1-1-2011

Homoglutathione Synthetase And The Plant Thiol-Redox Proteome

Ashley Galant
Washington University in St. Louis

WASHINGTON UNIVERSITY

Division of Biology and Biomedical Sciences

Plant Biology

Dissertation Examination Committee

Joseph M. Jez, Chair

Robert E. Blankenship

Ursula W. Goodenough

Leslie M. Hicks

Tuan-Hua David Ho

Barbara N. Kunkel

Thomas J. Smith

HOMOGLUTATHIONE SYNTHETASE AND THE PLANT THIOL-REDOX PROTEOME

By

Ashley L. Galant

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2011

Saint Louis, Missouri

ABSTRACT OF THE DISSERTATION

Homoglutathione Synthetase and the Plant Thiol-Redox Proteome

by

Ashley Galant

Doctor of Philosophy in Biology and Biomedical Science (Plant Biology)

Washington University in St. Louis, 2011

Professor Joseph M. Jez, Chairperson

In the plant cell, redox regulation and redox responsiveness are governed by a series of mechanisms that hinge upon the use of small molecule redox-couples and reversible, thioredoxin- or glutaredoxin-mediated protein disulfides. This work examines first the structural basis for synthesis of these small molecules and second how plants are able to adapt and respond to changes in environmental redox state.

Among the major redox-couples, glutathione (GSH) is maintained at the highest cellular concentrations, and is furthermore employed in a protective capacity as an anti-xenobiotic and anti-oxidation protein thiol-modification. Almost all eukaryotes utilize GSH, but some legumes additionally synthesize homoglutathione (hGSH), which is a GSH analog that contains a terminal β-alanine residue instead of a terminal glycine. This alternate reaction is catalyzed by hGSH synthetase, which is related to GSH synthetase; however, the specific features that alter substrate specificity are unknown. To understand the molecular basis for the synthesis of the legume-specific molecule, the three-dimensional structure of hGSH synthetase from *Glycine* m*ax* (soybean) was solved by x-

ray crystallography in three forms - apoenzyme, bound to γ-glutamylcysteine, and with hGSH, ADP, and a sulfate ion bound in the active site. Comparison of these structures with those of GSH synthetase suggest that two residues - a leucine and a proline in the Ala-rich loop region of the enzymes - dictate the use of β-alanine instead of glycine in hGSH synthetase. Site-directed mutagenesis studies and kinetic analysis further support this conclusion.

As a means of regulating activity, many plant proteins limit access to their active sites and control the aggregation of catalytic oligomeric complexes through the formation of redox-reversible disulfide bonds. In order to identify plant proteins and pathways that utilize such bonds and/or thiol modifications to modulate oxidation state, an N-ethylmaleimide- and 5-iodoacetamidofluorescein-based dual-labeling strategy was employed in conjunction with 2D-gel electrophoresis and LC-MS/MS. Initial experiments with root protein extracts from *B. juncea* identified several new proteins that were differentially expressed and/or oxidized in response to exposure to the glutamate-cysteine ligase inhibitor buthionine sulfoxide or H_2O_2. A clear lack of overlap between the proteins altered by each condition was also noted. To assess oxidative changes to the plant thiol-redox proteome under agriculturally relevant conditions, soybean plants were field grown under ambient and elevated tropospheric ozone concentrations. Investigation into changes in protein expression and oxidation state again yielded numerous novel protein responses. Intriguingly, many of the largest changes were observed in pathways involved in core carbon metabolism, a sharp contrast to the changes in redox-centric pathways seen following acute ozone exposure. This observation, in conjunction with a

comparison of protein responses across several different ozone concentrations, led to the conclusion that ozone exposure is governed by a threshold effect: a concentration at which the plants transition from an active redox response toward maintenance of core processes and metabolism.

ACKNOWLEDGEMENTS

For everyone who believed I could.

TABLE OF CONTENTS

ABSTRACT.....ii

ACKNOWLEDGEMENTS.....v

TABLE OF CONTENTS.....vi

LIST OF FIGURES.....viii

LIST OF TABLES.....xi

LIST OF ABBREVIATIONS.....xiii

CHAPTER 1: INTRODUCTION.....1

REFERENCES.....28

CHAPTER 2 PREFACE.....51

REFERENCES.....54

CHAPTER 2: STRUCTURAL BASIS FOR EVOLUTION OF PRODUCT DIVERSITY IN SOYBEAN BIOSYNTHESIS.....55

REFERENCES.....62

CHAPTER 3 PREFACE.....65

REFERENCES.....68

CHAPTER 3: REDOX-REGULATORY MECHANISMS INDUCED BY OXIDATIVE STRESS IN *BRASSICA JUNCEA* ROOTS MONITORED BY 2-DE PROTEOMICS.....69

REFERENCES.....72

CHAPTER 4 PREFACE.....90

CHAPTER 4: FROM CLIMATE CHANGE TO PROTEINS: REDOX PROTEOMICS OF SOYBEAN OZONE RESPONSES.....95

REFERENCES.....116

CHAPTER 5: CONCLUSIONS AND FUTURE WORK.....162

REFERENCES.....191

APPENDIX I.....201

REFERENCES.....215

LIST OF FIGURES

CHAPTER 1

FIGURE 1: The three major redox couples found in plants.....46

FIGURE 2: The tripeptide glutathione.....47

FIGURE 3: Redox regulation of plant GCL.....48

FIGURE 4: GSH and hGSH bioynthesis.....49

FIGURE 5: Key reactions for synthesizing ozone in the stratosphere and troposphere.....50

CHAPTER 2

FIGURE 1: Diversity in GSH biosynthesis.....56

FIGURE 2: Structure of hGS.....58

FIGURE 3: Domain and loop movements of hGS.....58

FIGURE 4: Substrate binding sites in the open and closed forms of hGS.....59

FIGURE 5: Comparison of the Ala-rich loops in hGS and GS.....59

SUPPLEMENTAL FIGURE 1: Protein expression and purification analysis.....64

CHAPTER 3

FIGURE 1: Redox proteome labeling approach.....70

FIGURE 2: Differential protein redox changes.....71

SUPPLEMENTAL FIGURE 1: Differential protein expression changes.....76

SUPPLEMENTAL FIGURE 2: Example of 2D-gel images after IAF-labeling and SYPRO staining.....77

CHAPTER 4

FIGURE 1: Representative 2-DE Gels.....123

FIGURE 2: Venn diagram of proteins that differ in leaf and root tissues under 60 ppb and 115 ppb O_3 treatment compared to ambient conditions.....124

FIGURE 3: Summary of fold changes in total and redox proteomes of leaf tissue exposed to 115 ppb O_3.....125

FIGURE 4: Metabolic Overview of Total and Redox Proteome Changes in Soybean Leaf Tissue Exposed to 115 ppb O_3.....126

SUPPLEMENTAL FIGURE 1: Redox Proteome Labeling Approach.....128

SUPPLEMENTAL FIGURE 2: Additional Total and Redox Proteome Changes.....129

SUPPLEMENTAL FIGURE 3: Immunoblot analysis of RuBisCO large subunit.....130 expression

APPENDIX I

FIGURE 1: Overview of the methionine chain-elongation pathway of aliphatic glucosinolate biosynthesis in *Arabidopsis thaliana*.....221

FIGURE 2: Glucosinolate profiles in seeds and leaves from wild-type, *atipmdh1* mutant, and transgenic plants harboring each *AtIPMDH* driven by the *AtIPMDH* promoter.....222

FIGURE 3: Structure of AtIPMDH2.....223

FIGURE 4: Comparison of the catalytic efficiencies of wild-type and mutant AtIPMDH using 3-isopropylmalate and 3-(2'-methylthio)ethylmalate.....224

FIGURE 5: Glucosinolate profiles in seeds and leaves of wild-type, *atipmdh1* mutant, and transgenic plants expression AtIPMDH1-F137L, AtIPMDH2-L133F and AtIPMDP3-L134F driven by *AtIPMDH1* native promoter.....225

LIST OF TABLES

CHAPTER 2

TABLE 1: Comparison of kinetic parameters for *Arabidopsis* GS and soybean hGS.....56

TABLE 2: Crystallographic statistics.....57

TABLE 3: Substrate specificity of wild-type and mutant soybean hGS.....60

SUPPLEMENTAL TABLE 1: Oligonucleotide primers used for site-directed mutagenesis.....64

CHAPTER 3

TABLE 1: Total number of spots differentially expressed and oxidized ($p<0.05$) in response to H_2O_2 and BSO.....71

SUPPLEMENTAL TABLE 1: List of proteins identified as differentially expressed (from SYPRO stain) or oxidized (from IAF labeling) in response to H_2O_2 or BSO treatments.....78

SUPPLEMENTAL TABLE 2: List of the 52 proteins identified as single protein classified according to their biological process.....84

CHAPTER 4

TABLE 1: Comparison of enzyme activities in leaf tissues exposed to ambient (40 ppb) and high (115 ppb) O_3.....127

SUPPLEMENTAL TABLE 1: Total number of spots detected and identified as either differentially expressed or oxidized across all experimental conditions.....131

SUPPLEMENTAL TABLES 2-5: Data Summary Tables.....132

APPENDIX I

TABLE 1: Kinetic parameters of wild-type and mutant AtIMPDHs.....226

LIST OF ABBREVIATIONS

Abbreviation	Full Name
redox	reduction-oxidation
ROS	reactive oxygen species
NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
ATP	adenosine triphosphate
γEC	γ-glutamylcysteine
GSH	glutathione
GSSG	oxidized glutathione
hGSH	homoglutathione
GCL	glutamate cysteine ligase
GS	glutathione synthetase
hGS	homoglutathione synthetase
SAR	systemic acquired resistance
NEM	N-ethylmaleimide
IAM	iodoacetamide
IAA	iodoacetic acid
DTT	dithiothreitol
TCEP	tris[2-carboxyethyl]phosphine
IAF	5’-iodoacetomidofluorescein
HPLC	high performance liquid chromatography

FPLC	fast protein liquid chromatography
MS	mass spectrometry
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
MALDI	matrix-assisted laser desorption/ionization
BSO	buthionine sulfoximide
NMVOCS	non-methane volatile organic chemicals
ppb	parts per billion
FACE	free air concentration enrichment

CHAPTER 1

INTRODUCTION

In chemistry, the term "oxidation" describes a means by which an ion, atom, or molecule loses total possession of one or more of its electrons due to an increase in its overall oxidation state [IUPAC, 1997]. As a concept, the notion of oxidation arose out of the Phlogiston theory of the 17th century, which sought to explain the alchemical processes behind fire and rust, among other transformations. Following the 18th century work of Antoine Lavoisier, who identified the roles of oxygen and hydrogen in many chemical reactions, the tenets of modern oxidation theory began to gain traction among scientists and thinkers [Lavoisier, 1789]. In current parlance, the concept of oxidation exists in a duality with that of reduction, which (as the exact opposite of oxidation) describes the gain of total electron possession by an ion, atom or molecule. However, following Lavoisier's investigation of oxidation, more than 100 years passed before scientists in the field of electrochemistry drafted the ionic theory of dissociation, which defined the paired, yet opposite, nature of oxidation and reduction reactions [Jensen, 2007].

While redox (reduction-oxidation) reactions are traditionally associated with chemistry, they also perform a number of critical functions in biology, particularly in plants. For example, the successful production of glucose via photosynthesis relies on redox reactions for the oxidation of water to molecular oxygen, the reduction of $NADP^+$ to NADPH, and the reversible reduction and oxidation of the protein ferredoxin [Schurmann and Buchanen, 2008]. As the process complimentary to photosynthesis, cellular respiration in both plants and animals likewise relies on many of the same reactions, only driven in the reverse direction (i.e., the reduction of molecular oxygen to

water). Many other reactions utilized by plants, including those associated with xenobiotic metabolism, light-sensing, and herbivory signaling, also employ redox chemistry.

Although their reversible nature is a boon in terms of adaptability, the widespread adoption of redox reactions by biological systems has resulted in a number of difficulties, both large scale and small. From a macro-scale perspective, most of the earth's habitable spaces exist in an atmosphere composed of approximately 78% N_2, 21% O_2, and 1% argon gas. While both N_2 and argon are largely inert under the prevailing atmospheric conditions, O_2 is more reactive. In the presence of heat and light, O_2 can degrade and recombine into a number of free radical species, among them peroxide, superoxide, and ozone, all of which are effective oxidizing agents [Halliwell, 2006]. In the absence of either preventive or responsive mechanisms, these radicals, or reactive oxygen species (ROS), can rapidly permeate cell membranes, abolish redox reaction equilibrium though large scale generation of oxidized reactants/products, and irreversibly inactivate redox-sensitive proteins. It is estimated that approximately 1-2% of the O_2 that enters a plant will either consist of or be converted to ROS [Bhattachrjee, 2005]. On the micro-scale, many cellular processes also produce local concentrations of hydroxide and peroxide as reaction byproducts or for use in signaling pathways; without available containment and decontamination procedures, affected cells will likewise suffer the effects of widespread oxidative disruption and damage.

Fortunately, cells have evolved a number of different systems by which they control the relative levels of oxidized and reduced species present at any given time;

these processes are collectively referred to as redox homeostasis. In plants, three major molecules - NADPH, ascorbate, and glutathione - along with several secondary molecules, such as NADH, flavins, and quinones, are responsible for maintaining an appropriate redox state (Figure 1). In solution these compounds exist as redox couples with a fluctuating ratio of reduced to oxidized molecules. When local cellular conditions grow too oxidized, these compounds will react with the excess ROS in order to drive the system back towards homeostasis. In order to then restore the compound's redox equilibrium, additional reduced molecules will need to be synthesized or the oxidized molecules will be reduced via a compound-specific regeneration cycle.

Although all of the major redox couples in plants are capable of buffering against oxidative conditions, each compound also fulfills a more specialized redox-related role. NAD(P)H provides essential reducing power for numerous enzymatic reactions including those associated with photosynthesis; among the major redox couples it has the largest (most negative) midpoint reducing potential: -320mV at pH 7.0 [Noctor, 2006]. Ascorbate, as the smallest of the three major redox couples, can traverse the plasma membrane to assist in detoxification of the apoplasm. It also can be stored in an unconjugated form in the vacuole at relatively high concentrations for future use. Finally, glutathione, the only major redox player to contain a thiol group, acts as a sulfur sink and cysteine storage molecule. By virtue of its thiol group, glutathione is also highly effective at maintaining protein thiol-redox state through the modification of disulfide bonds, as discussed below.

PART I - HOMOGLUTATHIONE SYNTHETASE AND GLUTATHIONE BIOSYNTHESIS

Glutathione (GSH) is composed of glutamate, cysteine, and glycine, with a γ-linkage between glutamate and cysteine rendering the peptide immune to degradation by all but one class of protease, the γ-glutamyl transpeptidases (Figure 2). As described above, glutathione can be found in both a reduced (GSH) and an oxidized (GSSG) form, with the reduced form favored over the oxidized form by up to 200-fold [Noctor, 2006; Masip, et. al, 2006]. Due to multiple polar groups, glutathione is exceedingly soluble, and can be found at a foliar concentration of ~1 to 5mM, with local concentrations of 7 to 20mM reported [Mullineaux and Rausch, 2005]. These traits, in combination with a relatively high reduction potential (-240 mV at pH 7.0), make glutathione both effective and highly adaptable as a modulator of redox homeostasis [Rouhier, 2008]. Not only can glutathione spontaneously detoxify reactive oxygen and nitrogen species, but it also can directly protect proteins against irreversible oxidation through glutathionylation of critical residues [Gallogly and Mieyal, 2007]. Additional roles for glutathione include detoxification of peroxides through the ascorbate-glutathione cycle, conversion of toxic aldehydes like formaldehyde and methylglyoxyl to less harmful variants, and sequestration of toxic heavy metals such as cadmium [Potters et al., 2002; Dixon et. al, 1998; Rauser, 1995; Skipsey et. al, 2000].

In the event of protein glutathionylation or the formation of an undesirable disulfide bond, the disulfide-oxidoreductase glutaredoxin is capable, with the assistance

of two GSH molecules, of reducing the residue moiety via either a monothiol or a dithiol mechanism [Rouhier et. al, 2008]. This activity results in the formation of one GSSG molecule, which then must be reduced to maintain the buffering capacity of glutathione. Another oxidoreductase, glutathione reductase, utilizes the greater reducing potential of NADPH to convert GSSG back to GSH, producing $NADP^+$ in the process [Gill and Tuteja, 2010]. Finally, the regeneration cycle is completed when NADP re-enters the chloroplast stroma, and is reduced to NADPH by ferredoxin-$NADP^+$ reductase as part of the photosynthetic Z-scheme.

If additional or replacement reducing power is required by the plant cell, more glutathione may be synthesized in two ATP-dependent steps. In the first reaction, the enzyme glutamate cysteine ligase (GCL; also known as γ-glutamylcysteine synthetase; EC 6.3.2.2) catalyses the formation of γ-glutamylcysteine from glutamate and cysteine. In the second reaction the enzyme glutathione synthetase (GS; EC 6.3.2.3) adds glycine to γ-glutamylcysteine to produce the complete tripeptide. In *Arabidopsis thaliana*, GCL contains a chloroplast-localization sequence and is expressed solely in the plastid, while GS is primarily cytosolic [Meyer and Hell, 2005]. Expression analysis indicates that of the two enzymes, GCL functions as the rate-limiting step and is the target of multiple regulatory controls [Foyer et al., 1995; Arisi et al., 1997].

Because cellular glutathione levels can impact so many different facets of a plant's stress response, GCL activity remains at all times tightly regulated via at least three major mechanisms. 1.) At the substrate level, GCL activity is limited by the availability of glutamate and cysteine, with the latter being the more limiting of the two.

Because cysteine is derived from serine and acetyl-CoA, its use as a substrate for GCL drains from carbon reserves that could be instead spent to generate proteins and/or ATP [Youssefian et al., 2001]. Accordingly, in order to minimize the perturbations of other pathways, only those resources that are absolutely needed are drawn off in order to synthesize glutathione. 2.) At the level of redox control, GCL is regulated by a unique mechanism that was only recently deduced. It has been known for nearly twenty years that mammalian GCL is heterodimer composed of a larger (MW 70,000 kDa), catalytic subunit and a small (MW 30,000 kDa) regulatory subunit. Under reducing conditions and in the absence of the regulatory subunit, catalytic activity is only a fraction of that found when GCL is present as a holoenzyme [Huang et al, 1993a; Huang et al., 1993b, Chen et al., 2005]. This data in conjunction with more recent mutagenesis studies led to the conclusion that formation of a reversible disulfide bond between the catalytic subunit and the regulatory subunit prompts GCL activation [Fraser et al., 2003]. But while the human GCL has a regulatory subunit, no comparable protein subunit has been found to be encoded by a plant genome. Instead, a combination of mutagenic, kinetic, and crystallographic studies have shown that plant GCL is a homodimer whose activation and subsequent dimerization under oxidative conditions is controlled by a pair (one on each monomer) of disulfide bonds [Hicks et al., 2007] (Figure 3). Additionally, access of substrates to the active site is controlled by a second disulfide bond; only under sufficiently oxidative conditions does the β-hairpin flap over the active site entrance swing back to allow uninhibited access [Hothorn et al, 2006]. 3.) Finally, at the level of expression, transcription of the gene encoding GCL has been shown to increase when

plants are subjected to known sources or signals of oxidative stress, including jasmonic acid and the heavy metals copper and cadmium [Xiang and Oliver, 1998]. Although the human GCL promoter has been well characterized, little comparable work has been done in plants [Soltaninassab et al., 2000]. The Arabidopsis GCL promoter does contain several possible G-box elements similar to those found in the jasmonic acid sensing portion of the *Pin2* promoter, but no further studies have been done to verify their role [Xiang and Oliver, 1998]. With additional evidence also suggesting that GCL activity is regulated by light intensity and/or phosphorylation, (the latter being documented for the human GCL variant), it is clear that the enzyme requires still a great deal of further investigation [Ogawa et al., 2004; Sun et al., 1996].

While there remain many unanswered questions with regards to GCL, the second step in the glutathione synthesis pathway presents an entirely different set of unknowns. In most plants, γ-glutamylcysteine is converted to glutathione through the activity of GS. Like GCL, eukaryotic GS is catalytically active as a dimer; however, in *Escherichia coli,* and other prokaryotes, GS functions as a tetramer. To date, three x-ray crystal structures of the GS from *Saccharomyces cerevisiae*, *E. coli*, and *Homo sapiens*, as well as a "loopless" *E. coli* variant structure, have been solved [Yamaguchi et. al, 1993; Kato et. al, 1994; Polekhina et. al, 1999; Gogos and Shapiro, 2002]. These structures indicate that GS falls within the large ATP-grasp superfamily of protein structures [Galperin and Koonin, 1997]. Members of this protein family are defined by the presence of two sets of two anti-parallel β-sheets connected by a series of loops. This motif, known as an ATP-grasp or palmate-grasp for its provision of an ATP-binding pocket, comprises the

majority of the active site in proteins of this family, and also provides the family with its name. In addition to the shared motif, members of the ATP-grasp superfamily also rely on a similar mode of action: that is, the ATP-dependent ligation of a carboxyl group carbon from one reactant with the amino (or imino) group nitrogen of a second reactant via the formation of an acyl-phosphate intermediate [Meister and Anderson, 1983; Ogita and Knowles, 1988; Meister 1989; Fan et. al, 1995]. For the GS-catalyzed reaction, γ-glutamylcysteine provides the carboxyl group, while glycine provides the amino group. With regard to enzyme kinetics, the reaction likely proceeds via a random terreactant mechanism, with slight preference given to the order of substrate binding [Jez and Cahoon, 2004]. As an interesting side note, although members of the ATP-grasp superfamily are structurally and mechanistically similar, their amino acid sequences are quite divergent and range only from 10%-20% identical across the family [Galperin and Koonin, 1997]. Other examples of ATP-grasp family members include D-Ala:D-Ala ligase, biotin carboxylase, and carbamoyl phosphate synthase [Fan et. al, 1994; Artymiuk et. al, 1996, Thoden et. al, 1997].

Among the branches of life, synthesis of glutathione as a storage agent and redox buffer is extremely well conserved. Virtually all eukaryotes produce glutathione, as do many bacteria. Only among the archaea is glutathione synthesis sharply limited, with just halobacteria known to produce it. Despite its time-tested effectiveness, some eukaryotic parasites (such as those of the genus *Trypanosoma*), do not produce glutathione at all, while many plant species produce it in addition to one or more glutathione homologs [Muller et al., 2003]. For example, in many legumous plants homoglutathione – in which

the terminal glycine has been replaced by β-alanine, is found with or in place of glutathione in a tissue-dependent fashion [Moran et al, 2000]. In members of the Poaceae, except maize, a tripeptide called hydroxymethylglutathione, which contains a terminal serine instead of glycine, is produced [Klapheck et al, 1994]. In maize, yet another glutathione-like peptide is synthesized, this one with a terminal glutamate [Meuwly and Rauser, 1992]. Of these alternate forms of glutathione, the most is known about homoglutathione. Except for broadbean and lupine, every legume that has been thus far investigated produces homoglutathione in addition to glutathione [Moran et al., 2000]. In some tissues, such as alfalfa leaves, homoglutathione completely replaces glutathione as the dominant thiol compound, while in others they may be found more or less one-for-one. By contrast, in cowpea leaves glutathione is the dominant tripeptide and homoglutathione is almost completely absent [Matamoros et al., 1999]. In soybean, both homoglutathione and glutathione are present, with leaves and seeds containing 50- to 200-fold and 135-fold more homoglutathione than glutathione, respectively. [Klapheck, 1988; Matamoros et al., 1999].

For synthesis of both glutathione and homoglutathione, the first reaction - synthesis of γ-glutamylcysteine - is shared. However, while glutathione synthesis requires the activity of GS to add a terminal glycine to the tripeptide, homoglutathione synthesis relies on the homologous enzyme homoglutathione synthetase (hGS) for introduction of β-alanine (Figure 4). Because the genomes of numerous legumes, soybean included, show evidence for two rounds of genome duplication, it has been proposed that hGS arose from GS by divergent evolution after the first duplication event

[Shoemaker et al.,2006; Van et al., 2008; Gill et al., 2009, Frendo et. al, 2001]. In the soybean genome, as an example, there are two copies each of the genes encoding GS and hGS, with each pair sharing 87% and 93% sequence identity, respectively. The relative expression patterns of the various copies remain unknown.

Between legumous GS and hGS genes, the sequence identity is ~60-70% depending on the species assessed [Frendo et. al, 2001]. Given such a high degree of identity, it is curious that hGS has managed to evolve a unique, if parallel, function to GS. From a redox perspective, the benefit of synthesizing homoglutatione instead of, or in addition to, glutathione is unclear. Both molecules preserve the redox-reactive cysteine residue that allows conversion from GSH to GSSG; in fact the only obvious difference is that homoglutathione is effectively one carbon bond-length longer than glutathione courtesy of β-alanine. But while the difference in substrate size does not shed light on a defined role for homoglutathione, it does hint at structural differences that have arisen between hGS and GS. Based on the available GS structures, γ-glutamylcysteine occupies a pocket at one end of the cleft formed by the ATP-grasp motif, while ATP and glycine occupy the opposing end [Yamaguchi et. al, 1993; Polekhina et. al, 1999; Gogos and Shapiro, 2002]. The carboxyl tail of glycine contacts two Ala residues (Ala 462 and Ala 463 in the human GS), that are part of a larger alanine-rich loop domain. Based on a comparison of the yeast apoenzyme and ATP/γ-glutamylcysteine-containing GS structures, it is evident that this domain is capable of movement during the overall reaction cycle [Gogos and Shapiro, 2002]. However, kinetic studies indicate that AtGS will not accept β-alanine as a substrate, which thus implies that any domain movement

that occurs is simply not enough to accommodate the longer β-alanine molecule in the glycine binding pocket [Galant et. al, 2009, Jez and Cahoon, 2004].

In addition to the previously mentioned alanine-rich loop, a second domain of the GS enzyme appears mobile based upon the yeast crystal structure [Gogos and Shapiro, 2002]. This domain, termed the "lid domain" according to the nomeclature associated with the human GS structure, is composed of residues 355-417 of the yeast enzyme and makes up one corner of the roughly triangular-shaped overall GS structure. Within the lid domain is a subdomain, known as the glycine-rich loop, that appears to make extensive contacts with the bound ATP moiety. Based on the apoenzyme and reactant-bound yeast GS structures, it appears that this domain's purpose is to swing inward and lock ATP in place following its binding within the ATP-grasp cleft. Investigation into the kinetic mechanism of AtGS has indicated that the GS enzymes favor a semi-specific binding order for their three substrates: namely, either ATP or γ-glutamylcysteine first, and glycine (or an appropriate homolog) last [Jez and Cahoon, 2004]. Accordingly, the yeast structure bound with γ-glutamylcysteine and an ATP homolog represents the third stage in the reaction mechanism, with the apoenzyme and enzyme bound with either ATP or γ-glutamylcysteine representing the first and second stages, respectively. Among the yeast, *E. coli*, and human GS structures, no structure for the γ-glutamylcysteine-bound variant of the second stage exists; thus it is unclear if any movement of the lid domain is prompted by the binding of γ-glutamylcysteine alone. Furthermore, while cumulatively the available crystal structures provide a visual representation of the reaction mechanism, snapshots of more than two stages are not available for a single enzyme, leading to

difficulties in the comparison of domains of different sizes and of different numbering schemes across the various GS structures.

In order to fill in the gap in the GS-type reaction mechanism as well as provide a series of snapshots from a single enzyme, I solved three crystal structures of the homoglutathione synthetase from *Glycine max*. The three structures are of the apoenzyme, the enzyme with γ-glutamylcysteine bound, and the enzyme with ADP and homoglutathione bound, and represent respectively the first, second, and fifth stages of the reaction mechanism. This series of structures, along with their accompanying mutagenesis and kinetics data, identify the structural elements that are responsible for the differences in substrate specificity between GS and hGS. For more information, please refer to Chaper 2.

PART II - THE THIOL-REDOX PROTEOME - BUTHIONINE SULFOXIMINE (BSO), H_2O_2, AND *BRASSICA JUNCEA* (INDIAN MUSTARD)

The example of GCL from Part 1 illustrates that in some cases, the formation of disulfide bonds is desirable as a means of controlling enzyme activity. In plants, this type of redox regulation is relatively rare, not because few proteins utilize it, but rather because only a handful of redox-sensing candidates have been appropriately characterized. The majority of known redox-regulated proteins are controlled by one of a half-dozen or so different potential-sensing pairs, among them thioredoxin/ferredoxin, glutathione/glutaredoxin, and NADP/thioredoxin [Buchanan and Balmer, 2005]. The

protein components of these redox pairs - namely ferredoxin, thioredoxin, and glutaredoxin - all utilize a reduction mechanism that necessitates direct contact between them and the target protein that they are reducing. Accordingly, screens to identify potential targets of these regulatory proteins have typically been able to employ affinity chromatography followed by mass spectrometry or N-terminal sequencing; this method only pulls out proteins that strongly interact with the "bait", with no regard for the specific means of redox regulation (disulfide bond, glutathionylation, etc.) utilized by the individual protein [Balmer et al., 2003; Motohashi et al., 2001; Yano et al., 2001]. Currently, the total number of identified plant redox-regulated proteins identified via affinity chromatrography stands at several hundred, with the largest subsets shown to interact with thioredoxin and glutaredoxin, respectively [Buchanan and Balmer, 2005, Hisabori et al., 2005, Rouhier et al., 2005, Wormuth et al., 2007]. To date, very few proteins known to use disulfide bonds to regulate their activity have been identified via this methodology. Besides GCL, one example is NPR1, which was originally studied because of its essential role in plant systemic acquired resistance (SAR). Under normal conditions, NPR1 exists as an inactive oligomer; however, reduction of intermolecular disulfide bonds between the subunits following SAR initiation allows the protein to achieve its active monomeric form [Mou et al., 2003]. Another example is OxyR, a peroxide-sensing transcription factor from *E. coli* and *Salmonella typhimurium* that is activated through the formation of an intramolecular disulfide bond and deactivated via the activity of glutaredoxin [Zheng et al., 1998, Christman et al., 1989].

While affinity-based methodologies are well established for isolating certain subsets of the redox-regulated proteome, other more inclusive methodologies exist as well. One alternative is a two-dimensional (2D)-gel electrophoresis approach that allows all proteins within a given sample to be separated and fixed. Most 2D-gels rely on isoelectric point for separation in the first dimension, and denatured molecular weight in the second dimension, though native molecular weight may also be used [O'Farrell, 1975]. In order to utilize 2D separation for the detection of redox-sensitive proteins, the protein mixture under scrutiny must first be treated so as to distinguish redox-labile proteins from the remainder of the proteins in the sample. In a complex protein sample, many proteins will have solvent-exposed cysteine residues. While some of these residues may form disulfide bonds or contain thiol modifications such as glutathionylation, others will be present as free thiol groups that will not and will never be modified naturally. Because this latter class of cysteine residues is not redox-labile, it is important that they be chemically blocked to prevent reactions during the subsequent detection of modified and disulfide-bound cysteines. Thus, a soluble protein extract must first be treated with a thiol alkylating agent such as *N*-ethylmaleimide (NEM), iodoacetamide (IAM) or iodoacetic acid (IAA). Although all of these compounds form an adduct that is almost impossible to reverse, NEM is perhaps the best choice based on faster reaction speed and activity under a wider pH range relative to the other two compounds [Rogers et al., 2006].

Once the non-labile thiols in a complex protein mixture have been blocked, the next step is to detect the redox-reactive cysteine residues. Because these thiols are either

involved in disulfide bonds or are blocked by secondary modifications, the protein mixture must first be treated with a reducing agent such as dithiothreitol (DTT) or tris[2-carboxyethyl]phosphine (TCEP) so as to reduce the thiol groups. Because the previous reaction of free thiols with NEM was done via alkylation as opposed to thiol-disulfide exchange, reducing agents will have no effect on those adducts. With reduction of the previously oxidized thiols complete, they become available for chemical modification via one of several fluorescent compounds. The choice of which fluorescent compound to use depends upon the emission wavelength required, the pH of the protein mixture to be labeled, and the overall experimental design. Popular thiol-reactive dyes include those of the bromobimane family (most commonly monobromobimane), iodoacetomide-fluorescein conjugates (typically 5'-iodoacetomidofluorescein - IAF), and the cyanine dyes (Cy2, Cy3, and Cy5) [Timms, 2005; Fahey and Newton, 1987; Baty et al., 2002; Chen et al., 2008].

When labeling of the complex protein mixture is complete, the protein sample may be loaded onto a 2D gel and separated as previously described. Because the protein spots will not be visible to the naked eye, the resulting gel(s) must be imaged at the appropriate wavelength for whichever fluorescent dye was selected in order to see the spots that contain redox-reactive proteins. For comparison of the relative numbers of redox-reactive proteins to total proteins across a given pI/mW range, the gels may then be further stained with a total protein dye and imaged again. There are many sensitive dyes, including SYPRO Ruby, SYPRO Tangerine, and FOCUS FASTsilver, available; however, care must be taken to ensure that the dye chosen does not bind covalently (as

this may interfere with downstream mass spectrometry applications) and that its excitation and emission wavelengths do not overlap with those of the chosen thiol-reactive dye.

Depending on the complexity of the protein sample under scrutiny, several different methods for identifying individual spots from the 2D gels may be available. If a protein is particularly well characterized, whole protein extraction and verification of retention time via HPLC may be all that is required. In most cases however, the protein identification will not be so easy, and a mass spectrometry-based approach will be necessary. For this method, the protein will need to be digested with a predictable protease such as trypsin, and the resulting peptides separated via reverse-phase liquid chromatography. Those peptides will then further fragmented by MS/MS, and the resulting spectra will be identified via comparison against a known database (NCBI, SWISSPROT, etc.) of protein sequences.

As stated previously, the 2D gel-based approach, although more time consuming, presents an advantage over related affinity chromatography methodologies because it allows for the detection of proteins that do not react strongly with a “doxin” and/or contain one or more redox-labile disulfide bonds. In plants as well as other eukaryotes, a number of efforts have already been undertaken in order to identify novel redox-reactive proteins using 2D-SDS-PAGE. In Alvarez et al. [2009a], the authors dissected shoot tissue from *Arabidopsis thaliana* seedlings, and labelled with iodoacetamide and mBBr before separating the protein in two dimensions. Fifty resulting proteins were then identified via nano-LC-MS/MS as redox-reactive; five proteins were further identified as

new members of the thiol-redox proteome. In an effort to tease out the relationship between redox-regulation and dormancy control, another group isolated protein from hormone-treated wheat seeds, labeled with mBBr and iodoacetamide (the opposite order from above), and performed nano-LC-MS/MS and MALDI MS after 2D-gel separation. Their experiments resulted in the identification of 79 unique redox-modifiable proteins with possible roles in seed dormancy [Bykova et al., 2011]. Across the literature, there exist other example of studies seeking to identify redox-labile plant proteins under a given set of conditions [Zhou et. al, 2011; Tanou et al, 2010; Alvarez et al., 2009b]; Maeda et al., 2005; Rinalducci et al., 2008].

Importantly, compared to the sheer volume of redox-related studies in animal or microbial systems, the investigation of redox-reactive proteins in plants is still in its infancy. The majority of studies in plants to date have been furthermore very narrow in focus, concentrating solely on the effect of a specific compound or growth condition on the relative abundance or redox state of proteins. While such investigations are necessary to elucidate the mode of action of the relevant plant-response networks, they do not address what is happening between the stimuli and the protein that is changing. Namely, how do specific proteins respond to the ROS that serve as antagonizers and/or signaling molecules in response to a set change in conditions?

To begin to address this question, I have undertaken a series of experiments designed to address how different sources of ROS differentially antagonize a redox-responsive system. Using *Brassica juncea* (Indian mustard), I treated the roots of potted plants with buthionine sulfoximine (BSO), hydrogen peroxide, or water (as a control).

Because BSO, as an inhibitor of GCL, is an indirect source of endogenous ROS [Griffith and Meister, 1979], and hydrogen peroxide is a direct source of exogenous ROS, I anticipated that they would affect the expression and redox state of different, though possibly overlapping, sets of proteins. Furthermore, because only a relatively small body of work on redox-responsive proteins has originated in field of plant biology, I hoped that this series of experiments would add to the available body of knowledge, particularly if they allowed for the identification of new targets of redox-regulation. A more detailed description of the methodology and results of these preliminary 2D-gel experiments is presented in Chapter 3.

PART III - THE THIOL-REDOX PROTEOME - OZONE AND *GLYCINE MAX* (SOYBEAN)

The use of *Arabidopsis thaliana* as a model system began as as early as 1907, when Strasburger and his student Laibach suggested its value for studying chromosomes. In the mid-forties, the development of *Arabidopsis* as a platform for mutagenesis began, and by the 1970s the plant and its close relatives had been widely adopted by biology labs around the world [www.arabidopsis.org (TAIR); Redei, 1975]. Arabidopsis is an excellent model system for a number of reasons, including is fast growth rate, genome plasticity, and extensive family tree. While some members of the Brassicales are widely cultivated, Arabidopsis itself has no agricultural significance. And while ostensibly the plant and its cousins occur naturally throughout much of Europe and Asia, the modern

native habitat of Arabidopsis is very much a petri dish in a lab. Thus, when it comes time to study how a plant responds to and interacts with real world oxidative conditions and stimuli, Arabidopsis and its brethen are not necessarily the best choice for further investigation. But in order to select an appropriate plant, one must first understand the breadth and scale of the challenge that oxidative damage causes in a more natural environment.

As described briefly at the beginning of the introduction, one of the largest sources of naturally occurring ROS stems from the conversion of O_2 to various oxide radicals. While the availability of O_2 does not change significantly, the relative concentration of another ROS-producing molecule - ozone - is increasing. In its simplest state, ozone is composed of three charge-stable oxygen atoms; the chemistry that gives rise to it, however, is somewhat complicated and also depends upon where the ozone is being produced. In the stratosphere, ultraviolet energy in the form of a photon can split O_2 to yield monoatomic oxygen. Monoatomic oxygen is highly unstable, so it rapidly recombines with O_2 to yield O_3. In the troposphere, ozone production begins when carbon monoxide reacts with hydroxide, yielding a proton and carbon dioxide. The proton then further reacts with O_2 to produce the peroxy radical HO_2. HO_2 is also unstable, and will react quickly with any number of non-methane volatile organic chemicals (NMVOCs); their products will then react further with ultraviolet energy to produce monoatomic oxygen and subsequently ozone as previously described [Tang et al., 2011; Renaut et al., 2008] (Figure 5). Unlike stratospheric ozone production, the production of ozone in the troposphere is dependent upon the availability of NMVOCS;

these compounds may include nitrogen oxides, sulfur oxides, terpenes, and assorted aqueous solvents. Many NMVOCS are released into the atmosphere as part of the waste streams from various industrial processes; others are byproducts of the combustion of gasoline and diesel in vehicle engines. Unfortunately, as both global averaged industrial output and vehicle ownership, largely as a result of economic development in China and Southeast Asia, are rising and predicted to continue doing so, the available global tropospheric concentrations of NMVOCS, and in turn ozone, are likely to increase as well [Monks et al., 2009; van Aardenne et al., 2001; Fu et al., 2007; Meagher et al., 1998].

Historical data indicates that prior to the industrial revolution, tropospheric ozone concentrations in the northern hemisphere were quite low, averaging only 11 ppb (parts per billion) with deviation of 5 ppb depending on the season. Even as the pace of industrialization increased during the first half of the 19th century, ozone concentrations in the northern hemisphere remained modest at 15 ppb [Volz and Kley, 1988]. However, during the period between 1950 and 1980, ozone concentrations began to trend upward by approximately 0.35 ppb per year, and by the mid-1980s were increasing by up to 0.5 ppb per year [Tang et al., 2011; Fuhrer, 2009; Cooper et al., 2010; Hudman et al., 2008]. Perhaps by virtue of a balance in decreasing and increasing emissions between developed and developing countries respectively, the present day rate of increase is holding steady at approximately 1-2% of the ambient concentration per year [Morgan et al., 2006; Chameides et al., 1994]. Currently, the average annual ambient tropospheric ozone concentration ranges from 20 to 45 ppb over the mid-latitudes of the Northern

Hemisphere [Vingarzan, 2004; Booker et al., 2009]. As ozone synthesis is dependent upon available energy levels, ozone concentrations tend to follow a cyclical cycle though out the year. Thus, during the summer months (June-August depending upon latitude), local ozone concentrations may peak at an average of 60-80 ppb before tapering off again in the fall [Fowler et al., 1999; Mauzerall et al., 2000]. Furthermore, due to the earth's natural light-dark cycle, ozone concentrations are also diurnally cyclical, with the highest concentrations coinciding with the brightest/hottest parts of the day [Fuhrer et al., 1997].

Because both historical trends and present-day atmospheric profiling data support a situation in which tropospheric ozone concentrations will continue to rise over the course of the next 50-100 years, many different environmental models have been put forth to help predict likely ozone concentrations and the areas that might be most affected. One model, taking into account biomass emissions and emissions legislation, predicts that India and southeast Asia, including southern China, will see large increases (9-11 ppb) in surface ozone concentrations between now and 2030 while concentrations over North America remain steady [van Dingenen et al., 2009]. Another prediction, which averages the results of 10 different modeled scenarios, indicates that the Middle East, India, and China will see summertime concentrations rise by 45-55 ppb, and that the southern/eastern United States and Mexico will see increases of 25-35 ppb by 2100 [Prather et al., 2003]. The results of many other modeled scenarios have been published, and while they tend to utilize different sets of baseline data and differing predictive criteria, the world regions that are highlighted as being under threat from rising tropospheric ozone concentrations - namely China, India, and the eastern United States -

remain largely consistent [Murazaki and Hess, 2006; Liao et al., 2006; Bell et al., 2007; Ebi and McGregor, 2008; Racherla and Adams, 2006; Nolte et al., 2008]. This is problematic for several reasons. From a human health perspective, high tropospheric ozone concentrations are dangerous not only because they can lead to tissue oxidation and irreversible damage, but also because ozone is one of the primary components of smog. Because ozone is denser than air, it can trap pollutants close to the earth's surface, leading to a variety of respiratory problems. Local geography and population density can further exacerbate the problem; images of a smog-filled Los Angeles, which sits in a natural depression, and Beijing, which is home to more than 12 million people, have been etched into the public consciousness in recent years. Since China and India together currently account for roughly 45% of the world's population (a percentage that continues to rise every year), increasing ozone concentrations in those regions could be particularly catastrophic for the health of large numbers of people in the future [CIA World Factbook, https://www.cia.gov/library/publications/the-world-factbook].

Another reason why the predicted increases in ozone concentrations over China, India, and the United States are problematic stems from the ability of these countries to produce large volumes of crops for home use and export. While crop production of course partially correlates with country size (the United States, China, and India are ranked 3rd, 4th, and 7th respectively by total land area), these three countries are nonetheless consistently the top producers of many different grains and legumes including rice, maize, millet, soybeans, and wheat [FAOSTAT]. For these crops and many others, ozone exposure, much as it does for humans, can cause widespread

oxidative damage to essential tissues, which in turn negatively affects crop yield. From a food security perspective, it is easy to predict how widespread devastation of harvest yield as a result of oxidative damage (coupled of course to ongoing disasters such as droughts and floods) could rapidly lead to food shortages and famine in highly populated countries without implementation of preventive measures.

How might a country, and the world at large preemptively avoid widespread crop loss through oxidative damage? The most obvious option: implementation of widespread limitations to further industrial emissions, is both politically difficult and slow to yield fruitful results. Many crops, particularly those grown in regions where summertime ozone averages top 60 ppb, are currently suffering from oxidative damage and producing reduced yields. The second option: the utilization of ozone-resistant crop strains, is arguably much cheaper and politically favorable; however, there is a problem. Because high tropospheric ozone concentrations are a relatively new phenomenon there are very few cultivars, natural or engineered, which demonstrate ozone-tolerance or resistance. While a few crop species - namely plums and strawberries - seem to be partially resistant to the effects of ozone exposure, most of the major grain and legume crops show moderate to severe sensitivity [Mills et al., 2007]. Among those major crops, soybean is the most sensitive to ozone. Across all assayed soybean cultivars, the relationship between seed yield and seasonal daytime ozone concentration is largely linear, with ozone concentration in ppb inversely proportional to seed yield in kg/hectare [Betzelberger et al., 2010]. At concentrations as low as 40 ppb, soybean growth and seed yield begins to decrease; increasing the atmospheric concentration to 70 ppb accordingly

results in yield losses of 11-36% [Morgan et al., 2006; Emberson et al., 2009; Heck et al; 1983; Heagle et al., 1998].

While numerous transcriptomic, metabolomic, and proteomic studies have been undertaken in both soybeans and other crop species to elucidate the mechanism by which ozone exposure negatively impacts yield, a clear answer has yet to emerge. One possibility is that irreversible oxidation of key proteins, particularly those involved in photosynthesis, forces the plant to shift resources destined for starch storage and/or cell division toward supplemental amino acid and protein synthesis. Another possibility is that affected plants utilize their resources to upregulate ROS scavenging pathways with the hope of maintaining the status quo. In either case, one would expect that exposure to similar ozone concentrations would upregulate similar response pathways in the various affected crop species. However, there has been very little consensus across crops as to the transcripts and proteins that are differentially expressed following ozone exposure [Ahsan et al., 2010; Agrawal et al., 2002; Feng et al., 2002; Bohler et al., 2007; Bagard et al., 2008; Cho et al., 2008; Sarkar et al., 2010; Torres et al., 2007; Tosti et al., 2006; Gadjev et al., 2006]. This suggests that either the response mechanisms vary significantly between different crops (not an unexpected conclusion given the evolutionary distance between monocots, dicots, etc.), or that plants are very sensitive to variations in the sets of exposure conditions utilized across the various experiments. The limited availability of evidence to support either hypothesis indicates that, in order to shed further light on ozone response pathways and jumpstart the development of ozone-resistant crop cultivars, further experimentation is necessary.

In order to identify novel proteins which are differentially regulated and/or differentially expressed in response to elevated tropospheric ozone concentrations, I have undertaken a series of proteomics-based experiments based upon the methodology developed and described in Chapter 3. Instead of utilizing *Arabidopsis* or *Brassica juncea* (as described previously) as my system of inquiry, I opted to conduct my investigation with soybeans, for several reasons. First, although *Brassica juncea* is a minor crop plant, neither it nor *Arabidopsi*s are widely cultivated, meaning - in terms of sheer scale - neither will greatly contribute to food insecurity due to increasing ozone. Soybeans, on the other hand, are one of the most widely bred crop species, particularly in regions of the world most at-risk from rising ozone concentrations. In the United States, the majority of soybean cultivation occurs in the upper Midwest - specifically in Illinois, Indiana, and Iowa, as well as in Nebraska - which is within the region where ozone concentrations are predicted to increase the most in the coming decades (USDA-NASS; Fishman et al., 2010]. In addition to its localization and practical utility, the physiological response of soybeans to chronic ozone exposure has already been well characterized. For Arabidopsis, the majority of experiments have utilized acute exposure regiments which, while damaging, do not impart the same long-term effects as naturally-occuring chronic exposure cycles [Chen et al., 2009]. In order to maximize the value of physiological data imparted from chronic exposure experiments with soybeans, it ideally should be paired with more in-depth analysis of protein oxidative responses. Accordingly, in Chapter 4, I present the results of redox proteomics experiments

comparing protein expression and oxidiation profiles in soybean tissue grown in the field under ambient and elevated chronic ozone concentrations.

References

Agrawal, G.K., Rakwal, R., Yonekura, M., Kubo, A., Saji, H. (2002). Proteome analysis of differentially displayed proteins as a tool for investigating ozone stress in rice (*Oryza sativa* L.) seedlings. Proteomics 2: 947-959.

Ahson, N., Nanjo, Y., Sawada, H., Kohno, Y., Komatsu, S. (2010). Ozone stress-induced proteomic changes in leaf total soluble and chloroplast proteins of soybean reveal that carbon allocation is involved in adaptation in the early developmental stage. Proteomics 10: 2605-2619.

Alvarez, S., Wilson, G.H., Chen, S. (2009a). Determination of *in vivo* disulfide-bonded proteins in *Arabidopsis*. Journal of Chromatography B 877: 101-104.

Alvarez, S., Zhu, M., Chen, S. (2009b). Proteomic of Arabidopsis redox proteins in response to methyl jasmonate. Journal of Proteomics 73: 30-40.

Arisi, A.-C.M., Noctor, G., Foyer, C., Jouanin, L. (1997). Modification of thiol contents in poplars (*Populus tremula X P. alba*) overexpressing enzymes involved in glutathione synthesis. Planta 203: 362-372.

Artymiuk, P.J., Poirrette, A.R., Rice, D.W., Willet, P. (1996). Biotin carboxylase comes into the fold. Natural Structural Biology 2: 128-132.

Bagard, M., Le Thiec, D., Delacote, E., Hasenfratz-Suder, M.-P., Banvoy, J., Gerard, J., Dizengremel, P., Jolivet, Y. (2008). Ozone-induced changes in photosynthesis and photorespiration of hybrid poplar in relation to the developmental stage of the leaves. Physiologia Plantarum 134: 559-574.

Balmer, Y., Koller, A., del Val, G., Manieri, W., Schürmann, P., Buchanan, B.B. (2003). Proteomics gives insight into the regulatory function of chloroplast thioredoxins. PNAS 100: 370-375.

Baty, J.W., Hampton, M.B., Winterbourn, C.C. (2002). Detection of oxidant sensitive thiol proteins by fluorescence labeling and two-dimensional electrophoresis. Proteomics 2: 1261-1266.

Bell M.L, Goldberg R., Hogrefe C., Kinney P.L., Knowlton K., Lynn B., Rosenthal J., Rosenzweig C., Patz J.. (2007). Climate change, ambient ozone, and health in 50 US cities. Climate Change 82: 61–76.

Betzelberger, A.M., Gillespie, K.M., McGrath, J.M. Koester, R.P., Nelson, R.L., Ainsworth, E.A. (2010). Effects of chronic elevated ozone concentration on antioxidant capacity, photosynthesis and seed yield of 10 soybean cultivars. Plant, Cell and Environment 33: 1569-1581.

Bhattachrjee, S. (2005). Reactive oxygen species and oxidative burst: roles in stress, senescence and signal transduction in plants. Current Science 89: 1113–1121.

Bohler, S., Bagard, M., Oufir, M., Planchon, S., Hoffman, L., Jolivet, Y., Hausman, J.-F., Dizengremel, P., Renaut, J. (2007). A DIGE analysis of developing poplar leaves subjected to ozone reveals major changes in carbon metabolism. Proteomics 7: 1584-1599.

Booker, F., Muntifering, R., McGrath, M., Burkey, K., Decoteau, D., Fiscus, E., Manning, W., Krupa, S., Chappelka, A., Grantz, D. (2009). The ozone component of global change: potential effects on agricultural and horticultural plant yield,

product quality and interactions with invasive species. Journal of Integrative Plant Biology 51: 337-351.

Buchanan, B.B., Balmer, Y. (2005). Redox regulation: a broadening horizon. Annual Review of Plant Biology 56: 187-220.

Bykova, N.V., Hoehn, B., Rampitsch, C., Banks, T., Stebbing, J.-A., Fan, T., Knox, R. (2011). Redox-sensitive proteome and antioxidant strategies in wheat seed dormancy. Proteomics 11: 865-882.

Chameides W.L., Kasibhatla P.S., Yienger J., Levy H.L. (1994). Growth of continental-scale metro-agro-plexes, regional ozone pollution, and world food production. Science 264: 74–77.

Chen, C.P., Frank, T.D., Long, S. (2009). Is a short, sharp shock equivalent to long-term punishment? Contrasting the spatial pattern of acute and chronic ozone damage to soybean leaves via chlorophyll fluorescence imaging. Plant, Cell and Environment 32: 327-335.

Chen, S.-H., Hsu, J.-L., Lin, F.-S. (2008). Fluorescein as a versatile tag for enhanced selectivity in analyzing cystine-containing proteins/peptides using mass spectrometry. Analytical Chemistry 80: 5251-5259.

Chen, Y., Shertzer, H.G., Schneider, S.N., Nebert, D.W., Dalton, T.P. (2005). Glutamate cysteine ligase catalysis. Journal of Biological Chemistry 280: 33766-33774.

Cho, K., Shibato, J., Agrawal, G.K, Jung, Y.-H., Kubo, A., Jwa, N-S., Tamogami, S., Satoh, K., Kikuchi, S., Higashi, T., Kimura, S., Saii, H., Tanaka, Y., Iwahashi, H., Masuo, Y., Rakwal, R. (2008). Integrated transcriptomics, proteomics, and

metabolomics analyses to survey ozone responses in the leaves of rice seedlings. Journal of Proteome Research 7: 2980-2998.

Christman, M.F., Storz, G., Ames, B.N. (1989). OxyR, a positive regulator of hydrogen peroxide-inducible genes in Escherichia coli and Salmonella typhimurium, is homologous to a family of bacterial regulatory proteins. PNAS 86: 3484-3488.

Cooper, O.R., Parrish, D.D., Stohl, A., Trainer, M., Nedelee, P., Thouret, V., Cammas, J.P., Oltmans, S.J., Johnson, B.J., Tarasick, D., Leblanc, T., McDermid, I.S., Jaffe, D., Gao, R., Stith, J., Ryerson, T., Aiken, K., Campos, T., Weinheimer, A., Avery, M.A. (2010). Increasing springtime ozone mixing ratios in the free troposphere over western North America. Nature 463: 344-348.

Dixon, D.P., Cummins, I., Cole, D.J., Edwards, R. (1998). Glutathione-mediated detoxification systems in plants. Current Opinion in Plant Biology 1: 258-266.

Ebi, K., McGregor, G. (2008). Climate change, tropospheric ozone and particulate matter, and health impacts. Environmental Health Perspectives 116: 1449-1455.

Emberson, L.D., Buker, P., Ashmore, M.R., Mills, G., Jackson, L.S., Agrawal, M., Atikuzzman, M.D., Cinderby, S., Engardt, M., Jamir, C., Kobayashi, K., Oanh, N.T.K., Quadir, Q.F., Wahid, A. (2009). A comparison of North America and Asian exposure-response data for ozone effects on crop yields. Atmospheric Environment 43: 1945-1953.

Fahey, R.C., Newton, G.L. (1987). Determination of low-molecular-weight thiols using monobromobimane fluorescent labeling and high-performance liquid chromatography. Methods in Enzymology 143: 85-96.

Fan, C., Moews, P.C., Walsh, C.T., Knox, J.R. (1994). Vancomycin resistance: structure of D-alanine:D-alanine ligase at 2.3 A resolution. Science 266: 439-43.

Fan, C., Moews, P.C., Shi, Y., Walsh, C.T., Knox, J.R. (1995). A common fold for peptide synthetases cleaving ATP to ADP: Glutathione synthetase and D-alanine:D-alanine ligase of *Escherichia coli.* PNAS 92: 1172-1176.

Feng, Y.W., Komatsu, S., Furukawa, T., Koshiba, T., Kohno, Y. (2008). Proteome analysis of proteins responsive to ambient and elevated ozone in rice seedlings. Agriculture, Ecosystems and Environment 125: 255-265.

Fishman, J., Creilson, J.K., Parker, P.A., Ainsworth, E.A., Vinning, G.G., Szarka, J., Booker, F.L., Xu, X. (2010). An investigation of widespread ozone damage to the soybean crop in the upper Midwest determined from ground-based and satellite measurements. Atmospheric Environment 44: 2248-2256.

Fowler, D., Cape, J.N., Coyle, M., Flechard, C., Kuylenstierna, J., Hicks, K., Derwent, D., Johnson, C., Stevenson, D. (1999). The Global Exposure of Forests to Air Pollutants. Water, Air, and Soil Pollution 116: 5-32.

Foyer, C.H., Souriau N., Perret, S., Lelandais, M., Kunert, K.J., Pruvost, C., Jouanin, L. (1995). Overexpression of glutathione reductase but not glutathione synthetase leads to increases in antioxidant capacity and resistance to photoinhibition in poplar trees. Plant Physiology 109: 1047-57.

Fraser, J.A., Kansagra, P., Kotecki, C., Saunders, R.D.C., McLellan, L.I. (2003). The modifier subunit of drosophila glutamate-cysteine ligase regulates catalytic activity by covalent and noncovalent interactions and influences glutathione homeostasis *in vivo*. Journal of Biological Chemistry 278: 46369-46377.

Frendo, P., Jimenez, M.J., Mathieu, C., Duret, L., Gallesi, D., Van de Sype, G., Herouart, D., Puppo, A. (2001). A *Medicago truncatula* homoglutathione synthetase is derived from glutathione synthetase by gene duplication. Plant Physiology 126: 706-1715.

Fu, T.-M., Jacob, D.J., Palmer, P.I., Chance, K., Wang, Y.X., Barletta, B., Blake, D.R., Stanton, J.C., Pilling, M.J. (2007). Space-based formaldehyde measurements as constraints on volatile organic compound emissions in east and south Asia and implications for ozone. Journal of Geophysical Research 112: D06312.

Fuhrer, J. (2009). Ozone risk for crops and pastures in present and future climates. Naturwissenschaften 96: 173-194.

Fuhrer, J., Scarby, L., Ashmore, M.R. (1997). Critical Levels For Ozone Effects on Vegetation in Europe. Environmental Pollution 97: 91-106.

Gadjev, I., Vanderauwera, S., Gechev, T.S., Laloi, C., Minkov, I.N., Shulaev, V., Apel, K., Inze, D., Mittler, R., van Breusegem, F. (2006). Transcriptomic Footprints Disclose Specificity of Reactive Oxygen Species Signaling in *Arabidopsis*. Plant Physiology 141: 436-445.

Galant, A., Arkus, K.A.J., Zubieta, C., Cahoon, R.E., Jez., J.M. (2009). Structural Basis for Evolution of Product Diversity in Soybean Glutathione Biosynthesis. Plant Cell 21: 3450-3458.

Gallogly, M.M., Mieyal, J.J. (2007). Mechanisms of reversible protein glutathionylation in redox signaling and oxidative stress. Current Opinion in Pharmacology 7: 381-391.

Galperin, M.Y., Koonin, E.V. (1997). A diverse superfamily of enzymes with ATP-dependent carboxylate-amino/thiol ligase activity. Protein Science 6: 2639-2643.

Gill N, Findley S, Walling JG, Hans C, Ma J, Doyle J, Stacey G, Jackson SA (2009). Molecular and chromosomal evidence for allopolyploidy in soybean. Plant Physiology 151: 1167–1174.

Gill, S. S., Tuteja, N. (2010). Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. Plant Physiology and Biochemistry 48: 909-930.

Gogos, A., and Shapiro, L. (2002). Large conformational changes in the catalytic cycle of glutathione synthase. Structure 10: 1669-1676.

Griffith, O.W., Meister, A. (1979). Potent and specific inhibition of glutathione synthesis by buthionine sulfoximide (S-n-butyl homocystein sulfoxime). Journal of Biological Chemistry 254: 7558-7560.

Halliwell, B. (2006). Reactive species and antioxidants. Redox biology is a fundamental theme of aerobic life. Plant Physiology 141: 312-322.

Heagle, A.S, Miller, J.E., Pursley, W.A. (1988). Influence of ozone stress on soybean response to carbon dioxide enrichment: III. Yield and Seed Quality. Crop Science 1: 128-134.

Heck, W.W., Adams, R.M., Cure, W.W., Heagle, A.S., Heggestad, H.E., Kohut, R.J., Kress, L.W., Rawlings, J.O., Taylor, O.C. (1983). A reassessment of crop loss from ozone. Environmental Science and Technology 17: 572A-581A.

Hicks, L.M., Cahoon, R.E., Bonner, E.R., Rivard, R.S., Sheffield, J., Jez, J.M. (2007). Thiol-based regulation of the redox-active glutamate-cysteine ligase from *Arabidopsis thaliana*. Plant Cell 19: 2653-2661.

Hisabori, T., Hara, S., Fujii, T., Yamazaki, D., Hosoya-Matsuda, N., Motohashi, K. (2005). Thioredoxin affinity chromatography: a useful method for further understanding the thioredoxin network. Journal of Experimental Botany 56: 1463–1468.

Hothorn, M., Wachter, A., Gromes, R., Stuwe, T., Rausch, T., Scheffzek, K. (2006). Structural basis for the redox control of plant glutamate cysteine ligase. Journal of Biological Chemistry 281: 27557-27565.

Huang, C.-S., Anderson, M., Meister, A. (1993). Amino-acid-sequence and function of the light subunit of rat-kidney gamma-glutamylcysteine synthetase. Journal of Biological Chemistry 268: 20578-20583.

Huang, C.-S., Chang, L.-S., Anderson, M., Meister, A. (1993). Catalytic and regulatory properties of the heavy subunit of rat kidney γ-glutamylcysteine synthetase. Journal of Biological Chemistry. 268: 19675-19680.

Hudman, R.C., Murray, L.T., Jacob, D.J., Millet, B., Turquety, S., Wu, S., Blake, D.R., Goldstein, A.H., Holloway, J., Sachse, G.W. (2008). Biogenic *versus* anthropogenic sources of CO in the United States. Geophysical Research Letters 35: L04801.

IUPAC Compendium of Chemical Terminology. (1997). 2nd Edition. 1990, 62: 2204.

Jensen, W. B., (2007). The Origin of the Oxidation-State Concept. Journal of Chemistry Education 84: 14-18-1419.

Jez, J.M., Cahoon, R.E. (2004). Kinetic Mechanism of Glutathione Synthetase from *Arabidiosis thaliana*. Journal of Biological Chemistry 279(41): 42726-42731.

Kato, H., Tanaka, T., Yamaguchi, H., Hara, T., Nishioka, T., Katsube, Y. Oda, J. (1994). Flexible loop that is novel catalytic machinery in a ligase. Atomic structure and function of the loopless glutathione synthetase. Biochemistry 33: 4995-4999.

Klapheck, S. (1988). Homoglutathione: isolation, quantification, and occurrence in legumes. Physiologia Plantarum 74:727–732

Klapheck, S., Fliegner, W., Zimmer, I. (1994). Hydroxymethyl-phytochelatins [(γ-glutamylcysteine)$_n$-serine] Are metal-induced peptides of the Poaceae. Plant Physiology 104:1325-1332.

Lavoisier, A. (1789). Traité élémentaire de chimie, présenté dans un ordre nouveau et d'après les découvertes modernes. Paris: Chez Cuchet.

Liao H., Chen W.T., Seinfeld J.H. (2006). Role of climate change in global predictions of future tropospheric ozone and aerosols. Journal of Geophysical Research Atmospheres 111:D12304.

Maeda, K., Finnie, C., Svensson, B. (2005). Identification of thioredoxin h-reducible disulfides in proteomes by differential labeling of cysteines: Insight into recognition and regulation of proteins in barley seeds by thioredoxin h. Proteomics 5: 1634-1644.

Masip, L., Veeravalli, K., Georgiou, G. (2006). The many faces of glutathione in bacteria. Antioxidants and Redox Signaling 8: 753-763.

Matamoros, M.A., Moran, J.F., Iturbe-Ormaetxe, I., Rubio, M.C., Becana, M. (1999). Glutathione and homoglutathione synthesis in legume root nodules. Plant Physiology 121: 879-888.

Mauzerall, D.L., Narita, D., Akimoto, H., Horowitz, L., Walters, S., Hauglustaine, D.A., Brasseur, G. (2000). Seasonal characteristics of tropospheric ozone production and mixing ratios over East Asia: a global three-dimensional chemical transport model analysis. Journal of Geophysical Research 105: 17895-17910.

Meagher, J.F., Cowling, E.B., Fehnsenfeld, F.C., Parkhurst, W.J. (1998). Ozone formation and transport in southeastern United States: Overview of the SOS Nathville/ Middle Tennessee ozone study. Journal of Geophysical Research 103: 22213-22223.

Meister, A., (1989). Mechanism and regulation of the glutamine-dependent carbamoyl phosphate synthetase of *Escherichia coli*. Advanced Enzymology Related Area in Molecular Biology 62: 315-374.

Meister, A., Anderson M., E. (1983). Glutathione. Annual Review of Biochemistry 52: 711-760.

Meuwly, P., Rauser, W.E. (1992). Alteration of thiol pools in roots and shoots of maize seedlings exposed to cadmium. Plant Physiology 99: 8-15.

Meyer, A. J., Hell, R. (2005). Glutathione homeostasis and redox-regulation by sulfhydryl groups. Photosynthesis Research 86: 435-457.

Mills, G., Buse, A., Gimeno, B., Bermejo, V., Holland, M., Emberson, L., Pleijel, H. (2007). A synthesis of AOT40-based response functions and critical levels of ozone for agricultural and horticultural crops. Atmospheric Environment 41: 2630-2643.

Monks, P.S., Granier, C., Fuzzi, S., Stohl, A., Williams, M.L., Akimoto, H., Amann, M. Baklanov, A., Baltensperger, U., Bey, I., Blake, N., Blake, R.S., Carslaw, K., Cooper, O.R., Dentener, F., Fowler, D., Fragkou, E., Frost, G.J., Generoso, S., Ginoux, P., Grewe, V., Guenther, A., Hansson, H.C., Henne, S., Hjorth, J., Hofzumahaus, A., Huntrieser, H., Isaksen, I.S.A., Jenkin, M.E., Kaiser, J., Kanakidou, M., Klimont, Z., Kulmala, M., Laj, P., Lawrence, M.G., Lee, J.D., Liousse, C., Maione, M., McFiggans, G., Metzger, A., Mieville, A., Moussiopoulos, N., Orlando, J.J., O'Dowd, C.D., Palmer, P.I., Parrish, D.D., Petzold, A., Platt, U., Poeschl, U., Prevot, A.S.H., Reeves, C.E., Reimann, S., Rudich, Y., Sellegri, K., Steinbrecher, R., Simpson, D., ten Brink, H., Theloke, J., van der Werf, G.R., Vautard, R., Vestreng, V., Vlachokostas, C., von Glasow, R. (2009). Atmospheric composition change - global and regional air quality. Atmospheric Environment 43: 5268-5350.

Moran, J.F., Iturbe-Ormaetze, I., Matamoros, M.A., Rubio, M.C., Clemente, M.R., Brewin, N.J., Becana, M. (2000). Glutathione and homoglutathione synthetases of legume nodules. Cloning, expression, and subcellular localization. Plant Physiology 124: 1381-1392.

Morgan, P.B., Mies, T.A., Bollero, G.A., Nelson, R.L., Long, S.P. (2006). Season-long elevation of ozone concentration to projected 2050 levels under fully open-air conditions substantially decreases growth and production of soybean. New Phytologist 170: 333-343.

Motohashi K., Kondoh A., Stumpp M.T., Hisabori T. (2001). Comprehensive survey of proteins targeted by chloroplast thioredoxin. PNAS 98:11224-11229.

Mou, Z., Fan, W., Dong, X. (2003). Inducers of plant systemic acquired resistance regulate NPR1 function through redox changes. Cell 113: 935-944.

Muller, S., Liebau, E., Walter, R., Krauth-Siegel, R. (2003). Thiol-based redox metabolism of protozoan parasites. Trends in Parasitology. 19: 320-328.

Mullineaux, P.M., Rausch, T. (2005). Glutathione, photosynthesis and the redox regulation of stress-responsive gene expression. Photosynthesis Research 86: 459-474.

Murazaki K., Hess P. (2006). How does climate change contribute to surface ozone change over the United States? Journal of Geophysical Research 111:D05301.

Noctor, G. (2006). Metabolic signaling in defense and stress: the central roles of soluble redox couples. Plant, Cell and Environment. 29: 409-425.

Nolte, C.G., Gilliland, A.B., Hogrefe, C., Mickley, L.J. (2008). Linking global to regional models to assess future climate impacts on surface ozone levels in the United States. Journal of Geophysical Research 113: D14307.

O'Farrel, P.H. (1975). High resolution two-dimensional electrophoresis of proteins. Journal of Biological Chemistry 250: 4007-4021.

Ogawa, K., Hatano-Iwasaki, A., Yanagida, M., Iwabuchi, M. (2004). Level of glutathione is regulated by ATP-dependent ligation of glutamate and cysteine through photosynthesis in *Arabidopsis thaliana*: mechanism of strong interaction of light intensity with flowering. Plant Cell Physiology 45: 1-8.

Ogita, T., Knowles, J.R. (1988). On the interdependency of carboxyphosphate in biotin-dependent carboxylations. Biochemistry 27: 8028-8033.

Polekhina, G., Board, P., Gala, R., Rossjohn, J., Parker, M. (1999). Molecular basis of glutathione synthetase deficiency and a rare gene permutation event. EMBO Journal 18: 3204-3213.

Potters, G., De Gara, L., Asard, H., Horemans, N. (2002). Ascorbate and glutathione: guardians of the cell cycle, partners in crime? Plant Physiology and Biochemistry 40: 537-548.

Prather, M., Gauss, M., Berntsen, T., Isaksen, I., Sundet, J., Bey, I., Brasseur, G., Dentener, F., Derwent, R., Stevenson, D., Grenfell, L., Hauglustaine, D., Horowitz, L., Jacob, D., Mickley, L., Lawrence, M., von Kuhlmann, R., Muller, J.-F., Pitari, G., Rogers, H., Johnson, M., Pyle, J., Law, K., van Weele, N., Wild,

O. (2003). Fresh air in the 21st century? Geophysical Research Letters 30: 72(1-4).

Racherla P.N., Adams P.J. (2006). Sensitivity of global tropospheric ozone and fine particulate matter concentrations to climate change. Journal of Geophysical Research Atmospheres 111:D24103.

Rauser, W.E. (1995). Phytochelatins and related peptides. Plant Physiology 109: 1141-1149.

Redei, G.P. (1975). *Arabidopsis* as a genetic tool. Annual Review of Genetics 9: 111-127.

Rinalducci, S., Murgiano, L., Zolla, L. (2008). Redox proteomics: basic principles and future perspectives for the detection of protein oxidation in plants. Journal of Experimental Botany 59: 3781-3801.

Rogers, L.K., Leinweber, B.L., Smith, C.V. (2006). Detection of reversible protein thiol modifications in tissues. Analytical Biochemistry. 358: 171-184.

Rouhier, N., Lemaire, S., Jacquot, J.-P. (2008). The role of glutathione in photosynthetic organisms: emerging functions for glutaredoxins and glutathionylation. Annual Review of Plant Biology. 59: 143-166.

Rouhier, N., Villarejo, A., Srivastava, M., Gelhaye, E., Keech, O., Droux, M., Finkemeier, I., Samuelsson, G., Dietz, K.J., Jacquot, J.P., Wingsle, G., (2005). Identification of plant glutaredoxin targets. Antioxidants and Redox Signaling. 7: 919–929.

Sarkar, A., Rakwal, R., Agrawal, S.B., Shibato, J., Ogawa, Y., Yoshida, Y., Agrawal, G.K., Agrawal, M. (2010). Investigating the impact of elevated levels of ozone on tropical wheat using integrated phenotypical, physiological, biochemical, and proteomics approaches. Journal of Proteome Research 9: 4565-4584.

Schurmann, P., Buchanan, B. B. (2008). The ferredoxin/thioredoxin system of oxygenic photosynthesis. Antioxidants and Redox Signaling 10: 1235-1273.

Shoemaker R.C., Schlueter J., Doyle J.J. (2006). Paleopolyploidy and gene duplication in soybean and other legumes. Current Opinion in Plant Biology 9: 104–109.

Skipsey, M., Andrews, C. J., Townson, J. K., Jepson, I. & Edwards, R. (2000). Cloning and characterization of glyoxalase I from soybean. Archives of Biochemistry and Biophysics 374:261–268.

Soltaninassab, S.R., Sekhar, K.R., Meredith, M.J. Freeman, M.L. (2000). Multi-faceted regulation of γ-glutamylcysteine synthetase. Journal of Cellular Physiology 182: 163-170.

Sun, W.-M., Huang, Z.-Z., Lu, S.C. (1996). Regulation of γ-glutamylcysteine synthetase by protein phosphorylation. Biochemistry Journal 320: 321-328.

Tang, X., Wilson, S.R., Solomon, K.R., Shao, M., Madronich, S. (2011). Changes in air quality and tropospheric composition due to depletion of stratospheric ozone and interactions with climate. Photochemical and Photobiological Sciences 10: 280-291.

Tanou, G., Job, C., Belghazi, M., Molassiotis, A., Diamantidis, G., Job, D. (2010). Proteomic signatures uncover hydrogen peroxide and nitric oxide cross-talk signaling network in citrus plants. Journal of Proteome Research 9: 5994-6006.

Timms, J.F. (2005). Gel electrophoresis, 2D-difference. Proteins: from analytics to structural genomics, Volume 2. Edited by Robert A. Meyers. 22: 647-666.

Thoden, J.B., Holden, H.M., Wesenberg, G., Rauschel, F.M., Rayment, I. (1997). Structure of carbamoyl phosphate synthetase: a journey of 96 A from substrate to product. Biochemistry 36: 6305-6316.

Torres, N.L., Cho, K., Shibato, J., Hirano, M., Kubo, A., Masuo, Y., Iwahashi, H., Jwa, N.-S. Agrawal, G.K., Rakwal, R. (2007). Gel-based proteomics reveals potential novel protein markers of ozone stress in leaves of cultivated bean and maize species of Panama. Electrophoresis 28: 4369-4381.

Tosti, N., Pasqualini, S., Borgogni, A., Ederli, L., Falistocco, E., Crispi, S., Paolocci, F. (2006). Gene expression profiles of O_3-treated *Arabidopsis* plants. Plant, Cell and Environment 29: 1686-1702.

Van, K., Kim, D.H., Cai, C.M., Kim, M.Y., Shin, J.H., Graham, M.A., Shoemaker, R.C., Choi, B.S., Yang, T.J., Lee, S.H. (2008). Sequence level analysis of recently duplicated regions in soybean [Glycine max (L.) Merr.] genome. DNA Research 15: 93–102.

van Aardenne, J.A., Dentener, F.J., Olivier, G.J., Klein Goldewijk, C.M.G., Lelieveld, J. (2001). A 1° x 1° resolution data set of historical anthropogenic trace gas

emissions for the period 1890-1990. Global Biogeochemical Cycles 15: 909-928.

van Dingenen, R., Dentener, F.J., Raes, F., Krol, M.C., Emberson, L., Colfala, J. (2009). The global impact of ozone on agricultural crop yields under current and future air quality legislation. Atmospheric Environment 43: 604-618.

Vingarzan, R., (2004). A review of surface ozone background levels and trends. Atmospheric Environment 38: 3431-3442.

Volz, A., Kley, A. (1988). Evaluation of the Montsouris series of ozone measurements made in the nineteenth century. Nature 332: 240-242.

Wormuth, D., Heiber, I., Shaikali, J., Kandlbinder, A., Baier, M., Dietz., K.-J. (2007). Redox regulation and antioxidant defense in *Arabidopsis* leaves viewed from a systems biology approach. Journal of Biotechnology 129: 229-248.

Xiang, C., Oliver, D.J. (1998). Glutathione metabolic genes coordinately respond to heavy metals and jasmonic acid in *Arabidopsis*. Plant Cell 10: 1539-1550.

Yamaguchi, H., Kato, H., Hata, Y., Nishioka, T., Kimura, A., Oda, J., Katsube, Y. (1993). Three-dimensional structure of the glutathione synthetase from *Escherichia coli* at 2.0 Å resolution. Journal of Molecular Biology 229: 1083-1100.

Yano H., Wong J.H., Lee Y.M., Cho M.-J., Buchanan B.B. (2001). A strategy for the identification of proteins targeted by thioredoxin. PNAS 98: 4794–99.

Yi, H., Galant, A., Ravilious, G.E., Preuss, M.L., Jez, J.M. (2010). Sensing sulfur conditons: simple to complex protein regulatory mechanisms in plant thiol metabolism. Molecular Plant 3: 269-279.

Youssefian, S., Nakamura, M., Orudgev, E., Kondo, N. (2001). Increased cysteine biosynthesis capacity of transgenic tobacco overexpressing an o-acetylserine(thiol) lyase modifies plant responses to oxidative stress. Plant Physiology. 126: 1001-1011.

Zheng, M., Aslund, F., Storz, G. (1998). Activation of the OxyR transcription factor by reversible disulfide bond formation. Science 279: 1718-1721.

Zhou, L., Bokhari, S.A., Dong, C.-J., Liu, J.-Y. (2011). Comparative proteomics analysis of the root apoplasts of rice seedlings in response to hydrogen peroxide. PLoS ONE 6: e16723.

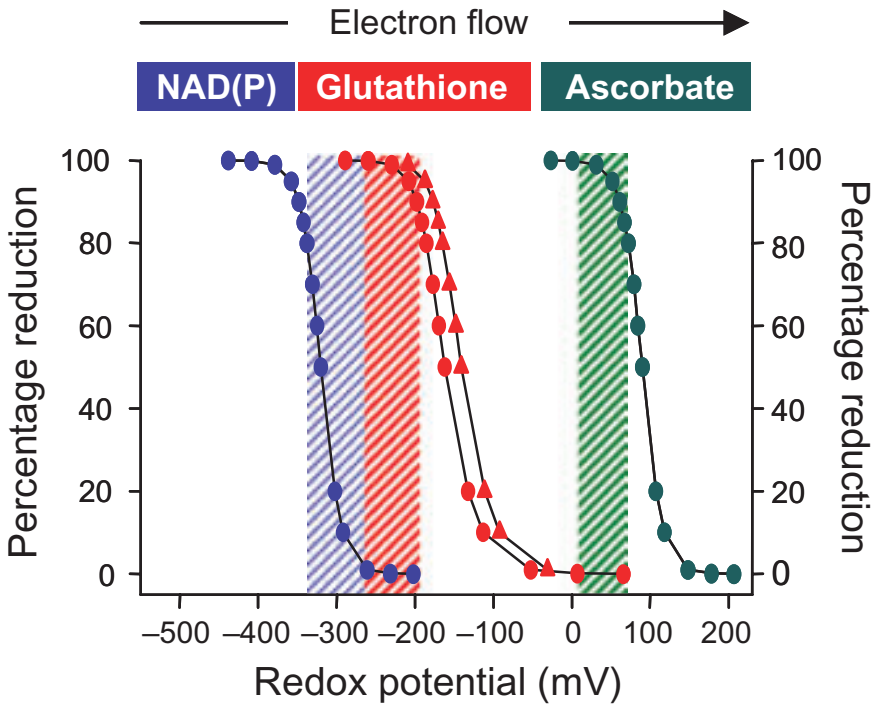


Figure 1. The three major redox couples found in plants. The relationship between percentage reduction and redox potential for each redox couple is shown.

From Noctor, 2006.

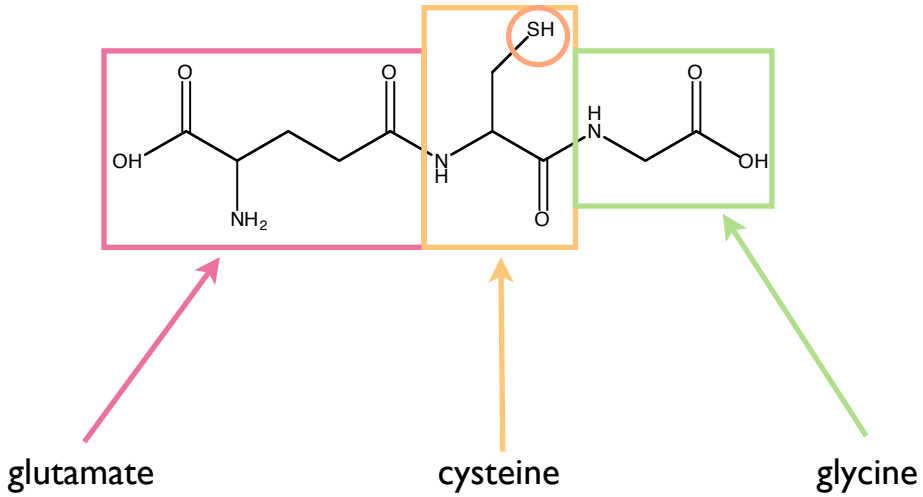


Figure 2. The tripeptide glutathione. Glutathione is composed of three amino acids: glutamate, cysteine, and glycine, as outlined. The critical sulfhydryl group that allows glutathione to act as an effective redox agent is circled.

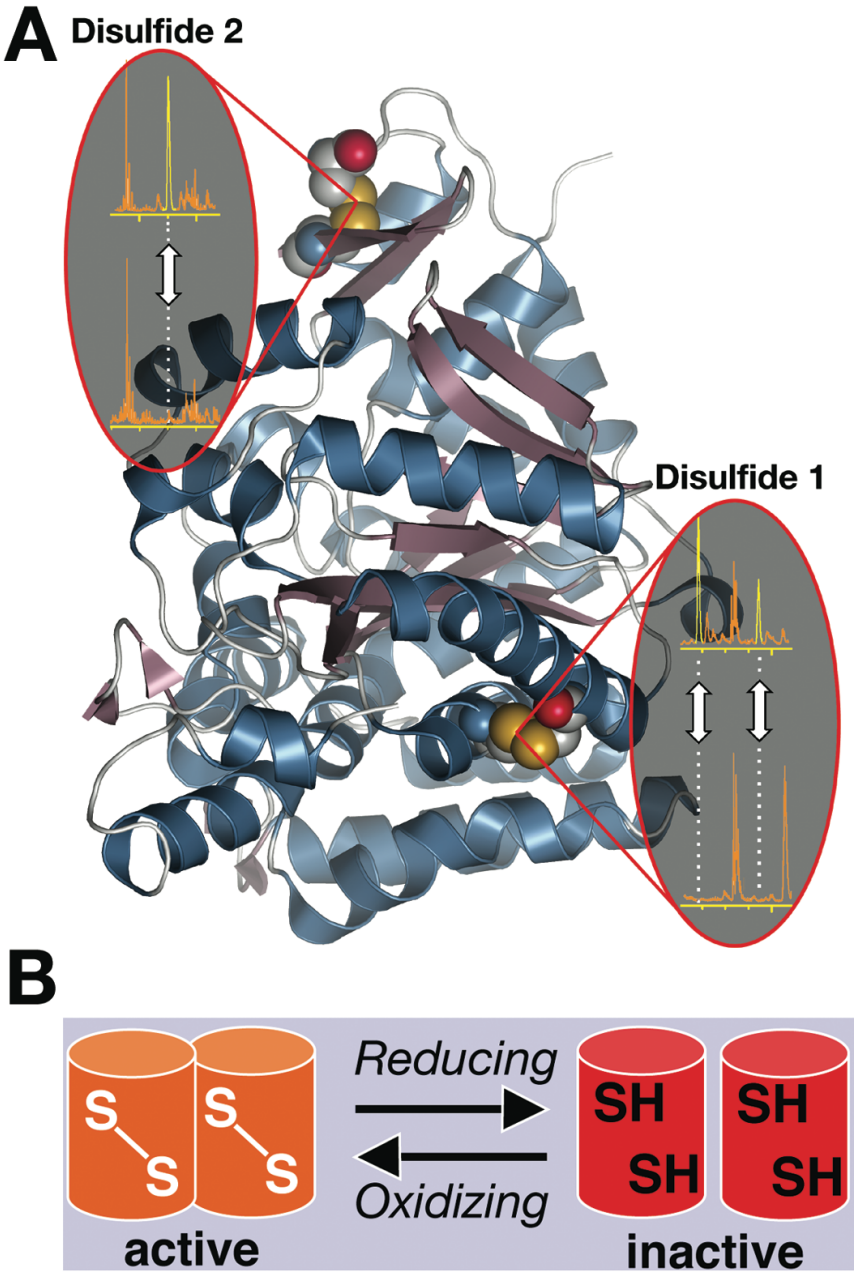


Figure 3. Redox regulation of plant GCL. A.) A monomer of the *Brassica juncea* (Indian Mustard) GCL enzyme with the locations and mass spectrometry profiles of the reduced versus oxidized disulfide bonds indicated. B.) Schematic detailing the transition from the inactive to active form of the GCL enzyme, and *vice versa*.

From Yi et al., 2010.

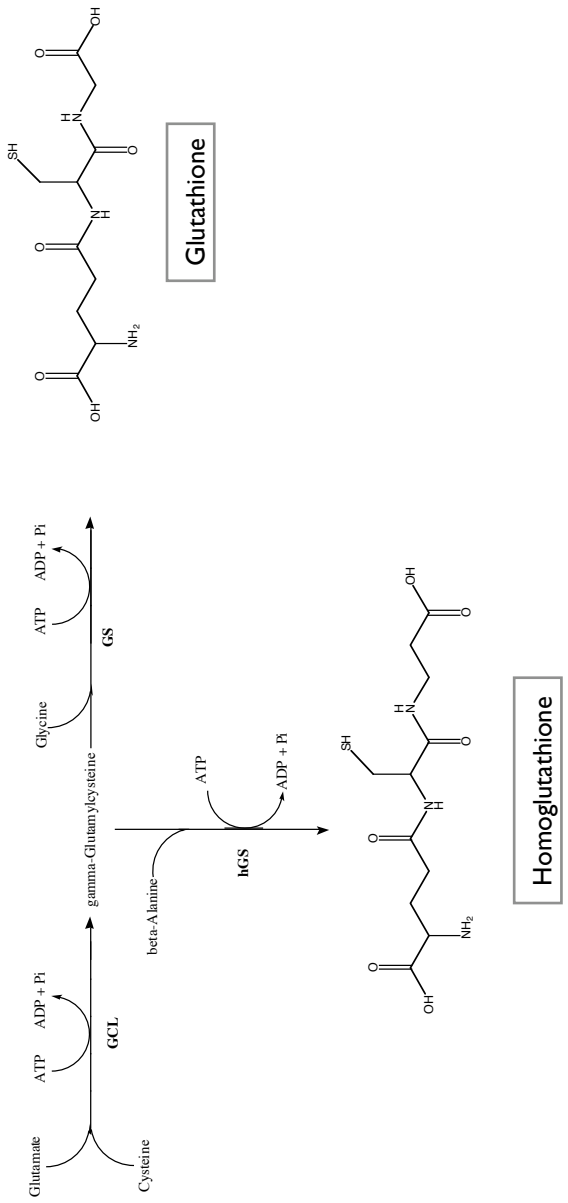


Figure 4. GSH and hGSH biosynthesis. hGSH and GSH are synthesized in two ATP-dependent steps. GCL activity is shared between the two pathways, while a committed enzyme (either GS or hGS) performs the second reaction.

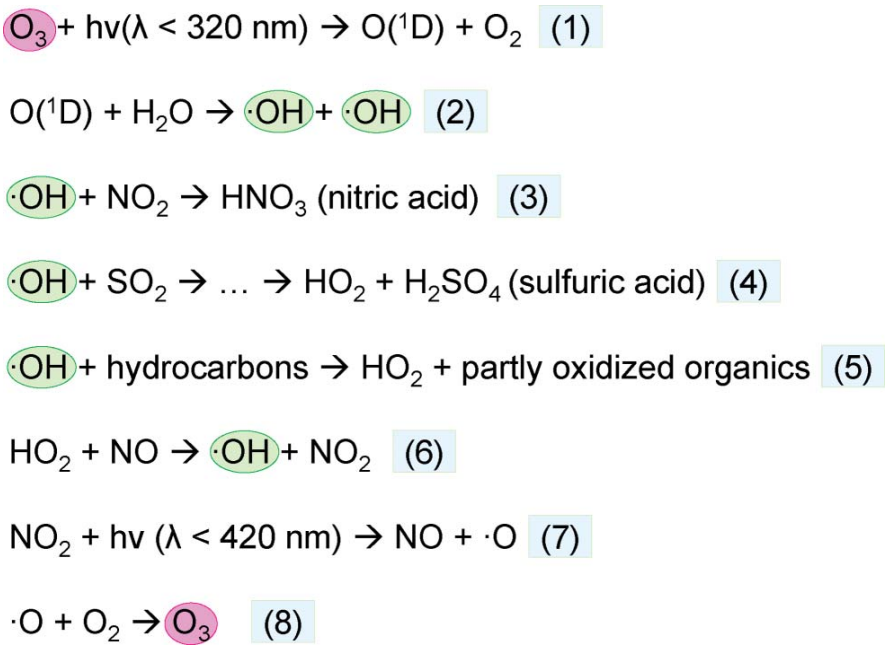


Figure 5. Key reactions for synthesizing ozone in the stratosphere and troposphere. From Tang et al., 2011.

CHAPTER 2

STRUCTURAL BASIS FOR EVOLUTION OF PRODUCT DIVERSITY IN SOYBEAN GLUTATHIONE BIOSYNTHESIS

PREFACE

As described in the introduction, homoglutathione (hGSH) is not the only redox-labile glutathione homolog that is produced in plants. To date, three such tripeptides - hGSH (in legumes), hydroxymethylglutathione (in grasses), and gamma-glutamylcysteinylglutamate (in cadmium-stressed maize), have been isolated. Of these three, glutamylcysteinylglutamate - isolated in the early 1990s - is the most recent discovery; by the mid-90s both hGSH and hydroxymethylglutathione (hmGSH) had been under study for nearly a decade [Meuwly et al., 1993]. While the enzyme activity behind hGSH synthesis was rapidly identified, early hmGSH work focused on its interactions with alcohol dehydrogenase and reactive oxygen species (ROS) [Macnicol, 1987; Zopes et al., 1993; Martinez et al., 1996]. In 2002, carboxypeptidase Y was identified as catalyzing the synthesis of hmGSH *in vitro*; however, it remains unknown if this activity has any physiological significance [Okumura et al., 2003]. To date, the source of glutamylcysteinylglutamate remains unclear. Kinetic characterization has confirmed that the maize glutathione synthetase is not capable of using glutamic acid in place of glycine [Skipsey et al., 2005]. Thus, it is likely that unidentified enzyme is responsible for glutamylcysteinylglutamate synthesis. The chemical diversity of the GSH homologs suggests that the substrate specificity of the glutathione synthetase (GS)-related enzymes in these plants differs from the canonical GS, but the origin and exact role of these enzymes remains obscured.

Author Contributions: JMJ designed research; AG, KAJA, CZ, and REC performed research; AG, REC, and JMJ analyzed data; AG and JMJ wrote the paper.

Preface References

Macnicol, P.K. (1987). Homoglutathione and glutathione synthetases of legume seedlings: partial purification and substrate specificity. Plant Science 53: 229-235.

Martinez, M.C., Achkor, H., Persson, B., Fernandez, M.R., Shafqat, J., Farres, J., Jornvall, H., Pares, X. (1996). Arabidopsis formaldehyde dehydrogenase - Molecular properties of plant class III alcohol dehydrogenase provide further insights into the origins, structure and function of plant class P and liver class I alcohol dehydrogenases. European Journal of Biochemistry 241: 849-857.

Meuwly, P., Thibault, P., Rauser, W.E. (1993). Gamma-glutamylcysteinylglutamic acid - a new homolog of glutathione in maize seedlings exposed to cadmium. FEBS Letters 336: 472-476.

Okumura, R., Koizuma, Y., Sekiya, J. (2003). Synthesis of hydroxymethylglutathione from glutathione and L-serine catalyzed by Carboxypeptidase Y. Bioscience Biotechnology and Biochemistry 67: 434-437.

Skipsey, M., Davis, B.G., Edwards, R. (2005). Diversification in substrate usage by glutathione synthetases from soya bean (*Glycine max*), wheat (*Triticum aestivum*) and maize (*Zea mays*). Biochemistry Journal 391: 567-574.

Zopes, H., Klapheck, S., Bergmann, L. (1993). The Function of Homoglutathione and Hydroxymethylglutathione for the scavenging of hydrogen-peroxide. Plant Cell and Physiology 34: 515-521.

The Plant Cell, Vol. 21: 3450–3458, November 2009, www.plantcell.org

Structural Basis for Evolution of Product Diversity in Soybean Glutathione Biosynthesis[W]

Ashley Galant,[a] Kiani A.J. Arkus,[a,b,1] Chloe Zubieta,[b] Rebecca E. Cahoon,[b] and Joseph M. Jez[a,b,2]

[a] Department of Biology, Washington University, St. Louis, Missouri 63130
[b] Donald Danforth Plant Science Center, St. Louis, Missouri 63132

The redox active peptide glutathione is ubiquitous in nature, but some plants also synthesize glutathione analogs in response to environmental stresses. To understand the evolution of chemical diversity in the closely related enzymes homoglutathione synthetase (hGS) and glutathione synthetase (GS), we determined the structures of soybean (*Glycine max*) hGS in three states: apoenzyme, bound to γ-glutamylcysteine (γEC), and with hGSH, ADP, and a sulfate ion bound in the active site. Domain movements and rearrangement of active site loops change the structure from an open active site form (apoenzyme and γEC complex) to a closed active site form (hGSH·ADP·SO_4^{2-} complex). The structure of hGS shows that two amino acid differences in an active site loop provide extra space to accommodate the longer β-Ala moiety of hGSH in comparison to the glycinyl group of glutathione. Mutation of either Leu-487 or Pro-488 to an Ala improves catalytic efficiency using Gly, but a double mutation (L487A/P488A) is required to convert the substrate preference of hGS from β-Ala to Gly. These structures, combined with site-directed mutagenesis, reveal the molecular changes that define the substrate preference of hGS, explain the product diversity within evolutionarily related GS-like enzymes, and reinforce the critical role of active site loops in the adaptation and diversification of enzyme function.

INTRODUCTION

The tripeptide glutathione (GSH) is found in nearly all eukaryotes and prokaryotes and functions as a key component in an array of redox-linked cellular systems (Meister, 1995). In plants, GSH maintains cellular redox homeostasis, detoxifies harmful xenobiotics and heavy metals, and can regulate enzyme activity through glutathionylation (May et al., 1998; Noctor and Foyer, 1998; Rouhier et al., 2008). Although GSH is the predominant thiol-containing tripeptide found in plants, various plant species produce glutathione homologs in which the terminal Gly is substituted with a different amino acid (Figure 1A). For example, legumes make GSH, in addition to producing homoglutathione (hGSH), in which β-Ala replaces Gly, in a tissue-specific manner (Klapheck et al., 1995; Matamoros et al., 1999). Synthesis of hGSH maintains redox balance in legume nodules (Moran et al., 2000) and is critical for rhizobia-legume nodulation in roots (Matamoros et al., 2003; Frendo et al., 2005; Loscos et al., 2008). Similarly, many grasses synthesize GSH and hydroxymethylglutathione, with Ser instead of Gly, and exposure to cadmium activates the production of γ-glutamylcysteinylglutamate in maize (*Zea mays*; Rauser et al., 1986; Klapheck et al., 1994; Meuwly et al., 1995). The molecular details of how these peptides are generated and the biological functions of GSH analogs in plants are poorly understood, but these specialized peptides likely provide for specific responses to various environmental stresses.

Although the biosynthetic routes for the Ser- and Glu-containing peptides are unclear, the two-step pathways leading to GSH and hGSH are similar and better understood at the metabolic level. In the first reaction of the pathway, Glu-Cys ligase catalyzes the formation of γ-glutamylcysteine (γEC) from Glu and Cys (Jez et al., 2004; Hicks et al., 2007). The second step in the synthesis of either GSH or hGSH depends on the specificity of the synthetase for the terminal substrate. In nearly all organisms, glutathione synthetase (GS) catalyzes the addition of Gly to γEC (Meister, 1995; Jez and Cahoon, 2004; Herrera et al., 2007). In legumes, homoglutathione synthetase (hGS) uses β-Ala instead of Gly to form hGSH (Matamoros et al., 1999; Frendo et al., 2001; Iturbe-Ormaetxe et al., 2002). Although GS and hGS share similar reaction mechanisms based on biochemical and structural studies, the molecular basis for the difference in substrate specificity is unclear due to no available structural data for any plant GS or hGS.

Based on sequence similarity, both GS and hGS are members of the ATP-grasp enzyme superfamily (Galperin and Koonin, 1997). All ATP-grasp family members catalyze the ATP-dependent ligation of the carboxyl group carbon of one substrate to the amino- or imino-nitrogen of another substrate. For example, hGS catalyzes the transfer of the γ-phosphate group of ATP to the C-terminal carboxylate of γEC to yield an acylphosphate intermediate (Figure 1B). Subsequent nucleophilic attack on this intermediate by β-Ala leads to formation of hGSH with release of ADP and inorganic phosphate (Figure 1B). The structurally characterized tetrameric GS from *Escherichia coli* (Yamaguchi

[1] Current address: Cell and Molecular Biology Program, Duke University, Durham, NC 27710.
[2] Address correspondence to jjez@wustl.edu.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Joseph M. Jez (jjez@wustl.edu).
[W] Online version contains Web-only data.
www.plantcell.org/cgi/doi/10.1105/tpc.109.071183

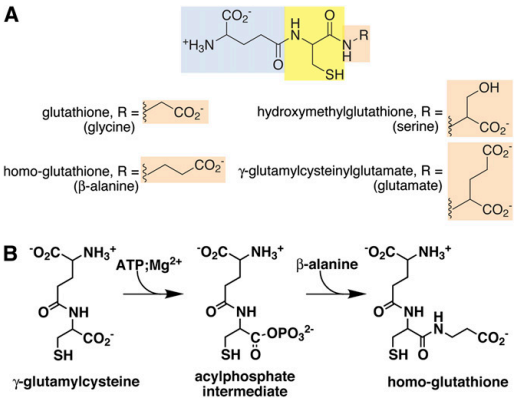


Figure 1. Diversity in GSH Biosynthesis.

(A) Chemical structures of GSH and related peptides from plants. Colors indicate the γ-glutamyl (blue), cysteinyl (yellow), and variable amino acid (peach) residues of each molecule.

(B) Overall reaction catalyzed by hGSH synthetase. This scheme shows the formation of the acylphosphate intermediate resulting from phosphorylation of γ-glutamylcysteine and the subsequent addition of β-Ala to yield hGSH. Note that the release of ADP and P_i from the reaction is not shown.

et al., 1993) and the dimeric GS from eukaryotes, such as humans, yeast, and plants (Polekhina et al., 1999; Gogos and Shapiro, 2002; Jez and Cahoon, 2004), are unrelated in sequence but share the common ATP-grasp fold. The GS from human, yeast, and *Arabidopsis thaliana* and the hGS from soybean (*Glycine max*) are related, with ~40% sequence identity.

The evolution of GS-related enzymes in plants led to greater product diversity; however, the molecular basis for this adaptation is unknown. To understand the structural evolution of hGS from GS, we determined the x-ray crystal structures of soybean hGS at three separate points during its reaction sequence: (1) the apoenzyme in an open active site conformation, (2) an open form with γEC bound, and (3) a closed form with hGSH, ADP, and a sulfate ion bound in the active site. These structures, combined with site-directed mutagenesis, reveal the structural features that define the substrate preference of hGS, explain the product diversity within evolutionarily related GS-like enzymes, and reinforce the critical role of active site loops in the adaptation and diversification of enzyme function.

RESULTS

Protein Expression and Kinetic Analysis of hGS

Soybean hGS was overexpressed in *E. coli* as a His-tagged fusion protein and purified using Ni^{2+}-affinity and size-exclusion chromatographies. Analysis of the protein by SDS-PAGE showed a monomeric molecular mass of 50 kD, which agrees with the predicted mass based on amino acid sequence (see Supplemental Figure 1 online). The protein eluted from the gel filtration column as a 102-kD species corresponding to a dimer (see Supplemental Figure 1 online). Other eukaryotic GS also are dimeric (Polekhina et al., 1999; Gogos and Shapiro, 2002; Jez and Cahoon, 2004). Purified recombinant hGS had a specific activity of 1.2 μmol min^{-1} mg $protein^{-1}$ and required Mg^{2+} for activity. Steady state kinetic parameters of hGS for γEC, ATP, and β-Ala were determined (Table 1). In comparison to the GS from *Arabidopsis* (Jez and Cahoon, 2004; Herrera et al., 2007), hGS displayed a turnover rate (V/E_t) fivefold lower but with comparable K_m values for both ATP and γEC. In contrast with GS, which shows no activity if Gly is substituted with β-Ala, Ser, or Glu (Jez and Cahoon, 2004), hGS exhibited a 700-fold preference for β-Ala over Gly as the terminal substrate. Estimates of the turnover rate and K_m values of hGS with Gly should be considered as approximate because higher concentrations of Gly, and higher amounts of protein were required to observe activity. hGS did not accept either Ser or Glu as a substrate.

Overall Structure of hGS

Soybean hGS crystallized under similar conditions in either the absence or presence of ligands (Table 2). The protein adopts either a closed active site form (bound with hGSH and ADP) or an open active site form (apoenzyme and γEC bound). In the closed form, the unit cell contained two crystallographically independent molecules with the physiological dimer formed by crystallographic symmetry. For each open-form structure, the asymmetric unit contains two monomers that represent the physiologic dimer.

The overall structure of the closed-form hGS homodimer is shown in Figure 2A. The core structure of each monomer is a triangular α/β-fold that is ~60 Å × 60 Å in length and width, in which binding of hGSH and ADP (Figure 2B) defines features of the active site. A smaller lid domain (residues 366 to 427) formed by an antiparallel β-sheet, two α-helices, and a Gly-rich loop (residues 390 to 398) undergoes major conformational changes

Table 1. Comparison of Kinetic Parameters for *Arabidopsis* GS and Soybean hGS

	V/E_t (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
	GS[a]		
γEC	12.2 ± 0.3	39 ± 5	312,800
ATP	12.1 ± 0.3	57 ± 10	212,300
Gly	12.6 ± 0.5	1,510 ± 88	8,340
βAla	–	–	–
	hGS		
γEC	2.5 ± 0.1	44 ± 6	56,820
ATP	1.7 ± 0.1	23 ± 4	73,910
Gly	<0.1	>100 mM	1
βAla	2.4 ± 0.1	3,390 ± 100	708

Values are expressed as a mean ± SE for $n = 3$.

[a]Kinetic parameters for *Arabidopsis* GS are from Jez and Cahoon (2004) and are provided here for comparison to soybean hGS.

Table 2. Crystallographic Statistics

Crystal	Open	Open + γEC	Closed + hGSH + ADP
Space Group	$P2_1$	$P2_1$	$P3_2$
Cell dimensions	*a* = 64.96 Å, *b* = 80.55 Å, *c* = 90.00 Å; α = γ = 90.0°, β = 96.9°	*a* = 64.88 Å, *b* = 80.95 Å, *c* = 89.12 Å; α = γ = 90.0°, β = 95.6°	*a* = *b* = 115.7 Å, *c* = 101.8 Å; α = β = 90.0°, γ = 120°
Data Collection			
Wavelength (Å)	0.979	0.979	0.979
Resolution range (Å) (highest shell resolution)	28.4–2.0 (2.05–2.0)	19.8–2.1 (2.16–2.1)	29.3–1.9 (1.95–1.9)
Reflections (total/unique)	132,097/59,099	191,449/51,043	445,529/116,705
Completeness (highest shell)	94.8% (86.8%)	96.5% (92.3%)	97.2% (93.1%)
$\langle I/\sigma \rangle$ (highest shell)	13.7 (3.4)	11.4 (2.4)	17.4 (2.9)
R_{sym}[a] (highest shell)	6.9% (37.9%)	10.8% (46.7%)	5.1% (45.8%)
Model and Refinement			
R_{cryst}[b]/R_{free}[c]	19.8/26.9	20.7/28.8	19.7/25.0
No. of protein atoms	7108	6992	7424
No. of water molecules	472	290	741
No. of ligand atoms	–	48	152
r.m.s. deviation, bond lengths (Å)	0.021	0.028	0.047
r.m.s. deviation, bond angles (°)	2.03	2.55	3.46
Average B-factor (Å^2)	26.0	29.3	36.1
Stereochemistry: most favored, allowed, generously allowed	88.4, 9.3, 2.3%	86.2, 12.2, 1.6%	91.1, 8.2, 0.7%

[a] $R_{sym} = \Sigma|I_h - \langle I_h \rangle|/\Sigma I_h$, where $\langle I_h \rangle$ is the average intensity over symmetry.
[b] $R_{cryst} = \Sigma||F_o| - \langle |F_c| \rangle|/\Sigma|F_o|$, where summation is over the data used for refinement.
[c] R_{free} is defined the same as R_{cryst} but was calculated using 5% of data excluded from refinement.

between the closed and open active site structures (Figures 3B and 3C). In addition, a second loop (residues 479 to 491) forms part of the active site. For consistency with the human and yeast GS structures (Polekhina et al., 1999; Gogos and Shapiro, 2002), this second loop is referred to as the Ala-rich loop, even though the corresponding Ala residues are replaced by a Leu and a Pro in hGS. Dimerization of hGS occurs through a pseudo-twofold axis between two helices (α2 and α9) and an antiparallel β-sheet (β1 and β2) of each monomer. The closed-form structure of soybean hGS is similar to those of other ATP-grasp family proteins, such as human GS (root mean square [r.m.s.] deviation of 1.6 Å^2 for 459 Cα atoms) (Figure 2C) and yeast GS (r.m.s. deviation of 2.2 Å^2 for 448 Cα atoms). The hGS structure is also related to the synthetase domains of the bifunctional glutathionylspermidine synthetase/amidase from *E. coli* (r.m.s. deviation of 3.8 Å^2; Pai et al., 2006) and trypanothione synthetase/amidase from *Leishmania* (r.m.s. deviation of 4.0 Å^2; Fyfe et al., 2008).

Domain Movements: Open and Closed Active Site Forms

The apoenzyme and γEC-bound structures are nearly identical, with an r.m.s. deviation of 0.5 Å^2. Disordered regions include most of the Gly-rich loop (residues 391 to 396) and other portions of the lid domain (residues 410 to 420) in each monomer of the dimer. For both open form structures, the Ala-rich loop is ordered in one monomer but disordered (residues 480 to 489) in the second monomer. In the open form, both the lid domain and Ala-rich loop are positioned away from the active site to reveal the binding sites for γEC and ATP and allow for substrate binding (Figures 3A and 3B).

In the closed form, the lid domain, including the Gly-rich loop, and the Ala-rich loop undergo major rearrangements compared with the open form (Figures 3A and 3C). With the exception of residues 410 to 416, the lid domain becomes ordered, with residues in the Gly-rich loop providing multiple interactions with the nucleotide. Likewise, the Ala-rich loop shifts to position residues for contact with both ADP and hGSH. As noted for the bacterial and eukaryotic GS (Polekhina et al., 1999; Gogos and Shapiro, 2002; Jez and Cahoon, 2004), nucleotide binding triggers movement of the lid domain and Ala-rich loop through multiple protein–ligand interactions. Based on the proposed reaction mechanism for GS (Herrera et al., 2007), enclosure of the active site likely prevents hydrolysis of the reactive acyl-phosphate intermediate (Figure 1B).

γ-Glutamylcysteine Binding Site in the Open Form

In the reactions catalyzed by hGS and GS, γEC is a common substrate, and its binding site is highly conserved in both sequence and structure between hGS and the GS from human, yeast, and *Arabidopsis*. In the γEC binding site of hGS, Ser-176, Arg-295, Glu-241, and Gln-238 interact with the glutamyl portion of the molecule (Figure 4A). Of these, the charge–charge interaction between the Arg and the carboxylate group is critical for γEC binding in GS (Herrera et al., 2007), suggesting an analogous role for this interaction in hGS. Similar to interactions observed in the structure of yeast GS complexed with γEC and an ATP analog (Gogos and Shapiro, 2002), Tyr-298 forms a hydrogen bond to the carbonyl of the glutamyl group and there is a bidentate charge–charge interaction between Arg-153 and the

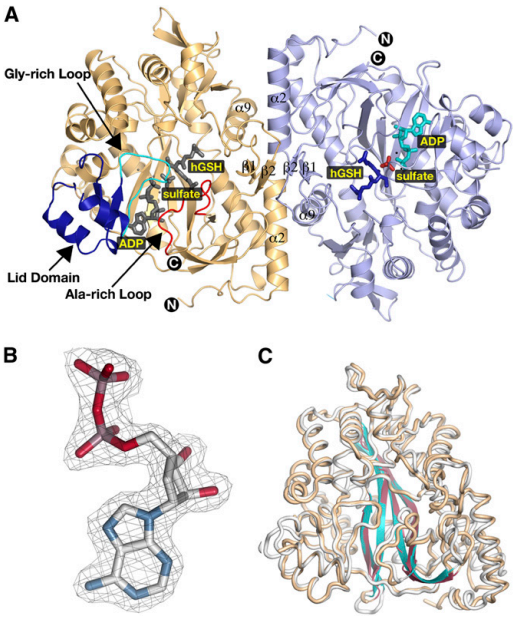


Figure 2. Structure of hGS.

(A) Ribbon diagram of the hGS dimer. Each monomer is colored either gold or blue. Secondary structure elements that form the dimer interface are labeled. The locations of the lid domain (dark blue), Gly-rich loop (cyan), and the Ala-rich loop (red) are highlighted in the gold monomer. The positions of bound ADP (cyan), sulfate (red), and hGSH (dark blue) are highlighted in the blue monomer with corresponding ligands colored gray in the gold monomer. The N- and C-terminal residues of each monomer observed in the electron density maps are indicated.
(B) Sample electron density. The $2F_o$-F_c omit map (1.0 σ) for ADP bound in the active site of the closed form.
(C) Structural overlay of human GS (tan) and soybean hGS (white). The ATP-grasp structural motifs in GS and hGS are colored magenta and cyan, respectively.

carboxylate of the cysteinyl moiety (Figure 4A). Nearly all of these interactions are conserved when hGSH is bound in the site.

Active Site and Ligand Binding in the Closed Form

To define the active site, hGS was cocrystallized in the presence of reaction products ADP and hGSH (Figure 4B). In addition to the reaction products, a sulfate ion and three magnesium ions were identified in the active site of the closed form structure. Clear tetragonal density for the sulfate, which mimics binding of the inorganic phosphate product, was observed (Polekhina et al., 1999). Based on the positional similarity with the yeast and human GS structures, coordination, and strong electron density (4σ), three atoms were modeled as Mg^{2+}.

As with the γEC binding site, the residues forming the nucleotide binding site between the lid domain and Ala-rich loop of hGS (Figures 3B and 4B) are structurally conserved with those in the structures of human and yeast GS (Polekhina et al., 1999; Gogos and Shapiro, 2002). The adenosine ring forms main-chain contacts with Ile-427 and Gln-425 and a hydrogen bond with Lys-388. The ribose hydroxyl groups interact with Lys-477 and Glu-450, respectively. A series of polar interactions occur between the diphosphate tail and Lys-334, Asn-397, and two Mg^{2+} ions. The α- and β-phosphate groups of the nucleotide and Glu-169 coordinate one Mg^{2+} with a second ion bound by the β-phosphate group, the sulfate, Glu-169, Asn-171, and Glu-392. Based on mechanistic studies of *Arabidopsis* GS, the magnesium ions and their coordinating residues play critical roles in stabilizing charges during catalysis (Herrera et al., 2007). The functional role of the third Mg^{2+} is unclear, as it does not interact with any of the bound ligands. This ion is coordinated by interactions with Glu-392 and main-chain contacts with Met-170 and Gly-332 that appear to help orient residues coordinated to the other Mg^{2+} ions.

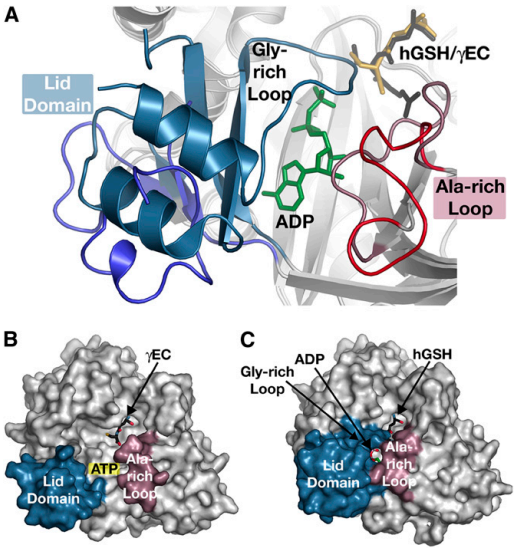


Figure 3. Domain and Loop Movements in hGS.

(A) Ribbon diagram comparing the open and closed active site forms. The active site regions of the γEC bound open form and the closed form are aligned. Stick drawings show γEC (gold) in the open form and ADP (green) and hGSH (black) in the closed form. The positions of the lid domain, including the Gly-rich loop, and the Ala-rich loop in the closed form are shown in blue and rose, respectively. The locations of the lid domain and Ala-rich loop in the open form are shown in lighter blue and red, respectively. In the open form, the Gly-rich loop region is disordered.
(B) Surface rendering of the open form bound with γEC. The lid domain (blue) and Ala-rich loop (rose) leave the nucleotide binding site open.
(C) Surface rendering of the closed form bound with ADP and hGSH. The lid domain (blue) and Ala-rich loop (rose) enclose the active site.

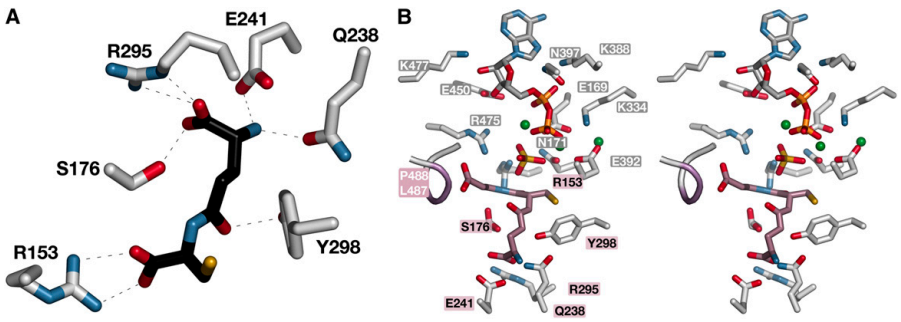


Figure 4. Substrate Binding Sites in the Open and Closed Forms of hGS.

(A) The γEC binding site. Side chains of residues that form hydrogen bonds (dotted lines) with the bound ligand are shown. Atom types are indicated by colors: red = oxygen; blue = nitrogen; yellow = sulfur.
(B) The active site of hGS. The positions of hGSH (rose), ADP (gray), sulfate (yellow), and magnesium ions (green) and the side chains of interacting residues are shown. A portion of the Ala-rich loop is shown as a tube with the residues that vary between hGS and GS colored. For clarity, main chain and water-mediated contacts are not shown. Residues in the nucleotide binding site, the tripeptide binding site, and the Ala-rich loop are labeled white on gray, black on rose, and white on rose, respectively. Alternate conformers for Ser-176 are shown. Atom types are indicated by colors: red = oxygen; blue = nitrogen; yellow = sulfur; orange = phosphorus.

Within the peptide binding site, all the interactions of the glutamyl portion of hGSH are identical to those observed in the γEC complex with minor differences in interactions with the cysteinyl group (Figure 4). Ser-176 is observed in alternate conformations. The side chains of Tyr-298 and Arg-153 are repositioned in the closed form complex. Tyr-298 rotates away from the tripeptide, and Arg-153 now interacts with the cysteinyl carbonyl group and the sulfate. The Arg is essential for catalyzing formation of the acylphosphate intermediate in the first part of the catalytic mechanism and in guiding nucleophilic attack in the second half of the reaction to yield the tripeptide product (Herrera et al., 2007). The carboxylate of the β-Ala moiety of hGSH forms a hydrogen bond with the backbone amide of Val-486 and an ionic interaction with the guanido group of Arg-475. Additional van der Waals contacts between the β-Ala–derived portion of hGSH are made with Leu-487 and Pro-488 in the Ala-rich loop. Interestingly, these two residues differ in hGS compared with GS.

Determinants of Substrate Specificity and Product Diversity

In the active site of hGS, Leu-487 and Pro-488 are the only residues that differ from the characterized eukaryotic GS sequences (Figure 5). In GS, these residues are sequential Ala residues, which help give the Ala-rich loop its name. Structural comparison of hGS and human GS shows that the Ala-rich loop in hGS is shifted ~3 Å away from the corresponding position of the loop in the GS structure to accommodate the larger β-Ala moiety (Figure 5).

To test the functional significance of Leu-487 and Pro-488 in determining the specificity of hGS for β-Ala over Gly, we generated Ala substitutions at each position (L487A and P488A) and the corresponding double mutant (L487A/P488A). Each mutant protein was expressed, purified, and assayed to determined steady state kinetic parameters for β-Ala and Gly as substrates (Table 3). Wild-type hGS displays a specificity ratio 708:1 in preference of β-Ala. Each point mutation altered substrate preference to different degrees. Although the L487A mutant shows a 3.4-fold reduction in catalytic efficiency with β-Ala and a 46-fold improvement using Gly as a substrate, this enzyme still prefers the hGS substrate by nearly fivefold. The P488A mutation yields an enzyme with almost equal preference for either substrate, resulting from a 274-fold increase in efficiency with Gly and a minor 2.3-fold reduction in k_{cat}/K_m with β-Ala. The L487A/P488A mutant retains activity with β-Ala at a 10-fold reduction compared with the hGS, but this mutant is as effective with Gly as the parent enzyme is with β-Ala. The combination of substitutions in the double L487A/P488A mutant converts hGS into a GS with a 950-fold increase in k_{cat}/K_m with Gly.

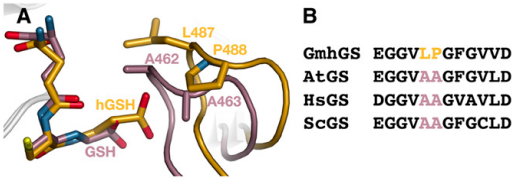


Figure 5. Comparison of the Ala-Rich Loops in hGS and GS.

(A) The overlaid structures of soybean hGS (gold) and human GS (rose) show the similarity of hGSH (gold) and GSH (rose) binding in each structure. The Ala-rich loop of each structure is shown with the side chains of the two amino acid differences between the structures shown.
(B) Sequence comparison of the Ala-rich loops of soybean hGS (GmhGS), *Arabidopsis* GS (AtGS), human GS (HsGS), and yeast GS (ScGS). The amino acid differences are highlighted using the color scheme from **(A)**.

Table 3. Substrate Specificity of Wild-type and Mutant Soybean hGS

	β-Ala		
	V/E_t (s^{-1})	K_m (mM)	k_{cat}/K_m (M^{-1} s^{-1})
hGS	2.4 ± 0.1	3.4 ± 0.1	708
L487A	0.8 ± 0.1	3.8 ± 0.6	211
P488A	2.5 ± 0.2	8.0 ± 1.4	313
LP/AA[a]	1.9 ± 0.2	24.8 ± 2.3	77
		Gly	
hGS	<0.1	>100	1
L487A	0.3 ± 0.1	6.5 ± 0.5	46
P488A	1.4 ± 0.1	5.1 ± 0.4	274
LP/AA[a]	2.1 ± 0.3	2.2 ± 0.1	950

Values are expressed as a mean ± SE for $n = 3$.
[a]Denotes the corresponding double mutant (L487A/P488A).

DISCUSSION

Functional diversity across enzyme families with shared three-dimensional structures and reaction chemistry is a hallmark in the evolution of metabolic pathways. Nearly all eukaryotes and prokaryotes synthesize the multifunctional peptide GSH (Meister, 1995); however, some plants also synthesize GSH analogs with substitutions of the terminal Gly (Figure 1A) (Rauser et al., 1986; Klapheck et al., 1994; Klapheck et al., 1995; Meuwly et al., 1995; Matamoros et al., 1999). In particular, many legumes produce hGSH for root nodulation (Matamoros et al., 2003; Frendo et al., 2005; Loscos et al., 2008). As hGS likely evolved from GS, we examined the structural basis for adaptation of product diversity in hGS. Crystallographic analysis of soybean hGS provides insight on structural changes during the catalytic cycle of both hGS and GS and, combined with site-directed mutagenesis, defines active site differences that govern substrate preference. This work reinforces the critical role of flexible loops in the adaptation and diversification of enzyme function.

Catalysis in hGS and GS requires the orchestration of binding multiple substrates and the rearrangement of active site features, including the lid domain, Gly-rich loop, and Ala-rich loop. Together with studies of the kinetic and chemical mechanisms of GS (Jez and Cahoon, 2004; Herrera et al., 2007), crystal structures of hGS (Figures 2 and 3) and GS (Polekhina et al., 1999; Gogos and Shapiro, 2002) now provide views of the progression through the catalytic cycle from apoenzyme (hGS and yeast GS) to first substrate complex (hGS•γEC complex) to second substrate complex (yeast GS in complex with γEC and an ATP analog) to product complex (hGS and human GS). Kinetic analysis of *Arabidopsis* GS indicates a mechanism in which γEC is the preferred first substrate followed by ATP (Jez and Cahoon, 2004). Within the active site, hGS shares common structural and chemical features with GS. In both enzymes, the γEC binding site is structurally static, whereas the ATP and β-Ala/Gly binding sites are dynamic. The structure of the hGS•γEC complex in the open active site conformation (Figures 3B and 4A) provides direct evidence for formation of this complex in agreement with the predicted mechanism for GS and hGS. Binding of ATP, which makes extensive contacts with residues in the lid domain, Gly-rich loop, and Ala-rich loop (Figures 3C and 4B), likely triggers transformation to the closed active site structure (Gogos and Shapiro, 2002; Gunasekaran et al., 2003). The closed active site conformation protects the reactive acylphosphate reaction intermediate from hydrolysis (Figure 1B) and orders the Ala-rich loop to form the binding site for either Gly or β-Ala (Figure 4B). Functionally, these conformational changes provide a cooperative linkage through the reaction cycle as binding of one substrate enhances binding of the next substrate, as suggested by the interaction factors in the kinetic mechanism (Jez and Cahoon, 2004). Structural/functional analysis of hGS and GS also suggests that the dynamic nature of the active site is important for catalysis and substrate recognition.

The structural conservation between the active sites of hGS and GS implies a shared reaction mechanism (Herrera et al., 2007). In the first half of the hGS reaction, formation of the acylphosphate intermediate occurs by transfer of the γ-phosphate of ATP to γ-glutamylcysteine. For this step, the Mg^{2+} ions in the active site orient the phosphate group and Arg-153 likely stabilizes the transition state. In the second half of the reaction, nucleophilic attack of the β-Ala amino group on the acylphosphate intermediate releases phosphate and yields hGSH. Positioning of Arg-153 and the Mg^{2+} bound by Glu-169 and Asn-171 would stabilize the transition state with the Ala-rich loop and Arg-475 orienting β-Ala for attack on the reaction intermediate to yield hGSH.

The major difference between hGS and GS is substrate specificity for β-Ala and Gly, respectively. In each enzyme, residues in the Ala-rich loop contact the terminal residue of the tripeptide product (Figure 5). A Leu and Pro in the hGS from soybean and other legumes replaces the invariant double Ala sequence of the eukaryotic GS (Moran et al., 2000; Frendo et al., 2001; Iturbe-Ormaetxe et al., 2002; Skipsey et al., 2005). Structurally, the Ala-rich loop of hGS shifts relative to the same loop in GS to allow space for binding of the larger hGSH product and β-Ala substrate (Figure 5A). Site-directed mutagenesis of Leu-487 and Pro-488 demonstrates that changes at both positions are necessary to convert hGS ($k_{cat}/K_m^{\beta\text{-ala}} = 708\ M^{-1}\ s^{-1}$) to a GS with comparable catalytic efficiency ($k_{cat}/K_m^{gly} = 950\ M^{-1}\ s^{-1}$) (Table 3). Interestingly, the L487A/P488A mutant retains limited activity with β-Ala ($k_{cat}/K_m^{\beta\text{-ala}} = 77\ M^{-1}\ s^{-1}$). This suggests that additional changes in the Ala-rich loop, or more subtle allosteric mutations, may be required to completely shift substrate preference and product specificity. The mobility of active site features in both hGS and GS (i.e., the lid domain and Ala-rich loop) likely plays a role in determining the rate of catalysis and for allowing evolutionary changes in these enzymes.

In both hGS and GS, structuring of the lid domain and Ala-rich loop appears linked to binding of ATP and the terminal substrate (i.e., β-Ala or Gly). Although the rate constants for each step in the catalytic cycle of either enzyme are unknown, the crystal structures of these enzymes suggest that dynamic active site structures may limit catalysis and explain the different turnover rates of GS (k_{cat} ~12 s^{-1}) and hGS (k_{cat} ~2 s^{-1}) (Table 1) (Gunasekaran et al., 2003; Tokuriki and Tawfik, 2009). Based on these results, it is possible that the nucleophilic attack of the terminal substrate is a limiting step in the reaction mechanism. Presumably, GS is a highly evolved enzyme in eukaryotes

because of the central role that glutathione plays in regulating intracellular redox state (Meister, 1995). By contrast, hGS likely evolved by gene duplication and subsequent mutation (Tokuriki and Tawfik, 2009), and additional sequence changes in the lid domain and/or Ala-rich loop may be needed to optimize interactions with substrates and the movement of active site features.

Active site loops are central in the evolution of enzyme functionality (Todd et al., 1999; Penning and Jez, 2001; Gunasekaran et al., 2003; Tokuriki and Tawfik, 2009). The flexible and mutable nature of loops allows for the sampling of the new sequences and localized structures that generate shifts in substrate specificity or new catalytic activity. Frendo et al. (2001) originally proposed that legumes evolved hGS from gene duplication of GS after the divergence of the order Fabales, which includes the legumes, from other flowering plants. Our results suggest a molecular mechanism underpinning the evolution of hGS from GS. Although hGS retains the γEC and ATP binding sites and maintains the positioning of catalytically essential Arg residues (Arg-153 and Arg-475) and key Mg^{2+} ions, two changes in the Ala-rich loop are sufficient to alter substrate specificity.

While this work helps illuminate the molecular basis for hGS evolution from an ancestral GS, many questions remain as to the role hGSH and other GS analogs in plants. Although the interplay between genomes, protein function, and a plant's environment shapes the evolution of new metabolism, it is unclear why legumes required evolution of hGS and hGSH production in nodules. Aside from the shared localization of hGS in nodules (Moran et al., 2000; Frendo et al., 2001; Iturbe-Ormaetxe et al., 2002; Skipsey et al., 2005), there appears to be no correlation between the presence of hGS in a legume species and the position of that species in the legume phylogeny (Wojciechowski et al., 2004). Nonetheless, given the conservation of hGS in the legumes examined so far, it seems likely that environmental factors, such as nodulation and/or habitat, contributed to the diversification of GSH metabolism. In addition, as suggested by the presence of Ser- and Glu-containing GSH analogs in other plants (Rauser et al., 1986; Klapheck et al., 1994; Meuwly et al., 1995), the adaptation of GSH biosynthesis for production of specialized tripeptides in response to environmental stresses may be more widespread. Continued genomic and biochemical explorations of legumes, and other plants, promise new insights on how these plants evolved more specialized environmental response systems.

METHODS

Materials

All oligonucleotides were synthesized by Integrated DNA Technologies. Ni^{2+}-nitrilotriacetic acid (NTA) was from Qiagen. Benzamidine-sepharose and the HiLoad 26/60 Superdex-200 FPLC column were purchased from GE/Amersham Health Sciences. The QuikChange site-directed mutagenesis kit was from Stratagene. hGSH was from Bachem. All other reagents were of ACS grade or better and were purchased from Sigma-Aldrich.

Protein Expression, Purification, and Mutagenesis

Soybean (*Glycine max*) hGS was PCR-amplified from a soybean seed cDNA library using 5′-dTTTCCATGGC**ATG**GCTCAACCTTTGACC-ACC-3′ as the forward primer (the *Nco*I site is underlined, and the start codon is in bold) and 5′-dTTTGCGGCCGC**TCA**AGTTAGGTATACAG-TATCTACCAC-3′ as the reverse primer (the *Not*I site is underlined, and the stop codon is in bold). The resulting PCR product was digested with *Nco*I and *Not*I and then subcloned into pHIS8 (Jez et al., 2000) for expression of an N-terminally octahistidine-tagged protein. Automated nucleotide sequencing confirmed the fidelity of the bacterial expression construct (Washington University Sequencing Facility).

Transformed *Escherichia coli* BL21(DE3) cells were grown at 37°C in Terrific broth containing 50 μg mL^{-1} kanamycin until A_{600} ~0.8. After induction with 1 mM isopropyl 1-thio-β-D-galactopyranoside, the cultures were grown at 20°C for 4 to 8 h. Cells were pelleted by centrifugation (10,000*g*; 10 min) and resuspended in 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 20 mM imidazole, 5 mM $MgCl_2$, 10% (v/v) glycerol, and 1% (v/v) Tween 20. Sonication was used to lyse cells. Following centrifugation (45,000*g*; 45 min), the supernatant was passed through a Ni^{2+}-NTA column. The column was then washed with the same buffer minus Tween 20. His-tagged protein was eluted with 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 250 mM imidazole, 5 mM $MgCl_2$, and 10% (v/v) glycerol. Incubation with thrombin (1/1000th the amount of hGS by weight) during overnight dialysis at 4°C against wash buffer removed the His tag. Dialyzed protein was reloaded on a mixed benzamidine-sepharose/Ni^{2+}-NTA column. The flow-through of this step was loaded onto a HiLoad 26/60 Superdex-200 FPLC column equilibrated with 25 mM HEPES, pH 7.5, 5 mM $MgCl_2$, and 100 mM NaCl. Fractions containing purified protein were pooled, concentrated to 10 to 12 mg mL^{-1}, and stored at −80°C. Protein concentration was determined by the Bradford method (Protein Assay; Bio-Rad) with BSA as standard.

Site-directed mutants of hGS (L487A, P488A, and L487A/P488A) were generated using oligonucleotides containing the desired mutations (see Supplemental Table 1 online) and the QuikChange PCR method with the pHIS8-hGS vector as template. Introduction of the desired mutation was confirmed by sequencing of the constructs. Expression and purification of each mutant protein was performed as described for the wild-type protein.

Enzyme Assays

The activity of hGS was determined spectrophotometrically at 25°C by measuring the rate of formation of ADP using a coupled assay with pyruvate kinase and lactate dehydrogenase. A standard reaction mixture (0.5 mL) contained 100 mM HEPES, pH 7.5, 150 mM NaCl, 20 mM $MgCl_2$, 2.5 mM γEC, 10 mM β-Ala (or Gly), 2.5 mM disodium ATP, 2 mM sodium phosphoenolpyruvate, 0.2 mM NADH, 5 units of type III rabbit muscle pyruvate kinase, and 10 units of type II rabbit muscle lactate dehydrogenase. The rate of decrease in A_{340} (Σ = 6270 M^{-1} cm^{-1}) was observed using a Beckman DU800 UV/Vis spectrophotometer. Steady state kinetic parameters were determined by initial velocity experiments in which concentrations for two substrates were fixed at saturating levels and the third substrate concentration varied (0.2 to 10 times the K_m value). Untransformed data was fit to the Michaelis-Menten equation, $v = k_{cat}[S]/(K_m + [S])$, using Kaleidagraph (Synergy Software).

Protein Crystallization and Structure Determination

Crystals of hGS were obtained by the vapor diffusion method in 4-μL hanging drops of a 1:1 mixture of protein and crystallization buffer (20% PEG3000, 0.1 M MOPSO, pH 7, and 0.2 M $MgSO_4$) at 4°C over a 0.5-mL reservoir. For cocrystallization with ligands, either 5 mM γEC or 2.5 mM ADP and 5 mM hGSH was added to the protein before crystallization. All crystals were stabilized in cryoprotectant (crystallization solution plus ligands with 15% [v/v] glycerol) before flash freezing in liquid nitrogen. Data collection (100K) was performed at the Stanford Synchrotron

Radiation Facility (SSRL) on monochromatic beamline 9-1. Diffraction data was integrated and reduced using XDS (Kabsch, 1993) and scaled with XSCALE (Kabsch, 1993). The structure of closed-form hGS in complex with ADP and hGSH was solved by molecular replacement performed with PHASER (McCoy et al., 2007) using a homology model of the soybean enzyme generated with SWISS-MODEL (Kopp and Schwede, 2003) from the structure of human GS (PDB: 2HGS; Polekhina et al., 1999). Model building was performed in O (Jones et al., 1993), and all refinements were performed with REFMAC (Murshudov et al., 1997). Waters were added using ARP (Lamzin and Wilson, 1993). Quality of the model was evaluated using PROCHECK (Laskowski et al., 1993). Structures of the open form hGS and open form hGS in complex with γEC were solved by molecular replacement using the final closed form hGS structure. Modeling building, refinement, and assessment were performed as above. Crystal parameters, data collection statistics, and refinement statistics for the three structures are summarized in Table 2. Atomic coordinates and structure factors have been deposited in the Protein Data Bank (www.rcsb.org). All structural figures were generated with PyMol (http://www.pymol.org).

Accession Numbers

Sequence data from this article can be found in the GenBank/EMBL database under the following accession numbers: soybean hGS (accession CAB91078), human GS (PDB: 2HGS; accession NP_000169), *Saccharomyces cerevisiae* GS (accession CAA74136), and *Arabidopsis thaliana* GS (U22359). Coordinates and structure factors for the soybean hGS apoenzyme (PDB: 3KAJ), γ-glutamylcysteine complex (PDB: 3KAK), and the hGSH•ADP complex (PDB: 3KAL) have been deposited in the RCSB Protein Data Bank.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure 1. Protein Expression and Purification Analysis.

Supplemental Table 1. Oligonucleotide Primers Used for Site-Directed Mutagenesis.

ACKNOWLEDGMENTS

This work was funded by an American Chemical Society Petroleum Research Fund grant (ACS-PRF-43012-AC4) and a National Science Foundation grant (MCB-0824492) to J.M.J. K.A.J.A was a recipient of an ACS-PRF Summer Underrepresented Minority Research Fellowship. The SSRL Biotechnology Program is supported by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the Department of Energy, Office of Biological and Environmental Research.

Received September 5, 2009; revised October 9, 2009; accepted November 5, 2009; published November 30, 2009.

REFERENCES

Frendo, P., Harrison, J., Norman, C., Hernández-Jiménez, M.J., Van de Sype, G., Gilabert, A., and Puppo, A. (2005). Glutathione and homoglutathione play a critical role in the nodulation process of *Medicago truncatula.* Mol. Plant Microbe Interact. **18:** 254–259.

Frendo, P., Jiménez, M.J., Mathieu, C., Duret, L., Gallesi, D., Van de Sype, G., Hérouart, D., and Puppo, A. (2001). A *Medicago truncatula* homoglutathione synthetase is derived from glutathione synthetase by gene duplication. Plant Physiol. **126:** 1706–1715.

Fyfe, P.K., Oza, S.L., Fairlamb, A.H., and Hunter, W.N. (2008). Leishmania trypanothione synthetase-amidase structure reveals a basis for regulation of conflicting synthetic and hydrolytic activities. J. Biol. Chem. **283:** 17672–17680.

Galperin, M.Y., and Koonin, E.V. (1997). A diverse superfamily of enzymes with ATP-dependent carboxylate-amine/thiol ligase activity. Protein Sci. **6:** 2639–2643.

Gogos, A., and Shapiro, L. (2002). Large conformational changes in the catalytic cycle of glutathione synthase. Structure **10:** 1669–1676.

Gunasekaran, K., Ma, B., and Nussinov, R. (2003). Triggering loops and enzyme function: identification of loops that trigger and modulate movements. J. Mol. Biol. **332:** 143–159.

Herrera, K., Cahoon, R.E., Kumaran, S., and Jez, J.M. (2007). Reaction mechanism of glutathione synthetase from *Arabidopsis thaliana*: Site-directed mutagenesis of active site residues. J. Biol. Chem. **282:** 17157–17165.

Hicks, L.M., Cahoon, R.E., Bonner, E.R., Rivard, R.S., Sheffield, J., and Jez, J.M. (2007). Thiol-based regulation of redox-active glutamate-cysteine ligase from *Arabidopsis thaliana.* Plant Cell **19:** 2653–2661.

Iturbe-Ormaetxe, I., Heras, B., Matamoros, M.A., Ramos, J., Moran, J.F., and Becana, M. (2002). Cloning and functional characterization of a homoglutathione synthetase from pea nodules. Physiol. Plant. **115:** 69–73.

Jez, J.M., and Cahoon, R.E. (2004). Kinetic mechanism of glutathione synthetase from *Arabidopsis thaliana.* J. Biol. Chem. **279:** 42726–42731.

Jez, J.M., Cahoon, R.E., and Chen, S. (2004). *Arabidopsis thaliana* glutamate-cysteine ligase: Functional properties, kinetic mechanism, and regulation of activity. J. Biol. Chem. **279:** 33463–33470.

Jez, J.M., Ferrer, J.L., Bowman, M.E., Dixon, R.A., and Noel, J.P. (2000). Dissection of malonyl-coenzyme A decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase. Biochemistry **39:** 890–902.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1993). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. D Biol. Crystallogr. **49:** 148–157.

Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Cryst. **26:** 795–800.

Klapheck, S., Fliegner, W., and Zimmer, I. (1994). Hydroxymethyl-phytochelatins [(gamma-glutamylcysteine)$_n$-serine] are metal-induced peptides of the *Poaceae.* Plant Physiol. **104:** 1325–1332.

Klapheck, S., Schlunz, S., and Bergmann, L. (1995). Synthesis of phytochelatins and homo-phytochelatins in *Pisum sativum* L. Plant Physiol. **107:** 515–521.

Kopp, J., and Schwede, T. (2003). The SWISS-MODEL repository: New features and functionalities. Nucleic Acids Res. **34:** D315–D318.

Lamzin, V.S., and Wilson, K.S. (1993). Automated refinement of protein models. Acta Crystallogr. D Biol. Crystallogr. **49:** 129–147.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. J. Appl. Cryst. **26:** 283–291.

Loscos, J., Matamoros, M.A., and Becana, M. (2008). Ascorbate and homoglutathione metabolism in common bean nodules under stress conditions and during natural senescence. Plant Physiol.**146:** 1282–1292.

Matamoros, M.A., Moran, J.F., Iturbe-Ormaetxe, I., Rubio, M.C., and Becana, M. (1999). Glutathione and homoglutathione synthesis in legume root nodules. Plant Physiol. **121:** 879–888.

Matamoros, M.A., Clemente, M.R., Sato, S., Asamizu, E., Tabata, S., Ramos, J., Moran, J.F., Stiller, J., Gresshoff, P.M., and Becana, M. (2003). Molecular analysis of the pathway for the synthesis of thiol tripeptides in the model legume *Lotus japonicus*. Mol. Plant Microbe Interact. **16:** 1039–1046.

May, M.J., Vernoux, T., Leaver, C., Van Montagu, M., and Inze, D. (1998). Glutathione homeostasis in plants: Implications for environmental sensing and plant development. J. Exp. Bot. **49:** 649–667.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Cryst. **40:** 658–674.

Meister, A. (1995). Glutathione metabolism. Methods Enzymol. **252:** 26–30.

Meuwly, P., Thibault, P., Schwan, A.L., and Rauser, W.E. (1995). Three families of thiol peptides are induced by cadmium in maize. Plant J. **7:** 391–400.

Moran, J.F., Iturbe-Ormaetxe, I., Matamoros, M.A., Rubio, M.C., Clemente, M.R., Brewin, N.J., and Becana, M. (2000). Glutathione and homoglutathione synthetases of legume nodules: Cloning, expression, and subcellular localization. Plant Physiol. **124:** 1381–1392.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. **53:** 240–255.

Noctor, G., and Foyer, C.H. (1998). Ascorbate and glutathione: Keeping active oxygen under control. Annu. Rev. Plant Physiol. Plant Mol. Biol. **49:** 249–279.

Pai, C.H., Chiang, B.Y., Ko, T.P., Chou, C.C., Chong, C.M., Yen, F.J., Chen, S., Coward, J.K., Wang, A.H., and Lin, C.H. (2006). Dual binding sites for translocation catalysis by *Escherichia coli* glutathionylspermidine synthetase. EMBO J. **25:** 5970–5982.

Penning, T.M., and Jez, J.M. (2001). Enzyme redesign. Chem. Rev. **101:** 3027–3046.

Polekhina, G., Board, P.G., Gali, R.R., Rossjohn, J., and Parker, M.W. (1999). Molecular basis of glutathione synthetase deficiency and a gene permutation event. EMBO J. **18:** 3204–3213.

Rauser, W.E., Hunziker, P.E., and Kagi, J.H.R. (1986). Reverse-phase high-performance chromatography of cadmium-binding proteins from the grass *Agrostis gigantea*. Plant Sci. **45:** 105–109.

Rouhier, N., Lemaire, S.D., and Jacquot, J.P. (2008). The role of glutathione in photosynthetic organisms: Emerging functions for glutaredoxins and glutathionylation. Annu. Rev. Plant Biol. **59:** 143–166.

Skipsey, M., Davis, B.G., and Edwards, R. (2005). Diversification in substrate usage by glutathione synthetases from soya bean (*Glycine max*), wheat (*Triticum aestivum*) and maize (*Zea mays*). Biochem. J. **391:** 567–574.

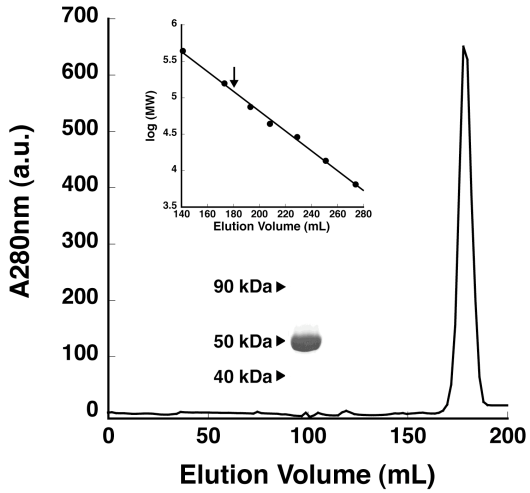
Todd, A.E., Orengo, C.A., and Thornton, J.M. (1999). Evolution of protein function, from a structural perspective. Curr. Opin. Chem. Biol. **3:** 548–556.

Tokuriki, N., and Tawfik, D.S. (2009). Protein dynamism and evolvability. Science **324:** 203–207.

Wojciechowski, M.F., Lavin, M., and Sanderson, M.J. (2004). A phylogeny of legumes (leguminosae) based on analysis of the plastid matK gene resolves many well-supported subclades within the family. Am. J. Bot. **91:** 1846–1862.

Yamaguchi, H., Kato, H., Hata, Y., Nishioka, T., Kimura, A., Oda, J., and Katsube, Y. (1993). Three-dimensional structure of the glutathione synthetase from *Escherichia coli* B at 2.0 Å resolution. J. Mol. Biol. **229:** 1083–1100.

Supplemental Data. Galant et al. (2009). Structural Basis for Evolution of Product Diversity in Soybean Glutathione Biosynthesis. Plant Cell. 10.1105/tpc.109.071183



Supplemental Figure 1. Protein expression and purification analysis. Size-exclusion chromatography of GmhGS. Purified hGS was chromatographed on a Superdex-200 26/60 FPLC column with 25 mM Hepes (pH 7.5), 5 mM $MgCl_2$, and 100 mM NaCl. The inset graph shows the molecular weight calibration of the column. The following standards were used: ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa), ovalbumin (44 kDa), carbonic anhydrase (24 kDa), ribonuclease A (13.7 kDa), and aprofinin (6.5 kDa). The arrow represents the elution volume of soybean hGS. The inset SDS-PAGE shows the purified protein stained with Coomassie Blue. Arrows correspond to molecular weight markers as indicated.

Supplemental Table 1. Oligonucleotide primers used for site-directed mutagenesis.

L487A	5'-dCTTATGAAGGAGGAGTT**GCG**CCTGGTTTTGGAGTGGTAG-3'
P488A	5'-dCTTATGAAGGAGGAGTTTTG**GCT**GGTTTTGGAGTGGTAG-3'
L487A/P488A	5-dCTTATGAAGGAGGAGTT**GCGGCT**GGTTTTGGAGTGGTAG-3'

For PCR-based mutagenesis, complementary sense and antisense primers were used. The table only shows the sense sequence. Codons encoding the mutations are in bold type.

CHAPTER 3

REDOX-REGULATORY MECHANISMS INDUCED BY OXIDATIVE STRESS IN *BRASSICA JUNCEA* ROOTS MONITORED BY 2-DE PROTEOMICS

PREFACE

In this chapter, I describe the application of an NEM and IAF-based thiol labeling strategy to protein extracts from *B. juncea* roots that had been treated with either 1mM H_2O_2 or 50μm BSO. The framework for this series of experiments arose from a 2008 NSF grant application submitted by Dr. Joseph Jez and Dr. Leslie Hicks; that grant was in turn based upon earlier experiments which determined that GCL utilizes intramolecular disulfide bonds as a means of redox regulation [Jez et al., 2004; Hicks et al., 2007]. Because at the time only a handful of proteins (only one of which was from plants) which utilized thiol-based regulatory switches had been identified, the grant, among other things, proposed the use of a 2D-SDS-PAGE/LC-MS/MS methodology for identifying additional candidate proteins. The methodology itself was not entirely new - variants had previously been used for identifying thiol-containing proteins in mammals, yeast, and bacteria [Yang et al., 2007; Le Moan et al., 2006; Dosanjh et al., 2005]. However, in plants to date use of the technique had been much more limited, and primarily focused on the identification of novel thioredoxin targets [Yano et al., 2002; Lee et al., 2004; Yano and Kuroda, 2005]. Instead of using thioredoxin to reduce protein extracts, in our approach we opted to use the general reductant DTT. The advantage of this choice was that we would be able to identify target proteins that are reduced by other "doxins" besides thioredoxin, or proteins for which the physiological reductant is unknown.

Because dataset briefs published in the journal Proteomics are limited to ~2500 words, the body of this chapter contains only an abbreviated description of the protein

extraction and labeling methodology employed. Thus, I have included a more detailed version at the end of the chapter.

Author Contributions: JMJ and LMH designed research; AG performed research; SA, AG, JMJ, and LMH analyzed data; AG, JMJ, LHM, and SA wrote the paper.

Preface References

Dosanjh, N.S., Rawat, M., Chung, J.H., Av-Gay, Y. (2005). Thiol specific oxidative stress response in Mycobacteria. FEMS Microbiology Letters 249: 87-94.

Hicks, L.M., Cahoon, R.E., Bonner, E.R., Rivard, R.S., Sheffield, J., Jez, J.M. (2007). Thiol-based regulation of redox-active glutamate-cysteine ligase from *Arabidopsis thaliana*. Plant Cell 19: 2653-2661.

Jez, J.M., Cahoon, R.E., Chen, S. (2004). *Arabidopsis thaliana* Glutamate-Cysteine Ligase. Journal of Biological Chemistry. 279: 33463-33470.

Lee, K., Lee, J., Kim, Y., Bae, D., Kang, K.Y., Yoon, S.C., Lim, D. (2004). Defining the plant disulfide proteome. Electrophoresis 25: 532-541.

Le Moan, N., Clement, G., Le Maout, S., Tacnet, F., Toledano, M.B. (2006). The Saccharomyces cerevisaiae proteome of oxidized protein thiols - Contrasted functions for the thioredoxin and glutathione pathways. Journal of Biological Chemistry 281: 10420-10430.

Yang, Y., Song, Y., Loscalzo, J. (2007). Regulation of the protein disulfide proteome by mitochondria in mammalian cells. PNAS 104: 10813-10817.

Yano, H., Kuroda, M. (2005). Disulfide proteome yields a detailed understanding of redox regulations: A model study of thioredoxin-linked reactions in seed germination. Proteomics 6: 294-300.

Yano, H., Kuroda, S., Buchanan, B.B. (2002). Disulfide proteome in the analysis of protein function and structure. Proteomics 2: 1090-1096.

 DOI 10.1002/pmic.201000450 *Proteomics* 2011, *11*, 1346–1350

DATASET BRIEF

Redox-regulatory mechanisms induced by oxidative stress in *Brassica juncea* roots monitored by 2-DE proteomics

Sophie Alvarez[1]*, *Ashley Galant*[2]*, *Joseph M. Jez*[2] and *Leslie M. Hicks*[1]

[1] Donald Danforth Plant Science Center, St. Louis, MO, USA
[2] Department of Biology, Washington University, St. Louis, MO, USA

ROS, including hydrogen peroxide (H_2O_2), can serve as cellular signaling molecules following oxidative stress. Analysis of the redox state of proteins in *Brassica juncea* roots by 2-DE proteomics following treatment with either exogenous H_2O_2 or buthionine sulfoximine, which depletes glutathione to cause accumulation of endogenous H_2O_2, led to the identification of different sets of proteins. These data suggest that exogenous and endogenous oxidative stresses trigger specialized responses.

Received: July 20, 2010
Revised: December 1, 2010
Accepted: December 20, 2010

Keywords:
2-DE / Buthionine sulfoximine / Hydrogen peroxide / Oxidative stress / Plant proteomics

ROS produced endogenously in response to environmental changes serve as signaling molecules in communications within and between cells [1]. Among ROS, hydrogen peroxide (H_2O_2) causes reversible and irreversible redox modifications to proteins during oxidative stress [2, 3]. Although many H_2O_2-induced protein modifications result in irreversible oxidative damage, reversible modification of cysteines (i.e., oxidation of thiols to disulfide bonds, glutathionylation, or *S*-nitrosylation) is an important mechanism for regulating protein function. To balance between deleterious effects and oxidative signaling, intracellular H_2O_2 levels are controlled by mechanisms, such as the glutathione-ascorbate system, that maintain concentrations of key reducing molecules [1]. As a consequence, H_2O_2 has long been used to elicit oxidative stress responses to study redox mechanisms and provide insight into the molecular physiology of adaptive responses.

Correspondence: Dr. Leslie M. Hicks, Proteomics and Mass Spectrometry Facility, Donald Danforth Plant Science Center, St. Louis, MO 63132, USA
E-mail: lhicks@danforthcenter.org
Fax: +1-314-587-1324

Abbreviations: BSO, buthionine sulfoximine; **DHAR,** dehydroascorbate reductase; **H_2O_2,** hydrogen peroxide; **IAF,** 5-iodoacetamidofluorescein; **TPI,** triose phosphate isomerase; **TRXh,** H-type thioredoxin

In this study, we examine the changes in the redox proteome of *Brassica juncea* (Indian mustard) roots using specific labeling of cysteines by 5-iodoacetamidofluorescein (IAF) in response to exogenous and endogenous H_2O_2-induced oxidative stresses. Application of H_2O_2 to plant roots provides an exogenous stress and application of buthionine sulfoximine (BSO), which depletes glutathione, produces an accumulation of endogenous H_2O_2 [4]. Largely different sets of proteins regulated by H_2O_2 were identified for each treatment at the redox and abundance levels. Interestingly, proteins involved in similar biological processes, such as the brassinosteroid signaling pathway, were differentially regulated by each H_2O_2 source in *B. juncea* roots.

Wild-type *B. juncea* seeds were germinated in a growth chamber at 22°C, 200 μmol/m^2 light intensity, 50% relative humidity, during a 16-h light/8-h dark cycle. After 3 wk, seedlings were transplanted to 3.8 L pots in the greenhouse (same light/dark cycle). After 6 wk, plants were treated with 2 L of distilled water, 1 mM H_2O_2, or 50 μM BSO, positioned to allow rapid draining, and after 1 h treated again with 1 L of solution. Following draining (2 h), roots were washed to remove soil, flash-frozen in liquid nitrogen, and stored at −80°C. Concentrations of H_2O_2 and BSO were chosen

*These authors have contributed equally to this study.

Colour Online: See the article online to view Fig. 1 in colour.

based on a previous experiment, showing that these compounds alter oxidation state of a redox-sensitive protein in planta [5].

For each treatment, three biological replicate samples from three different plants were obtained for processing. Root tissue (~800 mg FW) was ground and suspended in extraction buffer (100 mM Tris-HCl, pH 8.0; 100 mM *N*-ethylmaleimide (NEM); 1% CHAPS; 1% protease inhibitor cocktail (Sigma, St. Louis, USA) to 200 mg/mL for protein extraction and alkylation of free sulfhydryl groups (Fig. 1, step 1). Samples were centrifuged and the soluble protein fraction was removed, precipitated with methanol 3×, resuspended in 150 µL of reduction buffer (50 mM Tris-HCl, pH 8.0, 7 M urea, 2 M thiourea, 50 mM DTT), and incubated for 15 min (25°C) to reduce disulfide bonds (Fig. 1, step 2). Proteins were next precipitated with methanol 3×, resuspended in 150 µL of labeling buffer (40 mM HEPES, pH 7.5; 50 mM NaCl; 200 µM IAF), and incubated for 10 min (25°C) for the labeling reaction (Fig. 1, step 3). Proteins were precipitated with methanol 3× and resuspended in destreak rehydration buffer (GE Healthcare, Waukesha, WI, USA). Protein concentrations were determined by CBX protein assay (G-Biosciences, St. Louis, USA).

Extracted protein (200 µg) was loaded onto pH strips 4-7 (Bio-Rad, Hercules, CA, USA) and 2-DE performed as described previously [6]. Gels were imaged with a Typhoon 9410 (GE Healthcare) to detect IAF-labeled proteins (λ_{ex} = 488 nm and λ_{em} = 526 nm). Gels were then stained with Sypro Ruby and imaged to detect total proteins (λ_{ex} = 457 nm and λ_{em} = 610 nm). Image analysis, including gel alignment, spot averaging and normalization, and multivariate statistics, employed SameSpots software (Nonlinear Dynamics, Durham, NC, USA) to determine which protein spots changed in protein abundance and oxidation in response to H_2O_2 and BSO treatments relative to controls. Means and standard deviations were calculated from three replicates and compared between control and treatments using ANOVA. Spots with a *p*-value of <0.05 were picked for protein identification via trypsin digestion and LC-MS/MS as described previously [6]. The peptide tandem mass spectra were processed using Analyst QS v1.1 (AB Sciex, Foster City, CA, USA) and searched against the NCBInr database (July 2010, 11 368 323 sequences) using an in-house version of MASCOT v2.20 (Matrix Science, Boston, MA, USA) with the following parameters: tryptic peptides with ≤ 1 missed cleavage site; precursor and MS/MS fragment ion mass tolerances of 0.8 and 0.8 Da, respectively; variable carbamidomethylation and fluoresceination of cysteine; and variable oxidation of methionine. The data were filtered using Scaffold 3 (Proteome Software, Portland, OR, USA). Positive identification criteria were ≥ 2 peptide sequences, protein probability of 99.9%, and peptide probability of 80%.

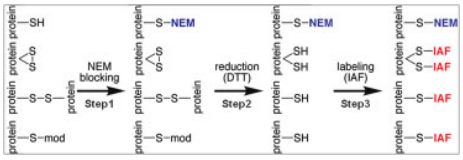


Figure 1. Redox Proteome Labeling Approach. Proteins with free thiols (–SH), disulfide bonds (–S–S–), or modified cysteines (–S–mod) are incubated with *N*-ethylmaleimide to block free sulfhydryl groups. Oxidized thiols are reduced with DTT. The resulting free thiols are fluorescently labeled with IAF and the proteins separated by 2-DE and identified by LC-MS/MS.

A total of 59 and 50 spots showed significant changes ($p<0.05$) in redox-state after H_2O_2 and BSO treatments, respectively, and 27 and 40 spots differed significantly in total protein abundance for the H_2O_2 and BSO treatments, respectively (Table 1 and Supporting Information Table 1). For the four comparisons (Table 1), the *q*-values ranged from 12 to 37% and from 55 to 61% of the spots were confidently identified by LC-MS/MS using the criteria described. Of the 103 spots confidently identified, only the 52 spots containing a single protein were used for further analysis of redox and abundance changes. These proteins were categorized according to their biological process (Supporting Information Table 2). The 29 proteins that change in redox state after H_2O_2 and BSO treatments are most represented in amino acid biosynthesis, redox homeostasis, and glycolysis (Fig. 2). The two main biological processes in which the 23 proteins change in abundance are redox homeostasis and defense response (Supporting Information Table 2 and Supporting Information Fig. 1). Images from IAF labeling and Sypro staining were overlaid to identify possible co-regulation of redox and protein abundance changes (Supporting Information Fig. 2). A significant number of the protein spots do not overlap and none of the protein spots showing changes in abundance overlapped with the ones identified as redox regulated. Thus, specificity of the post-translational redox change is largely independent of changes in total protein abundance. Only the redox changes will be discussed further.

Multiple proteins, such as dehydroascorbate reductase (DHAR), glutathione-*S*-transferases (GST), and H-type thioredoxins (TRXh), involved in redox homeostasis were identified as changed in oxidation state following each treatment. In response to H_2O_2 application, DHAR, which is essential for the glutathione–ascorbate cycle, showed decreased IAF spot intensity, indicating greater reduction of the enzyme compared with the control (Fig. 2). Spinach DHAR contains a thiol group required for reduction of oxidized glutathione [7]. Thus, a change in the redox state of DHAR may increase the regeneration of ascorbate from dehydroascorbate and enhance detoxification of H_2O_2. Two GST isoforms showed increased oxidation in response to BSO and H_2O_2 and one isoform was more reduced only following BSO treatment (Fig. 2). GSTs catalyze the conjugation of reduced glutathione to sulfhydryl groups of

Table 1. Total number of spots differentially expressed and oxidized ($p<0.05$) in response to H_2O_2 and BSO

	H_2O_2/Sypro	H_2O_2/IAF	BSO/Sypro	BSO/IAF	Total
Total number of spots detected	235	243	288	250	–
Number of spots differentially expressed or oxidized	27 (37%)	59 (12%)	40 (29%)	50 (16%)	176
Number of spots identified as one protein ID	11	17	12	12	52
Number of spots identified with multiple proteins IDs	4	19	12	16	51
Total number of spots identified	15 (55%)	36 (61%)	24 (60%)	28 (56%)	103

The *q*-values for each experiment for the number of spots significantly different are indicated in parentheses.

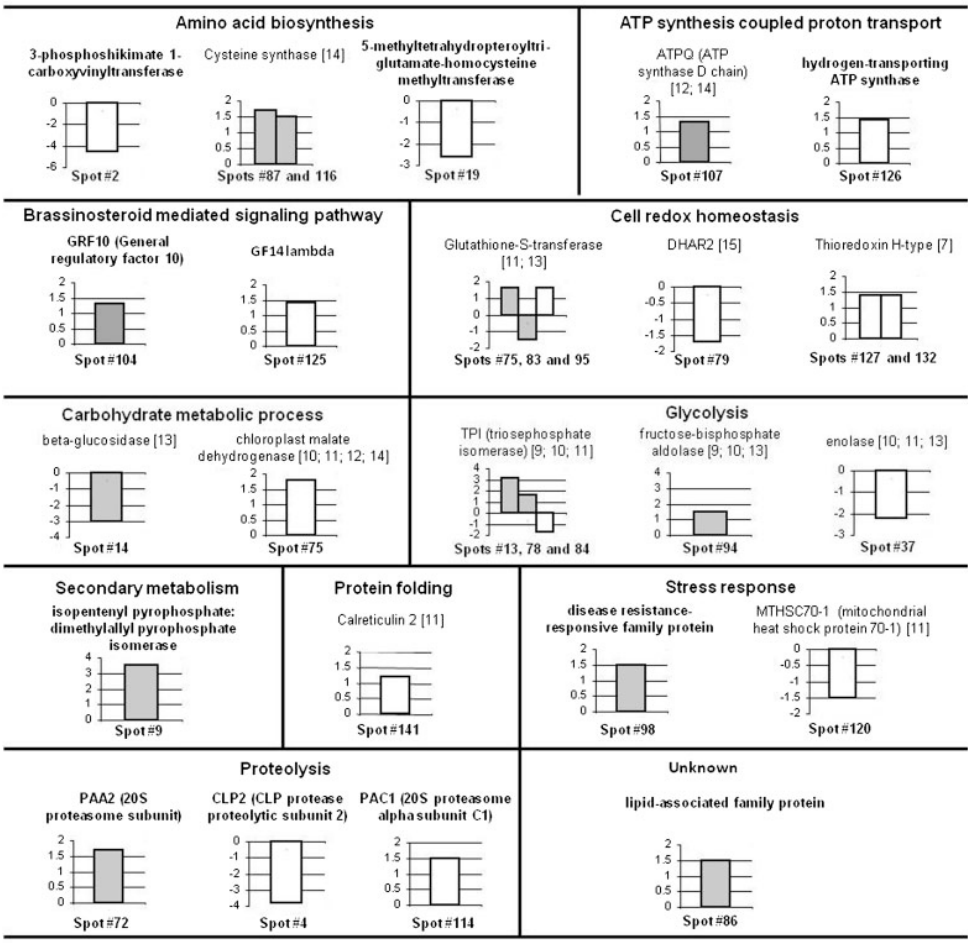


Figure 2. Differential Protein Redox Changes. Proteins with a change in redox state were grouped by biological function with fold change in oxidation shown in the bar graph. White and grey bars correspond to H_2O_2 and BSO treatments, respectively. Proteins described for the first time as redox-altered proteins are indicated in bold. For the proteins previously described as either disulfide-bonded proteins or *S*-thiolated protein, the reference numbers are indicated.

proteins and small molecules; in the case where BSO inhibits the biosynthesis of glutathione, GST activity also likely decreases in the absence of substrate. The increase in oxidized state of GST is mainly due to the oxidative conditions from H_2O_2 accumulation. Two TRXh isoforms also showed greater oxidation following H_2O_2 application, indicating increased disulfide formation and/or thiol modification (Fig. 2). TRXhs reduce disulfide bonds in a range of proteins to provide a mechanism for regulating redox imbalance [8]. In poplar, mitochondrial TRXh2 contains a glutathionylation site that modifies the redox potential of TRXh2 to decrease its activity [7], but in pea TRXh isoforms can differentially effect redox imbalance regulation [9]. Increased oxidation of TRXh in *B. juncea* may result from elevated demand to modulate H_2O_2 effects. Overall, exogenous H_2O_2 yields more changes on cellular antioxidant mechanisms such as the glutathione–ascorbate cycle and the thioredoxin system than BSO treatment.

Proteins in glycolysis, stress response, carbohydrate metabolism, and proteolysis also exhibited redox changes (Fig. 2). As relatively little information (as compared with mammalian systems) on redox regulation of proteins in plants is available, it is difficult to define the redox effect on the biological process according to the treatment since the reduction/oxidation of a particular protein can cause activation and/or repression of the protein activity [3]. Several proteins identified in this study as redox-sensitive are known targets of thioredoxins and/or glutathionylation (Fig. 2). For example, triose phosphate isomerase (TPI), a glycolytic enzyme, was first identified as a target for glutathionylation in *Arabidopsis* [10]. TPI requires glutathionylation for maintaining activity and oxidized glutathione inhibits the enzyme. TPI is also regulated by thioredoxin in the endosperm during germination of cereal grains [11] and *Medicago truncatula* seeds [12]. Here, several isoforms of TPI were identified as changing in redox state following BSO and H_2O_2 treatment. Two isoforms showed increased oxidation after BSO treatment, whereas one isoform showed greater reduction only following H_2O_2 treatment. Decreased glutathione levels after BSO treatment increases the oxidation state of cells and may trigger the specific oxidation of TPI to maintain energy production through the thioredoxin system. Although a specific modification may result from a treatment, we were not able to determine if the modification was either formation of a disulfide bridge or glutathionylation. BSO treatment also increased the oxidation of a second glycolytic enzyme, fructose-biphosphate aldolase. The previous studies demonstrate that this enzyme is glutathionylated in *Arabidopsis* [10] and is a thioredoxin target during germination of wheat grains [11]. On the contrary, enolase showed increased reduction in response to H_2O_2 (Fig. 2). Enolase is also a thioredoxin target and is redox regulated during germination of wheat grains and *M. truncatula* seeds [11, 12]. Additional proteins identified from carbohydrate metabolism and ATP-coupled proton transport are known to be redox regulated (Fig. 2) [13–16].

Potential new disulfide-containing proteins in amino acid synthesis and proteolytic processing were also identified in this study, including 3-phosphoshikimate 1-carboxyvinyltransferase, cobalamin-independent methionine synthase, the PAA2 20S proteasome subunit, the CLP protease proteolytic subunit 2, and 20S proteasome α-subunit C1. More interestingly, two 14-3-3 proteins involved in brassinosteroid signaling, general regulatory factor 10 (GRF10 or GFε) and GF14λ, were identified as increased in oxidation state in response to BSO and H_2O_2 treatments, respectively (Fig. 2) [17]. Protein phosphorylation mediates the interaction of 14-3-3 proteins with target proteins. Redox modification of 14-3-3 proteins may change protein conformation, thus impairing protein–protein interaction and inactivating signaling pathways. Brassinosteroids are plant hormones involved in a range of cellular and physiological processes including plant growth and tolerance to a variety of abiotic and biotic stresses [18, 19]. Brassinosteroids induce H_2O_2 in cucumber leaves and increase oxidative tolerance [20]. In this study, the application of H_2O_2 and the induction of endogenous H_2O_2 may have different effects on 14-3-3 proteins and possibly alter brassinosteroid signaling involved in the induction of oxidative stress tolerance.

In conclusion, several new oxidative stress redox-regulated proteins were identified using a specialized 2-DE proteomics approach. These results showed that specific redox and protein induction occurred when H_2O_2 was applied directly, including changes of specific protein isoforms, and that different mechanisms can be induced if redox regulation mechanisms, such as the glutathione–ascorbate cycle, are blocked to increase endogenous H_2O_2 levels. By resolving different protein isoforms either from the same gene family or from differential post-translational modifications, 2-DE proteomics has proven its utility to decipher the complexity of redox regulation mechanisms in plants. This approach is directly applicable to examine biologically relevant stress situations on agronomic crops, and could significantly impact the understanding of redox regulation both generally and specifically to facilitate crop improvement efforts.

Supporting data are accessible in the PRIDE database, login review33615, password hTXrNqWY, direct link http://www.ebi.ac.uk/pride/login.do.

This research was funded by a National Science Foundation grant (MCB-0824492) to J.M.J. and L.M.H. A.G. was supported by an American Society of Plant Biologists Pioneer Hi-Bred Graduate Research Fellowship.

The authors have declared no conflict of interest.

References

[1] Neill, S., Desikan, R., Hancock, J., Hydrogen peroxide signaling. *Curr. Opin. Plant Biol.* 2002, *5*, 388–395.

[2] Gilbert, H. F., Redox control of enzyme activities by thiol/disulfide exchange. *Meth. Enzym.* 1984, *107*, 330–351.

[3] Ziegler, D. M., Role of reversible oxidation-reduction of enzyme thiols-disulfides in metabolic regulation. *Annu. Rev. Biochem.* 1985, *54*, 305–329.

[4] Griffith, O. W., Meister, A., Potent and specific inhibition of glutathione synthesis by buthionine sulfoximine (S-n-butyl homocystein sulfoxime). *J. Biol. Chem.* 1979, *254*, 7558–7560.

[5] Hicks, L. M., Cahoon, R. E., Bonner, E. R., Rivard, R. S. et al., Thiol-based regulation of redox-active glutamate-cysteine ligase from *Arabidopsis thaliana*. *Plant Cell* 2007, *19*, 2653–2661.

[6] Alvarez, S., Berla, B. M., Sheffield, J., Cahoon, R. E. et al., Comprehensive analysis of the *Brassica juncea* root proteome in response to cadmium exposure by complementary proteomic approaches. *Proteomics* 2009, *9*, 2419–2431.

[7] Hossain, M. A., Asada, K., Purification of dehydroascorbate reductase from spinach and its characterization as a thiol enzyme. *Plant Cell Physiol.* 1984, *25*, 85–92.

[8] Gelhaye, E., Rouhier, N., Gerard, J., Jolivet, Y. et al., A specific form of thioredoxin h occurs in plant mitochondria and regulates the alternative oxidase. *Proc. Natl. Acad. Sci. USA* 2004, *101*, 14545–14550.

[9] Traverso, J. A., Vignols, F., Cazalis, R., Pulido, A. et al., PsTRXh1 and PsTRXh2 are both pea h-type thioredoxins with antagonistic behavior in redox imbalances. *Plant Physiol.* 2007, *143*, 300–311.

[10] Ito, H., Iwabuchi, M., Ogawa, K., The sugar-metabolic enzymes aldolase and triose-phosphate isomerase are targets of glutathionylation in Arabidopsis thaliana: detection using biotinylated glutathione. *Plant Cell Physiol.* 2003, *44*, 655–660.

[11] Wong, J. H., Balmer, Y., Cai, N., Tanaka, C. K. et al., Unraveling thioredoxin-linked metabolic processes of cereal starchy endosperm using proteomics. *FEBS Lett.* 2003, *547*, 151–156.

[12] Alkhalfioui, F., Renard, M., Vensel, W. H., Wong, J. et al., Thioredoxin-linked proteins are reduced during germination of *Medicago truncatula* seeds. *Plant Physiol.* 2007, *144*, 1559–1579.

[13] Alvarez, S., Wilson, G. H., Chen, S., Determination of in vivo disulfide-bonded proteins in Arabidopsis. *J. Chromatogr. B* 2009, *877*, 101–104.

[14] Alvarez, S., Zhu, M., Chen, S., Proteomics of Arabidopsis redox proteins in response to methyl jasmonate. *J. Proteomics* 2009, *73*, 30–40.

[15] Balmer, Y., Vensel, W. H., Tanaka, C. K., Hurkman, W. J. et al., Thioredoxin links redox to the regulation of fundamental processes of plant mitochondria. *Proc. Natl. Acad. Sci. USA* 2004, *101*, 2642–2647.

[16] Dixon, D. P., Skipsey, M., Grundy, N. M., Edwards, R., Stress-induced protein S-glutathionylation in Arabidopsis. *Plant Physiol.* 2005, *138*, 2233–2244.

[17] DeLille, J. M., Sehnke, P. C., Ferl, R. J., The Arabidopsis 14-3-3 family of signaling regulators. *Plant Physiol.* 2001, *126*, 35–38.

[18] Kagale, S., Divi, U. K., Krochko, J. E., Keller, W. A., Krishna, P., Brassinosteroid confers tolerance in *Arabidopsis thaliana* and *Brassica napus* to a range of abiotic stresses. *Planta* 2007, *225*, 353–364.

[19] Nakashita, H., Yasuda, M., Nitta, T., Asami, T. et al., Brassinosteroid functions in a broad range of disease resistance in tobacco and rice. *Plant J.* 2003, *33*, 887–898.

[20] Xia, X. J., Wang, Y. J., Zhou, Y. H., Tao, Y. et al., Reactive oxygen species are involved in brassinosteroid-induced stress tolerance in cucumber. *Plant Physiol.* 2009, *150*, 801–814.

PROTEOMICS

Supporting Information for Proteomics

DOI 10.1002/pmic.201000450

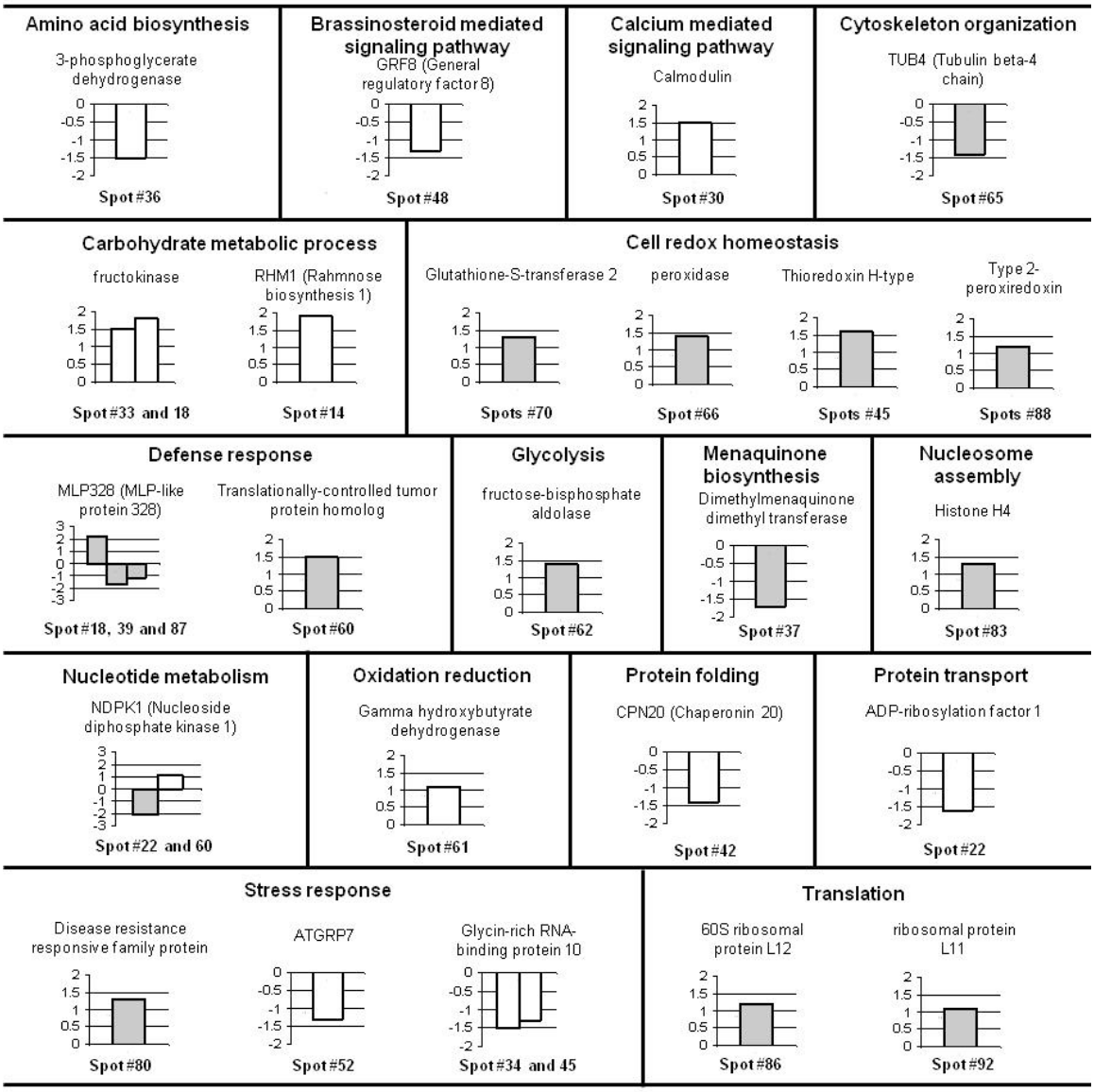
Sophie Alvarez, Ashley Galant, Joseph M. Jez and Leslie M. Hicks

Redox-regulatory mechanisms induced by oxidative stress in *Brassica juncea* roots monitored by 2-DE proteomics

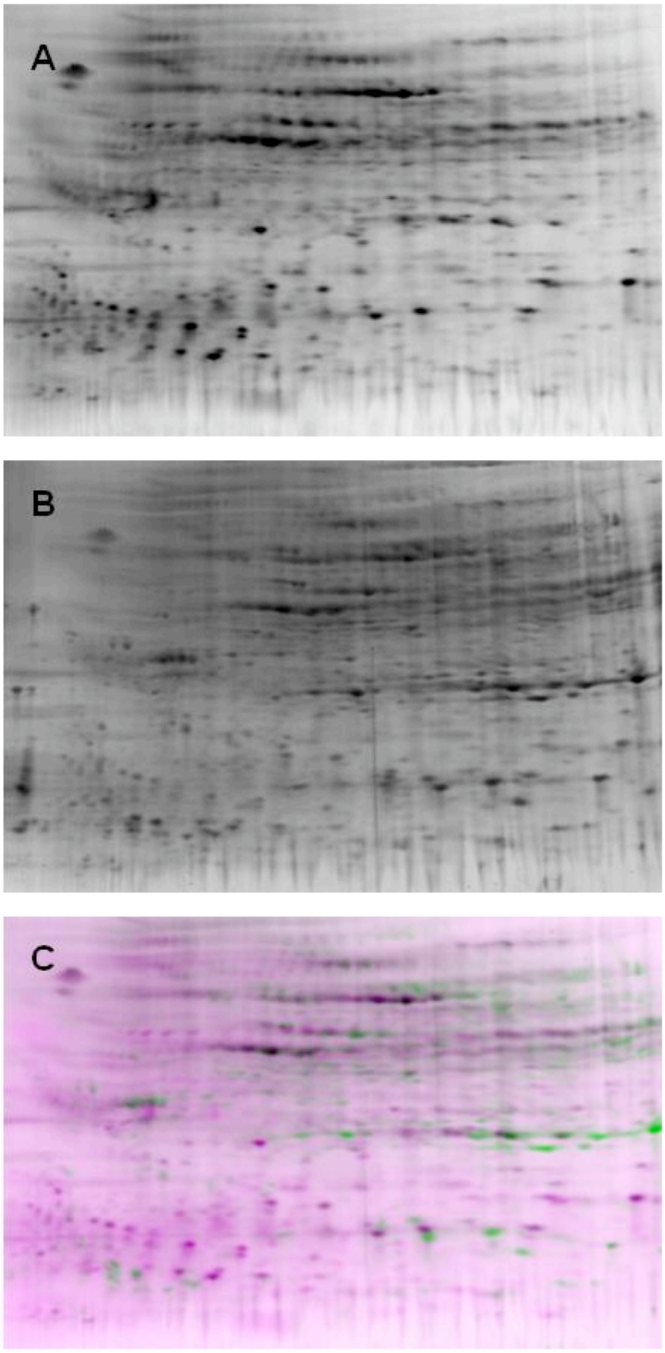
Supplemental Figure 1. Differential Protein Expression Changes. Proteins with a change in abundance were grouped by biological function with fold change in expression shown in the bar graph. White and grey bars correspond to H_2O_2 and BSO treatments, respectively.

Supplemental Figure 2. Example of 2D-gel images after IAF-labeling (A) and Sypro staining (B). The two pictures were overlapped using SameSpots (C) with color pink and green representing IAF and Sypro, respectively.

Supplemental Figure 1



Supplemental Figure 2



Supplemental Table 1. List of proteins identified as differentially expressed (from sypro stain) or oxidized (from IAF labeling) in response to H2O2 and BSO treatments.

Stain	Treatment	Spot #	Protein accession #	Protein name	Peptide sequence	Best Mascot ion score	Average of normalized volume in Control	Standard deviation	Average of normalized volume in Treatment	Standard deviation	Fold change*	P value
Sypro	H2O2	14	gi|195623672	RHM1	AMVEELLK EEDTPNFTGSFYSK MPISSDLSNPR TGWIGGLLGK	42 47 73 64	53.4	7.9	101.1	27	1.9	0.022
Sypro	H2O2	16	gi|14423528	putative fructokinase	EAGALLSYDPNLR LGDDEFGHMLAGILR LLLVTLGEK	116 49 66 66 73 79	37.5	7.3	69.2	4.5	1.8	0.023
Sypro	H2O2	19	gi|15219234	VHA-A; ATP binding / hydrogen ion transporting ATP synthase, rotational mechanism	DALAEGDKITLETAK FDPDFINIR ITYIAPAGQYSLK ITYTLIK LAADTPLLTGQR LAEMPADSGYPAYLAAR LGDLFYR TTLVANTSNMPVAAR	83 44 46 41 58 52 60 84 57	560.1	51.2	337.1	30.1	-1.7	0.002
Sypro	H2O2	19	gi|11034734	myrosinase	GYAIGTDAPGR IGPVMITR	47						
Sypro	H2O2	22	gi|105873020	ADP-ribosylation factor 1	DAVILVFANK ILMVGLDAAGK	48 48	461.4	45.6	280.3	65.9	-1.6	0.033
Sypro	H2O2	30	gi|15221284	calmodulin, putative	IAPSELGILMR SLGQNPTQAQLK	53 73 81 51 51	946.5	115.6	1461	112.4	1.5	0.007
Sypro	H2O2	33	gi|14423528	putative fructokinase	LLLVTLGEK TALAFVTLR	49 48	1014.7	108.1	1524.2	314.7	1.5	0.041
Sypro	H2O2	36	gi|18394525	PGDH (3-PHOSPHOGLYCERATE DEHYDROGENASE)	LAVQLASGGK YVGVSLVGK	51 41	1063.6	80.4	723.9	47.9	-1.5	0.003
Sypro	H2O2	40	gi|15224351	OASB (O-ACETYLSERINE (THIOL) LYASE B)	IGYSMITDAEEK LIAVVFPSFGER LILTMPASMSLER	80 85 46	815.7	107.7	1168.4	90	1.4	0.014
Sypro	H2O2	40	gi|145332399	malate dehydrogenase (NAD)	DDLFNINAGIVK	46 69						
Sypro	H2O2	42	gi|297824877	glutathione S-transferase	VLAALYEK VLDIYEAR	54 47	2016.5	170.4	1420.5	164.9	-1.4	0.013
Sypro	H2O2	46	gi|116786768	*S-adenosyl-L-homocystein hydrolase*	LVGVSEETTTGVK SGIIVLAEGR	79 86	2146.1	358.1	2867.3	74.7	1.3	0.043
Sypro	H2O2	46	gi|17939849	mitochondrial F1 ATP synthase beta subunit	EMIESGVIK ENINSFQGLLDGK FTQANSEVSALLGR IGLFGGAGVGK IMNVLGEPIDER VVDLLAPYQR VLNTGAPITVPVGR VGLTGLTVAEYFR TVLIMELINNVAK IPSAVGYQPTLASDLGALQER	56 43 56 95 55 67 62 76 90						
Sypro	H2O2	48	gi|30698122	GRF8 (GENERAL REGULATORY FACTOR 8)	AAQDVAVADLAPTHPIR DSTLIMQLLR TAAEDTMIAYK	76 60 44	950.1	52.5	725.6	59.1	-1.3	0.009
Sypro	H2O2	52	gi|15226605	CCR2 (COLD, CIRCADIAN RHYTHM, AND RNA BINDING 2)	ALETAFAQYGDVIDSK GFGFVTFK	57 44	2871.1	357.9	2244.3	179.2	-1.3	0.05
Sypro	H2O2	58	gi|15233272	TPI (TRIOSEPHOSPHATE ISOMERASE)	NVSADVAATTR VAYALAQGLK	47 62	1423.3	103.4	1716.7	136	1.2	0.04
Sypro	H2O2	58	gi|297792905	*quinone reductase*	FGMMAAQFK GGSPYGAGTFAGDGSR	47 107						
Sypro	H2O2	60	gi|13899069	*NDPK1*	GDFAIDIGR IIGATNPAASEPGTIR KIIGATNPAASEPGTIR	48 78 52	1836.9	89.9	2017	54.4	1.1	0.042
Sypro	H2O2	61	gi|15375068	gamma hydroxybutyrate dehydrogenase	FVEGPVSGSK GYIDMSTVDAETSLK KPAEDGQLIILAAGDK	64 74 73	624.6	22.4	680.9	16.5	1.1	0.025
IAF	H2O2	2	gi|114171	3-phosphoshikimate 1-carboxyvinyltransferase, chloroplastic	ASEIVLQPIR FAEVLEK LGLNVER	82 60 45	308.1	127	67.9	22.3	-4.5	0.012
IAF	H2O2	4	gi|16209712	*CLP2 (CLP PROTEASE PROTEOLYTIC SUBUNIT 2)*	GQADDIQNEAK IALQSPAGAAR	56 57	465.6	27.2	122.1	64.7	-3.8	0.007
IAF	H2O2	8	gi|15221107	enolase, putative	EGLVLLIDAIEK IEEELGNVR LIGVDVR YAGEAFR YDLNFKK	47 63 50 45 40	1201.1	340.3	375.8	227.5	-3.2	0.031
IAF	H2O2	8	gi|110612122	actin	AGFAGDDAPR	59 54						
IAF	H2O2	19	gi|15238686	5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase	GVTAFGFDLVR IVEVNALAK YGAGIGPGVYDIHSPR	66 50 59	759.7	15.7	287.8	25.5	-2.6	4.96E-05
IAF	H2O2	22	gi|15238762	GDH1 (GLUTAMATE DEHYDROGENASE 1)	DDGTLASFVGFR IVAVSDITGAIK LSISELER MGAFTLGVNR	70 63 55 43	47.7	8.2	120.4	28.6	2.5	0.007
IAF	H2O2	22	gi|14764532	monodehydroascorbate reductase	LLPESYK QGISFAAK SLVSAAGDVFK TSVPDVYAVGDVATFPLK	53 44 53 55 69 57						
IAF	H2O2	37	gi|15221107	enolase, putative	IEEELGNVR LIGVDVR NINELVAPK NQADVDALMLELDGTPNK	54 58 53 56 53 56 49	1999.5	344	902.9	71.4	-2.2	0.002
IAF	H2O2	40	gi|10334503	GTP-binding protein		44 51						
IAF	H2O2	40	gi|15224993	PAA2 (20S PROTEASOME SUBUNIT PAA2)	ATEIEVGVVR YLGLLATGMTADSR	67 58	711.1	182.9	325.2	30.4	-2.2	0.01
IAF	H2O2	60	gi|14764532	monodehydroascorbate reductase	EFASQGVKPGELAVISK SLVSAAGDVFK	45 52	2361.2	192.8	1211.5	216.5	-1.9	0.004
IAF	H2O2	60	gi|15225353	succinyl-CoA ligase (GDP-forming) beta-chain, mitochondrial	LNFDDNAAFR	68 47 42						
IAF	H2O2	60	gi|33149230	ThiJ-like protein	LVADVLLDEVAEK RGGADVTVASVEDK VVVDGNVITSR	90 53 52						

IAF	H2O2	64	gi|15225353	succinyl-CoA ligase (GDP-forming) beta-chain, mitochondrial, putative	LITADDLDDAAEK LNFDDNAAFR MLGQVLVTK VPIDVFAGITDEDAAK	78 60 57 66 50	836.9	168.9	436.8	122.6	-1.9	0.033
IAF	H2O2	64	gi|15219412	PGK (PHOSPHOGLYCERATE KINASE)	FANGTEAVAK GVTTIIGGGDSVAAVEK LVAGLPEGGVLLENVR	57 62 52						
IAF	H2O2	64	gi|18391442	DET3 (DE-ETIOLATED 3) [Arabidopsis thaliana]	EVVDNIQSQVAK LVQDQESLR SNSFVEGVSQK YPTMSPLK	62 58 43 53						
IAF	H2O2	66	gi|15219721	malate dehydrogenase, cytosolic, putative	EFAPSIPEK MELIDAAFPLLK	43 51	2878.2	285.7	1527.1	63.8	-1.9	5.21E-04
IAF	H2O2	66	gi|15231715	fructose-bisphosphate aldolase, putative	AAQEALYVR GILAADESTGTIGK	68 77						
IAF	H2O2	66	gi|115345735	annexin 2	ALDKELSSDFER AYSDDDFIR ELSSDFER	62 57 41 57						
IAF	H2O2	66	gi|120675	Glyceraldehyde-3-phosphate dehydrogenase, cytosolic	AGIALSDNFVK TLLFGEKPVTVFGIR	56 48						
IAF	H2O2	67	gi|297816838	*TPI (TRIOSEPHOSPHATE ISOMERASE)*	EAGSTMDVVAAQTK FFVGGNWK NVSADVAATTR VAYALAQGLK	58 48 59 62	261.5	25.2	490.6	101.6	1.9	0.011
IAF	H2O2	67	gi|15233268	PAC1 (20S proteasome alpha subunit C1)	AAAVGANNQAAQSILK DGVVLIGEK	54 69						
IAF	H2O2	68	gi|15004984	glutamate dehydrogenase	MGAFTLGVNR VVAVSDITGAVR	47 43	1080.9	278	2021.9	320.9	1.9	0.019
IAF	H2O2	68	gi|33149230	ThiJ-like protein	LVADVLLDEVAEK VVVDGNVITSR	75 39						
IAF	H2O2	70	gi|15219257	PAB1 (PROTEASOME SUBUNIT PAB1)	EGFEGEISSK KLPSILVDEASVQK LPSILVDEASVQK LYKEPIPVTQLVR	61 62 81 61 51	157.1	13.9	287.5	65.3	1.8	0.016
IAF	H2O2	70	gi|15242792	GAMMA CAL1 (GAMMA CARBONIC ANHYDRASE LIKE 1)	SILEAGSVVPPGR YVTVGAYSLLR	61 55						
IAF	H2O2	72	gi|15237947	UTP--glucose-1-phosphate uridylyltransferase	FFDNAIGVNVPR LDAFLSQGK SGFINLVSR SIPSIVELDSLK	67 44 66	1516.5	236.6	2754	216.2	1.8	0.004
IAF	H2O2	72	gi|15224101	cytosol aminopeptidase	FDMGGSAAVLGAAK TIEVNNTDAEGR	48 67						
IAF	H2O2	72	gi|15227987	*LOS2 (Low expression of osmotically responsive genes 1)*	AGAVVSGIPLYK AVGNVNNIIGPALIGK ISGDALKDLYK TYDLNFK VNQIGSVTESIEAVK	70 72 57 63						
IAF	H2O2	72	gi|110741046	*mitochondrial processing peptidase alpha subunit*	EIEAIGGNTSASASR SAVLMNLESR	47 45						
IAF	H2O2	74	gi|15238762	GDH1 (GLUTAMATE DEHYDROGENASE 1)	DDGTLASFVGFR LSISELER MGAFTLGVNR	61 47 48	1810.3	224.1	1016	138.2	-1.8	0.006
IAF	H2O2	74	gi|1149571	mitochondrial elongation factor Tu	AIAFDEIDKAPEEK FPGDDIPIIR LMDAVDEYIPDPVR	51 71 57						
IAF	H2O2	74	gi|14764532	monodehydroascorbate reductase	QGISPAAR	51						
IAF	H2O2	74	gi|1542941	Acetoacetyl-coenzyme A thiolase	ILITLLGILK TPMGGFLGSLSSLPATK	46 54						
IAF	H2O2	74	gi|6684341	alcohol dehydrogenase	FGVTEFVNPK KGQSVAIFGLGAVGLAAAEGAR TDIPGVVEK	64 57 48						
IAF	H2O2	75	gi|207667274	chloroplast malate dehydrogenase	DDLPWKDAGVK GVAFANKPAN IQNAGTEVVDAK KLFGVTTLDVVR	51 51 57	270.6	42.3	478.4	104	1.8	0.024
IAF	H2O2	76	gi|15240663	ADK1 (ADENYLATE KINASE 1)	GFILDGFPR LVFIGPPGSGK VLNFAIDDSVLEER	49 51 59	176.8	41.8	311.5	49.1	1.8	0.028
IAF	H2O2	76	gi|15146314	AT3g25530/MWL2_15	FVEGPVSGSK KPAEDGQLIILAAGDK	46 66						
IAF	H2O2	79	gi|15222163	DHAR2	TPPEFASVGSK VLLTLEEK	57 46	1982.6	87.3	1142.7	232.9	-1.7	0.011
IAF	H2O2	82	gi|30691626	MTHSC70-1 (mitochondrial heat shock protein 70-1)	AVVTVPAYFNDAQR SSGGLSDDEINR	45 48	3333.1	197.5	1960.2	276.1	-1.7	0.004
IAF	H2O2	82	gi|15219234	VHA-A	DADAISDAITEETAK EDDLNEIVQLVGK LAADTPLLTGQR LAEMPADSGYPAYLAAR LGDLFYR	58 78 54 54 60 61						
IAF	H2O2	84	gi|297816838	TPI (TRIOSEPHOSPHATE ISOMERASE)	AILNESNEFVGDK NVSADVAATTR VAYALAQGLK	43 50 62	724.1	55.6	428.6	102.9	-1.7	0.026
IAF	H2O2	87	gi|11131564	Cysteine synthase	IGFSMISDAEK	47	367.2	37.2	617.7	100.3	1.7	0.01
IAF	H2O2	88	gi|15231176	ATPQ (ATP SYNTHASE D CHAIN, MITOCHONDRIAL)	AFDEVNTQLQTK VLVTDEAR	59 54	705	154.3	1185.1	144.4	1.7	0.025
IAF	H2O2	88	gi|113205143	40S ribosomal protein S13, putative [Solanum demissum]	GISASALPYK GLTPSQIGVILR	56 52						
IAF	H2O2	88	gi|105873020	ADP-ribosylation factor 1	DAVILVFANK ILMVGLDAAGK	51 61						
IAF	H2O2	95	gi|297824877	glutathione S-transferase	GMFGMTTDPAAVQELEGK VLAALYEK VLDIYEAR	70 54 49	202.9	17.2	332.5	40.4	1.6	0.005
IAF	H2O2	104	gi|297816838	*TPI (TRIOSEPHOSPHATE ISOMERASE*	EAGSTMDVVAAQTK VAYALAQGLK	68 63	1467.6	230.6	923.2	171.9	-1.6	0.031
IAF	H2O2	104	gi|24421231	ascorbate peroxidase	ALLDDPVFRPLVEK EGLLQLVSDK	49 61						
IAF	H2O2	112	gi|4928472	type 2 peroxiredoxin	FVADGSGEYTK LLGLELDLK	64 63	712.1	25.1	1099.5	156.2	1.5	0.008
IAF	H2O2	112	gi|297832912	dimethylmenaquinone methyltransferase family protein	VFEDNVIVR VLVVDGGGSLR VPLNIAGTR	49 56 48						
IAF	H2O2	112	gi|119720786	hydrogen-transporting ATP synthase	LASASTDLEK LTVNFVLPYASELSGK	60 52						

IAF	H2O2	114	gi|15233268	PAC1 (20S proteasome alpha subunit C1)	AAAVGANNQAAQSILK DGVVLIGEK TTIFSPEGR	60 45 52	203	35.1	310.4	13.2	1.5	0.012
IAF	H2O2	116	gi|11131564	Cysteine synthase	IGFSMISDAEK	69 59	928.1	56.3	1407.2	192.1	1.5	0.009
IAF	H2O2	118	gi|110741046	hypothetical protein	MIAAEDIGR SAVLMNLESR	47 42						
IAF	H2O2	118	gi|118484871	hypothetical protein ARALYDRAFT_902771	AGAVVSGIPLYK AVGNVNSIIGPALIGK TYDLNFK TYDLNFKEENNDGSQK VNQIGSVTESIEAVK	53 65 47 61 55	634.2	83.6	958.6	125.6	1.5	0.018
IAF	H2O2	118	gi|110740927	eukaryotic protein synthesis initiation factor 4A	MLFDIQK VLITTDLLAR	42 59						
IAF	H2O2	118	gi|15224101	cytosol aminopeptidase	GDILVGVTEK	87 59						
IAF	H2O2	120	gi|30691626	MTHSC70-1 (mitochondrial heat shock protein 70-1); ATP binding / unfolded protein binding [Arabidopsis thaliana]	AVVTVPAYFNDAQR GELLVGTPAK HLNITLTR IAGLDVQR SSGGLSDDEINR TTPSVVAINQK	49 53 52 50 110 49	3331.6	139.7	2255.9	321.3	-1.5	0.009
IAF	H2O2	123	gi|24421231	ascorbate peroxidase	EGLLQLVSDK SYPTVSEDYQK	77 41	413.1	68.1	608.1	50.4	1.5	0.021
IAF	H2O2	123	gi|15224993	PAA2 (20S PROTEASOME SUBUNIT PAA2)	ATEIEVGVVR LFQVEYAFK YLGLLATGMTADSR	59 50 68						
IAF	H2O2	125	gi|13447108	GF14 lambda	AAQDIAAADMAPTHPIR DSTLIMQLLR IVSSIEQKEESR KAAAEDTMVAYK LLEENLIPSAAASESK YEEMVNFMEK	54 59 56 54 49 78 48 109	402.5	81.3	578.7	50.4	1.4	0.04
IAF	H2O2	126	gi|119720786	hydrogen-transporting ATP synthase	LASASTDLEK LTVNFVLPYASELSGK	83 45	1145.3	46.1	1645.7	146.8	1.4	0.003
IAF	H2O2	127	gi|11135129	Thioredoxin H-type	VDVDELATVAK	75 63	925.2	148.6	1325.8	185.8	1.4	0.043
IAF	H2O2	129	gi|15224351	OASB (O-ACETYLSERINE (THIOL) LYASE B)	AFGAELVLTEPAK IGYSMITDAEEK LIAVVFPSFGER	72 70 61 56						
IAF	H2O2	129	gi|15241111	CYSD2 (CYSTEINE SYNTHASE D2)	IGLKGMLEK LIVVVFPSGGER	50 42						
IAF	H2O2	129	gi|297794887	molybdenum cofactor sulfurase family protein	DSFLVVR LALVESELPK VGDTISVLR	46 66 55	242.5	18.9	337.7	60.1	1.4	0.006
IAF	H2O2	129	gi|22331875	OASC (O-ACETYLSERINE (THIOL) LYASE ISOFORM C)	AFGAELVLTDPAK IGYSMVTDAEQK IQGIGAGFIPK LIAVVFPSFGER YLSTPLFQSIR	77 46 47 61 51						
IAF	H2O2	129	gi|81176557	annexin-like protein	AGINATTNR	44						
IAF	H2O2	129	gi|1052973	fructokinase	APGGAPANVAIAVSR TALAFVTLK	72 45						
IAF	H2O2	132	gi|11135129	Thioredoxin H-type	VDVDELATVAK	50 77 64	1650.3	85.6	2236.3	158.7	1.4	0.004
IAF	H2O2	141	gi|297849222	calreticulin 2	FYAISAEFPEFSNK GIQTSEDYR LAEETWGK LLSGDVDQK	65 49 44 49	7732.6	347.8	9263.7	356.4	1.2	0.006
Sypro	BSO	9	gi|169862267	hypothetical protein CC1G_06970	IAVLGFAFK LAANAMLAQR	47						
Sypro	BSO	9	gi|15232888	6-phosphogluconate dehydrogenase family protein	FLSGLKDER LPANLVQAQR NAELANLLVDPEFAK VIIMLVK	47 75	157.3	44.7	57.2	20.5	-2.8	0.017
Sypro	BSO	9	gi|15237422	KAS I (3-KETOACYL-ACYL CARRIER PROTEIN SYNTHASE I)	ALESANLGGDKLNTIDK GFSSEGYDGK	56 60 119						
Sypro	BSO	17	gi|18399899	lipid-associated family protein	AGTDSIISAR	49 53	104.5	13.6	48.5	9.5	-2.2	0.004
Sypro	BSO	18	gi|18379240	MLP328 (MLP-LIKE PROTEIN 328)	ITMIWEK VYDVIFQFIQK	44 62	357.1	71.4	763	140.7	2.1	0.01
Sypro	BSO	22	gi|19570344	nucleoside diphosphate kinase 1	IIGATNPAASEPGTIR KIIGATNPAASEPGTIR NVIHGSDSVESANK	93 47 54	400	75	198.9	43.9	-2	0.013
Sypro	BSO	29	gi|18424620	TUB2 (Tubulin beta-2)	AVLMDLEPGTMDSLR FPGQLNSDLR LAVNLIPFPR MMLTFSVFPSPK VSEQFTAMFR YLTASAMFR	52 61 56 61	982.4	176.1	530.7	27.7	-1.9	0.006
Sypro	BSO	29	gi|15219345	ATMC4 (METACASPASE 4)	AVLIGINYPGTK IRPSLFDAFGDDSSPK LASGFLEGK	59 64 52						
Sypro	BSO	29	gi|32967699	S-adenosyl-L-homocystein hydrolase	LVGVSEETTTGVK LVGVSEETTTGVKR SGIIVLAEGR TEFGPAQPFK	113 63 69						
Sypro	BSO	29	gi|112144574	elongation factor-1 alpha	IGGIGTVPVGR IPLQDVYK	62 47						
Sypro	BSO	37	gi|15232963	dimethylmenaquinone methyltransferase family protein	QIFSGPVVTVK VFEDNGLIR VLVVDGGGSLR VPLNIAGTR	48 46 57 50	306.2	36.1	175.4	49.2	-1.7	0.28
Sypro	BSO	43	gi|15233272	TPI (TRIOSEPHOSPHATE ISOMERASE)	FFVGGNWK IIYGGSVNGGNCK NVSADVAATTR VASPAQAQEVHDELR VAYALAQGLK	43 46 61 60	396.3	58.5	644.2	131.5	1.6	0.027
Sypro	BSO	43	gi|115434516	Os01g0147900	ESGSTMDVVAAQTK FFVGGNWK VIACVGETLEQR	65 43 61						
Sypro	BSO	45	gi|11135129	Thioredoxin H-type	FIAPFVELAK VDVDELATVAK	45 77	477.3	21.9	768.3	205.1	1.6	3.80E-02
Sypro	BSO	47	gi|224981577	1433-3	DAAESTLVAYK DSTLIMQLLR KDAAESTLVAYK NLLSVAYK SVDNEELTVEER	77 65 68 54 81	906.7	106	584.5	82.7	-1.6	0.014
Sypro	BSO	47	gi|13447104	GF14 omega	DSTLIMQLLR IISSIEQKEESR KDAAENTLSAYK NLLSVAYK	65 61 85 54						
Sypro	BSO	47	gi|18395103	GRF10 (GENERAL REGULATORY FACTOR 10)	AAVAAAETGLAPTHPVR DSTLIMQLLR EAADQSLEAYK KEAADQSLEAYK VCNDILSVIDK YDEMVEAMK YLAEFSSGAER	72 65 83 75 56 54						

Sypro	BSO	47	gi|297811739	GRF5 (GENERAL REGULATORY FACTOR 5)	DSTLIMQLLR EENVYLAK NLLSVAYK SAQDIALADLPPTHPIR	49 65 54 71						
Sypro	BSO	50	gi|14423528	putative fructokinase	APGGAPAWAIAVSR LGDDEFGHMLAGILR	110 53 60						
Sypro	BSO	50	gi|1429207	annexin	[illegible]	67 67	270.3	5.9	175.9	47	-1.5	0.049
Sypro	BSO	50	gi|14422255	enoyl-[acyl-carrier protein] reductase	[illegible] IIPGYGGGMSSAK	52 67 48						
Sypro	BSO	60	gi|20140684	Translationally-controlled tumor protein homolog	LQEQPTYDK VVDIVDTFR	43 47	218.5	16.2	317.3	63	1.5	0.046
Sypro	BSO	63	gi|11034734	myrosinase	GEFIGPLFVEDK GYAIGTDAPGR IGPVMITR	58 58 59	1501.2	122.8	1078.8	207.6	-1.4	0.045
Sypro	BSO	63	gi|217074426	unknown	YAGMLEYDGELK YENDWGVVK	83 44						
Sypro	BSO	65	gi|18424620	TUB2	AVLMDLEPGTMDSLR FPGQLNSDLR LAVNLIPFPR VSEQFTAMFR	66 51 63 50	656.2	70.8	476.9	78.1	-1.4	0.044
Sypro	BSO	65	gi|15241472	TUB4 (tubulin beta-4 chain)	[illegible] FPGQLNSDLR IDVYFNEASGGK LAVNLIPFPR MMMTFSVFPSPK [illegible]	51 66 66 63 57 50 47						
Sypro	BSO	66	gi|166198115	peroxidase precursor	[illegible] IAASILR	89 63 76	1298.9	77	1778.7	317.8	1.4	0.043
Sypro	BSO	67	gi|297811059	inorganic pyrophosphatase family protein	AKMEVATDEDFTPIK IIHETNESWAK IVAISLDDPK MEVATDEDFTPIK	66 51 70 70 51	546.5	12.1	402.4	23.9	-1.4	0.001
Sypro	BSO	67	gi|297841835	predicted protein	IEAVEPEK IEAVEPEKNLITFR VIEGDLMK	50 53 39						
Sypro	BSO	70	gi|31790095	glutathione S-transferase 2	LATVLDVYEAR VPAFEDGDLK YEEEGTNLLPADSK	70 46 75	1416.7	272.9	1908.6	82.3	1.3	0.048
Sypro	BSO	72	gi|4928472	type 2 peroxiredoxin	FALLIDNLK FVADGSGEYTK	47 47						
Sypro	BSO	72	gi|15232963	dimethylmenaquinone methyltransferase family protein	VFEDNGLIR VLVVDGGGSLR VPLNIAGTR	59 55 42	864.8	112.5	1155.5	76.2	1.3	0.024
Sypro	BSO	72	gi|205364142	glutathione peroxidase	GNDVDLSTYK YAPTTSPLSIEK	54 45						
Sypro	BSO	77	gi|15237947	UTP--glucose-1-phosphate uridylyltransferase, putative / UDP-glucose pyrophosphorylase, putative / UGPase, putative [Arabidopsis thaliana], gi|12585448|sp|P57751.1|UGPA1_ARATH RecName: Full=UTP--glucose-1-phosphate uridylyltransferase 1; AltName: Full=UDP-glucose pyrophosphorylase 1; Short=UGPase 1; Short=UDPGP 1, gi|13430664|gb|AAK25954.1|AF360244_1 putative UDP-glucose pyrophosphorylase [Arabidopsis thaliana], gi|13605671|gb|AAK32829.1|AF361816_1 AT5g17310/MKP11_16 [Arabidopsis thaliana], gi|10177076|dbj|BAB10518.1| UDP-glucose pyrophosphorylase [Arabidopsis thaliana], gi|14532836|gb|AAK64100.1| putative UDP-glucose pyrophosphorylase [Arabidopsis thaliana]	FFDNAIGVNVPR LDAFLSQGK LVEADALK SGFINLVSR SIPSIVELDSLK VVADEFVPWPSK	79 51 57 69 57 43	1140.5	113	1499.3	32.5	1.3	0.009
Sypro	BSO	77	gi|17939849	mitochondrial F1 ATP synthase beta subunit	ENINSFQGLLDGK FTQANSEVSALLGR IMNVLGEPIDER TIAMDGTEGLVR TVLIMELINNVAK VGLTGLTVAEYFR	48 49 58 80 66 48						
Sypro	BSO	77	gi|116787286	unknown	FDMGGSAAVLGAAK GLTFDSGGYNIK TIEVNNTDAEGR	98 47 69						
Sypro	BSO	77	gi|297827139	hypothetical protein ARALYDRAFT_902771	AGAVVSGIPLYK AVGNVVNIIGPALIGK TYDLNFK VNQIGSVTESIEAVK	45 65 60 62						
Sypro	BSO	79	gi|15237947	UTP--glucose-1-phosphate uridylyltransferase, putative	FFDNAIGVNVPR LVEADALK SIPSIVELDSLK	66 48 53						
Sypro	BSO	79	gi|115472953	Os07g0574800	LISQIISSLTTSLR SLDIERPTYTNLNR	52 54						
Sypro	BSO	79	gi|116787286	unknown	FDMGGSAAVLGAAK TIEVNNTDAEGR	116 69	312.6	21.2	407.8	30.7	1.3	0.011
Sypro	BSO	79	gi|297827139	hypothetical protein ARALYDRAFT_902771	AGAVVSGIPLYK AVGNVVNIIGPALIGK TYDLNFKEENNDGSQK	47 61 56						
Sypro	BSO	82	gi|13447108	GF14 lambda	AAAEDTMVAYK DSTLIMQLLR LLEENLIPSAAASESK	61 59 67						
Sypro	BSO	82	gi|13447104	GF14 omega	DAAENTLSAYK DSTLIMQLLR EEFVYMAK KDAAENTLSAYK LVPAAASGDSK NLLSVAYK VSAAVDGDELTVEER	71 59 47 78 66 51 125						

Sypro	BSO	82	gi|115446909	Os02g0580300	DSTLIMQLLR NLLSVAYK TVDSEELTVEER YEEMVEFMEK	59 51 72 48 56	1652.3	79.6	1297	194.3	-1.3	0.048
Sypro	BSO	82	gi|18411901	GRF2 (GENERAL REGULATORY FACTOR 2)	DSTLIMQLLR EEFVYMAK NLLSVAYK SAQDIANAELAPTHPIR VSAAVDGDELTVEER YEEMVEFMEK YLAEFK	59 47 51 69 125 48 56						
Sypro	BSO	82	gi|224981569	1433-1	DSTLIMQLLR EAADGSLEAYK ILSSIEQKEESK YLAEFSSGAER	59 93 49 69						
Sypro	BSO	83	gi|110171851	histone H4	DNIQGITKPAIR IFLENVIR TVTAMDVVYALK	53 52 49	323.6	9.7	411.6	47.5	1.3	0.025
Sypro	BSO	86	gi|15228098	60S ribosomal protein L12 (RPL12A)	IGPLGLAPK VTGGEVGAASSLAPK	44 50	1093.04	93.2	1324.8	100.8	1.2	0.041
Sypro	BSO	87	gi|18379240	MLP328 (MLP-LIKE PROTEIN 328)	ITMMWEK VQQVVEQFIQK	48 57	3471.6	135.4	2905.3	283	-1.2	0.038
Sypro	BSO	88	gi|4928472	type 2 peroxiredoxin	FALLIDNLK FVADGSGEYTK LLGLELDLK	50 56 65	733.3	14.3	875.9	40.8	1.2	0.004
IAF	BSO	9	gi|7110585	isopentenyl pyrophosphate:dimethylallyl	DSELIEENALGVR GTLGEAVDMK LVVDNFLMK	46 48	172.8	80.1	602.9	223.5	3.5	0.023
IAF	BSO	11	gi|46093471	stromal ascorbate peroxidase	FDNSYFTEIK YAADQDAFFK YEIELK	46 75 43	682.9	334.8	206.6	96.5	3.1	0.037
IAF	BSO	11	gi|15240663	adenylate kinase, putative	LVFIGPPGSGK VLNFAIDDSVLEER	42 56						
IAF	BSO	13	gi|297816838	*TPI (TRIOSEPHOSPHATE ISOMERASE)*	NVSADVAATTR VAYALAQGLK	52 62	129	62.5	404.6	113.9	3.1	0.037
IAF	BSO	14	gi|1732572	beta-glucosidase	FGLYYIDFK NLNTDAFR	60 41	772.6	203	255.7	199.4	-3	0.046
IAF	BSO	30	gi|261343270	O-acetylserine(thiol)lyase isoform A6	IDGFVSGIGTGGTITGAGK IGFSMISDAEK LFVAVFPSFGER LIITMPASMSIER	60 50 44 46 82	328	82.6	710.1	78.9	2.2	0.009
IAF	BSO	30	gi|14423528	putative fructokinase	APGGAPAIVAIAVSR LLLVTLGEK VSDVELEFLTGSNK	121 68 59						
IAF	BSO	46	gi|15224993	PAA2 (20S PROTEASOME SUBUNIT PAA2)	ATEIEVGVVR YLGLLATGMTADSR	54 71	104.9	9	196.1	47.2	1.9	0.017
IAF	BSO	46	gi|15221463	coatomer protein epsilon subunit family protein / COPE family protein	AASAEDNFER LLAMYLSSPENK	83 53						
IAF	BSO	51	gi|11135129	Thioredoxin H-type	FIAPIFVELAK VDVDELATVAK	65 66	52.3	9.5	94.3	11.3	1.8	0.008
IAF	BSO	51	gi|110171851	histone H4	IFLENVIR ISGLIYEETR	52 64						
IAF	BSO	52	gi|297816838	predicted protein	EAGSTMDVVAAQTK NVSADVAATTR	53 43	1311.8	146.9	729.2	19.2	-1.8	0.001
IAF	BSO	52	gi|297824877	glutathione S-transferase	VLAALYEK VLDIYEAR	12 16						
IAF	BSO	61	gi|297816838	predicted protein	AILNESNEFVGDK EAGSTMDVVAAQTK NVSADVAATTR VAYALAQGLK	81 45 73 60	420.8	86.1	729.3	41.1	1.7	0.011
IAF	BSO	61	gi|11135129	Thioredoxin H-type	FIAPIFVELAK VDVDELATVAK VVGAAKEEIEAK	49 77 61						
IAF	BSO	61	gi|110171851	histone H4	ISGLIYEETR TVTAMDVVYALK	70 70						
IAF	BSO	61	gi|297828984	hypothetical protein ARALYDRAFT_896526	FIVVADDTK SLGIPLVALDTHPR	44 68						
IAF	BSO	65	gi|15224993	PAA2 (20S PROTEASOME SUBUNIT PAA2)	AAGITSIGVR ATEIEVGVVR YLGLLATGMTADSR	42 43 68	469.6	77.4	797.8	71.4	1.7	7.00E-03
IAF	BSO	65	gi|15233268	PAC1 (20S proteasome alpha subunit C1)	AAAVGANNQAAQSILK DGVVLIGEK TTIFSPEGR	59 43 51						
IAF	BSO	65	gi|24421231	ascorbate peroxidase	EGLLGLVSSK SYPTVSEDYQK	71 42						
IAF	BSO	68	gi|15242459	mtHSC70-2 (HEAT SHOCK PROTEIN 70)	AVVTVPAYFNDAQR EIEDAVADLR GELLVGTPAK VQSIVAEIFGK	92 53 50 45	609	7.1	361.5	70	-1.7	0.009
IAF	BSO	68	gi|2654210	heat shock 70 protein	AVVTVPAYFNDAQR GELLVGTPAK QAVTNPTNTIFGTK TTPSVVAFNPK	92 50 62						
IAF	BSO	68	gi|2493122	V-type proton ATPase catalytic subunit A	LGDLFYR SGDVYIPR TTLVANTSNMPVAAR TVISQALSK	68 60 61 71 54 78						
IAF	BSO	72	gi|15224993	PAA2 (20S PROTEASOME SUBUNIT PAA2)	ATEIEVGVVR LFQVEYAFK VLGLLATGMTADSR	54 46 53	462.9	61.3	766.3	127.1	1.7	0.018
IAF	BSO	75	gi|15224582	GSTF10	AVVTLVER	57	1172.5	171.4	1931.4	377.2	1.6	0.024
IAF	BSO	78	gi|297816838	*TPI (TRIOSEPHOSPHATE ISOMERASE)*	NVSADVAATTR VAYALAQGLK	47 48	182.4	16.4	296.6	43.6	1.6	0.008
IAF	BSO	81	gi|297816838	predicted protein	FFVGGNWK NVSADVAATTR	41 52 62	171.7	23.2	108.1	6.8	-1.6	0.006
IAF	BSO	81	gi|15219257	PAB1 (PROTEASOME SUBUNIT PAB1)	EGFEGEISSK LYKEPIPVTQLVR	68 54						
IAF	BSO	83	gi|87294807	glutathione S-transferase	FGNFSIEAESPK IVAYAAEFR	40 58	1033.8	222.3	668.1	44	-1.5	0.033
IAF	BSO	86	gi|18399899	lipid-associated family protein	GNLDIFSGK VGDVGIR	41 55	282.8	63.6	431.5	54	1.5	0.043
IAF	BSO	93	gi|11135129	Thioredoxin H-type	FIAPIFVELAK VDVDELATVAK VVGAAKEEIEAK	49 77 61	663.9	27.7	986.9	24.5	1.5	0.001
IAF	BSO	93	gi|110171851	histone H4	ISGLIYEETR TVTAMDVVYALK	70 70						
IAF	BSO	94	gi|15231715	fructose-bisphosphate aldolase, putative	ANSEATLGTYK GILAADESTGTIGK GILAADESTGTIGKR LASIMENVEENR	61 61 86 64 79	1476.2	204.3	2179.8	357.2	1.5	0.04

IAF	BSO	97	gi|15235213	caffeoyl-CoA 3-O-methyltransferase, putative	DFVLELNK ILAMDVNR	45 57	2183.4	281.9	1495.4	150.3	-1.5	0.017
IAF	BSO	97	gi|15224993	PAA2 (20S PROTEASOME SUBUNIT PAA2)	AAGITSIGVR ATEIEVGVVR LFQVEYAFK YLGLLATGMTADSR	88 59 46 60						
IAF	BSO	98	gi|15222633	disease resistance-responsive family protein / dirigent family protein	DLSVVGGTGDFFMSR GIVTFETDTFEGAK	52 77	436.2	37.6	636.6	107.3	1.5	0.032
IAF	BSO	101	gi|15239146	ATNADP-ME2 (NADP-MALIC ENZYME 2)	GIQVIVVTDGER YMALMDLQER	61 48	3239.5	112.9	2277.9	342.1	-1.4	0.018
IAF	BSO	101	gi|15229559	HSP60 (Heat shock protein 60)	GISMAVDAVVTNLK IGGASEAEVGEK IGVQIQNALK SAIELSTSDYDKEK	51 68 69						
IAF	BSO	101	gi|414103	myrosinase, thioglucoside glucohydrolase	GYAVGTDAPGR IGPVMITR	47 59						
IAF	BSO	101	gi|15010652	At1g09780/F21M12_17	ALEYEDFDKFDR LDQLQLLIK YLVSPPEIDR	55 47 53						
IAF	BSO	104	gi|18395103	GRF10 (GENERAL REGULATORY FACTOR 10)	ILSSIEQK YLAEFSSGAER	40 55	494.4	42.2	666.8	95.6	1.3	0.039
IAF	BSO	106	gi|15241286	MAB1 (MACCI-BOU)	LAEEGISAEVINLR VLAPYSAEDAR	78 53	990.3	100.7	765.7	6.4	-1.3	0.013
IAF	BSO	106	gi|14423528	putative fructokinase	APGGAPANVAIAVSR LLLVTLGEK TALAFVTLR	102 60 60						
IAF	BSO	107	gi|15231176	ATPQ (ATP SYNTHASE D CHAIN, MITOCHONDRIAL)	AFDEVNTQLQTK FDALLVELK	57 59	115.9	7.5	148.6	9.1	1.3	0.009
IAF	BSO	108	gi|15236385	SBP1 (selenium-binding protein 1)	GGPQMIQLSLDGK YLVLPSLISGR	47 63	1557.8	159.5	1217.4	56.6	-1.3	0.02
IAF	BSO	108	gi|118429132	vacuolar ATPase subunit B	AVVQVFEGTSGIDNK GQVLEVDGEK QIYPPINVLPSLSR TPVSQDMLGR TVSGVAGPLVILEK	75 50 47 64 61 45						
IAF	BSO	108	gi|297799542	hypothetical protein ARALYDRAFT_492381	AAAEVLALQK DSLLLFSR SVGALIALYER VLQQLVAVESLGK	49 45 51 46						
IAF	BSO	113	gi|297824877	glutathione S-transferase	VLAALYEK VLDIYEAR	54 52	1106.3	60.1	952.9	22.9	-1.2	0.012
IAF	BSO	113	gi|297816838	predicted protein	NVSADVAATTR VAYALAQGLK	69 64						
IAF	BSO	114	gi|33285914	putative dehydroascorbate reductase	ALFSLDSFEK VSAVDLSLAPK	59 75	1606.4	76.2	1399.1	77.9	-1.1	0.03
IAF	BSO	114	gi|15233272	ATCTIMC (CYTOSOLIC TRIOSE PHOSPHATE ISOMERASE)	AILNESNEFVGDK EAGSTMDVVAAQTK NVSADVAATTR VASPAQAQEVHDELR VASPAQAQEVHDELRK VAYALAQGLK	82 66 49 59 46 58						
IAF	BSO	114	gi|15219257	PAB1 (PROTEASOME SUBUNIT PAB1)	EGFEGEISSK KLPSILVDEASVQK	62 48						

* fold change indicates the direction of the change of the protein expression or oxidation in the treated sample.

Spot highlighted in yellow are spots for which more than one protein were identified.

Supplemental Table 2. List of the 52 proteins identified as single protein classified according to their biological process.

Stain	Treatment	Spot #	Protein accession #	Protein name	Biological pathway	Average of normalized volume in Control	Standard deviation	Average of normalized volume in Treatment	Standard deviation	Fold change*	P value
IAF	H2O2	2	gi|114171	3-phosphoshikimate 1-carboxyvinyltransferase, chloroplastic	Amino acid biosynthesis	308.1	127	67.9	22.3	-4.5	0.012
IAF	H2O2	19	gi|15238686	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	Amino acid biosynthesis	759.7	15.7	287.8	25.5	-2.6	4.96E-05
IAF	H2O2	87	gi|11131564	Cysteine synthase	Amino acid biosynthesis	367.2	37.2	617.7	100.3	1.7	0.01
IAF	H2O2	116	gi|11131564	Cysteine synthase	Amino acid biosynthesis	928.1	56.3	1407.2	192.1	1.5	0.009
IAF	BSO	107	gi|15231176	ATPQ (ATP SYNTHASE D CHAIN, MITOCHONDRIAL)	ATP synthesis coupled proton transport	115.9	7.5	148.6	9.1	1.3	0.009
IAF	H2O2	126	gi|119720786	hydrogen-transporting ATP synthase	ATP synthesis coupled proton transport	1145.3	46.1	1645.7	146.8	1.4	0.003
IAF	BSO	104	gi|18395103	GRF10 (GENERAL REGULATORY FACTOR 10)	Brassinosteroid mediated signaling path	494.4	42.2	666.8	95.6	1.3	0.039
IAF	H2O2	125	gi|13447108	GF14 lambda	Brassinosteroid mediated signaling path	402.5	81.3	578.7	50.4	1.4	0.04
IAF	BSO	14	gi|1732572	beta-glucosidase	Carbohydrate metabolic process	772.6	203	255.7	199.4	-3	0.046
IAF	H2O2	75	gi|207667274	chloroplast malate dehydrogenase	Carbohydrate metabolic process	270.6	42.3	478.4	104	1.8	0.024
IAF	BSO	75	gi|15224582	GSTF10	Cell redox homeostasis	1172.5	171.4	1931.4	377.2	1.6	0.024
IAF	BSO	83	gi|87294807	glutathione S-transferase	Cell redox homeostasis	1033.8	222.3	668.1	44	-1.5	0.033
IAF	H2O2	79	gi|15222163	DHAR2	Cell redox homeostasis	1982.6	87.3	1142.7	232.9	-1.7	0.011
IAF	H2O2	95	gi|297824877	glutathione S-transferase	Cell redox homeostasis	202.9	17.2	332.5	40.4	1.6	0.005
IAF	H2O2	127	gi|11135129	Thioredoxin H-type	Cell redox homeostasis	925.2	148.6	1325.8	185.8	1.4	0.043
IAF	H2O2	132	gi|11135129	Thioredoxin H-type	Cell redox homeostasis	1650.3	85.6	2236.3	158.7	1.4	0.004
IAF	BSO	13	gi|297816838	*TPI (TRIOSEPHOSPHATE ISOMERASE)*	Glycolysis	129	62.5	404.6	113.9	3.1	0.037
IAF	BSO	78	gi|297816838	*TPI (TRIOSEPHOSPHATE ISOMERASE)*	Glycolysis	182.4	16.4	296.6	43.6	1.6	0.008
IAF	BSO	94	gi|15231715	fructose-bisphosphate aldolase, putative	Glycolysis	1476.2	204.3	2179.8	357.2	1.5	0.04
IAF	H2O2	37	gi|15221107	enolase, putative	Glycolysis	1999.5	344	902.9	71.4	-2.2	0.002
IAF	H2O2	84	gi|297816838	TPI (TRIOSEPHOSPHATE ISOMERASE)	Glycolysis	724.1	55.6	428.6	102.9	-1.7	0.026
IAF	H2O2	141	gi|297849222	calreticulin 2	Protein folding	7732.6	347.8	9263.7	356.4	1.2	0.006
IAF	BSO	72	gi|15224993	PAA2 (20S PROTEASOME SUBUNIT PAA2)	Proteolysis	462.9	61.3	766.3	127.1	1.7	0.018
IAF	H2O2	4	gi|16209712	*CLP2 (CLP PROTEASE PROTEOLYTIC SUBUNIT 2)*	Proteolysis	465.6	27.2	122.1	64.7	-3.8	0.007
IAF	H2O2	114	gi|15233268	PAC1 (20S proteasome alpha subunit C1)	Proteolysis	203	35.1	310.4	13.2	1.5	0.012
IAF	BSO	9	gi|7110585	isopentenyl pyrophosphate:dimethylallyl pyrophosphate isomerase	Secondary metabolism	172.8	80.1	602.9	223.5	3.5	0.023
IAF	BSO	98	gi|15222633	disease resistance-responsive family protein / dirigent family protein	Stress response	436.2	37.6	636.6	107.3	1.5	0.032
IAF	H2O2	120	gi|30691626	MTHSC70-1 (mitochondrial heat shock protein 70-1); ATP binding / unfolded protein bindi	Stress response	3331.6	139.7	2255.9	321.3	-1.5	0.009
IAF	BSO	86	gi|18399899	lipid-associated family protein	unknown	282.8	63.6	431.5	54	1.5	0.043
Sypro	H2O2	36	gi|18394525	PGDH (3-PHOSPHOGLYCERATE DEHYDROGENASE)	Amino acid biosynthesis	1063.6	80.4	723.9	47.9	-1.5	0.003
Sypro	H2O2	48	gi|30698122	GRF8 (GENERAL REGULATORY FACTOR 8)	Brassinosteroid mediated signaling path	950.1	52.5	725.6	59.1	-1.3	0.009
Sypro	H2O2	30	gi|15221284	calmodulin, putative	Calcium-mediated signaling	946.5	115.6	1461	112.4	1.5	0.007
Sypro	H2O2	33	gi|14423528	putative fructokinase	Carbohydrate metabolic process	1014.7	108.1	1524.2	314.7	1.5	0.041
Sypro	H2O2	16	gi|14423528	putative fructokinase	Carbohydrate metabolic process	37.5	7.3	69.2	4.5	1.8	0.023
Sypro	H2O2	14	gi|195623672	RHM1	Carbohydrate metabolic process	53.4	7.9	101.1	27	1.9	0.022
Sypro	H2O2	42	gi|297824877	glutathione S-transferase	Cell redox homeostasis	2016.5	170.4	1420.5	164.9	-1.4	0.013
Sypro	BSO	70	gi|31790095	glutathione S-transferase 2	Cell redox homeostasis	1416.7	272.9	1908.6	82.3	1.3	0.048
Sypro	BSO	66	gi|166198115	peroxidase precursor	Cell redox homeostasis	1298.9	77	1778.7	317.8	1.4	0.043
Sypro	BSO	45	gi|11135129	Thioredoxin H-type	Cell redox homeostasis	477.3	21.9	768.3	205.1	1.6	3.80E-02
Sypro	BSO	88	gi|4928472	type 2 peroxiredoxin	Cell redox homeostasis	733.3	14.3	875.9	40.8	1.2	0.004
Sypro	BSO	18	gi|18379240	MLP328 (MLP-LIKE PROTEIN 328)	Defense response	357.1	71.4	763	140.7	2.1	0.01
Sypro	BSO	87	gi|18379240	MLP328 (MLP-LIKE PROTEIN 328)	Defense response	3471.6	135.4	2905.3	283	-1.2	0.038
Sypro	BSO	60	gi|20140684	Translationally-controlled tumor protein homolog	Defense response	218.5	16.2	317.3	63	1.5	0.046
Sypro	BSO	83	gi|110171851	histone H4	Nucleosome assembly	323.6	9.7	411.6	47.5	1.3	0.025
Sypro	H2O2	60	gi|13899069	*NDPK1*	Nucleotide metabolism	1836.9	89.9	2017	54.4	1.1	0.042
Sypro	BSO	22	gi|19570344	nucleoside diphosphate kinase 1	Nucleotide metabolism	400	75	198.9	43.9	-2	0.013
Sypro	H2O2	61	gi|15375068	gamma hydroxybutyrate dehydrogenase	Oxidation reduction	624.6	22.4	680.9	16.5	1.1	0.025
Sypro	H2O2	22	gi|105873020	ADP-ribosylation factor 1	Protein transport	461.4	45.6	280.3	65.9	-1.6	0.033
Sypro	BSO	37	gi|15232963	dimethylmenaquinone methyltransferase family protein	Secondary metabolism	306.2	36.1	175.4	49.2	-1.7	0.28
Sypro	H2O2	52	gi|15226605	CCR2 (COLD, CIRCADIAN RHYTHM, AND RNA BINDING 2)	Stress response	2871.1	357.9	2244.3	179.2	-1.3	0.05
Sypro	BSO	86	gi|15228098	60S ribosomal protein L12 (RPL12A)	Translation	1093	93.2	1324.8	100.8	1.2	0.041
Sypro	BSO	17	gi|18399899	lipid-associated family protein	unknown	104.5	13.6	48.5	9.5	-2.2	0.004

PROTEIN EXTRACTION AND LABELING PROTOCOL

Plant Growth

Wildtype *B. juncea* seeds were obtained from stocks maintained at the Donald Danforth Plant Science Center, and allowed to germinate in a growth chamber at 22°C, 200 mmol/m2 light intensity, 50% relative humidity during a 16-hour light/8-hour dark cycle. Once their second set of true leaves began to emerge (typically two-three weeks after planting), the seedlings were transplanted to one gallon pots and moved to a greenhouse with the same light/dark cycle. The plants were grown normally until they began to flower, at which point they were separated into groups: control, H_2O_2-treated, and BSO-treated. H_2O_2-treated and BSO-treated plants were pot-watered with 2 L of 1 mM H_2O_2 and 2 L of 50 μM BSO respectively, and positioned to allow rapid draining. After 1 hour, an additional 1 L of 1 mM H_2O_2 or 1 L of 50 μM BSO was respectively applied, and the plants again were allowed to drain for an additional hour. Control plants were irrigated with distilled water, and were otherwise treated identically. At the end of 2 hours, the plant roots were rapidly washed to remove excess soil, flash frozen in liquid nitrogen, and stored at -80°C

Protein Labeling

For the H_2O_2, BSO, and control treatments, soluble protein extraction was completed in triplicate using tissue from three different plants, and all steps were carried out at 4°C unless otherwise indicated. Approximately 800 mg of root tissue, or 400 mg

of leaf tissue were ground to a fine powder using liquid nitrogen and a mortar and pestle. The tissue was suspended in extraction buffer [100 mM Tris-HCl, pH 8.0; 100 mM NEM; 1% CHAPS; 1% plant protease inhibitors (Sigma, P9599)] to 200 mg/mL, and sonicated for 3 x 15 seconds. In between each sonication the tissue was vortexed briefly and allowed to sit on ice for 30 seconds. The samples were then centrifuged for 16.1k x g for 15 minutes to precipitate the insoluble debris. Following centrifugation, the supernatant was mixed with 4 volumes of pre-chilled methanol and stored on ice. In most cases, protein precipitation was observed almost immediately. After 30 minutes, the samples were again centrifuged at 16.1k x g for 15 minutes, with the supernatant being discarded once the spin was complete. The protein pellet was washed twice more (2 x 30 minutes) with the same volume of methanol, and centrifuged at 16.1k x g for 5 minutes after each wash. During the washing steps, care was taken to periodically disrupt the pellet by hand (using a sterile pipette tip) or by vortexing to ensure complete removal of excess NEM. After the final spin, the pellet was air-dried for several minutes, and then resuspended in 150µL of reduction buffer [50 mM Tris-HCl, pH 8.0, 7 M urea, 2 M thiourea, 50 mM DTT (added in just prior to use)] per 200 mg of starting tissue. The suspension was incubated for 15 minutes at room temperature, and then re-precipitated with 4 volumes of pre-chilled methanol. As described above, the protein was washed 3 times (3 x 30 minutes) with the same volume of methanol and centrifuged after each wash (1 x 15 minutes; 2 x 5 minutes), with care taken to disrupt the pellet. After the final spin, the pellet was air-dried briefly, and resuspended in 150µL of labeling buffer [40 mM HEPES, pH 7.5; 50 mM NaCl; 200 µM IAF (added in just prior to use)] per 200 mg of starting

tissue. Due to the light sensitivity of IAF, this and all subsequent steps were carried out under dim light. The suspension was incubated for 10 minutes at room temperature, and then re-precipitated with 4 volumes of pre-chilled methanol. Again, the protein was washed 3 times (3 x 30 minutes) with the same volume of methanol and centrifuged after each wash (1 x 15 minutes; 2 x 5 minutes), with care taken to disrupt the pellet. After the final spin, the pellet was air-dried briefly, and resuspended in a small (<60μL per 200 mg of starting tissue) volume of destreak buffer (GE Healthcare). The protein concentration was then determined by CB-X assay (G-Biosciences).

2D-SDS-PAGE

For 2D-SDS-PAGE, 200 μg of extracted protein was resuspended to a total volume of 180 μL in destreak buffer and absorbed into a pH 4-7 gel strip (Bio-rad) overnight. Isoelectric focusing in the first dimension was carried out at room temperature in a Proteon IEF cell using a four-step method: 1) 250 V(olts), linear increase, 30 minutes; 2) 500 V, linear increase, 1 hour; 3) 8000 V, linear increase, 2.5 hours; 4) 8000 V, rapid increase, 35,000 Vhours. Separation in the second dimension was by molecular weight, and achieved using a standard gel box run at 150 V until the dye front reached the end of the gel. First dimension gel strips were secured in place relative to the second dimension gel using 1 mL of agarose. Following the second dimension separation, gels were removed from their cassettes and imaged (Ex: 488 nm: Em: 520 nm) using a Typhoon 9410 variable mode imager (Amersham Biosciences) to detect IAF-labeled proteins. After imaging, gels were bathed for 30 minutes with 100 mL of fixing solution

[10% methanol 7% glacial acetic acid] using an orbital shaker. The fixing solution was then poured off, and gels were then bathed overnight in 50 mL of Sypro Ruby protein stain (Bio-rad). The next day, the protein stain was poured off, and the gels were again bathed for 30 minutes with 100 mL of fixing solution. Following a washing step with MilliQ water to remove any excess stain, the gels were imaged again (Ex: 457 nm; Em: 610 nM) to detect total protein. For all gel replicates within a given set or sets to be compared, the same laser intensity (400 V for IAF images and 650-800 V for Sypro images) was used.

Spot Analysis and Excision

Replicate gels images were aligned using Progenesis Samespots (Nonlinear Dynamics). Further alignment of replicate control gel and replicate ozone treatment gel images was carried out for each for each pairwise comparison. In order to quantify expression and thiol composition differences between the control and treated samples, spot volume (as a function of intensity) was calculated and normalized for each spot in the aligned images. Those spots that different significantly in volume (ANOVA, $p<0.05$) between the averaged control and ozone treatment gels were then marked for excision.

Excision of significant spots from their respective gels was performed using a Gelpix System (Genetix) under high humidity (>85%) to prevent gel distortion or tearing. Gel plugs were dehydrated with 200 μL of acetonitrile (ACN) for 15 minutes at 900 rpm (revolutions per minute) using a room-temperature table-top shaker. The ACN was then removed, and, to remove the Sypro Ruby stain, the plugs were washed 5 x 15 minutes,

900 rpm, with 200 μL of 50mM NH_4HCO_3, 50% ACN, with the liquid discarded after every wash. Following the last NH_4HCO_3/ACN bath, the plugs further were washed for 5 minutes, 900 rpm, with 100 μL of ACN; when this step was complete the liquid was again discarded and the plugs were allowed to air-dry for several minutes. Once dry, the plugs were submerged in 20 μL of trypsin digestion buffer (50mM NH_4HCO_3 containing 6 ng/μL trypsin) and rehydrated overnight at 37°C. The next morning, 30 μL of 1% formic acid, 2% ACN was added to the digests, which were then shaken for 30 minutes at 900 rpm. Following the wash, the supernatant was collected from each plug and transfered to a new tube. Again, the plugs were shaken for 30 minutes at 900 rpm, this time in 24 μL of 60% ACN. After this final wash, the supernatant was removed and added to that collected during the previous step, and the plugs were discarded. Using a SpeedVac the combined digest from each gel plug was lyophilized to dryness, then finally resuspended in 7 μL of 1% formic acid, 5% ACN. Identification of the proteins contained in each digest was carried out by nano-LC-MS/MS as previously described [Alvarez et. al, 2009].

Alvarez, S., Berla, B., Sheffield, J., Cahoon, R.E., Jez, J.M., Hicks, L.M. (2009). Comprehensive analysis of the *Brassica juncea* root proteome in response to cadmium exposure by complementary proteomic approaches. Proteomics 9: 2419-2431.

CHAPTER 4

FROM CLIMATE CHANGE TO PROTEINS: REDOX PROTEOMICS OF OZONE-INDUCED RESPONSES IN SOYBEAN

PREFACE

This chapter describes the application of the dual-labeling methodology developed in chapter 2 to protein extracts exposed to either ambient or elevated tropospheric ozone concentrations. This work was completed in collaboration with the laboratory of Dr. Lisa Ainsworth at the University of Illinois at Urbana-Champaign (UIUC), USDA-ARS; without their expertise and facilities, the experiments described herein would not have been possible.

The SoyFACE Facility

The specific facilities at UIUC are referred to as SoyFACE (Soybean Free Air Concentration Enrichment), and are one of only a handful of FACE sites worldwide. The majority of FACE sites focus on the effects of elevated CO_2 concentrations on plant growth (in these cases the C in FACE actually stands for CO_2 instead of concentration); only SoyFACE and AspenFACE at the Harshaw Experimental Forest in Wisconsin have investigated the effects of tropospheric ozone in addition to CO_2. Regardless of the specific antagonists and/or species under inquiry however, the basic technology behind all FACE-type experiments remains the same. At SoyFACE, soybean plots are surrounded by octagonal rings composed of micropore tubing. The rings are approximately 16 meters in diameter, and are separated from one another in all directions by 100 meters of untreated soybean plants to avoid gas cross contamination. In 2009, 16 rings were in active use, and contained various soybean cultivars exposed to target

concentrations of ozone ranging from ambient (~40 ppb) to 200 ppb. 4 of the 16 rings were not exposed to ozone, but rather to elevated CO_2 at a target concentration of 585 ppm. The ozone used for elevated concentrations is produced on site with dedicated ozone generators, and pumped from the generator housings directly to the various rings, where its rate and direction of diffusion can be directly controlled. Wind direction, wind speed, temperature, humidity and host of other factors are monitored in real time for each ring; these factors will determine the rate at which and direction from which ozone is released so as to maintain the target tropospheric concentration. As described in the thesis introduction, natural ozone concentrations are cyclical, with the highest concentrations observed during the daylight hours and the lowest concentrations at night. At SoyFACE, this natural cycle is mimicked by only running the ozone generators during a 9-hour daytime period, and allowing the rings to settle back to the ambient concentration at night.

Unlike plants grown in a growth chamber or in open-top pots, the plants at the SoyFACE facility are exposed to all of the elements - including rain, hail, extreme temperature fluctuations and insect infestations - that a normal soybean crop would experience. While these competing factors can make the final experimental statistics more difficult to deconvolute, they nonetheless provide a more accurate picture of how the sum of expected elements, including ozone exposure, affects crop yield and protein expression.

Life Stages of a Soybean Plant

As for *Arabidopsis*, the life of a soybean plant has been divided into a series of defined stages [TAIR; http://www.arabidopsis.org/portals/education/growth.jsp]. These stages are divided into two sets: the "V" stages which mark periods of vegetative growth, and the "R" stages which chronicle the emergence of the reproductive organs [http://www.ag.ndsu.edu/pubs/plantsci/rowcrops/a1174/a1174w.htm#Growth]. An outline of the stages is provided below:

Vegetative Stages

VE - seedling emergence

VC - cotyledons unfold

V1 - first trifoliate unfolds

V2 - second trifoliate unfolds

V3 - third trifoliate unfolds

V4 - fourth trifoliate unfolds

V5 - fifth trioliate unfolds

V6 - 6th trifoliate unfolds; flowering will begin if it has not already

Reproductive Stages

R1 - the first flower opens

R2 - all flowers are open or have opened

R3 - the first pod develops

R4 - pod development extends to the top nodes of the plant

R5* - seed development beings

R6 - at least one full-size seed is present; Senescence of the lowest leaves begins

R7 - pod browning begins

R8 - 95% of pods are brown.

* Note that vegetative growth continues even after reproductive growth begins. Thus, a plant that is in stage R5 may also be simultaneously in stage V11 or higher.

Author contributions: AG, LMH, and JMJ designed research; AG and RPK performed research; RPK and EAA contributed new reagents/analytical tools; AG analyzed data; and AG, LMH, and JMJ wrote the paper.

Classification: Biological Studies - Plant Biology

FROM CLIMATE CHANGE TO PROTEINS: REDOX PROTEOMICS OF OZONE-INDUCED RESPONSES IN SOYBEAN

Ashley Galant[a], Robert P. Koester[b], Elizabeth A. Ainsworth[b,c], Leslie M. Hicks[d], Joseph M. Jez[a,*]

[a]Department of Biology, Washington University, One Brookings Drive, Campus Box 1137, St. Louis, MO 63130 USA; [b]Department of Plant Biology, University of Illinois, Urbana-Champaign, Urbana, IL 61801 USA; [c]USDA-ARS Global Change and Photosynthesis Research Unit, Urbana, IL 61801 USA; and [d]Donald Danforth Plant Science Center, 975 North Warson Road, St. Louis, MO 63132 USA

The authors declare no conflict of interest.

[*]To whom correspondence may be addressed. E-mail: jjez@biology2.wustl.edu

ABSTRACT

Ozone (O_3) is an important atmospheric pollutant with respect to agricultural losses. Although O_3 affects a range of crops, soybean yield is extremely sensitive to this environmental oxidative stress. To understand metabolic alterations in response to chronic O_3 exposure, changes in the total and redox proteomes of soybean plants grown in the field at the Soybean Free Air Concentration Enrichment (SoyFACE) facility under ambient (40 ppb), moderate (60 ppb), and high (115 ppb) O_3 levels were examined. The changes in the total and redox proteomes of soybean leaf tissue exposed to chronic high O_3 levels are more widespread and not the same as those resulting from short-term acute O_3 exposure. Compared to the ambient control, the 115 ppb O_3 leaf sample contained 35 proteins that increased up to 5-fold in expression level, 22 proteins that were up to 5-fold more oxidized without changes in expression levels, and 22 proteins that increased in total expression level and became 2- to 9-fold more oxidized. These changes occur in proteins across carbon metabolism, photosynthesis, amino acid metabolism, specialized metabolism of flavonoids and isoprenoids, signaling & homeostasis, antioxidant responses, protein degradation, and nucleic acid pathways. Our data directly demonstrates that O_3 exposure in plants changes the oxidation states of multiple proteins across metabolic pathways, and may provide a snapshot of metabolic adaptation to long-term field growth under chronic O_3 stress. Understanding how environmental O_3 affects redox-sensitive pathways will aid in the development of crops better adapted to global climate change.

INTRODUCTION

Global climate change and air pollution pose significant challenges to agriculture and food production worldwide (1). In the Northern hemisphere, tropospheric ozone (O_3) is a major pollutant that affects agriculture yields of multiple crops (2-4). Since the 19th century, ground O_3 levels have doubled with tropospheric concentrations in industrialized nations rising 0.5-2.5% per year, and major crop growing regions of the United States, India, and China facing more rapid changes of up to 10% per year (5-7). The current global mean O_3 level of ~60 ppb and higher localized concentrations are already above the established 40 ppb threshold for crop losses (2, 8-9). Climate models predict that mean surface O_3 concentrations may rise 20-25% globally by 2050 with levels in India and south Asia reaching comparable levels by 2020 (10-12). Understanding how crops respond to increasing O_3 pollution (and other environmental stresses) is essential for meeting the growing demands for sustainable food systems as the world faces increasing population, urbanization, and climate changes.

The negative effects of O_3 on crop yield are well documented from both short-term acute exposure studies and long-term chronic free-air concentration enrichment (FACE) experiments (2-4; 12-16). Among major food crops, soybean (*Glycine max*) is one of the most sensitive to atmospheric O_3 levels, which can vary between 50-120 ppb during summer days (17-18). At concentrations as low as 40 ppb, soybean growth and seed yield begin to decrease with even modest changes in O_3 levels significantly reducing crop production. For example, in FACE trials with soybean, a 13 ppb increase in O_3 from 56 to 69 ppb resulted in a 20% decrease in crop yield (13-14). Comparable reductions in

yield occur across multiple soybean varieties, suggesting that breeding for O_3 tolerance may be difficult. Economically, annual crop losses to O_3 damage at current tropospheric levels are estimated at $2-4 billion in the US and $3-5.5 billion in China, and will likely increase in the future (12).

As an environmental stress, O_3 acts as an oxidant in crop growth and results in visible necrotic damage, including chlorophyll loss and decreased seed yields in both mass and number (3, 14-21). At the molecular level, proteomic studies of rice, wheat, soy, and poplar exposed to acute, short-term O_3 stress in growth chambers reveal drastic reductions in the major leaf photosynthetic proteins and induction of defense/stress-related proteins (22-28). Although multiple physiological experiments indicate that acute and chronic ozone exposures do not induce the same damage mechanisms in plants (15-16), assessments of proteome changes have not examined crop plants grown in the field under chronic O_3 stress. Moreover, published studies do not probe the possible effect of O_3 on redox-sensitive proteins in plants, as these changes are not observable by standard proteomic methods (29-30).

Recently, we used a differential labeling and mass spectrometry-based approach (**Fig. S1**) to identify plant proteins that respond to changes in redox environment resulting from exogenous and endogenous oxidative stresses (30). Here we employ this method to assess the changes that occur in the total and redox proteomes of soybean in response to growth under chronic elevated O_3 levels in the field. Soybean plants were grown at the Soybean FACE (SoyFACE) facility (U. Illinois/USDA) under ambient (40 ppb), elevated (60 ppb), and high (115 ppb) atmospheric O_3 conditions. Soluble protein extracts from

root and leaf tissues were then isolated for analysis of changes in total and redox-proteomes using two-dimensional gel electrophoresis (2-DE), differential labeling, and nano-LC/MS/MS. The data presented here indicates that the changes in the total and redox proteomes of soybean leaf tissue resulting from chronic exposure to high O_3 levels are more widespread across metabolism than previously reported and are not necessarily the same as those resulting from short-term acute O_3 exposure. In addition, we provide the first direct demonstration that high O_3 exposure in leaf tissue alters the oxidation states of multiple proteins in different biochemical pathways. These changes may play a role in the metabolic adaptation to long-term field growth under chronic O_3 exposure.

RESULTS

Analysis of 2-DE spots in O_3-treated soybean

To identify O_3-responsive proteins in soybean, protein extracts of leaf and root tissue from plants grown at the SoyFACE facility under ambient, 60 ppb, and 115 ppb O_3 for were obtained (16). For each condition, protein extraction was performed in triplicate using tissue from three different plants. Extracted proteins were incubated with N-ethylmaleimide (NEM) to block free thiols, reduced with DTT, and then reacted with 5-iodoacetamidofluorescein (IAF) to label previously oxidized thiols (**Fig. S1**) (30). After 2-DE, gels were imaged for IAF signal, and then stained with SYPRO Ruby and imaged for total protein (**Fig. 1**). Three replicate gels for each of three independent samples were compared pairwise against the ambient gel images, and spots that significantly changed

in signal intensity identified. Across the 8 condition permutations (60 ppb or 115 ppb O_3; root or leaf tissue; and SYPRO or IAF), a total of 1455 significant spots were detected, of which 277 were differentially expressed and/or oxidized (**Table S1**). Spots were excised, trypsin digested, and analyzed by nano-LC/MS/MS. The resulting spectra were searched against the NCBInr database using an in-house version of MASCOT (**Tables S2-S5**). From this search, 57 spots contained a single protein match and 83 spots were identified as containing two or more proteins (**Table 1)**. The 115 ppb O_3 leaf sample had the largest numbers of identified proteins that changed in expression and/or oxidation state.

Identification of differentially expressed/oxidized proteins

In the identified spots, a total of 159 proteins were found to change in total expression and/or oxidation state (**Fig. 2A**). Of those proteins, 55, 27, 9, and 30 were unique to the 115 ppb O_3 leaf, 115 ppb O_3 root, 60 ppb O_3 leaf, and 60 ppb O_3 root samples, respectively. A further 38 proteins were found to change in multiple tissue-O_3 treatment combinations. Within each of the four tissue-ozone combinations varied numbers of proteins changed in total expression and/or oxidation state (**Figs. 2B-2E**). For example, the 115 ppb O_3 leaf sample contained a total of 79 unique proteins, of which 35 changed in total expression, 22 displayed altered oxidation state, and 22 changed in both expression and oxidation. For each protein identified, total expression level and/or oxidation state either increased or decreased in the treated tissue relative to the control (**Tables S2-S5**). While many proteins were localized to a single spot, some other proteins could be found in multiple spots, suggesting the presence of multiple isoforms. Where

these isoforms could not be distinguished from one another via the database search, the fold change (and associated spot number) is reported for the parent protein.

High O_3-induced changes the total and redox proteomes of soybean leaf

Comparison between the four tissue-O_3 treatment combinations revealed two distinct trends in the high O_3 leaf sample compared to the other samples - increased expression levels and oxidation of the largest number of proteins (**Fig. 3 and Fig. S2**). Although analysis of the 115 ppb O_3 root and 60 ppb O_3 leaf and root samples showed multiple proteins either increasing or decreasing in expression level, these changes were generally less than 2-fold different (**Fig. S2**). Moreover, both the numbers of protein changes and the fold changes in these samples were generally less than those observed in the high O_3 leaf sample. For example, in the 115 ppb O_3 root sample, the expression of 6 proteins increased and 10 proteins decreased, and in the 60 ppb O_3 leaf and root samples, fewer proteins increased in expression than decreased. Overall, the fold changes in total protein levels observed in the high O_3 root, elevated O_3 leaf and root samples were comparable to those described in previous studies of plant proteomes following acute O_3 exposure in growth chambers (22-28).

In stark contrast, the 115 ppb O_3 leaf sample contained 35 proteins with up to 5-fold increased expression compared to ambient samples, and only 2 proteins with ~1.5-fold decreased levels (**Fig. 3**). Even more striking was the shift in proteins that increased in oxidation in the high O_3 leaf sample compared to the other three samples. 22 proteins increased in total expression level and became 2- to 9-fold more oxidized. In addition, 22

other proteins became up to 5-fold more oxidized without significant changes in expression levels (**Fig. 3**). In comparison, the high O_3 root sample had 11 proteins that increased in oxidation and 8 that were more reduced (**Fig. S2A**). In the 60 ppb O_3 tissue samples, only a handful of proteins were either more oxidized or reduced than controls (**Figs. S2B & S2C**). The observed changes in the total and redox proteomes of leaf tissue exposed to high O_3 occurred across a range of metabolic pathways (**Fig. 4**), including redox systems, carbon metabolism, photosynthesis, signaling & homeostasis systems, amino acid metabolism, specialized metabolism of flavonoids and isoprenoids, protein degradation, and nucleic-acid systems, and are discussed in more detail later.

Analysis of enzymatic activities in leaf tissue exposed to high O_3-treatment

To better examine the activity changes in the 115 ppb O_3 leaf sample, targeted assays of the glycolytic/Calvin cycle enzymes phosphoglycerate kinase (PGK), fructose 1,6-bisphosphate aldolase (FBA), and glyceraldehyde-3-phosphate dehydrogenase (GAPDH), malate dehydrogenase (MDH) in the citric acid cycle, and glutamine synthetase (GS) in amino acid synthesis were performed. All of these enzymes showed increased protein expression and ~5-fold oxidation changes in the 115 ppb O_3 leaf sample. Except for PGK, each enzyme exhibited increased activity in the high O_3 leaf sample compared to the ambient control with the fold change in enzymatic activity correlated to increased expression level (**Table 1**).

In addition to these enzymes, the activity and total expression of ribulose 1,5-bisphosphate carboxylase oxygenase (RuBisCO) and phosphoenolpyruvate carboxlase

(PEPC) were examined because O_3 exposure can alter levels of these proteins in plants (2-3; 13-17; 19-21). For RuBisCO, both activity assays (**Table 1**) and Western blot analysis of the large subunit (**Fig. S3**) showed no significant difference between the high O_3 leaf sample and the ambient control. Likewise, expression of PEPC in the ambient and high O_3 leaf samples, as determined by Western blot, was not altered (not shown).

In the 115 ppb O_3 leaf sample, a glycosyl hydrolase/chitinase showed 4- and 9.4-fold increases in expression and oxidation state, respectively. Glycosyl hydrolases are involved in the degradation of various sugars, but are also mechanistically related to chitinases, which cleave glycosidic bonds and are typically involved in pathogen responses to insects or herbivory. Because many glycosyl hydrolases display varied activities, a fluorescence-based assay was used to evaluate exochitinase, endochitinase, and chitobiosidase activity in control and high O_3 leaf tissues, which were shown to increase 1.6-, 4.1-, and 11.1-fold, respectively (**Table 2**). It is unclear if these changes result from the identified protein or from aggregate changes in multiple glycosyl hydrolases.

DISCUSSION

Understanding the molecular mechanisms and metabolic consequences of how global climate changes, such as elevated tropospheric O_3 levels, impact crop plants is essential for efforts to maintain crop performance under increasing environmental stresses. Although earlier growth-chamber studies describe the effects of acute O_3 exposure on the proteomes of different crops, including soybean (22-28), these reports have neither

reported on the consequences of chronic O_3 exposure under field conditions nor examined the effect of O_3 on redox-sensitive proteins. Previous studies, in which soybean, rice, or wheat were exposed to constant 120 to 200 ppb O_3 for 3 to 5 days in growth chambers, typically identified 20 to 50 proteins that changed in expression level, either up or down (22-28). Analysis of the total and redox proteomes of leaf and root tissues from soybean plants grown in the field at SoyFACE under elevated (60 ppb) chronic daytime exposure to O_3 showed similar changes in both number of proteins and expression levels to earlier growth-chamber experiments (**Figs. 2 & S2**). In addition, leaf and root samples of soybean grown under elevated O_3 stress exhibited less than 2-fold differences in redox state for only a handful of proteins (**Fig. S2**). In contrast, the high O_3 soybean leaf sample analyzed here displayed striking increases in both expression levels and/or oxidation of multiple proteins across different metabolic pathways (**Figs. 3 & 4**). This suggests that there is a cross-over point between 60 and 115 ppb O_3 at which the expression levels and oxidation state of multiple proteins in leaf tissue dramatically shift, potentially as a metabolic adaptation to long-term field growth under chronic O_3 exposure.

Tropospheric O_3 negatively affects soybean growth and yield at concentrations greater than 40 ppb with decreased shoot and pod biomass, fewer pods produced, and premature leaf senescence (2-4; 12-18). O_3 enters leaves through the stomata and produces reactive oxygen species (ROS) that subsequently oxidize the plasma membrane and photosystem components leading to degradation of chlorophyll (2-3). Physiologically, long-term O_3 stress leads to reduced photosynthesis and mobilization of

reserve energy stores by converting leaf starch to sugars (28). Accordingly, sugar catabolism increases and previous studies identified several primary metabolic proteins, including MDH and phosphoglycerate mutase, as highly expressed under acute O_3 stress (23). The data presented here indicates that the changes in total and redox proteomes of soybean leaf tissue resulting from chronic exposure to high O_3 levels are more widespread across metabolism than previously reported and not necessarily the same as those resulting from acute O_3 exposure (**Fig. 4**).

Decreased photosynthetic efficiency, reduced RuBisCO activity, and elevated PEPC activity are classic markers for O_3 damage and senescence; however, these effects vary with the length, concentration, and type of exposure (2-3; 13-17; 19-21). Here soybean tissues were harvested at the R3 stage before significant changes in photosynthesis were observed (13; 16). The increased expression of chlorophyll a/b-binding protein, ferredoxin reductase, and a chlorophyllase-like protein observed in the 115 ppb O_3 leaf sample (**Fig. 4**) may help maintain photosynthesis at this growth stage before ozone-induced senescence occurs. Similarly, RuBisCO (large and small subunits), RuBisCO activase, a RuBisCO-associated protein, and RuBisCO-binding protein displayed elevated expression and/or oxidation in the 115 ppb O_3 leaf sample. Moreover, proteins related to iron homeostasis (ferredoxin reductase and ferritin) also change in soybean leaf under high O_3 stress. Because the spots containing RuBisCO included multiple proteins, activity assays and Western blot analysis were used to further examine possible changes in activity. Both methods showed no significant alteration in RuBisCO at this stage of soybean growth (**Table 1** and **Fig. S3**). Likewise, the effect of chronic 115 ppb O_3

exposure on PEPC levels in leaf tissue was analyzed by Western blot, which indicated no significant change in expression compared to ambient O_3 exposure at the time of harvest. These results suggest that major alterations in soybean photosynthesis are likely linked to senescence and occur later in the growing season for plants under chronic O_3 exposure.

Multiple proteins (i.e., PGK, GAPDH, FBA, ribose-5-phosphate isomerase, phosphoribulokinase, triosephosphate isomerase, MDH, and isocitrate dehydrogenase) in the reduction and regeneration phases of the Calvin cycle, glycolysis, and the TCA cycle increase in expression and/or oxidation state in the high O_3 leaf sample (**Fig 4**). In addition, the total activity levels of FBA, GAPDH, and MDH increased in the 115 ppb O_3 leaf sample (**Table 1**). This is consistent with earlier proteomic studies (22-28), but the analysis here indicates a wider range of protein changes and for the first time identifies redox-state alterations resulting from an environmental oxidative stress. All of these proteins are known to interact with thioredoxin, which is essential for maintaining the protein redox-state in plants (29). Moreover, phosphoribulokinase and GAPDH form a protein complex via thioredoxin-mediated redox changes in response to light intensity (31). O_3-related changes in cellular oxidation state may affect this interaction. Similarly, MDH is a critical regulatory point in the TCA cycle; however, the cytosolic form of the enzyme is redox-regulated and inactivated under oxidizing conditions (32). In the 115 ppb O_3 leaf sample, MDH had 2- and 5-fold higher expression and oxidation compared to controls with a 1.3-fold increase in total activity (**Table 1**). It is possible that these changes reflect the need to maintain MDH in the leaf to supply metabolites to the TCA cycle.

In addition to changes in core carbon metabolism, the starch and sugar mobilization pathways (phosphohexomutase, glucanase, and a glycosyl hydrolase/acid chitinase), the glycerate and glycolate pathways, and the biosynthesis of isoprenoids, carotenoids, and (iso)flavonoids display increased expression and/or oxidation state in the high O_3 leaf sample (**Fig. 4**). The changes in enzymes involved with the conversion of starch to sugar are consistent with a shift in energy demands of crops under O_3 stress (23). Of the 79 proteins identified in the 115 ppb O_3 leaf sample, a protein annotated as a glycosyl hydrolase/acid chitinase undergoes the greatest fold changes in both expression and oxidation state (**Fig. 3**). Although it is unclear if this protein functions in cell wall degradation, pathogen response, or sugar mobilization, the overall activity of glycosyl hydrolases in soybean leaf increase 1.6- to 11-fold in the high O_3 samples (**Table 1**). These increases may be connected to the mobilization of starch for energy production.

Proteins in three specialized metabolic pathways related to O_3 stress were also identified in the 115 ppb O_3 leaf sample. In the isoprenoid synthesis pathway, deoxyxylulose phosphate (DXP) oxidoreductase and isopentenyl diphosphate (IPP) isomerase are oxidized (**Fig. 4**); however, the effect of oxidation of these proteins remains to be determined. Interestingly, volatile isoprenoid emissions, including isoprene and monoterpenes, may act as an ozone protection mechanism in plants (33). Moreover, changes in carotenoid and (iso)flavonoid pathways (cartenoid-associateed protein, chalcone isomerase, isoflavone reductase, and caffeoyl-CoA methyltransferase) suggest alterations in the synthesis of these compounds, which act as photoprotective compounds and anti-oxidants (34).

In leaves exposed to high O_3 levels, up-regulation and/or oxidation of proteins in amino acid biosynthesis and/or nitrogen homeostasis were also observed (**Fig. 4**). The cytosolic form of GS, a central player in nitrogen sensing, increased in expression and oxidation in the 115 ppb O_3 leaf sample, which also corresponded with increased total activity compared to controls (**Table 1**). Elevated expression of cytosolic GS is associated with leaf senescence and the recycling of ammonia during stress conditions (35). Aspartate-semialdehyde dehydrogenase is the primary control point for the biosynthesis of isoleucine, methionine, lysine, and threonine. Although redox-control has not been described for the plant enzyme, reversible oxidation of a catalytic cysteine in the bacterial homolog alters activity (36). Likewise, carbamoyl phosphate synthetase, which becomes more oxidized following high O_3 exposure in soybean, is also sensitive to changes in redox environment (37). Also related to nutrient metabolism, the observed expression and oxidation changes in 14-3-3 proteins may further modify the activities of enzymes across the carbon, nitrogen, and sulfur nutrient assimilation pathways and/or signal transduction systems linked to stress responses (38-39).

The proteomic analysis here supports studies demonstrating that redox-protection mechanisms play a critical role in plant responses to O_3 exposure, and for the first time directly demonstrates that O_3 exposure changes the oxidation states of multiple proteins in different metabolic pathways. High chronic O_3 exposure leads to an oxidative stress that activates redox protection mechanisms in plants and increases expression and/or oxidation of proteins in those systems, including ascorbate peroxidase, methionine sulfoxide reductase, and glutathione-S-transferases (GSTs) (**Figs. 3 and 4**). In plants, the

ascorbate-glutathione system is critical for maintaining redox homeostasis and for scavenging ROS produced by photosynthesis (40). As such, the increased expression and oxidation of ascorbate peroxidase, which is critical in this system, is directly linked to cellular responses to attenuate oxidative stress induced by high O_3 exposure in leaf tissue. This is also linked to increased mobilization of sugar stores, which can further enhance ascorbate synthesis (41). Thus, the up-regulation of glucose catabolism is linked to energy production and the generation of reducing equivalents for the detoxification of ROS. The nearly 3-fold increase in oxidation of methionine sulfoxide reductase, which targets oxidized methionine residues (42), suggests an important role for this protein in responding to O_3 stress. Moreover, the reaction mechanism for the fungal methionine sulfoxide reductase proceeds through the formation of a thioredoxin-mediated intramolecular disulfide bond (43); however, it is unclear if the activity of the plant enzyme is redox-responsive. Likewise, GST isoforms were detected as changing in expression and/or oxidation across several different tissue-type/ozone concentration combinations (**Tables S2-S5**). In plants, GSTs comprise a large family of enzymes that conjugate glutathione to either small molecules or proteins for xenobiotic detoxification and redox-modifications (44).

Although the role of O_3 as an oxidative environmental stress is well established (2-4, 7), the extent of redox-linked changes in crop plants in the field faced with chronic exposure to high O_3 concentrations has not been examined previously. Analysis of the soybean redox proteome in the 115 ppb O_3 leaf sample revealed 44 proteins with 2- to 9-fold higher oxidation than in control samples. This work is the first report that O_3-

exposure is directly linked to redox changes in plant proteins. Given that 2-DE methods were used, the observed changes in the redox proteome of soybean leaf are likely only a small fraction of the total number of proteins that change in oxidation state. Future targeted efforts using more sensitive isolation/detection strategies promise to reveal a greater extent of redox-linked changes resulting from O_3 stress. Because changes in redox-state of plant proteins can drastically alter activity in response to environmental and cellular stresses, further work is also required to examine how oxidative stresses modulate protein activity across plant metabolism. From a physiological perspective, a better understanding of how above and below ground metabolisms alter is also required. O_3 enters leaves via the stomata, but alterations in the expression and/or oxidation state of proteins in root tissues were observed, albeit not at the same intensity as in leaf tissue. Nevertheless, it is unclear if these result from systemic changes in metabolism or directly from O_3 exposure in roots. Ultimately, understanding how environmental ozone affects redox-sensitive pathways will aid in the development of crops better adapted to global climate change and provide information about how to target the engineering of ozone protection systems.

MATERIALS AND METHODS

Plant Growth and SoyFACE O_3 Treatment. Soybean (*G. max* (L.) Merr) were planted and exposed to ambient (40 ppb), elevated (60 ppb), and high (115 ppb) O_3, as described previously (16). All O_3 levels are seasonal 9-hour average concentrations. O_3 was not added at night or when leaves were wet.

Soluble Protein Extraction. Protein extraction was performed in triplicate using tissue from three different plants. All steps were carried out at 4 °C, unless otherwise indicated. Approximately 800 mg of root or 400 mg of leaf tissue were frozen in liquid nitrogen and ground to a fine powder using a mortar and pestle. The tissue was suspended in extraction buffer [100 mM Tris-HCl, pH 8.0; 100 mM NEM; 1% CHAPS; 1% plant protease inhibitors (Sigma)] to 200 mg mL^{-1}. For lysis, samples were sonicated (3 x 15 sec) with the tissue mixed and iced (30 sec) between sonications. Samples were then centrifuged (16,000 x g; 15 min) to precipitate insoluble debris. The resulting supernatant was mixed with 4 volumes of pre-chilled methanol and stored on ice for 30 min. Samples were again centrifuged (16,000 x g; 15 min) and the protein pellet harvested. The pellet was washed twice with methanol (2 x 30 min) and centrifuged (16,000 x g; 5 min) after each wash. During washing, care was taken to periodically disrupt the pellet to ensure complete removal of excess NEM. After the final spin, the pellet was air-dried and resuspended in 150 μL of reduction buffer (50 mM Tris-HCl, pH 8.0, 7 M urea, 2 M thiourea, 50 mM DTT) per 200 mg of starting tissue. The suspension

was incubated for 15 min at room temperature and then re-precipitated with 4 volumes of pre-chilled methanol. As described above, the protein was washed 3 times (3 x 30 min) with the same volume of methanol and centrifuged after each wash (1 x 15 min; 2 x 5 min). After the final spin, the pellet was air-dried and resuspended in 150 µL of labeling buffer [40 mM HEPES, pH 7.5; 50 mM NaCl; 200 µM IAF] per 200 mg of starting tissue. Due to the light sensitivity of IAF, this and all subsequent steps were carried out under dim light. The suspension was incubated for 10 minutes at room temperature, and then re-precipitated with 4 volumes of pre-chilled methanol. Again, the protein was washed 3 times (3 x 30 min) with the methanol and centrifuged after each wash (1 x 15 min; 2 x 5 min). After the final spin, the pellet was air-dried and resuspended in a small volume (~50 µL per 200 mg of starting tissue) of DeStreak buffer (GE Healthcare). Protein concentration was determined by CB-X assay (G-Biosciences).

Protein Separation by 2-DE. As above, all steps were carried out under dim light. To begin, 200 µg of extracted protein was resuspended to a total volume of 180 µL in DeStreak buffer and absorbed into a pH 4-7 gel strip. Isoelectric focusing in the first dimension was carried out at room temperature in a Proteon IEF cell using a four-step method: 1) 250 V, linear increase, 30 min; 2) 500 V, linear increase, 1 hr; 3) 8000 V, linear increase, 2.5 hr; 4) 8000 V, rapid increase to 35,000 until complete. Separation by molecular weight was in the second dimension at 150 V until the dye front reached the gel edge. Gels were then removed from their cassettes and imaged (λ_{ex}=488 nm and λ_{em}=520 nm) using a Typhoon 9410 (GE Healthcare) to detect IAF-labeled proteins.

After imaging, gels were bathed for 30 min in 100 mL of fixing solution [10% methanol; 7% glacial acetic acid] using an orbital shaker. The fixing solution was removed and the gels bathed overnight in 50 mL of Sypro Ruby (Biorad) protein stain. After staining, the gels were again bathed for 30 min in 100 mL of fixing solution. Following a wash step with MilliQ water to remove any excess stain, the gels were imaged again (λ_{ex}=457 nm and λ_{em}=610 nm) to detect total protein. For all gel replicates within a given set or sets to be compared, the same laser intensity (400 V for IAF images and 650-800 V for Sypro images) was used. After imaging was complete, gels were stored in MilliQ water at 4 °C until needed.

Gel Analysis, Spot Extraction and Mass Spectrometry. Replicate gels images were aligned using Progenesis Samespots (Nonlinear Dynamics). Further alignment of replicate control and ozone-treatment gel images was carried out for each pairwise comparison. In order to quantify expression and thiol composition differences between the control and treated samples, spot volume was calculated and normalized for each spot in the aligned images. Those spots that different significantly in volume (ANOVA, $p<0.05$) between the averaged control and ozone treatment gels were then marked for identification.

Excision of significant spots from their respective gels was performed using a Gelpix System (Genetix) under high humidity (>85%) to prevent gel distortion or tearing. Gel plugs were dehydrated with 200 μL of acetonitrile (ACN) using a room-temperature table-top shaker (15 min; 900 rpm). ACN was then removed and the plugs washed five

times with 200 μL of 50mM NH_4HCO_3; 50% ACN (15 min; 900 rpm) with the liquid discarded after every wash to remove the Sypro Ruby stain. Following the last NH_4HCO_3/ACN wash, the plugs further were washed with 100 μL of ACN (15 min; 900 rpm). After this, the liquid was discarded and the plugs air-dried. Once dry, the plugs were submerged in 20 μL of trypsin digestion buffer (50mM NH_4HCO_3 with 6 ng μL^{-1} trypsin) and rehydrated overnight at 37 °C. Next, 30 μL of 1% formic acid; 2% ACN was added to the digests, which were then shaken (30 min; 900 rpm). Following this wash, the supernatant was saved, the plug transfered to a new tube containing 24 μL of 60% ACN, and the tube shaken (30 min; 900 rpm). After this final wash, the supernatant was removed and added to that collected during the previous step, and the plugs discarded. The combined digest from each gel plug was lyophilized to dryness and then resuspended in 7 μL of 1% formic acid; 5% ACN.

Peptide separation and analysis were carried out as previously described (30). LC-MS/MS was conducted via an Eksigent nanoLC with a Dionex C18 PepMap100 column (75 μM id) coupled to a QSTARR XL MS/MS-TOF (Applied Biosystems) The peptide tandem mass spectra were processed using Analyst QS v1.1 (AB Sciex) and searched against the NCBInr database (July 2010, 11368323 sequences) using an in-house version of MASCOT v2.20 (Matrix Science) with the following parameters: tryptic peptides with ≤1 missed cleavage site; precursor and MS/MS fragment ion mass tolerances of 0.8 and 0.8 Da, respectively; variable carbamidomethylation and fluoresceination of cysteine; and variable oxidation of methionine. The data was filtered using Scaffold 3 (Proteome

Software). Positive identification criteria was ≥2 peptide sequences, protein probability of 99.9%, and peptide probability of 80%.

Enzyme Assays. Standard spectrophotometric assays were used to determine activity of PGK (45), MDH (46), GS (47), GAPDH (48), FBA (49), and RuBisCO (50). To measure chitinase activity, a fluorescence-based kit (Sigma, CS1030) was used. Tissue extracts were prepared as above, and equal amounts of control and treated tissue extracts were added to tubes containing assay-appropriate buffer plus 1% plant protease inhibitors (Sigma).

ACKNOWLEDGEMENTS

This research was funded by the National Science Foundation (MCB-0904215 to J.M.J. and L.M.H.). A.G. was supported by an American Society of Plant Biologists Pioneer Hi-Bred Graduate Research Fellowship.

REFERENCES

1. FAO (2005) The State of Food Insecurity in the World. Food and Agriculture Organization of the United Nations, Rome.
2. Heagle AS (1989) Ozone and Crop Yield. *Annu Rev Phytopathol* 27: 397-423
3. Fiscus EL, Booker FL, Burkey KO (2005) Crop responses to ozone: uptake, modes of action, carbon assimilation, and partitioning. *Plant Cell Environ* 28: 997-1011
4. Ainsworth EA, Rogers A, Leakey ADB (2008) Targets for crop biotechnology in a future high-CO_2 and high-O_3 world. *Plant Physiol* 147: 13-19
5. Staehelin J, Thudium J, Buehler R, Volzthomas A, Graber W (1994) Trends in surface ozone concentrations at Arosa (Switzerland). *Atmos Environ* 28: 75-87
6. Vingarzan R (2004) A review of surface ozone background levels and trends. *Atmos Environ* 38: 3431-3442
7. Ashmore M, Toet S, Emberson L (2006) Ozone – a significant threat to future world food production? *New Phytol* 170: 201-204
8. Fuhrer J, Skärby L, Ashmore MR (1997) Critical levels for ozone effects on vegetation in Europe. *Environ Pollut* 97: 91-106
9. Mills G, Hayes F, Buse A, Reynolds B (2000) in Annual Report 1999/2000 of UN/ ECE ICP Vegetation, Center for Ecology and Hydrology, Bangor, ME
10. Prather M, Ehhalt D, Dentener F, Derwent R, Dlugokencky E, Holland E, Isaksen I, Katima J, Kirchoff V, Matson P, Midgley P, Wang M (2001) in *Climate Change 2001: the Scientific Basis. Contribution of Working Group I to the Third Assessment Report*

of the Intergovernmental Panel on Climate Change, eds Houghton JT et al. (Cambridge, UK), pp 238-287

11. Dentener F, Stevenson M, Cofala J, Mechler R, Amann M, Bergamaschi P, Raes F, Derwent R (2005) The impact of air pollutants and methane emission controls on tropospheric ozone and radiative forcing: CTM calculations for the period of 1990-2030. *Atmos Chem Phys* 5: 1731-1755

12. Van Dingenen R, Dentener FJ, Raes F, Krol MC, Emberson L, Cofala J (2009) The global impact of ozone on agricultural crop yields under current and future air quality legislation. Atmos Environ 43, 604-618

13. Morgan PB, Bernacchi CJ, Ort DR, Long SP (2004) An in vivo analysis of the effect of season-long open-air elevation of ozone to anticipated 2050 levels on photosynthesis in soybean. *Plant Physiol* 135: 2348-2357

14. Morgan PB, Miles TA, Bollero GA, Nelson RA, Longs SP (2006) Season-long elevation of ozone concentration to projected 2050 levels under fully open-air conditions substantially decreases the growth and production of soybean. *New Phytol* 170: 333-343

15. Chen CP, Frank TD, Long SP (2009) Is a short, sharp shock equivalent to long-term punishment? Contrasting the spatial pattern of acute and chronic ozone damage to soybean leaves via chlorophyll fluorescence imaging. *Plant Cell Environ* 32: 327-335

16. Betzelberger AM, Gillespie KM, McGrath JM, Koester RP, Nelson RL, Ainsworth EA (2010) Effects of chronic elevated ozone concentration on antioxidant capacity,

photosynthesis and seed yield of 10 soybean cultivars. *Plant Cell Environ* 33: 1569-1581

17. Krupa S, McGrath MT, Anderson CP, Booker FL, Burkey KO, Chappelka AH, Chevone BI, Pell EJ, Zilinskas BA (2001) Ambient ozone and plant health. *Plant Dis* 85, 4-12

18. Emberson LD, Buker P, Ashmore MR, Mills G, Jackson L, Agrawal M, Atikuzzaman MD, Cinderby S, Engardt M, Jamir C, Kobayashi K, Oanh K, Quadir QF, Wahid A (2009) A comparison of North American and Asian exposure response data or ozone effects on crop yields. *Atmos Environ* 43: 1945-1953

19. Baier M, Kandlbinder A, Golldack D, Dietz KL (2005) Oxidative stress and ozone: perception, signalling, and response. *Plant Cell Environ* 28,1012-1020

20. Chernikova T, Robinson MJ, Lee EH, Mulchi CL (2000) Ozone tolerance and antioxidant enzyme activity in soybean cultivars. *Photosynth Res* 64: 15-26

21. Robinson JM, Britz SJ (2000) Tolerance of field grown soybean cultivar to elevated ozone levels is concurrent with hight leaflet ascorbic acid level, higher ascorbate-dehydroascorbate redox status, and long term photosynthetic productivity. *Photosynth Res* 64: 77-87

22. Agrawal GK, Rakwal R, Yonekura M, Kubo A, Saji H (2002) Proteome analysis of differentially displayed proteins as a tool for

investigating ozone stress in rice (Oryza sativa L.) seedlings. *Proteomics* 2: 947-959

23. Bohler S, Bagard M, Oufir M, Planchon S, Hoffmann L, Jolivet Y, Hausman JF, Dizengremel P, Renaut J (2007) A DIGE analysis of developing poplar leaves subjected

to ozone reveals major

changes in carbon metabolism. *Proteomics* 7: 1584-1599

24. Feng YW, Komatsu S, Furukawa T, Koshiba T, Kohno Y (2008) Proteome analysis of proteins responsive to ambient and elevated ozone in rice seedlings. *Agr Ecosys Environ* 125: 255-265

25. Cho K, Shibato J, Agrawal GK, Jung YH, Kubo A, Jwa NS, Tamogami S, Satoh K, Kikuchi S, Higashi T, Kimura S, Saji H, Tanaka Y, Iwahashi H, Masuo Y, Rakwal R (2008) Integrated transcriptomics, proteomics, and metabolomics analyses to survey ozone

responses in the leaves of rice seedling. *J Proteome Res* 7: 2980-2998

26. Renaut J, Bohler S, Hausman JF, Hoffmann L, Sergeant K, Ahsan N, Jolivet Y, Dizengremel P (2009) The impact of atmospheric composition on plants: a case study of ozone and poplar. Mass Spectrom Rev 28: 495-516

27. Sarkar A, Rakwal R, Bhushan Agrawal S, Shibato J, Ogawa Y, Yoshida Y, Kumar Agrawal G, Agrawal M (2010) Investigating the impact of elevated levels of ozone on tropical wheat using

integrated phenotypical, physiological, biochemical, and proteomics approaches. *J Proteome Res* 9: 4565-4584

28. Ahsan N, Nanjo Y, Sawada H, Kohno Y, Komatsu S (2010) Ozone stress-induced proteomic changes in leaf total soluble and chloroplast

proteins of soybean reveal that carbon allocation is involved in adaptation in

the early developmental stage. *Proteomics* 10: 2605-2619

29. Buchanan BB, Balmer Y (2005) Redox regulation: a broadening horizon. *Annu Rev Plant Biol* 56: 187-220

30. Alvarez S, Galant A, Jez JM, Hicks LM (2011) Redox-regulatory mechanisms induced by oxidative stress in *Brassica juncea* roots monitored by 2-DE proteomics. *Proteomics* 11: 1346-1350

31. Howard TP, Metodiev M, Lloyd JC, Raines CA (2008) Thioredoxin-mediated reversible dissociation of a stromal multiprotein complex in response to change in light availability. *Proc Natl Acad Sci USA* 105: 4056-4061

32. Hara S, Motohashi K, Arisaka F, Romano PG, Hosoya-Matsuda N, Kikuchi N, Fusada N, Hisabori T (2006) Thioredoxin-h1 reduces and reactivates the oxidized cytosolic malate dehydrogenase dimer in higher plants. J Biol Chem 281: 32065-32071

33. Loreto F, Fares S (2007) Is ozone flux inside leaves only a damage indicator? clues from volatile isoprenoid studies. *Plant Physiol* 143: 1096-1100

34. Middleton EM, Teramura AH (1993) The role of flavonol glycosides and carotenoids in protecting soybean from ultraviolet-B damage. *Plant Physiol* 103: 741-752

35. Brugière N, Dubois F, Masclaux C, Sangwan RS, Hirel B (2000) Immunolocalization of glutamine synthetase in senescing tobacco (Nicotiana tabacum L.) leaves suggests that ammonia assimilation is progressively shifted to the mesophyll cytosol. *Planta* 211: 519-527

36. Alvarez E, Ramón F, Magán C, Díez E (2004) L-cystine inhibits aspartate-beta-semialdehyde dehydrogenase by covalently binding to the essential 135Cys of the enzyme. *Biochim Biophys Acta* 1696: 23-29

37. Hart EJ, Powers-Lee SG. Role of Cys1327 and Cys1337 in redox sensitivity and allosteric monitoring in human carbamoyl phosphate synthetase. *J Biol Chem* 284: 5977-5985

38. Oecking C, Jaspert N (2009) Plant 14-3-3 proteins catch up with their mammalian orthologs. *Curr Opin Plant Biol* 12: 760-765

39. Shin R, Jez JM, Basra A, Zhang B, Schachtman DP (2011) 14-3-3 proteins fine-tune plant nutrient metabolism. *FEBS Lett* 585: 143-147

40. Ishikawa T, Shigeoka S (2008) Recent advances in ascorbate biosynthesis and the physiological significance of ascorbate peroxidase in photosynthesizing organisms. *Biosci Biotechnol Biochem* 72: 1143-1154

41. Nishikawa F, Kato M, Hyodo H, Ikoma Y, Sugiura M, Yano M (2005) Effect of sucrose on ascorbate level and expression of genes involved in the ascorbate biosynthesis and recycling pathway in harvested broccoli florets. *J Exp Bot* 56: 65-72

42. Rouhier N, Vieira Dos Santos C, Tarrago L, Rey P (2006) Plant methionine sulfoxide reductase A and B multigenic families. *Photosynth Res* 89: 247-262

43. Antoine M, Boschi-Muller S, Branlant G (2003) Kinetic characterization of the chemical steps involved in the catalytic mechanisom of methionine sulfoxide reductase a from *Neisseria meningitidis*. *J Biol Chem* 278: 45352-45357

44. Dixon DP, Lapthorn A, Edwards R (2002) Plant glutathione transferases. *Genome Biol* 3: 3004

45. De BK, Kirtley (1977) Interaction of phosphoglycerate kinase with erythrocyte membranes. *J Biol Chem* 252: 6715-6720

46. Zeikus JG, Fuchs G, Kenealy W, Thauer RK (1977) Oxidoreductases involved in cell carbon synthesis of Methanobacterium thermoautotrophicum. *J Bacteriol* 132: 604-613

47. Kingdon HS, Hubbard JS, Stadtman ER (1968) Regulation of glutamine synthetase. XI. The nature and implications of a lag phase in the E. coli glutamine synthetase reaction. *Biochemistry* 7: 2136-2142

48. Bergmeyer HU, Gawehn K, Grassl M (1974) in *Methods of Enzymatic Analysis*, ed Bergmeyer HU (Academic Press, NY), pp 466-467

49. Robertson P, Fridovich I (1980) Continuous colorimetric monitoring of the fructose bisphosphate aldolase reaction. *Anal Biochem* 108: 332-334

50. Jordan DB, Ogren WL (1981) A sensitive assay procedure for simultaneous determination of ribulose-1,5-bisphosphate carboxylase and oxygenase activities. Plant Physiol 67: 237-245

Figure 1. Representative 2-DE Gels. The gels shown contain proteins isolated from roots of soybean exposed to 60 ppb O_3. (**A**) 2-DE gel visualized for IAF-labeling of the redox proteome. Lines and corresponding numbers indicate spots which significantly differed (p<0.05) in degree of oxidation as compared to control. (**B**) The same 2-DE gel from panel A, but with total proteome visualized with SYPRO ruby. Lines and corresponding numbers indicate spots which significantly differed (p<0.05) in total expression as compared to control.

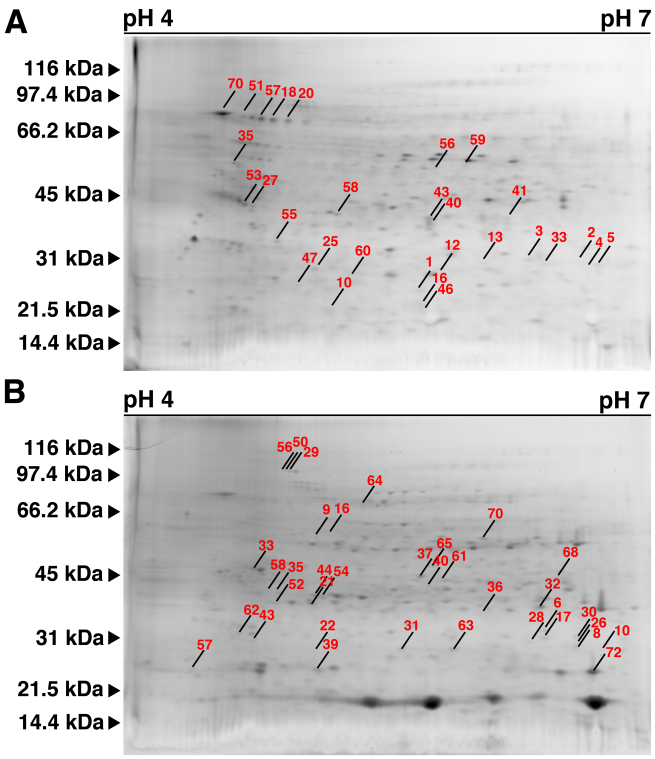


Figure 2. Venn diagram of proteins that differ in leaf and root tissues under 60 ppb and 115 ppb O_3 treatments compared to ambient conditions. (**A**) Distribution of proteins across all four combinations of tissue and O_3 concentration. Numbers in overlapping regions of the lobes indicate proteins found in more than one set of conditions. (**B-E**) Detailed breakdown of numbers of differentially oxidized (IAF) and/or expressed (SYPRO ruby) proteins between treated samples and controls. In each panel, the miniature Venn diagram in the top left corner indicates which lobe from panel A is analyzed.

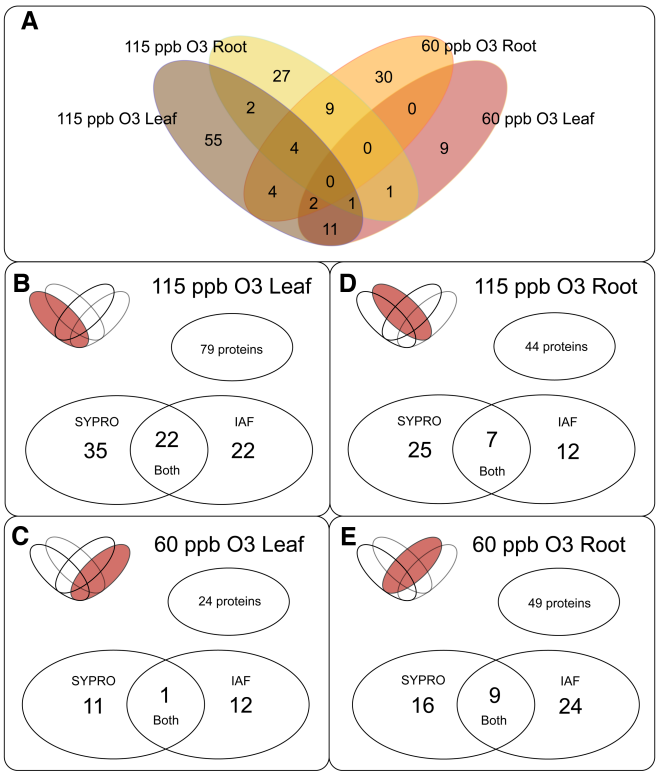


Figure 3. Summary of fold changes in total and redox proteomes of leaf tissue exposed to 115 ppb O_3. Fold changes, relative to ambient O_3 control, in oxidation state (IAF - fold change) and expression level (Sypro - fold change) for identified proteins identified are plotted. Names of representative proteins are shown with highly oxidized (orange box) and oxidized/expressed proteins (red) indicated. Additional information about the identified proteins is provided in Tables S2.

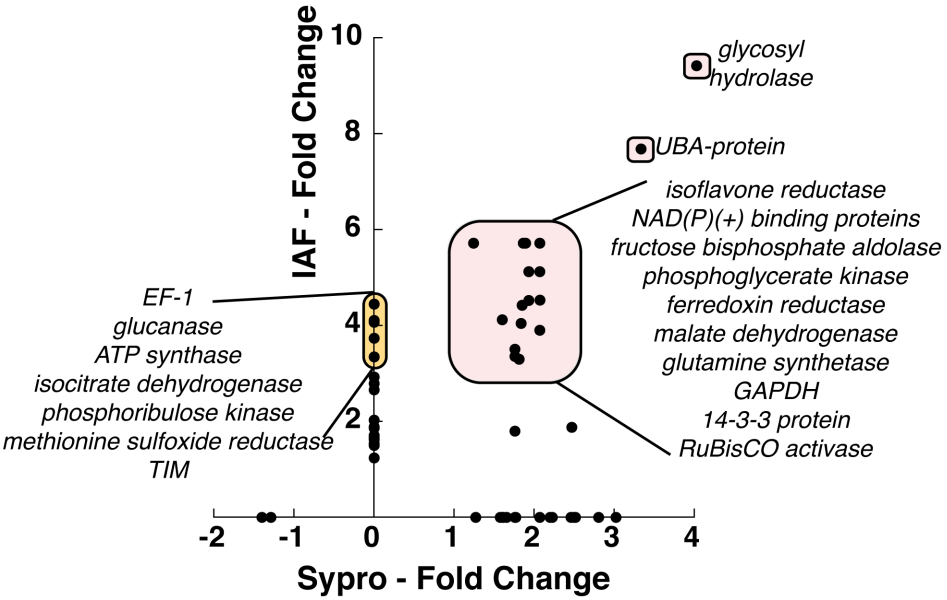


Figure 4. Metabolic Overview of Total and Redox Proteome Changes in Soybean Leaf Tissue Exposed to 115 ppb O_3. A schematic view of the different metabolic pathways identified is shown. Proteins that change in oxidation state (orange), expression level (red), or both oxidation state and expression level (red with black outline) are shown. Detailed information about the identified proteins is provided in Table S2.

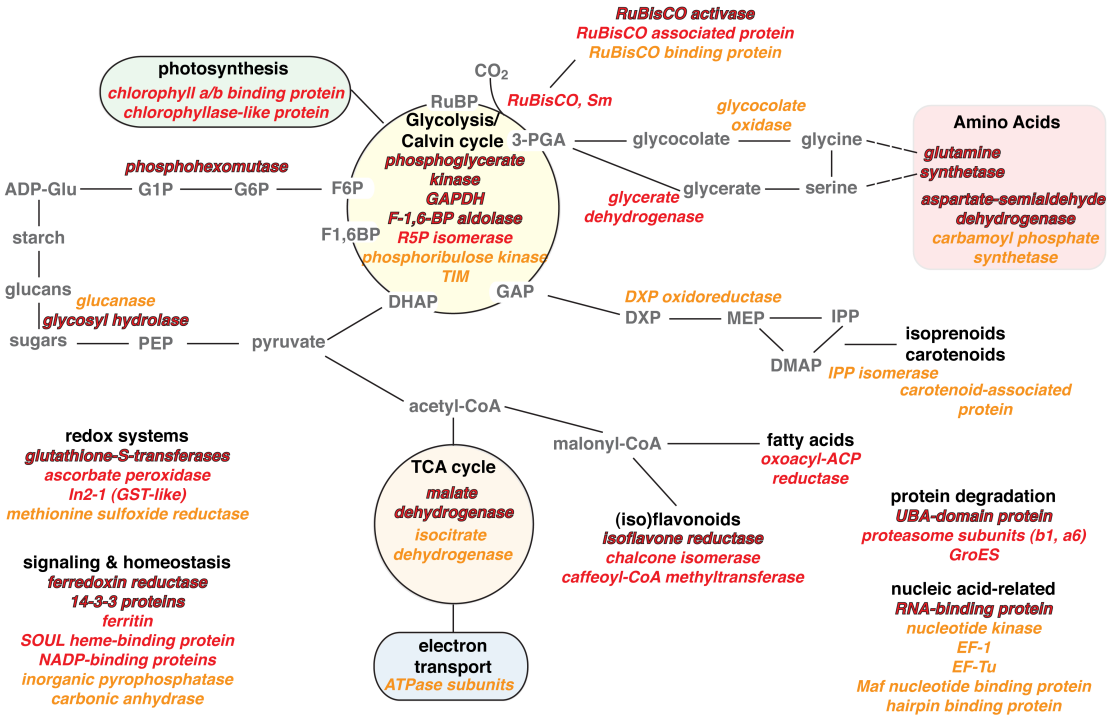


Table 1. Comparison of enzyme activities in leaf tissues exposed to ambient (40 ppb) and high (115 ppb) O_3. All assays were performed as described in the methods section. Values are averages ± standard deviations for n = 4-8. ND - no detected changes. Abbreviations are as used in the text.

	ambient ozone activity (μmol min^{-1} g FW^{-1})	high ozone activity (μmol min^{-1} g FW^{-1})	activity fold change	total protein (oxidation) fold change
PGK	715 ± 82	629 ± 68	0.9	1.8 (5.7)
FBA	31.6 ± 6.5	72.1 ± 9.9	2.3	1.8 (4.8)
GAPDH	183 ± 75	693 ± 301	3.8	1.7 (4.4)
MDH	72.0 ± 3.3	93.9 ± 10.4	1.3	2.0 (5.4)
GS	106 ± 9	246 ± 25	2.3	1.9 (4.5)
RuBisCO	11.1 ± 2.6	10.0 ± 4.1	0.9	ND (ND)
exochitinase	10.5 ± 0.8	17.2 ± 0.3	1.6	4.0 (9.4)
endochitinase	0.34 ± 0.02	1.40 ± 0.08	4.1	4.0 (9.4)
chitobiosidase	0.012 ± 0.046	0.133 ± 0.023	11.1	4.0 (9.4)

Supporting Information - Galant et al.

Figure S1. Redox Proteome Labeling Approach. Proteins with free thiols (-SH), disulfide bonds (-S-S-), or modified cysteines (-S-mod) are incubated with N-ethylmaleimide (NEM) to block free sulfhydryl groups. Oxidized thiols are reduced with dithiothreitol (DTT). The resulting free thiols are labeled with 5-iodoacetamidofluorescein (IAF), and then the proteins are separated by 2-DE and identified by LC-MS/MS.

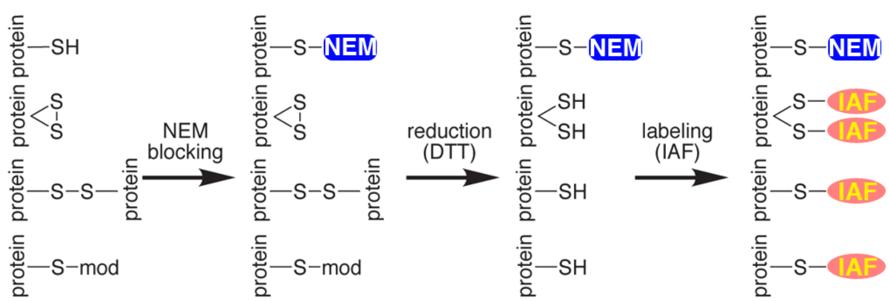


Figure S2. Additional Total and Redox Proteome Changes. Panels A-C show the fold changes in oxidation (IAF - fold change) and expression level (Sypro - fold change) relative to ambient controls for proteins identified by mass spectrometry in root and leaf tissues exposed to high and elevated O_3. Detailed information about the identified proteins is provided in Tables S2-S5.

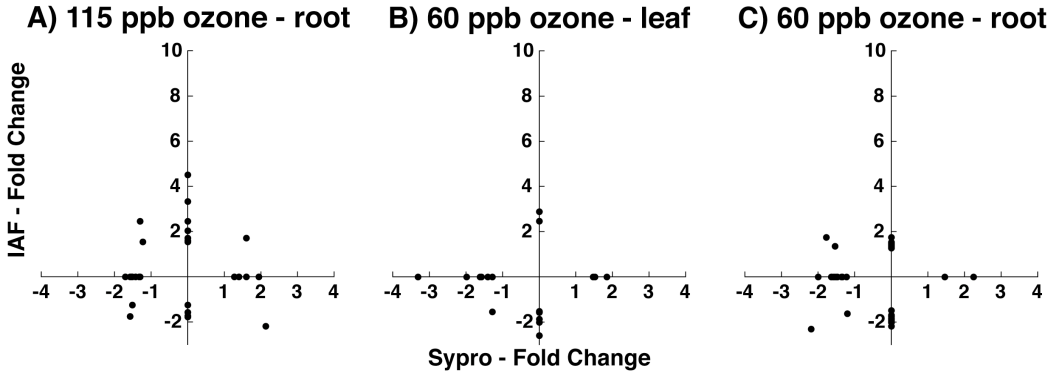


Figure S3. Immunoblot analysis of RuBisCO large subunit expression. Protein extracts from leaf tissue exposed to ambient (40 ppb) and high (115 ppb) O_3 were probed using anti-RuBisCO large subunit antibody. Lanes 1 and 3 contain 10 μg of total protein extract and lanes 2 and 4 contain 5 μg of total protein extract.

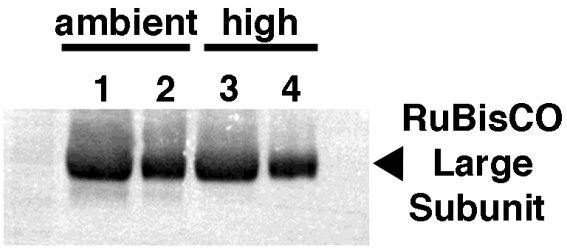


Table S1. Total number of spots detected and identified as either differentially expressed or oxidized across all experimental conditions. Differentially expressed/oxidized spots are further broken down into those with either single or multiple proteins.

O_3 Level, Tissue, and Signal	Spots Detected	Spots Differentially Expressed/Oxidized	Spots with One Protein	Spots with Multiple Proteins	Total Spots Identified
115 ppb, leaf Sypro	154	29	9	17	26
115 ppb, leaf IAF	171	47	9	18	27
115 ppb, root Sypro	208	36	4	10	14
115 ppb, root IAF	196	29	5	6	11
60 ppb, leaf Sypro	196	44	7	4	11
60 ppb, leaf IAF	195	43	7	3	10
60 ppb, root Sypro	158	21	7	10	17
60 ppb, root IAF	177	28	9	15	24
Total	1455	277	57	83	140

Tables S2-S5. Data Summary Tables. These tables list peptides that are differentially expressed and/or oxidized from 115 ppb O_3 leaf tissue (Table S2), 115 ppb O_3 root tissue (Table S3), 60 ppb O_3 leaf tissue (Table S4), and 60 ppb O_3 root tissue (Table S5). The accession number of the protein, molecular weight, fold changes, identified peptides, and mascot ion scores are indicated. For each protein, the fold change and spot number are listed as [(spot number) fold change]. The magnitude of the fold change associated with each protein is indicated by color in the box, as follows: 1.2- to 3-fold, pale green; 3- to 5-fold, medium green; >5-fold, dark green; -1.2- to -3-fold, pale blue; and -3- to -5-fold medium blue.

Supporting Information Table 1. List of Peptides Identified from 115 ppb Leaf Proteins that are Differentially Expressed/Oxidized

Protein	Accession Number(s)	Mol Weight	Fold Change IAF (spot number)	Fold Change SYPRO (spot number)	Classification	IAF Peptides (spot number)	Best Mascot Ion Score	SYPRO Peptides (spot number)	Best Mascot Ion Score
TIM phosphate binding protein (Glycine max)	gil255640328	14 kDa	[(35)3.347],[(55)2.375],[(59)2.260]		glycolysis	[(35)] AAQDALLFR SNSLAQLGK YTAEGESEEATR [(55)] AAQDALLFR ATPLQVADYTLK SNSLAQLGK YTAEGESEEATR [(59)] AAQDALLFR SNSLAQLGK	 56.3 60.8 70.3 67.8 64.4 66.1 87.9 71.2 59.3		
Nucleoside/nucleotide kinase (Glycine max)	gil255639590	50 kDa	[(35)3.347]		nucleotide metabolism	[(35)] EANPGNALLELR VVDPQLETVNK	 55.6 59.1		
Rossmann-fold NAD(P)(+)-binding protein; 3-oxoacyl-(acyl-carrier-protein) reductase (Populus triocarpa)	gil224100059 gil255547878 gil255638092	28 kDa		[(13)2.805]	fatty acid metabolism			[(13)] ILETIPLGR VESPVVVVTGASR	 57.7 60.2
Rossmann-fold NAD(P)(+)-binding protein (Glycine max)	gil255647108	42 kDa	[(21)4.533]	[(24)2.072]	unknown	[(21)] EADFSTDDIILGK FIGVFLSR	 76.1 54.8	[(24)] EADFSTDDIILGK FIGVFLSR	 76.1 54.8
Rossmann-fold NAD(P)(+)-binding protein (Glycine max)	gil255635535	36 kDa	[(10)5.718],[(21)4.533]	[(40)1.655], [(25)2.070], [(24)2.072]	unknown	[(10)] DVSFLTNLPGASEK LLEDGYAVNTTIR LVDAGFEFK [(21)] DVSFLTNLPGASEK LLEDGYAVNTTIR	 54.8 55.7 57.7 58.5 90	[(40)] DVSFLTNLPGASEK LLEDGYAVNTTIR [(25)] DVSFLTNLPGASEK LLEDGYAVNTTIR LVDAGFEFK [(24)] DVSFLTNLPGASEK LLEDGYAVNTTIR	 59 66.6 54.8 55.7 57.7 58.5 90
Rossmann-fold NAD(P)(+)-binding protein (Glycine max)	gil255642211	36 kDa		[(13)2.805]	unknown			[(13)] ALFSQITTR QNIGAADDVIVGDIR	 52.2 93.4
Rossmann-fold NAD(P)(+)-binding protein; possible isoflavone reductase (Glycine max)	gil255648230 gil2687724	36 kDa	[(10)5.718]	[(40)1.655], [(25)2.070]	secondary metabolism	[(10)] FFPSEFGLDVDR NLAQIDITVPPR VFIQGDGNVK	 51.7 60 55.4	[(40)] ILVLGPTGAIGR VFIQGDGNVK [(25)] FFPSEFGLDVDR NLAQIDITVPPR VFIQGDGNVK	 66.9 57.1 51.7 60 55.4
14-3-3 protein [Nicotiana tabacum]	gil15778154 gil26454611	29 kDa	[(89)1.667]		signal transduction	[(89)] ENFVYVAK ERENFVYVAK	 56.6 55.8		

glutathione-s-transferase [Glycine max]	gil2052029 gil220683633 gil255631159 gil38679415	25 kDa	[(96)1.545]		redox homeostasis	[(96)] FWADFVDNK LLEEQLGDK	54.2 75.8		
glutathione-s-transferase [Glycine max]	gil255637207	26 kDa		[(10)3.022]	redox homeostasis			[(10)] GVEYEYKEENLR KVYDYVVAVTK VYDYVVAVTK	84 49.9 58.1
proteasome beta type-1 subunit (Vitis vinifera)	gil225453909 gil255541320 gil255627685 gil255640620	25 kDa		[(20)2.226]	protein degradation			[(20)] MSTGYNILTR TVFASATER	69.7 50.1
ATP synthase beta subunit [Dioscorea communis]	gil17224743 gil21684883 gil37720945 gil42559020	52 kDa	[(26)4.107]		ATP-coupled proton transport	[(26)] ESGVINEKNIAESK LAIFETGIK	68.1 70.4		
RuBisCO-associated protein (Glycine max)	gil730692	31 kDa		[(16)2.460]	Calvin cycle			[(16)] ILAVANEYK VISDATASFPAK VLVSIGNK	63.5 47.4 48.4
RuBisCO subunit binding-protein alpha subunit, putative chloroplast precursor (Oryza sativa japonica)	gil108712217 gil125546535 gil134101 gil225436538 gil255587664 gil297734943 gil31193919 gil5650721	50 kDa	[(87)1.711]		Calvin cycle	[(87)] GILNVAAIK GYISPQFVTNPEK VGAATETELEDR	49.4 50.4 94.5		
beta-form rubisco activase (Glycine max)	gil255635315	49 kDa		[(46)1.600], [(51)1.549]	Calvin cycle			[(46)] GLAYDISDDQQDITR LLLYGNMLVQEQENVK LVDTFPGQSIDFFGALR MCALFINDLDAGAGR MDGFYIAPAFMDK MGINPIMMSAGELESGNAGEPAK TDGIPEQDIVK VPIIVTGNDFSTLYAPLIR VPLILGIWGGK VYDDEVR WISGVGVDSVGK YLNEAALGNANEDAIQR	93.4 101 51.7 55.7 53.8 114 62.6 85.9 60.5 57.4 66.6 93.1

beta-form rubisco activase (Glycine max)	gil255635315	49 kDa		[(46)1.600], [(51)1.549]	Calvin cycle			[(51)] GLAYDISDDQQDITR LLLYGNMLVQEQENVK LVDTFPGQSIDFFGALR MCALFINDLDAGAGR MDGFYIAPAFMDK MGINPIMMSAGELESGNAGEPAK VPIIVTGNDFSTLYAPLIR VYDDEVR WISGVGVDSVGK YLNEAALGNANEDAIQR	 96.4 103 91.1 78.9 76.3 97.1 96.2 56.3 66.2 96.2
Elongation factor 1 (Glycine max)	gil255627339	25 kDa	[(22)4.449]		translation	[(22)] KLDEYLLPR LDEYLLPR SVQMEGLLWGASK	 45.6 56.8 72.5		
Elongation factor Tu, chloroplastic (Glycine max)	gil2494261	52 kDa	[(26)4.107]		translation	[(26)] ILDEALAGDNVGLLLR KYDEIDAAPEER NTTVTGVEMFQK VGETVDLVGLR	 59.9 78.6 51 76.1		
phosphoglycerate kinase, putative (Ricinus communis)	gil255544584	50 kDa		[(35)1.771]	Calvin cycle/glycolysis			[(35)] ELDYLVGAVSSPK FAVGTEAIAK GVTTIIGGGDSVAAVEK IGVIESLLEK LDLATSLLAK LSELLGIQVVK RPFAAIVGGSK	 65 64.1 85.1 71.1 62.1 61.3 50.9
phosphoglycerate kinase, putative (Populus triocarpa)	gil224109060	50 kDa		[(35)1.771]	Calvin cycle/glycolysis			[(35)] ADLNVPLDDNQNITDDTR FAVGTEAIAK GVTTIIGGGDSVAAVEK IGVIESLLEK LSELLGIQVVK LVASLPDGGVLLLENVR RPFAAIVGGSK	 75.3 64.1 85.1 71.1 61.3 70.1 50.9
eukaryotic ferritin (Glycine max)	gil255637227	29 kDa		[(19)2.226]	redox homeostasis			[(19)] GDALYAMELALSLEK IAEYVTQLR KIAEYVTQLR NNDPQLADFIESEFLYEQVK	 61.1 57.2 67 67.4
fructose-bisphosphate aldolase, putative (Ricinus communis)	gil255575381 gil255647273	39 kDa		[(46)1.600]	Calvin cycle/glycolysis			[(46)] ANSEATLGTYK AQEALLVR GILAADESTGTIGK	 57.2 58.3 82.4
fructose-bisphosphate aldolase (Glycine max)	gil40457267	38 kDa	[(10)5.718]	[(40)1.655], [(25)2.070]	Calvin cycle/glycolysis	[(10)] GILAADESTGTIGK YVGGSGSGLTSESLYVK	 62.1 84.4	[(40)] GILAADESTGTIGK LASINVENIEANR YVGGSGSGLTSESLYVK [(25)] GILAADESTGTIGK YVGGSGSGLTSESLYVK	 82.6 55.7 107 62.1 84.4
inorganic pyrophosphatase (Glycine max)	gil255640293	32 kDa	[(74)1.885]		ATP-coupled proton transport	[(74)] APFVNDVDDVEK IVAISLDDPK MEVATDESFTPIK	 64.5 62.3 52.7		
Glyceraldehyde 3-phosphate dehydrogenase (Glycine max)	gil255636463	48 kDa	[(25)4.123]	[(46)1.600]	Calvin cycle/glycolysis	[(25)] ILDNETITVDGK YDSMLGTFK	 66.7 59.2	[(46)] GLTAEDVNAAFR ILDNETITVDGK YDSMLGTFK	 60.5 81.8 55.4

RRM (RNA recognition motif), also known as RBD (RNA binding domain) or RNP (ribonucleoprotein domain) protein (Glycine max)	gil255645259	28 kDa	[(93)1.621],[(66)2.105]		unknown	[(93)] GNVSDGYGGGNYSR YGEVVDAR [(66)] GNVSDGYGGGNYSR YGEVVDAR	75.8 55.4 55.8 54.5		
RRM (RNA recognition motif), also known as RBD (RNA binding domain) or RNP (ribonucleoprotein domain) protein (Glycine max)	gil255639723	30 kDa		[(10)3.022]	unknown			[(10)] GFGFVTFGSPDEVK LFVGNLPFSVDSAR QFNGYELDGR SAIQSLDGVDLNGR	61.5 56.8 53.5 62.9
RRM (RNA recognition motif), also known as RBD (RNA binding domain) or RNP (ribonucleoprotein domain) protein (Glycine max)	gil255636284	31 kDa	[(50)2.497],[(33)3.357]		unknown	[(50)] AVEMFSGYELNGR LEQIFSEHGK VYVGNLPWEVDDAR [(33)] AVEMFSGYELNGR VYVGNLPWEVDDAR	63.8 53 61.8 62.3 77.2		
predicted phosphatase/phosphohexomutase (Populus trichocarpa x Populus deltoides)	gil118488927 gil224136009 gil255540407 gil255640791	35 kDa	[(74)1.885],[(87)1.711]	[(36)1.760]	carbohydrate metabolism	[(74)] IFAGDVVPR TELFMALIEK [(87)] IFAGDVVPR TELFMALIEK	65.3 53.6 65.2 55.8	[(36)] IFAGDVVPR TELFMALIEK	64.9 51.3
plant basic secretory protein (Glycine max)	gil255633510	27 kDa	[(68)2.030]		unknown	[(68)] INLFVEDR WDQGYDVTAR YVNDYGGDVK	59.5 56.2 67.1		
Chromoplast-specific carotenoid-associated protein, precursor (Cucumis sativus)	gil62899808	35 kDa	[(92)1.627]		secondary metabolism - carotenoid	[(92)] ALVDSFYGTDR GDGGSVFVLLK	61.3 51.8		
NADPH-specific isocitrate dehydrogenase (Glycine max)	gil169989 gil1708401 gil255635311 gil44921641 gil479386	49 kDa	[(26)4.107],[(28)4.034]		TCA cycle	[(26)] LIDDMVAYALK LLDFTEK SFAEASMATALEK [(28)] LIDDMVAYALK SFAEASMATALEK	61.1 53.2 54.2 53.4 73		
1-deoxy-D-xylulose 5-phosphate reductoisomerase (Pueraria montana var. lobata)	gil35187000	51 kDa	[(26)4.107]		secondary metabolism - isoprenoid	[(26)] AGGTMTGVLSAANEK AVEMFVEEK	82.4 50		

phosphoribulokinase, putative (Glycine max)	gil255646270	45 kDa	[(35)3.347],[(25)4.123]		Calvin cycle	[(35)] ANDFDLMYEQVK DLYEQLIASK FYGEVTQQMLK GVTALDPR KPDFEAYIDPQK LTSVFGGAAEPPK [(25)] ANDFDLMYEQVK FYGEVTQQMLK GVTALDPR KPDFEAYIDPQK LTSVFGGAAEPPK	76.2 49.5 71.8 61.5 55.7 69.1 71.8 73.8 61.8 58 75.2		
beta-1,3-glucanase 3 (Glycine max)	gil2921317	26 kDa	[(22)4.449]		carbohydrate metabolism	[(22)] GSGIELIMDVAK IYYPDEEALQALR	68.9 71.8		
cytosolic glutamine synthetase GSbeta1 (Glycine max)	gil10946357 gil121336	39 kDa	[(35)3.347],[(10)5.718],[(21)4.513]	[(40)1.655],[(25)2.070],[(24)2.072]	amino acid metabolism	[(35)] EDGGYEVIK TLPGPVSDPSELPK [(10)5.718] EDGGYEVIK GDWNGAGAHTNYSTK TLPGPVSDPSELPK LDLTAEELSEEK MELVDAAFPLLK VLVVANPANTNALILK (21)] EDGGYEVIK GDWNGAGAHTNYSTK TLPGPVSDPSELPK	49.3 62 57.7 65.9 59.9 73.8 52.3 63.2 58.2 80.2 59.6	[(40)] EDGGYEVIK GDWNGAGAHTNYSTK TLPGPVSDPSELPK [(25)] EDGGYEVIK GDWNGAGAHTNYSTK TLPGPVSDPSELPK LDLTAEELSEEK MELVDAAFPLLK VLVVANPANTNALILK [(24)] EDGGYEVIK GDWNGAGAHTNYSTK TLPGPVSDPSELPK	60.8 53.1 57.1 57.7 65.9 59.9 73.8 52.3 63.2 58.2 80.2 59.6
peptide methionine sulfoxide reductase (Glycine max)	gil255631520	22 kDa	[(45)2.791]		redox homeostasis	[(45)] IVTEILPAK SGIYYYNDTQAR	59.9 58.7		
Carbamoyl-phosphate synthase L chain, ATP binding domain (Glycine max)	gil255633622	19 kDa	[(103)1.240]		amino acid metabolism	[(103)] GVAVSSVEEAR TDQVEDIAGK	64.6 60.1		
glycolate oxidase (Vitis vinifera)	gil147789493 gil167961875 gil189418957 gil225462096 gil255557255	41 kDa	[(98)1.502]		carbon metabolism	[(98)] AIALTVDTPR GVLTAEDTR VPVFLDGGVR	57.9 54.2 59.7		
Nucleotide binding protein Maf (Glycine max)	gil255642112	22 kDa	[(45)2.791]		nucleotide metabolism	[(45)] IILGSSSVAR LPVDDYLK	58.7 52.7		
isopentenyl diphosphate (IPP) isomerase (Glycine max)	gil255635611 gil255647098 gil262036858 gil6856554	27 kDa	[(89)1.667]		secondary metabolism - isoprenoid	[(89)] ESELIDENALGVR YELLLQQR	55.3 54		

S-adenosylmethionine-dependent methyltransferase (Glycine max)	gil255638590	27 kDa	[(75)1.880]	[(15)2.471]	unknown	[(75)] IDFIESPALPILDK ITAIDINR KTYEVGLPVIK NPVILQSEDLTK TYEVGLPVIK YILETAVYPR	81.7 51 52.2 66.2 49.6 67.8	[(15)] IDFIESPALPILDK ITAIDINR KTYEVGLPVIK NPVILQSEDLTK TYEVGLPVIK YILETAVYPR	81.7 51 52.2 66.2 49.6 67.8
In2-1 protein (Glycine max)	gil11385579 gil255637642	27 kDa		[(20)2.226]	redox homeostasis			[(20)] LATWFEELNK VDPQEIVDLFK YVDANFEGTPLFPSDPAK	59.3 60.3 62.4
SOUL heme-binding protein (Glycine max)	gil255627213 gil255640875	26 kDa		[(10)3.022]	metal trafficking			[(10)] DSSVGEEAAALK LFDYIQGK	59.4 47.9
putative chalcone isomerase 4 (Glycine max)	gil51039630	23 kDa		[(42)1.615]	secondary metabolism - flavonoid			[(42)1.615] GAQYGVQIETAVR YVIENANVVEAIK	60.3 81
chalcone isomerase A [Glycine max]	gil114199183 gil122725493 gil255625955 gil5921724 gil75305825	23 kDa		[(19)2.226]	secondary metabolism - flavonoid			[(19)] FTGIGVYLEDK SVGTYGDAEAAAIEK	64.7 74.1
UBA-domain containing protein, Nascent polypeptide-associated complex NAC (Medicago trunculata)	gil217075454 gil255628877 gil87241023	22 kDa	[(5/6)8.149/7.223]	[(8/11)3.66/3.007]	unknown	[(5/6)] IEDLSSQLQTQAAEQFK NILFVISKPDVFK SPTSDTYIIFGEAK	60.7 47.8 84.9	[(8/11)] IEDLSSQLQTQAAEQFK NILFVISKPDVFK SPTSDTYIIFGEAK	60.7 47.8 84.9
GTP-binding protein (Cicer arietinum)	gil10334503 (+61)	25 kDa		[(59)-1.401]	unknown			[(59)] LVIVGDGGTGK NLQYYEISAK	62.8 48.9
chlorophyll a/b-binding protein (Glycine max)	gil1053216 gil255646685	28 kDa		[(42)1.615]	photosynthesis			[(42)] FGEAVWFK QASSGSPWYGPDR	59.6 49.3
Glycerate dehydrogenase (Cucumis sativus)	gil118564 gil118723307 gil1304042 gil1304044 gil13873334 gil147805559 gil167963236 gil225428049 gil255581842 gil255639618 gil29293057 gil29293059 gil90761112	42 kDa		[(35)1.771]	carbon metabolism			[(35)] EGMATLAALNVLGK GQTVGVIGAGR IVEADEFMR	53.8 57 49.5
putative aldo/keto reductase (Glycine max)	gil224555758 gil255637828	36 kDa		[(40)1.655]	unknown			[(40)] NLDQNIGALAVK YIGLSEASPDTIR	59.9 59.5
caffeoyl-CoA 3-O-methyltransferase (Amorpha fruticosa)	gil146762454 gil255565252	25 kDa		[(13)2.805]	secondary metabolism - flavonoid			[(13)] ENYELGLPVIK ILAMDINR	50.4 59.5

aspartate-semialdehyde dehydrogenase (Glycine max)	gil255689397	38 kDa		[(25)2.070]	amino acid biosynthesis			[(25)] GTVVVDNSSAFR NAPGVVVIDDR	 51.8 63.1
unknown (Glycine max)	gil255644538	31 kDa		[(20)2.226]	unknown			[(20)] DVADSVLADR YSSAAPLSPDAR	 49.6 54.1
phosphatidylinositol/phosphatidylglycerol transfer protein (PG/PI-TP) (Glycine max)	gil255627393	17 kDa		[(65)-1.289]	lipid related			[(65)-1.289] GQPATFSIAATTGK GVEISPDPIAR	 82.5 59.2
Glycosyl hydrolase family 18, acidic chitinase (Glycine max)	gil255642487 gil4835584	32 kDa	[(6/5)7.233/8.149],[(1/6)12.651/7.233][(4/1)8.613/12.651]	[(5/3)4.445/4.600],[(3/8)4.600/3.660],[(8/11)3.660/3.007]	carbohydrate metabolism	[(6/5)] ALNSFSSQR* YGGVMLWNR [(1/6)] ALNSFSSQR* YGGVMLWNR [(4/1)] ALNSFSSQR* YGGVMLWNR	 75 61.5 64.8 51.2 64.9 52.8	[(5/3)] ALNSFSSQR* YGGVMLWNR [(3/8)] ALNSFSSQR* YGGVMLWNR [(8/11)] ALNSFSSQR* YGGVMLWNR	 64.9 52.8 64.8 51.2 75 61.5
ribulose 1,5-bisphosphate carboxylase small subunit precursor (Glycine soja)[i]	gil1079736 gil10946375 gil10946377 gil10946379 gil132086 gil132113 gil255625881 gil255630492 gil3914590	20 kDa		[(36)1.760]	Calvin cycle			[(36)] EVEYLLR IIGFDNVR	 57.1 58.6
Stress-induced protein SAM22 (Glycine max)[i]	gil134194 gil22218276 gil229597555	17 kDa		[(6)3.781],[(65)-1.289],	stress response			[(6)] AIEAYLLAHPDYN ALVTDADNVIPK GDAEPNQDELK GVFTFEDEINSPVAPATLYK ITFLEDGETK SVENVEGNGGPGTIK YETKGDAEPNQDELK [(65)] ITFLEDGETK SVENVEGNGGPGTIK	 76.2 71.2 56.5 85.6 51.3 89.6 59.1 51.3 63.7
ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (Orothamnus zeyheri)[i]	gil1045644 (+1547)	50 kDa	[(74)1.885],[(26)4.107][(10)5.718]	[(25)2.070],	Calvin cycle	[(74)] DDENVNSQPFMR DTDILAAFR EITLGFVDLLR [(26)] DTDILAAFR LTYYTPDYETK [(10)] EITLGFVDLLR LTYYTPDYETK	 49.8 48.3 50.2 52.1 59.8 52.8 55.2	[(25)] EITLGFVDLLR LTYYTPDYETK	 52.8 55.2

rubisco activase (Glycine max)[i]	gil290766485	49 kDa	[(25)4.123],[(103)1.240][(21)4.533]	[(24)2.072],[(51)1.549]	Calvin cycle	[(25)] GLAYDISDDQQDITR YLNEAALGNANEDAINR [(103)] GLAYDISDDQQDITR LLTYGNMLVQEQENVK WISGVGVDGIGK YLNEAALGNANEDAINR [(21)] GLAYDISDDQQDITR LLTYGNMLVQEQENVK YLNEAALGNANEDAINR	80.8 85.1 88.7 66.1 57.6 106 84.4 77.6 117	[(24)] GLAYDISDDQQDITR LLTYGNMLVQEQENVK YLNEAALGNANEDAINR [(51)] GLAYDISDDQQDITR IVDTFPGQSIDFFGALR LLTYGNMLVQEQENVK MCALFINDLDAGAGR MGINPIMMSAGELESGNAGEPAK QYLDNNMDGFYIAPAFMDK TDSIPEQDVVK VPIIVTGNDFSTLYAPLIR VYDDEVR WISGVGVDGIGK YLNEAALGNANEDAINR	84.4 77.6 117 96.4 91.1 102 78.9 97.1 66.3 59.1 96.2 56.3 66.9 124
peroxiredoxin 5-like (Glycine max)[i]	gil255640689	25 kDa		[(65)-1.289]	redox homeostasis			[(65)] VNEEVLLLSDGNGTFTK YALLAEDGVVK	77.3 68.9
carbonic anhydrase (Glycine max)[i]	gil255630357 gil270342124	24 kDa	[(75)1.880]	[(15)2.471],[(55)-1.475]	carbon metabolism	[(75)] EYEQAIEELQK NVANIVPPYDQSK	52.4 56.3	[(15)] EYEQAIEELQK NVANIVPPYDQSK [(55)] EDGGYEVIK GDWNGAGAHTNYSTK TLPGPVSDPSELPK	52.4 56.3 58.2 80.2 59.6
Phosphoglycerate kinase, chloroplastic (Nicotiana tabacum)[i]	gil2499497	50 kDa	[(10)5.718]	[(25)2.070],[(35)1.771]	Calvin cycle/glycolysis	[(10)] ELDYLVGAVSNPK GVTTIIGGGDSVAAVEK IGVIESLLEK YSLAPLVPR	64.1 54.5 64.1 53.6	[(25)] ELDYLVGAVSNPK GVTTIIGGGDSVAAVEK IGVIESLLEK YSLAPLVPR [(35)] ADLNVPLDDNQNITDDTR ELDYLVGAVSNPK FAVGTEAIAK GVSLLLPSDVVIADK GVTTIIGGGDSVAAVEK IGVIESLLEK LSELLGIQVVK RPFAAIVGGSK YSLAPLVPR	64.1 54.5 64.1 53.6 75.3 72.6 64.1 59.8 85.1 71.1 61.3 50.9 59.8
CYPOR-like ferridoxin reductase (Populus triocarpa)[i]	gil224074257	40 kDa	[(10)5.718]	[(25)2.070]	redox homeostasis	[(10)] EGQSIGVIPDGIDK GIDDIMVSLAAK LDFAVSR LYSIASSAIGDFGDSK	50.1 64.7 57.6 79.6	[(25)] EGQSIGVIPDGIDK GIDDIMVSLAAK LDFAVSR LYSIASSAIGDFGDSK	50.1 64.7 57.6 79.6
fructose-bisphosphate aldolase, chloroplastic (Glycine max)[i]	gil1168411 gil217072476 gil22633 gil84468410	42 kDa	[(55)2.375],[(10)5.718]	[(46)1.600],[(25)2.070]	Calvin cycle/glycolysis	[(55)] EAAWGLAR LASIGLENTEANR SNSIAQLGK [(10)] ALFSQITTR QNIGAADDVIVGDIR	54.4 95.7 66.1 52.2 93.4	[(46)] IVDVLIEQNIVPGIK LASIGLENTEANR [(25)] ALFSQITTR QNIGAADDVIVGDIR	59.2 75 52.2 93.4

glyceraldehyde-3-phosphate dehydrogenase A subunit (Glycine max)[f]	gil77540210	43 kDa	[(35)3.347],[(25)4.123],[(10)5.718],[(21)4.513]	[(46)1.600],[(40)1.655],[(25)2.070],[(24)2.072]	Calvin cycle/glycolysis	[(35)] DSPLDVIAINDTGGVK TFAEEVNAAFR VVDLADIVANK [(25)] DSPLDVIAINDTGGVK VVDLADIVANK [(10)] AVALVLPTLK DSPLDVIAINDTGGVK TFAEEVNAAFR VVDLADIVANK [(21)] TFAEEVNAAFR VVDLADIVANK	51.5 49.7 64.2 63.3 68.8 61.7 71.6 56.6 70.3 62 59.2	[(46)] DSPLDVIAINDTGGVK TFAEEVNAAFR VVDLADIVANK [(40)] AVALVLPTLK TFAEEVNAAFR VVDLADIVANK [(25)] AVALVLPTLK DSPLDVIAINDTGGVK TFAEEVNAAFR VVDLADIVANK [(24)] TFAEEVNAAFR VVDLADIVANK	56.3 59.7 55.5 56.5 62.3 72.4 61.7 71.6 56.6 70.3 62 59.2
pentapeptide repeats-containing protein (Medicago truncatula)[i]	gil217071608	26 kDa		[(65)-1.289]	unknown			[(65)] GVDFSNAVLDR NTVLSGSTFDDAK	77.6 70.9
14-3-3-like protein, protein kinase C inhibitor homologue (Oenothera elata subsp. hookeri)[o]	gil1168195 gil15778152 gil255641583 gil26454607 gil2879818 gil3912948	29 kDa	[(50)2.497],[(33)3.577],[(22)4.449]	[(36)1.760]	signal transduction	[(50)] EAAESTLSAYK IISSIEQK [(33)] EAAESTLLAYK IISSIEQK IISSIEQKEESR TVEVEELTVEER [(22)] EAAESTLSAYK IISSIEQK	55.9 61 58.2 54.8 56.5 68.1 65.9 56.2	[(36)] DSTLIMQLLR EAAESTLLAYK IISSIEQKEESR TVEVEELTVEER YEEMVEFMEK	58.2 81.6 63.7 69.7 67.7
glutathione-s-transferase (Glycine max)[g]	gil255627415 gil255640468	23 kDa		[(19)2.226],[(20)2.226],[(10)3.022]	redox homeostasis			[(19)] VPVVLFDGK VTAVDLSLAPK [(20)] VPVVLFDGK VTAVDLSLAPK [(10)] VPVVLFDGK VTAVDLSLAPK	59.4 59 51.1 58.6 50.8 56.2
ascorbate peroxidase 2 (Glycine max)[h]	gil1336082 gil1420938 gil16304410 gil187962070 gil217072458 gil255638684 gil37196687 gil4406539	27 kDa		[(13)2.805],[(20)2.226]	redox homeostasis			[(13)] EGLLQLPSDK TGGPFGTIK [(20)] ALLSDPVFRPLVEK EGLLQLPSDK TGGPFGTIK	56.6 56.7 60.9 52.3 48.5

ribose 5-phosphate isomerase type A (RPI_A) subfamily protein (Glycine max)[h]	gil255640161			[(66)1.271]	Calvin cycle			[(66)] FVVVVDDTK LAADKAVESVK LGALLASGQLSDIVGVPTSK LQELFKEEGVEAK SGMVLGLGTGSTAAFVVAK SLGIPLSVLDDNPR	 56.1 71.1 92.7 71.2 89.4 101
14-3-3 family protein (Vitis vinifera)[h]	gil225451995 gil255638346 gil3023194	30 kDa	[(33)3.357]	[(36)1.760]	signal transduction	[(33)] EAAESTLLAYK IISSIEQK IISSIEQKEESR	 58.2 54.8 56.5	[(36)] DSTLIMQLLR EAAESTLLAYK IISSIEQKEESR TVEVEELTVEER YEEMVEFMEK	 58.2 81.6 63.7 69.7 67.7
dienelactone hydrolase family protein; chlorophyllase-like (Glycine max)[h]	gil255629231 gil255645019	17 kDa		[(10)3.022]	chlorophyll metabolism; photosynthesis			[(10)] DDTTFDAYVVGK ISQLGSGFK	 58.9 64
polyketide cyclase 2 superfamily protein (Glycine max)[i]	gil255628305	17 kDa		[(65)-1.289]	secondary metabolism			[(65)] AIVLDASNVFPK LTLAEGLGYVK LVATPDGGSIVK SVETIEGDGGPGTIK	 53.5 61.8 51.7 84.7
proteasome subunit alpha type-6 (Glycine max)[i]	gil12229897 gil255648341	27 kDa		[(13)2.805]	protein degradation			[(13)] FTYGYEMPVDVLAK LFQVEYAFK YLGLLATGMTADAR	 63.1 56.5 54.3
malate dehydrogenase (Glycine max)[i]	gil5929964	36 kDa	[(10)5.718],[(21)4.533]	[(25)2.070],[(24)2.072]	TCA cycle	[(10)] ALEGADVVIIPAGVPR LFGVTTLDVVR VAVLGAAGGIGQPLSLLMK [(21)] ALEGADVVIIPAGVPR LFGVTTLDVVR VAVLGAAGGIGQPLSLLMK	 72.5 83.1 69.3 67.3 74.4 71.1	[(25)] ALEGADVVIIPAGVPR LFGVTTLDVVR VAVLGAAGGIGQPLSLLMK [(24)] ALEGADVVIIPAGVPR LFGVTTLDVVR VAVLGAAGGIGQPLSLLMK	 72.5 83.1 69.3 67.3 74.4 71.1
cytosolic malate dehydrogenase (Cicer arietinum)[i]	gil10334493 gil217073248 gil27462762 gil27462764 gil77999077 gil83283965	35 kDa	[(10)5.718]	[(25)2.070]	TCA cycle	[(10)] LDLTAEELSEEK MELVDAAFPLLK VLVVANPANTNALILK	 73.8 52.3 63.2	[(25)] LDLTAEELSEEK MELVDAAFPLLK VLVVANPANTNALILK	 73.8 52.3 63.2

triosephosphate isomerase (Glycine max)[k]	gil77540216	27 kDa	[(33)3.577],[(96)1.545],[(104)-1.222][(75)1.880]	[(15)2.471],[(36)1.760],[(31)-1.88]	alvin cycle/glycolys	[(33)] EAGTTTAVVAEQTK IVTTLNEAK [(96)] EAGTTTAVVAEQTK IVTTLNEAK QLLNESNEFVGDK VAYALQQGLK [(104)] AEFVDIINAATVK EAGTTTAVVAEQTK IVTTLNEAK QLLNESNEFVGDK VAYALQQGLK [(75)] AEFVDIINAATVK IVTTLNEAK QLLNESNEFVGDK VAYALQQGLK	62 57.8 65.8 51.9 72.1 56.2 57.4 61.2 65.6 68 56.4 70 65.5 69 49.1	[(15)] AEFVDIINAATVK IVTTLNEAK QLLNESNEFVGDK VAYALQQGLK [(36)] IVTTLNEAK VAYALQQGLK [(31)-1.88] AEFVDIINAATVK QLLNESNEFVGDK VAYALQQGLK VIACIGETLEQR	70 65.5 69 49.1 59.2 56 75.7 74.2 52.4 72.9
Chaperonin 10 Kd subunit (cpn10 or GroES) (Glycine max)[k]	gil255645102	27 kDa		[(28)1.958],[(42)1.615],[(10)3.022]	protein folding			[(28)] DDDIVGILETDDIK DGSDYITLR TSGGLLLTEATK YAGTEVDFDGTK [(42)] DGSDYITLR TSGGLLLTEATK YAGTEVDFDGTK [(10)] DGSDYITLR TSGGLLLTEATK	56.3 56.1 83.7 87 45.8 60.9 59.8 56.4 60.8
polyketide cyclase 2 superfamily protein (Glycine max)[l]	gil18643 gil255640867	17 kDa		[(6)3.781],[(65)-1.289]	secondary metabolism			[(6)] ALVTDADNVIPK GIFTFEDETTSPVAPATLYK ITFVEDGESK SVENLEGNGGPGTIK [(65)] ITFVEDGESK SVENLEGNGGPGTIK	71.2 52.4 51.2 83.3 54.8 76.9

Supporting Information Table 2. List of Peptides Identified from 115 ppb Root Proteins that are Differentially Expressed/Oxidized

Protein	Accession Number(s)	Mol Weight	Fold Change IAF (spot number)	Fold Change SYPRO (spot number)	Classification	IAF Peptides (spot number)	Best Mascot Ion Score	SYPRO Peptides (spot number)	Best Mascot Ion Score
TIM phosphate binding protein (Glycine max)	gil255634120	20 kDa		[(40)-1.573]	Calvin cycle/glycolysis			[(40)] EAGTTTAVVSEQTK IVTTLNEAK QLLNESNEFVGDK	 64.6 64.8 90.2
polyketide cyclase 2 superfamily protein (Glycine max)	gil255631546	17kDa	[(47)1.717]		secondary metabolism	[(47)] ALVTDADNIIPK ITFVEDGETK SVENVEGNGGPGTIK	 71.8 63 61.8		
polyketide cyclase 2 superfamily protein (Glycine max)	gil255627117	18 kDa	[(13)2.461]	[(63)-1.303]	secondary metabolism	[(13)] FIFQAIDDNDHGGTIIK HWTYTIDGK LFSGDIDHNYK	 40.7 59.7 62.9	[(63)] ITTEIGVHATATK LFSGDIDHNYK	 52.2 52.3
Rossmann-fold NAD(P)(+)-binding protein, isoflavone reductase-like (Glycine max)	gil255637531	34 kDa	[(59)1.544]		secondary metabolism - flavonoid	[(59)] AGNPTFALVR VIILGDGNPK	 60.3 63		
similar to heat shock cognate 70 kDa protein 1 (Vitis vinifera)	gil225434984 gil225434986 gil225434994 gil225434996	71 kDa		[(44)-1.514]	protein folding			[(44)] FSDSSVQSDIK NQVAMNPINTVFDAK TTPSYVAFTDTER VEIIANDQGNR	 69.1 83.1 58.7 50
chaperonin GroEL (Bradyrhizobium japonicum USDA 110)	gil27377170 gil27380737 gil27382090	58 kDa		[(64)1.273]	protein folding			[(64)] LAGGVAVIR LENVTLNMLGR SVAAGMNPMDLK VGGATEVEVK	 52.9 67.6 55.2 55.7
glutathione-s-transferase (Glycine max)	gil255629025 gil255646535	24 kDa	[(53)-1.573]		redox homeostasis	[(53)] FSGAAAPAEAAPAK SIEMPGLLWGASK	 58.7 59.7		
F1 ATPase (Pisum sativum)	gil2116558	60 kDa		[(64)1.273]	ATP-coupled proton transport			[(64)] ITDEFTGK TIAMDATEGVVR	 58.4 71.4
F0-F1 ATP synthase subunit beta [Rhodopseudomonas palustris BisA53]	gil115522308 gil27375551 gil283839557 gil39933253 gil75674631 gil85713719 gil86747391 gil91975039 gil92116150	51 kDa		[(64)1.273]	ATP-coupled proton transport			[(64)] FTQAGSEVSALLGR FVDLADTIK VVDLLAPYAK	 68.8 56.3 66
ferric leghemoglobin reductase-2 precursor (Glycine max)	gil3309269	53 kDa	[(70)-1.250]	[(44)-1.514]	nodule metabolism	[(70)] AIDNAEGLVK VVGVDTSGDGVK VVSSTGALALTEIPK	 55.8 74.9 72.9	[(44)] AIDNAEGLVK VVGVDTSGDGVK	 52.1 57.4

cyclophilin (Phaseolus vulgaris)	gil145049729 gil254047060 gil289780455 gil829119	18 kDa	[(13)2.461]		protein folding	[(13)] IVFELYADVTPR VFFDMTIGGQPAGR	67.2 62.5		
RuBisCO large subunit-binding protein subunit alpha (Brassica napus)	gil1351030 gil464727	58 kDa	[(20)-2.187]	[(9)2.132]	Calvin cycle/glycolysis	[(20)] LLVEFENAR VGAATETELEDR	58.9 75.9	[(9)] LLVEFENAR VGAATETELEDR	49.2 69.6
dirigent-like protein (Glycine max)	gil255629177	21 kDa		[(63)-1.303]	secondary metabolism - lignin			[(63)] DNENPTIVQIVDTPK ELPIVGGTGVFR	58.1 79.4
plant basic secretory protein (Glycine max)	gil255627793	25 kDa		[(52)-1.423]	unknown			[(52)] FRDAIGDDYAK SGFVAQLNQQMR TVDQLWQDYK YVNDYSGDLK	40.6 60.7 53.9 62.2
ascorbate peroxidase (Glycine max)	gil110590276 gil110591017 gil161761102 gil161761104 gil29726917 gil310561 gil37196683	28 kDa		[(44)-1.514]	redox homeostasis			[(44)] ALLSDPVFRPLVDK EGLLQLPSDK TGGPFGTIK	62.6 74.2 56.1
NADPH-dependent FMN reductase (Glycine max)	gil255630927	22 kDa		[(32)-1.704]	redox homeostasis			[(32)] GAASVEGVEAK GGSPYGAGTYAGDGSR VYIVYYSTYGHVEK	54.8 100 77
NADPH-dependent oxidoreductase, putative (Glycine max)	gil255644585	38 kDa	[(59)1.544]	[(70)-1.228]	redox homeostasis	[(59)] AYTPGSPLIGYGVSK ESDMNIVESTITLK LPQGSNDVLLK	58.1 64.5 55	[(70)] AYTPGSPLIGYGVSK TLDAVLLNMR	75.2 57.3
calreticulin-1 (Glycine max)	gil117165712	48 kDa	[(70)-1.250]		protein folding	[(70) FYAISAEYPEFSNK KPEGYDDIPK QTGSLYSDWDLLPPK SGTLFDNVLITDDPEYAK TEAGEDTKEEGVHDEL YVGIELWQVK	78.1 66.2 65.6 111 58.7 56.5		
transaldolase-like (Glycine max)	gil255646850	48 kDa	[(53)-1.573],[(27)-1.983]		secondary metabolism	[(53)] AISSSNAYNDQFR LADDTEGTIEAAK [(27)] AISSSNAYNDQFR GVTSNPAIFQK LADDTEGTIEAAK	80.5 83.2 65.6 58.7 89.8		
reversibly glycosylated protein (Phaseolus vulgaris)	gil38194918	40 kDa		[(16)1.942]	unknown			[(16)] VPEGFDYELYNR YVDAILTIPK	55.3 57.8
reversibly glycosylated protein (Glycine max)	gil255645037	42 kDa		[(16)1.942]	unknown			[(16)] LGAVDPYFTK YVDAVLTIPK	64.6 60.3

peptidase M17 (Oryza sativa japonica)	gil115449199 gil238013218 gil242066822 gil293332992 gil75261364	33 kDa		[(64)1.273]	proteolysis			[(64)] FDMGGSAAVFGAAK TIEVNNTDAEGR	57.6 64.6
ADP-Ribosylation Factor 1 Complexed With GDP, Full Length Non-Myristoylated (Homo sapiens)	gil1065361 (+113)	21 kDa		[(62)-1.308]	vesicle transport			[(62)] ILMVGLDAAGK QDLPNAMNAAEITDK	58.9 58.3
pyruvate dehydrogenase (acetyl-transferring)	gil255635250 gil255635914	39 kDa		[(16)1.942]	TCA cycle			[(16)] EGISAEVINLR VLSPYSSEDAR	58.4 51.9
alcohol dehydrogenase 1 (Pisum sativum)°	gil113361 gil113366 gil229464655 gil229464657 gil229464659 gil237512195 gil255638965 gil29373061 gil4039115 gil452767 gil452769 gil51587338 gil61651638	41 kDa		[(16)1.942], [(9)2.132]	carbon metabolism			[(16)] FGVNEFVNPK IIGVDLVSSR [(9)] FGVNEFVNPK IIGVDLVSSR	59.1 72.5 58.4 66.6
molecular chaperone DnaK, provisional (Vitis vinifera)°	gil118488840 gil147860809 gil166919372 gil224120086 gil224129424 gil225456004	74 kDa		[(56)1.397]	protein folding			[(16)] FGVNEFVNPK IIGVDLVSSR [(9)] FGVNEFVNPK IIGVDLVSSR	59.1 72.5 58.4 66.6

proteasome subunit alpha type-5 (Glycine max)[e]	gil12229923 gil217071324 gil217071540 gil224119566 gil225441983 gil255584432 gil255626321 gil255641541 gil255647525 gil255647791 gil297742930	26 kDa	[(53)-1.573]		protein degradation	[(53)] AIGSGSEGADSSLQEQYNK GVNTFSPEGR ITSPLLEPSSVEK	135 62.9 68.1		
copper chaperone homolog CCH (Glycine max)[e]	gil6525011	14 kDa		[(62)-1.308]	metal trafficking	[(53)] AIGSGSEGADSSLQEQYNK GVNTFSPEGR ITSPLLEPSSVEK	135 62.9 68.1		
copper chaperone homolog CCH (Glycine max)[e]	gil6525011	14 kDa		[(62)-1.308]	metal trafficking			[(62)] GNVEPDEVLQAVSK MEGVESFDIDLK SSQTVVLK TAFWVDEAPQSK	61.2 71.2 58 51
eukaryotic UGPase (Glycine max)[e]	gil255635072	51 kDa		[(64)1.273]	nucleotide metabolism			[(64)] ANPENPSIELGPEFK SAVAGLNEISESEK VLQLETAAGAAIR	88.4 80.1 72.3
cystathionine beta-synthase domain-containing protein (Glycine max)[e]	gil255631750	23 kDa	[(2)4.512]		amino acid biosynthesis	[(2)] AIAGIITER GMVGMVSIGDVVR LITVTPDTK SMTQNNVGALVVVK VLQAMQLMTDK	62.9 65 54.6 81.1 58.8		
cyclophilin (Glycine max)[e]	gil17981611 gil255628137	18 kDa	[(13)2.461], [(4)4.207]		protein folding	[(13)] IVMELYADVTPR VFFDMTIGGQSAGR [(4)] IVMELYADVTPR VFFDMTIGGQSAGR	67.2 53.7 74.8 65.5		
YjgF YER057c UK114 family[e]	gil255629976 gil255632612	20 kDa	[(39)1.600]		unknown	[(39)] AGGASYSSVVK TTIMLADLK	66.9 55.5		

ATP synthase subunit alpha, mitochondrial (Helianthus annuus)[e]	gil114404 gil114411 gil15429015 gil224365668 gil22742 gil231585 gil2645893 gil282767712 gil296040811 gil357982 gil5305369 gil543866 gil74181742 gil903732	55 kDa	[(43)-1.730]		ATP-coupled proton transport	[(43)] AAELTTLLESR TAIAIDTILNQK VVSVGDGIAR	61 61.9 50.6		
14-3-3-like protein, protein kinase C inhibitor homologue (Oenothera elata subsp. hookeri)[g]	gil1168195 gil15778152 gil255641583 gil26454607 gil2879818 gil3912948	29 kDa		[(33)1.659],[(52)-1.423]	signal transduction			[(33)] DSTLIMQLLR EAAESTLSAYK IISSIEQKEESR KEAAESTLSAYK LVPSAASGDSK VSASAESEELTVEER YEEMVEFMEK [(52)] EAAESTLSAYK IISSIEQK	58.1 69.1 72.9 56.8 61.3 90.2 56.2 75.5 50.3
glutathione-s-transferase (Glycine max)[g]	gil255627415 gil255640468	23 kDa	[(43)-1.753]	[(40)-1.573]	redox homeostasis	[(43)] VPVVLFDGK VTAVDLSLAPK	57.4 67	[(40)] VPVVLFDGK VTAVDLSLAPK	66.3 72.4
translationally controlled tumor protein (Elaeis guineensis)[i]	gil192910898 gil192912974 gil192912976 gil20140683 gil75222628	19 kDa		[(62)-1.308]	cell expansion			[(62)] LDAEQQELFK LQEQPAFDKK QFVTFMK VVDIVDTFR	47.6 57 48.5 49.3
dirigent-like protein (Glycine max)	gil255626867	21 kDa		[(63)-1.303]	secondary metabolism - lignin			[(63)] AQGLFGLASLEDR EMPIVGGTGVFR IIEPSASEVR	82.3 60 75.8
Universal stress protein family protein (Glycine max)	gil255626103 gil255628729	18 kDa		[(63)-1.303]	stress response			[(63)] EKIVEAVGDLK IGVALDFSK IVEAVGDLK KIGVALDFSK	50.7 53.2 64.7 51.4
polyketide cyclase 2 superfamily protein (Glycine max)[i]	gil255628305	17 kDa	[(47)1.717]	[(39)1.600]	secondary metabolism	[(47)] AIVLDASNVFPK GDEQLAEEYVK LTLAEGLGYVK SVETIEGDGGPGTIK	66.4 56.2 57.8 70.5	[(39)] AIVLDASNVFPK GDEQLAEEYVK LTLAEGLGYVK SVETIEGDGGPGTIK	70.9 50.4 59.3 53.3
proteasome subunit alpha type-6 (Glycine max)[i]	gil12229897 gil255648341	27 kDa		[(61)-1.326]	protein degradation			[(61)] AAGITSIGVR ATEIEVGVVR YLGLLATGMTADAR	68.1 51.4 84.7

malate dehydrogenase (Glycine max)[i]	gi|5929964	36 kDa	[(59)1.544]	[(70)-1.228]	TCA cycle	[(59)] ALEGADVVIIPAGVPR SEVVGYQGDEELGK	66.1 78.8	[(70)] ALEGADVVIIPAGVPR DDLFNINAGIVK LFGVTTLDVVR SEVVGYQGDEELGK TQDGGTEVVEAK	57 60.6 77.1 87 54.9
cytosolic malate dehydrogenase (Cicer arietinum)[i]	gi|10334493 gi|217073248 gi|27462762 gi|27462764 gi|77999077 gi|83283965	35 kDa	[(59)1.544]		TCA cycle	[(59)] ALGQISER EFAPSIPEK LDLTAEELSEEK MELVDAAFPLLK VLVVANPANTNALILK	62.8 56.1 78.3 55.3 55.4		
polyketide cyclase 2 superfamily protein (Glycine max)[i]	gi|18643 gi|255640867	17 kDa		[(39)1.600]	secondary metabolism			[(39)] ALVTDADNVIPK AVEAYLLANPHYN ITFVEDGESK KITFVEDGESK SVENLEGNGGPGTIK	49 62.5 59 61.5 71.7

Supporting Information Table 3. List of Peptides Identified from 60 ppb Leaf Proteins that are Differentially Expressed/Oxidized

Protein	Accession Number(s)	Mol Weight	Fold Change IAF (spot number)	Fold Change SYPRO (spot number)	Classification	IAF Peptides (spot number)	Best Mascot Ion Score	SYPRO Peptides (spot number)	Best Mascot Ion Score
polyketide cyclase 2 superfamily protein (Glycine max)	gil255630540	17 kDa	[(79)2.467]		secondary metabolism	[(79)] LTFVEDGQTK LVADPNGGSIAK	 50.7 69.7		
Malate dehydrogenase [NADP], chloroplastic (Pisum sativum)	gil462579	48 kDa	[(128)-1.547]		TCA cycle	[(128)] LASGEVFGPDQPIALK TEAELLAEK	 69.6 55.8		
ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (Ostryopsis davidiana)	gil5731964	52 kDa		[(53)1.520]	Calvin cycle			[(53)] DTDILAAFR XKDTDILAAFR	 46.1 54.4
peroxiredoxin (PRX) family protein, 2-Cys (Glycine max)	gil255641409	22 kDa		[(50)-1.620]	redox homeostasis			[(50)] SGGLGDLNYPLISDVTK SYGVLIPDQGIALR	 92 54.8
translation elongation factor-TU (Glycine max)	gil2546952	27 kDa		[(77)-1.281]	translation			[(77)] ILDEALAGDNVGLLLR LMDEVDDYIPIPQR NTTVTGVEMFQK VGETVDLVGLR	 79.4 63.4 49.1 83.6
plastid high chlorophyll fluorescence 136 precursor (Zea mays)	gil148251625 gil218198925 gil222636266 gil224096552 gil225423755 gil242097170 gil75252730	43 kDa		[(64)-1.407]	photosynthesis			[(64)] AADNIAANLYSVK SIPSAEDEDFNYR	 71.8 62.8
thiamine pyrophosphate (TPP) family protein (Populus triocarpa)	gil224063766	81 kDa	[(127)-1.575],[(135)-1.468]		cofactor metabolism	[(127)] ALPTYTPESPADATR FLAIDAVEK KYSEEAAELK NGNTGYDEIR NLSQQNLNALVK YSEEAAELK [(135)] KYSEEAAELK NGNTGYDEIR NLSQQNLNALVK VLPGLLGGSADLASSNMTLLK YSEEAAELK	 46.1 52.3 81.4 57 48.1 49.5 69.9 66.7 75.8 74.9 65.9		

23 kDa OEC protein (Salicornia veneta)	gil148535011 gil197691941 gil255635846	22 kDa	[(75)-2.599]		photosynthesis	[(75)] EVEYPGQVLR QYYSITVLTR	51.9 48.4		
glutamine synthetase precursor (Glycine max)	gil13877511 gil255648131	48 kDa	[(128)-1.574]		amino acid metabolism	[(128)] AAEIFSNPK EDGGFEVIKK LEGLLNLDITPFTDK	72.1 56 58.6		
ribulose 1,5-bisphosphate carboxylase small subunit precursor (Glycine soja)'	gil1079736 gil10946375 gil10946377 gil10946379 gil132086 gil132113 gil255625881 gil255630492 gil3914590	20 kDa		[(37)1.845]	Calvin cycle			[(37)] FETLSYLPDLDDAQLAK IIGFDNVR SPGYYDGR	87 67.1 53.1
Stress-induced protein SAM22 (Glycine max)'	gil134194 gil22218276 gil229597555	17 kDa		[(30)-1.984]	stress response			[(30)] AIEAYLLAHPDYN ALVTDADNVIPK GDAEPNQDELK IESIDEANLGYSYSVVG GAALPDTAEK ITFLEDGETK SVENVEGNGGPGTIK YETKGDAEPNQDELK	62.3 68.5 49.1 49 51.9 60.3 77.3
ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (Orothamnus zeyheri)'	gil1045644 (+1547)	50 kDa		[(61)1.421],[(53)1.520]	Calvin cycle			[(61)] DTDILAAFR XKDTDILAAFR [(53)] DTDILAAFR XKDTDILAAFR	56.9 60.1 46.1 54.4
rubisco activase (Glycine max)'	gil290766485	49 kDa	[(128)-1.547]		Calvin cycle	[(128)] GLAYDISDDQQDITR LLTYGNMLVQEQENVK MGINPIMMSAGELESGNAGEP AK VPLILGIWGGK YLNEAALGNANEDAINR	98.2 85.2 89.6 56.4 107		
peroxiredoxin 5-like (Glycine max)'	gil255640689	25 kDa	[(79)2.467]		redox homeostasis	[(79)] VNEEVLLLSDGNGTFTK YALLAEDGVVK	56.2 67.1		
carbonic anhydrase (Glycine max)'	gil255630357 gil270342124	24 kDa		[(8)-3.313]	carbon metabolism			[(8)] EYEQAIEELQK NVANIVPPYDQSK	49.7 47.9

Phosphoglycerate kinase, chloroplastic (Nicotiana tabacum)[i]	gil2499497	50 kDa	[(128)-1.574],[(133)-1.513]	[(77)-1.281]	Calvin cycle/glycolysis	[(128)] ADLNVPLDDNQNITDDTR ELDYLVGAVSNPK GVTTIIGGGDSVAAVEK IGVIESLLEK LSELLGIQVVK [(133)] ADLNVPLDDNQNITDDTR ELDYLVGAVSNPK GVTTIIGGGDSVAAVEK IGVIESLLEK RPFAAIVGGSK YSLAPLVPR	68.5 72.4 73.2 71.3 54.2 57.5 52.9 54.5 68.5 49.1 59.6	[(77)] ELDYLVGAVSNPK GVTTIIGGGDSVAAVEK YSLAPLVPR	56.3 62.7 58.5
CYPOR-like ferridoxin reductase (Populus trioarpa)[i]	gil224074257	40 kDa		[(64)-1.407]	redox homeostasis			[(64)] GIDDIMVSLAAK LYSIASSAIGDFGDSK	78 87.4
fructose-bisphosphate aldolase, chloroplastic (Glycine max)[i]	gil1168411 gil217072476 gil22633 gil84468410	42 kDa	[(109)-1.869]		Calvin cycle/glycolysis	[(109)] EAAWGLAR LASIGLENTEANR SNSLAQLGK	57 91.3 68.4		
glyceraldehyde-3-phosphate dehydrogenase A subunit (Glycine max)[i]	gil77540210	43 kDa	[(112)-1.811],[(92)-2.207]		Calvin cycle/glycolysis	[(112)] AVALVLPTLK VVDLADIVANK [(92)] TFAEEVNAAFR VVDLADIVANK	53.5 77.2 54.4 76.5		
pentapeptide repeats-containing protein (Medicago truncula)[i]	gil217071608	26 kDa	[(79)2.467]		unknown	[(79)] FDGADMTEVVMSK GVDFSNAVLDR	60.7 68.3		
translationally controlled tumor protein (Elaeis guineensis)[i]	gil192910898 gil192912974 gil192912976 gil20140683 gil75222628	19 kDa		[(52)-1.581]	cell expansion			[(52)] LDAEQQELFK VVDIVDTFR	66.1 69.6
triosephosphate isomerase (Glycine max)[k]	gil77540216	27 kDa		[(38)1.834],[(8)-3.313],[(82)1.210]	Calvin cycle/glycolysis			[(38)] EAGTTTAVVAEQTK IVTTLNEAK QLLNESNEFVGDK VAYALQQGLK [(8)] IVTTLNEAK QLLNESNEFVGDK VAYALQQGLK [(82)] IVTTLNEAK QLLNESNEFVGDK VAYALQQGLK	48.8 49.7 61.7 53.4 52.3 53.6 61.9 63.7 78.3 52
Chaperonin 10 Kd subunit (cpn10 or GroES) (Glycine max)[k]	gil255645102	27 kDa	[(75)-2.599]		protein folding	[(75)] DGSDYITLR TSGGLLLTEATK YAGTEVDFDGTK	74.7 70.3 67.7		

polyketide cyclase 2 superfamily protein (Glycine max)[i]	gil18643 gil255640867	17 kDa	[(67)2.888]		secondary metabolism	[(67)] ALVTDADNVIPK AVEAYLLANPHYN ITFVEDGESK	66.4 46.3 51.7		

Supporting Information Table 4. List of Peptides Identified from 60 ppb Root Proteins that are Differentially Expressed/Oxidized

Protein	Accession Number(s)	Mol Weight	Fold Change IAF (spot number)	Fold Change SYPRO (spot number)	Pathway	IAF Peptides (spot number)	Best Mascot Ion Score	SYPRO Peptides (spot number)	Best Mascot Ion Score
triosephosphate isomerase (Glycine max)	gil255645535	33 kDa	[(44)1.750]	[(33)-1.776]	Calvin cycle/glycolysis	[(44)] GGAFTGEISAEQLK GPEFATIVNSVTSK IEISAQNSWVGK IIYGGSVNGGNSAELAK LVADLNSAK	 75.2 82.3 57.3 56.6 68.5	[(33)] AEFVDIINAATVK EAGTTTAVVAEQTK QLLNESNEFVGDK VAYALQQGLK	 74.4 60.8 82.3 54.7
methionine synthase (Glycine max)	gil33325957	84 kDa	[(13)-3.129],[(67)1.276]	[(58)-1.223],[(48)-1.539]	amino acid biosynthesis	[(13)] IPPTEEIADR YLFAGVVDGR [(67)] IPPTEEIADR IVEVNALAK YLFAGVVDGR	 66.4 65.3 58.8 54.2 52.8	[(58)] IPPTEEIADR YLFAGVVDGR [(48)] IVEVNALAK YLFAGVVDGR	 50.3 71.7 57.5 84.3
putative fructokinase 2 (Vitis vinifera)	gil225433918	35 kDa		[(58)-1.223]	carbohydrate metabolism			[(58)] LPLWPSAEEAR TALAFVTLR	 70.6 75.9
nucleoside diphosphate kinase 1 (Glycine max)	gil2498078 gil26245395	16 kDa	[(3)4.559]	[(36)-1.719],[(34)1.748]	nucleotide metabolism	[(3)] GLIGEIISR IIGATNPAQSEPGTIR	 64.2 60.9	[(36)] GLIGEIISR IIGATNPAQSEPGTIR KIIGATNPAQSEPGTIR LVTVDRPFAEK [(34)] GDFAIDIGR GLIGEIISR IIGATNPAQSEPGTIR KIIGATNPAQSEPGTIR LVTVDRPFAEK	 68.6 87.2 57.6 57.6 57.3 68.3 80.5 91.9 66.7
Rossmann-fold NAD(P)(+)-binding protein, possible isoflavone reductase homolog 2 (Glycine max)	gil255637547 gil255640090 gil6573171	34 kDa	[(58)1.471]		secondary metabolism - flavonoid	[(58)1.471] FIVEASAK ILFIGGTGYIGK	 52.7 63		
dimethylmenaquinone methyltransferase (Glycine max)	gil255630950	18 kDa	[(27)-2.309]	[(16)-2.191]	secondary metabolism	[(27)-2.309] QVFSGPIVTLK QVPINIAGTR VFEDNVLVR VLVVDGGASLR	 48 56.8 62.2 69	[(16)] QVFSGPIVTLK VFEDNVLVR VLVVDGGASLR	 48 65.3 53.7

transaldolase-like protein (Glycine max)	gil255646850	48 kDa	[(46)-1.736],[(31)-2.181],[(33)-2.034]		secondary metabolism	[(46)] AISSSNAYNDQFR LADDTEGTIEAAK VTSVASFFVSR [(31)] AISSSNAYNDQFR GVTSNPAIFQK LADDTEGTIEAAK VTSVASFFVSR [(33)] AISSSNAYNDQFR GVTSNPAIFQK IGTPVALNLR LADDTEGTIEAAK	85.7 107 60.8 83.1 54.2 66.6 73.3 90.3 54.4 49 92		
cell-autonomous heat shock cognate protein 70 (Cucurbita maxima)	gil26985223	71 kDa	[(56)1.493],[(47)-1.696],[(62)1.385],[(54)-1.558]	[(60)-1.204],[(57)-1.283]	protein folding	[(56)] NQVAMNPINTVFDAK TTPSYVGFTDTER [(47)] FSDSSVQSDIK NQVAMNPINTVFDAK TTPSYVGFTDTER [(62)] FSDSSVQSDIK TTPSYVGFTDTER [(54)] FSDSSVQSDIK NQVAMNPINTVFDAK TTPSYVGFTDTER	54.6 55.5 51.4 57.5 50.3 63.5 51.4 64.3 83 62.8	[(60)] FSDSSVQSDIK NQVAMNPINTVFDAK TTPSYVGFTDTER VEIIANDQGNR [(57)] FSDSSVQSDIK NQVAMNPINTVFDAK TTPSYVGFTDTER	83.1 80.7 60.3 67.2 71.6 57.2 54.6
14-3-3-like protein B; SGF14B (Glycine max)	gil3023195	28 kDa	[(46)-1.736]		signal transduction	[(46)] IVSSIEQKEEGR VVVGSTPASELTVEER	63.1 90.2		
DHAR class glutathione transferase DHAR2 (Glycine max)	gil255627415 gil255640468	23 kDa		[(50)-1.514]	redox homeostasis			[(50)] VPVVLFDGK VTAVDLSLAPK	51.5 62.9
glutathione-s-transferase (Glycine max)	gil255625731	25 kDa		[(60)-1.204],[(50)-1.514]	redox homeostasis			[(60)] VLDVYEER VYGPTYGSPK [(50)] VLDVYEER VYGPTYGSPK	53.8 52.2 57.7 65.2
proteasome beta type-6 subunit (Vitis vinifera)	gil225429850 gil255574159 gil255625747 gil296081796	25 kDa	[(47)-1.696]		protein degradation	[(47)] SGSAADSQIVSDYVR TVIINSEGVTR	76.5 60		
proteasome alpha type 4 subunit (Glycine max)	gil255637272	28 kDa		[(51)-1.473],[(33)-1.776]	protein degradation			[(51)] AAAIGANNQAAQSILK LELAEVFLSPSGK [(33)] AAAIGANNQAAQSILK LELAEVFLSPSGK	105 57.2 58.2 67.3

S-adenosyl-L-homocystein hydrolase (Petunia x hybrida)	gil32967697 gil464734	53 kDa	[(57)-1.485],[(34)-2.014]		amino acid biosynthesis	[(57)] LVGVSEETTTGVK VAVVAGYGDVGK [(34)] LVGVSEETTTGVK VAVVAGYGDVGK	70.1 49.2 62.8 62.7		
protein disulfide isomerase (Glycine max)	gil171854980 gil49257109	59 kDa	[(57)-1.485]		redox homeostasis	[(57)] SADEATAFIGENK VAIVGVFPK	67 65.2		
peroxidase 1 precursor (Glycine max)	gil255645056 gil5002342	37 kDa	[(57)-1.485]		redox homeostasis	[(57)] GLDVVNQIK MGNIGVLTGSQGEIR	50.8 86.8		
Aspartate aminotransferase P2, mitochondrial (Lupinus angustifolius)	gil112979 gil169915 gil255648095 gil25990362 gil2605932 gil3378163 gil7548843	50 kDa		[(23)-1.998]	amino acid biosynthesis			[(23)-1.998] ATAELLLGADNPAIK LNLGVGAYR VATVQGLSGTGSLR	68.3 52.5 70.2
chain A, mutant beta-amylase (W55r) (Glycine max)	gil149241163 gil157830459 gil157830496 gil169913 gil231541 gil46015332 gil46015333 gil46015334 gil46015336 gil46015825 gil46015829 gil50513924 gil62122629 gil62122631 gil62122633 gil62122635 gil62738229 gil62738230 gil62738231 gil71673371 gil902938	56 kDa	[(49)-1.676],[(34)-2.014]		carbohydrate metabolism	[(49)] LSDDLLQK LSMFGVTYLR [(34)] LSDDLLQK LSMFGVTYLR VAGENALPR	57.5 76.5 61 54.1 61.2		
vacuolar ATPase subunit B1 (Triticum aestivum)	gil125744990 gil2493131 gil2493132 gil255560497 gil6715512	54 kDa		[(23)-1.998]	ATP-coupled proton transport			[(23)-1.998] AVVQVFEGTSGIDNK TPVSLDMLGR TVSGVAGPLVILDK	74.9 56.3 54.3
leghemoglobin C1; nodulin-50 (Glycine max)	gil126236	15 kDa	[(62)1.385]		nodule metabolism	[(62)] AVTDPQFVVVK QEALVSSSFEAFK	53.6 85.3		
NADPH:quinone reductase and related Zn-dependent oxidoreductase	gil255638934	35 kDa		[(51)-1.473]	redox homeostasis			[(51)-1.473] VGDEVYGDINVK VIGSLAEYTAAEER	61.2 67.3
TRX family protein, possibly m-type (Glycine max)	gil255631161	20 kDa	[(56)1.493]		redox homeostasis	[(56)1.493] LNTDESPSTATR TTLTSSIEK	56.9 55		

pyridine nucleotide-disulphide oxidoreductase (Glycine max)	gil255636578	44 kDa		[(56)-1.338]	redox homeostasis			[(56)-1.338] GTVAVGFTSNSDGEVK LTDFGVEGADAK	63.5 69.8
actin (Oryza sativa japonica)	gil115484337 gil168472715 gil217072994 gil225448323 gil296085677 gil9965319	42 kDa		[(56)-1.338]	cell structure			[(56)-1.338] AGFAGDDAPR GEYDESGPAIVHR	63.3 60.9
secretory peroxidase (Glycine max)	gil255638280	38 kDa	[(67)1.276],[(61)1.387]		redox homeostasis	[(67)] EIVNQFASDQK QLGGPDFDVPLGR VSQLDVITDR [(61)1.387] QGVFTSDQDIAGSPK VSQLDVITDR	57.2 51.7 63.7 66.2 62.3		
Glo EDI BRP like 8 family (Glycine max)	gil255628635	17 kDa	[(55)1.531]		unknown	[(55)] AAFGAEEVGR ANDAILFFK	66.8 59.4		
unknown (Glycine max)	gil255647164	12 kDa	[(59)1.424]		unknown	[(59)] LVAEAAQSALK VADAAGDLLDAAGK	59.3 85.9		
Phosphoglucomutase, cytoplasmic (Pisum sativum)	gil12585296 gil12585316 gil12585330 gil12643355 gil15220668 gil15223226 gil224099523 gil224111476 gil238479031 gil238479033 gil255573724 gil26185954 gil297841819 gil297845376 gil40233152 gil62321043 gil9295686	63 kDa	[(62)1.385]		carbohydrate metabolism	[(62)] LSGTGSEGATIR YLFEDGSR	68.1 51.6		
predicted translation initiation factor 2B subunit	gil255647295	23 kDa	[(67)1.276]		translation	[(67)] AIGSYGASIIQQQTEK GSLQLLDQR	72.9 51		
glyoxalase I	gil255637721	32 kDa	[(39)-1.905]		carbon metabolism	[(39)] EPGPVQGGTTVIAFVK SAEVVNQVIK TTSFLDPDGWK	65.2 53.8 57.2		

alcohol dehydrogenase 1 (Pisum sativum)°	gil113361 gil113366 gil229464655 gil229464657 gil229464659 gil237512195 gil255638965 gil29373061 gil4039115 gil452767 gil452769 gil51587338 gil61651638	41 kDa	[(67)1.276],[(66)1.298],[(58)1.471],[(39)-1.905]	[(48)-1.539]	carbon metabolism	[(67)] FGVNEFVNPK IIGVDLVSSR [(66)] FGVNEFVNPK IIGVDLVSSR [(58)] FGVNEFVNPK IIGVDLVSSR [(39)] FGVNEFVNPK IIGVDLVSSR	57.8 54.2 52.3 58.8 69.9 66.7 57.6 58.2	[(48)-1.539] FGVNEFVNPK IIGVDLVSSR	63.3 54.8
molecular chaperone DnaK, provisional (Vitis vinifera)°	gil118488840 gil147860809 gil166919372 gil224120086 gil224129424 gil225456004	74 kDa		[(52)1.460]	protein folding			[(52)] GPDGDVIDADFTDSK IAGLEVLR NQADSVVYQTEK QFAAEEISAQVLR	101 74.8 72.2 83.7
proteasome subunit alpha type-5 (Glycine max)°	gil12229923 gil217071324 gil217071540 gil224119566 gil225441983 gil255584432 gil255626321 gil255641541 gil255647525 gil255647791 gil297742930	26 kDa	[(31)-2.181]		protein degradation	[(31)] GVNTFSPEGR ITSPLLEPSSVEK LGSTAIGLK	63.6 68.8 53		
copper chaperone homolog CCH (Glycine max)°	gil6525011	14 kDa	[(55)1.531]		metal trafficking	[(55)] AAFGAEEVGR ANDAILFFK	66.8 59.4		
eukaryotic UGPase (Glycine max)°	gil255635072	51 kDa		[(41)-1.629]	nucleotide metabolism			[(41)] ISIVANPGVK LDALLSQGK SAVAGLNEISESEK	56 53.2 78.5
cystathionine beta-synthase domain-containing protein (Glycine max)°	gil255631750	23 kDa		[(14)2.246]	amino acid biosynthesis			[(14)] AIAGIITER GMVGMVSIGDVVR SMTQNNVGALVVVK	60.2 58.2 96.5

cyclophilin (Glycine max)[e]	gil17981611 gil255628137	18 kDa		[(14)2.246]	protein folding			[(14)] IVMELYADVTPR VFFDMTIGGQSAGR	 56.6 80.9
YjgF YER057c UK114 family[e]	gil255629976 gil255632612	20 kDa	[(62)1.385]		unknown	[(62)] AGGASYSSVVK STYQVAALPLDAK	 73.5 51		
ATP synthase subunit alpha, mitochondrial (Helianthus annuus)[e]	gil114404 gil114411 gil15429015 gil224365668 gil22742 gil231585 gil2645893 gil282767712 gil296040811 gil357982 gil5305369 gil543866 gil74181742 gil903732	55 kDa	[(34)-2.014]		ATP-coupled proton transport	[(34)] AAELTTLLESR AVDSLVPIGR VVDALGVPIDGR VVSVGDGIAR	 88.4 58.1 67.1 63.8		
ascorbate peroxidase 2 (Glycine max)[h]	gil1336082 gil1420938 gil16304410 gil187962070 gil217072458 gil255638684 gil37196687 gil4406539	27 kDa	[(47)-1.696],[(54)-1.558]	[(60)-1.204]	redox homeostasis	[(47)] ALLSDPVFRPLVEK EGLLQLPSDK [(54)] ALLSDPVFRPLVEK EGLLQLPSDK	 58.5 63.8 64.2 52.2	[(60)] ALLSDPVFRPLVEK EGLLQLPSDK TGGPFGTIK	 52.6 64.1 53.7
ribose 5-phosphate isomerase type A (RPI_A) subfamily protein (Glycine max)[h]	gil255640161			[(57)-1.283],[(23)-1.998]	Calvin cycle			[(57)] FVVVVDDTK LQELFKEEGVEAK SGMVLGLGTGSTAAFVVAK SLGIPLSVLDDNPR [(23)] LGALLASGQLSDIVGVPTSK SGMVLGLGTGSTAAFVVAK SLGIPLSVLDDNPR	 49.6 51.3 58.4 57 64.4 50 77.7
14-3-3 family protein (Vitis vinifera)[h]	gil225451995 gil255638346 gil3023194	30 kDa	[(46)-1.736]		signal transduction	[(46)] EAAESTLLAYK IISSIEQK IISSIEQKEESR TVEVEELTVEER	 76.1 56.3 66.7 66.2		
dienelactone hydrolase family protein; chlorophyllase-like (Glycine max)[h]	gil255629231 gil255645019	17 kDa	[(44)1.750]		photosynthesis related	[(44)] DDTTFDAYVVGK ISQLGSGFK	 67.8 51.5		

polyketide cyclase 2 superfamily protein (Glycine max)[i]	gil255628305	17 kDa	[(27)-2.309],[(32)-2.215],[(56)1.493]		secondary metabolism	[(27)] AIVLDASNVFPK GDEQLAEEYVK SVETIEGDGGPGTIK [(32)] AIEDFIQANPDYN AIVLDASNVFPK GDEQLAEEYVK SVETIEGDGGPGTIK SVETIEGDGGPGTIKK [(56)] GDEQLAEEYVK SVETIEGDGGPGTIK	54.8 55.7 71.6 52 54.6 67.5 84.7 62.4 64.6 69.3		
proteasome subunit alpha type-6 (Glycine max)[i]	gil12229897 gil255648341	27 kDa		[(44)-1.568]	protein degradation			[(44)] AAGITSIGVR ATEIEVGVVR	 74.9 66.7
malate dehydrogenase (Glycine max)[i]	gil5929964	36 kDa	[(67)1.276],,[(58)1.471],[(61)1.387][(66)1.298]	[(48)-1.539]	TCA cycle	[(67)] ALEGADVVIIPAGVPR DDLFNINAGIVK LFGVTTLDVVR [(58)] ALEGADVVIIPAGVPR DDLFNINAGIVK LFGVTTLDVVR SEVVGYQGDEELGK VAVLGAAGGIGQPLSLLMK [(61)] ALEGADVVIIPAGVPR ANLDDDVIK LFGVTTLDVVR SEVVGYQGDEELGK [(66)1.298] DDLFNINAGIVK SEVVGYQGDEELGK	83.2 54.2 78.2 90.8 64.6 88 75.1 57 81.3 51.5 72.4 80.6 70.1 86.9	[(48)] ALEGADVVIIPAGVPR ANLDDDVIK LFGVTTLDVVR	 64 51.1 65.9
cytosolic malate dehydrogenase (Cicer arietinum)[i]	gil10334493 gil217073248 gil27462762 gil27462764 gil77999077 gil83283965	35 kDa	[(58)1.471]		TCA cycle	[(58)1.471] KLDLTAEELSEEK LDLTAEELSEEK MELVDAAFPLLK VLVVANPANTNALILK	62.9 79.1 60 63.8		

triosephosphate isomerase (Glycine max)[k]	gil77540216	27 kDa	[(66)1.298],[(54)-1.558],[(44)1.750],[(60)-1.419]	[(57)-1.283],[(51)-1.473],[(33)-1.776],[(47)-1.539]	alvin cycle/glycolys	[(66)] AEFVDIINAATVK EAGTTTAVVAEQTK IVTTLNEAK QLLNESNEFVGDK [(54)] EAGTTTAVVAEQTK IVTTLNEAK QLLNESNEFVGDK [(44)] EAGTTTAVVAEQTK IVTTLNEAK QLLNESNEFVGDK [(60)] EAGTTTAVVAEQTK QLLNESNEFVGDK	55.7 54.4 49.3 83.3 70.8 52.1 84.1 72.9 52.8 71.5 81.2 64.9	[(57)] EAGTTTAVVAEQTK VAYALQQGLK [(51)] AEFVDIINAATVK EAGTTTAVVAEQTK QLLNESNEFVGDK VAYALQQGLK [(33)] AEFVDIINAATVK EAGTTTAVVAEQTK QLLNESNEFVGDK VAYALQQGLK [(47)] EAGTTTAVVAEQTK IVTTLNEAK QLLNESNEFVGDK VAYALQQGLK	52.9 51.9 71.5 70.8 86.3 57.5 74.4 60.8 82.3 54.7 62.4 60.5 62.8 55.6
Chaperonin 10 Kd subunit (cpn10 or GroES) (Glycine max)[k]	gil255645102	27 kDa		[(60)-1.204],[(50)-1.514]	protein folding			[(60)-1.204] DGSDYITLR TSGGLLLTEATK [(50)-1.514] DGSDYITLR GKDGSDYITLR KPLSVTPGNTVLYSK TSGGLLLTEATK YAGTEVDFDGTK YTAIKPLGDR	61.2 55.9 58 56.3 77.6 80 80.6 56.3

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

The goal of my thesis work was to obtain new insights into plant redox biology, with an emphasis on novel forms of redox-response due to oxidative stress. Specifically, I first sought to shed new light on the use of the glutathione homolog homoglutathone in legumes through structural and kinetic analyses of the substrate specificity responsible for the different modes of synthesis. Secondly, I have examined the regulation of the soybean thiol-redox proteome in response to changes in real-world oxidative conditions, namely field-exposure to increasing concentrations of tropospheric ozone.

In the preceding chapters, I described the methodologies used and results obtained in my efforts to meet my experimental goals. Based on my research results, I determined that synthesis of homoglutathione (hGSH), instead of glutathione (GSH), by the enzyme homoglutathione synthetase is largely specified by the replacement of two alanine residues in the alanine-rich loop with a leucine and proline (Chapter 2). In Chapter 3, preliminary efforts to test the methodology necessary for detecting differences in protein expression and thiol redox-state led to the conclusion that it was sound for application in future larger-scale experiments. Additionally, by comparing the changes in protein expression and thiol redox-state between *B. juncea* roots treated with BSO and H_2O_2, I was able to conclude that differing sources of ROS (exogenous versus endogenous) led to different protein redox responses. The tested methodology was then applied to tissue from soybean plants grown under various concentrations of tropospheric ozone in natural field setting (Chapter 4). From the results obtained, I was able to identify widespread and large-fold changes in the expression of key proteins which, contrary to what has been shown in acute exposure experiments, are largely involved in carbon fixation and flux

(glycolysis, Calvin cycle, and the TCA cycle). As indicated by their formatting, Chapter 2 and Chapter 3 were previously published in the journals Plant Cell and Proteomics, respectively. Final adjustments to Chapter 4 are currently underway, and it will be submitted to PNAS shortly.

Homoglutathione Synthetase and Molecular Diversity of Plant Glutathione Biosynthesis

As described in the preface to Chapter 2, it has been known for more than 20 years that some plants - specifically legumes and grasses - produce thiol-containing tripeptides besides GSH. Of these tripeptides, hGSH (from legumes) has been the most thoroughly investigated. Metabolite studies on the localization of hGSH indicate that its distribution varies across different tissue types in a species dependent manner [Moran et al, 2000; Matamoros et al., 1999]. In nodules, the most common site of localization, hGSH content further varies as a function of time till senescence and in response to changes in stressor concentrations [Loscos et al., 2008].

Like GSH, hGSH is synthesized in two ATP-dependent steps. While the first step's enzyme, γ-glutamylcysteine synthetase, is shared between the GSH and hGSH biosynthesis pathways, each pathway has a dedicated second enzyme. Although the activity necessary for hGSH synthesis was easily isolated, very little was known about the homoglutathione synthetase (hGS) enzyme itself [Macinol, 1987]. Beginning in 1993, a series of glutathione synthetase (GS) crystal structures were solved and published, indicating that the enzymes fell within the ATP-grasp superfamily of proteins [Yamaguchi et. al, 1993; Kato et. al, 1994; Polekhina et. al, 1999; Gogos and Shapiro,

2002]. Because of high homology between hGS and GS sequences, as well strong structural identity between ATP-grasp family members, it became apparent that hGS would look quite similar to GS overall. However, hGS used a different substrate (β-alanine) than GS (glycine); the subtle differences in residues and residue placement required for this shift in substrate specificity could not be determined by homology modeling and necessitated the acquisition of an hGS crystal structure.

Chapter 2, "Structural Basis for Evolution of Product Diversity in Soybean Glutathione Biosynthesis", describes three x-ray crystal structures of hGS: the open-form apoenzyme, the open-form enzyme with γ-glutamylcysteine bound, and the closed-form enzyme with hGSH, ADP, and a sulfate ion bound. From these structures, in conjunction with the available hGS and GS sequences, it was concluded that two residues in the active site are primarily responsible for dictating substrate specificity between the two types of enzymes. In GS, terminal contacts with the carboxyl tail of glycine are provided by two alanine residues as part of the larger alanine-rich loop region. However, in soybean (and most other sequenced species) hGS, these two alanines are replaced by a leucine and proline; based on the hGS structure, these replacement residues pull the alanine-rich loop outwards by several Å, allowing the larger β-alanine molecule to be accommodated in the active site. The critical role of the leucine and proline residues in conferring substrate specificity to hGS was further confirmed by way of site-directed mutagenesis. Normally, hGS is not capable of catalyzing the synthesis of GSH at a physiologically relevant rate. However, the presence of one of two mutations - either leucine to alanine or proline to alanine, improves GSH catalysis by a factor of 10 or 100,

respectively. With the simultaneous inclusion of both mutations, GSH catalysis improves by nearly 1000-fold, demonstrating that the two mutations act synergistically to improve enzyme efficiency. While a 1000-fold improvement is significant, kinetic analysis of the *Arabidopsis* GS enzyme reveals a rate of glycine turnover nearly 10-fold higher than that of the mutated hGS. Thus, while the AA/LP site is clearly important for dictating substrate specificity, it can be concluded that other, as yet unknown residues also play a small but significant role in optimizing turnover rate.

A further conclusion obtained from the three solved hGS crystal structures, centers on the apparent domain movements required for enzyme catalysis. Based upon "snapshots" of various enzymes with various substrates bound, it was suspected that GS-type proteins went through some degree of domain movement as part of their catalytic cycle. However, as structures from a complete cycle were not available for a single enzyme, this suspicion could not be confirmed. By solving the structure of hGS in three forms - the apo enzyme, the intermediate enzyme-γ-glutamylcysteine complex, and the post-reaction complex - a more complete picture of the GS-type domain-movement cycle has been provided. Two regions - the lid domain (which includes the glycine-rich loop) and the previously described alanine-rich loop, appear to be the most dynamic and mobile portions of the hGS enzyme. Prior to substrate binding, these two domains are pulled back and away, allowing exposure of the active site to the surrounding environment. No change in domain structure is evident following γ-glutamylcysteine binding. When all three of γ-glutamylcysteine, ATP (represented here by ADP and a sulfate), and are in the active site, the lid domain swings inward to cover the ATP moiety, while the alanine rich

loops provides stabilizing contacts with both β-alanine and ATP. Once the reaction is complete, ADP and hGSH are released from the active site, and the enzyme resets itself into the open apoenzyme form.

As described above, x-ray crystal structures of hGS were solved in three forms. While these three structures represent key points in the reaction cycle, they do not portray every step in the mechanism. Because GS, and by extension hGS, utilize a random Ter-reactant mechanism, there are actually nine different possible combinations of reactants and products [Jez and Cahoon, 2004]. While some of these combinations are energetically unfavorable and unlikely to occur, others are part of the most-likely mechanism. In particular, the formation of an enzyme-ATP complex, and an enzyme-ATP-γ-glutamylcysteine complex is strongly favored and, in the case of the latter, necessary for completion of the reaction. In order to fill in the missing pieces of the reaction mechanism and provide insight into additional domain movements, it would be beneficial to obtain crystal structures of both of these hGS complexes in addition to those already solved. Specifically, a structure of the enzyme•non-hydrolyzable ATP analog•γ-glutamylcysteine complex could confirm whether, as indicated for the yeast GS, movement of the lid domain occurs before introduction of β-alanine into the active site pocket [Gogos and Shapiro, 2002].

There also exists the potential for a number of follow-up experiments involving the hGSH. Because hGSH has only been found in six legume species - none of which are closely related to one another - an obvious question to ask is whether hGSH is found in other legumes [Wojciechowski. et al., 2004]. This line of inquiry would necessitate the

use of HPLC-based profiling methods for various legume tissue types, and leads to questions regarding the evolutionary origins of hGSH. Because the available legume genome sequences show signs of at least one duplication event, it is likely that many other legumes contain multiple GS genes. However, there is no guarantee that these copies evolved into hGS in every species. For example, in broad bean, which does not produce hGSH, the evolution may never have occurred and/or the hGS gene(s) may be silenced. If only some legumes contain hGS genes, then it is likely that either hGS evolution was an independent event in each species or hGS evolved and then was subsequently lost in many legumes. The conclusion that hGS evolution was independent of course leads to a new query: why do all of the independently-evolved hGS enzymes utilize β-alanine over some other amino acid? This question would require an answer in two parts. First, it would be necessary to consider the availability of β-alanine (produced from uracil degradation or from spermine/spermidine) as compared to other amino acids [plantcyc.org]. Second, the specific properties, such as solubility, effectiveness as a redox buffer, and transportability, of hGSH versus other (artificially) synthesized tripeptides would need to be assessed. These experiments will also hopefully shed light on hGSH's role relative to GSH in legumes.

Redox Proteomics: Platform development with Brassica juncea (Indian Mustard)

Based on previous experiments [Jez et al., 2004; Hicks et al., 2007], it was determined that glutamate-cysteine ligase (GCL), the first enzyme in the GSH and hGSH biosynthesis pathways, is redox-regulated through the formation of two reversible

intramolecular disulfide bonds. One disulfide bond controls dimerization, while the second controls access to the active site; together under reducing conditions they inactivate the enzyme. Although GCL is not the only enzyme to utilize a thiol-based redox regulatory mechanism, it is one of the first - besides NPR1 - to be identified in plants. Prior work has largely focused on targets of the redox-regulatory proteins thioredoxin and glutaredoxin; while the column chromatography-based techniques employed are sound, they select for proteins with strong interactions, meaning that low abundance targets or targets that bind to a different redox-regulatory protein are missed [Balmer et al., 2003; Motohashi et al., 2001; Yano et al., 2001]. To circumvent this problem, an alternative non-column-based dual redox labeling strategy was developed. In short: the free thiol groups of proteins from a plant extract are first labeled with an alkylating agent. Next, oxidized thiols are reduced and labelled with a thiol-labile fluorescein derivative. The protein mixture is then separated by isoelectric point and molecular weight via 2D-SDS-PAGE. Protein spots containing oxidized thiols can be visually identified, as fluorescein derivatives fluoresce under a narrow range of wavelengths. The gel spots can then be excised, proteins digested, peptides analyzed by LC-MS/MS, and spectra matched via MASCOT to yield a likely peptide identity. Using this approach, the effect of different oxidizing treatments upon protein thiol state can be compared. If the gels are additionally exposed to a total protein stain, then changes in total protein expression as a result of treatment can be compared as well.

Chapter 3, “Redox-regulatory mechanisms induced by oxidative stress in *Brassica juncea* roots monitored by 2-DE proteomics”, describes a series of experiments in which,

B. juncea roots were exposed to different oxidizing agents, and the resulting redox and total protein expression changes were measured. Following 2D-SDS-PAGE, treatment of roots with either BSO or H_2O_2 yielded 50 and 59 gel spots, respectively, that differed significantly in thiol redox-status between the control and treated samples. A further 40 and 27 spots, respectively, displayed significant and non-overlapping changes in protein expression between the control and treated samples. From these initial results, it was concluded that both BSO and H_2O_2 directly affect both thiol-redox status and expression of root proteins. A closer look at the significant spots detected revealed that a subset of the detected spots - approximately 11-17 depending on the treatment combination - contained a single protein, while the remainder contained multiple proteins that co-migrated on the 2D-gel. Because of the difficulty in assigning definitive redox and expression fold change values to a single protein within a multi-protein spot, only spots containing a single protein were analyzed in greater detail. Of those 52 single protein spots, a large number were found to contain proteins involved in maintaining redox homeostasis. Given that BSO and H_2O_2 are a source of endogenous and exogenous ROS, respectively, this result was not surprising. Perhaps more surprising however was the simultaneous identification of proteins involved in pathways less commonly associated with ROS management: namely glycolysis, carbohydrate metabolism, and amino acid biosynthesis, among others. Widespread upregulation/increased oxidation of metabolic enzymes such as O-acetylserine sulfhydrylase and malate dehydrogenase suggest that existing front-line redox mechanisms may be insufficient to combat ROS exposure, and

that metabolic reallocation may be required to support the synthesis of additional redox compounds to maintain cellular state.

An additional conclusion from these experiments stems from the identification of proteins associated with the H_2O_2 and BSO tissue treatments: of the proteins from the 52 single protein spots, not one was identified in more than one combination of treatments (H_2O_2/BSO) and detection methods (SYPRO/IAF). This result indicates that, in *B. juncea* at least, ROS initiates different redox-response mechanisms depending on whether the source is endogenous or exogenous. The lack of overlap between SYPRO- and IAF-significant proteins further indicates a disconnect between protein expression changes and redox state: for the proteins identified, a change in redox-regulation or a change in expression - never both - may alter flux through their respective redox response pathways.

Among the proteins identified in this series of experiments, several - including 3-phosphoshikimate 1-carboxyvinyltransferase (a component of the shikimate pathway leading to tyrosine, tryptophan, and phenylalanine biosynthesis) and 5-methyltetrahydropteroyltri-glutamate-homocysteine S-methyltransferase (from the methionine biosynthesis pathway), have not been previously identified as containing redox-sensitive thiol groups. Due to the nature of the dual labeling strategy, it is not possible to determine from the existing data whether these proteins contain reversible disulfides or are modified by glutathione, etc. As these proteins may represent regulatory control points in their respective pathways, it would be beneficial to identify possible redox-regulatory mechanisms through the use of a more targeted functional studies.

One of the goals of this set of experiments was to develop a methodology that could be used to probe the thiol-redox proteome in a variety of plant species and tissue types under varying conditions. Based on the results as previously described, that goal has been met. The next step, application to different plants under other oxidative stress conditions has already been initiated, and the first round of results from those experiments are described in Chapter 4.

Connecting Proteomes and Climate Change: Ozone-Induced Changes in the Total and Redox Proteomes of Glycine max (Soybean)

Once a suitable thiol-labeling strategy (described in Chapter 3), had been developed, the next step to employ it with a more agriculturally relevant and redox-sensitive crop, such as soybean. To do this, we established a collaborative effort with the laboratory of Lisa Ainsworth in the USDA-ARS group at the University of Illinois-Urbana-Champaigne. The Soybean Free Air Concentration Enrichment (SoyFACE) facility at UI-UC uses a ring-based ozone exposure setup to maintain concentrations over crops by way of a computer controlled system. Because the soybean plants are sown directly in the soil, problems obtaining sufficient light and water while being exposed are avoided, and as an added benefit the plants are able to grow in a natural, exposure-realistic environment.

As described in Chapter 4, "From crops to climate change: redox proteomics of soybean ozone responses," soybean plants were field grown under three different ozone concentrations using FACE technology, and differences in expression and thiol-content

assessed via 2D-SDS-PAGE and LC-MS/MS. To confirm that the observed expression changes translated into appreciable changes in protein activity, activity assays were also conducted for a number of implicated proteins.

Although previous studies have investigated the effects of acute ozone on the soybean proteome, none have reported on the consequences of chronic exposure over a growing season. Moreover, this work is the first to use plants grown in the field with relevant day-night ozone exposure cycles versus growth chamber experiments in which plants are constantly exposed to high ozone levels for a few days. This work indicates that a number of proteins, particularly those involved in primary metabolism, redox homeostasis, and amino acid metabolism, experience large shifts in their expression and redox-state. While a subset of these proteins have been previously identified in other ozone experiments, many represent new additions to the redox-responsive proteome in plants. In this section, I will focus on the protein classifications in which some of the largest and/or most abundant fold changes in expression and redox-state took place. I will also discuss observations of both redox-regulation and redox-response.

i. Ozone induces changes in primary metabolism and amino acid biosynthesis

At concentrations greater than 40 ppb, tropospheric ozone begins to negatively affect soybean growth and yield. Above ground, these changes visibly manifest themselves as decreased shoot and pod biomass, fewer pods produced, and premature leaf senescence. Within the leaves themselves, ozone entering through the stomata is converted to ROS, which irreversibly oxidizes the plasma membrane and photosystem

components, and results in the degradation of chlorophyll. Although no comparable work has been done with soybeans, acute exposure of clover to 75 ppb ozone resulted in slower root tip formation and elongation [Vollsnes et al., 2010]. Likewise, in potatoes, chronic exposure to 80-120 ppb O3 resulted in necrosis and vascular damage above ground, and decreased tuber size and yield below ground [Asensi-Fabado et al., 2010].

Given that the primary site of plant O3 exposure is the leaves, it follows that many of the largest fold changes in the proteome occur there. Typically, in response to oxidative damage, photosynthetic output and Calvin cycle activity are downregulated. To compensate, plants draw on reserve energy stores, resulting in a decrease in leaf starch concentrations and a corresponding increase in the sucrose concentration [Ahsan et al., 2010]. Accordingly, enzymes involved in sugar catabolism must also be upregulated; previous studies have identified several primary metabolism proteins - among them malate dehydrogenase and phosphoglycerate mutase - that are increasingly expressed under ozone stress [Bohler et al., 2007]. My results are consistent with this view; in leaves exposed to 115 ppb O3, there is upregulation and/or increased oxidation of multiple primary metabolism enzymes including malate dehydrogenase, phosphoribulokinase (PRK), and glyceraldehyde-3-phosphate dehydrogenase (G3PDH). The latter two enzymes are thought to form a complex that is redox-regulated by thioredoxin in response to changes in light intensity [Howard et al., 2008]. Although ozone-produced ROS damage cellular light-harvesting capacity, they also lend themselves to an increasingly oxidizing environment. In this work, PRK and G3PDH are both approximately 4-fold more oxidized in the 115 ppb-treated tissue relative to the

control. Likewise, activity assays indicate that G3PDH is approximately 3.7-fold more active in the 115 ppb-treated tissue relative to the control. Like PRK and G3PDH, malate dehydrogenase (MD) is an important regulatory control point, serving as the rate-limiting step of the TCA cycle. Previous studies have shown that cytosolic MD is redox-regulated, with reversible inactivation occuring under oxidizing conditions [Hara et al., 2006]. In this study, MD was detected in a number of different tissue-O3 concentration combinations, with the highest upregulation (2-fold) and oxidation (5-fold) occurring in leaves exposed to 115 ppb ozone. Though oxidation is thought to decrease MD activity, this study sees greater activity in leaves exposed to 115 ppb ozone as compared to leaves exposed to ambient conditions. This is likely due to the fact that, even under oxidizing conditions, reduced protein makes up a much higher percentage of the total protein than oxidized protein, and thus a 2-fold change in expression results in a proportionally larger percentage of reduced - and in this case active - protein being available.

As expression of key proteins increases in response to oxidative stress, the demand for amino acids used in protein synthesis also increases. In leaves exposed to 115 ppb ozone, the upregulation and/or increased oxidation of a number of proteins involved in amino acid biosynthesis, including aspartate-semialdehyde dehydrogenase (ASADH), glutamine synthetase (GS), and phosphoglyerate kinase (PK), was observed.

The first enzyme, ASADH, is of particular importance in plants, as it represents the primary control point in the biosynthetic pathways responsible for the production of isoleucine, methionine, lysine, and threonine. Experiments conducted in *E. coli* suggest that one of the enzyme's catalytic cysteine residues can be reversibly reduced to control

enzyme activity; however no comparable redox analysis has been conducted using a plant variant of the enzyme [Alvarez et al., 2003]. This study is the first to identify the soybean enzyme as a possible redox-regulation target, and as a component of the plant oxidative-response mechanism.

Like ASADH, GS represents a major metabolic control point in plants. The enzyme, along with its partner enzyme glutamate synthetase, is part of the GS-GOGAT cycle and is necessary for assimilation of ammonium in plants. GS itself catalyzes the ATP-dependent fixation of ammonium to glutamate to form glutamine, while glutamate synthetase catalyzes the synthesis of two molecules of glutamate from one molecule of glutamine and one molecule of 2-oxoglutarate [Bernard and Habash, 2009]. The active sites of both cytosolic and plastidic isoforms of GS are known to contain one or more reaction-critical cysteine residues; redox modification of these residues presents a likely means for attenuating GS enzyme activity [Bernard and Habash, 2009; Choi et al., 1999].

In the 2D-gel analysis, PK was observed to be upregulated and more oxidized in response to oxidative stress. Although this enzyme is a key component in the serine biosynthesis pathway (which leads to production of the redox-critical amino acid cysteine), I was unable to detect a difference in enzyme activity between the control and 115 ppb ozone-treated extracts. This is likely due to the fact that the isoforms of the enzyme were all identified in spots containing several other proteins, including fructose-bisphosphate aldolase (FBA). Since activity assays confirmed a difference in FBA activity between the control and treated extracts, it is likely that this (and other enzymes) account for the activity in those spots and that PK is a false positive.

ii. Ozone affects redox homeostasis and induces stress response proteins

Given that tropospheric ozone is a source of ROS, it is not surprising a number of proteins involved in maintaining cellular redox homeostasis were differentially regulated and/or oxidized in response to ozone exposure. Across several different tissue-type/ ozone concentration combinations, multiple different isoforms of glutathione-s-transferase (GST) were detected. In plants, GSTs comprise a large family of enzymes that, as their name suggests, conjugate glutathione to a variety of substrates to aid in xenobiotic detoxification or to serve a redox-protective function (Dixon et al., 1998). In our study different isoforms were both up or down regulated, more oxidized or more reduced in the treated tissues relative to the controls, but this is to be expected given the diversity of reactions that GSTs catalyze. In the roots of soybean plants treated with 60 ppb ozone, I detected an uncharacterized isoform of thioredoxin that was upregulated relative to the control tissue (Table 2d). By homology (BLAST), this thioredoxin is likely an m-type, meaning that it is localized to the chloroplasts. Given that the ROS generated from ozone degradation arguably have their greatest effect against redox-sensitive chloroplast components, the upregulation of such a thioredoxin is logical. This thioredoxin was not observed as being differentially regulated/expressed in the roots of plants exposed to 115 ppb ozone, nor was it observed in leaf tissue, suggesting that its expression may be tied to the degree of oxidative damage experienced by components of the "dark reactions" of carbon fixation.

Methionine sulfoxide reductase (MSR) also appears to be an important component of the soybean ozone response and was observed to be 2.7-fold more oxidized

in this study following exposure to 115 ppb ozone. This enzyme is responsible for recycling the limiting amino acid methionine through the reduction of methionine sulfoxide. The reaction mechanism for the bacterial enzyme proceeds through the formation of a intramolecular, thioredoxin-mediated, disulfide bond [Antoine et al., 2003]. Although it follows that the enzyme would be more oxidized in an ozone-rich environment, additional work is needed with the plant enzyme to confirm that the same reaction mechanism applies.

Like MSR, the enzyme ascorbate peroxidase (AP) also plays an important redox-protective/corrective role. Ascorbate, as described in the introduction, is one of the three major redox couples found in plants; AP, as part of the larger ascorbate-glutathione cycle, utilizes ascorbate as a substrate in order to detoxify hydrogen peroxide and other peroxides produced from oxidative bursts [Noctor and Foyer, 1998]. In this study, total AP (isoform 2) expression was up-regulated approximately 2.5-fold in leaves exposure to 115 ppb ozone, while AP (isoform 1) was down-regulated 1.5-fold in 115 ppb-exposed roots. The cause of these differing expression profiles in unclear, though it may relate to the redistribution of resources from more moderately oxidized (roots) to more severely oxidized (leaves) tissue types.

iii. Ozone induces both a redox response and changes in redox-regulation

Because ozone, as a denser-than-air gas, is ground-hugging but not ground-penetrating, it follows that the largest effects of high tropospheric ozone concentrations would be found in aerial plant tissues. In this study, the largest fold-changes in protein

expression and redox state were found in leaves exposed to 115 ppb ozone, and overall more than half of the significant proteins were detected in leaf tissue. However, a sizable minority of proteins were identified exclusively in root tissue. Given that this tissue was not exposed to ozone, several hypotheses exist concerning why local changes in expression and redox state were observed. One possibility is that ozone-induced ROS act as a propagated signaling molecules that directly prompt changes in protein redox state and regulation. Ozone, besides its capacity for degradation into multiple ROS, is also known to prompt the rapid release of H_2O_2 into the cell apoplast as a result of even short term exposure [Rao and Davis, 2001]. H_2O_2 is able to diffuse into cells through oxidative-gated aquaporins, and can function as an intercellular signal for activating the hypersentitive response and other plant defense mechanisms [Henzler et al., 2004; Henzler and Steudle, 2000]. In this situation, ROS can not only directly regulate the activity of anti-oxidation enzymes, but can also instigate a secondary redox response through the activation of transcription factors and other protein expression machinery. However, given their high reactivity, it is unlikely that most ROS would be able to diffuse more than one or two cells from their site of origin before being consumed [Murphy et al., 2001]. Although it is theoretically possible that ROS could serve as a long-distance signaling molecule by propagating from cell at a time, the majority of evidence indicates that it is confined to localized intracellular responses.

Another hypothesis for explaining the observed root response to ozone is that another small molecule besides the various ROS is responsible for signal transduction. Out of the other numerous hormones and peptides known to play a role the plant defense

response, calcium is the most likely candidate. Besides the advantage it gains by being easily diffused, calcium is also normally maintained in cells for controlling channel flux, and has previously been implicated as a secondary messenger in several other signal transduction networks. Prior work in *Arabidopsis* has indicated that roots and shoots independently undergo large changes in calcium flux as a result of ozone exposure; however, it is unclear how much of that flux crosses the shoot/root boundary [Evans et al., 2005]. Assuming that signal transduction does at least in part proceed from the exposed shoots to the roots, then the changes in expression we see in soybean root tissue are likely a secondary redox response, with changes in protein thiol status brought on by secondary local calcium-induced oxidative bursts.

A third possibility for explaining the observed root response to ozone, is that the roots themselves are sensing the small amount of ozone able to diffuse from the surface, and are responding independently of the shoot tissue. Previous studies have confirmed that exposed roots can mount their own response to ROS independent of the aerial tissue if the two portions are physically separated [Rentel and Knight, 2004]. If this is the case, then our observed changes expression and thiol status may again be a result of primary redox regulation of local proteins, with secondary redox responses occurring beyond the sites of exposure. In all actuality though, it is likely that a combination of all three methods of signal propagation are occurring in soybeans as a result of ozone exposure; further local real-time analysis may be required to fully tease out the differing degrees of regulation and response observed, particularly in the below-ground tissues.

iv. Future Work

In chapter 4, 159 proteins were identified that differentially respond and/or are regulated by changes in tropospheric ozone concentration. Seventy-nine of these proteins saw significant changes in their expression and/or oxidation state following exposure to ozone at a concentration of 115 ppb. Many of these proteins (Figure 4 of Chapter 4) represent critical control points for the maintenance of carbon metabolism. The identities and distribution of proteins identified in soybean plants exposed to 115 ppb ozone, as compared to plants exposed to a lower ozone concentration, suggests that the observed metabolic changes may result from acclimation of the plant to chronic ozone exposure. Although altered total and redox proteomes were examined in soybean, it is unclear if these changes correspond to changes in carbon flux and sugar mobilization. That is to say, are more metabolites being pushed through the system, and from where is the carbon being obtained?

To examine the effect of chronic ozone exposure on soybean metabolism it would be beneficial if subsequent analyses of the soybean ozone-redox proteome were carried out in conjunction with analyses of the soybean metabolome. Using standard GC- and HPLC-based methodologies, it would be possible to compare the relative pool sizes of key metabolites (glycolysis, TCA cycle, and amino acid precursors, etc.) from ozone-exposed and ambient-exposed soybean tissues. To quantify changes in storage metabolites, fatty acids from crude soybean extract could be converted to fatty acid methyl esters (so as to make them more polar and volatile), and separated by GC [www.gerstel.com]. Metabolite identity and quantity could then be determined via

comparison to the retention time and known make-up of a standard mixture; identity could also be determined via GS-MS [Lehmann et al., 2009]. For other primary and secondary metabolites (including amino acids), both polar and non-polar extractions from the initial tissue would be carried out. Metabolites in the polar phase would be separated by HILIC (Hydrophobic Interaction Liquid Chromatography) so as to improve their volatility for downstream electrospray ionization, while metabolites in the non-polar phase would be separated by traditional reverse phase chromatography [Grumbach et al., 2004]. As above, metabolite identity and quantity could be determined in both cases using a standard injection mixture; identifications could also be confirmed via LC-MS. The resulting data from this series of experiments would indicate whether observed upregulation of protein expression corresponded to changes in carbon metabolism, and at what stage of metabolism reallocation of resources was occurring to fuel *de novo* protein synthesis.

To complement the above approach, a C^{14}-based feeding study could also be carried out. In short, leaf disks from ambient-exposed and high ozone-exposed soybeans would be soaked in buffer containing one of several C^{14}-labeled metabolites. At pre-determined time-points, the leaf disks would then be flash frozen, and leaf metabolites extracted and fractionated by TLC. The radioactive plates would then be imaged using x-ray film, and radioactive counts for individual spots would be determined by liquid scintillation [Katahira and Ashihara, 2009]. Alternately, C^{13} could used in place of C^{14} in conjunction with fractionation and separation via GS-MS to improve both metabolite separation and overall sensitivity [Feng et al., 2010, Tang et al., 2010]. Because C13

occurs naturally at approximately 1 out of every 100 possible carbon atoms, care must be taken to avoid confusing background C13 with that resulting from labeling. To obtain absolute quantification, GC metabolite profiles for both ambient- and high ozone-treated tissues could be compared against a standard injection containing known quantities of metabolites of interest. By repeating this experiment for different lengths of time and with different starting metabolites (in particular different sugars, starches, and photosynthetic precursors), it should be possible to determine not only which metabolic pools are changing, but also from where the carbon necessary for observed metabolite and protein changes is being sourced.

For the work described in Chapter 4, tissue from soybean exposed to three different ozone concentrations (115, 60, and 40 ppb) was harvested at the R3 stage of development for further analysis; however, acclimation to chronic ozone exposure is likely a time-dependent process and further analysis should involve investigation of changes at various soybean developmental stages across the growing season. While the initial experiments (Chapter 4) provide new insights into ozone-induced changes to the soybean proteome, in particular the existence of a response threshold between 60 and 115 ppb, the results only reflect a snapshot along a continuum of time- and dose-dependent responses. By the R3 stage of development, the expression and activity of a number of metabolomic proteins has been altered, while other ozone-damage marker proteins (in particular RuBisCO and PEPC) remain as yet unchanged. However, it is unclear how early during soybean growth and exposure the observed changes are manifested, and whether or not the changes remain consistent as the growing season progresses. By

harvesting and analyzing tissue from additional timepoints, a more complete picture of the soybean temporal redox response, including changes in the specific activities of key proteins, could be established. As with the additional timepoints, the collection and analysis of tissues exposed to alternate ozone concentrations, in particular those above 115 ppb and those between 60 and 115 ppb, would provide new and valuable insights. While the existing data implies the presence of a ozone response threshold between 60 and 115 ppb, it is unclear at what ozone concentration the transition occurs. Better knowledge of the threshold for changes in the ozone proteomic response may allow future farmers to enact preventative measures as per predictions for daytime and seasonal tropospheric ozone concentrations.

Analysis of tissues exposed to ozone concentrations above 115 ppb would allow for better understanding of how ozone affects the proteome under future-predicted conditions. Growth chamber studies examining continous or acute exposure to have indicated that acute exposure to ozone concentrations in excess of 100 ppb results programmed cell death [Pell et al., 1997; Chen et al., 2009; Overmyer et al., 2003]. A better understanding of physiological effects resulting from the natural diurnal ozone cycle and the presence of variable weather conditions may allow field grown soybean to better mitigate the toxic effects of ozone. Moreover, proteomic and metabolic strategies may also help in the identification of molecular targets for engineering of ozone tolerant strains of soybean.

v. To Build a Better Soybean Plant ...

Because current soybean cultivars are not ozone tolerant, and because traditional breeding takes more time than can be afforded, biotechnology to engineer soybeans presents the most reasonable avenue for generating ozone resistant varieties, but this requires significant understanding of the molecular responses to this abiotic stress. Based on the completed experiments, it is clear that not just individual proteins, but rather whole pathways are involved in maintaining the soybean ozone response. While it would be arguably advantageous to re-engineer all of proteins of glycolysis to increase their substrate affinities under high stress conditions, it is much more so within the realm of practicality to focus on one or two proteins at a time. Because many of the proteins identified in this work are redox regulated, or may yet be identified as such, one possible avenue is to increase the availability of reducing equivalents to help redox-regulated proteins maintain an active redox state during exposure to ROS.

The majority of proteins identified as differentially oxidized following ozone exposure are known targets of thioredoxin or glutaredoxin; thus the "doxins" may represent a starting point for manipulating the soybean ozone response. Unfortunately, both thioredoxins and glutaredoxins exist as protein families within a given species (for example, *A. thaliana* has at least 19 and 22 different thioredoxins and glutaredoxins, respectively), making the choice of one or two to alter difficult in the absence of detailed interaction data [Buchanan and Balmer, 2005]. Because both glutaredoxins and thioredoxins ultimately derive their reducing power from NADPH, another option for increasing the availability of reducing equivalents is to target chloroplast ferredoxin.

Like the other doxins, ferredoxin is a member of a large protein family; however, plant species tend to contain relatively few ferredoxin-encoding genes (*A. thaliana* has 4), a subset of which encode chloroplast-specific ferredoxins [Hanke et al., 2004; Fukuyama, 2004]. Although NADPH production is ultimately tied to light intensity, the presence of additional ferredoxin activity may aid in the scavenging of elections that would otherwise be lost from photosystem I under high light, and additionally help present the formation of additional ROS.

As an alternative to providing cells with additional reducing equivalents to combat ozone stress, it may be possible to engineer plants to better neutralize ROS at their point of entry - prior to the induction of damaging oxidative cascades. Ozone enters a plant primarily via the stomota, and in the apoplastic space it is rapidly converted to various toxic ROS. Prior research in broad bean has indicated that ascorbate and dehydroascorbate concentrations rise rapidly following ozone exposure [Luwe and Heber, 1995; Kangasjarvi et al., 2005]. As ascorbate is one of the three major redox couples employed by plants, this observation suggests an effort by the plants to quickly stave off the creation of damaging ROS. In line with this observation, we also noted the upregulation of ascorbate perodixase - an enzyme that utilizes ascorbate to detoxify peroxides - in leaves exposed to 115 ppb ozone. While ascorbate has been implicated as a major player in apoplasmic detoxification and remains a focus of ongoing studies [Burkey et al., 2003; Ainsworth et al., 2008], other redox reactive molecules may provide more subtle contributions. Many other compounds, including flavonols, flavones, isoflavonoids, and anthocyanins, also display ROS scavenging capabilities, and have

been found to be localized in part to the apoplast. For example, the presence of the flavonol quercetin has been noted in the oxidized outer scales of brown onion, and the isoflavonoids daidzein and sojagol have been identified in ozone-stressed soybean leaves [Takahama and Hirota, 2000; Keen and Taylor, 1975]. Despite clear benefits associated with many of these compounds in response to oxidative damage, work in the area has been slow, due in part to inherent metabolite and metabolite-derivative toxicity at higher concentrations [Didyk and Blum, 2006; Bais et al., 2003; Parvez et al., 2004].

In the work described in Chapter 4, we noted that expression of two isoforms of chalcone isomerase were upregulated relative to the control in leaves exposed to 115 ppb ozone. This enzyme presents an interesting target for improving plant ozone tolerance in that it is directly upstream of the enzymes responsible for isoflavone, flavonone, and anthocyanin biosynthesis, among other ROS scavenging metabolites. While based on our work and available crystal structures chalcone isomerase does not appear to be redox regulated at the protein level, it is clear that the enzyme responds in some capacity to oxidative cues [Ferrer et al., 2008]. Limited investigations in duckweed (*Lenma gibba*) have indicated that chalcone isomerase trancript accumulation is tied to the inhibition of chloroplastic electron transport, though it remains unclear whether other signaling factors are involved [Akhtar et al., 2010]. Accordingly, further investigation of the chalcone isomerase promotor and its redox responsiveness is warranted. Depending on the motifs and modes of regulation found during study of the chalcone isomerase promoter, it may be possible to adjust expression of the enzyme so that its production is more tightly linked to increases in atmospheric ozone concentration. In that eventuality, the effect of

greater chalcone isomerase availability on downstream redox metabolite accumulation, and in turn their ability to disarm invading ozone-based ROS, would also be prime targets for further investigation.

An additional strategy could be to target the enzymes of carbon metabolism in an effort to maintain core metabolism at levels that maintain sufficient plant growth and seed yield. As another target for improving soybean ozone tolerance, RuBisCO at first glance appears to be an excellent candidate. The enzyme, as the lynchpin of the Calvin cycle, is responsible for the carboxylation of ribulose-1,5-bisphosphate, and accounts for 30-50% of the soluble protein present in leaves [Feller et al., 2008]. While higher soluble RuBisCO concentrations are unlikely to be achievable do to limiting nitrogen, common sense would dictate that by increasing the enzyme's catalytic efficiency, it should be possible to increase overall carbon flux [Parry et al., 2003]. Extensive efforts have been undertaken to increase CO_2 specificity via directed evolution, as well as through the formation of hybrid cross-species holozymes; however, it appears that CO_2 specificity is closely tied to O_2 specificity (RuBisCO also displays alternate oxygenase activity), and thus no significant improvements have been made [Parry et al., 2003; Mueller-Cajar and Whitney, 2008].

In light of these difficulties, efforts to improve carbon flux have also been directed at key RuBisCO-interactor: RuBisCO activase (RA). RA catalyzes the removal of the inhibitor carboxyarabinitol 1-phosphate (CA1P) from the RuBisCO active site in the absence of a key carbamylation modification required for catalysis. As a regulator of RuBisCO activity, RA is in turn regulated via several distinct mechanisms. Because the

enzyme displays ATPase activity (as ATP hydrolysis is required for CA1P removal), it is sensitive to the ratio of ADP:ATP present in the chloroplast. At a ratio of 1:1, common during the nighttime hours, RA activity is minimal. As the light reactions of photosynthesis become active however, this ratio increases to 1:2 or 1:3, and RA activity increases accordingly [Zhang and Portis, 1999; Kallis et al., 2000]. In addition to ADP:ATP ratio, some isoforms of RA are also regulated by temperature and the formation of a reversible disulfide bond. In plants, some species contain a gene encoding only a single, short form of RA, while others produce an additional long form splice variant or express a longer isoform from a distinct gene [Portis, 2003]. While the larger isoform appears more sensitive to the ADP:ATP ratio, it is also generally more thermo-stabile as well [Shen et al. 1991]. The larger isoform additionally is regulated through the formation of a thioredoxin-mediated disulfide bond at the C-terminus of the protein [Zhang and Portis, 1999]. The isoforms appear to be co-expressed, and their is limited evidence to suggest that redox regulation of the larger isoform can effect the activity of the smaller isoform, though the extent of their interaction remains unclear [Zhang et al., 2001].

In this study, two different isoforms of soybean RA, (one of which was differentially oxidized), were identified in leaves treated with 115 ppb ozone. Because the large isoform is already redox regulated, it presents an excellent target for improving carbon flux and oxidation tolerance in crop plants. In *Arabidopsis*, directed evolution of the small isoform has resulted in several variants that display increased thermo-tolerance and secondarily improve CO_2 assimilation rates and seed yield [Kurek et al., 2007].

Since tropospheric ozone concentrations are tied to temperature, improvements to the thermo-tolerance alone of the soybean large RA isoform could confer some benefit. Using additional directed evolution-based strategies, such as the blending of isoforms from various species, it may be possible to generate an RA variant that improves soybean ozone tolerance and helps maintain seed yield.

References

Ahsan, N., Nanjo, Y., Sawada, H., Kohno, Y., Komatsu, S. (2010). Ozone stress-induced proteomic changes in leaf total soluble and chloroplast proteins of soybean reveal that carbon allocation is involved in adaptation in the early developmental stage. Proteomics 10: 2605-2619.

Akhtar, T.A., Lees, H.A., Lampi, M.A., Enstone, D., Brain, R.A., Greenberg, B.M. (2010). Photosynthetic redox imbalance influences flavonoid biosynthesis in *Lemna gibba*. Plant, Cell, and Environment 33: 1205-1219.

Ainsworth, E.A., Rogers, A., Leakey, A.D.B. (2008). Targets for crop biotechnology in a future high-CO_2 and high-O_3 world. Plant Physiology 147: 13-19.

Alvarez, E., Ramon, F., Magan, C., Diez, E. (2003). L-cystine inhibits aspartate-beta-semialdehyde dehydrogenase by covalently binding to the essential ^{135}Cys of the enzyme. Biochimica et Biophysica Acta 1696: 23-29.

Antoine, M., Boschi-Muller, S., Branlant, G. (2003). Kinetic characterization of the chemical steps involved in the catalytic mechanisom of methionine sulfoxide reductase a from *Neisseria meningitidis*. Journal of Biological Chemistry 278: 45352-45357.

Asensi-Fabado, A., Garcia-Breijo, F.J., Reig-Arminana, J. (2010). Ozone-induced reductions in below-ground biomass: an anatomical approach in potato. Plant, Cell, and Environment 33: 1070-1083.

Bais H.P., Walker T.S, Kennan A.J., Stermitz F.R., Vivanco J.M. (2003). Structure-dependent phytotoxicity of catechins and other flavonoids: flavonoid conversions by cell-free protein extracts of *Centaurea maculosa* (spotted knapweed) roots. Journal of Agricultural Food Chemistry 51:897–901.

Balmer, Y., Koller, A., del Val, G., Manieri, W., Schürmann, P., Buchanan, B.B. (2003). Proteomics gives insight into the regulatory function of chloroplast thioredoxins. PNAS 100: 370-375.

Bernard, S.M., Habash, D.Z. (2009). The importance of cytosolic glutamine synthetase in nitrogen assimilation and recycling. New Phytologist 182: 608-620.

Bohler, S., Bagard, M., Oufir, M., Planchon, S., Hoffman, L., Jolivet, Y., Hausman, J.-F., Dizengremel, P., Renaut, J. (2007). A DIGE analysis of developing poplar leaves subjected to ozone reveals major changes in carbon metabolism. Proteomics 7: 1584-1599.

Buchanan, B.B., Balmer, Y. (2005). Redox regulation: a broadening horizon. Annual Review of Plant Biology 56: 187-220.

Burkey, K.O., Eason, G., Fiscus, E.L. (2003). Factors that affect leaf extracellular ascorbic acid content and redox status. Plant Physiology 117: 51-57.

Chen, C.,P, Frank T.,D., Long S.,P. (2009) Is a short, sharp shock equivalent to long-term punishment? Contrasting the spatial pattern of acute and chronic ozone damage to soybean leaves via chlorophyll fluorescence imaging. Plant Cell and Environment 32: 327-335

Cho, K., Shibato, J., Agrawal, G.K, Jung, Y.-H., Kubo, A., Jwa, N-S., Tamogami, S., Satoh, K., Kikuchi, S., Higashi, T., Kimura, S., Saii, H., Tanaka, Y., Iwahashi, H., Masuo, Y., Rakwal, R. (2008). Integrated Transcriptomics, Proteomics, and Metabolomics Analyses to Survey Ozone Responses in the Leaves of Rice Seedlings. Journal of Proteome Research 7: 2980-2998.

Choi, Y.A., Kim, S.G., Kwon, Y.M. (1999). The plastidic glutamine synthetase activity is directly modulated by means of redox change at two unique cysteine residues. Plant Science 149: 175-182.

Didyk, N.P., Blum, O.B. (2011). Natural antioxidants of plant origin against ozone damage of sensitive crops. Acta Physiol Plant 33: 25-34.

Dixon, D.P., Cummins, I., Cole, D.J., Edwards, R. (1998). Glutathione-mediated detoxification systems in plants. Current Opinion in Plant Biology 1: 258-266.

Gogos, A., and Shapiro, L. (2002). Large Conformational Changes in the C a t a l y t i c Cycle of Glutathione Synthase. Structure 10: 1669-1676.

Grumbach, E.S., Wagrowski-Diehl, D.M., Mazzeo, J.R., Alden, B., Iraneta, P.C. (2004). Hydrophilic interaction chromatograpjy using silica columns for the retention of polar analytes and enhanced ESI-MS sensitivity. LCGS North America 22: 1010-1023.

Guidi, L., Degl'Innocenti, E. (2008). Ozone effects on high light-induced photoinhibition in *Phaseolus vulgaris* Plant Science 174: 590-596.

Evans, N.H., McAinsh, M.R., Hetherington, A.M., Knight, M.R. (2005). ROS perception in *Arabidopsis thaliana*: the ozone-induced calcium response. The Plant Journal 41: 615-626.

Feller, U., Anders, I., Mae, T. (2008). Rubiscolytics: fate of Rubisco after its enzymatic function in a fell is terminated. Journal of Experimental Botany 59: 1615-1624.

Feng, X., Tang, K.-H., Blankenship, R.E., Tang, Y.J. (2010). Metabolic flux analysis of the mixotrophic metabolisms in the green sulfur bacterium *Chlorobaculum tepidum*. Journal of Biological Chemistry 285: 39544-39550.

Ferrer, J.L., Austin, M.B., Stewart, C. Jr, Noel, J.P. (2008). Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. Plant Physiology and Biochemistry 46: 356–370

Fukuyama, K. (2004). Structure and function of plant-type ferredoxins. Photosynthesis Research 81: 289-301.

Hanke, G.T., Kimata-Ariga, Y.,m Taniguchi, I., Hase, T. (2004). A post genomic characterization of *Arabidopsis* ferredoxins. Plant Physiology 134: 255-264.

Hara, S., Motohashi, K., Arisaka, F., Romano, P.G.N, Hosoya-Matsuda, N., Kikuchi, N., Fusada, N., Hisabori, T. (2006). Thioredoxin-h1 reduces and reactivates the oxidized cytosolic malate hydydrogenase dimer in higher plants. Journal of Biological Chemistry 281: 32065-32071.

Henzler, T., Ye, Q., Steudle, E. (2004). Oxidative gating of water channels (aquaporins) in *Chara* by hydroxyl radicals. Plant Cell and Environment 27: 1184-1195.

Henzler, T., Steudle, E. (2000). Transport and metabolic degradation of hydrogen peroxide in *Chara corallina*: model calculations and measurements with the pressure probe suggest transport of H_2O_2 across water channels. Journal of Experimental Botany 51: 2053-2066.

Hicks, L.M., Cahoon, R.E., Bonner, E.R., Rivard, R.S., Sheffield, J., Jez, J.M. (2007). Thiol-based regulation of the redox-active glutamate-cysteine ligase from *Arabidopsis thaliana*. Plant Cell 19: 2653-2661.

Howard, T.P., Metodiev, M., Lloyd, J.C., Raines, C.A. (2008). Thioredoxin-mediated reversible dissociation of a stromal multiprotein complex in response to changes in light availability. PNAS 105: 4056-4061.

Jez, J. Cahoon, R. (2004). Kinetic Mechanism of Glutathione Synthetase from *Arabidiosis thaliana*. Journal of Biological Chemistry 279(41): 42726-42731.

Kallis, R.P., Ewy, R.G., and Portis, A.R. Jr. (2000). Alteration of the adenine nucleotide response and increased Rubisco activation activity of *Arabidopsis* Rubisco activase by site-directed mutagenesis. Plant Physiology 123: 1077–1086.

Kangasjarvi, J., Jaspers, P., Kollist, H. (2005). Signalling and cell death in ozone-exposed plants. Plant, Cell, and Environment. 28: 1021-1036.

Katahira, R., Ashihara, H. (2009). Profiles of the biosynthesis and metabolism of pyridine nucleotides in potatoes (*Solanum tuberosum L.*). Planta 231: 35-45.

Kato, H., Tanaka, T., Yamaguchi, H., Hara, T., Nishioka, T., Katsube, Y. Oda, J. (1994). Flexible Loop that is Novel Catalytic Machinery in a Ligase. Atomic Structure

and Function of the Loopless Glutathione Synthetase. Biochemistry 33 (17): 4995-4999.

Keen, N.T., Taylor, O.C. (1975). Ozone injury in soybeans. Plant physiology 55: 731-733.

Kurek, I., Chang, T.K., Bertain, S.M., Madrigal, A., Liu, L., Lassner, M.W., Zhu, G. (2007). Enhanced, Thermostability of *Arabidopsis* Rubisco activase improves photosynthesis and growth rates under moderate heat stress. The Plant Cell DOI 10.1105/tpc.107.054171

Lehmann, M., Schwarzlander, M., Obata, T., Sirkantaramas, S., Burow, M., Olsen, C.E., Tohge, T., Fricker, M.D., Moller, B.L., Fernie, A.R., Sweetlove, L.J., Laxa, M. (2009). The metabolic response of *Arabidopsis* roots to oxidative stress is distinct from that of heterotrophic cells in culture and highlights a complex relationship between the levels of transcripts, metabolites, and flux. Molecular Plant 2: 390-406.

Loscos, J., Matamoros, M.A., Becana, M. (2008). Ascorbate and homoglutathione metabolism in common bean nodules under stress conditions and during natural senescence. Plant Physiology 146: 1282-1292.

Luwe, M.W. F., Heber, U. (1995). Ozone detoxification in the apoplasm and symplasm of spinach, broad bean and beech leaves at ambient and elevated concentrations of ozone in air. Planta 197: 448-455.

Macnicol, P.K. (1987). Homoglutathione and glutathione synthetases of legume seedlings: partial purification and substrate specificity. Plant Science 53: 229-235.

Matamoros, M.A., Moran, J.F., Iturbe-Ormaetxe, I., Rubio, M.C., Becana, M. (1999). Glutathione and Homoglutathione Synthesis in Legume Root Nodules. Plant Physiology 121: 879-888.

Moran, J.F., Iturbe-Ormaetze, I., Matamoros, M.A., Rubio, M.C., Clemente, M.R., Brewin, N.J., Becana, M. (2000). Glutathione and Homoglutathione Synthetases of Legume Nodules. Cloning, Expression, and Subcellular Localization. P l a n t Physiology 124: 1381-1392.

Motohashi K., Kondoh A., Stumpp M.T., Hisabori T. (2001). Comprehensive survey of proteins targeted by chloroplast thioredoxin. PNAS 98:11224-29

Mueller-Cajar, O., Whitney, S.M. (2008). Evolving improved *Synechococcus* Rubisco functional expression in *Escherichia coli.* Biochemistry Journal 414: 205-214.

Murphy, M.P., Holmgren, A., Larsson, N.-G., Halliwell, B., Chang, C.J., Kalyanaraman, B., Rhee, S.G., Thornalley, P.J., Partridge, L., Gems, D., Nystrom, T., Belousov, V., Schumacker, P.T., Winterbourn, C.C. (2011). Unraveling the biological roles of reactive oxygen species. Cell Metabolism 13: 361-366.

Noctor, G., Foyer, C.H. (1998). Ascorbate and Glutathione: Keeping active oxygen under control. Annual Review of Plant Physiology and Plant Molecular Biology. 49: 249-279.

Overmyer, K., Brosche, M., Pellinen, R., Kuittinen, T., Tuominen, H., Ahlfors, R., Keinanen, M., Scheel, D., Kangasjarvi, J. (2003). Ozone-induced programmed cell death in the *Arabidopsis* radical-incused cell death 1 mutant. Plant Physiology 137: 1092-1104.

Parry, M.A.J., Andralojc, P.J., Mitchell, R.A.C., Madgwick, P.J., Keys, A.J. (2003). Manipulation of Rubisco: the amount, activity, function and regulation. Journal of Experimental Botany 54: 1321-1333.

Parvez, M.M., Tomita-Yokotani, K., Fujii, Y., Konishi, T., Iwashina, T. (2004). Effects of quercetin and its seven derivatives on the growth of *Arabidopsis thaliana* and *Neurospora crassa*. Biochemical Systematics Ecology 32:631-635.

Passarinho, P.A., de Vries, S.C. (2002). Arabidopsis chitinases: a genomic survey. The Arabidopsis Book.

Pell, E.J., Schlagnhaufer, C.D., Arteca, R.N. (1997). Ozone-induced oxidative stress: mechanisms of action and reaction. Physiologia Planatarum 100: 264-273.

Polekhina, G., Board, P., Gala, R., Rossjohn, J., Parker, M. (1999). Molecular basis of glutathione synthetase deficiency and a rare gene permutation event. EMBO 18: 3204-3213.

Portis, A.R., Jr. (2003). Rubisco activase - Rubisco's catalytic chaperone. Photosynthesis Research 75: 11-27.

Rao, M.V., Davis, K.R. (2001). The physiology of ozone induced cell death. Planta 213: 682-690.

Rentel, M.C., Knight, M.R. (2004). Oxidative stress-induced calcium signaling in *Arabidopsis* Plant Physiology 135: 1471-1479.

Shen, J.B., Orozco, E.M., and Ogren, W.L. (1991). Expression of the two isoforms of spinach ribulose 1,5-bisphosphate carboxylase activase and essentiality of the conserved lysine in the consensus nucleotide-binding domain. Journal of Biological Chemistry 266: 8963–8968

Takahama, U., Hirota, S. (2000). Deglucosidation of quercetin glucosides to the aglycone and formation of antifungal agents by peroxidase-dependent oxidation of quercetin on browning of onion scales. Plant Cell Physiology 41: 1021-1029.

Tang, K.-H., Feng, X., Zhuang, W.-Q., Alvarez-Cohen, L., Blankenship, R.E., Tang, Y.J. (2010). Carbon flow of heliobacteria is related more to *Clostridia* than to the green sulfur bacteria. Journal of Biological Chemistry. 285: 35104-35112.

Vollsnes, A.V., Kruse, O.M.O.; Eriksen, A.B.; Oxaal, U.; Futsaether, C.M. (2010). *In vivo* root growth dynamics of ozone exposed *Trifolium subterraneum*. Environmental and Experimental Botany 69: 183-188.

Wojciechowski, M.F., Lavin, M., and Sanderson, M.J. (2004). A phylogeny of legumes (leguminosae) based on analysis of the plastid matK gene resolves many well- supported subclades within the family. American Journal of Botany 91: 1846–1862.

Yamaguchi, H., Kato, H., Hata, Y., Nishioka, T., Kimura, A., Oda, J., Katsube, Y. (1993). Three-dimensional Structure of the Glutathione Synthetase from *Escherichia coli* at 2.0 Å Resolution. Journal of Molecular Biology 229: 1083-1100.

Yano H., Wong J.H., Lee Y.M., Cho M.-J., Buchanan B.B. (2001). A strategy for the identification of proteins targeted by thioredoxin. PNAS 98:4794–99.

Zhang, N., Schürmann, P., and Portis, A.R., Jr. (2001). Characterization of the regulatory function of the 46-kDa isoform of Rubisco activase from *Arabidopsis*. Photosynthetic Research 68: 29–37.

Zhang, N., and Portis, A. R., Jr. (1999). Mechanism of light regulation of Rubisco: a specific role for the larger Rubisco activase isoform involving reductive activation by thioredoxin-f. PNAS USA 96: 9438–9443

APPENDIX I

STRUCTURAL AND FUNCTIONAL EVOLUTION OF ISOPROPYLMALATE DEHYDROGENASES IN LEUCINE AND GLUCOSINOLATE PATHWAYS OF *ARABIDOPSIS THALIANA*

Classification: Biological Sciences - Plant Biology, Biochemistry

Structural and Functional Evolution of Isopropylmalate Dehydrogenases in Leucine and Glucosinolate Pathways of *Arabidopsis thaliana*

Yan He[a], Ashley Galant[b], Qiuying Pang[a], Johanna M. Strul[a], Sherifat F. Balogun[b], Joseph M Jez[b], Sixue Chen[a*]

[a]Department of Biology, Genetics Institute, Plant Molecular and Cellular Biology Program, University of Florida, Gainesville, FL 32610, USA.

[2]Department of Biology, Washington University, St. Louis, MO 63130, USA.

*Corresponding author: Department of Biology, University of Florida, Gainesville, FL 32610, USA. Voice: (352) 273-8330; Facsimile: (352) 273-8070; E-mail: schen@ufl.edu

Running Title: Structure and function of Arabidopsis isopropylmalate dehydrogenases

Keywords: crystal structure, functional evolution, isopropylmalate dehydrogenases, leucine biosynthesis, glucosinolate biosynthesis, plant

Data deposition: The atomic coordinates and structure factors for AtIPMDH2 have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID code 3R8W).

Author contributions: JMJ and SC designed research; YH, AG, QY, JMS and SB performed research; YH, AG, JMJ and SC analyzed data; and YH, JMJ and SC wrote paper.

The authors declare no conflict of interest.

Abstract

The methionine chain-elongation pathway required for aliphatic glucosinolate biosynthesis in plants is thought to have evolved from leucine biosynthesis. In *Arabidopsis thaliana*, three 3-isopropylmalate dehydrogenases (AtIPMDHs) play key roles in either methionine chain-elongation for aliphatic glucosinolate biosynthesis (AtIPMDH1) or in leucine synthesis (AtIPMDH2 and AtIPMDH3). Here we elucidate the molecular basis underlying the metabolic specialization of these enzymes. The crystal structure of AtIPMDH2 was solved to provide the first detailed molecular architecture of a plant IPMDH. Modeling of 3-isopropylmalate binding in active site of the crystal structure and sequence comparisons of prokaryotic and eukaryotic IPMDH suggest that substitution of one active site residue in AtIPMDH1 may lead to altered metabolic function. Site-directed mutagenesis of Phe137 to a leucine in AtIPMDH1 (AtIPMDH1-F137L) reduced the enzyme activity toward 3-(2'-methylthio)ethylmalate by 200-fold, but enhanced catalytic efficiency with 3-isopropylmalate to levels observed with AtIPMDH2 and AtIPMDH3. Conversely, the AtIPMDH2-L134F and AtIPMDH3-L133F mutants enhanced catalytic efficiency with 3-(2'-methylthio)ethylmalate ~100-fold and reduced activity for 3-isopropylmalate. Furthermore, the altered *in vivo* glucosinolate profile of an Arabidopsis *ipmdh1* T-DNA knockout mutant could be restored to wild-type levels by constructs expressing AtIPMDH1, AtIPMDH2-L134F, or AtIPMDH3-L133F, but not by AtIPMDH1-F137L. These results demonstrate that a single amino acid substitution results in functional divergence of IPMDH *in planta* to affect substrate specificity and contribute to the evolution of specialized glucosinolate biosynthesis from ancestral leucine biosynthesis.

\body

Introduction

To compensate for their sessile nature, plants evolved mechanisms to cope with rapid environmental changes and challenges (1). The production of specialized metabolites is one of the important mechanisms for the survival and fitness of plants (2). The molecular diversity of these specialized compounds arises from differential modification of common backbone structures, which necessitates the evolution of homologous enzymes with varied specificities (1). In plants, glucosinolates constitute a diverse group of sulfur-containing specialized metabolites (3-4). Biosynthesis of methionine-derived glucosinolates is initiated by the sequential addition of methylene groups to produce chain-elongated methionine derivatives via an iterative three-step chain-elongation process that mimics the chemistry of leucine synthesis (Fig. 1A).

To date, all the genes involved in the methionine chain-elongation process have been identified and characterized in *Arabidopsis thaliana* (5-14). The different enzymes of the methionine chain-elongation pathway for glucosinolate synthesis appear to have evolved from leucine synthesis either by gene duplication and neo-functionalization of one of the duplicated genes or by sub-functionalization via differential temporal and spatial expression of gene copies (14-15). For example, four genes in Arabidopsis encode isopropylmalate synthases (IPMS) with two (IPMS1 and IPMS2) serving in leucine biosynthesis and the other two genes encoding methylthioalkylmalate (MAM) synthases (MAM1 and MAM3) catalyzing the committed step in methionine chain-elongation (5-6, 16). A recent study showed that loss of a C-terminal regulatory domain and a few amino acid exchanges can covert IPMS into MAM (14). Specialization of the Arabidopsis isopropylmalate isomerases (IPMI) for different metabolisms occurs by changes in the oligomeric composition of these enzymes. IPMI are heterodimeric

enzymes consisting of a large subunit encoded by a single gene and a small subunit encoded by one of three genes (8-9, 12). Metabolic profiling of the large subunit mutant revealed accumulation of intermediates in both the leucine pathway and the methionine chain-elongation pathway, demonstrating the dual function of this subunit in both leucine and glucosinolate biosynthesis (10). In contrast, the small subunits are specialized to either leucine biosynthesis or methionine chain-elongation (2, 10, 12). Furthermore, among the six branch-chain aminotransferases (BCATs) in Arabidopsis, BCAT4 in the cytosol is specifically involved in glucosinolate biosynthesis, whereas BCAT3 in the plastids functions in both amino acid and glucosinolate biosynthesis (7, 9). However, the changes that tailor BCAT activity are unclear.

Previously, we showed that *Arabidopsis thaliana* isopropylmalate dehydrogenase 1 (AtIPMDH1) catalyzes the oxidative decarboxylation step in the methionine chain-elongation of glucosinolate biosynthesis and that AtIPMDH2 and AtIPMDH3 are primarily involved in leucine biosynthesis (Fig. 1B) (11, 13). These studies highlight the functional specialization of these isoforms, but do not reveal how these activities evolved. Here we examine the molecular basis for the functional evolution of the IPMDH family in Arabidopsis. The crystal structure of AtIPMDH2, the first determined for a plant IPMDH, reveals an active site structure similar to that of the bacterial enzymes and provides a template for modeling substrate binding in the active site. Analysis of the AtIPMDH2 structure, sequence comparisons, and site-directed mutagenesis demonstrate that a single residue difference in the active site drastically alters substrate specificity of the AtIPMDH isoforms both *in vitro* and *in vivo*. This work demonstrates the basis for functional divergence of an AtIPMDH isoform for glucosinolate biosynthesis from those involved in leucine biosynthesis.

Results

Sub-functionalization and AtIPMDH Metabolic Specialization. The three *IPMDH* genes in Arabidopsis have overlapping, yet distinct expression patterns. *AtIPMDH1* (At5g14200) is highly expressed in leaves and roots; *AtIPMDH2* (At1g80560) is weakly expressed throughout the plant; and *AtIPMDH3* (At1g31180) is constitutively expressed at high levels in all tissues (11, 13, 17). To test the possible contribution of differential expression to the specialization of *AtIPMDHs*, each gene was placed under control of the native *AtIPMDH1* promoter and then transformed into an *atipmdh1* mutant line (11). As shown in Fig. 2, the altered glucosinolate profile of the *atipmdh1* mutant could only be rescued by expression of *AtIPMDH1*. The results indicate that subfunctionalization may not be the cause of AtIPMDH specialization.

Structure of AtIPMDH2. To determine the molecular architecture of a plant IPMDH, the 2.25 Å resolution x-ray crystal structure of AtIPMDH2 was solved by molecular replacement (Table S1). AtIPMDH2 is a dimeric protein with each monomer consisting of two domains (Fig. 3A). Domain 1 contains seven α-helices (α1-4 and α9-11) and five β-strands (β1-3 and β11-12), along with the N- and C-termini. Four α-helices (α5-8) and seven β-strands (β4-10) comprise domain 2. Between the two domains, β4 and β5 form the interdomain region. The second domain also serves as the dimerization interface with β6 and β7 of each monomer as part of an inter-subunit β-sheet and α7 and α8 of each monomer forming a four-helix bundle at the dimer interface. The overall structure of AtIPMDH2 is similar to those of the IPMDH from various bacteria, including *Salmonella typhimurium* and *Thermus thermophilus* (18-20), with a root mean square deviation of 1.3-1.7 $Å^2$ over ~350 residues. Because the plant and bacterial IPMDH share ~50% sequence identity, conservation of key residues defines the active site region situated in a cleft between the two domains of each monomer (Fig. 3A).

The active site (Fig. 3B) is roughly delineated by α8 at the bottom and with α4 of one monomer and α7 of the adjacent monomer forming opposite sides of the site. Within the active

site, all of the residues previously identified in structures of bacterial IPMDH in complex with isopropylmalate and Mg^{2+} are also conserved in AtIPMDH2 (18, 20). Because efforts to obtain a structure of AtIPMDH2 in complex with ligands did not yield crystals, 3-isopropylmalate and Mg^{2+} were modeled into the plant enzyme based on the positions of these ligands observed in the bacterial structures (Fig. 3B and C) (18, 20). This comparison shows that Asp264* (asterisk denotes adjacent monomer), Asp288, and Asp292 are positioned to interact with a catalytically essential divalent metal (i.e., Mg^{2+} or Mn^{2+}) and that a trio of arginines (Arg136, Arg146, and Arg174) is poised to form charge-charge interactions with the carboxylate groups of the substrate. Residues corresponding to Leu132, Leu133, Tyr181, Lys232*, Asn234*, and Val235* form a largely hydrophobic region around the isopropyl group of the substrate.

Although all of these amino acids are invariant in the bacterial and plant IPMDH involved in leucine biosynthesis, the side-chain corresponding to Leu133 is replaced with a phenylalanine in AtIPMDH1 (Fig. 3D), which is the isoform previously shown to be primarily involved in glucosinolate synthesis in Arabidopsis (11). Mechanistically, the conversion of 3-isopropylmalate to 4-methyl-2-oxovalerate in leucine synthesis and the conversion of 3-malate derivatives (e.g., 3-(2'-methylthio)ethylmalate) to 2-oxo acids (e.g., 5-methylthio-2-oxopentoate) in glucosinolate synthesis likely use a common metal-dependent reaction (Fig. S1); however, different substrate side-chains of 3-malate derivatives (Fig. 1B) must fit in the plant IPMDH active site for production of aliphatic glucosinolates with six different chain lengths (C3-C8). Thus, we hypothesize that this single amino acid exchange from the leucine found in AtIPMDH2 and AtIPMDH3 to the phenylalanine in the active site of AtIPMDH1 may contribute to the functional divergence of this isoform for glucosinolate biosynthesis.

Biochemical Analysis of Wild-Type and Mutant AtIPMDH. Previous studies on the AtIPMDH demonstrate that each isoform accepts 3-isopropylmalate as a substrate (11, 13), but

a kinetic comparison with a glucosinolate pathway substrate has not been reported. Using both 3-isopropylmalate and 3-(2'-methylthio)ethylmalate, the steady-state kinetic parameters for each AtIPMDH were determined (Table 1 and Fig. 4). Comparison of the catalytic efficiencies shows that AtIPMDH2 and AtIPMDH3 favor 3-isopropylmalate over 3-(2'-methylthio)ethylmalate by 14,900- and 29,600-fold, respectively. Moreover, these isoforms were ~20-fold more active with the leucine biosynthesis substrate than AtIPMDH1. In comparison, AtIPMDH1 accepts both substrates with comparable k_{cat}/K_m values, but was ~500-fold more efficient with the glucosinolate substrate than the other two isoforms. These catalytic efficiencies agree with the observed *in vivo* roles of the AtIPMDH isoforms in glucosinolate and leucine synthesis pathways (11-13).

To investigate the significance of the active site difference in the AtIPMDH, a series of point mutants (AtIPMDH1-F137L, AtIPMDH2-L133F, and AtIPMDH3-L134F) were generated. Kinetic analysis of these mutants demonstrates the critical role of this active site change in determining substrate specificity (Table 1 and Fig. 4). In AtIPMDH1, substitution of Phe137 with a leucine reduced the k_{cat}/K_m of the mutant for 3-(2'-methylthio)ethylmalate to values comparable to those observed for AtIPMDH2 and AtIPMDH3. This was also accompanied by improved catalytic efficiency with 3-isopropylmalate, as the AtIPMDH1-F137L mutant was only 2- to 3-fold less efficient with this substrate than AtIPMDH2 and AtIPMDH3. The complementary mutation in either AtIPMDH2 (L133F) or AtIPMDH3 (L134F) yields mutant enzymes that were ~30-fold less active with 3-isopropylmalate than the corresponding wild-type proteins, but still comparable to wild-type AtIPMDH1. Moreover, AtIPMDH2-L133F and AtIPMDH3-L134F displayed nearly a 100-fold improvement in activity with 3-(2'-methylthio)ethylmalate as a substrate to k_{cat}/K_m values that were 4- and 7-fold less than those observed with AtIPMDH1. These results demonstrate the critical role of the residue at position 133 (AtIPMDH2 numbering)

in the evolution of AtIPMDH1 for the methionine chain-elongation reactions of glucosinolate biosynthesis.

***In vivo* Analysis of AtIPMDH Mutant Function.** To test whether the amino acid substitution that occurred in AtIPMDH1 contributes to its specific function *in vivo*, *atipmdh1* mutant plants were transformed with each of the mutant *AtIPMDH* genes driven by the *AtIPMDH1* promoter. After isolation of homozygous lines, the glucosinolate profile in each mutant was examined. In comparison to the results shown in Fig. 1, the pronounced glucosinolate phenotype in the *atipmdh1* mutant could not be rescued by AtIPMDH1-F133L (Fig. 5), indicating that the active site substitution impaired AtIPMDH1 function for glucosinolate synthesis *in vivo*. In contrast, the glucosinolate phenotype could be restored to the wild-type profile by expression of either AtIPMDH2-L133F or AtIPMDH3-L134F (Fig. 5). The *in planta* findings corroborate the conclusion drawn from the biochemical analysis of recombinant proteins and provide evidence for the evolution of AtIPMDH1 by gene duplication and a single critical amino acid substitution.

Discussion

The evolution of specialized metabolism from primary metabolism is a common theme across biochemical pathways in plants (and microbes). Here we explored the molecular basis underlying the divergence of biological function in the IPMDHs of Arabidopsis. Although all three AtIPMDH accept 3-isopropylmalate, AtIPMDH1 is less efficient than the other isoforms (11, 13). Previous work also showed that knockout mutants of *AtIPMDH1* result in reduced levels of C4 to C8 aliphatic glucosinolates (11). In contrast, knockout mutations of the other isoforms did not alter glucosinolate levels but reduced leucine content (11, 13). Interestingly, a double mutation

of *AtIPMDH2* and *AtIPMDH3* in Arabidopsis plants led to defects in pollen and embryo sac development, suggesting that leucine synthesis is essential for gametophyte formation. Using a combination of structural and functional analysis, this work demonstrates that a single amino acid change in the AtIPMDH active site leads to functional specialization of these enzymes in leucine synthesis (primary metabolism) and aliphatic glucosinolate synthesis (specialized metabolism).

Possible sub- and neo-functionalization processes can drive the evolution of specialized metabolism (14-15). To evaluate if altered expression of *AtIPMDH* isoforms underlies functional specialization, each isoform gene was expressed under control of the *AtIPMDH1* promoter in an *atimpdh1* mutant background (Fig. 2). Because the glucosinolate profile in the mutant was rescued only by expression of *AtIPMDH1*, it appears that neo-functionalization, involving gene duplication and subsequent mutation to a new function, may be the underlying evolutionary mechanism.

The three-dimensional structure of AtIPMDH2 (Fig. 3) and functional analysis (Table 1 and Fig. 4) of the AtIPMDH provides insight on the specific changes required to alter the metabolic roles of these enzymes. A common chemical transformation is required to convert 3-isopropylmalate to 4-methyl-2-oxovalerate in leucine synthesis and 3-malate derivatives to 2-oxo acids in glucosinolate synthesis (Fig. 1). The AtIPMDH active site includes invariant residues for binding of either Mg^{2+} or Mn^{2+} (Asp288, Asp292, Asp264*) and for charge-charge interactions with the substrate carboxylate groups (Arg136, Arg146, and Arg174). Likewise, Tyr181 and Lys232*, which are proposed to perform general acid-base chemistry in the reaction mechanism (21), are conserved. For both 3-isopropylmalate (leucine synthesis) and 3-malate derivatives (glucosinolate synthesis), the overall reaction (Fig. S1) involves oxidation of the alcohol by deprotonation and hydride transfer to NAD^+. This is followed by spontaneous

decarboxylation, stabilization of the resulting enolate by the metal ion, and protonation to yield the final product.

Leucine and glucosinolate synthesis requires the same chemistry, but the AtIPMDH active site must accommodate reactants with different side-chains (i.e., isopropyl versus elongated methionine side-chain groups). The AtIPMDH2 structure and sequence analysis reveals a single amino acid difference of a leucine (AtIPMDH2 and AtIPMDH3) versus a phenylalanine (AtIMPDH1) in the active site. This difference occurs in the set of residues proposed to form the substrate interaction surface in the bacterial and plant IPMDH (18-20). Both *in vitro* and *in vivo* functional analysis of AtIPMDH1-F137L, AtIPMDH2-L133F, and AtIPMDH3-L134F demonstrates that switching this amino acid in each isoform is sufficient to interconvert catalytic efficiency (Table 1 and Fig. 4) and to change the aliphatic glucosinolate profiles in transgenic plants (Fig. 5). These results suggest that gene duplication of AtIPMDH followed by mutation of one active site residue in AtIPMDH1 leads to its specialized role for glucosinolate synthesis in Arabidopsis.

The structure-function analysis of the AtIPMDH provides insight on the molecular basis for altered function, but it is unclear how the leucine to phenylalanine mutation allows AtIPMDH1 to accommodate the growing methionine chain in subsequent iterations of the glucosinolate synthesis reactions (Fig. 1A). Multiple structures of IPMDH from bacteria indicate that the structural features around the active site are flexible and that active site dynamics likely plays a potential role in substrate recognition and catalysis (22). Moreover, the effect of the longer side-chain on the kinetics of the various glucosinolate biosynthesis pathway enzymes (i.e., BCAT, MAM, IPMI, and IPMDH) has not been explored. In Arabidopsis, multiple lines of evidence strongly support the evolution of methionine chain-elongation process of glucosinolate biosynthesis from leucine biosynthesis (5-8, 11); however, the molecular underpinnings for this

evolution are only beginning to be understood. For example, the substrate specialization of the heterodimeric IPMI is determined by which small subunit associates with the large subunit (2, 8, 10, 12). More recently, the changes needed to convert IPMS from leucine synthesis into a MAM was demonstrated to involve the loss of a C-terminal regulatory domain responsible for feedback inhibition by leucine and a series of amino acid mutations (14). In contrast to large remodeling of protein structure in IPMS and MAM, the substrate specificity of IPMDH requires one amino acid difference.

Interactions between Arabidopsis and its environment may have driven the co-evolution of the pathways needed to synthesize the core glucosinolate structure and the elongation of the methionine side-chain. The biosynthesis of the glucosinolates has been suggested to have evolved from the prevalent system of cyanogenic glucoside biosynthesis (23-24). Evidence for this includes the wide distribution of cyanogenic glucosides in plants and arthropods, and the conservation of cytochrome P450s in the biosynthesis of glucosinolates and cyanogenic glucosides. In addition, metabolic engineering using cytochromes P450 involved in cyanogenic glycoside biosynthesis allows for the generation of acyanogenic plants that also display altered glucosinolate profiles (24-25). It is evident that when environmental challenges such as insect herbivores present themselves, specialization of enzymes from different pathways contributes to the evolution of methionine-derived glucosinolates for plant survival.

In summary, we have determined the molecular changes responsible for the recruitment of AtIPMDH from leucine biosynthesis for the specialized synthesis of glucosinolates. Future studies need to explore protein level changes in other glucosinolate enzymes to understand how the entire glucosinolate pathway evolved.

Materials and Methods

Plants and Growth. Seeds of *Arabidopsis thaliana* ecotype Columbia (Col-0) and SALK mutant *atipmdh1* (Salk_063423C) were obtained from the Arabidopsis Biological Resource Center (ABRC). Seed germination and plant growth conditions were as previously described (11, 13).

Plasmid Construction and Plant Transformation. Oligonucleotides used in this study are listed in Table S2. The full-length coding sequences of *AtIPMDH1*, *AtIPMDH2* and *AtIPMDH3* were amplified using the Platinum Pfx DNA Polymerase (Invitrogen) with appropriate primer pairs. PCR products were firstly cloned into pSC-B-amp/kan vector using StrataClone Blunt PCR Cloning Kit (StrataClone), and then sequenced. Correct fragments were subcloned into the *AtIPMDH1*pro::GUS vector (11) to generate constructs for each isoform and/or mutant under control of the *AtIPMDH1* promoter. The resulting constructs were introduced into *Agrobacterium tumefaciens* strain C58C1 followed by transformation into *atipmdh1* plants. Transgenic plants were selected for hygromycin resistance and homozygous plants used for subsequent analysis.

Glucosinolate Analysis. Rosette leaves of 4-week-old plants and mature seeds were used for glucosinolate analysis. Glucosinolates were analyzed using HPLC–mass spectrometry as previously described (11, 13).

Protein Expression, Purification, Assays, Crystallization, and Structure Determination. Expression and purification of wild-type and mutant AtIPMDH as histidine-tagged proteins for functional analysis was performed as previously described (11). IPMDH assay conditions using either 3-isopropylmalate or 3-(2'-methylthio)ethylmalate as a substrate and the analysis of steady-state kinetic parameters were as previously described (11). For crystallization of AtIPMDH2, the histidine-tag was removed by thrombin digestion and the

protein further purified using size-exclusion chromatography (27). Crystals of AtIPMDH2 were obtained in 5 μL hanging drops of a 1:1 mixture of protein and crystallization buffer (0.16 M ammonium sulfate, 0.08 M sodium acetate trihydrate, 20% PEG 4000, 20% glycerol) at 4 °C over a 0.7 mL reservoir. Data collection (100 K) was performed at the beamline 19-ID at the Advanced Photon Source Argonne National Laboratory. Diffraction data was integrated and reduced using HKL3000 (28). The structure of AtIPMDH2 was solved by molecular replacement performed with PHASER 29) using the structure of IPMDH from *Salmonella typhimurium* (19) as a search model. Model building was performed in COOT (30) and all refinements were performed with Phenix (31).

Site-Directed Mutagenesis. Site-directed mutagenesis was performed using the QuikChange PCR method (Stratagene). Bacterial expression vectors for each *AtIPMDH* (11, 13) were used as templates with specific oligonucleotide pairs (Table S2). Protein expression, purification, and assays were performed as described above for wild-type protein.

Acknowledgements

We thank Ms. Bing Chen for technical support and the staff at UF ICBR for DNA and protein sequencing. This work was supported by the National Science Foundation (MCB-0845162 to S.C. and MCB-0824492 to J.M.J.). A.G. was supported by an American Society of Plant Biologists Pioneer Hi-Bred Graduate Research Fellowship and S.B. was a recipient of an NIH-MARC uSTAR Undergraduate Fellowship. Portions of this research were carried out at the Argonne National Laboratory Structural Biology Center at the Advanced Photon Source, a national user facility operated by the University of Chicago for the Department of Energy Office of Biological and Environmental Research (DE-AC02-06CH11357).

References

1. Kliebenstein DJ, Lambrix VM, Reichelt M, Gershenzon J, Mitchell-Olds T (2001) Gene duplication in the diversification of secondary metabolism: tandem 2-oxoglutarate-dependent dioxygenases control glucosinolate biosynthesis in Arabidopsis. *Plant Cell* 13: 681-693.
2. Sawada Y, Kuwahara A, Nagano M, Narisawa T, Sakata A, Saito K, Hirai MY (2009) Omics-based approaches to methionine side chain elongation in Arabidopsis: characterization of the genes encoding methylthioalkylmalate isomerase and methylthioalkylmalate dehydrogenase. *Plant Cell Physiol* 50: 1181-1190.
3. Halkier BA, Gershenzon J (2006) Biology and biochemistry of glucosinolates. *Ann Rev Plant Biol* 57: 303-333.
4. Yan XF, Chen S (2007) Regulation of plant glucosinolate metabolism. *Planta* 226: 1343-1352.
5. Field B, Cardon G, Traka M, Botterman J, Vancanneyt G, Mithen R (2004) Glucosinolate and amino acid biosynthesis in Arabidopsis. *Plant Physiol* 135: 828-839.
6. Textor S, Bartram S, Kroymann J, Falk KL, Hick A, Pickett JA, Gershenzon J (2004) Biosynthesis of methionine-derived glucosinolates in *Arabidopsis thaliana*: recombinant expression and characterization of methylthioalkylmalate synthase, the condensing enzyme of the chain-elongation cycle. *Planta* 218: 1026-1035.
7. Schuster J, Knill T, Reichelt M, Gershenzon J, Binder S (2006) Branched-chain aminotransferase4 is part of the chain elongation pathway in the biosynthesis of methionine-derived glucosinolates in Arabidopsis. *Plant Cell* 18: 2664-2679.
8. Binder S, Knill T, Schuster J (2007) Branched-chain amino acid metabolism in higher plants. *Physiol Plant* 129: 68-78.
9. Knill T, Schuster J, Reichelt M, Gershenzon J, Binder S (2008) Arabidopsis branched-chain aminotransferase 3 functions in both amino acid and glucosinolate biosynthesis. *Plant Physiol* 146: 1028-1039.

10. Knill T, Reichelt M, Paetz C, Gershenzon J, Binder S (2009) *Arabidopsis thaliana* encodes a bacterial-type heterodimeric isopropylmalate isomerase involved in both Leu biosynthesis and the Met chain elongation pathway of glucosinolate formation. *Plant Mol Bio*l 71: 227-239.

11. He Y, Mawhinney TP, Preuss ML, Schroeder AC, Chen B, Abraham L, Jez JM, Chen S (2009) A redox-active isopropylmalate dehydrogenase functions in the biosynthesis of glucosinolates and leucine in Arabidopsis. *Plant J* 60: 679-690.

12. He Y, Chen B, Pang Q, Strul JM, Chen S (2010) Functional specification of Arabidopsis isopropylmalate isomerases in glucosinolate and leucine biosynthesis. *Plant Cell Physiol* 51: 1480-1487.

13. He Y, Chen L, Zhou Y, Mawhinney TP, Chen B, Kang BH, Hauser BA, Chen S (2011) Functional characterization of *Arabidopsis thaliana* isopropylmalate dehydrogenases reveals their important roles in gametophyte development. *New Phytol* 189: 160-175.

14. deKraker JW, Gershenzon J (2011) From amino acid to glucosinolate biosynthesis: protein sequence changes in the evolution of methylthioalkylmalate synthase in Arabidopsis. *Plant Cell* 23: 38-53.

15. Sankoff D (2001) Gene and genome duplication. *Curr Opin Genet Dev* 11: 681-684.

16. Textor S, de Kraker JW, Hause B, Gershenzon J, Tokuhisa JG (2007) MAM3 catalyzes the formation of all aliphatic glucosinolate chain lengths in Arabidopsis. *Plant Physiol* 144: 60-71.

17. Nozawa A, Takano J, Miwa K, Nakagawa Y, Fujiwara T (2005) Cloning of cDNAs encoding isopropylmalate dehydrogenase from *Arabidopsis thaliana* and accumulation patterns of their transcripts. *Biosci Biotech Biochem* 69: 806-810.

18. Kadono S, Sakurai M, Moriyama H, Sato M, Hayashi Y, Oshima T, Tanaka N (1995) Ligand-induced changes in the conformation of 3-isopropylmalate dehydrogenase from *Thermus thermophilus*. *J Biochem* 118: 745-752.

19. Wallon G, Kryger G, Lovett ST, Oshima T, Ringe D, Petsko GA (1997) Crystal structures of *Escherichia coli* and *Salmonella typhimurium* 3-isopropylmalate dehydrogenase and comparison with their thermophilic counterpart from *Thermus thermophilus*. *J Mol Biol* 266: 1016-1031.

20. Imada K, Inagaki K, Matsunami H, Kawaguchi H, Tanaka H, Tanaka N, Namba K (1998) Structure of 3-isopropylmalate dehydrogenase in complex with 3-isopropylmalate at 2.0 A resolution: the role of Glu88 in the unique substrate-recognition mechanism. *Structure* 6: 971-982.

21. Aktas DF, Cook PF (2009) A lysine-tyrosine pair carries out acid-base chemistry in the metal ion-dependent pyridine dinucleotide-linked β-hydoxyacid oxidative decarboxylases. *Biochemistry* 48, 3565-3577.

22. Singh RK, Kefala G, Janowski R, Mueller-Dieckmann C, von Kries JP, Weiss MS (2005) The high-resolution structure of LeuB (Rv2995c) from *Mycobacterium tuberculosis*. *J Mol Biol* 346: 1-11.

23. Chen S, Andereason E (2001) Update on glucosinolate metabolism and transport. *Plant Physiol Biochem* 39: 743-758.

24. Bak S, Paquette SM, Morant M, Morant AV, Saito S, Bjarnholt N, Zagrobelny M, Jørgensen K, Osmani S, Simonsen HT, Perez RS, van Heeswijck TB, Jorgensen B, Moller BL (2006) Cyanogenic glycosides: a case study for evolution and application of cytochromes P450. *Phytochem Rev* 5: 309-329.

25. Chen S, Glawishnig E, Jorgensen K, Naur P, Jorgensen B, Olsen CE, Rasmussen H, Pickett JA, Halkier BA (2003) CYP79F1 and CYP79F2 have distinct functions in the biosynthesis of aliphatic glucosinolates in Arabidopsis. *Plant J* 33: 923-937.

26. Alvarez S, He Y, Chen S (2008) Comparative investigations of glucosinolate-myrosinase system in Arabidopsis suspension cells and hypocotyls. *Plant Cell Physiol* 49: 324-333.

27. Galant A, Arkus KAJ, Zubieta C, Cahoon RE, Jez JM (2009) Structural basis for evolution of product diversity in soybean glutathione synthesis. *Plant Cell* 21: 3450-3458.

28. Otwinowski Z, Minor W (1997) in *Methods in Enzymology, Volume 276: Macromolecular Crystallography, part A*, eds Carter CW, Sweet RM (Academic Press, NY), pp 307-326.

29. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. *J Appl Cryst* 40: 658-674.

30. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D* 66: 486-501.

31. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D* 66: 213-221.

Figure Legends

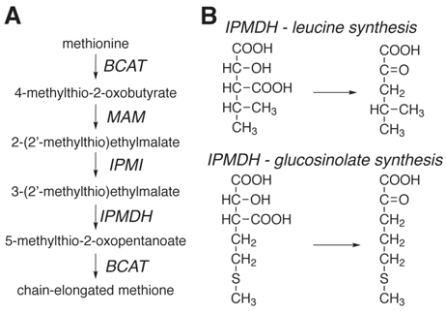
Fig. 1. (**A**) Overview of the methionine chain-elongation pathway of aliphatic glucosinolate biosynthesis in *Arabidopsis thaliana*. Note that a chain-elongated 2-oxo acid can serve as a substrate for MAM1 and MAM3 in subsequent rounds through the pathway to yield longer side-chain products. (**B**) IPMDHs catalyze the conversion of 3-isopropylmalate to 4-methyl-2-oxovalerate in leucine synthesis and the conversion of 3-(2'-methylthio)ethylmalate to 5-methylthio-2-oxopentoate in glucosinolate synthesis.

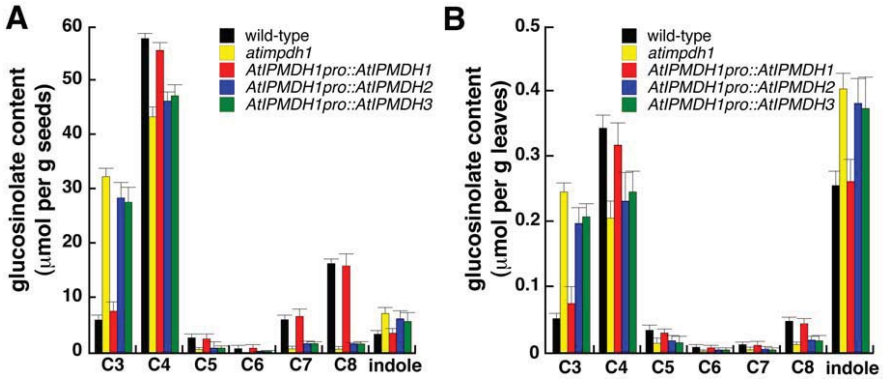
Fig. 2. Glucosinolate profiles in seeds (**A**) and leaves (**B**) from wild-type, *atipmdh1* mutant, and transgenic plants harboring each *AtIPMDH* driven by the *AtIPMDH1* promoter. Levels of aliphatic glucosinolates with varied methylene chain length (C3-C8) are shown. All indole glucosinolates are combined into a single group. Data are mean ± standard deviation (n=3).

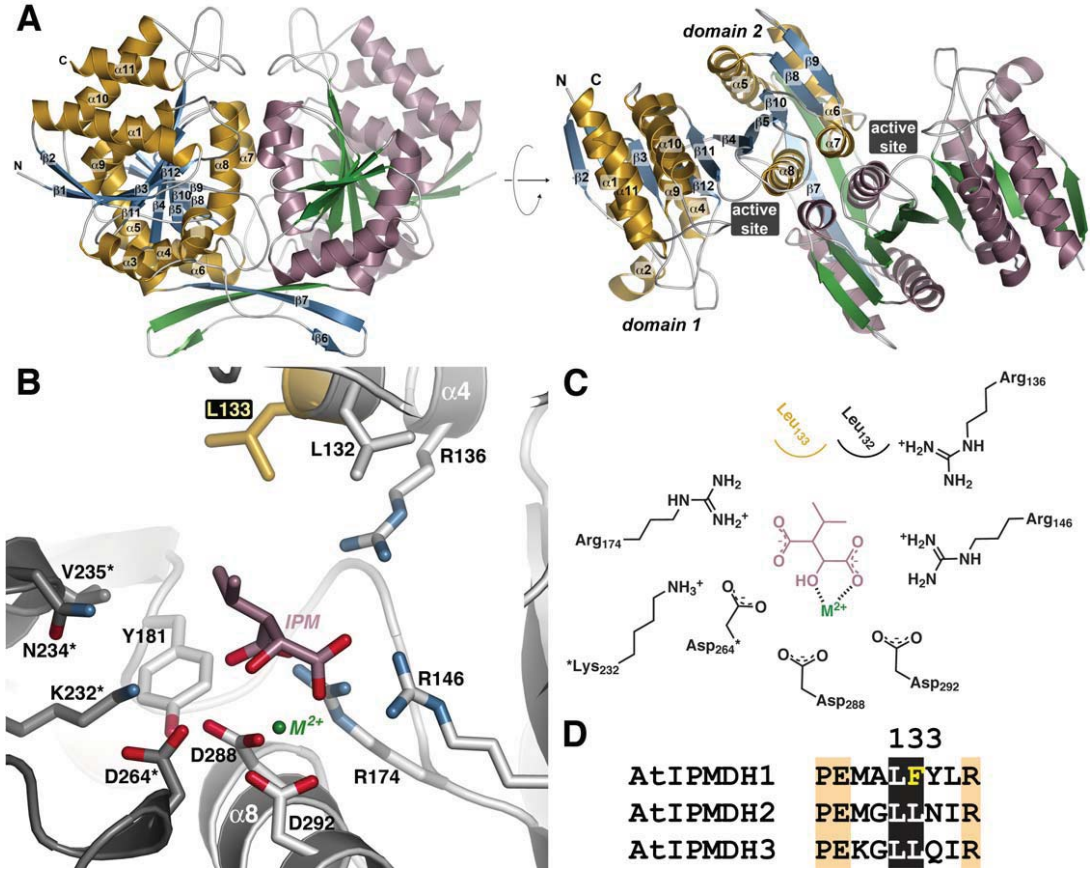
Fig. 3. Structure of AtIPMDH2. **(A**) Ribbon diagrams of the AtIPMDH2 dimer. Monomer A is shown with gold α-helices and blue β-strands and monomer B is drawn with rose α-helices and green β-strands. Secondary structure features are labeled on the A monomer. The left view shows the dimer down the 2-fold axis. The right view is rotated 90° to show the two domains of each monomer. The position of the active site cleft is indicated. (**B**) Active site view and model of 3-isopropylmalate (IPM) and divalent metal (M^{2+}). Side-chains of active site residues are shown with those from the adjacent monomer (grey) indicated by an asterisk. The positions of the substrate and metal are modeled based on the bacterial structures (18-20). The active site difference among the AtIPMDH isoforms is highlighted in gold. (**C**) Schematic of the active site model. (**D**) Sequence comparison of the region including residue 133 (AtIPMDH2 numbering).

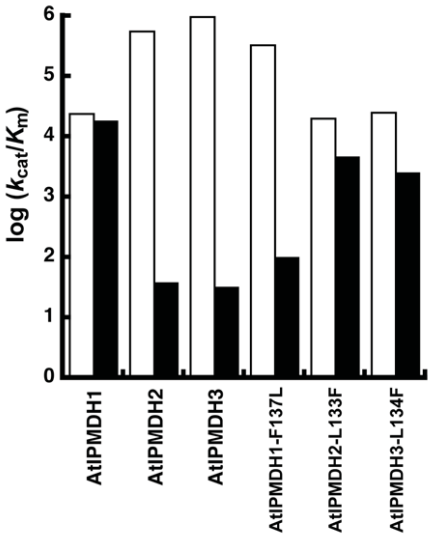
Fig. 4. Comparison of the catalytic efficiencies (k_{cat}/K_m) of wild-type and mutant AtIPMDH using 3-isopropylmalate (white bars) and 3-(2'-methylthio)ethylmalate (black bars).

Fig. 5. Glucosinolate profiles in seeds (**A**) and leaves (**B**) of wild-type, *atipmdh1* mutant, and transgenic plants expressing AtIPMDH1-F137L, AtIPMDH2-L133F and AtIPMDH3-L134F driven by *AtIPMDH1* native promoter. Levels of aliphatic glucosinolates with varied methylene chain length (C3-C8) are shown. All indole glucosinolates are combined into a single group. Data are mean ± standard deviation (n=3).


A
methionine
BCAT
4-methylthio-2-oxobutyrate
MAM
2-(2'-methylthio)ethylmalate
IPMI
3-(2'-methylthio)ethylmalate
IPMDH
5-methylthio-2-oxopentanoate
BCAT
chain-elongated methione
B
IPMDH - leucine synthesis
IPMDH - glucosinolate synthesis


A
glucosinolate content
(µmol per g seeds)
60
50
40
30
20
10
0
wild-type
atimpdh1
AtIPMDH1pro::AtIPMDH1
AtIPMDH1pro::AtIPMDH2
AtIPMDH1pro::AtIPMDH3
C3
C4
C5
C6
C7
C8
indole
B
glucosinolate content
(µmol per g leaves)
0.5
0.4
0.3
0.2
0.1
0
wild-type
atimpdh1
AtIPMDH1pro::AtIPMDH1
AtIPMDH1pro::AtIPMDH2
AtIPMDH1pro::AtIPMDH3
C3
C4
C5
C6
C7
C8
indole


A
domain 2
domain 1
active site
active site
B
L133
L132
R136
α4
V235*
N234*
Y181
IPM
K232*
R146
M2+
D288
D264*
R174
α8
D292
C
Leu133
Leu132
Arg136
Arg174
Arg146
M2+
*Lys232
Asp264*
Asp288
Asp292
D
133
AtIPMDH1 PEMALFYLR
AtIPMDH2 PEMGLLNIR
AtIPMDH3 PEKGLLQIR


log (k_{cat}/K_m)
0
1
2
3
4
5
6
AtIPMDH1
AtIPMDH2
AtIPMDH3
AtIPMDH1-F137L
AtIPMDH2-L133F
AtIPMDH3-L134F

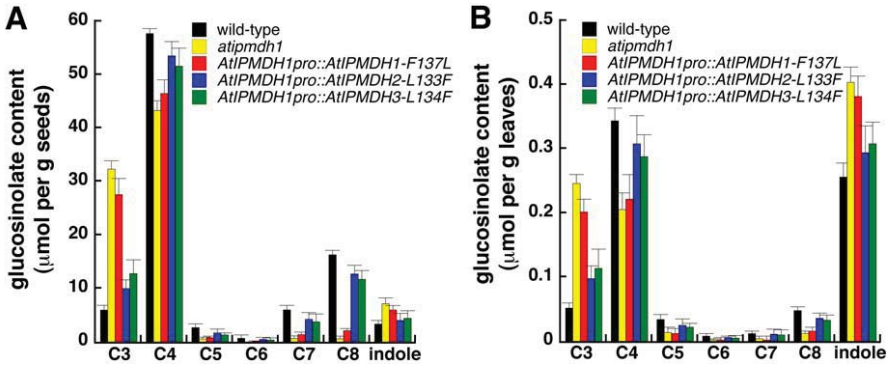

A
glucosinolate content
(µmol per g seeds)
60
50
40
30
20
10
0
wild-type
atipmdh1
AtIPMDH1pro::AtIPMDH1-F137L
AtIPMDH1pro::AtIPMDH2-L133F
AtIPMDH1pro::AtIPMDH3-L134F
C3
C4
C5
C6
C7
C8
indole
B
glucosinolate content
(µmol per g leaves)
0.5
0.4
0.3
0.2
0.1
0
wild-type
atipmdh1
AtIPMDH1pro::AtIPMDH1-F137L
AtIPMDH1pro::AtIPMDH2-L133F
AtIPMDH1pro::AtIPMDH3-L134F
C3
C4
C5
C6
C7
C8
indole

Table 1. Kinetic parameters of wild-type and mutant AtIPMDHs.

	3-isopropylmalate			*3-(2'-methylthio)malate*		
	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
AtIPMDH1	37 ± 4	25.2 ± 2.3	24,471	51 ± 5	45.3 ± 3.6	18,763
AtIPMDH1-F137L	230 ± 14	11.4 ± 1.7	336,257	2.0 ± 0.2	323 ± 21	103
AtIPMDH2	373 ± 33	10.9 ± 1.3	570,336	1.0 ± 0.2	435 ± 32	38.3
AtIPMDH2-L133F	37 ± 5	30.3 ± 2.5	20,352	22 ± 1	77.0 ± 8.5	4,761
AtIPMDH3	543 ± 36	9.2 ± 1.4	983,696	1.0 ± 0.1	502 ± 35	33.2
AtIPMDH3-L134F	44 ± 5	28.5 ± 1.5	25,731	16 ± 1	103 ± 10	2,589

Data are means ± standard error (n = 3).