The Pennsylvania State University

The Graduate School

PROGRESS TOWARDS UNDERSTANDING THE MECHANISMS OF SPECIFIER PROTEINS IN THE GLUCOSINOLATE-MYROSINASE DEFENSE SYSTEM OF BRASSICAS

A Dissertation in

Chemistry

by

Leanne E. Mocniak

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2021

The dissertation of Leanne E. Mocniak was reviewed and approved by the following:

Ray Bryant
Adjunct Professor of Agronomy, USDA-ARS
Dissertation Co-Adviser
Co-Chair of Committee

J. Martin Bollinger, Jr.
Professor of Chemistry and of Biochemistry and Molecular Biology
Dissertation Co-Adviser
Co-Chair of Committee

Amie K. Boal
Professor of Chemistry and of Biochemistry and Molecular Biology

Carsten Krebs
Professor of Chemistry and of Biochemistry and Molecular Biology

Dan G. Sykes
Teaching Professor, Department of Chemistry

Marvin Hall
Professor Emeritus of Forage Management

Kyle Elkin
Special Signatory

Philip Bevilacqua
Distinguished Professor of Chemistry and of Biochemistry and Molecular Biology
Head of the Department of Chemistry

ABSTRACT

μThe glucosinolate-myrosinase system (GLS-MYR) is a natural plant defense mechanism unique to plants in the order of Brassicales. The plants most commonly studied belong to the family Brassicaceae (brassicas) that includes economically and agriculturally important crops such as mustards, turnips, kale, cabbages, rapeseed, etc. Glucosinolates (GLSs) are a family of *S*-(β-D-glucopyranosyl)thiohydroximate-*O*-sulfonate secondary metabolites, each with a distinguishing side chain linked to the central carbon. Myrosinase (MYR) is a retaining β-thioglucosidase with the primary function to hydrolyze the thioglucosidic bond of the GLS. The GLS-MYR reaction mechanism is initiated *in vivo* by tissue damage (i.e., insect chewing, livestock mastication), where the compartmentalized MYR is released and hydrolyzes the GLS substrate. The result of MYR hydrolysis is an unstable thiohydroximate-*O*-sulfonate (aglycone) intermediate that non-enzymatically rearranges via a Lossen-like rearrangement to form an isothiocyanate (RNCS). RNCSs are biologically active compounds that are toxic to generalist insects, bacteria, fungi, and nematodes, and are therefore essential components of the innate biochemical defense system in brassica plants. Additionally, the biological activity of RNCSs has also been linked to health promoting effects in humans (i.e., anticancer) as well as negative health conditions observed in livestock grazed on a brassica diet (i.e., goiter).

Several brassica species also contain a secondary class of enzymes, specifier proteins (SPs), that divert the course of the GLS-MYR reaction away from generation of RNCSs to form alternative, less toxic hydrolysis products. SPs are Kelch-domain-repeat, β-propeller proteins that catalyze formation of these alternative products, including nitriles (nitrile specifier protein, NSP), epithionitriles (epithionitrile specifier protein, ESP), and thiocyanates

(thiocyanate forming protein, TFP). There is little information available regarding SP function, including the chemical mechanisms by which SPs generate the alternative hydrolysis products. The lack of mechanistic information presents a substantial barrier to research progress towards harnessing the GLS-MYR system for agricultural or human health applications. Therefore, the goal of this dissertation is to advance the current understanding of SP function and mechanism of action to enable targeted agricultural applications utilizing the GLS-MYR system. For example, if SP mechanism of action is properly characterized, it may be possible to use that information to artificially influence the innate GLS-MYR system in brassicas to produce elevated or reduced levels of RNCSs with the goal to increase brassica toxicity to insects or decrease negative health effects observed in animals grazed on brassicas, respectively.

In **Chapter 1**, the current knowledge of the chemical mechanisms involved the GLS-MYR defense system are summarized. This includes an outline of GLS biosynthesis and characterization, the details of MYR-catalyzed GLS hydrolysis, and a description of the formation and bioactivity of RNCS products. The current status of SP characterization is also outlined, including the shortcomings and areas of improvement to be discussed in later chapters. The chapter concludes with a summary of the known applications of the GLS-MYR system relevant to both human health and agriculture.

Chapter 2 highlights the importance of accurate determination of a GLS profile for agriculturally relevant brassica species. In this study, three forage brassica varieties [turnip (*B. rapa* L.), canola (*B. napus* L.), and rapeseed (*B. napus* L.)] were analyzed for individual GLS content. The average GLS content across three herbage collection dates was 2.9 ± 0.9 mg g^{-1}, 6.4 ± 1.3 mg g^{-1}, turnip, and 14 ± 3.4 mg g^{-1} for canola, rapeseed, and turnip, respectively (units are reported in milligrams of glucosinolate, calculated as sinigrin equivalents, per gram of dry

plant material). Our semi-quantitative screening of various forage brassicas demonstrates the high variation of individual GLS content and the total GLS profile from species to species. The use of this efficient, simple method delivers a nearly complete GLS profile for forage brassicas, including GLS that were not identified using prior methods. Accurate GLS profiles are necessary for efficient use of specific forage varieties in targeted agricultural applications such as a healthy diet for livestock and effective soil biofumigation.

In **Chapter 3,** we discuss the challenges involved in proper experimental analysis of SP mechanisms due to the inherent instability of the aglycone intermediates. Knowledge of the aglycone lifetimes would facilitate SP analysis, therefore, we developed a spectrophotometric method used to monitor the Lossen-like rearrangements of MYR-generated aglycones from nine GLSs. We discovered that their half-lives ($t_{1/2}$) vary by more than a factor of 50, from < 3 s to 150 s (22 °C). The $t_{1/2}$ of the sinigrin-derived allyl aglycone (34 s), which can form the epithionitrile product (1-cyano-2,3-epithiopropane) in the presence of ESP, proved sufficient to enable spatial and temporal separation of the MYR and ESP reactions. Our results confirm recent proposals that ESP is an autonomous iron-dependent enzyme that intercepts the unstable aglycone rather than a direct effector of MYR. Our characterization of aglycone lifetimes lead to analysis of SPs that had not been possible previously.

Chapter 4 presents novel characterization of the epithionitrile specifier protein (ESP) and nitrile specifier protein (NSP) from *Arabidopsis thaliana* that, to our knowledge, has not yet been proven experimentally. We provide evidence that ESP and NSP employ the +II oxidation state of their iron cofactors. The analysis of SP iron dependence performed in Chapter 4 affords more robust evidence than that presented in Chapter 3. Additionally, we further characterized ESP and NSP by challenging their activity with diverse aglycones from

varied GLSs. From this analysis, we found that the catalytic efficiency of ESP ranged from 1 – 3.4 $\mu M^{-1}s^{-1}$ when challenged with six different aglycones. Our results indicate that ESP is surprisingly unselective for olefin-containing aglycones that it can convert to its namesake product. We also found that the catalytic efficiency of NSP is 10-fold lower than that of ESP when challenged with allyl aglycone and 4-methylthiobutyl aglycone (0.28, and 0.14 $\mu M^{-1}s^{-1}$, respectively). Taken together, these results imply that SP activity does not discriminate against the aglycone side-chain composition. Our emerging understanding of SP function could inform development of agricultural applications based on strategically bred or genetically altered brassica plants with optimized GLS and GLS-catabolite profiles.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, thank you to my co-advisors, Kyle and Marty, for giving me the opportunity to merge your expertise and build my own project. Your persistent encouragement has instilled confidence in myself, and your support and guidance were essential for my success. For this I am incredibly grateful. Thank you to Ray Bryant, for your willingness to bear the administrative responsibility as the chair of my committee. Without your support, I would not have been able to create this unique and rewarding experience. Thank you to my entire doctoral committee. Creating my own project was challenging in many ways, but especially because I had very limited resources. I appreciate your intellectual support and the time you invested in making sure I stayed on track. Without your exceptional leadership, I would not be where I am today. Thank you to my colleagues in the Bollinger-Krebs group and in the USDA-ARS facility for your constant support of my unique situation and willingness to help me when I needed it. Lastly, I would like to thank my family, friends, and fiancé. Graduate school is a long and difficult process, and I wouldn't have been able to get through it without the confidence you showed in me and your emotional support. This work was supported in part by NIGMS grant 1R35 GM127079-03 to Professor Carsten Krebs and by USDA-NIFA-OREI. Any opinions, findings, and conclusions or recommendations expressed in this work are those of the author and do not necessarily reflect the views of the NIGMS or USDA-NIFA-OREI.

Chapter 1
Introduction

1.1. GLUCOSINOLATES

Glucosinolates (GLSs) are a large family of secondary metabolites synthesized by plants in the Brassicales order. The core structure of a GLS consists of a *S*-β-D-glucopyrano unit linked to a (*Z*) thiohydroximate-*O*-sulfate moiety at the anomeric carbon, and a variable side chain (R group) connected to the central carbon of the thiohydroximate moiety (Figure 1-1, Table 1-1).[1] To date, there over 150 proposed GLS structures, approximately two-thirds of which have been satisfactorily characterized (verification of side chain composition).[1] This high level of GLS diversity is caused by the variable chemical composition of the side chain, the identity of which is derived from the amino acid used for GLS biosynthesis. Aliphatic side chains are typically derived from methionine, alanine, valine, leucine, or isoleucine, while benzenic GLS are synthesized from phenylalanine or tyrosine, and indole GLS are derived from tryptophan. These amino acid precursors can undergo additional modifications to their structure before GLS biosynthesis (i.e. chain elongation) or the GLS side chain can be altered post biosynthesis (i.e. hydroxylation, methylation, or oxidation of sulfur to sulfoxides).[2] An example pathway for GLS biosynthesis from methionine is shown in Figure 1-2. The identity and quantity of the individual GLSs (GLS profile) synthesized in brassicas varies widely from species to species and is regulated by natural genetic factors.[3] However, the GLS profile can also be affected by external influences. For example, when treated with the plant hormone methyl jasmonate, broccoli plants respond by elevating levels of indole GLS (glucobrassicin, gluconasturtiin) while levels of aliphatic GLS (sinigrin, glucoraphanin) are unaffected.[4] Although the influence of environmental factors has been extensively studied, the underlying

biochemical mechanisms responsible for modulation of the GLS profile remain unknown.[3] A summary of GLS used in this work is shown in Table 1-1.

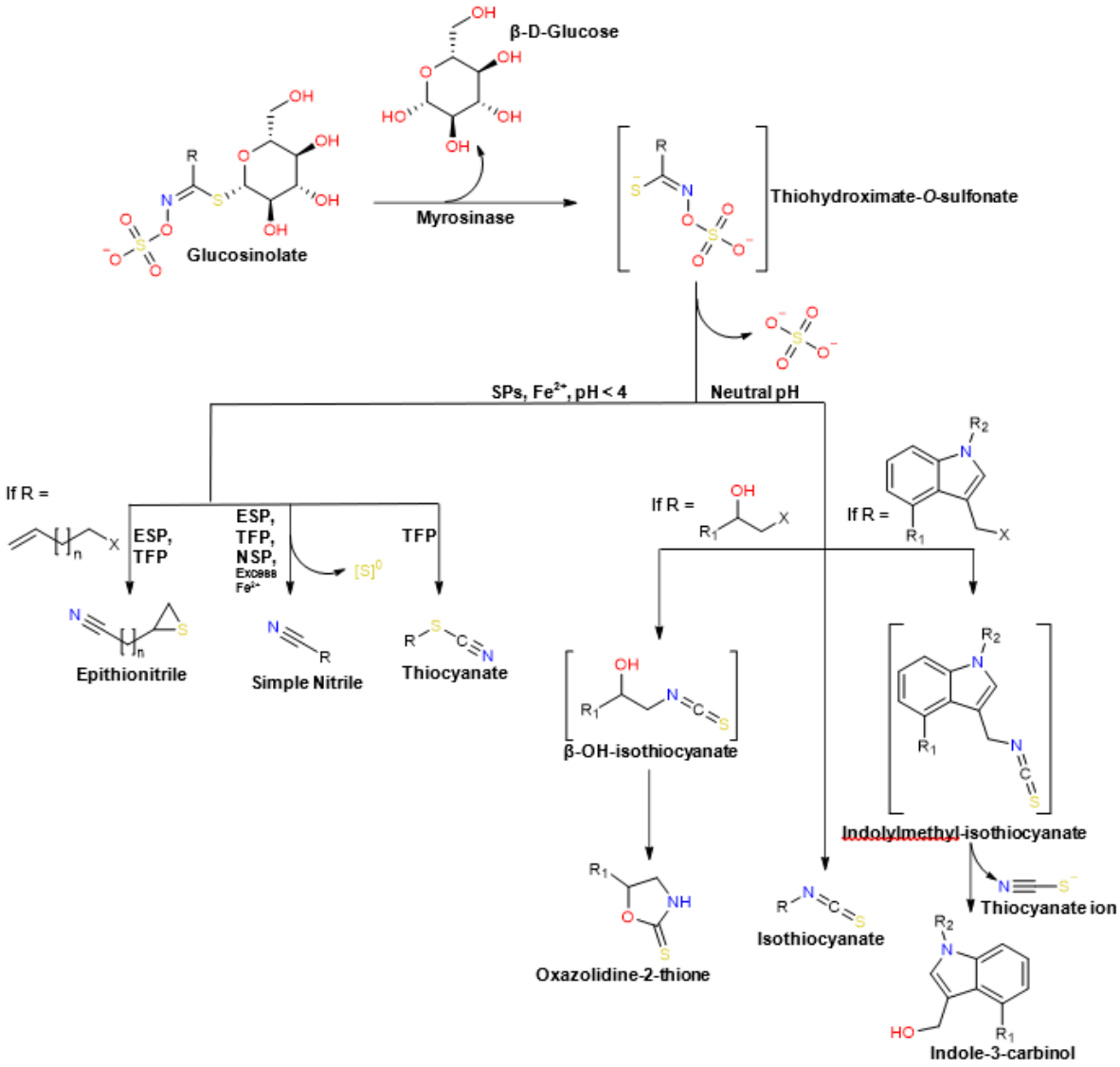


Figure 1-1. Summary of GLS breakdown pathways. MYR mediated GLS hydrolysis promotes formation of thiohydroximate-*O*-sulfonate (aglycone) intermediate. At neutral pH, the aglycone will form RNCSs via a Lossen-like rearrangement. If the side chain is indolic. i.e., glucobrassicin, the RNCS undergoes further decomposition to indole-3-carbinol. If the side chain contains a hydroxyl group on an alkyl chain, i.e., progoitrin, the hydroxy-NCS formed will cyclize to form oxazolidine-2-thione. Under alternative conditions (i.e., pH < 4, in the presence of excess Fe(II), or in the presence of SPs, the aglycone intermediate will form RCN, ETN, or TCN instead of RNCSs (SPs are detailed below). Adapted from Barba et al.[5]

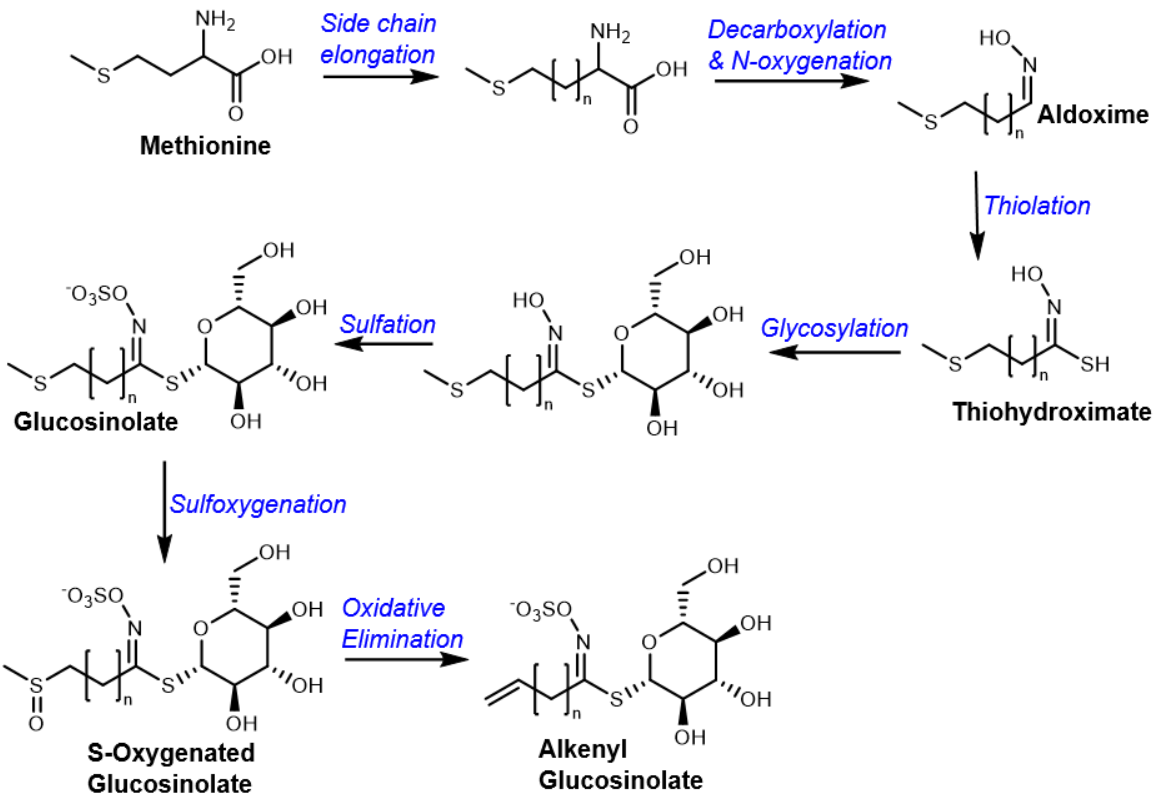


Figure 1-2. Example biosynthetic pathway of aliphatic GLSs from methionine in *A. thaliana*. Blue lettering above the reaction arrows represents enzyme-catalyzed steps. Methionine may undergo chain elongation (potentially adding up to 11 methylene groups) prior to GLS biosynthesis. Post-synthetic modifications such as sulfoxygenation and oxidative elimination of the methylsulfinyl group to form alkenyl side chains are also possible. Adapted from Ishida et al.[6] and Sønderby et al.[7]

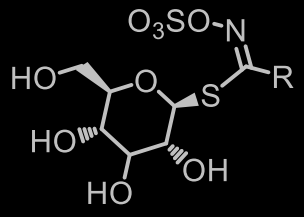
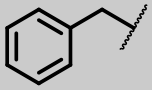
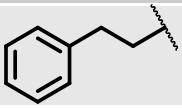
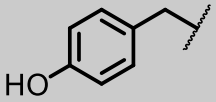
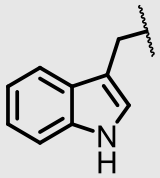
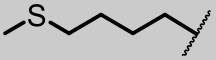
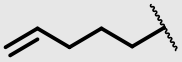
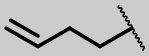
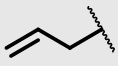
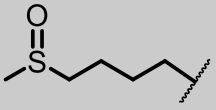
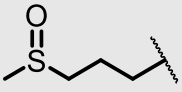
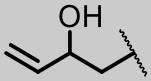
Glucosinolate	Common Name	R Group (side chain)
		
Benzyl GLS	Glucotropaeolin	
Phenethyl GLS	Gluconasturtiin	
p-hydroxybenzyl GLS	Sinalbin	
3-indolylmethylGLS	Glucobrassicin	
4-methythiobutyl GLS	Glucoerucin	
4-pentenyl GLS	Glucobrassicanapin	
3-butenyl GLS	Gluconapin	
Allyl GLS	Sinigrin	
4-methylsulfinylbutyl GLS	Glucoraphanin	
3-methylsulfinyl propyl GLS	Glucoiberin	
2-hydroxy-3-butenyl GLS	Progoitrin	

Table 1-1. List of GLS and their corresponding side chains used in this work.

1.2. MYROSINASE STRUCTURE AND FUNCTION

The initial discovery of myrosinase dates back to 1840, when a protein deemed Myrosine was found to be necessary for liberating essential oils from mustard seeds.[8] Since then, studies on MYR have improved substantially to include advances such as discovery in numerous additional brassica species,[8] physiological localization,[9] the discovery of ascorbate as a MYR coenzyme,[10,11] the first crystal structure of MYR,[12] and further subcellular characterizations.[13–15] Studies to investigate the cellular localization of MYR and GLSs revealed that they are stored separately, but the location varies from species to species. For example, *Arabidopsis thaliana* localizes MYR in myrosin cells and GLSs in S-cells, but both cell types are specific to *A. thaliana* only.[16] Additionally, myrosinases from different species have highly variable substrate specificity and multiple isoforms per species, which is expected with such a high number of possible GLS substrates. This work utilizes MYR isolated from *Sinapis alba* (white mustard) because it is easily obtained from mustard seed, is highly stable, and is to our knowledge, the best-characterized isoform.[1,11,12]

Functionally, myrosinase is a retaining thioglucoside glucohydrolase (EC 3.2.3.1), and is the only known *S*-glucosidase belonging to glycoside hydrolase family 1 (GH1). It is structurally similar to the classic *O*-glucosidases in GH1 with a $(\beta/\alpha)_8$ barrel fold configuration.[12,17] MYR is also heavily glycosylated (up to 20% of its weight) and naturally occurs as a homodimer held together by a tetrahedrally-coordinated Zn (II) ion.[12] These structural characteristics ensure MYR is stable in extreme environments, such as mustard seeds, whose dehydrated nature is hostile for most proteins. Classic *O*-glucosidases require a nucleophile and a proton acceptor to catalyze the hydrolysis of a glucoside. In most *O*-glucosidases, these are two glutamate (Glu) residues, where the first Glu residue cleaves the

O-glucosidic bond via nucleophilic attack and the second Glu residue activates a water molecule to hydrolyze the glucosyl-enzyme intermediate.[18]

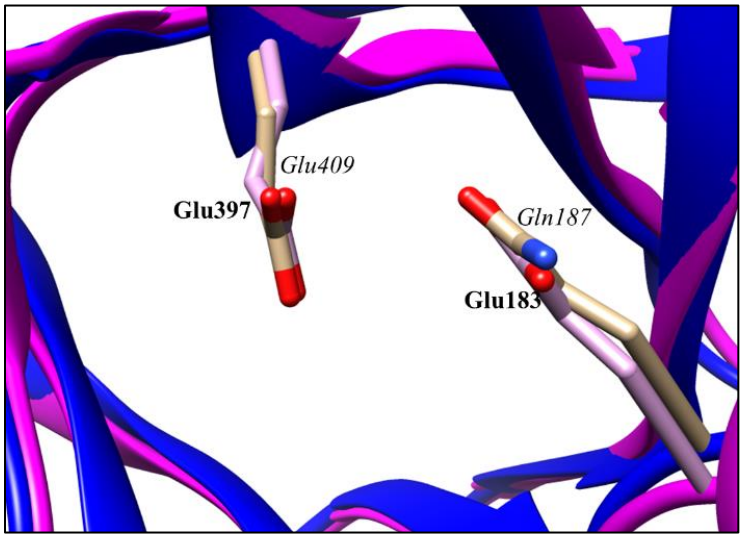


Figure 1-3. Comparison of β-*S*-glucosidase (from white mustard seed, PDB 1MYR) and β-*O*-glucosidase (from white clover, PDB 1CBG) active sites. The sites are highly similar, representative of their similar mechanism of action (Figure 1-4). In β-S-glucosidase, the second Glu residue (Glu183) is replaces by a Gln residue (Gln187).

Despite the structural similarities, the mechanism of action for GLS hydrolysis by *S*-glucosidase (MYR) differs from that of *O*-glucosidases (Figure 1-4). Hydrolysis is initiated by nucleophilic attack of the thioglucosidic bond at the anomeric carbon by a Glu residue in a retaining configuration (this step is identical to classic *O*-glucosidases). This cleavage forms an unstable aglycone intermediate and a glucosyl-Glu complex. The aglycone intermediate spontaneously rearranges to an isothiocyanate (RNCS), via a Lossen-like rearrangement, and the glucosyl-Glu complex is hydrolyzed to open the active site for another GLS.[8] A detailed description of the Lossen rearrangement is presented in section 1.3. The hydrolysis of the glucosyl-enzyme complex in MYR differs from *O*-glucosidases because, instead of the acid/base catalysis with Glu, an ascorbate molecule bound in the MYR catalytic site acts as a reversibly dissociable base to abstract a proton from (activate) a water molecule (Figure 1-3).[11]

In MYR, the Gln that replaces the typical Glu residue is used to position the activated water molecule at the anomeric carbon for hydrolysis (Figure 1-4).[12] The biochemical reasoning for the replacement of the Glu residue with a Gln residue is that the nitrogen of the amide group is able to form a hydrogen bond with the sulfate group on the GLS, which would otherwise result in unfavorable electrostatic interactions if a Glu residue were present.[8]

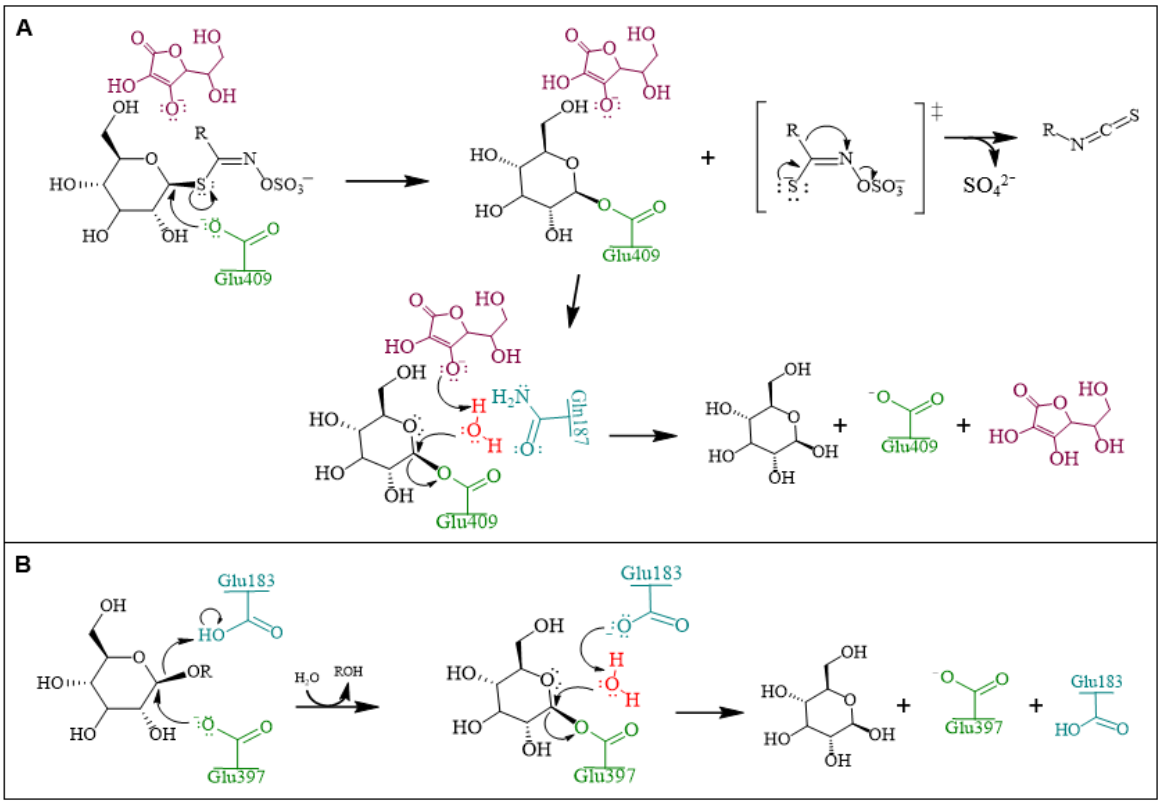


Figure 1-4. Comparison of *S*-glucosidase and *O*-glucosidase. (A) MYR (*S*-glucosidase) reaction mechanism. Glu409 cleaves the thioglucosidic bond at the anomeric carbon via nucleophilic attack. The resulting aglycone intermediate spontaneously rearranges via a Lossen-like rearrangement to an RNCS. An ascorbate molecule bound in the MYR active site activates a water molecule, positioned by Gln187 to hydrolyze the Glu409-glucose complex. (B) *O*-glucosidase reaction mechanism. The first step, nucleophilic attack of the anomeric carbon, is the same as MYR. For *O*-glucosidase, the hydrolysis of the Glu397-glucose complex is initiated by a second glutamate residue, Glu183.

It has been repeatedly shown that ascorbate significantly increases MYR activity,[11,19] but the evolutionary replacement of the second acid/base residue (Glu) in the active site with

Gln suggests that the initial nucleophilic attack with Glu would still allow GLS hydrolysis to proceed without the presence of ascorbate (albeit much slower, demonstrating a 10-fold decrease in activity without ascorbate present).[8,11] The final products of GLS hydrolysis are an aglycone intermediate (which spontaneously decays to RNCS), a sulfate ion, and a glucose molecule. Although RNCSs are the direct product of aglycone decay, other environmental influences can non-enzymatically interact with the aglycone intermediate and alter the final products. Factors such as low pH (< 4) or the presence of excess Fe(II) in solution will promote formation of simple nitriles (RCNs) instead of RNCSs.[20,21]

Current knowledge of the MYR active site is inferred from the crystal structure of MYR from *S. alba*, despite the lack of a solved structure with a glucosinolate or aglycone bound (Figure 1-5).[12] The active site has a hydrophobic pocket (Phe473, Phe371, Phe331, Tyr330, Ile257) that recognizes the hydrophobic portion of the glucosinolate (the variable R group side chain). However, substrate specificity is thought to be controlled by the sulfate recognition site at Ser190 and Arg259, and weakly Arg194, as evidenced by the fact that desulfoglucosinolates are poor MYR inhibitors.[22] It is also suggested that the thiohydroximate portion of a GLS contributes to MYR substrate recognition due to the ability of the polar N and O atoms to form hydrogen bonds to the active site amino acid residues. The ascorbate (Ser190) and glucose (Gln187, Trp457, Phe465) binding sites are commonly positioned, so intact GLS and ascorbate cannot bind to the active site at the same time. [11,12] This notion supports the suggestion that ascorbate is not necessary for MYR function, but does act as a coenzyme for faster catalysis.

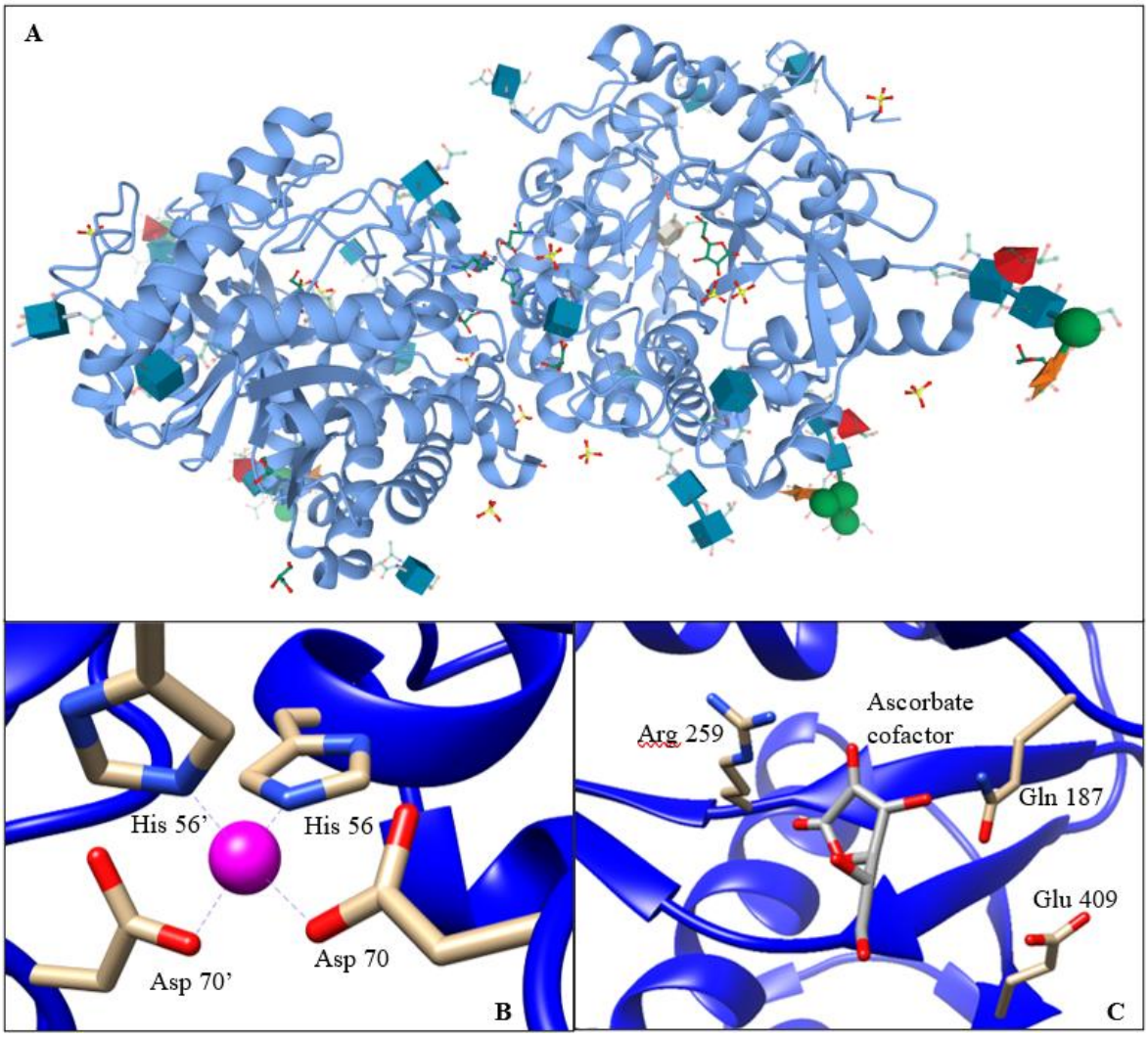


Figure 1-5. Crystal structure of *S. alba* MYR. (A) PDB 1E73, depicts the glycosylation sites (yellow star, D-xylopyranose; green circle, D-mannose; blue square, N-acetyl-D-glucosamine; red triangle, L-fucose. (B) PDB 1MYR, Zinc binding site. (C) PDB 1MYR, active site with ascorbate cofactor. (Note: there are no available structures with a glucosinolate bound).

1.3. THE LOSSEN REARRANGEMENT

The efficacy of the GLS-MYR system depends on the formation of isothiocyanates from the aglycone intermediate. The spontaneous rearrangement of the aglycone to RNCSs has been routinely classified as a Lossen-like rearrangement because the migration of the aglycone R group to the N atom is highly reminiscent of the classical Lossen reaction. The Lossen rearrangement was discovered in 1872 by W. Lossen, who studied the synthesis of hydroxylamine hydrochloride and its subsequent condensation with benzoyl chloride.[23] He

was able to isolate benzoyl benzohydraxamate from this reaction and investigated the thermolysis to phenyl isocyanate and benzoic acid (Figure 1-6).[24] The reaction by which benzoyl benzohydraxamate decays to phenyl isocyanate had never been observed prior, and the phenomenon was named after W. Lossen. In modern chemistry, a "Lossen rearrangement" most often refers to the conversion of *O*-activated hydroxamic acids to their corresponding isocyanate.

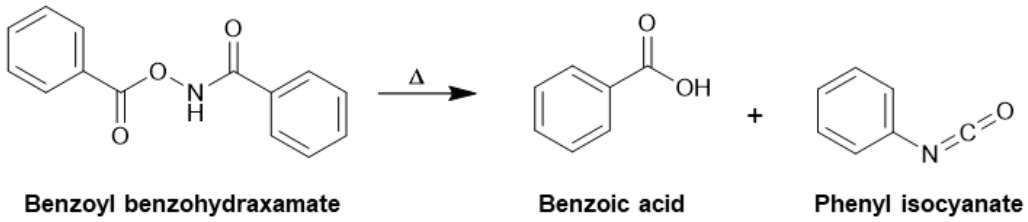


Figure 1-6. Schematic of Lossen discovery.

Hydroxamic acids typically require activation to undergo a Lossen rearrangement because hydroxide is a poor leaving group. The activation can be achieved via dehydration-substitution of the hydroxide proton (Figure 1-7, step 1).[23] The Lossen rearrangement is often promoted by deprotonation of the nitrogen with a base, which triggers the α-elimination of a leaving group (Figure 1-7 step 2).[25] The loss of the leaving group generates an acyl nitrene (which is highly reactive, Figure 1-7 step 3) and the isocyanate is produce by the C to N migration of the R group (Figure 1-7 step 4). It should be noted that a nitrene intermediate has never been isolated to date; therefore the formation of the nitrene and R group migration could be either concerted or occur separate steps.[23] The rate of the Lossen rearrangement is also affected by the nature of the leaving group, such that leaving groups with low pK_a values result in faster rearrangements.[25]

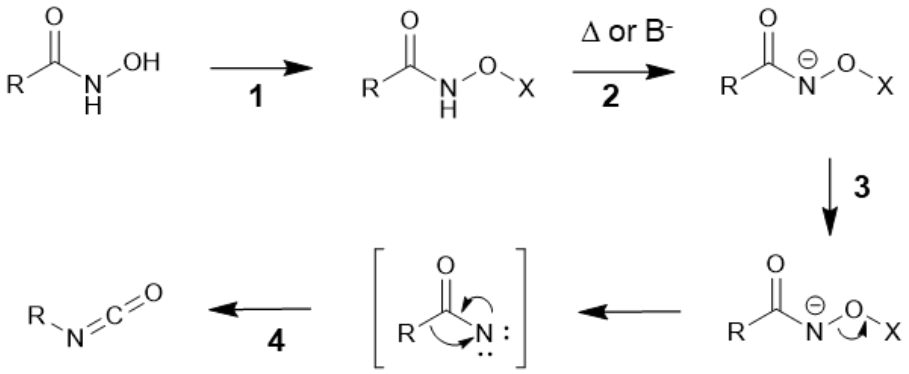


Figure 1-7. Mechanism of classical Lossen rearrangement, depicted stepwise.

The Lossen rearrangement is emerging in modern methods for synthesis of bioactive compounds for medicinal applications and as safer alternative for upscaled industrial syntheses.[26,27] Other common approaches that facilitate carbon to nitrogen rearrangements such as the Hofmann, the Schmidt, or the Curtius rearrangements typically occur under hazardous conditions. For example, the Hofmann rearrangement often uses highly corrosive hypervalent iodine or bromite,[28] the Curtius rearrangement is based on dangerous azide chemistry,[29] and the Schmidt rearrangement requires use of explosive hydrazoic acid.[30] The Lossen rearrangement achieves the same migration under milder, less hazardous conditions.

The details of the classical Lossen rearrangement are mirrored in the decay of an aglycone intermediate in the GLS-MYR system. The core structural element of a GLS is a thiohydroximate moiety, which is a thiolate tautomer of the hydroxamates observed in typical Lossen reactions.[1] MYR-mediated production of the aglycone is analogous to the activated hydroxamic acid typically utilized in classical Lossen reactions. As stated above, the loss of the sulfate leaving group on the aglycone drives the decay of the aglycone to the RNCS, which includes an identical migration of the R group from C to N (Figure 1-4). These similarities result in the classification of aglycone decay to RNCS as “Lossen-like.”

1.4. ISOTHIOCYANTES

The bioactivity of RNCSs results from their lipophilic and electrophilic chemical nature which arises from the N=C=S moiety (Figure 1-1). The lipophilic nature of RNCSs facilitates passive diffusion into the intracellular environments of susceptible organisms, and the electrophilic nature ensures RNCSs rapidly react with biological nucleophiles (amino acids, glutathione, etc.) which detrimentally disrupts normal physiological conditions.[2] A review by Hanschen et al.[31] presents a summary of the reactivity of specific RNCSs with biological nucleophiles (i.e. hydroxy, amino, or thiol groups), but none of the described experiments were performed *in vivo*. Therefore, the mechanisms by which RNCSs cause both health-promoting effects in humans and deleterious effects in grazing animals are not well understood (see the Human and Agricultural applications sections below for more detail). Interestingly, the RNCSs produced from GLS hydrolysis are responsible for the spicey, pungent flavor that is characteristic of brassicas like mustard and horseradish.

The composition of the R group has a significant impact on the biological activity of the resulting GLS hydrolysis products in terms of reactivity and subsequently, toxicity.[32] It has been reported that RNCS toxicity increases with greater electrophilic character: RNCSs with electron withdrawing substituents are significantly more reactive, and therefore, more toxic. It is believed that short-chain, aliphatic GLS (and their corresponding RNCS) have the greatest potential biocidal reactivity,[33] an observation that is important for harnessing the GLS-MYR system for agricultural or medicinal applications. Seeking brassicas with the highest GLS content, preferably aliphatic in nature, seems the ideal choice especially for human health applications (i.e., anticancer, outlined in section 1.6 *Applications for the GLS-MYR system in Human Health).* However, recent studies have demonstrated that the internal maintenance of

the GLS-MYR system *in vivo* is costly for plants as each GLS requires two sulfur donors and an amino acid for biosynthesis.[34,35] For example, Bekaert et al.[36] showed that biosynthesis of constitutive GLS levels consume more than 15% of the total energy consumed by the leaf tissue of *A. Thaliana* during photosynthesis, and such requirements continue to rise under biological attack, as GLS biosynthesis is induced. The implications of the energy needed for GLS biosynthesis manifest in a wide range of GLS profiles in plants, with both genetic and environmental influences.[2] The importance of organism choice for agricultural or medicinal applications due to the wide variation in GLS content is clear, but is complicated by the many factors that affect the GLS-MYR system.

1.5. SPECIFIER PROTEINS

The GLS-MYR system is mediated in a subset of brassicas by a family of Kelch-like specifier proteins (SPs), introducing an additional layer of complexity to brassica defense, both chemically and evolutionarily. When present, SPs alter the products of GLS hydrolysis by diverting the rearrangement of the aglycone intermediate (substrate of SPs) produced by MYR hydrolysis towards alternative products that are less toxic to insects. There are three types of SPs currently known, and are classified by the products they generate. Epithiospecifier protein (ESP) catalyzes the formation of epithionitriles, thiocyanate forming protein (TFP) produces thiocyanates, and nitrile specifier proteins (NSP) generate simple nitriles and convert the thiohydroximate sulfur of the aglycone to an unknown zero-valent form. To date, only 21 specifier protein isoforms have been identified from 11 brassica species (Table 1-2).[37–39] The evolutionary purpose of SPs is unknown, especially considering that they reduce the innate toxicity of the GLS-MYR defense system of brassica plants against herbivory. However, it is postulated that these alternative products are related to plant signaling. For example, it has

been shown that an increase in simple nitrile production deters oviposition (laying of eggs) on brassicas by specialist insects (organisms that display immunity to RNCSs).[40]

Abbreviation	Species	UniProtID
ApTFP	*Alliaria petiolata*	J7FPI6
AtESP	*Arabidopsis thaliana*	Q8Ry71
AtNSP1	*Arabidopsis thaliana*	Q9SDM9
AtNSP2	*Arabidopsis thaliana*	O49326
AtNSP3	*Arabidopsis thaliana*	O04318
AtNSP4	*Arabidopsis thaliana*	O04316
AtNSP5	*Arabidopsis thaliana*	Q93XW5
BoESP1	*Brassica oleracea var. italica*	Q4TU02
BoESP2	*Brassica oleracea var. italica*	A0A0D3D8Y4
BoESP3	*Brassica oleracea var. italica*	A0A0D3AU36
ChESP	*Cardamine hirsuta*	J7FLJ0
ChNSP	*Cardamine hirsuta*	J7FR70
CiESP	*Cardamine impatiens*	J7FRY5
DaESP	*Draba aurea*	J7FMU8
DlESP	*Draba lanceolata*	J7FU88
IsESP	*Isatis tinctoria*	J7FPI9
ItNSP	*Isatis tinctoria*	J7FMV0
LsTFP	*Lepidium sativum*	A1XLE2
SpESP	*Schouwia purpurea*	J7FLJ4
SpNSP	*Schouwia purpurea*	J7FU93
TaTFP	*Thlaspi arvense*	G1FNI6

Table 1-2. List of all specifier proteins identified to date from brassica species with UniProt identification numbers.

The aforementioned lack of available information concerning SPs manifests a number of unanswered, fundamental questions regarding SP function and mechanism of action that has prevented proper investigation and characterization:

i. Do specifier proteins depend on Fe(II) as a cofactor?

ii. Are specifier proteins autonomous enzymes or allosteric myrosinase effectors?

iii. Is it possible to characterize specifier proteins kinetically and mechanistically?

(i). All three SPs are suspected to be Fe(II) dependent (although, as stated previously, simple nitrile formation can occur in the presence of excess Fe(II) ions or at low pH). Very few studies that report iron dependence supplied Fe(II) in stoichiometric amounts with respect to the SP[41–43] (commonly added in high excess[21,44,45]), and each SP has the capability of producing simple nitriles when presented with an aglycone that do not promote their namesake products. Further discussion of SP substrate specificity is detailed in Chapter 4. Moreover, several studies attempted to examine the effects of Fe(II) compared to Fe(III) on SP activity, but fail to verify the oxidation state of the iron present.[41] Therefore, the presumption of SP dependence on Fe(II) is inconclusive without stoichiometric ratios of iron to SP and proper assurance of oxidation state. This work provides conclusive evidence that verifies SP dependence on iron in its reduced form, Fe(II) (Chapters 3 and 4).

(ii). The literature suggests two potential modes of action for SPs.[41] The first hypothesis (A) posits that SPs are autonomous proteins that directly bind aglycones to chemically alter GLS breakdown. The alternative hypothesis (B) is that SPs act as MYR effectors, requiring a physical interaction between SP and MYR to affect GLS hydrolysis. This work provides confirmation that SPs are autonomous proteins (hypothesis A) and, to our knowledge, definitive validation of hypothesis A has not yet been reported elsewhere. This verification is imperative for SP characterization because it informs experimental procedures that facilitates answering the last fundamental inquiry (Chapters 3 and 4).

(iii). Many aspects of SP function remain poorly understood due to the reportedly short-lived nature of the aglycone compounds, which are considered the SP substrates. These compounds have been repeatedly characterized in the literature as “extremely unstable,”[46–50] but, to the

best of our knowledge, no quantitative assessment of their intrinsic lifetimes has been reported previously. Accordingly, experimental analyses of SP reactions have all been carried out by including MYR to produce the unstable aglycone *in situ*. This experimental configuration does not allow a distinction to be made between the impact of the SP directly on the outcome of the MYR reaction and the presumed sequential action of the two proteins individually on the small molecules.[37,51] Therefore, the dependence of SP activity on MYR for production of the unstable aglycone substrate has become a substantial barrier that prevents SP characterization. To address the difficulties presented by the MYR dependence and the inherent instability of the aglycones, this work demonstrated a quantitative assessment of aglycone stability and its application for kinetic and mechanistic characterization of SPs (Chapter 3).

The lack of experimental capabilities to study SP mechanism of action has resulted in a heavy focus on SP structure as a pathway to understanding SP function. Only one crystal structure exists for each type of SP (TFP[47], ESP[48], NSP[52]) and they all lack a metallic cofactor and an aglycone bound in the active site (Figure 1-8), which is particularly problematic for mechanistic studies given their inherent Fe(II) dependence. Most structural information available is based on molecular modeling experiments,[38,41,53,54] so the characterization of SP active sites reported in the literature is inferred from mutagenesis experiments with the amino acid residues identified in the models (see Figures 1-9, 1-11 and 1-13 for a summary of known SP structural analysis).[47,53] There are also several significant differences in the modelled versus solved crystal structures, which are mostly attributed to conformational changes to the active site when the aglycone binds, but these differences further challenge proper SP active site characterization and mechanistic elucidation. The results of the crystallography experiments show that each SP isoform consists of multiple repeating Kelch motifs, which have a

characteristic 4-straded β-sheet structure.[47] TFP and ESP are classified as homodimers with a six-bladed β-propeller fold (β_1- β_6), and each β-sheet consists of 4 antiparallel β-strands.[47,48,55] In contrast, NSP has a monomer structure (that does follow the same β-propeller fold as TFP and ESP), with an N-terminal, jacalin-like lectin domain instead of second β-propeller monomer.[52,55] The active site for all three SPs is located at the bottom of the β-propeller structure. The current knowledge of SP function and mechanism of action is detailed below.

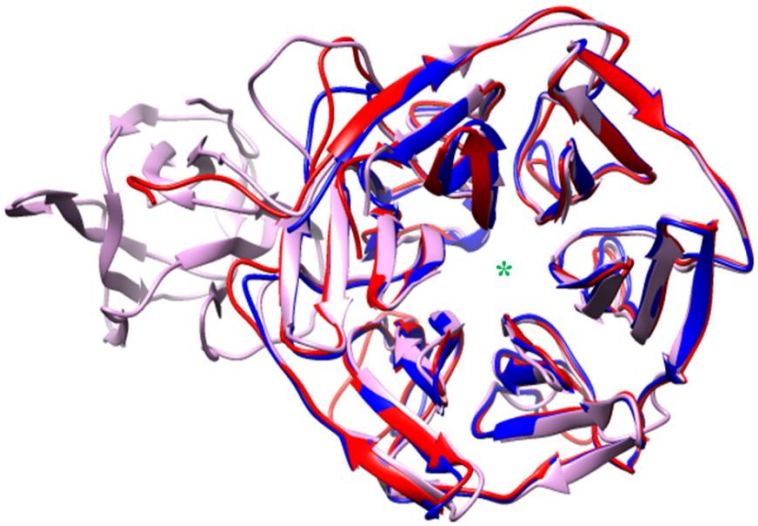


Figure 1-8. Overlay of all three specifier protein structures. Only one monomer is shown for ESP and TFP. ESP, red (PDB 5QG0); NSP, pink (PDB 5GQT); TFP, blue (PDB 5A10). The active site for all three SPs is centrally located under the β-propeller structure (green star *).

1.5.1. Epithiospecifier protein (ESP): Epithiospecifier proteins redirect the aglycone intermediate rearrangement to form epithionitriles (ETNs, nitriles with a thiirane moiety) instead of RNCSs. An interesting constraint to ESP function is the formation of ETNs will only occur from aglycones whose R group has a terminal double bond (i.e., sinigrin), all other aglycones will end up as simple nitriles (RCNs). The mechanism of action for all SPs is still under scrutiny, but studies using isotopically labeled GLS show that the sulfur atom in the thiirane ring of ETNs originates from the sulfur in the thioglucosidic bond of GLS.[56] Brocker

and Benn suggest the sulfur atom is abstracted from the central carbon of the thiohydroximate moiety and transferred to the end of the unsaturated side chain where it breaks the double bond and cyclizes to form a thiirane ring via an Fe(III)-sulfur radical mechanism.[56] To date, there is no existing experimental evidence to support this hypothesis, or any other, although there have been mechanisms proposed based on the action of other enzymes, or predicted with modeling software (Figure 1-10).[41,48] The largest inconsistency in mechanistic knowledge, for all three SPs, is a complete lack of empirical evidence for a radical based mechanism (i.e., formation of an Fe(III)-sulfur radical). To our knowledge, there have not been any spectroscopic studies to verify the formation of an Fe(III)-sulfur radical, and anoxic studies that include radical scavengers do not result in SP inhibition, which would be expected if the SP mechanism requires formation of the Fe(III)-sulfur radical for activity.[57]

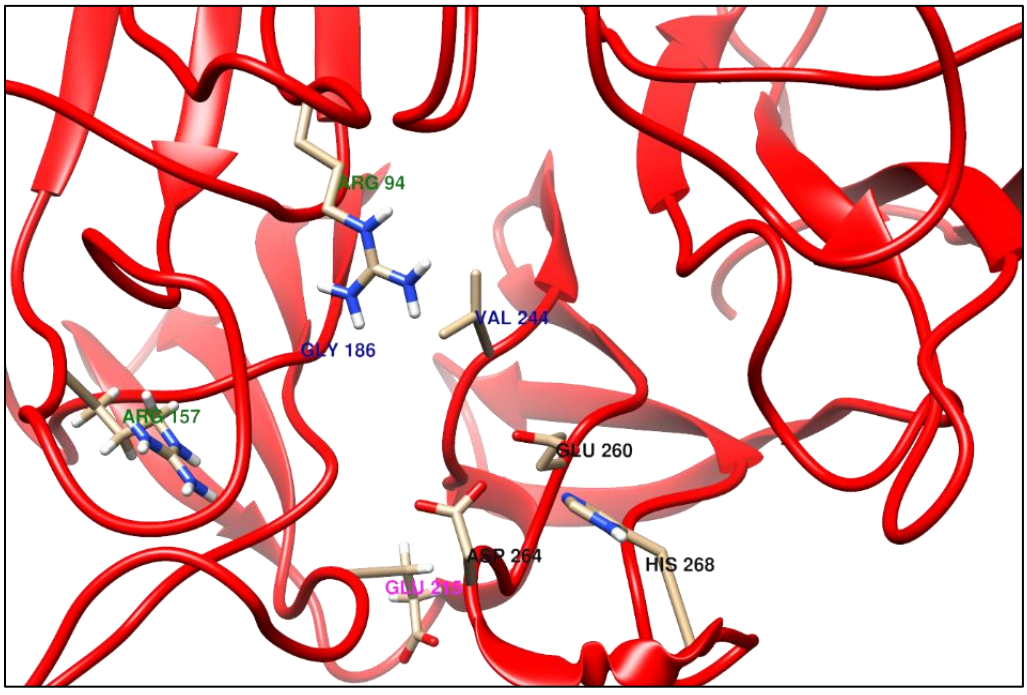


Figure 1-9. Proposed binding ligands for AtESP suggested binding residues for ESP. Iron binding sites include E260, D264, H268 (verified with mutants E260A, E260Q, D264A, and D264N that showed complete loss of ESP activity). Sulfate binding residues are R94 and R157 (R94K mutant showed highly reduced ESP activity). Substrate stabilization sites G186

and V244 were also confirmed with mutational analysis (G186M and V244C result in complete knockout of ESP activity).[48,53]

Eisenschmidt-Bönn et al.[38] report the most recent attempt to describe the intricate details involved in the catalytic rearrangement of aglycone intermediates by SPs to non-isothiocyanate products. Using molecular modeling and intermediate state stability calculations ("*In silico* loop structure and semiempirical quantum mechanical calculations"), Eisenschmidt-Bönn et al. mapped the formation of thiocyanate (TCN) and ETN from *Ta* TFP (TFP from *Thlaspi arvense)*, and simple nitriles (RCN) from *At* NSP (Figures 1-10, 1-12 and 1-14). As *Ta* TFP has nearly 70% sequence identity to *At* ESP, it is plausible that the mechanism of action for the formation of ETN are synonymous in *Ta* TFP and *At* ESP.[47] In the proposed mechanism for ETN formation from allyl aglycone, the iron cofactor bound in the active site coordinates to the thiolate on the aglycone (Figure 1-10). Protonation of the sulfate group by an Arg residue is followed by its spontaneous release (Figure 1-10, step 1). The intermediate is stabilized by a single electron transfer to the N from the oxidation of Fe(II) to Fe(III) (Figure 1-10, step 2). The first transition state forms by electron distribution between C2, S and C. Close conformation of the sulfur atom to C2 and C-S promotes the homolytic opening of the C-S and C2-C3 bonds (Figure 1-10, step 3). The N atom then appears to lose its radical nature and forms a nitrile group while the C2-S bond forms (Figure 1-10, step 4). Transition state 2 catalyzes convergence of the C3 radical and the sulfur atom to form the thiirane ring (Figure 1-10, step 5), and the single electron reduces Fe(III) back to Fe(II) (Figure 1-10, step 6). The deprotonation of a Glu residue rebuilds the metal complex to enable another turnover.

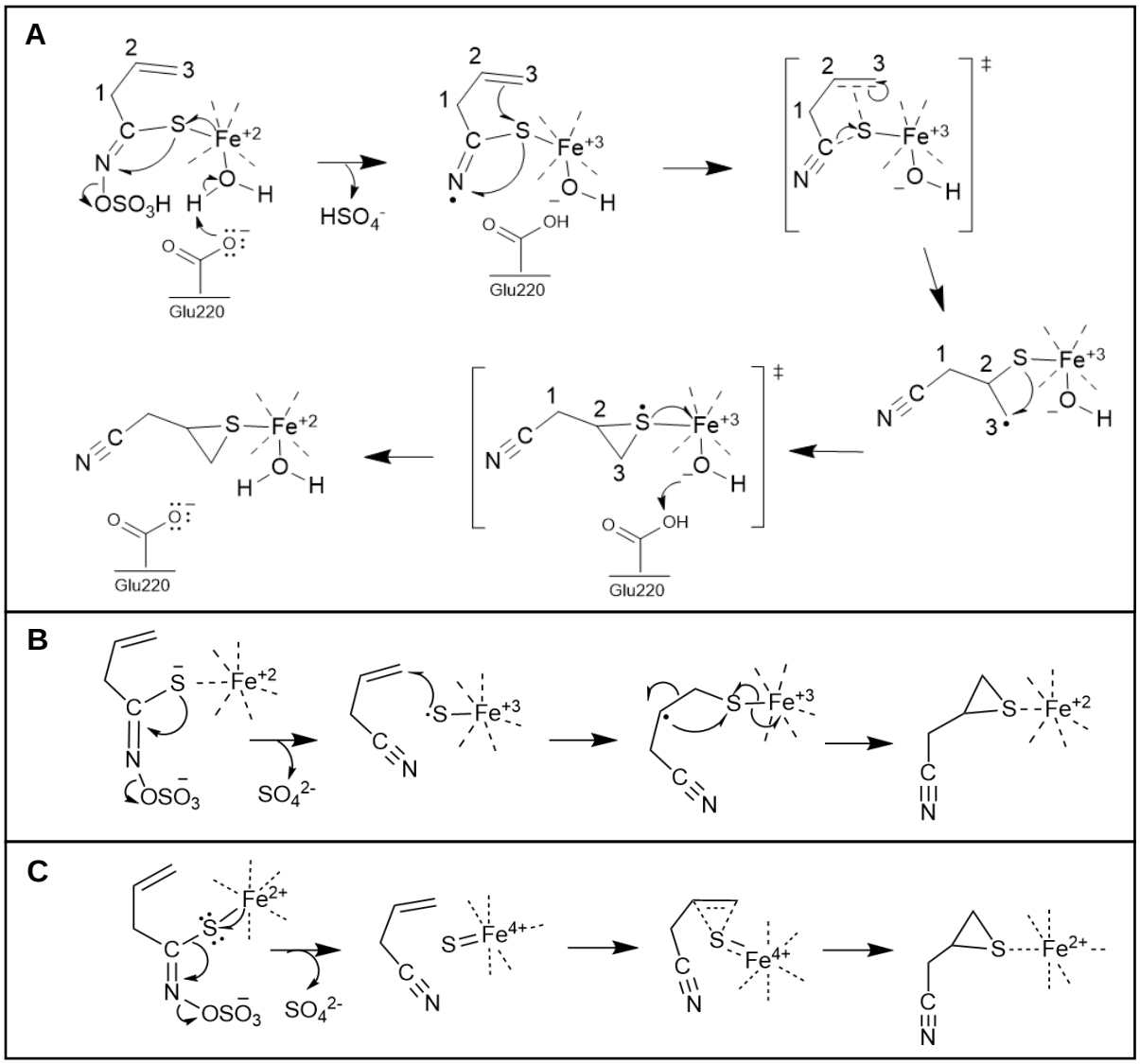


Figure 1-10. Proposed reaction mechanisms of ESP: (A) Scheme proposed by Eisenschmidt-Bönn et al.[38] (B) Mechanism adapted from Backenköhler et al.[41] (C) Mechanism proposed in this work.

1.5.2. Thiocyanate Forming Protein (TFP): The discovery of high sequence similarity of *Ta* TFP to *At* ESP confirmed that TFPs are members the SP family. However, TFP is the least abundant SP in brassicas with only three isoforms identified to date (Table 1-2), from three different brassica species. Additionally, TFP catalyzes formation thiocyanates (TCN) from the aglycones of only three GLS, sinigrin (allyl glucosinolate), glucotropaeolin (benzyl glucosinolate), and glucoerucin (4-methylthiobutyl glucosinolate). TFP also has the ability to

form all three of the possible alternate hydrolysis products (TCN, ETN and RCN) depending on the GLS side chain.[44] The limited capacity for TCN formation from any other aglycone side-chain composition is speculated to be due to the small size of the TFP active site. According to this hypothesis, the aglycone binding position is highly constrained and the conformation of the bound aglycone may determine the reaction outcome (TCN, ETN, or RCN, Figure 1-11).[38]

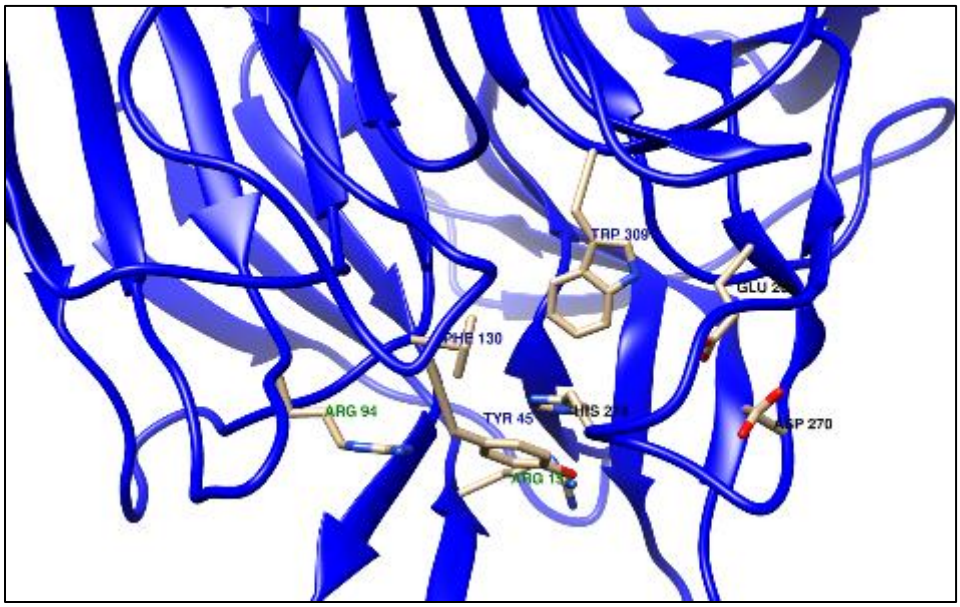


Figure 1-11. Proposed binding characteristics of *Ta* TFP. The iron coordinating residues are D270, H274, and E266 (verified by H274C, H274G, H274E, and E266Q mutants that showed complete knockout of TFP activity). Sulfate binding residues R94 and R157 were confirmed with R94A and R157A mutants that resulted in high reduction in and complete lack of ESP activity, respectively. Aglycone recognition sites are W309, Y45, and F130 (confirmed by W309A, W309F, Y45N, and F130A mutants that show no ESP activity).[47,53]

The modeling experiments by Eisenschmidt-Bönn et al. using allyl aglycone suggests that the catalytic properties of TFP facilitate the proper positioning the Fe(II) cofactor, as well as provide the appropriate amino acid residue structure to support specific aglycone conformations, and protonation/ deprotonation of intermediates (Figure 1-12). Their central claim is the Fe(II) is oxidized to Fe(III) for intermediate stabilization after the dissociation of

the sulfate group for ETN and TCN formation, but not for RCN formation. It is important to note that the molecular modeling in this experiment suggest that the iron cofactor is actually coordinated to the terminal double bond of the aglycone for TCN, instead of the iron coordinating to the sulfur as predicted by others and as is shown for ETN formation.[38] However, given that other studies have proven the capability of TCN formation with benzyl and 4-methylthiobutyl aglycones, their proposed mechanism is only applicable to allyl aglycone. To our knowledge, there are no models for TFP-catalyzed TCN formation from benzyl or 4-methylthiobutyl aglycones.

The first two steps of the Eisenschmidt-Bönn et al. model for TCN formation by *Ta* TFP using allyl aglycone is the same as ETN formation, protonation of the sulfate moiety triggers its spontaneous release (Figure 1-12, step 1) and the oxidation of Fe(II) to Fe(III) stabilizes the intermediate by a single electron transfer to the N atom (Figure 1-12, step 2). Dissociation of an Fe(III) coordinated water molecule protonates a Glu residue, which is supposedly linked to facilitating departure of the sulfate group. The first intermediate transition state forms when the aglycone thiolate converges on C3 (the terminal carbon in the double bond), arranging the sulfur atom perpendicular to the partial-positive side of the double bond (Figure 1-12, step 3). The resulting orbital overlap leads to the formation of a cyclic intermediate with a C2 carbanion (Figure 1-12, step 4). The C2 carbanion donates the free electron pair to the Fe(III) complex, and the Glu residue is deprotonated to restore the coordinated water molecule. The second transition state, which is considered the rate limiting step, involves homolytic cleavage of the C – C1 bond and results in TCN bound to Fe(III) (Figure 1-12, step 5). The final step is the dissociation of TCN from the metal complex, where

the electron pair donated to the Fe(III) is used for formation of the double bond, and the single electron reduces Fe(III) back to Fe(II) (Figure 1-12, step 6).

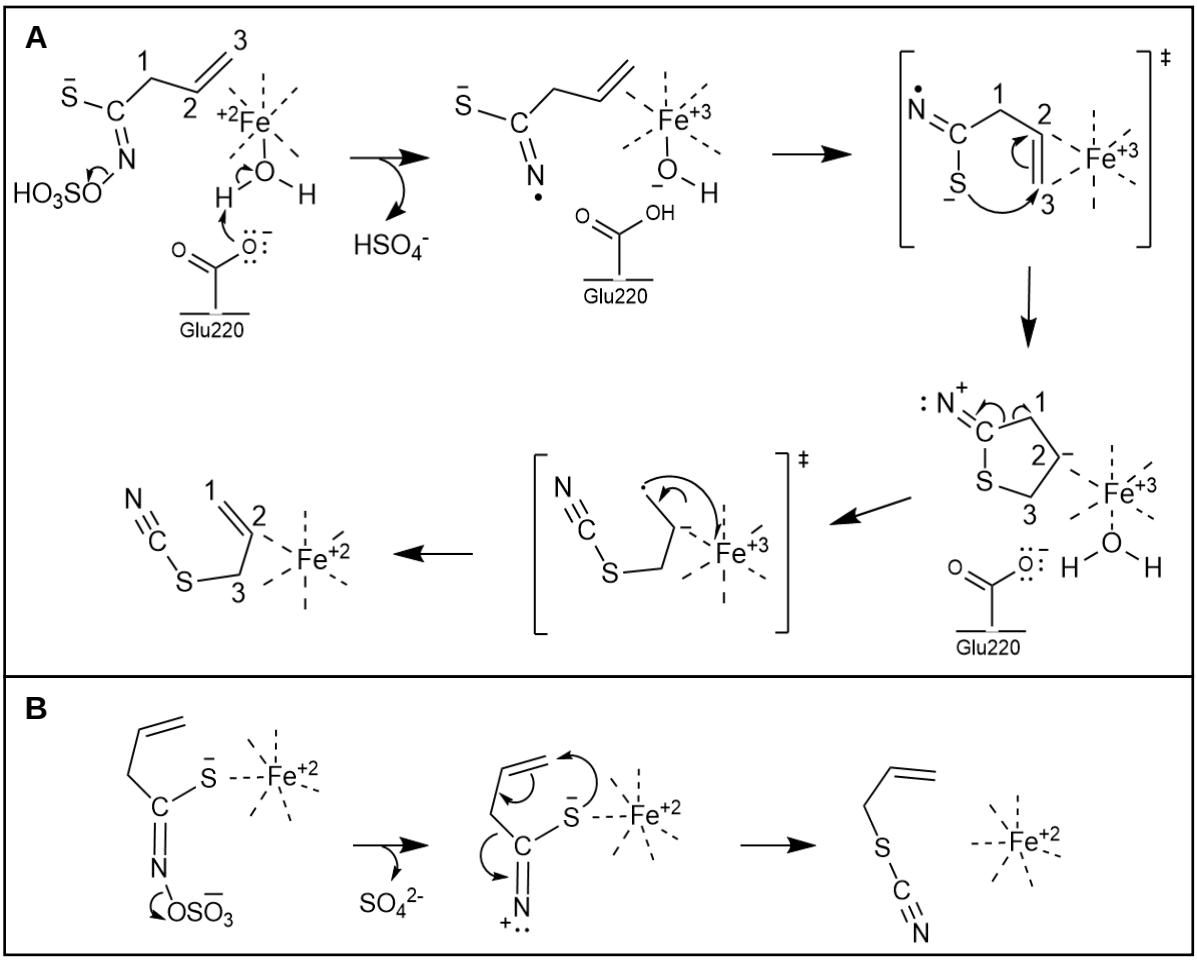


Figure 1-12. Proposed reaction mechanisms of TFP. (A) Scheme proposed by Eisenschmidt-Bönn et al.[38] (B) Mechanism adapted from Backenköhler et al.[41]

1.5.3. Nitrile Specifier Protein: Although NSP does not exist as a homodimer in nature, its monomer does share high sequence identity with the β – propeller structure of the other SPs (> 70%). The NSPs have very low substrate specificity compared to TFP and ESP, it will form RCNs from any aglycone.[52] The current assumption with the formation of RCNs involves releasing the abstracted sulfur atom as elemental sulfur.[52,55] However, the only known evidence for this was presented in 1977[58] but hasn't been shown since, and both the solved crystal

structure and the modeled structures did not identify any acceptor for the sulfur atom to support this (Figure 1-13).

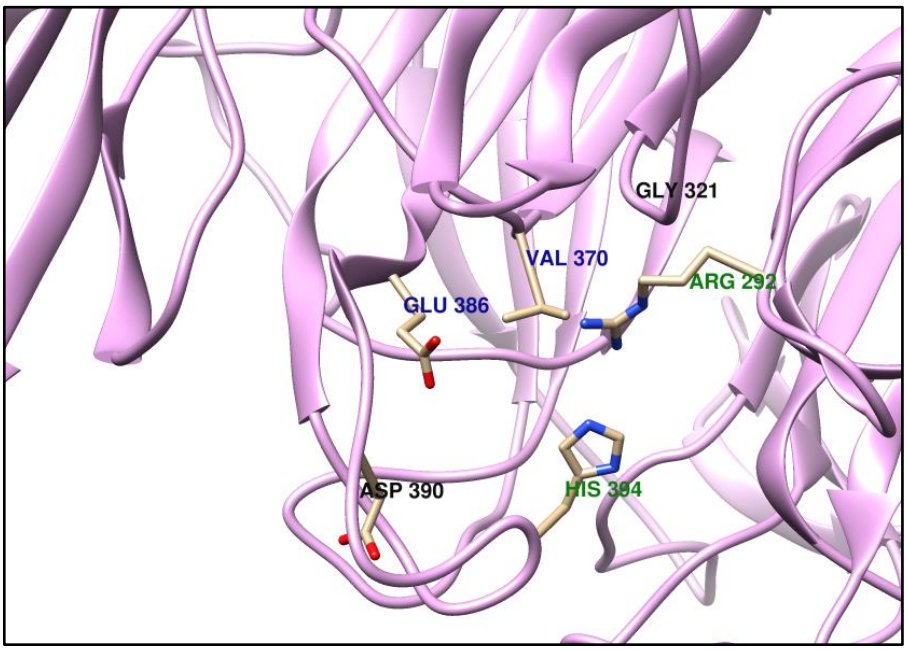


Figure 1-13. Proposed binding sites of *At* NSP. Iron coordination sites are E386 and D390. The sulfate binding residues are R292 and H394. Aglycone substrate stabilization is by G321 and V370. There has been no mutational analysis to confirm these active site characteristics for NSP to our knowledge.

The molecular modeling experiment by Eisenschmidt-Bönn et al. describes the formation of RCN from AtNSP3 as much simpler than the TCN/ETN formation by TFP. The formation of RCN is initiated by the removal of the sulfate group from the aglycone, but in the case of NSP, the C-S bond undergoes heterolytic cleavage during sulfate dissociation (Figure 1-14). Their model predicts a lack of charge transfer during this process, suggesting the Fe(II) does not participate in a redox capacity like it does for the other SP reactions. Alternatively, the Fe(II) may only facilitate the charge displacement necessary for sulfur abstraction.

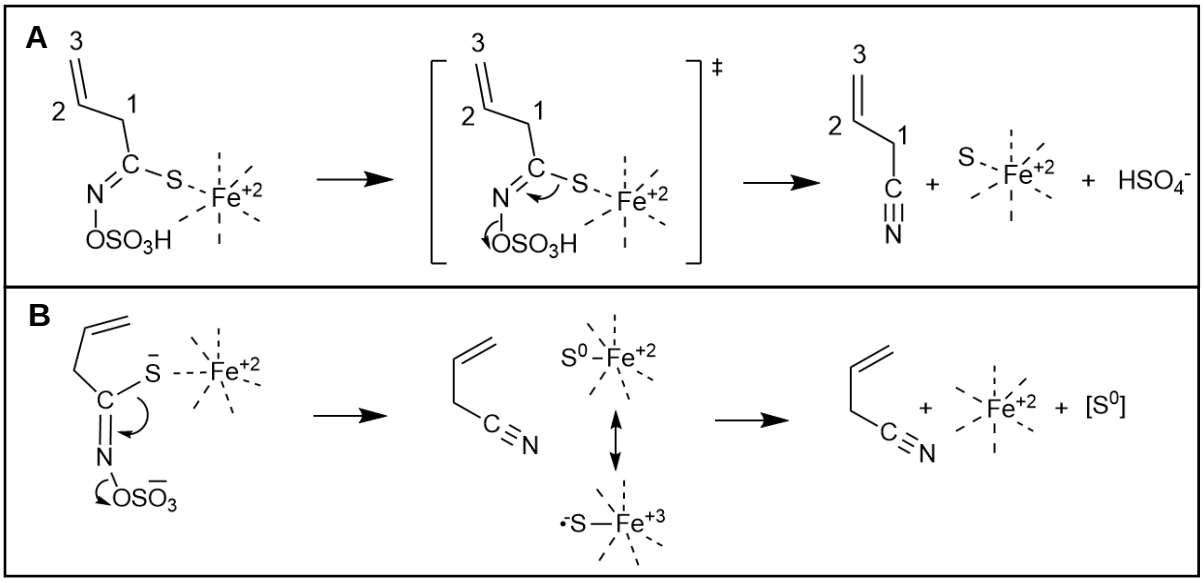


Figure 1-14. Proposed reaction mechanisms of NSP. (A) Scheme proposed by Eisenschmidt-Bönn et al.[38] (B) Mechanism adapted from Backenköhler et al.[41]

1.6. APPLICATIONS FOR THE GLS-MYR SYSTEM IN HUMAN HEALTH

Brassica vegetables are an important component to the human diet as they contain several health-promoting phytochemicals in addition to glucosinolates, such as carotenoids, tocopherols, and polyphenols.[59] Although all of these phytochemicals have an impact on human health, most current studies are focused on the therapeutic effects of RNCSs, specifically their anticancer (chemopreventative) properties. The most prolific and promising RNCS for cancer treatment is 4-methylsulfinylbutyl isothiocyanate, or sulforaphane (SFN), the RNCS hydrolysis product of glucoraphanin (4-methylsulfinylbutyl-glucosinolate). Because broccoli is a commercially relevant crop that is typically high in glucoraphanin content, it is the most commonly studied plant for chemotherapeutic benefits.[2]

To date, there are several comprehensive reviews that discuss the efficacy of SFN for cancer treatments, summarize the results of clinical trials, and outline the current progress in RNCS research in human health.[59–61] The chemopreventative properties of SFN result from

several modes of action, of which the most significant and well-characterized are summarized below:

(i) SFN influences the modulation of Phase I and II metabolism enzymes by targeting the NrF2/Keap1/ARE-signaling pathway in cancerous cells. SFN promotes the inhibition of Phase I enzymes, preventing the activation of carcinogenic compounds, and promotes the induction of Phase II enzymes which augments the elimination of carcinogens.[2,62,63]

(ii) SFN promotes downregulation of cyclooxygenase-2 and Bcl-2/ Bcl-xL protein expression, which inhibit cancer cell proliferation and induces apoptosis, respectively. The downregulation ultimately results in elimination of the cancer cell clonal expansion.[60,63]

(iii) SFN inhibits metastasis and invasion of cancer cells through modulation of mitogen-activated protein kinase and cyclin-dependent kinase signal transduction pathways.[63,64]

Studies on SFN are numerous and progress seems to be rapid, but there are other RNCSs that have anticancer characteristics as well, including allyl isothiocyanate (AITC, from sinigrin), benzyl isothiocyanate (BITC, from glucotropaeolin) and phenylethyl isothiocyanate (PEITC, from gluconasturtiin). Each ITC exhibits slightly different efficacies, which are summarized by Fofaria et al.[62] AITC appears to have short-term effects on cancer cells by reversibly damaging cancer cell DNA leading to apoptosis; BITC causes cancer cell cycle arrest and disruption of the mitochondrial membrane potential through generation of ROS (reactive oxygen species); PEITC also influences the generation of ROS, and inhibits ROS detoxification mechanisms in cancer cells. Interestingly, Fofaria et al. also describe the

synergistic effects of using RNCSs in combination with other chemotherapeutic agents (cisplatin, paclitaxel, metformin, etc.), which ultimately results in enhancement of the antitumor effect on various types of cancer.

Another interesting application for RNCSs in human health is their use in antimicrobial food packaging. Recent advances in food spoilage prevention show that activation of food packaging surfaces with volatile antimicrobial agents results in inhibition of the growth of spoiling bacteria on food surfaces.[65] RNCSs are effective antimicrobial agents against multiple microorganisms that are responsible for food spoilage such as *A. parasiticus* (mold) and *E. coli* (bacteria).[66,67] This application of RNCSs increases shelf life and reduces food waste, which supports food sustainability and security. As RNCSs are highly volatile, activated food packaging will only remain antimicrobial for a short period of time. To combat this problem, a recent study by Bahmid et al.[65] determined the efficacy of mustard seeds (which contain MYR and are typically high in GLS content) as an avenue for the slow release of RNCS into the headspace of food packaging. Their results indicate that the size and lipid content of the mustard seeds are directly responsible for the rate of RNCS release from the seeds. The smaller the seed particle, the faster the release of RNCS, and higher levels of lipids slow the release of RNCSs. This new application for RNCSs shows the utility of the GLS-MYR system in our everyday life in addition to cancer prevention.

1.7. AGRICULTURAL APPLICATIONS OF THE GLS-MYR SYSTEM

The use of brassica crops for soil biofumigation dates back past the 1920s, when is was discovered that the presence of mustard plants resulted in a reduction of cyst nematodes on potato roots in the United Kingdom.[68,69] Biofumigation is an agronomic approach to soil-borne pest management by utilizing the biocidal nature of RNCSs that are produced from grinding

brassica plants.[70] As the type, distribution, and concentration of GLS in different plant parts and species tends to vary significantly, the identity/bioactivity of the RNCSs that are generated also vary accordingly.[71] A review by Matthiessen and Kirkegaard[33] compares the attributes and shortcomings to using brassicas for biofumigation, and emphasizes the necessity of providing an accurate assessment of the biofumigation potential of a specific brassica cultivar when performing empirical studies. This assessment is most easily described by reporting the GLS content in the plant of interest, but can be a slight over estimation due to any inherent inefficiencies by which GLS are converted to RNCS *in vivo*.[59] Additionally, many environmental factors affect the GLS content in brassicas (i.e. sunlight,[72] available nutrients,[73] temperature,[74] water content,[75] etc.), meaning the most accurate estimation of biofumigation potential must be measured by cultivating the plants in the environment of interest.[33] Successful soil biofumigation with brassica crops is a complex task due to the highly variable nature of the processes involved, but research continues to progress as more information about the GLS-MYR system becomes available. A recent review by dos Santos et al. summarizes the results of the current studies on biofumigation with brassicas to control fungi, oomycetes, nematodes, bacteria, and insects.[76]

The use of forage quality brassica crops for cool-season livestock grazing has also been studied and applied for many years. Feeding animals on forage brassicas naturally extends the grazing season into the winter months while continuing to provide sufficient levels of crude protein.[77] However, numerous experiments have emphasized the potential negative health side-effects of a high brassica diet (typically > 25% brassica, but side-effects are dependent on the GLS profile of the brassica) on grazing ruminants. Studies indicate that the byproducts from GLS hydrolysis, initiated during grazing (mastication and consumption), can cause health

issues such as low fertility[78], decreased weight gain and milk production[79], and thyroid problems that lead to goiter.[80,81] A review by Tripathi et al.[82] outlines the specific effects of a brassica diet on several species of grazing animals (including pigs, steers, calves, dairy cows, goats, and lambs) and reports the maximum GLS content livestock can tolerate without deleterious effects (Table 1-3).

Animal	GLS tolerance level (μmol g^{-1} diet)
Goat	16
Dairy Cow	11
Growing calve	7.7
Growing steer	10-15
Rabbit	7
Poultry	2-4
Fish	1.4
Pig	0.8
Rat	0.5

Table 1-3. Summary of different animal species maximum tolerance to GLS before negative side effects associated with the GLS-MYR system begin.[83–90] Adapted from Tripathi et al.[82]

Brassicas that contain a particularly high concentration of GLS whose side chain has a hydroxyl group at the C2 position (i.e. progoitrin, which is reportedly high in some turnip and rapeseed varieties[91,92]) have been linked to thyroid conditions, such as goiter, in ruminant animals fed on a high brassica diet. The thyroid regulates metabolism through secretion of T3 and T4 hormones (triiodothyronines and tetraiodothyronines), whose biosynthesis are dependent on the levels of iodide in the blood stream. Therefore, goiter, or enlargement of the

thyroid gland, is primarily caused by iodine deficiencies, and goitrogens are any compound that interfere with thyroxinogenesis (the biosynthesis of the T3 and T4 hormones).[81] Brassicas that contain high levels of progoitrin have been linked to goiter in grazing livestock because the product of progoitrin hydrolysis, 2-hydroxyl-3-butenyl-NCS, will undergo non-enzymatic cyclization to form 5-vinyloxazolidine-2-thione, otherwise known as goitrin (a highly goitrogenic compound, shown in Figure 1-1).[51] The thiocarbamide group of goitrin most likely contributes to the inhibition of thyroxinogenesis by blocking the incorporation of iodide into T3/T4 hormones. Any RNCSs with similar hydroxylation at C2 will result in a goitrogen, but progoitrin is the most prominent in brassicas.[80]

It is clear that the hydrolysis products of the GLS-MYR system play in a pivotal role in many applications from human cancer treatment to livestock diet and crop production. It is reasonable to assume that improvement to a plant's innate ability to produce RNCSs would result in greater biocidal efficacy. Alternatively, a reduction in natural RNCS production, perhaps by promoting the formation of other less-toxic degradation products or by downregulating GLS biosynthesis, would increase brassica palatability and reduce health risks to livestock. However, the mechanism of formation and evolutionary implications of the alternate hydrolysis products are unknown and these improvements to the GLS-MYR system are not yet possible. Therefore, future studies should continue to focus on the GLS profiles of brassica crops, and how to harness those profiles for practical applications. To address the need build GLS profiles of agriculturally relevant brassicas, this work demonstrates a simple, cost-effective method for generating a semi-quantitative GLS profile of individual brassica species (Chapter 2).

Chapter 2

Simple identification and accurate mass confirmation of individual glucosinolates in brassica varieties by mass spectrometry

2.1. INTRODUCTION

Glucosinolates (GLS) are a class of secondary metabolites unique to plants of the Brassicaceae family. They are an essential component of the glucosinolate-myrosinase system, the innate defense mechanism in brassica plants.[50,93] The Brassicaceae family includes economically important crops such as canola (*Brassica napus* L.), cabbages (*B. oleracea* L.), and other important forages and livestock feedstuffs such as rapeseed (*B. napus* L.), kale (*B. oleracea*), radish (*Raphanus sativus* L.), turnip (*B. rapa* L.), and various oilseed meals. Glucosinolates can be found in all parts of the plant, but the occurrence and concentration of individual GLS vary according to species, variety within species, environmental factors, age, and type of plant tissue.[94]

Glucosinolates maintain a common core structure of a thiohydroximate-*O*-sulfate, with an *S*-β-D-glucopyrano unit anomerically linked to the sulfur atom.[1] Individual GLS are differentiated from one another by a single substituent, a variable "R-group" side chain attached to the carbon center of the hydroximate. The distinguishing side chain is determined by the amino acid precursor for GLS biosynthesis, the most common of which are methionine, phenylalanine, and tryptophan. Each amino acid-derived R-group can undergo additional modifications to its structure both pre- (i.e., chain elongation) and post- (i.e., hydroxylation, methylation, oxidation of sulfur to sulfoxides) GLS biosynthesis. High variability in the R-group chemical composition results in 150+ unique GLS structures.[1,95–97] Approximately two-thirds of these have been structurally verified, and the remaining third are inferred but remain

without proper evidence.[1,98] Despite the large variety of existing glucosinolates, only 10 pure GLS standards are available to purchase, which is prohibitive to conventional quantitative analysis of GLSs. Additionally, the total GLS profile in most brassicas is dominated by only 4 – 5 specific GLS that are different from species to species.[95] The available research related to GLS in livestock nutrition has focused on oilseed by-products (i.e., canola and rapeseed meals) which, due to the oil extraction process, have a greater concentration of GLS than would be expected in leaf and stem tissues of plants used as forages such as those used in this study.[99]

Intact GLS in the ungrazed plant are biologically inert. However, when tissue damage occurs (i.e. insect or animal mastication), GLS catabolism is initiated by release of the compartmentalized β-thioglucosidase enzyme, myrosinase (MYR, EC 3.2.1.147).[9] Myrosinase hydrolyzes the thioglucosidic bond of a GLS, producing an unstable aglycone (thiohydroximate-*O*-sulfonate) intermediate. The transient intermediate spontaneously degrades via a Lossen-like rearrangement to form an isothiocyanate (RNCS), which is highly reactive towards biological nucleophiles.[31] The RNCSs produced from the glucosinolate-myrosinase system are responsible for the biochemical defense capabilities of brassicas due to their electrophilic nature. The bioactivity of RNCSs has been harnessed for biofumigation in agriculture for many years, ensuring protection against soilborne pests without intensive use of pesticides.[71] However, RNCSs have also been shown to be the source of several health disorders in livestock grazed on forage brassicas, including goiter,[80] infertility,[78] and reduced feed intake and growth.[100] Because brassicas can be incorporated as a considerable portion of livestock diet, it is prudent to determine the GLS profile of these crops to avoid the associated health concerns and maintain animal productivity. The specific biological effects of individual RNCSs vary considerably and therefore have diverse effects on both soil and animal health.[31]

As more chemical information on individual GLS becomes available, it is evident that the presence of individual GLS is species/variety-specific, and the different chemical characteristics of their corresponding RNCSs imply variation in their potency for soil biofumigation and risk of adverse health effects in animals.[101–104] Therefore, determination of individual GLS in species used for cover crops and as feedstuffs for livestock is critical.[71,105] Previous methods have described successful GLS analyses, but require completion of complex or unconventional procedures, which hinders efficiency of sample preparation and analysis. Examples of inefficient steps include desulfatation of intact GLS,[106,107] boiling solvents,[107–109] lengthy chromatographic separations,[110] multiple extraction steps,[111] and quantification of GLS that is dependent on the purchase of authentic GLS standards. Other successful methods that employ more conventional techniques typically include a straightforward extraction of intact GLS at room temperature[112] and efficient detection using liquid chromatography (LC) with electrospray ionization (ESI) coupled to a mass spectrometer (MS).[108,109,113,114]

The existing methods for GLS analysis have inefficiencies that increase processing time, complexity, and cost. Additionally, the current methods for quantification of individual GLS in forage brassicas are lacking and there is opportunity to streamline the analyses. The objectives of this study are to develop a method for quick and accurate identification of the individual GLS present in leaf and stem tissues of forage brassicas; and semi-quantitatively assess the total GLS profile of three relevant forage brassica species frequently grown in the northeastern U.S. The purpose of the method described herein is a simple laboratory technique to enable researchers to build an accurate survey of the GLS profile for a specific brassica plant. The intended application of an accurate GLS profile is to supplement the information needed for proper species selection for targeted agricultural uses, such as biofumigation and

livestock feed. Additionally, our proposed method can also be used to determine whether a GLS profile requires more in-depth characterization or analysis, such as high abundance of an apparently unknown GLS.

2.2. RESULTS AND DISCUSSION

2.2.1. Method optimization.

The LC-MS method was optimized using a sulfonamide embedded stationary phase to enhance the separation of the polar GLS. Compared to methods using a C18 stationary phase,[107,109,112,115] it was found that GLS with indole, aromatic or polar side chains were poorly retained, demonstrated by co-elution, shifting retention time (RT) or poor peak shape (Appendix A, Figure S2-1). The reproducibility of the extraction technique was verified by six replicate preparations and analyses of a rapeseed herbage sample, and five replicated preparations of a turnip herbage sample (Appendix A, Table S2-2). The standard error for the total GLS content in both the turnip and rapeseed sample was < 3%. Extraction efficiency was determined by performing a second extraction step on a turnip sample (in duplicate) and measuring the GLS content in both the first and second extraction (Appendix A, Table S2-3). Our results showed that less than 9% of the sinigrin internal standard and other observable GLS remained in the plant material after the first extraction step.

Sinigrin (0.5 ppm) was used to perform standard addition analysis. Peak area reproducibility of sinigrin from 8 replicates (0.5 ppm sinigrin in water) using this gradient method showed less than 5% variation (Appendix A, Table S2-4). The average value of sinigrin internal standard detected in blank samples was 0.46 ppm (8% error from concentration spiked before extraction). The signal-to-noise (S/N) ratio was no greater than 2 for any GLS, and the variation in retention time for all GLS, both inter- and intraday

experiments, was between 0.73 and 3.7 seconds. The mass errors from the measured *m/z* values compared to the precursor *m/z* values were ≤ 3.6 ppm (Table 2-1).

We performed serial dilutions in duplicate using extracts from turnip, rapeseed, and canola plant material. All methanol extracts are diluted 1:10 with water before analysis for cleaner samples (to protect the equipment). To evaluate the potential effects of dilution and sample matrix on the ionization efficiency of GLSs, we also analyzed 1:100 and 1:1000 dilutions of the methanol extracts. Serial dilutions were also performed on a GLS standard solution containing glucobrassicanapin, gluconapin, sinigrin, sinalbin, progoitrin, glucoraphanin, and glucoiberin (from 0.001 – 10 ppm) (Appendix A, Figure S2-3). Our results show that all GLS exhibit a linear response to serial dilutions and the matrix (extracted plant material) does not interfere with GLS ionization (Appendix A, Table S2-6).

Glucosinolate (semi-systematic name)	RT (min)	Precursor (*m/z*)	Mass Measured (*m/z*)	Mass Error (ppm)	PRM Confirmation Fragment (*m/z*) ($R+CNHOS^-$)	PRM Confirmation Fragment 2 (*m/z*) ($R+C_2NHO4S^-$)or($R+CS_2O_3^-$)
2-propenyl	3.94	358.02720	358.02788	-1.8993	116.0176	161.9861
1-methylethyl	4.45	360.04285	360.04406	-3.3607		
3-butenyl	5.27	372.04285	372.04342	-1.5321	130.0332	176.0018
1-methylpropyl	6.30	374.05850	374.05926	-2.0318	132.0489	176.0018
2-methylpropyl	-	374.05850	-	-		
n-butyl[2]	-	374.05850	-	-		
4-pentenyl	7.66	386.05850	386.05922	-1.8650	144.0494	192.9999
2(R)-2-hydroxyl-3-butenyl	-	388.03776	-	-		
2(S)-2-hydroxyl-3-butenyl	3.79	388.03776	388.03845	-1.7782	146.0291	191.9972
2-methylbutyl	-	388.07415	-	-		
3-methylbutyl[2]	-	388.07415	-	-		
n-pentyl[2]	9.00	388.07415	388.07434	-0.4896		
1-(hydroxymethyl)propyl	3.50	390.05341	390.05434	-2.3843		
2-hydroxy-2-methylpropyl	3.60	390.05341	390.05390	-1.2562		
3-hydroxybutyl[2]	4.15	390.05341	390.05382	-1.0511		

4-hydroxybutyl[2]	4.42	390.05341	390.05439	-2.5125		
5-hexenyl	9.60	400.07415	400.07426	-0.2749		
2(R)-hydroxy-4-pentenyl	4.20	402.05341	402.05428	-2.1639		
2(S)-hydroxy-4-pentenyl	4.82	402.05341	402.05454	-2.8106	160.0443	208.9948
4-oxopentyl	-	402.05341	-	-		
2-hydroxy-2-methylbutyl	4.00	404.06906	404.06947	-1.0147		
Benzyl	8.20	408.04285	408.04326	-1.0048		
4-(methylthio)butyl	8.09	420.04622	420.04706	-1.9998	178.0366	226.9871
2-phenylethyl	9.80	422.05850	422.05931	-1.9192	180.0494	228.9999
4-hydroxybenzyl	4.90	424.03776	424.03653	2.9007	182.0281	230.9786
5-(methylthio)pentyl	9.73	434.06187	434.06254	-1.5436		
4-(methylsulfinyl)butyl	3.38	436.04113	436.04231	-2.7062	194.0315	240.0001
2-(3-hydroxyphenyl)ethyl	7.42	438.05341	438.05412	-1.6208		
2(R)-hydroxy-2-phenylethyl	7.65	438.05341	438.05455	-2.6024		
4-hydroxyphenethyl	9.06	438.05341	438.05444	-2.3513		
3-methoxybenzyl	9.24	438.05341	438.05429	-2.0089		
4-methoxybenzyl	9.70	438.05341	438.05426	-1.9404		
3,4-dihydroxybenzyl	4.10	440.03267	440.03144	2.7952		
Indol-3-ylmethyl	9.40	447.05374	447.05479	-2.3487	205.0441	253.9946
3-hydroxy-5-(methylthio)pentyl[2]	-	450.05678	-	-		
5-(methylsulfinyl)pentyl	3.87	450.05678	450.05748	-1.5554	208.0472	256.9976
4-(methylsulfonyl)butyl[2]	3.30	452.03605	452.03767	-3.5838		
(R)-2-hydroxy-2-(3-hydroxyphenyl)ethyl	3.90	454.04833	454.04792	0.9030		
(R)-2-hydroxy-2-(4-hydroxyphenyl)ethyl	-	454.04833	-	-		
(S)-2-hydroxy-2-(4-hydroxyphenyl)ethyl	-	454.04833	-	-		
3-hydroxy-4-methoxybenzyl	5.30	454.04833	454.04864	-0.6827		
4-hydroxy-3-methoxybenzyl	-	454.04833	-	-		
4-hydroxyindol-3-ylmethyl	8.25	463.04866	463.04922	-1.2094	221.0396	269.9901
4-methoxyindolyl	9.50	463.04866	463.04921	-1.1878		
3-hydroxy-6-(methylthio)hexyl[2]	-	464.07243	-	-		
6-(methylsulfinyl)hexyl	4.90	464.07243	464.07311	-1.4653		
3-hydroxy-5-(methylsulfinyl)pentyl[2]	3.37	466.05117	466.05274	-3.3687		

?-methoxyindol-?-ylmethyl*	9.42	477.06431	477.06474	-0.9013	235.0552	284.0057
4-methoxyindol-3-ylmethyl	10.78	477.06431	477.06468	-0.7756	235.0552	284.0057
1-methoxyindol-3-ylmethyl	12.03	477.06431	477.06479	-1.0062	235.0552	284.0057
4-hydroxy-3,5-dimethoxybenzyl	7.63	484.03834	484.03870	-0.7437		
2-benzoylprogoitrin[2]	9.41	492.06343	492.06314	0.5894		
3,4,5-trimethoxybenzyl	7.75	498.07454	498.07339	2.3089		
1,4-dimethoxyindol-3-ylmethyl	10.80	507.07487	507.07555	-1.3410		

Table 2-1. Summary of individual glucosinolates (GLS) and their isomers identified in 'Appin' forage turnip, 'Barisca' forage rapeseed, 'Inspiration' winter canola. We provide the mass error between the measured and precursor *m/z* values and detected PRM confirmation fragments to demonstrate the accuracy of the current method. [1]GLS that were identified with an authentic standard. [2]GLS that have not been fully characterized or lack proper evidence for existence. GLS that cannot be definitively identified by retention time (RT) are listed in this table in alphanumeric order for simplicity in data presentation. *An unknown isomer of neoglucobrassicin appears to be present, but does not have a systematic or common name.

2.2.2. Limits of detection and quantification.

With only one authentic standard with reliable purity (sinigrin), it was determined that the LOD and LOQ for the other GLS would be set to one order of magnitude greater than what the limits were for the authentic standard. The LOQ and LOD for sinigrin were determined by S/N ratio as 0.4 and 0.1 ng/ml respectively. The area counts of these concentrations were compared to the area counts of the other GLS and were set to a threshold value giving a LOQ and LOD of 4 and 1 ng/ml respectively. The linear calibration range for sinigrin, which was analyzed in triplicate, had a maximum relative standard deviation of 5% (at the low end) and a dynamic range of six orders of magnitude. Additionally, blank samples were analyzed between every calibration standard and unknown sample. The maximum amount of sinigrin carried over to the blank sample was seen for the highest calibration standard (100 µg/ml sinigrin), and was less than 0.1% (100 ng/ml).

2.2.3. Uncertainty in GLS identification.

A number of GLS have masses that are similar or equal to other GLS, making them difficult to distinguish using LC-MS. Several GLS have masses that are isobaric, such as 2(S)-2-hydroxyl-3-butenyl-GLS (progoitrin, *m/z* 388.0377, RT 3.79 min) and n-pentenyl-GLS (*m/z* 388.0741, RT 9.00 min). Other GLS have isomers with the same exact-mass, such as the constitutional isomers 4-hydroxyindol-3-ylmethyl-GLS (RT 8.25 min) and 4-methoxyindolyl-GLS (RT 9.50 min), with a parent ion mass of 463.0486 *m/z*. Our HRAM-MS data is sufficient to unassailably differentiate isobaric GLS, but there are many GLS with one or more constitutional isomers that cannot be verified without authentic standards, application of an isomer-selective chromatographic method, or supporting Nuclear Magnetic Resonance (NMR) analysis.

Analysis of brassica samples for GLS content was completed using extracted ion chromatograms (XIC), where a specific *m/z* value was isolated from the total ion chromatogram (TIC) (Figure 2-1 shows example chromatograms for the major GLS identified using our current method). GLSs that were present in the plant samples in significant amounts (contributing > 1% to the total GLS content of the sample) were subjected to PRM analysis, which confirms the presence of two GLS fragments that are unique to individual GLS: Fragment 1 $[M-242-H]^-$corresponds to $R+CNHOS^-$, Fragment 2A $[M-196-H]^-$ corresponds to $R+C_2NHO_4S^-$, and Fragment 2B $[M-193-H]^-$ corresponds to $R+CS_2O_3^-$.[116,117] See Table 2-1 for a list of confirming fragments, the structures of which are shown in Figure 2-2.

For XICs that present more than one peak per *m/z* value (i.e., isomers with the same exact-mass) in a single sample, tentative GLS identification was made using a combination of information available from recent literature, the predicted retention behavior of GLS isomers,

and, in a few cases, an authentic standard. There also are multiple *m/z* fragments that are unanimously present in all GLS, i.e., *m/z* 74.99 ($C_2H_3OS^-$), 96.96 ($HOSO_3^-$) and 259.01 ($C_6H_{11}O_5SO_4^-$).[98,118] The presence of these common *m/z* fragments provided additional confirmation that these additional peaks represent GLS isomers. We did not do PRM analysis on all GLS with precursors identified because their concentrations in these particular samples are too low, but it would be interesting to investigate whether the rest of the GLS follow the same fragmentation patterns and can use the those shown in Figure 2-2 for further confirmation of identity.

The literature referenced to assist with GLS identification (ID) include a review by Agerbirk and Olsen, 2012 and a recent review by Blažević et al., 2020 which summarize the current knowledge of existing GLS and the common techniques used for detection and characterization.[1,96] Blažević et al., 2020 also lists which GLS are fully characterized, those that require further characterization, and those that have been "discontinued" due to lack of proper evidence. All GLS assignments presented here are tentative where there are no authentic standards available. Appendix A, Table S2-5 summarized all GLS identified in this study with their corresponding semi-systematic names, common names (where applicable), molecular formula, and numeric assignment based on Fahey et al., 2001.

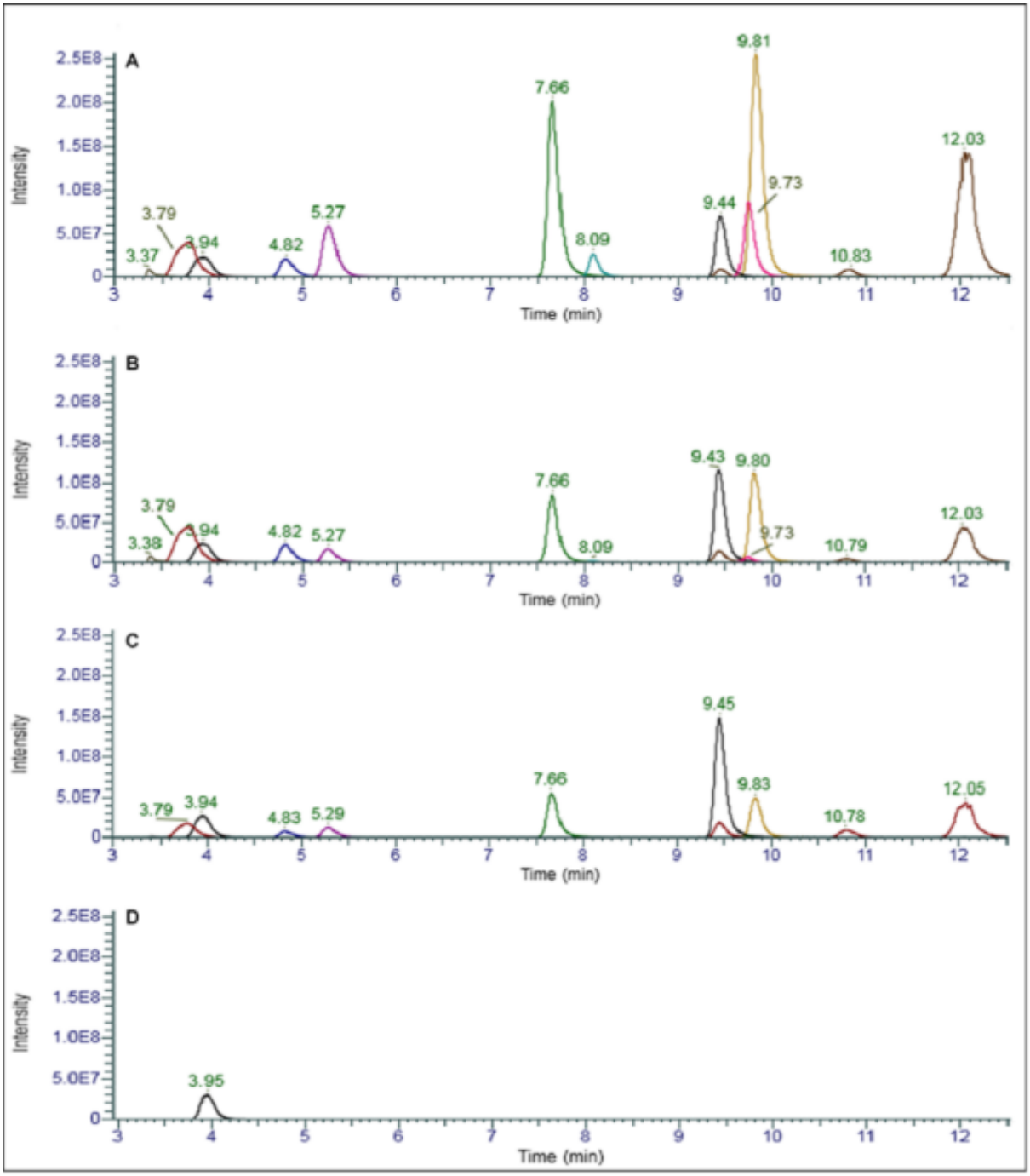


Figure 2-1. Example extracted ion chromatograms showing some of the major GLS identified in each species. (A) turnip, (B) rapeseed, (C) canola, (D) annual ryegrass control (spiked with sinigrin). GLS shown are glucoraphanin (grey, 3.38 min), progoitrin (red, 3.79 min), sinigrin (black, 3.94 min), gluconapoleiferin (blue, 4.83 min), gluconapin (purple, 5.27 min), glucobrassicanapin (green, 7.65 min), glucoerucin (light blue, 8.09 min), glucobrassicin (black, 9.44 min), glucoberteroin (pink, 9.73 min), gluconasturtiin (yellow, 9.80 min), neoglucobrassicin (brown, 12.03 min), 4-methoxyglucobrassicin (brown, 10.78 min), potential third neoglucobrassicin isomer (brown, 9.42 min).

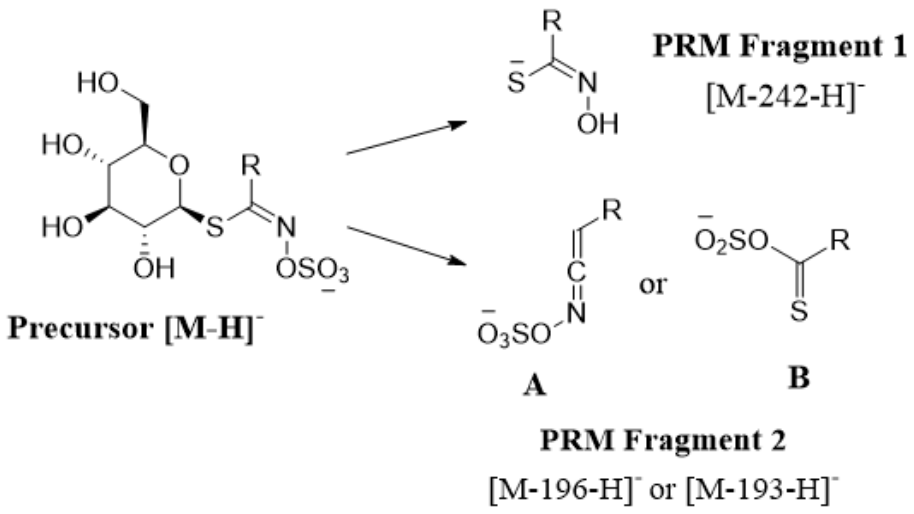


Figure 2-2. Structures of GLS fragments analyzed. The precursor *m/z* value is used for preliminary identification of GLS. PRM analysis of the precursor produces fragments 1 and 2, which are specific to individual GLS (differentiated by the R group). The fragmentation pattern that produces fragment 1 is ubiquitous to all GLS. The fragmentation patterns that produce fragment 2A and fragment 2B are not ubiquitous among all GLS. Therefore, the second PRM confirming fragment (2A or 2B) is dependent on the GLS side chain. Identification of GLSs is confirmed when PRM fragments 1 and 2 are present within 20% of the target ratio.

2.2.4. Assignment of minor GLS without authentic standards.

The GLS profile of individual brassica species typically contains 4 – 5 GLS in significant quantities (> 80% of the total GLS content), but it is common to observe several "minor" GLS in smaller amounts (< 1% of the total GLS content). As the bioactivity of ITCs varies significantly with the identity of the corresponding GLS side chain, it is possible that the GLS present in small quantities may still have a large impact on the overall bioactivity of that specific brassica plant. Here, sample chromatograms were analyzed by comparison to chromatograms of a GLS standard mixture. Peaks that were identified as GLS in a sample that did not match the authentic standards were tentatively assigned based on their exact mass, predicted retention behavior, and whether it has been previously identified in the specific brassica plant.

2.2.4.1. Tentative GLS assignments:

(i) The identification of 1-methylpropyl-GLS (glucocochlearin, retention time 6.30 min; *m/z* 374.05850) was presumed because it has been previously identified in turnips,[92] and the RT suggested a chain length in between 3-butenyl-GLS (gluconapin, RT 5.27 min) and 4-pentenyl-GLS (glucobrassicanapin, RT 7.66 min). The other possible isomers, isobutyl-GLS or n-butyl-GLS (the latter has not yet been satisfactorily characterized), have branched chain R-groups and would likely elute earlier if present.

(ii) The peak corresponding to *m/z* 388.07415 at RT 9.00 min was assigned to n-pentyl-GLS, which has a straight chain R-group. The possible isomers (2-methylbutyl-GLS and 3-methylbutyl-GLS) are also both branched chain and would likely elute earlier than 9.00 min.

(iii) There are four peaks for *m/z* 390.05341 (RT 3.50, 3.60, 4.15 and 4.42 min), and there are four possible GLS isomers: 3-hydroxybutyl-GLS, 4-hydroxybutyl-GLS, 1-(hydroxymethyl)propyl-GLS (glucosisaustricin) and 2-hydroxy-2-methylpropyl-GLS (glucoconringiin). These isomers are incredibly similar and lack authentic standards, therefore, individual GLS cannot be assigned to a single peak.

(iv) Identification of the *m/z* 438.05341 peaks (one large at RT 9.24 min and several small at RT 7.45, 7.62, 9.06, and 9.73 min, see Appendix A, Figure S2-2A) are similarly speculative as there are five possible isomers: (2(R)-hydroxy-2-phenylethyl-GLS (glucobarbarin), 3-methoxybenzyl-GLS (glucolimnanthin), 4-methoxybenzyl-GLS (glucoaubrietin), 4-hydroxyphenethyl-GLS (homosinalbin) and *m*-hydroxyphenethyl-GLS). Three of these side chains a have hydroxyl group, which based on the elution order of progoitrin (RT 3.79 min) compared to gluconapin (RT 5.27 min) and

glucosinalbin (RT 4.90 min) compared to glucotropaeolin (RT 8.20 min), reduces retention on the column. Therefore, the two earlier peaks at 7.45 and 7.62 min are most likely either glucobarbarin, homosinalbin or *m*-hydroxyphenethyl-GLS (order of elution is unknown). The largest peak at 9.24 min is most likely to be glucolimnanthin as it has been reported in oilseed meal previously.[119] If this is the case, then it is likely the peak at 9.06 min is one of the hydroxylated GLS isomers, and the peak at 9.73 min could be glucoaubrietin.

(v) 5-(methylsulfinyl)pentyl-GLS (glucoalyssin) was assigned to the peak at 3.87 min (*m/z* 450.05678), because the only known isomer (3-hydroxy-5-(methylthio)pentyl-GLS) is only partially characterized, and glucoalyssin has been reported previously in turnips [120].

(vi) The XIC of *m/z* 454.04833 show two peaks at RT 3.86 min and 5.28 min (Appendix A, Figure S2-2B). There are five possible GLS isomers with the corresponding exact-mass: (R)-2-hydroxy-2-(3-hydroxyphenyl)ethyl-GLS (*m*-hydroxyepiglucobarbarin), (S)-2-hydroxy-2-(4-hydroxyphenyl)ethyl-GLS (*p*-hydroxyglucobarbarin), (R)-2-hydroxy-2-(4-hydroxyphenyl)ethyl-GLS (*p*-hydroxyepiglucobarbarin), 4-hydroxy-3-methoxybenzyl-GLS (3-methoxysinalbin) and 3-hydroxy-4-methoxybenzyl-GLS (glucobretschneiderin). These are all indistinguishable by HPLC-MS, most likely even with authentic standards. However, since 3-methoxysinalbin and glucobretschneiderin have a methoxy group in addition to a hydroxyl group, they will most likely be retained longer on the column than the other set of isomers with only hydroxyl groups. Therefore, the peak at 3.90 min is most likely one of the hydroxy-glucobarbarin

isomers, and the peak at 5.30 min is most likely 3-methoxysinalbin or glucobretschneiderin.

(vii) The peak at 4.90 min for *m/z* 464.07243 was assigned to 6-(methylsulfinyl)hexyl- GLS (glucohespirin) instead of its isomer (3-hydroxy-6-(methylthio)hexyl-GLS). The isomer is not fully characterized, and the retention behavior is more similar to an increase in RT due to the longer chain length (i.e., compared to 4-(methylsulfinyl)butyl-GLS) than a decrease in RT due to a hydroxyl group (i.e., compared to 5-(methylthio)pentyl-GLS).

(viii) 1-methoxyindol-3-ylmethyl-GLS (neoglucobrassicin, *m/z* 477.06431) has one known isomer, 4-methoxyindol-3-ylmethyl-GLS (4-methoxyglucobrassicin). However, our data shows three peaks with the same exact-mass (RT 9.42 min, 10.78 min, and 12.03 min). A second isomer of neoglucobrassicin appears to be present in our samples that has not been previously described. We believe it is most likely a third orientation of the methoxy group. Based on previous reports that indicate that neoglucobrassicin is typically the most abundant of the known isomers, we assigned it to the peak at 12.03 min.[120,121] Since we do not know the nature of the second isomer, we cannot speculate the elution order of the other two peaks.

Interestingly, we may have identified a few GLS that have been "disproven" according to Blažević et al., 2020 HRAM-MS data shows a peak corresponding to 5-hexenyl-GLS (RT 9.6 min, *m/z* 400.07415), with a mass errors of 0.2749 ppm. Several other GLS with straight-chain aliphatic R-groups have been verified with authentic standards (sinigrin, RT 3.94 min; gluconapin, RT 5.27 min; glucobrassicanapin, RT 7.66 min). The retention pattern of these GLS shows that an increase to the R group chain length increases retention. 5-hexenyl-GLS

has a longer carbon chain than glucobrassicanapin, and we saw a corresponding increase in RT. This pattern, combined with the accuracy of the HRAM-MS data, is compelling enough to infer that 5-hexenyl-GLS does exist despite the lack of experimental evidence.

Our data also shows two peaks corresponding to *m/z* 402.0541, at RT 4.20 min and 4.81 min, with mass errors of 2.1639 ppm and 2.8106 ppm, respectively (Appendix A, Figure S2-2C). The larger peak at 4.81 min was assigned as (2S)-2-hydroxy-4-pentenyl-GLS (gluconapoleiferin), which has been identified previously in turnip[120] and rapeseed[122]. There are two anticipated isomers of gluconapoleiferin, 4-oxopentyl-GLS and (2R)-2-hydroxy-4-pentenyl-GLS. There has been no proof of either to exist, and there is only one report of 4-oxopentyl-GLS in any brassica.[123] Our results suggest the existence of an isomer to gluconapoleiferin, but we cannot accurately identify the peak at 4.20 min without further structural analysis.

The potential identification of these unknown GLS or GLS isomers highlights the extreme variability of the GLS side chain structure in nature and emphasizes that there is still much to learn about the glucosinolate-myrosinase defense system and its implication in agricultural applications. Although most of these identified GLS isomers are not present in significant amounts (< 1 % of the total GLS profile), their detection in canola, rapeseed, or turnip plants establishes it is likely that they exist in other brassica species at concentrations that would affect the corresponding bioactivity.

2.2.5. GLS content of forages used in this study.

Our results show that the GLS profile (identity of individual GLSs) and total GLS content did not depend on the date of herbage collection. The variation in total GLS content

was less than 15 % for rapeseed between all three collection dates, and the first two collection dates for turnip. The turnip plant saw an apparent 45 % increase of total GLS content between the first (7 October) and third (3 November) collection date. This significant increase in GLS production is presumably due to upregulation in GLS biosynthesis as the plant matures to strengthen its biochemical defense system. The variation in total GLS content for canola was less than 30 % between all three collection dates. This high variability is expected due to the minimal GLS content typical in canola plants.

Our results support previously reported GLS profiles in turnip[91,120], canola[124] and rapeseed[92,125] crops. We have also tentatively identified several GLS that have not been reported, including GLS that were recently discovered[98], summarized in Tables 2-1 and 2-2. The small differences observed in GLS concentrations reported in this study and other previous reports are most likely due to expected natural variation among brassica species, varieties and/or environmental factors such as temperature, light exposure, and humidity, all of which can impact GLS production.[126] Turnip had the greatest ($P < 0.0001$) total GLS concentration followed by rapeseed ($P \leq 0.0005$) and canola ($P < 0.0001$). Total GLS concentration for rapeseed was 6.4 mg g^{-1}, which on the lower end of the range reported in previous studies (3 – 20 mg g^{-1} total GLS).[127,128] Cartea and Velasco, 2008 described the primary GLS (>80% of the total) found in rapeseed leaves to be glucobrassicanapin (~50%), progoitrin (~25%), and gluconapin (~10%). Our results show that the primary GLS in rapeseed crops are progoitrin (30.3%), glucobrassicanapin (20.5%), glucobrassicin (10.7%), gluconapin (10.1%), and gluconasturtiin (9.5%).

The total GLS concentration of the turnips was 14 mg g^{-1}, which is consistent with other reports.[129,130] Padilla et al., 2007 reported the primary GLS in turnips were gluconapin

(~74%), glucobrassicanapin (~7%), and progoitrin (~4%). In the current study, the primary GLS found in turnip leaves were glucobrassicanapin (34.0%), progoitrin (15.3%), glucoberteroin (10.5%), gluconasturtiin (9.5%), gluconapin (8.8%), and neoglucobrassicin (8.0%).

Canola was designed to be a low-GLS variety of rapeseed, and showed 55% lower ($P < 0.0001$) GLS than rapeseed, 2.9 mg g^{-1}. Most reports of GLS in canola are focused on the seed content, but our results appear to be consistent with the few reports of GLS in canola tissue [131,132]. Potter et al., 1998 reported the primary GLS in canola are glucobrassicin (~46%), progoitrin (~18%), gluconasturtiin (~12%), and glucobrassicanapin (~9%). We found the primary GLS in canola tissue to be glucobrassicanapin (31.0%), progoitrin (20.5%), glucobrassicin (15.0%), gluconapin (11.4%), gluconasturtiin (7.7%), and neoglucobrassicin (7.6%).

We also identified all of the minor GLS previously reported in rapeseed [133,134] and most of the minor GLS previously reported in turnip[91,120,128] (exceptions include glucoiberverin and glucoiberin that were reported by Cartea and Velesco, 2008, but were not detected with the current method). The additional GLS tentatively identified in the current study are summarized in Table 2-2. A few GLS that have not yet been reported in rapeseed include glucoraphanin and glucoberteroin, and glucosinalbin, 1,4-dimethoxyindol-3-ylmethyl-GLS. Two additional neoglucobrassicin isomers were observed in turnip. We also found a number of minor GLS in canola in very small amounts (Table 2-2). As expected, there were no GLS observed in the non-brassica control (annual ryegrass). The identification of GLS reported herein without an authentic standard are tentatively inferred based on retention behavior and HRAM-MS data. The uncertainties in GLS identification can potentially be resolved using MS/MS techniques,

but the low concentrations of these minor GLS severely limits the sensitivity of MS/MS analysis and would likely not provide conclusive information.

Glucosinolate	***Response Factor*[g]**	***Turnip***	**Sinigrin Equiv[c]**	**Total Profile %**	***Rapeseed***	**Sinigrin Equiv[c]**	**Total Profile %**	***Canola***	**Sinigrin Equiv[c]**	**Total Profile %**
Glucoputranjivin[a]		+			+*			-		
Gluconapin[a,e,f]	0.981	+	0.918	6.44	+	0.4477	7.47	+	0.234	7.95
Glucocochlearin[a,e]		+			+			+		
Glucosisymbrin[a,e]		+			-			-		
Glucobrassicanapin[a,e,f]	1.05	+	4.46	31.3	+	1.22	19.1	+	0.800	27.2
Progoitrin[a,f]	0.989	+	1.55	10.9	+	1.38	21.6	+	0.407	13.9
n-pentyl[a,b,d,e]		+*			+*			-		
Glucoconringiin[a,e]		+*			-			-		
Glucosisaustricin[a,d,e]		+			-			-		
5-hexenyl[d]		+*			+*			+*		
Gluconapoleiferin[a,b]	0.989	+	0.63	4.49	+	0.299	4.67	+	0.084	2.86
4-oxopentyl[a,b,d]		+			+			+		
Glucocleomin[a,e]		+*			-			-		
Glucotropaeolin[b]		+			+*			+*		
Glucoerucin[f]	0.993	+	0.269	1.89	+			+		
Gluconasturtiin[b,f]	0.579	+	1.96	13.7	+	0.879	13.7	+	0.310	10.5
Glucosinalbin[a,e,f]	0.861	+*			+*			-		
Glucoberteroin[b]	0.993	+	1.52	10.6	+	0.117	1.82	+	0.033	1.14
Glucoraphanin[b,f]	0.971	+			+*	0.083	1.30	+*	0.029	0.981
Glucobarbarin[a,b,e]		+*			+			+*		
Glucolimnanthin[a,b,e]		+			+			-		
Glucomatronalin		+*			+*			+*		
Glucobrassicin[f]	0.703	+	1.03	7.23	+	0.973	15.2	+	0.591	20.1
Glucoalyssin[a,b,e]		+			+			+		
Glucoerysolin[b]		+			-			-		
(R)-2-hydroxy-2-(4-hydroxyphenyl)ethyl[a]		+*			+*			+*		

4-hydroxy-3-methoxybenzyl[a]		+*			+*			-		
4-hydroxyglucobrassicin[a]		+			+*			+*		
Glucorapassasin A[a,d]		+*			+*			+*		
Glucohespirin[a,e]		+*			+*			+*		
3-hydroxy-5-(methylsulfinyl)pentyl[a,b]		+*			+*			+		
Neoglucobrassicin[a]	0.703	+	1.60	11.2	+	0.775	12.1	+	0.298	10.1
4-Methoxyglucobrassicin[a]	0.703	+	0.169	1.19	+	0.075	1.18	+	0.083	2.82
?-methoxyindol-?-ylmethyl[a]	0.703	+			+	0.094	1.46	+	0.066	2.23
4-hydroxy-3,5-dimethoxybenzyl		+*			+*			+*		
2-benzoylprogoitrin[b]		+*			-			-		
3,4,5-trimethoxybenzyl		+*			+*			-		
1,4-dimethoxyindol-3-ylmethyl		+*			+*			+*		
Total GLS		**14 mg g^{-1}**			**6.4 mg g^{-1}**			**2.9 mg g^{-1}**		

Table 2-2. List of GLS identified in 'Appin' forage turnip, 'Barsica' forage rapeseed, and 'Inspiration' winter canola leaf/stem tissues. (+) represents GLS present, (-) represents not present. *denotes GLS not previously reported to our knowledge. [a]GLS that have 1 or more isomers; peak ID based on commonality with other reports or retention characteristics. [b]GLS that have 1 or more isobars; peak ID based on HRAM-MS data. [c]calculated for GLS that make up >1% total GLS content; expressed as sinigrin equivalents (mg GLS/ g dry weight). [d]GLS has a contested structure. [e]GLS that have 1 or more isomer with a contested structure. [f]Authentic standard used for peak ID. [g]Response factor calculated using available GLS standard, see Appendix A, Table S2-1 for details. Equiv (equivalents).

2.3. CONCLUSIONS

The methodology developed in this study simplifies sample preparation and GLS analysis, which will enable researchers to build a more robust GLS profile for forage brassica species that may include GLS that were not previously considered. Rapid and accurate identification of the major individual GLS in a specific species will allow for improved selection of brassicas to minimize animal health and productivity issues while providing a

highly nutritious forage source. We also demonstrate that this method is useful for building a preliminary, rapid GLS profile to determine the direction for more detailed GLS analysis (i.e., profiling the minor GLS). Due to the large array of R-substituents, each GLS has a different efficacy for biological activity, making the characterization of individual GLS in forage varieties important for future research regarding enteric methane mitigation, livestock health, and soil biofumigation.[105,135,136]

2.4. MATERIALS AND METHODS

2.4.1. Plant material.

'Appin' forage turnip, 'Barisca' forage rapeseed, 'Inspiration' winter canola, and a non-brassica control species 'KB Supreme' annual ryegrass (*Lolium multiflorum* Lam.) were planted into a prepared seedbed on 26 August, 2015 at The Pennsylvania State University Russell E. Larson Agronomy Farm in Rock Springs, PA [Hagerstown silt loam (fine, mixed, semiactive, mesic Typic Hapludalfs)]. The experiment was conducted as a randomized complete block design (n = 4). Forages were planted using a no-till drill (Plot Seed XL, Wintersteiger Inc., Salt Lake City, UT) at the recommended agronomic rate of 5.6 kg ha^{-1} for turnip, rapeseed, and canola, and 22 kg ha^{-1} for annual ryegrass into 5.5 × 9.1 m plots. Plots were fertilized with 72 kg N ha^{-1} at the time of seeding, and lime, potash, and phosphorus were applied according to soil test results. An additional 57 kg N ha^{-1} was applied on 18 April, 2016.

Herbage was collected during the fall of 2015 on 7 October, 21 October, and 3 November. Herbage (leaf and stem) was randomly collected from four locations within each plot and fresh herbage samples were placed in a freezer (-4 °C) until being lyophilized (Ultra 35 Super ES; Virtis Co. Inc., Gardiner, NY). The lyophilized herbage was ground to pass

through a 1-mm (18 mesh) sieve, and then further ground using a mortar and pestle until samples were a fine powder (200-400 μm mesh). The pulverized samples were stored at -4 °C prior to extraction (~1 week).

2.4.2. Reagents and materials.

Sinigrin (2-propenyl-GLS) hydrate standard with a purity of 99.0% was purchased from Sigma Aldrich (Milwaukee, United States). Ten additional GLS standards were purchased from Phytoplan (Diehm and Neuberger GmbH Heidelberg, Germany): glucobrassicin, glucobrassicanapin, glucoerucin, glucoiberin, gluconapin, gluconasturtiin, glucoraphanin, glucoraphenin, sinalbin, and progoitrin. Optima grade methanol and acetic acid (Fisher scientific, Hampton, NH) were used for separations and extractions. Polyethersulfone 0.45 μm syringe filters (25mm) and 1.8 mL glass vials (VWR International, Radnor, PA) were used for the sample preparation and analyses. Ultrapure deionized water (18.2 MΩ) was produced in-house via a RO-75, 8-stage water purification system adapted from APEC water (City of Industry, CA).

2.4.3. Standard solutions.

A sinigrin standard solution was prepared by dissolving 5 mg in 50 mL of ultrapure water to a concentration of 100 μg mL^{-1}. Stock solutions of the Phytoplan standards were prepared by dissolving 10 mg of each standard in 1.5 mL methanol. Combined calibration standards were made by diluting the stock solutions to 75, 50, 25, 10, 5, 1, 0.1, 0.01, and 0.001 μg mL^{-1}. Samples were spiked with 100 μL of 5 μg mL^{-1} of sinigrin standard to assess extraction efficiency and matrix interferences. Matrix blanks were used to determine instrument baseline, detection limits, quantification limits, and the presence of carry-over

between samples. Calibration linearity was assessed by fit of the calibration curve (forced through zero) from 100-0.001 µg mL^{-1}. The standard curve had a R^2 of 0.9968 or better over the calibration range.

2.4.4. Glucosinolate extraction.

Glucosinolates were extracted from 0.1 g of dry matter (DM) using a modified method by Doheny-Adams et al.[112] The herbage from each sample was weighed into a 16×100 mm borosilicate test tube (VWR International, Radnor, PA). Ten mL of 70% methanol and 0.5 ppm internal standard (sinigrin) were added to each tube, then vortexed for 30 seconds. The tubes were incubated undisturbed at room temperature for 30 min prior to an additional 30 min of agitation on an end-over-end shaker at 15 rpm. After shaking, each tube was centrifuged at 1600 ×g for 5 min, and an aliquot of the supernatant was removed using a 3 mL syringe (VWR International, Radnor, PA). The aliquot was passed through a polyethersulfone 0.45 µm syringe filter (VWR International, Radnor, PA) into 1.8 mL glass vials (VWR International, Radnor, PA). Each sample was then diluted 1:10 in ultrapure water before analysis by HPLC-MS.

2.4.5. HPLC-MS of extracted samples.

An ICS-5000+ chromatography system (Thermo-Fisher Scientific, Sunnyvale CA) was interfaced to a Q Exactive orbitrap MS (Thermo-Fisher Scientific, Bremen, Germany) using a heated electrospray injection (HESI) source operated in negative ion mode. The chromatography system was equipped with a four-channel proportioning pump, autosampler and temperature regulated column compartment (set at 30 ˚C). The analytical column was a sulfonamide-embedded C16 (Acclaim RSLC Polar Advantage; 2.1 × 150 mm; 2.2µm particle

size; ThermoFisher). The eluent system used was a mixture of (A) acetic acid in water (0.1% v/v), and (B) methanol at a flow rate of 0.2 mL min^{-1} using a gradient of: (i) -4 to 1.5 min (A and B 80:20, v/v); (ii) 1.5-5.25 min (B, 100% v/v); (iii) 5.25-12.5 min (B 100% v/v). The sample was injected at 0.0 min and the injection loop volume was 2 µl. The mass spectrometer was operated in parallel reaction monitoring (PRM) mode with a scheduled target list. The list contains the exact *m/z* values for the precursors and their corresponding retention times, and the exact *m/z* values of the expected fragments of interest (confirming fragments). The GLS included in the list are shown in Table 2-1 with their corresponding PRM confirming fragments. The instrument simultaneously performs a full scan (from 325-525 *m/z*; resolution, 70,000 full width at half maximum (FWHM)) and MS^2 analysis with a normalized collision energy (NCE) of 25 on the precursor *m/z* values in the target list at the indicated RT ± 90s (mass isolation of ± 1 *m/z* of the indicated confirming fragments). GLS identification was confirmed when these fragments were present within 20% relative intensity of the target ratio. The max inject time (IT) was 200 ms until the automatic gain control (AGC) target, 1×10^6 is reached. The ESI parameters were sheath gas, 30 psi; auxiliary gas, 15 psi; sweep gas, 2 psi; spray voltage, 3.5 kV; capillary temperature, 200 ˚C; S-lens radio frequency, 70; auxiliary gas heater, 300 ˚C. To minimize the matrix constituents from contaminating the source, a two-place divert valve was used to direct LC flow to waste from -4.0 to 0.0 min. During integration, a three-point Gaussian smoothing algorithm was applied to the data. The LC-MS system was controlled by Chromeleon 7.2 software (Thermo-Fisher Scientific, Sunnyvale CA), which collected, stored and analyzed the data.

2.4.6. Sinigrin equivalents.

A solution containing 1 ppm of GLS standards (sinigrin, glucobrassicin, glucobrassicanapin, glucoerucin, glucoiberin, gluconapin, gluconasturtiin, glucoraphanin, glucoraphenin, sinalbin, progoitrin) was analyzed under 8 different isocratic conditions (from 30 – 80% A) to evaluate changes in GLS ionization efficiency as the mobile phase composition changed. The response from sinigrin at each isocratic condition (calculated in ppm, using the calibration curve for the current gradient method) was compared to the corresponding value for sinigrin in our gradient method, producing a ratio (Factor S). The same comparison (response under isocratic conditions vs. gradient method) was completed for each additional GLS in the standard solution. Each ratio represents a GLS factor. By dividing these ratios, a response factor (RF) was calculated for each GLS in the standard solution: $\frac{\text{GLS Factor}}{\text{Factor S}}$. An example calculation for the glucoerucin RF with raw data is shown in Appendix A, Table S2-1. The RF adjusts the response of each GLS in a sample so that the final concentration for each GLS can be calculated using a single calibration curve for sinigrin. The RF values calculated for each GLS standard yields corrected concentrations using the current gradient method that are within 15% standard error or better from the known spiked concentrations, except gluconasturtiin, which is within 20% standard error. This semi-quantitative analysis provides the GLS profile of each sample, expressed as sinigrin equivalents. For GLS that we did not have standards for, the RF of the standard that best matched its chemical structure was used instead (similar ionization efficiencies were assumed based on retention time and structural similarities).

2.4.7. Statistical analysis.

Processed analytical data from Chromeleon was imported into PROC GLIMMIX of SAS (SAS Inst., Cary, NC) for the analysis of variance with sampling date treated as a repeated measure and treatment as the main effect. Differences among forage brassica varieties were determined using least squares means analysis with $\alpha = 0.05$.

2.5. ACKNOWLEDGMENTS

The authors thank J. Everhart (USDA-ARS), C. Kasmer (Pennsylvania State University) and A. Rigdon (Restek Corp.) for field and laboratory expertise and time contributed to conducting this experiment. This work was partially funded by USDA-NIFA-OREI.

Chapter 3

Lifetimes of the Aglycone Substrates of Specifier Proteins, the Autonomous Iron Enzymes that Dictate the Products of the Glucosinolate-Myrosinase Defense System in Brassica Plants

3.1. INTRODUCTION

Forage brassicas have been used in agriculture for decades, both as biofumigants to control pests[69] and as food for grazing livestock during winter months, when cool-season grasses are dormant.[77] Unfortunately, brassicas can also both directly harm grazers, causing goiter,[80] infertility,[78] decreased weight gain,[100] and other health problems, and adversely affect the flavor of their consumable products (milk, meat).[137] As a result, there is intense interest in understanding brassica physiology and (in relevant species) modulating production of the compounds responsible for the desired and deleterious properties of the plants. The antifeedant activity and toxicity of brassicas arise from a common set of related compounds, organoisothiocyanates (RNCSs), which are produced by the glucosinolate-myrosinase system, an innate defense mechanism of the plants (Figure 3-1). Glucosinolates (GLSs) are a large family of *S*-(β-D-glucopyranosyl)thiohydroximate-*O*-sulfonate compounds with different, characteristic, carbon-linked sidechains (R in Figure 3-1). More than 150 GLSs have been identified. Examples of side chains include allyl in sinigrin, 3-indolylmethyl in glucobrassicin, and 2-hydroxybut-3-en-yl in progoitrin. Catabolism of GLSs is triggered by injury to plant tissue (e.g., by mastication, pathogen colonization), which releases the compartmentalized β-thioglucosidase enzyme, myrosinase (MYR, EC 3.2.1.147; UniProt ID P29736).[50] MYR hydrolyzes the thioglucosidic bonds of GLSs – promiscuously with regard to their side chains – producing unstable thiohydroximate-*O*-sulfonate (aglycone) intermediates.

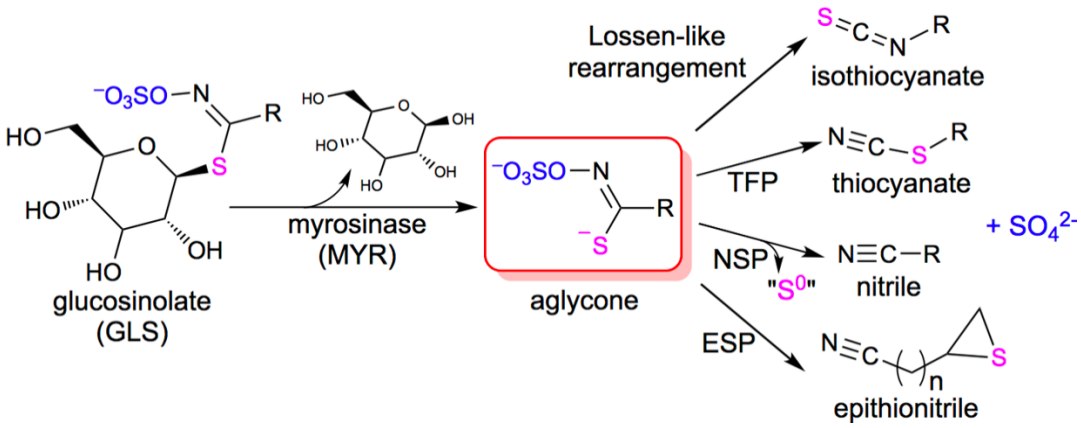


Figure 3-1. The Glucosinolate-Myrosinase Defense System of Brassica Plants and Activities of Specifier Proteins (ESP, NSP, TFP) Therein.

The diverse aglycones can readily undergo a Lossen-like[23] rearrangement to the corresponding RNCSs, which have the aforementioned adverse physiological effects on both generalist insects and livestock.[8,49] Many brassica species also contain specifier proteins (SPs), a group of structurally homologous proteins that can redirect GLS degradation to different, less toxic epithionitrile (Epithionitrile Specifier Protein, ESP), thiocyanate (Thiocyanate Forming Protein, TFP), and nitrile (Nitrile Specifier Protein, NSP) byproducts (Figure 3-1). SPs have a conserved β-propeller core architecture created from so-called "Kelch" sequence motifs. Mechanisms for their redirection of aglycone degradation have been advanced in the literature, but direct experimental interrogation has been limited.[38,41] Given the desire of agriculturalists to control the distribution of GLS breakdown products in food and forage brassicas, a deeper understanding of SP biochemistry could have considerable economic value.

Recent studies have suggested that SPs are autonomous, iron-dependent lyase or lyase-isomerase enzymes that act directly upon the unstable aglycone products of MYR to preempt their rearrangement to RNCSs.[38,41] To our knowledge, definitive validation of this hypothesis, and a firm basis to exclude the possibility that SPs instead interact directly with MYR to alter the course of its reaction, have not yet been reported. Indeed, many aspects of SP function

remain poorly understood, owing in large measure to the reportedly fleeting nature of the aglycone compounds, which are, according to the prevailing hypothesis, the SP substrates. These compounds have repeatedly been characterized in the literature as "extremely unstable,"[46–50] but, to the best of our knowledge, no quantitative assessment of their intrinsic lifetimes has been reported. Accordingly, experimental analyses of SP reactions have all been carried out by including MYR to produce the unstable aglycone *in situ*. This experimental configuration does not allow a distinction to be made between an impact of the SP directly on the outcome of the MYR reaction and the presumed sequential action of the two proteins separately on the small molecules.[37,51]

This study was motivated by the goals of (i) definitively resolving the general mode of action of the SPs and (ii) acquiring information to enable the dissection of their chemical mechanisms by the suite of rapid-kinetic (stopped-flow, freeze-quench) and spectroscopic (Mössbauer, EPR, x-ray absorption) methods that we have applied to other iron enzymes.[138–140] Toward these objectives, we sought kinetic parameters to inform design of experiments in which an unstable aglycone intermediate would be rapidly accumulated in the absence of an SP (by complete MYR-catalyzed hydrolysis of a GLS) before being mixed with an SP. In this configuration, an SP acting autonomously as an enzyme, with the unstable intermediate as its substrate, would intercept the aglycone and redirect its breakdown to a product other than the RNCS. By contrast, an SP functioning as an effector of MYR would be ineffectual when added after completion of the MYR reaction. Moreover, for the case of an SP acting autonomously, pre-accumulation of the aglycone would permit initiation of the SP reaction without the lag phase operant in a two-step (coupled) sequence. This experimental configuration would thereby favor accumulation of intermediates in the SP reaction cycle, potentially enabling their

direct spectroscopic characterization, as we have achieved for other iron-dependent enzymes.[138–140] Successful design of such a sequential-mixing experiment would require knowledge of both the catalytic parameters of MYR with a given GLS, so that the time for its complete hydrolysis might be known, and the half-life of the resultant aglycone, so that a concentration of MYR sufficient to make GLS hydrolysis much faster than decay of the aglycone to the RNCS might be selected. To determine aglycone half-lives, we developed a generally applicable colorimetric assay to monitor disappearance of the nucleophilic thiohydroximate moiety, as described below.

Chromogenic reporters such as 5,5'-dithiobis-(2-nitrobenzoic acid) (Ellman's reagent) have been used extensively in quantification of free, reduced cysteines in proteins and other biological thiols. Attack of a thiol(ate) upon the Ellman's reagent disulfide produces a mixed-disulfide and 2-nitrobenzoate-5-thiophenolate ion, which absorbs visible light (λ_{max} = 412 nm, ε_{412} = 14 $mM^{-1}cm^{-1}$). We postulated (i) that the thioiminate moiety common to all aglycones would be similarly nucleophilic toward such reagents and (ii) that the RNCS resulting from the Lossen-like aglycone rearrangement would lack this reactivity, thus affording a reporter of the progress of the spontaneous rearrangement. For technical reasons, we selected 2,2'-dithiodipyridine (2-PDS) as the chromogenic sulfur-reactive agent; attack upon its disulfide produces 2-thiopyridone (2-TP; Figure 3-2), which absorbs light maximally at 342 nm (ε_{342} = 7.9 $mM^{-1}cm^{-1}$; Appendix B, Figure S3-1).

Here, we illustrate the use of this reagent to define the timescales of the Lossen-like rearrangements of the aglycone intermediates derived from nine different GLSs. Knowledge of the lifetime of the sinigrin-derived allyl-aglycone and the kinetic parameters of *Sinapis alba* (*Sa*) MYR revealed that the enzyme can, under appropriate conditions, fully consume the GLS

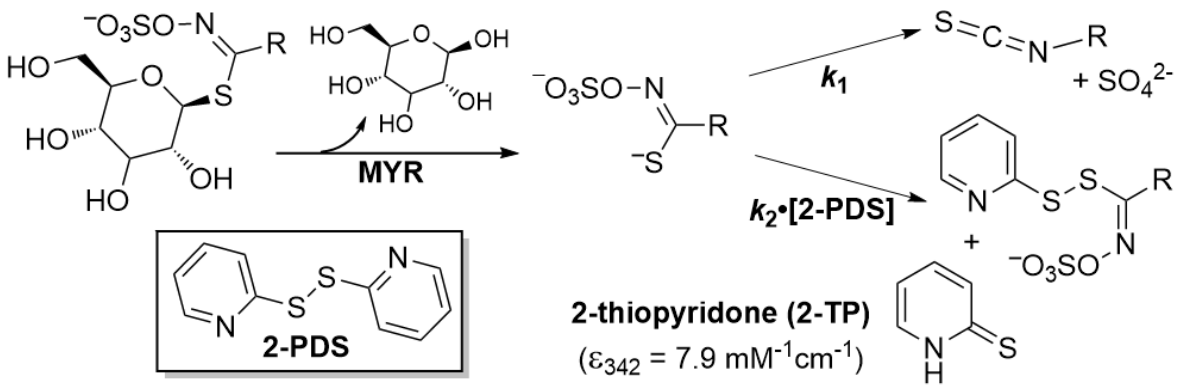


Figure 3-2. Use of 2-Thiopyridine Disulfide (2-PDS) to Trap and Detect the Unstable Aglycone Intermediates: Framework to Interpret the Partition Between the Unimolecular Lossen-like Rearrangement (k_1) and Biomolecular Trapping (k_2) Pathways.

rapidly compared to rearrangement of its aglycone. This result enabled demonstration that sinigrin hydrolysis and Fe(II)-activated interception of the aglycone by recombinant *Arabidopsis thaliana* (*At*) epithionitrile specifier protein (ESP) (UniProt ID A0A1P8ARA7) can be separated in time and space, thus firmly establishing that this SP functions independently of MYR, by autonomous action on the aglycone. The kinetic parameters were further leveraged to design competition assays to demonstrate that *At* ESP operates catalytically with an efficiency parameter (k_{cat}/K_M) of 10^4-10^5 $M^{-1}s^{-1}$. These studies provide the quantitative underpinnings for a direct analysis of SP reactions in the transient state.

3.2. RESULTS

3.2.1. Preparation of Affinity-tagged At ESP and Sa Seed Myrosinase (MYR). According to the prevailing hypothesis that SPs act as autonomous enzymes upon the aglycones produced by MYR rather than as binding partners of MYR that affect its reaction outcome, heterologous MYR-SP pairs would be expected to function just as well as homologous pairs from a common source. Accordingly, we opted to obtain (i) the MYR glycoprotein from inexpensive, commercially available white mustard (*Sinapis alba*, *Sa*) seeds by extractions followed by

carbohydrate-affinity chromatography on Concanavalin A Sepharose 4B resin and (ii) an N-terminally SUMO-His_6-tagged form of *At* ESP by purification on Ni(II)-nitrilotriacetate affinity resin after over-expression in *Escherichia coli*, thus obviating any growth of the brassica plants themselves. By these procedures, we obtained typical yields of 60 mg MYR from 300 g of seeds and 80 mg *At* ESP from 45 g of the frozen *E. coli* cell paste (harvested from 12 L of culture). In SDS-PAGE analysis with Coomassie Brilliant Blue staining, the preparations of MYR were characterized by a pair of prominent bands with smeared intensity between them (Appendix B, Figure S3-2A), most likely reflecting heterogeneous glycosylation.[12,141] Preparations of *At* ESP were estimated to be ~ 90% pure by this analysis (Appendix B, Figure S3-2B), with one of the major contaminants having apparent molecular masses suggestive of oligomers of the predominant *At* ESP monomer and others likely representing endogenous *E. coli* proteins. Iron is reportedly an essential cofactor of SPs,[38,41,142] and so we measured the Fe:ESP stoichiometry of our preparations by inductively coupled plasma mass spectrometry (ICP-MS). We found our preparations to contain detectable but sub-stoichiometric levels of its putative cofactor (0.15 ± 0.05 equiv), implying that they should require additional Fe(II) to be fully activated.

3.2.2. Steady-state Kinetic Characterization of MYR Preparations. Intending to focus our initial interrogation of SP mechanisms on ESP, we first characterized the activity of our preparations of MYR on sinigrin, a GLS with an allyl side chain capable of undergoing ESP-directed epithiolation. Sinigrin absorbs ultraviolet light with a peak (λ_{max}) at 227 nm (ε_{227} = 6.5 $mM^{-1}cm^{-1}$). We used this feature to monitor progress of MYR-catalyzed sinigrin hydrolysis. The high absorptivity made it challenging to access high substrate concentrations ($\geq 5 \bullet K_M$) in these assays, but our intended applications required knowledge primarily of the

value of k_{cat}/K_M, which we were able to determine with adequate precision by this method (Appendix B, Figure S3-3). The parameters determined in three replicate experiments – k_{cat} = 8.2 (± 0.4) s^{-1}, K_M = 1.1 (± 0.2) mM, and k_{cat}/K_M = 7.3 (± 0.3) $mM^{-1}s^{-1}$ – agree well with values previously reported for *S. alba* MYR.[143,144] The value of k_{cat}/K_M implies (and kinetic monitoring confirmed; see below) that, at a readily accessible enzyme concentration of 10 µM, our preparations of MYR are capable of consuming ~ 90% of sub-saturating (< 0.2•K_M) sinigrin within 30 s (Appendix B, Figure S3-7). Increased [MYR] would, of course, allow for faster hydrolysis, important for the case that an aglycone would undergo the Lossen-like rearrangement on a shorter timescale.

3.2.3. Use of 2,2'-Dipyridyl Disulfide (2-PDS) to Trap the Sinigrin (Allyl) Aglycone. We next assessed whether a thiol-reactive colorimetric indicator, 2,2'-dithiodipyridine (2-PDS), could rapidly trap and stabilize the thioiminate moiety of the sinigrin aglycone, distinguishing it from both the GLS substrate and the allyl-NCS final product of its putatively rapid Lossen-like rearrangement. In a reaction solution containing 0.01 mM MYR, 0.10 mM sinigrin, and 1.5 mM 2-PDS, the absorption spectrum characteristic of 2-thiopyridone (2-TP), the product of nucleophilic attack upon the 2-PDS disulfide, developed on the timescale (~ 30 s) expected for MYR-catalyzed hydrolysis of sinigrin (Appendix B, Figure S3-4, *black, pink and cyan*) and was stable for minutes (but not indefinitely) thereafter. The final ΔA_{342nm} for the reaction indicated production of 0.090 ± 0.003 mM 2-TP (Figure 3-3, *red*). The spectrum of 2-TP failed to develop in the absence of MYR (Appendix B, Figure S3-4, *blue*) or sinigrin (*red*) or upon incubation of 2-PDS with a standard of the allyl-NCS product from the Lossen-like rearrangement of the sinigrin aglycone (*purple*). Similarly, addition of 2-PDS only after a prior incubation of MYR and sinigrin that was sufficiently long to allow both hydrolysis of the GLS

and decay of the aglycone to reach completion also resulted in negligible 2-TP production (*see below*).

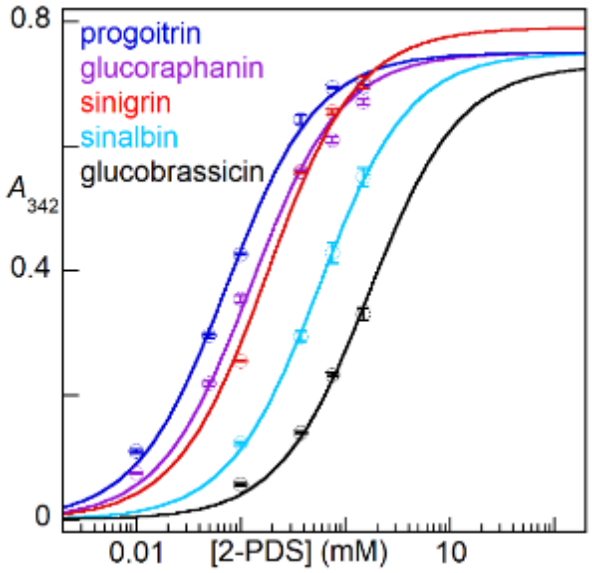


Figure 3-3. Efficiencies of trapping of the MYR-generated, unstable aglycone intermediates from five GLSs as functions of the concentration of the 2-thiopyridine disulfide (2-PDS) trap. The bimolecular trapping reaction produces 2-thiopyridone (2-TP), which gives the detected absorbance at 342 nm (ε_{342} = 7.9 $mM^{-1}cm^{-1}$). It competes with the spontaneous unimolecular Lossen-like rearrangement, which does not change A_{342}. In each case, the GLS was present at a concentration of ~ 0.10 mM, giving a maximum absorbance of ~ 0.79 for the case of 100 % trapping, the asymptotic limit in the fits of Eq. 1 to the data (solid lines). The number of replicates for each set of data is given in Table 3-1. Details are provided in the *Experimental Procedures*, and the structures of the GLSs are given in Table 3-1.

These observations confirm that 2-PDS reacts with the sulfur of the aglycone generated by MYR with sufficient selectivity (over the GLS substrate and allyl-NCS final product) and efficiency to support accurate quantification of the unstable intermediate. The 10% shortfall in the yield of 2-TP from the theoretical value of 0.10 mM (the concentration of the limiting reactant, sinigrin) suggested that the Lossen-like rearrangement competes to some extent at this concentration of 2-PDS (Figure 3-2), an inference that we verified by explicit variation of [2-PDS] in the sinigrin/MYR reaction (Figure 3-3, *red*). The yield of 2-TP, as reported by the final value of A_{342}, decreases in the expected way with decreasing [2-PDS]. Fitting the equation

appropriate for Figure 3-2 (Eq. 1) to the plot of A_{342}-versus-[2-PDS] allowed a value of k_1/k_2 to be determined (0.18 mM). A concentration of 2-PDS sufficiently high to fully suppress competition from the Lossen-like rearrangement (> 20 k_1/k_2 or ~ 4 mM) was inaccessible due to its modest solubility. More soluble disulfide-based thiol-trapping compounds (e.g., Ellman's reagent) react too slowly to be useful in these experiments.

3.2.4. Use of the Trap to Determine the Lifetime of the Sinigrin (Allyl) Aglycone. We next carried out sinigrin hydrolysis and aglycone trapping in separate steps, in order to resolve the kinetics of the trapping reaction and (more importantly) the Lossen-like rearrangement of the allyl aglycone to the corresponding allyl-NCS. The above coupled reaction, in which the 2-PDS trap was already present when the sinigrin was added to MYR, served to define the time required for complete hydrolysis. In the experiment, we allowed hydrolysis to proceed for this same time in the absence of PDS, and we then varied the time *after completion of hydrolysis* before addition of the 2-PDS trap (Δt). Upon addition of 2-PDS, the observed rate constant for development of the 342-nm absorption feature of 2-TP (k_{obs}) should be the sum of the effective first-order rate constants for the competing trapping and rearrangement reactions (Figure 3-2; Eq. 2), now resolved from MYR-catalyzed sinigrin hydrolysis. Fitting the equation for an exponential growth (Eq. 3) to the A_{342}-versus-time traces from four independent trials at short Δt values (e.g., Appendix B, Figure S3-5) gave $k_{obs} = 0.216 \pm 0.004$ s^{-1}. The ΔA_{342} for each time point (each different value of Δt) after complete trapping served to report the progress of the spontaneous Lossen-like rearrangement of the aglycone to the allyl-NCS during that time. The plot of A_{342}-versus-Δt (Figure 2, *red*) maps a first-order (exponential) decay process (Eq. 4) with a half-life ($t_{1/2}$) of 34 s, corresponding to a rate constant of 0.021 s^{-1} (*red solid line*). This value corresponds directly to k_1 in Figure 3-2. From these two measured values, k_2 was

estimated as 0.13 $mM^{-1}s^{-1}$. The resultant ratio, k_1/k_2, of 0.16 mM is similar to the value (0.18 mM) determined by analysis of the dependence of the trapping efficiency on [2-PDS] (Figure 3-3, *red*).

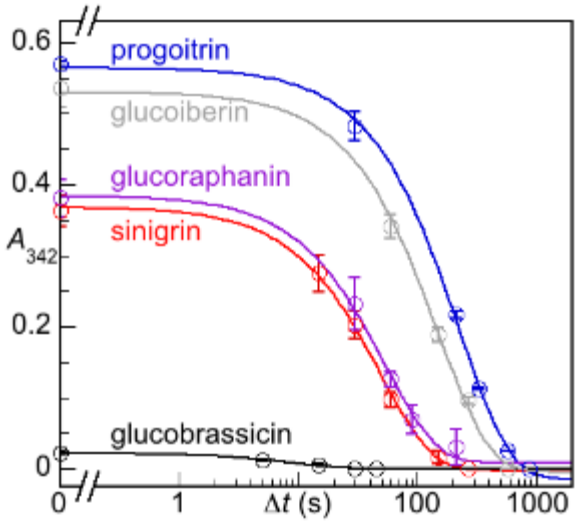


Figure 3-4. Direct monitoring of decay of the unstable aglycone intermediates from five GLSs by variation of the delay time (Δt) between completion of the MYR-catalyzed hydrolysis reaction and addition of the 2-PDS trap. Details are provided in *Experimental Procedures*.

3.2.5. Use of 2-PDS Trap to Determine Lifetimes of Other GLSs. We performed this same set of measurements – (1) direct monitoring of 2-TP formation in a coupled reaction containing 2-PDS, MYR, and a GLS to define the timescale of hydrolysis, (2) variation of [2-PDS] to determine k_1/k_2 (Figure 3-2), and (3) variation of Δt between completion of hydrolysis and addition of 2-PDS to resolve k_1 ($t_{1/2}$) for the Lossen-like rearrangement of the aglycone – for a series of GLSs. The data are shown in Figures 3-3, 3-4, and S3-6. Parameters are compiled in Table 3-1. Consistent with literature reports, MYR is active against all GLSs tested. The time needed for complete hydrolysis at a given [MYR] (10 μM) varied by less than a factor of four (Appendix B, Figure S3-7). By contrast, the values of $t_{1/2}$ for the aglycones varied more markedly (by more than 50-fold for all GLSs tested), ranging from much less than, to much

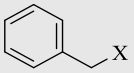
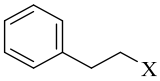
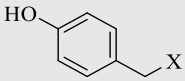
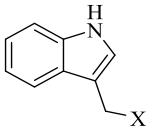
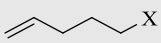
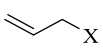
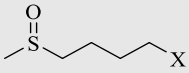
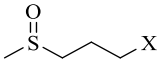
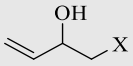
GLS	Aglycone[a]	k_1 (s^{-1})[b]	$t_{1/2}$ (s)	k_2 ($mM^{-1}s^{-1}$)[c]	k_1/k_2 (mM)[d,f]
Glucotropaeolin		0.35 ± 0.03[g]	2	n.d.	n.d.
Gluconasturtiin		0.33 ± 0.04[f]	2	n.d.	n.d.
Sinalbin		0.21 ± 0.04[g]	3	n.d.	0.72 ± 0.005
Glucobrassicin		0.11 ± 0.05[f]	7	n.d.	1.9 ± 0.04
Glucobrassicanapin		0.088 ± 0.003[f]	8	0.30[e]	n.d.
Sinigrin		0.021 ± 0.0007[i]	33	0.13 ± 0.01[h]	0.18 ± 0.0009
Glucoraphanin		0.019 ± 0.006[g]	37	0.23[e]	0.18 ± 0.0002
Glucoiberin		0.0068 ± 0.002[f]	100	0.17[e]	n.d.
Progoitrin		0.0045 ± 0.0002[g]	150	0.16[e]	0.11 ± 0.015

Table 3-1. Summary of Kinetic Parameters in Partitioning of the Aglycone Intermediates Between Trapping by 2-PDS and the Lossen-like Rearrangement. [a]X represents the common thiohydroximate-*O*-sulfonate aglycone core; [b]By variation of Δ*t*; [c]By mixing pre-formed aglycone with 2-PDS; [d]In coupled reactions by variation of 2-PDS; [e]1 trial; [f]2 trials; [g]3 trials; [h]4 trials; [i]5 trials.

greater than, the 34-s $t_{1/2}$ determined for the allyl aglycone from sinigrin (Figures 3-4 and Appendix B, S3-6; Table 3-1). For example, we measured half-lives of ~ 7 s and ~ 150 s for the 3-indolylmethyl aglycone from glucobrassicin and the 2-hydroxybut-3-en-yl aglycone from progoitrin, respectively. The larger values of k_1 were challenging to measure and are less precise, especially for the case of glucobrassicin, which is a poor substrate for the *Sa*

MYR (thus requiring longer time for complete hydrolysis and diminishing the value of A_{342} at $\Delta t = 0$) and produces a relatively labile aglycone (Figures S3-6 and S3-7). In some of these cases, agreement between k_1/k_2 ratios determined by variation of [2-PDS] and the individual values of k_1 and k_2 determined kinetically is not as good as for sinigrin, but, qualitatively, aglycones determined to have short half-lives (large values of k_1) are associated with large k_1/k_2 ratios. The data imply that the aglycones react with 2-PDS with similar rate constants (k_2 in Table 3-1), consistent with the expectation that the nucleophilicity of the sulfur in the common thiohydroximate-*O*-sulfonate moiety should be relatively insensitive to the remote substitutions defining the different aglycone side chains.

3.2.6. Verification of Iron-Dependent Activity of Recombinant At ESP. We assessed the capacity of our recombinant *At* ESP to redirect decay of the unstable sinigrin allyl aglycone from the allyl-NCS to the epithionitrile product (i.e., its activity). Detection of the two products was accomplished by gas chromatography coupled to mass spectrometry (GC-MS). Inclusion of ≥ 10 μM *At* ESP along with 2 molar equivalents Fe(II) in anoxic, MYR-catalyzed, sinigrin-hydrolysis reactions abolished the peak for the allyl-NCS product of the Lossen-like rearrangement and gave rise to an intense new peak for the epithionitrile product (Appendix B, Figure S3-8B), an assignment that was corroborated by observation of the expected fragment ions.[145,146]

We verified the dependence of this ESP activity on iron by allowing 2 μM MYR to hydrolyze 0.50 mM sinigrin in an anoxic solution containing 10 μM ESP and varying concentrations of added Fe(II). As isolated, our recombinant *At* ESP had 0.2 equiv iron already bound (in this preparation) and, therefore, had detectable activity even without addition of Fe(II). However, GC-MS analysis of products revealed both allyl-NCS and epithionitrile in

reactions under these conditions (in the absence of added iron), whereas inclusion of as little as 1.0 additional equivalents Fe(II) fully suppressed the peak of the allyl-NCS and gave maximal intensity for the peak of the epithionitrile (Appendix B, Figure S3-8B, S3-9). The observed activation of ESP by stoichiometric Fe(II) is consistent with the published conclusion that SPs use iron as their essential cofactor.[41,142]

3.2.7. Experimental Verification of the Autonomous Activity of At ESP. The kinetic and analytical results summarized above positioned us to definitively test the prevailing hypothesis that SPs are iron-dependent enzymes acting autonomously on alkyl thiohydroximate-*O*-sulfonate (aglycone) substrates rather than effectors of MYR. The insight afforded by the MYR assays, monitoring either consumption of sinigrin or (in the presence of 2-PDS) formation of 2-TP, allowed us to time the pre-incubation of sinigrin (250 µM) and MYR (25 µM) to allow the hydrolysis but not the Lossen-like rearrangement to proceed to completion (30 s) before adding 100 µM ESP together with 200 µM Fe(II). GC-MS chromatograms of control samples from reactions lacking any ESP or already containing ESP [10 µM with 20 µM Fe(II)] at the initiation of the MYR reaction showed the expected formation of either solely allyl-NCS or solely epithionitrile, respectively (Appendix B, Figure S3-8A, B). The chromatogram for the reaction to which ESP was added only after sinigrin hydrolysis had reached completion exhibited peaks for both compounds with areas of ~ 50 % compared to those of the corresponding peaks in the control samples (Appendix B, Figure S3-8C). Thus, approximately half of the sinigrin was processed to epithionitrile despite complete production of the aglycone in the absence of ESP (Figure 3-5). This observation formally rules out the possibility that the *At* ESP acts as an effector of MYR to change the outcome of its reaction. The observed production of ~ 50 % allyl-NCS in the experimental sample reflects the fact that the incubation

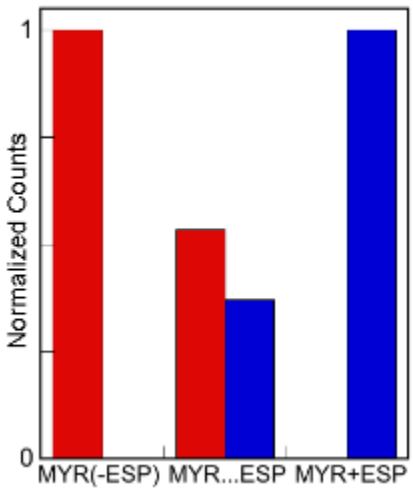


Figure 3-5. Plot of the intensities of GC-MS peaks for the allyl-NCS (*red*) and epithionitrile (*blue*) products from hydrolysis of sinigrin (0.25 mM) by MYR (25 µM) in the absence of *At* ESP (*left*), in the presence of 10 µM ESP and 20 µM Fe(II) (*right*), or with 0.10 mM ESP and 0.20 mM Fe(II) added after a 30-s incubation (*middle*). Details of the reactions and GC-MS analysis are provided in *Experimental Procedures*, and chromatograms are provided as Appendix B, Figure S3-8.

time for MYR-mediated hydrolysis (30 s) was comparable to the intrinsic half-life for decay of the allyl aglycone by the uncatalyzed Lossen-like rearrangement (34 s); presumably, a higher yield of the epithionitrile could be obtained by increasing [MYR] and decreasing the time for the hydrolysis step.

Because MYR remained present with ESP in the experiment of Figure 3-5, the results do not rule out the possibility that the ESP must engage with MYR to catalyze epithionitrile formation. In other words, rather than ESP serving as an effector of MYR, it would remain formally possible that MYR could act as an activator of ESP. To rule out this possibility, we immobilized MYR on the concanavalin A resin, exposed sinigrin to the MYR-adsorbed resin to hydrolyze it, and collected the resultant aglycone into a reservoir containing the Fe(II)•ESP. Even with this spatial separation of MYR and ESP, the epithionitrile was the predominant product (Appendix B, Figure S3-10A). By contrast, only the allyl-NCS product was detected

in the control sample prepared identically but with only buffer in the collection reservoir (Appendix B, Figure S3-10B). The results unequivocally prove that *At* ESP acts autonomously to catalyze epithionitrile formation.

3.2.8. Estimate of k_{cat}/K_M for Recombinant At ESP. Variation of the concentration of Fe(II)•ESP that was present with [MYR] (0.5 μM) at initiation of the hydrolysis reaction confirmed the expected increase in allyl-NCS yield and decrease in epithionitrile yield with decreasing [ESP] (Figure 3-6A). If it is assumed that the concentration of aglycone to accumulate in this MYR-ESP coupled reaction does not approach the (unknown) K_M of the *At* ESP for its substrate, then this dependence can be understood to result from competition between (1) the Lossen-like rearrangement, with instantaneous rate (flux) equal to the product of its first-order rate constant (k_1 = 0.021 s^{-1}) and the unknown concentration of aglycone at any time, and (2) capture by ESP, with rate equal to the product of its concentration, the effective second-order rate constant for its productive combination with the aglycone (k_{cat}/K_M), and the same unknown [aglycone] (Figure 3-7).

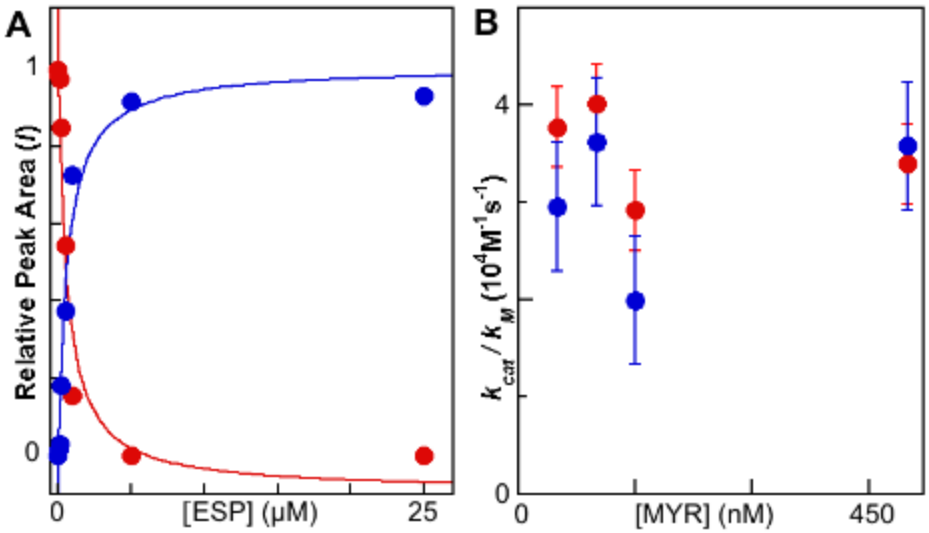


Figure 3-6. Estimates of k_{cat}/K_M (22 °C) of *At* ESP for its allyl aglycone substrate. (***A***) Plot of the relative areas of the GC-MS peaks for the allyl-NCS (*red*) and epithionitrile (*blue*) products from hydrolysis of sinigrin by 0.50 μM MYR versus [ESP] [with 2 molar equiv Fe(II)]. The solid lines are fits of Eqs. 5 (*blue*) and 6 (*red*) to the data. (***B***) Values of k_{cat}/K_M estimated from the best fit parameters obtained in panel *A* and additional, otherwise identical experiments with varying [MYR].

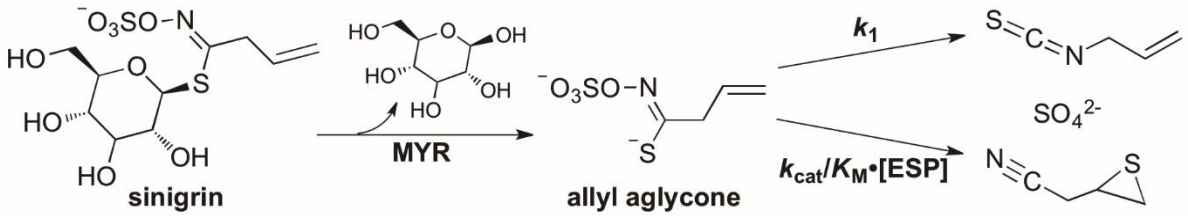


Figure 3-7. Competition Between the Lossen-like Rearrangement and Epithiolation by ESP Analyzed to Estimate k_{cat}/K_M of *At* ESP for the Allyl Aglycone of Sinigrin.

Thus, at the [ESP] yielding, for example, equal quantities of the two products, k_{cat}/K_M•[ESP]•[aglycone] = k_1•[aglycone], giving $k_{cat}/K_M = k_1$/[ESP]. Fitting plots of the peak integrals for the epithionitrile and allyl-NCS products versus [Fe(II)•ESP] (Figure 3-6A) to the appropriate analytical expressions (Eqs. 5 and 6) yielded an estimate of 3.5 x 10^4 $M^{-1}s^{-1}$ for the k_{cat}/K_M of the recombinant *At* ESP. Use of lesser MYR concentrations – which should disfavor accumulation of the aglycone and thereby diminish any error introduced by the assumption that its concentration remains much less than the K_M of the *At* ESP – yielded estimates of k_{cat}/K_M in the range of 2-4 x 10^4 $M^{-1}s^{-1}$, with no discernable correlation between the [MYR] and the estimated value of k_{cat}/K_M (Figure 3-6B). This analysis suggests that the estimates obtained are in the correct order of magnitude, i.e., that $k_{cat}/K_M = 10^4$-10^5 $M^{-1}s^{-1}$.

3.3. DISCUSSION

3.3.1. Lifetimes of the Aglycones from a Series of GLSs. The agricultural importance of the MYR-GLS defense system of crop and forage brassicas has long been appreciated,[147–149] and there is growing interest in understanding the effects of dietary GLSs, their breakdown products, and their interactions with intestinal microbiota in humans.[150] Because SPs can, in

some brassicas, modulate the distribution of breakdown products of this system, a mechanistic understanding of their function might also have value. One frequently cited barrier to biochemical and biophysical interrogation of SPs and their reactions is the instability of the MYR-generated aglycone intermediates, which, according to the prevailing hypothesis, are the SP substrates.[46–50] However, as we embarked upon an experimental mechanistic analysis of SPs, we could not, to our surprise, find any report of the lifetime of a GLS-derived aglycone, making it unclear how high a barrier the oft-quoted instability might actually present.

We reasoned that the alkyl thiohydroximate-*O*-sulfonate (aglycone) species might act as thiolate-like nucleophiles toward well-known, disulfide-based, colorimetric thiol indicators (e.g., Ellman's reagent), whereas the alkyl isothiocyanate (RNCS) products of their putatively rapid Lossen rearrangements would not be similarly nucleophilic, and that this difference might provide a means to define the kinetics of the rearrangement reactions. After verifying this expectation for one such indicator, 2-PDS, which had the best balance of rapid reactivity and solubility, we used it to determine lifetimes of the aglycones derived from nine GLSs. The $t_{1/2}$ values vary by almost two orders of magnitude from < 3 s to 150 s (22 °C). Curiously, there is no obvious correlation between the steric or electronic properties of the migrating carbon center and aglycone stability. The Lossen rearrangement is thought to proceed via a nitrene-like transition state formed by incipient departure of the activated leaving group, in this case sulfate (Scheme S1). The nitrogen to which the alkyl group migrates should become electron deficient (while remaining formally neutral), potentially favoring more electron-rich migrating alkyl groups. However, the anticipated concerted nature of the rearrangement renders such analysis based on local polarity overly simplistic, and the data confirm the need for a more

sophisticated analysis that accounts for the likely pericyclic (rather than dipolar) or free-radical nature of the reaction.

3.3.2. Demonstration of Autonomous, Iron-Dependent Enzymatic Activity of At ESP. The varying lifetimes of the aglycones span the time required for their enzymatic production at experimentally accessible MYR and GLS concentrations, implying that the more stable species are amenable to sequential-mixing experiments, in which a useful concentration of the aglycone would first be accumulated prior to initiation of the SP reaction by a second mix. Indeed, by this protocol, we showed for the case of *At* ESP and sinigrin that interaction of the MYR and ESP are not required for epithionitrile formation, thus removing any remaining doubt that SPs are autonomous enzymes that catalyze diverse transformations upon the array of known aglycones. The specificity parameter, k_{cat}/K_M, of ~ 10^4-10^5 $M^{-1}s^{-1}$ estimated for our recombinant *At* ESP is quite respectable, albeit not close to the limit of diffusion approached by "perfect" enzymes. Although it will be technically challenging, it should be possible to determine specificity parameters for other SPs and, in favorable cases, perhaps to parse values of k_{cat}/K_M into the contributions from each parameter.

3.3.3. Value of Kinetic Information for Resolution of SP Mechanisms. The kinetic parameters obtained in this study will inform more advanced experiments to interrogate SP reactions in the single-turnover regime and thereby to elucidate their mechanisms. Several fascinating questions raised by published mechanistic hypotheses can now potentially be addressed. For the case of epithionitrile formation, for example, it has been suggested that the thiolate anion of the aglycone coordinates to the Fe(II) cofactor (Appendix B, Figure S3-1). From this state, mechanisms involving (A) a discrete, separated Fe-S complex [Fe(IV)-S^{2-} or Fe(III)-S•] that transfers sulfur to the olefin of the side chain[41] or (B) direct transfer not

involving a separated Fe-S species[38] can be envisaged. Direct detection of intermediates in the reaction could potentially provide dispositive tests of these hypotheses. In addition, it is unclear to what extent SPs recognize and discriminate on the basis of the aglycone side chains. Failure to do so could, for example, lead to initiation by ESP of the breakdown of an aglycone lacking an olefin to accept the sulfur. Would the enzyme then be inhibited/inactivated or, alternatively, might there be a facile route to release the simple organonitrile and inorganic sulfur for additional turnovers? The results reported in this study provide the kinetic underpinnings for design of experiments to address these interesting questions regarding the brassica MYR-GLS defense system and its specifier proteins.

3.4. CONCLUSIONS

We showed that the lifetimes of the aglycones produced from a series of glucosinolates vary widely and, for the more stable compounds, are sufficiently long to separate their production by MYR from their spontaneous Lossen-like rearrangement to the corresponding isothiocyanate or their chemical trapping by the colorimetric indicator, 2-PDS. This finding enabled the temporal and spatial separation of the MYR reaction from that of one specifier protein, recombinant *At* ESP, to show conclusively that the latter protein is an autonomous, iron-dependent lyase/isomerase. We presume that, as has repeatedly been proposed in recent literature, other SPs will be found also to act autonomously. The kinetic information that we obtained sets the stage for biophysical studies to resolve the divergent mechanisms of the SPs and the role(s) of the iron cofactor therein.

3.5. MATERIALS AND METHODS

Materials. We purchased *Sinapis Alba* seeds from Allied Seed LLC, *Escherichia coli* DH5α and BL21(DE3) competent cells from New England Biolabs, Ni(II)-nitrilotriacetic acid

agarose (Ni-NTA) resin from MCLAB, and Concanavalin A Sepharose 4B resin from General Electric Healthcare. We obtained glucosinolate Phyproof reference substances (sinigrin hydrate, glucoiberin potassium salt, gluconasturtiin potassium salt, glucoraphanin potassium salt, glucotropaeolin potassium salt, progoitrin potassium salt, sinalbin potassium salt, glucobrassicin potassium salt, and glucobrassicanapin potassium salt) from Supelco. We purchased allyl isothiocyanate, benzyl isothiocyanate, 2,2'-dithiodipyridine, methyl-α-D-mannopyranoside, hydroxocobalamin acetate, ethanolamine, amino acids (20 common), ascorbic acid, and ethylenediaminetetraacetic acid (EDTA) from Sigma-Aldrich. We purchased spectinomycin, L-(+)-arabinose, and disodium phosphate from DOT scientific, isopropyl β-D-1-thiogalactopyranoside (IPTG) and kanamycin from Gemini Bioproducts, ammonium chloride, monobasic potassium phosphate, and ferrous ammonium sulfate from EMD Millipore Corp, 4-(2-Hydroxyethyl)-1-piperazinepropanesulfonic acid (EPPS) from Chem-Impex International, Inc., and sodium chloride from Hawkins Pharmaceutical Group.

Purification of Sinapis Alba (Sa) Myrosinase (MYR). We purified MYR from *Sa* seeds by a procedure that we adapted from one reported by Pessina et al.[151] We ground 300 g of seeds with a mortar and pestle in liquid N_2 and removed lipids by extraction with 3 L of hexanes. We homogenized the dried seed meal in 0.9 L buffer (20 mM Tris-HCl, 0.25 M NaCl, pH 7.4) at room temperature (22 °C) for 30 min and filtered the homogenate through cheese cloth before centrifuging it at 22,000*g* for 20 min. We passed the supernatant through 11 μm filter paper (Whatman Grade 1) to remove suspended particles and then subjected it to a batch adsorption onto 50 mL Concanavalin A Sepharose 4B affinity resin. We incubated the suspension at 4 °C with gentle shaking for 40 min to allow for adsorption. We then washed the resin extensively with the same buffer in a chromatography column under gravity flow. We

eluted the protein in 50 mL fractions by washing with 0.4 L of 0.25 M methyl-α-D-mannopyranoside dissolved in the same buffer. We combined the fractions that contained MYR, which we detected by SDS-PAGE with Coomassie Brilliant Blue staining. We reduced the volume of the pooled fractions to 0.5 mL in a 30 kDa molecular weight cut off filtration device (PALL Corporation). We then dialyzed the purified protein against 4 L of storage buffer [25 mM sodium EPPS, 100 mM NaCl, 5% (w/v) glycerol, pH 7.9] to remove the sugar eluant.

Preparation of Arabidopsis thaliana (At) Epithionitrile Specifier Protein (ESP). We used two vectors – pBA28, which encodes an N-terminally SUMO-His_6 tagged *At* ESP and confers resistance to kanamycin (sequence provided in the *Supporting Information*), and pBAD42, which encodes membrane-bound transporter proteins for cobalamin and confers resistance to spectinomycin[152] – to transform *E. coli* BL21 (DE3) cells. Although they do not yet understand the mechanism, Prof. Squire Booker's group has determined that co-expression of the genes for the cobalamin system can (as we found for *At* ESP) result in an increased yield of the soluble form of a recombinant protein. We selected the co-transformants on LB-agar plates supplemented with kanamycin and spectinomycin (50 μg/mL each). We grew the transformants in 1-L cultures of M9 medium supplemented with 40 mM ethanolamine, 2.5 μM hydroxocobalamin, all 20 amino acids, and 50 μg/mL each of kanamycin and spectinomycin. We incubated the cultures in air at 37 °C with shaking at 200 rpm until they reached an optical density at 600 nm (OD_{600}) of 0.3, at which point we induced expression of the cobalamin-related genes by addition of arabinose to 0.2% (w/v). We incubated the cultures at 37 °C until they reached OD_{600} ~ 0.7 and then induced expression of *At* ESP by addition of IPTG and ferrous ammonium sulfate to 0.15 mM and 0.05 mM, respectively. We set the temperature of the incubator to 18 °C and incubated the cultures for an additional 17 h before harvesting the

cells by centrifugation at 6500*g*. We resuspended these cells in purification buffer (20 mM sodium phosphate, 250 mM NaCl, 15 mM imidazole, and 80 μg/mL phenylmethylsulfonyl fluoride at pH 7.5) at a ratio of 3 mL per g of cells. We lysed the cells by passing them through a French pressure cell (ThermoFisher) at 1200 psi. We removed the insoluble debris by centrifugation at 22,000*g* and then loaded the supernatant under gravity flow onto a 40-mL affinity chromatography column of Ni(II)-nitrilotriacetate agarose resin. We washed the column with ten volumes of the purification buffer and then eluted the bound protein by washing with an otherwise identical buffer containing 300 mM instead of 15 mM imidazole. We combined the fractions that contained ESP (identified by SDS-PAGE with Coomassie Brilliant Blue staining) and reduced the volume to 1 mL using a 30 kDa molecular weight cut off filtration device (PALL Corporation). We dialyzed the concentrated sample twice: first, against 25 mM sodium EPPS (pH 7.9), 150 mM NaCl, 10% [v/v] glycerol, and 5 mM ethylenediaminetetraacetic acid (EDTA) to remove free Ni(II) and, second, against the same buffer lacking EDTA to remove the chelator. We then rapidly froze the protein in liquid N_2 and stored it at - 80 °C.

Steady state Kinetic characterization of MYR preparations. We conducted MYR activity assays in 20 mM Tris-HCl (pH 7.4) buffer at room temperature (22 °C) in 0.5 mL quartz cuvettes with either 1 cm (for the two lowest sinigrin concentrations) or 0.2 cm (for the four highest sinigrin concentrations) pathlength. We incubated purified *Sa* MYR (2 μM) with 0.2 mM ascorbate and sinigrin at concentrations ranging from 0.080 mM to 1.5 mM. We initiated reactions by addition of a concentrated solution of sinigrin and monitored hydrolysis for 10 min at 227 nm using a Cary 60 UV-Visible Spectrophotometer, performing triplicate reactions for each sinigrin concentration. We calculated initial velocities (v_0) as $(\Delta A_{227})/(\Delta t \bullet \varepsilon_{227,\text{sinigrin}} \bullet$

pathlength) using slopes ($\Delta A_{227}/\Delta t$) from the early segments of the traces. We divided these values of v_0 by [MYR] to get turnover frequencies and plotted these values against [sinigrin]. We fit the data by the equation for a hyperbola to estimate the maximum turnover frequency (k_{cat}) and Michaelis-Menten constant (K_M). We determined the efficiency parameter, k_{cat}/K_M, separately from k_{cat} and K_M by dividing the slope of the tangent to the hyperbolic fit at [sinigrin] = 0 (which is equal to k_{cat} • [MYR]/K_M) by the known [MYR] of 2 μM. We used the programs Excel (Microsoft) and KaleidaGraph (Synergy Software) for this regression and statistical analysis.

Validation of 2,2'-Dipyridyl Disulfide (2-PDS) as Reporter of MYR-Generated Aglycones. Initial experiments to validate the use of 2-PDS as a selective reporter of the MYR-generated aglycones employed sinigrin in coupled reactions with all other required components. We carried out reactions in 100 mM sodium phosphate (pH 6) reaction buffer at 22 °C in a 0.5 mL quartz cuvette with a 1 cm pathlength, monitoring with an Agilent HP8453 diode array spectrophotometer. Assays contained 0.01 mM MYR, 0.25 mM ascorbate, and 1.5 mM 2-PDS and were initiated by addition of 0.10 mM sinigrin. Development of the absorption feature of 2-TP at 342 nm (ε_{342} = 7.9 $mM^{-1}cm^{-1}$) confirmed its time-dependent production in these reactions. The failure of this spectral signature to develop in control reactions lacking (i) sinigrin, (ii) MYR, or (iii) 2-PDS, or (iv) having sinigrin replaced by 0.10 mM of the allyl-NCS product of the Lossen-like rearrangement of its aglycone confirmed that 2-PDS reacts only with the aglycone. Use of other GLSs (0.10 mM) in place of sinigrin also generated 2-TP (*see below*).

Variation of [2-PDS] to Define Relative Rates of Trapping and Rearrangement. We carried out assays to map the dependence on [2-PDS] of the partition ratio between the bimolecular

aglycone trapping and the unimolecular rearrangement (Figure 3-2) as above, with 0.01 mM MYR, 0.25 mM ascorbate, and 0.1 mM glucosinolate in the same reaction buffer. The concentrations of 2-PDS tested were 0.1, 0.375, 0.75 or 1.5 mM. We plotted the final values of A_{342} from these reactions versus [2-PDS]. We fit these plots by Eq. 1 to obtain the value of k_1/k_2.

$$A_{342} = \underset{(\varepsilon_{342} \text{ of 2-TP})}{[\text{GLS}](\text{mM}) \bullet 7.9\ \text{mM}^{-1}\text{cm}^{-1}} \bullet 1\ \text{cm} \bullet \left(\frac{[\text{2-PDS}]}{\frac{k_1}{k_2} + [\text{2-PDS}]} \right) \quad (1)$$

Use of 2-PDS in Coupled Reactions to Define Timescale for Complete Hydrolysis of GLSs by MYR. Use of the 2-PDS indicator in this coupled configuration (with all components present at $t = 0$) allowed the time required for complete hydrolysis to be defined for sinigrin and the other GLSs. In these experiments, we used a Cary 60 UV-Visible absorption spectrophotometer equipped with kinetics capability and magnetic stirring. We performed all assays in a 3.5 mL quartz cuvette with a 1 cm path length and a micro-stir bar to ensure efficient mixing. We initiated the reactions by addition of a GLS. The cessation of growth of the 342-nm absorption feature of 2-TP signified completion of MYR-mediated hydrolysis.

Temporal Separation of the MYR and Trapping Reactions to Resolve the Rearrangement and Trapping Rate Constants. In these experiments, we omitted 2-PDS from the initial incubation (of 0.01 mM MYR, 0.25 mM ascorbate, and 0.1 mM sinigrin in reaction buffer at 22 °C), which we timed to be just long enough for the hydrolysis to reach completion (10-25 s, depending on the GLS). Addition of 2-PDS to 1.5 mM then allowed the trapping reaction to proceed, now resolved from the MYR reaction but still in competition with the rearrangement reaction. We fit the A_{342}-versus-time traces by the equation for an exponential growth (Eq. 3),

in which the extracted rate constant, k_{obs}, is equal to the sum of the effective first-order rate constants for the trapping reaction (k_2 • [2-PDS]) and the rearrangement reaction (k_1), as depicted in Figure 3-2 (Eq. 2).

$$k_{obs} = k_2 \bullet [\text{2-PDS}] + k_1 \quad (2)$$

$$A_{342}(t) = A_{342}(0) + \Delta A_{342} \bullet (1 - e^{-k_{obs} \bullet t}) \quad (3)$$

A varied delay between completion of GLS hydrolysis and addition of 2-PDS (Δt) allowed rearrangement of the aglycone to the RNCS to occur without competition from trapping. We plotted the final A_{342} values at completion of each trapping reaction versus Δt and fit the data as an exponential decay (Eq. 4) with rate constant (k_1) equivalent to that in Figure 3-2.

$$A_{342}(\Delta t) = A_{342}(\infty) + \Delta A_{342} \bullet (e^{-k_1 \bullet \Delta t}) \quad (4)$$

Coupled Reactions of MYR and Recombinant At ESP Demonstrating Fe(II)-Activated Epithiolation. We carried out reactions to verify the activity of the recombinant *At* ESP and its dependence on Fe(II) in a final volume of 0.2 mL at 22 °C for 15 min. Reactions included 2 µM MYR, 10 µM ESP, 0.25 mM ascorbate, and 0.5 mM sinigrin. Fe(II) (from ferrous ammonium sulfate) was present at 0-5 equiv relative to ESP. The Fe(II) and ascorbic acid stocks were freshly prepared in anoxic 2 mN H_2SO_4 and H_2O, respectively. We added benzyl isothiocyanate internal standard (0.5 mM) after incubation. We then extracted volatile products (isothiocyanates and epithionitriles) from the aqueous portion in 0.2 mL dichloromethane and analyzed products on a Shimadzu QP5000 GC-MS. We used a Rxi-1ms 30 m x 0.32 µm column and an injection temperature of 250 °C in split mode (ratio 1:5) with 2 µL injection

volume. The temperature program was: 75 °C held for 1 min, ramped to 275 °C at a rate of 25 °C/min and held for 1 min. We operated the mass spectrometer with an interface temperature of 280 °C in SIM mode using electron impact ionization. We detected the products and standard by their peaks at *m*/*z* = 99, 72, and 39 for allyl-NCS; 99, 72, and 59 for 1-cyano-2,3-epithiopropane (epithionitrile from sinigrin); and 145, 91, and 65 for benzyl-NCS. For the products, identification was supported by both their retention times and the match of their mass spectra to those in the NIST database.

Verification of Autonomous Enzymatic Activity of At ESP. To establish that ESP acts catalytically and independently of MYR, we separated the MYR and ESP reactions, first in time and then also in space. In the first experiment, we incubated a 0.2-mL reaction containing 25 μM MYR, 0.25 mM sinigrin, and 0.5 mM ascorbate in anoxic 50 mM sodium EPPS buffer (pH 7.9) at 22 °C for 30 s (to completion) and then added ESP and Fe(II) [$Fe(NH_4SO_4)_2$] simultaneously to concentrations of 0.1 mM and 0.2 mM, respectively. After an additional 4-min incubation, we added the benzyl isothiocyanate internal standard to a concentration of 0.25 mM, extracted the volatile products from the aqueous solution as described above, and analyzed them by GC-MS (as described). We also performed control experiments (1) with 0.01 mM ESP present at the initiation of the MYR reaction to demonstrate complete suppression of allyl-NCS formation and (2) with ESP omitted entirely. Finally, we quenched a sample after the initial 30-s reaction by incubation in boiling water to verify that sinigrin hydrolysis was complete. The aqueous portion of the last control was analyzed by direct injection into a Q Exactive orbitrap mass spectrometer (Thermo-Fisher Scientific) with heated electrospray sample introduction and mass resolution set to 70,000 FWHM at 200 *m*/*z*.

In the second experiment, we ruled out the remaining possibility that *At* ESP might associate with and be activated by the heterologous *Sa* MYR by also spatially separating the MYR and ESP. We loaded 0.25 mL of Concanavalin A Sepharose 4B resin into a 1.5 mL Micro Bio-Spin Chromatography column (Bio Rad) and washed the column five times with two volumes (each time) of 50 mM sodium EPPS buffer (pH 8) by centrifugation at 4000*g* for 15 s. We added ~ 5 mg of MYR in 0.3 mL of the same buffer onto the column and then resuspended the resin by repeated gentle inversion at 4 °C for 15 min. We washed the resin a second time with 5 x 2 column volumes of buffer to ensure that any unadsorbed MYR was removed. We placed the column in a clean 2 mL collection tube containing 0.20 mL of 0.30 mM ESP and 0.60 mM Fe(II) [$Fe(NH_4SO_4)_2$] in the same buffer. In the control sample, we used a column prepared as above and a collection tube containing 0.20 mL of only the buffer. We loaded 0.30 mL of 0.30 mM sinigrin and 0.70 mM ascorbate onto each column, inverted once, and then centrifuged at 4000*g* for 15 s to elute the allyl aglycone into the reservoir containing either the ESP and Fe(II) or the buffer. We added the benzyl-NCS internal standard to 0.20 mM to each sample and then extracted the products into 0.5 mL dichloromethane. We analyzed the products by GC-MS, as described in previous sections.

Estimation of k_{cat}/K_M of Recombinant At ESP for Allyl Aglycone. We performed coupled MYR-ESP reactions in 0.2 mL total volume at 22 °C in the anoxic reaction buffer described above. We fixed the concentration of sinigrin at 0.25 mM, varied [ESP] from 0.050-25 µM [with 2 equiv Fe(II) in each case], and fixed [MYR] at one of four concentrations: 0.050, 0.10, 0.15, or 0.50 µM. In each case, ascorbate was present at 20 equiv relative to MYR. We varied the incubation time to ensure complete sinigrin hydrolysis: 3 h, 2 h, 1 h, and 20 min for the reactions with 0.050, 0.10, 0.15, and 0.50 µM MYR, respectively. For each reaction, we added

benzyl-NCS internal standard to a concentration of 0.25 mM after this incubation, extracted the volatile products into 0.2 mL dichloromethane, and analyzed the products by GC-MS, as described above. We also carried out control reactions with ESP or MYR omitted. We plotted the relative peak areas for the allyl-NCS and epithionitrile products against [ESP] and fit the data by the expressions appropriate for the competition shown in Figure 3-7 (Eqs. 5 and 6, where *A* and *B* are amplitude and offset fit parameters). From the fit parameters and the value of k_1 determined above, we obtained the estimates for the catalytic efficiency (k_{cat}/K_M) of *At* ESP given in the text.

$$I_{\text{epithionitrile}} = A \bullet \left(\frac{[\text{ESP}]}{\frac{k_1}{k_{\text{cat}}/K_{\text{M}}} + [\text{ESP}]} \right) + B \quad (5)$$

$$I_{\text{isothiocyanate}} = A \bullet \left\{ 1 - \left(\frac{[\text{ESP}]}{\frac{k_1}{k_{\text{cat}}/K_{\text{M}}} + [\text{ESP}]} \right) \right\} + B \quad (6)$$

3.6. ACKNOWLEDGMENTS

We thank the reviewer for suggesting the experiment to spatially separate MYR and ESP.

Chapter 4

Assessment of the Side-chain Selectivities and Catalytic Efficiencies of Two Specifier Proteins from the Glucosinolate-Myrosinase Defense System *of Arabidopsis thaliana*

4.1. INTRODUCTION

The glucosinolate-myrosinase system (GLS-MYR) is the innate, biochemical defense mechanism specific to plants in the order of Brassicales (Figure 4-1). Plants of the family Brassicaceae (brassicas) have elicited the most research attention because they include such economically and agriculturally important crops as mustards, turnips, kale, cabbages, and rapeseed. The GLS-MYR system is initiated by tissue damage such as that caused by insect feeding or animal mastication. Such damage releases the compartmentalized β-thioglucosidase, myrosinase (MYR, EC 3.2.1.147; UniProt ID P29736), and results in its contact with the *S*-(β-D-glucopyranosyl)organothiohydroximate-*O*-sulfonate secondary metabolites known as glucosinolates (GLSs). MYR hydrolyzes the thioglucosidic bond of the GLS, producing an unstable organothiohydroximate-*O*-sulfonate (aglycone) intermediate.[50] This aglycone will non-enzymatically decay via a Lossen-like rearrangement to form an organoisothiocyanate (RNCS), which, depending on the identity of its side-chain (R in Figure 4-1), may be toxic to insects and thereby provide the plant with protection from herbivory.[49] Side-chains of known GLS are derived from amino acids and, to date, number more than 150.[1] Examples include allyl in sinigrin, 4-(methylthio)butyl in glucoerucin, and 2-phenylethyl in gluconasturtiin. The structure of the side-chain affects the bioactivity – including toxicity to insects – of the corresponding RNCS.

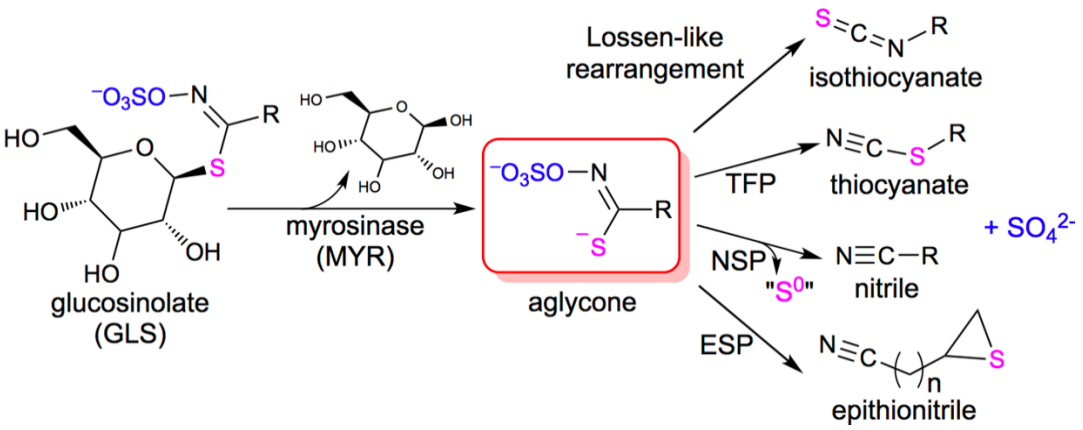


Figure 4-1. The Glucosinolate-Myrosinase Defense System of Brassica Plants and Activities of Specifier Proteins (ESP, NSP, TFP) Therein. Reprinted with Permission from Mocniak et al.[42]

Many brassicas have additional operatives in the GLS-MYR system, called specifier proteins (SPs) that can redirect glucosinolate breakdown to different, often less toxic, products. The known SPs share a common protein architecture, known as the Kelch-domain-repeat or β-propeller fold. Despite their similar structures, they direct formation of distinct products, including nitriles (nitrile specifier proteins, NSP), epithionitriles (epithionitrile specifier protein, ESP), and thiocyanates (thiocyanate forming protein, TFP). Although these SP isoforms are structurally homologous, each SP has the ability to accept multiple aglycones with varied side-chain compositions, which dictates the outcome of SP catalysis. For example, ESP produces its namesake product, epithionitrile, only from aglycones that contain a terminal olefin group on the side-chain.[48] Additionally, thiocyanate formation by TFP has, to date, been demonstrated only for three aglycones: allyl, benzyl, and 4-(methylthio)butyl.[47] Interestingly, when ESP and TFP encounter an aglycone with a side -chain composition that does not match these requirements, a simple nitrile is produced instead.[9] Accordingly, NSP forms simple

nitriles (RCN) from all aglycone side-chain compositions. The presence and specific isoform of SPs is highly varied among brassica species, and some SPs are even present in several brassica specialist insect species. Insects like the diamond back moth, *Plutella xylostella*, have developed a natural immunity to the GLS-MYR system and thrive by feeding on brassicas.[20,40,153] However, the evolutionary purpose of SPs in brassicas remains obscure since they effectively reduce the toxicity of the GLS-MYR defense mechanism, and their byproducts have not been proven to have significant biological implications.

Brassicas are often used in agriculture for a variety of applications, the most common of which are for animal grazing and soil biofumigation. Brassicas are routinely utilized for grazing animals because they are cool season crops that naturally extend the grazing season, allowing livestock to continue feeding on fresh plant material longer.[77] However, the corresponding RNCSs of certain GLS (i.e., progoitrin) have been shown to cause harmful side effects on animals grazed on a high brassica diet.[154] Alternatively, brassicas that are high in GLS content are often utilized for biofumigation which is a natural way to control soil borne pests without heavy use of pesticides.[103] The presence of SPs has a significant effect on the outcome of both these agriculturally important applications. If the mechanism by which SPs alter the aglycone rearrangement is known, it opens up the possibility to influence a plants ability to employ SPs to detoxify the GLS-MYR system and eliminate the negative side effects observed in animals grazed on a diet that is high in brassica content. Alternatively, it would also be possible to influence suppression of plant SP activity to improve its biofumigation efficiency. Knowledge of SP kinetics and mechanism of action is crucial for the success of these applications: if it is understood how these SPs chemically alter the aglycone, it will be possible improve the use of the GLS-MYR system for more targeted applications.

The utility of SPs in agriculture would have a positive impact, but routine use of SPs for these applications remains minimal due to the limited characterization of SPs to date. The mechanism by which SPs preempt the Lossen-like rearrangement of any aglycone remains unknown. It is also unclear whether, and to what extent, those SPs that mediate reactions requiring a specific side-chain structure (ESP and TFP) recognize and select for the competent side-chain or react promiscuously – but leading towards different outcomes – with diverse aglycones without regard to side-chain structure. The current knowledge of SPs mechanism of action and kinetic characterization has been severely restricted by the challenges involved in proper SP analysis, and the available information is limited to estimations based on computer modeling experiments (Figure 4-2 shows our proposed mechanisms of ESP and NSP mechanism of action).[38,53,155–157] The unstable nature of the aglycone products of MYR hydrolysis initially made it challenging to discern whether SPs act as autonomous lyase (or lyase/isomerase) enzymes or, alternatively, interact directly with MYR to influence the outcome of its hydrolysis reaction.[37,38,55] Our recent determination of the lifetimes of several aglycones enabled the physical and temporal separation of the MYR and ESP reactions to establish that the latter is, in fact, an autonomous enzyme.[42] Confirming past reports, we found iron to be required for maximal activity, but did not address the oxidation state of the cofactor.[38,41,142] Moreover, we did not address the autonomy of other SPs nor the substrate selectivity of any of these proteins. For ESP, in particular, the question of substrate selectivity would seem to be central to its function, given that formation of the namesake product requires the presence in the side chain of an olefinic sulfur acceptor. Does ESP select for aglycones with such a side chain? If not, does initiation of breakdown of an ETN-incompetent aglycone lead to accumulation of an inhibited, sulfurylated enzyme state?

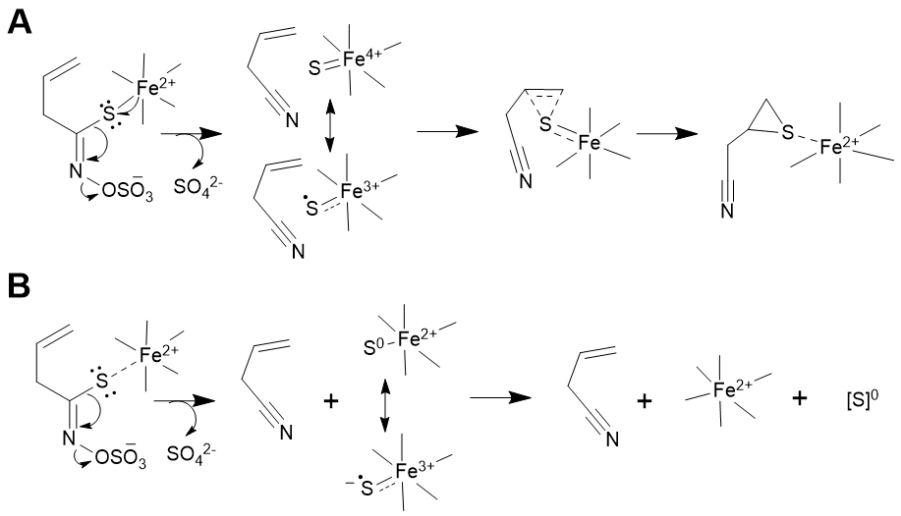


Figure 4-2. (A) Our proposed reaction mechanism of ETN formation from allyl aglycone in *At* ESP and (B) our proposed reaction mechanism of RCN formation from allyl aglycone in *At* NSP.

Here, we examined the substrate selectivity of ESP and compared it to that of NSP, which, we anticipated, might have a more efficient pathway for sulfur release and be less selective for an olefinic side chain in its substrate. Our experiments provide an estimated value for the catalytic efficiency of ESP with six substrates (aglycones from sinigrin, gluconasturtiin, glucotropaeolin, glucoerucin, gluconapin, and glucobrassicanapin), and our results indicate that ESP activity does not depend on the chemical composition of the aglycone side-chain. A similar analysis was performed with NSP with aglycones from sinigrin and glucoerucin, the results of which are in agreement with the that from ESP, NSP activity also does not depend on the aglycones side-chain composition. We also addressed the oxidation state of the iron cofactors in both proteins. Our analysis confirms previous reports that both SPs use the Fe(II) form of their cofactors. More importantly, it shows that ESP, despite its ability to catalyze transfer of the thiohydroximate sulfur to the allyl side chain of the sinigrin-derived aglycone without detectable loss to solvent, nevertheless efficiently degrades other aglycones that lack sulfur-accepting functionality to the corresponding simple nitriles. Lastly, we report the

efficacy of several compounds as stable aglycone structural mimics as ESP inhibitors to be used in future x-ray crystallography studies for SP active site elucidation.

4.2. RESULTS

4.2.1. Verification of NSP Autonomous Activity. In our prior study, we used MYR from *S. alba* seeds and recombinant *At* ESP (UniProt ID Q8RY71) expressed in *Escherichia coli* to show that the two enzymes could be separated in space, or their reactions separated in time, without loss of ESP function. Here, we sought demonstrate that the *At* NSP (UniProt ID Q9SDM9) also functions in this fashion. MYR immobilized on concanavalin A Sepharose 4B resin retains its ability to hydrolyze GLSs which enables complete spatial separation of MYR from NSP. An aliquot containing sinigrin was incubated with the MYR-adsorbed resin for 45 seconds and the resultant aglycone solution was collected in a reservoir containing Fe(II) and NSP. GC-MS chromatograms of the reaction mixture (Appendix C, Figure S4-2) show that allyl-CN was formed despite the physical separation of MYR and NSP. An identically prepared control without NSP (only buffer in the reservoir) showed only RNCS formation. These results show that, as for *At* ESP, *At* NSP activity is also fully autonomous and required no physical interaction with *Sa* MYR.

4.2.2. Investigation of Iron Cofactor Oxidation State. Although we previously confirmed that maximal activity of *At* ESP requires supplementation with Fe(II), we did not fully interrogate the oxidation state, Fe(II) or Fe(III). The reduced form of the protein was prepared by addition of stoichiometric Fe(II) to a solution of pure SPs in an anoxic chamber and treated with 5 mM sodium dithionite for 15 min at 23 °C. The oxidized, Fe(III) form was prepared the same but treated with 5 mM potassium ferricyanide for 15 min at 23 °C. Before the activity assays, both

were separated from small molecules by a desalting, chromatography step was added to remove any excess small molecules from the addition of iron. GC-MS assays monitoring the fractions of the allyl-isothiocyanate non-enzymatic and enzymatic epithionitrile/simple nitrile products show that, for both ESP and NSP, the Fe(II)-SP complex is more active (Figure 4-3). For ESP, reconstitution with Fe(II) activated by 22 fold relative to no addition of iron, and oxidation with ferricyanide diminished activity by 7 fold relative to the Fe(II). For NSP, Fe(II) activated by 19 fold and oxidation diminished activity by 3 fold. Although the SP purification protocol includes a dialysis step with EDTA to remove all divalent metals in solution, the chelation is not strong enough to fully strip SPs of iron incorporated during protein expression in *E. coli*. Therefore, SP preps typically retain approximately 10% iron occupancy, which is responsible for the minimal activity observed in the no-iron-added control (black traces in Figure 4-3).

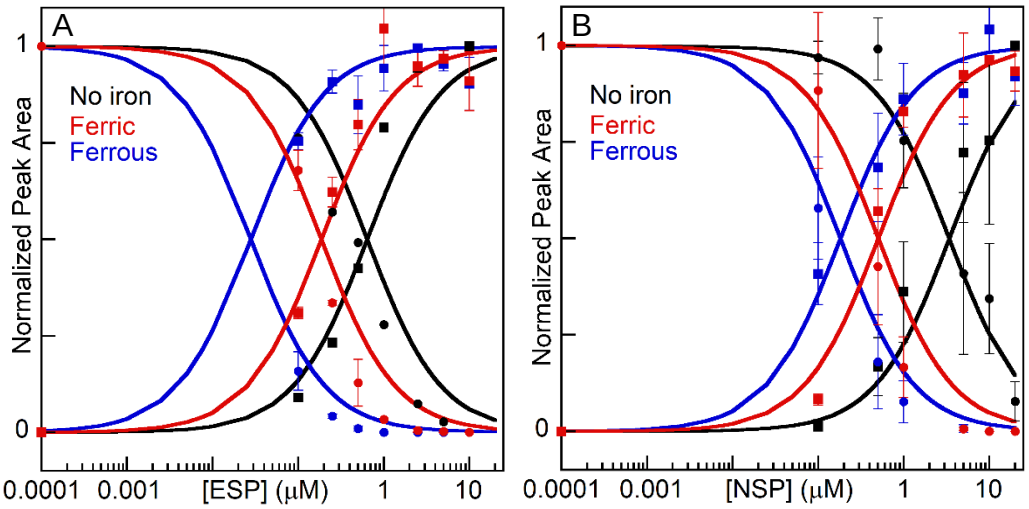


Figure 4-3. Effect of iron oxidation state on ESP (A) and NSP (B) catalytic efficiency. The average peak areas of volatile products (RNCS, ETN or RCN) were plotted against the corresponding [SP] (points with error bars). Each set of data was fitted according to Eqs 1 and 2 using a global analysis with linked parameters. The solid lines represent the best fit curve. The halfway points (where the curves cross) are visual representations of the k_{cat}/K_M values for SPs loaded with different forms of iron. I.e., the halfway point corresponds to [SP] where 50%

RNCS vs ETN/RCN is produced. The higher this concentration, the lower the $k_{cat}/K_{M,}$ or better catalytic efficiency. Black traces represent no-iron-added controls. The ferric trace is in red, and the ferrous trace is in blue.

4.2.3. Assessment of Specifier Protein Catalytic Efficiency and Substrate Specificity. We sought to determine the catalytic efficiency *(k_{cat}/K_M)* of *At* ESP using six aglycone substrates (from sinigrin, gluconasturtiin, glucotropaeolin, glucoerucin, gluconapin, and glucobrassicanapin) and of *At* NSP using two different aglycone substrates (from sinigrin and glucoerucin). As we described previously, variation of [SP] at initiation of GLS hydrolysis by MYR resulted in a predictable decrease in RNCS yield and corresponding increase in RCN or ETN yield as [SP] was increased. Provided that the assumption that the concentration of accumulated aglycone in the reaction does not the approach the K_M of either NSP or ESP (currently unknown), the dependence on [SP] is reasoned to result from competition between the Lossen-like rearrangement of the aglycone with the first-order rate constant k_1 and its capture by an SP with effective first-order rate constant [SP]•k_{cat}/K_M (Figure 4-4). According to this analysis, when equal amounts of RNCS and ETN are produced, k_{cat}/K_M•[SP]•[aglycone] = k_1•[aglycone], and $k_{cat}/K_M = k_1/$[SP]. The peak integrals of each product were plotted against [SP] and the traces were fitted according to Eqs 1 and 2, with the modification that a global analysis was used with linked parameters (Figure 4-5). Reactions were performed using multiple GLS (aglycones) to investigate whether the identity of the aglycone side-chain has an effect on the efficiency of SP activity. Results are summarized in Table 4-1.

Because ESP requires an aglycone with an olefinic side chain to produce its namesake product, it seemed plausible that it might exhibit selectivity for such substrates. To the contrary, we found that ESP can efficiently convert non-olefinic side chains of other aglycones to the corresponding RCNs. The results summarized in Table 4-1 show that the catalytic efficiency

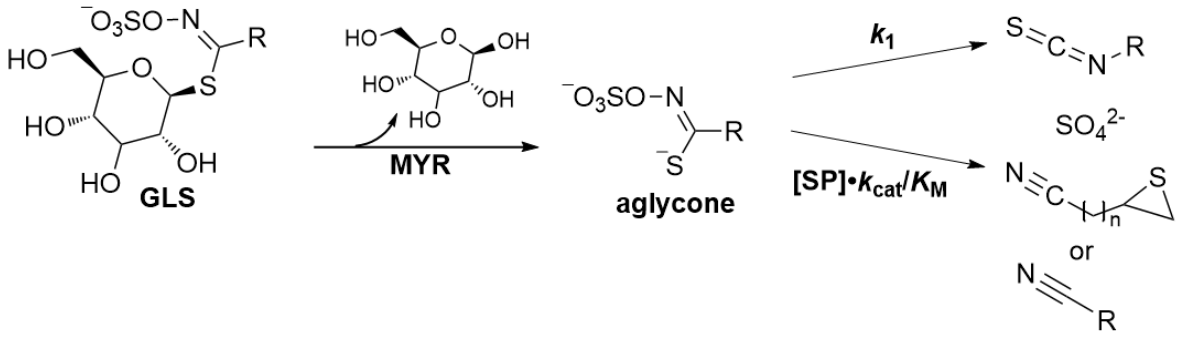


Figure 4-4. Competition between aglycone Lossen-like rearrangement of aglycone capture by SP used to analyze the catalytic efficiency of At ESP and At NSP for varied aglycones.

(k_{cat}/K_M) of ESP varies only modestly (3-fold) across six aglycones tested (Figure 4-5). Notably, improvements in the method of activation and assay resulted in values of k_{cat}/K_M for ESP of 10^6 $M^{-1}s^{-1}$, whereas our previous study had estimated ~ 3 x 10^4 $M^{-1}s^{-1}$ for ESP for the sinigrin aglycone. More importantly, the similarity of the values for the other aglycones implies that ESP (1) does not select (e.g., by virtue of greater affinity) for aglycone substrates that it can convert to epithionitriles and (2) does not become significantly inhibited upon initiating catalysis on an aglycone that cannot subsequently accept the thiohydroximate sulfur as epithiol. Similarly, NSP does not demonstrably discriminate between the allyl aglycone from sinigrin and the methylthiobutyl aglycone from glucoerucin.

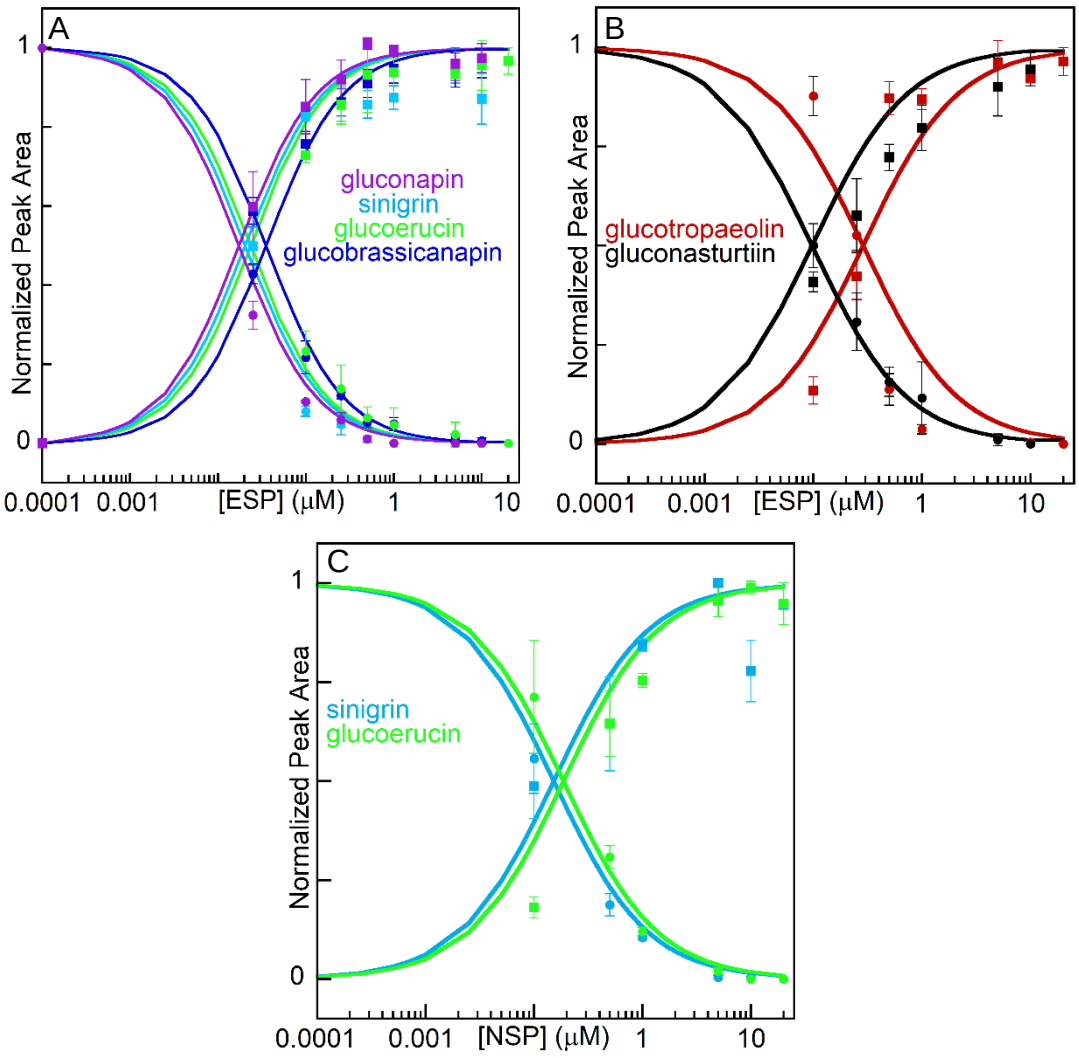


Figure 4-5. Evaluation of aglycone side-chain composition on SP activity. The average peak areas of volatile products (RNCS, ETN or RCN) were plotted against the corresponding [SP] (points with error bars). Each set of data was fitted according to Eq 1 and 2 using a global analysis with linked parameters. The solid lines represent the best fit curve. The halfway points (where the curves cross) are visual representations of the k_{cat}/K_M values. I.e., the halfway corresponds to [ESP] where 50% RNCS vs ETN/RCN is produced. The higher this concentration, the lower the $k_{cat}/K_{M,}$ or better catalytic efficiency. (A) Curve fits for ESP using aglycones from gluconapin (purple), sinigrin (light blue), glucoerucin (green), and glucobrassicanapin (dark blue). (B) Curve fits for ESP using aglycones from glucotropaeolin (red) and gluconasturtiin (black). (C) Curve fits for NSP using aglycones from sinigrin (light blue) and glucoerucin (green).

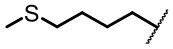
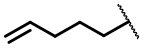
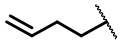
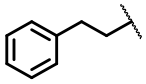
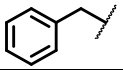
Glucosinolate/ Treatment	ESP $\frac{k_{cat}}{K_M}$ ($\mu M^{-1} s^{-1}$)	NSP $\frac{k_{cat}}{K_M}$ ($\mu M^{-1} s^{-1}$)	Aglycone R group
Low iron control [~20% load]	0.033 ± 0.004	0.006 ± 0.0008	Allyl Aglycone
Fe(II) [sodium dithionite]	0.75 ± 0.2	0.12 ± 0.01	Allyl Aglycone
Fe(III) [potassium ferricyanide]	0.11 ± 0.01	0.042 ± 0.006	Allyl Aglycone
Sinigrin (allyl)	1.0 ± 0.4	0.14 ± 0.02	
Glucoerucin	2.2 ± 0.4	0.25 ± 0.04	
Glucobrassicanapin	2.5 ± 0.3	n.d.	
Gluconapin	2.0 ± 0.4	n.d.	
Gluconasturtiin	3.4 ± 0.3	n.d.	
Glucotropaeolin	1.2 ± 0.2	n.d.	

Table 4-1. Summary of k_{cat}/K_M values for *At* ESP and *At* NSP. Experiments to analyze the oxidation state of the iron cofactor were all performed with sinigrin (allyl aglycone).

4.2.4. Effects of Synthetic Aglycone Mimics on Epithiospecifier Protein Activity A major deficit in current understanding of SP mechanism is the absence of an experimental structure with cofactor and substrate bound in the active site. Solution of a structure with an unreactive analogue of an aglycone might reveal details of enzyme-substrate interactions. As a first step toward design of such an inhibitor, we evaluated sulfate (SO_4^{2-}), phosphate (PO_4^{3-}), and (phenylthioamidomethyl)phosphonic acid for their abilities to inhibit ESP, presumably by occupying portions or all of the aglycone binding site (Figure 4-6). Sulfate is a poor ESP inhibitor, requiring 50 ± 2 mM to cause 50% inhibition ($[I]_{0.5}$), while phosphate showed increased affinity, $[I]_{0.5} = 3 \pm 0.7$ mM. On the basis of this insight, we synthesized a compound with a structure that mimics the aglycone core structure with a stabilizing replacement of the *N*-linked sulfate, predicting that it might prevent aglycone binding. This compound, (phenylthioamidomethyl)phosphonic acid, is a reasonably tight-binding inhibitor, with $[I]_{0.5} =$

0.2 ± 0.06 mM. The compound may thus be useful in obtaining crystals of the enzyme in a complex mimicking Michaelis complex.

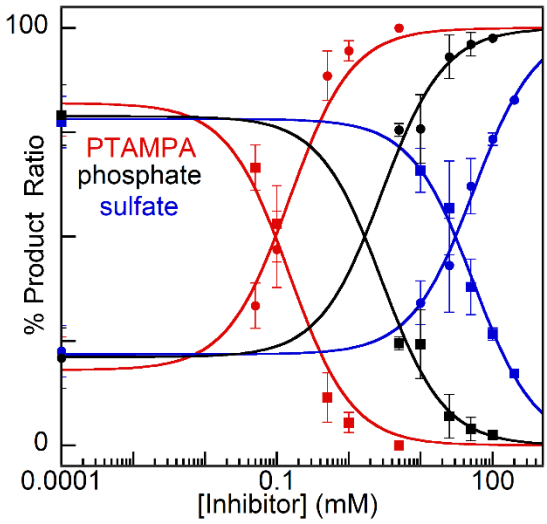


Figure 4-6. Evaluation of inhibitors on SP activity. The ratio of volatile products produced (RNCS, ETN or RCN) were plotted against the corresponding [inhibitor] (points with error bars). Each set of data was fitted according to Eq 3 and 4. The solid lines represent the best fit curve. The halfway points (where the curves cross) are visual representations of the $[I]_{50\%}$ values. High $[I]_{50\%}$ values indicate the compound is a poor inhibitor. PTAMPA = (phenylthioamidomethyl)phosphonic acid.

4.3. DISCUSSION

The results presented here and in our previous report have confirmed the favored hypothesis that SPs are autonomous iron-dependent enzymes that need not interact with MYR to exert their activities.[38,41,54] As both *At* ESP and *At* NSP are autonomous, we presume that TFP is as well, though it has not yet been confirmed. Additionally, the optimization of the SP reaction conditions by adding in a desalting step after activation increased the lower limit we can set for the k_{cat}/K_M of ESP by 100 fold compared to our previous report. A general comparison of the catalytic efficiencies of *At* ESP to *At* NSP shows that NSP is 10 fold less

active than ESP when presented with aglycones from glucoerucin and sinigrin. Although it is possible that the same reaction conditions (i.e., pH 8 reaction buffer) are less ideal for NSP than ESP, it is unlikely because ESP and NSP are both present in and occasionally localized together in aerial epidermal cells of *A. thaliana*, and therefore would have the same physiological reaction conditions.[43,158]

To our knowledge, there is very little information available regarding SP substrate recognition, and the available literature are reports from computational modeling studies that have not been experimentally supported. Assuming that the role of the Fe(II) cofactor in ESP catalysis is to facilitate the abstraction of the glycosidic-sulfur, it follows that the sulfur atom must be released to free the active site for subsequent turnovers. Therefore, the hypothesis that ESP will demonstrate lower activity (smaller k_{cat}/K_M values) with non-alkenyl aglycone substrates due to a lack of an available functional group for reincorporation of the abstracted sulfur to the aglycone is reasonable. However, the similar k_{cat}/K_M values for ESP with the various aglycones presented here implies that, when presented with a non-alkenyl aglycone, the rate of ESP catalysis does not depend on release of the abstracted glycosidic-sulfur.

The fate of the Fe(II)-abstracted sulfur atom is currently under scrutiny and remains unknown at this time, but our results do suggest that the sulfur abstraction and presumed release are not rate-limiting steps of the ESP mechanism of action. The recent hypotheses in literature suggest that the sulfur is released as elemental sulfur (S_8^0) or other forms of sulfur (HS^-, S_n^{2-}, H_2S).[38,41] To investigate this hypothesis, we tested several conventional colorimetric methods for free elemental sulfur/thiol/thiolate detection (methylene blue and 3,3'-dithiobis(6-nitrobenzoic) acid), but were unable to detect anything, despite adjusting the reaction conditions to promote unambiguous detection, if present. This suggests that the sulfur is not

"released" as expected. We postulate an alternative mechanism for sulfur removal that involves the formation of a persulfide with another aglycone (Figure 4-7), but may be otherwise incorporated to other inherent biochemical pathways for sulfur-containing compounds *in vivo*.[38]

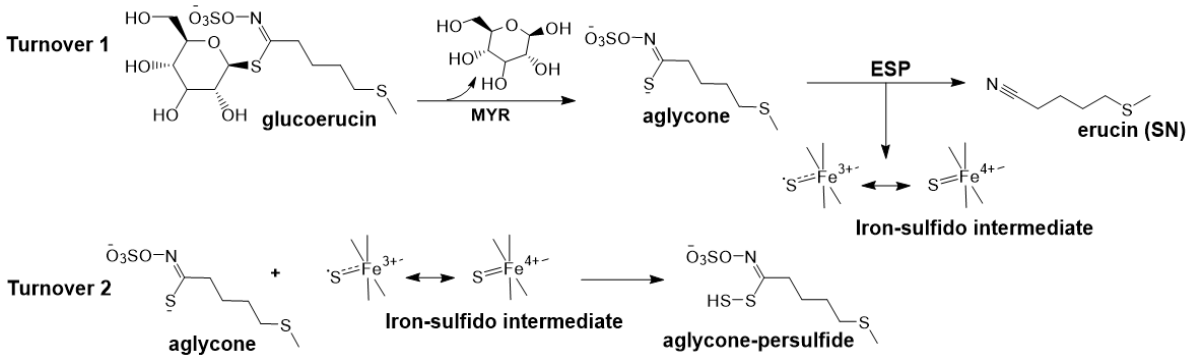


Figure 4-7. Formation of aglycone-persulfide from subsequent turnovers of aglycones and abstracted glycosidic-sulfur. Instead of the sulfur being spontaneously released from the SP active site, we propose that the sulfur is removed via persulfidation reaction with the free thiol on the next aglycone to enter the active site.

Identification of a stable SP inhibitor with a core structure that mimics the aglycone is useful for x-ray crystallography experiments to determine the structure of the SP active site. The current literature describing SP structure only includes one solved crystal structure for each SP, but all three are apo with no substrate bound. Therefore, if the SPs are crystallized with a stable substrate mimic bound in the active site, then x-ray crystallography will confirm the identity of the amino acid residues that interact with each portion of the aglycone mimic. This information is essential for full elucidation of SP mechanisms of action. The synthesized (phenylthioamidomethyl)phosphonic acid compound turned out to be a highly effective inhibitor of ESP activity. Future experiments will attempt to crystallize ESP with the

(phenylthioamidomethyl)phosphonic acid aglycone mimic bound in the active site. If successful, this will be the first experimental structural characterization of the ESP active site.

4.4. CONCLUSIONS

We demonstrated that the catalytic efficiencies of *At* NSP and *At* ESP do not depend on the chemical composition of the aglycone side chain. Therefore, SPs do not select for the specific aglycones that result in the SP namesake product (epithionitrile, thiocyanate, simple nitrile). We additionally present evidence to confirm the favored hypothesis that SPs are autonomous, Fe(II)-dependent enzymes. Future work to expand on the results presented here include identical experiments to update the characterization of TFP, and x-ray crystallography experiments for all three SPs with (phenylthioamidomethyl)phosphonic acid bound in the active site for elucidation of the SP active sites.

4.5. MATERIALS AND METHODS

Materials. Phyproof glucosinolate reference substances (sinigrin hydrate, gluconasturtiin potassium salt, glucotropaeolin potassium salt, glucoerucin potassium salt, gluconapin potassium salt, and glucobrassicanapin potassium salt) were obtained from Supelco. We purchased Concanavalin A Sepharose 4B resin from General Electric Healthcare, Ni(II)-nitrilotriacetic acid agarose (Ni-NTA) resin from MCLAB, *Sinapis alba* seeds from Allied Seed LLC, and *Escherichia coli* DH5α and BL21(DE3) competent cells from New England Biolabs. Benzyl isothiocyanate, allyl isothiocyanate, ethylenediaminetetraacetic acid (EDTA), aminomethylphosphonic acid, benzothioylchloride, methyl-a-D-mannopyranoside, hydroxocobalamin acetate, ascorbic acid, ethanolamine, sodium dithionite, potassium ferricyanide, and amino acids (20 common) were purchased from Sigma-Aldrich. We obtained

L-(+)-arabinose, disodium phosphate, and spectinomycin from DOT Scientific, kanamycin and β-D-1-thiogalactopyranoside (IPTG) from Gemini Bioproducts, ferrous ammonium sulfate, ammonium chloride, and monobasic potassium phosphate from EMD Millipore Corp, sodium chloride from Hawkins Pharmaceutical Group, and 4-(2-hydroxyethyl)-1-piperazinepropanesulfonic acid (EPPS) from Chem-Impex International, Inc.

Purification of Sinapis Alba (Sa) Myrosinase (MYR). MYR from *Sa* seeds was purified as reported previously.[42] Steady state kinetics of *Sa* MYR were also analyzed as previously reported.

Preparation of Arabidopsis thaliana (At) Epithionitrile Specifier Protein (ESP). *At* ESP was purified according to the procedure reported previously.[42]

Preparation of Arabidopsis thaliana (At) Nitrile Specifier Protein (NSP). *At* NSP prepared as described previously for *At* ESP[42] with the following modifications: a pBA28 vector was used to encode an N-terminally SUMO-His6 tagged *At* NSP and confer resistance to kanamycin (sequence provided in *Supporting Information*). Transformants were selected on LB-agar plates supplemented with kanamycin (50μg/mL) and grown in 1 L cultures of LB rich medium supplemented with 50μg/mL kanamycin. Cultures were incubated in air at 37 °C with shaking at 200 rpm until an optical density at 600 nm (OD600) of ~ 0.7 and then induced expression of *At* NSP by addition of IPTG and ferrous ammonium sulfate at 0.20 mM and 0.05 mM, respectively. The cultures were further incubated at 18 °C for 17 h and cells harvested by centrifugation at 6500g. *At* NSP was purified identical to *At* ESP, outlined previously (Appendix C, Figure S4-1).[42]

Verification of Autonomous Enzymatic Activity of At NSP. The reaction to verify the autonomous activity of At NSP was performed as previously described.[42]

Reactions of Recombinant At ESP and At NSP for Investigating the Oxidation State of the Iron Cofactor. Reactions were carried out in triplicate to verify the oxidation of the iron (ferrous or ferric) cofactor needed for recombinant *At* ESP and *At* NSP activity. Assays were performed in an anoxic chamber with a final volume of 0.2 mL EPPS buffer (pH 8) at room temperature (22 °C) for 15 min incubation. All reactions included 0.5 μM MYR, 0.015 mM ascorbate, and 0.25 mM sinigrin. ESP was present from 0 – 10 μM (NSP present from 0 – 20 μM,) added to each reaction using a 0.5 mM stock loaded with iron in its reduced (ferrous) or oxidized (ferric) state. A 25 mM iron stock was prepared from ferrous ammonium sulfate dissolved in anoxic 2 mN H_2SO_4. A strong reducing agent, sodium dithionite, was added in excess (5 mM) to 0.5 mM ESP (or NSP), then loaded with 0.75 mM iron (1.5 equivalents) from the 25 mM ferrous ammonium sulfate stock. The sodium dithionite ensures the iron remains in its reduced form. The SP stock treated with sodium dithionite and iron was allowed to equilibrate for 10 min, then all excess reducing agent and iron were removed by passing SP stock over a Zeba Spin Desalting Column with 7K MWCO and 0.5 mL volume (Thermo Scientific). An identical preparation was done with 5 mM potassium ferricyanide, a strong oxidizing agent. Benzyl isothiocyanate (0.25 mM) internal standard was added to each sample and the volatile products were extracted with 0.2 mL methylene chloride. Extracts were analyzed by GC-MS as described below.

Reactions of Recombinant At ESP and At NSP for Determination of Catalytic Efficiency. Reactions were performed as described above with the following adjustments to determine the

catalytic efficiency of *At* ESP and *At* NSP: the type of glucosinolate substrate was varied (each one added 0.25 mM, see Table 4-1 for all substrates tested, and the SP stocks (0.5 mM) were loaded with 1.5 equivalents of Fe(II) as described above, without additional reducing agent.

Synthesis of (phenylthioamidomethyl)phosphonic acid. (Phenylthioamidomethyl)phosphonic acid was synthesized by adding 177mg (1.6 mmol) (aminomethyl)phosphonic acid to 4 mL THF and cooled to 0 °C on ice. NaOH (2.8 mL of 2M, 5.6 mmol) was added dropwise and the solution was stirred at 0 °C until (aminomethyl)phosphonic acid was fully dissolved. 250 mg (1.6 mmol) benzothioylchloride chloride was added dropwise and the solution was stirred for 2 hr at 0 °C. After 2 hr, the ice bath was removed and the reaction was allowed to stir overnight at room temperature (22 °C). THF was removed with a rotary evaporator and the product was precipitated from solution with addition of 1 mL 6M HCl. The precipitate was washed extensively with deionized H_2O and briefly with ether to remove impurities. Total yield of 100 mg.

Reactions to Evaluate Synthesized Inhibitors as Aglycone Mimics. Reactions were performed as described above to determine the efficacy of the synthesized materials as ESP inhibitors with the following modifications: ESP was present at 50 nM and added last to each reaction, and the inhibitor concentration was titrated from 0 mM – 200 mM (sodium sulfate), 0 mM – 100 mM (sodium phosphate buffer at pH 8), and 0 mM – 50 mM ((phenylthioamidomethyl)phosphonic acid).

GC-MS Parameters. For GC-MS analysis of volatile products (isothiocyanates, epithionitriles, and simple nitriles), a benzyl isothiocyanate internal standard (0.5 mM) was added to each sample after reaction completion. For assays performed with glucotropaeolin (corresponding

RNCS is benzyl), allyl isothiocyanate (0.25 mM) was added. The volatile products were extracted from the aqueous portion in 0.2 mL dichloromethane and analyzed on a Shimadzu QP5000 GC-MS. We used a Rxi-1ms 30 m x 0.32 µm column, injection temperature of 250 °C in split mode (ratio 1:5) and 2 µL injection volume. The temperature program was: 75 °C held for 1 min, ramped to 275 °C at a rate of 25 °C/min and held for 1 min. The MS was operated with an interface temperature of 280 °C in SIM mode using electron impact ionization. The products and internal standard were detected by their expected fragmentation patterns. For the products, identification was supported by both their retention times and the match of their mass spectra to those in the NIST database. See Appendix C, Table S4-1 for corresponding *m/z* values.

***Global Analysis of SP Catalytic Efficiency and ESP Inhibition*.** The relative peak areas for the RNCS and ETN/RCN products were plotted against the corresponding [SP] and fit the data according to Eq 1 and 2 as reported previously for k_{cat}/K_M, with the modification that we used a global analysis (Igor Pro; WaveMetrics, Lake Oswego, OR, USA) with linked parameters k_1, k_{cat}/K_M, and A, where A is a parameter adjusting for the [GLS] at any given time. For the inhibition studies, the peak areas of RNCS or ETN produced were normalized to percent of total product and fit according to Eqs 3 and 4, where the term $k_i/(k_i + [I])$ adjusts for the dependence on the concentration of inhibitor.

$$I_{\text{epithionitrile}} = A \bullet \left(\frac{[\text{ESP}]}{\frac{k_1}{k_{\text{cat}}/K_{\text{M}}} + [\text{ESP}]} \right) + B \quad (1)$$

$$I_{\text{allyl-NCS}} = A \bullet \left\{ 1 - \left(\frac{[\text{ESP}]}{\frac{k_1}{\frac{k_{\text{cat}}}{K_{\text{M}}}} + [\text{ESP}]} \right) \right\} + B \quad (2)$$

$$I_{\text{allyl-NCS}} = A \bullet \left\{ 100 - \left(\frac{[\text{ESP}] \bullet \frac{k_i}{k_i + [I]}}{\frac{k_1}{\frac{k_{\text{cat}}}{K_{\text{M}}}} + \left([\text{ESP}] \bullet \frac{k_i}{k_i + [I]} \right)} \right) \right\} \quad (3)$$

$$I_{\text{epithionitrile}} = A \bullet \left(\frac{[\text{ESP}] \bullet \frac{k_i}{k_i + [I]}}{\frac{k_1}{\frac{k_{\text{cat}}}{K_{\text{M}}}} + \left([\text{ESP}] \bullet \frac{k_i}{k_i + [I]} \right)} \right) \quad (4)$$

4.6. ACKNOWLEDGEMENTS

This work was supported in part by grant 1R35 GM127079-03 to Professor Carsten Krebs.

Chapter 5
Current and Future Directions

5.1. CURRENT DIRECTION: Biological Assays to Investigate the Effects of ESP on the toxicity of the GLS-MYR system to insect larvae

5.1.1. Background and Methods. Isothiocyanates (RNCSs) are well known for their bioactivity (toxicity) towards pests (insects, bacteria, fungi, etc.). However, there have been no *in vivo* experiments performed to directly examine the biological effects of specifier protein (SP) activity on insect health and response to SP-catalyzed alteration to the glucosinolate-myrosinase (GLS-MYR) system. Therefore, we designed a preliminary experiment to investigate the ability of the epithionitrile specifier protein from *Arabidopsis thaliana* (*At* ESP) to detoxify the effects of the GLS-MYS system of white mustard (*Sinapis alba*) seeds on three types of insect larvae, cabbage looper (*Trichoplusia ni*), beet armyworm (*Spodoptera exigua*), and diamondback moth (*Plutella xylostella*). Two separate trials were performed, one with first instar larvae (early-stage development, newly hatched) and other with fourth instar larvae (late-stage development, close to cocooning). Larvae were obtained from Benzon Research. The bioassays were performed using white mustard seed (*Sinapis alba*) as a biologically relevant MYR source instead of adding MYR that was purified using affinity chromatography. Assays were performed in 1 L mason jars with airtight lids. Each larvae species was exposed to nine different treatments to examine the effects of three GLS concentrations (the observed effects are caused by the corresponding variation in RNCS produced) on the larvae with or without addition of ESP. Each treatment was performed in triplicate.

For the first trial, low, middle, and high concentrations of sinigrin (allyl glucosinolate) were calculated based on the reported toxicity of allyl-NCS to several species of insect larvae (LC_{50}

value ~5 μL/L RNCS)[159,160] with the assumption that 100% of sinigrin is converted to allyl-NCS over a 24 h period. Trial 1 utilized first instar larvae and included ~15 larvae per 1L mason jar, except the diamondback moth larvae which had 30+ per jar (they are very small and difficult to add individually). Trial 2 utilized fourth instar larvae and had 3-5 per 1L jar, depending on number of available larvae per species. 10mL diet (General Noctuid powder, which forms an agar gel when prepared in boiling water) was dispensed on the bottom of each jar. Insect larvae were placed directly on the solidified diet on the bottom of the jars using a paint brush.

For each treatment (except the negative and positive controls), 7g of ground, defatted white mustard seed was placed in a 3 oz plastic cup. Ascorbic acid was added to each plastic cup for a final concentration of 40 μM to catalyze MYR hydrolysis of sinigrin. Treatments were added to dry seeds first (sinigrin, ESP, ascorbate) and 10 mL dilute buffer (10mM sodium phosphate, pH 8) was added last to mix the treatments and initiate MYR reaction (MYR in dried mustard seeds is active in the presence of water). The allyl RNCS standard was dissolved in 1mL vegetable oil, and the oil was dispersed on a cotton ball placed inside the plastic cup. The plastic cups with the treatments were carefully placed in the center of the jar with tweezers, taking care not to place cups on top of larvae. Jars were sealed with gastight lids and incubated for 24 h in a growth chamber set to 27 °C, 65% relative humidity, incandescent light for a photoperiod of 14 h. Insect health (mortality, motility, size and color) was observed after 24 hours. A cotton swab was used to as touch stimulus. The treatments were as follows:

1. RNCS positive control (10 μL allyl RNCS standard)
2. Baseline control (7g seeds, 10mL buffer, and 40 μM ascorbic acid)
3. Negative control (Larvae and diet only)
4. MYR only – low [GLS] (7g seeds, 10mL buffer, 40 μM ascorbic acid, and 20 μM sinigrin)

5. MYR only – mid [GLS] (7g seeds, 10mL buffer, 40 μM ascorbic acid, and 50 μM sinigrin)
6. MYR only– mid [GLS] (7g seeds, 10mL buffer, 40 μM ascorbic acid, and 100 μM sinigrin)
7. MYR + ESP – low [GLS] (7g seeds, 10mL buffer, 40 μM ascorbic acid, 20 μM sinigrin, and 9 μM *At* ESP)
8. MYR + ESP – mid [GLS] (7g seeds, 10mL buffer, 40 μM ascorbic acid, 50 μM sinigrin, and 9 μM ESP)
9. MYR + ESP – high [GLS] (7g seeds, 10mL buffer, 40 μM ascorbic acid, 100 μM sinigrin, and 9 μM ESP)

After the first trial we discovered that the mustard seeds contain a very high concentration of sinalbin (hydroxybenzyl glucosinolate), ~4 mg per 1g of mustard seed, which incidentally increased the concentration of RNCS produced in the first trial by > 50x the intended concentration. This resulted in each treatment being significantly more toxic than originally planned. Therefore, the following modifications were made to the second trial with fourth instar larvae: the weight of seeds added per treatment was adjusted to coordinate with the intended [GLS] in each treatment instead of adding sinigrin (0.1g seed for the low [GLS] treatment, 1 g for the middle [GLS] treatment, and 4 g for the high [GLS] treatment). Accordingly, the concentration of ESP and ascorbate were adjusted in each treatment to maintain a consistent ratio of 1:10 MYR to ESP and 1:20 MYR to ascorbate. Additionally, we removed seeds from the baseline control, instead adding only buffer and ascorbate to the plastic cup. Qualitative analysis after the 24 h incubation was done via direct comparison of treatments to the negative and positive controls.

5.1.2. Trial 1 – First Instar Larvae (Early Stage)

The results of the first trial were largely inconclusive because most treatments negatively affected the larvae. Although it is possible that the presence of SPs enabled larvae survival, there was large variation between replicates that prevented conclusive observations (up to 70%

deviation between replicates). Larvae health was qualitatively analyzed by observing a large percentage of larvae did not survive or had very low response to touch stimulus, with the exception of the negative control (all larvae were healthy). No larvae survived the positive control.

5.1.3. Trial 2 – Fourth Instar Larvae (Late Stage)

5.1.3.1. Diamondback Moth. Diamondback moths are considered brassica specialists, a species of insect that has evolved to not only survive, but thrive by feeding solely on brassica plants.[161,162] It has also been shown that diamondback moths have a preference for *S. alba* plants,[163] so it was unsurprising that the larvae were attracted to the mustard seeds. More larvae made their way into the treatment cup/into the seed solution than remaining on the agar diet after 24 h. For comparison, very few beet armyworm or cabbage looper larvae ended up in the treatment cup. Healthy diamondback moth larvae (negative control) displayed very high motility in response to touch stimulation (moved very quickly away from stimulation). On average, the diamondback moth larvae had the highest tolerance to the GLS-MYR system. The adverse effects were mostly observed as changes to larvae motility, and was visibly reduced in the high [GLS]-MYR only treatment compared to the negative control. The larvae in the high [GLS]-MYR only treatment were also typically smaller than the negative control. Therefore, the presence of ESP improved the larvae motility and prevented size reduction at high [GLS].

5.1.3.2. Cabbage Looper. Interestingly, the Cabbage looper also has a preference for brassica plants,[164] but it is not considered a specialist insect and does not have its own innate protection against the GLS-MYR system. The larvae in the negative control were very responsive to touch stimuli, typical movement was rapid wiggling or “looping” (curled into balls). The larvae in the middle [GLS] treatment showed a reduced response to touch stimuli

(lowered motility). Unfortunately, there were not enough healthy larvae supplied by Benzon to execute the high [GLS] treatment. However, the larvae in the middle [GLS] with ESP treatment were as healthy as the larvae in the negative control treatment. The larvae in the high [GLS] with ESP treatment displayed reduced motility in response to touch stimuli. Therefore, the presence of SPs was beneficial to insect health during the 24 h period.

5.1.3.3. Beet Armyworm. The beet armyworm is a generalist insect with no preference for brassica plants. The larvae in the negative control demonstrated a quick response to touch stimuli (they typically rolled over a few times in a hurried manner). In the middle and high [GLS]-MYR only treatments, the larvae were, on average, smaller than the negative control larvae. They also displayed a reduced response to touch stimuli, generally slower movement. However, the larvae in the middle and high [GLS] treatments with ESP present displayed no adverse effects. These results support the hypothesis that the presence of SPs enables insect survival when exposed to the GLS-MYR system.

5.1.4. Conclusions and Future Directions

Overall, the general results from these biological trials indicate that insect larvae are able to better endure the toxic nature of the GLS-MYR defense system of brassica plants when SPs are present. To our knowledge, this is the first *in vivo* evidence for SPs ability to reduce the natural toxicity of the GLS-MYR system, and demonstrates the encouraging prospect for SP use in targeted *in vivo* applications. As stated previously, the overarching, long term goal of this project is to be able to manipulate the natural GLS-MYR mechanism and SP protein activity for more efficient agricultural applications like livestock feed or biofumigation. Moving forward, optimization of the experimental conditions used in this study would be useful for comparison of the innate physiological conditions of brassicas against the bioassay

conditions. If the most efficient, optimized experimental conditions align with the physiological conditions in brassicas, then largescale field trials can be carried out. If the optimal and physiological conditions are largely different, more focus can be placed on further development of the processes involved in executing the bioassays.

5.2. FUTURE DIRECTION: Investigation of the fate of the SP catalyzed abstraction of the glycosidic-sulfur

One of the most important pieces of information that is essential for SP mechanism elucidation is the fate of the sulfur abstracted during SP catalysis with non-specific aglycones (i.e., NSP activity with any aglycone, or ESP activity with non-olefinic aglycones). The current hypotheses suggest it is released in various forms of sulfur (S_8^0, HS^-, S_n^{2-}, H_2S), but there is a complete lack of experimental evidence to support this. I have done some preliminary work in this dissertation in an attempt to investigate these claims using two conventional colorimetric methods for detection of these sulfur species, the Cline method (via formation of methylene blue), and a method that induces the reactivation of disulfide-inhibited papain enzyme.[165,166] Neither method yielded successful results. It has also been suggested (both by us and others) that the sulfur is actually recycled *in vivo* back into other metabolic processes by attaching to a biological sulfur acceptor.[167] Especially for ESP, this could be investigated by supplying allyl cyanide to an ESP reaction with a non-alkenyl aglycone as a sulfur acceptor. The results would be straightforward to interpret, if any allyl ETN is detected with the products, then the ESP is using the added allyl cyanide as a sulfur sink. This would also be an interesting test for SP substrate binding abilities. Determination of the fate of the abstracted sulfur during SP catalysis should be the next area of inquiry to continue this project.

5.3. FUTURE DIRECTION: In-depth Characterization of *Ta* TFP

The same in-depth characterization of TFP from *Thlaspi arvense* (TFP is not present in *A. thaliana*) should be performed as was demonstrated for *At* NSP and *At* ESP, to confirm the hypotheses that all three SPs are autonomous, Fe(II) dependent enzymes. Given that our results for the catalytic efficiencies for ESP and NSP show their catalytic activity is not dependent on the aglycone side chain, we presume the same to be true for TFP. However, TFP has been suggested to have a smaller active site with more substrate constraint than ESP or NSP, which results in extreme substrate selectivity. TFP has only been shown to catalyze formation of TCN from three aglycones- allyl, benzyl, 4-methylthio. When presented with any other aglycone, TFP forms RCNs. Therefore, it would be interesting to verify whether TFP follows the same trend as ESP and does not demonstrate a bias for catalysis with alternative aglycone side chain compositions.[53]

5.4. FUTURE DIRECTION: Specifier Protein Inhibition and X-ray Crystallography

The current knowledge of SP structure is limited because the inherent instability of the aglycone substrate prevents crystallization of SPs with substrate bound in the active site. Additionally, the current crystal structures of all three SPs have also been solved without the iron cofactor.[47,48,52] In Chapter 4, I present some preliminary results for the identification of a SP inhibitor, (phenylthioamidomethyl)phosphonic acid (PTAMPA), that is a stable structural mimic of the aglycone. Analysis of PTAMPA for ESP inhibition revealed that it is an efficient inhibitor. If we assume that efficient SP inhibition by any aglycone mimic is achieved through competitive binding in the active site, then PTAMPA would be very useful in X-ray crystallography experiments. Successful crystallization of SPs with PTAMPA bound will enable detailed analysis of the amino acid residues present in the active site, which is

imperative information for SP mechanism elucidation. The recent literature remains dependent on computational models,[38,41] and although some mutagenesis studies[53] have identified several amino acid residues for SP function, a crystal structure with a substrate mimic bound in the active site would be the first empirical evidence to confirm or refute the proposed SP mechanisms of action and proposed iron/ substrate binding sites.

5.5. FUTURE DIRECTION: Site-Directed Mutagenesis of SPs for Mechanism Elucidation

Mutational analyses have been performed to verify that several amino acid residues are essential for proper SP function (outlined in more detail in Chapter 1).[53] However, mutagenesis experiments that target the SP ability to re-apply the abstracted sulfur have not been investigated. As stated above, successful elucidation of a SP crystal structure with both the iron cofactor and substrate mimic bound would provide details on the amino acid residues present in the active site and what their function would be (substrate recognition, catalysis, etc.). Using this information, it could be possible to create SP variants that affect the abstraction of the glycosidic sulfur, but not inhibit general SP function. For example, when the successful ESP variant is provided with allyl aglycone, it catalyzes the formation of RCN instead of the usual ETN. The variant will continue to prevent the aglycone from the Lossen-like rearrangement to an RNCS, but cannot perform the epithiolation step. Identification of the ESP variant that achieves this result would provide essential information regarding the role of the iron cofactor in SP catalysis.

REFERENCES

(1) Blažević, I.; Montaut, S.; Burčul, F.; Olsen, C. E.; Burow, M.; Rollin, P.; Agerbirk, N. Glucosinolate Structural Diversity, Identification, Chemical Synthesis and Metabolism in Plants. *Phytochemistry* **2020**, *169*, 1–57. https://doi.org/10.1016/j.phytochem.2019.112100.

(2) Jeschke, V.; Burow, M. Glucosinolates. *eLS* **2018**, 1–8. https://doi.org/10.1002/9780470015902.a0027968.

(3) Miao, H.; Wang, J.; Cai, C.; Chang, J.; Wang, Q. Accumulation of Glucosinolates in Broccoli. In *Glucosinolates, Reference Series in Phytochemistry*; 2016; pp 1–30. https://doi.org/10.1007/978-3-319-26479-0.

(4) Liu, A. G.; Juvik, J. A.; Jeffery, E. H.; Berman-Booty, L. D.; Clinton, S. K.; Erdman, J. W. Enhancement of Broccoli Indole Glucosinolates by Methyl Jasmonate Treatment and Effects on Prostate Carcinogenesis. *J. Med. Food* **2014**, *17* (11), 1177–1182. https://doi.org/10.1089/jmf.2013.0145.

(5) Barba, F.; Nikmaram, N.; Roohinejad, S.; Khelfa, A.; Zhu, Z.; Koubaa, M. Bioavailability of Glucosinolates and Their Breakdown Products : Impact of Processing. *Front. Nutr.* **2016**, *3*. https://doi.org/10.3389/fnut.2016.00024.

(6) Ishida, M.; Hara, M.; Fukino, N.; Kakizaki, T.; Morimitsu, Y. Glucosinolate Metabolism, Functionality and Breeding for the Improvement of Brassicaceae Vegetables. *Breed. Sci.* **2014**, *64* (1), 48–59. https://doi.org/10.1270/jsbbs.64.48.

(7) Sønderby, I. E.; Geu-flores, F.; Halkier, B. A. Biosynthesis of Glucosinolates – Gene Discovery and Beyond. *Trends Plant Sci.* **2010**, *15* (5), 283–290. https://doi.org/10.1016/j.tplants.2010.02.005.

(8) Bhat, R.; Vyas, D. Myrosinase: Insights on Structural, Catalytic, Regulatory, and

Environmental Interactions. *Crit. Rev. Biotechnol.* **2019**, *39* (4), 508–523. https://doi.org/10.1080/07388551.2019.1576024.

(9) Wittstock, U.; Kurzbach, E.; Herfurth, A. M.; Stauber, E. J. Glucosinolate Breakdown. In *Advances in Botanical Research*; Kopriva, S., Ed.; Elsevier Ltd, 2016; Vol. 80, pp 125–169. https://doi.org/10.1016/bs.abr.2016.06.006.

(10) Ettlinger, M. G.; Dateo, George P., J.; Harrison, B. W.; Mabry, T. J.; Thompson, C. P. Vitamin C as a Coenzyme: The Hydrolysis of Mustard Oil Glucosides. *PNAS* **1961**, *47*, 1875–1880.

(11) Burmeister, W. P.; Cottaz, S.; Rollin, P.; Vasella, A.; Henrissat, B. High Resolution X-Ray Crystallography Shows That Ascorbate Is a Cofactor for Myrosinase and Substitutes for the Function of the Catalytic Base. *J. Biol. Chem.* **2000**, *275* (50), 39385–39393. https://doi.org/10.1074/jbc.M006796200.

(12) Burmeister, W. P.; Cottaz, S.; Driguez, H.; Iori, R.; Palmieri, S.; Henrissat, B. The Crystal Structures of Sinapis Alba Myrosinase and a Covalent Glycosyl-Enzyme Intermediate Provide Insights into the Substrate Recognition and Active-Site Machinery of an S-Glycosidase. *Structure* **1997**, *5* (5), 663–676.

(13) Bhat, R.; Kaur, T.; Khajuria, M.; Vyas, R.; Vyas, D. Purification and Characterization of a Novel Redox-Regulated Isoform of Myrosinase (β-Thioglucoside Glucohydrolase) from *Lepidium Latifolium* L. *J. Agric. Food Chem.* **2015**, *63* (47), 10218–10226. https://doi.org/10.1021/acs.jafc.5b04468.

(14) Ahuja, I.; de Vos, R. C. H. De; Rohloff, J.; Stoopen, G. M.; Halle, K. K.; Jam, S.; Ahmad, N.; Hoang, L.; Hall, R. D.; Bones, A. M. Arabidopsis Myrosinases Link the Glucosinolate-Myrosinase System and the Cuticle. *Nat. Publ. Gr.* **2016**.

https://doi.org/10.1038/srep38990.

(15) Sehrawat, A.; Sougrakpam, Y.; Deswal, R. Cold Modulated Nuclear S-Nitrosoproteome Analysis Indicates Redox Modulation of Novel Brassicaceae Specific, Myrosinase and Napin in *Brassica Juncea*. *Environ. Exp. Bot.* **2019**, *161* (July 2018), 312–333. https://doi.org/10.1016/j.envexpbot.2018.10.010.

(16) Andreasson, E.; Jorgensen, L. B. Localization of Plant Myrosinases and Glucosinolates. In *Recent Advances in Phytochemistry*; 2003; pp 79–99.

(17) Nong, H.; Zhang, J.; Li, D.; Wang, M.; Sun, X.; Zhu, Y. J.; Meijer, J.; Wang, Q. Characterization of a Novel β-Thioglucosidase CpTGG1 in Carica Papaya and Its Substrate-Dependent and Ascorbic Acid-Independent O-β-Glucosidase Activity. *J. Integr. Plant Biol.* **2010**, *52*, 879–890. https://doi.org/10.1111/j.1744-7909.2010.00988.x.

(18) Barrett, T.; Suresh, C. G.; Tolley, S. P.; Dodson, E. J.; Hughes, M. A. The Crystal Structure of a Cyanogenic P-Glucosidase from White Clover, a Family 1 Glycosyl Hydrolase. *Structure* **1995**.

(19) Ohtsuru, M.; Hata, T. The Interaction of L-Ascorbic Acid with the Active Center of Myrosinase. *Biochim. Biocphysica Acta* **1979**, *567*, 384–391.

(20) Williams, D. J.; Critchley, C.; Pun, S.; Chaliha, M.; O'Hare, T. J. Differing Mechanisms of Simple Nitrile Formation on Glucosinolate Degradation in *Lepidium Sativum*and *Nasturtium Officinale* Seeds. *Phytochemistry* **2009**, *70* (11–12), 1401–1409. https://doi.org/10.1016/j.phytochem.2009.07.035.

(21) Matusheski, N. V.; Swarup, R.; Juvik, J. A.; Mithen, R.; Bennet, M.; Jeffery, E. H.; Bennett, M.; Jeffery, E. H. Epithiospecifier Protein from Broccoli (Brassica Oleracea

L. Ssp. Italica) Inhibits Formation of the Anticancer Agent Sulforaphane. *J. Agric. Food Chem.* **2006**, *54* (6), 2069–2076. https://doi.org/10.1021/jf0525277.

(22) Shikita, M.; Fahey, J. W.; Golden, T. R.; Holtzclaw, W. D.; Talalay, P. An Unusual Case of "uncompetitive Activation" by Ascorbic Acid: Purification and Kinetic Properties of a Myrosinase from Raphanus Sativus Seedlings. *Biochem. J.* **1999**, 725–732.

(23) Thomas, M.; Alsarraf, J.; Araji, N.; Tranoy-Opalinski, I.; Renoux, B.; Papot, S. The Lossen Rearrangement from Free Hydroxamic Acids. *Org. Biomol. Chem.* **2019**, *17* (22), 5420–5427. https://doi.org/10.1039/c9ob00789j.

(24) Lossen, W. Ueber Benzoylderivate Des Hydroxylamins. *Liebigs Ann. Chem.* **1872**, *161*, 347–362.

(25) Bright, R. D.; Hauser, C. R. The Influence of Substituents on the Rates of Decomposition of the Potassium Salts of Dihydroxamic Acids. The Lossen Rearrangement. *J. Am. Chem. Soc.* **1939**, *61*, 618–629.

(26) Ohmoto, K.; Yamamoto, Tetsuya Horiuchi, Toshihide Kojima, T.; Hachiya, Katsutoshi Hashimoto, Shinsuke Kawamura, Masanori Nakai, H.; Toda, M. Improved Synthesis of a New Nonpeptidic Inhibitor of Human Neutrophil Elastase. *Synlett* **2001**, *2*, 299–301.

(27) Zhao, J.; Gimi, R.; Katti, S.; Reardon, M.; Nivorozhkin, V.; Konowicz, P.; Lee, E.; Sole, L.; Green, J.; Siegel, C. S. Process Development of a GCS Inhibitor Including Demonstration of Lossen Rearrangement on Kilogram Scale. *Org. Process Res. Dev.* **2015**, *19* (5), 576–581. https://doi.org/10.1021/op500379a.

(28) McDermott, T. S.; Bhagavatula, L.; Borchardt, T. B.; Engstrom, K. M.; Gandarilla, J.;

Kotecki, B. J.; Kruger, A. W.; Rozema, M. J.; Sheikh, A. Y.; Wagaw, S. H.; Wittenberger, S. J. Development of a Scalable Synthesis of Dipeptidyl Peptidase-4 Inhibitor ABT-279. *Org. Process Res. Dev.* **2009**, *13* (6), 1145–1155. https://doi.org/10.1021/op900197r.

(29) Lwowski, W. Acyl Azides and Nitrenes. In *Azides and Nitrenes Reactivity and Utility*; Scriven, E. F. V., Ed.; Indianapolis, 1984; pp 205–242.

(30) Schmidt, K. F. Über Den Imin-Rest. *Ber.* **1924**, *57* (4), 704–706.

(31) Hanschen, F. S.; Lamy, E.; Schreiner, M.; Rohn, S. Reactivity and Stability of Glucosinolates and Their Breakdown Products in Foods Angewandte. *Angew. Chem. Int. Ed.* **2014**, *53*, 11430–11450. https://doi.org/10.1002/anie.201402639.

(32) Ľ. Drobnica and J. Augustin. Reactions of Isothiocyanates with Amino Acids, Peptides and Proteins. II. Kinetics of the Reaction of Aromatic Isothiocyanates with Amino Acids and Cyclization of the Addition Products. *Collect. Czech. Chem. Commun.* **1965**, *30* (4), 1221–1228.

(33) Matthiessen, J. N.; Kirkegaard, J. A. Biofumigation and Enhanced Biodegradation: Opportunity and Challenge in Soilborne Pest and Disease Management. *CRC. Crit. Rev. Plant Sci.* **2006**, *25* (3), 235–265. https://doi.org/10.1080/07352680600611543.

(34) Züst, T.; Agrawal, A. A. Trade-Offs Between Plant Growth and Defense Against Insect Herbivory: An Emerging Mechanistic Synthesis. *Annu. Rev. Plant Biol.* **2017**, *68* (1), 513–534. https://doi.org/10.1146/annurev-arplant-042916-040856.

(35) Kopriva, S.; Gigolashvili, T. Glucosinolate Synthesis in the Context of Plant Metabolism. In *Advances in Botanical Research*; Elsevier Ltd, 2016; Vol. 80, pp 99–124. https://doi.org/10.1016/bs.abr.2016.07.002.

(36) Bekaert, M.; Edger, P. P.; Hudson, C. M.; Pires, J. C.; Conant, G. C. Metabolic and Evolutionary Costs of Herbivory Defense: Systems Biology of Glucosinolate Synthesis. *New Phytol.* **2012**, *196*.

(37) Kuchernig, J. C.; Burow, M.; Wittstock, U. Evolution of Specifier Proteins in Glucosinolate-Containing Plants. *BMC Evol. Biol.* **2012**, *12* (127), 1–14.

(38) Eisenschmidt-Bönn, D.; Schneegans, N.; Backenköhler, A.; Wittstock, U.; Brandt, W. Structural Diversification during Glucosinolate Breakdown: Mechanisms of Thiocyanate, Epithionitrile and Simple Nitrile Formation. *Plant J.* **2019**, *99* (2), 329–343. https://doi.org/10.1111/tpj.14327.

(39) Witzel, K.; Abu Risha, M.; Albers, P.; Börnke, F.; Hanschen, F. S. Identification and Characterization of Three Epithiospecifier Protein Isoforms in *Brassica Oleracea*. *Front. Plant Sci.* **2019**, *10* (December), 1–14. https://doi.org/10.3389/fpls.2019.01552.

(40) Jeschke, V.; Gershenzon, J.; Vassão, D. G. Insect Detoxification of Glucosinolates and Their Hydrolysis Products. In *Advances in Botanical Research*; 2016; Vol. 80, pp 199–245. https://doi.org/10.1016/bs.abr.2016.06.003.

(41) Backenköhler, A.; Eisenschmidt, D.; Schneegans, N.; Strieker, M.; Brandt, W.; Witt. Iron Is a Centrally Bound Cofactor of Specifier Proteins Involved in Glucosinolate Breakdown. *PLoS One* **2018**, *13* (11), 1–18.

(42) Mocniak, L. E.; Elkin, K.; Bollinger, J. M. Lifetimes of the Aglycone Substrates of Specifier Proteins, the Autonomous Iron Enzymes That Dictate the Products of the Glucosinolate-Myrosinase Defense System in Brassica Plants. *Biochemistry* **2020**, *59*, 2432–2441. https://doi.org/10.1021/acs.biochem.0c00358.

(43) Kissen, R.; Bones, A. M. Nitrile-Specifier Proteins Involved in Glucosinolate

Hydrolysis in *Arabidopsis Thaliana*. *J. Biol. Chem.* **2009**, *284* (18), 12057–12070. https://doi.org/10.1074/jbc.M807500200.

(44) Burow, M.; Bergner, A.; Gershenzon, J.; Wittstock, U. Glucosinolate Hydrolysis in Lepidium Sativum — Identification of the Thiocyanate-Forming Protein. *Plant Mol Biol* **2007**, 49–61. https://doi.org/10.1007/s11103-006-9071-5.

(45) Kuchernig, J.; Backenköhler, A.; Lübbecke, M.; Burow, M.; Wittstock, U. A Thiocyanate-Forming Protein Generates Multiple Products upon Allylglucosinolate Breakdown in Thlaspi Arvense. *Phytochemistry* **2011**, *72*, 1699–1709. https://doi.org/10.1016/j.phytochem.2011.06.013.

(46) Angelino, D.; Jeffery, E. Glucosinolate Hydrolysis and Bioavailability of Resulting Isothiocyanates: Focus on Glucoraphanin. *J. Funct. Foods* **2014**, *7*, 67–76.

(47) Gumz, F.; Krausze, J.; Eisenschmidt, D.; Backenko, A.; Barleben, L.; Brandt, W.; Wittstock, U.; Backenköhler, A.; Barleben, L.; Brandt, W.; Wittstock, U. The Crystal Structure of the Thiocyanate-Forming Protein from *Thlaspi Arvense*, a Kelch Protein Involved in Glucosinolate Breakdown. *Plant Mol. Biol.* **2015**, *89* (1–2), 67–81. https://doi.org/10.1007/s11103-015-0351-9.

(48) Zhang, W.; Wang, W.; Liu, Z.; Xie, Y.; Wang, H.; Mu, Y.; Huang, Y. Crystal Structure of the Epithiospecifier Protein, ESP from *Arabidopsis Thaliana* Provides Insights into Its Product Specificity. *Biochem. Biophys. Res. Commun.* **2016**, *478* (2), 746–751. https://doi.org/10.1016/j.bbrc.2016.08.019.

(49) Augustine, R.; Bisht, N. C. Regulation of Glucosinolate Metabolism: From Model Plant *Arabidopsis Thaliana* to *Brassica* Crops. In *Glucosinolates. Reference Series in Phytochemistry*; Mérillon, J. M., K., R., Eds.; Springer, Cham: Switzerland, 2017; pp

163–199. https://doi.org/10.1007/978-3-319-25462-3.

(50) Chhajed, S.; Misra, B. B.; Tello, N.; Chen, S. Chemodiversity of the Glucosinolate-Myrosinase System at the Single Cell Type Resolution. *Front. Plant Sci.* **2019**, *10*, 1–11. https://doi.org/10.3389/fpls.2019.00618.

(51) Wittstock, U.; Burow, M. Glucosinolate Breakdown in Arabidopsis: Mechanism, Regulation and Biological Significance. *Arab. B.* **2010**, *8*, 1–14. https://doi.org/10.1199/tab.0134.

(52) Zhang, W.; Zhou, Y.; Wang, K.; Dong, Y.; Wang, W.; Feng, Y. Crystal Structure of the Nitrile-Specifier Protein NSP1 from *Arabidopsis Thaliana*. *Biochem. Biophys. Res. Commun.* **2017**, *488* (1), 147–152. https://doi.org/10.1016/j.bbrc.2017.05.027.

(53) Brandt, W.; Backenkohler, A.; Schulze, E.; Plock, A.; Herberg, T. Molecular Models and Mutational Analyses of Plant Specifier Proteins Suggest Active Site Residues and Reaction Mechanism. *Plant Mol. Biol.* **2014**, *84*, 173–188. https://doi.org/10.1007/s11103-013-0126-0.

(54) Román, J.; González, D.; Inostroza, M.; Mahn, A. Molecular Modeling of Epithiospecifier and Nitrile-Specifier Proteins of Broccoli and Their Interaction with Aglycones. *Molecules* **2020**, *25*, 1–16. https://doi.org/10.3390/molecules25040772.

(55) Wittstock, U.; Burow, M. Tipping the Scales - Specifier Proteins in Glucosinolate Hydrolysis. *IUBMB Life* **2007**, *59* (12), 744–751. https://doi.org/10.1080/15216540701736277.

(56) Brocker, E.; Benn, M. H. The Intramolecular Formation of Epithioalkanenitriles from Alkenylglucosinolates by Crambe Abyssinica Seed Flour. *Phyt* **1983**, *22* (3).

(57) Burow, M.; Markert, J.; Gershenzon, J.; Wittstock, U. Comparative Biochemical

Characterization of Nitrile-Forming Proteins from Plants and Insects That Alter Myrosinase-Catalysed Hydrolysis of Glucosinolates. *FEBS* **2006**, *273*, 2432–2446. https://doi.org/10.1111/j.1742-4658.2006.05252.x.

(58) Luthy, J.; Benn, M. H. Thiocyanate Formation from Glucosinolates: A Study of the Autolysis of Allylglucosinolate in *Thlaspi Arvense* L. Seed Flour Extracts. *Can. J. Biochem.* **1977**, *55* (10), 1028–1031. https://doi.org/10.1139/o77-153.

(59) Björkman, M.; Klingen, I.; Birch, A. N. E.; Bones, A. M.; Bruce, T. J. A.; Johansen, T. J.; Meadow, R.; Mølmann, J.; Seljåsen, R.; Smart, L. E.; Stewart, D. Phytochemicals of Brassicaceae in Plant Protection and Human Health - Influences of Climate, Environment and Agronomic Practice. *Phytochemistry* **2011**, *72* (7), 538–556. https://doi.org/10.1016/j.phytochem.2011.01.014.

(60) Mitsiogianni, M.; Koutsidis, G.; Mavroudis, N.; Trafalis, D. T.; Botaitis, S.; Franco, R.; Zoumpourlis, V.; Amery, T.; Galanis, A.; Pappa, A.; Panayiotidis, M. I. The Role of Isothiocyanates as Cancer Chemo-Preventive, Chemo-Therapeutic and Anti-Melanoma Agents. *Antioxidants* **2019**, *8* (4), 1–35. https://doi.org/10.3390/antiox8040106.

(61) Houghton, C. A. Sulforaphane: Its "Coming of Age" as a Clinically Relevant Nutraceutical in the Prevention and Treatment of Chronic Disease. *Oxid. Med. Cell. Longev.* **2019**, *2019*, 1–27. https://doi.org/10.1155/2019/2716870.

(62) Fofaria, N. M.; Ranjan, A.; Kim, S.-H.; Srivastava, S. K. Mechanisms of the Anticancer Effects of Isothiocyanates. In *The Enzymes*; Elsevier Inc., 2015; Vol. 37, pp 111–137. https://doi.org/10.1016/bs.enz.2015.06.001.

(63) Banerjee, S.; Paruthy, S. Therapeutic Paradigm Underscoring Glucosinolate

Sulforaphane in Chemo- and Radiosensitization of Cancer: Preclinical and Clinical Perspective. In *Glucosinolates*; Mérillon, J.-M., Ramawat, K. G., Eds.; Springer International Publishing, 2017; pp 339–379. https://doi.org/10.1007/978-3-319-25462-3.

(64) Navarro, S.; Li, F.; Lampe, J. W. Mechanisms of Action of Isothiocyanates in Cancer Chemoprevention: An Update. *Food Funct.* **2011**, *2* (10), 579–587. https://doi.org/10.1039/c1fo10114e.Mechanisms.

(65) Bahmid, N. A.; Pepping, L.; Dekker, M.; Fogliano, V.; Heising, J. Using Particle Size and Fat Content to Control the Release of Allyl Isothiocyanate from Ground Mustard Seeds for Its Application in Antimicrobial Packaging. *Food Chem.* **2020**, *308* (September 2019), 1–8. https://doi.org/10.1016/j.foodchem.2019.125573.

(66) Kurek, M.; Laridon, Y.; Torrieri, E.; Guillard, V.; Pant, A.; Stramm, C.; Gontard, N.; Guillaume, C. A Mathematical Model for Tailoring Antimicrobial Packaging Material Containing Encapsulated Volatile Compounds. *Innov. Food Sci. Emerg. Technol.* **2017**, *42*, 64–72.

(67) Romeo, L.; Iori, R.; Rollin, P.; Bramanti, P.; Mazzon, E. Isothiocyanates: An Overview of Their Antimicrobial Activity against Human Infections. *Molecules* **2018**, *23* (3), 1–18. https://doi.org/10.3390/molecules23030624.

(68) Ngala, B. M.; Haydock, P. P.; Woods, S.; Back, M. A. Biofumigation with Brassica Juncea, Raphanus Sativus and Eruca Sativa for the Management of Field Populations of the Potato Cyst Nematode Globodera Pallida. *Pest Manag. Sci.* **2015**, *71* (5), 759–769. https://doi.org/10.1002/ps.3849.

(69) Lord, J. S.; Lazzeri, L.; Atkinson, H. J.; Urwin, P. E. Biofumigation for Control of

Pale Potato Cyst Nematodes: Activity of Brassica Leaf Extracts and Green Manures on *Globodera Pallida* in Vitro and in Soil. *J. Agric. Food Chem.* **2011**, *59* (14), 7882–7890. https://doi.org/10.1021/jf200925k.

(70) Szczygłowska, M.; Piekarska, A.; Konieczka, P.; Namieśnik, J. Use of Brassica Plants in the Phytoremediation and Biofumigation Processes. *Int. J. Mol. Sci.* **2011**, *12* (11), 7760–7771. https://doi.org/10.3390/ijms12117760.

(71) Bangarwa, S. K.; Norsworthy, J. K.; Mattice, J. D.; Gbur, E. E. Glucosinolate and Isothiocyanate Production from Brassicaceae Cover Crops in a Plasticulture Production System. *Weed Sci.* **2011**, *59* (2), 247–254. https://doi.org/10.1614/ws-d-10-00137.1.

(72) Linn, A.; Steindal, H.; Rødven, R.; Hansen, E.; Mølmann, J. Effects of Photoperiod, Growth Temperature and Cold Acclimatisation on Glucosinolates, Sugars and Fatty Acids in Kale. *Food Chem.* **2015**, *174*, 44–51.

(73) Almuziny, M.; Decker, C.; Wang, D.; Gerard, P. Nutrient Supply and Simulated Herbivory Differentially Alter the Metabolite Pools and the Efficacy of the Glucosinolate-Based Defense System in Brassica Species. *J Chem Ecol* **2017**, *43*, 129–142. https://doi.org/10.1007/s10886-016-0811-y.

(74) Mølmann, J. A. B.; Steindal, A. L. H.; Bengtsson, G. B.; Seljåsen, R.; Lea, P.; Skaret, J.; Johansen, T. J. Effects of Temperature and Photoperiod on Sensory Quality and Contents of Glucosinolates, Flavonols and Vitamin C in Broccoli Florets. *Food Chem.* **2015**, *172*, 47–55.

(75) Oliviero, T.; Verkerk, R.; Dekker, M. Effect of Water Content and Temperature on Glucosinolate Degradation Kinetics in Broccoli (Brassica Oleracea Var. Italica). *Food*

Chem. **2012**, *132*, 2037–2045. https://doi.org/10.1016/j.foodchem.2011.12.045.

(76) dos Santos, C. A.; Abboud, A. C. de S.; Do Carmo, M. G. F. Biofumigation with Species of the Brassicaceae Family: A Review. *Cienc. Rural* **2021**, *51* (1), 1–17. https://doi.org/10.1590/0103-8478cr2020040.

(77) Hall, M. H.; Jung, G. A. Use of Brassica Crops to Extend the Grazing Season. *Agron. Facts 33*, **2008**, 1–4.

(78) Mawson, R.; Heaney, R. K.; Zdunczyk, Z.; Kozlowska, H. Rapeseed Meal-glucosinolates and Their Antinutritional Effects Part 5. Animal Reproduction. *Nahrung* **1994**, *6*, 588–598. https://doi.org/10.1002/food.19930370405.

(79) Ravichandiran, S.; Kusumakar, S.; Dutta, N.; Pattanaik, A. K.; Chauhan, J. .; Agnihotri, A. Comparative Assessment of Soybean Meal with High and Low Glucosinolate Rapeseed–Mustard Cake as Protein Supplement on Performance of Growing Crossbred Calves. *J. Sci. Food Agric.* **2008**, *88*, 832–838. https://doi.org/10.1002/jsfa.3159.

(80) Gaitan, E. Goitrogens In Food And Water. *Annu. Rev. Nutr.* **1990**, *10* (1), 21–39. https://doi.org/10.1146/annurev.nutr.10.1.21.

(81) Engelking, L. R. Iodine and Cobalt. In *Textbook of Veterinary Physiological Chemistry*; Elsevier Inc.: Waltham, 2015; pp 325–329.

(82) Tripathi, M. K.; AS, M. Prospects and Problems of Dietary Glucosinolates in Animal Feeding. *Adv. Dairy Res.* **2017**, *05* (03), 5–8. https://doi.org/10.4172/2329-888x.1000180.

(83) Durge, S. M.; Tripathi, M. K.; Tripathi, P.; Dutta, N.; Rout, P. K.; Chaudhary, U. B. Intake, Nutrient Utilization, Rumen Fermentation, Microbial Hydrolytic Enzymes and

Hemato-Biochemical Attributes of Lactating Goats Fed Concentrate Containing <i>Brassica Juncea<i\> Oil Meal. *Small Rumin. Res.* **2014**, *121*, 300–307. https://doi.org/10.1016/j.smallrumres.2014.09.006.

(84) Laarveld, B.; Brockman, R. P.; Christensen, D. A. The Goitrogenic Potential of Tower and Midas Rapeseed Meal in Dairy Cows Determined by Thyrotropin-Releasing Hormone Test. *Can. J. Anim. Sci.* **1980**, *61*, 141–149.

(85) Mawson, R.; Heaney, R. K.; Zdunczyk, Z.; Kozlowska, H. Rapeseed Meal-Glucosinolates and Their Antinutritional Effects Part 3. Animal Growth and Performance. *Nahrung* **1994**, *38* (2), 167–177.

(86) Bush, R. .; Nicholson, W. G.; Macintyre, T. M.; McQueen, R. E. A Comparison of Candle and Tower Rapeseed Meals in Lamb, Sheep and Beef Steer Rations. *Can. J. Anim. Sci.* **1978**, *58* (3), 369–376. https://doi.org/10.4141/cjas78-048.

(87) Tripathi, M. K.; Mishra, A. S.; R., P.; Mondal, D.; Jakhmola, R. C. Effect of Graded Levels of High Glucosinolate Mustard (Brassica Juncea Meal Inclusion on Nutrient Utilization, Growth, Performance, Organ Weight, and Carcass Compositiom of Growing Rabbits. *World Rabbit Sci.* **2003**, *11*, 211–226.

(88) Glencross, B.; Hawkins, W.; Curnow, J. Nutritional Assessment of Australian Canola Meals. II. Evaluation of the Influence of the Canola Oil Extraction Method on the Protein Value of Canola Meals Fed to the Red Seabream (<i>Pagrus Auratus<i\>, Paulin). *Aquac. Res.* **2004**, *35* (1), 25–34. https://doi.org/10.1111/j.1365-2109.2004.00975.x.

(89) Opalka, M.; Dusza, L.; Koziorowski, M.; Staszkiewicz, J.; Lipiñski, K.; Tywoñczuk, J. Effect of Long-Term Feeding with Graded Levels of Low Glucosinolate Rapeseed

Meal on Endocrine Status of Gilts and Their Piglets. *Livest. Prod. Sci.* **2001**, *69* (3), 233–243. https://doi.org/10.1016/S0301-6226(01)00171-3.

(90) Bjerg, B.; Eggum, B. O.; Jacobsen, I.; Otte, J.; Sørensen, H. Antinutritional and Toxic Effects in Rats of Individual Glucosinolates (± Myrosinases) Added to a Standard Diet (2). *J. Anim. Physiol. Anim. Nutr. (Berl).* **1989**, *61*, 227–244. https://doi.org/10.1111/j.1439-0396.1989.tb00105.x.

(91) Vieites-Outes, C.; López-Hernández, J.; Lage-Yusty, M. A. Modification of Glucosinolates in Turnip Greens (*Brassica Rapa* Subsp. *Rapa* L.) Subjected to Culinaryheat Processes. *CYTA - J. Food* **2016**, *14* (4), 536–540. https://doi.org/10.1080/19476337.2016.1154609.

(92) Yang, B.; Quiros, C. F. Survey of Glucosinolate Variation in Leaves of *Brassica Rapa* Crops. *Genet. Resour. Crop Evol.* **2010**, *57* (7), 1079–1089. https://doi.org/10.1007/s10722-010-9549-5.

(93) Mohn, T.; Cutting, B.; Ernst, B.; Hamburger, M. Extraction and Analysis of Intact Glucosinolates — A Validated Pressurized Liquid Extraction/Liquid Chromatography – Mass Spectrometry Protocol for Isatis Tinctoria, and Qualitative Analysis of Other Cruciferous Plants. *J. Chromatogr. A* **2007**, *1166*, 142–151. https://doi.org/10.1016/j.chroma.2007.08.028.

(94) Holst, B.; Williamson, G. A Critical Review of the Bioavailability of Glucosinolates and Related Compounds. *Nat. Prod. Rep.* **2004**, *21* (3), 425–447. https://doi.org/10.1039/b204039p.

(95) Fahey, J. W.; Zalcmann, A. T.; Talalay, P. The Chemical Diversity and Distribution of Glucosinolates and Isothiocyanates among Plants. *Phytochemistry* **2001**, *56*, 5–51.

https://doi.org/10.1016/s0031-9422(00)00316-2.

(96) Agerbirk, N.; Olsen, C. E. Glucosinolate Structures in Evolution. *Phytochemistry* **2012**, *77*, 16–45. https://doi.org/10.1016/j.phytochem.2012.02.005.

(97) Hahn, C.; Müller, A.; Kuhnert, N.; Albach, D. Diversity of Kale (*Brassica Oleracea* Var. *Sabellica*): Glucosinolate Content and Phylogenetic Relationships. *J. Agric. Food Chem.* **2016**, *64* (16), 3215–3225. https://doi.org/10.1021/acs.jafc.6b01000.

(98) Olsen, E. C.; Huang, X.; Hansen, C. I. C.; Cipollini, D.; Ørgaard, M.; Matthes, A.; Geu-flores, F.; Koch, M. A.; Agerbirk, N. Glucosinolate Diversity within a Phylogenetic Framework of the Tribe Cardamineae (Brassicaceae) Unraveled with HPLC-MS/MS and NMR-Based Analytical Distinction of 70 Desulfoglucosinolates. *Phytochemistry* **2016**, *132*, 33–56. https://doi.org/10.1016/j.phytochem.2016.09.013.

(99) Shahidi, F.; Daun, J. K.; DeClercq, D. Glucosinolates in Brassica Oilseeds: Processing Effects and Extraction. In *Antinutrients and Phytochemicals in Food*; Shahidi, F., Ed.; American Chemical Society, 1997; pp 152–170. https://doi.org/10.1021/bk-1997-0662.ch009.

(100) Tripathi, M. K.; Mishra, A. S. Glucosinolates in Animal Nutrition: A Review. *Anim. Feed Sci. Technol.* **2007**, *132* (1–2), 1–27. https://doi.org/10.1016/j.anifeedsci.2006.03.003.

(101) Fales, S. L.; Gustine, D. L.; Bosworth, S. C.; Hoover, R. J. Concentrations of Glucosinolates and S-Methylcysteine Sulfoxide in Ensiled Rape (*Brassica Napus* L.). *J. Dairy Sci.* **1987**, *70* (11), 2402–2405. https://doi.org/10.3168/jds.S0022-0302(87)80301-6.

(102) Larkin, R. P.; Griffin, T. S. Control of Soilborne Potato Diseases Using Brassica

Green Manures. *Crop Prot.* **2007**, *26* (7), 1067–1077. https://doi.org/10.1016/j.cropro.2006.10.004.

(103) Gimsing, A. L.; Kirkegaard, J. A. Glucosinolates and Biofumigation: Fate of Glucosinolates and Their Hydrolysis Products in Soil. *Phytochem. Rev.* **2009**, *8* (1), 299–310. https://doi.org/10.1007/s11101-008-9105-5.

(104) Hanschen, F. S.; Yim, B.; Winkelmann, T.; Smalla, K. Degradation of Biofumigant Isothiocyanates and Allyl Glucosinolate in Soil and Their Effects on the Microbial Community Composition. *PLoS One* **2015**, 1–18. https://doi.org/10.1371/journal.pone.0132931.

(105) Sun, X. Z.; Waghorn, G. C.; Hoskin, S. O.; Harrison, S. J.; Muetzel, S.; Pacheco, D. Methane Emissions from Sheep Fed Fresh Brassicas (*Brassica* Spp.) Compared to Perennial Ryegrass (*Lolium Perenne*). *Anim. Feed Sci. Technol.* **2012**, *176*, 107–116. https://doi.org/10.1016/j.anifeedsci.2012.07.013.

(106) Aires, A.; Carvalho, R. Rapid Separation of Indole Glucosinolates in Roots of Chinese Cabbage (*Brassica Rapa* Subsp. Pekinensis) by High-Performance Liquid Chromatography with Diode Array Detection. *Int. J. Anal. Chem.* **2017**, 1–7. https://doi.org/https://doi.org/10.1155/2017/5125329.

(107) Grosser, K.; van Dam, N. M. Van. A Straightforward Method for Glucosinolate Extraction and Analysis with High-Pressure Liquid Chromatography (HPLC). *Jove* **2017**, 1–9. https://doi.org/10.3791/55425.

(108) Cataldi, T. R. I.; Rubino, A.; Lelario, F.; Bufo, S. A. Naturally Occurring Glucosinolates in Plant Extracts of Rocket Salad (*Eruca Sativa* L.) Identified by Liquid Chromatography Coupled with Negative Ion Electrospray Ionization and

Quadrupole Ion-Trap Mass Spectrometry. *RCM* **2007**, *21*, 2374–2388. https://doi.org/10.1002/rcm.

(109) Thomas, M.; Badr, A.; Desjardins, Y.; Gosselin, A.; Angers, P. Characterization of Industrial Broccoli Discards (*Brassica Oleracea* Var. Italica) for Their Glucosinolate, Polyphenol and Flavonoid Contents Using UPLC MS/MS and Spectrophotometric Methods. *Food Chem.* **2018**, *245*, 1204–1211. https://doi.org/10.1016/j.foodchem.2017.11.021.

(110) Tolra, R.; Alonso, R.; Poschenrieder, C.; Barcelo, J. Determination of Glucosinolates in Rapeseed and *Thlaspi Caerulescens* Plants by Liquid Chromatography – Atmospheric Pressure Chemical Ionization Mass Spectrometry. *J. Chromatogr. A* **2000**, *889*, 75–81. https://doi.org/10.1016/s0021-9673(00)00373-3.

(111) Tian, Q.; Rosselot, R. A.; Schwartz, S. J. Quantitative Determination of Intact Glucosinolates in Broccoli, Broccoli Sprouts, Brussels Sprouts, and Cauliflower by High-Performance Liquid Chromatography – Electrospray Ionization – Tandem Mass Spectrometry. *Anal. Biochem.* **2005**, *343*, 93–99. https://doi.org/10.1016/j.ab.2005.04.045.

(112) Doheny-Adams, T.; Redeker, K.; Kittipol, V.; Bancroft, I.; Hartley, S. E. Development of an Efficient Glucosinolate Extraction Method. *Plant Methods* **2017**, 1–14. https://doi.org/10.1186/s13007-017-0164-8.

(113) Rochfort, S. J.; Trenerry, V. C.; Imsic, M.; Panozzo, J.; Jones, R. Class Targeted Metabolomics: ESI Ion Trap Screening Methods for Glucosinolates Based on MS^n Fragmentation. *Phytochemistry* **2008**, *69*, 1671–1679. https://doi.org/10.1016/j.phytochem.2008.02.010.

(114) Ares, A. M.; Nozal, M. J.; Bernal, J. L.; Bernal, J. Optimized Extraction, Separation and Quantification of Twelve Intact Glucosinolates in Broccoli Leaves. *Food Chem.* **2014**, *152*, 66–74. https://doi.org/http://dx.doi.org/10.1016/j.foodchem.2013.11.125.

(115) Ares, A. M.; Nozal, M. J.; Bernal, J. Development and Validation of a Liquid Chromatography-Tandem Mass Spectrometry Method to Determine Intact Glucosinolates in Bee Pollen. *J. Chromatogr. B* **2015**, *1000*, 49–56. https://doi.org/10.1016/j.jchromb.2015.07.017.

(116) Bianco, G.; Pascale, R.; Lelario, F.; Bufo, S. A.; Cataldi, T. R. I. *The Investigation of Glucosinolates by Mass Spectrometry*; 2016. https://doi.org/10.1007/978-3-319-26479-0_12-1.

(117) Bialecki, J. B.; Ruzicka, J.; Weisbecker, C. S.; Attygalle, A. B. Collision-Induced Dissociation Mass Spectra of Glucosinolate Anions. *J. Mass Spectrom.* **2010**, *2009*, 272–283. https://doi.org/10.1002/jms.1711.

(118) Bianco, G.; Agerbirk, N.; Losito, I.; Cataldi, T. R. I. Acylated Glucosinolates with Diverse Acyl Groups Investigated by High Resolution Mass Spectrometry and Infrared Multiphoton Dissociation. *Phytochemistry* **2014**, *100*, 92–102. https://doi.org/10.1016/j.phytochem.2014.01.010.

(119) Stevens, J. F.; Reed, R. L.; Alber, S.; Pritchett, L.; Machado, S. Herbicidal Activity of Glucosinolate Degradation Products in Fermented Meadowfoam (*Limnanthes Alba*) Seed Meal. *J. Agric. Food Chem.* **2009**, *57* (5), 1821–1826. https://doi.org/10.1021/jf8033732.

(120) Padilla, G.; Cartea, M. E.; Velasco, P.; de Haro, A.; Ordas, A. Variation of Glucosinolates in Vegetable Crops of *Brassica Rapa*. *Phytochemistry* **2007**, No. 68,

536–545. https://doi.org/10.1016/j.phytochem.2006.11.017.

(121) Schnitzler, W. H. Studies for Growth Adaptation and Identification of Glucosinolates in Various Chinese Brassicas, Technische Universität München, 1999. https://doi.org/10.13140/RG.2.1.2283.2406.

(122) El-Beltagi, H.; Mohamed, A. A. Variations in Fatty Acid Composition, Glucosinolate Profile and Some Phytochemical Contents in Selected Oil Seed Rape (*Brassica Napus* L.) Cultivars. *Grasas y Aceites* **2010**, *61* (2), 143–150. https://doi.org/10.3989/gya.087009.

(123) Al-Gendy, A. A.; El-gindi, O. D.; Hafez, A. S.; Ateya, A. M. Glucosinolates, Volatile Constituents and Biological Activities of *Erysimum Corinthium* Boiss. (Brassicaceae). *Food Chem.* **2010**, *118* (3), 519–524. https://doi.org/10.1016/j.foodchem.2009.05.009.

(124) Theodoridou, K.; Yu, P. Effect of Processing Conditions on the Nutritive Value of Canola Meal and Presscake. Comparison of the Yellow and Brown-Seeded Canola Meal with the Brown-Seeded Canola Presscake. *J. Sci. Food Agric.* **2013**, *93* (8), 1986–1995. https://doi.org/10.1002/jsfa.6004.

(125) Raikos, V.; Neacsu, M.; Duthie, G.; Nicol, F.; Reid, M.; Cantlay, L. L.; Ranawana, V. Proteomic and Glucosinolate Profiling of Rapeseed Isolates from Meals Produced by Different Oil Extraction Processes. *J. Food Process. Preserv.* **2017**, *41* (4), 1–8. https://doi.org/10.1111/jfpp.13060.

(126) Bohinc, T.; Trdan, S. Environmental Factors Affecting the Glucosinolate Content in Brassicaceae. *J. Food, Agric. Environ.* **2012**, *10* (2), 357–360. https://doi.org/10.1021/jf0624897.

(127) Font, R.; Del Río-Celestino, M.; Rosa, E.; Aires, A.; De Haro-Bailón, A.

Glucosinolate Assessment in *Brassica Oleracea* Leaves by near-Infrared Spectroscopy. *J. Agric. Sci.* **2005**, *143* (1), 65–73. https://doi.org/10.1017/S0021859605004806.

(128) Velasco, P.; Soengas, P.; Vilar, M.; Cartea, M. E.; Del Rio, M. Comparison of Glucosinolate Profiles in Leaf and Seed Tissues of Different *Brassica Napus* Crops. *J. Am. Soc. Hortic. Sci.* **2008**, *133* (4), 551–558. https://doi.org/10.21273/jashs.133.4.551.

(129) Forss, D. A.; Barry, T. N. Observations on Nitrile Production during Autolysis of Kale and Swedes, and Their Stability during Incubation with Rumen Fluid. *J. Sci. Food Agric.* **1983**, *34* (10), 1077–1084. https://doi.org/10.1002/jsfa.2740341007.

(130) Barry, T. N. The Feeding Value of Forage Brassica Plants for Grazing Ruminant Livestock. *Anim. Feed Sci. Technol.* **2013**, *181* (1–4), 15–25. https://doi.org/10.1016/j.anifeedsci.2013.01.012.

(131) Chavadej, S.; Brisson, N.; McNeil, J. N.; De Luca, V. Redirection of Tryptophan Leads to Production of Low Indole Glucosinolate Canola. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91* (6), 2166–2170. https://doi.org/10.1073/pnas.91.6.2166.

(132) Potter, M. J.; Davies, K.; Rathjen, A. J. Suppressive Impact of Glucosinolates in *Brassica* Vegetative Tissues on Root Lesion Nematode *Pratylenchus Neglectus*. *J. Chem. Ecol.* **1998**, *24* (1), 67–80. https://doi.org/10.1023/A:1022336812240.

(133) Cartea, M. E.; Velasco, P. Glucosinolates in Brassica Foods: Bioavailability in Food and Significance for Human Health. *Phytochem. Rev.* **2008**, *7*, 213–229. https://doi.org/10.1007/s11101-007-9072-2.

(134) Yasumoto, S.; Matsuzaki, M.; Hirokane, H.; Okada, K. Glucosinolate Content in

Rapeseed in Relation to Suppression of Subsequent Crop. *Plant Prod. Sci.* **2010**, *13* (2), 150–155. https://doi.org/10.1626/pps.13.150.

(135) Müller, R.; de Vos, M.; Sun, J. Y.; Sønderby, I. E.; Halkier, B. A.; Wittstock, U.; Jander, G. Differential Effects of Indole and Aliphatic Glucosinolates on Lepidopteran Herbivores. *J. Chem. Ecol.* **2010**, *36* (8), 905–913. https://doi.org/10.1007/s10886-010-9825-z.

(136) Beauchemin, K. A.; McGinn, S. M. Methane Emissions from Beef Cattle: Effects of Monensin, Sunflower Oil, Enzymes, Yeast, and Fumaric Acid. *J. Anim. Sci.* **2004**, *82* (11), 3346–3356. https://doi.org/10.2527/2004.82113346x.

(137) Watkins, P. J.; Frank, D.; Singh, T. K.; Young, O. A.; Warner, R. D. Sheepmeat Flavor and the Effect of Different Feeding Systems: A Review. *J. Agric. Food Chem.* **2013**, *61* (15), 3561–3579. https://doi.org/10.1021/jf303768e.

(138) Bollinger, J. M., J.; Tong, W. H.; Ravi, N.; Huynh, B. H.; Edmondson, D. E.; Stubbe, J. Use of Rapid Kinetics Methods to Study the Assembly of the Diferric-Tyrosyl Radical Cofactor of *E. Coli* Ribonucleotide Reductase. *Meth. Enzym.* **1995**, *258*, 278–303.

(139) Krebs, C.; Price, J. C.; Baldwin, J.; Saleh, L.; Green, M. T.; Bollinger, J. M., J. Rapid Freeze-Quench 57Fe Mössbauer Spectroscopy: Monitoring Changes of an Iron-Containing Active Site during a Biochemical Reaction. *Inorg. Chem.* **2005**, *44* (4), 742–757. https://doi.org/10.1021/ic048523l.

(140) Bollinger, J. M., J.; Krebs, C. Stalking Intermediates in Oxygen Activation by Iron Enzymes: Motivation and Method. *J. Inorg. Biochem.* **2006**, *100* (4), 586–605. https://doi.org/10.1016/j.jinorgbio.2006.01.022.

(141) Liebminger, E.; Grass, J.; Jez, J.; Neumann, L.; Altmann, F.; Strasser, R. Myrosinases TGG1 and TGG2 from Arabidopsis Thaliana Contain Exclusively Oligomannosidic N-Glycans. *Phytochemistry* **2012**, *84*, 24–30. https://doi.org/10.1016/j.phytochem.2012.08.023.

(142) Williams, D.; Critchley, C.; Pun, S.; Chaliha, M.; O'Hare, T. J. Key Role of Fe(II) in Epithiospecifier Protein Activity. *J. Agric. Food Chem.* **2010**, *58*, 8512–8521. https://doi.org/10.1021/jf904532n.

(143) Eylen, D. Van; Hendrickx, M.; Loey, A. Van. Temperature and Pressure Stability of Mustard Seed (*Sinapis Alba* L.) Myrosinase. *Food Chem.* **2006**, *97*, 263–271. https://doi.org/10.1016/j.foodchem.2005.03.046.

(144) Okunade, O. A.; Ghawi, S. K.; Methven, L.; Niranjan, K. Thermal and Pressure Stability of Myrosinase Enzymes from Black Mustard (*Brassica Nigra* L. W.D.J. Koch. Var. *Nigra*), Brown Mustard (*Sinapsis Alba* L. Subsp. *Maire*) Seeds. *Food Chem.* **2015**, *187*, 485–490.

(145) Spencer, G. F.; Daxenbichler, M. E. Gas Chromatography-Mass Spectrometry of Nitriles, Isothiocyanates and Oxazolidinethiones Derived from Cruciferous Glucosinolates. *J. Sci. Food Agric.* **1980**, *31*, 359–367.

(146) Lu, M.; Hashimoto, K.; Uda, Y. Rat Intestinal Microbiota Digest Desulfosinigrin to Form Allyl Cyanide and 1-Cyano-2,3-Epithiopropane. *Food Res. Int.* **2011**, *44* (4), 1023–1028. https://doi.org/10.1016/j.foodres.2011.03.001.

(147) Matthiessen, J. N.; Kirkegaard, J. A. Biofumigation and Enhanced Biodegradation: Opportunity and Challenge in Soilborne Pest and Disease Management. *CRC. Crit. Rev. Plant Sci.* **2006**, *25*, 235–265. https://doi.org/10.1080/07352680600611543.

(148) Kirkegaard, J. A.; Sarwar, M. Biofumigation Potential of Brassicas: I. Variation in Glucosinolate Profiles of Diverse Field-Grown Brassicas. *Plant Soil* **1998**, *201* (1), 71–89. https://doi.org/10.1023/A:1004364713152.

(149) Birch, A. N. E.; Wynne Griffiths, D.; Hopkins, R. J.; Macfarlane Smith, W. H.; McKinlay, R. G. Glucosinolate Responses of Swede, Kale, Forage and Oilseed Rape to Root Damage by Turnip Root Fly (*Delia Floralis*) Larvae. *J. Sci. Food Agric.* **1992**, *60* (1), 1–9. https://doi.org/10.1002/jsfa.2740600102.

(150) Liou, C. S.; Sirk, S. J.; Diaz, C. A. C.; Klein, A. P.; Fischer, C. R.; Higginbottom, S. K.; Erez, A.; Donia, M. S.; Sonnenburg, J. L.; Sattely, E. S. A Metabolic Pathway for Activation of Dietary Glucosinolates by a Human Gut Symbiont. *Cell* **2020**, *180* (4), 717–728. https://doi.org/10.1016/j.cell.2020.01.023.

(151) Pessina, A.; Thomas, R. M.; Palmieri, S.; Luisi, P. L. An Improved Method for the Purification of Myrosinase and Its Physicochemical Characterization. *Arch. Biochem. Biophys.* **1990**, *280* (2), 383–389. https://doi.org/10.1016/0003-9861(90)90346-Z.

(152) Lanz, N. D.; Blaszczyk, A. J.; McCarthy, E. L.; Wang, B.; Wang, R. X.; Jones, B. S.; Booker, S. J. Enhanced Solubilization of Class B Radical *S*-Adenosylmethionine Methylases by Improved Cobalamin Uptake in Escherichia Coli. *Biochemistry* **2018**, *57* (9), 1475–1490. https://doi.org/10.1021/acs.biochem.7b01205.

(153) Burow, M.; Müller, R.; Gershenzon, J.; Wittstock, U. Altered Glucosinolate Hydrolysis in Genetically Engineered *Arabidopsis Thaliana* and Its Influence on the Larval Development of *Spodoptera Littoralis*. *J. Chem. Ecol.* **2006**, *32* (11), 2333–2349. https://doi.org/10.1007/s10886-006-9149-1.

(154) Radulović, N. S.; Todorovska, M. M.; Zlatković, D. B.; Stojanović, N. M.;

Randjelović, P. J. Two Goitrogenic 1,3-Oxazolidine-2-Thione Derivatives from Brassicales Taxa: Challenging Identification, Occurrence and Immunomodulatory Effects. *Food Chem. Toxicol.* **2017**, *110* (September), 94–108. https://doi.org/10.1016/j.fct.2017.10.005.

(155) Hanschen, F. S.; Brüggemann, N.; Brodehl, A.; Mewis, I.; Schreiner, M.; Rohn, S.; Kroh, L. W. Characterization of Products from the Reaction of Glucosinolate-Derived Isothiocyanates with Cysteine and Lysine Derivatives Formed in Either Model Systems or Broccoli Sprouts. *J. Agric. Food Chem.* **2012**, *60*, 7735–7745. https://doi.org/10.1021/jf301718g.

(156) Vik, D.; Mitarai, N.; Wulff, N.; Halkier, B. A.; Burow, M. Dynamic Modeling of Indole Glucosinolate Hydrolysis and Its Impact on Auxin Signaling. *Front. Plant Sci.* **2018**, *9*. https://doi.org/10.3389/fpls.2018.00550.

(157) Brudenell, A. J. P.; Griffiths, H.; Rossiter, J. T.; Baker, D. A. The Phloem Mobility of Glucosinolates. *J. Exp. Bot.* **1999**, *50* (335), 745–756.

(158) Burow, M.; Rice, M.; Hause, B.; Gershenzon, J.; Wittstock, U. Cell- and Tissue-Specific Localization and Regulation of the Epithiospecifier Protein in *Arabidopsis Thaliana*. *Plant Mol Biol* **2007**, *64*, 173–185. https://doi.org/10.1007/s11103-007-9143-1.

(159) Shi, C. H.; Hu, J. R.; Xie, W.; Yang, Y. T.; Wang, S. L.; Zhang, Y. J. Control of *Bradysia Odoriphaga* (Diptera: Sciaridae) With Allyl Isothiocyanate Under Field and Greenhouse Conditions. *J. Econ. Entomol.* **2017**, *110* (3), 1127–1132. https://doi.org/10.1093/jee/tow303.

(160) Işikber, A. A.; Tunaz, H.; Er, M. K.; Sağlam, Ö. Fumigant Toxicity of Mustard

Essential Oil and Its Main Compound Alone and Combinations with Modified Atmosphere Treatments against *Tribolium Confusum* Du Val., 1863 (Coleoptera: Tenebrionidae)[1]. *Turkiye Entomoloji Derg.* **2019**, *43* (2), 179–187. https://doi.org/10.16970/entoted.533655.

(161) Alan, J.; Renwick, A.; Haribal, M.; Gouinguené, S.; Städler, E. Isothiocyanates Stimulating Oviposition by the Diamondback Moth, *Plutella Xylostella*. *J. Chem. Ecol.* **2006**, *32* (4), 755–766. https://doi.org/10.1007/s10886-006-9036-9.

(162) Grzywacz, D.; Rossbach, A.; Rauf, A.; Russell, D. A.; Srinivasan, R.; Shelton, A. M. Current Control Methods for Diamondback Moth and Other Brassica Insect Pests and the Prospects for Improved Management with Lepidopteran-Resistant Bt Vegetable Brassicas in Asia and Africa. *Crop Prot.* **2010**, *29* (1), 68–79. https://doi.org/10.1016/j.cropro.2009.08.009.

(163) Sarfraz, M.; Dosdall, L. M.; Keddie, B. A. Resistance of Some Cultivated Brassicaceae to Infestations by *Plutella Xylostella* (Lepidoptera: Plutellidae). *J. Econ. Entomol.* **2007**, *100* (1), 215–224. https://doi.org/10.1603/0022-0493(2007)100[215:ROSCBT]2.0.CO;2.

(164) Coapio, G. G.; Cruz-López, L.; Guerenstein, P.; Malo, E. A.; Rojas, J. C. Oviposition Preference and Larval Performance and Behavior of *Trichoplusia Ni* (Lepidoptera: Noctuidae) on Host and Nonhost Plants. *Arthropod. Plant. Interact.* **2018**, *12* (2), 267–276. https://doi.org/10.1007/s11829-017-9566-9.

(165) CLINE, J. D. Spectrophotometric Determination of Hydrogen Sulfide in Natural Waters. *Limnol. Oceanogr.* **1969**, *14* (3), 454–458. https://doi.org/10.4319/lo.1969.14.3.0454.

(166) Singh, R.; Blattler, W. A.; Collinson, A. R. Assay for Thiols Based on Reactivation of Papain. *Methods Enzymol.* **1995**, *251* (1975), 229–237.

(167) Falk, K. L.; Tokuhisa, J. G.; Gershenzon, J. The Effect of Sulfur Nutrition on Plant Glucosinolate Content: Physiology and Molecular Mechansims. *Plant Biol.* **2007**, *9*, 573–581. https://doi.org/10.1055/s-2007-965431.

Appendix A

Chapter 2 Supporting Information

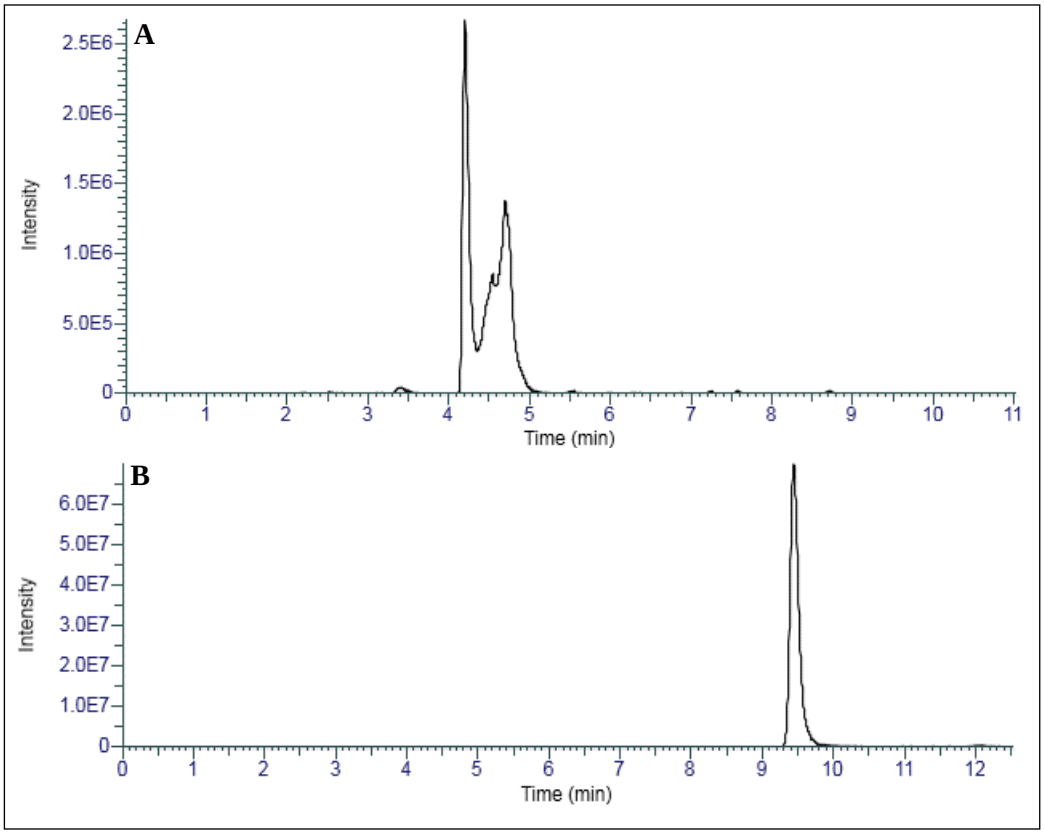


Figure S2-1. Comparison of the XIC for glucobrassicin (*m/z* 447.05374) on a traditional C18 column (A) versus the sulfonamide-embedded column used in the current study (B).

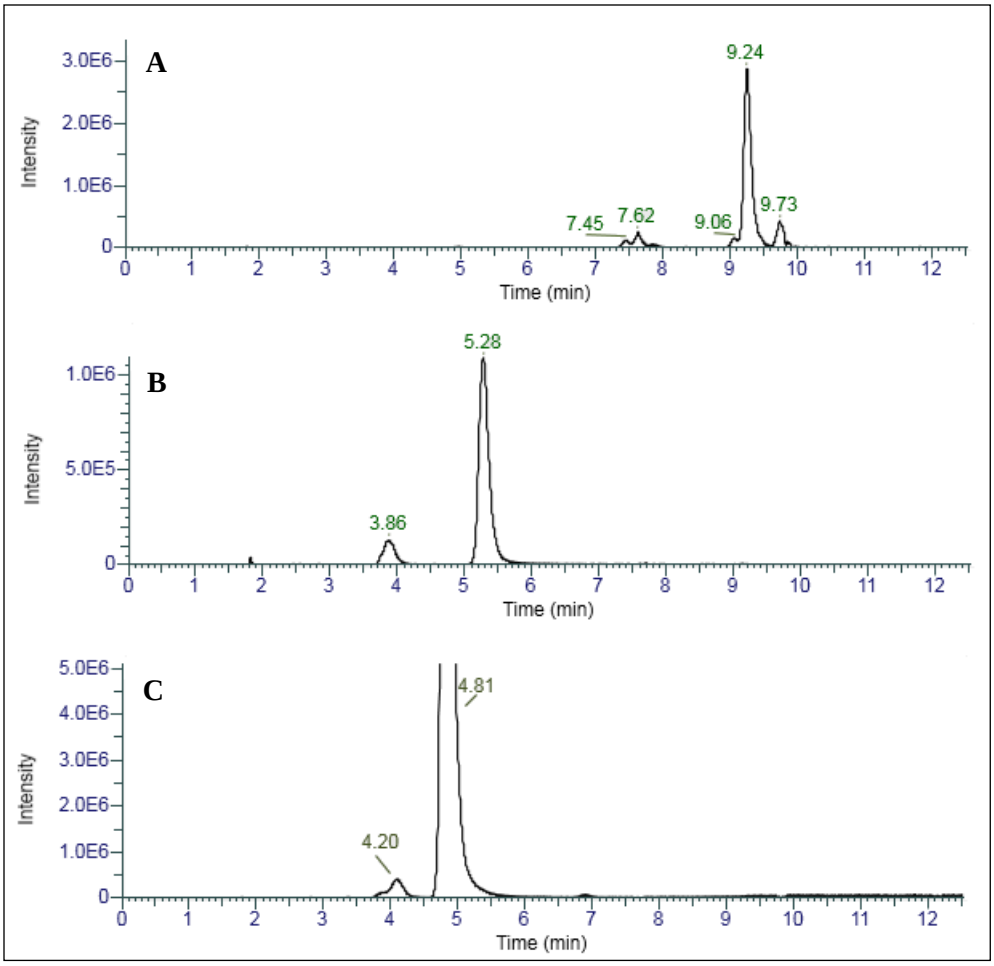


Figure S2-2. Example XICs of minor GLS with isomers. (A) Glucobarbarin (*m/z* 438.05341) from rapeseed. (B) Hydroxyglucobarbarin (*m/z* 454.04833) from turnip. (C) Gluconapoleiferin (*m/z* 402.05341) from turnip.

HPLC Method	Raw Data (ppm)		Isocratic vs Current Gradient Method		Glucobrassicin vs Sinigrin
	Sinigrin	Glucobrassicin	Sinigrin (Factor S)	Glucobrassicin (GLS Factor)	Response Factor (RF)
30% HOAc/ 70% MeOH	1.63	1.35	1.41	1.02	0.720
40% HOAc/ 60% MeOH	2.04	1.51	1.77	1.14	0.642
45% HOAc/ 55% MeOH	1.99	1.50	1.73	1.13	0.652
50% HOAc/ 50% MeOH	1.68	1.35	1.45	1.01	0.695
55% HOAc/ 45% MeOH	1.70	1.38	1.48	1.04	0.702
60% HOAc/ 40% MeOH	1.90	1.53	1.65	1.15	0.696
70% HOAc/ 30% MeOH	1.68	1.62	1.46	1.21	0.833
80% HOAc/ 20% MeOH	1.52	1.20	1.32	0.904	0.687
90% HOAc/ 10% MeOH	1.24	1.05	1.07	0.790	0.733
Current Gradient Method	*1.15*	*1.33*	-	-	Avg. 0.703 ± 0.05

Table S2-1. Example data and calculation of a response factor using glucoerucin. [GLS] = response of GLS in ppm given from the calibration curve. *Iso* = [GLS] at given under the individual isocratic method, and *gra* = [GLS] under the current gradient method. Actual amounts of sinigrin and glucobrassicin spiked are 1.15 ppm and 0.921 ppm, respectively.

1. $Factor\ S = \frac{[Sinigrin]_{iso}}{[Sinigrin]_{gra}}$

2. $GLS\ Factor = \frac{[GLS]_{iso}}{[GLS]_{gra}}$

3. $RF = \frac{GLS\ Factor}{Factor\ S}$

4. $RF \times [GLS]_{gra} = [GLS]_{actual}$

Example calculation using 55% HOAc glucobrassicin data (shown above in Table S2-1):

1. $Factor\ S = \frac{1.70\ ppm}{1.15\ ppm} = 1.48$

2. $GLS\ Factor = \frac{1.38\ ppm}{1.33\ ppm} = 1.04$

3. $RF = \frac{1.04}{1.48} = 0.703$

4. $0.703 \times 1.33\ ppm = 0.935\ ppm$

Actual glucobrassicin = 0.921 ppm (1.5% error)

Sample	Sinigrin (ISTD)	Gluco-napin	Glucobrassi-canapin	Progoitrin	Gluco-erucin	Glucora-phanin	Gluconapo-leiferin	Gluco-brassicin	Gluconas-turtiin	4-methoxy-glucobrassicin	Neogluco-brassicin	Total
Rapeseed 1	0.46	0.77	1.33	1.65	0.03	0.05	0.21	0.44	1.01	0.21	0.50	6.20
Rapeseed 2	0.46	0.72	1.23	1.61	0.03	0.05	0.20	0.43	0.98	0.21	0.54	6.00
Rapeseed 3	0.48	0.79	1.37	1.68	0.03	0.05	0.21	0.46	1.06	0.22	0.49	6.36
Rapeseed 4	0.48	0.75	1.33	1.68	0.03	0.05	0.21	0.46	1.03	0.21	0.54	6.29
Rapeseed 5*	0.50	0.74	1.24	1.62	0.02	0.05	0.20	0.42	0.95	0.20	0.38	5.83
Rapeseed 6*	0.52	0.77	1.34	1.67	0.03	0.05	0.21	0.44	1.00	0.20	0.45	6.16
Average	***0.48***	***0.76***	***1.31***	***1.65***	***0.03***	***0.05***	***0.21***	***0.44***	***1.01***	***0.21***	***0.48***	***6.14***
St. Dev.	***0.024***	***0.023***	***0.052***	***0.027***	***0.00092***	***0.0010***	***0.0056***	***0.014***	***0.036***	***0.0060***	***0.054***	***0.18***
% RSD	***4.9***	***3.1***	***4.0***	***1.7***	***3.5***	***2.0***	***2.7***	***3.2***	***3.6***	***2.9***	***11.3***	***2.9***
Turnip 1	0.43	0.58	2.47	1.76	0.64	0.10	0.56	0.48	2.46	0.38	1.82	11.25
Turnip 2	0.49	0.59	2.49	1.86	0.59	0.11	0.59	0.49	2.45	0.38	1.78	11.33
Turnip 3	0.46	0.61	2.44	1.87	0.65	0.11	0.59	0.51	2.66	0.38	1.83	11.63
Turnip 4	0.47	0.63	2.57	1.90	0.65	0.11	0.61	0.53	2.65	0.40	1.93	11.98
Turnip 5*	0.51	0.63	2.70	1.91	0.63	0.11	0.62	0.54	2.67	0.39	1.83	12.03
Average	***0.47***	***0.61***	***2.53***	***1.86***	***0.63***	***0.11***	***0.59***	***0.51***	***2.58***	***0.39***	***1.84***	***11.65***
St. Dev.	***0.026***	***0.021***	***0.094***	***0.053***	***0.023***	***0.0046***	***0.020***	***0.021***	***0.10***	***0.0067***	***0.048***	***0.32***
% RSD	***5.4***	***3.4***	***3.7***	***2.8***	***3.6***	***4.2***	***3.4***	***4.1***	***3.9***	***1.7***	***2.6***	***2.7***

Table S2-2. Extraction reproducibility. Values are reported in ppm (parts per million) since adjustment to sinigrin equivalents was not needed for statistical analyses of RF values. *Sinigrin was spiked into sample after extraction.

Sample	Sinigrin (spiked)	Gluco-napin	Glucobrassi-canapin	Pro-goitrin	Gluco-erucin	Glucora-phanin	Glucona-poleiferin	Gluco-brassicin	Glucona-sturtiin	4-methoxy-glucobrassicin	Neogluco-brassicin	Total
Turnip 1st Extraction_1	0.49	0.65	2.68	1.98	0.71	0.09	0.64	0.55	2.80	0.39	1.92	12.40
Turnip 1st Extraction_2	0.49	0.66	2.74	2.03	0.73	0.10	0.64	0.55	2.80	0.40	1.96	12.62
Turnip 2nd Extraction_1*	0.52	0.05	0.24	0.17	0.06	0.01	0.05	0.05	0.27	0.04	0.18	1.12
Turnip 2nd Extraction_2*	0.49	0.05	0.23	0.16	0.06	0.01	0.05	0.05	0.26	0.03	0.17	1.08

Table S2-3. Extraction efficiency. Residual GLS from second extraction_1 = 9.03 %; Residual GLS from second extraction_2 = 8.56 %

Sample	Peak Area (counts*sec)	Amount (ppm)
0.5 ppm sinigrin_spiked_1	2.19e8	0.454
0.5 ppm sinigrin_spiked_2	2.26e8	0.468
0.5 ppm sinigrin_spiked_3	2.29e8	0.475
Average	2.27e8	0.465
Standard Deviation	4.65e6	0.0087
%RSD	2.05%	2.05%
5 ppm sinigrin_spiked_1	2.34e9	4.84
5 ppm sinigrin_spiked_2	2.36e9	4.87
5 ppm sinigrin_spiked_3	2.45e9	5.07
Average	2.38e7	4.93
Standard Deviation	4.98e7	0.103
%RSD	2.09%	2.09%

Table S2-4. Peak area reproducibility of 0.5 ppm/5 ppm sinigrin control samples in triplicate.

Glucosinolate (semi systematic name)	Common Name	Molecular Formula	Fahey ID
2-propenyl	Sinigrin[1]	C10H17NO9S2	D-107
1-methylethyl	Glucoputranjivin	C10H19NO9S2	C-56
3-butenyl	Gluconapin[1]	C11H19NO9S2	D-12
1-methylpropyl	Glucococchlearin	C11H21NO9S2	C-61
2-methylpropyl	Isobutyl	C11H21NO9S2	C-62
n-butyl[2]	-	C11H21NO9S2	B-13
4-pentenyl	Glucobrassicanapin[1]	C12H21NO9S2	D-101
2(R)-2-hydroxyl-3-butenyl	Epiprogoitrin	C11H19NO10S2	D-24b
2(S)-2-hydroxyl-3-butenyl	Progoitrin[1]	C11H19NO10S2	D-24a
2-methylbutyl	-	C12H23NO9S2	C-54
3-methylbutyl[2]	-	C12H23NO9S3	C-55
n-pentyl[2]	-	C12H23NO9S2	B-102
1-(hydroxymethyl)propyl	Glucosisaustricin	C11H21NO10S2	E-30
2-hydroxy-2-methylpropyl	Glucoconringiin	C11H21NO10S2	E-31
3-hydroxybutyl[2]	-	C11H21NO10S2	E-25
4-hydroxybutyl[2]	-	C11H21NO10S2	E-26
5-hexenyl	-	C13H23NO9S2	D-19
2(R)-hydroxy-4-pentenyl	Epigluconapoleiferin	C12H21NO10S3	D-38b
2(S)-hydroxy-4-pentenyl	Gluconapoleiferin	C12H21NO10S2	D-38a
4-oxopentyl	-	C12H21NO10S2	F-99
2-hydroxy-2-methylbutyl	Glucocleomin	C12H23NO10S2	E-29
Benzyl	Glucotropaeolin	C14H19NO9S2	G-11
4-(methylthio)butyl	Glucoerucin[1]	C12H23NO9S3	A-84
2-phenylethyl	Gluconasturtiin[1]	C15H21NO9S2	G-105
4-hydroxybenzyl	Glucosinalbin[1]	C14H19NO10S2	G-23
5-(methylthio)pentyl	Glucoberteroin	C13H25NO9S3	A-94
4-(methylsulfinyl)butyl	Glucoraphanin[1]	C12H23NO10S3	A-64
2-(3-hydroxyphenyl)ethyl	-	C15H21NO10S2	148
2(R)-hydroxy-2-phenylethyl	Glucobarbarin	C15H21NO10S2	G-40
4-hydroxyphenethyl	Homosinalbin	C15H21NO10S2	140
3-methoxybenzyl	Glucolimnanthin	C15H21NO10S2	G-45
4-methoxybenzyl	Glucoaubrietin	C15H21NO10S2	G-46
3,4-dihydroxybenzyl	Glucomatronalin	C14H19NO11S2	G-14
Indol-3-ylmethyl	Glucobrassicin[1]	C16H20N2O9S2	I-43
3-hydroxy-5-(methylthio)pentyl[2]	-	C13H25NO10S3	A-37
5-(methylsulfinyl)pentyl	Glucoalyssin	C13H25NO10S3	A-72
4-(methylsulfonyl)butyl[2]	Glucoerysolin	C12H23NO11S3	A-76
(R)-2-hydroxy-2-(3-hydroxyphenyl)ethyl	m-Hydroxyepiglucobarbarin	C15H21NO11S2	142R
(R)-2-hydroxy-2-(4-hydroxyphenyl)ethyl	p-Hydroxyepiglucobarbarin	C15H21NO11S2	139R
(S)-2-hydroxy-2-(4-hydroxyphenyl)ethyl	p-Hydroxyglucobarbarin	C15H21NO11S2	139S
3-hydroxy-4-methoxybenzyl	Glucobretschneiderin	C15H21NO11S2	153
4-hydroxy-3-methoxybenzyl	3-Methoxysinalbin	C15H21NO11S2	151
4-hydroxyindol-3-ylmethyl	4-hydroxyglucobrassicin	C16H20N2O10S2	I-28
4-methoxyindolyl	Glucorapassasin A	C16H20N2O10S2	147
3-hydroxy-6-(methylthio)hexyl[2]	-	C14H27NO10S3	A-36
6-(methylsulfinyl)hexyl	Glucohespirin	C14H27NO10S3	A-67
3-hydroxy-5-(methylsulfinyl)pentyl[2]	-	C13H25NO11S3	A-33
?-methoxyindol-?-ylmethyl*	-	C17H22N2O10S2	I?
4-methoxyindol-3-ylmethyl	4-Methoxyglucobrassicin	C17H22N2O10S2	I-48
1-methoxyindol-3-ylmethyl	Neoglucobrassicin	C17H22N2O10S2	I-47
4-hydroxy-3,5-dimethoxybenzyl	-	C16H22NO12S2	152
2-benzoylprogoitrin[2]	-	C18H22NO11S2	123
3,4,5-trimethoxybenzyl	-	C17H25NO12S2	G-114
1,4-dimethoxyindol-3-ylmethyl	-	C18H24N2O11S2	138

Table S2-5. Summary of GLS used in this study. Since all GLS do not have a common name, so the classification given by Fahey et al.[95] and Agerbirk and Olsen [96] are shown (Fahey ID) with the semi-systematic chemical name.

Sample (dilution factor)*	Gluconapin (ppm)	Glucobrassicanapin (ppm)	Gluconasturtiin (ppm)	Glucobrassicin (ppm)
Rapeseed 1 10x	0.265	0.262	0.451	0.119
Rapeseed 1 100x	0.022	0.021	0.041	0.011
Rapeseed 1 1000x	0.002	0.002	0.004	0.001
Rapeseed 2 10x	0.230	0.223	0.339	0.100
Rapeseed 2 100x	0.018	0.018	0.038	0.010
Rapeseed 2 1000x	0.002	0.002	0.004	0.001
Canola 1 10x	0.148	0.122	0.170	0.087
Canola 1 100x	0.011	0.009	0.016	0.008
Canola 1 1000x	0.001	0.001	0.001	0.001
Canola 2 10x	0.127	0.101	0.154	0.078
Canola 2 100x	0.013	0.011	0.020	0.010
Canola 2 1000x	0.001	0.001	0.002	0.001
Turnip 1 10x	0.453	1.416	0.934	0.188
Turnip 1 100x	0.042	0.134	0.095	0.019
Turnip 1 1000x	0.005	0.013	0.011	0.002
Turnip 2 10x	0.458	1.414	0.984	0.194
Turnip 2 100x	0.040	0.132	0.095	0.019
Turnip 2 1000x	0.004	0.013	0.010	0.002

Table S2-6. Study of dilution and matrix effects on GLS response. Values are reported in ppm (parts per million) since adjustment to sinigrin equivalents was not needed to examine dilution effects. Duplicate samples were extracted and diluted at 1:10, 1:100 and 1:100x the original extract. * All samples are diluted 1:10 after extraction, before injection for cleaner samples (to preserve the HPLC equipment). These four GLS were chosen to represent GLS linear response to sample dilution.

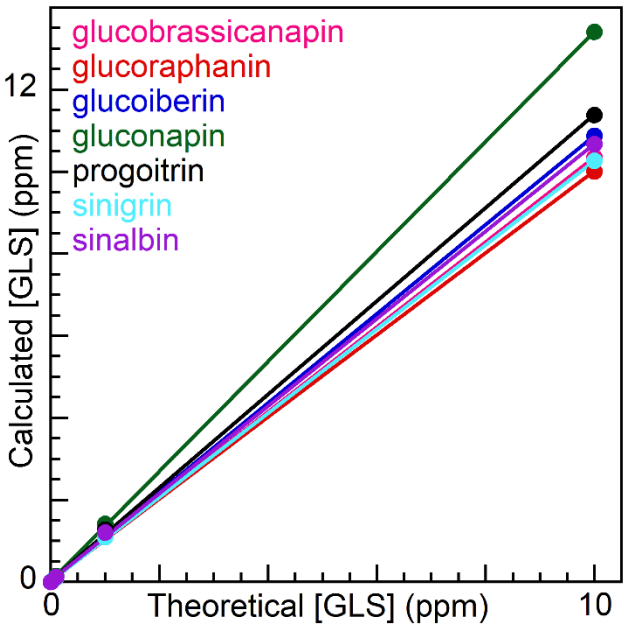


Figure S2-3. Serial dilution of 7 GLS standards from 0.001 – 10 ppm. Calculated values for [GLS] are reported in ppm, not sinigrin equivalents. $R^2 = 0.999$ for all 7 GLS, demonstrating a linear response to dilution.

Appendix B

Chapter 3 Supporting Information

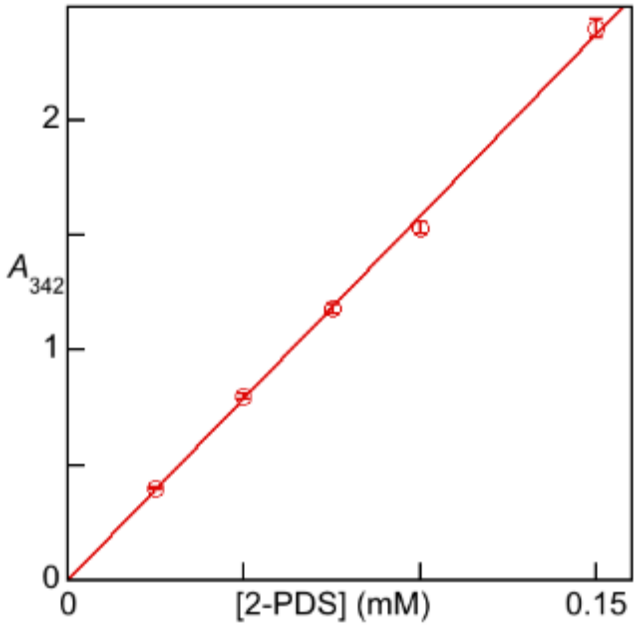


Figure S3-1. Determination of the effective molar absorptivity of 2-thiopyridone (2-TP). 2-PDS (0.025, 0.050, 0.075, 0.100, and 0.150 mM) in 0.5 mL reaction buffer (100 mM sodium phosphate, pH 6) was reduced by addition of 1 mM tris(2-carboxyethyl)phosphine (TCEP). The *red line* is a fit of the equation for a line to the plot of A_{342} versus [2-PDS]. Because each mole of 2-PDS yields two moles of 2-TP, ε_{342}(2-TP) is the slope divided by two.

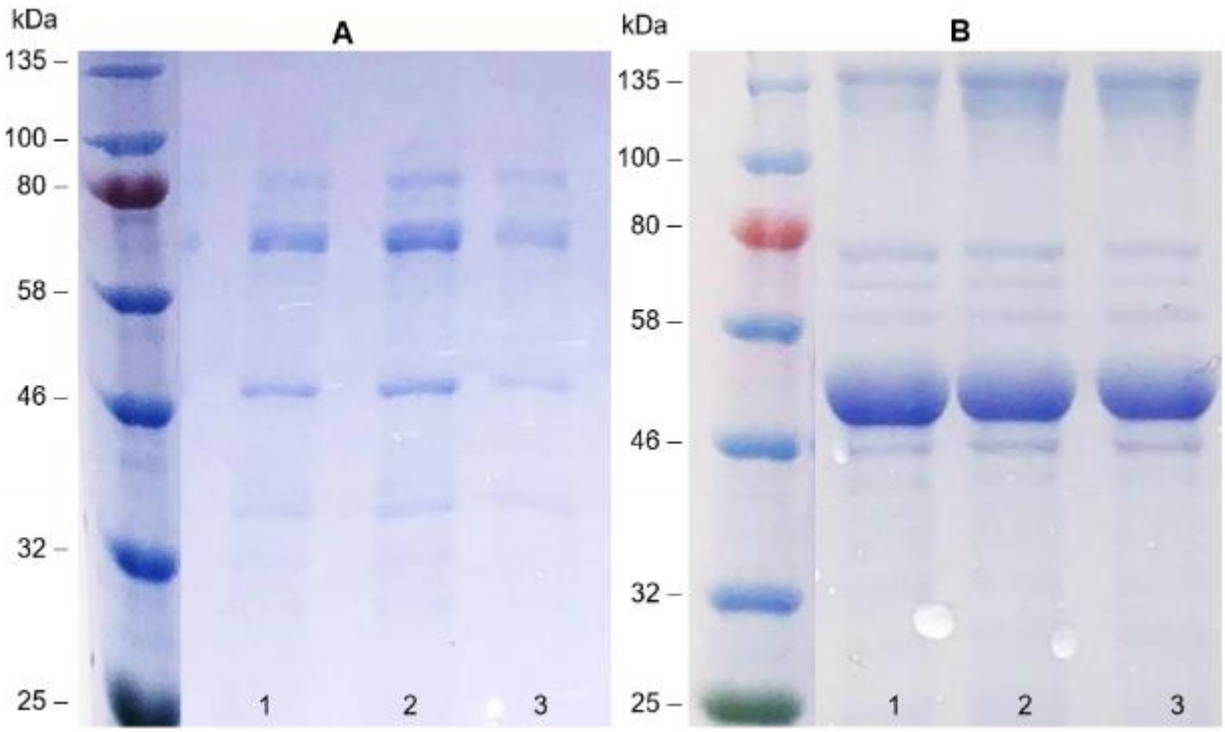


Figure S3-2. SDS PAGE analysis of *Sa* MYR and recombinant *At* ESP. (***A***) Fractions eluting from the Concanavalin A Sepharose 4B affinity column in purification of *Sa* MYR. The enzyme is associated with the pair of smeared, prominent bands at ~ 65 kDa. The breadth/multiplicity of bands is most likely a result of heterogenous glycosylation. (***B***) Fractions eluting from the Ni(II)-nitrilotriacetate IMAC column in purification of recombinant SUMO-His_6-tagged *At* ESP. The most prominent band at ~ 52 kDa represents the ESP monomer with the SUMO-His_6 tag. The apparent contaminants are most likely oligomers of the ESP monomer.

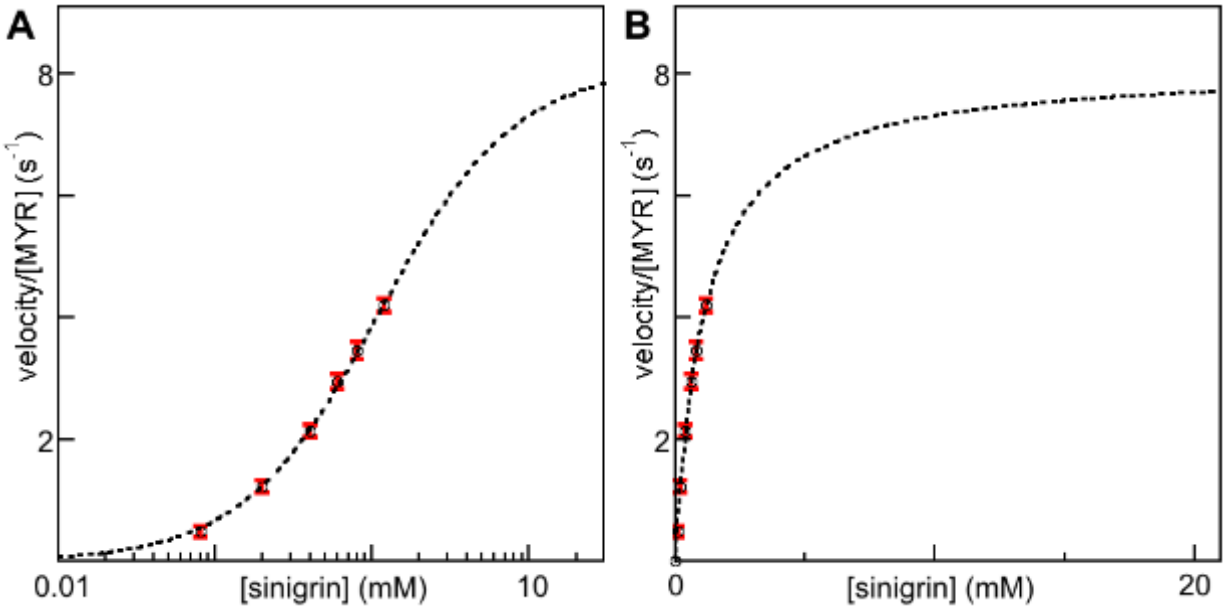


Figure S3-3. Steady-state kinetic characterization of purified *Sa* MYR. The initial velocities of the MYR hydrolysis reaction monitored by absorbance at 227 nm [ε_{227}(sinigrin) = 6.5 $mM^{-1}cm^{-1}$] were calculated using the slopes of the early segments of the traces. These values were divided by [MYR] to get the plotted frequencies. The hyperbolic fit gave estimates for the maximum turnover efficiency (k_{cat} = 8.2 ± 0.4 s^{-1}) and Michaelis-Menten constant (K_M = 1.1 ± 0.2 mM), and k_{cat}/K_M (7.3 ± 0.3 $mM^{-1}s^{-1}$) for the *Sa* MYR. The red bars reflect twice the standard deviations of the values measured in three replicate experiments from the mean values. The data are plotted with both logarithmic (***A***) and linear (***B***) x-axes for ease of visualization.

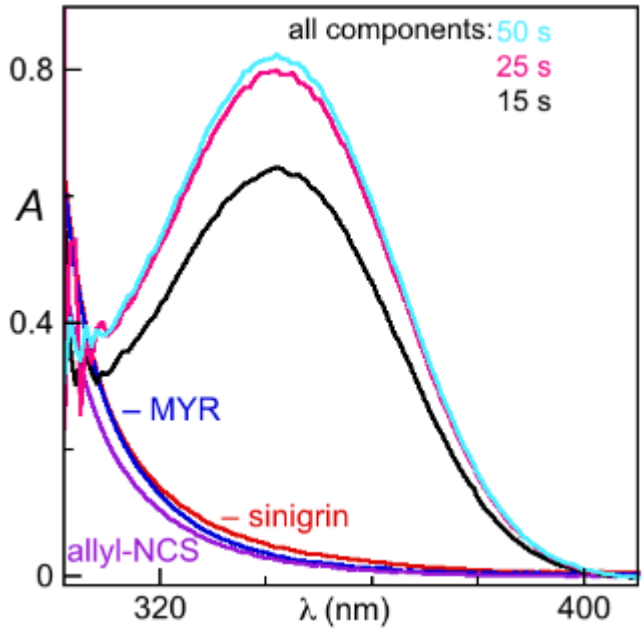


Figure S3-4. Spectra establishing the specificity of 2-PDS as reporter of aglycones. The *black*, *pink*, and *cyan* traces are from an experiment with all components (MYR, sinigrin, 2-PDS) combined at $t = 0$. The *dark blue* and *red* traces are from control experiments in which the indicated component was omitted from the assay. The *purple* trace is from a control experiment in which sinigrin was replaced by the product of the Lossen-like rearrangement of the sinigrin aglycone, the allyl isothiocyanate.

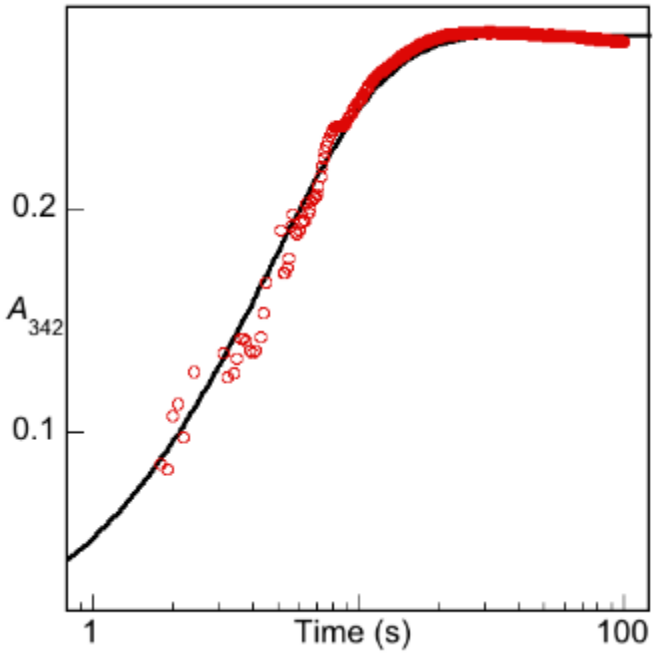


Figure S3-5. Kinetics of aglycone capture by 2-PDS. MYR was incubated with sinigrin until hydrolysis reached completion, at which point 2-PDS was added to a final concentration of 1.5 mM and the absorbance at 342 nm was monitored in time. The solid line is the fit of Eq. 3 to the data. It gave an effective first-order rate constant for approach to completion (k_{obs}), from which k_2 was calculated according to Eq. 2.

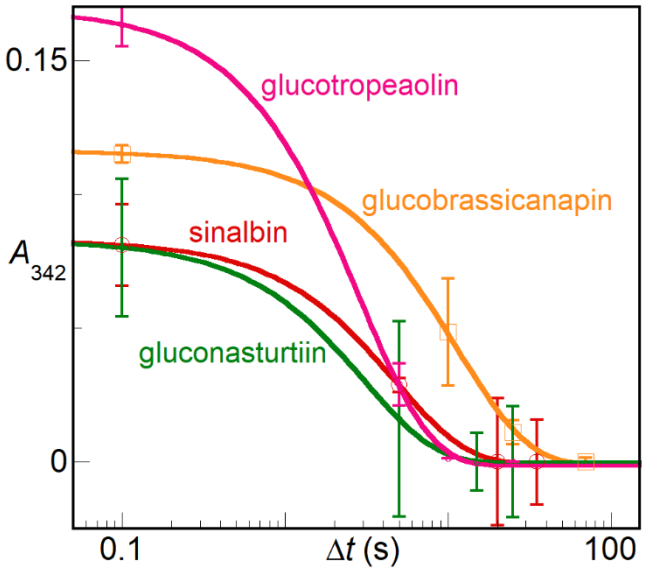


Figure S3-6**.** Kinetics of the Lossen-like rearrangements of the aglycones from four additional GLSs (those not shown in Figure 3-4). The solid lines are fits of Eq 4 to the data. Experimental details are provided in the *Materials and Methods* section of the main text.

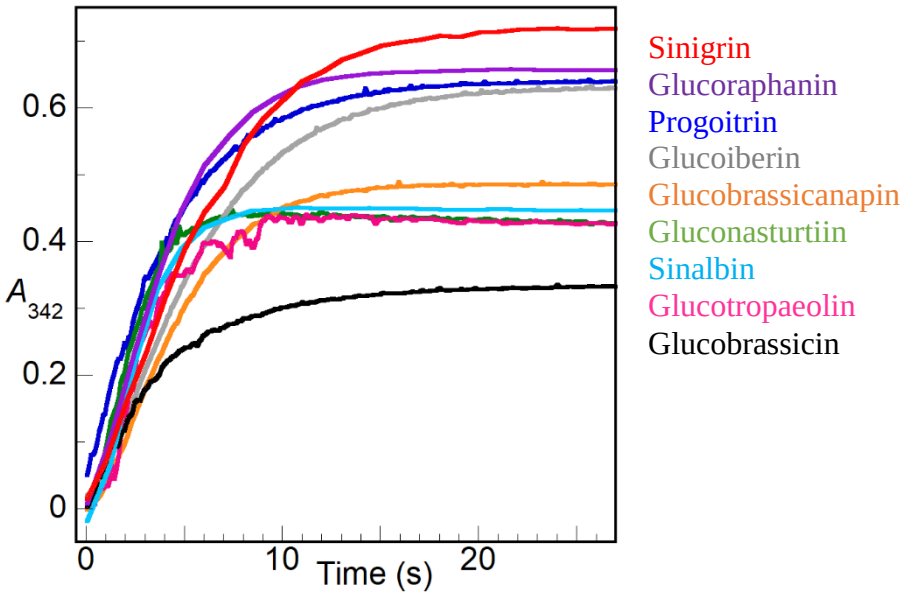


Figure S3-7. Use of 2-PDS in coupled assays to define the times required for complete hydrolysis of each GLS by MYR. Experimental details are provided in the *Materials and Methods* section of the main text.

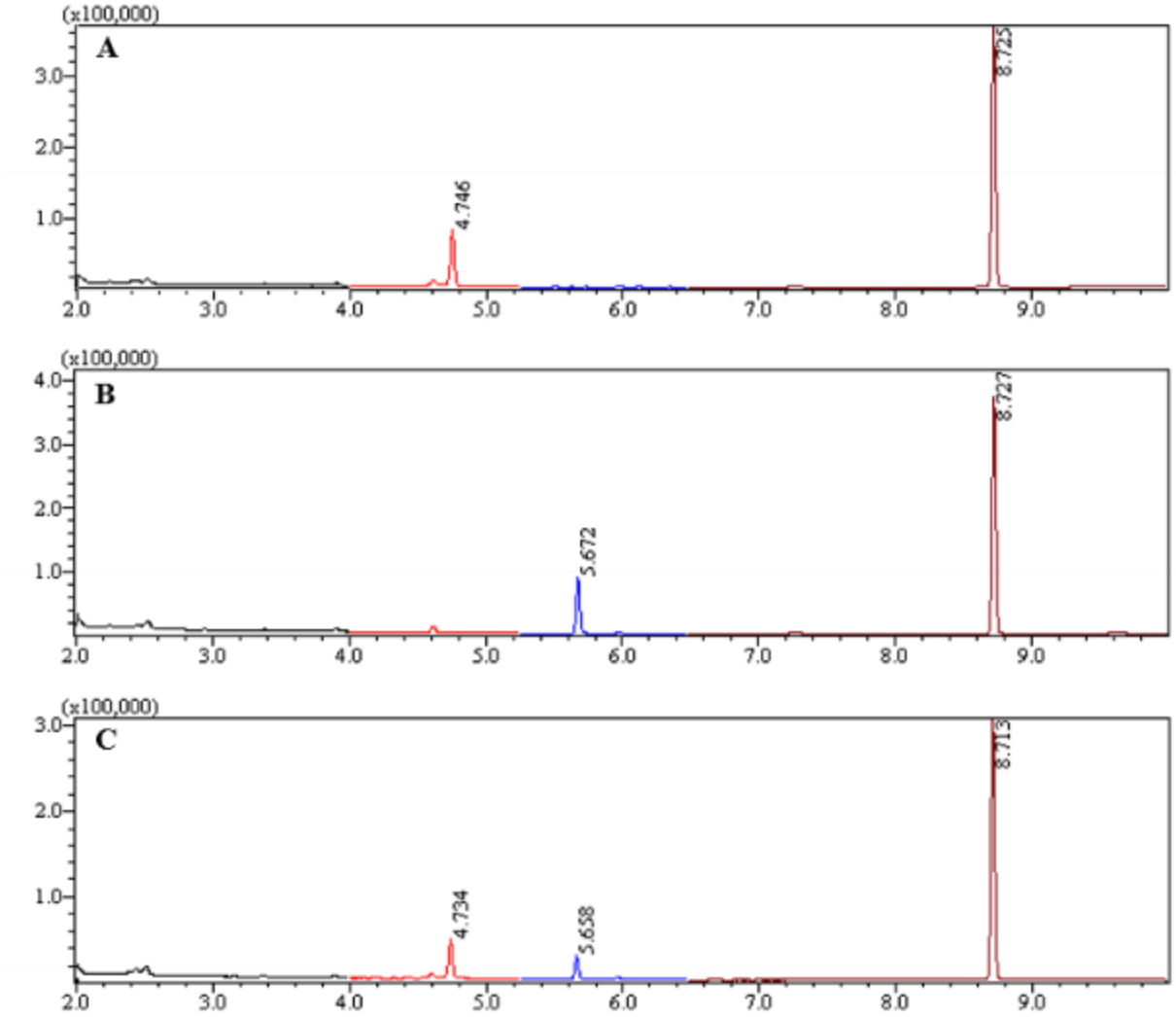


Figure S3-8. GC-MS chromatograms from analysis for the allyl-NCS and epithionitrile products in assays of MYR and sinigrin (± ESP). Bar graphs of the same data are presented in Figure 3-5. (***A***) Control reaction, in which hydrolysis of 0.25 mM sinigrin by 25 µM MYR was carried out in the absence of ESP, yielding only allyl -NCS (*red*, 4.7 min). (***B***) Control reaction, in which ESP was present (10 µM) upon initiation of sinigrin hydrolysis by MYR, yielding only the epithionitrile (*blue*, 5.6 min). (***C***) Experimental reaction, in which ESP was added after complete hydrolysis of sinigrin by a 30-s incubation with MYR prior to addition of ESP to 0.10 mM [with 0.20 mM Fe(II)], yielding similar quantities of the two products. The *brown* peak at 8.7 min is from the benzyl-NCS internal standard.

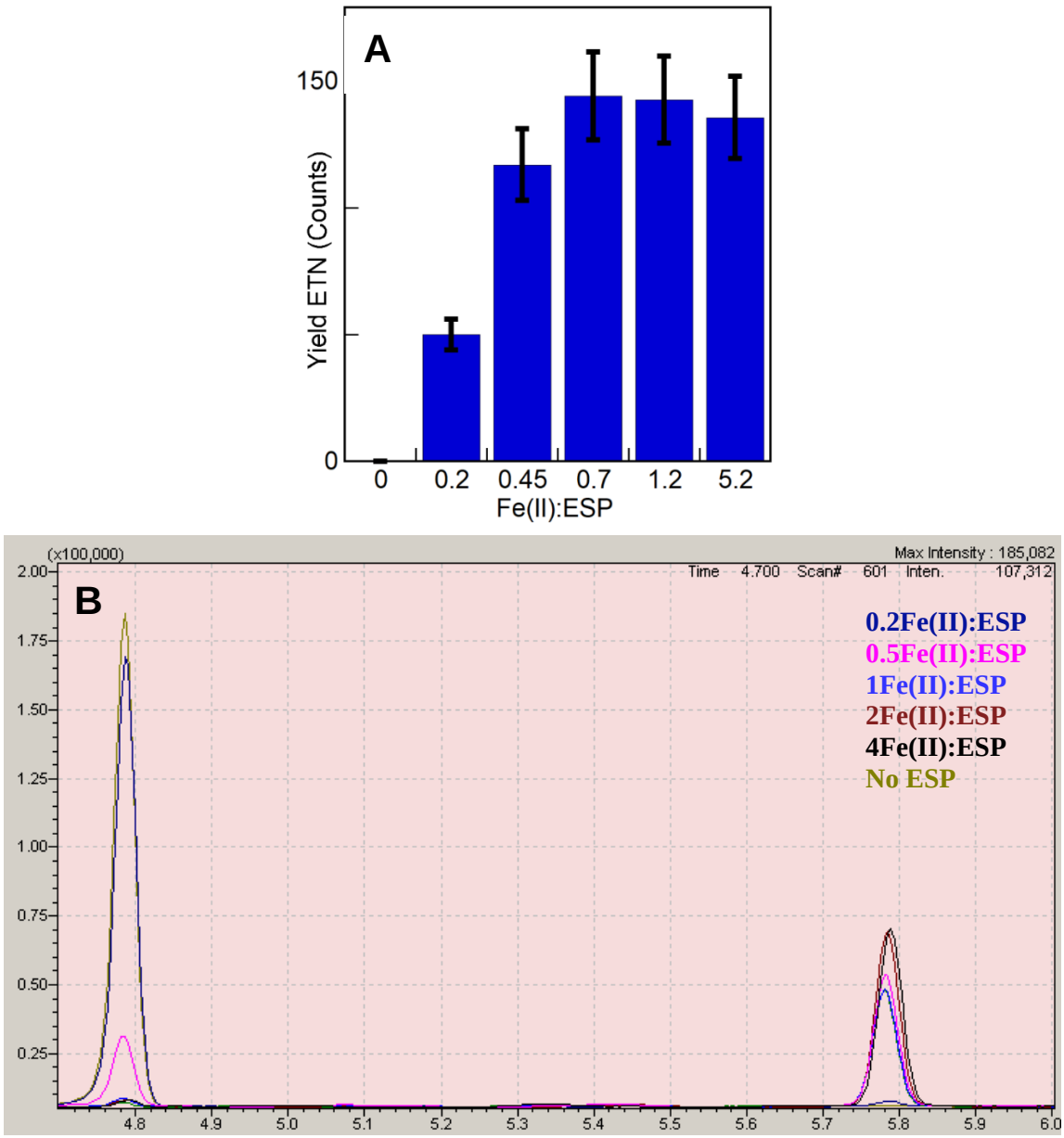


Figure S3-9. Dependence of ESP activity on added Fe(II). Epithionitrile formation was monitored by the intensity of its GC-MS signature (peak at ~ 5. 8 min elution time). ESP was present (10 µM) at initiation of MYR-catalyzed (2 µM) hydrolysis of sinigrin (0.50 mM). (***A***) The x-axis gives the equivalency of total iron, which includes the 0.2 equiv present in this preparation of ESP and the Fe(II) added to activate the protein, from an experiment conducted in the absence of O_2. The Fe(II):ESP = 0 point is not experimental but assumes complete inactivity in the (hypothetical) true absence of Fe(II). (***B***) Raw GC-MS chromatograms from assays containing 15 µM ESP, 3 µM MYR, 0.50 mM sinigrin, and Fe(II). The chromatograms show an increase in ETN formation as Fe(II) increases and a corresponding decrease in allyl-NCS formation.

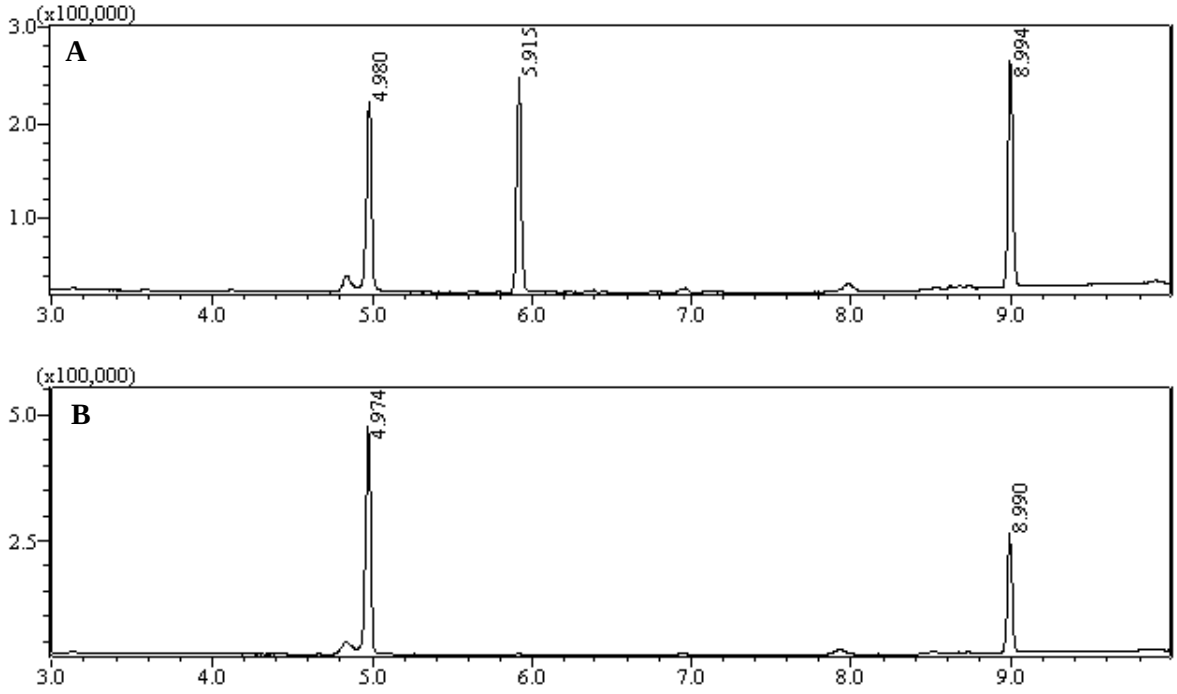
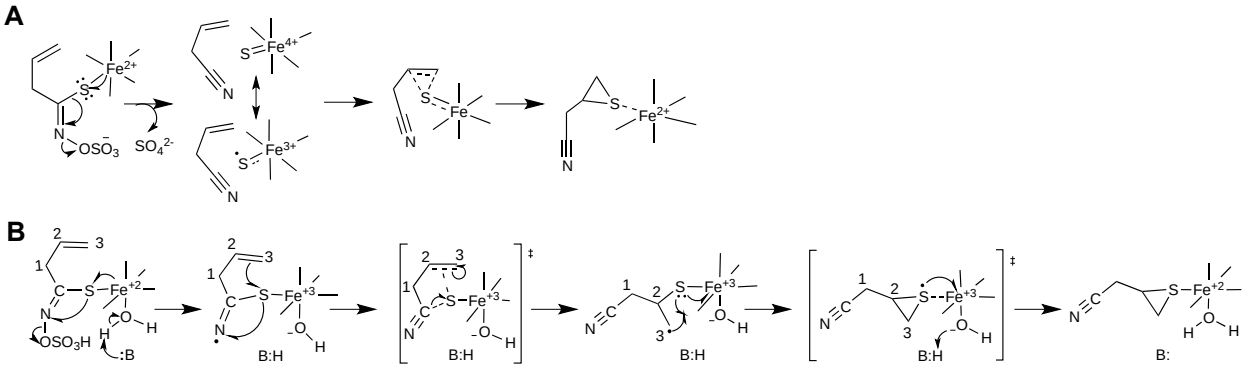


Figure S3-10. GC-MS chromatograms of the experiment in which resin-bound MYR hydrolyzed sinigrin before the aglycone was eluted into a solution of Fe(II)•ESP. (***A***) Assay with ESP added, showing a combination of allyl-NCS (4.9 min) and allyl epithionitrile (5.9 min). (***B***) Corresponding control reaction in which only buffer was added to the reservoir, showing only allyl-NCS formed. The peak at 8.9 min is from the benzyl-NCS internal standard.



Scheme S1**.** Possible mechanisms of epithionitrile formation from sinigrin by ESP. (***A***) Mechanism involving initial transfer of sulfur to iron to form separated Fe-S species (depicted as a pair of resonance tautomers) that then attacks olefin to transfer sulfur and cyclize. An example of this mechanistic type can be found in reference 10. (***B***) Mechanism involving iron-promoted direct migration of sulfur from central carbon of thiohydroximate intermediate to olefin without separated Fe-S intermediate (adapted from Eisenschmidt-Bönn et al.[38]).

At ESP pBA0028 plasmid DNA sequence (ESP is capitalized). 5' to 3'

gcgtagaggatcgagatctcgatcccgcgaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagaaataattttg
tttaactttaagaaggagatataccatgcatcatcaccatcaccacggttcttctatggctagcatgtcggactcagaagtcaatcaagaagcta
agccagaggtcaagccagaagtcaagcctgagactcacatcaatttaaaggtgtccgatggatcttcagagatcttcttcaagatcaaaaagac
cactcctttaagaaggctgatggaagcgttcgctaaaagacagggtaaggaaatggactccttaagattcttgtacgacggtattagaattcaa
gctgatcagacccctgaagatttggacatggaggataacgatattattgaggctcacagagaacagattggtgggatcgaggaaaacctgtact
tccaatccaatattctcgagATGGCCCCAACCTTGCAAGGTCAATGGATCAAGGTCGGTCAGAAAGGCGGCACCGGC
CCTGGCCCGCGTTCGTCACATGGTATCGCTGCGGTGGGTGACAAGCTGTATAGCTTTGGCGGTGAGCTGACC
CCTAACAAACACATTGACAAGGACCTGTATGTTTTCGATTTCAACACTCAGACCTGGTCGATTGCACAGCCGA
AAGGCGACGCGCCGACGGTTTCCTGCCTGGGCGTCCGCATGGTGGCGGTGGGCACCAAAATCTACATCTTTG
GTGGTAGAGATGAGAATCGTAACTTCGAGAATTTCCGTTCTTATGATACCGTGACTTCGGAGTGGACGTTCCT
TACCAAGCTGGATGAAGTGGGTGGTCCAGAAGCTCGCACCTTCCATTCTATGGCAAGCGACGAGAATCATGT
CTACGTCTTCGGTGGTGTTTCCAAAGGCGGTACTATGAATACGCCGACCCGCTTCCGTACCATTGAAGCGTAC
AACATTGCGGATGGTAAGTGGGCACAGTTGCCGGACCCAGGCGATAACTTCGAGAAACGCGGTGGTGCGGG
TTTTGCGGTAGTTCAAGGCAAAATTTGGGTGGTGTACGGCTTTGCGACGAGCATTGTGCCGGGCGGTAAGGA
CGATTACGAAAGCAATGCGGTTCAATTTTACGACCCGGCGAGCAAGAAGTGGACCGAAGTAGAAACCACGG
GTGCCAAACCGAGCGCACGTTCGGTGTTCGCGCACGCTGTTGTTGGCAAGTACATCATTATCTTTGCCGGTGA
AGTGTGGCCGGATCTGAATGGTCACTATGGTCCGGGCACCCTGTCCAACGAGGGCTATGCGTTAGACACGGA
AACCCTGGTCTGGGAGAAGTTGGGTGAAGAAGGCGCTCCGGCGATCCCGCGTGGTTGGACGGCCTACACCG
CGGCTACCGTTGACGGCAAGAACGGTTTGCTGATGCACGGTGGTAAACTCCCGACCAACGAGCGCACCGACG
ACCTGTACTTTTATGCCGTCAATTCCGCGtaaggtacctgagatccggctgctaacaaagcccgaaaggaagctgagttggctgct
gccaccgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaaggaggaactatatccggattggc
gaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgccctagcgcccg
ctcctttcgctttcttcccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgggggctccctttagggttccgatttagtgcttt
acggcacctcgaccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgatagacggtttttcgccctttgacgttggagtc
cacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatctcggtctattcttttgatttataagggattttgccgatttcggc
ctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaacaaaatattaacgtttacaatttcaggtggcacttttcggggaaa
tgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgaattaattcttagaaaaactcatcgagcatcaaatga
aactgcaatttattcatatcaggattatcaataccatatttttgaaaaagccgtttctgtaatgaaggagaaaactcaccgaggcagttccatagg
atggcaagatcctggtatcggtctgcgattccgactcgtccaacatcaatacaacctattaatttcccctcgtcaaaaataaggttatcaagtgag
aaatcaccatgagtgacgactgaatccggtgagaatggcaaaagtttatgcatttctttccagacttgttcaacaggccagccattacgctcgtc
atcaaaatcactcgcatcaaccaaaccgttattcattcgtgattgcgcctgagcgagacgaaatacgcgatcgctgttaaaaggacaattacaa
acaggaatcgaatgcaaccggcgcaggaacactgccagcgcatcaacaatattttcacctgaatcaggatattcttctaatacctggaatgctgt
tttcccggggatcgcagtggtgagtaaccatgcatcatcaggagtacggataaaatgcttgatggtcggaagaggcataaattccgtcagccag
tttagtctgaccatctcatctgtaacatcattggcaacgctacctttgccatgtttcagaaacaactctggcgcatcgggcttcccatacaatcgat
agattgtcgcacctgattgcccgacattatcgcgagcccatttatacccatataaatcagcatccatgttggaatttaatcgcggcctagagcaag
acgtttcccgttgaatatggctcataacaccccttgtattactgtttatgtaagcagacagttttattgttcatgaccaaaatcccttaacgtgagtt
ttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaa
accaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaata
ctgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgc
tgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgc
acacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaa
ggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgt
cgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacg

gttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgatacc
gctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcctgatgcggtattttctccttacgcatctgtgc
ggtatttcacaccgcatatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagtatacactccgctatcgctacgtgactgg
gtcatggctgcgccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgt
ctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgaggcagctgcggtaaagctcatcagcgtggtcgtgaagcgat
tcacagatgtctgcctgttcatccgcgtccagctcgttgagtttctccagaagcgttaatgtctggcttctgataaagcgggccatgttaagggcgg
ttttttcctgtttggtcactgatgcctccgtgtaagggggatttctgttcatgggggtaatgataccgatgaaacgagagaggatgctcacgatac
gggttactgatgatgaacatgcccggttactggaacgttgtgagggtaaacaactggcggtatggatgcggcgggaccagagaaaaatcactc
agggtcaatgccagcgcttcgttaatacagatgtaggtgttccacagggtagccagcagcatcctgcgatgcagatccggaacataatggtgca
gggcgctgacttccgcgtttccagactttacgaaacacggaaaccgaagaccattcatgttgttgctcaggtcgcagacgttttgcagcagcagt
cgcttcacgttcgctcgcgtatcggtgattcattctgctaaccagtaaggcaaccccgccagcctagccgggtcctcaacgacaggagcacgatc
atgcgcacccgtggggccgccatgccggcgataatggcctgcttctcgccgaaacgtttggtggcgggaccagtgacgaaggcttgagcgagg
gcgtgcaagattccgaataccgcaagcgacaggccgatcatcgtcgcgctccagcgaaagcggtcctcgccgaaaatgacccagagcgctgcc
ggcacctgtcctacgagttgcatgataaagaagacagtcataagtgcggcgacgatagtcatgccccgcgcccaccggaaggagctgactggg
ttgaaggctctcaagggcatcggtcgagatcccggtgcctaatgagtgagctaacttacattaattgcgttgcgctcactgcccgctttccagtcg
ggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgccagggtggtttttcttttcacca
gtgagacgggcaacagctgattgcccttcaccgcctggccctgagagagttgcagcaagcggtccacgctggtttgccccagcaggcgaaaatc
ctgtttgatggtggttaacggcgggatataacatgagctgtcttcggtatcgtcgtatcccactaccgagatatccgcaccaacgcgcagcccgg
actcggtaatggcgcgcattgcgcccagcgccatctgatcgttggcaaccagcatcgcagtgggaacgatgccctcattcagcatttgcatggttt
gttgaaaaccggacatggcactccagtcgccttcccgttccgctatcggctgaatttgattgcgagtgagatatttatgccagccagccagacgc
agacgcgccgagacagaacttaatgggcccgctaacagcgcgatttgctggtgacccaatgcgaccagatgctccacgcccagtcgcgtaccg
tcttcatgggagaaaataatactgttgatgggtgtctggtcagagacatcaagaaataacgccggaacattagtgcaggcagcttccacagcaa
tggcatcctggtcatccagcggatagttaatgatcagcccactgacgcgttgcgcgagaagattgtgcaccgccgctttacaggcttcgacgccg
cttcgttctaccatcgacaccaccacgctggcacccagttgatcggcgcgagatttaatcgccgcgacaatttgcgacggcgcgtgcagggccag
actggaggtggcaacgccaatcagcaacgactgtttgcccgccagttgttgtgccacgcggttgggaatgtaattcagctccgccatcgccgctt
ccactttttcccgcgttttcgcagaaacgtggctggcctggttcaccacgcgggaaacggtctgataagagacaccggcatactctgcgacatcg
tataacgttactggtttcacattcaccaccctgaattgactctcttccgggcgctatcatgccataccgcgaaaggttttgcgccattcgatggtgtc
cgggatctcgacgctctcccttatgcgactcctgcattaggaagcagcccagtagtaggttgaggccgttgagcaccgccgccgcaaggaatgg
tgcatgcaaggagatggcgcccaacagtcccccggccacggggcctgccaccatacccacgccgaaacaagcgctcatgagcccgaagtggc
gagcccgatcttccccatcggtgatgtcggcgatataggcgccagcaaccgcacctgtggcgccggtgatgccggccacgatgcgtccg

Appendix C

Supporting Information for Chapter 4

Parent GLS	Isothiocyanate Fragments (*m/z*)	Epithionitrile Fragments (*m/z*)	Nitrile Fragments (*m/z*)
Sinigrin (allyl)	99 (M), 72, 59, 45	99 (M), 72, 59, 45	67 (M), 45, 41, 39
Glucotropaeolin (benzyl)	149 (M), 91, 72, 65	NA	117 (M), 116, 90, 63
Gluconasturtiin (phenethyl)	163 (M), 105, 91, 65	NA	131 (M), 91, 65, 51
Glucobrassicanapin (4-pentenyl)	127 (M), 85, 67, 41	127 (M), 85, 67, 55	*95 (M), 55, 41
Glucoerucin (4-methylthiobuytl)	161 (M), 115, 72, 61	NA	129 (M), 115, 82, 61
Gluconapin (3-butenyl)	113 (M), 72, 55, 45	113 (M), 73, 53, 45	*81 (M), 53, 41, 39

Table S4-1. Summary of m/z values used for volatile product detection. * these exist but were not tested for in this work. NA means it does not exist. Those whose frag patterns are very similar were ID with NIST database and then differentiated by RT. Frags found in NIST database.

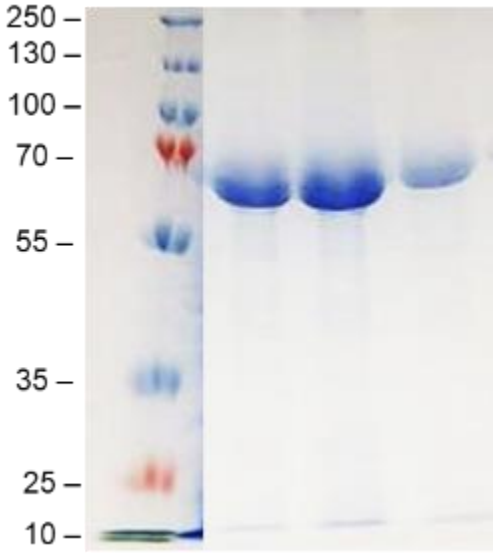
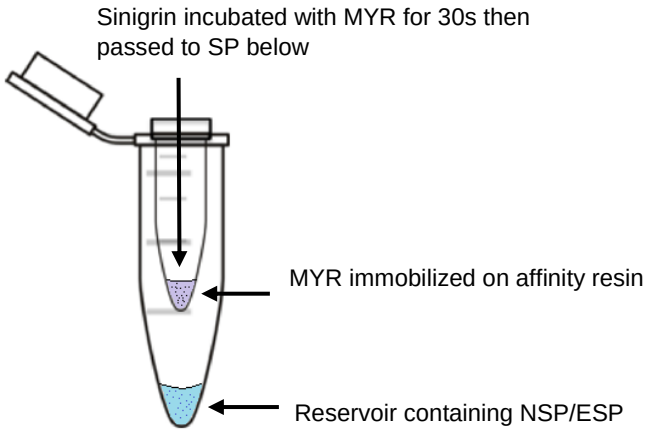


Figure S4-1. SDS PAGE analysis of *At* NSP. Fractions eluting from the Ni(II)-nitriloacetate IMAC column in purification of recombinant SUMO-His_6 tagged At NSP. The most prominent at ~ 65 kDa represents the NSP monomer with the SUMO-His_6 tag.



Scheme S4-1. Schematic of SP autonomy experiment.

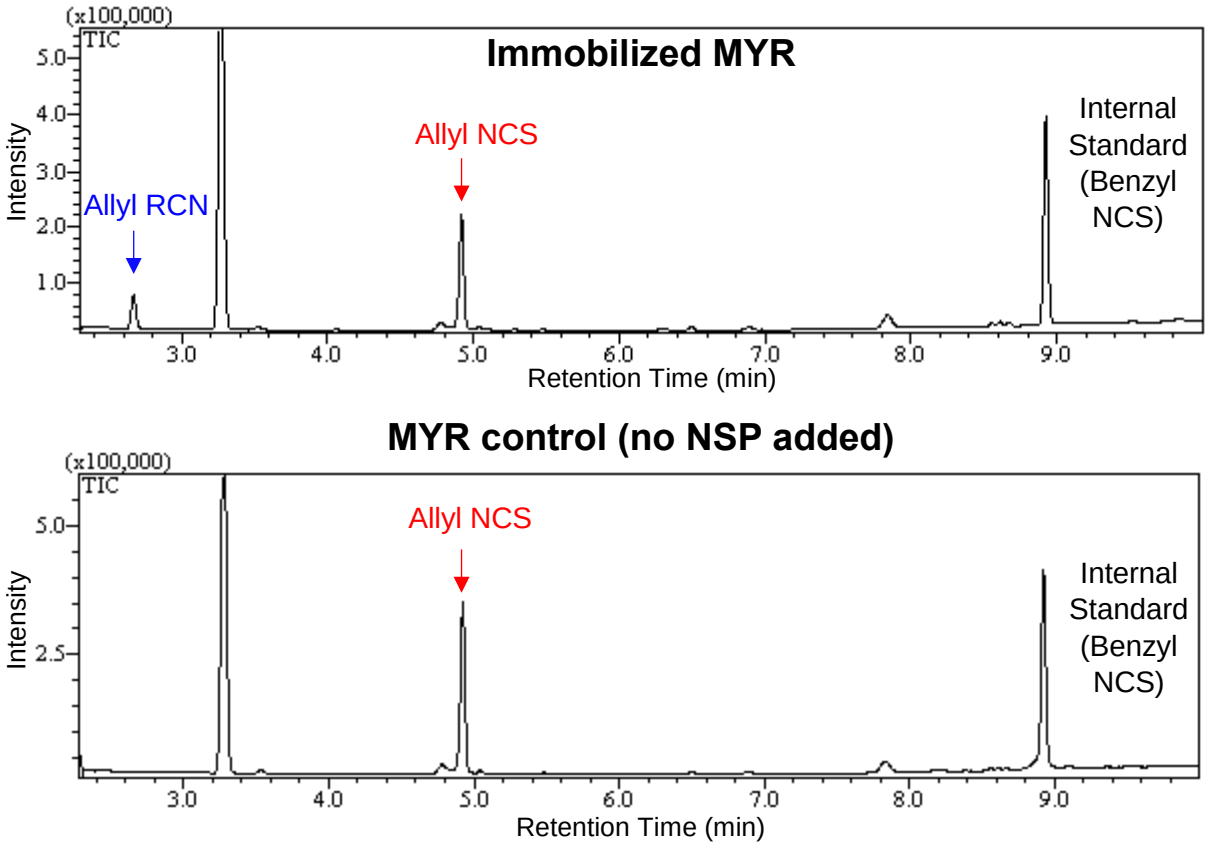


Figure S4-2. Chromatograms of NSP autonomy. (Large peak at 3.27 min is chlorocyclohexane contamination in the DCM used for sample extraction).

AtNSP1 pBA0028 plasmid DNA Sequence (NSP is capitalized). 5’ to 3’

gcgtagaggatcgagatctcgatcccgcgaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagaaataattttg
tttaactttaagaaggagatataccatgcatcatcaccatcaccacggttcttctatggctagcatgtcggactcagaagtcaatcaagaagcta
agccagaggtcaagccagaagtcaagcctgagactcacatcaatttaaaggtgtccgatggatcttcagagatcttcttcaagatcaaaaagac
cactcctttaagaaggctgatggaagcgttcgctaaaagacagggtaaggaaatggactccttaagattcttgtacgacggtattagaattcaa
gctgatcagacccctgaagatttggacatggaggataacgatattattgaggctcacagagaacagattggtgggatcgaggaaaacctgtact
tccaatccaatattctcgagATGGCCCAGAAATTGGAGGCGAAGGGCGGCGAAATGGGCGACGTTTGGGATGACG
GTGTTTATGAGAATGTGCGTAAAGTTTATGTTGGTCAGGCGCAGTACGGTATTGCCTTCGTGAAGTTCGAGTA
CGTCAATGGTTCCCAAGTTGTGGTCGGTGACGAGCACGGTAAGAAAACTGAGTTAGGTGTTGAGGAATTTGA
AATTGATGCGGACGACTACATTGTGTACGTGGAGGGTTATCGTGAGAAGGTCAATGACATGACCAGCGAGAT
GATTACGTTCCTCTCCATCAAAACTTTCAAAGGCAAAACCTCCCACCCAATCGAGAAACGTCCGGGCGTTAAG
TTTGTTTTGCACGGTGGTAAGATCGTGGGTTTTCATGGCCGCAGCACGGACGTATTGCACTCTTTGGGCGCTT
ACGTGTCGCTGTCGTCTACCATCAAACTGTTAGGCAAATGGATCAAGGTGGAACAAAAGGGCGAAGGTCCTG
GCCTGCGTTGCTCGCACGGCATTGCTCAGGTGGGCAACAAGATTTACTCCTTCGGCGGTGAATTTACCCCGAA
TCAACCGATTGACAAACATCTGTATGTGTTTGATCTGGAAACCCGCACCTGGTCCATTAGCCCGGCGACCGGC
GACGTTCCGCACCTGTCGTGCCTGGGTGTCAGAATGGTCAGCGTCGGTTCTACCCTGTATGTATTCGGCGGTA
GAGATGCGAGCCGCCAATACAACGGCTTCTATTCCTTTGATACGACCACCAATGAGTGGAAGTTGTTGACGCC
GGTAGAGGAAGGTCCGACTCCACGTAGCTTCCATAGCATGGCTGCGGATGAGGAAAACGTCTACGTGTTCGG
TGGCGTCAGCGCCACGGCACGCCTGAACACCCTGGACTCCTACAACATCGTTGATAAGAAATGGTTTCATTGC
TCGACGCCGGGTGACAGCCTGACGGCGCGTGGCGGTGCGGGTCTTGAGGTGGTTCAGGGCAAGGTGTGGGT
GGTTTACGGTTTTAACGGTTGTGAAGTTGACGACGTTCACTACTATGATCCGGTCCAAGATAAATGGACCCAG
GTCGAGACTTTCGGTGTGCGCCCGAGCGAGCGTTCTGTGTTTGCGAGCGCGGCAATCGGTAAGCATATCGTG
ATCTTCGGCGGTGAGATTGCAATGGACCCACTGGCTCATGTGGGTCCTGGTCAACTGACCGATGGGACCTTTG
CGCTGGACACCGAAACCTTGCAATGGGAGCGCCTGGATAAATTCGGCGGTGAGGAAGAAACGCCGTCGAGC
CGTGGTTGGACCGCGTCCACCACCGCAACCATTGACGGTAAGAAGGGTCTGGTAATGCACGGTGGCAAGGC
CCCGACCAACGACCGCTTCGATGACCTGTTCTTCTATGGCATTGATTCGGCGtaaggtacctgagatccggctgctaaca
aagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactagcataacccccttggggcctctaaacgggtcttgaggggttttttg
ctgaaaggaggaactatatccggattggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcgtga
ccgctacacttgccagcgccctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcggg
ggctccctttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgat
agacggtttttcgcccctttgacgttggagtccacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatctcggtctattct
tttgatttataagggattttgccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaacaaaatattaacgt
ttacaatttcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgaattaa
ttcttagaaaaactcatcgagcatcaaatgaaactgcaatttattcatatcaggattatcaataccatatttttgaaaaagccgtttctgtaatgaa
ggagaaaactcaccgaggcagttccataggatggcaagatcctggtatcggtctgcgattccgactcgtccaacatcaatacaacctattaattt
cccctcgtcaaaaataaggttatcaagtgagaaatcaccatgagtgacgactgaatccggtgagaatggcaaaagtttatgcatttctttccaga
cttgttcaacaggccagccattacgctcgtcatcaaaatcactcgcatcaaccaaaccgttattcattcgtgattgcgcctgagcgagacgaaat
acgcgatcgctgttaaaaggacaattacaaacaggaatcgaatgcaaccggcgcaggaacactgccagcgcatcaacaatattttcacctgaa
tcaggatattcttctaatacctggaatgctgttttcccggggatcgcagtggtgagtaaccatgcatcatcaggagtacggataaaatgcttgatg
gtcggaagaggcataaattccgtcagccagtttagtctgaccatctcatctgtaacatcattggcaacgctacctttgccatgtttcagaaacaac
tctggcgcatcgggcttcccatacaatcgatagattgtcgcacctgattgcccgacattatcgcgagcccatttatacccatataaatcagcatcc
atgttggaatttaatcgcggcctagagcaagacgtttcccgttgaatatggctcataacacccttgtattactgtttatgtaagcagacagtttta
ttgttcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttttct
gcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaa

ctggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatac
ctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcg
cagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatga
gaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccag
ggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatg
gaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataa
ccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcct
gatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatatggtgcactctcagtacaatctgctctgatgccgcatagttaagcca
gtatacactccgctatcgctacgtgactgggtcatggctgcgccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctccc
ggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgaggcagctgcgg
taaagctcatcagcgtggtcgtgaagcgattcacagatgtctgcctgttcatccgcgtccagctcgttgagtttctccagaagcgttaatgtctggc
ttctgataaagcgggccatgttaagggcggttttttcctgtttggtcactgatgcctccgtgtaagggggatttctgttcatgggggtaatgatacc
gatgaaacgagagaggatgctcacgatacgggttactgatgatgaacatgcccggttactggaacgttgtgagggtaaacaactggcggtatg
gatgcggcgggaccagagaaaaatcactcagggtcaatgccagcgcttcgttaatacagatgtaggtgttccacagggtagccagcagcatcc
tgcgatgcagatccggaacataatggtgcagggcgctgacttccgcgtttccagactttacgaaacacggaaaccgaagaccattcatgttgttg
ctcaggtcgcagacgttttgcagcagcagtcgcttcacgttcgctcgcgtatcggtgattcattctgctaaccagtaaggcaaccccgccagccta
gccgggtcctcaacgacaggagcacgatcatgcgcacccgtggggccgccatgccggcgataatggcctgcttctcgccgaaacgtttggtggc
gggaccagtgacgaaggcttgagcgagggcgtgcaagattccgaataccgcaagcgacaggccgatcatcgtcgcgctccagcgaaagcggt
cctcgccgaaaatgacccagagcgctgccggcacctgtcctacgagttgcatgataaagaagacagtcataagtgcggcgacgatagtcatgc
cccgcgcccaccggaaggagctgactgggttgaaggctctcaagggcatcggtcgagatcccggtgcctaatgagtgagctaacttacattaat
tgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgta
ttgggcgccagggtggtttttctttcaccagtgagacgggcaacagctgattgcccttcaccgcctggccctgagagagttgcagcaagcggtcc
acgctggtttgccccagcaggcgaaaatcctgtttgatggtggttaacggcgggatataacatgagctgtcttcggtatcgtcgtatcccactacc
gagatatccgcaccaacgcgcagcccggactcggtaatggcgcgcattgcgcccagcgccatctgatcgttggcaaccagcatcgcagtggga
acgatgccctcattcagcatttgcatggtttgttgaaaaccggacatggcactccagtcgccttcccgttccgctatcggctgaatttgattgcgag
tgagatatttatgccagccagccagacgcagacgcgccgagacagaacttaatgggcccgctaacagcgcgatttgctggtgacccaatgcga
ccagatgctccacgcccagtcgcgtaccgtcttcatgggagaaaataatactgttgatgggtgtctggtcagagacatcaagaaataacgccgg
aacattagtgcaggcagcttccacagcaatggcatcctggtcatccagcggatagttaatgatcagcccactgacgcgttgcgcgagaagattg
tgcaccgccgctttacaggcttcgacgccgcttcgttctaccatcgacaccaccacgctggcacccagttgatcggcgcgagatttaatcgccgcg
acaatttgcgacggcgcgtgcagggccagactggaggtggcaacgccaatcagcaacgactgtttgcccgccagttgttgtgccacgcggttgg
gaatgtaattcagctccgccatcgccgcttccacttttcccgcgttttcgcagaaacgtggctggcctggttcaccacgcgggaaacggtctgat
aagagacaccggcatactctgcgacatcgtataacgttactggtttcacattcaccaccctgaattgactctcttccgggcgctatcatgccatac
cgcgaaaggttttgcgccattcgatggtgtccgggatctcgacgctctcccttatgcgactcctgcattaggaagcagcccagtagtaggttgagg
ccgttgagcaccgccgccgcaaggaatggtgcatgcaaggagatggcgcccaacagtcccccggccacggggcctgccaccatacccacgcc
gaaacaagcgctcatgagcccgaagtggcgagcccgatcttccccatcggtgatgtcggcgatataggcgccagcaaccgcacctgtggcgcc
ggtgatgccggccacgatgcgtccg

VITA

Leanne E. Mocniak

EDUCATION

2016 – 2021 Ph.D. in Chemistry
The Pennsylvania State University, University Park, PA
Co-Advisor: Adjunct Professor Ray Bryant
Co-Advisor: Professor J. Martin Bollinger, Jr.

2012 – 2016 B.S. in Biochemistry, *magna cum laude*
Seton Hall University, South Orange, NJ
Advisor: Professor Nicholas Snow

HONORS & AWARDS

2016 NJ-ACS Undergraduate Research Conference, Third Place
2014 – 2016 Clare Boothe Luce Undergraduate Scholarship
2013 – 2015 Clare Boothe Luce Summer Research Fellowship

PROFESSIONAL EXPERIENCE

2019 – 2021 GWIS Conference Committee Member, *The Pennsylvania State University*
2016 – 2017 Graduate Teaching Assistant, *The Pennsylvania State University*
2015 – 2016 Undergraduate Teaching Assistant, *Seton Hall University*
2013 – 2016 Student Instructor, *Seton Hall University*
2015 – 2016 Special Intern, *State of New Jersey Division of Criminal Justice, Northern Regional Medical Examiner*

PUBLICATIONS

Mocniak, L. E., Slater, J., Elkin, K.R., Bollinger, J. M. Jr. Assessment of the Side-chain Selectivities and Catalytic Efficiencies of Two Specifier Proteins from the Glucosinolate-Myrosinase Defense System of *Arabidopsis Thaliana*. *In preparation.*

Mocniak, L.E., Elkin, K.R., Dillard, S.L., Bryant, R.B., Soder, J.K. Simple identification and accurate mass confirmation of individual glucosinolates in brassica varieties by mass spectrometry. *In preparation*.

Mocniak, L. E., Elkin, K.R., Bollinger, J. M. Jr. (2020). Lifetimes of the Aglycone Substrates of Specifier Proteins, the Autonomous Iron Enzymes That Dictate the Products of the Glucosinolate-Myrosinase Defense System in Brassica Plants. *Biochemistry 59*, 2432–2441.