Friedrich Schiller Universität Jena
Biologisch-Pharmazeutische Fakultät

Max Planck Institut für Chemische Ökologie
Abteilung für Biochemie

Identification and characterization of glucosinolate sulfatase performing the detoxification in the phloem feeding insect- *Bemisia tabaci*

Masterarbeit
zur Erlangung des Grades eines
Master of Science

vorgelegt von

Abinaya Manivannan

aus Indien

Jena, January 2018

Betreuer:
Prof. Dr. Jonathan Gershenzon
Prof. Dr. Wilhelm Boland

List of Abbreviations

ACN	Acetonitrile
BLAST	Basic Local Alignment Search Tool
cDNA	complementary DNA
ddH2O	Double distilled water
ddNTP	Didesoxyribonucleotide triposphate
DMSO	Dimethyl sulfoxide
DNA	Desoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
EIC	Extracted ion chromatogram
FA	Formic acid
GSH	Glutathione
GSS	Glucosinolate sulfatase
GST	Glutathione S-transferase
HPLC	High pressure liquid chromatography
i3m	indol-3-ylmethyl
ITC	Isothiocyanates
LB	Lysogeny broth
LC-MS	Liquid chromatography–mass spectrometry
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
MEAM1	Middle East–Asia Minor 1
MED	Mediterranean
4moi3m	4-methoxyindol-3-ylmethyl
MRM	Multiple Reaction Monitoring
4mtb	4-methylthiobutyl
4msob	4-methylsulfinylbutyl
7msoh	7 methylsulfinylheptyl

8msoo	8-methylsulfinyloctyl
5msop	5-methylsulfinylpentyl
3msop	3-methylsulfinylpropyl
Ni-NTA	Nickel- Nitrilotriacetic acid
NMR	Nuclear magnetic resonance
NSP	Nitrile specifying protein
PCR	Polymerase chain reaction
pOHBz	*p*-hydroxybenzyl
PVDF	Polyvinylidene flouride
RNA	Ribonucleic acid
RNAi	RNA interference
rpl-13	Ribosomal protein L-13
rpm	Revolutions per minute
RT-qPCR	Real-time quantitative PCR
SDS	Sodium dodecyl sulfate
SG	Sinigrin
SUMF1	Sulfatase-modifying factor
U	Units

List of Tables

List of Figures

Contents

1. Zusammenfassung

Kreuzblütler gehören zur Ordnung Brassicales und stellen weltweit relevante Vertreter gemüseartiger Nahrungspflanzen dar. Vertreter dieser Ordnung nutzen zur chemischen Abwehr Glukosinolate, eine Gruppe pflanzlicher Abwehrmetabolite, die als glucosylierte Protoxine gespeichert werden. Diese werden räumlich getrennt von denen sie aktivierenden β-Thioglucosidasen, genannt Myrosinasen, gelagert. Gewebeverletzungen, hervorgerufen durch z. B. Herbivorie, führen zur Aufhebung der Kompartimentierung der zwei Komponenten, was eine Hydrolyse der Glucosinolate in Downstream-Produkte zur Folge hat, welche toxisch für Insekten, Pilze, Nematoden und Bakterien sein können. Trotz dieses effizienten Zwei-Komponenten-Abwehrsystems, haben einige Herbivoren Strategien entwickelt, um Kreuzblütler als Wirtspflanzen nutzen zu können. Ein für kauende herbivore Insekten bekannter Entgiftungsmechanismen basiert auf Glucosinolat Sulfatase (GSS) Enzymen, die Glucosinolate in Desulfoglucosinolate umwandeln, die daraufhin effizient ausgeschieden werden können. Kürzlich wurde beschrieben, dass *Bemisia tabaci*, ein phloem-fressender Insektenschädling, Desulfoglucosinolate produziert und ausscheidet. Zuvor wurde angenommen, dass stechend-saugende Herbivore nur minimale Gewebeverletzungen hervorrufen können. Eine wachsende Datenlage weist jedoch darauf hin, dass Zwei-Komponenten-Abwehrsysteme trotz allem durch Fraß dieser Insekten aktiviert werden können.

Ziel der vorliegenden Arbeit war die Identifizierung und Charakterisierung der GSS, die für den Entgiftungsmechanismus bei *B. tabaci* verantwortlich ist. Putative Genkandidaten wurden durch phylogenetische Analysen and anhand von Expressionsspiegeln in ganzen Insekten und Darmproben identifiziert und heterolog in Insektenzellen exprimiert. Es stellte sich heraus, dass die identifizierte GSS eine explizite Bevorzugung von indolischen Glucosinolaten gegenüber aliphatischen und benzenischen Glucosinolaten zeigte. Dies impliziert, dass *B. tabaci* GSS-Entgiftungsmechanismen nutzt, um den negativen Effekten des Myrosinase-unabhängigen Abbaus von indolischen Glucosinolaten entgegenzuwirken. Die Ergebnisse der vorliegenden Arbeit erweitern das Wissen zu Glucosinolat-Entgiftungsmechanismen bei phloem-fressenden Insekten.

2. Abstract

Cruciferous plants belonging to the order Brassicales are a dominant vegetable food crop worldwide. Plants from this order defend themselves chemically using glucosinolates, a class of plant defensive metabolites that are stored as glucosylated pro-toxins segregated from their activating β-thioglucosidases, known as myrosinases. Upon tissue rupture, such as during herbivory, the compartmentalisation of these two components is compromised, leading to the hydrolysis of glucosinolates into downstream products that are toxic to insects, fungi, nematodes and bacteria. Despite this efficient two-component defence system, some herbivores have developed strategies to circumvent it and successfully feed on cruciferous plants. One such detoxification mechanism previously reported in chewing insect herbivores relies on glucosinolate sulfatase (GSS) enzymes that convert glucosinolates into desulfoglucosinolates, which are subsequently excreted. Recent studies revealed that *Bemisia tabaci*, a phloem-feeding insect that is a major agricultural pest, also produces and excretes desulfoglucosinolates. This was initially surprising, as the piercing-sucking herbivory mode has long been thought to cause only minimal rupture of the plant tissues, but a growing amount of evidence suggests that two-component defences are still activated during feeding by these insects.

In this study, we aim to identify and characterise a GSS performing glucosinolate detoxification in *B. tabaci*. Putative gene candidates were selected by phylogenetic analysis and expression levels in whole insects and gut tissues, and were heterologously expressed in insect cells. We found a GSS enzyme showing an explicit preference towards indolic glucosinolates in comparison to aliphatic and benzenic glucosinolates, implying that *B. tabaci* employs this GSS detoxification mechanism to cope with the negative effects of myrosinase-independent breakdown of indolic glucosinolates. Our findings expand the knowledge on the biochemistry of glucosinolate detoxification mechanisms present in phloem-feeding insects.

3. Introduction

Plants rely on a diverse arsenal of chemicals to defend themselves against herbivores and pathogens. One particularly successful strategy to safely accumulate large concentrations of defensive compounds while preventing auto-toxicity problems is to use two-component activated defences, for example glucosinolates. These glucosylated substances are stable and only become toxic after enzymatic activation upon attack. Such a complex multi-component strategy, however, offers multiple targets for counter-adaptations from potential herbivores, who can either prevent or redirect activation, or simply limit toxicity metabolically. On the other hand, another effective adaptation against activated defences is best employed by piercing-sucking, phloem-feeding insects such as aphids and whiteflies. These herbivores stealthily penetrate plant tissues to feed on the sugar-rich phloem sap, causing only minimal damage to surrounding cells and therefore limiting plant defensive responses. Nevertheless, even this clever feeding behaviour still leads to some activation of glucosinolates to form toxic isothiocyanates (ITC), so that detoxification strategies preventing activation can still be greatly advantageous to phloem feeding insects. In this thesis, we examine in further detail one such biochemical mechanism, the desulfation of glucosinolates performed by the whitefly *Bemisia tabaci*, a generalist pest of multiple food and ornamental plant crops that can feed on glucosinolate-producing crucifer plants.

3.1 Glucosinolates

Glucosinolates are sulphur containing specialized ("secondary") defensive metabolites restricted to the order Brassicales, which includes the plant families Brassicaceae and Capparaceae (Ishida et al. 2004). The Capparaceae family consists of nearly 40–45 genera and 700–900 species (Jocelyn et al. 2002). The Brassicaceae plant family (mustard family) includes many economically important oilseed, vegetable, condiment, and fodder crops such as *Brassica rapa* (Chinese cabbage, Chinese mustard, bok choy and turnip), *B. oleracea* (cabbage, broccoli, cauliflower, kale, Brussels sprouts and kohlrabi), *B. napus* (rapeseed and rutabaga), *B. juncea* (mustard green), and *Raphanus sativus* (radish), as well as the model plant *Arabidopsis thaliana*. These plants have peculiar tastes and smells, which are conferred in great part by hydrolysis products from glucosinolates, also called mustard oils (Berry 2015; Engel et al. 2006). These mustard oil-containing plants find use not only as food for humans, but also as biofumigants in plant-parasitic management and as bioherbicides targeting weed seeds (Brown and Morra 1995; Zasada et al. 2004). Apart from their agricultural significance,

glucosinolates and their hydrolysis products are also recognized as cancer-preventive agents in various animal models (Ávila et al. 2013). Studies have additionally shown that the ITC deriving from glucosinolates hydrolysis possess antimicrobial activities (Fahey et al. 2013).

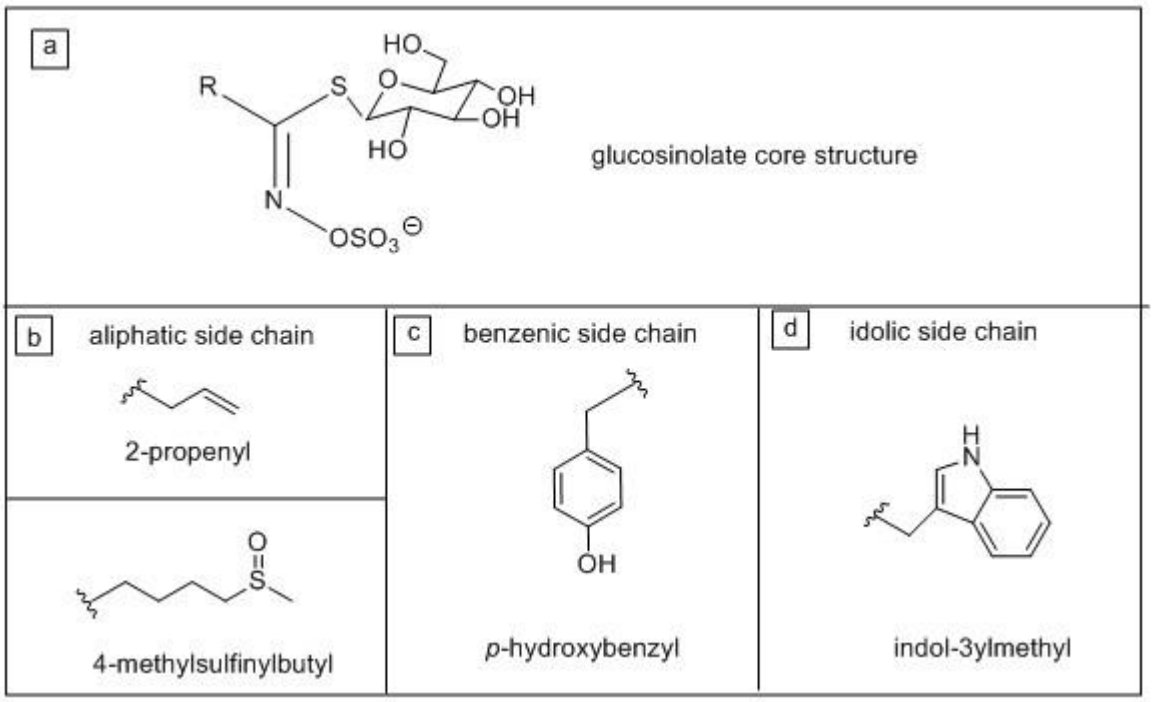


Figure 1: a) Core chemical structure of the glucosinolates b) Chemical structure of aliphatic glucosinolates: 2-propenyl glucosinolate and 4-methylsulfinylbutyl glucosinolate c) Chemical structure of benzenic glucosinolate: p-hydroxybenzyl glucosinolate d) Chemical structure of indolic glucosinolate: indol-3-ylmethyl glucosinolate

The core chemical structure of the glucosinolates (Figure 1) consists of a β-thioglucose residue, a sulfonated oxime moiety, and a variable R group derived from one of 8 amino acids (Ala, Leu, Ile, Met, Val, Phe, Tyr and Trp) (Fahey et al. 2002; Halkier and Gershenzon 2006). Approximately 200 glucosinolates have been described based on the variability at the R group, and are classified as aliphatic, benzenic and indolic glucosinolates according to the precursor amino acid from which they are derived (Ishida et al. 2004). Glucosinolates derived from Ala, Leu, Ile, Met, or Val are called aliphatic glucosinolates, those which are derived from Phe or Tyr are called benzenic glucosinolates, and the ones derived from Trp are called indolic glucosinolates (Halkier and Gershenzon 2006). The side chains of aliphatic and benzenic glucosinolates can undergo elongation by a series of steps of 1-carbon addition. The biosynthesis of the glucosinolate core structure occurs by the following steps: conversion of amino acids to aldoximes by cytochrome P450s of the CYP79 family, followed by formation of activated, oxidised aldoximes by CYP83; transformation of activated aldoximes to thiohydroximic acids mediated by GSTs and a *C-S* lyase; conversion of the thiohydroximic

acids to desulfoglucosinolates mediated by *S*-glucosyltransferase belonging to the UGT74 family, and finally sulfation by sulfotransferases (Ishida et al. 2004; Halkier and Gershenzon 2006; Fahey et al. 2002). The glucosinolates thus produced can be subsequently modified by oxygenation, alkylation, acylation, and glucosylation, giving rise to increased structural variation (Ishida et al. 2004).

3.2 Glucosinolate–myrosinase system

In addition to glucosinolates, cruciferous plants contain thioglucosidases known as myrosinases, which catalyse the hydrolysis of glucosinolates (Rask et al. 2000). Both entities are physically segregated from one another in the plant system by compartmentalisation (Höglund et al. 1991; Grob and Matile 1979). While the localization of each component in the intact plant is still controversial, myrosinases have been reported as being localised in protein-accumulating idioblasts called myrosin cells, present in the phloem parenchyma. Glucosinolates, on the other hand, are enriched in sulphur-containing S-cells, and would thus be separated from the myrosinases at cellular level (Koroleva and Cramer 2011). However, in other spatial separation models the glucosinolates are confined into vacuoles, namely myrosin grains, and the myrosinases are present in cytoplasm of the same myrosin cells, enabling sub-cellular compartmentalisation of the two components (Lüthy and Matile 1984).

Upon tissue disruption, such as during herbivory, the compartmentalisation of the two components of the glucosinolate–myrosinase system is compromised, and the myrosinase actively hydrolyses the glucose moiety of the glucosinolate (Rask et al. 2000). The products formed are a free glucose and a highly unstable aglucone, which then readily rearranges to form toxic products such as ITC, thiocyanates, nitriles and others (Figure 2). The toxicity of the ITC is attributed by the lipophilic nature of the side chain and the electrophilic nature of the ITC group. Because of these two toxicity attributes, ITCs easily cross the lipid bilayer membranes in the insect's gut epithelium and reach the intracellular environment and the hemolymph, reacting with nucleophilic biological targets (Kawakishi and Kaneko 1987). Thus, the two-component glucosinolate-myrosinase system, often referred to as "mustard oil bomb", serves as an important chemical defence system against herbivores (Matile 1980; Mauricio and Rausher 1997; Siemens and Mitchell-Olds 1998). Plants can accumulate large (millimolar) concentrations of intact, pro-toxic glucosinolates that are only detonated upon herbivory, causing negative to the consuming insects.

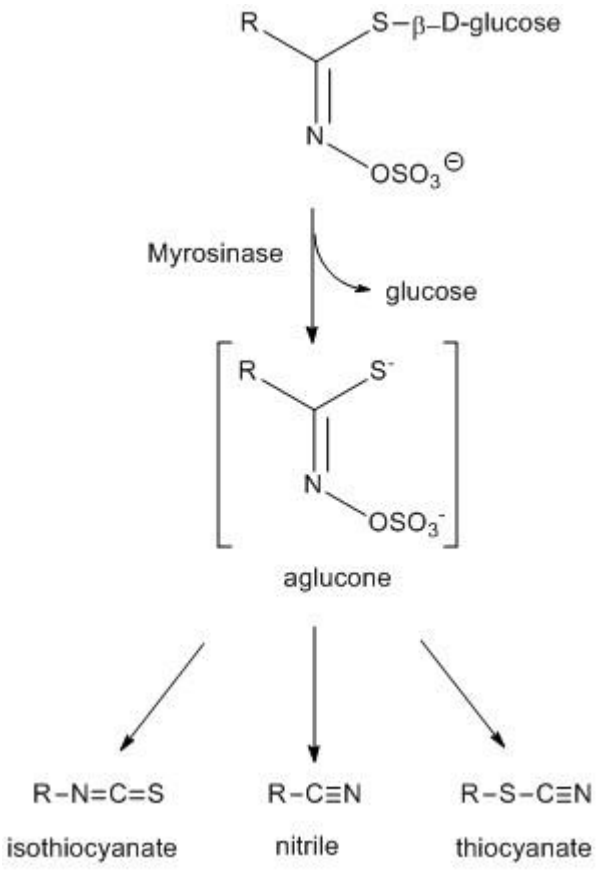


Figure 2: Reaction catalysed by plant myrosinases. Myrosinases remove glucose from glucosinolates forming an unstable aglucone, which rearranges forming products such as isothiocyanates, nitriles and thiocyanates.

3.3 Insect counteradaptation to glucosinolate-myrosinase system

Several herbivores, however, have developed their own means to detoxify plant defensive compounds and successfully feed on the plants. A few mechanisms have been described for the detoxification of glucosinolates by certain chewing herbivores. The generalist *Spodoptera littoralis* uses glutathione *S*-transferases to form hydrophilic glutathione-ITC conjugates, enabling efficient excretion of the toxic compounds (Jeschke et al. 2016). The larvae of *Pieris rapae*, a crucifer-specialist feeder, use a "nitrile specifying protein" (NSP) for directing the formation of nitriles during glucosinolate hydrolysis in their gut, thus preventing the formation of ITC, which are more toxic (Wittstock et al. 2004). Another enzymatic detoxification mechanism used by another crucifer-specialist, the diamondback moth larvae (*Plutella xylostella*), as well as the generalist desert locust (*Schistocerca gregaria*) and silverleaf whitefly (*B. tabaci*), is the enzyme glucosinolate sulfatase (GSS) that cleaves off the sulfate group of glucosinolates to form desulfoglucosinolates, essentially reversing the last step in the biosynthesis of the glucosinolate core structure. Since the desulfoglucosinolates can no longer

be used as substrates by myrosinases, no toxic downstream products are produced (Falk and Gershenzon 2007; Ratzka et al. 2002) (Figure 3).

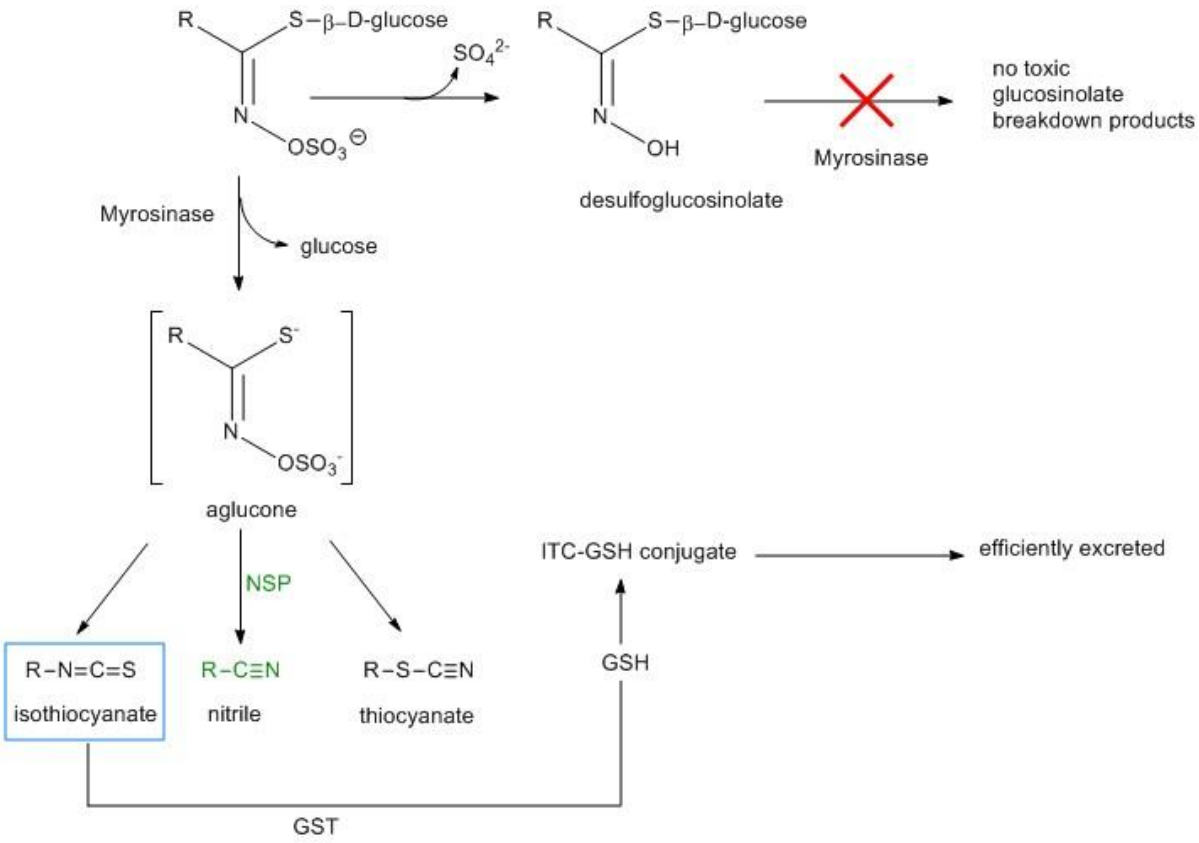


Figure 3: Detoxification mechanisms of glucosinolate-myrosinase defence system in specific insect herbivores. (ITC- isothiocyanate; GST- glutathione S-transferase; GSH- glutathione).

3.4 *Bemisia tabaci*

The silverleaf whitefly (*B. tabaci*, order Hemiptera), also referred to as sweet potato whitefly, is a phloem-feeding insect that thrives mainly in tropical and subtropical regions, and less predominantly in temperate regions. It has 6 life stages: the egg, four nymphal stages, and the adult stage (Brown et al. 1995).

B. tabaci is a cryptic species complex, existing as a conglomerate of nearly 36 previously identified biotypes (Perring 2001). A recent phylogenetic analysis based on their genetic variability delineated them to 24 morphologically indistinguishable species (Perring 2001), but the molecular markers and parameters used for species differentiation in this complex are still not widely accepted. The most invasive species are the MED biotype (Mediterranean, formerly named Q), indigenous to Mediterranean region, and the MEAM1 biotype (Middle East–Asia Minor 1, formerly B), originating in the Middle East–Asia Minor region (McKenzie et al. 2012;

Perring 2001). The biotype MED has emerged as a more serious pest species due to its greater virus transmission efficiency and higher insecticide resistance (Luo et al. 2010; Liu et al. 2013).

It is an extreme generalist with a wide host plant range, attacking and damaging nearly 500 plant species from over 60 plant families (Greathead 1986). *B. tabaci* is one of the most serious agricultural pests in the world, causing three types of damage to the crop plants: direct damage, indirect damage, and virus transmission. The silverleaf whitefly feeds on the crop plants by piercing and sucking the phloem sap content. Such a feeding mechanism inflicts a direct damage to the host plant, draining nutrients and causing weakening and early plant wilting, leaf chlorosis and withering, premature dropping of leaves and even plant death (Berlinger 1986). The damage reduces plant growth rates, and thereby decreases the crop yield. Additionally, during feeding on the nutritionally imbalanced (extremely sugar-rich) phloem sap, the whiteflies excrete excess carbohydrates as a sticky-sugary substance called "honeydew" that deposits on the plant surface. The accumulation of whitefly honeydew serves as a thriving nutrient source for black sooty mold (*Capnodium*), forming a black coating on the plant surface (Berlinger 1986). The fungus does not cause any direct harm to the plant; instead, as a result of blocking incoming sunlight, it inflicts an indirect damage by reducing photosynthesis. The black sooty deposition also decreases the market value of the vegetables. The third type of damage derives from *B. tabaci* serving as a potent vector for plant pathogenic viruses from the geminivirus group (Jones 2003). They are known to cause approximately 40 different plant diseases, imparting considerable damage to the crop plants (Cohen and Berlinger 1986).

With such a wide host range, high potential to develop insecticide resistance and being a serious agricultural pest, the study of *B. tabaci* in detail has quickly increased in importance, motivated by the need for efficient pest control methods.

3.5 Glucosinolate desulfation in *B. tabaci*

One of the biochemical mechanisms reported in the crucifer specialist diamondback moth (*P. xylostella*) and the generalist desert locust (*Schistocerca gregaria*) to detoxify the glucosinolate-myrosinase plant defence system is an enzyme called glucosinolate sulfatase (GSS). This enzyme converts glucosinolates into desulfoglucosinolates, thereby rendering it not activatable by the plant myrosinases, and therefore useless as a plant defence. A recent study in our research group has revealed that the *B. tabaci* also desulfates glucosinolates, with the resulting desulfoglucosinolate products excreted in the honeydew (Malka et al. 2016). This transformation was not observed in other phloem feeders examined (the green peach aphid

Myzus persicae and the cabbage whitefly *Aleyrodes proletella* – unpublished data), suggesting it is not widespread through all hemipteran insects. A point to be noted here is that, unlike the previously reported chewing insects (diamondback moth larvae and the desert locust) which use this enzymatic desulfation mechanism, *B. tabaci* is a piercing-sucking phloem-feeding insect. The phloem-feeding insects have a specialized mode of feeding. They pierce the plant surface using their stylets and carefully penetrate the tissues to reach to the phloem sieve elements, avoiding major activation of wound responses in the host plant (Walling 2008). Because of only minimal mechanical damage, it has been argued that they do not strongly activate the plant β-glucosidases associated with glucosylated defence metabolites (Pentzold et al. 2014). For instance, some glucosinolates are also excreted intact in the honeydew from the green peach aphid (*M. persicae*) (Barth and Jander 2006) and from *B. tabaci* (Malka et al. 2016). However, recent evidence indicates that the piercing-sucking feeding strategy still leads to detectable glucosinolate activation (Danner et al., 2017), and glucosinolates ingested by aphids can still be activated independently of the classic plant myrosinases (Kim and Jander 2007; Kim et al. 2008). Therefore, pre-emptive deactivation of glucosinolates, such as via GSS activity, might still confer benefits to phloem-feeding insects as well.

3.6 Glucosinolate sulfatase (GSS)

The GSS encoding gene and enzyme in *P. xylostella* have been successfully identified and characterised (Ratzka et al. 2002). GSS belongs to the aryl sulfatase enzyme family (Ratzka et al. 2002). This class of enzymes have a highly conserved consensus motif C/S-X-P-X-A-X4-T-G (Kertesz et al. 2000). The lead residue is always a cysteine in eukaryotic systems, while in the bacterial system it is either a cysteine or a serine residue. In order to be catalytically active, sulfatases need to undergo a post-translational modification by a sulfatase-modifying factor (encoded by *SUMF1*) (Buono and Cosma 2010). This gene encodes for a formylglycine-generating enzyme, which specifically binds and modifies the conserved cysteine or serine to a catalytically active aldehyde, α-formylglycine (Dierks et al. 1999). In *P. xylostella*, the *GSS* and the *SUMF1* are tightly co-expressed, and help this insect to be a successful herbivore of cruciferous plants (You et al. 2013). The *P. xylostella* GSS-encoding gene is reportedly expressed only in the gut tissues and is also under developmental control, with expression observed only in the larval instars and absent in eggs, pupae and adults (Ratzka et al. 2002).

4. Objective

In the plant order Brassicales, the glucosinolate-myrosinase two-component chemical defence system is used to repel attacking pathogens and predators. Among the strategies developed by crucifer-feeding insects is the GSS enzyme, that converts glucosinolates into desulfated products which are no longer used as a substrate by myrosinases. Only recently, the phloem-feeding insect silverleaf whitefly (*B. tabaci*) was found to efficiently excrete desulfoglucosinolates when reared on glucosinolate-containing diets.

This study aims to identify and characterise the GSS enzyme(s) performing glucosinolate detoxification in *B. tabaci*, by cloning, heterologously expressing and biochemically characterising putative *B. tabaci* arylsulfatases. A substrate screening with glucosinolate mixtures will be used to investigate whether the *B. tabaci* GSS acts preferentially on distinct glucosinolate classes. Additionally, once the *B. tabaci* GSS-encoding gene is identified, we will examine the evolutionary relationship between the chewing insect *P. xylostella* GSS and the phloem-feeding insect *B. tabaci* GSS.

5.Materials and methods

5.1 Used donor and recipient organisms

Bemisia tabaci	Strain Middle East-Asia Minor 1 (MEAM1, formerly known as biotype B), maintained in the Dept. of Entomology, The Hebrew University of Jerusalem, Israel
Escherichia coli	NEB® 10-beta Competent *E. coli* (High Efficiency), NewEngland Biolabs® Inc.
Spodoptera frugiperda	Sf9™ cells in Sf-900 II SFM (1X), Gibco® by Life Technologies (Darmstadt, Germany)

5.2 Plants and insects

Bemisia tabaci (MEAM1) was reared on eggplant (*Solanum melongena*) and Brussels sprout (*Brassica oleraceae*) plants. *Arabidopsis thaliana* Col-0 was cultivated in a controlled-environment growth chamber under short day conditions (9.5:13.5 h, L:D, 100% light intensity, 21°C, 50-60% relative humidity).

5.3 Primers

All primers were obtained from Sigma-Aldrich and 100 µM of stock solutions were prepared in ddH2O according to the manufacturer's instruction. Working concentration of the primers (10 µM) were prepared by diluting the stock solutions. All primers used are listed in section (9.2).

5.4 Chemicals

All chemicals used in this work are listed in section (9.1). MilliQ grade water (ddH2O) was obtained using the Milli-Q® Integral Water Purification System (Merck Millipore, Darmstadt, Germany).

5.5 Sf9 cell culture

Spodoptera frugiperda Sf9 cells (Gibco R by Life Technologies,Darmstadt, Germany) were cultured in Sf- 900 II serum-free medium. The adherent cell cultures were maintained at 27°C and subcultured every 3 - 4 days.

5.6 Primer designing

The gene sequences of the putative *B. tabaci* MEAM1 glucosinolate sulfatases (Bta02222, Bta04774, Bta03550, Bta06756, Bta14666, Bta14667) were obtained from the Whitefly genome database (whiteflygenomics.org) and were used for the design of primers for the amplification of the respective genes. Forward primers included the Kozak consensus sequence (Kozak, 1987; Kozak, 1991; Kozak, 1990) prior to the start codon. Reverse primers were devoid of stop codon to have a C-terminal fusion with V5 epitope and His-tag using the pIB/V5-His TOPO® TA vector system (Invitrogen™). The design of the primers was carried out using the Geneious software (version 10.0.5).

5.7 Amplification of genes by Polymerase Chain Reaction (PCR)

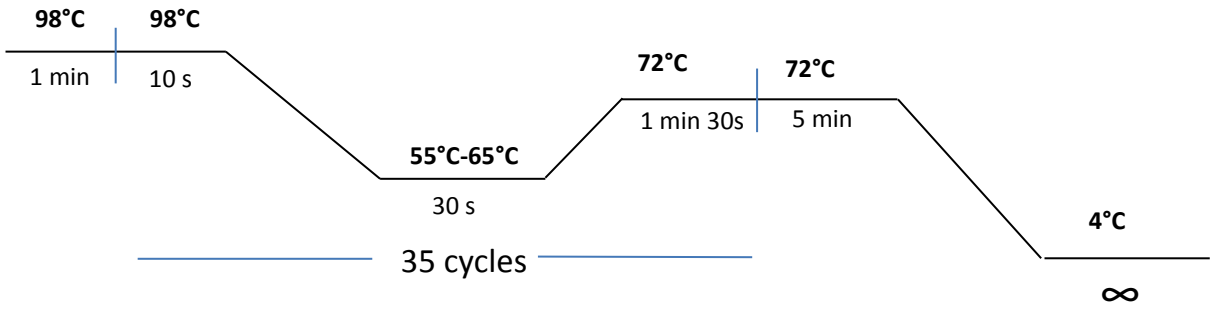
Interesting *B. tabaci* arylsulfatase gene candidates were picked based on the BLAST and phylogenetic analysis results (Bta02222, Bta04774, Bta03550, Bta06756, Bta14666, Bta14667) and were amplified using Phusion HF Polymerase (Thermo Scientific, Schwerte, Germany). The *B. tabaci* cDNA (prepared by Dr. Osnat Malka, THUJ, Israel) was used as a template for the amplification of the full-length above-mentioned genes. The list of primers is available in (Table S1).

The following protocol was used for the PCR:

Component	**Volume (µL)**
ddH_2O	13.25
5X HF Phusion buffer	5.0
dNTP mix (2.5 mM)	0.5
Forward primer (10 µM)	2.5
Reverse Primer (10 µM)	2.5

cDNA template (50 ng/µL)	1.0
Phusion DNA Polymerase	0.25
Total	**25.0**

The program used in the Biometra TGradient Thermocycler (Goettingen, Germany) for the amplification was as follows:



The annealing temperatures for each of the genes were determined by performing gradient PCR using GoTaq® DNA Polymerase (Promega, Mannheim, Germany).

5.8 Gel extraction of DNA fragments

The amplified PCR products were resolved on 1 % (w/v) agarose gel in 1x TAE buffer (40 mM Tris, 150 mM EDTA, 20 mM glacial acetic acid), containing 0.005 % (v/v) Midori Green (Nippon Genetics Europe GmbH, Dueren, Germany) and visualised at 254 nm using the GeneGenius Bio-Imaging System (Syngene, Cambridge, UK). The DNA fragments corresponding to the expected sizes were excised from the gel and transferred to a microcentrifuge tube. The DNA fragments were then extracted from the gel using either QIAquick gel extraction kit (Qiagen, Germany) or the Zymoclean™Gel DNA Recovery Kit (ZymoResearch, Freiburg im Breisgau, Germany) as per the manufacturer's instruction.

5.9 Determination of DNA concentrations

DNA concentrations were determined by measuring the absorbance at 260 nm using a NanoDrop 2000c Spectrophotometer. The NanoDrop 2000c Operating Software (version 1.6) calculated the DNA concentrations as A260 nm (absorbance unit) = 1 corresponding to 50 ng/µl of DNA. The purity of the DNA samples was confirmed by ensuring the absorbance ratios 260/280 and 260/230 were close to 1.8 and 2.0 respectively.

5.10 Cloning of putative glucosinolate sulfatase genes in *Escherichia coli*

3' A-overhangs were added to the purified gene amplicons (Bta02222, Bta04774, Bta03550, Bta06756, Bta14666, Bta14667) using Go®Taq DNA Polymerase (Promega, Mannheim, Germany) to enable cloning into the pIB/V5-His TOPO® TA expression vector (Invitrogen™). The following reaction mixture was made and incubated at 72°C for 20 min for the addition of 3' A-overhangs:

Components	**Volume (µL)**
Purified PCR product	15
5x GoTaq Colourless Reaction Buffer	4.0
dNTP mix (2.5 mM)	0.6
GoTaq DNA polymerase	0.4
Total	**20.0**

10 ng of the gene amplicon with 3' A-overhangs was added to 0.5 µL of salt solution and 0.5 µL of the pIB/V5-His TOPO® TA expression vector and the total volume was adjusted to 3 µL using ddH2O, followed by overnight incubation at room temperature. The reaction mixture was then added to 50 µL of NEB 10-beta competent *Escherichia coli* cells and was incubated on ice for 30min. The incubation was followed by a heat-shock treatment, where the competent cells were put in a 42°C water bath for 30 s and snap cooled on ice for 5 min. Later, 950 µL of room temperature NEB 10-beta/Stable Outgrowth Medium was added to the mixture and the whole was incubated at 37°C, for 1 h, at 250 rpm. The recovered cells were then plated on LB-agar plates with ampicillin (10 g tryptone, 5 g, yeast extract, 5 g NaCl in 1 L ddH2O, containing

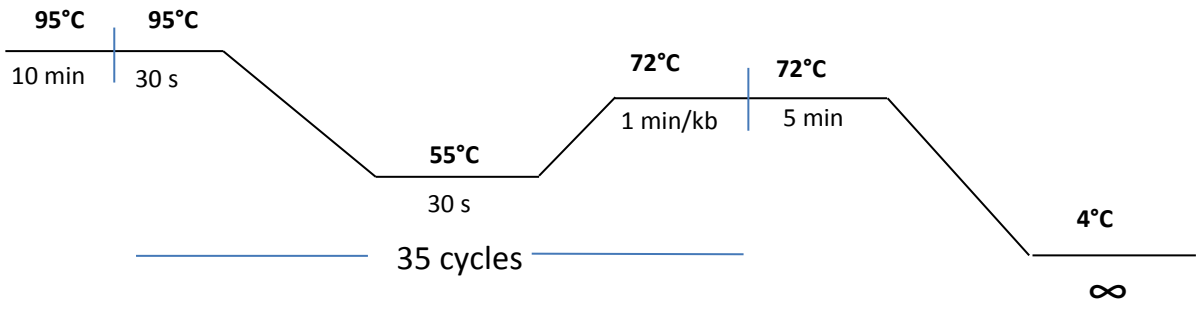
1.2 % agar and 50 µg/mL ampicillin) for selection of the transformed cells. These plates were incubated at 37°C for 16-20 h for the bacterial colonies to appear and grow sufficiently in size. The colonies were then screened for positive ones containing the correct construct by colony PCR.

5.11 Colony PCR

To identify the *E. coli* colonies containing the expression vector with the desired gene insert, single colonies were picked randomLy, suspended in 10 µL of distilled water and were used as templates for colony PCR. Vector-specific primers were used (Table S3) for amplification. The following protocol was used for the PCR:

Component	**Volume (µL)**
ddH_2O	12.3
5x GoTaq Green® Reaction Buffer	5.0
dNTP mix (2.5 mM)	0.5
Forward primer (10 µM)	2.5
Reverse Primer (10 µM)	2.5
Template (*E. coli* colonies in water)	2.0
GoTaq® DNA Polymerase	0.2
Total	**25.0**

The program used in the Biometra TGradient Thermocycler (Goettingen, Germany) for the amplification was as follows:



The samples were then resolved on a 1 % (w/v) agarose gel to identify the positive colonies.

5.12 Plasmid isolation

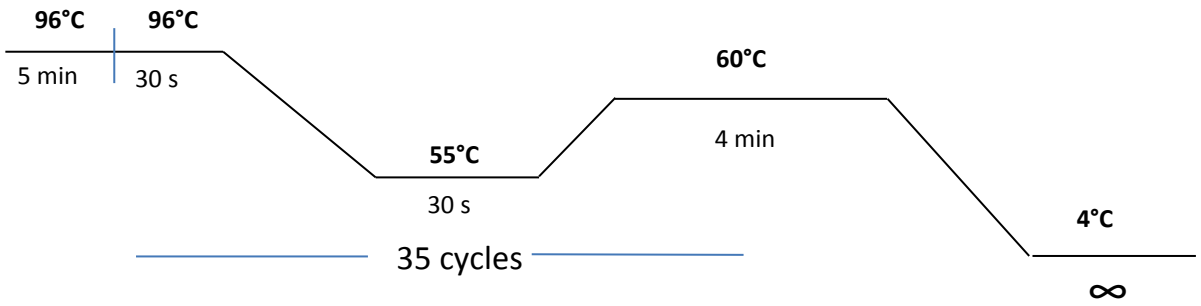
The positive colonies bearing the expression vector with the desired gene insert identified by the colony PCR were inoculated in 4 mL of LB liquid medium containing ampicillin (10 g tryptone, 5 g, yeast extract, 5 g NaCl in 1 L ddH2O, 50 µg/mL ampicillin) and were incubated at 37°C for 16 h at 250 rpm. Later, plasmid isolation was carried out using the NucleoSpin® Plasmid DNA Purification Kit (Macherey-Nagel, Dueren, Germany) as per the manufacturer's instructions. The DNA concentrations of isolated plasmids were determined using a NanoDrop 2000c Spectrophotometer.

5.13 Sanger Sequencing

The sequences of the gene amplicons cloned in *E. coli* were determined by Sanger sequencing using the BigDye™ Terminator v3.1 Sequencing Kit (Thermo Scientific, Schwerte, Germany). The protocol used for the reaction set up was as follows:

Component	**Volume (µL)**
200 ng plasmid DNA	x
BigDyeTM Terminator v3.1 Ready Reaction Mix	4.0
BigDyeTM Terminator v3.1 5x Sequencing Buffer	4.0
Vector-specific forward primer or reverse primer (10 µM)	1.0
ddH_2O	11-x
Total	**20.0**

The program used in the Biometra TGradient Thermocycler (Goettingen, Germany) for the amplification was as follows:



The elongation process was terminated by the incorporation of ddNTPs, thus leading to DNA fragments of varying lengths with fluorescence-labelled 3' ends. These DNA fragments were purified using DyeEx® Spin Column Kit (Qiagen, Hilden, Germany) and were subsequently separated by capillary electrophoresis. The fluorescently labelled 3'-end ddNTPs were detected using ABI PrismDyeEx ® Genetic Analyzer 3130xl (Applied Biosystems™ by Thermo Scientific, Schwerte, Germany).

The sequencing results of the gene amplicons obtained by Sanger sequencing were aligned and verified with the original sequences from the Whitefly genome database using Geneious software (version 10.0.5).

5.14 Heterologous expression of *B. tabaci* glucosinolate arylsulfatase genes in Sf9 cell line

After successful cloning of the genes of interest into the pIB/V5-His TOPO® TA expression vector, the predicted *B. tabaci* arylsulfatase genes were introduced for recombinant expression into *Spodoptera frugiperda* Sf9 cell lines by lipid-mediated transfection using FuGENE® Transfection Reagent (Promega, Mannheim, Germany).

Preparation of Sf9 cells for transfection

1 mL of 90 % confluent Sf9 cell culture was transferred onto a 6-well culture plate containing 2 mL of Sf-900 II SFM (1X) and was incubated overnight at 27°C. 50~60 % cell confluency was ensured prior to transfection.

Preparation of transfection mixture

5 µL of FuGENE® Transfection Reagent was carefully pipetted into a microcentrifuge tube containing 1.7 µg of purified plasmid containing the gene of interest and the total volume was completed to 150 µL using Sf-900 II SFM (1X) medium. The transfection mixture was mixed gently and incubated at room temperature for 15 min.

Transfection procedure

The old cell medium was carefully removed, and 2.85 mL of fresh medium was added to the cells. The entire transfection mixture was then added to the cells and mixed gently by horizontal shaking. The cells were then incubated for 48 h at 27°C.

Selection of stable transformants

48 h post-transfection, the medium containing the transfection mixture was replaced with 3 mL of fresh medium and the cells were split (1:2) and incubated overnight at 27°C. After overnight incubation, the medium was replaced with fresh medium containing blasticidin (Sigma-Aldrich, Germany) at a concentration 50 µg/mL. The cells were incubated at 27°C and the medium containing 50 µg/mL blasticidin was replaced every 3 – 4 days until the cells reached 100 % confluency. This selection procedure took approximately 2 – 3 weeks, depending on the transfection efficiency. Non-transfected Sf9 cells, with and without treatment with blasticidin, were used as controls for the selection procedure. The resistant cells were expanded by moving them to a T-75 flask and the concentration of blasticidin was dropped to 10 µg/mL for maintenance with subculturing (“splitting”) every 3 – 4 days.

5.15 SDS – PAGE and western blot

Western blotting was performed to verify the heterologous expression of the *B. tabaci* arylsulfatase genes.

Sample preparation

An aliquot of the extracellular medium was taken prior to harvesting the cells, to check if the proteins were secreted into the extracellular medium. Cells were then harvested from a 100 % confluent T-75 flask by centrifuging at 500x g for 5 min. The cell pellet was subsequently resuspended in 1 mL of hypotonic buffer (20 mM Tris, 5 mM EDTA, 1 mM DTT, 20 % glycerol, pH 7.5) and incubated on ice for 5 min. The cells were then lysed by using a Potter-Elvehjem tissue grinder (Kontes Glass Co., Vineland, USA). The homogenised cell contents were centrifuged at 1200x g for 5 min. The supernatant (cytosolic fraction), the pellet (the membrane-containing fraction) and the extracellular medium were then resolved by SDS-PAGE.

SDS-PAGE

16 µL of the protein sample was mixed with 4 µL of 5x LaemmLi sample buffer (0.25 M Tris, pH 6.8, 10 % SDS, 50 % glycerol, 5 % β-mercaptoethanol, 0.02 % bromophenol blue) and was incubated at 95°C for 10 min for denaturing the protein sample. The samples were then resolved on a polyacrylamide gel (Bio-Rad Mini-PROTEAN® Precast Gels) at 120 V for 75 min using 1x Tris/Glycine Buffer (25 mM Tris, 192 mM Glycine, 0.1 % SDS in ddH2O, pH 8.3). PageRuler Plus Prestained Protein Ladder, 10-180 kDa range (Thermo Scientific) was used as a size standard.

Western blot

A PVDF membrane (BioRad) was activated in 100 % methanol for 2 min. The activated PVDF membrane, the protein resolved-gel and 2 blotting papers were incubated in 1x Towbin Transfer Buffer (25 mM Tris, 192 mM Glycine, 20 % MeOH in ddH2O, pH 8.3) for 15 min under continuous shaking. The proteins resolved on the gel were transferred to the activated PVDF membrane by semi-dry blotting procedure at 13 V for 25 min. The membrane was subsequently blocked with blocking buffer (5 % (w/v) skim- milk powder in 1X TBST (20 mM

Tris, 150 mM NaCl, 0.1 % Tween 20, pH 7.4)) for 1 h at 4°C under continuous shaking. The membrane was then washed once with 1X TBST for 10 min and incubated overnight with blocking buffer containing the anti-V5-HRP antibody (1: 5.000 (v/v) (Novex® by Life Technologies, Germany) at 4°C under continuous shaking. The membrane was then washed twice with 1X TBST for 10 min and once with 1X TBS (20 mM Tris, 150 mM NaCl, pH 7.4) for 10 min.

ECL detection

The membrane was incubated with 5 mL of Tris-HCl (100 mM, pH-8.5) containing 22 µL of *p*-coumaric acid (90 mM in DMSO) and another 5 mL of Tris-HCl (100 mM, pH-8.5) containing 50 µL of luminol (250 mM in DMSO) for 5 min in the dark. Subsequently, the membrane was sandwiched between transparent polythene sheets and an Amersham™ Hyperfilm™ ECL (GE Healthcare, Germany) was mounted on the sandwich for 8 min in an Amersham™ Hypercassette™. The film was then developed using Carestream® autoradiography GBX developer and fixer solutions (Sigma-Aldrich). The entire ECL detection procedure was carried out in dark.

5.16 *In vitro* enzymatic assay for arylsulfatase activity

To verify the general arylsulfatase activity of the heterologously expressed *B. tabaci* arylsulfatases, *in vitro* enzyme assays were performed using 4-nitrocatechol sulfate (Sigma-Aldrich) as substrate. The reaction set was as follows:

Component	**Volume (µL)**	**End concentration**
Crude protein extract in hypotonic buffer (pH 7.5) / extracellular culture medium	150	–
4-nitrocatechol sulfate (100 mM)	25	5 mM
Tris-HCl (20 mM) pH 6.3	250	10 mM
EDTA (100 mM)	25	5 mM
KCl (1 M)	50	100 mM
Total	**500.0**	

After overnight incubation at 25°C, the reaction was stopped by addition of 0.5 mL of 5 N NaOH, and the absorbance of the resulting *p*-nitrocatechol was measured at 515 nm.

5.17 Screening *B. tabaci* arylsulfatase candidates for glucosinolate sulfatase activity

To identify the *B. tabaci* glucosinolate sulfatase, *in vitro* enzyme assays were performed with the crude protein extracts of cells heterologously expressing *B. tabaci* arylsulfatases. The reaction set up was as follows

Component	**Volume (µL)**	**End concentration**
Crude protein extract in hypotonic buffer (pH 7.5)	80	–
Glucosinolate substrate (25 mM)	20	5 mM
Total	**100.0**	

The reactions were performed using allyl glucosinolate (sinigrin, Sigma-Aldrich), *p*-hydroxybenzyl (*p*OHBn) glucosinolate (sinalbin, purified from *Sinapis alba* seeds) and 4-methylsulfinylbutyl (4-msob, purified from *Brassica oleraceae*) as substrates. The reaction components were pipetted into a microcentrifuge tube, on an ice platform and subsequently incubated at 25°C for 30 min. The enzyme reactions were stopped using acetic acid (10 % reaction volume). Each of the enzymatic assays were performed in duplicates. Enzymatic assays with *Helix pomatia* sulfatase were used a positive control (Roy et al. 1987). The desulfated glucosinolate product formation was analysed by HPLC-MS/MS.

5.18 Polyhistidine tagged protein purification

The polyhistidine tagged, heterologously expressed and extracellularly secreted *B. tabaci* glucosinolate sulfatase (Bta14666) was affinity purified over Ni-NTA agarose resin (Qiagen, Germany). The extracellular culture medium was collected and was concentrated to 1 mL

using centrifugal filter units (Amicon® Ultra Centrifugal Filters) with a molecular weight cut-off 10000.

Equilibration of Ni-NTA agarose resin

1 mL of distilled water was added to a 1.5 mL microcentrifuge tube containing 400 µL of Ni-NTA agarose beads. The supernatant was discarded after centrifuging for 30 s at 13000 rpm. Subsequently 1 mL of lysis buffer (50 mM Tris pH 8.0, 20 mM imidazole, 500 mM NaCl, 10% glycerol, 1 % Tween20) was added to the beads and was mixed by gentle inversion. The supernatant was again discarded after centrifuging for 30 s at 13000 rpm.

Binding of the histidine tagged protein to Ni-NTA agarose resin

The extracellular culture medium-concentrate was added to the equilibrated resin, and mixed under circular rotation at 4°C for 1 h and then centrifuged at 13000 rpm for 5 min at 4°C. Before discarding the supernatant, an aliquot of it was collected for analysis by SDS-PAGE. To the protein-bound resin, 1 mL of wash buffer (50 mM Tris pH 8.0, 20 mM imidazole, 500 mM NaCl, 10 % glycerol) was added, mixed gently and centrifuged at 13000 rpm for 5 min at 4°C. An aliquot of the supernatant was collected for analysis by SDS-PAGE prior discarding. The wash step was repeated twice.

Elution of histidine tagged protein bound to Ni-NTA agarose resin

200 µL of elution buffer (50 mM Tris pH 8.0, 250 mM imidazole, 500 mM NaCl, 10 % glycerol) was added to the protein-bound resin and then mixed gently by inversion. After centrifuging at 13000 rpm for 5 min at 4°C, the supernatant was collected and stored at -20°C. The elution step was repeated four times.

5.19 Bradford assay for determination of protein concentration

Bradford assay (Bradford 1976) was performed to determine the protein concentration in the crude and purified protein samples. Bovine serum albumin BSA (SERVA, Germany) was used as a reference protein to generate a calibration curve. BSA standards of concentrations 12.5, 25, 50, 100, 250, 500, 750, 1000, 1500 and 2000 µg/mL were prepared in ddH_2O. 5 µL of protein sample whose concentration had to be determined was added to 200 µL of 1X Bradford

reagent (SERVA) and incubated at room temperature for 5 min. The assay was performed in triplicates using a 96-well microplate. After ensuring that there were no air bubbles, absorbance was measured at 595 nm using the Molecular Devices SpectraMAX 250 Plate Reader (Marshall Scientific LLC, California, USA).

5.20 *In vitro* glucosinolate sulfatase assay

1 µL of the purified enzyme was added to 91 µL of 100 mM of Tris (pH 7.5) and incubated with 8 µL of aqueous solution of 25 mM allyl glucosinolate (end substrate concentration - 2 mM) for 10 min at 25°C. The enzyme reaction was stopped using acetic acid (10 % reaction volume). Each of the enzymatic assays were performed in duplicates. The desulfated allyl glucosinolate product formed was quantified by HPLC-MS/MS.

5.21 Determination of the linearity of *B. tabaci* glucosinolate sulfatase reaction and optimal assay time

10 µL of the purified enzyme was added to 910 µL of 100 mM of Tris buffer (pH 7.5) and was incubated with 80 µL of aqueous solution of 25 mM allyl glucosinolate (end substrate concentration - 2 mM) at 25°C. Aliquots of 100 µL of the reaction solution were taken out at 0 min, 5 min, 10 min, 15 min, 20 min, 25 min, 30 min, 45 min and 60 min time points, and the enzyme reaction of each aliquot was immediately stopped using acetic acid (10% reaction volume). The mother reaction solution was vortexed before and after aliquoting. Each of the enzymatic assays were performed in duplicates. The desulfated allyl glucosinolate product formed was quantified by HPLC-MS/MS.

5.22 Determination of *B. tabaci* glucosinolate sulfatase optimal temperature

1 µL of the purified enzyme was added to 91 µL of 100 mM of Tris (pH-7.5) and was incubated with 8 µL of aqueous solution of 25 mM allyl glucosinolate (end substrate concentration - 2 mM) for 10 min at 10°C, 15°C, 25°C, 35°C, 45°C, 55°C, 65°C and 75°C. The enzyme reaction was stopped using acetic acid (10 % reaction volume). Each of the enzymatic assays were performed in duplicates. The desulfated allyl glucosinolate product formed was quantified by HPLC-MS/MS.

5.23 Determination of *B. tabaci* glucosinolate sulfatase optimal pH

Experiment 1: broad pH range

1 µL of the purified enzyme was added to 91 µL of each of the following buffer solutions:
Phosphate citrate buffer (0.2 M dibasic sodium phosphate; 0.1 M citric acid; pH 3.0, 4.0, 5.0, 6.0, 7.0);
0.1 M Tris-HCl buffer (pH 7.0, 8.0);
0.2 mM Glycine-NaOH buffer (pH 8.0, 9.0);
To the mixture were added 8 µL of aqueous solution of 25 mM allyl glucosinolate (end substrate concentration - 2 mM) with incubation for 10 min at 25°C. The enzyme reaction was stopped using acetic acid (10% reaction volume). Each of the enzymatic assays were performed in duplicates. The desulfated allyl glucosinolate product formed was quantified by HPLC-MS/MS.

Experiment 2: narrow pH range

1 µL of the purified enzyme was added to 91 µL of 0.1 M sodium phosphate buffer solutions with pH- 6.25, 6.50, 6.75, 7.00, 7.25, 7.50, 7.75 and 8.00, and was incubated with 8 µL of aqueous solution of 25 mM allyl glucosinolate (end substrate concentration - 2 mM) for 10 min at 25°C. The enzyme reaction was stopped using acetic acid (10% reaction volume). Each of the enzymatic assays were performed in duplicates. The desulfated allyl glucosinolate product formed was quantified by HPLC-MS/MS.

5.24 Glucosinolate extraction from *Arabidopsis thaliana*

Leaves were harvested from *A. thaliana* Col-0 rosette stage plants and were flash frozen using liquid N_2 to prevent the activation of plant myrosinase. Subsequently, the leaves were freeze dried overnight using Alpha 1-4 LDplus freeze dryer (Martin Christ, Osterode am Harz, Germany). The fresh weight and the dry weight of the leaves used for the extraction were calculated before and after freeze drying respectively. The extraction of the leaves for intact glucosinolates was carried out by incubation with 10 mL 80 % (v/v) methanol in water per g of the dry weight, on ice under continuous shaking for 5 min, followed by the addition of metal beads (3 mm) and vigorous shaking for 10 min in a paint shaker. The supernatant collected after centrifugation (4000x g at 4°C for 15 min), was then processed through centrifugal filter

units **(**Amicon® Ultra Centrifugal Filters) of molecular weight cut-off 10000 to remove the plant myrosinase. The flow through was then treated on a rotatory evaporator for removal of the 80 % (v/v) methanol in water and was resuspended in ddH_2O.

5.25 Substrate preference of *B. tabaci* glucosinolate sulfatase

To test the substrate preference of the *B. tabaci* glucosinolate sulfatase, enzyme assays were performed with glucosinolate extract obtained from *A. thaliana* Col-0. 45 µL of the purified enzyme was added to 910 µL of 100 mM of Tris buffer (pH-7.5) and was incubated with 45 µL glucosinolate extract at 25°C. An aliquot of 100 µL of the reaction solution was taken out at 0 min, 10 min, 20 min, 15 min, 20 min, 30 min, 60 min, 120 min, 180 min and 300 min time points, and the enzyme reaction of each aliquot was immediately stopped using acetic acid (10 % reaction volume). The mother reaction solution was vortexed before and after aliquoting. Each of the enzymatic assays were performed in duplicates. No-enzyme reaction set up served as a control for non-enzymatic degradation of glucosinolates. The intact glucosinolates in each reaction sample were quantified by HPLC-MS/MS.

5.26 Desulfoglucosinolate standard preparation

100 µL purified *Helix pomatia* sulfatase was added to 1820 µL of 100 mM of Tris buffer (pH 7.5) and was incubated with 80 µL of aqueous solution of 25 mM of allyl glucosinolate (end substrate concentration - 1 mM)**.** Assuming quantitative substrate conversion into desulfoglucosinolates only, this 1 mM desulfated glucosinolate solution was then used as a stock solution for the preparation of desulfoglucosinolate standards as shown below:

Desulfoglucosinolate standard concentration (µM)	**Concentration(mM) of stock solution used**	**Volume of stock solution (µL)**	**Volume of water (µL)**
250 µM	1 mM	250	750
200 µM	1 mM	200	800
150 µM	1 mM	150	850
100 µM	1 mM	100	900
75 µM	1 mM	75	925

50 µM	1 mM	50	950
25 µM	1 mM	25	975
10 µM	1 mM	10	990
5 µM	1 mM	5	995

5.27 HPLC-MS/MS

Analyses of desulfated and intact glucosinolates were performed on an Agilent Technologies (Santa Clara, CA, USA) 1200 Series HPLC using a Nucleodur Sphinx RP column (250 × 4.6 mm × 5 µm, Macherey-Nagel, Düren, Germany) coupled to an API 3200 triple-quadrupole mass spectrometer (Applied Biosystems, Darmstadt, Germany). Formic acid (0.2 %) in water and acetonitrile were employed as mobile phases A and B, respectively. The flow rate was 1.1 mL/min. The elution profile was as follows: 0–2.5 min, 1.5 % B; 2.5–5 min, 1.5–10 % B; 5–12.5 min, 10–40 % B; 12.5–17.5 min, 40–70 % B; 17.6–20 min, 100 % B; and 20.1–24 min, 1.5 % B. The ion spray voltage was maintained at 4500 eV. The turbo gas temperature was set to 700 °C. Nebulizing gas was set at 70 psi, curtain gas at 20 psi, heating gas at 60 psi, and collision gas at 10 psi. Multiple reaction monitoring (MRM) was used to monitor parent ion to fragment ion conversion (Malka et al. 2016). Parameters for MRM analysis of different glucosinolates are available in (Table S4 and Table S5). Intact glucosinolates were detected in the negative ionization mode, while desulfoglucosinolates were detected in the positive mode. Analyst 1.5 software (Applied Biosystems) was used for data acquisition and processing.

5.28 Phylogenetic analysis

The amino acid sequence of the *P. xylostella* glucosinolate sulfatase (Px018104) was used as a query sequence to obtain the putative *B. tabaci* glucosinolate sulfatases by doing a BLAST against the whitefly genome database using the Geneious software (version 10.0.5). All the shortlisted protein sequences were aligned using MUSCLE and a Neighbour Joining tree was constructed using MEGA software (version 5.2). Reliability of the constructed Neighbour Joining tree was assessed using bootstrapping method (1000 bootstrap replicates).

5.29 Quantitative real-time PCR (qPCR) analysis

To check if the *Bta14666* was expressed constitutively or induced only upon glucosinolate consumption, a qPCR analysis was performed. For this purpose, cDNAs were generated from total RNA isolated from *B. tabaci* that were reared on two different diet conditions: eggplant (non-glucosinolate diet) and brussels sprout (glucosinolate diet). cDNA from four biological replicates of each diet condition was obtained from Dept. of Entomology (The Hebrew University of Jerusalem). cDNA concentration of 5 ng/µL was used for the qPCR analysis.

Ribosomal protein L-13 (*Bta04282*) was used as a reference gene to normalize the expression levels of *Bta14666*. Primers used for the qPCR analysis of *Bta14666* and the chosen reference gene were designed using the Geneious software (version 10.0.5). All the primers used were designed with an optimal melting temperature of 60°C. The list of primers is available in Table S2. To ensure primer specificity, the qPCR amplicons were analysed by agarose gel electrophoresis, purified, cloned into the pCR*TM*4-TOPO® vector and were sequence confirmed.

The following protocol was used for the qPCR:

Component	**Volume (µL)**
2x Brilliant® III SYBR Master Mix	10.0
Forward primer (10 µM)	1.0
Reverse Primer (10 µM)	1.0
cDNA (5 ng/µL)	1.0
Sterile water	7.0
Total	**20.0**

The program used in the CFX Connect™ Real-Time PCR Detection System (BioRad, Munich, Germany) for the amplification was as follows:

1. Initial Denaturation (95 °C) 3 min
2. Denaturation (95 °C) 10 s

3. Primer annealing/ extension (60 °C) 20 s

4. Plate read (fluorescence detection)

Cycle: repeat steps 2.-4. 39x

5. Denaturation 10 s

Melt Curve (60 – 95 °C, increment 0.5 °C) 5 s

Each reaction was done in triplicates to account for manual errors, and non-template control reactions were set up to ensure there was no DNA contamination during pipetting. Melting curve analysis was conducted at the end of qPCR run by gradually heating the PCR reaction mixture from 50°C to 95°C, to detect any genomic DNA contamination and to detect the formation of primer dimers.

Expression levels of *Bta14666* was relatively quantified using the ΔΔCP method (Pfaffl 2001). The expression levels of the gene of interest (GOI) relative to the reference gene (Ref) were calculated, based on its real-time PCR efficiencies (E), and the crossing point difference (ΔCP) of one unknown sample (treatment) versus one control (Pfaffl 2001) (see equation below).

$$ratio = \frac{(E_{GOI})^{\Delta CP_{GOI}(control-sample)}}{(E_{Ref})^{\Delta CP_{Ref}(control-sample)}}$$

Expression values of the biological replicates were averaged, and standard error was calculated.

6. Results

6.1 Identification of putative *Bemisia tabaci* GSS

The sequences of the *Plutella xylostella* GSS (Px018104, obtained from Diamondback moth Genome Database**;** http://iae.fafu.edu.cn/DBM/ on 15/01/2017), which have been previously identified and characterised (Ratzka et al. 2002), were used as query to identify predicted *B. tabaci* sulfatases (BtSulf) that could serve as candidates for GSS (BtGSS).

Table 1: Shortlisted putative *B. tabaci* sulfatase (BtSulf) candidates based on BLAST hit analyses using *P. xylostella* GSSs as the query sequences (cut-off e-value <1.0 e^{-5}), and their corresponding expression levels in gut tissues and whole insect bodies calculated from publicly available data; the candidates picked for expression in Sf9 cells and screening for GSS activity are highlighted in blue. The gene IDs are those of the whitefly genome database (*http://www.whiteflygenomics.org/*).

Gene ID	Name	E value	Annotation	FPKM MEAMB midgut	FPKM MEAMB adults	FPKM MEAMB gut/adults
Bta06756	BtSulf1	2.24E-120	Arylsulfatase B	0	244	1690.074
Bta03550	BtSulf2	6.48E-102	Arylsulfatase	0	81	648.561
Bta02222	BtSulf3	4.03E-101	Arylsulfatase	557	747	1.289
Bta04774	BtSulf4	2.03E-97	Arylsulfatase	17	296	10.064
Bta14667	BtSulf5	3.25E-88	Arylsulfatase B	14	0	196.656
Bta04900	BtSulf6	6.95E-84	Arylsulfatase B	0	187	985.108
Bta14665	BtSulf7	1.23E-81	Arylsulfatase B	62	303	2.824
Bta01141	BtSulf8	1.63E-79	Arylsulfatase B	14	40	1.651
Bta14669	BtSulf9	3.86E-67	Arylsulfatase B	0	11	149.470
Bta04898	BtSulf10	3.63E-52	Arylsulfatase B	0	67	669.999
Bta14666	BtSulf11	1.86E-49	Arylsulfatase B	8292	1141	12.572
Bta08750	BtSulf12	1.61E-16	N-acetylglucosamine-6-sulfatase	3223	10161	1.822
Bta06730	BtSulf13	2.39E-13	Extracellular sulfatase SULF-1-like protein	21	488	13.432
Bta01054	BtSulf14	8.21E-10	N-sulfoglucosamine sulfohydrolase	50	226	2.612
Bta04899	BtSulf15	2.60E-08	Arylsulfatase B	0	85	1072.950

Full-length predicted sequences of BtSulf candidates were found by performing a peptide BLAST against the whitefly genome database (International Whitefly Genome Initiative (IWGI), 2015, http://www.whiteflygenomics.org/) (Table 1). Most of the 15 shortlisted protein candidates had been previously annotated as arylsulfatases.

The top candidates (BtSulf1; BtSulf2; BtSulf3; BtSulf4; BtSulf5) were picked for cloning from *B. tabaci* whole-insect cDNA (MEAM1, formerly B) and expression in Sf9 cells, and subsequently screened for GSS activity. Based on the expression levels measured according to the public Illumina data from the NCBI (Bemisia tabaci MEAM1 Silverleaf whitefly B biotype SRR835757, SRR059302, SRR1523522), BtSulf11 was also shortlisted for screening, as it was found to be highly expressed in the gut. BtSulf11 was also chosen for expression and screening based on the phylogenetic analysis performed (see section 6.9), as BtSulf11was closely related to BtSulf5 (Figure 16).

6.2 Heterologous expression of putative *B. tabaci* GSS in insect cells

Previously published reports showed that *P. xylostella* GSS expressed in *E. coli* had no sulfatase activity (Ratzka et al. 2002), suggesting the absence of post-translational modifications necessary to have an active sulfatase. Hence insect cells (*Spodoptera frugiperda* Sf9 cells) were used for expressing the putative *B. tabaci* GSS candidates. The full-length ORFs of BtSulf1, BtSulf2, BtSulf3, BtSulf4, BtSulf5 and BtSulf11 were successfully PCR-amplified from *B. tabaci* cDNA (MEAM1), and cloned into pCR® 4Blunt-TOPO® vector. The clone was then transformed with NEB 10-beta competent *Escherichia coli* cells by heat-shock treatment, and the cells were then plated on LB-agar plates with ampicillin for selection of the transformed cells, and subsequently the recovered colonies were analysed by colony PCR for positive ones containing the gene insert in the right orientation. The alignment of the cloned gene amplicon sequences (obtained by Sanger sequencing) with the gene candidate sequences obtained from the whitefly genome database was done for sequence confirmation. BtSulf11 was observed to have a small insertion of 54 bp in length in comparison to the gene sequence available in the whitefly genome database. This observed deviation indicates either the presence of isoforms or the misassembly of the indel region. Confirmed amplified ORFs were reamplified with A-overhangs and finally cloned into the pIB/V5-His-TOPO® TA expression vector for expression in Sf9 cells. Transfected cells were harvested and expression of recombinant *B. tabaci* arylsulfatases fused to a V5-epitope at the C-terminus was analysed by SDS-PAGE and subsequent western blot using a V5-specific antibody. *P. xylostella* GSS

cloned into the pIB/V5-His-TOPO® TA expression vector was a generous gift from Dr. Roy Kirsch, Dept. Entomology, MPI-CE Jena and was used as a positive control for the recombinant expression of active arylsulfatases in Sf9 cells. All selected candidates, except for BtSulf5, were successfully expressed (Figure 4). BtSulf11 and *P. xylostella* GSS were found to be secreted into the extracellular culture medium. The molecular weights of the heterologously expressed proteins are listed in Table S6.

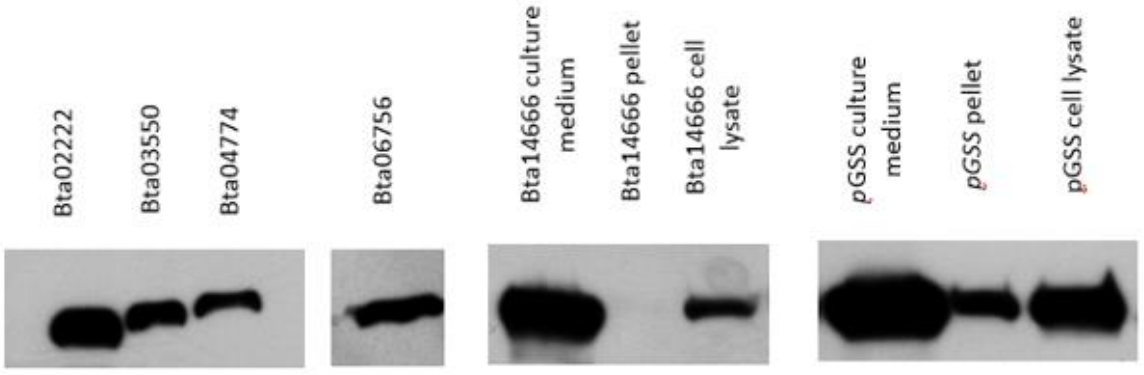


Figure 4: Immunoblot showing the successful expression of *B. tabaci* arylsulfatases in Sf9 cells. Equal volumes (20 µL) of each of the crude protein samples were loaded for analysis by SDS-PAGE. (pGSS- *P. xylostella* GSS; the gene ID is as per whitefly genome database).

6.3 Sulfatase activity of expressed *B. tabaci* arylsulfatases

Since arylsulfatases catalyse desulfonation of sulfate esters (Kertesz et al. 2000; Buono andCosma 2010), we used 4-nitrocatechol sulfate as a substrate to examine the general sulfatase activity of the expressed *B. tabaci* arylsulfatases. The resulting desulfated product, 4-nitrocatechol, gives a pink coloration under high pH (Figure 5), which was measured spectrophotometrically at an absorbance of 515 nm (Figure 6).

Recombinantly expressed arylsulfatases actively desulfated 4-nitrocatechol sulfate. *P. xylostella* GSS and *H. pomatia* sulfatase were appropriate active positive controls for these enzymatic assays. No direct comparison of specific activities of the different sulfatases can be made, as the PxGSS and BtSulf enzymes were not purified and their amounts were not normalized before the assays.

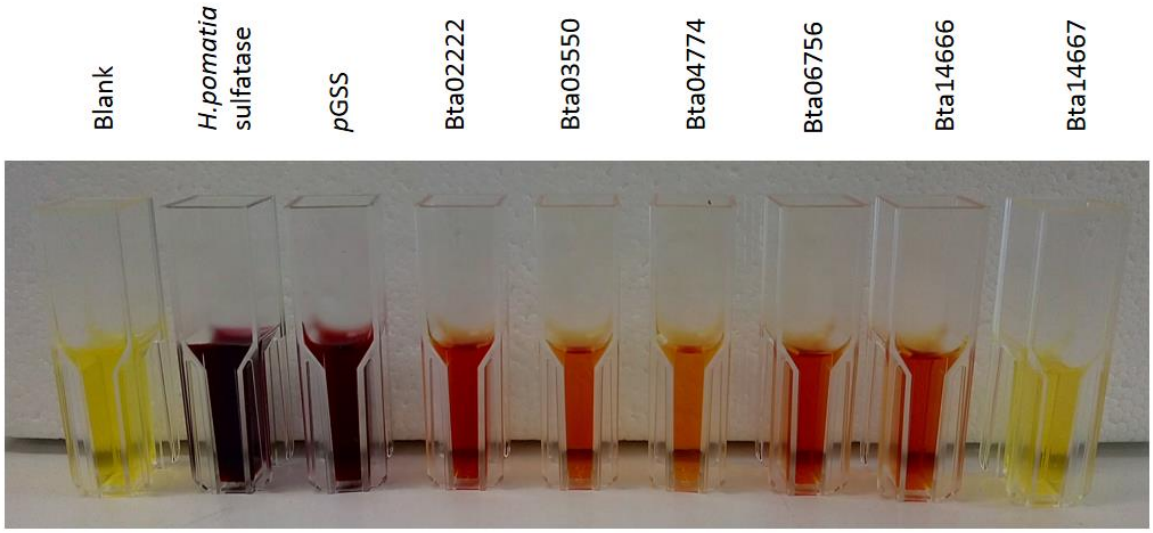


Figure 5: In vitro enzymatic assays using 4-nitrocatechol sulfate as substrate, showing general arylsulfatase activity of the recombinantly expressed *B. tabaci* arylsulfatases. Desulfated product, 4-nitrocatechol, gives a pink coloration under high pH. *P. xylostella* GSS (pGSS) and *H. pomatia* sulfatase served as positive controls.

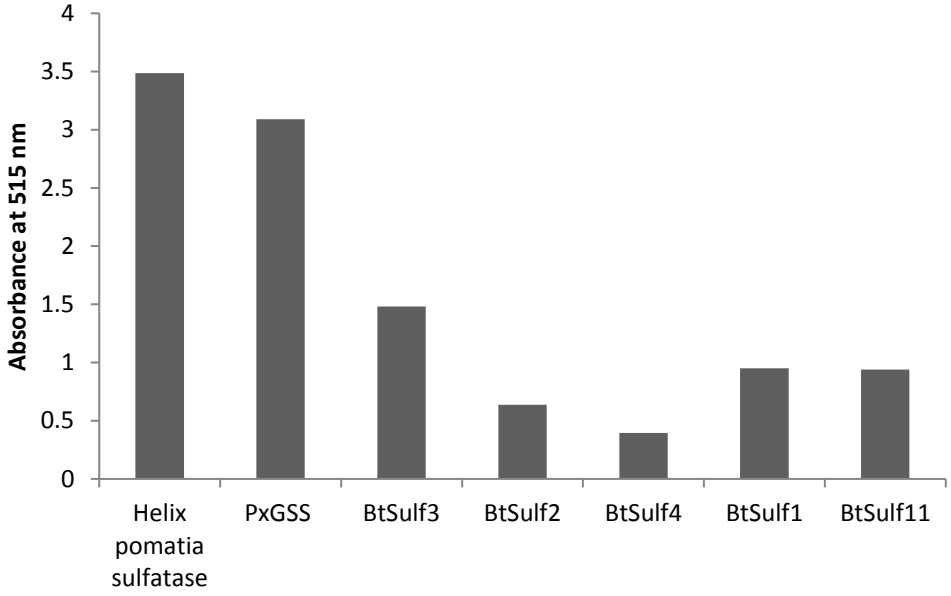


Figure 6: In vitro enzymatic assay, using 4-nitrocatechol sulfate as substrate, showing general arylsulfatase activity of the recombinantly expressed *B. tabaci* arylsulfatases. Desulfated product, 4-nitrocatechol, was measured at an absorbance of 515 nm. *P. xylostella* GSS and *H. pomatia* sulfatase served as positive controls.

6.4 *B. tabaci* encodes a GSS gene

To identify *B. tabaci* GSS activities, the expressed *B. tabaci* arylsulfatases were screened for their activity towards pure glucosinolates (allyl (sinigrin), *p*-hydroxybenzyl (*p*OHBz, sinalbin) and 4-methylsulfinylbutyl (4-msob, glucoraphanin) glucosinolates). The desulfated glucosinolates were detected by positive mode mass spectrometry. Intact glucosinolates also gave rise to masses corresponding to the desulfated glucosinolates due to in-source fragmentation in the spectrometer. However, these in-source fragmented desulfated glucosinolates were readily distinguished from proper desulfoglucosinolates due to different retention times. The desulfated glucosinolates eluted much before the intact glucosinolates under the chromatographic conditions applied, allowing unambiguous detection by comparison to authentic standards of intact glucosinolates and to corresponding pure desulfoglucosinolates produced enzymatically by incubation with *H. pomatia* sulfatase (following the protocol of REF). HPLC-MS/MS analyses detecting desulfated glucosinolates revealed a GSS activity in BtaSulf11 (BtGSS1) towards all tested pure glucosinolates (Figure 7).

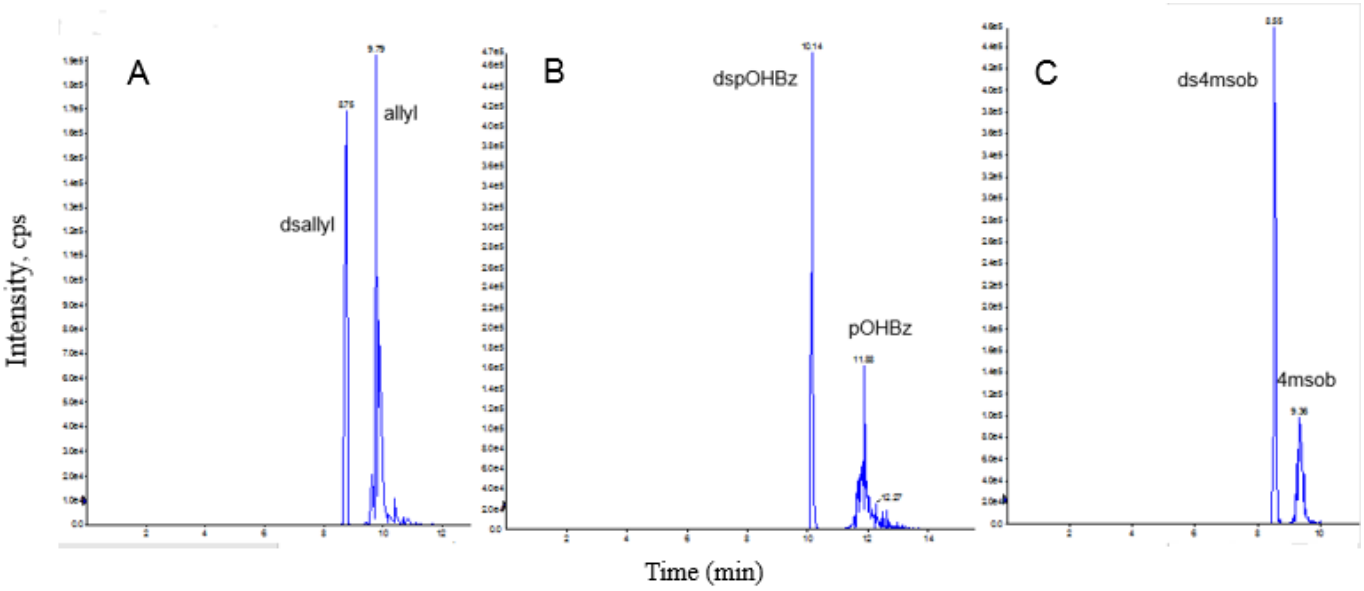


Figure 7*:* LC-MS MRM chromatograms of desulfated glucosinolates produced from reaction with BtSulf11 (allyl = 2-propenyl; pOHBz = *p*-hydroxybenzyl; 4msob = 4-methylsulfinylbutyl; dsallyl- desulfated allyl; dspOHBz- desulfated pOHBz; ds4msob- desulfated 4msob glucosinolates)

6.5 Affinity-based purification BtGSS1

BtGSS1 was secreted into the extracellular culture medium (Figure 4) by Sf9 cells. Prior to purification, this extracellular medium was therefore concentrated, and the recombinantly expressed BtGSS1 (fused to a polyhistidine tag at the C-terminus) was affinity-purified over Ni-NTA agarose resin for further characterisation.

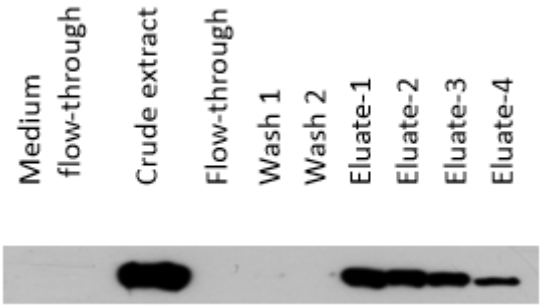


Figure 8: Immunoblot showing the purification of recombinant BtGSS1 using Ni-NTA agarose. 16 µL of each of the 1 mL crude extract, flow-through and wash fractions, and 3.2 µL of each of the 200 µL eluate fractions were loaded for analyses by SDS-PAGE. (medium flow-through: flow-through during processing with centrifugal filter units to get rid of small molecules and proteins; crude extract: concentrated culture medium containing BtGSS1 after centrifugal filter units; flow-through, wash1,2, eluate 1,2,3,4: flow-through, wash and eluate fractions during purification with Ni-NTA agarose resin).

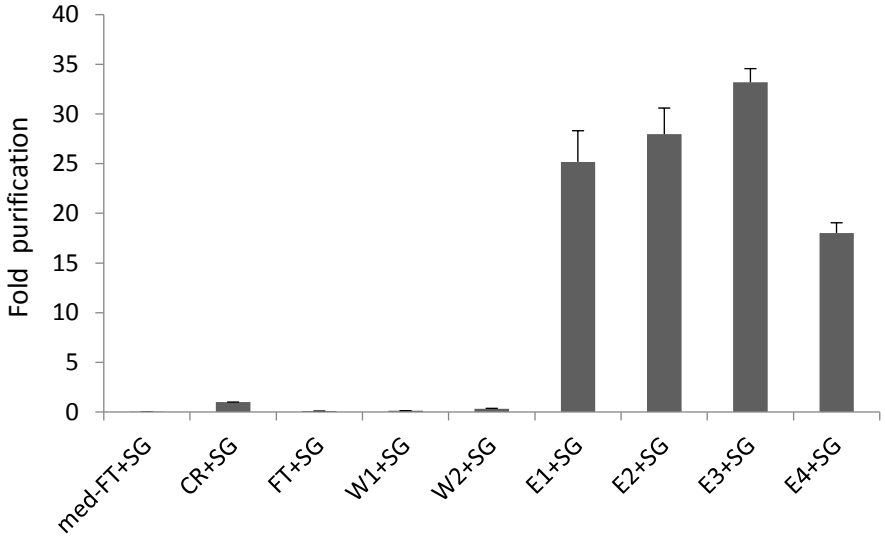


Figure 9: Fold purification of his-tag purified BtGSS1 (SG- Sinigrin; med-FT- medium flow-through; CR- crude extract; FT- flow-through; W1,2- wash1,2; E1,2,3,4- eluate 1,2,3,4).

Enzyme activities of the crude, flow-through, wash and eluate fractions were calculated as moles of product formed per time (pmol/s) [unit – picokatal, pkat]. Subsequently, the specific activity was calculated as enzyme activity per mass of total protein present [unit - pkat/mg of total protein], as determined by Bradford measurements. Finally, fold purifications were determined by dividing the specific activities of the eluate by that of the crude fraction (Table 2). The calculated activities for the flow-through and wash fractions ensured that the loss of the recombinant protein in those steps were minimal.

Table 2 Fold purification of his-tag purified BtGSS1 (SG- Sinigrin; med-FT- medium flow-through; CR- crude extract; FT- flow-through; W1,2- wash1,2; E1,2,3,4- eluate 1,2,3,4

Sample	Enzyme activity (pkat)	Total protein conc. (µg/mL)	Specific activity (pkat/mg of total protein)	Fold purification
med-FT+SG	0.000269307	200	0.269306524	0.000601899
CR+SG	15.26225881	6822.222222	447.4279286	1
FT+SG	0.840124338	5877.777778	28.58649925	0.063890735
W1+SG	0.276774296	1262.222222	43.85515931	0.098016142
W2+SG	0.22916725	406.6666667	112.7070526	0.251899905
E1+SG	5.834196794	577.7777778	10097.78423	22.5685157
E2+SG	3.463179217	321.1111111	10785.0245	24.10449552
E3+SG	1.913616549	140	13668.68963	30.54947794
E4+SG	1.124833142	125.5555556	8959.24446	20.02388293

6.6 Enzymatic characterisation of BtGSS1

The linearity of the assay was assessed over a time range of 0 min to 60 min, after BtGSS1 had been subjected to concentration, purification and storage. The assay was found to be linear, with most of the substrate (>95%) remaining unreacted within the assay times, and the purified enzyme was found to retain its activity over the assessed time range, at 25 °C and pH-7.5 (Figure 10).

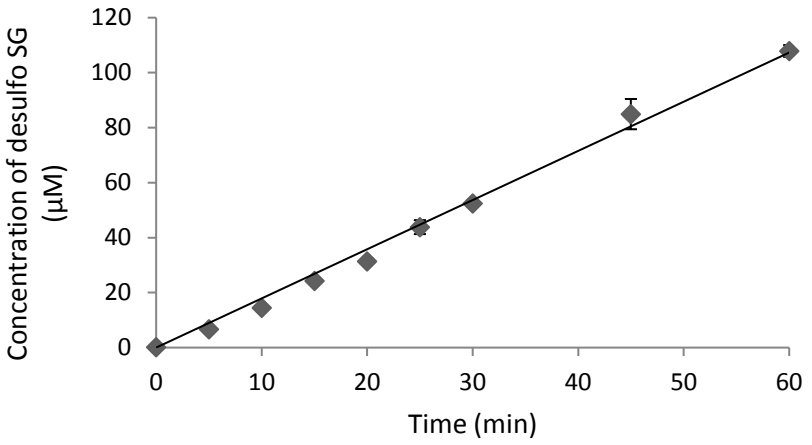


Figure 10: BtGSS1 stability over a time range of 0 min to 60 min, after enzyme had been subjected to concentration, purification and storage (SG- Sinigrin).

The temperature dependence of the BtGSS1 glucosinolate sulfatase activity was probed in the range of 10°C to 75°C. The optimum temperature was found to be around 50-55°C (Figure 11).

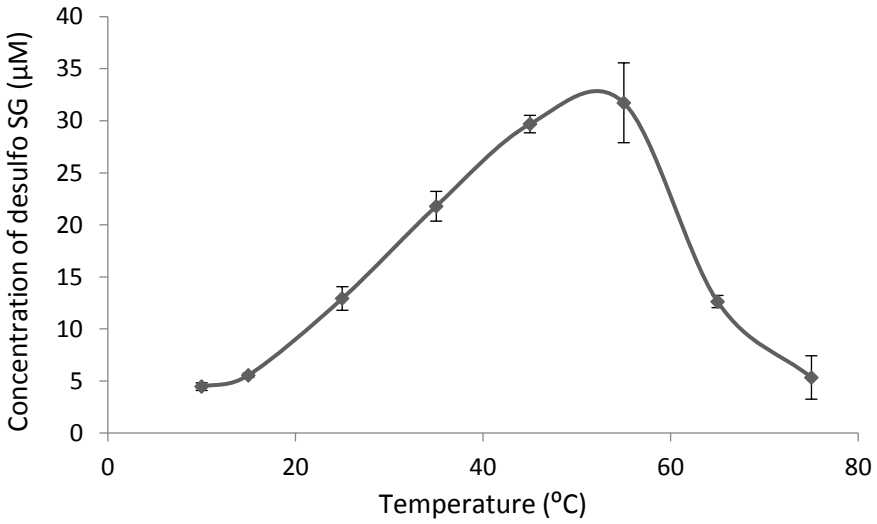


Figure 11: Determination of optimal temperature of BtGSS1, found to be around 50-55°C.

The influence of pH on GSS activity of BtGSS1 was tested in the range of pH 3 to 8. Initially, the experiment was performed in three different buffer systems and intervals of one pH unit: Phosphate citrate buffer (pH 3 to 7); Tris-HCl buffer (pH 7 to 9); Glycine-NaOH buffer (pH 9 and 10). The enzyme was found to be more active in the pH range 6 to 8 (Figure 12), with Tris-HCl buffer apparently leading to higher product formation compared to phosphate. Subsequently, the activity of the enzyme was monitored within this narrow pH range (pH 6 to

8) using a single buffer system (phosphate buffer). The pH optimum for BtGSS1 was found to be around 7.75 (Figure 13).

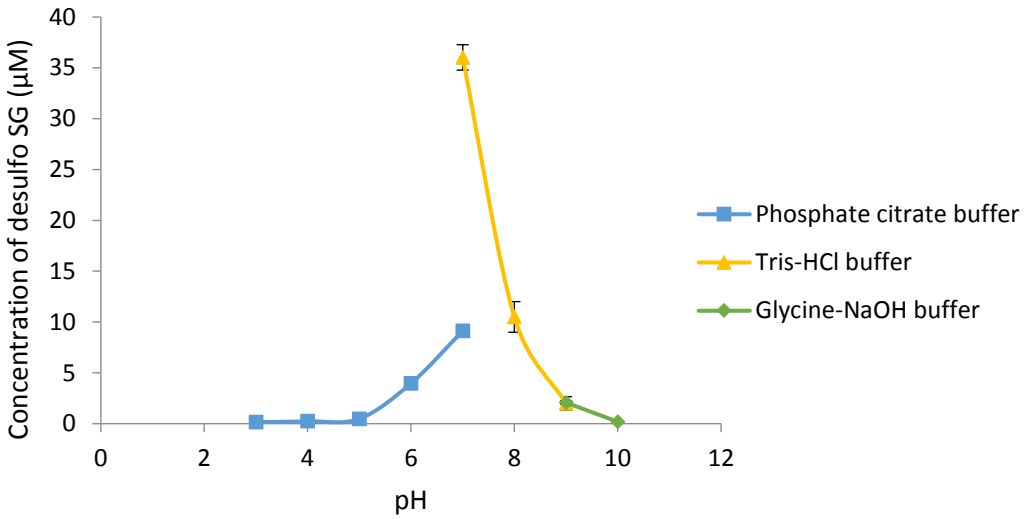


Figure 12: Determination of rough optimal pH range of BtGSS1, found to be in the pH range 6-8.

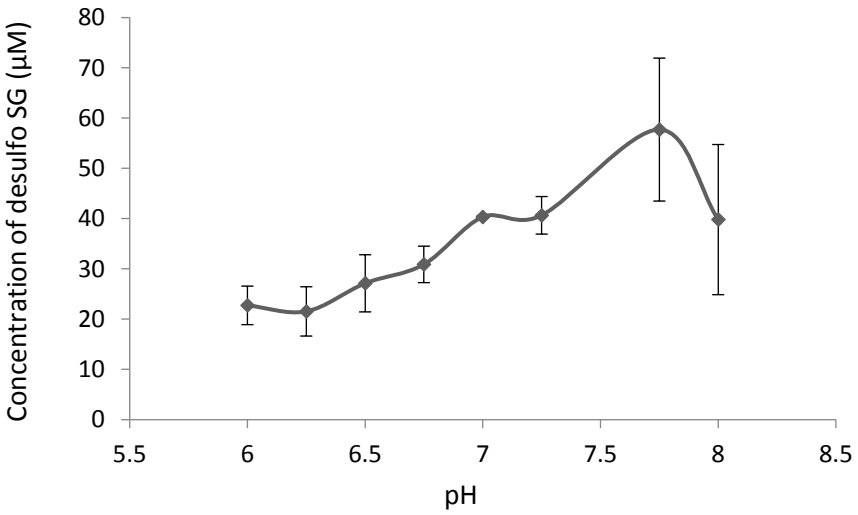


Figure 13: Narrow determination of optimal pH of BtGSS1, which was found to be around 7.75.

6.7 BtGSS1 shows a preference towards indolic glucosinolates

To evaluate whether the identified *B. tabaci* glucosinolate sulfatase (BtGSS1) acts on different glucosinolate classes and whether it shows preference to specific substrates, enzyme assays were performed with glucosinolate extract obtained from *A. thaliana* Col-0.

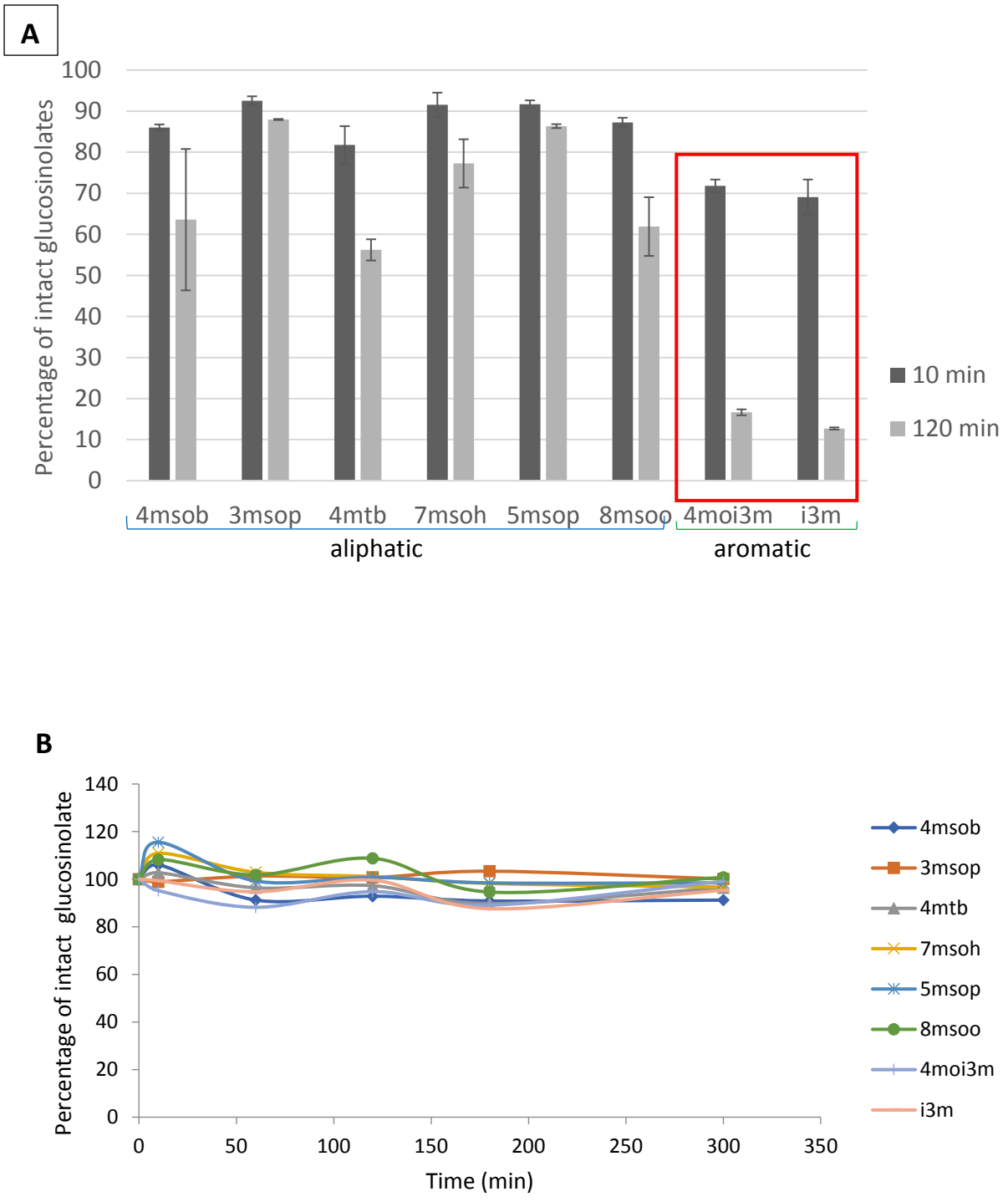


Figure 14: A. Rate of glucosinolate desulfation of different glucosinolates by BtGSS1. Crude glucosinolate extract obtained from *A. thaliana* Col-0 was used. Amount of each glucosinolate present at the 0 min was taken as 100%, and in reference with this, subsequent amounts unreacted glucosinolates over the reaction time was calculated in percentage. **B.** No-enzyme reaction set up served as a control for non-enzymatic degradation of glucosinolates. (4msob = 4-methylsulfinylbutyl; 3msop = 3-methylsulfinylpropyl; 4mtb = 4-methylthiobutyl; 7msoh = 7 methylsulfinylheptyl; 5msop = 5-methylsulfinylpentyl; 8msoo = 8-methylsulfinyloctyl; 4moi3m = 4-methoxyindol-3-ylmethyl; i3m = indol-3-ylmethyl).

For this purpose, enzymatic assays were set up with the plant extract containing glucosinolates (semi-quantitatively – 4msob and i3m being the major aliphatic and indolic glucosinolates respectively) and the individual glucosinolate depletion rates were determined over a time range of 0 min to 300 min. The intact glucosinolates were detected by negative mode mass spectrometry. BtGSS1 exhibited preference towards indolic glucosinolates, shown as a quicker depletion of these compounds relative to aliphatic glucosinolates (Figure 14.A). No-enzyme reactions were tested as a control for non-enzymatic degradation of glucosinolates, and showed that their concentrations were all stable during the assay time range (Figure 14.B), i.e. spontaneous glucosinolate degradation was not detected under these assay conditions. This experiment suggested that BtGSS1 could act on different glucosinolate classes and likely has different substrate affinities/catalytic efficiencies towards them.

6.8 *BtGSS1* is constitutively expressed in *B. tabaci* irrespective of glucosinolate consumption

To compare the relative expression levels of *BtGSS1* in *B. tabaci* reared on non-glucosinolate diet (eggplant) and those reared on glucosinolate rich diet (Brussels sprouts), quantitative real-time PCR (qPCR) analysis was performed. cDNA was obtained from *B. tabaci* MEAM1 adult insects (four biological replicates of each diet condition, a generous gift from Dr. Osnat Malka, Dept. of Entomology, The Hebrew University of Jerusalem, Israel) and used for the qPCR analysis. Ribosomal protein L-13 (Bta04282; *rpl-13*) was used as a reference gene to normalize the expression levels of BtGSS1 (based on the information obtained from Dr. Osnat Malka).

Melting curve analyses were conducted at the end of qPCR runs by gradually heating the PCR reaction mixture from 50°C to 95°C, to detect genomic DNA contamination and the formation of primer dimers. Additionally, to ensure primer specificity, the qPCR amplicons were analysed by agarose gel electrophoresis, purified, cloned into the pCR*TM*4-TOPO® vector and were sequence confirmed. Expression levels of the *BtGSS1* were quantified relatively using the ΔΔCP method (Pfaffl 2001).

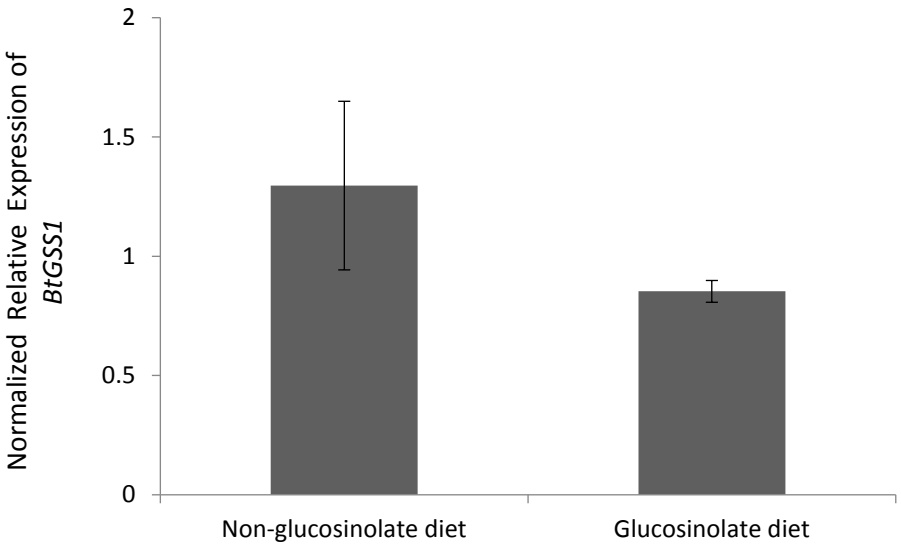


Figure 15: Comparison of relative expression levels of *BtGSS1* in *B. tabaci* adult insects that consumed non-glucosinolate diet (eggplant) and those which consumed glucosinolate diet (Brussels sprouts).

BtGSS1 was observed to be constitutively expressed irrespective of whether the insects consumed glucosinolate or non-glucosinolate diet (Figure 15). The expression of *BtGSS1* was observed to be higher, relative to the expression of *rpl-13*, suggesting that in general *BtGSS1* is highly expressed in MEAM1 *B. tabaci* species.

6.9 Phylogenetic analysis

To examine whether PxGSS and the identified GSS in *B. tabaci* (BtGSS1) share a common evolutionary origin, a phylogenetic analysis was performed. The shortlisted putative BtSulf candidates (Table 1) and the PxGSS amino acid sequences were used for the phylogenetic analysis.

The phylogenetic analysis revealed that PxGSS and the identified GSS in *B. tabaci* (BtGSS1) were not closely related, suggesting that *B. tabaci* might have evolved its GSS after speciation, from a different ancestral gene.

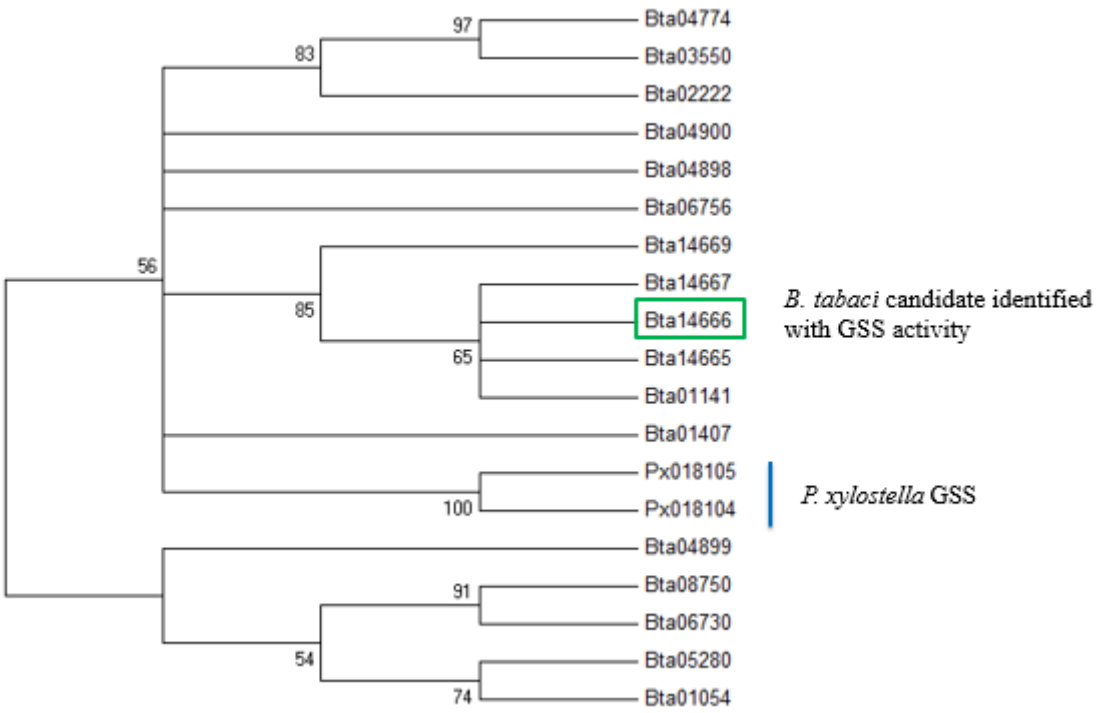


Figure 16: Phylogenetic tree of shortlisted putative *B. tabaci* GSS candidates and the *P. xylostella* GSS amino acid sequences. Protein sequences were aligned using MUSCLE algorithm, a consensus phylogenetic tree was constructed using the neighbor-joining method and assessed using bootstrapping method (1000 bootstrap replicates); branches corresponding to partitions reproduced in less than 50% bootstrap replicates were collapsed.

7. Discussion

Although cruciferous plants possess a sophisticated glucosinolate-myrosinase defence system to thwart their herbivorous enemies, several insects have devised their own mechanisms to detoxify these plant defensive metabolites. One such strategy is to employ a GSS (glucosinolate sulfatase) enzymatic system to convert glucosinolates to their corresponding desulfated forms, and thus prevent them from being broken down to toxic downstream products by myrosinases (Ratzka et al. 2002; Falk and Gershenzon 2007). It was interesting that this detoxification mechanism is also present in the phloem-feeding insect *B. tabaci*, as this activity had been observed only in leaf-chewing insects until recently (Malka et al. 2016). As *B. tabaci* is considered to be a major agricultural pest, it was of interest to identify the GSS(s) performing the detoxification in this insect system.

We successfully identified one GSS (BtGSS1) in *B. tabaci* which actively desulfated different glucosinolate classes when expressed heterologously. The identified enzyme was found to have a temperature optimum around 55°C and pH optimum around 7.7. The enzyme was secreted into the extracellular culture medium by Sf9 cells. Transcriptomic data suggested an elevated expression of this enzyme in the *B. tabaci* gut tissues relative to the rest of the body. These data together suggest that, *in vivo*, this enzyme could be secreted into the insect's gut lumen, where it can efficiently detoxify the potentially harmful glucosinolates after ingestion. It is commonly believed that phloem-feeding insects, by ensuring careful piercing of plant tissues and stealthy sucking of the plant sap, do not strongly activate the glucosinolate-myrosinase system (Walling 2008; Pentzold et al. 2014). Indeed, upon *B. tabaci* infestation on *A. thaliana*, no induced expression of myrosinase-encoding genes was observed by microarray analysis (Kempema et al. 2007). However, a recent study shows that phloem-feeding indeed leads to glucosinolate activation and production of isothiocyanates (Danner et al. 2017). In addition, accumulation of transcripts that encode for β-glucosidase–like protein (SLW) has been reported in response to *B. tabaci* feeding, in squash (*Curcurbita pepo* cv Chefini) (Van de Ven et al. 2000). These studies together justify the need for an efficient detoxification mechanism to cope with glucosinolate toxicity, even in piercing-sucking insects such as whiteflies.

Our findings show that the identified *B. tabaci* GSS (BtGSS1) exhibits a clear preference towards indolic glucosinolates compared to their aliphatic counterparts. This result shows us that the *B. tabaci* probably employs the GSS detoxification machinery to combat the myrosinase-independent breakdown of indolic glucosinolates (Barth and Jander 2006). Also,

induced production of indolic glucosinolate has been reported in *A. thaliana* in response to *B. tabaci* feeding (Mikkelsen et al 2000; Kempema et al. 2007), suggesting these compounds in particular could be active against this herbivore. Moreover, transcriptomic evidence in few phloem-feeding insect infestations in *A. thaliana*, such as in *B. tabaci, M. persicae* and *Brevicoryne brassicae,* reports an up-regulation of *myb51*, a positive regulator of indolic glucosinolate accumulation (Foyer at al., 2014). All these findings taken together suggests that the plant specifically induces the production of indolic glucosinolates in order to target phloem-feeding insects, as these insects normally escape the general glucosinolate-myrosinase chemical defence system involving two-component activation. Thus, the plants impose a requirement of specialised strategies on the phloem-feeding insects, to handle the non-enzymatic breakdown of indolic glucosinolates.

Another possibility could be that the indolic glucosinolates could be broken down by untypical β-thioglucosidases, some of which might be more specific to indolic glucosinolates. The enzymatic breakdown products of indolic glucosinolates are said to have strong antifeedant and toxic effects on the crucifer-consuming insects (Agerbirk et al., 2009), and plants can also convert indolyl-3-methyl glucosinolate into more bioactive methoxylated forms upon aphid attack (Wiesner et al. 2013).

Quantitative real-time PCR analysis comparing gene expression levels showed that the gene encoding the identified *B. tabaci* GSS (BtGSS1) is constitutively expressed irrespective of whether the insect consumed glucosinolates or not. This suggests that this GSS could have another (innate) function for the insect itself as a more general arylsulfatase. While the in-vivo functions of arylsulfatases are still not well-defined, one of their proposed roles is in the moulting process of southern armyworm larvae (Pellicia et al. 1987). It is known that the insects store ecdysone, the steroidal moulting hormone, as inactive sulfated conjugates (Koolman 1973). Arylsulfatases could therefore be involved in converting the ecdysone to their active form. However, the ecdysteroid localisation is limited to the ovaries, haemolymph, and eggs, and not in the gut (King and Siddall 1974). According to previous research studies of the *P. xylostella* GSS and *S. gregaria* GSS, and also in accordance with what we observed with the identified *B. tabaci*, the GSS is secreted into the gut lumen (Ratzka et al. 2002; Falk and Gershenzon 2007). So, it seems unlikely that BtGSS1 is involved in ecdysone metabolism. However, we have tested the BtGSS1 expression levels only in MEAM1 insects, and could be that MED and other *B. tabaci* species might either have lower levels of GSS expression, and are therefore less adapted to feeding on crucifers.

So far, in *B. tabaci*, the only identified glucosinolate detoxification mechanism is desulfation using a GSS. Additional strategies could also be employed by this insect to circumvent glucosinolate toxicity. In another generalist phloem-feeding insect, *M. persicae*, several glucosinolate detoxification mechanisms have been reported, such as rapid excretion, efficient metabolism, and repression of toxicity by forming glucosinolate conjugates (De vos et al. 2005; Elzinga et al. 2014; Kim et al. 2008).

So far, we have identified only one GSS in *B. tabaci*. Future experiments should attempt to confirm whether other phylogenetically close sulfatases (e.g. BtSulf5) have GSS activity., If BtGSS1 is shown to be the most important GSS enzyme in vivo, it could serve as a potential target for *B. tabaci* pest management. This could be achieved for example by generating transgenic plants producing dsRNA targeting the *B. tabaci* GSS and accumulating these in the phloem sap, thereby silencing the GSS and impairing the insect's detoxification system. Alternatively, *B. tabaci* mutants lacking GSS activity could be raised and released in nature to mate with the wild-type. The homozygous mutants could be made incapable of producing off-springs, and thus ensuring a control over the spread of transgenic *B. tabaci*. However, all these would be possible only after we silence the identified GSS gene(s) using RNA interference (RNAi) and see how the *B. tabaci* performs when reared on glucosinolate rich diet. This would provide insights whether the insect has other strategies to cope with glucosinolate toxicity, and tells us how important is this detoxification mechanism, for *B. tabaci* to successfully feed on glucosinolate containing plants.

8. Conclusion

The glucosinolate-myrosinase system is a sophisticated two-component chemical defence strategy present in plants of the order Brassicales. The constituents of this binary defence system are kept spatially segregated, but during herbivory, myrosinases catalyse the hydrolysis of glucosinolates forming toxic breakdown products. Faced with this efficient defence system, some insects have developed strategies to cope with glucosinolate toxicity and successfully feed on cruciferous plants. But while a long-standing paradigm assumes that phloem-feeding insects don't cause tissue disruption and therefore do not activate the glucosinolate-myrosinase system, a growing amount of evidence suggests the contrary. Plants induce the production of glucosinolates (especially indolic ones) upon infestation by phloem feeders; levels of glucosinolates influence the performance of these insects; glucosinolate hydrolysis products have been detected during attack by aphids; and these hydrolysis products negatively affect phloem-feeding insects. However, the strategies used by phloem-feeding insects that feed on glucosinolate-containing plants to circumvent these defensive metabolites are not yet well understood.

Bemisia tabaci is a phloem-feeding insect that efficiently detoxifies glucosinolates by conversion to their corresponding desulfated forms, which are unsusceptible to hydrolysis by myrosinases. *B. tabaci* is considered a serious agricultural pest causing extensive damage to crops including cruciferous vegetables, and also vectors many plant pathogenic viruses. However, while the desulfated products had been reported previously in the insect honeydew, the enzyme(s) responsible for this conversion were not studied.

In this thesis, we have successfully identified and initially characterised a *B. tabaci* sulfatase performing glucosinolate metabolism. The identified glucosinolate sulfatase (GSS) exhibited a strong preference in vitro towards indolic glucosinolates in comparison to aliphatic and benzenic glucosinolates, implying that *B. tabaci* employs this GSS detoxification mechanism to cope with the negative effects of indolic glucosinolate breakdown, a process that can happen even independently of myrosinase activity. The gene encoding the *B. tabaci* GSS was expressed both in insects that consumed glucosinolates and those that did not, suggesting that the GSS could have an innate role in *B. tabaci*. However, the exact function of sulfatases and their biological substrates in insects remain unclear. A phylogenetic analysis revealed that *P. xylostella* GSS and *B. tabaci* GSS are not the closest

relatives within these two species, suggesting that *B. tabaci* could have gained the GSS function much later in evolution, possibly by gene duplication during speciation.

So far, in *B. tabaci*, the only identified glucosinolate detoxification mechanism is desulfation. Other additional strategies may also be employed by the insect to circumvent glucosinolate toxicity (especially against the abundant aliphatic glucosinolates), which need to be investigated further via chemical analysis. These studies should also be followed by heterologous expression and mutagenesis strategies. Gene silencing efforts would also help provide more information on how important these detoxification mechanisms are for *B. tabaci* species that feed on cruciferous plants, and how they might help this insect colonize and become a successful pest on these crop plants.

References

Agerbirk, N., De Vos, M., Kim, J. H., & Jander, G. (2008). Indole glucosinolate breakdown and its biological effects. *Phytochemistry Reviews*, *8*(1), 101. https://doi.org/10.1007/s11101-008-9098-0

Ávila, F. W., Faquin, V., Yang, Y., Ramos, S. J., Guilherme, L. R. G., Thannhauser, T. W., & Li, L. (2013). Assessment of the Anticancer Compounds Se-Methylselenocysteine and Glucosinolates in Se-Biofortified Broccoli (Brassica oleracea L. var. italica) Sprouts and Florets. *Journal of Agricultural and Food Chemistry*, *61*(26), 6216–6223. https://doi.org/10.1021/jf4016834

Barth, C., & Jander, G. (2006). Arabidopsis myrosinases TGG1 and TGG2 have redundant function in glucosinolate breakdown and insect defense. *The Plant Journal*, *46*(4), 549–562. https://doi.org/10.1111/j.1365-313X.2006.02716.x

Berlinger, M. J. (1986). *Host plant resistance to Bemisia tabaci*. *Agriculture, Ecosystems & Environment* (Vol. 17). https://doi.org/10.1016/0167-8809(86)90028-9

Berry, P. E. (2015). Brassicales. In *Encyclopaedia Britannica*. Encyclopædia Britannica, inc. Retrieved from https://www.britannica.com/plant/Brassicales

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, *72*(1–2), 248–254. https://doi.org/10.1016/0003-2697(76)90527-3

Brown, J. K., Frohlich, D. R., & Rosell, R. C. (1995). The Sweetpotato or Silverleaf Whiteflies: Biotypes of Bemisia tabaci or a Species Complex? *Annual Review of Entomology*, *40*(1), 511–534. https://doi.org/10.1146/annurev.en.40.010195.002455

Brown, P. D., & Morra, M. J. (1995). Glucosinolate-containing plant tissues as bioherbicides. *Journal of Agricultural and Food Chemistry*, *43*(12), 3070–3074. https://doi.org/10.1021/jf00060a015

Buono, M., & Cosma, M. P. (2010). Sulfatase activities towards the regulation of cell metabolism and signaling in mammals. *Cellular and Molecular Life Sciences*, *67*(5), 769–780. https://doi.org/10.1007/s00018-009-0203-3

Chen, W., Hasegawa, D. K., Kaur, N., Kliot, A., Pinheiro, P. V., Luan, J., … Fei, Z. (2016). The draft genome of whitefly Bemisia tabaci MEAM1, a global crop pest, provides novel insights into virus transmission, host adaptation, and insecticide resistance. *BMC Biology*, *14*(1), 110. https://doi.org/10.1186/s12915-016-0321-y

Cohen, S., & Berlinger, M. J. (1986). Transmission and cultural control of whitefly-borne viruses. *Agriculture, Ecosystems & Environment*, *17*(1), 89–97. https://doi.org/https://doi.org/10.1016/0167-8809(86)90030-7

Danner, H., Desurmont, G. A., Cristescu, S. M., & van Dam, N. M. (2017). Herbivore-induced plant volatiles accurately predict history of coexistence, diet breadth, and feeding mode of herbivores. *New Phytologist*, n/a-n/a. https://doi.org/10.1111/nph.14428

De Barro, P. J., Liu, S.-S., Boykin, L. M., & Dinsdale, A. B. (2010). Bemisia tabaci: A Statement of Species Status. *Annual Review of Entomology*, *56*(1), 1–19. https://doi.org/10.1146/annurev-ento-112408-085504

De Vos, M., Van Oosten, V. R., Van Poecke, R. M. P., Van Pelt, J. A., Pozo, M. J., Mueller, M. J., … Pieterse, C. M. J. (2005). Signal Signature and Transcriptome Changes of Arabidopsis During Pathogen and Insect Attack. *Molecular Plant-Microbe Interactions*, *18*(9), 923–937. https://doi.org/10.1094/MPMI-18-0923

Dierks, T., Lecca, M. R., Schlotterhose, P., Schmidt, B., & von Figura, K. (1999). Sequence determinants directing conversion of cysteine to formylglycine in eukaryotic sulfatases. *The EMBO Journal*, *18*(8), 2084–2091. https://doi.org/10.1093/emboj/18.8.2084

Elzinga, D. A., De Vos, M., & Jander, G. (2014). Suppression of plant defenses by a Myzus persicae (green peach aphid) salivary effector protein. *Molecular Plant-Microbe Interactions : MPMI*, *27*(7), 747–756. https://doi.org/10.1094/MPMI-01-14-0018-R

Engel, E., Baty, C., le Corre, D., Souchon, I., & Martin, N. (2002). Flavor-Active Compounds Potentially Implicated in Cooked Cauliflower Acceptance. *Journal of Agricultural and Food Chemistry*, *50*(22), 6459–6467. https://doi.org/10.1021/jf025579u

Fahey, J. W., Stephenson, K. K., Wade, K. L., & Talalay, P. (2013). Urease from Helicobacter pylori is inactivated by sulforaphane and other isothiocyanates. *Biochemical and Biophysical Research Communications*, *435*(1), 1–7. https://doi.org/10.1016/j.bbrc.2013.03.126

Fahey, J. W., Zalcmann, A. T., & Talalay, P. (2001). The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry*, *56*(1), 5–51. https://doi.org/https://doi.org/10.1016/S0031-9422(00)00316-2

Falk, K. L., & Gershenzon, J. (2007). The Desert Locust, Schistocerca gregaria, Detoxifies the Glucosinolates of Schouwia purpurea by Desulfation. *Journal of Chemical Ecology*, *33*(8), 1542–1555. https://doi.org/10.1007/s10886-007-9331-0

Foyer, C. H., Verrall, S. R., & Hancock, R. D. (2015). Systematic analysis of phloem-feeding insect-induced transcriptional reprogramming in Arabidopsis highlights common features and reveals distinct responses to specialist and generalist insects. *Journal of Experimental Botany*, *66*(2), 495–512. Retrieved from http://dx.doi.org/10.1093/jxb/eru491

Greathead, D. J., & Greathead, A. H. (1992). Biological control of insect pests by parasitoids and predators: the BIOCAT database. *Biocontrol News and Information*.

Halkier, B. A., & Gershenzon, J. (2006). BIOLOGY AND BIOCHEMISTRY OF GLUCOSINOLATES. *Annual Review of Plant Biology*, *57*(1), 303–333. https://doi.org/10.1146/annurev.arplant.57.032905.105228

Hall, J. C., Sytsma, K. J., & Iltis, H. H. (2002). Phylogeny of Capparaceae and Brassicaceae based on chloroplast sequence data. *American Journal of Botany* , *89*(11), 1826–1842. https://doi.org/10.3732/ajb.89.11.1826

Höglund, a S., Lenman, M., Falk, a, & Rask, L. (1991). Distribution of myrosinase in rapeseed tissues. *Plant Physiology*, *95*(1), 213–21. https://doi.org/Doi 10.1104/Pp.95.1.213

Ishida, M., Hara, M., Fukino, N., Kakizaki, T., & Morimitsu, Y. (2014). Glucosinolate metabolism, functionality and breeding for the improvement of Brassicaceae vegetables. *Breeding Science*, *64*(1), 48–59. https://doi.org/10.1270/jsbbs.64.48

Jeschke, V., Gershenzon, J., & Vassão, D. G. (2016). A mode of action of glucosinolate-derived isothiocyanates: Detoxification depletes glutathione and cysteine levels with ramifications on protein metabolism in Spodoptera littoralis. *Insect Biochemistry and Molecular Biology*, *71*, 37–48. https://doi.org/https://doi.org/10.1016/j.ibmb.2016.02.002

Jones, D. R. (2003). Plant Viruses Transmitted by Whiteflies. *European Journal of Plant Pathology*, *109*(3), 195–219. https://doi.org/10.1023/A:1022846630513

K., G., & P., M. (1979). Vacuolar location of glucosinolates in horseradish root cells. *Plant Science Letters*.

Karlson, P., & Koolman, J. (1973). On the metabolic fate of ecdysone and 3-dehydroecdysone in Calliphora vicina. *Insect Biochemistry*, *3*(12), 409–417. https://doi.org/https://doi.org/10.1016/0020-1790(73)90074-7

Kawakishi, S., & Kaneko, T. (1987). Interaction of proteins with allyl isothiocyanate. *Journal of Agricultural and Food Chemistry (USA)*.

Kazana, E., Pope, T. W., Tibbles, L., Bridges, M., Pickett, J. A., Bones, A. M., … Rossiter, J. T. (2007). The cabbage aphid: a walking mustard oil bomb. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1623), 2271–2277. https://doi.org/10.1098/rspb.2007.0237

Kempema, L. A., Cui, X., Holzer, F. M., & Walling, L. L. (2007). Arabidopsis Transcriptome Changes in Response to Phloem-Feeding Silverleaf Whitefly Nymphs. Similarities and Distinctions in Responses to Aphids. *Plant Physiology*, *143*(2), 849–865. https://doi.org/10.1104/pp.106.090662

Kertesz, M. A. (2000). Riding the sulfur cycle – metabolism of sulfonates and sulfate esters in Gram-negative bacteria. *FEMS Microbiology Reviews*, *24*(2), 135–175. https://doi.org/10.1111/j.1574-6976.2000.tb00537.x

Kim, J. H., Lee, B. W., Schroeder, F. C., & Jander, G. (2008). Identification of indole glucosinolate breakdown products with antifeedant effects on Myzus persicae (green peach aphid). *The Plant Journal*, *54*(6), 1015–1026. https://doi.org/10.1111/j.1365-313X.2008.03476.x

King, D. S., & Siddall, J. B. (1974). Biosynthesis and Inactivation of Ecdysone BT - Invertebrate Endocrinology and Hormonal Heterophylly. In W. J. Burdette (Ed.) (pp. 147–152). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-65769-6_12

Koroleva, O. A., & Cramer, R. (2011). Single-cell proteomic analysis of glucosinolate-rich S-cells in Arabidopsis thaliana. *Methods*, *54*(4), 413–423. https://doi.org/https://doi.org/10.1016/j.ymeth.2011.06.005

Liu, B., Preisser, E. L., Chu, D., Pan, H., Xie, W., Wang, S., … Zhang, Y. (2013). Multiple Forms of Vector Manipulation by a Plant-Infecting Virus: Bemisia tabaci and Tomato

Yellow Leaf Curl Virus. *Journal of Virology*, *87*(9), 4929–4937. https://doi.org/10.1128/JVI.03571-12

Luo, C., Jones, C., Devine, G., Zhang, F., Denholm, I., & Gorman, K. (2010). *Insecticide resistance in Bemisia tabaci biotype Q (Hemiptera: Aleyrodidae) from China*. *Crop Protection* (Vol. 29). https://doi.org/10.1016/j.cropro.2009.10.001

Lüthy, B., & Matile, P. (1984). The mustard oil bomb: Rectified analysis of the subcellular organisation of the myrosinase system. *Biochemie Und Physiologie Der Pflanzen*, *179*(1–2), 5–12. https://doi.org/10.1016/S0015-3796(84)80059-1

Malka, O., Shekhov, A., Reichelt, M., Gershenzon, J., Vassão, D. G., & Morin, S. (2016). Glucosinolate Desulfation by the Phloem-Feeding Insect Bemisia tabaci. *Journal of Chemical Ecology*, *42*(3), 230–235. https://doi.org/10.1007/s10886-016-0675-1

Matile, P. (1980). Die Senfolbombe: Zur Kompartimentierung des Myrosinasesystems. . *Biochem. Physiol. Pflanzen*, *175*, 722–731. https://doi.org/10.1016/S0015-3796(80)80059-X

Mauricio, R., & Rausher, M. D. (1997). Experimental Manipulation of Putative Selective Agents Provides Evidence for the Role of Natural Enemies in the Evolution of Plant Defense. *Evolution*. https://doi.org/10.2307/2411196

McKenzie, C. L., Bethke, J. A., Byrne, F. J., Chamberlin, J. R., Dennehy, T. J., Dickey, A. M., … Shatters, R. G. (2012). Distribution of Bemisia tabaci (Hemiptera: Aleyrodidae) Biotypes in North America After the Q Invasion. *Journal of Economic Entomology*, *105*(3), 753–766. https://doi.org/10.1603/EC11337

Mikkelsen, M. D., Hansen, C. H., Wittstock, U., & Halkier, B. A. (2000). Cytochrome P450 CYP79B2 from Arabidopsis catalyzes the conversion of tryptophan to indole-3-acetaldoxime, a precursor of indole glucosinolates and indole-3-acetic acid. *Journal of Biological Chemistry*, *275*(43), 33712–33717. https://doi.org/10.1074/jbc.M001667200

Pelliccia, J. G., Freshler, A., Ladd, C., & Richards, S. (1987). There are two distinct arylsulfatase activities inDrosophila. *Biochemical Genetics*, *25*(7), 459–464. https://doi.org/10.1007/BF00554348

Pentzold, S., Zagrobelny, M., Rook, F., & Bak, S. (2014). How insects overcome two-component plant chemical defence: plant β-glucosidases as the main target for herbivore adaptation. *Biological Reviews*, *89*(3), 531–551. https://doi.org/10.1111/brv.12066

Perring, T. M. (2001). The Bemisia tabaci species complex. *Crop Protection*, *20*(9), 725–737. https://doi.org/10.1016/S0261-2194(01)00109-0

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Research*, *29*(9), e45–e45. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC55695/

Rask, L., Andreasson, E., Ekbom, B., Eriksson, S., Pontoppidan, B., & Meijer, J. (2000). *Myrosinase: Gene Family Evolution and Herbivore Defense in Brassicaceae*. *Plant Molecular Biology* (Vol. 42). https://doi.org/10.1023/A:1006380021658

Ratzka, A., Vogel, H., Kliebenstein, D. J., Mitchell-Olds, T., & Kroymann, J. (2002). Disarming the mustard oil bomb. *Proceedings of the National Academy of Sciences* , *99*(17), 11223–11228. https://doi.org/10.1073/pnas.172112899

Roy, A. B. B. T.-M. in E. (1987). Sulfatases from Helix pomatia. In *Sulfur and Sulfur Amino Acids* (Vol. 143, pp. 361–366). Academic Press. https://doi.org/https://doi.org/10.1016/0076-6879(87)43064-4

Siemens, D. H., & Mitchell-Olds, T. (1998). Evolution of Pest-Induced Defenses in Brassica Plants: Tests of Theory. *Ecology*, *79*(2), 632–646. https://doi.org/10.2307/176959

van de Ven, W. T. G., LeVesque, C. S., Perring, T. M., & Walling, L. L. (2000). Local and Systemic Changes in Squash Gene Expression in Response to Silverleaf Whitefly Feeding. *The Plant Cell*, *12*(8), 1409–1424. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC149112/

Walling, L. L. (2008). Avoiding Effective Defenses: Strategies Employed by Phloem-Feeding Insects. *Plant Physiology*, *146*(3), 859–866. https://doi.org/10.1104/pp.107.113142

Wiesner, M., Hanschen, F. S., Schreiner, M., Glatt, H., & Zrenner, R. (2013). Induced Production of 1-Methoxy-indol-3-ylmethyl Glucosinolate by Jasmonic Acid and Methyl Jasmonate in Sprouts and Leaves of Pak Choi (Brassica rapa ssp. chinensis). *International Journal of Molecular Sciences*, *14*(7), 14996–15016. https://doi.org/10.3390/ijms140714996

Wittstock, U., Agerbirk, N., Stauber, E. J., Olsen, C. E., Hippler, M., Mitchell-Olds, T., … Vogel, H. (2004). Successful herbivore attack due to metabolic diversion of a plant chemical defense. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(14), 4859–4864. https://doi.org/10.1073/pnas.0308007101

You, M., Yue, Z., He, W., Yang, X., Yang, G., Xie, M., … Wang, J. (2013). A heterozygous moth genome provides insights into herbivory and detoxification. *Nature Genetics*, *45*(2), 220–225. https://doi.org/10.1038/ng.2524

Zasada, I. A., & Ferris, H. (2004). Nematode suppression with brassicaceous amendments: Application based upon glucosinolate profiles. *Soil Biology and Biochemistry*, *36*(7), 1017–1024. https://doi.org/10.1016/j.soilbio.2003.12.014

9. Supplements

9.1 List of chemicals

Acetonitrile	VWR Chemicals (Radnor, USA)
Agarose	Carl Roth (Karlsruhe, Germany)
Ampicillin	Carl Roth (Karlsruhe, Germany)
Blasticidin S	Sigma-Aldrich (Steinheim, Germany)
β-mercaptoethanol	Sigma-Aldrich (Steinheim, Germany)
Bromphenol blue	Sigma-Aldrich (Steinheim, Germany)
p-coumaric acid	Sigma-Aldrich (Steinheim, Germany)
DMSO	Acros Organics (New Jersey, USA)
DTT (dithiotreitol)	Carl Roth (Karlsruhe, Germany)
EDTA disodium salt	Carl Roth (Karlsruhe, Germany)
Ethanol	Merck Millipore (Darmstadt, Germany)
Formic acid	Fischer Chemical (Geef, Belgium)
Glacial acetic acid	Carl Roth (Karlsruhe, Germany)
Glycerol	FLUKA, Sigma-Aldrich (Steinheim, Germany)
Glycine	Carl Roth (Karlsruhe, Germany)
H_2O_2 (hydrogen peroxide) 30% (w/w) in H_2O	FLUKA, Sigma-Aldrich (Steinheim, Germany)
Imidazole	Sigma-Aldrich (Steinheim, Germany)
K_2HPO_4 (dipotassium phosphate)	Sigma-Aldrich (Steinheim, Germany)
Kanamycin	Duchefa (Haarlem, Netherlands)

KCl (potassium chloride)	Carl Roth (Karlsruhe, Germany)
KCl (potassium chloride)	Carl Roth (Karlsruhe, Germany)
KH_2PO_4 (monopotassium phosphate)	Sigma-Aldrich (Steinheim, Germany)
LB-Agar (Luria/Miller)	Carl Roth (Karlsruhe, Germany)
low-fat powdered milk	FLUKA, Sigma-Aldrich (Steinheim, Germany)
Luminol	FLUKA, Sigma-Aldrich (Steinheim, Germany)
Methanol	Merck Millipore (Darmstadt, Germany)
Na_2HPO_4 (disodium phosphate)	Carl Roth (Karlsruhe, Germany)
NaCl (sodium chloride)	Carl Roth (Karlsruhe, Germany)
NaH_2PO_4 (monosodium phosphate)	Carl Roth (Karlsruhe, Germany)
4-nitrocatechol sulfate	FLUKA, Sigma-Aldrich (Steinheim, Germany)
SDS (sodium dodecyl sulfate)	Carl Roth (Karlsruhe, Germany)
Sinigrin	Sigma-Aldrich (Steinheim, Germany)
Tris (tris(hydroxymethyl)-aminomethane)	Carl Roth (Karlsruhe, Germany)
Tween®20	Sigma-Aldrich (Steinheim, Germany)
yeast extract	Carl Roth (Karlsruhe, Germany)

9.2 List of Primers

9.2.1 Primers used for the amplification of *B. tabaci* arylsulfatase gene candidates

Table S 1: Primers used for the amplification of *B. tabaci* arylsulfatase gene candidates. The list includes the sequences (5'-3' direction) and the amplicon size of the respective genes.

Gene	Name	Sequence (5'-3')	Tm (°C)	Amplicon size (bp)
Bta02222	Bta02222_FP	ACCATGATCACTTCGGCCGGAT	68.7	1500
	Bta02222_RP	CATAACATCATTTTCAGTGATGAGGTC	65.6	
Bta04774	Bta04774_FP	ACCATGGTCTCAAATATCTGGACCGC	66.2	1854
	Bta04774_RP	AATAACTCGTCCGCCGTTTC	66	
Bta03550	Bta03550_FP	ACCATGATAAAATTGAGGGCTTGGG	65	1731
	Bta03550_RP	AAAAAGTTTTCCTCCGTTTCTC	66	
Bta06756	Bta06756_FP	ACCATGACACGCCGCCAAGAA	67.9	2001
	Bta06756_RP	CGTGACAACAGAGCCTTTCTG	65.3	
Bta14666	Bta14666_FP	'ACCATGTTTTGCAATAATTTTATCAAAACAGTG	64.9	1434
	Bta14666_RP	GTCCAACCATGGAGCCCA	67.7	
Bta14667	Bta14667_FP	ACCATGTTTTCTCACGGAAAGATCAAAA	66.4	1707
	Bta14667_RP	AGACATGCCCCTGCCCTC	68.1	

9.2.2 Primers used for quantitative real-time PCR (qPCR)

Table S 2: Primers used for qRT-PCR analysis of BtGSS11 responsible for glucosinolate desulfation and housekeeping gene from *B. tabaci*. The list includes the sequences (5'-3' direction).

Gene	Name	Sequence
Bta14666	14666-qPCR-p3-FP	5'-GATCCCTGCGAGTTCCACAA-3'
	14666-qPCR-p3-RP	5'-CTTGCGATTCGACCAACTGC-3'
rpl-13	rpl-FP	5'-CATTCCACTACAGAGCTCCA-3'
	rpl-RP	5'-TTTCAGGTTTCGGATGGCTT-3'

9.2.3 Primers used for Sanger Sequencing

Table S 3: Primers used for Sanger sequencing to confirm the sequence and right orientation of qPCR amplicons in the respective expression or sequencing vector. The list includes the sequences (5'-3' direction).

Vector	Name	Sequence
pIB/V5-His-TOPO®	OpIE2 Fwd	5'-CGCAACGATCTGGTAAACAC-3'
	OpIE2 Rev	5'-GACAATACAAACTAAGATTTAGTCAG-3'
pCR® 4Blunt-TOPO®	T7 primer	5'-CCCTATAGTGAGTCGTATTA-3'
	M13 Rev	5'-CAGGAAACAGCTATGAC-3'

9.3 Parameters for multiple reaction monitoring (MRM)

9.3.1 Multiple reaction monitoring (MRM) analyses of intact glucosinolates (negative ionization mode)

Table S 4: LC-MS parameters used for the multiple reaction monitoring (MRM) analyses of intact glucosinolates (negative ionization mode). CE: collision energy; CEP: collision cell entrance potential; CXP: collision cell exit potential; DP: declustering potential; EP: entrance potential; 4msob = 4-methylsulfinylbutyl; 3msop = 3-methylsulfinylpropyl; 4mtb = 4-methylthiobutyl; 7msoh = 7 methylsulfinylheptyl; 5msop = 5-methylsulfinylpentyl; 8msoo = 8-methylsulfinyloctyl; 4moi3m = 4-methoxyindol-3-ylmethyl; i3m = indol-3-ylmethyl.

Analyte	Q1 (m/z)	Q3 (m/z)	time (min)	DP (V)	EP (V)	CEP (V)	CE (V)	CXP (V)
3msop	421.8	95.9	7.6	-65	-4.5	-18	-60	0
4msob	435.9	95.8	9.5	-65	-5	-16	-60	0
8msoo	492.1	95.8	13.6	-75	-4.5	-24	-58	0
5msop	449.9	95.8	10.5	-65	-5	-16	-60	0
7msoh	477.9	95.8	12.5	-65	-5	-16	-60	0
4mtb	419.9	95.9	14.1	-60	-11	-16	-58	0
i3m	447	95.8	15.8	-65	-12	-18	-50	0
4moi3m	477.1	95.8	16.3	-65	-12	-18	-50	0

9.3.2 Multiple reaction monitoring (MRM) analyses of desulfated glucosinolates (positive ionization mode)

Table S 5: LC-MS parameters used for the multiple reaction monitoring (MRM) analyses of desulfated glucosinolates (positive ionization mode). CE: collision energy; CEP: collision cell entrance potential; CXP: collision cell exit potential; DP: declustering potential; EP: entrance potential; 4msob = 4-methylsulfinylbutyl; 3msop = 3-methylsulfinylpropyl; 4mtb = 4-methylthiobutyl; 7msoh = 7 methylsulfinylheptyl; 5msop = 5-methylsulfinylpentyl; 8msoo = 8-methylsulfinyloctyl; 4moi3m = 4-methoxyindol-3-ylmethyl; i3m = indol-3-ylmethyl; p-OHBenz = p-hydroxybenzyl

Analyte	Q1 (m/z)	Q3 (m/z)	time (min)	DP (V)	EP (V)	CEP (V)	CE (V)	CXP (V)
dsAllyl	280	118	9	30	5	15	15	5
ds-3msop	344	182	7	30	5	15	15	5
ds-4msob	358	196	8.5	30	5	15	15	5
ds-5msop	372	210	9.5	30	5	15	15	5
ds-7msoh	400	238	11.5	30	5	15	15	5
ds-8msoo	414	252	12	30	5	15	15	5
ds-4mtb	342	180	12.5	30	5	15	15	5
ds-i3m	369	207	13	30	5	15	15	5
ds-pOHBenz	346	184	10.5	30	5	15	15	5
ds-4moi3m	399.1	237	13.5	30	5	15	15	5

9.4 Molecular mass of the heterologously expressed GSS

Table S 6: Molecular mass of the heterologously expressed GSS. The list includes the protein sequence length and molecular mass.

Gene ID	Name	Sequence length (aa)	Molecular mass (kDa)
Bta06756	BtSulf1	666	74.6
Bta03550	BtSulf2	576	64.8
Bta02222	BtSulf3	549	61.47
Bta04774	BtSulf4	617	69.48
Bta14667	BtSulf5	568	63.7
Bta14666	BtSulf11	477	54.07
Px018104	PxGSS	547	60.79

10. Acknowledgements

I would like to thank **Prof. Dr. Jonathan Gershenzon** for the opportunity and the support to conduct my Master's thesis at the Max Planck Institute for Chemical Ecology, Department of Biochemistry. I would also like to thank my second supervisor, **Prof. Dr. Wilhelm Boland** for his support.

I owe my sincere gratitude to my project guide **Dr. Daniel Giddings Vassão** for his guidance and encouragement. I am thankful to him for believing in me and giving me an opportunity to work under his guidance. A big thanks to him for all the time he had invested in guiding me during my project work, which has absolutely helped me to enhance my knowledge and my professional conduct.

A very huge thanks to **Katrin Luck** for helping me with molecular techniques and cell culturing techniques, and for all the tight hugs and encouraging words when it was needed the most. I would like to thank all my group members **Felipe**, **Verena**, **Anton**, **Micheal**, **Elena** and **Anna** for the wonderful working environment, and for all the helpful scientific discussions. I would like to thank **Dr. Michael Reichelt** for helping me with the analytical techniques.

I thank **Thomas** and **Laura** for their immense emotional support and companionship. Thank you so much for being there for me every single time.

Finally, I want to thank my family for the constant love, support and motivation. I acknowledge my debt of thanks to my dad and mom for constantly persuading me to follow my passion and for their infinite love and encouragement. I also thank my brother for his huge moral support, despite being miles far away.

11. Selbständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

_____ _____

Ort, Datum Abinaya Manivannan