Homoglutathione Synthetase And The Plant Thiol-Redox Proteome

Ashley Galant
Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/etd

Recommended Citation
Galant, Ashley, "Homoglutathione Synthetase And The Plant Thiol-Redox Proteome" (2011). All Theses and Dissertations (ETDs). 576.
https://openscholarship.wustl.edu/etd/576

This Dissertation is brought to you for free and open access by Washington University Open Scholarship. It has been accepted for inclusion in All Theses and Dissertations (ETDs) by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.
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 max (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-idoacetamidofluorescein-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$_2$O$_2$. 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
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 O₃ treatment compared to ambient conditions.................................124
FIGURE 3: Summary of fold changes in total and redox proteomes of leaf tissue exposed to 115 ppb O₃...........................................................................................................125
FIGURE 4: Metabolic Overview of Total and Redox Proteome Changes in Soybean Leaf Tissue Exposed to 115 ppb O₃...........................................................................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
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$_2$O$_2$ 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$_2$O$_2$ 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
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>redox</td>
<td>reduction-oxidation</td>
</tr>
<tr>
<td>ROS</td>
<td>reactive oxygen species</td>
</tr>
<tr>
<td>NAD(P)H</td>
<td>nicotinamide adenine dinucleotide (phosphate)</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>γEC</td>
<td>γ-glutamylcysteine</td>
</tr>
<tr>
<td>GSH</td>
<td>glutathione</td>
</tr>
<tr>
<td>GSSG</td>
<td>oxidized glutathione</td>
</tr>
<tr>
<td>hGSH</td>
<td>homoglutathione</td>
</tr>
<tr>
<td>GCL</td>
<td>glutamate cysteine ligase</td>
</tr>
<tr>
<td>GS</td>
<td>glutathione synthetase</td>
</tr>
<tr>
<td>hGS</td>
<td>homoglutathione synthetase</td>
</tr>
<tr>
<td>SAR</td>
<td>systemic acquired resistance</td>
</tr>
<tr>
<td>NEM</td>
<td>N-ethylmaleimide</td>
</tr>
<tr>
<td>IAM</td>
<td>iodoacetamide</td>
</tr>
<tr>
<td>IAA</td>
<td>iodoacetic acid</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>TCEP</td>
<td>tris[2-carboxyethyl]phosphine</td>
</tr>
<tr>
<td>IAF</td>
<td>5′-iodoacetomidofluorescein</td>
</tr>
<tr>
<td>HPLC</td>
<td>high performance liquid chromatography</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>FPLC</td>
<td>fast protein liquid chromatography</td>
</tr>
<tr>
<td>MS</td>
<td>mass spectrometry</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>MALDI</td>
<td>matrix-assisted laser desorption/ionization</td>
</tr>
<tr>
<td>BSO</td>
<td>buthionine sulfoximide</td>
</tr>
<tr>
<td>NMVOCs</td>
<td>non-methane volatile organic chemicals</td>
</tr>
<tr>
<td>ppb</td>
<td>parts per billion</td>
</tr>
<tr>
<td>FACE</td>
<td>free air concentration enrichment</td>
</tr>
</tbody>
</table>
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 $\gamma$-glutamylcysteine synthetase; EC 6.3.2.2) catalyses the formation of $\gamma$-glutamylcysteine from glutamate and cysteine. In the second reaction the enzyme glutathione synthetase (GS; EC 6.3.2.3) adds glycine to $\gamma$-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 a 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, \(\gamma\)-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 homogluthathione – 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.
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 homoglutathione 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 Chapter 2.

PART II - THE THIOL-REDOX PROTEOME - BUTHIONINE SULFOXIMINE (BSO), H₂O₂, 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 chromatography 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), iodoacetamide-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 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 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 O₂ to various oxide radicals. While the availability of O₂ 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 O₂ to yield monoatomic oxygen. Monoatomic oxygen is highly unstable, so it rapidly recombines with O₂ to yield O₃. In the troposphere, ozone production begins when carbon monoxide reacts with hydroxide, yielding a proton and carbon dioxide. The proton then further reacts with O₂ to produce the peroxo radical HO₂. HO₂ 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 throughout 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 *Arabidopsis* 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-occurring 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 oxidation profiles in soybean tissue grown in the field under ambient and elevated chronic ozone concentrations.
References


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,


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


Hudman, R.C., Murray, L.T., Jacob, D.J., Millet, B., Turquety, S., Wu, S., Blake, D.R.,
anthropogenic sources of CO in the United States. Geophysical Research Letters
35: L04801.


Education 84: 14-18-1419.


Flexible loop that is novel catalytic machinery in a ligase. Atomic structure and


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.


**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.
\[
\begin{align*}
\text{O}_3 + \text{hv}(\lambda < 320 \text{ nm}) & \rightarrow \text{O}('D) + \text{O}_2 \quad (1) \\
\text{O}('D) + \text{H}_2\text{O} & \rightarrow \text{OH} + \text{OH} \quad (2) \\
\text{OH} + \text{NO}_2 & \rightarrow \text{HNO}_3 \text{ (nitric acid)} \quad (3) \\
\text{OH} + \text{SO}_2 & \rightarrow \ldots \rightarrow \text{HO}_2 + \text{H}_2\text{SO}_4 \text{ (sulfuric acid)} \quad (4) \\
\text{OH} + \text{hydrocarbons} & \rightarrow \text{HO}_2 + \text{partly oxidized organics} \quad (5) \\
\text{HO}_2 + \text{NO} & \rightarrow \text{OH} + \text{NO}_2 \quad (6) \\
\text{NO}_2 + \text{hv} (\lambda < 420 \text{ nm}) & \rightarrow \text{NO} + \cdot \text{O} \quad (7) \\
\cdot \text{O} + \text{O}_2 & \rightarrow \text{O}_3 \quad (8)
\end{align*}
\]

**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
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 \textit{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


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 to Glycine max to a closed active site form (γEC complex) to a closed active site form (hGSH+ADP+SO₄²⁻ 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 glycyl 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...
RESULTS

Protein Expression and Kinetic Analysis of hGS

Soybean hGS was overexpressed in E. coli as a His-tagged fusion protein and purified using Ni²⁺-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 Mg²⁺ 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/EC) fivefold lower but with comparable Kₘ 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 Kₘ 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.

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

<table>
<thead>
<tr>
<th>Substrate</th>
<th>γEC/Kₘ (μM)</th>
<th>γEC/Kₘ (M⁻¹ s⁻¹)</th>
<th>ATP/Kₘ (μM)</th>
<th>ATP/Kₘ (M⁻¹ s⁻¹)</th>
<th>Gly/Kₘ (μM)</th>
<th>Gly/Kₘ (M⁻¹ s⁻¹)</th>
<th>β-Ala/Kₘ (μM)</th>
<th>β-Ala/Kₘ (M⁻¹ s⁻¹)</th>
<th>hGS/Kₘ (μM)</th>
<th>hGS/Kₘ (M⁻¹ s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>γEC</td>
<td>12.2 ± 0.3</td>
<td>39 ± 5</td>
<td>312,800</td>
<td>12.2 ± 0.3</td>
<td>39 ± 5</td>
<td>12.2 ± 0.3</td>
<td>312,800</td>
<td>12.2 ± 0.3</td>
<td>39 ± 5</td>
<td>12.2 ± 0.3</td>
</tr>
<tr>
<td>ATP</td>
<td>12.1 ± 0.3</td>
<td>57 ± 10</td>
<td>212,300</td>
<td>12.1 ± 0.3</td>
<td>57 ± 10</td>
<td>12.1 ± 0.3</td>
<td>212,300</td>
<td>12.1 ± 0.3</td>
<td>57 ± 10</td>
<td>12.1 ± 0.3</td>
</tr>
<tr>
<td>Gly</td>
<td>12.6 ± 0.5</td>
<td>1,510 ± 88</td>
<td>8,340</td>
<td>12.6 ± 0.5</td>
<td>1,510 ± 88</td>
<td>12.6 ± 0.5</td>
<td>8,340</td>
<td>12.6 ± 0.5</td>
<td>1,510 ± 88</td>
<td>12.6 ± 0.5</td>
</tr>
<tr>
<td>β-Ala</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>β-Ala</td>
<td>–</td>
<td>β-Ala</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Values are expressed as a mean ± SE for n = 3. *Kinetic parameters for Arabidopsis GS are from Jez and Cahoon (2004) and are provided here for comparison to soybean hGS.
of 1.6 Å in proteins, such as human GS (root mean square [r.m.s.] deviation of 2.2 Å for soybean hGS is similar to those of other ATP-grasp family members). The second loop, involving residues 410 to 416, the lid domain becomes ordered, with the Ala-rich loop undergoing major rearrangements compared to the corresponding Ala residues in the Gly-rich loop providing multiple interactions with the nucleotide. Likewise, the Ala-rich loop shifts to position 480 to 489 in the second monomer. In the open form, both the lid domain and Ala-rich loop are ordered 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 to 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 480 to 489 in the second monomer. In the open form, both the lid domain and Ala-rich loop are ordered from the active site to reveal the binding sites for γEC and ATP and allow for substrate binding (Figures 3A and 3B).

### 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. 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 to 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 480 to 489 in the second monomer. In the open form, both the lid domain and Ala-rich loop are ordered from the active site to reveal the binding sites for γEC and ATP and allow for substrate binding (Figures 3A and 3B).

<table>
<thead>
<tr>
<th>Space Group</th>
<th>crystal</th>
<th>crystal + γEC</th>
<th>Closed + hGSH + ADP</th>
</tr>
</thead>
<tbody>
<tr>
<td>P2₁</td>
<td>P2₁</td>
<td>P2₁</td>
<td>P2₁</td>
</tr>
<tr>
<td>a = 64.96 Å, b = 80.55 Å, c = 90.00 Å; α = γ = 90.0°, β = 96.9°</td>
<td>a = 64.88 Å, b = 80.95 Å, c = 89.12 Å; α = γ = 90.0°, β = 95.6°</td>
<td>a = b = 115.7 Å, c = 101.8 Å; α = β = 90.0°, γ = 120°</td>
<td></td>
</tr>
</tbody>
</table>

### Data Collection

- **Resolution range (Å)**: 28.4–2.0
- **Reflections (total/unique)**: 132,097/59,099
- **Completeness (highest shell)**: 94.8% (86.8%)
- **Average B-factor (Å²)**: 88.4, 9.3, 2.3%
- **Stereochemistry: most favored, allowed, generously allowed**

### Table 2. Crystallographic Statistics

<table>
<thead>
<tr>
<th>Crystal</th>
<th>Open</th>
<th>Open + γEC</th>
<th>Closed + hGSH + ADP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space Group</td>
<td>P2₁</td>
<td>P2₁</td>
<td>P2₁</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td>a = 64.96 Å, b = 80.55 Å, c = 90.00 Å; α = γ = 90.0°, β = 96.9°</td>
<td>a = 64.88 Å, b = 80.95 Å, c = 89.12 Å; α = γ = 90.0°, β = 95.6°</td>
<td>a = b = 115.7 Å, c = 101.8 Å; α = β = 90.0°, γ = 120°</td>
</tr>
</tbody>
</table>

### Data Collection

- **Wavelength (Å)**: 0.979
- **Resolution range (Å) (highest shell)**: 28.4–2.0
- **Completeness (highest shell)**: 94.8% (86.8%)
- **Average B-factor (Å²)**: 88.4, 9.3, 2.3%
- **Stereochemistry: most favored, allowed, generously allowed**
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 GSH 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 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 (1.0σ) 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.

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.
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.
to first substrate complex (hGS together with studies of the kinetic and chemical mechanisms of including the lid domain, Gly-rich loop, and Ala-rich loop. To-multiple substrates and the rearrangement of active site features, substrate preference. This work reinforces the critical role of flexible mutagenesis, defines active site differences that govern sub-product diversity in hGS. Crystallographic analysis of soybean from GS, we examined the structural basis for adaptation of Frendo et al., 2005; Loscos et al., 2008). As hGS likely evolved to produce hGSH for root nodulation (Matamoros et al., 2003; et al., 1986; Klapheck et al., 1994; Klapheck et al., 1995; Meuwly (Meister, 1995); however, some plants also synthesize GSH and prokaryotes synthesize the multifunctional peptide GSH due to the evolution of metabolic pathways. Nearly all eukaryotes dimensional structures and reaction chemistry is a hallmark in Functional diversity across enzyme families with shared three-

**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 and prokaryotes synthesize the multifunctional peptide GSH. 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=EC 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 γ-EC 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 γ-EC 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 GS 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 Arg-153 likely stabilizes the transition state with the Ala-rich loop and Arg-475 orienting β-Ala for attack on the reaction intermediate to yield hGSH.

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

<table>
<thead>
<tr>
<th>Substrate</th>
<th>V_max (μM/min)</th>
<th>K_m (μM)</th>
<th>kcat/K_m (M^{-1} s^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>β-Ala</td>
<td>2.4 ± 0.1</td>
<td>3.4 ± 0.1</td>
<td>708</td>
</tr>
<tr>
<td>L487A</td>
<td>0.8 ± 0.1</td>
<td>3.8 ± 0.6</td>
<td>211</td>
</tr>
<tr>
<td>P488A</td>
<td>2.5 ± 0.2</td>
<td>8.0 ± 1.4</td>
<td>313</td>
</tr>
<tr>
<td>LP/AA</td>
<td>1.9 ± 0.2</td>
<td>24.8 ± 2.3</td>
<td>77</td>
</tr>
<tr>
<td>hGS</td>
<td>&lt;0.1</td>
<td>&gt;100</td>
<td></td>
</tr>
<tr>
<td>L487A</td>
<td>0.3 ± 0.1</td>
<td>6.5 ± 0.5</td>
<td>46</td>
</tr>
<tr>
<td>P488A</td>
<td>1.4 ± 0.1</td>
<td>5.1 ± 0.4</td>
<td>274</td>
</tr>
<tr>
<td>LP/AA</td>
<td>2.1 ± 0.3</td>
<td>2.2 ± 0.1</td>
<td>960</td>
</tr>
</tbody>
</table>

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

Denotes the corresponding double mutant (L487A/P488A).

In both hGS and GS, structuring of the lid domain and Ala-rich loop appears linked to binding of ATP and the terminal substrate recognition 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; Tokunki 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 Mg2+ 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; Hurbe-Ormaeche et al., 2002; Spigey et al., 2002), there appears to be a 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 GS 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., 1996), 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. NTA+ 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 QuickChange site-directed mutagenesis kit was from Stratagene. hGSH was from Bachem. All other reagents were 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'-dTTCAGCATGATGGCTGACCCTTTGACCC-ACC-3' as the forward primer (the Ncol site is underlined, and the start codon is in bold) and 5'-dTTCAGCATGATGGCTGACCCTTTGACCC-3' as the reverse primer (the Ncol site is underlined, and the stop codon is in bold). The resulting PCR product was digested with Ncol and NotI and then subcloned into pHiSH (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 A600 = ~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 (10,000g; 10 min) and resuspended in 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 25 mM imidazole, 5 mM MgCl2, 10% (v/v) glycerol, and 1% (v/v) Tween 20. Sonication was used to lyse cells. Following centrifugation (45,000g; 45 min), the supernatant was passed through a NTA+–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 MgCl2, 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. Diluted protein was loaded on a mixed benzamidine-sepharose/NTA+–NTA column. The flow-through of this step was loaded onto an HiLoad 26/60 Superdex-200 FPLC column equilibrated with 25 mM HEPES, pH 7.5, 5 mM MgCl2 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 QuickChange PCR method with the pHiSH-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 containing 100 μM HEPES, pH 7.5, 150 mM NaCl, 20 mM MgCl2, 2.5 mM γEC, 10 mM L-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 A600 (= 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 Km value). Untransformed data was fit to the Michaelis-Menten equation, v = vmax[S]/(Km + [S]), using Kaleidagraph (Synergy Software).

Protein Crystallization and Structure Determination

Crystals of hGS were obtained by the vapor diffusion method in 4-% L hanging drops of a 1:1 mixture of protein and crystallization buffer (20% PEG3000, 0.1 M MOPSO, pH 7, and 0.2 M MgSO4) at 4°C over a 0.5-mL reservoir. For cocrystallization with ligands, either 6 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 XSSCALE (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 v-EC 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).

Supplemental Data

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


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 MgCl₂, 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 aprofin (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.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sense Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>L487A</td>
<td>5'-dTCTATGAAGGAGGAGTTGGCGCTGGTGCTTTGAGTGAGTAG-3'</td>
</tr>
<tr>
<td>P488A</td>
<td>5'-dTCTATGAAGGAGGAGTTGGCTGTGGCTTTGGAGTGAGTAG-3'</td>
</tr>
<tr>
<td>L487A/P488A</td>
<td>5'-dCTTATGAAGGAGGAGTTGGCGCTGTGGCTTTGGAGTGAGTAG-3'</td>
</tr>
</tbody>
</table>

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.

Copyright American Society of Plant Biologists.

www.plantcell.org
CHAPTER 3

REDOX-REGULATORY MECHANISMS INDUCED BY OXIDATIVE STRESS IN

BRASSICA JUNCEA ROOTS MONITORED BY 2-DE PROTEOMICS
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$_2$O$_2$ 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


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

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

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$_2$O$_2$), 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$_2$O$_2$ or buthionine sulfoximine, which depletes glutathione to cause accumulation of endogenous H$_2$O$_2$, led to the identification of different sets of proteins. These data suggest that exogenous and endogenous oxidative stresses trigger specialized responses.

**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$_2$O$_2$) causes reversible and irreversible redox modifications to proteins during oxidative stress [2, 3]. Although many H$_2$O$_2$-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$_2$O$_2$ levels are controlled by mechanisms, such as the glutathione-ascorbate system, that maintain concentrations of key reducing molecules [1]. As a consequence, H$_2$O$_2$ has long been used to elicit oxidative stress responses to study redox mechanisms and provide insight into the molecular physiology of adaptive responses.

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$_2$O$_2$-induced oxidative stresses. Application of H$_2$O$_2$ to plant roots provides an exogenous stress and application of buthionine sulfoximine (BSO), which depletes glutathione, produces an accumulation of endogenous H$_2$O$_2$ [4]. Largely different sets of proteins regulated by H$_2$O$_2$ 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$_2$O$_2$ source in *B. juncea* roots.

Wild-type *B. juncea* seeds were germinated in a growth chamber at 22°C, 200 μmol/m$^2$/s 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$_2$O$_2$, 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$_2$O$_2$ and BSO were chosen...
based on a previous experiment, showing that these compounds alter oxidation state of a redox-sensitive protein in plants.

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 thioserea, 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 destack 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 (described previously [6]. Gels were imaged with a Typhoon (Bio-Rad, Hercules, CA, USA) and 2-DE performed as 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$_2$O$_2$ and BSO treatments, respectively, and 27 and 40 spots differed significantly in total protein abundance for the H$_2$O$_2$ 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 spots were overlaid with the ones identified as redox regulated. 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$_2$O$_2$ 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$_2$O$_2$ Two GST isoforms showed increased oxidation in response to BSO and H$_2$O$_2$ 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$_2$O$_2$ and BSO

<table>
<thead>
<tr>
<th></th>
<th>H$_2$O$_2$/Sypro</th>
<th>H$_2$O$_2$/IAF</th>
<th>BSO/Sypro</th>
<th>BSO/IAF</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of spots detected</td>
<td>235</td>
<td>243</td>
<td>288</td>
<td>250</td>
<td>176</td>
</tr>
<tr>
<td>Number of spots differentially expressed or oxidized</td>
<td>27 (37%)</td>
<td>59 (12%)</td>
<td>49 (29%)</td>
<td>50 (16%)</td>
<td>176</td>
</tr>
<tr>
<td>Number of spots identified as one protein ID</td>
<td>11</td>
<td>17</td>
<td>12</td>
<td>12</td>
<td>52</td>
</tr>
<tr>
<td>Number of spots identified with multiple proteins IDs</td>
<td>4</td>
<td>19</td>
<td>12</td>
<td>16</td>
<td>51</td>
</tr>
<tr>
<td>Total number of spots identified</td>
<td>15 (55%)</td>
<td>36 (61%)</td>
<td>24 (60%)</td>
<td>28 (56%)</td>
<td>103</td>
</tr>
</tbody>
</table>

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$_2$O$_2$ 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 \( \text{H}_2\text{O}_2 \) accumulation. Two TRXh isoforms also showed greater oxidation following \( \text{H}_2\text{O}_2 \) 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 imbalance [8]. 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). Proteins in glycolysis, stress response, carbohydrate metabolism, and proteolysis also exhibited redox changes (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 \( \text{H}_2\text{O}_2 \) treatment. Two isoforms showed increased oxidation after BSO treatment, whereas one isoform showed greater reduction only following \( \text{H}_2\text{O}_2 \) treatment. Decreased glutathione levels after BSO treatment increases the oxidized 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-biphosphatase 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 \( \text{H}_2\text{O}_2 \) (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 metabolism 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 GF2c) and GF14o, were identified as increased in oxidation state in response to BSO and \( \text{H}_2\text{O}_2 \) 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 \( \text{H}_2\text{O}_2 \) in cucumber leaves and increase oxidative tolerance [20]. In this study, the application of \( \text{H}_2\text{O}_2 \) and the induction of endogenous \( \text{H}_2\text{O}_2 \) 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 \( \text{H}_2\text{O}_2 \) 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 \( \text{H}_2\text{O}_2 \) 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 hTnXNgWFy, 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


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$_2$O$_2$ 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

<table>
<thead>
<tr>
<th>Amino acid biosynthesis</th>
<th>Brassinosteroid mediated signaling pathway</th>
<th>Calcium mediated signaling pathway</th>
<th>Cytoskeleton organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-phosphoglycerate</td>
<td>GRF3 (Genes regulatory factor 3)</td>
<td>CaM (Calmodulin)</td>
<td>TUB4 (Tubulin beta-4 chain)</td>
</tr>
<tr>
<td>dehydrogenase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spot #36</td>
<td>Spot #48</td>
<td>Spot #30</td>
<td>Spot #65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Carbohydrate metabolic process</th>
<th>Cell redox homeostasis</th>
<th>Defense response</th>
<th>Glycolysis</th>
<th>Menaquinone biosynthesis</th>
<th>Nucleosome assembly</th>
</tr>
</thead>
<tbody>
<tr>
<td>fructose-1,6-bisphosphatase</td>
<td>GSH1 (Glutathione S-transferase 2)</td>
<td>MLP10 (MLP-like protein 3.4B)</td>
<td>fructose-1,6-bisphosphatase aldehyde</td>
<td>Dimethylmenaquinone dimethyltransferase</td>
<td>Histone H4</td>
</tr>
<tr>
<td>Spot #33 and 38</td>
<td>Spot #14</td>
<td>Spot #29 and 67</td>
<td>Spot #62</td>
<td>Spot #57</td>
<td>Spot #83</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nucleotide metabolism</th>
<th>Oxidation reduction</th>
<th>Protein folding</th>
<th>Protein transport</th>
<th>Stress response</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDPK1 (Nucleoside diphosphate kinase 1)</td>
<td>Gamma hydroxybutyrate dehydrogenase</td>
<td>CPR20 (Chaperonin 20)</td>
<td>ADF-ribosylation factor 1</td>
<td>Disease resistance responsive family protein</td>
<td>60S ribosomal protein L12</td>
</tr>
<tr>
<td>Spot #22 and 60</td>
<td>Spot #61</td>
<td>Spot #42</td>
<td>Spot #22</td>
<td>Spot #80</td>
<td>Spot #92</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spot #92</td>
</tr>
</tbody>
</table>
### Supplemental Table 1

<table>
<thead>
<tr>
<th>Spot #</th>
<th>Peptide sequence</th>
<th>Average of normalized volume in Control</th>
<th>Standard deviation</th>
<th>Average of normalized volume in Treatment</th>
<th>Standard deviation</th>
<th>P value</th>
<th>Fold change</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 (19562 3672)</td>
<td>AMVEELLK            EEDTPNFTGSFYSK</td>
<td>101.1</td>
<td>4.5</td>
<td>53.4</td>
<td>7.8</td>
<td>1.0</td>
<td>27</td>
<td>1.9</td>
</tr>
<tr>
<td>16 (14423 3528)</td>
<td>AGFAGDDAPR</td>
<td>481.4</td>
<td>26.2</td>
<td>265.3</td>
<td>26.2</td>
<td>-1.9</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td>17 (15219 324)</td>
<td>TNWIGGLLGK</td>
<td>946.5</td>
<td>11.6</td>
<td>1461</td>
<td>112.4</td>
<td>1.5</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>18 (15203 326)</td>
<td>LLKVTLGEK</td>
<td>1014.7</td>
<td>116.1</td>
<td>1524.2</td>
<td>314.7</td>
<td>1.5</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>19 (15204 325)</td>
<td>LDAEMPADSGYPAYLAAR</td>
<td>1063.6</td>
<td>17.1</td>
<td>733.9</td>
<td>47.9</td>
<td>-1.5</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>20 (15224 391)</td>
<td>DAVILVFANK</td>
<td>315</td>
<td>10.5</td>
<td>1461</td>
<td>112.4</td>
<td>1.5</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>21 (15222 399)</td>
<td>VMSADVAATTR</td>
<td>950.1</td>
<td>15.7</td>
<td>2244.3</td>
<td>1716.7</td>
<td>1.3</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>22 (26792 4017)</td>
<td>EDDTPNFTGSFYSK</td>
<td>2146.1</td>
<td>328.1</td>
<td>2867.3</td>
<td>74.7</td>
<td>1.3</td>
<td>0.033</td>
<td></td>
</tr>
<tr>
<td>23 (26790 3965)</td>
<td>EDDTPNFTGSFYSK</td>
<td>2871.1</td>
<td>357.9</td>
<td>2344.3</td>
<td>179.2</td>
<td>-1.3</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>24 (13895 092)</td>
<td>KPLK</td>
<td>308.1</td>
<td>45.5</td>
<td>679</td>
<td>22.3</td>
<td>-4.5</td>
<td>0.012</td>
<td></td>
</tr>
</tbody>
</table>
| 25 (15375 068) | KQDEGRVLA 
AND REGU 
AT  | 836.9 | 48.5 | 2117 | 364.4 | -1.2 | 0.001 |
| 26 (2609 712) | CLPZ  | 465.6 | 27.4 | 122.1 | 64.7 | -3.8 | 0.007 |
| 27 (2621 107) | PIIQ 
AND REGU 
AT  | 1201.1 | 340.3 | 375.8 | 227.5 | -3.2 | 0.031 |
| 28 (11607 2175) | ADN | 1251.1 | 46.5 | 375.8 | 227.5 | -3.2 | 0.031 |
| 29 (15239 696) | 5-methyltetrahydrofolate 
and 
reductase | 759.7 | 15.7 | 287.8 | 25.5 | -2.6 | 4.9E-03 |
| 30 (15238 762) | OGDH 
GLUTAMATE 
DEHYDROGENASE 1 | 759.7 | 15.7 | 287.8 | 25.5 | -2.6 | 4.9E-03 |
| 31 (14794 732) | 5-methyltetrahydrofolate 
and 
reductase | 47.7 | 8.2 | 120.4 | 28.6 | 2.5 | 0.007 |
| 32 (15211 107) | 5-methyltetrahydrofolate 
and 
reductase | 1999.5 | 344.5 | 902.9 | 71.4 | -2.2 | 0.002 |
| 33 (15334 823) | AATP | 711.1 | 182.9 | 325.2 | 364.4 | -2.2 | 0.011 |
| 34 (15234 993) | AATP | 711.1 | 182.9 | 325.2 | 364.4 | -2.2 | 0.011 |
| 35 (14794 732) | 5-methyltetrahydrofolate 
and 
reductase | 2361.2 | 192.8 | 1211.5 | 216.5 | -1.9 | 0.004 |
<table>
<thead>
<tr>
<th>Accession</th>
<th>Description</th>
<th>gi</th>
<th>Score</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi</td>
<td>18391442</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15231715</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>120675</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15233268</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15219257</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15242792</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>110741046</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>292816838</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15227987</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15219234</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15219234</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15242792</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>1149571</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>113205143</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15231176</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>113205143</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>113205143</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>297816838</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>297824877</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>30691626</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>297824877</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>30691626</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>297816838</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15240663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>152316938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein Name</td>
<td>Accession Number</td>
<td>Description</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------</td>
<td>-------------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td>PAC1 (20S proteasome alpha subunit 1)</td>
<td>gi</td>
<td>115233288</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Cystathionine beta-synthase</td>
<td>gi</td>
<td>111131694</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Hypothetical protein</td>
<td>gi</td>
<td>1150745108</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Hypothetical protein</td>
<td>gi</td>
<td>1184554771</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Hypothetical protein</td>
<td>gi</td>
<td>1150745229</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Cytochrome c</td>
<td>gi</td>
<td>115224101</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Hypothetical protein (1-34 kDa)</td>
<td>gi</td>
<td>110994926</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Cysteine synthase</td>
<td>gi</td>
<td>119424131</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Hypothetical protein</td>
<td>gi</td>
<td>1150523493</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Hypothetical protein</td>
<td>gi</td>
<td>114341708</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Hypothetical protein</td>
<td>gi</td>
<td>111857278</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Hypothetical protein</td>
<td>gi</td>
<td>1150523493</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Hypothetical protein</td>
<td>gi</td>
<td>119772968</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Hypothetical protein</td>
<td>gi</td>
<td>1217231875</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Hypothetical protein</td>
<td>gi</td>
<td>117152637</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Hypothetical protein</td>
<td>gi</td>
<td>117152873</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Protein S-adenosyl-L-homocystein hydrolase</td>
<td>gi</td>
<td>224981577</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Dimethylmenaquinone methyltransferase family protein</td>
<td>gi</td>
<td>112144574</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>ISOMERASE</td>
<td>gi</td>
<td>117557956</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>BSO</td>
<td>gi</td>
<td>116172888</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>BSO</td>
<td>gi</td>
<td>115045616</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>BSO</td>
<td>gi</td>
<td>1131130129</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>BSO</td>
<td>gi</td>
<td>1217674585</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>BSO</td>
<td>gi</td>
<td>1150329663</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>BSO</td>
<td>gi</td>
<td>1150133254</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>BSO</td>
<td>gi</td>
<td>1131130129</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>BSO</td>
<td>gi</td>
<td>1150329663</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td>Score</td>
<td>E-value</td>
<td>q-value</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------------------------------------------------</td>
<td>-------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>gi</td>
<td>14423528</td>
<td>putative fructokinase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>1429207</td>
<td>gi</td>
<td>14422255</td>
<td>enoyl-[acyl-carrier protein] reductase</td>
</tr>
<tr>
<td>gi</td>
<td>11034734</td>
<td>gi</td>
<td>14423528</td>
<td>gi</td>
</tr>
</tbody>
</table>

*(Note: The table contains data on various proteins and their corresponding scores, E-values, and q-values. The data is presented in a tabular format with columns for Accession, Description, Score, E-value, q-value, MZ, MR, MS, MW, and PI.)*
<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Protein Name</th>
<th>Description</th>
<th>Log2 Ratio</th>
<th>Fold Change</th>
<th>P-Value</th>
<th>BSO Treatment</th>
<th>Control Treatment</th>
<th>BSO Change</th>
<th>Control Change</th>
<th>BSO Fold Change</th>
<th>Control Fold Change</th>
<th>BSO Treatment Change</th>
<th>Control Treatment Change</th>
<th>BSO P-Value</th>
<th>Control P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi</td>
<td>115499139</td>
<td>DSTLIMQLLR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>19411961</td>
<td>GRF2</td>
<td>General Regulatory Factor 2</td>
<td>-1.0</td>
<td>0.84</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>21825899</td>
<td>GRF2</td>
<td>General Regulatory Factor 2</td>
<td>-1.2</td>
<td>0.12</td>
<td>0.013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>26898672</td>
<td>GRF2</td>
<td>General Regulatory Factor 2</td>
<td>-4.3</td>
<td>0.001</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>111071815</td>
<td>H-protease</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>198967982</td>
<td>BSO</td>
<td>Proline-rich protein</td>
<td>-0.5</td>
<td>0.41</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>297819837</td>
<td>MLP328</td>
<td>MLP-Like Protein 328</td>
<td>1.4</td>
<td>0.12</td>
<td>0.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>329480633</td>
<td>MLP328</td>
<td>MLP-Like Protein 328</td>
<td>-1.3</td>
<td>0.03</td>
<td>0.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>265524099</td>
<td>Sypro 82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>123969570</td>
<td>Sypro 82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>114720685</td>
<td>Sypro 82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>116183425</td>
<td>Sypro 82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td>M</td>
<td>W</td>
<td>V</td>
<td>U</td>
<td>S</td>
<td>F</td>
<td>E</td>
<td>R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15235213</td>
<td>caffeoyl-CoA 3-O-methyltransferase, putative</td>
<td>45</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15224993</td>
<td>PAA2 (20S PROTEASOME SUBUNIT PAA2)</td>
<td>45</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15222633</td>
<td>disease resistance-responsive family protein / dirigent family protein</td>
<td>52</td>
<td>77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15239146</td>
<td>ATNADP-ME2 (NADP-MALIC ENZYME 2)</td>
<td>45</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15229559</td>
<td>HSP60 (Heat shock protein 60)</td>
<td>51</td>
<td>68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>414103</td>
<td>myrosinase, thioglucoside glucohydrolase</td>
<td>47</td>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15010652</td>
<td>At1g09780/F21M12_17</td>
<td>55</td>
<td>47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>18395103</td>
<td>GRF10 (GENERAL REGULATORY FACTOR 10)</td>
<td>40</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15241286</td>
<td>MAB1 (MACCI-BOU)</td>
<td>78</td>
<td>53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>14423528</td>
<td>putative fructokinase</td>
<td>102</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15231176</td>
<td>ATPQ (ATP SYNTHASE D CHAIN, MITOCHONDRIAL)</td>
<td>57</td>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15236385</td>
<td>SBP1 (selenium-binding protein 1)</td>
<td>47</td>
<td>63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>118429132</td>
<td>vacuolar ATPase subunit B</td>
<td>47</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>297799542</td>
<td>hypothetical protein</td>
<td>49</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>297824877</td>
<td>glutathione S-transferase</td>
<td>54</td>
<td>52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>297816838</td>
<td>predicted protein</td>
<td>69</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>33285914</td>
<td>putative dehydroascorbate reductase</td>
<td>59</td>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>15219257</td>
<td>PAB1 (PROTEASOME SUBUNIT PAB1)</td>
<td>62</td>
<td>48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* 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.
<table>
<thead>
<tr>
<th>Protein accession</th>
<th>Treatment</th>
<th>Biological process</th>
<th>Average of normalized volume</th>
<th>Standard deviation</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi</td>
<td>114171</td>
<td>IAF</td>
<td>Stain</td>
<td>22.3</td>
<td>-4.5</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>Amino acid biosynthesis</td>
<td>617.7</td>
<td>100.3</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>Amino acid biosynthesis</td>
<td>56.3</td>
<td>1407.2</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>9.1</td>
<td>1.3</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>1645.7</td>
<td>1.4</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>494.4</td>
<td>42.2</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>171.4</td>
<td>192.1</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>1645.7</td>
<td>1.4</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>494.4</td>
<td>42.2</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>171.4</td>
<td>192.1</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>1645.7</td>
<td>1.4</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>494.4</td>
<td>42.2</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>171.4</td>
<td>192.1</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>1645.7</td>
<td>1.4</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>494.4</td>
<td>42.2</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>171.4</td>
<td>192.1</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>1645.7</td>
<td>1.4</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>494.4</td>
<td>42.2</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>171.4</td>
<td>192.1</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>1645.7</td>
<td>1.4</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>494.4</td>
<td>42.2</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>171.4</td>
<td>192.1</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>1645.7</td>
<td>1.4</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>494.4</td>
<td>42.2</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>171.4</td>
<td>192.1</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>1645.7</td>
<td>1.4</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>494.4</td>
<td>42.2</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>171.4</td>
<td>192.1</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>1645.7</td>
<td>1.4</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>494.4</td>
<td>42.2</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>171.4</td>
<td>192.1</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>1645.7</td>
<td>1.4</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>494.4</td>
<td>42.2</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>171.4</td>
<td>192.1</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>1645.7</td>
<td>1.4</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>494.4</td>
<td>42.2</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>171.4</td>
<td>192.1</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>1645.7</td>
<td>1.4</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>494.4</td>
<td>42.2</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>171.4</td>
<td>192.1</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>1645.7</td>
<td>1.4</td>
</tr>
<tr>
<td>gi</td>
<td>11131564</td>
<td>IAF</td>
<td>ATP synthesis coupled proton transport</td>
<td>494.4</td>
<td>42.2</td>
</tr>
</tbody>
</table>
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$_2$O$_2$-treated, and BSO-treated. H$_2$O$_2$-treated and BSO-treated plants were pot-watered with 2 L of 1 mM H$_2$O$_2$ 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$_2$O$_2$ 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$_2$O$_2$, 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 (<60 µL 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 differ 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 NH₄HCO₃, 50% ACN, with the liquid discarded after every wash. Following the last NH₄HCO₃/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 NH₄HCO₃ 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 transferred 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].

CHAPTER 4

FROM CLIMATE CHANGE TO PROTEINS: REDOX PROTEOMICS OF

OZONE-INDUCED RESPONSES IN SOYBEAN
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 trifoliate 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 begins

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

Ashley Galant\textsuperscript{a}, Robert P. Koester\textsuperscript{b}, Elizabeth A. Ainsworth\textsuperscript{b,c}, Leslie M. Hicks\textsuperscript{d}, Joseph M. Jez\textsuperscript{a,*}

\textsuperscript{a}Department of Biology, Washington University, One Brookings Drive, Campus Box 1137, St. Louis, MO 63130 USA; \textsuperscript{b}Department of Plant Biology, University of Illinois, Urbana-Champaign, Urbana, IL 61801 USA; \textsuperscript{c}USDA-ARS Global Change and Photosynthesis Research Unit, Urbana, IL 61801 USA; and \textsuperscript{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 O₃ levels
are more widespread across metabolism than previously reported and are not necessarily
the same as those resulting from short-term acute O₃ exposure. In addition, we provide
the first direct demonstration that high O₃ 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 O₃ exposure.

RESULTS

Analysis of 2-DE spots in O₃-treated soybean

To identify O₃-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 O₃ 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 O₃; 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 O₃ 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 O₃ leaf, 115 ppb O₃ root, 60 ppb O₃ leaf, and 60 ppb O₃ root samples, respectively. A further 38 proteins were found to change in multiple tissue-O₃ 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 O₃ 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 O₃ root sample had 11 proteins that increased in oxidation and 8 that were more reduced (Fig. S2A). In the 60 ppb O₃ 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 O₃ 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 O₃-treatment**

To better examine the activity changes in the 115 ppb O₃ 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 O₃ leaf sample. Except for PGK, each enzyme exhibited increased activity in the high O₃ 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 O₃ 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 O₃ leaf sample and the ambient control. Likewise, expression of PEPC in the ambient and high O₃ leaf samples, as determined by Western blot, was not altered (not shown).

In the 115 ppb O₃ 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 O₃ 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 O₃ 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 O₃ exposure on the proteomes of different crops, including soybean (22-28), these reports have neither
reported on the consequences of chronic O₃ exposure under field conditions nor examined the effect of O₃ on redox-sensitive proteins. Previous studies, in which soybean, rice, or wheat were exposed to constant 120 to 200 ppb O₃ 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 O₃ 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 O₃ stress exhibited less than 2-fold differences in redox state for only a handful of proteins (Fig. S2). In contrast, the high O₃ 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 O₃ 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 O₃ exposure.

Tropospheric O₃ 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). O₃ 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 O₃ 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\textsubscript{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\textsubscript{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\textsubscript{3} leaf sample (Fig 4). In addition, the total activity levels of FBA, GAPDH, and MDH increased in the 115 ppb O\textsubscript{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\textsubscript{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\textsubscript{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\textsubscript{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\textsubscript{3} stress (23). Of the 79 proteins identified in the 115 ppb O\textsubscript{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\textsubscript{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\textsubscript{3} stress were also identified in the 115 ppb O\textsubscript{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-associateeed 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\textsubscript{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\textsubscript{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\textsubscript{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\textsubscript{3} concentrations has not been examined previously. Analysis of the soybean redox proteome in the 115 ppb O\textsubscript{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\textsubscript{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 O₃ 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. O₃ 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 O₃ 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 O₃ Treatment. Soybean (G. max (L.) Merr) were planted and exposed to ambient (40 ppb), elevated (60 ppb), and high (115 ppb) O₃, as described previously (16). All O₃ levels are seasonal 9-hour average concentrations. O₃ 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<sub>ex</sub>=488 nm and λ<sub>em</sub>=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 (\( \lambda_{\text{ex}} = 457 \) nm and \( \lambda_{\text{em}} = 610 \) 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 NH₄HCO₃; 50% ACN (15 min; 900 rpm) with the liquid discarded after every wash to remove the Sypro Ruby stain. Following the last NH₄HCO₃/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 NH₄HCO₃ 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 transferred 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
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


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


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


to ozone reveals major changes in carbon metabolism. *Proteomics* 7: 1584-1599


**Figure 1.** Representative 2-DE Gels. The gels shown contain proteins isolated from roots of soybean exposed to 60 ppb O₃. (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 O₃ treatments compared to ambient conditions. (A) Distribution of proteins across all four combinations of tissue and O₃ 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 O₃. Fold changes, relative to ambient O₃ 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 O₃. 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.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Ambient Ozone Activity ($\mu$mol min$^{-1}$ g FW$^{-1}$)</th>
<th>High Ozone Activity ($\mu$mol min$^{-1}$ g FW$^{-1}$)</th>
<th>Activity Fold Change</th>
<th>Total Protein (Oxidation) Fold Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>PGK</td>
<td>715 ± 82</td>
<td>629 ± 68</td>
<td>0.9</td>
<td>1.8 (5.7)</td>
</tr>
<tr>
<td>FBA</td>
<td>31.6 ± 6.5</td>
<td>72.1 ± 9.9</td>
<td>2.3</td>
<td>1.8 (4.8)</td>
</tr>
<tr>
<td>GAPDH</td>
<td>183 ± 75</td>
<td>693 ± 301</td>
<td>3.8</td>
<td>1.7 (4.4)</td>
</tr>
<tr>
<td>MDH</td>
<td>72.0 ± 3.3</td>
<td>93.9 ± 10.4</td>
<td>1.3</td>
<td>2.0 (5.4)</td>
</tr>
<tr>
<td>GS</td>
<td>106 ± 9</td>
<td>246 ± 25</td>
<td>2.3</td>
<td>1.9 (4.5)</td>
</tr>
<tr>
<td>RuBisCO</td>
<td>11.1 ± 2.6</td>
<td>10.0 ± 4.1</td>
<td>0.9</td>
<td>ND (ND)</td>
</tr>
<tr>
<td>exochitinase</td>
<td>10.5 ± 0.8</td>
<td>17.2 ± 0.3</td>
<td>1.6</td>
<td>4.0 (9.4)</td>
</tr>
<tr>
<td>endochitinase</td>
<td>0.34 ± 0.02</td>
<td>1.40 ± 0.08</td>
<td>4.1</td>
<td>4.0 (9.4)</td>
</tr>
<tr>
<td>chitobiosidase</td>
<td>0.012 ± 0.046</td>
<td>0.133 ± 0.023</td>
<td>11.1</td>
<td>4.0 (9.4)</td>
</tr>
</tbody>
</table>
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 O₃. 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.

<table>
<thead>
<tr>
<th>O₃ Level, Tissue, and Signal</th>
<th>Spots Detected</th>
<th>Spots Differentially Expressed/Oxidized</th>
<th>Spots with One Protein</th>
<th>Spots with Multiple Proteins</th>
<th>Total Spots Identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>115 ppb, leaf Sypro</td>
<td>154</td>
<td>29</td>
<td>9</td>
<td>17</td>
<td>26</td>
</tr>
<tr>
<td>115 ppb, leaf IAF</td>
<td>171</td>
<td>47</td>
<td>9</td>
<td>18</td>
<td>27</td>
</tr>
<tr>
<td>115 ppb, root Sypro</td>
<td>208</td>
<td>36</td>
<td>4</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>115 ppb, root IAF</td>
<td>196</td>
<td>29</td>
<td>5</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>60 ppb, leaf Sypro</td>
<td>196</td>
<td>44</td>
<td>7</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>60 ppb, leaf IAF</td>
<td>195</td>
<td>43</td>
<td>7</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>60 ppb, root Sypro</td>
<td>158</td>
<td>21</td>
<td>7</td>
<td>10</td>
<td>17</td>
</tr>
<tr>
<td>60 ppb, root IAF</td>
<td>177</td>
<td>28</td>
<td>9</td>
<td>15</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td>1455</td>
<td>277</td>
<td>57</td>
<td>83</td>
<td>140</td>
</tr>
</tbody>
</table>
Tables S2-S5. Data Summary Tables. These tables list peptides that are differentially expressed and/or oxidized from 115 ppb \( \text{O}_3 \) leaf tissue (Table S2), 115 ppb \( \text{O}_3 \) root tissue (Table S3), 60 ppb \( \text{O}_3 \) leaf tissue (Table S4), and 60 ppb \( \text{O}_3 \) 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

<table>
<thead>
<tr>
<th>Protein</th>
<th>Accession Numbers(s)</th>
<th>Mol Weight</th>
<th>Fold Change IAF (spot number)</th>
<th>Fold Change SYPRO (spot number)</th>
<th>Classification</th>
<th>IAF Peptides (spot number)</th>
<th>Best Mascot Ion Score</th>
<th>SYPRO Peptides (spot number)</th>
<th>Best Mascot Ion Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIM phosphate binding protein (Glycine max)</td>
<td>gi</td>
<td>255640328</td>
<td>14 kDa</td>
<td>[(35)3.347] [(55)2.375] [(59)2.260]</td>
<td>glycolysis</td>
<td>[S35] AAGDALLFR SNSLAQLGK YTAEGESEEATR</td>
<td>56.3</td>
<td>60.8</td>
<td>70.3</td>
</tr>
<tr>
<td>Nucleoside/nucleotide kinase (Glycine max)</td>
<td>gi</td>
<td>255630590</td>
<td>50 kDa</td>
<td>[(35)3.347]</td>
<td>nucleotide metabolism</td>
<td>[S35] AAQDALLFR SNSLAQLGK YTAEGESEEATR</td>
<td>56.3</td>
<td>60.8</td>
<td>70.3</td>
</tr>
<tr>
<td>Rossmann-fold NAD(P)(+)-binding protein; 3-oxoacyl-(acyl-carrier-protein) reductase (Populus trichocarpa)</td>
<td>gi</td>
<td>224100009</td>
<td>28 kDa</td>
<td>[(13)2.805]</td>
<td>fatty acid metabolism</td>
<td>[S13] EADFSTDDILGK FKGVFLSR</td>
<td>76.1</td>
<td>54.8</td>
<td>[S13] EADFSTDDILGK FKGVFLSR</td>
</tr>
<tr>
<td>Rossmann-fold NAD(P)(+)-binding protein (Glycine max)</td>
<td>gi</td>
<td>255647108</td>
<td>42 kDa</td>
<td>[(14)2.072]</td>
<td>unknown</td>
<td>[S14] EADFSTDDILGK FKGVFLSR</td>
<td>76.1</td>
<td>54.8</td>
<td>[S14] EADFSTDDILGK FKGVFLSR</td>
</tr>
<tr>
<td>Rossmann-fold NAD(P)(+)-binding protein (Glycine max)</td>
<td>gi</td>
<td>224100009</td>
<td>36 kDa</td>
<td>[(10)1.665], [(25)2.070]</td>
<td>secondary metabolism</td>
<td>[S10] FfpsEFGLDVDR NLACQDTYPPR VFIQGDGNVK</td>
<td>51.7</td>
<td>60.4</td>
<td>55.4</td>
</tr>
<tr>
<td>14-3-3 protein [Nicotiana tabacum]</td>
<td>gi</td>
<td>11577915</td>
<td>25 kDa</td>
<td>[(89)1.667]</td>
<td>signal transduction</td>
<td>[S89] ENFVYYAK ERENFVYYAK</td>
<td>56.6</td>
<td>55.8</td>
<td>[S89] ENFVYYAK ERENFVYYAK</td>
</tr>
<tr>
<td>Protein Name</td>
<td>Accession Numbers</td>
<td>Molecular Weight</td>
<td>pI</td>
<td>Mw</td>
<td>Function</td>
<td>Accession Numbers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------</td>
<td>------------------</td>
<td>----</td>
<td>----</td>
<td>----------</td>
<td>------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glutathione-S-Transferase (Glycine max)</td>
<td>gi</td>
<td>255633331 gi</td>
<td>2255633119 gi</td>
<td>388679415</td>
<td>25 kDa</td>
<td>[46]</td>
<td>[46] 1.54b</td>
<td>Redox homeostasis</td>
<td>[50]</td>
</tr>
<tr>
<td>Glutathione-S-Transferase (Glycine max)</td>
<td>gi</td>
<td>255633207</td>
<td>26 kDa</td>
<td>[10]</td>
<td>[10] 3.022</td>
<td>Redox homeostasis</td>
<td>[10]</td>
<td>GVEY YKEENLR VDYDDYVAVTK</td>
<td>84</td>
</tr>
<tr>
<td>Proteasome beta type-1 subunit (Vitis vinifera)</td>
<td>gi</td>
<td>255638009 gi</td>
<td>255541230 gi</td>
<td>255640620</td>
<td>26 kDa</td>
<td>[20]</td>
<td>[20] 3.226</td>
<td>Protein degradation</td>
<td>[20]</td>
</tr>
<tr>
<td>ATP synthase beta subunit [Dioscorea communis]</td>
<td>gi</td>
<td>17224743 gi</td>
<td>216848863 gi</td>
<td>377200445 gi</td>
<td>42559020</td>
<td>52 kDa</td>
<td>[26]</td>
<td>[26] 4.107</td>
<td>ATP-coupled proton transport</td>
</tr>
<tr>
<td>Rubisco-associated protein (Glycine max)</td>
<td>gi</td>
<td>12556532 gi</td>
<td>125563325 gi</td>
<td>12556334 gi</td>
<td>125564693 gi</td>
<td>12556356</td>
<td>50 kDa</td>
<td>[51]</td>
<td>[51] 2.111</td>
</tr>
<tr>
<td>Rubisco subunit binding protein alpha subunit, putative chloroplast precursor (Oryza sativa japonica)</td>
<td>gi</td>
<td>125563325 gi</td>
<td>12556334 gi</td>
<td>12556356 gi</td>
<td>12556376 gi</td>
<td>12556399 gi</td>
<td>125564693 gi</td>
<td>12556472 gi</td>
<td>50 kDa</td>
</tr>
<tr>
<td>Beta-form ribulose activase (Glycine max)</td>
<td>gi</td>
<td>255633215</td>
<td>49 kDa</td>
<td>[46]</td>
<td>[46] 1.60b</td>
<td>Calvin cycle</td>
<td>[46]</td>
<td>GLAYDAWAKDQDIYTR LLLYSMMVQEQVENNK LVPDTP SQSDFGGALR MCALFINDLDAGAGR MDDFYAPARKK MGIPIMMSGASEL ESRONAGEPAAK TDGPRQDIYK VPIRVTQNGDPFLYAPLIR VPIQGWGQK VYQDEVR WSGVQVDSVGK YLNEAEALGNANEDAQKR</td>
<td>93.4</td>
</tr>
</tbody>
</table>

93.4 | 101.0 | 51.7 | 55.7 | 53.8 | 114 | 62.0 | 85.9 | 60.5 | 57.4 | 66.6 | 93.1 |
<p>| Protein Name                                      | Accession   | Mw   | Description                        | gc | E value | Taxonomic ID | SeqID | 1st 20 AAs | 2nd 20 AAs | 3rd 20 AAs | 4th 20 AAs | 5th 20 AAs | 6th 20 AAs | 7th 20 AAs | 8th 20 AAs | 9th 20 AAs | 10th 20 AAs |
|--------------------------------------------------|-------------|------|------------------------------------|----|---------|--------------|-------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-------------|
| beta-form rubisco activase (Glycine max)         | gi|2529353324 | 49 kDa | Calvin cycle                      | 66 | 0.000   | GLAYDISDDQQDITR LLYGNMLVQEQENVK LVTTP0DS86FFSAUR MICALFINDLDAAGR MDSFYAPAPLMD NQIMPSMAGELESSNGAPEK VPIIVTGVDFSTLVPAPLR YVYDEVR WISGVDVSVGK YLNEAALGANGAEAGR | 96.4 | 103        | 91.1        | 76.8        | 76.3        | 97.1        | 96.2        | 56.3        | 66.2        | 96.2        |
| Elongation factor 1 (Glycine max)                | gi|252962739  | 25 kDa | translation                       | 45.6 | 56.8 | 72.5       | [22] | KLYDEYLLPR LDEYLLPR SVOMEGILWGASK | | | | | | | | | | | |
| Elongation factor 1u, chloroplastic (Glycine max)| gi|24944261   | 52 kDa | translation                       | 64.8 | 76.0 | 60.1       | [26] | ILDEALAGDNVLDAGAGR MDGFYIAPAFMDK MGINPIMMSAGELESGNAGEPAK | 96.2 | 61.3        | 56.3        | 59.9        | 58.6        | 76.1        | | | | | |
| phosphoglycerate kinase, putative (Ricinus communis) | gi|252954784  | 50 kDa | Calvin cycle/glycolysis           | [35] | 1.777 | [28] ELDYLYGAVSSPK FAIVTEALAK GVTI0GGDGSAVEK IGVIESLEK LDLATSLAK LGSELLGQWK RPFSAVGGSK | 65 | 64.1        | 85.1        | 71.1        | 62.1        | 61.3        | 50.0        | 61.3        | 70.1        | 50.0        |
| phosphoglycerate kinase, putative (Populus triocarpa) | gi|234109080  | 50 kDa | Calvin cycle/glycolysis           | [35] | 1.777 | [29] ADLVNPLDQNTDDTR FAVGEAALAK GVTI0GGDGSAVEK IGVIESLEK LGSELLGQWK LVASLPQGALLNVGK RPFSAVGGSK | 75.3 | 64.1        | 85.1        | 71.1        | 61.3        | 70.1        | 50.0        | 61.3        | 70.1        | 50.0        |
| eukaryotic ferritin (Glycine max)                 | gi|252963727  | 25 kDa | Callophagy/glycolysis             | 61.1 | 25.7 | 67.4       | [19] | GLAYDISDDQQDITR LLYGNMLVQEQENVK LVTTP0DS86FFSAUR MICALFINDLDAAGR MDSFYAPAPLMD NQIMPSMAGELESSNGAPEK VPIIVTGVDFSTLVPAPLR YVYDEVR WISGVDVSVGK YLNEAALGANGAEAGR | 65 | 64.1        | 85.1        | 71.1        | 62.1        | 61.3        | 50.0        | 61.3        | 70.1        | 50.0        |
| fructose-bisphosphate aldolase (Glycine max)      | gi|252975081  | 39 kDa | Calvin cycle/glycolysis           | [46] | 1.600 | [46] ANBEALTYK AGEALLVR GILADESTGTGK | 57.2 | 56.3        | 62.4        | | | | | | | |
| fructose-bisphosphate aldolase (Ricinus communis) | gi|252964773  | 39 kDa | Calvin cycle/glycolysis           | [46] | 1.600 | [46] ANBEALTYK AGEALLVR GILADESTGTGK | 57.2 | 56.3        | 62.4        | | | | | | | |
| fructose-bisphosphate aldolase (Glycine max)      | gi|404587367  | 38 kDa | Calvin cycle/glycolysis           | [46] | 1.600 | [46] ANBEALTYK AGEALLVR GILADESTGTGK | 57.2 | 56.3        | 62.4        | | | | | | | |
| fructose-bisphosphate aldolase (Glycine max)      | gi|404587367  | 38 kDa | Calvin cycle/glycolysis           | [46] | 1.600 | [46] ANBEALTYK AGEALLVR GILADESTGTGK | 57.2 | 56.3        | 62.4        | | | | | | | |
| inorganic pyrophosphatase (Glycine max)            | gi|252948233  | 32 kDa | ATP-coupled proton transport      | | | | [44] | APPVIN0DDOEVEK NNASLDDPK MEVA10DSFFTPK [25] | 64.5 | 62.3        | 52.7        | | | | | | | |
| Glycerinaldehyde 3-phosphate dehydrogenase (Glycine max) | gi|252963643  | 48 kDa | Calvin cycle/glycolysis           | [46] | 1.600 | [46] GLTAEDSTGTGK YVQGSGQGSLTESLYK | 62.1 | 84.4        | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th><strong>Protein</strong></th>
<th><strong>Gene ID</strong></th>
<th><strong>Molecular Weight</strong></th>
<th><strong>Accession Numbers</strong></th>
<th><strong>Description</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>RRM (RNA recognition motif), also known as RBD (RNA binding domain) or RNP (ribonucleoprotein domain) protein (Glycine max)</td>
<td>gi</td>
<td>255639723</td>
<td>30 kDa</td>
<td>gi</td>
</tr>
<tr>
<td>RRM (RNA recognition motif), also known as RBD (RNA binding domain) or RNP (ribonucleoprotein domain) protein (Glycine max)</td>
<td>gi</td>
<td>255636284</td>
<td>31 kDa</td>
<td>gi</td>
</tr>
<tr>
<td>RRM (RNA recognition motif), also known as RBD (RNA binding domain) or RNP (ribonucleoprotein domain) protein (Glycine max)</td>
<td>gi</td>
<td>255633513</td>
<td>27 kDa</td>
<td>gi</td>
</tr>
<tr>
<td>RRM (RNA recognition motif), also known as RBD (RNA binding domain) or RNP (ribonucleoprotein domain) protein (Glycine max)</td>
<td>gi</td>
<td>255634528</td>
<td>35 kDa</td>
<td>gi</td>
</tr>
<tr>
<td>RRM (RNA recognition motif), also known as RBD (RNA binding domain) or RNP (ribonucleoprotein domain) protein (Glycine max)</td>
<td>gi</td>
<td>255634791</td>
<td>49 kDa</td>
<td>gi</td>
</tr>
</tbody>
</table>

**Phosphatase/phosphohexomutase (Populus trichocarpa x Populus deltoides)**

**Plant basic secretory protein (Glycine max)**

**Plant basic secretory protein (Glycine max)**

**NADPH-specific isocitrate dehydrogenase (Glycine max)**

**1-deoxy-D-xylulose 5-phosphate reductoisomerase (Pueraria montana var. lobata)**
<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Gene ID(s)</th>
<th>SWATH (Da)</th>
<th>Description</th>
<th>MASCOT P-Values</th>
<th>Peptide Sequences</th>
<th>Function/Pathway</th>
<th>pI</th>
<th>MW</th>
<th>Z Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phosphoribulokinase, putative (Glycine max)</td>
<td>gi</td>
<td>2556</td>
<td>462/17</td>
<td>45 kDa</td>
<td>(35) 3.347, (25) 4.123</td>
<td>ANDFILMYEOLK DLYEOQASK FYGEYTOQMLK GVTALDPR KPDFEADPOQK LTSVFGGAAEPKK</td>
<td>Calvin cycle</td>
<td>76.2</td>
<td>49.5</td>
</tr>
<tr>
<td>Beta-1,3-glucanase 3 (Glycine max)</td>
<td>gi</td>
<td>2021</td>
<td>317</td>
<td>26 kDa</td>
<td>(22) 4.449</td>
<td>ANDFDLMYEQVK DLYEQLIASK FYGEVTQQMLK GVTALDPR KPDFEAYIDPQK LTSVFGGAAEPKK</td>
<td>Carbohydrate metabolism</td>
<td>68.9</td>
<td>71.8</td>
</tr>
<tr>
<td>Cytoxic glutamine synthetase Gb-1 (Glycine max)</td>
<td>gi</td>
<td>1034</td>
<td>465/7</td>
<td>39 kDa</td>
<td>(10) 5.718</td>
<td>EDGGYEVIK TLGPVSOPSEP</td>
<td>Amino acid metabolism</td>
<td>49.3</td>
<td>62</td>
</tr>
<tr>
<td>Cytosolic glutamine synthetase GSbeta1 (Glycine max)</td>
<td>gi</td>
<td>1094</td>
<td>6357gi</td>
<td>12136</td>
<td>39 kDa</td>
<td>(35) 3.347, (10) 5.718, (21) 4.513</td>
<td>EDGGYEVIK TLGPVSOPSELPK</td>
<td>Amino acid metabolism</td>
<td>52.3</td>
</tr>
<tr>
<td>Peptide methionine sulfoxide reductase (Glycine max)</td>
<td>gi</td>
<td>2556</td>
<td>31520</td>
<td>22 kDa</td>
<td>(45) 2.791</td>
<td>IVTEILPAK SGI3YNTQAR</td>
<td>Redox homeostasis</td>
<td>59.9</td>
<td>58.7</td>
</tr>
<tr>
<td>Cytosolic phosphatase synthetase L chain, ATP binding domain (Glycine max)</td>
<td>gi</td>
<td>2556</td>
<td>30652</td>
<td>19 kDa</td>
<td>(103) 1.240</td>
<td>GIASSVEER TQQVEAK</td>
<td>Amino acid metabolism</td>
<td>64.6</td>
<td>60.1</td>
</tr>
<tr>
<td>Glycolate oxidase (Vitis vinifera)</td>
<td>gi</td>
<td>147/184/80gi</td>
<td>167881875gi</td>
<td>188418927gi</td>
<td>225462096gi</td>
<td>205507255</td>
<td>41 kDa</td>
<td>(45) 1.054</td>
<td>AIALSTVDTPR GVLTAEDTR VPVFLDGQVR</td>
</tr>
<tr>
<td>Nucleotide binding protein Maf (Glycine max)</td>
<td>gi</td>
<td>2556</td>
<td>46212</td>
<td>22 kDa</td>
<td>(45) 2.791</td>
<td>ILGSSSAVAR LPVVDYLK</td>
<td>Nucleotide metabolism</td>
<td>58.7</td>
<td>52.7</td>
</tr>
<tr>
<td>Isopentenyl diphosphate (IPP) isomerase (Glycine max)</td>
<td>gi</td>
<td>2556</td>
<td>35611gi</td>
<td>255647086gi</td>
<td>203306565gi</td>
<td>88560554</td>
<td>27 kDa</td>
<td>(89) 1.667</td>
<td>ESELDENALGV YELQQQR</td>
</tr>
<tr>
<td>Protein Name</td>
<td>Accession</td>
<td>Molecular Weight (kDa)</td>
<td>Sequences (ID)</td>
<td>Function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td>------------------------</td>
<td>---------------</td>
<td>----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S-adenosylmethionine-dependent methyltransferase (Glycine max)</td>
<td>gi</td>
<td>12596395A</td>
<td>27 kDa</td>
<td>[([75])1.880]</td>
<td>unknown</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In2-1 protein (Glycine max)</td>
<td>gi</td>
<td>113855579</td>
<td>27 kDa</td>
<td>[([20])2.226]</td>
<td>redox homeostasis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOUL home binding protein (Glycine max)</td>
<td>gi</td>
<td>255627213</td>
<td>26 kDa</td>
<td>[([10])3.022]</td>
<td>metal trafficking</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative chalcone isomerase 4 (Glycine max)</td>
<td>gi</td>
<td>31039630</td>
<td>23 kDa</td>
<td>[([42])1.615]</td>
<td>secondary metabolism - flavonoid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chalcone isomerase A (Glycine max)</td>
<td>gi</td>
<td>114199181</td>
<td>23 kDa</td>
<td>[([19])2.226]</td>
<td>secondary metabolism - flavonoid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UBA-domain containing protein, Nascent polypeptide associated complex NAC (Medicago trunculata)</td>
<td>gi</td>
<td>217075454</td>
<td>22 kDa</td>
<td>[([59])1.401]</td>
<td>unknown</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GTP-binding protein (Cicer arietinum)</td>
<td>gi</td>
<td>10334503</td>
<td>25 kDa</td>
<td>[([59])1.401]</td>
<td>unknown</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chlorophyll a/b binding protein (Glycine max)</td>
<td>gi</td>
<td>1093216</td>
<td>28 kDa</td>
<td>[([42])1.615]</td>
<td>photosynthesis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dihydrosorbutate dehydrogenase (Cucumis sativus)</td>
<td>gi</td>
<td>118564</td>
<td>42 kDa</td>
<td>[([35])1.771]</td>
<td>carbon metabolism</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative aldol/keto reductase (Glycine max)</td>
<td>gi</td>
<td>224555575</td>
<td>36 kDa</td>
<td>[([40])1.655]</td>
<td>unknown</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>caffeoyl-CoA 3-O-methyltransferase (Amorpha fruticosa)</td>
<td>gi</td>
<td>146766245</td>
<td>25 kDa</td>
<td>[([13])2.805]</td>
<td>secondary metabolism - flavonoid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein Name</td>
<td>Accession</td>
<td>MW (kDa)</td>
<td>E-values</td>
<td>P-value</td>
<td>GO Terms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>-------------</td>
<td>----------</td>
<td>----------</td>
<td>---------</td>
<td>--------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspartate-semialdehyde dehydrogenase (Glycine max)</td>
<td>gi</td>
<td>255695398</td>
<td>38</td>
<td>3e-05</td>
<td>[G02200]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown (Glycine max)</td>
<td>gi</td>
<td>255644536</td>
<td>31</td>
<td>1e-08</td>
<td>[G02220]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphatidylinositol transfer protein (PG/PI-TP) (Glycine max)</td>
<td>gi</td>
<td>255627393</td>
<td>17</td>
<td>1e-01</td>
<td>[G02220]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycosyl hydrolase family 1, acidic chitinase (Glycine max)</td>
<td>gi</td>
<td>255642487</td>
<td>32</td>
<td>1e-01</td>
<td>[G02220]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ribulose-1,5-bisphosphate carboxylase small subunit precursor (Glycine soja)</td>
<td>gi</td>
<td>1079736</td>
<td>20</td>
<td>3e-01</td>
<td>[G02220]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stress-induced protein SAM22 (Glycine max)</td>
<td>gi</td>
<td>134194</td>
<td>17</td>
<td>4e-01</td>
<td>[G02220]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (Brachynamyces zeyheri)</td>
<td>gi</td>
<td>1045644</td>
<td>50</td>
<td>1e-01</td>
<td>[G02220]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**GO Terms**

- [G02200]: Amino acid biosynthesis
- [G02220]: Lipid related
- [G02220]: Calvin cycle
- [G02220]: Stress response
- [G02220]: Carbohydrate metabolism
- [G02220]: Calvin cycle
- [G02220]: Stress response
- [G02220]: Calvin cycle
- [G02220]: Stress response
- [G02220]: Calvin cycle
- [G02220]: Stress response
- [G02220]: Calvin cycle
- [G02220]: Stress response
- [G02220]: Calvin cycle
- [G02220]: Stress response
- [G02220]: Calvin cycle
- [G02220]: Stress response
- [G02220]: Calvin cycle
- [G02220]: Stress response
<table>
<thead>
<tr>
<th>Enzyme/Protein Family</th>
<th>Accession(s)</th>
<th>Molecular Weight</th>
<th>Homology Clusters</th>
<th>Enzymatic Function(s)</th>
<th>PDB Code(s)</th>
<th>Percentage Coverage</th>
<th>Amino Acid Similarity</th>
<th>Protein Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rubisco activase (Glycine max)</td>
<td>gi</td>
<td>250476648</td>
<td>49 kDa</td>
<td>(204/105</td>
<td>(103/145</td>
<td>(51/1.549)</td>
<td>Calvin cycle</td>
<td>GLAYDISDDQQDITRLTTYGMNLVQEDENVKYLNEAALGNAEDAINR</td>
</tr>
<tr>
<td>Peroxiredoxin-like (Glycine max)</td>
<td>gi</td>
<td>225684868</td>
<td>25 kDa</td>
<td>[4695/1.156]</td>
<td>[15/2.411]</td>
<td>[55/1.475]</td>
<td>Redox homeostasis</td>
<td>GLAYDISDDQQDITRLTTYGMNLVQEDENVKYLNEAALGNAEDAINR</td>
</tr>
<tr>
<td>Carbonic anhydrase (Glycine max)</td>
<td>gi</td>
<td>2300504307</td>
<td>24 kDa</td>
<td>[75/1.885]</td>
<td>[15/2.411]</td>
<td>[55/1.475]</td>
<td>Carbon metabolism</td>
<td>GLAYDISDDQQDITRLTTYGMNLVQEDENVKYLNEAALGNAEDAINR</td>
</tr>
<tr>
<td>Phosphoglycerate kinase, chloroplastic (Nicotiana tabacum)</td>
<td>gi</td>
<td>224994897</td>
<td>50 kDa</td>
<td>[103/8.718]</td>
<td>[25/2.070]</td>
<td>[35/1.771]</td>
<td>Calvin cycle/glycolysis</td>
<td>GLAYDISDDQQDITRLTTYGMNLVQEDENVKYLNEAALGNAEDAINR</td>
</tr>
<tr>
<td>CYPOR-like ferredoxin reductase (Populus trichocarpa)</td>
<td>gi</td>
<td>2224074327</td>
<td>40 kDa</td>
<td>[103/8.718]</td>
<td>[25/2.070]</td>
<td></td>
<td>Redox homeostasis</td>
<td>GLAYDISDDQQDITRLTTYGMNLVQEDENVKYLNEAALGNAEDAINR</td>
</tr>
<tr>
<td>Fructose-bisphosphate aldolase, chloroplastic (Glycine max)</td>
<td>gi</td>
<td>1164411</td>
<td>42 kDa</td>
<td>[592/8.76</td>
<td>105/6.718</td>
<td>[25/2.070]</td>
<td>Calvin cycle</td>
<td>GLAYDISDDQQDITRLTTYGMNLVQEDENVKYLNEAALGNAEDAINR</td>
</tr>
</tbody>
</table>

**Notes:**
- The table lists enzymes and proteins with their associated molecular weights and a variety of functions.
- Homology clusters are indicated by the number of clusters and their respective sizes.
- The table includes information on the percentage coverage, amino acid similarity, and protein descriptions associated with each enzyme or protein.
- Redox homeostasis and Calvin cycle/glycolysis are specific functions highlighted in the table.
<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Accession Numbers</th>
<th>M.W. (kDa)</th>
<th>PDB IDs</th>
<th>Thermostability (°C)</th>
<th>Function/Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glyceraldehyde-3-phosphate dehydrogenase A subunit (Glycine max)</td>
<td>gi</td>
<td>77540210</td>
<td>43</td>
<td>(35)3.347, (25)4.123, (10)5.718, (21)4.513</td>
<td>Calvin cycle/glycolysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pentapeptide repeats-containing protein (Medicago trunculata)</td>
<td>gi</td>
<td>217071606</td>
<td>26</td>
<td>(65)-1.289</td>
<td>Unknown</td>
</tr>
<tr>
<td>14-3-3 like protein, protein kinase C inhibitor homologue (Oenothera elata subsp. hookeri)</td>
<td>gi</td>
<td>217071606</td>
<td>26</td>
<td>(65)-1.289</td>
<td>Unknown</td>
</tr>
<tr>
<td>Glutathione-γ-transferase (Glycine max)</td>
<td>gi</td>
<td>255627415</td>
<td>23</td>
<td>(19)2.226, (20)2.226, (10)3.022</td>
<td>Redox homeostasis</td>
</tr>
<tr>
<td>Ascorbate peroxidase 2 (Glycine max)</td>
<td>gi</td>
<td>13360822</td>
<td>27</td>
<td>(13)2.206, (20)2.226</td>
<td>Redox homeostasis</td>
</tr>
<tr>
<td>Accession</td>
<td>Name</td>
<td>Description</td>
<td>Function</td>
<td>Identity</td>
<td>E-value</td>
</tr>
<tr>
<td>-----------</td>
<td>------</td>
<td>-------------</td>
<td>----------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>gi</td>
<td>255640161</td>
<td>Ribo5-phosphate isomerase type A (RPI_A) subfamily protein (Glycine max)</td>
<td>Calvin cycle</td>
<td>[36]1.271</td>
<td>(66)</td>
</tr>
<tr>
<td>gi</td>
<td>255638346</td>
<td>14-3-3 family protein (Vitis vinifera)</td>
<td>Protein degradation</td>
<td>[36]1.760</td>
<td>(36)</td>
</tr>
<tr>
<td>gi</td>
<td>255629882</td>
<td>Dienelactone hydrolase family protein, chlorophyllase-like (Glycine max)</td>
<td>Chlorophyll metabolism, photosynthesis</td>
<td>LQELFKEEGVEAK</td>
<td>(33)</td>
</tr>
<tr>
<td>gi</td>
<td>255628305</td>
<td>Polyketide cyclase 2 superfamily protein (Glycine max)</td>
<td>Protein degradation</td>
<td>[36]1.265</td>
<td>(36)</td>
</tr>
<tr>
<td>gi</td>
<td>12229897</td>
<td>Proteasome subunit alpha type-6 (Glycine max)</td>
<td>Protein degradation</td>
<td>[13]2.805</td>
<td>(13)</td>
</tr>
<tr>
<td>gi</td>
<td>5929964</td>
<td>Malate dehydrogenase (Glycine max)</td>
<td>TCA cycle</td>
<td>ALEGADVVIIPAGVPR</td>
<td>(10)</td>
</tr>
<tr>
<td>gi</td>
<td>10334493</td>
<td>Cytosolic malate dehydrogenase (Cicer arietinum)</td>
<td>TCA cycle</td>
<td>ALEGADVVIIPAGVPR</td>
<td>(25)</td>
</tr>
</tbody>
</table>

**Calvin cycle**

- **Signal transduction**
- **Chlorophyll metabolism, photosynthesis**
- **Secondary metabolism**
- **Protein degradation**
- **TCA cycle**
<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Accession</th>
<th>Mass (kDa)</th>
<th>N-terminal Peptide</th>
<th>C-terminal Peptide</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Triosephosphate isomerase (Glycine max)</td>
<td>gi</td>
<td>77540216</td>
<td>27</td>
<td>[a</td>
<td>2</td>
</tr>
<tr>
<td>Chaperonin 10 kDa subunit (cpn10 or GroES) (Glycine max)</td>
<td>gi</td>
<td>255645102</td>
<td>27</td>
<td>[a</td>
<td>2</td>
</tr>
<tr>
<td>Polyketide cyclase 2 superfamily protein (Glycine max)</td>
<td>gi</td>
<td>18643</td>
<td>17</td>
<td>[a</td>
<td>6</td>
</tr>
</tbody>
</table>
### Supporting Information Table 2. List of Peptides Identified from 115 ppb Root Proteins that are Differentially Expressed/Oxidized

<table>
<thead>
<tr>
<th>Protein</th>
<th>Accession Number(s)</th>
<th>Mol Weight</th>
<th>Fold Change IAF (spot number)</th>
<th>Fold Change SYPRO (spot number)</th>
<th>Classification</th>
<th>IAF Peptides (spot number)</th>
<th>SYPRO Peptides (spot number)</th>
<th>Best Mascot Ion Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIM phosphate binding protein (Glycine max)</td>
<td>gi</td>
<td>255634120</td>
<td>20 kDa</td>
<td>[40]-1.573</td>
<td>Calvin cycle/glycolysis</td>
<td>[EAGTTTAVVSGETK IVTTLEAK QLLNESNFYGTK]</td>
<td>64.6</td>
<td>64.8</td>
</tr>
<tr>
<td>polyketide cyclase 2 family protein (Glycine max)</td>
<td>gi</td>
<td>255631546</td>
<td>17 kDa</td>
<td>[47]-1.717</td>
<td>secondary metabolism</td>
<td>[ALVTDADNIIPK ITFVEDGETK SVENVENDGGGPTK]</td>
<td>71.8</td>
<td>63</td>
</tr>
<tr>
<td>polyketide cyclase 2 family protein (Glycine max)</td>
<td>gi</td>
<td>255627117</td>
<td>18 kDa</td>
<td>[13]-2.461</td>
<td>secondary metabolism</td>
<td>[FIFQAIDDNDHGGTIIK HWTYTIDGK LFSGDIDHNYK]</td>
<td>40.7</td>
<td>59.7</td>
</tr>
<tr>
<td>Rossmann-fold NAD(P)(+) binding protein, isoflavone reductase-like (Glycine max)</td>
<td>gi</td>
<td>255637531</td>
<td>34 kDa</td>
<td>[59]-1.544</td>
<td>secondary metabolism - flavonoid</td>
<td>[AGNPTIALVR VIILGDNPK]</td>
<td>60.3</td>
<td>63</td>
</tr>
<tr>
<td>similar to heat shock cognate 70 kDa protein 1 (Vitis vinifera)</td>
<td>gi</td>
<td>225434984</td>
<td>71 kDa</td>
<td>[44]-1.514</td>
<td>protein folding</td>
<td>[FSIDSSVSQDIIQIAMANPINTVQDART TPSYVATDTER VEIIANDQGNR]</td>
<td>69.1</td>
<td>83.1</td>
</tr>
<tr>
<td>chaperonin GroEL (Bradyrhizobium japonicum USDA 110)</td>
<td>gi</td>
<td>27377170</td>
<td>38 kDa</td>
<td>[89]-1.273</td>
<td>protein folding</td>
<td>[LAGGVAVIR LENVTLNMLGR SVAAGMNPMDLK VGGATEVEVK]</td>
<td>52.9</td>
<td>67.6</td>
</tr>
<tr>
<td>glutathione-S-transferase (Glycine max)</td>
<td>gi</td>
<td>225434986</td>
<td>24 kDa</td>
<td>[89]-1.563</td>
<td>redox homeostasis</td>
<td>[FSGAAAPAEAAPAK SIEMPGLLWGASK]</td>
<td>58.7</td>
<td>59.7</td>
</tr>
<tr>
<td>F1 ATPase (Pisum sativum)</td>
<td>gi</td>
<td>2116558</td>
<td>60 kDa</td>
<td>[64]-1.273</td>
<td>ATP-coupled proton transport</td>
<td>[ITDEFTGK TIAMDATEGVR]</td>
<td>58.4</td>
<td>71.4</td>
</tr>
<tr>
<td>F0-F1 ATP synthase subunit beta [Rhodospseudomonas palustris B5a3]</td>
<td>gi</td>
<td>115022038</td>
<td>51 kDa</td>
<td>[64]-1.273</td>
<td>ATP-coupled proton transport</td>
<td>[FLOQGVSVEALLGR FYVLADTQK VVDAPAK]</td>
<td>68.8</td>
<td>56.3</td>
</tr>
<tr>
<td>ferric leghemoglobin reductase-2 precursor (Glycine max)</td>
<td>gi</td>
<td>32009269</td>
<td>53 kDa</td>
<td>[70]-1.250</td>
<td>noduline metabolism</td>
<td>[ADNACGLYK VVGSQTSQDQVK VVSSTGALTEIPK]</td>
<td>55.8</td>
<td>74.9</td>
</tr>
<tr>
<td>Protein Name</td>
<td>Accession Numbers</td>
<td>Mass (kDa)</td>
<td>PDB ID(s)</td>
<td>Function/Property</td>
<td>Protein Folding score</td>
<td>IVFELYADVTPR</td>
<td>LFVQYKTVQMGQPGQR</td>
<td>LLVEFENARVGAAETELED</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>-------------------------------------------</td>
<td>------------</td>
<td>-----------</td>
<td>------------------------------------------------------</td>
<td>-----------------------</td>
<td>-------------</td>
<td>----------------</td>
<td>---------------------</td>
</tr>
<tr>
<td><strong>cyclophilin (Phaseolus vulgaris)</strong></td>
<td>gi</td>
<td>145047/256</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
<td>gi</td>
<td>829119</td>
</tr>
<tr>
<td>** RuBisCO large subunit-binding protein subunit alpha (Brassica napus)**</td>
<td>gi</td>
<td>12030/30</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>254047/900</td>
</tr>
<tr>
<td>** dirigent-like protein (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>** plant basic secretory protein (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>** ascorbate peroxidase (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>** NADPH-dependent FMN reductase (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>** NADPH-dependent FMN reductase (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>** oxoacid oxidase (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>** oxoacid oxidase (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>** oxoacid oxidase (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>** oxoacid oxidase (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>** oxoacid oxidase (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>** oxoacid oxidase (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>** oxoacid oxidase (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>** oxoacid oxidase (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>** oxoacid oxidase (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>** oxoacid oxidase (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>** oxoacid oxidase (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>** oxoacid oxidase (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>** oxoacid oxidase (Glycine max)**</td>
<td>gi</td>
<td>25965/2917</td>
<td>gi</td>
<td>4647/27</td>
<td>gi</td>
<td>254047/900</td>
<td>gi</td>
<td>208970/456</td>
</tr>
<tr>
<td>Peptidase M17 (Oryza sativa japonica)</td>
<td>33 kDa</td>
<td>[641]</td>
<td>Proteolysis</td>
<td>[641]</td>
<td>FDGMGSAAVPAK TIEVNTDAEGR</td>
<td>5.7 64.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADP-Ribosylation Factor 1 Complexed With GDP, Full Length Non-Myristoylated (Homo sapiens)</td>
<td>21 kDa</td>
<td>[62]</td>
<td>Vesicle transport</td>
<td>[62]</td>
<td>ILMVGLDAAGK QDLPNAMNAEITDK</td>
<td>5.8 5.8 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyruvate dehydrogenase (acetyl-transferring)</td>
<td>39 kDa</td>
<td>[16]</td>
<td>TCA cycle</td>
<td>[16]</td>
<td>EGISAEVINLR VLSpySSeA</td>
<td>5.8 5.1 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol dehydrogenase 1 (Pharus sativum)</td>
<td>41 kDa</td>
<td>[16]</td>
<td>Carbon metabolism</td>
<td>[16]</td>
<td>FGVNEVFNPK IGGVLVSSR</td>
<td>5.9 7.2 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Molecular chaperone DnaK, provisional (Vitis vinifera)</td>
<td>74 kDa</td>
<td>[66]</td>
<td>Protein folding</td>
<td>[66]</td>
<td>FGVNEVFNPK IGGVLVSSR</td>
<td>5.9 7.2 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein Name</td>
<td>Accession Numbers</td>
<td>Molecular Weight (Da)</td>
<td>Features</td>
<td>Protein Degradation</td>
<td>Metal Trafficking</td>
<td>Nucleotide Metabolism</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------</td>
<td>-----------------------</td>
<td>----------</td>
<td>---------------------</td>
<td>------------------</td>
<td>----------------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Proteasome subunit alpha type-5 (Glycine max)</td>
<td>gi</td>
<td>122</td>
<td>234653, gi</td>
<td>121707</td>
<td>13364, gi</td>
<td>121707</td>
<td>2540, gi</td>
<td>1224</td>
</tr>
<tr>
<td>Copper chaperone homolog CCH (Glycine max)</td>
<td>gi</td>
<td>6</td>
<td>256</td>
<td>011</td>
<td>14 kDa</td>
<td>[62]-1.308</td>
<td>AIGSGSEGADSSLQEYNYK GVNTSPEGR ITSPLEPSSVEK</td>
<td>135</td>
</tr>
<tr>
<td>Copper chaperone homolog CCH (Glycine max)</td>
<td>gi</td>
<td>6</td>
<td>256</td>
<td>011</td>
<td>14 kDa</td>
<td>[62]-1.308</td>
<td>AIGSGSEGADSSLQEYNYK GVNTSPEGR ITSPLEPSSVEK</td>
<td>135</td>
</tr>
<tr>
<td>Eukaryotic UDPase (Glycine max)</td>
<td>gi</td>
<td>25556</td>
<td>35072</td>
<td>51 kDa</td>
<td>[64]-1.273</td>
<td>AIGSGSEGADSSLQEYNYK GVNTSPEGR ITSPLEPSSVEK</td>
<td>135</td>
<td>62.9</td>
</tr>
<tr>
<td>Cystathionine beta-synthase domain-containing protein (Glycine max)</td>
<td>gi</td>
<td>25556</td>
<td>31750</td>
<td>23 kDa</td>
<td>[2]-4.512</td>
<td>AIGSGSEGADSSLQEYNYK GVNTSPEGR ITSPLEPSSVEK</td>
<td>135</td>
<td>62.9</td>
</tr>
<tr>
<td>Cyclophilin (Glycine max)</td>
<td>gi</td>
<td>17</td>
<td>981</td>
<td>911, gi</td>
<td>255</td>
<td>6</td>
<td>28137</td>
<td>18 kDa</td>
</tr>
<tr>
<td>YgiF YER057c UK14 family</td>
<td>gi</td>
<td>25562</td>
<td>9976, gi</td>
<td>25556</td>
<td>32812</td>
<td>20 kDa</td>
<td>[39]-1.600</td>
<td>AIGSGSEGADSSLQEYNYK GVNTSPEGR ITSPLEPSSVEK</td>
</tr>
<tr>
<td>Protein Name</td>
<td>Accession Numbers</td>
<td>Molecular Weight (kDa)</td>
<td>Description</td>
<td>Sequence (Peptide)</td>
<td>Mw (kDa)</td>
<td>PSSM Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------</td>
<td>------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>--------------------</td>
<td>----------</td>
<td>------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP synthase subunit alpha, mitochondrial (Helianthus annuus)</td>
<td>gi</td>
<td>114404 gi</td>
<td>114411 gi</td>
<td>1245015 gi</td>
<td>124365668 gi</td>
<td>221142 gi</td>
<td>231385 gi</td>
<td>24054933 gi</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ATP-coupled proton transport</td>
<td>AEELTTLESRR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14-3-3-like protein, protein kinase C inhibitor homologue (Oenothera elata subsp. hookeri)</td>
<td>gi</td>
<td>1168195 gi</td>
<td>11778152 gi</td>
<td>125641583 gi</td>
<td>256454007 gi</td>
<td>2879818 gi</td>
<td>3912948 gi</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>signal transduction</td>
<td>DSTLIMQLR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>glutathione-s-transferase (Glycine max)</td>
<td>gi</td>
<td>255607415 gi</td>
<td>255640466 gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>redox homeostasis</td>
<td>VPVLFDGK VTAVDSLAPK</td>
<td>57.4</td>
<td>6/7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dirigent-like protein (Glycine max)</td>
<td>gi</td>
<td>255639867 gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>secondary metabolism - lignin</td>
<td>AQGLFGLASLEDR IIEPSASEVR</td>
<td>82.3</td>
<td>6/6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Universal stress protein family protein (Glycine max)</td>
<td>gi</td>
<td>255626103 gi</td>
<td>255628920 gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>stress response</td>
<td>EKNILQLK QVQTRMK VVQIVDTRF</td>
<td>50.7</td>
<td>5/5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>polyketide cyclase 2 superfamily protein (Glycine max)</td>
<td>gi</td>
<td>255628305 gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>secondary metabolism</td>
<td>AVLDSASVFFPK</td>
<td>66.4</td>
<td>5/5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pro teaseosome subunit alpha type-6 (Glycine max)</td>
<td>gi</td>
<td>12523889 gi</td>
<td>1256645041 gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>protein degradation</td>
<td>AAAGITSDOR</td>
<td>68.1</td>
<td>5/5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein Name</td>
<td>Accession Numbers</td>
<td>Molecular Weight (kDa)</td>
<td>pI Value</td>
<td>TCA Cycle Peptide Sequence</td>
<td>pI Value</td>
<td>TCA Cycle Peptide Sequence</td>
<td>Molecular Weight (kDa)</td>
<td>pI Value</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>--------------------</td>
<td>------------------------</td>
<td>----------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>----------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Malate Dehydrogenase (Glycine max)</td>
<td>gi</td>
<td>5065964</td>
<td>36</td>
<td>ALEGADVVIIPAGVPR SEVGYQGDEELGK</td>
<td>78.8</td>
<td>ALEGADVVIIPAGVPR SEVGYQGDEELGK</td>
<td>78.8</td>
<td></td>
</tr>
<tr>
<td>Cytosolic Malate Dehydrogenase (Cicer arietinum)</td>
<td>gi</td>
<td>10334493 gi</td>
<td>12170732486 gi</td>
<td>27462762 gi</td>
<td>27462764 gi</td>
<td>276969277 gi</td>
<td>283839565</td>
<td>35</td>
</tr>
<tr>
<td>Polyketide Cyclase 2 Superfamily Protein (Glycine max)</td>
<td>gi</td>
<td>186643 gi</td>
<td>255649567</td>
<td>17</td>
<td>ALEGADVVIIPAGVPR SEVGYQGDEELGK</td>
<td>78.8</td>
<td>ALEGADVVIIPAGVPR SEVGYQGDEELGK</td>
<td>78.8</td>
</tr>
<tr>
<td>Protein</td>
<td>Accession Number(s)</td>
<td>Mol Weight</td>
<td>Fold Change IAF (spot number)</td>
<td>Fold Change SYPRO (spot number)</td>
<td>Classification</td>
<td>IAF Peptides (spot number)</td>
<td>SYPRO Peptides (spot number)</td>
<td>Best Mascot Score IAF</td>
</tr>
<tr>
<td>---------</td>
<td>---------------------</td>
<td>------------</td>
<td>-----------------------------</td>
<td>-------------------------------</td>
<td>-----------------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>polyketide cyclase 2 superfamily protein (Glycine max)</td>
<td>gi</td>
<td>2556</td>
<td>30540</td>
<td>17 kDa</td>
<td>(70)-2.467</td>
<td>secondary metabolism</td>
<td>LTFVDEGGQGTK</td>
<td>50.7</td>
</tr>
<tr>
<td>Malate dehydrogenase (NADP)+ chloroplastic (Pisum sativum)</td>
<td>gi</td>
<td>4625</td>
<td>79</td>
<td>48 kDa</td>
<td>(128)-1.547</td>
<td>TCA cycle</td>
<td>LASGEVFQPDQIALK</td>
<td>69.6</td>
</tr>
<tr>
<td>ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (Ostryopsis davidiana) peroxidoxin (PRX) family protein, 2-Cys (Glycine max)</td>
<td>gi</td>
<td>5731</td>
<td>964</td>
<td>52 kDa</td>
<td>(50)-1.520</td>
<td>Calvin cycle</td>
<td>ILDEALAGDNVGLLLR</td>
<td>57.1</td>
</tr>
<tr>
<td>translation elongation factor-TU (Glycine max)</td>
<td>gi</td>
<td>2556</td>
<td>41400</td>
<td>22 kDa</td>
<td>(50)-1.650</td>
<td>redox homeostasis</td>
<td>AADNIAANLYSVK</td>
<td>71.8</td>
</tr>
<tr>
<td>plastid high chlorophyll fluorescence 136 precursor (Zea mays)</td>
<td>gi</td>
<td>1482</td>
<td>51</td>
<td>625</td>
<td>43 kDa</td>
<td>(54)-1.407</td>
<td>photosynthesis</td>
<td>ILDEALAGDNVGLLYLVR</td>
</tr>
<tr>
<td>thiamine pyrophosphate (TPP) family protein (Populus trichocarpa)</td>
<td>gi</td>
<td>2254</td>
<td>03766</td>
<td>81 kDa</td>
<td>(127)-1.575</td>
<td>co-factor metabolism</td>
<td>ALPTYTPESPADATRFLAIDAVEK</td>
<td>46.1</td>
</tr>
</tbody>
</table>

**Supporting Information Table 3**: List of Peptides Identified from 60 ppb Leaf Proteins that are Differentially Expressed/Oxidized.
<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Full Name</th>
<th>Accession Numbers</th>
<th>Mass (kDa)</th>
<th>Protein Description</th>
<th>PSMs</th>
<th>P-values</th>
<th>q-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>23 kDa OEC protein</td>
<td>(Salicornia veneta)</td>
<td>gi</td>
<td>148535011</td>
<td>22</td>
<td>photosynthesis</td>
<td>239</td>
<td>2.599</td>
</tr>
<tr>
<td>glutamine synthetase precursor</td>
<td>(Glycine max)</td>
<td>gi</td>
<td>131977011</td>
<td>48</td>
<td>amino acid metabolism</td>
<td>128</td>
<td>1.574</td>
</tr>
<tr>
<td>ribulose 1,5-bisphosphate carboxylase small subunit precursor</td>
<td>(Glycine soja)</td>
<td>gi</td>
<td>10797396</td>
<td>20</td>
<td>Calvin cycle</td>
<td>37</td>
<td>1.845</td>
</tr>
<tr>
<td>stress-induced protein SAM22</td>
<td>(Glycine max)</td>
<td>gi</td>
<td>134194</td>
<td>17</td>
<td>stress response</td>
<td>30</td>
<td>1.984</td>
</tr>
<tr>
<td>ribulose-1,5-bisphosphatase carboxylase/oxygenase large subunit</td>
<td>(Orothamnus zeyheri)</td>
<td>gi</td>
<td>104564.4 (+1547)</td>
<td>50</td>
<td>Calvin cycle</td>
<td>61</td>
<td>1.421</td>
</tr>
<tr>
<td>rubisco activase</td>
<td>(Glycine max)</td>
<td>gi</td>
<td>2525604089</td>
<td>25</td>
<td>Calvin cycle</td>
<td>79</td>
<td>2.467</td>
</tr>
<tr>
<td>peroxiredoxin 5-like</td>
<td>(Glycine max)</td>
<td>gi</td>
<td>255630307</td>
<td>24</td>
<td>carbon metabolism</td>
<td>88</td>
<td>3.113</td>
</tr>
<tr>
<td>Protein Name</td>
<td>Accession Numbers</td>
<td>Molecular Weight</td>
<td>Predicted Peptide(s)</td>
<td>Predicted Protein(s)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------------------</td>
<td>----------------------------------------</td>
<td>------------------</td>
<td>----------------------</td>
<td>----------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphoglycerate kinase, chloroplastic (Nicotiana tabacum)</td>
<td>gi</td>
<td>2499497</td>
<td>50 kDa</td>
<td>[([128])-1.574],[([133])-1.513]</td>
<td>Calvin cycle/glycolysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYPOR-like ferridoxin reductase (Populus trichocarpa)</td>
<td>gi</td>
<td>224074257</td>
<td>40 kDa</td>
<td>[([64])-1.407]</td>
<td>Redox homeostasis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fructose-bisphosphate aldolase, chloroplastic (Glycine max)</td>
<td>gi</td>
<td>216701671, gi</td>
<td>217070476, gi</td>
<td>8446841</td>
<td>42 kDa</td>
<td>[([110])-1.289]</td>
<td>Calvin cycle/glycolysis</td>
</tr>
<tr>
<td>glyceraldehyde-3-phosphate dehydrogenase A subunit (Glycine max)</td>
<td>gi</td>
<td>73040210</td>
<td>43 kDa</td>
<td>[([112])-1.911],[([202])-2.207]</td>
<td>Calvin cycle/glycolysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pentapeptide repeats-containing protein (Medicago truncatula)/translationally controlled tumor protein (Elaeis guineensis)</td>
<td>gi</td>
<td>192910898, gi</td>
<td>192912974, gi</td>
<td>201404883, gi</td>
<td>792250269</td>
<td>26 kDa</td>
<td>[([79])-2.487]</td>
</tr>
<tr>
<td>triosephosphate isomerase (Glycine max)*</td>
<td>gi</td>
<td>77540216</td>
<td>27 kDa</td>
<td>[([29])-1.624],[([3])-3.313],[([601])-2.120]</td>
<td>Calvin cycle/glycolysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTP synthase (GroES) subunit (cpn10 or GroEL) (Glycine max)*</td>
<td>gi</td>
<td>255646010C</td>
<td>27 kDa</td>
<td>[([79])-2.955]</td>
<td>Protein folding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chaperonin 10 Kd subunit (cpn10 or GroEL) (Glycine max)*</td>
<td>gi</td>
<td>255646010C</td>
<td>27 kDa</td>
<td>[([79])-2.955]</td>
<td>Protein folding</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **gi**: Gene identifier
- **kDa**: Kilodaltons
- **Calvin cycle/glycolysis**: Indicates the protein is involved in Calvin cycle/glycolysis pathway
- **Redox homeostasis**: Indicates the protein is involved in redox homeostasis
- **Unknown**: Indicates the protein function is unknown
- **Cell expansion**: Indicates the protein is involved in cell expansion
- **Protein folding**: Indicates the protein is involved in protein folding
<p>| <strong>polyketide cyclase 2</strong> superclass protein (Glycine max) | gi|186143 | gi|255640867 | gi|1672888 | secondary metabolism | [T6/G2.888] ALVDADNVIPK AVEYLLANPHYN ITFVEDGESK | 66.4 | 46.3 | 51.7 |</p>
<table>
<thead>
<tr>
<th>Protein</th>
<th>Accession Number(s)</th>
<th>Mol Weight</th>
<th>Fold Change IAF (spot number)</th>
<th>Fold Change SYPRO (spot number)</th>
<th>Pathway</th>
<th>IAF Peptides (spot number)</th>
<th>Best Mascot Ion Score</th>
<th>SYPRO Peptides (spot number)</th>
<th>Best Mascot Ion Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Triosephosphate isomerase (Glycine max)</td>
<td>gi</td>
<td>255645035</td>
<td>33 kDa</td>
<td>[(44)-1.750]</td>
<td>[(53)-1.778]</td>
<td>Calvin cycle/glycolysis</td>
<td>[44] GQAFTGDESAEQLK GPEFAVTNVSYSK IEISAGNNSWIGK IYGGSGVNSGAELAK LVQDNLAK</td>
<td>75.2</td>
<td>[130] AEPVDMAATVK EAGTTTAVAYEQTK GLINESNWFYDQK VAYAQLQKL</td>
</tr>
<tr>
<td>Putative fructokinase 2 (Vitis vinifera)</td>
<td>gi</td>
<td>2225433918</td>
<td>35 kDa</td>
<td>[(58)-1.223]</td>
<td></td>
<td>Carbohydrate metabolism</td>
<td>[58] LPLWPSAEAR TALAFYTLR</td>
<td>70.6</td>
<td>[36] LPLWPSAEAR TALAFYTLR</td>
</tr>
<tr>
<td>Nucleoside diphosphate kinase 1 (Glycine max)</td>
<td>gi</td>
<td>2049078 gi</td>
<td>2045395</td>
<td>16 kDa</td>
<td>[(4)-1.556]</td>
<td>Nucleotide metabolism</td>
<td>[39] GLIGEIISR KGATNPQSEPQGTR</td>
<td>64.2</td>
<td>[34] GLIGEIISR KGATNPQSEPQGTR</td>
</tr>
<tr>
<td>Protein Name</td>
<td>GI Accession</td>
<td>M.Wt</td>
<td>ECIS</td>
<td>ECD</td>
<td>Protein Function</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>--------------</td>
<td>------</td>
<td>------</td>
<td>-----</td>
<td>------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transaldolase-like protein (Glycine max)</td>
<td>gi</td>
<td>255640604</td>
<td>48 kDa</td>
<td>[(46)-1.736],[(31)-2.191],[(33)-2.034]</td>
<td>Secorynary metabolism</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell autonomous heat shock cognate protein 70 (Cucurbita maxima)</td>
<td>gi</td>
<td>26985223</td>
<td>71 kDa</td>
<td>[(60)-1.204],[(57)-1.283]</td>
<td>Protein folding</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14-3-3-like protein B, SGF14B (Glycine max)</td>
<td>gi</td>
<td>2693215b</td>
<td>28 kDa</td>
<td>[(46)-1.736]</td>
<td>Signal transduction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UMAH class glutathione transferase DHAR2 (Glycine max)</td>
<td>gi</td>
<td>25067210</td>
<td>23 kDa</td>
<td>[50]-1.514</td>
<td>Redox homeostasis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proteasome beta type-6 subunit (Vitis vinifera)</td>
<td>gi</td>
<td>255429890</td>
<td>25 kDa</td>
<td>[(47)-1.696]</td>
<td>Protein degradation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proteasome alpha type 4 subunit (Glycine max)</td>
<td>gi</td>
<td>255637272</td>
<td>28 kDa</td>
<td>[(51)-1.473],[(33)-1.776]</td>
<td>Protein degradation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein Name</td>
<td>Accession</td>
<td>Molecular Weight</td>
<td>Reference Mass</td>
<td>Function</td>
<td>Description</td>
<td>Mass</td>
<td>FDR</td>
<td>Reference</td>
<td>Protein Name</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td>------------------</td>
<td>---------------</td>
<td>-----------</td>
<td>-------------</td>
<td>-------</td>
<td>-----</td>
<td>-----------</td>
<td>--------------</td>
</tr>
<tr>
<td>S-adenosyl-L-homocysteine hydrolase (Petunia x hybrida)</td>
<td>gi</td>
<td>32967697</td>
<td>53 kDa</td>
<td>[(57)-1.485],[(34)-2.014]</td>
<td>amino acid biosynthesis</td>
<td>70.1</td>
<td>49.2</td>
<td>[(57)] LVQVSEETTTVK VAVAGYGDVGK</td>
<td>protein disulfide isomerase (Glycine max)</td>
</tr>
<tr>
<td></td>
<td>gi</td>
<td>49023142</td>
<td>37 kDa</td>
<td>[(57)-1.485]</td>
<td>redox homeostasis</td>
<td>67</td>
<td>65.2</td>
<td>[(57)] GLUVNQIK MNGVLTGSQGEIR</td>
<td></td>
</tr>
<tr>
<td>Aspartate aminotransferase P2, mitochondrial</td>
<td>gi</td>
<td>112379</td>
<td>50 kDa</td>
<td>[(23)-1.998]</td>
<td>amino acid biosynthesis</td>
<td>68.3</td>
<td>52.5</td>
<td>[(23)]-1.998</td>
<td>Aspartate aminotransferase P2, mitochondrial</td>
</tr>
<tr>
<td>chain A, mutant beta-amylase (W55r) (Glycine max)</td>
<td>gi</td>
<td>149241163</td>
<td>56 kDa</td>
<td>[(49)-1.676],[(34)-2.014]</td>
<td>carbohydrate metabolism</td>
<td>57.5</td>
<td>76.5</td>
<td>[(49)] LSDDLLQK VAGENALPR</td>
<td>chain A, mutant beta-amylase (W55r) (Glycine max)</td>
</tr>
<tr>
<td></td>
<td>gi</td>
<td>125744990</td>
<td>54 kDa</td>
<td>[(23)-1.998]</td>
<td>ATP-coupled proton transport</td>
<td>74.9</td>
<td>56.3</td>
<td>[(23)-1.998]</td>
<td>leghemoglobin C1; nodulin-50 (Glycine max)</td>
</tr>
<tr>
<td></td>
<td>gi</td>
<td>255631161</td>
<td>20 kDa</td>
<td>[(56)-1.490]</td>
<td>redox homeostasis</td>
<td>56.9</td>
<td>55</td>
<td>[(56)-1.490]</td>
<td>TRX family protein, possibly m-type (Glycine max)</td>
</tr>
<tr>
<td>vacuolar ATPase subunit B1 (Triticum aestivum)</td>
<td>gi</td>
<td>125744990</td>
<td>54 kDa</td>
<td>[(23)-1.998]</td>
<td>ATP-coupled proton transport</td>
<td>74.9</td>
<td>56.3</td>
<td>[(23)-1.998]</td>
<td>vacuolar ATPase subunit B1 (Triticum aestivum)</td>
</tr>
<tr>
<td>leghemoglobin C1; nodulin-50 (Glycine max)</td>
<td>gi</td>
<td>126236</td>
<td>15 kDa</td>
<td>[(62)-1.385]</td>
<td>nodule metabolism</td>
<td>53.6</td>
<td>55.3</td>
<td>[(62)]-1.385</td>
<td>leghemoglobin C1; nodulin-50 (Glycine max)</td>
</tr>
<tr>
<td></td>
<td>gi</td>
<td>255631161</td>
<td>20 kDa</td>
<td>[(56)-1.490]</td>
<td>redox homeostasis</td>
<td>56.9</td>
<td>55</td>
<td>[(56)-1.490]</td>
<td>TRX family protein, possibly m-type (Glycine max)</td>
</tr>
<tr>
<td>Protein Description</td>
<td>Accession Numbers</td>
<td>Molecular Weight</td>
<td>GO Terms</td>
<td>Secondary Structure</td>
<td>Other Information</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>-------------------</td>
<td>------------------</td>
<td>----------------------------------------------</td>
<td>---------------------</td>
<td>-------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyridine nucleotide-disulphide oxidoreductase (Glycine max)</td>
<td>gi</td>
<td>255638280</td>
<td>44 kDa</td>
<td>[67]-1.276</td>
<td>redox homeostasis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>actin (Oryza sativa japonica)</td>
<td>gi</td>
<td>115484337, gi</td>
<td>168472715, gi</td>
<td>217072944, gi</td>
<td>255448323, gi</td>
<td>296068567, gi</td>
<td>9565319</td>
<td>42 kDa</td>
<td>[56]-1.338</td>
</tr>
<tr>
<td>Secretory peroxidase (Glycine max)</td>
<td>gi</td>
<td>255638280</td>
<td>38 kDa</td>
<td>[57]-1.267</td>
<td>redox homeostasis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glo EDI BHR-like 8 family (Glycine max)</td>
<td>gi</td>
<td>255632802</td>
<td>17 kDa</td>
<td>[50]-1.531</td>
<td>unknown</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unknown (Glycine max)</td>
<td>gi</td>
<td>255647164</td>
<td>12 kDa</td>
<td>[59]-1.404</td>
<td>unknown</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphoglyceromutase, cytoplasmic (Psammos sahnam)</td>
<td>gi</td>
<td>129857319, gi</td>
<td>125855330, gi</td>
<td>12643355, gi</td>
<td>152056688, gi</td>
<td>152122326, gi</td>
<td>1240000852, gi</td>
<td>224114175, gi</td>
<td>2384795134, gi</td>
</tr>
<tr>
<td>predicted translation initiation factor 26 subunit</td>
<td>gi</td>
<td>255647295</td>
<td>23 kDa</td>
<td>[67]-1.276</td>
<td>translation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>glycolate II</td>
<td>gi</td>
<td>255637721</td>
<td>32 kDa</td>
<td>[59]-1.404</td>
<td>carbon metabolism</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Alcohol Dehydrogenase 1

*Pisum sativum*

<table>
<thead>
<tr>
<th>Accession Numbers</th>
<th>Molecular Mass (kDa)</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi</td>
<td>113361</td>
<td>gi</td>
</tr>
</tbody>
</table>

## Molecular Chaperone DnaK, Provisional

*Vitis vinifera*

<table>
<thead>
<tr>
<th>Accession Numbers</th>
<th>Molecular Mass (kDa)</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi</td>
<td>118489843</td>
<td>74 kDa</td>
</tr>
</tbody>
</table>

## Proteasome Subunit Alpha Type-5

*Glycine max*

<table>
<thead>
<tr>
<th>Accession Numbers</th>
<th>Molecular Mass (kDa)</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi</td>
<td>12229923</td>
<td>gi</td>
</tr>
</tbody>
</table>

## Copper Chaperone Homolog QDH

*Glycine max*

<table>
<thead>
<tr>
<th>Accession Numbers</th>
<th>Molecular Mass (kDa)</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi</td>
<td>6252011</td>
<td>14 kDa</td>
</tr>
</tbody>
</table>

## Eukaryotic UDPase

*Glycine max*

<table>
<thead>
<tr>
<th>Accession Numbers</th>
<th>Molecular Mass (kDa)</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi</td>
<td>256330072</td>
<td>51 kDa</td>
</tr>
</tbody>
</table>

## Cystathionine Beta-Synthase Domain-Containing Protein

*Glycine max*

<table>
<thead>
<tr>
<th>Accession Numbers</th>
<th>Molecular Mass (kDa)</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi</td>
<td>256331760</td>
<td>23 kDa</td>
</tr>
<tr>
<td>Name</td>
<td>Accession Numbers</td>
<td>Molecular Mass (kDa)</td>
</tr>
<tr>
<td>------</td>
<td>-------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Cyclophilin (Glycine max)</td>
<td>gi</td>
<td>125561449</td>
</tr>
<tr>
<td>YF2 YEN057c UK114 family</td>
<td>gi</td>
<td>125565498</td>
</tr>
<tr>
<td>ATP synthase subunit alpha, mitochondrial (Helianthus annuus)</td>
<td>gi</td>
<td>11144044</td>
</tr>
<tr>
<td>Ascorbate peroxidase 2 (Glycine max)</td>
<td>gi</td>
<td>113900375</td>
</tr>
<tr>
<td>Ribose 5-phosphate isomerase type A (RPI_A) subfamily protein (Glycine max)</td>
<td>gi</td>
<td>2555640161</td>
</tr>
<tr>
<td>14-3-3 family protein (Vitis vinifera)</td>
<td>gi</td>
<td>2255410495</td>
</tr>
<tr>
<td>Dione lactone hydrolase family protein, chlorophyllase-like (Glycine max)</td>
<td>gi</td>
<td>255565231</td>
</tr>
<tr>
<td>Protein Name</td>
<td>GI/Enzyme ID</td>
<td>Molecular Weight</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>--------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Polyketide cyclase 2 superfamily protein (Glycine max)</td>
<td>gi</td>
<td>255670505</td>
</tr>
<tr>
<td>Proteasome subunit alpha type-6 (Glycine max)</td>
<td>gi</td>
<td>12229897</td>
</tr>
<tr>
<td>Malate dehydrogenase (Glycine max)</td>
<td>gi</td>
<td>15669984</td>
</tr>
<tr>
<td>Cytosolic malate dehydrogenase (Cicer arietinum)</td>
<td>gi</td>
<td>10334493</td>
</tr>
<tr>
<td>Protein Name</td>
<td>Accession</td>
<td>MW (kDa)</td>
</tr>
<tr>
<td>----------------------------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>Triosephosphate isomerase</td>
<td>gi</td>
<td>7540219</td>
</tr>
<tr>
<td></td>
<td>AEFVDIINAATVK</td>
<td>55.7</td>
</tr>
<tr>
<td></td>
<td>EAGTTTAVAEOTK</td>
<td>49.3</td>
</tr>
<tr>
<td></td>
<td>IVTTLNNEAK</td>
<td>70.8</td>
</tr>
<tr>
<td></td>
<td>QLIESNPVGDK</td>
<td>84.1</td>
</tr>
<tr>
<td></td>
<td>DGSDYITLR</td>
<td>52.9</td>
</tr>
<tr>
<td></td>
<td>GKDSDYITLR</td>
<td>77.6</td>
</tr>
<tr>
<td></td>
<td>KPLSYPQNYLVSK</td>
<td>80.6</td>
</tr>
<tr>
<td></td>
<td>TSGQGLTEATK</td>
<td>61.2</td>
</tr>
<tr>
<td></td>
<td>VATGEVDFDGTK</td>
<td>55.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Accession</th>
<th>MW (kDa)</th>
<th>pI Values</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chaperonin 10 Kd subunit</td>
<td>gi</td>
<td>205540102</td>
<td>27</td>
<td>Protein folding</td>
</tr>
<tr>
<td></td>
<td>GKDSDYITLR</td>
<td>58</td>
<td>58.3</td>
<td>58.3</td>
</tr>
<tr>
<td></td>
<td>KPLSYPQNYLVSK</td>
<td>77.6</td>
<td>77.6</td>
<td>77.6</td>
</tr>
<tr>
<td></td>
<td>TSGQGLTEATK</td>
<td>80.6</td>
<td>80.6</td>
<td>80.6</td>
</tr>
<tr>
<td></td>
<td>VATGEVDFDGTK</td>
<td>61.2</td>
<td>61.2</td>
<td>61.2</td>
</tr>
</tbody>
</table>

MW = Molecular Weight; pI = isoelectric point
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 homoglutathione 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\textsubscript{2}O\textsubscript{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.
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,
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 alkylation 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$_2$O$_2$ 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$_2$O$_2$ 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$_2$O$_2$ 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$_2$O$_2$ 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$_2$O$_2$/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 occurring 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 phosphoglycerate 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 an 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 is 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$_2$O$_2$ into the cell apoplast as a result of even short term exposure [Rao and Davis, 2001]. H$_2$O$_2$ is able to diffuse into cells through oxidative-gated aquaporins, and can function as an intercellular signal for activating the hypersensitive 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\textsuperscript{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\textsuperscript{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\textsuperscript{13} could used in place of C\textsuperscript{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 electrons 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 stomata, 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 peroxidase - 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 apoplastic 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 transcript 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 promotor, 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 due 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 CO2 specificity via directed evolution, as well as through the formation of hybrid cross-species holozymes; however, it appears that CO2 specificity is closely tied to O2 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 thermostable 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 CO₂ 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


Overmyer, K., Brosche, M., Pellinen, R., Kuittinen, T., Tuominen, H., Ahlfors, R.,
death in the Arabidopsis radical-incused cell death 1 mutant. Plant Physiology 137:
1092-1104.
Manipulation of Rubisco: the amount, activity, function and regulation. Journal of
Experimental Botany 54: 1321-1333.
of quercetin and its seven derivatives on the growth of Arabidopsis thaliana and
Arabidopsis Book.
glutathione synthetase deficiency and a rare gene permutation event. EMBO 18:
3204-3213.
Photosynthesis Research 75: 11-27.
682-690.


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 He⁶, Ashley Galant⁷, Qiuying Pang⁶, Johanna M. Strul⁸, Sherifat F. Balogun⁸, Joseph M Jez⁹, Sixue Chen*¹

¹Department of Biology, Genetics Institute, Plant Molecular and Cellular Biology Program, University of Florida, Gainesville, FL 32610, USA.

²Department 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 AIPMDH2 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.
\textbf{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 \textit{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 Å2 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-oxopentanoate)
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 $k_{\text{cat}}/K_m$ 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 $k_{\text{cat}}/K_m$ 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 $k_{\text{cat}}/K_m$ 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 AtPMDH 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 AtIPMDH1pro::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 ATPMDH2 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 ATPMDH2 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 ATPMDH (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


Escherichia coli and Salmonella typhimurium 3-isopropylmalate dehydrogenase and comparison
with their thermophilic counterpart from Thermus thermophilus. J Mol Biol 266: 1016-1031.

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.

ion-dependent pyridine dinucleotide-linked ω-hydroxyacid oxidative decarboxylases. Biochemistry
48, 3565-3577.

high-resolution structure of LeuB (Rv2995c) from Mycobacterium tuberculosis. J Mol Biol 346: 1-
11.


Cyanogenic glycosides: a case study for evolution and application of cytochromes P450.

JA, Halkier BA (2003) CYP79F1 and CYP79F2 have distinct functions in the biosynthesis of


27. Galant A, Arkus KAJ, Zubieta C, Cahoon RE, Jez JM (2009) Structural basis for evolution of


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-oxopentaoate 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 AtlPMDH 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 AtlPMDH1-F137L, AtlPMDH2-L133F and AtlPMDH3-L134F driven by AtlPMDH1 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

glucosinolate content (μmol per g seeds)

wild-type
allpmdh
AllPMDH1
AllPMDH2-L33F
AllPMDH3-L134F

B

glucosinolate content (μmol per g leaves)

c3 c4 c5 c6 c7 c8 indole
Table 1. Kinetic parameters of wild-type and mutant AtIPMDHs.

<table>
<thead>
<tr>
<th></th>
<th>3-isopropylmalate</th>
<th>3-(2′-methylthio)malate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k_{cat}$ (min⁻¹)</td>
<td>$K_m$ (µM)</td>
</tr>
<tr>
<td>AtIPMDH1</td>
<td>37 ± 4</td>
<td>25.2 ± 2.3</td>
</tr>
<tr>
<td>AtIPMDH1-F137L</td>
<td>230 ± 14</td>
<td>11.4 ± 1.7</td>
</tr>
<tr>
<td>AtIPMDH2</td>
<td>373 ± 33</td>
<td>10.9 ± 1.3</td>
</tr>
<tr>
<td>AtIPMDH2-L133F</td>
<td>37 ± 5</td>
<td>30.3 ± 2.5</td>
</tr>
<tr>
<td>AtIPMDH3</td>
<td>543 ± 36</td>
<td>9.2 ± 1.4</td>
</tr>
<tr>
<td>AtIPMDH3-L134F</td>
<td>44 ± 5</td>
<td>28.5 ± 1.5</td>
</tr>
</tbody>
</table>

Data are means ± standard error (n = 3).