

# Crystallization Studies of Pictet-Spenglerases

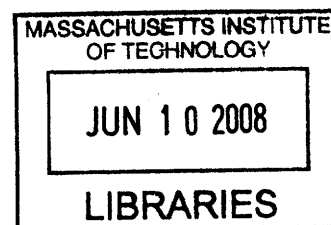
By  
William C. Hillmann

B.S., Biochemistry (2006)  
Boston College

Submitted to the Department of Chemistry in June 2008 in Partial Fulfillment of  
the Requirements for the Degree of

Master of Science  
in Biological Chemistry

At the  
Massachusetts Institute of Technology  
June 2008



© 2008 Massachusetts Institute of Technology  
All rights reserved

**ARCHIVES**

Signature of Author \_\_\_\_\_  
Department of Chemistry  
May 2008

Certified by \_\_\_\_\_  
Sarah E. O'Connor  
Assistant Professor  
Thesis supervisor

Accepted by \_\_\_\_\_  
Robert W. Field  
Chairman, Departmental Committee on Graduate Students

# Crystallization Studies of Pictet-Spenglerases

By

William C. Hillmann

Submitted to the Department of Chemistry on May 23, 2008 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Biological Chemistry

## ABSTRACT

Natural products are a rich source of medicinally important molecules. Monoterpene indole alkaloids from plants are an especially important source of therapeutic molecules. Due to the complexity of these molecules, biosynthesis of derivatives is an attractive way of obtaining molecules with potentially new or improved functionality. The rational design of mutants with altered/expanded substrate scope is an important step in engineering organisms to produce such compounds.

In monoterpene indole alkaloid biosynthesis, the enzyme strictosidine synthase catalyzes the first committed reaction. This reaction is a Pictet-Spengler coupling between tryptamine and secologanin and produces the biosynthetic intermediate strictosidine, common to all monoterpene indole alkaloids. To better understand the structural features that impart binding selectivity, crystallization studies of this enzyme were performed. The native enzyme and several interesting mutants were studied; co-crystallization experiments with inhibitors and substrates were also performed. Diffraction quality crystals of the native enzyme were obtained following optimization by grid screening, additive screens, and macroseeding. Data on the optimized crystals was collected at the Argonne National Labs synchrotron radiation source.

In addition to monoterpene indole alkaloids, the benzyloquinoline alkaloids are another class of medicinally important plant derived natural products. In a reaction analogous to that catalyzed by strictosidine synthase, the first committed step of benzyloquinoline biosynthesis is a Pictet-Spengler reaction between 4-hydroxyphenylacetaldehyde and dopamine, catalyzed by the enzyme norcoclaurine synthase. Two different forms of this enzyme have been identified, neither of which shows any homology to strictosidine synthase. Structural information for these enzymes could provide general structural features required for enzymatic Pictet-Spengler reactions. Before crystallization, the enzymes were expressed and tested for activity. Once active preparations of protein were available, crystallization studies were performed and crystals were obtained.

Thesis supervisor: Sarah E O'Connor  
Title: Assistant Professor

## **Acknowledgements**

First, thank you to my research advisors Professor Sarah O'Connor and Professor Cathy Drennan for all their help and support during my time. The effect of their help and advice, regarding the directions of my project and my life, has been invaluable.

Thanks also to the Drennan and O'Connor lab members for always taking the time to help out or answer questions. They helped make my experiences both enjoyable and interesting.

Thanks to Caroline for her unwavering support over the years, for always being willing to listen, and for keeping me sane and grounded when I most needed it.

Finally, thanks to my family for all their love and support in everything I chose to pursue.

## **Table of Contents**

Abstract.....	2
Acknowledgements.....	3
Table of Contents.....	4
List of Abbreviations.....	6
<b>1. Introduction</b>	
1.1 Medicinal Natural Products.....	7
1.1.1 Biosynthetic pathways to modified natural products	
1.2 Monoterpene Indole Alkaloids.....	9
1.2.1 Biosynthesis of monoterpene indole alkaloids	
1.3 Strictosidine Synthase.....	13
1.3.1 Reaction mechanism of strictosidine synthase	
1.3.2 Crystal structure of <i>R. serpentina</i> strictosidine synthase	
1.3.3 Obtaining modified monoterpene indole alkaloids	
1.4 Benzyloquinoline Alkaloids.....	21
1.4.1 Biosynthesis of benzyloquinoline alkaloids	
1.5 Norcoclaurine Synthase.....	23
1.5.1 Multiple forms and their relation to strictosidine synthase	
1.5.2 Reaction mechanism of <i>T. flavum</i> norcoclaurine synthase	
1.5.3 NMR based homology model of <i>T. flavum</i> norcoclaurine synthase	
1.6 Summary and goals.....	27
1.7 References.....	30
<b>2. Strictosidine synthase crystallization studies</b>	
2.1 <i>C. roseus</i> strictosidine synthase.....	32
2.1.1 Kinetic characterization of wild-type STS	
2.2 Crystallization experiments with <i>C. roseus</i> STS mutants.....	35
2.2.1 Sparse matrix screening with D177A and F226L	
2.2.2 Optimizing crystallization of F226L	
2.3 Crystallization experiments with wild-type <i>C. roseus</i> STS.....	39
2.3.1 Sparse matrix screening for initial conditions	
2.3.2 Optimizing crystallization of wild-type STS	
2.4 Data processing with <i>C. roseus</i> STS.....	49
2.4.1 Indexing with Denzo	
2.4.2 Integration and scaling with Denzo/Scalepack	
2.5 <i>R. serpentina</i> strictosidine synthase.....	57
2.5.1 Preparation of <i>R. serpentina</i> STS for crystallization	
2.5.2 <i>R. serpentina</i> STS activity assays	
2.6 Summary and future work.....	60
2.7 Materials and Methods.....	61
2.8 References.....	74



<b>3. Norcoclaurine synthase crystallization studies</b>	
3.1 Assay design and substrate synthesis.....	75
3.1.1 Synthesis of 4-hydroxyphenylacetaldehyde	
3.1.2 Monitoring reaction progress	
3.2 <i>T. flavum</i> norcoclaurine synthase.....	80
3.2.1 Expression of <i>T. flavum</i> NCS	
3.2.2 Dynamic light scattering	
3.3 Crystallization experiments with $\Delta 19$ <i>T. flavum</i> NCS.....	84
3.3.1 Sparse matrix screening for initial conditions	
3.3.2 Optimization of initial hits	
3.3.3 Co-crystallization experiments	
3.4 <i>C. japonica</i> norcoclaurine synthase.....	89
3.4.1 Expression and thrombin cleavage of <i>C. japonica</i> NCS	
3.4.2 Enzyme activity assays	
3.5 Summary and future work.....	94
3.6 Materials and Methods.....	96
3.7 References.....	108
<b>Appendix 1 Curriculum vitae.....</b>	<b>109</b>

## **List of Abbreviations**

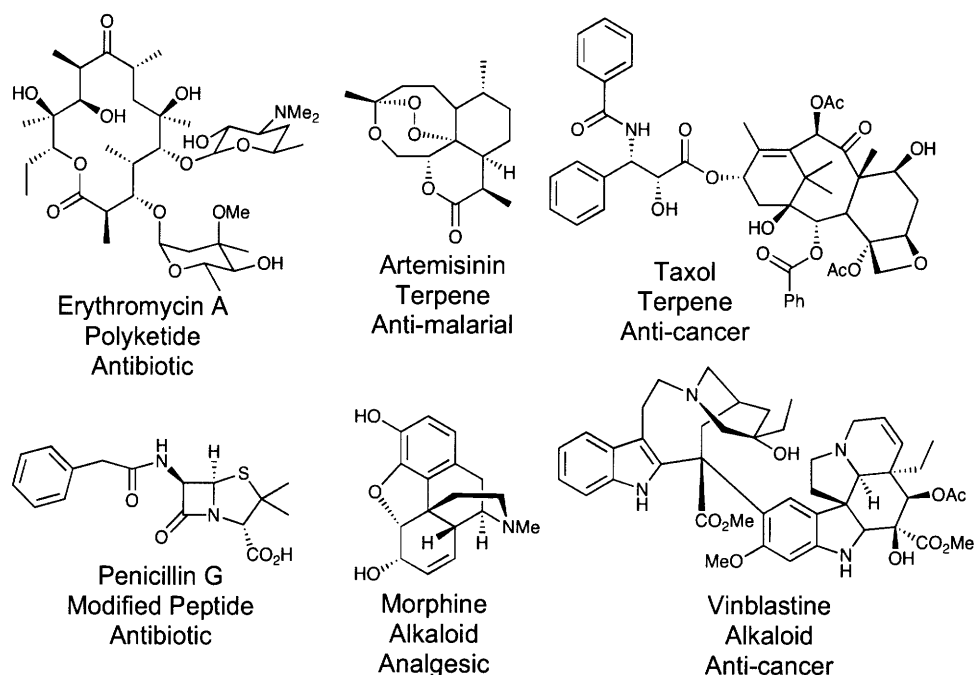
BIA	Benzylisoquinoline alkaloid
CjNCS	<i>C. japonica</i> norcoclaurine synthase
CrSTS	<i>C. roseus</i> strictosidine synthase
HCOONa	Sodium formate
4-HPAA	4-hydroxyphenylacetaldehyde
HPE	2-(4-hydroxyphenyl)ethanol
MBS	Mutational biosynthesis
MIA	Monoterpene indole alkaloid
MPD	(+/-)-2-methyl-2,4-pentandiol
NCE	New chemical entity
NCS	Norcoclaurine synthase
PDB	Precursor directed biosynthesis
PEG	Polyethylene glycol
RsSTS	<i>R. serpentina</i> strictosidine synthase
STS	Strictosidine synthase
TfNCS	<i>T. flavum</i> norcoclaurine synthase

## 1. Introduction

### 1.1 Medicinal Natural Products

Natural products provide a rich source of bioactive molecules used for treating a wide range of different human diseases.<sup>1</sup> Even in light of new methodologies for screening large, diversity oriented small molecule libraries, natural products still provide a large number of lead compounds used for developing new drugs. In a testament to the continuing utility of natural products, most combinatorial libraries rely on the combination of structural scaffolds and modifiers derived from natural products.<sup>2</sup> In an analysis of 1010 new chemical entities (NCEs) from 1981-2006, Newman and Cragg found that 43 were natural products (4.2%), 232 were derived from natural products by semi-synthesis (22.9%), 108 were synthetically made natural

Figure 1.1 – Medicinally relevant natural products.



product mimics (10.6%), 47 were synthetic with the pharmacophore derived from a natural product (4.6%), and 107 were synthetic natural product mimics with the pharmacophore from a

natural product (10.5%).<sup>1</sup> In total, more than half of the NCEs were derived from or inspired by natural products.

Medicinal natural products come from a wide array of natural product classes and exhibit a staggering array of structural diversity. Some of the classes that have yielded natural product drugs include aromatic polyketides, polyethers, coumarins, flavonoids, terpenoids, alkaloids, nonribosomal peptides, and aminoglycosides.<sup>3</sup> A sampling of relevant medicinal natural products can be seen in Figure 1.1.

### **1.1.1 Biosynthetic Pathways to Modified Natural Products**

Making modified natural products is an important way of obtaining new drugs. However, synthesis of these modified compounds is in many cases a technically challenging task due to the structural complexity of the molecules involved. Total synthesis of modified compounds is not an optimal solution, since these syntheses are often complicated and give very low yields of product. In view of these difficulties, biosynthetic pathways to modified natural products seem to be an attractive way to investigate new functionality with potential to scale up the production to industrial scales. In a recent review, several different biosynthetic approaches to modified natural products were detailed, including precursor directed biosynthesis and mutational biosynthesis.<sup>4</sup>

In precursor directed biosynthesis (PDB), modified starting materials are fed to wild-type organisms. Any modified natural products are then isolated and characterized. While no genetic manipulations are necessary, this strategy is limited by the ability of the producer organism biosynthetic pathway to accept the modified starting material and incorporate the modified products in downstream steps in the production pathway. Another biosynthetic approach to modified natural products is mutational biosynthesis (MBS), or mutasynthesis, to use the

terminology of Rinehart.<sup>5</sup> This strategy involves the creation of mutant enzymes which are then supplemented to the producing organism. These mutants are selected so that they incorporate non-native substrates into the biosynthetic pathway. As with PDB, any modified natural products produced can be isolated from cultures, characterized, and assayed for new or improved activity.

## 1.2 Monoterpene Indole Alkaloids

Alkaloids compose one of the largest and most structurally diverse classes of secondary metabolite natural products. These compounds are defined by the presence of a basic nitrogen atom (the term “alkaloid” comes from “alkali”); the different alkaloid classes can be differentiated based on the source of the nitrogen atom or the structural scaffold present in the final product.<sup>6</sup> The source is usually an amino acid, including both proteinogenic amino acids

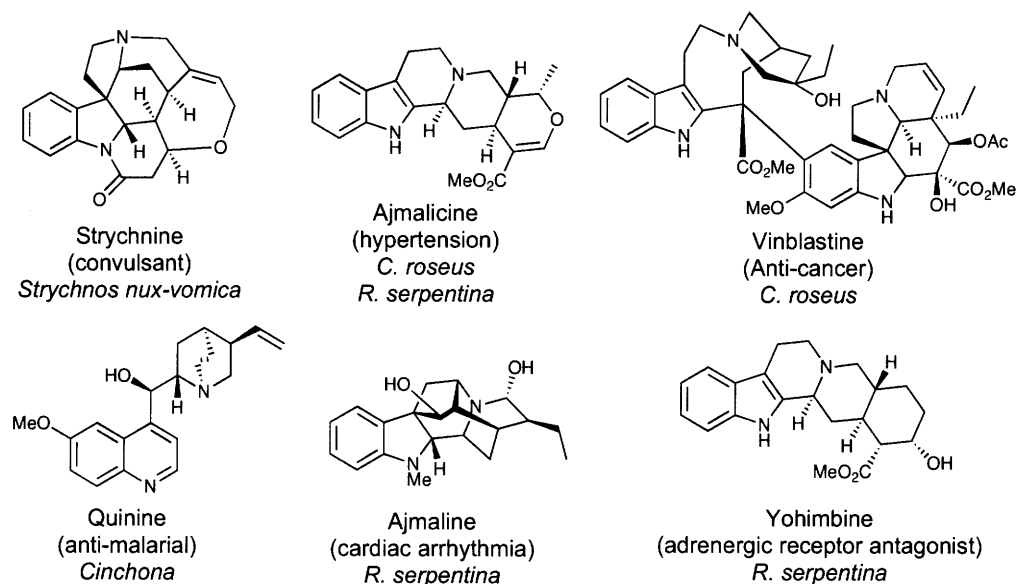


Figure 1.2 – A selection of biologically active monoterpene indole alkaloids

(i.e. lysine, tryptophan, and tyrosine) and non-proteinogenic amino acids (i.e. ornithine).<sup>3</sup>

Alkaloids are produced by an array of different organisms, including plants, animals, bacteria,

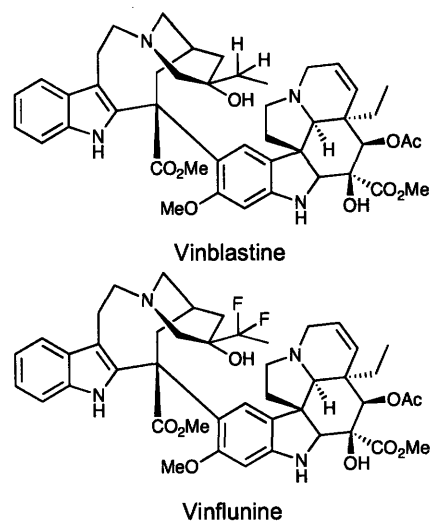
and fungi. Due to the prevalence of the use of plants in traditional therapies for a large number of illnesses, plant alkaloids include some of the most historically important pharmaceutical compounds.<sup>7</sup>

The monoterpene indole alkaloids (MIAs) are one major class of alkaloids, so named due to the presence of an indole (derived from tryptophan) and a monoterpene (secologanin, itself a complex secondary metabolite).<sup>8</sup> There are approximately 2000 members of this class of natural products, and they display remarkable structural diversity and a myriad of different biological activities.<sup>9,10</sup> Figure 1.2 shows a selection of MIAs and their associated medicinal uses.

Most notable in this selection are the *Vinca* alkaloids vincristine and vinblastine, anti-cancer compounds from *Catharanthus roseus*. The *Vinca* alkaloids are used to treat leukemia, lymphomas, and several solid tumors.<sup>11</sup> They have been shown to arrest the cell cycle by destabilizing microtubules, and have

been implicated in apoptosis mediated by an NF- $\kappa$ B mediated signaling pathway.<sup>11</sup> Also notable is that a recent semi-synthetic derivative, vinflunine (shown in Figure 1.3), has been shown to have improved activity and act through a novel mechanism.<sup>12</sup> In addition to demonstrating the importance of MIAs in medical treatment, the new findings regarding vinflunine demonstrate the ability of modification to impart new and interesting activities.

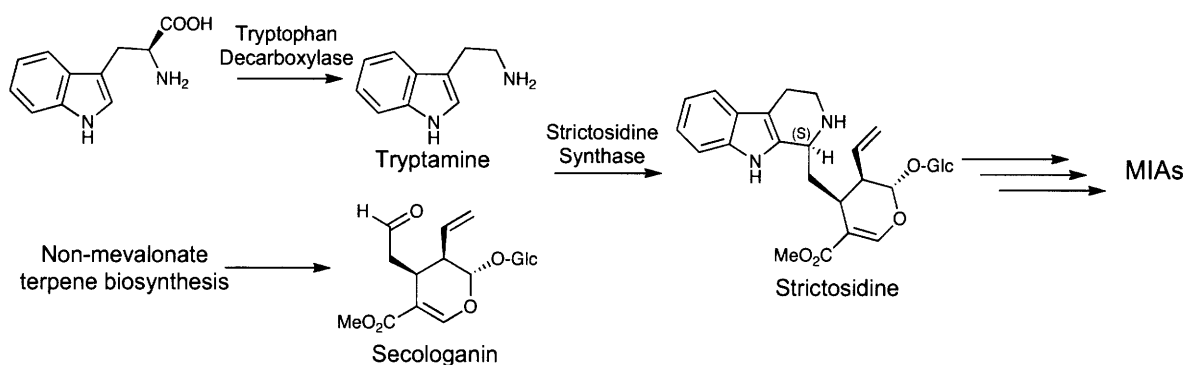
Figure 1.3 – Vinblastine and vinflunine.



### 1.2.1 Biosynthesis of Monoterpene Indole Alkaloids

In many ways, studying the biosynthesis of secondary metabolites in plants is a daunting prospect. Unlike bacteria and fungi, the biosynthetic genes encoding plant natural products are not clustered. Instead of identifying a single gene and being able to find the rest of the enzymes in a particular pathway by sequencing genomic DNA on either side of the initial hit, each enzyme in a plant biosynthetic pathway must be identified, cloned, and isolated individually.<sup>13</sup>

Scheme 1.1 – Initial steps in MIA biosynthesis.

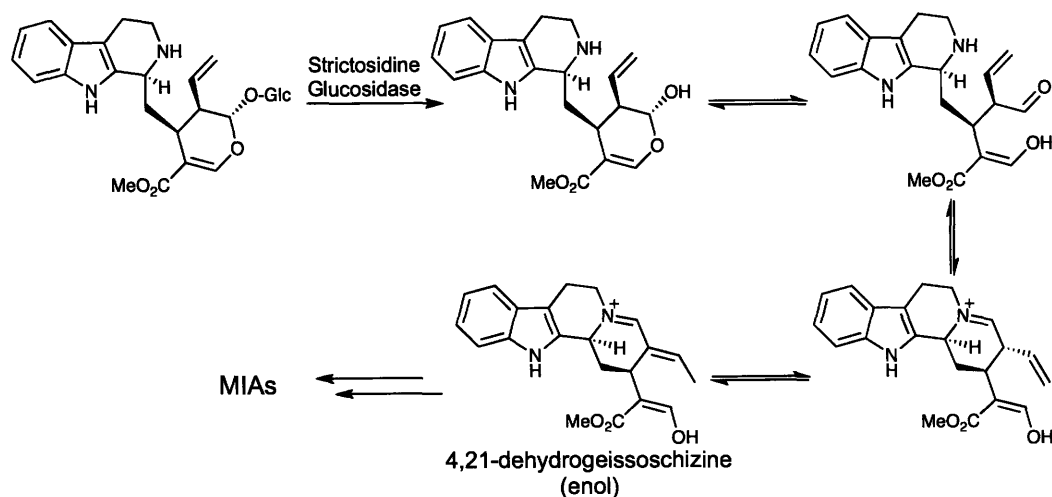


Within the group of MIAs, there are several different subclasses of alkaloids that have common structural features. The abundance of these different subtypes varies between producer organisms that include *Rauwolfia serpentina*, *Catharanthus roseus*, *Ophiorrhiza pumila*, and the *Cinchona* species.<sup>8</sup> From these diverse subclasses, it is clear that in different organisms the biosynthesis of MIAs can follow various different branching pathways. However, the MIAs are united by their common precursors (tryptamine and secologanin) and a common committed enzymatic step. This step is the coupling of tryptamine and secologanin through a stereospecific Pictet-Spengler reaction catalyzed by the enzyme strictosidine synthase (STS).<sup>14</sup> Tryptamine is derived from the amino acid tryptophan by an enzymatic decarboxylation catalyzed by tryptophan decarboxylase and secologanin is itself a terpenoid natural product thought to be derived from the non-mevalonate pathway.<sup>15,16</sup> The  $\beta$ -carboline produced by STS is known as

strictosidine, and is the common intermediate in MIA biosynthesis. This first reaction step is common of all MIAs, regardless of their eventual final structures. These initial steps are illustrated in Scheme 1.1.

After this initial coupling step, another common (but not ubiquitous) feature of MIA biosynthesis is the deglycosylation of strictosidine by the enzyme strictosidine glucosidase (SGD).<sup>17</sup> The glucose moiety serves multiple functions in secologanin and strictosidine. These functions include recognition and proper positioning of secologanin in the STS active site. The glucose also acts as a protecting group in the strictosidine product. Removal of the glucose reveals a reactive intermediate that can undergo spontaneous rearrangement to give 4,21-dehydrogeissoschizine.<sup>18</sup> This intermediate can be converted into a range of alkaloid structures from the different subclasses. An interesting exception to this general pathway can be seen in the biosynthesis of the quinoline alkaloid camptothecin.<sup>19</sup> In addition to a rearrangement of the indole ring system

Scheme 1.2 – Deglycosylation reveals a reactive intermediate.



into a quinoline system, following STS the first step of camptothecin biosynthesis is formation of strictosamide by attack of the  $\beta$ -carboline nitrogen on the methyl ester of secologanin. It is



currently not known whether the biosynthetic pathway for the other quinoline MIA, quinine, features a similar first step.<sup>8</sup> The degree to which the enzymes involved in the downstream biosynthetic steps are characterized varies based on the particular pathway, although feeding studies have been useful in identifying plausible intermediates.

### 1.3 Strictosidine Synthase

Strictosidine synthase is a Pictet-Spenglerase that catalyzes the stereoselective coupling of tryptamine and secologanin to give the  $\beta$ -carboline product strictosidine. It is a 344 amino acid protein and contains no metal or cofactors. As described above, STS represents the first committed step in MIA biosynthesis. Due to its central role in the synthesis of a number of medicinally important alkaloids, early efforts were made to identify and clone STS. These efforts are described in an excellent review by T.M. Kutchan.<sup>20</sup> STS was identified by traditional reverse genetic methods, meaning that the isolation of the enzyme preceded the isolation of the gene. STS was initially purified and characterized using cell-free extracts from *C. roseus* and *R. serpentina* cell suspension cultures.<sup>21,22</sup> Once cDNA libraries for producing organism became available, the gene for STS was isolated from both *C. roseus* and *R. serpentina*.<sup>23,24</sup> STS has since been cultured and expressed in a variety of heterologous systems, including *E. coli* and *S. cerevisiae*.<sup>25</sup> Notably, STS from *C. roseus* and *R. serpentina* features a predicted signal peptide used for intracellular localization. This must be removed before active protein may be obtained from heterologous systems.

#### 1.3.1 Reaction mechanism of strictosidine synthase

The Pictet-Spengler reaction is a general reaction between aldehydes and electron rich aromatic amines that involves two basic steps.<sup>26,27</sup> The first is the condensation of the amine and the aldehyde to form an iminium species. The electrophilic carbon of the iminium is then subject

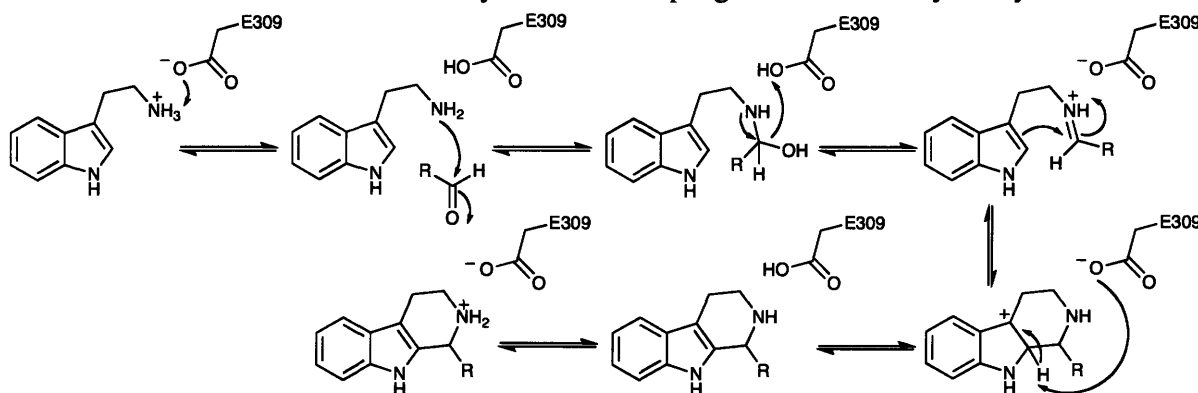
to attack by the aromatic ring in a classical electrophilic aromatic substitution reaction. The final intermediate species is deprotonated to restore aromaticity and complete the reaction sequence. The enzymatic Pictet-Spengler reaction has been kinetically characterized from several different species (including *R. serpentina* and *C. roseus*), and the mechanism of *C. roseus* STS has been studied in detail.<sup>28</sup>

In the Pictet-Spengler reaction between tryptamine and secologanin, a chiral center is formed in the product and as such there are two possible stereoisomeric products – strictosidine (S configuration) and vincoside (R configuration). In model chemical reaction systems, both strictosidine and vincoside are produced, however in the enzyme catalyzed system strictosidine is produced asymmetrically. The initially proposed mechanism for STS involved several acid/base catalyzed steps. An X-ray crystal structure of STS from *R. serpentina* suggested that Glu309 (*R. serpentina* sequence numbering) was the most likely residue for this catalytic proton shuttle.<sup>29</sup> Indeed, mutagenesis of this residue to an alanine resulted in the greatest loss of activity seen for a number of different active site mutants which were tested.<sup>29</sup>

Kinetic analysis of the mechanism of STS was performed by using kinetic isotope effects determined with deuterium labeled tryptamine substrates.<sup>28</sup> The data from these reactions showed that the rate-determining step of the reaction was the final deprotonation to reform aromaticity. While this may seem counter-intuitive at first, it is important to note that following the electrophilic aromatic substitution, both the forward reaction (to  $\beta$ -carboline product) and the reverse reaction (to the iminium precursor) would accomplish the energetically favorable reformation of aromaticity. As such, the rate of the reverse reaction is substantial, slowing the rate of the forward reaction enough to make it the rate-limiting step. This is a rare occurrence, but is not unprecedented.<sup>30,31</sup> In addition to the kinetic isotope effects, detailed analysis of the

pH dependence of the reaction catalyzed by STS shows multiple different acid and base catalyzed steps, thought to be the different proton transfers mediated by Glu309.<sup>28</sup> The current mechanistic proposal for the mechanism of STS is shown in Scheme 1.3.

Scheme 1.3 – Mechanism of the enzymatic Pictet-Spengler reaction catalyzed by STS.



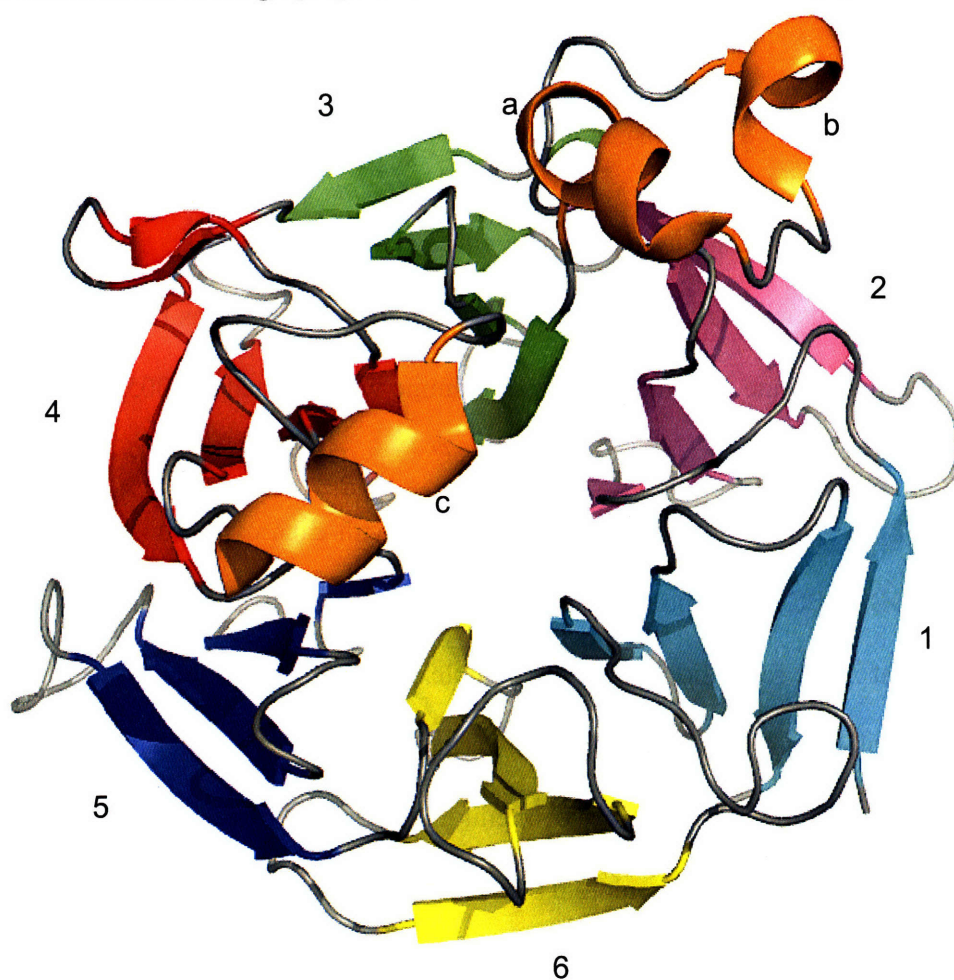
### 1.3.2 Crystal structure of *R. serpentina* strictosidine synthase

The crystallization of strictosidine synthase from *R. serpentina* was first reported in 2004, crystallization of the enzyme in complex with tryptamine was reported in 2005, and the structure of *R. serpentina* STS (rsSTS) was reported in 2006.<sup>29, 32-33</sup> Three structures were reported – the native protein (PDB 2FP8) and the protein in complex with each of its substrates individually (PDB 2FPB, PDB 2FPC for tryptamine and secologanin, respectively). The structure of the native protein is at 2.3 Å resolution, while the complexes with tryptamine and secologanin are at 2.8 and 3.0 Å resolution, respectively. Phases were obtained using Se-Met derived protein and MAD phasing. Due to the small number of methionine residues present in rsSTS (only two in the native protein), suitable phases could only be obtained after several Leu and Ile residues were mutated to Met. The best results were obtained from a 4-Met derivative (L116M, I190M).

The crystal structure of rsSTS revealed a six-bladed four-stranded  $\beta$ -propeller fold, shown in Figure 1.4.<sup>29</sup> The six blades of the propeller are roughly arranged in a six-fold axis around the center of the protein. The blades are composed of four twisted antiparallel  $\beta$ -sheets.

The structure is primarily  $\beta$ -sheet, but there are a few small  $\alpha$ -helices, one of which is joined by a disulfide bond between two cysteine residues. There are two molecules in the asymmetric unit and the interfacial area between the two molecules is  $1937 \text{ \AA}^2$ . This contact surface is fairly substantial and may be suggestive of a dimer, however there is evidence from size exclusion chromatography that STS is stable and acts as a monomer in solution.<sup>22</sup>

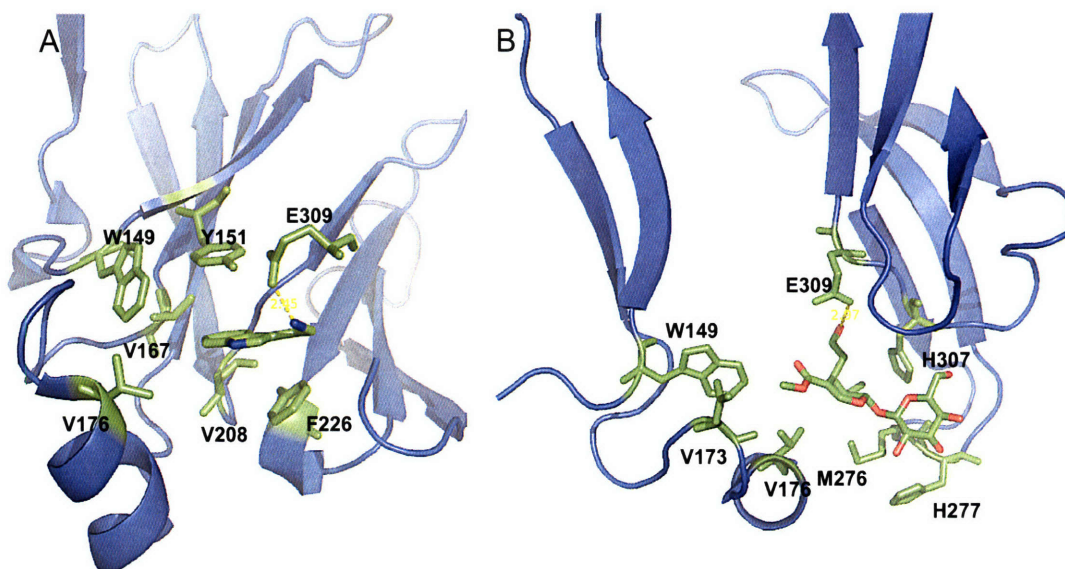
Figure 1.4 – X-ray crystal structure of STS. Each  $\beta$ -sheet is shown in a different color (1-6);  $\alpha$ -helices are shown in orange (a-c). The active site is in the central cavity (PDB 2FP8).



The active site of rsSTS is found in the center of the rough barrel formed by the six  $\beta$ -propellers, and is a pocket lined mostly with hydrophobic residues. Tryptamine binds deeper in the pocket than secologanin with the amino group positioned proximally to Glu309, one of the

few hydrophilic residues and the essential catalytic residue involved in acid/base chemistry in the active site. In addition to being surrounded by hydrophobic residues, the aromatic indole ring of tryptamine is surrounded by the aromatic rings of Phe226 and Tyr151. These aromatic residues may help to stabilize the positively charged intermediates formed during the course of the reaction. Secologanin binds in an extended conformation, with the aldehyde pointing in to the pocket located near the catalytic Glu309. The terpenoid ring is in the middle and the glucose is sticking out of the pocket. The glucose is solvent accessible and makes several important contacts with His307.

1.5 – Active site contacts made by substrates. Tryptamine binds deeper in the pocket than secologanin. a) Tryptamine binding (PDB 2FPB) b) Secologanin binding (PDB 2FPC)



This initial structure suggested important roles for several active site residues in binding and catalysis. These roles were probed by making mutants at the residues in questions and performing steady-state kinetic analysis of the resulting enzymes.<sup>29</sup> Three mutants were made – Y151F, H307A, and E309A. As expected, the Y151F mutant showed decreased affinity for tryptamine (~3 fold less than wild type), unchanged affinity for secologanin, and only slightly reduced catalysis. The H307A mutant showed unchanged affinity for tryptamine, but



dramatically reduced affinity for secologanin (~130 fold less than wild type) and catalysis. The drop in catalysis may be due to conformation changes in the active site caused by making such a substantial mutation so close to the catalytic residue. Finally, the E309A mutant showed unchanged affinity for tryptamine, slightly decreased affinity for secologanin, and almost no catalytic activity. This makes sense in light of its validated mechanistic role in the acid/base catalysis of STS.<sup>28</sup>

One of the ways of further probing the active site of enzymes is through crystallization and analysis of structures with substrate analogs or inhibitors bound. Such an inhibitor was synthesized by reducing the iminium intermediate formed in the non-enzymatic Pictet-Spengler reaction between tryptamine and

secologanin using sodium cyanoborohydride.<sup>28</sup> The crystal structure of the inhibitor bound complex shows that the glucose ring and the terpenoid ring overlay fairly well with the secologanin bound structure (PDB 2VAQ).<sup>28</sup> However, the indole ring of the inhibitor is twisted almost 90° with

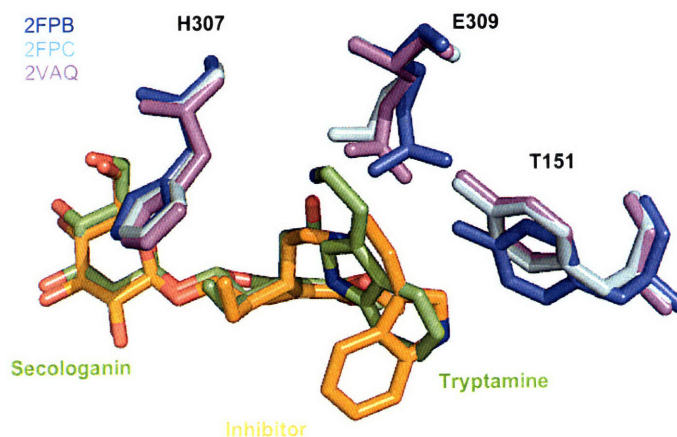


Figure 1.6 – Inhibitor bound structure. The overlay shows the tryptamine structure in blue, the secologanin structure in cyan, and the inhibitor structure in lavender. Secologanin and tryptamine are in green, and the inhibitor is in orange.

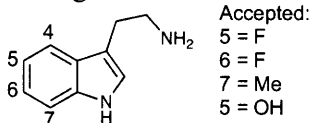
respect to the native tryptamine structure, and neither of the nucleophilic indole carbons is in the proper position for cyclization.<sup>28</sup> An overlay of the inhibitor structure with the tryptamine and secologanin bound structures is shown in Figure 1.6. It appears that the binding pocket of STS can accept substrate in a variety of different binding modes, not all of which are catalytically

competent. This is an intriguing result, and suggests that further structural characterization with different inhibitors/substrates bound could provide information pertaining to the scope of different binding conformations allowed in strictosidine synthase.

### 1.3.3 Obtaining modified monoterpene indole alkaloids

As previously discussed, there are several different ways to obtain biosynthetically derived modified natural products. Two of these approaches, precursor directed biosynthesis (PDB) and mutational biosynthesis (MBS) have yielded promising results for obtaining modified monoterpene indole alkaloids from *C. roseus*. Initial efforts towards PDB focused on determining the substrate scope of STS.<sup>34</sup> As the first committed step in MIA biosynthesis, it was possible that an “unnatural” substrate accepted by this enzyme could be accepted by downstream biosynthetic enzymes to produce complex modified MIAs. Results from this work showed that for tryptamine, indole analogs (benzofurans and benzothiophenes) could be accepted, as could a variety of substitutions on the benzene ring of the indole. For secologanin, several different esters were accepted, but no other changes were tolerated. Kinetic

Figure 1.7 – Tryptamine analogs used in PDB.



from equivalent to native substrate to orders of magnitude worse, as well as a wide range of activities and catalytic efficiencies. These promising results were followed by feeding studies in which the modified tryptamine substrates were fed to *C. roseus* cell cultures. These feeding studies showed that numerous unnatural substrates could be incorporated into downstream steps, including those shown in Figure 1.7 albeit with varying degrees of incorporation based on the particular tryptamine derivative used.<sup>35</sup>

In addition to work towards PDB of modified alkaloids, progress has been made towards the biosynthesis of modified alkaloids by MBS. Several mutants have been made of *C. roseus* crSTS that feature altered or expanded substrate specificity. One such mutant is crSTS D177A (*C. roseus* numbering used; subtract six to obtain the *R. serpentina* numbering), which has been found to accept a pentynyl secologanin derivative with greater catalytic efficiency than the natural substrate.<sup>36</sup> This mutant was identified from a rationally generated library of active site mutants. Several different ester derivatives were tested, but the best results were achieved for the pentynyl ester. This

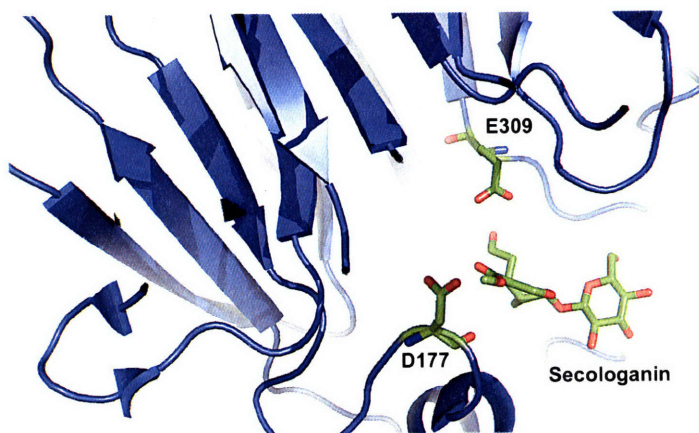


Figure 1.8 – Residue D177 was mutated, allowing STS to accept secologanin ester derivatives. The ester points out towards D177, while the aldehyde projects back into the active site towards E309.

is especially useful as a tag for further derivatization through copper-catalyzed cycloadditions between azides and alkynes, or “click” chemistry.<sup>37</sup> This can be used to add new functionality to alkaloid products, or to add affinity purification tags to small molecules (i.e. biotin).

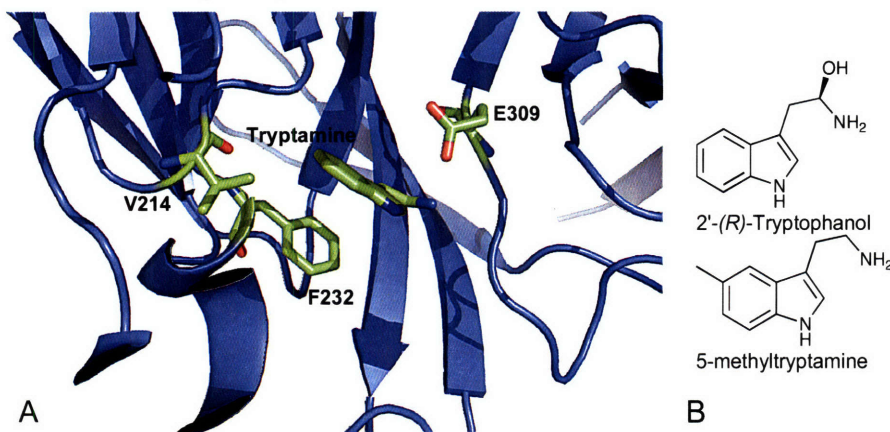
Other studies have shown that the substrate specificity of crSTS can be further altered. Several mutants with expanded substrate scopes were identified from a saturation-mutagenesis library of active site residues.<sup>38</sup> The residues to be mutagenized were selected based on their position in the crystal structure of rsSTS. Following expression and purification of the mutants, a high-throughput assay identified variants that accepted unnatural substrates. These include F232L and V214M (*C. roseus* numbering), which accept 2’-(R)-tryptanol and 5-methyltryptamine respectively. The wild-type enzyme does not accept either of these



compounds. Kinetic analysis of the mutants showed that they showed much lower affinity for both tryptamine and the modified substrates than the wild type enzyme showed for tryptamine.

Also, the mutants were not as catalytically active as wild type (with either tryptamine or the

Figure 1.9 – a) STS active site with mutated residues; b) Unnatural substrates accepted by STS mutants



unnatural substrate).<sup>38</sup> Despite these limitations, these mutants provide the potential for incorporating even more modifications into

the MIA biosynthetic pathway. Transgenic plant cell cultures including these mutants should be able to produce modified alkaloids.

#### 1.4 Benzyloisoquinoline alkaloids

The benzyloisoquinoline alkaloids comprise another diverse and medically important class of alkaloid natural products, a selection of which are shown in Figure 1.10. Benzyloisoquinoline alkaloids (BIAs) are primarily produced by five different plant families, including *Papaveraceae*, *Fumariaceae*, *Ranunculaceae*, *Berberidaceae*, and *Menispermaceae*.<sup>39</sup> One of the most important and most well studied of these is the opium poppy (*Papaver somniferum*). In addition to morphine, a molecule for which the medical and historical importance cannot be understated, this plant produces the analgesic codeine, the muscle relaxant papaverine, the antibiotic sanguinarine, and the antineoplastic noscapine.<sup>39</sup> The number of biologically active molecules isolated from this single species is staggering. Another benzyloisoquinoline alkaloid of

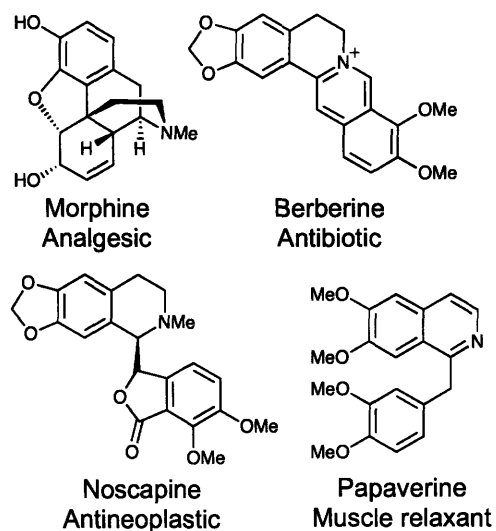
note is berberine, a potent antimicrobial compound.<sup>40</sup> Outside of the five main plant families, there are several other species that have been found to produce BIAs in appreciable amounts.

Most interesting among these is *Thalictrum flavum*, common meadow rue, which has been found to produce berberine and a potential anti-HIV drug magnoflorine.<sup>41</sup> As with MIAs, the benzyloquinoline alkaloids are classified on the basis of the structural feature that the final products exhibit, which can be very different. A clear contrast of these different structural classes is drawn between morphine and berberine, seen in Figure 1.10.

#### 1.4.1 Biosynthesis of benzyloquinoline alkaloids

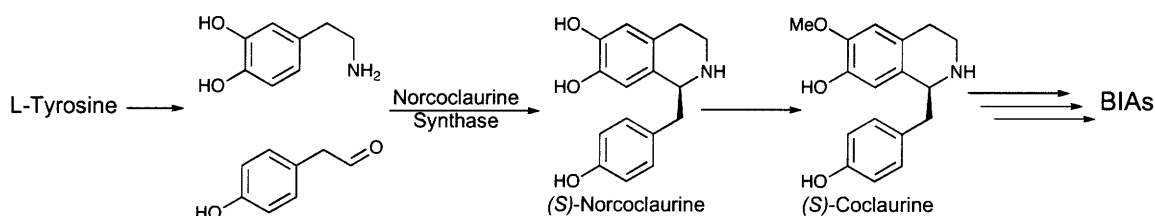
Elucidating the biosynthesis of BIAs shares the same challenges as studying any other plant biosynthetic pathway. However, due to the historical importance of morphine and the other medicinally relevant BIA alkaloids, much is known about these pathways. The first committed step in BIA biosynthesis is a Pictet-Spengler reaction between dopamine (the aromatic amine) and 4-hydroxyphenylacetaldehyde (4-HPAA, the aldehyde).<sup>42</sup> Both of these precursors are derived from the amino acid tyrosine by a series of decarboxylation and oxidation reactions.<sup>43</sup> This reaction is stereoselective, and is catalyzed by the enzyme norcoclaurine synthase (NCS) to give the product (*S*)-norcoclaurine. Following this reaction, a series of methylations and oxidations occur to give a variety of different intermediate molecules.<sup>3</sup> The biosynthetic pathway of papaverine branches off early in the sequence from the intermediate (*S*)-coclaurine. After several more steps, the berberine-type BIAs and morphine-type BIAs diverge from the

Figure 1.10 – Medicinally important benzyloquinoline alkaloids.



intermediate reticuline. Interestingly, berberine synthesis proceeds from the (*S*)-reticuline enantiomer, while morphine biosynthesis proceeds from the (*R*)-reticuline form.<sup>3</sup> In the morphine biosynthetic pathway, two coupling reactions between the two aromatic rings produce the intermediate thebaine, which is further converted into codeine and morphine.<sup>3</sup>

Scheme 1.3 – Early steps of BIA biosynthesis.



There are several parallels that can be drawn between monoterpene indole alkaloid and benzyloquinoline alkaloid biosynthesis. Both biosynthetic pathways begin with a stereoselective, enzyme catalyzed Pictet-Spengler reaction in the first committed step. In one step, a highly functionalized molecule is synthesized for transformation into a wide range of different product structures. Also, both pathways are highly divergent, producing a large number of compounds with many different structural types. However, while the (*S*)-norcoclaurine molecule produced by NCS is quite reactive, featuring two electron rich aromatic rings, there is no reactive intermediate similar to that produced upon the deglycosylation of strictosidine. The relationship between these two pathways illustrates both the parallels and the differences among plant alkaloid natural product biosynthetic pathways.

## 1.5 Norcoclaurine synthase

### 1.5.1 Multiple forms and their relation to strictosidine synthase

The structure of the common precursor to benzyloquinoline biosynthesis was not identified until recently, hampering the identification of the enzyme catalyzing this transformation. Norcoclaurine synthase was first isolated from cultures of *Eschscholzia*

*tenuifolia*, but its ability to accept incorrect substrates (3,4-dihydroxyphenylacetaldehyde in addition to 4-HPAA) led to the incorrect assignment of the common precursor.<sup>44</sup> Once the correct precursor molecule of (*S*)-norcoclaurine was identified in the early 1990s, this mistake was apparent.<sup>42</sup> However, little work was done on NCS throughout this decade. More detailed studies of this enzyme, including isolation from several plant species, activity assays, and kinetic characterization, were published in 2001 and 2002.<sup>39,45</sup> The enzyme was initially purified from *Thalictrum flavum*, and as such will be referred to as tfNCS. tfNCS is a ~25 kDa protein that feature a 19 amino acid N-terminal signal peptide with no associated metals or cofactors. The gene has been cloned and the protein has been expressed heterologously in *E. coli* (truncated to remove the signal sequence).<sup>46</sup> No sequence homology has been detected between tfNCS and strictosidine synthase (from any species), however tfNCS does have 50-60% homology with members of class 10 pathogenesis related proteins, which are a class of pollen and food allergens.<sup>47</sup> The relationship between tfNCS and STS appears to represent an example of convergent evolution.

In addition to enzymes characterized from *T. flavum*, enzymes involved in the initial steps of BIA biosynthesis have been characterized from another plant species, *Coptis japonica*.<sup>48</sup> In this work, screening of expressed sequence tag (EST) clones in a cell culture line selected for enhanced BIA production revealed a new form of NCS, bearing no homology to tfNCS or STS. This new enzyme, termed cjNCS, was isolated from cell cultures, purified to homogeneity through chromatographic methods, and sequenced. Following sequencing, comparison of the cjNCS sequence to other Pictet-Spenglerase sequences showed no homology. However, this enzyme did show homology to 2-oxoglutarate dependant non-heme iron dioxygenases. Interestingly, assays showed that the enzyme did not require oxygen or 2-oxoglutarate to

function, but assays performed in the presence of the chelator *o*-phenanthroline suggested that Fe<sup>2+</sup> is required for activity.<sup>48</sup> These results were rationalized in view of a sequence alignment of cjNCS with other plant 2-oxoglutarate-dependent non-heme iron enzymes, which showed that the iron binding residues of cjNCS were conserved, but the 2-oxoglutarate binding region, conserved in all other enzymes of this class, was not conserved in cjNCS.

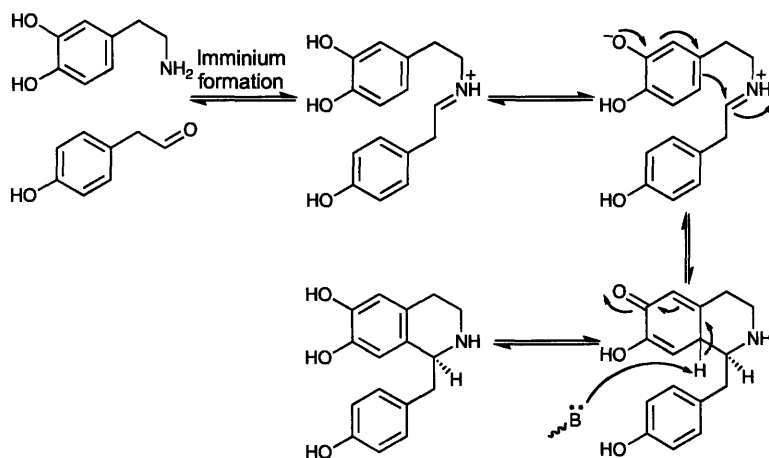
Along with cjNCS, another enzyme was identified from *C. japonica* on the basis of sequence homology with tfNCS and class 10 pathogenesis related proteins.<sup>48</sup> This enzyme, called cjPR10A, was also demonstrated to have Pictet-Spenglerase activity, although it was not up-regulated in the BIA producing cell cultures. cjPR10A was shown to have an expanded substrate scope compared to cjNCS, accepting 4-hydroxyphenylpyruvate in addition to the native substrate 4-HPAA.<sup>48</sup> Also, the product produced by cjPR10A had the same mass as that produced by cjNCS, but a different elution time by LC/MS. Finally, cjPR10A was found to have an N-terminal signal sequence like tfNCS, while cjNCS showed no such signal sequence. There are still several unresolved issues raised by this work, but further studies of a Pictet-Spenglerase like cjNCS with a potentially novel mechanism may lead to a broader understanding of the requirements for enzyme-catalyzed stereoselective Pictet-Spengler reactions.

### **1.5.2 Reaction mechanism of *T. flavum* norcoclaurine synthase**

Recently, a detailed mechanistic study has illuminated the mechanism of the Pictet-Spengler reaction catalyzed by tfNCS.<sup>49</sup> This work used alternative substrates and kinetic isotope effects to differentiate between two possible reaction mechanisms. Also, analysis of substituent effects helped to support the mechanistic conclusions. While this proposed mechanism is similar to that proposed for STS, the hydroxyl groups present on the aromatic ring of dopamine allow for the reaction intermediate to be charge neutral. Also similar to STS, the

deprotonation of the reaction intermediate appears to be the rate-limiting step. These seem to be common features for enzyme catalyzed Pictet-Spengler reactions. Also of note, this study determined that dopamine exhibited positive cooperative binding, with a Hill coefficient of 1.8. This supports the previous evidence that tfNCS functions as a dimer in solution.<sup>45</sup>

### 1.5.3 NMR based homology model of *T. flavum* norcoclaurine synthase



Scheme 1.4 – Proposed reaction mechanism of NCS.

Recently, a report describing the structural characterization of tfNCS was published.<sup>50</sup> A truncated version of the protein missing the N-terminal signal sequence was used; in total, 29 amino acids were removed from the N-terminus. Using NMR

spectroscopy, a semi-experimental homology model was constructed based on homology with the major birch pollen allergen Betv1. Betv1 was selected based on the significant homology between this protein and tfNCS, and the availability of an X-ray crystal structure.<sup>51</sup> CD spectroscopy verified that Betv1 and tfNCS had similar secondary structure and thermal stability properties, supporting the choice of this protein as a starting point for homology modeling.

Following the expression and purification of isotopically labeled protein, NMR measurements yielded signals for 86% of the tfNCS sequence. In addition to native measurements, the protein was also studied in the presence of substrates to attempt to localize the binding pocket. To investigate 4-HPAA binding, several different analogs were employed

including methyl(4-hydroxyphenyl)acetate and 2-(4-hydroxyphenyl)ethanol. Upon inclusion of these analogs, widespread changes in chemical shifts were observed, suggesting that 4-HPAA binding induces a significant structural rearrangement. For dopamine binding, a much smaller set of residues exhibited chemical shift changes. These included residues 69-72 and 150-155. Within these regions, several hydrophobic residues may be involved in interactions with the aromatic ring of dopamine, including F71 and M155. Near the putative dopamine binding site, there is one acidic residue, E75, that may play the same role in tfNCS as E309 plays in STS, namely that of a general acid/base catalyst. This seems plausible in light of the mechanistic studies showing similar mechanisms for these two enzyme catalyzed reactions.

While many new insights can be drawn from this structural characterization, there are still numerous unanswered questions. For instance, what are the specific interactions involved in substrate binding? What are the structural rearrangements caused by 4-HPAA binding? It has been shown that the two substrates show ordered binding, with 4-HPAA binding first.<sup>45</sup> How do the structural changes affect dopamine binding and catalysis? Is E75 really poised to be a catalytic residue, or is it just in the vicinity of the binding pocket without playing a role in the reaction? Further structural characterization would aid in answering these questions, in addition to suggesting residues for mutation for enzyme engineering.

## **1.6 Summary and goals**

Natural products are a rich source of bioactive molecules used for treating a wide range of human diseases. Alkaloids compose one of the largest and most structurally diverse classes of natural products. Two sub-classes of plant-derived alkaloids, the terpene indole alkaloids and the benzyloisoquinoline alkaloids, have provided several well-known therapeutics that are widely used today. The structural diversity and complexity of these compounds makes chemical

synthesis difficult and expensive. As such, the synthesis of analogs that may have modified or improved biological activity and the eventual large-scale production of these potentially improved compounds present a significant challenge.

An attractive solution is the production of modified alkaloids by utilizing mutational biosynthesis, or mutasynthesis. This methodology entails altering the enzymes in a biosynthetic pathway, feeding a reengineered producing organism a modified substrate, and isolating the “unnatural” product. Mutational biosynthesis requires a detailed understanding of the enzymes that function in alkaloid biosynthetic pathways, and an understanding of how mutations to these enzymes affect their function and substrate specificity. Structural characterization of enzymes involved in alkaloid biosynthesis by x-ray crystallography is critical for this task, as is structural characterization of mutants with expanded or altered substrate specificities.

The first committed step in terpene indole alkaloid biosynthesis is catalyzed by strictosidine synthase, an enzyme that catalyzes a Pictet-Spengler reaction coupling tryptamine and secologanin to form strictosidine. Several STS mutants have been identified with expanded substrate specificity, accepting tryptamine analogs that are not turned over by the wild type enzyme. One STS mutant has been found with altered substrate specificity, accepting a secologanin analog preferentially over the natural secologanin substrate. Understanding how these mutations confer different binding activities at atomic resolution will allow for catalysis of substrate analogs by maximizing productive binding modes through rational mutagenesis.

Analogous to terpene indole alkaloid biosynthesis, the first step in benzyloisoquinoline alkaloid biosynthesis is the Pictet-Spengler coupling of dopamine and 4-HPAA to form norcoclaurine catalyzed by norcoclaurine synthase. Interestingly, there is no sequence homology between STS and NCS, even though they catalyze the same reaction. Even more interesting is



the fact that the two enzymes (cjNCS and tfNCS) known to carry out norcoclaurine synthesis share no sequence homology. One of these enzymes (cjNCS) shows similarity to 2-oxoglutarate-dependant oxygenases and dependence on ferrous iron, but not 2-oxoglutarate or oxygen. Structural characterization of cjNCS, and further structural characterization of tfNCS, should provide insight into the mechanism, as well as suggest residues for mutation to change the substrate specificity of this enzyme and allow mutational biosynthesis of benzyloquinoline alkaloids.

## 1.7 References

1. Newman DJ, Cragg GM. (2007) *J. Nat. Prod.* 70, 461-477.
2. Reayi A, Arya P. (2005) *Curr. Opin. Chem. Biol.* 9, 240-247.
3. Dewick, PM. (2001) *Medicinal Natural Products: A Biosynthetic approach*. John Wiley and Sons, Ltd., Hoboken, NJ.
4. Kirschning A, Taft F, Knobloch T. (2007) *Org. Biomol. Chem.* 5, 3245-3259.
5. Rinehart KL. (1977) *Pure Appl. Chem.* 49, 1361-1384.
6. Cordell, GA. (1998) *The Alkaloids: Chemistry and Biology*. Academic Press, San Diego.
7. Tyler, VE. (1994) *Herbs of Choice*. Haworth Press, Inc., New York.
8. O'Connor SE, Maresh JM. (2006) *Nat. Prod. Rep.* 4, 532-547.
9. Leonard, J. (1999) *Nat. Prod. Rep.* 16, 319-338.
10. Saxton, EJ. (1997) *Nat. Prod. Rep.* 14, 559-590.
11. Huang Y, Fang Y, Wu J, Dziadyk JM, Zhu X, Sui M, Fan W. (2004) *Mol. Cancer Ther.* 3, 271-277.
12. Makarov, AA, Tsvetkov, PO, Villard C, Esquieu D, Pourroy B, Fahy J, Braguer D, Peyrot V, Lafitte D. (2007) *Biochemistry.* 46, 14899-14906.
13. Hashimoto T, Yamada Y. (2003) *Curr. Opin. Biotechnol.* 14, 163-168.
14. Treimer JF, Zenk MH. (1979) *FEBS Lett.* 97, 159-162.
15. Leete E. (1961) *Tetrahedron.* 14, 35-41.
16. Contin A, van der Heijden R, Lefeber AWM, Verpoorte R. (1998) *FEBS Lett.* 434, 413-416.
17. Gerasimenko I, Sheludko Y, Ma X, Stoeckigt J. (2002) *Eur. J. Biochem.* 269, 2204-2213.
18. Heinstein P, Holfe G, Stoeckigt J. (1979) *Planta Med.* 37, 349-357.
19. Lorence A, Nessler CL. (2004) *Phytochemistry.* 65, 2735-2749.
20. Kutchan TM. (1993) *Phytochemistry.* 32, 493-506.
21. Pfitzner U, Zenk MH. (1989) *Planta Med.* 55, 525-530.
22. Hamp N, Zenk MH (1988) *Phytochemistry.* 27, 3811-3815.
23. Kutchan TM, Hamp N, Lottspeich F, Beyreuther K, Zenk MH. (1998) *FEBS Lett.* 237, 40-44.
24. McKnight TD, Roessner CA, Devagupta R, Scott AI, Nessler CL. (1990) *Nucl. Acids Res.* 18, 4939.
25. Kutchan TM, Dittrich H, Bracher D, Zenk MH. (1991) *Tetrahedron.* 47, 5945-5954.
26. Cox ED, Cook JM. (1995) *Chem. Rev.* 95, 1797-1842.
27. Pictet A, Spengler T. (1911) *Ber. Dtsch. Chem. Ges.* 44, 2030-2036.
28. Marash JJ, Giddings L, Friedrich A, Loris EA, Panjekar S, Trout BL, Stoeckigt J, Peters B, O'Connor SE. (2008) *J. Am. Chem. Soc.* 130, 710-723.
29. Ma X, Panjekar S, Koepke J, Loris EA, Stoeckigt J. (2006) *The Plant Cell.* 18, 907-920.
30. Jackson AH, Lynch PP. (1987) *J. Chem. Soc., Perkin Trans. II.* 1483-1488.
31. Zollinger, H. (1955) *Helv. Chim. Acta.* 38, 1617-1631.
32. Ma X, Koepke J, Fritzs G, Diem R, Kutchan TM, Michel H, Stoeckigt J. (2004) *Biochim. Biophys. Acta.* 1702, 121-124.
33. Koepke J, Ma X, Fritzs G, Michel H, Stoeckigt J. (2005) *Acta Cryst.* D61, 690-693.
34. McCoy EM, Galan MC, O'Connor SE. (2006) *Bioorg. Med. Chem. Lett.* 16, 2475-2478.
35. McCoy EM, O'Connor SE. (2006) *J. Am. Chem. Soc.* 128, 14276-14277.
36. Chen S, Galan MC, Coltharp C, O'Connor SE. (2006) *Chemistry and Biology.* 13, 1137-1141.

37. Kolb HC, Sharpless KB. (2003) *Drug Discov. Today*. 8, 1128-1138.
38. Bernhardt P, McCoy EM, O'Connor SE. (2007) *Chemistry and Biology*. 14, 888-897.
39. Samanani N, Facchini PJ. (2001) *Planta*. 213, 898-906.
40. Birdsall TC, Kelly GS. (1997) *Alt. Med. Rev.* 2, 94-103.
41. Rashid MA, Gustafson KR, Kashman Y, Cardellina JH, McMahon JB, Boyd MR. (1995) *Nat. Prod. Lett.* 6,153-156.
42. Stadler R, Kutchan TM, Zenk M. (1989) *Phytochemistry*. 28, 1083-1086.
43. Rueffer M, Zenk MH. (1987). *Z. Naturforschteil. C*. 42, 319-332.
44. Rueffer M, El-Shagi H, Nagakura N, Zenk MH. (1981) *FEBS Lett.* 129, 5-9.
45. Samanani N, Facchini PJ. (2002) *J. Biol. Chem.* 277, 33878-33883.
46. Samanani N, Liscombe DK, Facchini PJ. (2004) *Plant J.* 40, 302-313.
47. Ebner C, Hoffmann-Sommergruber K, Breiteneder H. (2001) *Allergy*. 56 suppl. 67, 43-44.
48. Minami H, Dubouzet E, Iwasa K, Sato F. (2007) *J. Biol. Chem.* 282, 6274-6282.
49. Luk LYP, Bunn S, Liscombe DK, Facchini PJ, Tanner ME. (2007) *Biochemistry*. 46, 10153-10161.
50. Berkner H, Schweimer K, Matcko I, Rosch P. (2008) *Biochem. J.* Immediate publication, doi: 10.1042/BJ20080306.
51. Gajhede M, Osmark P, Poulsen FM, Ipsen H, Larsen JN, Joost van Neerven RJ, Schou C, Lowenstein H, Spangfort MD. (1996) *Nat. Struct. Biol.* 3, 1040-1045.

## **2. Strictosidine Synthase Crystallization Studies**

### **2.1 *C. roseus* strictosidine synthase**

Strictosidine synthase (STS) is the first committed enzymatic step towards making the wide variety of terpene indole alkaloid natural products.<sup>1</sup> As the first committed step, it serves as a gateway to the monoterpene indole alkaloids; modified starting materials for use in precursor directed biosynthesis or mutational biosynthesis must pass through STS before they can proceed to more complex alkaloid structures. Efforts towards precursor directed biosynthesis have shown that STS has some degree of promiscuity, however the scope of substrate analogs accepted will ultimately limit the modified alkaloids that can be produced.<sup>2,3</sup> Mutational biosynthesis provides a way around this limitation, providing the potential for an even broader spectrum of modified natural products.<sup>4</sup>

To identify mutants of STS with potentially expanded substrate specificity, a saturation mutagenesis library of relevant residues (as identified in the crystal structure of *R. serpentina* STS) was constructed.<sup>5</sup> Following expression and assay, several mutants were found which could accept substrate analogs unaccepted by native STS. Most notable among these analogs are V214M, F232L, and D177A.<sup>5,6</sup> Structural characterization of these analogs will provide insight into why these particular mutations accept these analogs, as well as provide rationale for further mutagenesis studies. Additionally, crystallography of the native enzyme with substrate analog inhibitors will allow us to probe the active site further and elucidate productive and unproductive binding conformations.

#### **2.1.1 Kinetic characterization of wild-type STS**

Previous experience in the O'Connor lab has shown that the activity of crSTS is highly dependent on the specific construct. For example, the native protein has an N-terminal signal

sequence that functions in cellular localization in the producing organisms. To obtain active protein from *E. coli*, the protein must be expressed as a truncated form without this leader peptide sequence. Also, several different constructs featuring a maltose binding protein (MBP) affinity purification tag at the N-terminus showed strongly reduced activity. For the purposes of crystallography, a much smaller affinity tag was desired. Because of the reduction of activity associated with substitution at the N-terminus, a C-terminal His-tag was selected. Since this was the first C-terminal His-tagged construct to be expressed and purified, its activity was verified with a brief kinetic characterization. Assays were only performed once so an analysis of error is not possible, however the results are consistent with previous kinetic characterizations of crSTS.<sup>7,8</sup>

crSTS was cloned, expressed, and purified as detailed in the Materials and Methods section. To determine the steady-state kinetic parameters for crSTS, assays were run in which the concentration of tryptamine was varied under pseudo-first order conditions ( $[\text{secologanin}] \gg K_{m,\text{sec}}$ ). Previous kinetic characterization of crSTS showed that the  $K_{m,\text{trp}}$  to be between 2-10  $\mu\text{M}$ . In light of this, the following tryptamine concentrations were used: 1, 2.5, 5, 10, 50  $\mu\text{M}$ .

Data were analyzed using the following formula:

$$\left(\left(\frac{A_{\text{strict}}}{A_{\text{NAA}}}\right) * A_{\text{NAA(Avg.)}}\right) / \epsilon_{\text{strict}} / \text{Time} = (v_o \mu\text{M}/\text{min}) / [\text{crSTS}] = k_{\text{cat}} (\text{min}^{-1})$$

$A_{\text{strict}}$  = integrated area under the curve for strictosidine

$A_{\text{NAA}}$  = integrated area under the curve for NAA

$A_{\text{NAA(Avg.)}}$  = average area under the curve for all NAA measurements

$\epsilon_{\text{strict}}$  = extinction coefficient for strictosidine at 280 nm =  $12000 \mu\text{M}^{-1}\text{cm}^{-1}$

$v_o$  = initial rate

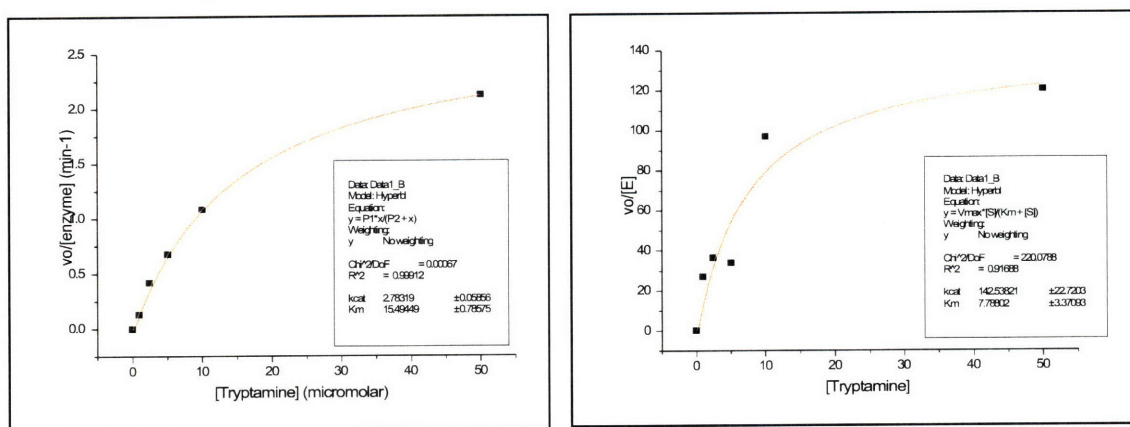
Time = time of quench in minutes

Plotting  $v_o$  vs. [tryptamine] produced a plot characteristic of Michaelis-Menton kinetics.

Non-linear regression was performed using the Origin software package (version 7.0) to give kinetic parameters.

Kinetic characterization was performed twice using two different preparations of crSTS. The first was done using partially purified (Ni-NTA) crSTS, the second was done using electrophoretically homogenous crSTS following size exclusion chromatography. The non-linear regression plots for these two trials can be seen in Figure 2.1. The different kinetic parameters can be seen in Table 2.1.

Figure 2.1 – Kinetic characterization of crSTS. The left plot shows Ni-NTA purified kinetics, the right plot shows pure kinetics.



	Ni-NTA purified	SEC purified
$k_{cat}$ (min <sup>-1</sup> )	27.8	142.5
$K_m$ , tryptamine (μM)	15.5	7.7
$k_{cat}/K_m$ (min <sup>-1</sup> μM <sup>-1</sup> )	1.8	18.5

As expected, the kinetic parameters were shown to improve upon further purification of crSTS. Data analysis for the fully purified crSTS was complicated by the fact that in some HPLC traces there was an unidentified peak that partially obscured the NAA standard. This peak did not show time dependence, nor did it seem to be related to the concentration of any of the reaction components. In some traces, the peak was distinct enough to be integrated separately, however in others it overlapped with the NAA peak to such a degree that independent integration was impossible. It is hypothesized that this peak could be due to polymerization of un-reacted

secologanin, but no characterization was done to investigate this. Because the kinetic parameters for fully purified crSTS were similar to those of the native enzyme and the main aim of this project is structural and not kinetic, these results were deemed satisfactory.

## 2.2 Crystallization experiments with *C. roseus* STS mutants

The crSTS mutants V214M, F226L, and D177A were obtained by site-directed mutagenesis performed on the gene in the pET-28a expression vector. Purification of the mutants proceeded as with the native protein. Following MonoQ chromatography, relevant fractions were collected and analyzed by SDS-PAGE. Even though the protein still contained some contamination, as shown in Figure 2.2, it was decided to attempt crystallization with the current preparation rather than perform further chromatography and potentially reduce the ability of STS to crystallize. Based on the success of subsequent crystallization screens, these mutants should be further purified before crystallization is attempted again. Also, future purifications should not include glycerol unless long-term freezer storage is anticipated.

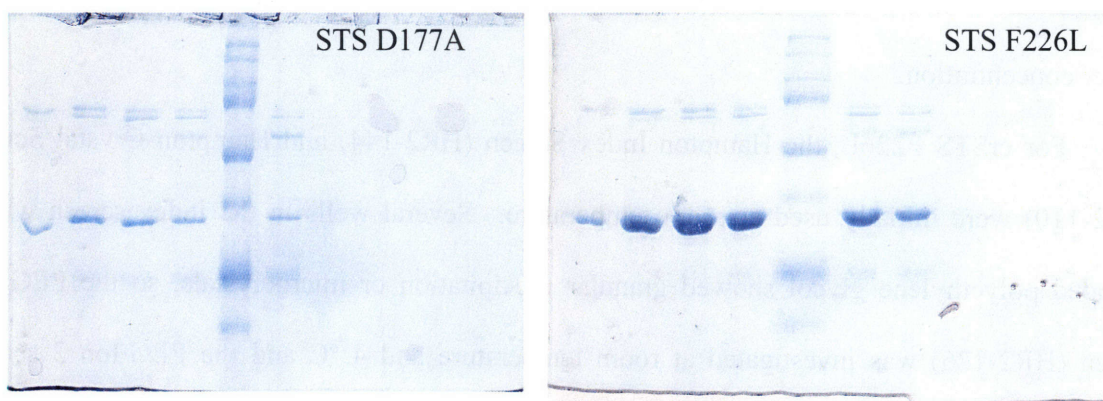


Figure 2.2 – SDS PAGE of crSTS mutants after Mono Q chromatography

Fractions from MonoQ chromatography were pooled and concentrated using an Amicon Ultra Centrifugal Filter device. crSTS F226L was concentrated to 15 mg/mL (from ~ 10 g cell pellet). Repeating the purification using ~30 g cells (wet weight) afforded enough protein to set

up screens at 7.5 mg/mL. Using crSTS F226L and D177A, initial crystallization screens were begun.

### **2.2.1 Sparse matrix screening with D177A and F226L**

All crystallization screens were set up in 24 well plates using siliconized glass cover slides and sealed with vacuum grease. For each well, 500  $\mu$ L of the pre-mixed condition was used as mother liquor. To construct the hanging drop, 1.5  $\mu$ L of protein was mixed with 1.5  $\mu$ L of mother liquor on the glass cover slide and then inverted and sealed on top of the well. Equilibration was allowed to occur, and wells were observed after 2, 4, and 6 days for scoring. Wells were observed at intermittent time periods after the six days, but were not officially scored.

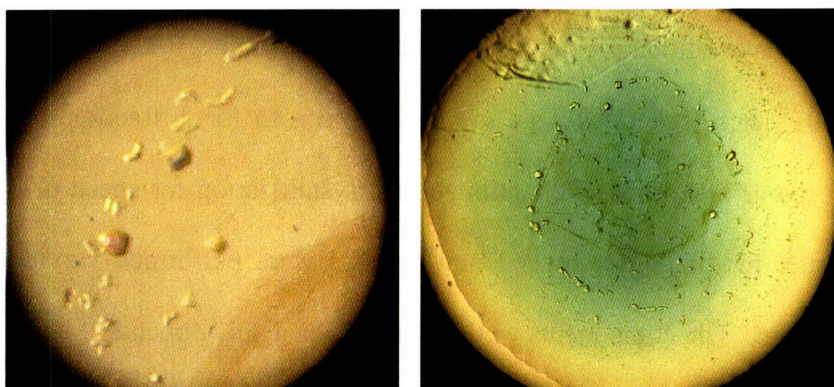
The Hampton Index Screen (HR2-144) was used for crSTS D177A. Numerous drops showed a light precipitate, and many others stayed clear. No hits appeared promising for optimization. Due to the faintness of even the amorphous precipitates, it is thought that the concentration was too low. Further work would include attempting these screens again at a higher concentration.

For crSTS F226L, the Hampton Index Screen (HR2-144) and Hampton Crystal Screen (HR2-110) were initially used at room temperature. Several wells in the index screen which included polyethylene glycol showed granular precipitation or microcrystals, so the PEG/Ion screen (HR2-126) was investigated at room temperature and 4 °C and the PEG/Ion 2 screen (HR2 – 098) was investigated at room temperature. Numerous wells in both screens showed heavy amorphous precipitation, although some drops remained clear for the length of the observation period. There were few granular precipitates, other than those in the PEG containing



conditions. However it seemed that the protein was well behaved and might eventually be optimized to produce crystals.

Figure 2.3 – Pictures of crystals in PEG/Ion condition 1 before and after “Izit” staining. No staining is observed.



A potential hit was seen in PEG/Ion condition 1 after 5 days at both room temperature and 4 °C. Pictures of these crystals are shown in Figure 2.3.

This condition contains 0.2 M NaF, 20% w/v PEG 3350; because of the NaF it is especially prone to forming salt crystals. However, several crystals were picked and analyzed following cryo-protection in 0.2 M NaF, 20% w/v PEG 3350, 20% ethylene glycol. One image was collected over 15 minutes with a 1° oscillation and a detector distance of 300 cm starting at 90°. This was repeated with a 20 minute exposure starting at 180°. The maximum resolution was to 3.3 Å, and very few spots were visible. A representative image may be seen in Figure 2.4. Based on the x-ray diffraction, it was concluded that these crystals were salt.

Despite the x-ray results, the crystals were investigated further. Addition of 1 µL of mother liquor resulted in the formation of new crystals, strongly supporting the case for salt. The crystals showed no

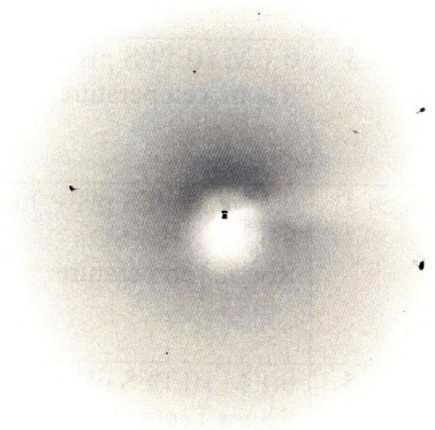


Figure 2.4 – X-ray diffraction of PEG/Ion Condition 1 crystals.

staining with IZIT dye and no protein was visible on SDS-PAGE, indicating that these were salt crystals.

### 2.2.2 Optimizing crystallization of F226L

At this point, first index screens set up were re-examined. Numerous small spherules were observed in condition 35. This condition included 1.0 M ammonium sulfate, 0.1 M HEPES pH = 7.0, 0.5% w/v PEG 8000. These spherules were too small to screen for diffraction, but they were well stained by IZIT dye (Figure 2.5). To optimize the crystallization for speed (1-2 weeks) and quality (larger, better morphology), several new screens were performed varying different components of the mixture. The components were pre-mixed in 1 mL Eppendorf tubes by vortexing. As before, 500  $\mu$ L was used as mother liquor and the drop size was 1.5  $\mu$ L.

Screen #	Constant	Varying	Results
Screen 1	0.1 M HEPES pH = 7.0 Room Temperature	PEG 8K – 0.5-20% (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> – 0.5-2 M	Potential hits within 5 days at 5% and 10% PEG.
Screen 2	0.1 M HEPES pH = 7.0 Room Temperature	PEG 8K – 0.5 – 16% (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> – 0.5-2 M	There is a problem with miscibility above 1 M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> and 2% PEG. The hits from screen one were simply immiscible droplets.
Screen 3	0.1 M HEPES pH = 7.0 Room Temperature	PEG 8K – 0-1.5% (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> – 0-1.25 M	Nothing seen within 1 week. Eventually, small spherules developed in some wells (over ~4 wks).
Screen 4	0.1 M HEPES pH = 7.0 0.5 M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> Room Temperature	PEG conc. – 0.5-20% PEG MW – 3350, 6K, 8K, 20K	After one week, a potential rod was seen in 5% PEG 20K, but it did not stain, was too thin for diffraction, and was not reproducible.
Screen 5	0.1 M HEPES pH = 7.0 Room Temperature	PEG 20K – 0.5-10% (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> – 0.4-0.6 M	No hits seen in an appreciable time period (after 2 weeks)

Future work towards crystallizing crSTS F226L would include repeating the initial screens at lower protein concentrations (i.e. 10 mg/mL). Also, glycerol should not be included in

the crystallization buffer. Finally, different protein buffers could be investigated to see which would produce the least amount of precipitation.

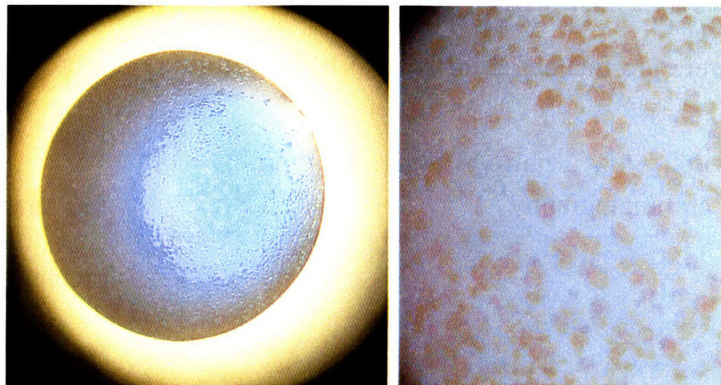


Figure 2.5 – IZIT staining of crSTS F226L Index 35

## 2.3 Crystallization experiments with wild-type *C. roseus* STS

### 2.3.1 Sparse matrix screening for initial conditions

It is essential to verify the activity of a protein before attempting crystallization screens. This ensures that any structural information obtained will be of a relevant form of the protein. A detailed kinetic characterization was not necessary; a qualitative assay was all that was required. To assay crSTS, a previously described colorimetric assay was performed.<sup>5,9,10</sup> The reaction mixture for this assay includes crSTS, both tryptamine and secologanin substrates, and strictosidine glucosidase (SG), the next step in the biosynthetic pathway of the terpene indole alkaloids. Deglycosylation of the strictosidine produced by STS yields a reactive intermediate that undergoes a series of rearrangements to eventually yield a yellow and insoluble precipitate. The yellow chromophore forms only in the presence of active STS when the reaction is incubated at 30 °C. Compared to a no enzyme control, crSTS showed a strong yellow color and substantial precipitation after a 10 minute incubation at 30 °C.

After verifying the activity, electrophoretically homogenous crSTS was used for initial sparse matrix screens. The Hampton Index Screen (HR2-144) was used with the following protein solutions– 10 mg/mL crSTS, 10 mg/mL crSTS + 1 mM tryptamine, 5 mg/mL crSTS, 5

mg/mL crSTS + 1 mM tryptamine. For trials including tryptamine, the protein and tryptamine were pre-incubated for several hours before trays were set up. For screening, the hanging drop

Figure 2.6 – Pure crSTS after SE200 chromatography



vapor diffusion method was used. Briefly, this method features a drop of protein + precipitant solution suspended hanging over a well of mother liquor (usually the precipitant solution). The drop is placed on a siliconized glass cover slip, which is then inverted over the well. The edges are sealed with vacuum grease to create a closed environment. As

water vapor diffuses from the drop into the mother liquor, it slowly concentrates. This increases the concentration of protein and precipitant, ideally resulting in gradual crystal nucleation and growth. All screens were performed at room temperature with 500  $\mu$ L wells and 3  $\mu$ L drops (1.5  $\mu$ L each protein and mother liquor). Trays were observed 2, 4, and 7 days after being set up for scoring.

In general, this protein preparation gave much improved results compared to previous trials with crSTS F226L. While some wells showed heavy amorphous precipitation, the other wells showed a range of results. A large number of wells showed granular precipitation, and many showed small microcrystals. While none of the wells showed crystals in 2 weeks, there seemed to be potential for optimizing crystal formation from one of the many drops with microcrystals. However, due to the large number of wells with microcrystals, it was reasoned that trying to crystallize the protein at 4  $^{\circ}$ C might promote formation of crystals from one of

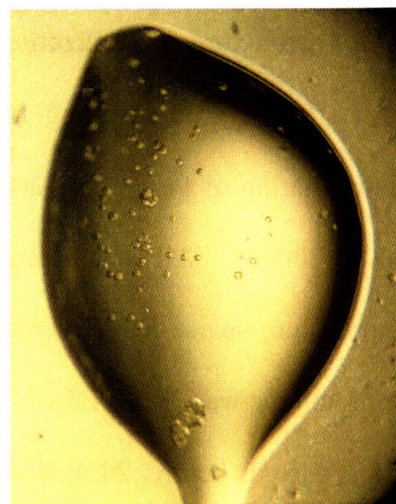


these drops. Before beginning grid screening around an arbitrarily chosen microcrystalline option, the Index screen was investigated at 4 °C.

Before setting up the trays for the reduced temperature Index screen, all buffers, trays, pipette tips, and cover slips were cooled in the cold room for ~3 hours. crSTS at 5 mg/mL was used to set up the trays. The previous trials had shown that 5 mg/mL protein worked as well as 10 mg/mL protein while giving fewer wells with heavy precipitation. Even though previous trials showed no improvement of crystal morphology/size with tryptamine, it was used as in co-crystallization in the hopes that it would give better crystals at 4 °C. After the screens were set up, they were observed at 1, 4, and 6 days. While waiting for results from these screens, Hampton Crystal Screen 1 (HR2-110) was set up at room temperature at 5 mg/mL with and without tryptamine.

On the fourth day, small spherulites were observed in Index screen condition 90 at 4 °C (shown in Figure 2.7). This conditions contained 0.2 M Sodium formate and 20% (w/v) PEG 3350. These crystals were seen in drops with and without tryptamine, however the tryptamine drop showed heavy precipitation on the sixth day while the drop without tryptamine stayed clear (except for the crystals, of course). Since this was the best hit so far, it was selected for optimization by grid screening.

Figure 2.7 – Initial Hit for crSTS (Index #90)



Screen	Concentration (mg/mL)	Tryptamine (+/-)	Temperature	Results
Index	10	-	r.t.	Range of results, from heavy ppt to microcrystals.
Index	10	+	r.t.	Range of results, from heavy ppt to microcrystals; a lot of drops with skin.
Index	5	-	r.t.	Better behaved than more concentrated screen, fewer wells with heavy ppt, more granular ppt and microcrystals.
Index	5	+	r.t.	Same as without tryptamine, but with more drops with skin.
Index	5	-	4 °C	Range of results <b>Hits:</b> Condition #90 (0.2 M HCOONa, 20% w/v PEG 3350)
Index	5	+	4 °C	Range of results <b>Hits:</b> Condition #90 (0.2 M HCOONa, 20% w/v PEG 3350); proceeding to heavy precipitate in several days.
Crystal Screen	5	-	r.t.	Range of results, no utilizable hits.
Crystal Screen	5	+	r.t.	Range of results, no utilizable hits.

### 2.3.2 Optimizing crystallization of wild-type STS

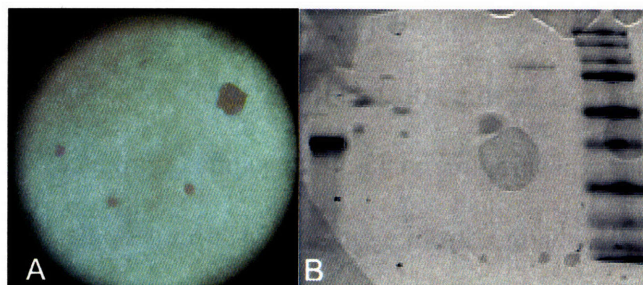
Initial grid screening focused on changing the concentration of sodium formate, the concentration of PEG 3350, and the molecular weight of the polyethylene glycol (3350, 6000, 8000, 20000). All grid screens were set up at 4 °C in the same manner as the reduced temperature sparse matrix screens. As before, the mother liquor volume was 500  $\mu$ L and the drop volume was 3  $\mu$ L (1.5  $\mu$ L mother liquor, 1.5  $\mu$ L 5 mg/mL crSTS in 10 mM Tris-HCl pH = 7.5). Also as before, 24 well plates were used and the wells were sealed with siliconized glass cover slips. An example of one of these sparse matrix screens is shown in Table 2.4.

0.1 M HCOONa 14% PEG3350	0.1 M HCOONa 16% PEG3350	0.1 M HCOONa 18% PEG3350	0.1 M HCOONa 20% PEG3350	0.1 M HCOONa 22% PEG3350	0.1 M HCOONa 24% PEG3350
0.2 M HCOONa 14% PEG3350	0.2 M HCOONa 16% PEG3350	0.2 M HCOONa 18% PEG3350	0.2 M HCOONa 20% PEG3350	0.2 M HCOONa 22% PEG3350	0.2 M HCOONa 24% PEG3350
0.3 M HCOONa 14% PEG3350	0.3 M HCOONa 16% PEG3350	0.3 M HCOONa 18% PEG3350	0.3 M HCOONa 20% PEG3350	0.3 M HCOONa 22% PEG3350	0.3 M HCOONa 24% PEG3350
0.4 M HCOONa 14% PEG3350	0.4 M HCOONa 16% PEG3350	0.4 M HCOONa 18% PEG3350	0.4 M HCOONa 20% PEG3350	0.4 M HCOONa 22% PEG3350	0.4 M HCOONa 24% PEG3350

As a result of multiple grid screens, it was discovered that changing the concentration of sodium formate did not affect crystal formation, but changing the concentration of PEG 3350 had a large effect. Crystals that formed from 12%-16% PEG 3350 were larger than the initial crystals, although they did not have markedly better morphology. Some of these larger crystals were sacrificed to characterize the nature

of the crystals. Some were stained with Izit dye, a proprietary dye sold by Hampton Research. To stain, 0.5  $\mu$ L of Izit dye was added to the 3  $\mu$ L drop. Protein crystals should stain blue because

Figure 2.8 – crSTS crystal characterization. a) Izit staining; b) SDS-PAGE analysis



the dye can diffuse into the crystal due to the larger spaces left between lattice contacts. Salt crystals should not stain because the spaces between lattice contacts should be too small to allow dye molecules to diffuse in. These crystals stained well and analysis by SDS-PAGE revealed a single protein band at the appropriate apparent molecular weight. At this point, it was concluded that the crystals were protein and suitable for diffraction screening.

Crystals of crSTS were screened for diffraction at the Drennan lab X-ray source under cryogenic conditions. The crystals to be screened were looped and frozen directly in the 100 K

cryostream. Initially, data was collected in a one degree oscillation ( $\Delta\phi = 1$ ) over fifteen minutes with the detector set to 300 mm. The first image collected was from protein crystallized in 0.2 M HCOONa, 16% PEG3350. It seemed to show irregular low resolution spots at very low intensity. An image collected on the same crystal at 90° from the first image showed no difference. After this, the detector was moved in to 250 mm and the exposure time was increased to 20 minutes (still for  $\Delta\phi = 1$ ). This gave no improvement over previous images taken.

After the poor results with the first crystal, several other crystals were screened for diffraction. For this second batch, cryo-protection was used to preclude lack of diffraction due to ice formation. Cryo-protection involved a five minute soak in the mother liquor solution with 20% glycerol added. After cryo-protection, the crystals were frozen in the cryo-stream as before. All crystals screened with cryo-protection gave the same general results as those screened without cryo-protection. Some images were taken using a  $\Delta\phi = 5$  to rule out salt crystals; these image showed no diffraction/weak diffraction and were clearly not salt patterns. Even increasing the exposure time to 45 minutes did not produce clear, unambiguous diffraction. The best results were obtained from a single crystal in a single orientation that gave weak, smeary spots at low (~10-15 Å) resolution. It was concluded that these crystals were too poor to produce good diffraction; further optimization of the crystallization conditions was performed.

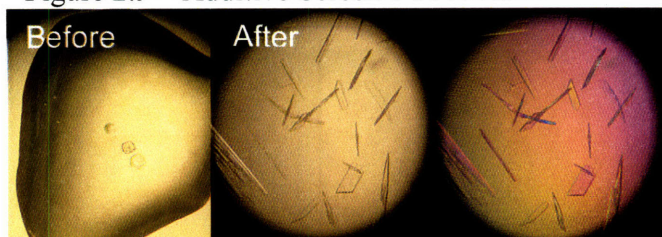
To obtain better crystals with improved morphology and, most importantly, better diffraction, additive and detergent screens were investigated. These screens are similar to sparse matrix screens, containing an array of different additives including salts, sugars, polymers, detergents, etc. Each screen contains 24 unique additives; detergent screen 1 and additive screens 1-3 were initially tried. For these screens, the initial conditions from Index Screen #90



were used (0.2 M HCOONa, 20% PEG3350) for the mother liquor. Trays were set up as described above with the only difference being the composition of the hanging drop. For the additive/detergent screens, the hanging drop included 1.5  $\mu$ L mother liquor, 0.3  $\mu$ L additive, and 1.2  $\mu$ L 6.25 mg/mL crSTS. The concentration of the protein was increased slightly to compensate for the added dilution and keep the final concentration of crSTS in the drop at 2.5 mg/mL. The hanging drop components were added in the order listed above (mother liquor followed by additive followed by protein) to ensure that the additive was fully ejected from the pipette tip into the drop. The screens were incubated at 4 °C and observed at 2, 4, and 6 days after being set up. They were scored on a simple scale of +, -, or 0 (+ = improved morphology, - = worse morphology, 0 = no visible change in morphology).

After 2 days, a greatly increased morphology was seen in additive screen 1, condition 14. This condition contains 30 % (v/v) (+/-)-2-methyl-2,4-pentandiol (MPD) at a final concentration

Figure 2.9 – Additive Screen 1 #14 results



of 3% (v/v) in the hanging drop. The crystal morphology changed from the rough, irregular spheres seen in previously optimized conditions to thin, rectangular plates.

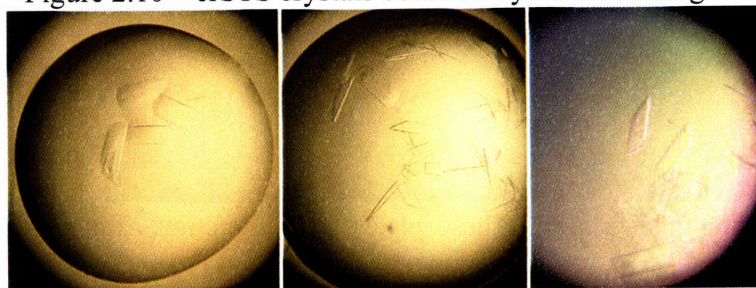
These plates were sometimes found in stacked clusters and sometimes as single plates. A few of the single plates appeared to be slightly thicker than others. Grid screening to further optimize the conditions revealed that better morphology (i.e. larger plates, more single plates) was obtained with higher MPD concentrations and lower PEG concentrations. Optimal crystals were obtained from 9%-14% MPD and 10%-16% PEG3350.

During this optimization, an inhibitor of crSTS was used in co-crystallization experiments to aid in crystal formation/growth. Also, one of the aims of this project is the use of

inhibitors to probe the enzyme active site, so using an inhibitor in co-crystallization would be a step towards this goal. Tryptamine and secologanin were linked through an amide linkage to form the inhibitor (synthesized by W. Runguphan, unpublished results). The inhibitor has an  $IC_{50}$  of 1  $\mu\text{M}$ ; this was converted using the Cheng-Prusoff Equation to give a  $K_i$  value of  $\sim 0.3 \mu\text{M}$ . Because of this, the inhibitor was used at a 10  $\mu\text{M}$  concentration in the hanging drop. Crystallization was tested over a range of PEG3350 concentrations (6%-16%) while keeping MPD concentration constant at 9%. Drops without the inhibitor included were set up as a control. After several days, the crystals in the inhibitor drops showed no improvement relative to the control. The inhibitor concentration was then varied between 5  $\mu\text{M}$  and 10  $\mu\text{M}$  (still vary PEG3350 in the same ranges and keeping MPD constant). Again, even after several days there was no improvement relative to the control drops. At this point, co-crystallization experiments were stopped. To obtain structures with inhibitors bound, further co-crystallization or soaking experiments will be necessary.

Figure 2.10 – crSTS crystals obtained by macroseeding.

Since many of the crystals in the optimized conditions were still very thin, macroseeding was used to obtain



thicker, more three-dimensional crystals. Macro seeding is a technique in which a pre-formed crystal is used as a seed for more crystal growth. The crystal is first “etched” by transferring it to a solution with a slightly lower precipitant concentration than that used to form the crystal. This allows the edges of the crystal to dissolve slightly. After etching, the crystal is transferred to a drop with fresh protein solution in it at a slightly lower concentration than that used to form the

seed. Upon equilibration, the protein in solution should grow onto the etched edges of the macroseed instead of forming new crystals. It is essential that the protein in the new drop be at a lower concentration. If not, microseeds transferred along with the macroseed can produce multiple nucleation sites and consume most of the protein in the drop in forming small and useless crystals. A selection of thicker crystals obtained using this method are pictured in Figure 2.10.

### **2.3.3 Screening crystals for diffraction**

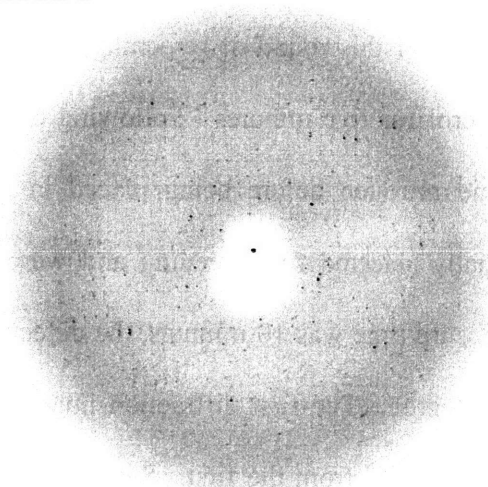
Optimized crystals from macroseeding were obtained approximately two weeks before a scheduled trip to the synchrotron at the Argonne National Laboratories Advanced Photon Source (APS). High resolution and high quality data sets are usually collected at synchrotron radiation sources. Synchrotron sources can produce much more intense and much more focused X-rays than those which can be produced at our home X-ray source. Several crystals were examined on a home X-ray source to judge which crystals to send to the synchrotron.

For this round of screening, all crystals were cryo-protected in mother liquor plus 20% glycerol for five minutes. Following cryo-protection, crystals were looped and flash frozen in liquid nitrogen before being placed in the 100 K cryo-stream. Large, single crystals were initially selected for screening and were looped using 0.3-0.4 mm microloops. The starting exposure time was 10 minutes, the detector distance was set at 300 mm, and the oscillation was one degree. The first diffraction image showed clearly that this was protein, as did an image taken at 90 ° from the first. Several different crystals were screened in the time available; the results for the different crystals are shown in Table 2.5.

Crystal	Results
crSTS-1	Showed clear protein diffraction to $\sim 3.6$ Å. The spots were strong for the home source, but they were split. There was a crack in the crystal when it was frozen which could be contributing to splitting. This crystal was thawed and a single fragment was broken off.
crSTS-2	Showed clear diffraction to $\sim 4.5$ Å, but the spots were weaker than those for crSTS-1. The spots appeared highly elliptical, so this crystal was annealed by blocking the cryo-stream for 10 seconds and re-screened.
crSTS-3	Showed diffraction to $\sim 4.5$ Å with decent looking spot morphology (most appeared circular).
crSTS-4	Showed diffraction to $\sim 3.3$ Å and had a set of good looking spots. However, some of the other spots appeared to be split, and there was a good deal of streaking. This crystal was probably not single.
crSTS-5	Showed diffraction to $\sim 4.0$ Å with decent looking spot morphology. This crystal gave the best results of any one screened, shown in Figure 2.11.
crSTS-6	This crystal showed poor diffraction and was discarded.
crSTS-1 Fragment	Following thawing and re-freezing of the chipped off fragment, very poor quality and streaky spots were seen. The fragment was discarded.
crSTS-2 Annealed	Annealing seemed to produce more spots than seen before, but they were streaky. Also, the quality of the previously observed reflections decreased after the annealing procedure. This crystal was also discarded.

The data collected during screening of these crystals was encouraging for several reasons. First of all, it showed that the optimized crystals were clearly protein and diffracted to moderate resolution (3.0-4.5 Å). Since they diffracted to this resolution at our home source, it was a good assumption that they would diffract to higher resolution with more intense X-rays at a synchrotron radiation source. Secondly, it seemed that a range of different crystals gave good diffraction, and that those sent to the synchrotron would give similar diffraction. However, the results of screening were not entirely satisfactory. The appearance of split spots and potential multiple lattices in some of the crystals was troubling,

Figure 2.11 – Diffraction results for crSTS-5



raising potential issues with later data processing. Also, the inability to index the crystal based on the diffraction data collected at home was less than optimal. Nonetheless, several crystals were sent to the synchrotron for data collection (crSTS-3, crSTS-4, and crSTS-5).

#### **2.4 Data processing with *C. roseus* STS**

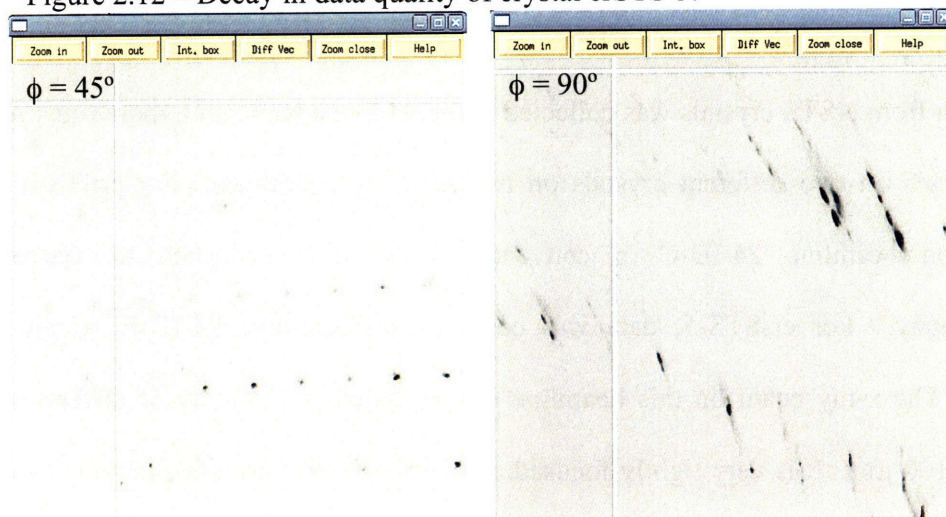
Data from crSTS crystals was collected at the Argonne National Laboratories APS. Data were collected on two different crystals on two different beamlines. For crSTS-3, data was collected on beamline 24-ID-C, a conventional beamline dedicated to macromolecular crystallography. For cr-STS-5, data was collected on beamline 24-ID-E, a micro-focused beamline. The x-ray beam on this beamline can be tuned to a variety of different diameters between 5-100  $\mu\text{m}$ . It is very tightly focused and very intense (more so than other synchrotron beams). Due to the intensity, crystal decay is a substantial problem. However, the small diameter of the beam allows for the crystal to be translated once decay begins to reduce data quality.

Because the space group was not known beforehand, 360° of data was collected for crSTS-3. Since crSTS-3 was shot on a regular beamline, the 360° was collected in a single spot. A set of the frames, from ~20°-65° look suitable for analysis; unfortunately, data outside of these regions is not usable. Before 20°, the diffraction is very weak and much fewer spots are visible. Above 65°, the data quality deteriorates rapidly. This is illustrated in Figure 2.12, with  $\phi = 45^\circ$  on the left and  $\phi = 90^\circ$  on the right. The figure shows the zoom view of the two images as displayed in the program Xdisp. The spots become very streaky, looking more like smears than actual spots. This problem continues through most of the rest of the data set. Due to the fact that so little of the data looked useful, it was reasoned that this data set would not provide enough information to solve the structure. Even if it did, the completeness would be so low as the render



the structure highly suspect. For this reason, no further work was done with the crSTS-3 data set and efforts were concentrated on crSTS-5.

Figure 2.12 – Decay in data quality of crystal crSTS-3.



The data set for crSTS-5 was collected on the micro-focused beamline. Because of the intensity of the beam, data was collected in six different wedges.

Wedge #	$\phi$ Covered	Comments
1	0°-31°	Initial frames look high quality, but data deteriorates towards the end of the wedge. Lunes are clearly visible, but stray spots appear in some frames. Spots morphology is good for some spots, less so for others.
2	27°-59°	Initial frames look good, but data deteriorates rapidly towards the end of the wedge. Lunes are clearly visible, but stray spots appear. Spot morphology is initially good, but becomes worse later in the wedge.
3	54°-90°	This is the best looking wedge. Diffraction is strongest here, and the spots look the best. Also, deterioration of data quality does not look as bad.
4	85°-109°	This wedge looked similar to wedge 1. Data seems to decay more rapidly, which is why fewer frames were collected in this wedge.
5	95°-110°	Ice rings are visible in this wedge, which is why fewer degrees were collected here. Spot morphology is poor.
6	120°-242°	Initial frames show very few spots. Lunes are not visible, and diffraction is weak compared to previous wedges. However, later in the wedge (after ~20 frames) the diffraction begins to look better, with stronger spots, clear lunes, and decent spot morphology. This deteriorates quickly, with the spots becoming increasingly streaky.

After each wedge was collected, the beam was translated down the length of the crystal.

The data for each of the wedges is summarized in Table 2.6.

In the initial frames of a wedge, the diffraction went past 1.9 Å, but in later frames, the decay reduced the resolution to less than

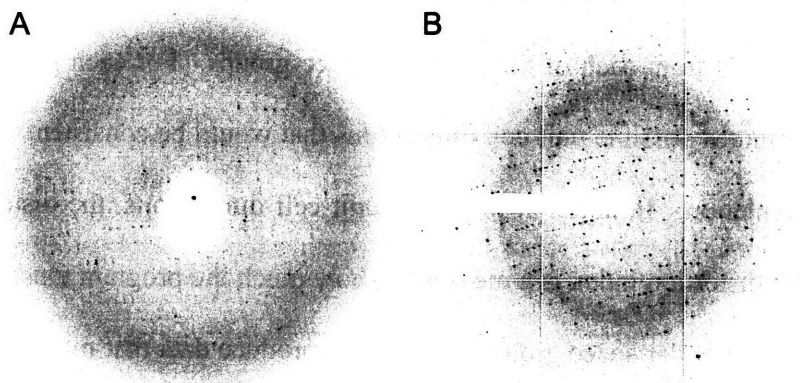


Figure 2.13 – Improved diffraction on the micro-focused beam.  
a) Home diffraction; b) Synchrotron diffraction.

2.6 Å. For indexing and integration, the data cut-off was set to 2.5 Å; for scaling the data was cut off at 3.0 Å. While the data did not look optimal, it was decided to attempt to process these data. From this point onward, discussion of data processing will refer only to the crSTS-5 data set collected on the 24-ID-E micro-focused beamline.

#### 2.4.1 Indexing with Denzo

The first step in processing data for any crystal is indexing. Indexing provides information regarding the space group and unit cell dimensions; this information is used as a guide when integrating reflection intensities and scaling data sets together. Knowing the space group is essential because it gives the reflections that are symmetry related. These reflections can be compared to judge the quality of the data set and the progress of data processing. Also, knowing how many of the reflections are symmetry related allows you to compute the completeness of your data set. If there are more symmetry related reflections, a data set with fewer reflections may still have a large degree of completeness and redundancy.

Using the data from the single frames, indexing was done using the program Denzo (part of the HKL2000 suite of programs). To index, the program selects a subset of reflections to use in its calculations. It then attempts to fit this selection of spots to the results that would be obtained from each of the potential space groups. For each potential space group, the program computes possible unit cell dimensions that would be consistent with the set of spots used in the calculation. In addition to giving unit cell dimensions, the distortion index is also calculated. The distortion index is a measure of how much the program must distort the unit cell dimensions of a particular space group to force it to produce data that matches the observed data. The space group with the highest symmetry and the lowest distortion index is selected for fitting and optimization.

In this process of fitting and optimization, Denzo predicts where spots should appear on the image. Fitting algorithms optimize a variety of values (crystal distance, detector distance, crystal rotation, unit cell dimensions, mosaicity etc.) to achieve the best relationship between spots and predictions. The quality of the fitting can be judged by visual inspection of the image (looking to see if predictions match with actual spots) or by looking at the value for the chi-squared goodness-of-fit test. The goal of fitting is to minimize the chi-squared value and maximize the overlap of predictions with actual reflections. If fitting proceeds well, it is reasonable to conclude that you have selected the proper space group and can proceed with data processing.

The indexing of the data set for crSTS-5 was not a straightforward process. Initial attempts focused on early frames of wedges 1-3 because they seemed to have the best diffraction. The automated peak search function would select between 300-600 peaks per image. Running the indexing script with large peak searches like this consistently gave unreasonable results. The

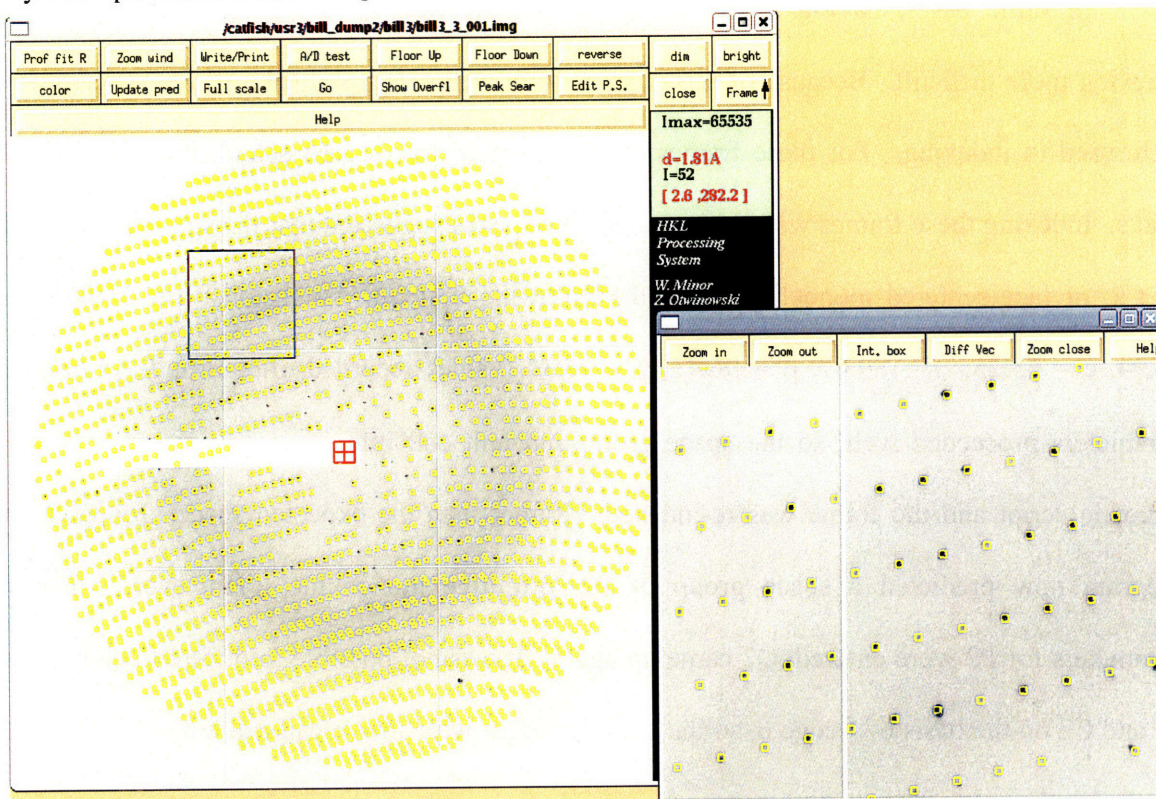


distortion indices for a number of different space groups were very low, however one of the unit cell dimensions in each of the space groups was impossibly low (below 1 Å). Reducing the number of peaks included in the indexing produced similar results, as did hand-selecting groups of peaks that seemed to be part of a single lattice. When the first several wedges did not work, indexing with wedge 6 was investigated.

In some cases, having too many reflections selected by the peak search can make indexing more difficult. Because of this, frames from wedge 6 that displayed fewer reflections were used in indexing. For these frames, the automated peak search selected fewer than 100 peaks. Indexing these frames with this smaller set of peaks suggested that the space group might be C2, or face-centered monoclinic, with the following unit cell dimensions:  $a = 218.53$ ,  $b = 46.94$ ,  $c = 75.25$ ;  $\alpha = 89.67$ ,  $\beta = 96.41$ ,  $\gamma = 91.34$  (values recorded after fitting). Fitting of parameters proceeded well, so the space group and unit cell information were entered in the indexing script and the frame was re-indexed. Rather than the expected space group C2, the program now predicted a space group of P2, primitive monoclinic. However, when the parameters for P2 were entered, C2 came up again. It was not possible to differentiate between P2 and C2 on the basis of wedge 6, so the earlier wedges were investigated again to conclusively determine the space group.

Since the frames in wedge 3 looked the best on the basis of visual inspection, it was selected first. Indexing with the C2 space group/unit cell information entered suggested C2 as expected. When the P2 information was entered and the frame was re-indexed, C2 came out as the space group with the lowest distortion index. This was a good indication that C2 was in fact

Figure 2.14 – Overlap between spots and predictions following parameter fitting. The yellow predictions indicate partial reflections.



the correct space group. Parameter fitting gave good overlap between the spots and predictions as shown in Figure 2.14. At this stage, it became apparent that the mosaicity was fairly high, possibly greater than 1.0. Even when the mosaicity was increased to 1.5, there were still some spots that were not covered by predictions (note the low resolution regions in Figure 2.14). This provided further evidence that there may be multiple lattices present in the crystal. Even though this is a significant issue in data processing, the spots seemed to be far enough apart so that the

lattices should be distinguishable. Using the information from the indexing, the first three wedges were integrated with Denzo, and the results scaled and analyzed using Scalepack.

#### **2.4.2 Integration and scaling with Denzo/Scalepack**

Integration is the process by which the intensities of the different reflections are measured. Based on the space group and unit cell, Denzo predicts where reflections should be and integrates the intensity of the peak in relation to the background. A weak level is set, and the program will reject reflections below this level. Also, the program will reject overlapping reflections. Several parameters can be optimized for integrations, including spot and background shape and size. It is important to have the right shapes and sizes to maximize the portion of the reflection that is measured while minimizing the reflections rejected due to overlap issues. Following integration, scaling allows you to analyze data from a single data set or merge data from multiple sets. Scaling assigns a scale factor to each frame included to account for weakening diffraction due to data decay over the course of data collection. Scalepack also performs several iterations of post-refinement, in which select integration parameters are refined. The most notable of these is the mosaicity. Importantly, Scalepack will reject reflections that don't seem to match and compare the symmetry related reflections. This is the first time when the data is split into multiple resolution shells and several crystallographic statistics are computed, including  $I/\sigma$ , completeness, redundancy, and  $R_{\text{sym}}$ . The quality of the data can be judged at the end of scaling based on the  $R_{\text{sym}}$  value and on the chi-squared goodness-of-fit test.

While scaling was performed on wedges 1-3, only wedge 3 will be discussed since it provided the highest quality data. Initially, scaling gave some questionable results. The scale factors assigned to the different frames varied a great deal, from unreasonably low to unreasonably high. The values for mosaicity also showed a great deal of variation, as did the

chi-squared test values. These results were indicative of a problem in the script used to run Scalepack, and were resolved when detector specific error parameters were included. Re-scaling following the inclusion of these parameters gave much more reasonable results without having to resort to using scale restrains. At this stage, it was discovered that the post-refinement was giving mosaicity values of ~1.0-1.1 while the integration was done with a mosaicity of 1.5. Reducing the mosaicity used in the integration to 1.0 and re-scaling gave post-refined mosaicity values that matched well with the initial value (ranging from 0.9-1.1).

Despite the success with the mosaicity, several issues with the scaling were not able to be resolved. First, the program consistently rejected close to 50% of the data. Optimizing the integration or scaling parameters was not able to reduce the number of rejected reflections. These problems with rejection are indicative of systemic data problems that may be caused by multiple lattices (crystal twinning) or by the decay in intensity of diffraction in later frames. Secondly, the chi-squared goodness-of-fit test gave very high values (none below 1.0, greater than 75 for one of the resolution shells). Related to this, the  $R_{\text{sym}}$  values were too large. Ideally,  $R_{\text{sym}}$  should be zero because it is comparing symmetry related reflections that should be identical. Practically this is not possible due to inherent experimental errors, so values of up to 7-8% are acceptable for overall  $R_{\text{sym}}$ . Overall, the  $R_{\text{sym}}$  for wedge 3 was 10.3% with much higher values for higher resolution shells.

It was thought that these problems might be due to the small amount of data present in the data set (only 36 frames in wedge 3), so data from multiple wedges was rescaled. A variety of permutations were attempted, including wedges 1+3, 2+3, 1+2, 1+2+3. For each one of these different tests, the scale factors, mosaicity values, and statistics were worse than wedge 3 alone. Due to these issues, it was concluded that the data from this crystal would not yield adequate

information to solve the structure. The work with this data set did produce some useful data, including the space group and unit cell dimensions for this crystal form. It also made clear the difficulty posed by multiple lattices that will need to be overcome before the structure can be solved.

Table 2.7 – Crystal parameters and statistics for data from wedge 3 and wedges 1-3 merged. Values in parentheses are for the highest resolution shell.

	Wedge 3		Wedges 1-3	
Space Group	C2		C2	
Unit cell parameters (Å)	a = 218.14 b = 46.35 c = 74.66	$\alpha$ = 90.00 $\beta$ = 96.79 $\gamma$ = 90.00	a = 218.22 b = 46.57 c = 74.82	$\alpha$ = 90.00 $\beta$ = 97.23 $\gamma$ = 90.00
Resolution (Å)	3.0		3.0	
Total observations	19223		53239	
Unique reflections	15126 (1500)		15486 (1547)	
Completeness (%)	52.6 (59.1)		71.4 (74.9)	
Redundancy	1.27 (1.18)		3.43 (1.89)	
R <sub>sym</sub> (%)	10.3 (26.5)		45.8 (59.5)	
I/ $\sigma$	9.6 (4.6)		12.5 (4.8)	

## 2.5 *R. serpentina* strictosidine synthase

The enzyme strictosidine synthase has been cloned from three species of plant – *Catharanthus roseus*, *Rauvolfia serpentina*, and *Ophiorrhiza pumila*.<sup>11</sup> Previous work in the O'Connor lab has focused on *C. roseus* because of its diverse and well characterized alkaloid profile; as such our biochemical and enzyme engineering studies have focused on crSTS. Previous crystallographic work has focused on the enzyme from *R. serpentina* (rsSTS), which is 79% identical to crSTS.<sup>12-14</sup> Since published crystallization conditions were available for rsSTS, this enzyme was initially selected to provide a rapid way to crystallographically characterize a variety of different mutants and substrate mimic inhibitors/wild-type enzyme complexes.

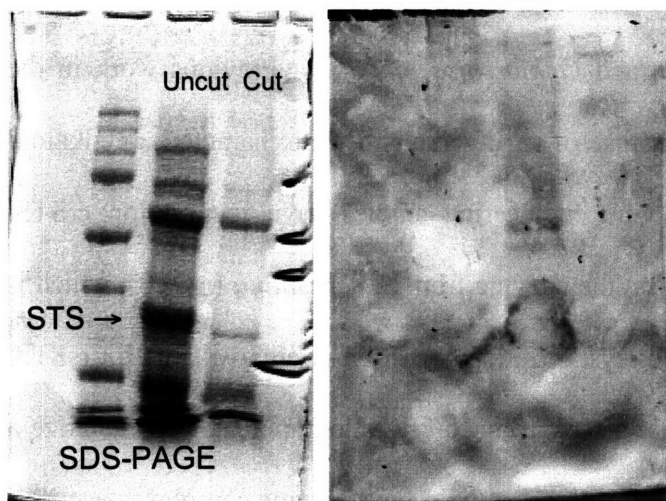
### 2.5.1 Preparation of rsSTS for crystallization

rsSTS was expressed and purified similarly to crSTS as detailed in the Materials and Methods section. One difference was that rsSTS was cloned into pET-28a with an N-terminal

His-tag instead of the C-terminal His-tag used for crSTS. This cloning and expression strategy was selected to follow the published procedure as closely as possible. This approach required that the His-tag be cleaved prior to crystallization because when rsSTS was previously crystallized, the N-terminal His tag had to be removed. In pET-28a, there is a thrombin cleavage site included between the His-tag and the N-terminus of the protein to allow for removal of the affinity tag. Thrombin is a site specific protease which functions in the blood clotting cascade; it recognizes the amino acid sequence Leu-Val-Pro-Arg-Gly-Ser and cuts on the C-terminal side of the Arg residue. Because of the specificity for this sequence, it should be possible to keep non-specific background cleavage to a minimum by optimizing the reaction time.

Initially, 20  $\mu$ L samples were taken at 1,2,4,6, and 24 hrs. These were analyzed by SDS-PAGE to determine the rate of cleavage, which would be indicated by the appearance of a protein with a lower apparent molecular weight. A sample of uncleaved protein was used as a

Figure 2.15 – SDS-PAGE and Western blot analysis of rsSTS thrombin cleavage



standard to show the starting molecular weight. After 24 hrs, minimal cleavage was apparent. This could be for any number of reasons. rsSTS tends to precipitate at pH values above 7.5, so the cleavage was run at this pH. However, thrombin's optimal activity is at a pH value of 8.2, so this could explain the lack of cleavage.

Additionally, running the reaction at 4 °C, while preferable to maintain protein stability, may be slowing the reaction to an unacceptable speed. Fortunately, this initial reaction was not

discarded and SDS-PAGE analysis after 7 days showed complete cleavage. A second cleavage reaction was set up and time points were taken at 1, 2, 3, 5, and 7 days. Cleavage appears to be complete after 2 days with minimal non-specific cleavage. This was verified by Western blot analysis using anti-His antibodies. It can be seen that there is a band in the uncut lane, indicating presence of the His tag, and no band in the cut lane, indicating cleavage. The blot is slightly over-developed, but it does appear that the thrombin cleavage has resulted in successful removal of the His tag.

### **2.5.2 *R. serpentina* STS activity assays**

Before further purification was attempted for rsSTS, activity assays were performed to make sure that the protein was functioning properly and would yield a relevant crystal structure. As with the crSTS assays performed before crystallization trials began, a simple qualitative assay would be sufficient to verify the activity, so the STS colorimetric assay was attempted.

Three reactions were initially prepared: – cleaved rsSTS as the experiment and uncleaved rsSTS/no enzyme as controls. No color was observed in any reaction following 30 minutes at 30 °C. The reactions were allowed to incubate for 48 hours; even after this long incubation period no color was observed. Since rsSTS was used directly from the thrombin cleavage without further purification, it was hypothesized that the Tris buffer and CaCl<sub>2</sub> present in the cleavage buffer were interfering in some way with the reaction. rsSTS was buffer exchanged into 50 mM NaH<sub>2</sub>PO<sub>4</sub> pH = 7.5 and the assays were repeated but no enzyme activity was apparent. The assays were repeated and analyzed by HPLC to preclude any issue with the colorimetric assay. No catalytic activity was observed, even after several hours incubation.

DNA sequencing verified the integrity of the gene, so the reason for this lack of activity is puzzling. There are slight differences between this construct of rsSTS and the previously



crystallized/characterized rsSTS due to differences in the expression vector used. However it is unlikely that changes in the two N-terminal amino acids left after cleavage would yield such a dramatic loss of activity. Because of the lack of activity efforts toward crystallization of rsSTS were halted and efforts towards the crystallization of *C. roseus* STS were begun. Not only would crSTS provide the most biochemically relevant crystal structure, it was known from previous work in the O'Connor lab that it could be easily obtained in a highly pure and active preparation.

## **2.6 Summary and future work**

Two different forms of strictosidine synthase have been studied with the ultimate goal of structural characterization of Pictet-Spenglerases. STS from *R. serpentina* was cloned, expressed, and partially purified in the hopes that reproducing the published crystallization conditions would provide a facile route to structures of mutants and structures with various inhibitors/substrates bound. Problems with enzyme activity greatly slowed this particular pathway of investigation, so *R. serpentina* STS was abandoned in favor of *C. roseus* STS.

*C. roseus* STS was cloned into an *E. coli* expression vector, expressed, and purified to electrophoretic homogeneity. Following kinetic characterization of the His-tagged construct to verify activity, it was subjected to sparse matrix crystal screens. The best hit from these screens was optimized through multiple rounds of grid screening, additive screens, and macroseeding. Data was collected for these crystals at the Argonne National Laboratories APS and processed using the HKL2000 suite, including Denzo and Scalepack. Several issues with the data set prevented the solving of a structure, including data decay and possible multiple lattices. It was possible to determine the space group and unit cell dimensions, and this information can be used to guide further data collection efforts.



Future work in this project can focus on two different areas. The first is getting a high quality data set that can be used to solve the structure by molecular replacement with the existing STS structure. This may be possible by simply screening more crystals for diffraction before sending them to the synchrotron to identify single crystals that diffract well. Another possible solution may be growing larger crystal using microseeding techniques. This may prevent or reduce the growth of multiple lattices and greatly simplify the data processing. Finally, the cryo-protection procedure can be further optimized, possibly including a gradual soak in increasing concentration of glycerol. This may minimize shock to the crystals upon flash freezing.

The second area is solving the structures of interesting mutants and solving structures with inhibitors bound. Since the mutants are all point mutations in the active site, the surface residues and lattice contacts should not be perturbed significantly. The mutants should crystallize in similar conditions as the wild type, although some grid screening may be necessary to fully optimize the conditions. As far as structures with inhibitors bound, co-crystallization has not been ruled out, even though including an inhibitor did not produce any difference in crystallization properties. In addition to further co-crystallization experiments, soaking experiments would be a worthwhile pursuit.

## **2.7 Materials and Methods**

### **Cloning crSTS into pET-28a**

The gene for STS was obtained from P. Bernhardt in a pGEM-T Easy (Roche, Amp<sup>R</sup>) sub-cloning vector with N-terminal NcoI and C-terminal XhoI restriction sites. The gene construct was cloned from the pGAL-MF yeast expression vector with a single nucleotide insertion following the gene to allow expression in pET-28a with a C-terminal His tag with a

minimal linker sequence (Gly,Ser). A restriction double digest was performed on pGEM-T\_STS and empty pET-28a using the following reaction mixture:

Restriction digest: 19  $\mu$ L sterile ddH<sub>2</sub>O  
5  $\mu$ L 10x NEB Buffer 2  
20  $\mu$ L plasmid DNA (pGEM-T\_STS or pET-28a)  
3  $\mu$ L NcoI  
3  $\mu$ L XhoI

The reactions were incubated at 37 °C for 3 hrs. Following incubation, 12.5  $\mu$ L 5x DNA loading dye was added and the entire reaction was loaded onto a 1% agarose gel and run at 100 V for 45 minutes. After staining with ethidium bromide for 1 hr, the STS gene fragment and the cut pET-28a plasmid were isolated using a gel extraction kit (Qiagen).

Following this, the gene fragment was ligated into pET-28a. Following the ligations, 1  $\mu$ L of the reaction mixture was added to 55  $\mu$ L electrocompetent *E. coli* Top 10 cells and incubated on ice for 2 minutes. Transformation was achieved using a Bio-rad Micropulser on the E2 setting. Following transformation, 250  $\mu$ L SOC media was added and cells were recovered on ice for ~5 min. After recovery, cells were grown shaking at 37 °C for 1 hr. before the entire 250  $\mu$ L was plated on LB Agar (Kan) plates. Plates were incubated for ~24 hrs. at 37 °C before colonies were picked and grown up in 5 mL liquid cultures with 30  $\mu$ g/mL Kan.

5 mL liquid cultures were grown overnight shaking vigorously at 37 °C. Plasmid DNA was isolated using a commercial miniprep kit (Qiagen). An analytical double digest using the following reaction mixture verified presence of the STS gene insert:

Double digest: 3  $\mu$ L sterile ddH<sub>2</sub>O  
1  $\mu$ L 10x NEB Buffer 2  
4  $\mu$ L pET-28a\_crSTS  
1  $\mu$ L NcoI  
1  $\mu$ L XhoI

This reaction was incubated at 37 °C for 1 hr. before analysis by gel electrophoresis. Following staining with ethidium bromide for ~1 hr., the STS insert (approximately 1 kb) could be seen in several samples. These were submitted for sequencing using the T7 promoter primer and T7 terminator primers with the following reaction mixture:

Sequencing: 6  $\mu$ L sterile ddH<sub>2</sub>O  
5  $\mu$ L pET-28a\_crSTS  
1  $\mu$ L *either* T7 promoter or T7 terminator primer

Sequencing verified the presence of the intact STS gene insert with the proper C-terminal hexahistidine tag.

### **Expression and purification for kinetic characterization**

The following buffers were used during the purification of crSTS for kinetic characterization:

Lysis Buffer: 50 mM NaH<sub>2</sub>PO<sub>4</sub>  
300 mM NaCl  
10 mM Imidazole  
pH = 8.0

Wash Buffer: 50 mM NaH<sub>2</sub>PO<sub>4</sub>  
300 mM NaCl  
20 mM Imidazole  
pH = 8.0

Elution Buffer: 50 mM NaH<sub>2</sub>PO<sub>4</sub>  
300 mM NaCl  
250 mM Imidazole  
pH = 8.0

SE30 Buffer: 50 mM NaH<sub>2</sub>PO<sub>4</sub>  
10% glycerol  
pH = 7.5

The vector pET-28a\_crSTS was transformed into *E. coli* BL21 (DE3) cells by electroporation. Following the pulse, 250  $\mu$ L SOC media was added and cells were recovered on ice for five minutes. After recovery, cells were grown shaking at 37 °C for 1 hr. before plating of 50  $\mu$ L on an LB-Agar (Kan) plate. This plate was incubated overnight at 37 °C. A 5 mL LB broth (1x Kan) starter culture was inoculated with cells from a single colony picked from the growth plate. This culture was grown overnight shaking at 37 °C. Following overnight starter culture growth, 1 mL of the 5 mL culture was added to 1 L LB broth (1x Kan). The 1 L culture

was incubated shaking at 37 °C until an O.D. of between 0.5-0.7 was reached (O.D. =  $A_{600}$  using fresh LB broth as a blank). Once O.D. had been reached, the temperature was dropped to 18 °C and the culture was induced with IPTG added to a final concentration of 1 mM. The cultures were grown overnight shaking at 18 °C. In the morning, cells were harvested by pelleting in the centrifuge at 4500 rpm for 15 min. Pellets were either used fresh or frozen and stored at -20 °C until future use.

To obtain crSTS, cells were lysed with a combination of lysozyme and sonication. The cell pellet was resuspended in 3 mL lysis buffer per gram pellet (wet weight). Lysozyme was added to a final concentration of 1 mg/mL from a 10 mg/mL stock solution in water stored at 4 °C. Stirring or mild shaking at 4 °C for 20-30 minutes gently agitated the lysozyme mixture. Following this, the lysis mixture was transferred to a 150 mL beaker packed in ice and sonication was achieved using a Branson Digital Sonifier set at 63% tip amplitude. 1s pulses were followed by 4s rests to prevent heating of the lysis mixture for a total of 2.5 minutes total pulse time. The lysis mixture was then transferred to 30 mL Oak Ridge tubes and centrifuged at 13,000 rpm for 50 minutes to pellet cellular debris. The supernatant was collected in 50 mL conical tubes and the pellet was discarded.

Purification by Ni-NTA affinity chromatography was performed in batch. Approximately 1 mL of Ni-NTA resin was used for each gram of cells lysed. Fresh resin was washed with 100-200 mL lysis buffer before use to remove ethanol from the storage buffer. The resin was then resuspended in ~1 mL lysis buffer per mL resin initially used. The resin was added to the supernatant in the 50 mL tubes and incubated with gentle shaking at 4 °C for 1 hr. Following this incubation period, the mixture was added to the column and the flow-through was collected. After collection of the flow-through, the column was washed with 2x 50 mL wash

buffer to remove weakly or non-specifically bound proteins. Finally, crSTS was eluted from the column with 3x 10 mL 250 mM imidazole elution buffer.

Following Ni-NTA affinity chromatography, crSTS was purified to electrophoretic homogeneity by FPLC. Size exclusion gel filtration chromatography using an SE30 column gave poor resolution, although later fractions were achieved in sufficiently pure state for use in kinetic assays. Using a higher resolution column (SE75 or SE200) should avoid this problem in the future. Following size exclusion chromatography, crSTS was further purified by MonoQ anion exchange chromatography with a gradient elution (0% B to 100% B). This gave pure crSTS (purity assessed by SDS-PAGE). The protein was buffer exchanged into SE30 buffer for kinetic characterization.

#### **Expression and purification for crystallization studies**

For crystallization studies, crSTS was expressed in *E. coli* as previously detailed and purified by Ni-NTA affinity chromatography in the same way as for kinetic characterization. After elution, fractions were pooled and buffer exchanged into 10 mM Tris HCl pH = 7.5 (the crystallization buffer previously used in the crystallization of rsSTS). It was purified by MonoQ anion exchange chromatography and size exclusion gel filtration chromatography using SE200 column (the highest resolution column available). The following buffers were used:

MonoQ A/SE200: 10 mM Tris HCl  
0 mM NaCl  
pH = 7.5

MonoQ B: 10 mM Tris HCl  
1 M NaCl  
pH = 7.5

After MonoQ and SE200, fractions containing pure crSTS were concentrated in an Amicon Ultra Centrifugal Filtration device to a final concentration of 12 mg/mL.

### **Assays for kinetic characterization of crSTS**

The assay mixture of crSTS consisted of the following components. crSTS was diluted to its final concentration from a 3.26  $\mu\text{M}$  stock solution. The proper dilution was determined empirically:

crSTS Kinetic Assays: 100  $\mu\text{L}$  200 mM  $\text{NaH}_2\text{PO}_4$  pH = 7.00  
50  $\mu\text{L}$  sterile ddH<sub>2</sub>O  
10  $\mu\text{L}$  40.3 mM secologanin (courtesy of P. Bernhardt)  
20  $\mu\text{L}$  10x tryptamine stock solution  
20  $\mu\text{L}$  crSTS

The reaction was incubated at 30 °C. At 30, 60, 120, and 180 minutes, 50  $\mu\text{L}$  aliquots were removed and quenched with 5.5  $\mu\text{L}$  2 M NaOH. Analysis was performed by HPLC using 0.1% TFA in H<sub>2</sub>O as buffer A and acetonitrile as buffer B. A gradient elution was used, proceeding from 10%B to 100%B over 15 minutes. To correct for injection error, 75  $\mu\text{M}$  naphthylacetic acid (NAA) was included in the assay buffer as an internal standard. Production of strictosidine was measured by UV absorbance at 280 and 228 nm. In actuality, the strictosidine lactam is produced under the basic quench conditions, however the conversion is rapid and total. For simplicity, it shall be referred to simply as strictosidine.

### **Site-directed mutagenesis with *C. roseus* STS**

Starting with pET28a\_crSTS, site-directed mutagenesis was used to obtain *E. coli* expression constructs for the relevant mutants (V214M, F226L, D177A). These mutants were previously identified and are potential targets for crystallographic analysis. Primers were obtained from Integrated DNA Technologies and were HPLC purified.

F226L: Forward – GTTGTTGTTGCTGAGCTGTTGTCCAACAGAATC  
Reverse – GATTCTGTTGGACAACAGCTCAGCAACAACAAC

V214M: Forward – GAAGGAGTTGCACATGCCAGGTGGTGCTG  
Reverse – CAGCACCACCTGGCATGTGCAACTCCTTC

D177A: Forward – CTCTATTCACGACGCTTCTCCAGAAGGTG  
Reverse – CACCTTCTGGAGAAGCGTCGTGAATAGAG

Site-directed mutagenesis was performed using a Stratagene QuickChange Site-Directed Mutagenesis kit. The following reaction mixtures were used for each reaction. The V214M and D177A reactions initially failed, so DMSO was included to prevent DNA secondary structure. With DMSO, both V214M and D177A reactions were successful.

<u>F226L</u> : 5 $\mu$ L 10x Pfu Turbo Buffer	<u>D177A/V214M</u> : 5 $\mu$ L 10x Pfu Turbo Buffer
1.3 $\mu$ L forward primer	1.3 $\mu$ L forward primer
1.3 $\mu$ L reverse primer	1.3 $\mu$ L reverse primer
1 $\mu$ L 10 mM dNTP mix	1 $\mu$ L 10 mM dNTP mix
0.6 $\mu$ L pET28a_crSTS	0.6 $\mu$ L pET28a_crSTS
40.8 $\mu$ L ddH <sub>2</sub> O	5 $\mu$ L DMSO
	35.8 $\mu$ L ddH <sub>2</sub> O

These components were mixed in individual PCR tubes. Immediately before reaction, 1  $\mu$ L Pfu Turbo DNA polymerase was added to each tube. PCR was done using a Stratagene RoboCycler Gradient 96 with the following program:

<u># of Cycles</u>	<u>Course</u>
1	30 s, 95 °C
1-17	30 s, 95 °C
	1 min, 55 °C (59 °C for V214M)
	6.4 min, 68 °C (1 min per kb plasma length)
Hold	Hold at room temperature

Following the completion of the PCR cycles, 1  $\mu$ L of the restriction enzyme DpnI was added to each reaction mixture to digest methylated DNA (the starting plasmid, but not the mutated product). After addition of DpnI, the reactions were incubated at 37 °C for 1 hr and then placed on ice. Freshly thawed XL1-Blue Supercompetent cells were aliquoted out into 50  $\mu$ L portions. 1  $\mu$ L of the PCR/DpnI digest mixture was added to each aliquot. These were incubated on ice for at least 5 minutes. Cells were then transformed by heat shock at 42 °C for

30 s, followed by addition of 500  $\mu$ L SOC media (pre-warmed to 37 °C). Cells were allowed to recover on ice for 2 min before incubation with shaking for 1 hr at 37 °C. After this initial growth phase, all 500  $\mu$ L was plated onto 2 LB Agar (Kan) plates and grown overnight. Colonies were picked and grown in liquid culture; DNA was isolated using a Qiagen miniprep kit. DNA sequencing using the T7 promoter and T7 terminator primers as previously described verified mutation.

### **Purification of crSTS mutants for crystallization studies**

Both mutants were expressed and purified by Ni-NTA chromatography as described for wild-type crSTS using the same methods and buffers. Following elution from the Ni-NTA column, the fractions containing crSTS F226L and D177A were concentrated and transferred into 50 mM Tris pH = 7.5, 10% glycerol (by buffer exchange or dialysis). They were then purified further by MonoQ anion exchange chromatography using the following buffers:

<u>Mono Q A</u> : 50 mM Tris pH = 7.5	<u>Mono Q B</u> : 50 mM Tris pH = 7.5
0 M NaCl	1 M NaCl
10% glycerol	10% glycerol

Following MonoQ chromatography, the relevant fractions were collected and analyzed by SDS-PAGE.

### **Colorimetric assay of STS**

This assay is a coupled assay that relies on strictosidine glucosidase. Deglycosylation of strictosidine produces an intermediate that spontaneously rearranges to give a yellow precipitate. The reaction mixture contains excess secologanin and tryptamine. The following mixture was used:



Colorimetric Assay: 10  $\mu$ L 10 mM tryptamine  
5  $\mu$ L 40.3 mM secologanin  
10  $\mu$ L 200 mM NaH<sub>2</sub>PO<sub>4</sub> pH = 7.00  
20  $\mu$ L SG  
10  $\mu$ L cleaved rsSTS/uncleaved rsSTS/ddH<sub>2</sub>O  
45  $\mu$ L ddH<sub>2</sub>O

Following incubation at 30 °C, the yellow precipitate should form between 10 and 30 minutes.

### **Cloning rsSTS into pET-28a**

Previous efforts at heterologous expression of rsSTS in *E. coli* gave very low yields. Because of this, the codon-optimized gene (lacking the N-terminal signal sequence) was purchased from Genart. It was obtained in a pGA18 plasmid with the ampicillin resistance gene. Following amplification and isolation (by miniprep) of this plasmid from *E. coli* Top10 cells, a double digest was performed to obtain the gene insert. The target vector pET28-a was digested simultaneously to prepare for the eventual ligation. The following reaction mixture was used:

Restriction digest: 19  $\mu$ L sterile ddH<sub>2</sub>O  
5  $\mu$ L 10x NEB Buffer 2  
20  $\mu$ L plasmid DNA (pGA18\_rsSTS or pET-28a)  
3  $\mu$ L NdeI  
3  $\mu$ L EcoRI

The reactions were incubated at 37 °C for 3 hrs. Following incubation, 12.5  $\mu$ L 5x DNA loading dye was added and the entire reaction was loaded onto a 1% agarose gel and run at 100 V for 45 minutes. After staining with ethidium bromide for 1 hr, the STS gene fragment and the cut pET-28a plasmid were isolated using a gel extraction kit (Qiagen).

Following this, the gene fragment was ligated into pET-28a. The following ligation mixture was successful after several trials. Following the ligations, 1  $\mu$ L of the reaction mixture was added to 55  $\mu$ L electrocompetent *E. coli* Top 10 cells and incubated on ice for 2 minutes.

Transformation was achieved using a Bio-rad Micropulser on the E2 setting. Following transformation, 250  $\mu$ L SOC media was added and cells were recovered on ice for ~5 min. After recovery, cells were grown shaking at 37 °C for 1 hr. before the entire 250  $\mu$ L was plated on LB Agar (Kan) plates. Plates were incubated for ~24 hrs. at 37 °C before colonies were picked and grown up in 5 mL liquid cultures with 30  $\mu$ g/mL Kan. The experimental plate showed ~50 colonies, while the control plate lacking the rsSTS gene insert showed no colonies.

5 mL liquid cultures were grown overnight shaking vigorously at 37 °C. Plasmid DNA was isolated using a commercial miniprep kit (Qiagen). An analytical double digest using the previously described reaction mixture verified presence of the rsSTS gene insert. The sequence was verified as previously described using the T7 promoter and T7 terminator primers present in the pET28-a vector.

### **Expression and Purification of rsSTS**

The following buffers were used during the purification of rsSTS:

Lysis Buffer: 50 mM NaH<sub>2</sub>PO<sub>4</sub>  
300 mM NaCl  
10 mM Imidazole  
pH = 8.0

Wash Buffer: 50 mM NaH<sub>2</sub>PO<sub>4</sub>  
300 mM NaCl  
20 mM Imidazole  
pH = 8.0

Elution Buffer: 50 mM NaH<sub>2</sub>PO<sub>4</sub>  
300 mM NaCl  
250 mM Imidazole  
pH = 8.0

*R. serpentina* STS was cultured and isolated similarly to *C. roseus* STS. Briefly, the vector pET-28a\_rsSTS was transformed into *E. coli* BL21 (DE3) cells by electroporation. Following the electroporation, cells were grown shaking at 37 °C for 1 hr. before plating on an LB-Agar (Kan) plate. This plate was incubated overnight at 37 °C. A 5 mL LB broth (1x Kan) starter culture was inoculated with cells from a single colony picked from the growth plate. This

culture was grown overnight shaking at 37 °C. Following overnight starter culture growth, 1 L LB broth (1x Kan) was inoculated. The 1 L culture was incubated shaking at 37 °C until an O.D. of between 0.5-0.7 was reached. Once O.D. had been reached, the temperature was dropped to 18 °C and the culture was induced with IPTG. The cultures were grown overnight shaking at 18 °C. In the morning, cells were harvested by pelleting in the centrifuge at 4500 rpm for 15 min. Pellets were either used fresh or frozen and stored at -20 °C until future use.

To obtain rsSTS, cells were lysed with a combination of lysozyme and sonication. The cell pellet was resuspended in 3 mL lysis buffer per gram pellet (wet weight). Lysozyme was added to a final concentration of 1 mg/mL from a 10 mg/mL stock solution in water stored at 4 °C. Stirring or mild shaking at 4 °C for 20-30 minutes gently agitated the lysozyme mixture. Following this, the lysis mixture was transferred to a 150 mL beaker packed in ice and sonication was achieved using a Branson Digital Sonifier set at 63% tip amplitude. 1s pulses were followed by 4s rests to prevent heating of the lysis mixture for a total of 2.5 minutes total pulse time. The lysis mixture was then transferred to 30 mL Oak Ridge tubes and centrifuged at 13,000 rpm for 50 minutes to pellet cellular debris. The supernatant was collected in 50 mL conical tubes and the pellet was discarded. After lysis, cell debris was pelleted by centrifugation at 13,000 rpm for 55 minutes.

Following lysis, rsSTS was purified by Ni-NTA metal affinity chromatography. Pre-charged resin (Qiagen) was washed with lysis buffer before use to remove ethanol and equilibrate the column. 1 mL of resin was used for each gram of lysed cells. The resin was added to the supernatant in the 50 mL tubes and incubated with gentle shaking at 4 °C for 1-2 hrs. Following this incubation period, the mixture was added to the column and the flow-through was collected. After collection of the flow-through, the column was washed with 100-

200 mL wash buffer to remove weakly or non-specifically bound proteins. Finally, rsSTS was eluted from the column with 3x 10 mL 250 mM imidazole elution buffer. Fractions were analyzed by SDS-PAGE. Elution fractions were concentrated and buffer exchanged into 10 mM Tris-HCl before further use.

### **Thrombin cleavage of rsSTS**

For the cleavage reaction, a Thrombin CleanCleave kit (Sigma) was used. This kit features thrombin bound to an agarose bead. Having the thrombin bound facilitates easy removal when the reaction is done; the beads are simply pelleted by gentle centrifugation. The kit was used as per the manufacturers instructions. The resin was obtained as a 50% (v/v) suspension of resin in 20 mM Tris-HCl pH = 8.2, 50% glycerol. 100  $\mu$ L of this slurry was aliquoted into a 1.5 mL Eppendorf tube and spun at 2500 rpm to pellet the resin. The supernatant was decanted, and the resin was resuspended in 500  $\mu$ L 1x cleavage buffer (10x cleavage buffer = 500 mM Tris-HCl pH = 8.0, 100 mM CaCl<sub>2</sub>). This was spun at 2500 rpm to pellet the resin and the supernatant was decanted. This step was then repeated. Following washing, the resin was resuspended in 100  $\mu$ L 10x cleavage buffer. 900  $\mu$ L rsSTS was added to this to give a final concentration of 1x cleavage buffer and ~2.5 mg/mL Ni-NTA purified rsSTS (as determined by A<sub>280</sub>). The cleavage reaction was incubated at 4 °C with gentle shaking. To monitor the reaction, the resin was pelleted and an aliquot was removed for SDS-PAGE analysis. The resin was then resuspended and the reaction was allowed to proceed.

### **Western blotting to verify thrombin cleavage**

To verify the success of the thrombin cleavage, a Western blot was performed with anti-His antibodies. Following SDS-PAGE analysis, the protein was transferred to a nitrocellulose membrane without Coomassie blue staining. This was done with Towbin's transfer buffer,

which includes 100 mL 10x TG (25 mM Tris, 192 mM glycine, pH = 8.3), 200 mL MeOH, and 700 mL ddH<sub>2</sub>O. Transfer was done at 4 °C at 100 V for one hour. After transfer, the nitrocellulose membrane was blocked overnight using 5% freeze-dried milk in TBS (Tris buffered saline).

Once blocking was complete, the membrane was washed with TBST (TBS, 1% Tween-20) two times for 10 minutes each. TBST was added to cover the membrane and the washes were performed with gentle agitation. Following this, the membrane was incubated with a 1:1000 dilution of anti-pentaHis antibody (Qiagen) in TBS for one hour. After this, the membrane was washed with TBST three times for 10 minutes each. The membrane was then incubated with a 1:1000 dilution of AP (alkaline phosphatase) tagged anti-mouse antibody in TBS for one hour. Subsequently, the membrane was washed with TBST two times for ten minutes and TBS once for ten minutes. Finally, the membrane was incubated with NBT/BCIP (Nitro-Blue tetrazolium chloride/5-bromo-4-chloro-3'-Indolyphosphate p-toluidine salt) that reacts with alkaline phosphatase to produce a purple/black insoluble precipitate. After staining was complete, the reaction was quenched with ~1 mM EDTA in TBS. The membrane was dried and fixed on parafilm and scanned to visualize.

## 2.8 References

1. Kutchan TM. (1993) *Phytochemistry*. 32, 493-506.
2. McCoy EM, Galan MC, O'Connor SE. (2006) *Bioorg. Med. Chem. Lett.* 16, 2475-2478.
3. McCoy EM, O'Connor SE. (2006) *J. Am. Chem. Soc.* 128, 14276-14277.
4. Kirschning A, Taft F, Knobloch T. (2007) *Org. Biomol. Chem.* 5, 3245-3259.
5. Bernhardt P, McCoy EM, O'Connor SE. (2007) *Chemistry and Biology*. 14, 888-897.
6. Chen S, Galan MC, Coltharp C, O'Connor SE. (2006) *Chemistry and Biology*. 13, 1137-1141.
7. Pfitzner U, Zenk MH. (1989) *Planta Med.* 55, 525-530.
8. Hampp N, Zenk MH. (1988) *Phytochemistry*. 27, 3811-3815.
9. Geerlings A, Redondo FJ, Memelink J, Contin A, van der Heijden R, Verpoorte R. (1999) *Biotechnol. Tech.* 13, 605-608.
10. Lujendijk TJC, Stevens LH, Verpoorte R. (1996) *Phytochem. Anal.* 7, 16-19.
11. O'Connor SE, Maresh JM. (2006) *Nat. Prod. Rep.* 4, 532-547.
12. Ma X, Panjikar S, Koepke J, Loris EA, Stoeckigt J. (2006) *The Plant Cell*. 18, 907-920.
13. Ma X, Koepke J, Fritsch G, Diem R, Kutchan TM, Michel H, Stoeckigt J. (2004) *Biochim. Biophys. Acta*. 1702, 121-124.
14. Koepke J, Ma X, Fritsch G, Michel H, Stoeckigt J. (2005) *Acta Cryst.* D61, 690-693.

### **3. Norcoclaurine synthase crystallization studies**

In addition to the monoterpene indole alkaloids, another broad class of plant derived alkaloid natural products is the benzyloquinoline alkaloids. This class of natural products includes morphine, one of the most prevalent opiate analgesics isolated from poppy, and papaverine, a vasodilator also isolated from poppy.<sup>1</sup> They are similar to terpene indole alkaloids in that the biosynthesis begins with a Pictet-Spengler reaction, however the substrates and enzymes are different. Benzyloquinoline biosynthesis begins with the coupling of dopamine and 4-hydroxyphenylacetaldehyde to yield norcoclaurine; this reaction is catalyzed by the enzyme norcoclaurine synthase (NCS).<sup>2-4</sup> Structural characterization of NCS would be interesting for several reasons, including the ability to illuminate general features of Pictet-Spenglerases, so efforts towards the crystallization of this enzyme were begun.

#### **3.1 Assay design and substrate synthesis**

Before crystallizing an enzyme, it is important to show that it is active to ensure that the structure will provide biologically relevant information. As such, the first step towards crystallizing norcoclaurine synthase was assaying it for function. Several different groups have studied NCS recently and have employed several different assays designs with unique detection methods. These include a radiolabeled TLC assay (Faccini), a chiral LC/MS assay (Sato), and a circular dichroism (CD) spectroscopy based assay (Tanner).<sup>3,5,6</sup> Because of availability of instrumentation and relative ease compared to the other assays (i.e. no radiolabels were necessary), the LC/MS assay was selected for development.

##### **3.1.1 Synthesis of 4-hydroxyphenylacetaldehyde**

The first step towards assaying the enzyme for function was the synthesis of the substrate 4-hydroxyphenylacetaldehyde (4-HPAA). This starting material is unstable and sensitive to

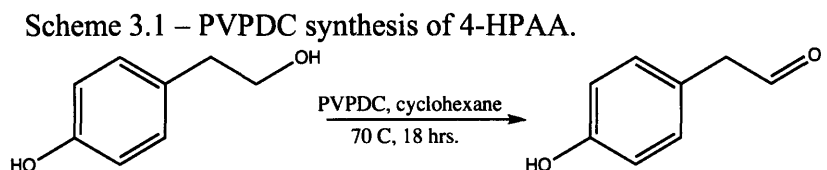


oxygen, which causes oxidation from the aldehyde to the carboxylic acid. There are several different ways of making 4-HPAA. In their assays, Facchini and Tanner use the Doering-Parikh oxidation to oxidize 4-hydroxyphenethyl alcohol. Sato, on the other hand uses an enzymatic method that oxidizes tyramine using commercially available tyramine oxidase from *Arthrobacter* (Sigma). Because of the apparent ease of this method and the simple work-up, Sato's enzymatic oxidation was attempted first to make 4-HPAA.

Using this preparation of 4-HPAA did not result in product formation, either from enzymatic reaction or from non-enzymatic chemical reaction. Since this enzymatic method for making 4-HPAA failed, chemical methods for the synthesis of 4-HPAA were investigated.

One of the traditional routes to the synthesis of aldehydes is oxidation using pyridinium dichromate (PDC). PDC is

a chromium based oxidizing reagent that is

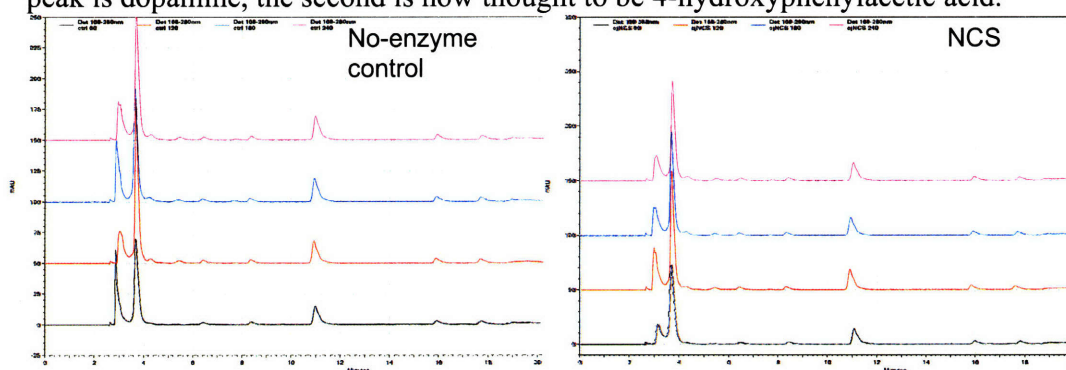


very similar to pyridinium chlorochromate. The drawback to synthesis of 4-HPAA by PDC oxidation is the extensive work-up required to remove the spent chromium reagent. This can be simplified by using a polyvinylpyridinium resin bound PDC, called PVPDC. Work-up for PVPDC oxidations consists of filtering and washing the resin because the product will be soluble in the correct solvents while the catalyst will remain bound to the resin.

Following PVPDC synthesis of the aldehyde, the enzymatic assays were still not yielding the expected results. As before, no reaction was seen, either as a decrease in one of the starting material peaks or a growth of one of the product peaks. The lack of background reaction is most troubling, because dopamine and 4-HPAA should react readily under the assay conditions (37 °C, pH = 7.0). A literature search provided an array of different conditions under which

dopamine and 4-HPAA should undergo chemical reaction.<sup>7-9</sup> Several attempts were made to show background reaction using these conditions, which involved changing the buffer from Tris to phosphate, varying the temperature, and varying the pH. None of the conditions tested showed chemical reaction.

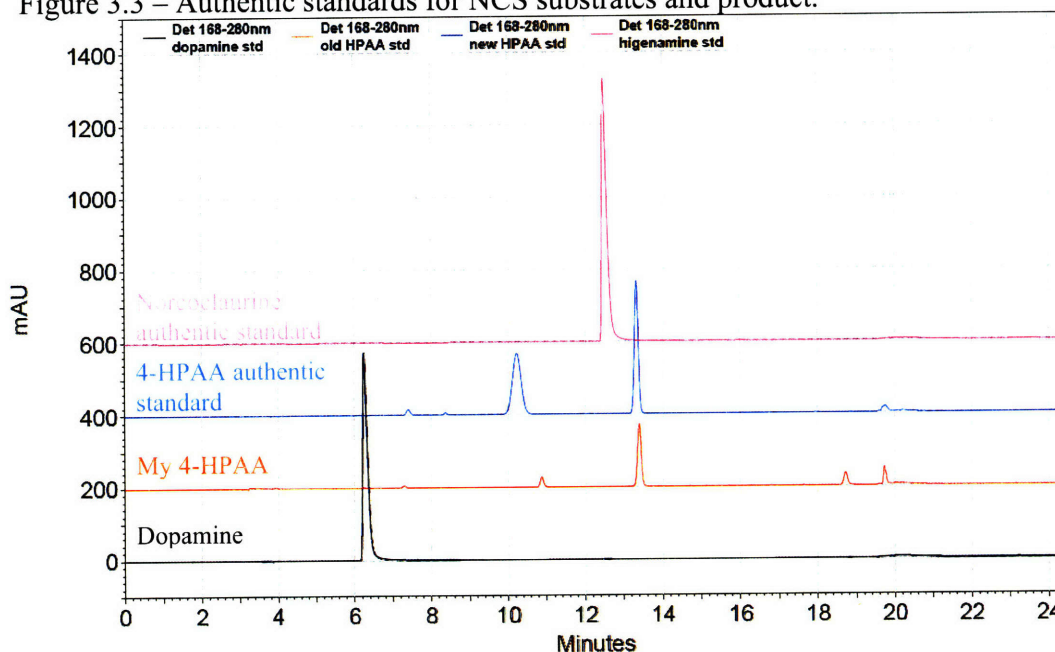
Figure 3.2 – Representative HPLC time course showing lack of activity. Each trace is a time point, progressing from 90 – 240 minutes moving upwards. The first large peak is dopamine, the second is now thought to be 4-hydroxyphenylacetic acid.



The authentic standard for 4-HPAA was purchased from Toronto Research Chemicals.

When obtained, it was dissolved in EtOAc to give a 20 mM stock solution, which was stored at -80 °C. It can be seen from HPLC analysis that there are two peaks, even in the standard. One of these is the aldehyde; the other is most probably the acid. The acid has not been characterized, but it is known that aldehydes such as 4-HPAA will spontaneously oxidize under ambient oxygen concentrations. It can be seen from the HPLC trace in Figure 3.3 that the single peak in my 4-HPAA corresponds to the second peak in the authentic standard. From reaction data, it appears that the first peak is the substrate (it is consumed), while the second peak is the acid (not consumed in the assays). With this new information, it was decided to use the authentic standard in all further assay reactions.

Figure 3.3 – Authentic standards for NCS substrates and product.



### 3.1.2 Monitoring the assay

Initially, it was thought that replicating Sato's LC/MS based assay would give the quickest results. This method uses a chiral CBH (cellobiohydrolase) column with mass spectrometric detection to monitor product formation ( $m/z = 272 [M+H]$ ). The CBH column is run with an isocratic gradient using 0.3% acetic acid buffer adjusted to  $pH = 4.7$ . It is important to use a volatile base when adjusting the  $pH$  of buffers intended for use on the LC/MS; the NaOAc salts formed when NaOH is used do not fly and result in a rapid deterioration of the mass spectrometer. Initial attempts did not show production of product on LC/MS, and it was thought that the reaction components were eluting with the solvent front. To remedy this, the  $pH$  of the running buffer was varied within the limits of the column ( $pH = 4.3 - 5.3$ ). When this did not remedy the problem, it was thought that the product might not fly well with the electrospray ionization method employed. It is now clear that the real problem was with the 4-HPAA starting material; a recently obtained norcoclaurine standard (Sequoia Natural Products) flies well on the

LC/MS. However, at the time it was thought best to transfer the assay to HPLC and employ UV/Vis detection to monitor product formation.

Following transfer of the assay to the HPLC, the first step was to optimize the elution gradient. Standard HPLC buffers were used (Buffer A = 0.1% TFA in H<sub>2</sub>O, Buffer B = Acetonitrile). The first gradient employed went from 10% B to 100% B over 15 minutes (the standard gradient used for separating tryptamine, secologanin, and strictosidine). With this gradient, there was poor resolution of the starting materials. The gradient was optimized to a much shallower 5% B to 30% B over 15 minutes to allow for good separation of the starting materials (and the norcoclaurine standard, once it was obtained). However, even with separation of the starting materials, no product formation was observed. It was at this point that the chemical test reactions were attempted and monitored by HPLC. The failure of these reactions revealed that the lack of product was due to the inability of the starting materials to react, not the monitoring method.

Once the standard for 4-HPAA was received, the enzymatic assays were redone using the two new His-tagged constructs of *T. flavum* NCS (N-terminal and C-terminal tags) and monitored by HPLC. For this new set of experiments, the disappearance of starting material and the appearance of product were clear. However, enzyme controls revealed a substantial background reaction that would complicate any characterization of the kinetic parameters performed using HPLC detection. Lowering substrate concentrations to the detection limit of the instrument did not give reduction in the background reaction as compared to the enzymatic reaction for either construct. It became apparent that the background reaction was too fast to be differentiated from the enzymatic reaction at substrate concentrations amenable to HPLC

visualization. As such, the assay has been moved back to the LC/MS and optimized to allow the enzymatic reaction to be successfully monitored (done by W. Runguphan).

### **3.2 *T. flavum* norcoclaurine synthase**

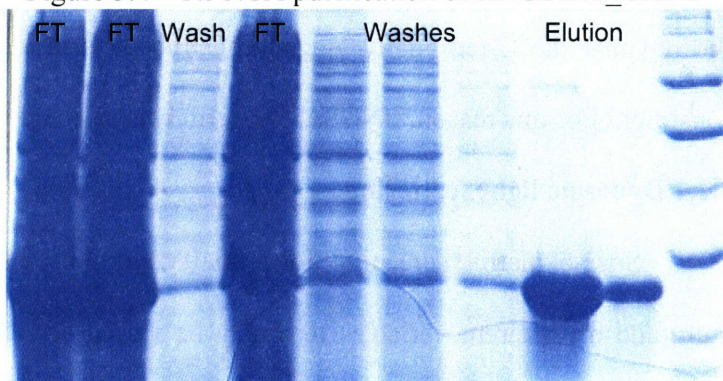
As discussed in the introduction, there are several different forms of NCS. The first NCS isolated comes from plant species *Thalictrum flavum* (meadow rue), and from here on will be referred to as tfNCS. tfNCS is a ~25 kDa proposed dimer with no dependence on any metals or cofactors. It shows no homology to cjNCS, STS, although it does show significant homology (50-60%) with class 10 pathogenesis related proteins (PR10) that seem to play a role in defense against pathogens.<sup>6,10</sup> Most interesting is the lack of homology to cjNCS because these two enzymes have the exact same substrates and function in the same biosynthetic pathway of benzyloisoquinoline alkaloid biosynthesis. A structural characterization of this enzyme could provide insight into general mechanisms of catalysis or general structural motifs for Pictet-Spenglerases. Also, a comparison of the structures of the two apparently unrelated enzymes tfNCS and cjNCS could provide interesting insight into why two radically different enzymes evolved to catalyze the same reaction in the same biosynthetic pathway.

#### **3.2.1 Expression of *T. flavum* NCS**

A synthetic gene encoding the full length tfNCS protein was cloned into the *E. coli* expression vector pET-28a. Following expression and cell lysis, purification was attempted by Ni-NTA affinity chromatography. However, tfNCS was not expressed in *E. coli*. It can be concluded that the 19 amino acid N-terminal signal peptide is an impediment to heterologous expression, even when masked by the N-terminal His tag. It probably would have presented a problem for crystallization as well; signal peptides are often unstructured regions of the protein. To remedy this situation, the signal peptide was removed by site-directed mutagenesis. In

addition to removing the N-terminal signal sequence, multiple other tfNCS constructs were made. These include and N-terminally and C-terminally tagged versions of the truncated protein (called

Figure 3.4 – Ni-NTA purification of tfNCSΔ19\_Cterm



tfNCSΔ19\_Nterm and tfNCSΔ19\_Cterm respectively). The new constructs were expressed, purified, and assayed by W. Runguphan. Following Ni-NTA purification using a Qiagen spin column kit highly pure protein was obtained, as shown in the gel in Figure 3.4.

This protein was assayed for function and monitored by either HPLC (W. Hillmann) or LC/MS (W. Runguphan). Attempts to characterize this enzyme using an HPLC based assay were not successful. With the correct substrate, the background reaction was substantial. Since the HPLC has a relatively poor detection limit ( $\mu\text{M}$ ), high concentrations of the substrates were

required to visualize the reaction.

At these high concentrations, the background reaction proceeded much more quickly than the enzymatic reaction and completely obscured activity.

Fortunately, the mass spectrometer of the LC/MS has a much lower detection limit (pM - nM), so lower concentrations of

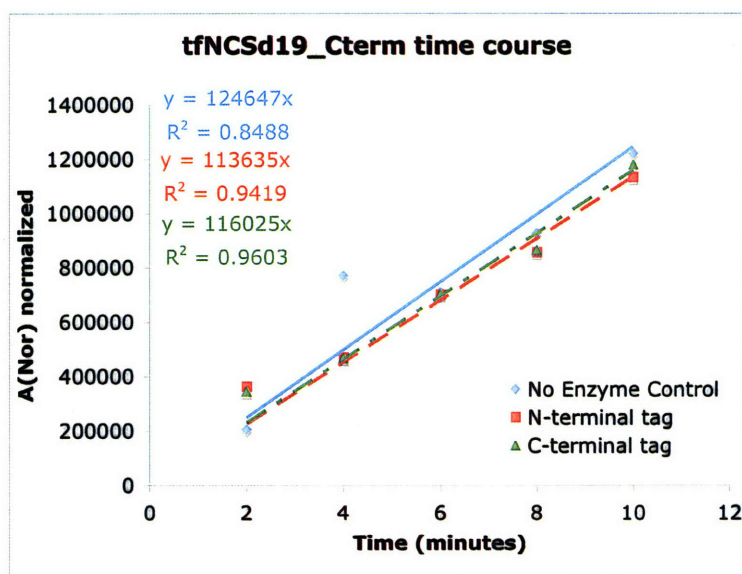


Figure 3.5 – Representative tfNCSΔ19 time course.

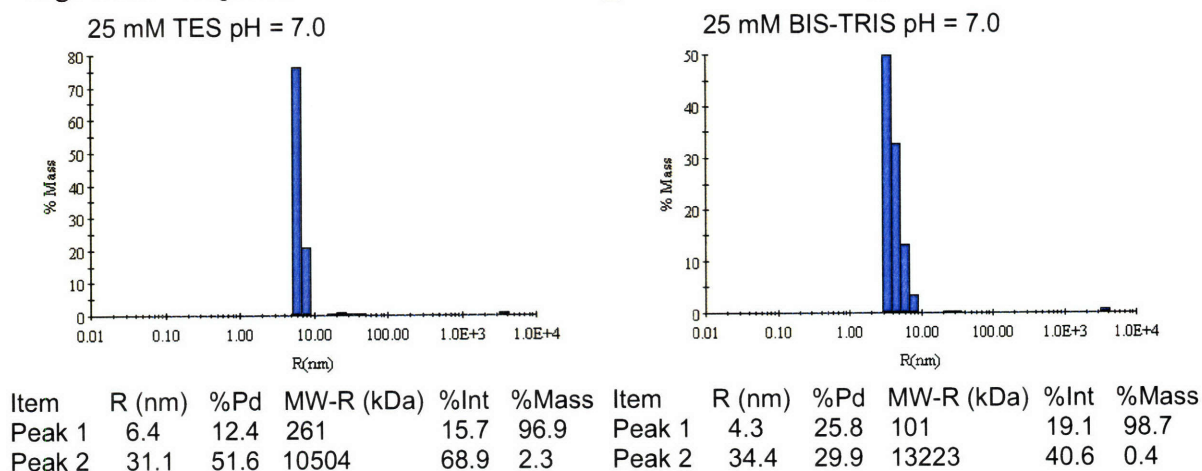


substrates were used. The assays of the two constructs using the optimized LC/MS assay showed that the C-terminally tagged construct was more active than the N-terminally tagged construct. For this reason, tfNCSΔ19\_Cterm was selected for crystallization experiments.

### 3.2.2 Dynamic light scattering

Sparse matrix screening for initial conditions can be a difficult process due to the multitude of different variables that may be varied and optimized. These include the protein concentration (from 2-30 mg/mL), pH of the storage buffer, screening temperature (r.t. or 4 °C), concentration of the commercial screening conditions (full or half strength), mother liquor volume, hanging drop volume, etc. One way to minimize this process is to screen the protein to be crystallized for optimal solubility using dynamic light scattering (DLS). In brief, this technique measures the aggregation properties of small particles in solution. Laser light is passed through a sample and a photomultiplier tube monitors the light scattered at 90°. Analysis of the results can provide several different values; the most important value for crystallography is the

Figure 3.6 – Representative data for tfNCSΔ19\_Cterm DLS screening.



polydispersity.<sup>11</sup> A lower polydispersity corresponds to a more uniform solution. Since more uniform protein preparations have been shown to crystallize better than polydisperse samples,

the goal of DLS is to find the buffer in which the protein has the lowest polydispersity value. For crystallographic purposes, ideal polydispersity values are less than 15%.<sup>12</sup>

In preparation for DLS screening, tfNCSΔ19\_Cterm was buffer exchanged from Ni-NTA elution buffer into 10 mM NaH<sub>2</sub>PO<sub>4</sub> at pH = 7.0 and concentrated to 27 mg/mL (measured by A<sub>280</sub>). This buffer has been used in previous protein preparations used for activity and mechanistic studies.<sup>4-6</sup> The protein was then buffer exchanged into an array of different buffers for screening. Each of these buffers was adjusted to an appropriate pH with the buffer range at which the protein should be stable. The final concentration of the protein was diluted to 0.68 mg/mL with each test buffer. Results of the DLS screening can be seen in Table 3.1

Buffer	pH	Polydispersity (%)	Mass of smallest radius peak
25 mM MOPS	7.0	13.9	97%
25 mM TES	7.0	15.7	96.9%
25 mM Bis-Tris	7.0	25.8	98.9%
25 mM Bicine	7.0	19.2	99.2%
25 mM MES	6.5	13.9	99.6%
25 mM HEPES	7.0	12.1	71.7%
10 mM TrisHCl, 100 mM NaCl	7.5	26.2	98.5%
10 mM NaH <sub>2</sub> PO <sub>4</sub>	7.0	55.8	100%

From these results, it can be seen that several different buffers have polydispersity values in the acceptable range (<15%), including MOPS, TES, MES, and HEPES. However, while all four buffers show a large mass percentage peak at the lowest radius, HEPES has a large higher radius peak which accounts for almost 30% of the mass (DLS cannot accurately assign values for the radii of proteins in solution, but it can provide relative measurements of species that are far enough apart in size). The first and smallest peak is presumed to be tfNCSΔ19\_Cterm and the second peak is presumed to be due to aggregation or microscopic precipitation of the protein. For this reason, only MOPS, TES, and MES were selected for further work. Protein samples were buffer exchanged into the appropriate buffers and concentrated. For MOPS and MES,



protein precipitated rapidly above 1 mg/mL. For TES, the protein was soluble and was ultimately concentrated to 10 mg/mL. This protein was used for crystallization trials.

### 3.3 Crystallization experiments with $\Delta 19$ *T. flavum* NCS

#### 3.3.1 Sparse matrix screening for initial conditions

The protein preparation used for crystallization screens contained Ni-NTA purified tfNCS $\Delta 19$ \_Cterm in 25 mM TES pH = 7.0. The Hampton Research Index Screen (HR-1440 and Hampton Research Crystal Screen 1 (HR-110) and Crystal Screen 2 (HR-112) were initially set up using the hanging drop vapor diffusion method as previously described. Protein at a concentration of 10 mg/mL was initially used. 500  $\mu$ L mother liquor was used in each well with 3  $\mu$ L hanging drops (1.5  $\mu$ L mother liquor, 1.5  $\mu$ L protein).

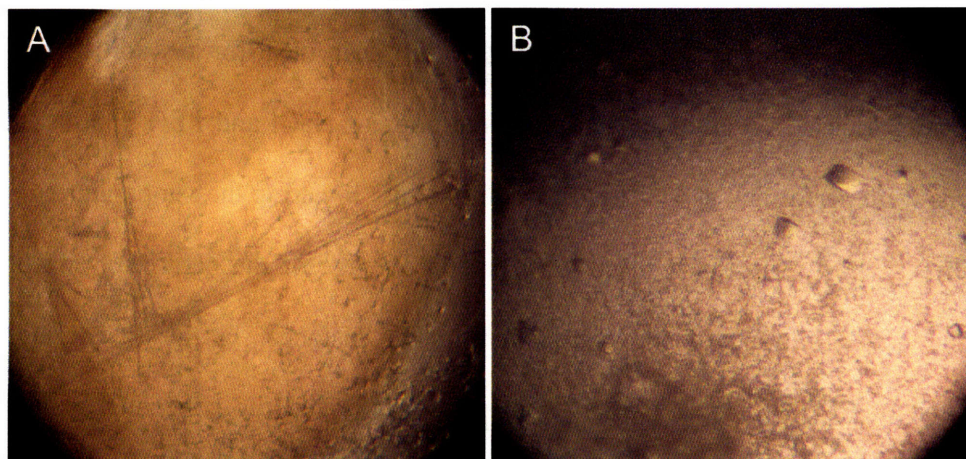
Screen	Concentration (mg/mL)	Temperature	Results
Index	10	r.t.	A large number of wells showed heavy precipitation almost immediately. For the remaining drops, phase separation and granular precipitants were the most common results. <b>Hit:</b> Condition #71 (0.2 M NaCl, 0.1 M Bis-Tris pH = 6.5, 25% w/v PEG 3350)
Crystal Screen 1	10	r.t.	Immediate heavy precipitation was seen in the majority of drops. Some granular precipitates and some clear drops.
Crystal Screen 2	10	r.t.	Immediate heavy precipitation was seen in many drops. Clear drops, phase separation, and granular precipitates dominated the rest.
Index	5	r.t.	Fewer drops showed heavy precipitate, although many still did. Phase separation was still common, as were granular precipitates. More drops remained clear. <b>Hit:</b> Condition #17 (1.26 M NaH <sub>2</sub> PO <sub>4</sub> •H <sub>2</sub> O, 0.14 M K <sub>2</sub> HPO <sub>4</sub> , pH = 5.6)

Immediately after setting up the initial screens, heavy precipitate was observed in a large number (~50%) of wells, so the Index screen was repeated at a lower protein concentration of 5 mg/mL.

Screens were monitored at 2, 6, and 9 days (Index and Crystal screens at 10 mg/mL) or 1, 4, and 6 days (Index screen at 5 mg/mL). Results of the initial screening are summarized in Table 3.2.

Several hits were obtained from the initial screening of tfNCSΔ19\_Cterm. The first was observed after six days in the 10 mg/mL Index screen. After two days, this well showed clear phase separation. Over the course of the next several days, the protein that was not consumed in the initial phase separation crystallized, giving crystals with a long needle-like morphology. One such needle cluster was observed after six days, with two more seen following observation on day nine. The other hit was seen after one day in the 5 mg/mL Index screen. Upon the first observation, small spherules were seen in the drop. Over the next several days, these grew larger and began to show a rectangular three-dimensional morphology (small prisms). Notably, the lower concentration Index screen did not show the same hit as the higher concentration screen in condition 71. The second screen showed phase separation as the first, but there was evidently not enough protein to continue on to form crystals at the lower concentration. Since these hits represented the best results from the initial screens, both were further optimized.

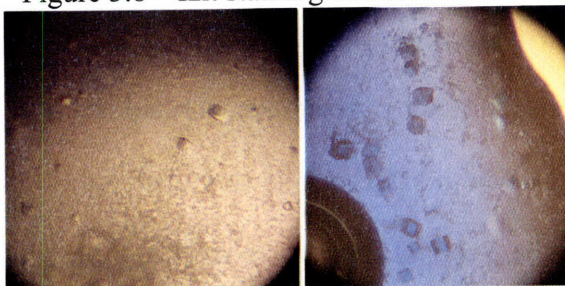
Figure 3.7 – Initial hits for tfNCSΔ19\_Cterm. a) Index #71, 10 mg/mL;  
b) Index 17, 5 mg/mL



### 3.3.2 Optimization of initial hits

Because of their superior morphology, initial grid screening efforts were focused around the hit from condition 17 at 5 mg/mL. The crystals produced in the initial hit did not stain well

Figure 3.8 – IZIT staining of Index 17



with IZIT dye (as shown in Figure 3.8), and they were not large enough to use for SDS-PAGE analysis. It was hoped that reproducing and optimizing the crystals would yield enough for analysis and characterization. A grid screen was

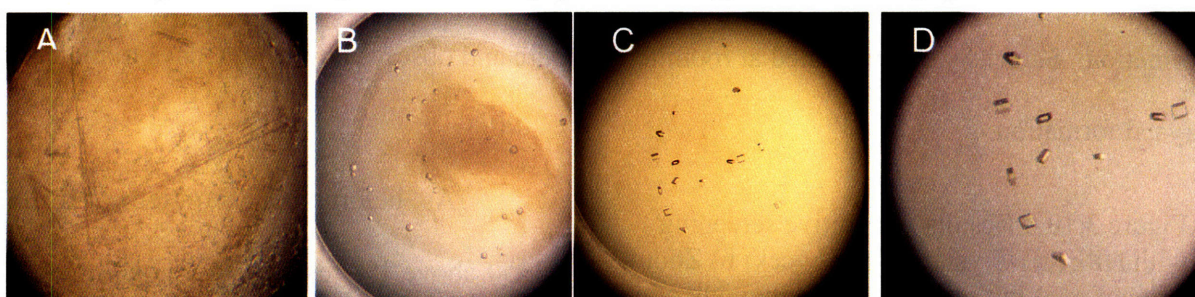
constructed in a 24 well plate and featured wells with the initial hit mother liquor and progressive dilutions of the mother liquor solution (90% mother liquor, 70%, 50%, 30%, and 10%). In addition to the range of dilutions, both 3  $\mu$ L total and 4  $\mu$ L total drops were used (4  $\mu$ L drops = 2  $\mu$ L protein solution, 2  $\mu$ L mother liquor). Larger drops have longer equilibration times and can sometimes lead to better and larger crystal growth. Unfortunately, none of the conditions resulted in reproduction of the crystal.

In view of the failure to reproduce the initial hit from condition 17, the hit from condition 71 at 10 mg/mL was investigated. A main goal of this optimization was to make the crystals larger and improve morphology. Thin needles are less than ideal for any sort of X-ray studies. Grid screening focused on adjusting the concentrations of NaCl and PEG 3350. Also, several grid screens were performed at 4 °C to test whether or not lowering the temperature would have a positive effect. Early screens focused on a broad coverage of the parameter space by widely varying the concentrations of the components. For instance, in the first grid screen the concentration of PEG 3350 was varied from 10% - 25% and the concentration of NaCl was varied from 50 mM – 200 mM. Results from this screen showed that lower concentrations of



both PEG 3350 and NaCl resulted in improved crystallization. Notably, at lower precipitant concentrations the amount of phase separation decreased until it no longer occurred. Also, at lower precipitant and lower salt concentrations the morphology of the crystals changed from thin needles to small, yet three dimensional, prisms. Conditions were optimized from 200 mM NaCl, 20% PEG 3350 to 0 mM NaCl, 1% PEG 3350. The optimized conditions yielded small crystals

Figure 3.9 – Improvement of tfNCSΔ19\_Cterm crystals by grid screening. a) Initial hit; b) 50 mM NaCl, 5% PEG 3350; c) 0 mM NaCl, 1% PEG 3350; d) Zoom view of c.



(~ 20  $\mu\text{m}$  x 10  $\mu\text{m}$  x 10  $\mu\text{m}$ ) with the predominant morphology being that of rectangular prisms.

The progression of the grid screening is illustrated in Figure 3.9.

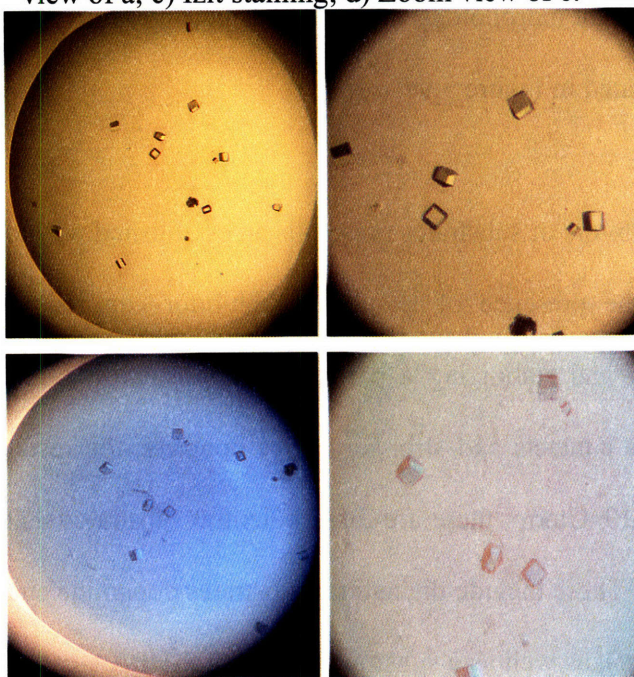
### 3.3.3 Co-crystallization experiments

Co-crystallization of proteins with inhibitors or substrates present is an established way to improve crystallization. This is because the presence of a substrate/inhibitor can reduce structural disorder by binding to the protein and stabilizing a particular conformation. The stabilized protein can then make more lattice contacts and stronger lattice contacts; the result being better and larger crystals. For tfNCSΔ19\_Cterm, there are three potential substrates that can be used in co-crystallization experiments. These include the natural substrates dopamine and 4-HPAA (in different drops to prevent turnover), as well as a substrate analog of 4-HPAA called 2-(4-hydroxyphenyl)-ethanol (HPE). This substrate analog is a reduced form of the aldehyde substrate 4-HPAA and is commercially available. In the recent NMR based homology model of

tfNCS, this substrate was shown to induce substantial chemical shifts, so could provide much needed stability to the protein.<sup>13</sup>

The co-crystallization experiments were performed in the same manner as previously described grid screens. Several hours before setting up trays, the protein was incubated with the desired additive (1 mM dopamine, 1 mM 4-HPAA, 1 mM HPE) for several hours at 4 °C. In addition to the new additives, a range of PEG 3350 concentrations were investigated, from 0.1% - 6 %. Finally, drops were also set up with native protein (no additive) as a control for crystal formation. Screens were incubated at room temperature and observed every few days to monitor crystal formation.

Figure 3.10 – Results of additive screening with HPE. a) Crystals after 1 week; b) Zoom view of a; c) IZIT staining; d) Zoom view of c.



As before, concentration of PEG 3350 above 1% resulted in a large number (~100) of very small crystals. PEG 3350 concentrations below 1% resulted in negligible crystal growth within one week (the time scale of previous crystal formation). It seems that 1% PEG 3350 is the most optimal concentration. As expected, the drops set up without any additives crystallized in the same way as before, giving small crystals. Drops with 4-HPAA showed heavy precipitation. This

is not surprising, since 4-HPAA reacts readily with oxygen and probably degraded rapidly in the well. Also, 4-HPAA is only soluble in organic solvent, so a small amount of ethyl acetate was

present in the drop. Since the 4-HPAA was added from a 20 mM stock solution, there was only ~0.5% EtOAc in the protein and ~0.25 % in the drop. However, this was obviously enough to negatively affect the crystallization properties.

In contrast to the drops with 4-HPAA, those with dopamine and HPE showed marked improvement in crystallization. Crystals formed first in the drops with HPE (Figure 3.10). These were much larger than the native crystals (approximate dimensions of 75  $\mu\text{m}$  x 50  $\mu\text{m}$  x 10  $\mu\text{m}$  including the additive). Like previous tfNCS $\Delta$ 19\_Cterm crystals, these crystals did not stain well with Izit dye, although the crystals did seem to develop a darker blue color after an extended incubation (several days). Unfortunately, following the addition of the Izit dye the drop showed no further crystal growth (either new crystals or existing crystals growing larger). For the dopamine containing drops, crystals took several additional days to form. These crystals were allowed to grow without disturbance and appeared to become slightly larger over the course of time. The dimensions after ~two weeks were approximately 100  $\mu\text{m}$  x 50  $\mu\text{m}$  x 20  $\mu\text{m}$ . Morphology of these crystals was consistent with the previously observed morphology for the native crystals and for the HPE crystals. At this time, characterization of the crystals by SDS-PAGE would confirm that they are in fact protein crystals. However, until high quality crystals of this nature are reproduced, sacrificing crystals for gel analysis would not be prudent. In light of these results, multiple drops were set up containing 0 mM NaCl, 1% PEG 3350, either 1 mM dopamine, 1 mM HPE, or 1 mM dopamine + 1 mM HPE. The results of this screen are pending.

### **3.4 C. japonica norcoclaurine synthase**

In 2007, a new version of NCS was isolated from *Coptis japonica* cell lines selected for their high alkaloid production.<sup>5</sup> This new form of NCS, referred to as cjNCS from here on, is a ~40 kDa monomer with some very interesting features. It showed no sequence homology to

previously characterized forms of NCS or to STS, however it did show homology to 2-oxoglutarate dependant non-heme iron monooxygenases. Assays of this enzyme (both native and heterologously expressed) showed that its activity was dependant on ferrous iron, but not on 2-oxoglutarate and oxygen. This is an extremely unusual and unprecedented result.

Structural characterization of cjNCS should yield a great deal of insight into its mechanism. For instance, what is the role of iron? Is it catalytic, structural, or does it facilitate proper binding of substrates? Since STS and NCS are the only known Pictet-Spenglerases, characterization of NCS should also allow for some generalizations to be made about Pictet-Spenglerases in general. Despite the different sequences are there common structural motifs in this class of enzymes shares? For these reasons, attempts towards a structural characterization of cjNCS were begun.

#### **3.4.1 Expression and thrombin cleavage of *C. japonica* NCS**

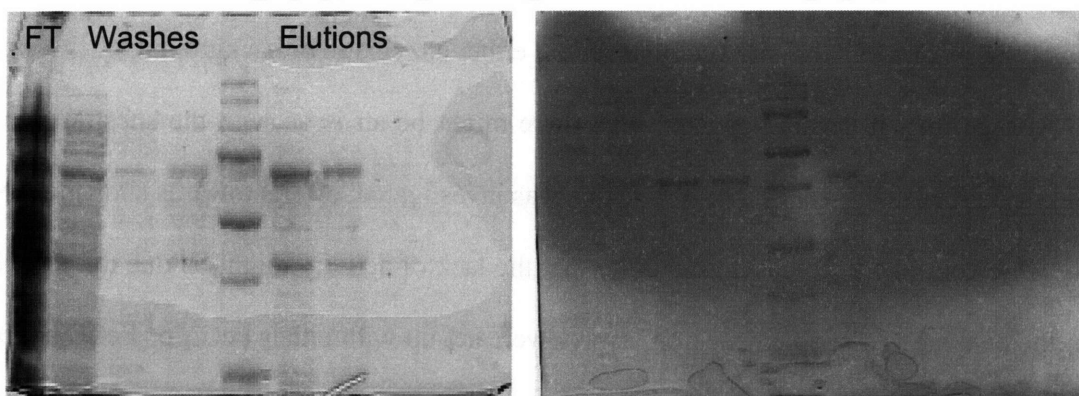
Unlike tfNCS, cjNCS did not feature a predicted N-terminal signal peptide. The full-length gene was ordered (Geneart) and cloned into the *E. coli* expression vector pET-28a with an N-terminal hexa-histidine tag and a thrombin cleavage site for affinity purification tag removal. Expression and purification of cjNCS proceeded uneventfully. The protein over-expressed well under IPTG induction and was purified to electrophoretic homogeneity by Ni-NTA metal affinity chromatography, MonoQ anion exchange chromatography, and SE200 size exclusion chromatography. Gels showing the various stages of purification are shown in Figure 3.11.

In preparation for activity assays and crystallization, the His-tag was removed by thrombin cleavage. This was accomplished using agarose bound thrombin kit as described for rsSTS. SDS-PAGE analysis of samples collected at 1, 3, 18, and 24 hours revealed incomplete cleavage of the tag (as measured by a decrease in apparent molecular weight of the cleaved



protein). The reaction was purified by passing the protein over an Ni-NTA column and collecting the flowthrough (flowthrough should contain cleaved protein, while any un-cleaved protein should remain bound to the column). This protein was concentrated to 2 mg/mL in 50

Figure 3.11 – Purification of cjNCS. The left gel shows fractions after Ni-NTA chromatography, the right after gel filtration chromatography.



mM Tris-HCl pH = 7.5 (verified by both  $A_{280}$  and Bradford assay) and used for activity assays.

### 3.4.2 Enzyme activity assays

Initially, it was thought that it would be a simple task to repeat the assays presented in by Sato.<sup>5</sup> Even though cjNCS is a non-heme iron protein, Sato did not perform any reconstitution steps following purification of heterologously expressed protein from *E. coli*. Also, the assay procedures seemed to be fairly straightforward. A full kinetic characterization was not the goal of these assays. Simply showing activity would be sufficient before moving on to crystallizing cjNCS. Following the purification procedure, both cleaved and un-cleaved cjNCS were subjected to enzymatic activity assays.

Initially, no activity was seen by LC/MS. Several different buffers were used as previously described. However, no appreciable difference was seen based on this change. It was at this time that the assay monitoring platform was switched from the LC/MS to the HPLC. After optimizing the elution gradient to separate the starting materials, there was still no product



formation seen. At this point, a variety of different quenches were investigated. These included a TFA quench, a methanol quench, and a temperature quench. Although there were no differences visible by HPLC, the temperature quench was selected as the least likely to perturb and obscure the system. Also, the methanol quench substantially diluted the reaction making product visualization more difficult. Assays with increased substrate concentrations (up to 5 mM for one of the substrates) also showed no reaction, either enzymatic or background.

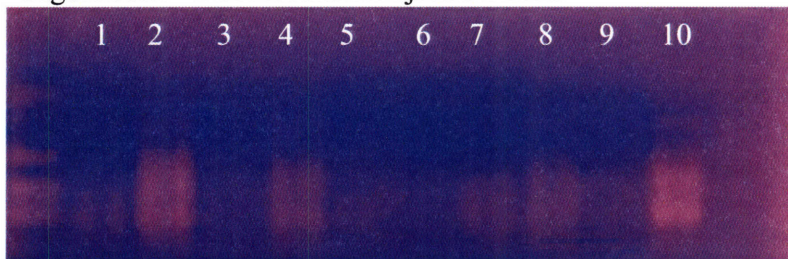
At this time, it became apparent that there might be an issue with the substrates in the reaction (specifically the 4-HPAA). While this was investigated, efforts towards showing cjNCS activity were continued. It was a possibility that the lack of activity might be due to the lack of iron in the enzyme. To supplement iron, assays were set up with either FeCl<sub>2</sub> or FeCl<sub>3</sub> added to supplement the enzyme. HPLC analysis showed a mixture of products, however nothing which corresponded to the norcoclaurine standard. It was concluded that if lack of iron was the problem, a more sophisticated reconstitution procedure would be required.

In retrospect, the reason why there was no reaction is that the 4-HPAA was actually the acid 4-hydroxyphenylacetic acid. Not knowing this, several new constructs of cjNCS were planned. These included an N-terminal and C-terminal FLAG tag (a non-metal binding affinity tag), as well as an untagged version of the protein (the initial published heterologous expression was of an untagged construct).

The new constructs were made by cloning the cjNCS gene out of the pET-28a plasmid using primers with the desired tag appended on to the desired end (N-terminal or C-terminal). After the PCR was complete, each sample was purified using a Qiagen PCR clean-up kit and analyzed by 1% agarose gel electrophoresis, shown in Figure 3.12. Samples that showed DNA fragments of the appropriate sized were selected and sub-cloned into a pGEM-T Easy transfer

vector. The pGEM-T Easy system includes a convenient blue/white screening method for selecting successful ligations. The pGEM-T vector contains the LacZ gene (beta-galactosidase).

Figure 3.12 – PCR results for cjNCS-FLAG

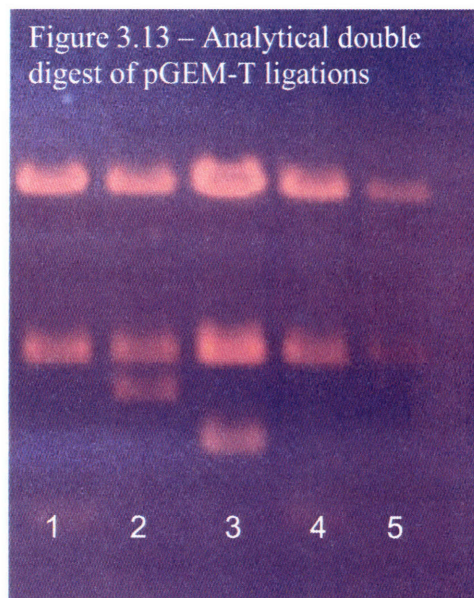


When ligation is successful, the gene insert disrupts LacZ; when the plasmid closes on itself, the LacZ gene remains intact. When grown on media

pre-treated with IPTG and X-Gal, colonies from successful ligations will be white while those with closed plasmid will be blue. White colonies were selected and cultured. The plasmid DNA was isolated by miniprep and analyzed by a restriction double digest, shown in Figure 3.13. Successful samples (lanes 1, 4, and 5 in the figure) were sequenced to verify the success of the cloning. A mutation in the FLAG tag sequence was identified and remedied by site-directed mutagenesis. Repeat sequencing verified the integrity of the gene.

The N-terminal FLAG tag, C-terminal FLAG tag, and untagged constructs were expressed and partially purified by W. Runguphan. All three of these constructs show catalytic activity above the background reaction, with the untagged construct and the C-terminally tagged construct showing slightly better activity than the N-terminally tagged construct.

Figure 3.13 – Analytical double digest of pGEM-T ligations



Monitoring the reaction with a CBH chiral column shows that the enzyme produces (S)-norcoclaurine stereoselectively. Interestingly, this stereoselectivity seems to decrease over time. Investigations into this phenomenon are on going.

### **3.5 Summary and future work**

Two different forms of norcoclaurine synthase have been studied with the ultimate goal of structural characterization of Pictet-Spenglerases. Each enzyme, tfNCS and cjNCS, has presented its own unique set of issues, and progress has been made towards crystallizing these enzymes. The first step was the design and optimization of an assay method. 4-HPAA, the aldehyde substrate, was difficult to obtain. Enzymatic and synthetic methods were attempted before a standard became commercially available. Many of the early difficulties experienced working with these enzymes may have been due to the inclusion of an oxidized substrate instead of the proper aldehyde in the reaction mixtures. Also important was determining a proper method to monitor the assay. The assays began on the LC/MS, moved to the HPLC, and then back to the LC/MS when it became apparent that the background reaction would obscure enzymatic activity on the HPLC. With a reliable assay, progress has moved forward much more rapidly on this project.

Initially, expression of tfNCS was attempted with the full-length gene. The presence of a 19 amino acid N-terminal signal peptide prevented expression in *E. coli* (and likely would have negatively impacted crystallization). After the removal of this signal peptide by site-directed mutagenesis, two different constructs were cloned featuring N- and C-terminal hexa-histidine tags. Assays revealed that the C-terminally tagged construct was more active, so it was selected for crystallization studies. Sparse matrix screening identified several hits, one of which was optimized through grid screening and co-crystallization to produce crystals with rectangular

prism morphology that are approximately 100 x 50 x 20  $\mu\text{m}$ . Future work with this project involves full characterization of these crystals and screening for diffraction. Once high quality diffraction data is obtained, molecular replacement may be performed with the recent NMR based homology model of tfNCS or birch pollen allergen Bet v1 (the template for the homology model). If molecular replacement does not work, phase information may be obtained through Se-Met derived protein for MAD/SAD experiments or using heavy atom derivatives for isomorphous replacement. Recently, a report detailing tfNCS crystallization has been published, however there is no structure available yet.<sup>14</sup> As such, work on this project will continue.

Following cloning, expression of cjNCS and the thrombin cleavage of the N-terminal His-tag were completed. Activity assays of the cleaved protein did not exhibit any catalytic activity. Despite trying to vary and optimize many different factors of the assay (monitoring method, chromatography conditions, quenching method, addition of exogenous iron), cjNCS did not appear to function at any time during the assays. Several new constructs were cloned and subjected to the assay procedures that were used to show tfNCS $\Delta$ 19 activity. All three of them show activity above the background reaction, but there are unresolved issues with the assay, namely the loss of stereoselectivity over time. Future work with this enzyme involves a more detailed study of its activity, including determining why the stereoselectivity decreases. Also, the purification needs to be optimized to yield crystallography quality protein. Once this is completed, sparse matrix screens can begin to identify initial conditions for crystallization. Since only one Fe is present per protein molecule, Fe MAD would probably not give sufficient phase information to solve the structure. However, Se-Met derived protein or heavy atom derivatives should give this information.

### **3.6 Materials and Methods**

#### **Enzymatic Synthesis of 4-HPAA**

Tyramine oxidase (Sigma) was stored as a 2 U/mL stock solution in 50 mM Tris-HCl, pH = 7.5 at 4 °C. An aliquot was also frozen and stored at -20 °C for long-term storage. To make 4-HPAA, 100 µL 2 U/mL tyramine oxidase was added to a 4 mM solution of tyramine. This was incubated at 37 °C for 2 hours. The reaction was then boiled for 5 minutes and centrifuged to remove the enzyme. The reaction was assumed to proceed with a quantitative yield; the final mixture (after boiling) was frozen at -80 °C to be used as a 2 mM 4-HPAA stock solution.

#### **Synthesis of 4-HPAA by PDC Oxidation**

PVPDC resin (Fluka) contains 2.2 mmol/g chromate. 1.9 g resin (4.2 mmol chromate) was suspended in water and stirred vigorously for 2-3 hours to properly wet the resin. The water was filtered off and the resin was washed with water to removed unbound chromate. The moist resin was added to a 50 mL round bottom flask and 15 mL cyclohexane and a Teflon coated stir bar were added. 143 mg (1.03 mmol) 4-hydroxyphenethyl alcohol was added to this. The flask was sealed under inert atmosphere and refluxed at 70 °C for 16 hrs. Resin was removed by filtration and washed with dichloromethane and ethyl acetate. This was concentrated to give a yellow oil that was further purified by flash silica gel chromatography in 9:1 chloroform:methanol. Three spots were evident on TLC and each was isolated. The TLC plates were visualized with UV light and by DNP aldehyde stain. Only one spot stained when exposed to DNP; this spot was isolated and the solvent removed under reduced pressure. This resulted in 89 mg product, for a yield of 64%. This product was dissolved in a 25% DMSO/water mixture

to give a 20 mM stock solution that was stored at -80 °C. A proton NMR shows what appears to be the correct product, although there are some unidentified peaks that are troubling.

### **General enzymatic assays of NCS**

Several different NCS assays were used over the course of these experiments. As procedures were optimized, the assays changed radically. The final optimized assay employed for HPLC detection is as follows:

General NCS assay: 100  $\mu$ L 200 mM NaH<sub>2</sub>PO<sub>4</sub> pH = 7.00  
10  $\mu$ L 20 mM Dopamine  
10  $\mu$ L 20 mM 4-HPAA  
20  $\mu$ L enzyme/buffer  
60  $\mu$ L ddH<sub>2</sub>O

These assays were run in 1.5 mL Eppendorf tubes and incubated at 37 °C until being quenched by reduced temperature. Following quenching, the results were analyzed by HPLC with a gradient elution from 5% - 30 % B over 15 minutes. For the optimized LC/MS assay, the reaction was quenched with MeOH and diluted to ~ 1 nM concentrations (in MeOH) for analysis (W. Runguphan).

### **Chemical test reactions of dopamine and 4-HPAA**

The continued failure of the assays prompted verification of the aldehyde starting material by reacting it with dopamine in a chemical test reaction. A literature search gave several different protocols for reacting dopamine and a range of different aldehydes in Pictet-Spengler reaction. Each test reaction had 1 mM dopamine and 1 mM 4-HPAA in addition to buffer or additional solvent. These buffers/solvents included the following:

1. MeOH
2. EtOH
3. 0.1 M NaH<sub>2</sub>PO<sub>4</sub>, pH = 6.0
4. 0.1 M NaH<sub>2</sub>PO<sub>4</sub>, pH = 6.5
5. 0.1 M NaH<sub>2</sub>PO<sub>4</sub>, pH = 7.0
6. 0.1 M NaH<sub>2</sub>PO<sub>4</sub>, pH = 7.5

The reactions were run overnight at room temperature and at 37 °C to give twelve different conditions. In addition, the MeOH, EtOH, 0.1 M NaH<sub>2</sub>PO<sub>4</sub>, pH = 6.5, and 0.1 M NaH<sub>2</sub>PO<sub>4</sub>, pH = 7.0 reactions were run at 95 °C for one hour to give sixteen total reaction conditions. Each reaction was analyzed by HPLC using a gradient elution of 5% B – 30% B over fifteen minutes. None of the test reactions showed norcoclaurine formation, and reactions run at higher temperature and higher pH turned brown and precipitated (presumably due to dopamine oxidation/polymerization).

### **Cloning tfNCS into pET28-a**

The *E. coli* codon-optimized gene for tfNCS was purchased from Genent. It was obtained in a pGA4 plasmid with the ampicillin resistance gene. A restriction double digest was performed to obtain the gene insert. The target vector pET28-a was also digested to prepare for ligation. The following reaction mixture was used:

Restriction digest: 19 µL sterile ddH<sub>2</sub>O  
5 µL 10x NEB Buffer 2  
20 µL plasmid DNA (pGA4\_tfNCS or pET-28a)  
3 µL NdeI  
3 µL EcoRI

The reactions were incubated at 37 °C for 3 hrs. Following incubation, 12.5 µL 5x DNA loading dye was added and the entire reaction was loaded onto a 1% agarose gel and run at 100 V for 45 minutes. After staining with ethidium bromide for 1 hr, the tfNCS gene fragment and the cut pET-28a plasmid were isolated using a gel extraction kit (Qiagen).

Following this, the gene fragment was ligated into pET-28a. The following ligation mixture was successful after several trials:

**Ligation:** 6  $\mu$ L tfNCS insert (or H<sub>2</sub>O in the control)  
2  $\mu$ L freshly cut pET28-a  
1  $\mu$ L 10x T4 DNA ligase buffer  
1  $\mu$ L T4 DNA ligase

This reaction mixture was incubated in a micro-cooler at 20 °C overnight. Following the ligations, 1  $\mu$ L of the reaction mixture was added to 55  $\mu$ L electrocompetent *E. coli* Top 10 cells and incubated on ice for 2 minutes. Transformation was achieved using a Bio-rad Micropulser on the E2 setting. Following transformation, 250  $\mu$ L SOC media was added and cells were recovered on ice for ~5 min. After recovery, cells were grown shaking at 37 °C for 1 hr. before the entire 250  $\mu$ L was plated on LB Agar (Kan) plates. Plates were incubated for ~18 hrs. at 37 °C before colonies were picked and grown up in 5 mL liquid cultures with 30  $\mu$ g/mL Kan.

5 mL liquid cultures were grown overnight shaking vigorously at 37 °C. Plasmid DNA was isolated using a commercial miniprep kit (Qiagen). An analytical double digest using the previously described reaction mixture verified presence of the rsSTS gene insert. The sequence was verified as previously described using the T7 promoter and T7 terminator primers present in the pET28-a vector.

### **Expression and purification of tfNCS**

The following buffers were used during the Ni-NTA purification of tfNCS:

**Lysis Buffer:** 50 mM NaH<sub>2</sub>PO<sub>4</sub>  
300 mM NaCl  
10 mM Imidazole  
pH = 8.0

**Wash Buffer:** 50 mM NaH<sub>2</sub>PO<sub>4</sub>  
300 mM NaCl  
20 mM Imidazole  
pH = 8.0

**Elution Buffer:** 50 mM NaH<sub>2</sub>PO<sub>4</sub>  
300 mM NaCl  
250 mM Imidazole  
pH = 8.0

tfNCS was cultured and isolated similarly to all previously discussed proteins. The vector pET-28a\_tfNCS was transformed into *E. coli* BL21 (DE3) cells by electroporation. Cells



were grown shaking at 37 °C for 1 hr. before plating on an LB-Agar (Kan) plate and incubation overnight at 37 °C. A 5 mL LB broth (1x Kan) starter culture was inoculated with cells from a single colony. This culture was grown overnight shaking at 37 °C. Following overnight starter culture growth, 1 L LB broth (1x Kan) was inoculated. The 1 L culture was incubated shaking at 37 °C until an O.D. of between 0.5-0.7 was reached. Once O.D. had been reached, the temperature was dropped to 16 °C and the culture was induced with IPTG. The cultures were grown overnight shaking at 16 °C. In the morning, cells were harvested by pelleting in the centrifuge at 4500 rpm for 15 min. Pellets were either used fresh or frozen and stored at -20 °C until future use.

To obtain tfNCS, cells were lysed using lysozyme and sonication in a manner similar to that described previously. After lysis, cell debris was pelleted by centrifugation at 13,000 rpm for 55 minutes. Purification by Ni-NTA affinity chromatography was performed by the standard procedure. After collection of the flow-through, the column was washed with 100-200 mL wash buffer to remove weakly or non-specifically bound proteins. Finally, tfSTS was eluted from the column with 3x 10 mL 250 mM imidazole elution buffer. Fractions were analyzed by SDS-PAGE.

#### **Activity assays of tfNCSΔ19\_Cterm and tfNCSΔ19\_Nterm**

The following reaction mixture was used: to assay the function of the new constructs of tfNCS:

tfNCSΔ19 Assays: 100 μL 200 mM NaH<sub>2</sub>PO<sub>4</sub> pH = 7.00  
10 μL 20 mM Dopamine  
10 μL 20 mM 4-HPAA (the newly purchased standard)  
20 μL tfNCSd19\_Nterm/tfNCSd19\_Cterm/buffer  
60 μL ddH<sub>2</sub>O

These assays were incubated at 37 °C until quenching. Reactions were stopped by temperature quench and analyzed by HPLC utilizing a gradient elution proceeding from 5% B to 30% B (A = 0.1% TFA in H<sub>2</sub>O, B = Acetonitrile) over 15 minutes.

To determine which enzyme was more active, a time course was run using the same reaction mixture described above. 33 μL aliquots were taken at 2, 4, 6, 8, and 10 minutes and cold quenched. To aid in quantitation, a naphthylacetic acid (NAA) internal standard was included. By normalizing to the NAA peak as done with the crSTS kinetic characterization, errors such as injection error could be corrected for. Plotting the normalized area under the curve for norcochlorine did not appear to show any difference between either enzymatic reaction or the background reactions. This could be remedied by reducing the substrate concentrations, however this was found to drop the assay below the level of reliable quantitation for the HPLC.

#### **Dynamic Light Scattering of tfNCSΔ19\_Cterm**

tfNCSΔ19\_Cterm was obtained in Ni-NTA elution buffer. It was buffer exchanged into 10 mM NaH<sub>2</sub>PO<sub>4</sub>, pH = 7.0 and concentrated to 27.5 mg/mL as measured by A<sub>280</sub>. Notably, no precipitation was present even at this high concentration. 1M stock solutions were made for each buffer to be tested and adjusted to the appropriate pH using 4 N NaOH. Before use, the 1 M stock buffers were diluted to 25 mM concentrations. Using the 25 mM test buffers, the tfNCSΔ19\_Cterm was diluted to a final concentration of 0.68 mg/mL (A<sub>280</sub>). No precipitation was observed in any of the samples. Before acquiring data, all protein/buffer solutions spun at 13,000 rpm for five minutes to remove any dust particles. Samples were equilibrated at room temperature for 10 minutes.

Data was collected using a DynaPro Titan Dynamic Light Scatterer. A 45 μL quartz cuvette was used for each 55 μL sample. Between collecting data on each sample, the cuvette

was washed with water and ethanol and dried with dry N<sub>2</sub>. The experimental parameters included 10 acquisitions over a 10 s acquisition period (1 s/acquisition). The laser power was modulated to obtain the optimal value of 800,000 counts per second (or as close as possible). Once a stable value was reached, data collection was begun. Due to the dilute nature of the buffers, a H<sub>2</sub>O solvent model was employed in data processing. The quality of each data collection was judged on the basis of the auto-correlation function. If the auto-correlation function was satisfactory, data processing was performed. Data was analyzed using the Dynamics v6.7.7.9 software. Values of interest were the polydispersity and %mass of the smallest peak (purity). Buffers with low polydispersity were selected for use in crystallization screens.

### **Cloning cjNCS into pET28-a**

The *E. coli* codon-optimized gene for cjNCS was purchased from Genart. There is no predicted signal sequence for cjNCS, so a full-length gene was ordered. It was obtained in a pGA4 plasmid with the ampicillin resistance gene. Following amplification and isolation (by miniprep) of this plasmid from *E. coli* Top10 cells, a double digest was performed to obtain the gene insert. The target vector pET28-a was digested simultaneously to prepare for the eventual ligation. The following reaction mixture was used:

Restriction digest: 19  $\mu$ L sterile ddH<sub>2</sub>O  
5  $\mu$ L 10x NEB Buffer 2  
20  $\mu$ L plasmid DNA (pGA4\_cjNCS or pET-28a)  
3  $\mu$ L NdeI  
3  $\mu$ L EcoRI

The reactions were incubated at 37 °C for 3 hrs. Following incubation, 12.5  $\mu$ L 5x DNA loading dye was added and the entire reaction was loaded onto a 1% agarose gel and run at 100

V for 45 minutes. After staining with ethidium bromide for 1 hr, the cjNCS gene fragment and the cut pET-28a plasmid were isolated using a gel extraction kit (Qiagen).

Following this, the gene fragment was ligated into pET-28a. The following ligation mixture was successful after several trials:

Ligation: 6  $\mu$ L cjNCS insert (or H<sub>2</sub>O in the control)  
2  $\mu$ L freshly cut pET28-a  
1  $\mu$ L 10x T4 DNA ligase buffer  
1  $\mu$ L T4 DNA ligase

This reaction mixture was incubated at 18 °C overnight. Following the ligations, 1  $\mu$ L of the reaction mixture was added to 55  $\mu$ L electrocompetent *E. coli* Top 10 cells and incubated on ice for 2 minutes. Transformation was achieved using a Bio-rad Micropulser on the E2 setting. Following transformation, 250  $\mu$ L SOC media was added and cells were recovered on ice for ~5 min. After recovery, cells were grown shaking at 37 °C for 1 hr. before the entire 250  $\mu$ L was plated on LB Agar (Kan) plates. Plates were incubated for ~18 hrs. at 37 °C before colonies were picked and grown up in 5 mL liquid cultures with 30  $\mu$ g/mL Kan.

5 mL liquid cultures were grown overnight shaking vigorously at 37 °C. Plasmid DNA was isolated using a commercial miniprep kit (Qiagen). An analytical double digest using the previously described reaction mixture verified presence of the rsSTS gene insert. The sequence was verified as previously described using the T7 promoter and T7 terminator primers present in the pET28-a vector.

### **Expression and purification of cjNCS**

The following buffers were used during the purification of cjNCS for assays and crystallization:

Lysis Buffer: 50 mM NaH<sub>2</sub>PO<sub>4</sub>  
300 mM NaCl  
10 mM Imidazole  
pH = 8.0

Wash Buffer: 50 mM NaH<sub>2</sub>PO<sub>4</sub>  
300 mM NaCl  
20 mM Imidazole  
pH = 8.0

Elution Buffer: 50 mM NaH<sub>2</sub>PO<sub>4</sub>  
300 mM NaCl  
250 mM Imidazole  
pH = 8.0

MonoQA/SE200: 50 mM Tris-HCl  
0 mM NaCl  
pH = 7.5

MonoQ Buffer B: 50 mM Tris-HCl  
1 M NaCl  
pH = 7.5

cjNCS was cultured and isolated similarly to crSTS and rsSTS. The vector pET-28a\_cjNCS was transformed into *E. coli* BL21 (DE3) cells by electroporation. Cells were grown shaking at 37 °C for 1 hr. before plating on an LB-Agar (Kan) plate and incubation overnight at 37 °C. A 5 mL LB broth (1x Kan) starter culture was inoculated with cells from a single colony. This culture was grown overnight shaking at 37 °C. Following overnight starter culture growth, 1 L LB broth (1x Kan) was inoculated. The 1 L culture was incubated shaking at 37 °C until an O.D. of between 0.5-0.7 was reached. Once O.D. had been reached, the temperature was dropped to 16 °C and the culture was induced with IPTG. The cultures were grown overnight shaking at 16 °C. In the morning, cells were harvested by pelleting in the centrifuge at 4500 rpm for 15 min. Pellets were either used fresh or frozen and stored at -20 °C until future use.

To obtain cjNCS, cells were lysed using lysozyme and sonication in a manner similar to that described. After lysis, cell debris was pelleted by centrifugation at 13,000 rpm for 55 minutes. Purification by Ni-NTA affinity chromatography was performed as described for crSTS. Resin was washed with lysis buffer before use to remove ethanol and equilibrate the column. The resin was added to the supernatant in the 50 mL tubes and incubated with gentle shaking at 4 °C for 1-2 hrs. Following this incubation period, the mixture was added to the

column and the flow-through was collected. After collection of the flow-through, the column was washed with 100-200 mL wash buffer to remove weakly or non-specifically bound proteins. Finally, cjSTS was eluted from the column with 3x 10 mL 250 mM imidazole elution buffer. Fractions were analyzed by SDS-PAGE. Elution fractions were concentrated and buffer exchanged into 50 mM Tris-HCl before further use.

cjNCS was further purified by MonoQ anion exchange chromatography and SE200 size exclusion chromatography using the FPLC. Fractions were collected and analyzed by SDS-PAGE.

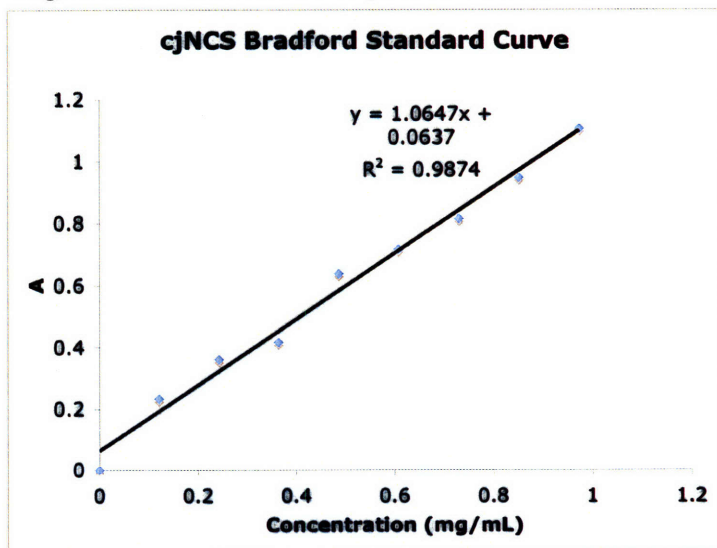
### Bradford assay of cjNCS

To verify the value for cjNCS concentration obtained by  $A_{280}$ , a Bradford assay was performed using the BioRad Bradford Assay Kit. A 1 mg/mL BSA standard solution was made. The actual concentration of this standard was shown to be 0.973 mg/mL (using  $\epsilon_{BSA} = 0.67$  (mg/mL)<sup>-1</sup>). To generate a standard curve, eight different serial dilutions were made from this

stock to give a total of nine data points for the curve. Three protein solutions were also analyzed – undiluted protein, 2x diluted protein, and 10x diluted protein.

To generate the curve and measure the concentrations, the concentrated reagent was diluted per the manufacturers instructions (one part reagent, four parts water).

Figure 3.14 – Bradford Assays Standard Curve



10  $\mu$ L of each sample was added to a 96 well microtiter plate (Nunc) followed by 200  $\mu$ L dilute reagent. The assay was developed in the dark for 10 minutes before the absorbencies were recorded using a Varian 50 MPR microplate reader.

### **Cloning new constructs of cjNCS**

The primers used to clone cjNCS with an N-terminal FLAG tag are as follows:

N-terminal: CCATGGGCGATATAAAAGATGATGATGATAAACCGATGAGCAAAAACCT  
GACCGGCGTG

C-terminal: TTGTCGACGGAGCTCGAATTCTTATTACAGTTTCAT

These primers were ordered and obtained from Integrated DNA Technologies. The following reaction mixture was used for the PCR reaction:

PCR reaction: 5  $\mu$ L 10x buffer  
1.5  $\mu$ L 10 mM dNTP mix  
2  $\mu$ L template DNA (pET28-a\_cjNCS)  
1  $\mu$ L N-terminal primer  
1  $\mu$ L C-terminal primer  
38.5  $\mu$ L dd H<sub>2</sub>O  
1  $\mu$ L High Fidelity Polymerase

These components were mixed in individual PCR tubes. PCR was done using a Stratagene RoboCycler Gradient 96 with the following program:

<u># of Cycles</u>	<u>Course</u>
1	120 s, 95 °C
33	30 s, 95 °C 1 min, 52-61 °C 1.75 min, 68 °C (1 min per kb plasma length)
Hold	Hold at room temperature

After the PCR was complete, each sample was purified using a Qiagen PCR clean-up kit and analyzed by 1% agarose gel electrophoresis. Samples that showed DNA fragments of the appropriate sized were selected and sub-cloned into a pGEM-T Easy transfer vector. The reaction mixture for pGEM-T Easy ligation is as follows:

pGEM-T Ligation: 5  $\mu$ L 2x ligation buffer  
1  $\mu$ L pGEM-T Easy vector  
3  $\mu$ L cjNCS-FLAG  
1  $\mu$ L T4 DNA ligase

The ligations were incubated at 4 °C overnight before transformation into electrocompetent *E. coli* Top 10 cells and plated as previously described. White colonies from the pGEM-T Easy ligation were selected and grown in liquid culture. Plasmid was harvested from the cells by miniprep and analyzed by a restriction double digest to verify the presence of the cjNCS-FLAG insert. Successful samples were submitted for sequencing using the T7 promoter and SP6 promoter primers. Sequencing revealed a mutation in the FLAG sequence.

The following primers were designed to remedy the mutation:

5'-3': ATGGGCGATTATAAAGATGATGATGATAAAGGCATGAGCAAAAAC  
3'-5': GTTTTTGCTCATGCCTTTATCATCATCATCTTTATAATCGCCAT

These primers were used with a Stratagene QuickChange Site-Directed Mutagenesis kit using procedures previously described.



### 3.7 References

1. Dewick, PM. (2001) *Medicinal Natural Products: A Biosynthetic approach*. John Wiley and Sons, Ltd., Hoboken, NJ.
2. Stadler R, Kutchan TM, Zenk M. (1989) *Phytochemistry*. 28, 1083-1086.
3. Samanani N, Facchini PJ. (2001) *Planta*. 213, 898-906.
4. Samanani N, Facchini PJ. (2002) *J. Biol. Chem.* 277, 33878-33883.
5. Minami H, Dubouzet E, Iwasa K, Sato F. (2007) *J. Biol. Chem.* 282, 6274-6282.
6. Luk LYP, Bunn S, Liscombe DK, Facchini PJ, Tanner ME. (2007) *Biochemistry*. 46, 10153-10161.
7. Czarnocki Z, MacLean DB, Szarek WA. (1985) *J. Am. Chem. Soc., Chem. Comm.* 19, 1318-1319
8. Pezzella A, Prota G. (2002) *Tet. Lett.* 43, 6719-6721.
9. Manini P, d'Ischia M, Lanzetta R, Parrilli M, Prota G. (1999) *Bioorg. Med. Chem. Lett.* 7, 2525-2530.
10. Samanani N, Liscombe DK, Facchini PJ. (2004) *Plant J.* 40, 302-313.
11. Chu, B. (1992) *Laser Light Scattering: Basic Principles and Practice*. Academic Press, Boston.
12. D'Arcy, A. (1994) *Acta. Cryst.* D50, 469-471.
13. Berkner H, Schweimer K, Matcko I, Rosch P. (2008) *Biochem. J.* Immediate publication, doi: 10.1042/BJ20080306.
14. Pasquo A, Bonamore A, Franceschini S, Macone A, Boffi A, Ilari A. (2008) *Acta Cryst.* F64, 281-283.

## Appendix 1 – Curriculum Vitae

### William Hillmann

18-125, 77 Massachusetts Ave, Cambridge, MA 02139  
bill.hillmann@gmail.com  
(908) 347-3412

#### EDUCATION

---

- 2006 - 2008    **MASSACHUSETTS INSTITUTE OF TECHNOLOGY**    Cambridge, MA  
*Master of Science in Biological Chemistry*
- 2002 – 2006    **BOSTON COLLEGE**    Chestnut Hill, MA  
*Bachelor of Science in Biochemistry, magna cum laude*
- Boston College Honors Program
  - Boston College Presidential Scholars Program

#### RESEARCH EXPERIENCE

---

- 2007 – 2008    **Massachusetts Institute of Technology**    Cambridge, MA  
Graduate Research Assistant, Department of Chemistry  
*Advisors: Sarah E. O'Connor, Ph.D. and Catherine L. Drennan, Ph.D.*
- 2006    **Massachusetts Institute of Technology**    Cambridge, MA  
Summer Research Assistant, Department of Chemistry  
*Advisor: Sarah E. O'Connor, Ph.D.*
- 2004 – 2006    **Boston College**    Chestnut Hill, MA  
Undergraduate Research Assistant, Department of Chemistry  
*Advisor: Steve D. Bruner, Ph.D.*

#### TEACHING EXPERIENCE

---

- 2007    **Massachusetts Institute of Technology**    Cambridge, MA  
*Teaching Assistant – 5.32 Intermediate Chemical Experimentation*
- 2006    **Massachusetts Institute of Technology**    Cambridge, MA  
*Teaching Assistant – 5.07 Biological Chemistry I*

#### AWARDS AND HONORS

---

Boston College Scholar of the College, 2006  
Phi Beta Kappa Honor Society, 2006  
Golden Key International Honor Society, 2006  
Northeastern Section of the American Chemical Society Norris-Richards Summer Research Scholarship, 2005  
Boston College Dean's Scholar in Biochemistry, 2005  
Boston College Presidential Scholarship, 2002-2006

#### FELLOWSHIPS

---

MIT/NIGMS Biotechnology Training Grant, 2007  
Boston College Undergraduate Research Fellowship, 2004, 2005