PLP, which is 247.1 g/mol; therefore, the compound with a m/z ratio of 277.1 must have come from a heavier compound, which would have to be NAD<sup>+</sup>/NADH.

Figure 27 shows that the separation of NAD+/NADH (peaks 3 and 4) from the reaction components could be achieved also on a preparative scale. The fractions corresponding to peaks 1 and 2 (Figure 27) were collected and frozen at -80°C. The identities of these two peaks have yet to be determined. If L-rogenate, or its decarboxylation product phenylalanine after acid treatment, is detected in one of these fractions, then rogenate can be readily prepared and purified by the methodology described in this study. The next step would be to scale up the reaction mixture volume, so that large scale synthesis and purification of rogenate can occur.

## **Chapter 5: Summary and Future Work**

We have found a number of active site residues that play a role in the function of *Synechocystis sp.* AD. In order decipher whether these residues play a role in catalysis and/or substrate binding,  $K_m$  and  $k_{cat}$  parameters must be obtained for the wild-type enzyme and all variants, and remains a priority with regard to future work. Nonetheless, some key insights have been made into the nature of ADs' active site. We confirmed that H112 is the catalytic residue, and is essential for enzyme function. H170 is important for enzyme function likely for an interaction it forms with H112, in order to help maintain it in a catalytically competent conformation. This interaction is observed in the crystal structure, and supported by previous mutagenesis studies on the equivalent residue in homologous enzymes. D171 likely plays a key role in maintaining enzyme stability, by forming backbone interactions with its side chain group, as well as its backbone atoms as well. Future mutagenesis studies could be worthwhile to determine whether factors such as charge and chain length are important determinants for these interactions. S178 appears to be one of the most important residues, and likely plays a role in binding arogenate to the active site, according to modeling and docking simulations. Furthermore, it may play a role in catalysis as well, due to its potential interaction with arogenates' ring carboxyl group. The conversion of H179 only modestly decreased enzyme activity, but when it was converted in unison with S178, an additive effect in enzyme activity was observed. H179 may affect enzyme activity by affecting NADP<sup>+</sup> binding through the existence of a hydrogen bonding network involving a water molecule and S92, rather than arogenate binding. Surprisingly, R213 and R217 were both found to affect enzyme activity, with R217 having a more severe effect than R213. Homology modeling and

docking simulations suggested that this is due to the residues' roles in binding arogenate to the active site. G221's short side chain appears to be important for enzyme function, and could be necessary for arogenate to efficiently make its way into the active site. It could also potentially affect catalysis by slowing down the process of product release. M228, on its own, was not found to be important for enzyme function upon its conversion to a polar and charged residue. Further mutagenesis studies involving the conversion of additional hydrophobic residues, previously hypothesized to compose an important hydrophobic pocket, to polar residues, could be interesting to investigate this hydrophobic pocket further. T117 was also confirmed being important for enzyme function despite not being highly conserved. Converting T117 to a glycine could be insightful regarding the importance of flexibility of the loop upon which T117 is located. While H264 appears to play a role in the enzymes' thermal stability, neither the single or triple-substituted variant was found to be tyrosine sensitive.

Arogenate appears to have been synthesized by coupled enzyme assay and separated from its major contaminants, NAD<sup>+</sup> and NADH, by HPLC on both the analytical and preparative levels. LC-MS analysis detected the presence of phenylalanine in the acidified mixture, which corresponds to arogenates' presence in the mixture prior to acidification. Fractions corresponding to two peaks that would contain arogenate were collected from preparative separation and need to be analyzed by direct injection onto the MS. Provided that the presence of arogenate is determined, the next step would be to scale up the size of the enzymatic arogenate synthesis reaction, so that sufficient amounts of purified arogenate can be obtained, as kinetic analysis of AD is a top priority.

Once suitable arogenate can be obtained and kinetic parameters can be determined, further mutagenesis studies can be pursued to further decipher AD's active site. In this study, the mutagenesis strategy that was adopted involved targeting a large number of residues within the active site, rather than focusing on a few residues and extensively characterizing them by making multiple conversions. Now that a number of important residues have been successfully identified, these key residues can be converted to other residues to determine the details about the interactions and roles they play. More extensive protein engineering strategies could also be explored in order to further investigate the issue of substrate specificity. Rational approaches, such as chimera construction, could be pursued. Rational design is often high risk, but the process can be improved considerably with the aid of predictive computer algorithms; the SCHEMA algorithm (79) calculates the likelihood of a chimera protein being able to fold into its native conformation, and hence helps increase the chance of finding functional chimeras. Random mutagenesis approaches could provide considerable insight into the determinants for substrate specificity by selecting for AD mutants that can turn over prephenate as a substrate. However, this requires the availability of a high-throughput selection assay. Such an assay is currently under development by exploiting the phenotype of the knockout *E. coli* strain, KB357, which is unable to synthesize L-Tyr or L-Phe.

Similarly, the issue of tyrosine inhibition could be further explored by performing mutagenesis on the active site of the AD from *Arabidopsis thaliana*, which is reported to be highly sensitive to tyrosine inhibition. Determining which residues participate in inhibition, as well as the mechanism of inhibition, would provide insight as to similarities

and differences between the two AD enzymes, and possibly an explanation as to why AD from *Synechocystis* is tyrosine insensitive.

## References

1. Kaneko, M., Hwang, E. I., Ohnishi, Y., and Horinouchi, S. (2003) Heterologous production of flavanones in *Escherichia coli* : potential for combinatorial biosynthesis of flavonoids in bacteria. *J Ind Microbiol Biotechnol* 30, 456-461.
2. Rescigno, A., Rinaldi, A. C., and Sanjust, E. (1998) Some aspects of tyrosine secondary metabolism - biochemical and immunohistochemical studies. *Biochem Pharmacol* 56, 1089-1096.
3. Meganathan, R. (2001) Ubiquinone biosynthesis in microorganisms. *FEMS Microbiol Lett* 203, 131-139.
4. Moller, B. L., and Conn, E. E. (1979) The biosynthesis of cyanogenic glucosides in higher plants. N-Hydroxytyrosine as an intermediate in the biosynthesis of dhurrin by *Sorghum bicolor* (Linn) Moench. *J Biol Chem* 254, 8575-8583.
5. Memelink, J. (2004) Putting the opium in poppy to sleep. *Nat Biotechnol* 22, 1526-1527.
6. Frick, S., and Kutchan, T. M. (1999) Molecular cloning and functional expression of *O*-methyltransferases common to isoquinoline alkaloid and phenylpropanoid biosynthesis. *Plant J* 17, 329-339.
7. Coggins JR, Abell C, Evans LB, Frederickson M, Robinson DA, Roszak AW and Lapthorn AP, (2003) Experiences with the shikimate-pathway enzymes as targets for rational drug design. *Biochem Soc Trans* 31, 548-552.
8. Haslam, E. (1974). *The Shikimate Pathway*, Wiley, New York.
9. Cabrera-Valladares N, Martínez A, Piñero S, Lagunas-Muñoz VH, Tinoco R, de Anda R, Vázquez-Duhalt R, Bolívar F and Gosset G, (2006) Expression of the melA gene from *Rhizobium etli* CFN42 in *Escherichia coli* and characterization of the encoded tyrosines. *Enzyme and Microbial Technology* 38, 772-779.
10. Seetharam G and Saville B, (2002) L-DOPA production from tyrosinase immobilized on zeolite. *Enzyme and Microbial Technology* 31, 747-753.
11. Ikram, U.-H., and Ali, S. (2002) Microbiological transformation of L-Tyr to 3,4-dihydroxyphenyl L-alanine (L-dopa) by a mutant strain of *Aspergillus oryzae* UV-7. *Curr Microbiol* 45, 88-93.
12. Bourke SL and Kohn J, (2003) Polymers derived from the amino acid L-tyrosine: polycarbonates, polyarylates and copolymers with poly(ethylene glycol). *Adv Drug Deliv Rev* 55, 447-66.
13. Davis, B. D. (1955) Intermediates in amino acid biosynthesis. *Adv Enzymol Relat Subj Biochem* 16, 247-312.
14. Sprinson, B. D. (1961) The Biosynthesis of Aromatic Compounds from D-Glucose. *Adv Carbohydr Chem* 15, 235-270.
15. Bonner, C. A., Jensen, R. A., Gander, J. E., and Keyhani, N. O. (2004) A core catalytic domain of the TyrA protein family: arogenate dehydrogenase from *Synechocystis*. *Biochem J* 382, 279-91.
16. Song, J., Bonner, C. A., Wolinsky, M., and Jensen, R. A. (2005) The TyrA family of aromatic-pathway dehydrogenases in phylogenetic context. *BMC Biol* 3, 13.
17. Sampathkumar P and Morrison JF, (1982) Chorismate mutase-prephenate dehydrogenase from *Escherichia coli*. Purification and properties of the bifunctional enzyme. *Biochim Biophys Acta* 702, 204-11.

18. Christopherson, R. I., and Morrison, J. F. (1985) Chorismate mutase-prephenate dehydrogenase from *Escherichia coli*: positive cooperativity with substrates and inhibitors. *Biochemistry* 24, 1116-21.
19. Turnbull J, Cleland WW and Morrison JF, (1990) Chorismate mutase-prephenate dehydrogenase from *Escherichia coli*. 1. Kinetic characterization of the dehydrogenase reaction by use of alternative substrates. *Biochemistry* 29, 10245-54.
20. Hassounah S. (2009) MSc. thesis, Department of Chemistry and Biochemistry, Concordia University, Montreal, Canada.
21. Lim, S., Springstead, J. R., Yu, M., Bartkowski, W., Shroder, I. and Monbouquette, H. G. (2009) Characterisation of a key trifunctional enzyme for aromatic amino acid biosynthesis. *Extremophiles*. 13,191-98.
22. Bonvin, J., Aponte, R. A., Marcantonio, M., Singh, S., Christendat, D., and Turnbull, J. L. (2006) Biochemical characterization of prephenate dehydrogenase from the hyperthermophilic bacterium *Aquifex aeolicus*. *Protein Sci* 15, 1417-32.
23. Xie, G., Bonner, C. A., and Jensen, R. A. (2000) Cyclohexadienyl dehydrogenase from *Pseudomonas stutzeri* exemplifies a widespread type of tyrosine-pathway dehydrogenase in the TyrA protein family. *Comp Biochem Physiol C Toxicol Pharmacol* 125, 65-83.
24. Gaines, C. G., Byng, G. S., Whitaker, R. J. and Jensen, R. A. (1982) L-Tyrosine regulation and biosynthesis via arogenate dehydrogenase in suspension-cultured cells of *Nicotiana silvestris* Speg. et Comes. *Planta* 156, 233-40.
25. Grinter, N. J. (1998) Developing an L-phenylalanine process. *ChemTech* 28, 33-37.
26. Zhao, G., Xia, T., Ingram, L. O., and Jensen, R. A. (1993) An allosterically insensitive class of cyclohexadienyl dehydrogenase from *Zymomonas mobilis*. *Eur J Biochem* 212, 157-65.
27. Bonner, C. A., Disz, T., Hwang, K., Song, J., Vonstein, V., Overbeek, R., and Jensen, R. A. (2008) Cohesion group approach for evolutionary analysis of TyrA, a protein family with wide-ranging substrate specificities. *Microbiol Mol Biol Rev* 72, 13-53.
28. Turnbull J, Morrison JF and Cleland WW, (1991) Kinetic studies on chorismate mutase-prephenate dehydrogenase from *Escherichia coli*: models for the feedback inhibition of prephenate dehydrogenase by L-tyrosine. *Biochemistry* 30, 7783-88
29. Turnbull, J., Cleland, W. W., and Morrison, J. F. (1991) pH dependency of the reactions catalyzed by chorismate mutase-prephenate dehydrogenase from *Escherichia coli*. *Biochemistry* 30, 7777-82.
30. Christendat, D., and Turnbull, J. L. (1996) Identification of active site residues of chorismate mutase-prephenate dehydrogenase from *Escherichia coli*. *Biochemistry* 35, 4468-79.
31. Sun, W., Shahinas, D., Bonvin, J., Hou, W., Kimber, M.S., Turnbull, J. and Christendat, D. (2009) The crystal structure of *Aquifex aeolicus* prephenate dehydrogenase reveals the mode of tyrosine inhibition. *J Biol Chem* 284(19), 13223-32.

32. Hyung-Keun Ku, H.K., Park, S.R., Yang, I. and Kim, S.K. (2010) Expression and functional characterization of prephenate dehydrogenase from *Streptococcus mutans*. *Process Biochem* 45, 607–12.
33. Xu, S., Yang, Y., Jin, R., Zhang, M. and Wang, H. (2006) Purification and characterization of a functionally active *Mycobacterium tuberculosis* prephenate dehydrogenase. *Protein Expr Purif* 49, 151-8.
34. Hudson, G. S., and Davidson, B. E. (1984) Nucleotide sequence and transcription of the phenylalanine and tyrosine operons of *Escherichia coli* K12. *J Mol Biol* 180, 1023-51.
35. Turnbull, J., and Morrison, J. F. (1990) Chorismate mutase-prephenate dehydrogenase from *Escherichia coli*. 2. Evidence for two different active sites. *Biochemistry* 29, 10255-61.
36. Christendat, D., Saridakis, V. C., and Turnbull, J. L. (1998) Use of site-directed mutagenesis to identify residues specific for each reaction catalyzed by chorismate mutase-prephenate dehydrogenase from *Escherichia coli*. *Biochemistry* 37, 15703-12.
37. Christendat, D., and Turnbull, J. L. (1999) Identifying groups involved in the binding of prephenate to prephenate dehydrogenase from *Escherichia coli*. *Biochemistry* 38, 4782-93.
38. Hermes, J. D., Tipton, P. A., Fisher, M. A., O'Leary, M. H., Morrison, J. F., and Cleland, W. W. (1984) Mechanisms of enzymatic and acid-catalyzed decarboxylations of prephenate. *Biochemistry* 23, 6263-75.
39. Deckert, G., Warren, P. V., Gaasterland, T., Young, W. G., Lenox, A. L., Graham, D. E., Overbeek, R., Snead, M. A., Keller, M., Aujay, M., Huber, R., Feldman, R. A., Short, J. M., Olsen, G. J., and Swanson, R. V. (1998) The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature* 392, 353-8.
40. Sun, W., Singh, S., Zhang, R., Turnbull, J. L., and Christendat, D. (2006) Crystal structure of prephenate dehydrogenase from *Aquifex aeolicus*. Insights into the catalytic mechanism. *J Biol Chem* 281, 12919-28.
41. Hudson, G. S., Howlett, G. J., and Davidson, B. E. (1983) The binding of tyrosine and NAD<sup>+</sup> to chorismate mutase/prephenate dehydrogenase from *Escherichia coli* K12 and the effects of these ligands on the activity and self-association of the enzyme. Analysis in terms of a model. *J Biol Chem* 258, 3114-20.
42. Hotz N. (2010) MSc. thesis, Department of Chemistry and Biochemistry, Concordia University, Montreal, Canada.
43. Hou, W. (2009) MSc. thesis. Department of Chemistry and Biochemistry, Concordia University, Montreal, Canada.
44. DeLano, W. L. (2002) in *The PyMOL Molecular Graphics System*, San Carlos, CA.
45. Christopherson R.I. (1985) Chorismate mutase-prephenate dehydrogenase from *Escherichia coli*: cooperative effects and inhibition by L-tyrosine. *Arch Biochem Biophys* 240(2), 646-654.
46. Lütke-Eversloh, T and Stephanopoulos, G, (2005) Feedback inhibition of chorismate mutase/prephenate dehydrogenase (TyrA) of *Escherichia coli*:

- generation and characterization of tyrosine-insensitive mutants *Appl Environ Microbiol* 71, 7224-7228.
47. Chiu H., Abdubek P., Astakhova T., Axelrod H. L., Carlton D., Clayton T., DasD. Deller M.C., Duan L., Feuerhelm J., Grant J. C., Grzechnik A., Han G. W., Jaroszewski L., Jin K. K., Klock H. E., Knuth M. W., Kozbial P., Krishna S. S., Kumar A., Marciano D., McMullan D., Miller M. D., Morse A.T., Nigoghossian E., Okach L., Reyes R., Tien H. J., Trame C. B., van den Bedem H., Weekes D., Xu Q., Hodgson K. O., Wooley J., Elsliger M., Deacon A. M., Godzik A., Lesley S.A., Wilsonb I. A. (2010) The structure of *Haemophilus influenza* prephenate dehydrogenase suggests unique features of bifunctional TyrA enzymes *Acta Crystallogr F F66*: 1317–25.
48. Legrand, P., Dumas, R., Seux, M., Rippert, P., Ravelli, R., Ferrer, J. L., and Matringe, M. (2006) Biochemical characterization and crystal structure of *Synechocystis* arogenate dehydrogenase provide insights into catalytic reaction. *Structure* 14, 767-76.
49. Quashie, P. K. (2010) MSc. Thesis, Department of Chemistry and Biochemistry, Concordia University, Montreal, Canada.
50. Jensen, R. A., Zamir, L., St. Pierre, M., Patel, N., Pierson, D.L. (1977) Isolation and Preparation of Pretyrosine, Accumulated as a Dead-End Metabolite by *Neurospora crassa*. *Journal of Bacteriology* 132, 896-903.
51. Zamir, L., Jensen, R.A., Arison, B.H., Douglas, A.W., Albers-Schoenberg, G., Bowen, J. R. Structure of arogenate (pretyrosine), an amino acid intermediate of aromatic biosynthesis. (1980) *JACS*, 13, 4499-4504.
52. Bonner, C.A., Fischer, R.S., Ahmad, S., Jensen, R.A. Remnants of an Ancient Pathway to L-Phenylalanine and L-Tyrosine in Enteric Bacteria: Evolutionary Implications and Biotechnological Impact. (1990) *Applied and Environmental Microbiology* 56, 3741-3747.
53. Zamir, L., Tiberio, R., Fiske, M., Berry, A., and R.A. Jensen. (1985) *Biochemistry* 24, 1607.
54. Maeda, H., Yoo, H., Dudareva, N. (2011) Prephenate aminotransferase directs plant phenylalanine biosynthesis via arogenate. *Nature Chemical Biology* 7, 19-21.
55. Danishefsky, S., Morris, J., Clizbe, L.A. (1981) Total Synthesis of Pretyrosine (Arogenate). *J. Am. Chem. Soc* 103, 1602-1604.
56. Rippert, P., Matringe, M. (2002) Molecular and biochemical characterization of an *Arabidopsis thaliana* arogenate dehydrogenase with two highly similar and active protein domains. *Plant Molecular Biology* 48, 361-368.
57. Stenmark, S., Pierson, D.L., Jensen, R.A. (1974) Blue-Green Bacteria synthesize L-Tyrosine by the Pretyrosine Pathway. *Nature*, 247, 290-293.
58. Bonner, C., Jensen, R.A. (1987) Arogenate Dehydrogenase. *Methods in Enzymology* 42, 488-494.
59. Arnold, F.H. and A.A. Volkov, *Directed evolution of biocatalysts*. *Curr Opin Chem Biol*, 1999. 3(1): p. 54-9.
60. Rieger, C. E., and Turnbull, J. L. (1996) Small scale biosynthesis and purification of gram quantities of chorismic acid. *Prep Biochem Biotechnol* 26, 67-76.

61. Dudzinski, P. K., Morrison, J. F. (1976) The Preparation and Purification of Sodium Prephenate. *Prep Biochem* 6, 113-21.
62. Dawson, C. M. R., Elliott, C. D., Elliot, H. W. and Jones, M. K. (1986) *Data for Biochemical Research*, 3rd edition ed., Oxford Science Publications, Clarendon Press.
63. Turnbull, J. L. (1988), PhD thesis, Australian National University, Canberra, Australia.
64. Rozen, S., Skaletsky, H.J. (2000) Primer3 on the WWW for general users and for biologist programmers, in *Bioinformatics Methods and Protocols: Methods in Molecular Biology* pp 365-386, Humana Press, Totowa, NJ.
65. Sambrook, J., Russell, D. (2001) *Molecular Cloning*, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring, New York.
66. Pace, N., Vajdos, F., Fee, L., Grimsley, G., Gray, T. (1995) How to measure and predict the molar absorbtion coefficient of a protein 4, 2411-2423.
67. Besir, H. (2008) Coomassie ohne Alkohol und Essig. *Laborjournal online*, <http://www.laborjournal.de/rubric/tricks/tricks/trick121.lasso>
68. Lamoureux, G. (2010) CHEM 498Q / CHEM 630Q: Molecular Modeling of Proteins TUTORIAL #5: Molecular Docking, Department of Chemistry and Biochemistry, Concordia University, Montréal, Canada
69. Biellmann, J.F., Lapinte, C., Haid, E., Weimann, G. (1979) Structure of Lactate Dehydrogenase Inhibitor Generated from Coenzyme. *Biochemistry* 79, 1212-1217.
70. Watson, D.H., Harvey, M.J., Dean, P.D.G. (1978) The Selective Retardation of NADP+-Dependent Dehydrogenases by Immobilized Procion Red HE-3B. *Biochem J.* 173, 591-596.
71. Schmid, F. X. (1998) *Spectral methods of characterizing protein conformation and conformational changes*, TE Creighton edition IRL press, Oxford, UK.
72. Kelly, S.M., Jess, T.J., Price, N.C. (2005) How to study proteins by circular dichroism *Biochimica et Biophysica Acta* 1751, 119-139.
73. Bonvin, J. (2008) PhD thesis, Department of Chemistry and Biochemistry, Concordia University, Montreal, Canada.
74. Sali, A. and Shen, M-y. Statistical potential for assessment and prediction of protein structures. *Protein Science* 15, 2507-2524, 2006.
75. Trott, O. and Olson, A.J. (2010) Auto-Dock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J Comp Chem.* 31, 455-61.
76. Prakash, P., Aruna, B., Sardesai, A.A., Hasnain, S.E. (2005) Purified Recombinant Hypothetical Protein Coded by Open Reading Frame Rv1885c of *Mycobacterium tuberculosis* Exhibits a Monofunctional AroQ Class of Preplasmic Chorismate Mutase Activity. *The Journal of Biological Chemistry*, 280, 19641-19648.
77. Khurana S, Powers DB, Anderson S, Blaber M. (1998) Crystal structure of 2,5-diketo-D-gluconic acid reductase A complexed with NADPH at 2.1-A resolution. *Proc Natl Acad Sci USA*, 12, 6768-73.
78. Sanli, G., Blaber, M. (2001) Structural assembly of the active site in an aldo-keto reductase by NADPH cofactor. *J Mol Biol.*, 5, 1209-18.

79. Voight, C.A., Martinez, C., Wang, Z.G., Mayo, S.L., Arnold, F.H. (2002) Protein building blocks preserved by recombination. *Nature structural biology*, 9, 553-558.
80. Manioudakis, J. (2004) Protocol for the Preparation of Sodium Prephenate from Chorismic Acid via the catalytic action of CM-PD. Department of Chemistry and Biochemistry, Concordia University, Montreal, Canada.
81. Stewart, J., Wilson, D.B., Ganem B. (1990) A Genetically Engineered Monofunctional Chorismate Mutase. *J. Am. Chem. Soc.*, 112, 4582-4584
82. Lee, A.Y., Stewart, J.D., Clardy, J., Ganem, B. (1995) New insight into the catalytic mechanism of chorismate mutases from structural studies *Current Biology*, 2, 195-203
83. Turnbull, J., Morrison, J.F., Cleland, W.W. (1991) Kinetic Studies on Chorismate Mutase-Prephenate Dehydrogenase from *Escherichia coli*: Models for the Feedback Inhibition of Prephenate Dehydrogenase by L-Tyrosine. *Biochemistry*, 30, 7783-7788
84. Manioudakis, J. (2009) PhD thesis. Concordia University, Montréal.
85. Bradford, M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding *Analytical Biochemistry* 72, 248-54.
86. NCCLS Approved Standard ACS-1 (1979) *Specification for Standardized Protein Solution (Bovine Serum Albumin)*, 2nd ed., National Committee for Clinical Laboratory Standards: Villanova, PA.