Purification of flavan-3-ol biosynthesis for co-localization onto nanocarriers and a multi-enzyme assay

Emily Kaitlin Davenport
Iowa State University
Purification of flavan-3-ol biosynthesis for co-localization onto nanocarriers and a multi-enzyme assay

by

Emily Kaitlin Davenport

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Chemical Engineering

Program of Study Committee:
Surya Mallapragada, Co-Major Professor
Reuben Peters, Co-Major Professor
Balaji Narasimhan

Iowa State University of Science and Technology
Ames, Iowa
2011

Copyright © Emily Kaitlin Davenport, 2011. All rights reserved.
Table of Contents

List of Figures.................................................................................................................. iv
List of Tables ...................................................................................................................... vi
Abstract............................................................................................................................... vii

Chapter 1: General Introduction ......................................................................................... 1
  Overview................................................................................................................................. 1
  Flavonoid Pathway................................................................................................................. 2
  Flavonoids.............................................................................................................................. 5
    Structure............................................................................................................................... 5
    Leucoanthocyanidins .......................................................................................................... 6
    Anthocyanidins ................................................................................................................... 6
    Flavan-3-ols ...................................................................................................................... 7
    Proanthocyanidins.............................................................................................................. 9
  DFR........................................................................................................................................ 10
  ANS....................................................................................................................................... 11
  ANR....................................................................................................................................... 13
  Enzyme Attachment.......................................................................................................... 14
Hypothesis............................................................................................................................ 16

Chapter 2: Purification of *Arabidopsis* DFR, ANS, and ANR for site-specific
attachment onto a chemically modified nanocarrier......................................................... 17
  Abstract................................................................................................................................. 17
  Introduction........................................................................................................................... 17
Methods ..................................................................................................................21

Materials .................................................................................................................21

DNA Manipulations ...............................................................................................21

Expression ...............................................................................................................23

Purification of AtANR ..............................................................................................24

Purification of AtDFR ..............................................................................................25

Purification of AtANS ..............................................................................................25

Activity Assays .......................................................................................................26

Thrombin Cleavage .................................................................................................28

Results .....................................................................................................................28

Discussion ...............................................................................................................34

Chapter 3: General Conclusions ............................................................................39

References .............................................................................................................40

Acknowledgements .................................................................................................47
List of Figures

Figure 1. Schematic of the flavonoid pathway. CHS: Chalcone Synthase, CHI: Chalcone Isomerase, F3(’)/H: Flavonol 3’-Hydroxylase, FLS: Flavonol Synthase, DFR: Dihydroflavonol 4-Reductase, ANS: Anthocyanidin Synthase, UFGT: UDP-glucose: flavonoid 3-O-glucosyltransferase, ANR: Anthocyanidin Reductase, LAR: Leucoanthocyanidin Reductase. Grayed ANS reactions are side products uninvolved in final product synthesis. LAR (italicized with dashed line) is not present in Arabidopsis thaliana.

Figure 2. General structures for each class of flavonoid

Figure 3. Both A-type and B-type (-)-epicatechin dimers

Figure 4. Crystal structure of ANS from Arabidopsis thaliana, complexed with narigenin (Turnbull, Prescott et al. 2001)

Figure 5. Different reactions catalyzed by ANS. Note that substrate stereochemistry may vary product ratios

Figure 6. Flavan-3-ol monomers. Substrates are italicized, enzymes are denoted in bold

Figure 7. Late flavonoid pathway in Arabidopsis thaliana. Grayed arrows are noted in vitro reactions

Figure 8. PCR product of ANS. Lane 1) 500 bp ladder, lane 2) PCR product to be purified from ANS/pDONR, lane 3) positive ANS control from original pUC57 vector

Figure 9. Left - DFR assay extracts after acid hydrolysis to convert leucocyanidin to cyanidin (pink). Right – control (clear)
Figure 10. Purification of DFR/pDest15 - (starting from left) lane 1) lysate, lanes 2-4) resin washes, lane 5) eluate, lane 6) empty, lane 7) standard ladder

Figure 11. Lysate kinetics for AtDFR as an untagged enzyme (pDest14) and with a GST tag (pDest15). Absorbance at 550 nm corresponds to the amount of cyanidin resulting from the acid hydrolysis of leucocyanidin produced by the enzyme.

Figure 12. FPLC chromatogram of purifying AtANR; purified protein peak between 3.5-4 mL.

Figure 13. Averaged ANR assay results showing increasing absorbance at 280 nm, the maximum absorbance wavelength for epicatechin. Data was taken from two different purification batches.

Figure 14. TLC with ANR assay extractions. Lane 1) ANR with no substrate, lane 2) ANR with cyanidin substrate, lane 3) epicatechin standard, lane 4) DHQ standard, lane 5) quercetin standard. Rf values for epicatechin in assay (0.643) and standard (0.610).

Figure 15. TLC of ANS assay extracts. Arrow is pointing to product (quercetin) band. Lane 1) ANS assay, lane 2) epicatechin standard, lane 3) DHQ standard, lane 4) quercetin standard. Rf values: for both quercetin generated in assay and the quercetin standard at 0.926.

Figure 16. Shows the predicted thrombin cleavage site between the GST tag and the ANS enzyme with glycine-serine linker and N-terminal DM3. The added 13 amino acids after cleavage are bolded and underlined.
List of Tables

Table 1. List of primers used for cloning PCR, all from 5’ to 3’................................. 22

Table 2. Specifications for each enzyme and its respective substrate in the collective assay buffer ........................................................................................................... 27
Abstract

Flavonoids are polyphenolic secondary metabolites that serve a variety of purposes, from assisting pollination in plants to providing anti-cancerous activity in humans. It is the wide range of functions that give these molecules the highly revered status they have today. Range in function can be linked with the range in structure, and with over 10,000 different flavonoids known, they make up one of the largest groups of natural products. One common sub-class of flavonoids is the flavan-3-ols, which are known for their health benefits and tendency to condense to form proanthocyanidins, or condensed tannins. Both of these classes of flavonoids represent the valuable end products of a highly investigated, yet still not fully understood, section of flavonoid biosynthesis. Furthermore, some of the intermediates are highly unstable and have never been isolated in planta. The enzymes that comprise this section are highly interconnected and regulated, suggesting that a possible multi-enzyme complex exists. With this evidence, we propose to create a synthetic complex using a polymer nanocarrier using anthocyanidin synthase and reductase (ANS and ANR, respectively), the final two enzymes in flavan-3-ol biosynthesis. Along with the preceding enzyme, dihydroflavonol 4-reductase (DFR), they have been separately and recombinantly expressed in E. coli. Both ANR and ANS are fused to a specific tag for binding to the nanocarrier, ANR – 6xHis and ANS – (monomeric) streptavidin. Specifying binding sites will allow adjustable enzyme concentration and distribution on the nanocarrier. DFR, which has a stable substrate, allows for generation of the highly unstable ANS substrate. All three enzymes have been designed and purified for co-localization on the nanocarrier and all have exhibited activity in an assay designed for the multi-enzyme reaction.
Chapter 1: General Introduction

Overview

One of the most investigated metabolic pathways in plants today is the flavonoid metabolic pathway. This pathway produces a variety of flavonoid metabolites, including proanthocyanidins and the well-know group of antioxidants: flavan-3-ols. These plant natural products have been found to take many forms and can exist as single monomers or form compounds as large as 20,000 (Haslam 1998). The amount of interest in this pathway began to increase in the 1960’s when scientists began exploring the biological function of these molecules, as little was known about the importance of their role and presence.

Over the years, the actions of flavonoids and their derivatives have been shown to be widespread. Generally, flavonoids and flavonoid derivatives have been credited with anti-viral, anti-cancer, and other preventative health benefits in animals and humans. They are also involved in assisting pollination, UV defense, and anti-fungal activity in plants (Harborne and Williams 2000; Beecher 2004).

Flavonoid applications, health benefits, separation, and natural occurrence have been reviewed (Gu, Kelm et al. 2003; Aron and Kennedy 2008). Additionally, many researchers have investigated and elucidated steps in the upstream flavonoid pathway (Burbulis and Winkel-Shirley 1999; Pelletier, Burbulis et al. 1999). Knowledge about the
later steps has been slowly accumulating, but full understanding the interactions of these enzymes and the final steps of proanthocyanidin synthesis are still lacking.

**Flavonoid Pathway**

The flavonoid pathway stems from the phenylpropanoid pathway. Shown in Figure 1, the products of this pathway are flavonols, anthocyanins, flavan-3-ols, and proanthocyanidins.

Accumulation of flavonoids and the enzymes that produce them first occurs in the seed coat when it is first exposed to UV light (Pelletier, Murrell et al. 1997) and can suddenly increase due to external stress at any time during the life of the plant (Winkel-Shirley 2002). Mutations resulting in clear to yellowish phenotypes, called *transparent testa* (*tt*), are missing essential enzymes within this pathway and have produced a number of disruptions in the production of pigments in the host plants (Koornneef 1990; Shirley, Kubasek et al. 1995). These mutants have given clues to identify the genes and to allow the mapping of their loci. Knockout studies, totaling to 24 mutants, have been done to analyze the anthocyanidin presence in seedlings in the various stages of plant growth and to characterize the pathway (Pelletier, Burbulis et al. 1999; Abrahams, Tanner et al. 2002; Abrahams, Lee et al. 2003).

While there has been a lot of research done on the association and characterization of the early steps of the flavonoid pathway, the late steps still lack elucidation.

Because of the interactive nature of the enzymes in this pathway and the knowledge that only a small amount of the enzymes present in the cell are free and soluble (Zalokar
speculation began of a flavonoid multi-enzyme complex (Burbulis and Winkel-Shirley 1999).

Figure 1. Schematic of the flavonoid pathway. CHS: Chalcone Synthase, CHI: Chalcone Isomerase, F3(‘)H: Flavonol 3(‘)-Hydroxylase, FLS: Flavonol Synthase, DFR: Dihydroflavonol 4-Reductase, ANS: Anthocyanidin Synthase, UFGT: UDP-glucose: flavonoid 3-O-glucosyltransferase, ANR: Anthocyanidin Reductase, LAR: Leucoanthocyanidin Reductase. Grayed ANS reactions are side products uninvolved in final product synthesis. LAR (italicized with dashed line) is not present in *A. thaliana*.

Multi-enzyme complexes, or metabolons, are physical macromolecular assemblies of enzymes inside the cell that usually co-exist in the same regulatory pathway. These
enzymes are co-localized often due to limitations in cooperative free-enzyme efficiency and product output. Together, in these metabolons, they tend to work with better efficiency and with more intra-enzyme cooperation (Winkel 2004). They have been mostly investigated in primary metabolism, but have been speculated and explored in secondary metabolism as well. Many pieces of evidence for a complex are present in the flavonoid pathway.

Many of the intermediates involved are highly unstable and possibly toxic. Rapid transport must occur to accumulate substantial amounts of the final product while intermediate levels remain very low (Mathews 1993; Ovadi and Srere 1996). Early experiments investigating the idea of a complex show significant evidence that the flavonoid pathway may have weak interactions with the ER membrane (Hrazdina and Wagner 1985). There is also direct evidence showing that the first two enzymes of the flavonoid pathway are co-localized within the cell (Saslowsky and Winkel-Shirley 2001; Saslowsky, Warek et al. 2005). There is some debate on the extent of the complex or whether it is more likely in a linear or a globular arrangement (Winkel-Shirley 1999).

Careful regulation of this pathway and its intermediates must also occur. Some products, such as quercetin, will inhibit activity in dihydroflavonol 4-reductase (DFR) and anthocyanidin reductase (ANR). These are two of the essential enzymes in proanthocyanidins and flavan-3-ol synthesis and can be seen in Figure 1. In addition, for polymerization of proanthocyanidins (PAs), the extension unit, a leucoanthocyanidin, must be more abundant than the initiator unit, a flavan-3-ol (Abrahams, Lee et al. 2003).
Many researchers have chosen to use *Arabidopsis thaliana* as a model organism to study flavan-3-ol and PA biosynthesis. Not only does *A. thaliana* have a fully sequenced genome, there are also two unique and simplistic attributes to the *Arabidopsis* model. First, it does not seem to encode an equivalent to leucoanthocyanidin reductase (LAR), the enzyme that catalyzes the formation of (+)-catechin. Therefore, the PA makeup and stereochemistry is restricted to that of (-)-epicatechin where the average degree of polymerization in *Arabidopsis* is 8 units (Abrahams, Lee et al. 2003). Second, all enzymes are encoded by only one gene, with the exception of FLS (Winkel-Shirley 2001). In *A. thaliana*, FLS has been found to have six gene copies, but only one of which is responsible for activity in the wild type (Pelletier, Murrell et al. 1997). This is very beneficial to genetic and metabolic engineering investigations.

**Flavonoids**

**Structure**

Flavonoids are polyphenolic compounds that are derived from phenylalanine and acetyl-CoA via the phenylpropanoid pathway. They can be categorized into 9 sub-classes (see Figure 2) and comprise one of the largest banks of natural
products, with over 10,000 naturally occurring (Tahara 2007). The abundance of these compounds varies widely depending on their source.

Most of the flavonoid synthesis occurs in the cytosol and its products are later transported to the vacuole, where many of water-soluble compounds are stored (Abrahams, Lee et al. 2003). The late steps of the flavonoid pathway contain many unstable compounds that are converted to compounds with higher function.

**Leucoanthocyanidins**

Leucoanthocyanidins are 3,4-diols that are highly reactive intermediates in the flavonoid pathway, shown in Figure 2 as the Flavanol backbone. They have never been isolated in nature (Forkmann and Martens 2001). This is possibly because of their tendency to polymerize even under the slightest perturbation in conditions or because of the highly efficient exchange between adjacent enzymes in the speculated flavonoid enzyme complex.

It has also been suggested that leucoanthocyanidins may convert to their previous form, dihydroflavonols, under storage conditions (Pang, Peel et al. 2007). Therefore, generation of these products requires immediate acidification or lyophilization. Additionally, leucoanthocyanidins are the extension units for building proanthocyanidins.

**Anthocyanidins**

Anthocyanidins (Figure 2) are the precursors to flavan-3-ol units, and are also the unglycosylated version of anthocyanins, their more prevalent and stable counterpart. Anthocyanidins possess a pH dependant color that varies from orange-red to blue.
Anthocyanidins exist as a flavylium cation, but very little of them are found in nature. They exist mainly as salts or anthocyanins, and what little are found are considered degradation products of anthocyanins (He, Mu et al. 2010).

Anthocyanidin is a substrate for two enzymes, marking the committed step in the downstream flavonoid pathway. Anthocyanidin reductase (ANR) will reduce anthocyanidin to a flavan-3-ol, which may condense to form proanthocyanidins. UDPglucose:flavonoid glucosyltransferase (UGFT) catalyze a glycosylation reaction to form anthocyanin.

Anthocyanins are best known for their pigmentation seen in plants. These compounds also aid in the recruitment of pollinators and protect plants against UV light (Harborne and Williams 2000). Furthermore, (Desai, Olsen et al. 2010) have shown that administering anthocyanin-rich blackberry extracts to rats will halt tumor growth and may even shrink tumor size.

**Flavan-3-ols**

Nine different variations on flavan-3-ols have been discovered. The basic structure is shown in Figure 2 as a Flavanol. The variations these variations depend on the placement of hydroxyl groups on the rings. In some plant species, these different molecules can be substituted for each other as a substrate; however, there are discrepancies in kinetic rates respective to the substrate and the plant species.

Flavan-3-ols are the most commonly consumed flavonoid in our diet (Scalbert and Williamson 2000). They are present my many foods such as tea, fruit, legumes,
vegetables, and cocoa. Foods rich in flavan-3-ols (monomeric or condensed) have been called “superfoods” and are highly encouraged in the diet for their anti-oxidant properties. Since then, a broader range of foods have been examined to determine flavonoid quantities. In one study, 56 foods were analyzed (de Pascual-Teresa, Santos-Buelga et al. 2000). Overall, (-)-epicatechin was shown to be the most abundant flavanol.

Extracts from plants rich in flavonoids have shown to have cancer-preventative and anti-cancerous properties (Ahmad, Cheng et al. 2000; Grotewold, Lin et al. 2006). These compounds have been shown to induce apoptosis in cancer cells (Cos, De Bruyne et al. 2004) thereby reducing tumor formation while also reducing oxidative stress (Bagchi, Bagchi et al. 2000).

These monomers are also capable of chelating transition metals, such as iron or lead, leading to an increased resistance to external concentrations. The presence of flavan-3-ol monomers (-)-epicatechin and (-)-catechin have been shown to slow or extinguish any invasive neighboring plant growth (Ahmad, Gupta et al. 2000; Bais, Vepachedu et al. 2003). Furthermore, this tactic can also be used to defend against invasive microorganisms (Scalbert, 1991 or 1999).

Flavan-3-ols have been shown to have anti-viral/microbial activity (Chung, Wong et al. 1998) and may also be able to suppress Staphylococcus aureus resistance to antibiotics (Vinson, Mandarano et al. 2002). They can also reduce rate of adhesion of the same infection in human epithelial cells (Janecki and Kolodziej 2010).
Flavan-3-ol monomers may condense to form dimers, oligomers, and polymers (Marles, Ray et al. 2003), these polymers are often referred to as proanthocyanidins.

**Proanthocyanidins**

Flavan-3-ol condensation products, called proanthocyanidins or condensed tannins (CTs), are widely known due to their presence in grapes and wine. In 1973 CTs were speculated to cause an astringency to render the host unpalatable to microbes and animals due to a precipitation of proteins (Bate-Smith 1973). In fact, CTs react with the proline-rich proteins in the saliva and these reactions results in an astringency in taste (Kallthrika, Bakker et al. 2001).

There are three types of proanthocyanidins. A-type proanthocyanidins carry two bonds between flavan-3-ol units. These are more structurally rigid than B-type proanthocyanidins. Examples of both types are shown in Figure 3. These dimers have one linkage that is either C4-C8 or C4-C6. The C-type contains at least three units: one initiating and two or more extension units (Hemingway, Karchesy et al. 1989).

Tannins have been shown to have anti-microbial and anti-fungal effects in plants (Beecher 2004), among

![A-type (-)-epicatechin dimer](image)

![B-type (+)-epicatechin dimer](image)

Figure 3. Both A-type and B-type (-)-epicatechin dimers
other protective abilities, but they also provide positive benefits for both animals and humans. Tannins can decrease the frequency of pasture bloat in grazing animals by binding reversibly to related proteins, thus slowing their degradation (L R McMahon 2000; Dixon and Sumner 2003). Despite their ability to form complexes with proteins and to exhibit anti-nutritive behavior, CTs are best known for their health benefits. These benefits include antioxidant action and cardio- and neuro-protective properties (Aron and Kennedy 2008).

Bioavailability of proanthocyanidins has been further explored recently with studies in vivo and in vitro. The interactions and digestion of both proanthocyanidins and their monomers is not yet fully understood. A study involving human dietary intake of these components determined that PAs were the least absorbed class of flavonoids through the digestive tract (Tsang, Auger et al. 2005). Small flavanol oligomers are also thought to be broken down into monomeric form and are much easier to metabolize than the larger peptides (Scalbert, Deprez et al. 2000; Deprez, Mila et al. 2001; Manach and Donovan 2004). It is thought that even flavan-3-ol monomers must be degraded into low molecular-weight metabolites before they can be absorbed by the body (Scalbert 1991; Rios, Gonthier et al. 2003).

DFR
Dihydroflavonol reductase, or DFR, is one of the last steps in flavan-3-ol biosynthesis and marks the beginning of a series of enzymatic reactions with highly unstable intermediates. It catalyzes the C-4 reduction of a dihydroflavonol to leucoanthocyanidin.
DFR is classified as a short-chain dehydrogenase/reductase and requires NADPH or NADH as a cofactor (Petit, Granier et al. 2007). It catalyzes a stereo-specific reaction to form a 3,4-cis diol from a dihydroflavonol (Stafford and Lester 1985). A crystal structure has only been obtained in grape (Petit, Granier et al. 2007). Excluding the C-terminal, both amino acid sequences are highly similar (77% identity). The crystal structure showed that the substrate is wrapped up by both the N- and C-terminals of the enzyme and that the NADPH binding site is located mainly in the N-terminal, which forms a characteristic Rossmann fold (Petit, Granier et al. 2007).

AtDFR has previously been recombinantly expressed in *E. coli*, also in other species, some of which have been characterized (Yan, Chemler et al. 2005; Leonard, Yan et al. 2008; Singh, Kumar et al. 2009).

**ANS**

Anthocyanidin synthase (ANS) follows DFR in the flavonoid pathway. It corresponds to the TDS4 gene in Arabidopsis and is also referred to as LDOX, for leucoanthocyanidin dioxygenase (Abrahams, Lee et al. 2003). ANS is classified as a 2-oxoglutarate dependant, non-haem dioxygenase, and therefore requires 2-oxoglutarate and Fe(II) as cofactors. ANS also requires the presence

---

*Figure 4. Crystal structure of ANS from A. thaliana, complexed with narigenin (Turnbull, Prescott et al. 2001)*
of oxygen and an unusually high concentration of ascorbate for activity (Saito, Kobayashi et al. 1999).

ANS from *A. thaliana* has been purified and crystallized, the structure can be seen in Figure 4 (Turnbull, Prescott et al. 2001). This enzyme catalyzes the next step towards flavan-3-ol biosynthesis: oxidizing flavan-3,4-diols to anthocyanidins. Here, both the substrate and product are relatively unstable. There have been different studies about the stereospecificity of this enzyme and the nature of its products. Leucoanthocyanidin stereochemistry has been shown *in vitro* to affect product outcome. This can be seen in Figure 5. It has been speculated that the 3,4-diol will become either an anthocyanidin or a dihydroflavonol after the first catalytic cycle from ANS. It is further speculated that the dihydroflavonol will further oxidize to a flavonol after a second catalytic cycle (Turnbull, Nagle et al. 2003).

![Figure 5](image_url)  
*Figure 5. Different reactions catalyzed by ANS. Note that substrate stereochemistry may vary product ratios*
Further characterization of this enzyme (Turnbull, Nakajima et al. 2004) experimented with the product outcomes of different substrate stereochemistry. In the direct flavan-3-ol pathway, ANS was found to react with the leucoanthocyanidin to form a cyanidin (Turnbull, Sobey et al. 2000). Complication due to the instability of substrates and products, further levels of oxidation, and the compounds forming complexes with the Fe(II) results in indirect evidence for the true mechanism and product(s). From the natural 2,3 cis-form of leucocyanidin, cyanidin was the least prevalent of the products from ANS, accounting for only 2%, while quercetin took 85% of the product profile (Turnbull, Sobey et al. 2000).

**ANR**

The last step, from anthocyanidin to flavan-3-ol, is catalyzed by another NADPH-dependent enzyme, anthocyanidin reductase (ANR). ANR resides in the BANYULS gene, which, when mutated, allowed accumulation of red anthocyanin pigments and loss of proanthocyanidins (Xie, Sharma et al. 2003). Since there were many similarities in sequence to DFR, the BAN gene was originally thought to encode an enzyme called leucoanthocyanidin reductase, or LAR, to form the flavan-3-ol (+)-catechin (Devic, Guilleminot et al. 1999). Later, BAN from *A. thaliana* was shown to

\[\text{Cyanidin} + \text{ANR} \rightarrow \text{(-)-Epicatechin}\]

\[\text{Leucocyanidin} + \text{LAR} \rightarrow \text{(-)-Catechin}\]

Figure 6. Flavan-3-ol monomers. Substrates are italicized, enzymes are denoted in bold.
catalyze the formation of (-)-epicatechin from cyanidin, leading to the discovery of ANR (Xie, Sharma et al. 2003). Both flavan-3-ol monomers are shown in Figure 6.

Along with the discovery of ANR came the discovery of the stereochemistry of PAs in *A. thaliana*. The lack of (+)-catechin suggested the dominance of ANR activity and eventually led to the lack of a gene encoding LAR (Abrahams, Lee et al. 2003).

Additionally, a full characterization of AtANR recombinantly expressed in *E. coli* has been performed alongside the ANR form *Medicago truncatula* (Xie, Sharma et al. 2004).

**Enzyme Attachment**

In the prospect of creating processes that are more commercially and industrially efficient, techniques have been developed for application in biochemistry. Enzyme attachment to a platform has been shown to provide many benefits for long-term or recovered processes (Sheldon 2007). By using immobilization techniques, biocatalysts can be used in continuous cascade reactions (Veum and Hanefeld 2006). This can also be used to aide stability of a biocatalyst (Cabral and Kennedy 1993), which may increase enzyme performance and percent recovery of the biocatalyst. Immobilization may also allow the enzyme to withstand more extreme temperatures or pH.

Concerning the biosynthesis of flavan-3-ols, there has been little related research. (Deffieux, Gaudrel-Grosay et al. 2009) has used attachment by immobilizing and further stabilizing leucoanthocyanidins instead of attaching the enzyme catalyzing the reaction.
(El-Zahab, Jia et al. 2004) used lactose dehydrogenase, glucose dehydrogenase, and the cofactor, NADH, all tethered to silica glass platforms. Variations in chain length showed differences in the use and generation of NADH between the two enzymes. This experiment was duplicated replacing glucose dehydrogenase with glutamate dehydrogenase (Liu, Zhang et al. 2009). Both showed very high regeneration and turnover of the cofactor.
Hypothesis

Our objective is to combine the concepts of the flavonoid multi-enzyme complex and biocatalyst immobilization to create a more efficient, synthetic flavan-3-ol biosynthesis complex. This complex will include the enzymes from the late steps in the flavonoid pathway leading to the biosynthesis of flavan-3-ols, molecules that have many beneficial attributes, especially in human health.

This will be achieved by creating site-specific attachments on a nanocarrier. The enzymes will be bound by their fusion tags to their respective sites and thus co-localizing around each other, creating a small environment in which to transport unstable intermediates and increase outflux of the desired product.

We propose that immobilization of the enzymes in the late flavonoid pathway to a nanocarrier platform will be beneficial to this scheme in multiple ways. First, it will provide the enzymes with a support and stability that they would not be able to attain in solution. As stated earlier, immobilization can also allow the enzyme to withstand more extreme conditions and resist denaturation. Second, with the immobilized enzymes co-localized onto the same nanocarrier, shuttling of the intermediates will thereby diminish transport losses and prevent unwanted side products.
Chapter 2: Purification of *Arabidopsis* DFR, ANS, and ANR for site-specific attachment onto a chemically modified nanocarrier

**Abstract**

Flavonoids and their derivatives are a group of highly researched secondary metabolites. They are well known for their beneficial and protective qualities in plants as well as in humans. Products of this pathway are highly revered as being supplemental to human health. Like many plant pathways, the close interactions and intricate design gives a strong argument for the existence of a multi-enzyme complex. Here, the enzymes involved in the late steps in the flavonoid pathway leading to flavan-3-ol, or epicatechin, biosynthesis are purified for co-localization onto a polymer. To do this DFR, ANS, and ANR from *A. thaliana* were recombinantly expressed in *E. coli*, purified for using specific affinities, and assayed for activity. In the future, they will work together to create a more efficient flux throughout the final steps in flavan-3-ol biosynthesis.

**Introduction**

Flavonoids comprise one of the largest groups of natural products in plants. There are at an estimated 10,000 different compounds known today (Tahara 2007). Derived from the phenylpropanoid pathway, these products carry an interesting variety of secondary-metabolic functions. They have a substantial impact on human health, showing anti-cancer, anti-obesity, anti-oxidant, and anti-microbial properties (Chung, Wong et al.
1998; Cos, De Bruyne et al. 2004; Tsuda, Ueno et al. 2006). Products of this pathway also show defensive properties in plants. They protect against UV light, attract pollinators, act as a defense mechanism, and can even assist in fertility of the plant (Harborne and Williams 2000). Apart from their many health and survival benefits, flavonoids also have a variety of functions. Condensed tannins, a polymerized form of flavonoid monomers, have an astringency and can be used in drinks such as wine to achieve a specific taste (Bate-Smith 1973). Anthocyanins, a glycosylated derivative of the same flavonoid monomer, are the largest group of natural water-soluble pigments. They provide the brilliant red, orange, and blue pigments in plants and can be used as an alternative for banned or artificial dyes (Yan, Chemler et al. 2005).
There are many outcomes to the flavonoid pathway (Figure 1). One specific path is the production of flavan-3-ols and their polymerized form, proanthocyanidins (PAs), otherwise known as condensed tannins (CTs). The branching point for this is marked by the reaction of dihydroflavonol reductase (DFR). The latter steps in this pathway can be seen in Figure 7 and contain highly unstable intermediates. Leucoanthocyanidin, the product of DFR, is a highly reactive compound and readily polymerizes even under the slightest change in conditions. This first intermediate is then used by anthocyanidin synthase (ANS) which catalyzes the reaction to form anthocyanidin. This flavylium cation is finally reduced by anthocyanidin reductase (ANR) to form the final flavan-3-ol.

While this pathway has been under investigation for over 50 years, the last steps to proanthocyanidins remain unknown.

Over the years, there has been much speculation about the nature of this pathway. Because of its unstable intermediates, enzyme-enzyme interactions, and high level of regulation, it has been proposed to exist in an arranged multi-enzyme complex (Burbulis and Winkel-Shirley 1999). These complexes, or metabolons, facilitate organized and efficient product output. They rely on some interaction with a surface, in this case the endoplasmic reticulum membrane, to provide stabilization and a platform on which to build (Hrazdina and Wagner 1985).

These multi-enzyme complexes can be seen as a way to provide more efficient and interactive activity inside the cell. Developing a way to synthetically create one of these complexes could provide many benefits to providing more efficiency in synthetically creating these end products, even in large-scale processes. Enzymes are currently being
used as biocatalysts in industry as more and more natural products and their derivatives are being used as pharmaceuticals or in the health market (Paterson and Anderson 2005). Enzymes can function with a higher level of substrate or stereo- specificity that is unattainable by traditional chemical processes. As always, there is a high demand in industry, especially in pharmaceutics, to create highly efficient, high output processes. Immobilizing enzymes as catalysts not only overcomes current setbacks in traditional chemical processes, but can increase stability and recovery of the biocatalyst, lowering production costs and increasing product yield. Harnessing protein engineering, and recombinant protein expression in species such as E. coli, allows for a lower cost of biocatalyst production (Sheldon 2007). Using these methods and applying them to these synthetic metabolons, would aid in the formation of compounds from pathways with unstable intermediates and low product yields in vitro, such as the lower flavonoid pathway.

To achieve co-localization on a platform, site-specific immobilization techniques will be used. Therefore, enzymes must carry the proper tags to bind to their respective sites while still retaining their own structure and function. Here, we develop a system on how to express and purify these tagged enzymes to complement the chemical modifications on the biopolymer platform.
Methods

Materials
Standards for TLC analysis ((-)-epicatechin, quercetin, cyanidin chloride, and dihydroquercetin) were purchased from Indofine Chemical Company Inc. (Hillsborough, NJ). Thrombin was purchased from VWR International (Radnor, PA). All other reagents, substrates, and chemicals used were purchased from Sigma-Aldrich (St. Louis, MO).

C41(DE3) chemically competent cells were purchased from Lucigen (Middleton, WI) and Arctic Express (DE3) cells were purchased from VWR Scientific for protein expression. Top10 competent cells were purchased from Invitrogen (Carlsbad, CA).

All PCR primers were synthesized by Invitrogen. Takara Primestar HS Polymerase was ordered from Fischer Scientific (Waltham, MA). Accuprime Pfx polymerase was ordered from Invitrogen.

DNA Manipulations
Native AtANR was obtained from the TAIR database. AtDFR and AtANS amino acid sequences were taken from TAIR and synthesized into pUC57 and codon optimized for growth in E. coli (Genscript USA Inc.). A (glycine$_3$-serine$_1$)$_4$ linker was placed on the N-terminus in between the amino acid sequence for AtANS and the sequence for a monomeric streptavidin tag.

PCR for AtANR was done with Accuprime Pfx polymerase (Invitrogen) using an initial melting temperature of 95°C for 5 minutes, then 34 cycles of 30 sec at 94°C, 30 seconds at 62°C, and 2 minutes at 72°C, followed by 7 minutes at 72°C and 10 minutes at 4°C.
AtDFR and AtANS were cloned using Takara Primestar HS polymerase. Both genes were cloned using the Primestar 2-step method with an elongation time of 1.75 minutes for AtDFR and 2.5 minutes for AtANS. Primers used to isolate these genes were ordered from Invitrogen and diluted to 5 nmol concentrations. They can be seen in Table 1.

Table 1. List of primers used for cloning PCR, all from 5’ to 3’

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>AtANR</td>
<td>CACCATGGACCAGACTTTACACACACCGGA</td>
<td>CACCTCATTAGCTTTGATCAATCCTTTTGACTC</td>
</tr>
<tr>
<td>AtDFR</td>
<td>CACCATGGTGTCCCAAAAAGAAACGGTG</td>
<td>TCAAGCACACATCTGTGCGCCAGC</td>
</tr>
<tr>
<td>AtANS</td>
<td>GGGACAAGTTTTGTACAAAAAAGCAGGCTTCG</td>
<td>GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAAT</td>
</tr>
<tr>
<td></td>
<td>AAGGAGATAGAACCATGGACCCGAGCAAAG</td>
<td>CATTTTTTCAGACA</td>
</tr>
</tbody>
</table>

Proper isolation of the correct gene was determined using electrophoresis in a 1.5% agarose gel. After confirmation by size, the correct DNA fragment was purified by a Millipore DNA gel purification kit and confirmed with full genetic sequencing.

AtANR and AtDFR fragments were inserted into pENTR using the SD/D TOPO enzyme from Invitrogen’s Gateway® recombination system by following protocol. AtANS was inserted into pDONR221 using the BP clonase enzyme from the Gateway® system following the manufacturer’s protocol. Both were transformed using Top10 cells (Invitrogen) and propagated on NZY medium plates with 1.5% (w/w) agarose and 50 µg/mL kanamycin.
All genes used were inserted into their respective destination vector (ANR: pDest17, DFR and ANS: pDest15) using the Gateway® LR clonase enzyme and following protocol. These reactions were transformed again into Top 10 cells for DNA propagation.

**Expression**

For protein expression of ANR and DFR, genes were transformed using C41 (DE3) cells with and without the chaperones Gro EL/ES. 5 mL cultures were inoculated with one colony from the transformation and were shaken at 37°C overnight. They were enlarged at 37°C to 50 mL of NZY media for 2 hours, and then transferred to 1 L NZY at 200 rpm. Cultures were grown until the media reached 0.6-0.8 Abs600. The cells were immediately moved to shake at 16°C for 1 hour. After equilibration at 16°C they were induced by 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and left to shake overnight.

ANS was expressed with both C41 (DE3) chemically competent cells and the Arctic Express RP cell line. After transformation, 5 mL cultures of both cells were grown overnight at 37°C from one colony. Cultures were enlarged to 50 mL of NZY media for 2 hours, and then transferred to 1 L of NZY. C41 cells were shaken at 200 rpm at 37°C until media reached 0.6-0.8 Abs600. All cultures containing 50 µg/mL carbenicillin. Cells were allowed to equilibrate at 16°C for 1 hour and induced by 1 mM IPTG for a total induction time of 14 hours. Arctic Express cells were grown in 5 mL NZY overnight at 37°C with 50 µg/mL carbenicillin. These were enlarged to 50 mL of NZY media with no antibiotic for 2 hours, shaking at 220 rpm at 30°C and then enlarged to 1 L NZY with no antibiotic. This was allowed to shake for 3 hours at 30°C, then moved to shake at 11°C for 1 hour before induced with 1 mM IPTG and shaken for 24 hours.
Cells were harvested by centrifuging media at 6,000 g for 20 minutes. The supernatant was poured off and the cell pellet resuspended in lysis buffer (50 mM Bis-Tris, pH 6.8, 150 mM KCl, 10 mM MgCl2, 10% Glycerol). The resuspended cells were lysed by sonication (Branson Sonifier 450: 3 × 8 s continuous output, at a setting of 4). The lysate was clarified by centrifugation at 16,000 g for 20 minutes.

Purification of AtANR

ANR/pDest17 (38,745 kDa) lysate was incubated on a rotisserie at 4˚C for 2 hours with 2mL Nickel-NTA resin (50% v/v in equilibration buffer: 50 mM Bis-Tris, pH 6.8, 1 mM DTT). The mixture was poured into a mini column under gravity flow. The beads were washed with four volumes of binding buffer I (50 mM Bis-Tris, pH 6.8, 10 mM imidazole, 1 mM DTT) then washed with a third volume of binding buffer II (50 mM Bis-Tris, pH 6.8, 20 mM imidazole, 1 mM DTT). The protein was eluted with 5 mL of elution buffer (50 mM Bis-Tris, pH 6.8, 250 mM imidazole, 1 mM DTT).

The imidazole was removed from the eluate by concentration through a Peirce 20 kDa MWCO protein concentrator to about 1 mL. The resulting protein was diluted 1:10 in dI water with 1 mM DTT and loaded onto a Mono Q HR 5/5 (GE Healthcare) column connected to an AKTA FPLC system. The column used buffer A as 50 mM Bis-Tris buffer and buffer B as 50 mM Bis-Tris and 3 M NaCl, both pH 6.8. The method is as follows: 0-4 minutes at 2% buffer B, 4-14 minutes gradient to 10% buffer B, 14-17 minutes at 100% buffer B. Flow rate was 1 mL/ min with 0.5 mL fractions collected, all at 4˚C.
Selected fractions were collected and analyzed by SDS-PAGE and quantified by \( A_{280} \) with an extinction coefficient of 31,750 M\(^{-1}\) cm\(^{-1}\).

### Purification of AtDFR

1 mM Triton®X-100 was added to the clarified lysate containing AtDFR/pDest15 (70,155 kDa) and was incubated at 4°C for 1 hour on a rotisserie with 2 mL glutathione-agarose resin (50% v/v in sterile dI water). Lyophilized resin had been allowed to swell by gentle mixing on a rotisserie for 24 hours in 13 mL sterile dI water at 4°C. 1 mM DTT, 0.1 mM EDTA, and 0.1 mM phenylmethylsulfonyl fluoride (PMSF) were added during incubation and to each following wash step.

After incubation, resin was collected by spinning at 1,000 rpms in a clinical centrifuge at 4°C. Supernatant was discarded and two volumes of equilibration buffer (10 mM phosphate, pH 7.4, 150 mM NaCl, 1% Triton®X-100) were added. This was mixed on a rotisserie at 4°C for 1 minute. The resin was collected as described previously. This was repeated 3 times.

Two elutions were performed, totaling 5 volumes of elution buffer (10 mM phosphate, pH 7.4, 150 mM NaCl, 1% Triton®X-100, 10 mM glutathione). Each elution was mixed at 4°C for 5 minutes. 1 mM DTT was added to final volume.

Eluate was analyzed by SDS-PAGE and quantified by \( A_{280} \) using an extinction coefficient for both DFR and the GST tag (85,930 M\(^{-1}\) cm\(^{-1}\)).

### Purification of AtANS

AtANS/pDest15 (85,282 kDa) followed the same affinity purification protocol as AtDFR.
Fractions were analyzed with SDS-PAGE, those containing ANS were quantified by $A_{280}$ using the molar extinction coefficient $126\,110\,M^{-1}\,cm^{-1}$ with fusion to GST and $84\,950\,M^{-1}\,cm^{-1}$ with no GST fusion tag.

**Activity Assays**

For initial AtANR assays, optimal conditions, from a previous characterization (Xie, Sharma et al. 2004), were given to ensure enzymatic activity. $100\,\mu g$ of enzyme (or $100\,\mu L$ of lysate, for lysate kinetics) was added to each assay. The ANR assay had a total volume of 1 mL and consisted of $100\,mM$ Tris-HCl, pH 8.0, $2\,mM$ NADPH, $50\,mM$ NaCl, 10% glycerol, and $100\,\mu M$ cyanidin chloride. Kinetics were performed at 55°C. Each was stopped by the addition of 1 mL of ethyl acetate and 1 minute of vortexing. The top layer was extracted and emulsions were broken by centrifugation at 3,000 rpm in a clinical centrifuge. The extraction was repeated two times for a total volume of 3 mL of ethyl acetate. Extracts were dried under nitrogen gas. For lysate kinetics they were re-dissolved in 1 mL methanol for UV analysis. For the purified protein the assays were re-dissolved in $25\,\mu L$ of methanol for TLC analysis. Cellulose TLC plates using a 1:3 methanol:methylene chloride solution were used as described in (Dobashi, Hirano et al. 2008)

AtDFR assays consisted of $100\,mM$ Tris-HCl, pH 7.0, $50\,mM$ NaCl, $2\,mM$ NADPH, and $100\,\mu M$ dihydroquercetin (DHQ) in 1 mL. $100\,\mu L$ of lysate was added and the reaction was incubated at 30°C. Extractions followed the same protocol as the ANR assay extraction. After drying with nitrogen gas the residual was dissolved in a n-butanol/HCl (95%/5% v/v) solution and incubated at 95°C for 30 minutes as described (Stafford and
Lester 1980). This process converts the unstable leucocyanidin to cyanidin (perhaps only at 40% efficiency (Stafford and Lester 1982)) and can be subject to UV analysis at 550 nm.

ANS assays contained 100 mM Tris-HCl, pH 7.0, 40 mM ascorbate, 5 mM Fe(II) sulfate, 5 mM 2-oxoglutarate, and 100 µM DHQ in 1 mL and used the same extraction procedure. This assay was to measure turnover from DHQ to quercetin. Because both substrate and product were colorless, qualitative analysis was performed on TLC with the same method as ANR.

Once activity was confirmed in the lysates, a collective assay buffer was used. The general assay buffer contained 100 mM Tri-HCl, pH 7.0, 40 mM ascorbate, 5 mM Fe(II), 5 mM 2-oxoglutarate, and 1 mM NADPH. These assays were performed in a 37°C water bath. Table 2 shows the conditions for enzyme and substrate concentrations for each enzyme’s respective assay. Total volume of each assay was 500 µL.

Table 2. Specifications for each enzyme and its respective substrate in the collective assay buffer

<table>
<thead>
<tr>
<th></th>
<th>ANR</th>
<th>ANS</th>
<th>DFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>150 µg</td>
<td>100 µg</td>
<td>50 µg</td>
</tr>
<tr>
<td>Substrate</td>
<td>100 µM</td>
<td>800 µM</td>
<td>500 µM</td>
</tr>
</tbody>
</table>
**Thrombin Cleavage**

Two different concentrations of thrombin were used to run the cleavage reaction. The first reaction was run at 4°C at a concentration of 0.5 NIH units/nmol enzyme. The second reaction was carried out at both room temperature and 4°C. This reaction was run at a concentration of 25.7 NIH units/nmol enzyme. Thrombin was reconstituted in 50 mM Tris-HCl, pH 7.0. The final reaction volume contained 150 mM NaCl and 2.5mM CaCl$_2$. Reactions were placed on a rotisserie for gentle mixing. Time points were taken during the reactions and analyzed by SDS-PAGE.

**Results**

The glycerol stock of the wild type *A. thaliana* ANR had been previously attained from the TAIR database and was a prime model for recombinant expression in this study. ANS needed a tag and linker and was planned to be synthesized at Genscript. DFR was also planned to be expressed because of the greater stability of its substrate, even though it has been shown to have low activity in *E. coli* (Martens, Teeri et al. 2002). DFR was also synthesized at Genscript. Both synthetic genes were codon optimized for expression in *E. coli*.

All genes were cloned using primers to make them accessible to the Gateway® system. ANR and DFR were readily cloned into the pENTR vector by the TOPO cloning enzyme. ANS, however, did not go into pENTR and instead used pDONR221 as an entry vector with BP clonase. Although only 51% GC content in nucleic acid makeup, ANS would only be amplified out of its original vector by Takara PrimeStar, a polymerase used especially in GC-rich genes.
Expression studies were done with multiple cell lines and chaperones. Each enzyme was overexpressed in three systems: 1) C41 (DE3), 2) C41 (DE3) – co-transformed with GroEL/ES, and 3) Arctic Express (DE3) RP competent cells. GroEL/ES didn’t assist overexpression in C41, therefore for both ANR and DFR the single transformation C41 cells were used. ANS was expressed and purified in both C41 and Arctic Express for comparison in the amount of in-tact DM3 tag on the protein.

AtDFR was expressed using the pDest15 vector which contains an N-terminal GST-tag, cleavable by thrombin and also in pDest14, a vector for expression of an untagged enzyme. Lysate kinetics with DFR in both pDest15 and pDest14 showed that both constructs were active. Also, the pDest15 construct retained activity after purification with glutathione resin.

Further purification produced about 90% purity as seen in Figure 10. Activity was tested by acid hydrolysis of the assay extracts producing a pink color corresponding to the

Figure 8. PCR product of ANS. Lane 1) 500 bp ladder, lane 2) PCR product to be purified from ANS/pDONR, lane 3) positive ANS control from original pUC57 vector

Figure 9. Left - DFR assay extracts after acid hydrolysis to convert leucocyanidin to cyanidin (pink). Right – control (clear)
formation of cyanidin from leucocyanidin. An example of these results can be seen in Figure 9.

![Figure 9](image)

**Figure 10.** Purification of DFR/pDest15 - (starting from left) lane 1) lysate, lanes 2-4) resin washes, lane 5) eluate, lane 6) empty, lane 7) standard ladder

![Figure 11](image)

**Figure 11.** Lysate kinetics for AtDFR as an untagged enzyme (pDest14) and with a GST tag (pDest15). Absorbance at 550 nm corresponds to the amount of cyanidin resulting from the acid hydrolysis of leucocyanidin produced by the enzyme
The 6xHis-Ni purification system is a widely used purification method and the 6xHis fusion tag exists in the Gateway® system as pDest17. The Gateway® cloning and expression system was used to achieve an N-terminal 6xHis tag on ANR. Recombinantly expressed AtANR in *E. coli* has previously been characterized (Xie, Sharma et al. 2004). ANR was twice purified, once with Ni-NTA resin and then again with anion exchange FPLC. Figure 12 shows a purified ANR peak at fractions A2-A4.

![FPLC chromatogram of purifying AtANR; purified protein peak between 3.5-4 mL](image)

Activity of AtANR lysate was tested by using UV analysis. Its substrate, cyanidin, has a red color and has a maximum absorbance at 550 nm. Assay extracts were expected to shift in absorption from 550 nm to 280 nm, the maximum absorbance of epicatechin. Using a negative controlled assay, the absorbance at 280 nm was monitored over time, representing the production of epicatechin from cyanidin, the results of which can be seen in Figure 13.
Figure 13. Averaged ANR assay results showing increasing absorbance at 280 nm, the maximum absorbance wavelength for epicatechin. Data was taken from two different purification batches.

To qualitatively show ANR activity, extracts were dried down and re-dissolved in 25 µL of methanol. 2.5 µL was used on a silica TLC plate against the product standard, shown in Figure 14.

Figure 14. TLC with ANR assay extractions. Lane 1) ANR with no substrate, lane 2) ANR with cyanidin substrate, lane 3) epicatechin standard, lane 4) DHQ standard, lane 5) quercetin standard. Rf values for epicatechin in assay (0.643) and standard (0.610)
For the attachment of ANS to the biotin on the polymer platform, an avidin or streptavidin tag was needed. A single-stranded streptavidin double-mutant was developed by (Qureshi and Wong 2002) titled DM3. This mutant contained two single-point mutations in the monomeric version of streptavidin that allowed the single strand to exist solely as a monomer even at room temperatures. It also gave the mutant a higher binding constant for biotin than the other double mutants (1.28 x 10^{-8} M), though not as high as the tetrameric form of streptavidin (≈ 10^{-14} M).

Because of the need to perform an affinity purification with ANS, the gene was inserted into pDest15 which contains an N-terminal GST tag. This also helped to distinguish between the enzyme (58.4 kDa) and the Cpn 60 chaperone present in the Arctic Express cells (60 kDa).

TLC was used to find ANS activity, shown by oxidizing DHQ to quercetin. A 1 hour assay extract was dried under nitrogen gas and re-dissolved in 25 µL of methanol. 5 µL were run next to DHQ and quercetin standards, shown in Figure 15.

![Figure 15. TLC of ANS assay extracts. Arrow is pointing to product (quercetin) band. Lane 1) ANS assay, lane 2) epicatechin standard, lane 3) DHQ standard, lane 4) quercetin standard. Rf values: for both quercetin generated in assay and the quercetin standard at 0.926]
To remove the GST tag before attachment to the polymer, a thrombin cleavage was performed. The cleavage site was predicted by PeptideCutter (SIB Swiss Institute of Bioinformatics (Gasteiger E., Gattiker A. et al. 2003)). The predicted cleavage site leaves 13 amino acids to the N-terminal of the sequence as displayed in Figure 6. The amino acids were not predicted to affect binding to biotin.

\[
\begin{array}{c}
\text{Thrombin Cleavage Site} \\
\text{GST Tag-} \\
\text{...ATFGGDDHPPKSDLPR} \quad \text{PWSNQTSLYKKA}_{\text{DM3}}-(\text{GSSS})_{\text{4}}-\text{ANS} \\
\end{array}
\]

Figure 16. Shows the predicted thrombin cleavage site between the GST tag and the ANS enzyme with glycine-serine linker and N-terminal DM3. The added 13 amino acids after cleavage are bolded and underlined

**Discussion**

Many studies have been done to elucidate the flavonoid pathway and to characterize the enzymes and phenotypic mutants. Products from this pathway are gaining notoriety as a group of highly functional compounds. To obtain these products, unstable intermediates are sequestered through a highly regulated series of enzymes. Co-localization and cell fractionation experiments show that many enzymes occupy the same physical area within the cell (Hrazdina and Wagner 1985; Saslowsky and Winkel-Shirley 2001; Saslowsky, Warek et al. 2005). Also, metabolic research has shown high levels of regulation throughout most of the pathway (Burbulis and Winkel-Shirley 1999; Abrahams, Lee et al. 2003). This evidence indicates that a multi-enzyme complex is likely for the flavonoid pathway enzymes. Products of the flavonoid pathway are involved in plant survival,
defense, and proliferation and have proven useful in human health and disease prevention (Cos, De Bruyne et al. 2004; Aron and Kennedy 2008). Their wide scope of applications has made the flavonoid pathway a target of many research teams. Here, protein expression and purification has been performed in pursuit of creating a site-specific method of attachment for these enzymes in multi-enzyme complexes.

With recombinant expression in E. coli, we have produced three functional flavonoid enzymes from the model plant Arabidopsis thaliana: dihydroflavonol 4-reductase, anthocyanidin synthase, and anthocyanidin reductase. These enzymes catalyze the final steps in the flavonoid pathway and produce epicatechin, a flavan-3-ol monomer. These compounds have a variety of health benefits and applications as well as being the building blocks for proanthocyanidins, a class of condensed flavonoids with similar functions (Chung, Wong et al. 1998; Cos, De Bruyne et al. 2004).

ANR has previously been characterized in A. thaliana (Xie, Sharma et al. 2004) and ANS has been purified and crystallized (Turnbull, Prescott et al. 2001). Until recent years, recombinant expression of DFR in E. coli has been unsuccessful in yielding soluble and functional enzymes in vitro (Leonard, Yan et al. 2008). Now, under different expression and purification systems, a unique form of each enzyme has been expressed, purified, and recovered retaining enzyme activity.

The goal for choosing an expression and purification system was to obtain a unique tag for ANS and ANR so that each could be attached to a chemically modified polymer surface. These attachments are separate and unrelated in that the location and spatial density of each enzyme could be manipulated by the attachment modification procedure.
The two chemical surface modifications that were chosen were nickel and biotin. Therefore, the enzyme tags would become 6xHis and streptavidin, respectively, due to their high binding affinity.

A covalent attachment of the enzyme to the polymer could be performed, but because of lack of specificity of a binding site on the protein this could cause deactivation of the enzyme as well as blockage to the active site. This could render a portion of the attached enzyme denatured or without activity. Thus, a tag or fusion protein tag would be more beneficial to predict binding patterns.

Although DFR would not be attached to the polymer platform initially, affinity purification would allow for better control and purity of the leucocyanidin products after assay extraction. With the N-terminal GST tag near the location of the NADPH binding site (Rossmann fold), activity might have been hindered by blocking or reducing accessibility of this site. DFR was expressed in both the pDest14 (untagged) and pDest15 (N-terminal GST tag) expression vectors. Activity was shown in both systems and did not decrease when expressed as a fusion protein to GST, suggesting that the N-terminal tag did not adversely affect the ability of the enzyme to utilize NADPH. Although GST is used a transporter for some flavonoid compounds, for example anthocyanins, using a GST tag hasn’t shown any binding or inhibition with DFR substrates or products (Hayashi, Takahashi et al. 2005).

Using a GST tag to purify DFR would allow for two options concerning leucocyanidin. For production and storage of the leucocyanidin substrate, the DFR-GST fusion protein could be purified and used to create leucocyanidin which could then be lyophilized and
stored at -80°C for use in future assays. Although, leucocyanidin has been shown to convert to DHQ under storage (Pang, Peel et al. 2007). Otherwise, the DFR-GST fusion protein could remain bound to the glutathione resin. After washing the resin thoroughly, DFR could be introduced to the assay, incubated for leucocyanidin production, and then removed with the resin by centrifugation. Because DHQ is both a substrate from DFR and a product of ANS, this would eliminate circulation of DHQ throughout the system. Cleavage of the tag may not be necessary.

Expression of ANS in pDest14 produced very little expression in C41 cells. Arctic Express cells or C41 cells co-transformed with a Gro EL/ES chaperone were employed to facilitate expression. These cells, however, express a folding chaperone approximately the same size at ANS. Protease activity was also hypothesized due to a large overexpression band at approximately 40 kDa, possibly corresponding to the size of ANS without the DM3 (monomeric streptavidin) tag. To create an affinity purification procedure for ANS, it was expressed in pDest15, which altered the size of the protein to 84 kDa. This allowed for a differentiation between ANS and the 60 kDa folding chaperone. This expression and purification was performed in both Arctic Express and C41 yielding comparable amounts of protein. Both retained activity as a double fusion protein with DM3 and GST on the N-terminal. Purity of ANS after GST elution was reduced due to poor recovery of the resin during batch purification.

The DM3 tag was chosen for simplicity in design and its ability to bind biotin with a sufficiently low dissociation constant. Although the affinity is not as high as the wild-type streptavidin ($K_d \approx 10^{-14} \text{ M}$), it still maintains a dissociation constant of $1.28 \times 10^{-8} \text{ M}$,
which is higher than that of a monomeric avidin ($K_d = 1.8 \times 10^{-7}$ M) used for commercial affinity purposes. This DM3 sequence is a double-point mutation of the single stranded streptavidin (T90A and D128A). It has proven to exist solely as a monomer (Qureshi and Wong 2002) and therefore it is unlikely to cause tetrameric forms of DM3-ANS. This sequence was synthesized at the N-terminal of the ANS gene, linked by a flexible (Gly$_3$-Ser)$_4$ linker. This was to allow for proper folding of both proteins. Correct folding was confirmed by GST affinity to glutathione during purification and confirming ANS activity. The N-terminal attachment to ANS was chosen due to the fact that there were no residues near the N-terminal that were involved in substrate or cofactor binding (Wilmouth, Turnbull et al. 2002).

Thrombin cleavage of the GST tag from the remaining DM3-ANS protein was predicted to only have one cleavage site, ultimately adding 13 amino acids to the final N-terminal. The cleavage reaction was carried out at two different concentrations at both room temperature and 4°C. Running time points for the reaction up to 24 hours on SDS-PAGE showed that the reaction was not near completion at 4°C, and also that there may have been a second cleavage site. The second site was speculated to reside near the C-terminal, due to similarities in the sequence to that of the cleavage point. This would cause ANS to lose the last 24 amino acids, some of which are involved in substrate binding. A small sample of the post-cleavage enzyme was run on the FPLC using anion exchange. Two possible bands were present and co-eluted. Furthermore, running a multi-day cleavage reaction at 4°C could significantly reduce enzyme activity. Thus, the decision was made to attempt streptavidin affinity to biotin without cleaving the GST tag.
Chapter 3: General Conclusions

In this work, we have demonstrated recombinant expression and purification of three enzymes in the late flavonoid pathway with the required tag specificities. Maintaining activity of these sequential flavonoid enzymes and their fusion tags through purification was the first step in building a flavonoid biosynthetic multi-enzyme complex. This will allow site-specific binding and spatial control of the enzymes when bound to a chemically modified surface. Steps have been taken to ensure that the fusion tags for affinity purification do not interfere with the activity of the enzyme. Additionally, the use of a monomeric streptavidin tag provides a simple design method for the attachment of enzymes onto biotinylated surfaces. These methodologies could allow for scale up biocatalyst processes and could be applied to any enzymatic coupling reaction. This provides an alternative to free enzyme reactions, where product generation may be low and the intermediates are unstable. By facilitating to the enzyme and reaction specificities, this method will allow for better regulation and control over these biocatalyst processes.
References


I would like to take a bit of space and express my gratitude for everyone that has helped me throughout my time at Iowa State. First and foremost, I would like to thank the Dr. Reuben Peters for giving me the wonderful opportunity to work in his lab and for taking a chance on hiring me. You are an exceptional mentor and have created a great environment to work in. I would also like to thank Dr. Surya Mallapragada and Dr. Balaji Narasimhan for their support and patience and for letting me take part in such an inspiring project.

To Dr. Frannie Mann, I can’t even begin to think of all the ways you have helped me grow as a person and as a scientist, you have been an amazing friend and mentor, and I know that you are going to do wonderful things. To Jill Thomas, thank you for always being someone that let me talk incessantly about my work and whom I can bounce all my crazy ideas off of; you helped me learn and brightened many days. You two have been my support system and I can’t thank you enough. If the opportunity arises to start The Thomas Mannport School of Higher Learning, I’m in.

Thanks to my parents, for supporting me and always pretending to be interested in the intricate details of my research no matter what. And to the rest of the Peters Lab, for functioning as a big family - it has been so great working with all of you for the past four years.