An Analysis of Thioredoxins *h* in the Grasses

by

Juan Juttner

A thesis submitted for the Degree of Doctor of Philosophy in the Faculty of Sciences at the University of Adelaide

Department of Plant Science
Waite Agricultural Research Institute
University of Adelaide

December 2002

ERRATA

Page 16 – 1st paragraph, 4th line
replace "Ecklund 1991" with "Eklund 1991"

Page 16 – 2nd paragraph, 10th line
replace "Val et al. 1999" with "del Val et al. 1999"

Page 19 – 4th paragraph, 2nd line
replace "Figure 1.2B" with "Figure 1.2b"

Page 42 – 2nd paragraph, 9th line
"dependent"

Page 62 – 1st paragraph, 9th line
replace "Lalio et al. 2001" with "Laloi et al. 2001"

Page 63 – 1st paragraph, 3rd line
"dependent"

Page 77 – Figure 4.6
"thioredoxin"

Page 79 – 2nd paragraph, 11th line and page 80 – 1st paragraph, 10th line
replace "Jacqout et al. 1994" with "Jacquot et al. 1994"

Page 84 – 4th paragraph, 2nd line
"Whatman"

Page 90 – 2nd paragraph, 2nd line
"patterns"

Page 93 – 1st paragraph, 2nd line
"expression profile is"

Page 115 - 1st paragraph, 7th line

replace "only three proteins from *Populus balsamifera* (Pb1), *Thellungiella salsuginea* (Ts1) and a hybrid poplar (Ph1) containing this form" with "only five proteins from *Populus balsamifera* (Pb1), *Thellungiella salsuginea* (Ts1), *Zea mays* (Zm2), *Sorghum bicolor* (Sb2) and a hybrid poplar (Ph1) containing this form"

Page 115 – 2nd paragraph, 7th line

"data are"

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma in any University and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference is made in the text.

I give consent to this copy of my thesis, when deposited in the University library, being available for loan and photocopying.

Juan Juttner

December 2002

Acknowledgements

I wish to thank my principal supervisor, Prof Peter Langridge for his support, advice and constructive criticism during the course of this study. I am particularly indebted to my co-supervisor Dr Ute Baumann for her encouragement, direction, considerable patience and friendship during all stages of this research. I also thank Dr Jason Able and Dr Klaus Oldach for reading the manuscript and providing helpful suggestions.

Many thanks to all members of the Langridge laboratory and for their assistance, advice and good humour. Special thanks to Ms Jan Neild for always knowing what and where it is, to Ms Lisa Pietsch for administrative assistance and to Mr Lachlan Tailby for rescuing me from several computing nightmares. Thanks also to my friends and family for support and keeping me sane.

Finally, and most importantly, I would like to thank my partner Rachel for traversing this rocky road with me and providing tireless support, encouragement and understanding.

This work was supported by a Grains Research and Development Corporation Junior Research Fellowship.

List of Abbreviations

aa	amino acid
ATP	adenosine 5'-triphosphate
BLAST	Basic Local Alignment Research Tool
bp	base pair
BSA	bovine serum albumin
°C	degree centigrade
cDNA	complementary deoxyribonucleic acid
Ci	curie
Da	dalton
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
DEPC	diethylpyrocarbonate
dGTP	deoxyguanosine triphosphate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleoside triphosphate
DTT	dithiothreitol
dTTP	deoxythymidine triphosphate
EDTA	ethylenediaminetetraacetic acid
g	gram
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid
hr	hour
IPTG	isopropyl-1-thio-b-D-galactoside
kb	kilobase
kDa	kiloDalton
mA	milliampere
mg	milligram/s
min	minute/s
µg	microgram/s
ml	milliliter/s

µl	microliter/s
M	molar
MOPS	3-(N-morpholino)propane-sulfonic acid
mRNA	messenger ribonucleic acid
NbS_2	5,5'-dithio-bis-(2-nitrobenzoic acid)
nm	nanometer
OD_{260}	optical density at nm
oligo(dT)	oligodeoxythymidylic acid
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
Pers. comm.	personal communication
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
PIPES	piperazine-N,N'-bis(2-ethane-sulfonic acid)
poly(A)	polyadenylic acid
RFLP	restriction-fragment-length polymorphism
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute
Sarkosyl	N-lauroylsarcosine
SDS	sodium dodecyl sulfate
SSC	sodium chloride/sodium citrate buffer
TAE	tris/acetate buffer
Taq	Thermus aquaticus DNA polymerase
TBS	tris-buffered saline
TE	tris/EDTA buffer
Tris-Cl	tris(hydroxymethyl)aminomethane hydrochloride
UV	ultraviolet
v/v	volume/volume
w/v	weight/volume
X-gal	5-bromo-4-chloro-3-indolyl-b-D-galactoside

Summary

The reductive environment of a cell is created by the movement of electrons from oxidisable organic molecules to oxygen and is maintained by thiol-disulfide redox systems such as the thioredoxin system. Thioredoxin systems are most complex in plants with five multigenic thioredoxin classes (*h, m, f, o* and x) participating in the redox reactions of at least three distinct organellar thioredoxin systems (chloroplastic, cytosolic and mitochondrial). With the exception of chloroplastic thioredoxins *m* and *f*, few of the putative functions ascribed to thioredoxins in plants have been verified. This is particularly the case within the largest class, the *h* class. One of the fundamental problems underlying functional analyses of thioredoxins in plants is a lack of an inclusive set of genomic and biochemical data of all thioredoxins for one plant species. The identification and characterisation of thioredoxins *h* in this thesis generated a comprehensive profile of this class and provides a platform for future functional analyses of *h*-class thioredoxins in cereals and other plants.

The study commenced with an investigation of the structure of *Bm2*, the putative pollen *S*-gene of *Phalaris coerulescens*. *Bm2* was found not to represent the male self-incompatibility determinant but rather encoded a novel and highly conserved thioredoxin of the *h*-class that was labelled thioredoxin *n*. EST-database screening and interspecific sequence comparisons revealed that thioredoxins *n* are present in a diverse range of plant species representing angiosperms and gymnosperms and form a distinct subclass of biochemically active thioredoxins.

To gain further insight into the thioredoxin *h*-class, a comprehensive molecular profile of all thioredoxins *h* in the grasses was compiled using public EST databases as a source of gene sequences. These analyses demonstrated that grasses encode four thioredoxins *h* each representing a phylogenetically distinct subclass. The involvement of each *h* subclass in the stress response of grasses was addressed by examining the expression profile of subclass representatives in wheat plants subjected to a range of oxidative stresses. The results showed that

under oxidative stress the transcript levels of all wheat thioredoxins *h* were altered. Differences in the expression profile of each mRNA suggest that the biochemical processes incorporating each thioredoxin were affected in a different manner and are indicative of each thioredoxin having a separate function. Possible stress-related functions for thioredoxins of each subclass are discussed.

To investigate how grass thioredoxins might be transcriptionally modulated under stress an analysis of the promoters of two thioredoxins *h* in rice was performed. Putative stress-related transcriptional response elements were located upstream of both genes. These findings are discussed in relation to the potential role each element could play in the stress-responsive modulation of two thioredoxin subclasses.

The presence of distinct thioredoxin *h* subclasses in the grasses raised questions regarding evolutionary relationships within the thioredoxin *h*-class of all plants. Phylogenetic analyses were performed using all plant thioredoxin *h* sequences present in the public database. This study found that the *h* class of angiosperms is comprised of three monophyletic subclasses of which the *n*-subclass is the oldest and most highly conserved. An examination of the topological profile of individual species and whole gene families indicated that the subclasses are likely to have arisen early in angiosperm evolution. The genomic processes involved in thioredoxin evolution and the chronology of evolutionary events are considered.

Table of Contents

Chapter 1 Literature review

1.1 Introduction

This thesis deals with two relatively distinct topics, those of grass self-incompatibility and plant thioredoxins that marginally overlap. Consequently, the following review is divided into two sections. The first deals with self-incompatibility and the molecular aspects of the various self-incompatibility systems studied to date. The second section covers the current state of knowledge regarding plant thioredoxins. This section is the more comprehensively reviewed since the greater proportion of research reported herein relates to plant thioredoxins.

1.2 Self-Incompatibility

Fertilization in flowering plants is a complex series of events beginning with the hydration and germination of pollen on the stigma surface and growth of the pollen tube down the stylar transmitting tract. Numerous mechanisms have evolved that allow flowering plants to select suitable pollen from the genetically diverse mix that may alight on the stigma. These mechanisms include barriers to interspecific pollination and intraspecific barriers, such as self-incompatibility (SI). SI is a process that enables the pistil of a plant to distinguish between genetically identical (self) and genetically distinct (non-self) pollen and results in the rejection of self but not non-self pollen. Consequently, SI fosters genetic diversity through the prevention of inbreeding.

The molecular components of SI are encoded by a multiallelic locus, called the *S*-locus. In the simplest, and best-studied systems, SI is governed by a single *S*-locus and occurs when the *S*-allele carried by the haploid pollen is identical to one of the *S*-alleles present in the stigma upon which it lands. However, more complex systems involving several loci, such as the two-locus system in grasses (Lundqvist 1954; Hayman 1956) and four-locus system in *Beta vulgaris* (Larsen 1977) have also been identified. Genetic studies of single loci systems have shown that SI behaves in a simple Mendelian manner indicating that the male and female *S*-determinants reside at a common locus or are very closely linked (Bateman 1955).

SI systems are categorised as two broad groups (Figure 1.1): gametophytic systems, such as those present in the plant families *Solanaceae, Rosaceae, Papaveraceae, Scrophulariaceae* and *Poaceae*, in which a pollen grain's own genotype determines its incompatibility phenotype; and sporophytic systems, such as found in the *Brassicaceae*, in which the genotype of the pollen producing parent governs the incompatibility phenotype of the resulting pollen. Within these broad categories the molecular components of SI in different families can vary greatly, and are summarised in the following sections.

1.2.1. Self-Incompatibility in the *Brassicaceae*

The sporophytic SI system of the *Brassicaceae* was the first to be investigated at the molecular level and represents the only system for which both the stylar and pollen components have been identified. SI in Brassica is a complex mechanism controlled by several closely linked, multiallelic and highly polymorphic genes. Consequently, the title '*S*-haplotype' rather than *S*-locus has been applied to the region at which these genes reside.

The first SI genes to be determined were two controlling the *S*-haplotype specificity of the pistil, the *S*-locus glycoprotein (SLG) (Nasrallah et al. 1987) and the *S*-receptor kinase (SRK) (Stein et al. 1991). The SLG gene encodes a secreted glycoprotein of approximately 400 amino acids that is targeted to the stigmatic papillar cell wall. SRK encodes a transmembrane protein kinase with serine/threonine kinase activity (Goring and Rothstein, 1992; Stein and Nasrallah, 1993) that is located in the stigmatic papillar plasma membrane. The extracellular domain of SRK shows extensive amino acid sequence identity with SLG of the same haplotype and is commonly called the S-domain (Watanabe et al. 1994). Several lines of evidence have been presented that indicate SLG and SRK represent stigma *S* genes. Firstly, both genes show *S*-haplotype-specific restriction fragment length polymorphisms (Chen and Nasrallah, 1990; Stein et al. 1991). Secondly, the temporal and spatial expression patterns of SLG and SRK are also consistent with a role in SI. The genes are expressed predominantly in stigma papillar cells and expression correlates with the onset of SI (Nasrallah et al. 1985; Nasrallah et al. 1988; Kumar and Trick 1994; Delorme et al. 1995). Thirdly, SLG

and SRK are closely linked having been demonstrated to be only 20 and 25 kb apart in *Brassica napus* and *Brassica campestris* respectively (Yu et al.1996; Boyes et al. 1997). Lastly, plants with mutated SRK genes, and those in which SLG transcripts are downregulated, fail to reject self-pollen (Nasrallah et al. 1992; Goring et al. 1993; Nasrallah et al. 1994; Shiba et al. 1995).

Until recently, attempts to introduce new SLG or SRK specificities to plants by transformation have been unsuccessful because the transgene induces homology-dependant gene silencing of the endogenous SLG and SRK (Conner et al. 1997). However, analysis of these transformants has shown that only the stigma phenotype is affected, demonstrating that the pollen component is encoded by a separate gene (Stahl et al. 1998). Takasaki et al. (2000) have since introduced SLG and SRK alleles into plants with S-haplotypes containing SLG/SRK's of low homology to the transgene. Significantly, the results of this study revealed that SRK alone is sufficient to determine *S*-haplotype specificity in the pistil, although SLG appeared to enhance the strength of the SI reaction conferred by the SRK transgene. Dixit et al. (2000) have provided evidence suggesting that SRK is inherently unstable and that the role of SLG may be to facilitate the accumulation of SRK to biologically required levels in the stigma.

Recently, two groups have reported the cloning of identical genes encoding the pollen component of Brassica SI. The two genes called, SCR (*S*-locus cysteine rich) and SP11 (*S*-locus protein 11) were both located by physical mapping and genomic sequencing in the *S*-haplotype region of *Brassica campestris* (Suzuki et al. 1999; Schopfer et al. 1999) and display features consistent with the pollen determinant: linkage to the *S*-haplotype, allelic polymorphism and expression in anthers or pollen (Takayama et al. 2000). SCR/SP11 encode small secreted proteins of approximately 60 amino acids that contain eight cysteine residues at conserved positions. Consistent with the sporophytic nature of Brassica SI, Takayama et al. (2000) demonstrated, by *in situ* hybridisation, that SP11 mRNA is present in both sporophytic (tapetal cells) and gametophytic (microspores) tissues. Loss-of-function and gain-of-function transformation experiments with SCR alleles have confirmed that the product of the SCR gene is capable of determining pollen SI specificity (Schopfer et al. 1999).

Based on the existing molecular data the current model for SI in the *Brassicaceae* proposes that, upon contact of the pollen grain with the stigma surface, SCR proteins are released from the pollen coat and taken into the wall of the stigma papillar cells. If an SCR protein encounters the extracellular domain of an SRK protein of the same haplotype, it binds and sets off a signal transduction cascade that leads to pollen rejection (Kao and McCubbin, 2000).

Screening of a yeast two-hybrid library with the kinase domain of SRK has identified a putative substrate for SRK and potentially the first component of the signal transduction cascade that leads to pollen tube arrest (Gu et al. 1998). The protein, named ARC1 (Arm Repeat Containing), was found to interact, in a phosphorylation dependant manner, specifically with SRK's but not with other plant receptor-like kinases. Significantly, antisense suppression of ARC1 messenger RNA levels, in a transformed incompatible *Brassica napus* line, resulted in partial self-compatibility (Stone et al. 1999).

1.2.2 Self-Incompatibility in *Solanaceae, Rosaceae* and *Scrophulariaceae*

Studies into the molecular process of gametophytic SI in species derived from three plant families, the *Solanaceae, Rosaceae* and *Scrophulariaceae*, have identified an identical stigma component and suggest a common mechanism.

The stigma component was first identified in *Nicotiana alata*, with the demonstration of major style proteins that co-segregated with specific pollen rejection phenotypes (Bredemeijer and Blass, 1981). The proteins were found to be basic glycoproteins of approximately 30 kDa exhibiting characteristics consistent with a role in SI. In summary; the proteins were located to the region of the stylar transmitting tract that coincided with the zone of pollen tube inhibition, they accumulated rapidly during the transition of the developing flower from self-compatible to self-incompatible and inhibited pollen tube growth, *in vitro*, in a *S*-dependant manner (Cornish et al. 1987; Jahnen et al. 1989).

The discovery of *S* proteins led to the cloning and sequencing of corresponding cDNAs, initially in *Nicotiana* (Anderson et al. 1986) and later in species representing all families (reviewed in McCubbin and Kao, 1999). Analysis of the *S*-alleles revealed that conserved regions of all sequences shared similarity with fungal RNases (Kawata et al. 1988). *S*-proteins derived from several species have since been shown to possess RNase activity (McClure et al. 1989; Singh and Kao, 1991; Sassa et al. 1992; Broothaerts et al. 1995). Evidence for the involvement of ribonuclease activity in the SI response was provided by McClure et al. (1990) and Gray et al. (1991) when they demonstrated that pollen RNA is degraded in self-incompatible, but not self-compatible, pollinations and that intact *S*-RNases are capable of entering and inhibiting growing pollen tubes *in vitro*. Consequently, *S*-proteins are now ascribed the title of *S*-RNases.

Plant transformation studies have confirmed that *S*-RNases have both a recognition function and biochemical involvement in SI. In *Nicotiana, Solanum* and *Petunia* transformation with a cloned *S*-RNase gene was sufficient to introduce a new pollen rejection specificity (Murfett et al. 1994; Matton et al. 1997; Lee et al. 1994). In addition, transformation of *Petunia inflata* with a S_3 allele, containing a mutation of the catalytic His-93 residue, rendered the transgenic plant incapable of rejecting S_3 pollen.

It is likely that the pollen and pistil functions of SI in *S*-RNase based systems are controlled by distinct closely linked genes such as has been determined for the *Brassica* system. Evidence for this has come from the analysis of a self-compatible cultivar of *Pyrus serotina* (*Rosaceae*) (Sassa et al. 1997). In this cultivar, self-compatibility was due to the deletion of the S_4-RNase gene, however, this deletion affected only pistil, but not pollen, function in SI interactions.

Two models have been proposed to explain how stigma *S*-RNases and pollen *S*-products might interact to cause an SI response (Kao and McCubbin, 1996). One model postulates that pollen *S*-alleles encode membrane-bound receptors, located in the wall of pollen tubes that only allow self *S*-RNases to enter. Once inside the pollen tube the ribonucleases degrade RNA leading to the arrest of pollen tube growth. The alternative model suggests that all *S*-RNases enter the pollen tube

where they encounter the product of the pollen *S*-gene, a ribonuclease inhibitor molecule. In this model only non-self *S*-RNases are inhibited whereas self *S*-RNases are able to degrade RNA and inhibit growth. To date, no receptor-like protein of pollen has been shown to interact with *S*-RNases nor has any RNase inhibitory activity been reported in the pollen.

1.2.3 Self-Incompatibility in the *Papaveraceae*

Genetically, the SI system in the field poppy (*Papaver rhoeas*) is identical to other single locus gametophytic systems (Lawrence et al. 1978), yet *Papaver* displays self-incompatible behaviour that is physiologically and morphologically distinct. In contrast to the previously mentioned gametophytic species, inhibition of pollen tube growth in *Papaver* occurs during or immediately after germination on the stigma surface (Franklin-Tong and Franklin, 1992), indicating a different mechanism for SI.

Molecular studies have confirmed that SI in *Papaver* is mediated by a mechanism unlike that of other single locus systems. The stigma component of SI is a basic glycoprotein of approximately 17 kDa (Franklin-Tong et al. 1989) that does not possess RNase features or display RNase activity (Franklin-Tong et al. 1991; Foote et al. 1994). Several *S* alleles have been expressed in *E. coli* and the recombinant proteins purified. When applied to germinating pollen, *in vitro,* the recombinant *S*-proteins inhibited pollen tubes in an *S*-allele-specific manner, confirming that they represent the stigmatic *S*-locus component (Franklin-Tong et al. 1988).

Franklin-Tong et al. (1990) demonstrated that *de novo* transcription of *Papaver* pollen genes occurred as a result of an incompatible reaction, and that the expression of these genes was a prerequisite for SI since transcription inhibitors alleviated the SI response. It has since been shown that pollen gene expression during an SI reaction is correlated with a transient increase in the level of cytosolic free calcium in the nuclei region of the pollen tube (Franklin-Tong et al. 1993). The SI phenotype has also been correlated with an increase in the phosphorylation state of two proteins, p26 and p68 (Rudd et al. 1996, 1997). p26 is composed of two phosphoproteins, p26.1 and p26.2, and was isolated from

soluble and microsomal protein fractions while p68 represents a single protein and was isolated from pollen tubes. Significantly, phosphorylation of p26.1 was Ca^{2+} and calmodulin dependent suggesting a Ca^{2+}-dependant protein kinase may be involved in SI signalling. By contrast, phosphorylation of p68 was Ca^{2+}-independent, indicating that both Ca^{2+}-dependent and Ca^{2+}-independent signal transduction pathways are involved in the pollen SI response in *Papaver*.

The actin cytoskeleton of pollen tubes has been proposed as a target for the signalling pathways involved in the SI response (Franklin-Tong 1999a; 1999b). Geitmann et al. (2000) recently demonstrated that the actin cytoskeleton of *Papaver* pollen tubes is dramatically rearranged in incompatible pollinations. The pattern of rearrangement was found to be SI specific as compatible pollen tubes displayed an unchanged actin configuration. Moreover, the rearrangements associated with SI were not an effect of the consequent arrest of growth since stopping tube growth by chemical means did not duplicate the SI-induced alterations.

SI signalling has also been implicated in post inhibition processes. Jordan et al. (2000) provided evidence that stigma *S*-proteins trigger DNA fragmentation in incompatible pollen tubes. Interestingly, increases in the levels of free Ca^{2+} also resulted in DNA fragmentation, implicating the Ca^{2+} signalling pathway in this process.

1.2.4 Allele Number and Sequence Diversity at the S-locus

Genetic models (Wright 1939; Paxman 1963; Yokoyama and Hetherington, 1982) and theoretical analyses (Takahata 1990; Vekemans and Slatkin, 1994) of self-incompatibility have predicted that, due to the action of frequency dependent selection, a greater number of more diverse and much older, alleles will be maintained at a self-incompatibility locus than would be found at a strictly neutral locus. Consistent with these predictions, large numbers of *S*-alleles have been recorded in natural populations of self-incompatible plant species (Table 1.1). In addition, *S*-RNases from the *Solanaceae* (Richman et al. 1996) and *Rosaceae* (Ushijima et al. 1998) and SRK/SLG's from the *Brassicaceae* (Boyes et al. 1997) commonly present higher levels of interspecific sequence identity than that found

between alleles from the same species, a feature consistent with the proposition that *S*-alleles are of ancient origin. The corollary of long coalescence is that S alleles display a greater level of sequence diversity than alleles present at a neutral locus. Sequence comparisons of *S*-alleles has revealed amino acid sequence similarities of between 29.2%-82.2% for rosaceous *S*-RNases (Ushijima et al. 1998), 65%-80% for the Brassica SRK's and SLG's (Kusaba et al. 1997) and 30%-74% for Brassica SP11/SCR's (Watanabe et al. 2000).

Sequence variation between *S*-alleles is often highest in defined, highly variable regions. The S-RNases are composed of five highly conserved regions (C1 to C5) and two highly variable regions (Hva and Hvb) (Ioerger et al. 1991; Tsai et al. 1992). Similarly, the extracellular S-domain region of the Brassica SRK contains three hypervariable regions. Transformation studies incorporating mutated *S*-RNase alleles have demonstrated that such hypervariable regions are important for allelic recognition (Matton et al. 1997; Matton et al. 1999). Matton et al. (1999) replaced amino acids in the Hva and Hvb regions of a *Solanum chacoense* S_{13} allele. Three substitutions generated an allele with dual S_{13}, S_{11} specificity, whilst four substitutions changed the specificity from S_{13} to S_{11}.

Table 1.1 Summary of the number of *S*-alleles recorded in populations of self- incompatible species

Species	Number of *S*-alleles
Gametophytic SI	
Papaver rhoeas	34-42
Phlox drummondii	45
Trifolium repens	>36
Trifolium pratensis	>35
Oenothera organensis	45
Physalis crassifolia	44
Solanum carolinense	14
Sporophytic SI	
Brassica campestris	20-30
Iberis amara	>22
Raphanus raphanistrum	>22
Sinapsis arvensis	52

After Richman and Kohn (1996)

1.2.5 Self-Incompatibility in the *Poaceae*

SI is widespread in the grass family *Poaceae*, having been identified in genera representing most subfamilies (Connor 1979; Hayman 1992). The most detailed studies have centred on species representing the subfamily *Pooideae*, and have identified self-incompatible species in all tribes (for a summary see Baumann et al. 2000). Typically, both self-incompatible and self-compatible species are found in the same genus with the frequency of SI being highest in perennial species (Beddows 1931).

1.2.5.1 The Genetics of SI in the *Poaceae*

The genetics of SI in the grasses has been most comprehensively studied in *Secale cereale* (Lundqvist 1954) and *Phalaris coerulescens* (Hayman 1956). In both species SI is gametophytically controlled by two unlinked, multiallelic genes, *S* and *Z*. A pollen grain elicits an incompatible response when both the S and Z alleles it carries are also present in the stigma upon which it lands (Figure 1.1).

It is likely all grasses share the same two locus SI system since eight species representing 3 tribes have been demonstrated, genetically, to contain a two locus SI system (Hayman 1992). Linkage studies have also provided evidence to support this conclusion. For example, the isozyme phosphoglucoisomerase PGI-2 has been shown to be linked to the S-locus in six self-incompatible species (Cornish et al. 1980; Leach and Hayman, 1987).

1.2.5.2 Molecular Studies into SI in the *Poaceae*: *Phalaris coerulescens*

The only *Poaceous* species for which SI has been investigated at the molecular level is the mediterranean grass *Phalaris coerulescens*. Li et al. (1994) used differential screening to identify a putative *S* gene from a pollen cDNA library of the grass *Phalaris coerulescens*. The cDNA clone, named *Bm2*, displayed features consistent with a pollen SI gene: it detected an RFLP that co-segregated with the *S* genotype in a population of 120 plants and was strongly expressed in pollen. Since *Bm2* was not full length, the structure of the gene was predicted from the genomic sequence of two alleles, S_1 and S_2. In contrast to *S*-alleles in other systems, the deduced amino acid sequences of S_1 and S_2 were found to be very similar, sharing

92% identity, with all variation between the sequences located in the amino-terminal region. All variation was restricted to the predicted second exon with most attributable to a frameshift caused by a single base deletion and insertion. The carboxyl half of the predicted proteins were completely conserved and shared homology with plant thioredoxin h proteins. The thioredoxin-like region was subsequently expressed in *E. coli* and found to possess thioredoxin activity (Li et al. 1995). Thus, the *Bm2* gene was reported as being composed of two domains, an N-terminal allele-specificity-determining domain and a C-terminal catalytic thioredoxin domain. However, unlike the S-RNases of the *Solanaceous* gametophytic SI systems sequence variation was not contained in identifiable hypervariable domains.

Interestingly, *in situ* hybridisation experiments have revealed that a self-incompatible pollination induces considerable *Bm2* expression at the site of pollen-stigma contact while a self-compatible pollination does not (Baumann et al. 2000). However, since *Bm2* is highly expressed in pollen and inhibited pollen tubes regularly rupture, it is possible that the *Bm2* mRNA detected came from the pollen tube itself. Moreover, a similar level of induction occurs upon wounding of the stigma (Baumann et al. 2000) a result suggesting that *Bm2* induction in stigma may represent an effect of SI rather than the cause of it.

Attempts to isolate *Bm2* homologues from other self-incompatible grasses, by RT-PCR, have provided inconclusive results (Li et al. 1997). Primers targeted to the thioredoxin domain of *Bm2* successfully amplified a small, corresponding region from pollen cDNA of other grasses. These PCR products were found to share 94-97% nucleotide identity with the corresponding region of *Bm2*. In contrast, primers constructed against the allelic region failed to amplify a product. Thus the presence of an allelic domain upstream of the thioredoxin domain-like product, such as predicted for S_1 and S_2, could not be confirmed for any of the grasses investigated.

1.3 Plant Thioredoxins

Thioredoxins have been studied in many organisms resulting in a large and varied body of literature. The following section deals, almost exclusively, with plant thioredoxins since these are the primary focus of this thesis. Literature regarding thioredoxins from organisms other than plants has only been included if relevant to subsequent results and discussion chapters. For recent reviews covering thioredoxins in other organisms the reader is directed to the following references; Arnér and Holmgren, (2000), Grant (2001), Powis and Montfort, (2001), Ritz and Beckwith, (2001).

1.3.1 Introduction

Thioredoxins are small (approximately 12-kDa) proteins that appear to be ubiquitous in all organisms. Within the predominantly reductive environment of a cell, thioredoxins are involved in a range of biochemical processes. These include the regulation of enzymes (Holmgren 1989; Arnér and Holmgren, 2000), the modulation of transcription factors (Schenk et al. 1994; Hirota et al. 1999), as hydrogen donors (Holmgren 1989) and in oxidative protection (Chae et al. 1994). Most, but not all, roles depend upon the capacity of thioredoxins to effectively reduce disulfide bonds in target proteins.

All thioredoxins share a conserved active site motif, Cys-X-Pro-Cys (where X is generally Gly but can also be Pro or Ala), and conserved amino acids at structurally important positions (Eklund et al. 1991). In the oxidised state of the thioredoxin protein, the cysteines of the active centre can form a redox-active disulfide bridge. The disulfide can be reduced to a dithiol by electrons from NADPH or ferredoxin via the enzyme thioredoxin reductase (Holmgren 1989). Reduced thioredoxins have redox potentials (-270 mV for *Escherichia coli* thioredoxin) similar to other cellular reductants such as glutathione (Follmann and Häberlein, 1995). The low redox potential enables thioredoxins to reduce disulfide bonds in target proteins or to transfer reducing equivalents to substrates such as methionine sulfoxide residues (Follmann and Häberlein, 1995).

Plants possess two thioredoxin systems that can be distinguished by the electron donor and the enzyme that catalyses thioredoxin reduction.

1.3.2 The Ferredoxin /Ferredoxin-Thioredoxin Reductase /Thioredoxin (FTR/Trx) System

The FTR/Trx system, which is composed of ferredoxin, ferredoxin-thioredoxin reductase (FTR) and two nuclear-encoded thioredoxins, *m* and *f* (Figure 1.2A), is found only in photosynthetic eukaryotes and cyanobacteria. In plants the system is located in the chloroplast and regulates the redox state of enzymes involved in carbon metabolism and further biochemical processes (Scheibe and Anderson 1981; Buchanan 1991). The thioredoxins have been designated *m* and *f* in reference to the enzymes they were initially found to interact with, NADP-malate dehydrogenase (NADP-MDH) and fructose-1,6-bisphosphatase (FBPase), respectively (Schürmann et al. 1976; Jacquot et al. 1976; Jacquot et al. 1978; Buchanan 1991). The reducing power for the chloroplastic thioredoxins is derived from the electron transport system of illuminated thylakoid membranes. Electrons are transferred from photosystem I to the iron-sulphur protein ferredoxin-thioredoxin reductase (FTR) via ferredoxin. Subsequently, FTR reduces the thioredoxins, which in turn change the redox status of their target proteins, thereby conferring a strict, light-sensitive control of both the assimilatory and dissimilatory pathways.

1.3.3 The NADPH/NADP-Thioredoxin Reductase/Thioredoxin System

In contrast to the chloroplastic thioredoxin system, the NTR/Trx system has been identified in all organisms with the exception of cyanobacteria. In this system NADPH is the source of reducing power for the NADP-thioredoxin reductase (NTR) catalysed reduction of thioredoxin (Arnér and Holmgren 2000) (Figure 1.2B). Plants and animals cells contain two distinct NTR/Trx systems, one located in the cytosol and the other present only in the mitochondria (Pedras et al. 1996; Lee et al. 1999; Miranda-Vizuete et al. 1999; Banze and Follmann, 2000; Laloi et al. 2001). In plants, thioredoxins regulated by the cytosolic system are commonly termed thioredoxins *h* in reference to their initial discovery in heterotrophic tissue whilst thioredoxins that comprise the mitochondrial NTR/Trx system have been ascribed the title thioredoxins *o* (for organelle).

1.3.4 Plant Thioredoxins are a Diverse Multigene Family

Current knowledge supports the view that plants possess the greatest compliment of thioredoxins found in all organisms. Multiple thioredoxins have been isolated from algae (Crawford et al. 1986; Langlotz et al. 1986) and several species of flowering plant (Berstemann et al. 1983; Vogt and Follmann, 1986). Elucidation of the total number of thioredoxins in a given species, their tissue specificity and subcellular localisation has been the focus of several subsequent studies. Until recently, the most comprehensive picture to emerge has been in soybean where protein and organelle fractionation techniques were used in conjunction with a range of enzymatic assays to identify no less than six distinct thioredoxin proteins in both leaves (Häberlein 1991) and seeds (Häberlein et al. 1995). The multiplicity of thioredoxin proteins as detected by biochemical methods is substantiated by the identification of multiple thioredoxin genes in plants (Rivera-Madrid et al. 1995). The genome of *Arabidopsis thaliana* contains no less than seventeen distinct thioredoxin genes (Figure 1.3A), eight of which encode thioredoxin proteins of the *h* class (Rivera-Madrid et al. 1995; Juttner et al. 2000; Laloi et al. 2001). The other genes comprise two thioredoxins *f*, four thioredoxins *m*, two thioredoxins *o* and a novel prokaryotic-like thioredoxin, thioredoxin *x* (Mestres-Ortega and Meyer, 1999; Laloi et al. 2001). By contrast, the genomes of *E. coli*, yeast and humans contain two, three and three thioredoxin genes respectively. The number and diversity of thioredoxins in *Arabidopsis* is probably representative of many higher plants. Database searches performed during the preparation of this thesis reveal a similar number of all thioredoxins including thioredoxin *x* in expressed sequence tag (EST) libraries of other species such as tomato and soybean.

1.3.5 Plant Thioredoxins: Expression Characteristics and Subcellular Locations

Thioredoxin transcripts have been detected in many plant tissues although the temporal and developmental regulation of most remains to be determined. Similarly, the subcellular location of relatively few of the proteins comprising each class of plant thioredoxins has been confirmed.

Thioredoxins *m* and *f* genes are highly expressed in green tissue and encode the only thioredoxin proteins isolated from chloroplasts (Jacquot et al. 1976; Wolosiuk et al. 1979). A single thioredoxin f and m protein has been isolated from the chloroplast of both spinach and pea. However, the expression and transcription of thioredoxins *m* and *f* may not be restricted to photosynthetic tissue. Chromatographically separated thioredoxins from soybean seed showed a virtually identical protein profile to that of soybean leaf, including two thioredoxins that have been assigned to the chloroplast (Häberlein et al. 1995). Correspondingly, the transcripts of thioredoxin *m* genes have since been detected in the seed and root of *Arabidopsis* (Mestres-Ortega and Meyer, 1999). Thioredoxin *f* genes are also present in EST libraries of tomato seed (Genbank clone accession numbers AW036169, AW036163). The transcription of thioredoxins *m* and *f* in non-photosynthetic tissue raises the question of whether proteins encoded by genes belonging to these classes perform additional, non-photosynthetic related functions. It has been suggested that chloroplastic-type thioredoxins present in seed are stored as precursor molecules (Häberlein et al. 1995). However their presence in root remains to be explained.

As has been observed for thioredoxins *m*, *h*-type thioredoxins are also expressed in a range of tissues. Plant thioredoxin *h* mRNAs have the highest expression level in rapidly growing cells (Marty et al. 1993; Rivera-Madrid et al. 1995). Proteins encoded by different members of the thioredoxin *h* class have been purified from the endoplasmic reticulum (Marcus et al. 1991) and the plasma membrane (Shi and Bhattacharya, 1996), and located in the nucleus by *in situ* hybridisation (Serrato et al. 2001). In addition, thioredoxin h proteins have been identified in the sieve tubes of several plants despite not having a signal peptide (Ishiwitari et al. 1995; Schobert et al. 1998).

The prokaryotic-like thioredoxin *x* of *Arabidopsis* has an expression profile similar to that of *Arabidopsis* thioredoxins *m* with highest transcript levels in leaf, seedlings and cultured cells (Mestres-Ortega and Meyer, 1999), however, the subcellular location of the corresponding proteins is as yet undetermined.

The most recently identified plant thioredoxin, thioredoxin *o*, was demonstrated by Western blot and cellular fractionation experiments to be located exclusively in the mitochondria of cultured *Arabidopsis* protoplasts (Laloi et al. 2001).

1.3.6 Thioredoxin Structure

Plant thioredoxins share most of the amino acid residues determined as being essential for the maintenance of protein structure and catalysis in *E. coli* thioredoxin (Eklund et al. 1991). The alignment of all available thioredoxins has also identified conserved residues, additional to those of known structural importance, that appear to be specific to each thioredoxin class (Figure 1.3A) (Jacquot et al. 1997).

Thioredoxins *m*, *f* and *o* from plants are most easily differentiated from *h* class thioredoxins by the presence of an N-terminal transit peptide of varying size (Lepiniec et al. 1992; Wedel et al. 1992). Additionally, *f*-type thioredoxins differ from thioredoxins *m* and *o* by the presence of a third conserved cysteine in the C-terminal part of the proteins, a feature shared with animal thioredoxins (Hartman et al. 1990). The recently identified thioredoxin *x* also appears to encode a signal peptide although the corresponding protein has yet to be located in any specific organelle. Thioredoxin *x* can be distinguished from m, *f* and o types by several sequence differences including a two-residue insertion in the C-terminus of the protein (Mestres-Ortega and Meyer, 1999). A further distinguishing feature of thioredoxins *m* and *x* is their closer sequence similarity to thioredoxins from prokaryotes than to other plant thioredoxins. Phylogenetic analyses have provided strong evidence to advocate a prokaryotic origin for *m* and *x* thioredoxins and a eukaryotic origin for thioredoxins *h* and *f* (Sahrawy et al. 1996; Mestres-Ortega and Meyer, 1999).

Thioredoxins have a similar three-dimensional structure. Nuclear magnetic resonance spectroscopy of *E. coli* (Jeng et al. 1994) and *Chlamydomonas* (Mittard et al. 1997) thioredoxins and X-ray crystallography of human (Weichsel et al. 1996) and spinach (Capitani et al. 2000) thioredoxins has demonstrated a conserved arrangement of five β sheets surrounded by four α helices, a structure

commonly referred to as the thioredoxin fold yet not restricted to thioredoxins (Martin 1995) (Figure 1.3B). Within the protein, the active site amino acids are located in a protrusion at the end of the second β sheet and the beginning of the second α helix (Ecklund 1991). The N-terminal active site cysteine is the attacking nucleophile in disulfide reduction and is the more exposed of the active site cysteines (Holmgren 1995). Reduction of the active centre disulfide bridge is accompanied by localised structural changes including an increase in distance between the two sulphur molecules and rotation of the side chain of the N-terminal cysteine (Jeng et al. 1994; Qin et al. 1994).

Thioredoxins *m* and *f* from spinach are the only higher plant thioredoxin proteins for which a crystal structure has been described (Capitani et al. 2000). The tertiary structure of thioredoxin *m* is more similar to that of *E. coli* than thioredoxin *f*. The most significant difference between thioredoxins *m* and *f* is found in the surface topography around the active site. The active site of thioredoxin *f* is surrounded by positive charges, which may be important in the interaction with target proteins (Dai et al. 2000). Also exposed on the surface is the conserved C-terminal cysteine which is located 9.7 Å from the N-terminal active-site cysteine. Mutagenesis studies suggest a role for the C-terminal cysteine in interactions with target enzymes (Val et al. 1999).

1.3.7 Thioredoxin Evolution

The evolution of photosynthetic eukaryotes appears to have been accompanied by an increase in the number and diversity of thioredoxins. Sequence and phylogenetic analyses of higher plant thioredoxins have been performed by many researchers and have determined that plants contain five distinct classes; thioredoxins *m, f, h, x* and *o* (Jacquot et al. 1997; Laloi et al. 2001). Through comparison with thioredoxins derived from other organisms, it has also become clear that many classes have different evolutionary origins. Thioredoxins *m* are more closely related, at the sequence level, to thioredoxins from prokaryotic sources than to other eukaryotic thioredoxins (Hartman et al. 1990; Jacquot et al. 1997) indicating a prokaryotic origin for this class. Moreover, Mestres-Ortega and Meyer (1999) have shown that of the four thioredoxins identified in the genome of

the cyanobacterium *Synechocystis* sp PCC6803 one (open reading frame (ORF) Slr0623) is found on the branch containing all eukaryote thioredoxins *m*. Given that *Synechocystis* is considered to represent a modern relative of the progenitor of plant chloroplasts, it has been suggested that all plant thioredoxins *m* may have evolved from an ancestral thioredoxin similar to Slr0623. The branch containing ORF Slr0623 also includes several red algae species. In red algae, thioredoxin *m* is encoded by the chloroplast genome (Reynolds et al. 1994) in contrast to the green algae, *Chlamydomonas*, where thioredoxin *m* is nuclear encoded as in higher plants. Since thioredoxins *m* are encoded by the chloroplast genome of red algae it has been suggested that the red algae represent organisms situated, evolutionally, just before the transfer of thioredoxins *m* from the chloroplast to the nuclear genome (Jacquot et al. 1997).

Sequence comparisons also support a prokaryotic origin for the recently discovered thioredoxin *x* of *Arabidopsis*. However, unlike thioredoxins of the *m* class, thioredoxin *x* has been shown to form a separate group with a different gene from *Synechocystis* sp PCC6803 (Slr1139) suggesting a separate origin for this plant thioredoxin (Mestres-Ortega and Meyer 1999).

Members of the most recently discovered class, thioredoxins *o*, are distinct from all other plant thioredoxins in the number (five) and position of the introns present in the coding region (Laloi et al. 2001). These features suggest a distinct evolutionary origin for this class. Within a phylogenetic tree comprising plant thioredoxins of all classes, thioredoxins *o* formed a clade more closely associated to those of thioredoxins *m* and *x*, implying that the class is also derived from a prokaryotic ancestor (Laloi et al. 2001). By contrast, thioredoxins of the *f* and *h* classes are likely to be of eukaryotic origin as they share highest sequence identity with thioredoxins from eukaryotic organisms and contain an intron at a position conserved between thioredoxins from plants and animals (Sahrawy et al. 1996).

The *h* class represents the largest and most divergent group of thioredoxins found in plants. Thioredoxins *h* of higher plants and *Chlamydomonas* are likely to have evolved from a common ancestor since the position of two introns is conserved across these species (Sahrawy et al. 1996). Duplication and divergence of

thioredoxin *h* genes appears to be an ancient event as multiple *h*-type thioredoxins are present in all higher plants and thioredoxin *h* sequences typically display high levels of interspecific sequence identity (Rivera-Madrid et al. 1995).

1.3.8 Thioredoxin Reductases

In all organisms, the action of thioredoxins is modulated by dimeric enzymes called thioredoxin reductases. As their name suggests, thioredoxin reductases are able to reduce oxidised thioredoxin; however, this is probably not their exclusive function as they have been shown to interact with other cellular substrates (Andersson et al. 1996; May et al. 1997). The sequencing of thioredoxin reductase genes and crystallization of thioredoxin reductase proteins from several species (Waksman et al. 1994; Williams 1995; Dai et al. 1996) has revealed that two broad classes exist in nature: The FTR that uses ferredoxin as an electron source and NTR, which uses NADPH as the source of reducing equivalents.

NTRs of plants are homodimeric flavoproteins that belong to a family of pyridine nucleotide disulfide oxidoreductase enzymes (Williams 1995). The protein subunits have a molecular weight of 35 kDa and contain an FAD- and NADPH-binding domain as well as a conserved redox active disulfide (Jacquot et al. 1994). Electrons are transferred from NADPH to FAD and then to the disulfide. The tertiary structure of NTR from *Arabidopsis thaliana* has been determined and has revealed that the plant NTR enzyme is more closely related to prokaryotic thioredoxin reductases than to NTR from higher eukaryotes such as mammals (Dai et al. 1996; Williams et al. 2000). Within the protein, the active disulfide site faces the isoalloxazin ring of the flavin (FAD) thus allowing the transfer of electrons from FAD to the disulfide bridge (Wang et al. 1996). As determined for *E. coli* (Dai et al. 1996), binding of thioredoxin by plant NTRs also requires a structural change in the NTR molecule. The NADPH domain rotates away from the flavin to accommodate the binding and subsequent reduction of thioredoxin (Lennon et al. 2000).

NTRs display varying levels of affinity for thioredoxins from different sources (Jacquot et al. 1994). Interestingly, several thioredoxins from divergent species have been shown to be better substrates for a particular NTR than some

thioredoxins derived from the same or more closely related organisms (Jacquot et al. 1994; Rivera-Madrid et al. 1995). Given that plants contain several distinct NTR proteins (Banze and Follmann, 2000), there is likely to be a level of specificity in the interaction between NTRs and thioredoxins.

FTR of higher plants is a nuclear encoded heterodimer, composed of two subunits, a variable subunit and a larger catalytic subunit. The catalytic subunit is conserved between species and contains a [4Fe-4S] centre and the reductive disulfide (Droux et al. 1987; Huppe et al. 1990). It sits on top of the variable subunit forming a thin molecule that resembles a concave disc (Dai et al. 2000a). The [4Fe-4S] cluster and the reductive disulfide are located at the centre of this disc. Dai et al. (2000b) suggested that both the electron donor (ferredoxin) and the electron acceptor (thioredoxin) simultaneously dock on opposite sides of the FTR molecule. Thioredoxin forms a mixed disulfide bond with the regulatory cysteine of FTR. This bond is resolved after an electron from a second ferredoxin has been received and the reduced thioredoxin is released.

1.3.9 Chloroplast Thioredoxins

The most extensively studied plant thioredoxins are those involved in the redox regulation of chloroplastic enzymes. However, since this thesis deals primarily with non-chloroplastic, *h*-type, thioredoxins only a brief summary of the functions of chloroplastic thioredoxin isoforms is provided. Recent reviews (Ruelland and Miginiac-Maslow, 1999; Schürmann and Jacquot, 2000; Baumann and Juttner, 2002) have covered this area of the literature in detail and can be consulted for specific information regarding the thioredoxin-dependant regulation of individual enzymes.

The FTR/thioredoxin *m*, *f* system represents one mechanism by which light regulates the activity of various chloroplastic proteins (Buchanan 1991) and thereby a range of biochemical processes (Figure 1.2B). Through the interaction with key enzymes, the FTR/thioredoxin system acts like a switch between anabolic and catabolic pathways in the chloroplast, thus preventing futile cycling (Buchanan 1991; Jacquot et al. 1997; Dai et al. 2000a).

No less than seven chloroplastic enzymes have been demonstrated to be modulated by thioredoxins *m* and *f* (Table 1.2), most of which are involved in carbon metabolism. Three enzymes, FBPase, sedoheptulose-1,7-bisphosphatase (SBPase) and phosphoribulokinase (PRK) are constituents of the Calvin cycle. Glucose-6-phosphate dehydrogenase (G6PDH) is involved in the oxidative pentose-phosphate pathway (glycolysis). NADP-MDH is a key enzyme of carbon assimilation in C4 plants and participates in the exportation of reducing equivalents from the chloroplast to the cytosol in C3 plants (Scheibe 1991). Rubisco activase regulates the activity of ribulose 1,5-bisphosphate carboxylase/oxygenase, an enzyme that initiates photosynthetic carbon assimilation in C3 plants (Zhang and Portis, 1999), whilst the CF_1 γ subunit of ATP synthase provides ATP for the chloroplast (Schwarz et al. 1997). Thioredoxins *m* and *f* coordinate the activity of these enzymes through the reduction of intramolecular disulfide bridges, a process that results in conformational changes (Ruelland and Miginiac-Maslow, 1999). Disulfide reduction causes the activation of all enzymes except G6PDH, which is deactivated when reduced (Scheibe 1991).

For most of the thioredoxin-regulated chloroplast enzymes, unregulated cytosolic forms also exist. Sequence comparisons of both forms have allowed the identification of cysteines specific for the chloroplastic isoforms (for a review see Ruelland and Miginiac-Maslow, 1999). Site directed mutagenesis studies have determined which cysteines form the redox-active disulfide for all of the enzymes described above (Table 1.2) (Brandes et al. 1996; Jacquot et al. 1997; Rodriguez-Suarez et al. 1997; Schwartz et al. 1997; Wenderoth et al. 1997; Duford et al. 1998; Zhang and Portis, 1999; Hirasawa et al. 2000). In addition, structural data from X-ray crystallography have verified the position of the disulfide bonds in FBPase (Villeret et al. 1995; Chiadmi et al. 1999) and NADP-MDH (Johansson et al. 1999). Comparative analysis of the protein sequences of these chloroplastic enzymes has revealed no common cysteine-containing consensus motif. Nor are the regulatory cysteines necessarily part of the active site of the enzymes. Presently, PRK is the only example of a thioredoxin-regulated enzyme containing regulatory cysteines at its active site (Hirasawa et al. 1998).

Table 1.2 Characteristics of chloroplast enzymes modulated by thioredoxins. The cysteines of the regulatory site are in bold.

Target Enzyme	Biochemical Pathway	Activator	Regulatory Site	Plant
FBPase	Calvin cycle	Trx *f*	E**C**X$_{19}$**C**IVNVCQ 153 173	Pea
SBPase	Calvin cycle	Trx *f*	S**C**GGTA**C**V 52 57	Wheat
PRK	Calvin cycle	Trx *m/f*	G**C**X$_{38}$**C**L 16 55	Spinach
G6PDH	Oxidative pentose-phosphate pathway	Inactivated by Trx *m*	**C**RIDKRED**C** 149 157	Potato
NADP-MDH	C3 plants: redox shuttle C4 plants: carbon assimilation	Trx *f/m*	E**C**FGVF**C**T 24 29 K**C**X$_{11}$**C**I 365 377	Sorghum
Rubisco-activase	Calvin cycle (indirectly)	Trx *f*	G**C**X$_{18}$**C**V 392 411	*Arabidopsis*
CF_1-ATP synthase γ subunit	photophosphorylation	Trx *f*	I**C**DINGN**C**V 198 204	Pea

From Baumann and Juttner (2002)

Initial biochemical studies into thioredoxins *m* and *f* indicated that they interacted preferentially with certain chloroplastic enzymes. Thioredoxin *f* was reported as activating FBPase, SBPase and PRK whereas thioredoxin *m* was thought to reduce NADP-MDH and G6PDH (Schürmann et al. 1976; Jacquot et al. 1978; Schürmann et al. 1981; Buchanan 1991). Recent analyses have, however, demonstrated that the specificity of interactions is less strict than first proposed. PRK activation has been reported to be twice as effective with thioredoxin *m* than with thioredoxin *f* (Geck and Hartmann, 2000) and NADP-MDH has since been shown to be more efficiently reduced by thioredoxin *f* than *m* (Geck et al. 1996). López Jaramillo et al. (1997) also showed that thioredoxin *m* of pea is capable of activating FBPase, albeit not very efficiently. To date, thioredoxin *f* has demonstrated the biochemical capacity to reduce all the aforementioned enzymes, except G6PDH, which currently represents the only enzyme exclusively reduced by thioredoxin *m*.

Site directed mutagenesis studies have provided some insight into the structural features of thioredoxins *f* and *m* that are relevant for protein interaction. The importance of the conserved third cysteine of thioredoxins *f* was demonstrated by substitution mutants that had the cysteine changed to a serine or alanine (Val et al. 1999). Both mutants displayed impaired interaction with FBPase and NADP-MDH, resulting in reduced activation of the enzymes. Charge topography around the active site also appears to be an important component of the interaction between chloroplast thioredoxins and target enzymes. Both Wangensteen et al. (2001) and Duek and Wolosiuk (2001) demonstrated that changing the charge of amino acids in thioredoxins *m* and *f* influenced the strength of interaction with FBPase. Additionally, the substitution of thioredoxin *f* amino acids around the active site for residues more closely resembling those of thioredoxin *m* resulted in a mutant *f* protein displaying more *m*-like reactivity with various target enzymes (Geck et al. 1996; Geck and Hartmann, 2000).

Several other redox-regulated chloroplast enzymes have been proposed as potential targets for thioredoxin reduction. These include Glyceraldehyde 3-phosphate dehydrogenase (Baalmann et al. 1995; Reichert et al. 2000), Acetyl CoA carboxylase (Sasaki et al. 1997; Kozaki et al. 2000), ADP-glucose pyrophosphorylase (Ballicora et al. 2000) and the light harvesting complex II kinase (Rintamäki et al. 2000).

1.3.10 Mitochondrial Thioredoxins

The mitochondria of plants and animals contain a distinct thioredoxin system comprising an NADPH-dependant thioredoxin reductase and specific thioredoxins (Pedrajas et al. 1999; Lee et al. 1999; Miranda-Vizuete et al. 1999; Banze and Follmann, 2000). Evidence for this system in plants came from the isolation of mitochondrial thioredoxin proteins (Konrad et al. 1996), and the identification of a mitochondrial-specific protein utilizing NADPH and plant thioredoxins as substrates (Banze and Follmann, 2000). Recently, genes encoding a mitochondrial specific thioredoxin and NTR were identified in *Arabidopsis* (Laloi et al. 2001). Both genes were found to possess an N-terminal signal sequence consistent with

organellar targeting. In addition, Western blots of protoplast protein fractions demonstrated that the protein products of both genes are located exclusively in the mitochondrial fraction.

In animals, thioredoxins of mitochondria have been shown to be involved in the regulation of mitochondrial 2-oxoacid dehydrogenase complexes (Bunik et al. 1995; Bunik et al. 1997) and thioredoxin-dependant peroxidases (Watabe et al. 1997; Araki et al. 1999), functions that could also apply to mitochondrial thioredoxins in plants. In addition, plant mitochondria contain biochemical pathways and enzymes not found in the mitochondria of other organisms, for example the alternative oxidase, some of which have been proffered as possible regulatory targets for mitochondrial thioredoxins (Vanlerberghe et al. 1995).

1.3.11 Cytosolic Thioredoxins: Thioredoxins *h*

1.3.11.1 Thioredoxins *h* in Seed Germination

Germination of cereal seeds involves the reduction of seed storage proteins (Lozano et al. 1996; Yano et al. 2001). Evidence for the involvement of thioredoxins in seed germination is based largely on the observation that thioredoxins from plant and bacterial sources have the biochemical capacity to reduce seed storage proteins and inactivate amylolytic enzyme inhibitors as determined by *in vitro* studies (Kobrehel et al. 1992). For example, wheat and *E. coli* thioredoxins were found to reduce and inactivate α-amylase and trypsin inhibitor proteins from several plant species (Kobrehel et al. 1991; Jiao et al. 1992). Thioredoxin also reduced a seed-specific serine protease, thiocalsin, that following reduction was activated by calcium (Besse et al. 1996). Reduced thiocalsin was able to cleave thioredoxin reduced gliadins and glutenins. More direct evidence for thioredoxin-mediated regulation of seed enzymes has come from a transformation experiment in which barley transgenic lines, overexpressing thioredoxin *h* in seed, displayed a four-fold increase in the activity of pullanase (limit dextrinase), an enzyme that cleaves α-1, 6 linkages in starch (Cho et al. 1999).

Examination of the level of thioredoxin protein in the endosperm of germinating wheat has revealed a progressive reduction during germination and that this process is enhanced by giberellic acid (GA_3) (Lozano et al. 1996). More recently, Serrato et al. (2001) reported the cloning of two thioredoxins from wheat that were predominantly localised in the nucleus of scutellum and aleurone cells and at a lower level in endosperm. In contrast to the endosperm, the level of thioredoxin protein in aleurone cells remained unchanged during germination. Moreover, GA_3 did not have an effect on the level of thioredoxin in this locality, indicating the presence of multiple differentially regulated thioredoxins in wheat seed. Alternatively, it is possible that the thioredoxin proteins identified in both studies are the same and that the GA_3-associated disappearance of thioredoxin in the endosperm described by Lozano et al. (1996) is due to the action of GA_3 regulated proteases.

Thioredoxin *h* is not the only native protein with a capacity to reduce seed proteins *in-vitro*. Kobrehel et al. (1991) reported that the α-amylase inhibitors DSG-1 (from durum wheat) and CM-1 (from bread wheat) as well as a corn kernel trypsin inhibitor and two purothionin proteins (α-1 and β from durum wheat) were all more effectively reduced by *E. coli* glutaredoxin than the *E. coli* thioredoxin system used. Further, Kobrehel et al. (1992) found that wheat gliadins and glutenins can also be reduced by *E. coli* glutaredoxin, albeit less effectively than thioredoxin. Since the publication of this work glutaredoxins have been cloned from the seeds of rice (Minakuchi et al. 1994) and cotyledons of castor bean (Szederkényi et al. 1997) and can be found to be represented in EST seed libraries from wheat and maize.

1.3.11.2 Thioredoxins *h* in the Companion Cell, Sieve Element Complex

Thioredoxins *h* represent a common (Schobert et al. 1998), and often predominant (Ishiwitari et al. 1995), component of the large number of proteins present in the phloem sap of plants (Fisher et al. 1992; Sakuth et al. 1993). Glutaredoxin has also been identified in the sieve tube exudate of four dicotyledonous species (Szederkényi et al. 1997; Schobert et al. 1998), suggesting that redox systems are an integral component of functional sieve tube complexes.

Mature sieve tubes are enucleate and contain no cellular organelles yet require a functional plasma membrane (Sjölund 1997). Consequently, sieve tubes are thought to be supported by proteins synthesised in neighbouring companion cells and transferred via plasmodesmata (Schobert et al. 1998). Microinjection studies with mesophyll cells from rice (Ishiwitari et al. 1998), *Cucurbita* and *Ricinus* (Balachandran et al. 1997) have revealed that recombinant thioredoxin h and glutaredoxin have the capacity to mediate their own transport through plasmodesmata. To determine the structural motifs required for cell-cell transport Ishiwitari et al. (1998) generated two thioredoxin *h* mutants incapable of plasmodesmatal movement. In both cases mutations involved the modification of charged residues predicted to project from the protein surface implying that surface charge topology is important for the binding and or transport of thioredoxins through plasmodesmata.

A possible role for thioredoxins in sieve tubes is as a hydrogen donor for peroxiredoxin (Prx). Prx was recently identified as an abundant protein in sieve tubes of poplar (*Populus trichocarpa*) (Rouhier et al. 2001). Moreover, *in vitro* peroxidase assays demonstrated that poplar Prx was able to reduce H_2O_2 using either the poplar thioredoxin in combination with *Arabidopsis* NTR or glutaredoxin/glutathione/glutathione-reductase as a proton donor system.

1.3.11.3 Thioredoxins *h* in Self-Incompatibility

A role for thioredoxins in the self-incompatibility (SI) response of plants was first reported when two *h*-like thioredoxin proteins (THL-1 and THL-2) were found to interact with the kinase domain of the SI locus receptor kinase (SRK), the female component of the *Brassica* SI system (Bower et al. 1996). Mutational studies have demonstrated that the interaction is dependent upon a functional thioredoxin active site and the presence of a conserved cysteine in the N-terminus of the SRK (Mazzurco et al. 2001). Recently, it was shown that SI in *Brassica oleracea* involves the phosphorylation of SRK and that a protein present in stigma extract is able to inhibit SRK phosphorylation (Cabrillac et al. 2001). Exposure of SRK and stigma extract to the SI protein component of incompatible pollen results in phosphorylation suggesting that the inhibitor protein prevents spontaneous

activation of the SI signalling pathway. Thioredoxin has been proposed as a candidate for the stigma inhibitor since THL-1 and *Spirulina* thioredoxin were found to inhibit SRK phosphorylation *in-vitro* (Cabrillac et al. 2001). Furthermore, the depletion of thioredoxins from stigma extract was reported to abolish the inhibitory effect of SRK phosphorylation. However, an interesting finding arising from this research has been the observation that a functional, reduced active site is required for the thioredoxin-mediated inhibition of SRK phosphorylation, yet SRK inhibition appears to involve direct binding of the inhibitor. How thioredoxin might form an inhibitory complex with SRK through the active site remains as yet unexplained.

1.3.12 Thioredoxins in Oxidative Stress Protection

A common role for thioredoxins in many organisms is oxidative stress protection. In yeast, studies on thioredoxin deletion mutants implicated both cytoplasmic thioredoxins, *trx1* and *trx2*, in oxidative protection (for review see Grant 2001). *trx2* deletion mutants were sensitive to hydrogen peroxide (H_2O_2) (Kuge and Jones, 1994) and resistant to the thiol oxidant diamide (Muller 1996). Additionally, *trx2* expression was strongly induced in yeast exposed to oxidants. The H_2O_2-dependant activation of *trx2*, was shown to be controlled by the yeast AP-1 (Yap1p) transcriptional factor, a regulator of several stress-responsive genes (Kuge and Jones, 1994). Moreover, Yap1p has been shown to be constitutively active in a *trx1/trx2* double mutant (Izawa et al. 1999). Similarly, mutational studies in the bacterium *E. coli* demonstrated that the mutation of specific components on the thioredoxin system resulted in increased H_2O_2 sensitivity (Takemoto et al. 1998). Expression of *E. coli* thioredoxin 2 (*trxC*) has since been shown to be regulated by the transcriptional activator OxyR in response to oxidative stress (Ritz et al. 2000).

One mechanism by which thioredoxins of these organisms participate in oxidative stress defence is by providing reducing equivalents to specific antioxidant proteins such as peroxiredoxin (Prx). When reduced, Prxs catalyse the reduction of hydrogen peroxide or alkyl hydroperoxides to water and the corresponding alcohol (Rouhier et al. 2001). Five Prx isoforms have been detected in yeast, three in *E. coli* and six in mammalian cells (Zhou et al. 2000). Analysis of yeast mutants

lacking different isoforms has revealed that there is a level of specificity in the interaction of these enzymes with certain types of peroxides (Chae et al. 1993; Jeong et al. 1999; Lee et al. 1999).

Although an oxidative stress protection role has yet to be conclusively demonstrated for thioredoxins in higher plants, several studies have provided indirect evidence advocating the presence of thioredoxin/Prx antioxidant system in plants. As for other organisms, higher plants also encode Prxs (Baier and Dietz, 1997; Rhee et al. 1999; Genot et al. 2001) several of which have been demonstrated to be reduced by the NTR/thioredoxin system (Choi et al. 1999; Verdoucq et al. 1999; Rouhier et al. 2001). Significantly, complementation assays have shown that *Arabidopsis* thioredoxins *h3, m1, m2, m4* and *x* are able to restore H_2O_2 tolerance to yeast thioredoxin mutants (Mouaheb et al. 1998; Issakidis-Bourguet et al. 2001). Moreover, an *Arabidopsis h3* C35S mutant, designed to trap target proteins, formed a complex with the YLR109 Prx of yeast (Bréhélin et al. 2000). Most recently, Goyer et al. (2002) cloned a thioredoxin-dependant peroxidase from the green algae *Chlamydomonas* that functioned as an antioxidant towards reactive oxygen species and protected DNA against ROS-induced degradation.

Unlike classical thioredoxins, an oxidative protective role has been attributed to *CDSP32*, a thioredoxin-like protein located in the chloroplast of potato (Rey et al. 1998). The C-terminus of the *CDSP32* gene encodes a typical thioredoxin sequence including the active centre consensus sequence and structural residues at appropriate positions. Expression of the thioredoxin region also demonstrated that the recombinant protein possessed thioredoxin activity. Under oxidative stress, *CDSP32* expression is upregulated and the corresponding protein accumulates in the stroma of stressed plants where it is suggested to preserve the redox potential of chloroplast proteins (Broin et al. 2000).

1.3.13 Objectives of this Thesis

The initial objective of this study was to resolve uncertainty regarding the structure of *Bm2* and to investigate the level of *S*-gene diversity in the two-locus SI system of grasses. To this end *Bm2* homologues were isolated from a range of

self-incompatible grass species. This analysis revealed that *Bm2* does not represent the male SI determinant but rather encodes a highly conserved and novel *h*-class thioredoxin. Consequently, the focus of this study shifted to the characterisation of these thioredoxins in the grasses and other plants.

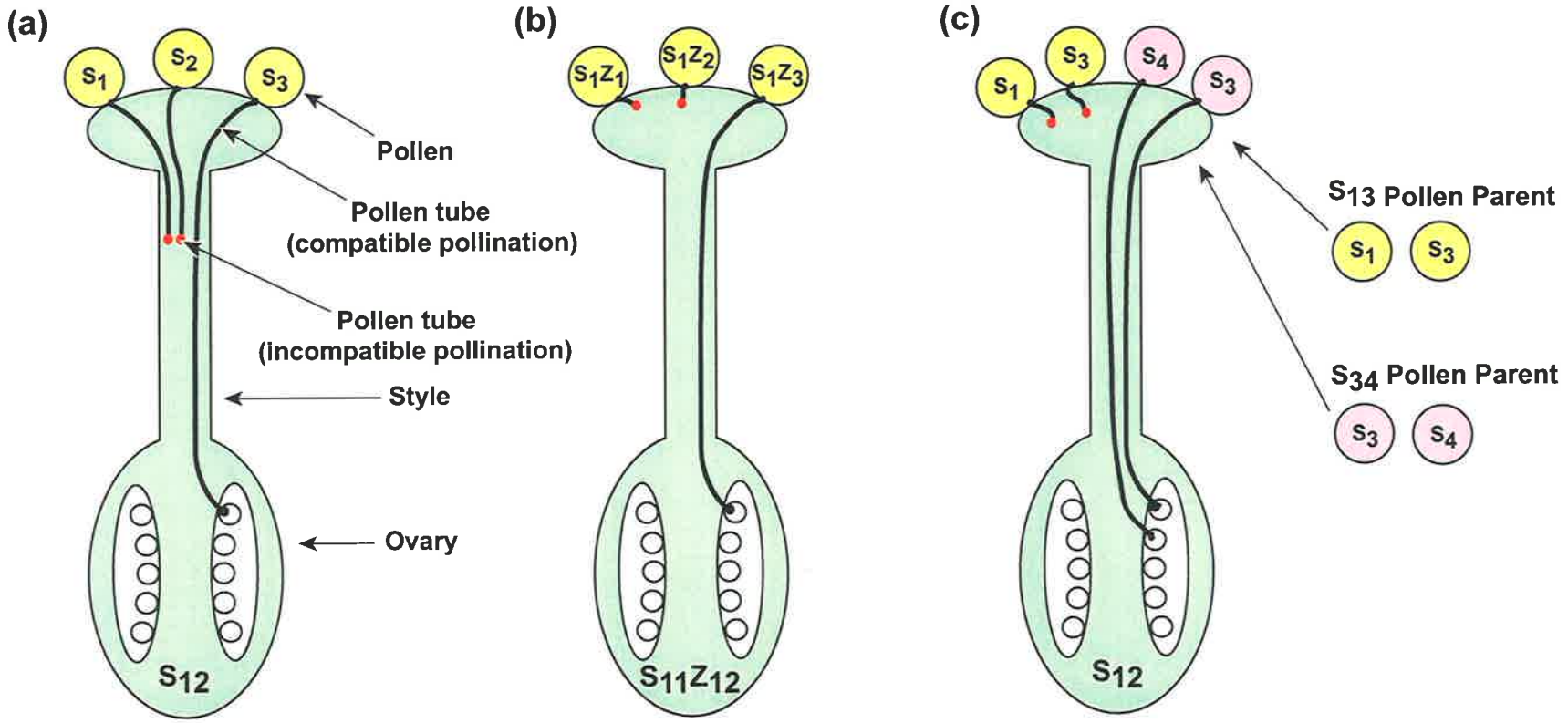


Figure 1.1 Schematic illustration of the genetics of different self-incompatibility systems.
a) The single locus gametophytic system found in the *Solanaceae, Rosaceae, Papaveraceae* and *Scrophulariaceae*. Self-incompatibility is controlled by a single multiallelic *S*-locus. Self-incompatibility arises when the *S*-allele carried by the pollen is matched in the pistil; **b)** The two locus gametophytic system present in the *Poaceae*. Self-incompatibility is controlled by two unlinked, multiallelic loci, *S* and *Z*. A pollen grain is incompatible only when both of its *S* and *Z* alleles are matched in the recipient pistil; **c)** The sporophytic system of the *Brassicaceae*. Self-incompatibility is controlled by a single multiallelic *S*-locus. However, pollen rejection occurs when the pollen parent shares alleles in common with the recipient pistil. For example, the S_3 allele derived from the S_{13} pollen parent is rejected as the parent shares an allele (S_1) in common with the recipient pistil. By contrast, the S_3 allele from the S_{34} pollen parent is compatible and able to effect fertilisation.

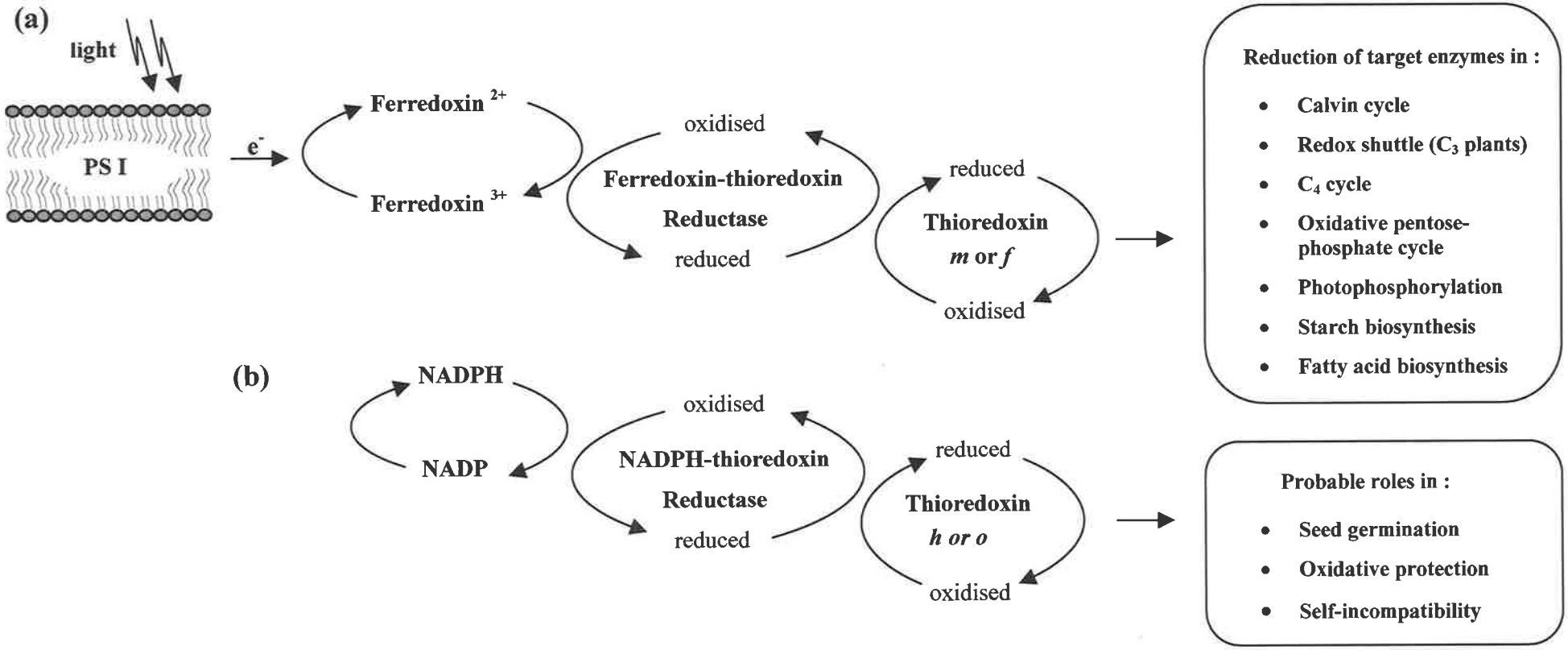


Figure 1.2 Schematic representation of the two thioredoxin systems found in plants. (a) Light-induced regulation of chloroplastic enzymes through the ferredoxin-thioredoxin reductase system. In the light, photosystem I of the photosynthetic electron transport chain reduces ferredoxin. Ferredoxin can then serve as an electron donor to ferredoxin-thioredoxin reductase, which reduces thioredoxins *m* and *f*. The thioredoxins in turn reduce the disulfide bonds of several target enzymes, thereby modifying their activity. **(b)** The NTR/Trx system. The reduction of plant thioredoxins *h* and *o* is catalyzed by NADPH-thioredoxin reductase using NADPH as a source of reducing equivalents. When the thioredoxin cysteine disulfide is reduced to a dithiol, the thioredoxin is able to reduce disulfides in target proteins.

(a)

```
Ato1    MKGNWSIVRKVLHRQFSTLRSSTPSSRLSTSIRPLVLAPNSISSLIARNSLFTASNI  57
Atf1                  MPLSLRLSPSPTALSPTTGGFGPSRKQCRIPYSGVPTTKIG  41
Atm1             MAAYTCTSRPPISIRSEMRIASSPTGSFSTRQMFSVLPESSGLRTRV  47
Atx                            MRSYLTPPVRSCSPATSVSVKPLSSVQVTS  30

Ath1                        MASEEGQVIACHTVETWNEQLQKANESKTLVVVDFT  36
Ato1    GPSIDFNFSNTSLPHRRSLCSEAGGENGVVLVKSEEEFINAMSKAQDGSLPSVFYFT 114
Atf1    FCSLDSRKRGDSSVVRCSLETVNVSVGQVTEVDKDTFWP--IVKAA-GEKLVVLDMY  95
Atm1    SLSSLSKNSRVSRLRRGVICEAQDTATGIPVVN-DSTWDSLVLKADE---PVFVDFW 100
Atx     VAANRHLLSLSSGARRTRKSSSSVIRCGGIKEIGESEFSSTVLESAQ---PVLVEFV  84
E.coli                        MSDKIIHLTDDSFDTDVLKA-DGA--ILVDFW  29

Ath1    ASWCGPCRFIAPFFADLAKKLPN-VLFLKVDTDE--LKSVASDWAIQAMPTFMFLKE  90
Ato1    AAWCGPCRFISPVIVELSKQYPD-VTTYKVDIDEGGISNTISKLNITAVPTLHFFKG 170
Atf1    TQWCGPCKVIAPKYKALSEKYDD-VVFLKLDCNPD-NRPLPKELGIRVVPTFKILKD 150
Atm1    APWCGPCKMIDPIVNELAQKYAGQFKFYKLNTDE--SPATPGQYGVRSIPTIMIFVN 155
Atx     ATWCGPCKLIYPAMEALSQEVGDKLTIVKIDHDA--NPKLIAEFKVYGLPHFILFKD 149
E.coli  AEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQ--NPGTAPKYGIRGIPTLLLFKN  84

Ath1    GKIL--DKVVGAKKDELQSTIAKHLA          114
Ato1    GSKK--GEVVGADVTKLKNLMEQLYK          194
Atf1    NKVV--KEVTGAKYDDLVAAIETARSAASG      178
Atm1    GEKK--DTIIGAVSKDTLATSINKFL          179
Atx     GKEVPGSRREGAITKAKLKEYIDGLLNSISVA    171
E.coli  GEVA--ATKVGALSKGQLKEFLDANLA         109
```

(b)

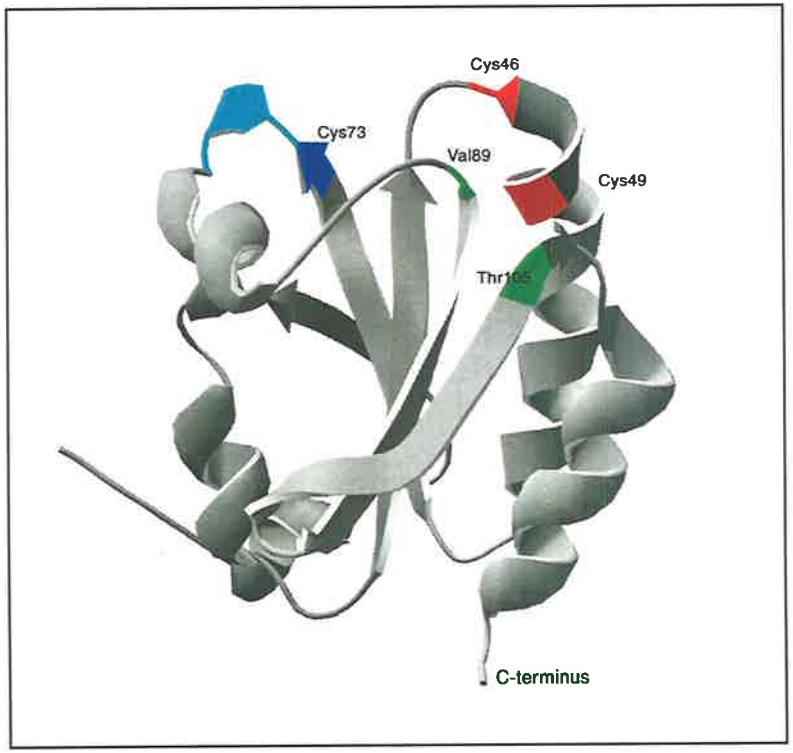


Figure 1.3(a) Comparative alignment of *Arabidopsis* thioredoxin protein sequences representing all classes of plant thioredoxins and *E. coli* thioredoxin: Sequences and accession numbers are: *Arabidopsis* h1 (Z14084), *o1* (AAC12840), f1 (AF144386), m1 (AF095749), x (AF095753), *E. coli* (AE000344). Structurally important amino acids are in blue while the active centre is boxed and in bold. Amino acids representing the predicted N-terminal signal sequence of thioredoxins o, f, m and x are in italics. The conserved C-terminal cysteine of thioredoxin f is in red and the two-amino-acid insertion of thioredoxin x is in green.

(b) Three-dimensional structure of spinach thioredoxin f showing the 'thioredoxin fold': The positions of the active site cysteines are indicated. The position of the third conserved cysteine, present in thioredoxins of the f class, is highlighted in blue. Regions targeted in mutational studies and mentioned in the text are: amino acids 74-77 (light blue); Thr105, Val89 (in green).

Chapter 2 General Materials and Methods

2.1 Plant Material

Phalaris coerulescens, Lolium perenne, Secale cereale, Hordeum bulbosum, Hordeum vulgare (cv Sloop), *Triticum aestivum* (cv Chinese spring) and *Zea mays* plant material was obtained from a glasshouse collection maintained at the Waite Campus, University of Adelaide, Australia. *Nicotiana tabacum* material was provided by Dr Ute Baumann of the University of Adelaide. Rice (*Oryza sativa* cv 'Amaroo') leaf tissue was obtained from Dr Arun Aryan, University of Adelaide.

2.2 RNA Isolation from Plant Material

With the exception of purchased reagents or unless otherwise indicated all solutions used for RNA extractions were prepared using double autoclaved nanopure H_2O then autoclaved again. All non-plastic material required for RNA preparation (mortars, pestles, glassware, spatulas etc) were wrapped in aluminium foil and sterilized by baking at 160^0C for 4 hrs. Autoclavable plasticware was sterilized by autoclaving twice whilst other plastic material (eg electrophoresis gel tanks and material used in RNA transfers) was soaked in 0.4 M KOH for 10 min then rinsed three times with double autoclaved H_2O.

2.2.1 DEPC Treatment of H_2O

H_2O was drawn into an RNase-free (baked 160^0C for 4 hrs) glass bottle and diethylpyrocarbonate (DEPC) added to 0.01% (v/v). The solution was allowed to stand for 12 hrs, then autoclaved.

2.2.2 Small Scale RNA Isolation

10-100 mg of plant material was placed in a 2 ml microcentrifuge tube together with one 5 mm and three 3.5 mm sterilized ball bearings and immersed in liquid nitrogen until frozen. The material was ground to a fine powder by 5 rounds of vortexing for 30 seconds then refreezing. RNA was extracted using 1 ml of Trizol (Gibco BRL) reagent following the manufacturers instructions then resuspended in DEPC-treated H_2O (2.2.1) and stored at -80^0C.

2.2.3 Large Scale RNA Isolation

Plant material was frozen in liquid nitrogen and ground to a fine powder using a mortar and pestle. The powder was homogenised in 3 to 4 ml of REB-buffer (10 mM Tris-HCL pH 8.0, 4% (v/v) sarcosyl, 10 mM Na_2EDTA) and the ice slurry was transferred to pre-chilled 10 ml Corex tubes and immediately centrifuged at 6000 rpm (JA20 rotor, Beckman J2-21) at 4^0C for 10 min. The supernatant was transferred to a fresh tube and caesium chloride added (1g CsCl per ml supernatant). The supernatant solution was pipetted onto 3 ml of a CsCl cushion (9.65 g CsCl dissolved in TE (10 mM Tris-HCL, 1 mM EDTA, pH 7.5) to a final volume of 10 ml) in ultracentrifuge tubes (Oakridge Polycarbonate for Ti70.1 and Ti65 rotors). The samples were centrifuged at 30,000 rpm for 16 hrs at 4^0C (rotor Ti70.1, Beckman L8-70). The supernatant was transferred to a 1.5 ml microcentrifuge tube and the RNA pellet dissolved in 400 µl REB-buffer then extracted with 400 µl of phenol/chloroform/isoamyl alcohol (25:24:1) and gentle mixing for 2 min. The two phases were separated by centrifugation at 10,000 rpm for 5 min in a bench top centrifuge. The upper phase was transferred to a fresh microcentrifuge tube and the RNA precipitated by the addition of 40 µl of 3 M sodium acetate (pH 4.8) and 1 ml of ethanol and placing at -20^0C for 2 hrs to overnight. The pellet was washed with 70% ethanol and resuspended in DEPC-treated H_2O (2.2.1) and stored at -80^0C.

2.2.4 Nucleic Acid Quantitation and Quality Determination

The concentration of RNA and DNA was determined spectrophotometrically (Shimadzu UV-160A spectrophotometer) at 260 nm. The quality of RNA and DNA was assessed visually following gel fractionation. RNA in a denaturing agarose gel (2.9.1) and DNA in a TAE agarose gel (2.12.1).

2.3 RT-PCR

First strand cDNA was synthesized from 2 µg of total RNA using Superscript II Reverse Transcriptase (Gibco-BRL) and either oligo dT, R2 or T2 as the initiation primer. The composition of each reaction was as follows; 20 mM Tris-HCL (pH 8.4), 50 mM KCl, 2.5 mM $MgCl_2$, 10 mM DTT, 30 ng R2 primer, 0.5 M of each dNTP and 200 units Superscript II RT. Reverse transcription was carried out at

45°C for 1 hr and terminated by heating to 75°C for 15 min. RNA was degraded by the addition of 2 units of RNase H and RNase T1 (Gibco-BRL) and incubation at 37°C for 10 min. cDNA was recovered by ethanol precipitation and the concentration determined by spectrometric measurement (Shimadzu UV-160A spectrophotometer). 0.1 μg cDNA was used per PCR reaction. Products of PCR reactions were visualised by agarose gel electrophoresis (2.12.1).

2.4 5'-RACE

Reverse transcription was performed as described previously (2.3) with the following exceptions; Thermoscript Reverse Transcriptase (Gibco-BRL) was the enzyme used, both R2 and T2 were used to prime first strand synthesis and the reactions were performed at 55°C and 60°C for R2 and T2 respectively. cDNA was purified using a Qiagen PCR Purification kit according to the manufacturers instructions and eluted in 30 μl of water. A homopolymeric C-tail was added to the 3' end of the purified cDNA by the addition of 1 μl of Terminal deoxynucleotidyl Transferase (Gibco-BRL) to a 24 μl reaction mix of the following composition; 10 mM Tris-HCl (pH 8.4), 25 mM KCl, 1.5 mM $MgCl_2$, 0.2 M dCTP and 10 μl of the purified cDNA. Five microliters of the tailed cDNA was used for PCR. Products of PCR reactions were visualised by agarose gel electrophoresis (2.12.1).

2.5 Preparation of Competent *Escherichia coli* Cells

Escherichia coli, strain DH5α, cells were streaked out onto an LB agar plate (1% (w/v) Bacto-tryptone, 0.5% (w/v) Bacto-yeast extract, 0.5% (w/v) NaCl, 1.5% (w/v) Agar, pH 7.5) and incubated overnight at 37°C. A single colony was collected and used to inoculate 250 ml of SOB medium (2% (w/v) Bact-tryptone, 0.5% (w/v) Bacto-yeast extract, 0.1 M NaCl, 2.5 mM KCl) in a 2-litre flask. Cells were grown to an OD_{600} of 0.6 at 18°C with vigorous shaking (200-250 rpm). The flask was chilled on ice for 10 min then the cells were transferred to pre-chilled 250 ml centrifuge bottles. Cells were pelleted by centrifugation at 2,500 rpm and 4°C for 10 min (JA14 rotor, Beckman J2-21) then resuspended in 80 ml TB buffer (0.01 M Pipes (1,4-piperazinediethanesulfonic acid), 0.055 M $MnCl_2$, 0.015 M $CaCl_2$, 0.25 M KCl, pH 6.7). Cells were allowed to sit in an ice

bath for 10 min then harvested as above. Cells were gently resuspended in 20 ml TB containing 7% (v/v) DMSO and incubated in an ice bath for a further 10 min. The suspension was then dispensed into 2 ml microcentrifuge tubes, frozen immediately in liquid nitrogen and stored at -80^{0}C.

2.6 Ligation of DNA-Sequences into Plasmid Vectors and Transformation of *Escherichia coli*

PCR fragments (2.3 and 2.4) were purified by affinity chromatography 'spinclean' microcolumns (Qiagen) and cloned into either the pGemT or pGemT EASY vector (Promega) according to the manufacturers instructions. 100 µl of competent cells (2.5) were added to the ligation mix, gently mixed and then placed on ice for 30 min. The cells were heat shocked at 42^{0}C for 1 min then 400 µl of SOC medium (SOB media (2.5) containing 0.01 M $MgSO_4$, 0.01 M $mgCl_2$ and 0.35%(w/v) glucose) was added and the tube was incubated at 37^{0}C for 1 hr with gentle shaking. Cells were plated out onto LB agar plates (2.5) containing 50 µg/ml ampicillin, 0.004% (w/v) X-gal and 0.1 mM IPTG and incubated at 37^{0}C overnight. Recombinant bacteria were identified as white colonies.

2.7 Isolation of Plasmid DNA

Recombinant bacterial colonies were picked from plates with sterile toothpicks and were inoculated into 10 ml of LB media (1% (w/v) Bacto-tryptone, 5% (w/v) Bacto-yeast extract, 1% (w/v) NaCl, pH 7.5) containing 50 µg/ml ampicillin. Cells were grown overnight at 37^{0}C with vigorous shaking then pelleted by centrifugation. Plasmids were purified from the bacterial pellet by either the alkaline lysis method of Birnboim and Doly (1979) or with a commercial plasmid purification kit (Qiagen) following the manufacturers instructions. Plasmids were resuspended in TE buffer or nanopure H_2O.

2.8 DNA Sequencing and Analysis

Sequencing of inserts cloned into the pGemT or pGemT EASY vectors (2.7) was performed by the Nucleic Acid and Protein Chemistry Unit, University of Adelaide using an Applied Biosystems model 373A DNA Sequencer. Sequencing reactions were performed as prescribed by the manufacturer. DNA sequences

were aligned by eye using the SeqED (Applied Biosystems) nucleotide-editing program.

2.9 RNA Gel Blots and Northern Hybridisation

2.9.1 Formaldehyde Gel Electrophoresis

Total RNA was fractionated in a denaturating agarose gel (1-1.2 % agarose, 0.02 M MOPS pH 7.0, 5 mM sodium acetate, 1 mM Na_2EDTA, 2.2 M formaldehyde). The RNA was prepared for loading by drying under vacuum, resuspending in 2.0 µl of Buffer A (0.5 M MOPS, 0.01 M Na_2EDTA pH 7.0) and 13.5 µl of formaldehyde/formamide/ H_2O (3.5/10/3.5), incubating at 70^0C for 10 min then chilling on ice. Just prior to loading 1.0 µl of RNA-loading buffer (322 µl Buffer A mixed with 178 µl 37 % formaldehyde, 500 µl of deionised formamide, 5 mg xylene cyanol, 5 mg bromocresol green and 400 mg sucrose) was added. Electrophoresis was performed in MOPS buffer (0.02 M MOPS pH 7.0, 5 mM sodium acetate, 1 mM Na_2EDTA). After electrophoresis the gel was stained in 2 mg/ml ethidium bromide for 10 min, destained for 1 hr in DEPC-treated H_2O and photographed using an electronic 'GelCam' documentation system (Bresatech Aust.)

2.9.2 Transfer of RNA onto Nylon Membranes

Gels were soaked in 20X SSC (3 M NaCl, 0.3 M trisodium citrate) for 15 min and the RNA transferred to a nylon membrane (Hybond-N^+, Amersham, UK) by capillary blotting for 12 hrs using 20X SSC as the transfer solution. The membrane was rinsed in 2X SSC for 2 min, air dried and baked at 70^0C for 30 min. RNA was cross linked to the membrane by exposing the membrane to 50 mJoules of short-wave UV light. Membranes were sealed in plastic bags and stored at -20^0C until used.

2.9.3 Probe Labelling

cDNA sequences generated by either PCR or excised from cloning vectors were separated on 1% TAE agarose gels. The sequence was excised under long wavelength UV light and purified using a Qiagen Gel Extraction Kit according to the manufacturers instructions then eluted in nanopure H_2O. Approximately 50 ng

of the probe was boiled with 0.9 μg of 9-mer random labelling primers (total volume 5 μ) for 5 min to denature then rapidly cooled in an ice bath. The probe/primer mix was added to 12.5 μl of ice-cold oligolabelling buffer (40 μM d(ATP, TTP, GTP), 0.1 M Tris-HCl pH 7.6, 0.1 M NaCl, 0.02 M $MgCl_2$, 200 μg/ml acetylated DNase-free bovine serum albumin (Sigma), 4 μl [α-^{32}P]-dCTP (Amersham), 2 units of DNA polymerase Klenow fragment (Boehringer Mannheim) and made up to 25 μl with ice cold nanopure H_2O. The labelling mix was incubated at 37^{0}C for 1 hr and the labelled DNA separated from unincorporated nucleotides by passing through a G-100 Sephadex column (Sambrook et al 1989).

2.9.4 Northern Hybridisation

Membranes were placed in a hybridisation bottle (Hybaid) together with 20 ml of hybridisation solution (50% (v/v) formamide, 5X SSPE (0.9 M NaCl, 0.05 M sodium phosphate, 5 mM EDTA pH7.7), 1X Denhardt's (Sambrook et al. 1989), 1% SDS, 100 μg/ml yeast RNA) and prehybridised overnight at 42^{0}C in a hybridisation oven (Hybaid). The hybridisation solution was replaced the following day before the labelled probe was added. The labelled probe was denatured in boiling water for 5 min and chilled on ice then added directly to the hybridisation bottle containing the membranes and hybridisation solution. Hybridisation was allowed to proceed for 24 hrs.

Membranes were then washed in 1X SSC, 0.1% SDS at 65^{0}C for 20 min, followed by 20 min washes at increasingly higher stringencies (0.5X SSC, 0.1% SDS and 0.2X SSC, 0.1% SDS and 0.1X SSC, 0.1% SDS). After washing the membranes were sealed in plastic and exposed to X-ray film (Fuji, Super HR-G30) at -80^{0}C for varying lengths depending on signal intensity. Films were developed in an AGFA CP1000 developer.

2.10 DNA Isolation from Plant Material

2.10.1 Small Scale DNA Isolation

Approximately 500 mg of young leaf material was placed in a 2 ml microcentrifuge tube together with one 5 mm and three 3.5 mm ball bearings and

immersed in liquid nitrogen until frozen. The material was ground to a fine powder by 5 rounds of vortexing for 30 seconds then refreezing. 750 µl of DNA extraction buffer (1% (w/v) sarcosyl, 0.1 M Tris-HCl, 0.1 M NaCl, 0.01 M Na_2EDTA, 0.1 M Na_2SO_3, pH 8.5) was added to the tube and mixed by hand inversion for 2 min. The slurry was extracted with 750 µl of phenol/chloroform/isoamyl alcohol (25:24:1) and gently mixed on an orbital rotor for 15 min. The two phases were separated by centrifugation at 10,000 rpm for 5 min in a bench top centrifuge. The aqueous phase was transferred to a fresh microcentrifuge tube and the phenol/chloroform/isoamyl alcohol extraction repeated. 500 µl of the aqueous phase was transferred to a fresh microcentrifuge tube and the DNA precipitated by the addition of 50 µl of 3 M sodium acetate (pH 4.8) and 1.5 ml of ethanol and incubation at room temperature for 1 hr. The precipitation solution was carefully removed by pipette and the DNA washed with 1 ml of 70% (v/v) ethanol. The DNA was pelleted by centrifugation (10,000 rpm for 5 min), the ethanol removed by pipette and the pellet allowed to air dry. The DNA was resuspended in 30-50 µl of TE (2.2.3) containing 40 µg/ml of RNase A then stored at -20^{0}C until used.

2.10.2 Large Scale DNA Isolation

Approximately 2g of leaf material was frozen in liquid nitrogen and ground to a fine powder using a mortar and pestle. The powder was homogenised in 4 ml of DNA extraction buffer (2.6.1) and the ice slurry was transferred to a 10 ml screw-cap tube (Sarstedt). Phenol/chloroform/isoamyl alcohol (25:24:1) (4 ml) was added and the tube placed on an orbital mixer for 15 min. The tube was then centrifuged at 5000 rpm (JA20.1 rotor, Beckman J2-21) for 10 min. The supernatant was transferred to 10 ml serum gel tube (Sarstedt) and re-extracted with another 4 ml of phenol/chloroform/isoamyl alcohol (10 min on an orbital mixer) then centrifuged at 5000 rpm for 5 min. The supernatant was decanted into a fresh 10 ml tube and DNA precipitated by the addition of 400 µl of 3 M sodium acetate (pH 4.8) and 4 ml of isopropanol. The DNA was spooled around the end of a sterile pasteur pipette and transferred to a 2 ml microcentrifuge tube. Excess supernatant was removed and the pellet washed with 2 ml of 70% (v/v) ethanol.

The pellet allowed to air dry and the DNA was resuspended in 300 µl of TE (2.2.2) containing 40 µg/ml of RNase A then stored at -20^{0}C until used.

2.11 Digestion of Genomic DNA with Restriction Endonucleases

DNA was digested in a solution of the following composition; 15-20 µg DNA, 2 µl of 0.04 M spermidine, 2 µl of 1 mg/ml acetylated bovine serum albumin (Boehringer Mannheim), 10 units of restriction endonuclease (Promega), 2 µl of the corresponding 10X restriction buffer (Promega) and H_2O to 20 µl. Digests were incubated at 37^{0}C for 3 hrs then an additional 10 units of restriction enzyme added and the digest allowed to proceed for a minimum of 2 hrs.

2.12 DNA Gel Blots and Southern Hybridisation

2.12.1 Agarose Gel Electrophoresis

DNA samples were mixed with 1/10 vol of 10X Ficoll dye (0.1 M Tris-HCl, 0.2 M Na_2EDTA, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol, 30% (w/v) ficoll type 4000, pH 8.0) and fractionated on a 1% agarose gel in TAE buffer (0.04 M Tris-acetate, 1 mM Na_2EDTA, pH8.0) at 20-100 mA for 4-16 hrs. After electrophoresis, gels were stained in 10 µg/ml ethidium bromide for 15 min, visualised under UV light and photographed.

2.12.2 Transfer of DNA onto Nylon Membranes

DNA was transferred to a nylon membrane (Hybond-N$^+$, Amersham, UK) by capillary blotting (Southern 1975) for 5-12 hrs using 0.4 M NaOH as the transfer solution. Following transfer the membrane was rinsed in 2X SSC (2.5.2) and dried by careful blotting on filter paper (Whatman 3MM). Membranes were sealed in plastic bags and stored at 4^{0}C until used.

2.12.3 Southern Hybridisation

Membranes were prehybridised for 12-20 hrs at 65^{0}C in a hybridisation bottle (Hybaid) containing 2 ml 5X HSB (3M NaCl, 0.1 M PIPES, 0.025 M Na_2EDTA pH 6.8), 1 ml Denhardts III (10% (w/v) SDS, 5% (w/v) tetrasodium pyrophosphate, 2% (w/v) gelatine, 2% (w/v) ficoll, 2% (w/v) polyvinyl pyrollidone), 1 ml salmon sperm DNA (5 mg/ml) and 6 ml H_2O. The

prehybridisation solution was replaced with 10 ml of hybridisation solution (4 ml H_2O, 2 ml 5XHSB, 2 ml Denhardts III, 1 ml 25% (w/v) dextran sulfate and 1 ml salmon sperm DNA (5 mg/ml)) and allowed to warm up to 65^0C before the probe was added. The labelled probe, prepared as described previously (2.9.3), was denatured in boiling water for 5 min and chilled on ice then added directly to the hybridisation bottle. Hybridisation was performed at 65^0C overnight. Membranes were then washed in 2X SSC, 0.1% SDS at 65^0C for 20 min, followed by 20 min washes at increasingly higher stringencies (1X SSC, 0.1% SDS and 0.5X SSC, 0.1% SDS and 0.2X SSC, 0.1% SDS). After washing the membranes were sealed in plastic and exposed to X-ray film (Fuji, Super HR-G30) at -80^0C for varying lengths depending on signal intensity. Films were developed in an AGFA CP1000 developer.

2.13 Removal of Radioactive Probes from Membranes

Southern membranes were stripped by pouring hot (90^0C) stripping solution (0.1% SDS, 2 mM Na_2EDTA, pH 8.0) onto them then incubating at 65^0C for 30 min. Northern membranes were stripped by pouring boiling 0.1% SDS onto them then incubating them at room temperature for 30 min. All membranes were re-exposed to X-ray film for 5 days to ensure that the probe had been removed.

Chapter 3 Clarification of the Structure of *Bm2* and its Role in SI

3.1 Introduction

Self incompatibility (SI) is widespread in the grass family (Graminae) and is controlled by two multiallelic loci, *S* and *Z*, which segregate independently (reviewed in Hayman 1992). Previous reports suggested that a cDNA called *Bm2*, isolated by differential screening of a mature pollen library of *Phalaris coerulescens*, represents a putative *S*-gene (Li et al. 1994). Since *Bm2* was not full length the first exon and intron of the gene were predicted from the genomic sequence of two alleles (S_1 and S_2). Based upon these sequence analyses it was proposed that the gene was composed of two domains. The 5' half of the gene was found to contain most of the nucleotide differences and was labelled the allelic domain, whilst the 3' region of the gene sequence was completely conserved and was named the thioredoxin domain as it shared similarity with plant thioredoxin *h* genes. In a subsequent study, a sequence orthologous to part of the conserved thioredoxin domain was amplified by RT-PCR from several SI and self-compatible (SC) grass species suggesting the presence of similar genes in these species (Li et al. 1997). Conversely, RT-PCR with primers designed to the allelic domain failed to amplify a corresponding product in other grasses, a result interpreted as supporting the premise that this region of the gene determined allelic identity.

Allelic variability is crucial for the recognition process in SI. Not surprisingly, proteins shown to be involved in single-locus SI systems, such as the S-RNases in the gametophytic system of the *Solanaceae, Rosaceae and Scrophulariaceae* as well as the S-receptor kinase (SRK) and S-locus glycoprotein (SLG) in the sporophytic SI reaction of the *Brassicaceae* show a high degree of amino acid sequence polymorphism. For example, within the *Rosaceae* sequence similarities ranged from 29.2% to 82.2% (Ushijima et al. 1998). For the S-RNases and SRK/SLG proteins it was found that sequence variations were clustered and several hypervariable and conserved sequence blocks have been identified. Clustering of amino acid variation is, however, not necessarily the rule since the S-proteins involved in the gametophytic SI response of the *Papaveraceae* show

amino acid variation (51.3% to 63.7% identity) that is dispersed throughout the protein (Kurup et al. 1998). Mechanisms responsible for the generation of *S*-allele polymorphism also differ between SI systems. *S*-allele diversity in the sporophytic Brassica system appears to be the result of both point mutations and intragenic recombination (Nasrallah 1997; Kusaba et al. 1997), whilst point mutations alone appear to be the primary generator of new alleles in gametophytic systems (Saba-El-Leil et al. 1994; Matton et al. 1997; Ishimizu et al. 1998). In contrast to *S*-alleles from other gametophytic SI systems, allelic variation between the S_1 and S_2 alleles of *Phalaris* was found to be comparatively low (92% predicted amino acid identity) and was due, almost entirely, to a frameshift caused by a single base deletion and insertion in the predicted allelic domain region (Li et al. 1994). Moreover, unlike many of the *S*-genes published to date, the *Phalaris S*-alleles did not display readily distinguishable, allele determining, hypervariable domains.

An additional contrasting feature of the *Phalaris Bm2* gene and *S*-genes from other species is their spatial expression pattern. All *S*-genes studied to date are expressed in either pollen or stigma tissue depending upon whether they encode the male or female SI determinant respectively (Sims 1993; Foote et al. 1994; Schopfer et al. 1999). By contrast, *Bm2* is expressed in both pollen and stigma tissue (Li et al. 1994; Baumann et al. 2000). Previous *in-situ* hybridisation experiments have reported that *Bm2* mRNA is highly expressed in response to a self-incompatible, but not self-compatible, pollination (Baumann et al. 2000). The apparent self-incompatible-dependant induction of *Bm2* mRNA has been interpreted as indicating a role for *Bm2* in SI. However, there are alternative explanations that have not as yet been tested. SI in *Phalaris* is regularly associated with the rupture of incompatible pollen tubes, thus the *Bm2* mRNA detected at the site of SI may have been released by the burst pollen tube rather than expressed by the surrounding stigma tissue. Furthermore, during the preparation of stigmas for *in-situ* hybridisation it was discovered that mechanical damage also induced *Bm2* expression (Baumann et al. 2000). In addition, oxidative stress has since been implicated in the gametophytic SI response of lily (Tezuka et al. 1997). Should SI in *Phalaris* similarly elicit a stress response it is possible that *Bm2* is expressed in response to this stress.

In order to clarify the structure of *Bm2* and determine the level of *S*-gene diversity in a two-locus SI system, full-length *Bm2* homologues were isolated from three SI grass species known to have the *S-Z* system and two further alleles from *Phalaris coerulescens*. In addition, the induction of *Bm2* expression in response to mechanical damage and SI was examined further.

3.2 Materials and Methods

3.2.1 Isolation of *Bm2* Homologues

RT-PCR was utilised to isolate *Bm2* homologues from the SI grass species *Secale cereale, Hordeum bulbosum, Lolium perenne* and two additional *Bm2* alleles from *Phalaris coerulescens*. The *Phalaris* S_1 and S_2 sequences published by Li et al. (1994) were used as a template for the design of PCR primers (Figure 3.2a). PCR primer sequences are given in Appendix A.

3.2.1.1 RNA Extraction

Total RNA was extracted from the mature pollen of all species and purified by CsCl gradient (general materials and methods 2.2.3).

3.2.1.2 PCR Reactions

Four RT-PCR (Reactions 1 to 4) and two 5'RACE (Reactions 5 and 6) reactions were performed on the RNA of each species (see general materials and methods 2.3 and 2.4):

Reaction 1; first strand cDNA was synthesised using oligo-dT as the initiation primer followed by PCR with the A1/R2 primer pair (PCR Protocol 1, Appendix B).

Reaction 2; reverse transcription with R2 as the initiation primer followed by PCR with the A1/T2 primer pair (PCR Protocol 2, Appendix B).

Reaction 3; reverse transcription with oligo-dT as the initiation primer followed by PCR with the R1/R2 primer pair (PCR Protocol 1, Appendix B)

Reaction 4; reverse transcription with R2 as the initiation primer followed by PCR with the R1/T2 primer pair (PCR Protocol 3, Appendix B).

Reaction 5; reverse transcription with R2 as the initiation primer followed by the addition of a homopolymeric C-tail then PCR with the T2 primer and an anchored primer AAP (specific to the homopolymeric tail) (PCR Protocol 4, Appendix B).

Reaction 6; reverse transcription with T2 as the initiation primer followed by the addition of a homopolymeric C-tail then PCR with the T1 primer and AAP (PCR Protocol 4, Appendix B).

Reactions 1 and 2 were designed to amplify full-length cDNA's and were based upon the assumption that the predicted first exon sequence of the S_1 and S_2 alleles was correct. Reactions 3 and 4 were designed to amplify the majority of the thioredoxin domain region of the gene using primers targeted to the *Bm2* cDNA sequence. Reactions 5 and 6 were 5'RACE reactions employing nested primers targeted to the *Bm2* cDNA sequence, and therefore did not require the first exon to have been correctly predicted. The possibility of amplifying contaminating genomic DNA in the 5'RACE reactions was avoided by designing the T1 primer such that it covered an intron-exon junction.

3.2.1.3 cDNA Cloning and Sequencing

The PCR products amplified from each grass were subcloned into pGemT vectors (2.6) and competent *E. coli* cells transformed with the recombinant plasmids (2.6). Recombinant bacteria were grown and plasmids harvested (2.7). Two independent clones of each product were sequenced from each direction (2.8). To ensure that sequence accuracy was achieved, reactions were repeated from the point of reverse transcription. Therefore, each sequence is the result of four individually generated sequences.

3.2.2 Analysis of *Bm2* Expression

Two expression studies were conducted to clarify the expression characteristics of *Bm2* in response to mechanical damage and self-incompatible pollination. Previous *in-situ* hybridisations were performed on stigmas subjected to an *in-vitro* pollination assay developed by Hayman (1956). For the assay, intact mature stigmas were removed from *Phalaris* florets and planted onto 2% agar plates containing 10% (w/v) sucrose and 100 ppm boric acid. Stigmas were then pollinated with *Phalaris* pollen sourced from plants of different *S* and *Z* genotypes and the SI phenotype assessed by viewing the pollinated stigmas under a light microscope. To ensure that the results obtained in this study could be compared with previous *in-situ* results, the stigmas used for pollination experiments were treated in an identical manner (Figure 3.9).

3.2.2.1 Mechanical Damage

Phalaris stigmas (1 day pre-anthesis) were mechanically damaged *in-planta*. The top of the floret was cut off and the stigma damaged by gentle squeezing with forceps. Stigmas were left for 3 hrs then harvested into 2 ml microcentrifuge tubes and frozen in liquid nitrogen. Control (undamaged) stigma material was derived from florets treated in an identical manner with the exception of forcep damage. Approximately 50 stigmas were harvested for each treatment.

3.2.2.2 Self-incompatible and Self-compatible Pollinations

Mature *Phalaris* stigmas (1 day pre-anthesis) were carefully excised from florets and planted onto agar plates (approximately 50 stigmas per plate). Plates were sealed and stored at 4°C overnight. The following day, plates were allowed to warm up to room temperature (approximately 22 °C) then the stigmas were pollinated with fresh pollen from a *Phalaris* plant of either the same *S* and *Z* genotype (self-incompatible) or a different genotype (self-compatible) or with pollen from wheat (cross-incompatible). Pollinated stigmas were held at room temperature for 3 hrs then harvested into 2 ml microcentrifuge tubes and frozen in liquid nitrogen. Control stigmas were unpollinated.

3.2.2.3 RNA Extractions

Total RNA was extracted from stigmas and mature anthers (1 day pre anthesis) using the small-scale RNA isolation method (2.2.2).

3.2.3.4 Amplification of *Phalaris* Peroxiredoxin (*PcPrx*)

Peroxiredoxin was amplified by RT-PCR from *Phalaris* anther RNA. The sequences of rice (Genbank accession number AU031625), wheat (BE427212) and barley (BQ468113) type-C peroxiredoxins (Rouhier et al. 2001) were used as the templates for the design of primers. First strand cDNA was synthesised using oligo-dT as the initiation primer (2.3) followed by PCR with the PrxF/PrxR primer pair (PCR Protocol 5, Appendix B). Subcloning and sequencing of the PCR product were carried out as described previously. PCR primer sequences are given in Appendix A.

3.2.3.5 RNA Transfer and Northern Hybridisation

RNA concentrations were determined spectrophotometrically (2.2.4) and 12 μg of RNA fractionated in a 1.2% denaturating agarose gel (2.9.1). RNA was then transferred to nylon membranes by capillary blotting (2.9.2) The *Phalaris PTrx1* and *PcPrx* cDNAs amplified by RT-PCR were radioactively labelled (2.9.3) and used as the probes for hybridisation (2.9.4). Membranes were washed three times to a final stringency of 0.2X SSC, 0.1% SDS (2.9.4).

3.3 Results

3.3.1 Isolation of *Bm2* Homologues

3.3.1.1 Analysis of PCR's

PCR's incorporating primer A1 (Reactions 1 and 2) failed to amplify a cDNA product regardless of the PCR conditions, even with *Phalaris* cDNA as the template. By contrast, PCR Reactions 3 to 6 yielded cDNA products. PCR's 3 and 4 amplified products of identical size from the cDNA of all species whilst the size of the product amplified in 5'RACE Reactions 5 and 6 varied slightly between species. PCR's 3, 4, 5 and 6 were designed to generate four overlapping products (Figure 3.1b). The sequences reported here represent contigs generated from the overlapping PCR products.

3.3.1.2 Nucleotide Sequence Analysis

Entire sequences were generated for all species and are presented as full sequence alignments in Appendix C. The cDNA's range in size from 939 bp for *Secale cereale* to 1024 bp for *Lolium perenne* and displayed 77.3-91.5% nucleotide identity. The two *Phalaris* alleles show the highest level of homology. *PTrx1* differs from *PTrx2* by 6 nucleotides; a three-base deletion 199 bases upstream of position +1 and three substitutions located 117 and 106 bases upstream and at position +339. Across all species, the most striking observation is that the 3' half, defined by nucleotides +1 to +396 (Appendix C), displays a very high level of conservation (95.9-98.7% nucleotide identity). A small number of single base substitutions constitute the only differences observed. This finding is consistent with the high level of homology reported by Li et al. (1997) when partial sequences from this region were compared. Conversely, within the 5' region, a

higher level of divergence is evident. Sequence homology is restricted to completely conserved blocks separated by extensive insertions, deletions and areas of multiple single base substitutions. The 5' divergence is most evident in the *Secale* sequence, which displays the highest level of variability in this section when compared to the other grasses investigated here.

3.3.1.3 Analysis of ORF's and Protein Alignments

The nucleotide sequences were analysed for the presence of open reading frames (ORF's) using Lasergene (DNASTAR Inc.). In all cases a single ORF was identified. The ORF starts at a conserved initiation codon (given as position +1 in the sequence alignment, Appendix C) and translates into a protein of 131 amino acids. Upstream of the initiation codon seven in-frame stop codons are found in *Secale*, three in *Hordeum,* two in *PTrx1* and *PTrx2* and one in *Lolium* (Figure 3.2). These confirm the position of the initiation codon and demonstrate that the 5' untranslated leader sequence contributes approximately 60% of the overall length of these mRNA's.

The deduced protein sequences for the five genes are shown in Figure 3.3. As indicated by the nucleotide homology, the level of sequence identity across these proteins was found to be very high. The two *Phalaris* sequences are identical, whereas *Secale, Hordeum* and *Lolium* show 95.4%, 97.7% and 95.4% amino acid identity with the two *Phalaris* proteins respectively. Although some nucleotide variation was detected in the coding region, as high as 23 differences in the *Secale* sequence when compared with the *Phalaris* sequences, very few differences result in amino acid changes (six in the case of *Secale*, of which three are nonsynonymous).

Comparing the data obtained in this study with the previously published *Phalaris* S_1 and S_2 sequences, a discrepancy in the region of the first intron and second exon was found (Figure 3.4). The cDNA sequence of the 5' end of the genes isolated in this study identified regions of homology to, and terminating within, the first intron of the *Phalaris* S_1 and S_2 sequences reported by Li et al. (1994). While the genes amplified in this experiment represent full length sequences, the first exon and intron of the S_1 and S_2 sequences were predicted and as such would

appear to be incorrect. This observation is further supported by the inability of PCR's employing the A1 primer (targeted to the predicted first exon) to amplify a product. Significantly, the S_1 and S_2 sequences contain a methionine residue in an identical position to the initiation codon of the genes isolated here. When this residue is taken as the start codon, the S_1 and S_2 sequences translate into proteins of identical size and sequence to those of *PTrx1* and *PTrx2*.

3.3.1.4 Database Searches

Database searches using BLAST (Altschul et al. 1997) were performed on the sequences and revealed that all proteins shared amino acid similarity with the *h* class of plant thioredoxins. The thioredoxin *h* proteins identified in the database as having the highest sequence identity were from wheat (Gautier et al. 1998) and rice (Ishiwatari et al. 1995) with 36% and 34% identity, respectively. A comparative alignment of the *Phalaris* PTrx1 protein with the thioredoxin *h* sequences of wheat and rice is given in Figure 3.5. Significantly four thioredoxin-like genes, three plant expressed sequence tags (ESTs) and one gene located on a bacterial artificial chromosome (BAC) clone, were also identified from the database as having high levels of sequence similarity to the genes isolated in this study. The ESTs from *Lycopersicon esculentum, Glycine max* and *Pinus taeda* and the BAC gene from *Arabidopsis thaliana* were predicted to share 62%, 62%, 59% and 65% identity with the *Phalaris* PTrx1 protein, respectively (Figure 3.6). In the case of *G. max* and *A. thaliana* this represents a higher level of identity than that detected when these genes are compared with the two *G. max* and five *A. thaliana* thioredoxin *h* sequences that have been published (Shi and Bhattacharyya, 1996; Rivera-Madrid et al. 1995). Therefore, it would appear that the EST and BAC sequences represent other plant homologues of the grass genes reported here.

3.3.2 Analysis of *Bm2* Expression

The *in-situ* observation of *Bm2* expression in response to damage (Baumann et al. 2000), together with the finding that SI in lily is associated with oxidative stress (Tezuka et al. 1997), raises the question of whether *Bm2* expression was induced by oxidative stress. In order to determine whether the tissue used for Northerns was under oxidative stress, membranes were also probed with a *Phalaris* type-C

peroxiredoxin cDNA fragment amplified from stigma RNA. The rationale for choosing the type-C peroxiredoxin was: 1) peroxiredoxins have been shown to be involved in the response to oxidative stress in several organisms (Baier and Dietz, 1999; Rhee et al. 1999; Lewis et al. 2000); 2) the type-C peroxiredoxin was found to be present in reproductive tissue EST libraries of several grass species; 3) type-C peroxiredoxins isolated from three plant species have been shown to reduce H_2O_2 in the presence of the NTR/thioredoxin system (Choi et al. 1999; Verdoucq et al. 1999; Rouhier et al. 2001).

3.3.2.1 Amplification of a *Phalaris* Type-C Peroxiredoxin cDNA

The *Phalaris* peroxiredoxin cDNA probe (*PcPrx*) was amplified from stigma RNA by RT-PCR. Visualisation of the PCR product in an agarose gel (Figure 3.7a) revealed a band of the expected size (327 bp). The identity of the PCR product was confirmed by sequencing and comparison with the cDNA sequence of wheat type-C peroxiredoxin (Figure 3.7b). The *Phalaris* sequence displayed 96% nucleotide similarity with the corresponding region of wheat and 86%, 97% and 84% with the same region of rice, barley and maize, respectively (Appendix D) indicating that this type of peroxiredoxin is highly conserved in the grasses.

3.3.2.2 Expression of *Bm2* (*PTrx1*) in Response to Mechanical Damage

Phalaris stigmas were mechanically damaged *in-planta* and the expression profile of PTrx1 and *PcPrx* determined by Northern blotting. Mechanical damage of *Phalaris* stigmas was found to result in a large increase in the expression level of both *PTrx1* and *PcPrx* mRNAs (Figure 3.8). The induction of *PcPrx* expression indicates that the damage was sufficient to cause oxidative stress. Moreover, both *PTrx1* and *PcPrx* were rapidly induced since the time between the application of damage and the sampling of tissue was relatively short (3hrs). These results confirm the previous *in-situ* hybridisation findings and suggest that the damage of *Phalaris* stigmas, both *in-planta* and in the plate assay, results in oxidative stress.

3.3.2.3 Expression of *Bm2* (*PTrx1*) in Response to SI

Self-incompatible and self-compatible pollinations were recreated in an *in-vitro* assay (Figure 3.9) and the expression patterns of *PTrx1* and *PcPrx* in response to both types of pollination assessed by Northern hybridisation (Figure 3.10).

Hybridisation with the *PTrx1* and *PcPrx* cDNA probes revealed that both transcripts were strongly upregulated in stigmas derived from plate assays (Panel ii and iii) when compared with their endogenous expression level (Panel i). Significantly, there was found to be no difference in the expression level of *PTrx1* in stigmas pollinated with either incompatible or compatible pollen. Therefore, the *in-situ* findings of massive *Bm2* expression at the site of a self-incompatible pollination do not indicate the induction of expression in the surrounding stigma tissue.

The level of *PTrx1* and *PcPrx* expression in the control (unpollinated) stigmas from the plate assay was marginally lower than that of pollinated stigmas. However, as both transcripts are highly expressed in mature anthers (Panel i), the slightly higher level of expression detected in pollinated tissue is most likely due to the additional mRNA from the pollen in these samples. Moreover, since the pollination contributed little to the expression level of both transcripts in stigma, it is likely that the *in-vitro* assay method was the primary cause of transcript accumulation. The substantial expression of *PTrx1* and *PcPrx* in stigmas derived from the plate assay is not dissimilar to that recorded for the mechanical damage Northerns (Figure 3.8). The increased expression of *PcPrx* in particular, suggests that the plate assay method similarly causes oxidative stress in *Phalaris* stigmas.

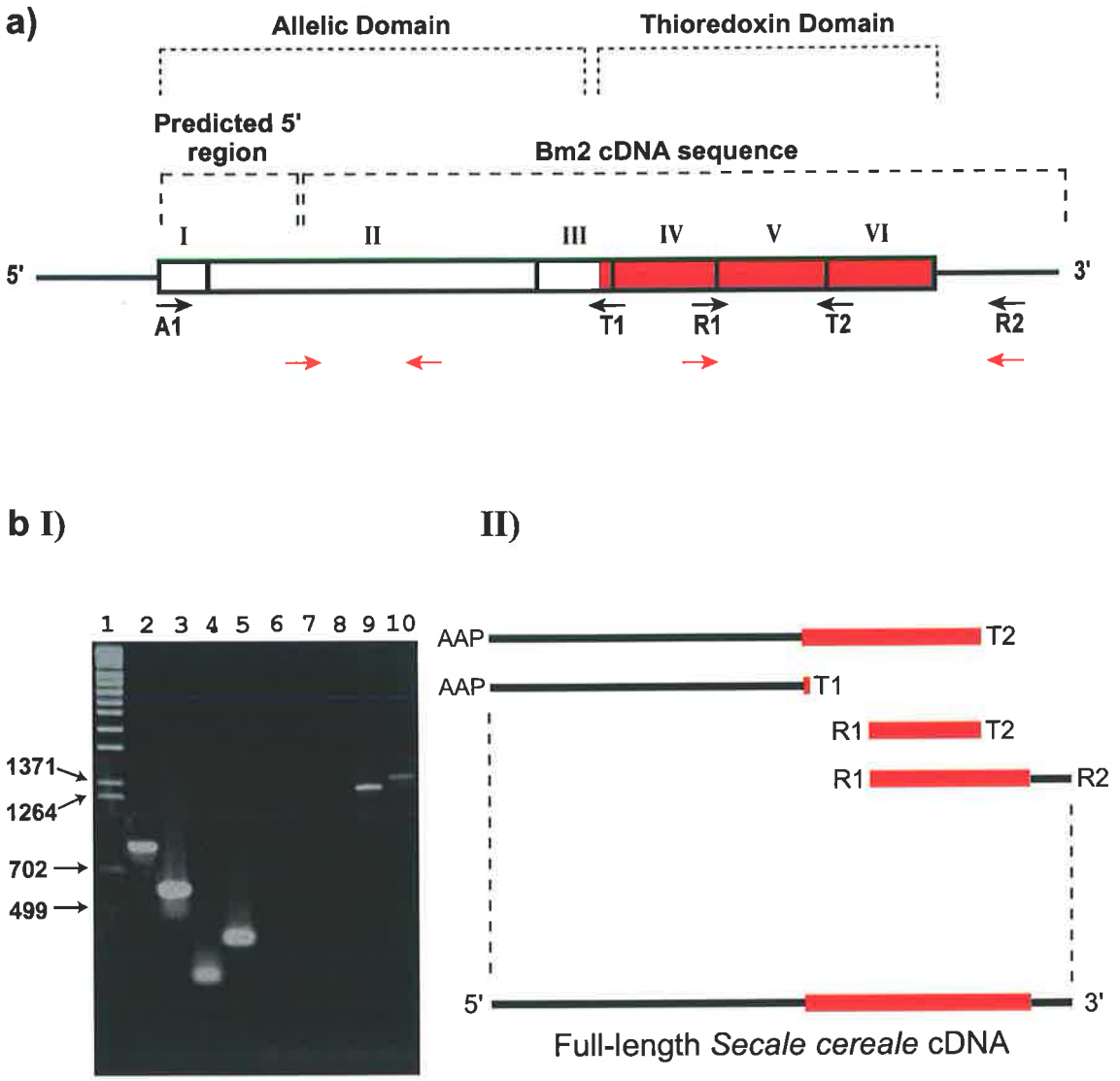


Figure 3.1 **a)** Location of PCR primers used in this study with respect to the published *Phalaris* S_1 sequence. Roman numerals represent exons and the positions of the proposed domains are indicated (the thioredoxin domain is also in red). Primers T1, T2 and R1 correspond to the putative thioredoxin domain, R2 is located in the 3' UTR and the upstream primer A1 was based upon the predicted first exon sequence. The known and predicted regions of the gene are indicated. Red arrows correspond to the PCR primers employed by Li et al. (1997) that are mentioned in the Introduction. **b)** The PCR strategy employed to amplify the *Bm2* homologue from *Secale cereale*. **I)** Agarose gel of the PCR products amplified using the following primer sets: Lane 2, AAP/T2; Lane 3, AAP/T1; Lane 4, R1/T2; Lane 5, R1/R2. Lanes 6-10 are controls; Lane 6, AAP/T2 (no template cDNA); Lane 7, U2/T2 (no template cDNA); Lane 8, R1/R2 (no template cDNA); Lane 9, R1/T2 (genomic DNA as template); Lane 10, R1/R2 (genomic DNA as template). Lane 1 is a size marker (band sizes in bp are indicated). **II)** Schematic illustration of the PCR products of lanes 2 to 5 and their positions in the full length cDNA contig. Red bars designate the thioredoxin domain region.

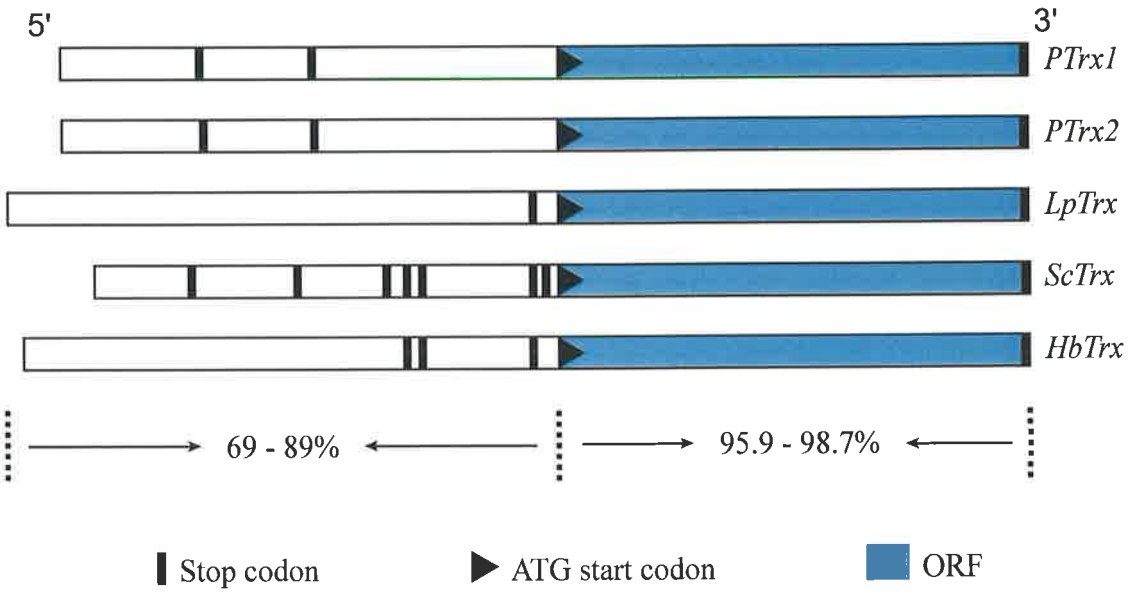


Figure 3.2
Schematic representation of the alignment of cDNA sequences from *Phalaris coerulescens* (*PTrx1, PTrx2*), *Lolium perenne* (*LpTrx*), *Secale cereale* (*ScTrx*) and *Hordeum bulbosum* (*HbTrx*). The positions of the common open reading frame and upstream stop codons are indicated. The range of nucleotide sequence homology for areas mentioned in the text is also given.

```
PTrx1,2   MGGCVGKDRGIVEDKLDFKGGNVHVITTKEDWDQKIAEANKDGKI 46
HbTrx     ---------S-------------------------V---------
LpTrx     ---------S-------------------------V---------
ScTrx     -------G-S---E----------------------E--------

PTrx1,2   VVANFSASWCGPCRVIAPVYAEMSKTYPQLMFLTIDVDDLVDFS  92
HbTrx     ------------------------------------M---
LpTrx     ------------------------------------M---
ScTrx     ----------------V-----------------------M---

PTrx1,2   STWDIRATPTFFFLKNGQQIDKLVGANKPELEKKVQALGDGS   131
HbTrx     ----------------------------------------
LpTrx     ----------------L--------R---------I----
ScTrx     ----------------------------------------
```

Figure 3.3
Alignment of the deduced amino acid sequences from *Phalaris coerulescens* (PTrx), *Hordeum bulbosum* (HbTrx), *Lolium perenne* (LpTrx) and *Secale cereale* (ScTrx). Dashes represent identical amino acids.

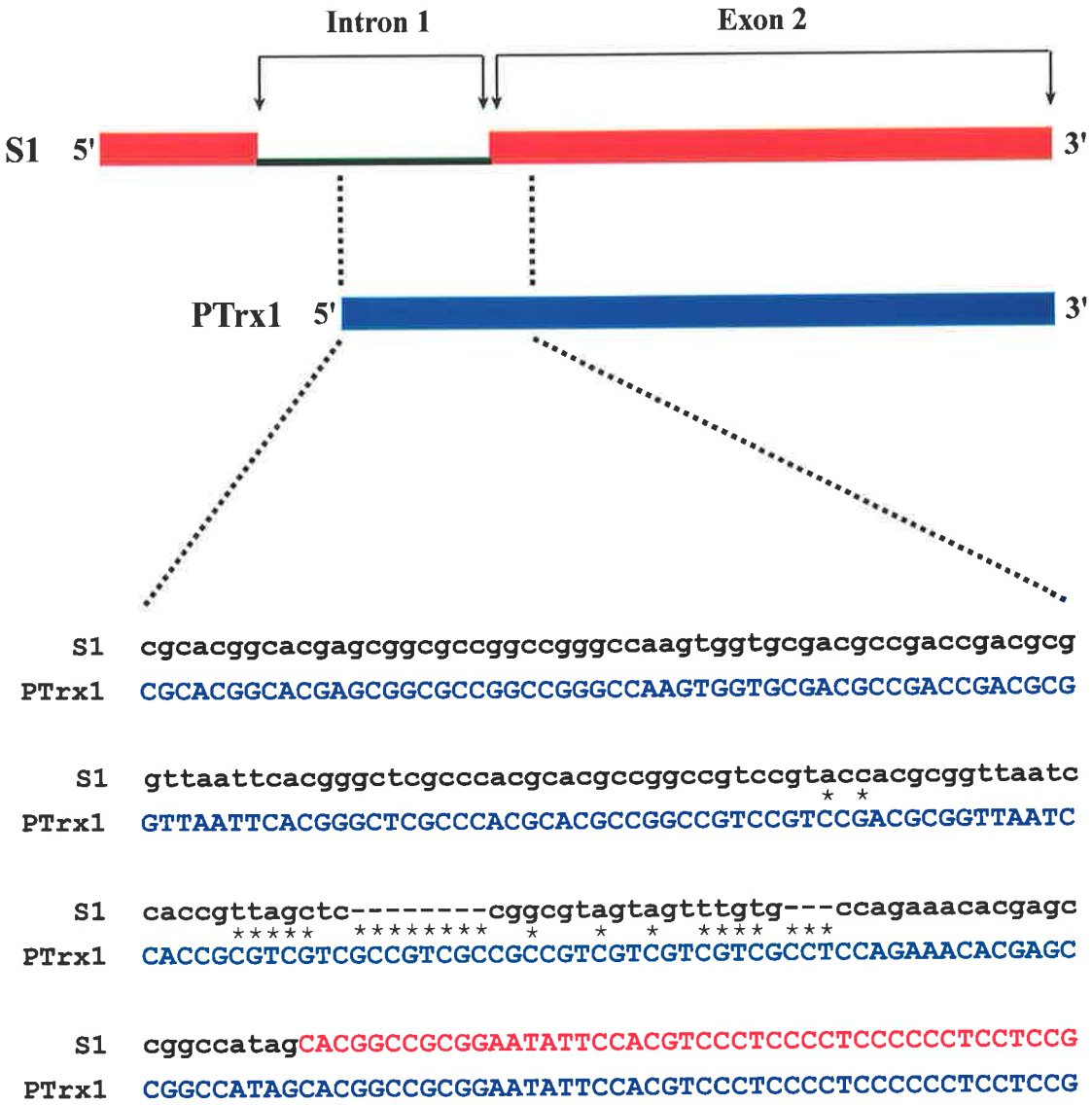


Figure 3.4
Comparative alignment of the 5' end of the *PTrx1* sequence with a homologous region covering the predicted first intron/second exon boundary of the *Phalaris* S_1 sequence. The sequence corresponding to the predicted first intron of S_1 is in lowercase letters whilst the predicted second exon is in uppercase and red. Different nucleotides are identified by an asterix. Gaps in the alignment are indicated by dashes.

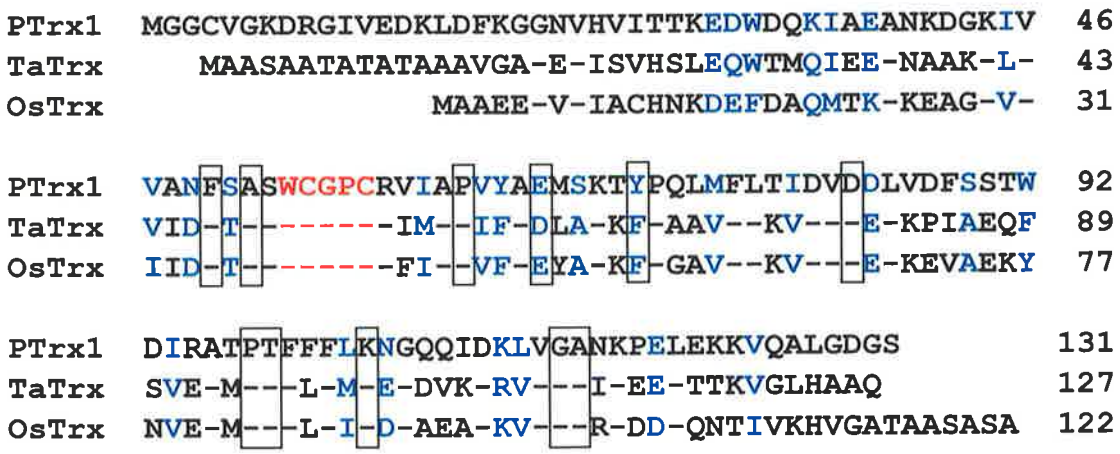


Figure 3.5
Alignment of the *Phalaris coerulescens* (PTrx1) amino acid sequence with the thioredoxin *h* sequences of wheat (TaTrx) and rice (OsTrx). Dashes indicate identical residues. The thioredoxin active site motif is in red and amino acids reported to be important for the maintenance of thioredoxin protein structure (Eklund et al. 1991) are boxed. Residues that are chemically similar across the three sequences are highlighted in blue.

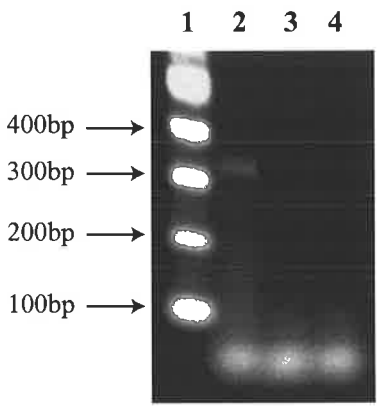
```
PTrx1                                MGGCVGKDRGIVEDKLDFKGGNVHVITTKEDWD   33
LeTrx   MGISDTVRSLFPCIKSHSTSDGDDSTHNVEFAG--VSL--T-ES--                   46
GmTrx                MGNCLRKAHADDDSDHIVELAS--VQL--T-ES--                 35
AtTrx                MGSCVSKGKGDDDSVHNVEFSG--VHL--T-ES--                 35
PtTrx                                       MAVAH--MHV--S-QE--           18

PTrx1   QKIAEANKDGKIVVANFSASWCGPCRVIAPVYAEMSKTYPQLMFLT                   79
LeTrx   Q-LAE-KKE---VIAN-S-S-----RMIS-FYC-L-EKYLSLM--T                   92
GmTrx   Q-LDQ-RKE---VIAN-S-T-----KVIA-HYC-L-VKYPSIM--L                   81
AtTrx   D-LAE-DRD---VVAN-S-T-----KIVA-FFI-L-EKHSSLM--L                   81
PtTrx   A-IFE-NTN---IVVD-W-S-----KMIA-FYA-L-EKYPQLV--K                   64

PTrx1   IDVDDLVDFSSTWDIRATPTFFFLKNGQQIDKLVGANKPELEKKVQ                  125
LeTrx   V---ELTEF-SS--IK-T---F-L-D-EQID----A-K---Q--IT                  138
GmTrx   V---ELADF-TS--IK-T---F-L-D-KEVD----A-K---E--IV                  127
AtTrx   V---ELSDF-SS--IK-T---F-L-N-QQIG----A-K---Q--VT                  127
PtTrx   V---EMAEV-AE--VR-M---I-I-D-KQID----L-Q---E--VL                  110

PTrx1   ALGDGS              131
LeTrx   TVADRHVACEPQPQPQ    154   (62%/78%)
GmTrx   VVNDVVPHKQ          137   (62%/78%)
AtTrx   SIIDSVPESPQRP       140   (65%/82%)
PtTrx   NCVAMTQSA           119   (59%/80%)
```

Figure 3.6
Comparative alignment of the deduced amino-acid sequence of *Phalaris* PTrx1 with homologous EST sequences from *Lysopersicon esculentum* (LeTrx; GenBank accession number AW092362), *Glycine max* (GmTrx; AW101975), *Pinus taeda* (PtTrx; AW043121) and an *Arabidopsis thaliana* homologue (AtTrx; located on BAC clone AC012562). Dashes indicate identical amino acids, the thioredoxin active site sequence is in red and chemically similar amino acids are in blue. The percentage protein identity and similarity of each gene, relative to the PTrx1 sequence, is given in brackets at the end of the sequence (%identity / %similarity).



(b)

```
TaPrx ATGGCTCCGATTGGCGTGGGCAGCACCCTCCCCGACGGCCAGCTCGCGTGGTTCGACGA  59

TaPrx GAGCGAACAGATGCAGCAGGTCTCCATCCACTCCCTGGCCGCCGGCAAGAAGGTCATCC 118

                                                    PrxF
TaPrx TCTTCGGCGTCCCTGGCGCCTTCACCCCCACCTGCAGCAATCAGCATGTACCAGGCTTC 177
PcPrx                        ACCCCCACCTGCAGCAATCAGCATGTACCAGGCTTC  36

TaPrx ATTACTCAGGCCGAGGAGCTCAAAGCCAAGGGTGTAGATGAGATCCTGCTTGTCAGCGT 236
PcPrx ATTACTCAAGCTGAGGATCTCAAAGCCAAGGGTGTAGAGGAGATTCTTCTTGTCAGCGT  95

TaPrx TAATGACCCCTTTGTCATGAAAGCATGGGCGAAGACATACCCAGATAACAAGCATGTGA 295
PcPrx TAATGACCCCTTTGTCATGAAGGCATGGGCAAAGACATACCCAGAGAACAAGCATGTGA 154

TaPrx AGTTCCTTGCTGATGGAGCGGCAGCATACACAAAAGCACTTGGTCTTGAGCTTGATCTT 354
PcPrx AGTTCCTTGCTGATGGAGCGGCAGCATACACAAAAGCACTTGGTCTTGAGCTTGATCTC 213

TaPrx AGTGAGAAAGGATTGGGTCTCCGTTCGAGGCGGTTTGCTCTCCTTGCTGACGACCTCAA 413
PcPrx ACGGAGAAGGGATTGGGTCTTCGTTCGAAGCGCTTTGCTCTCCTTGCTGACGACCTCAA 272

                                                    PrxR
TaPrx GGTCACCGTCGCAAACGTCGAGGAAGGTGGCCAGTTCACAATCTCTGGTGCCGAGGAGA 472
PcPrx GGTCACCGTCGCAAACATCGAGGAAGGAGGCCAGTTCACAATCTCTGGTGCCGAG      327

TaPrx TCCTCAAGGCACTGTAG 489
```

Figure 3.7
Amplification of the *Phalaris* C-type peroxiredoxin (*PcPrx*) cDNA probe. (a) 2% agarose gel of the PCR fragment amplified using the PrxF/PrxR primer set: Lane 1, 100bp DNA ladder (sizes indicated); Lane 2, RT-PCR product; Lane 3, control (no template cDNA); Lane 4, control (genomic DNA as template). (b) Alignment of the wheat C-type peroxiredoxin (*TaPrx*) cDNA sequence (GenBank EST clone BE427212) with the sequence of the *Phalaris* RT-PCR product. The primer sequences are highlighted in bold and different nucleotides are in blue.

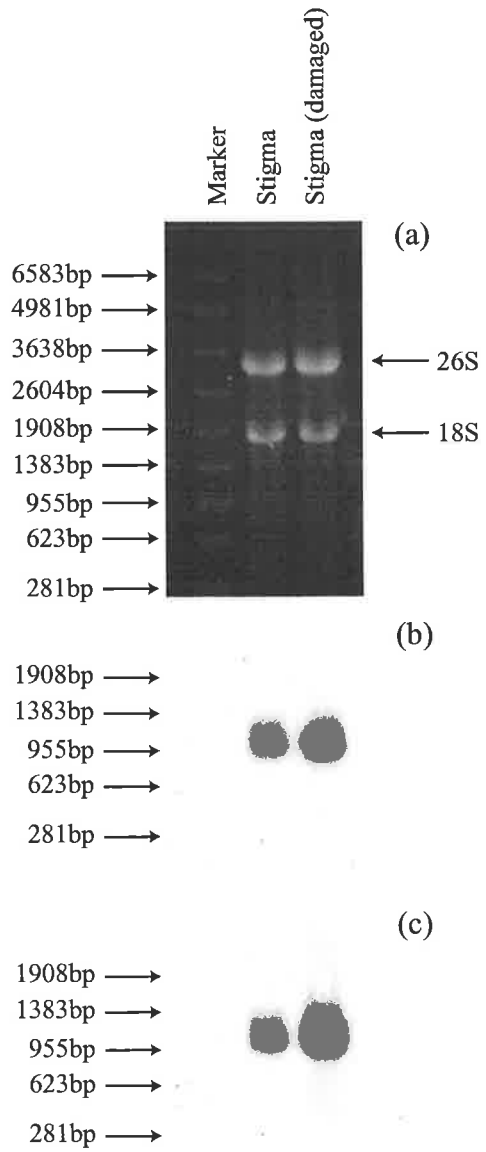


Figure 3.8
The effect of mechanical damage on the *PTrx1* and *PcPrx* mRNA levels in *Phalaris* stigma. (a) Denaturing agarose gel of stigma RNA. Sizes of RNA marker bands and the position of the 26S and 18S ribosomal RNA bands are indicated. (b) Northern blot with the *Phalaris PTrx1* cDNA probe. (c) Northern blot with the *Phalaris* peroxiredoxin cDNA probe. Both blots were exposed for 5 days

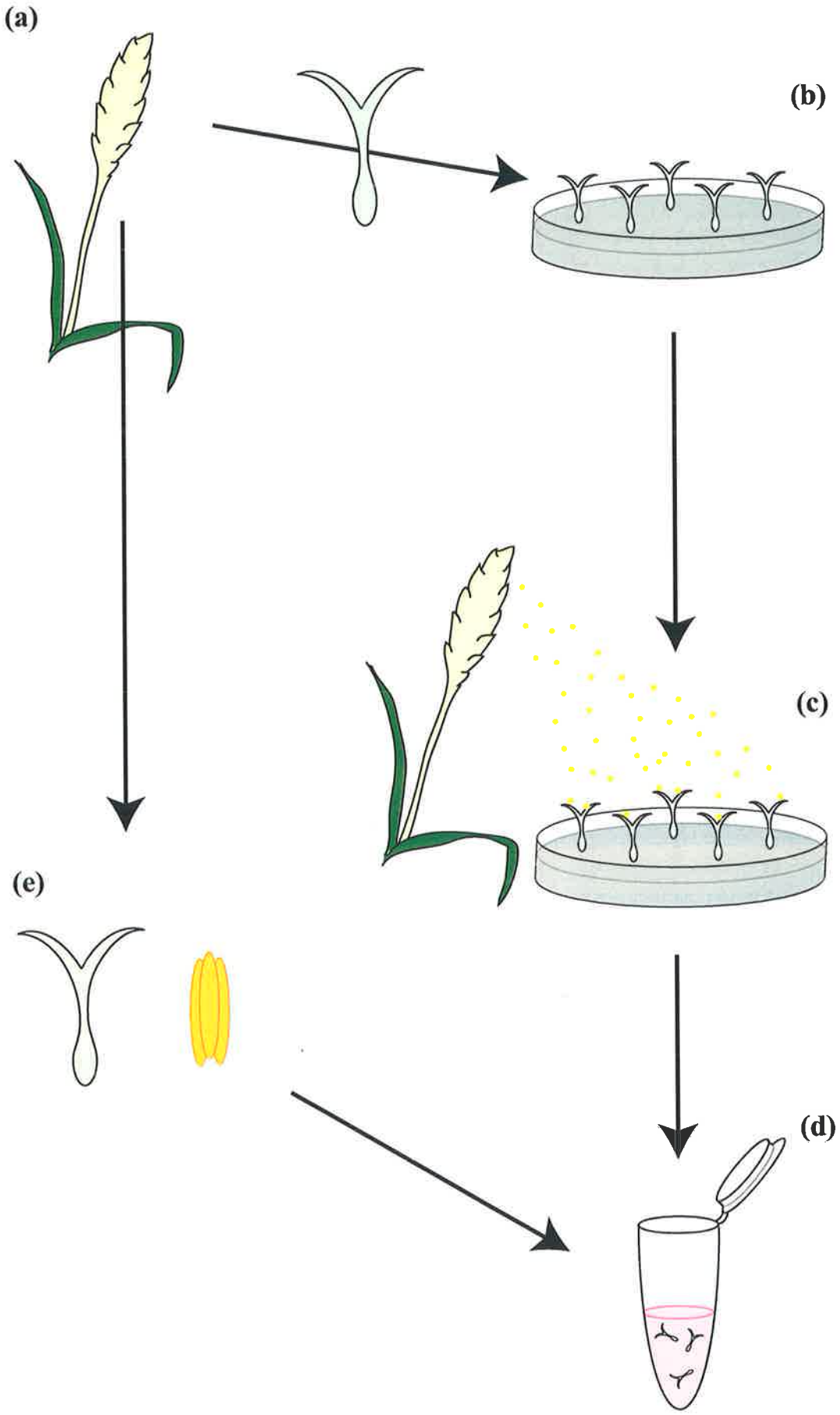

(a)
(b)
(c)
(d)
(e)

Figure 3.9 Schematic illustration of the *Phalaris coerulescens in-vitro* pollination assay

a) Intact, mature stigmas (1 day pre-anthesis) were removed from florets of a *Phalaris* plant of known genotype.

b) Stigmas were plated onto 2% agar plates containing 10% (w/v) sucrose and 100 ppm boric acid. Plates were then stored at 4°C for 24 hrs.

c) Plates were warmed to room temperature and pollinated with pollen from either:

    1) the stigma donor plant (self-incompatible pollination)
    2) a plant of different *S* and *Z* genotype (self-compatible pollination)
    3) wheat variety Chinese Spring (cross-incompatible pollination)

d) Plates were incubated at room temperature for 3 hrs, stigmas harvested and RNA extracted

e) As a control, RNA was also extracted from intact stigmas and anthers excised directly from the stigma donor plant.

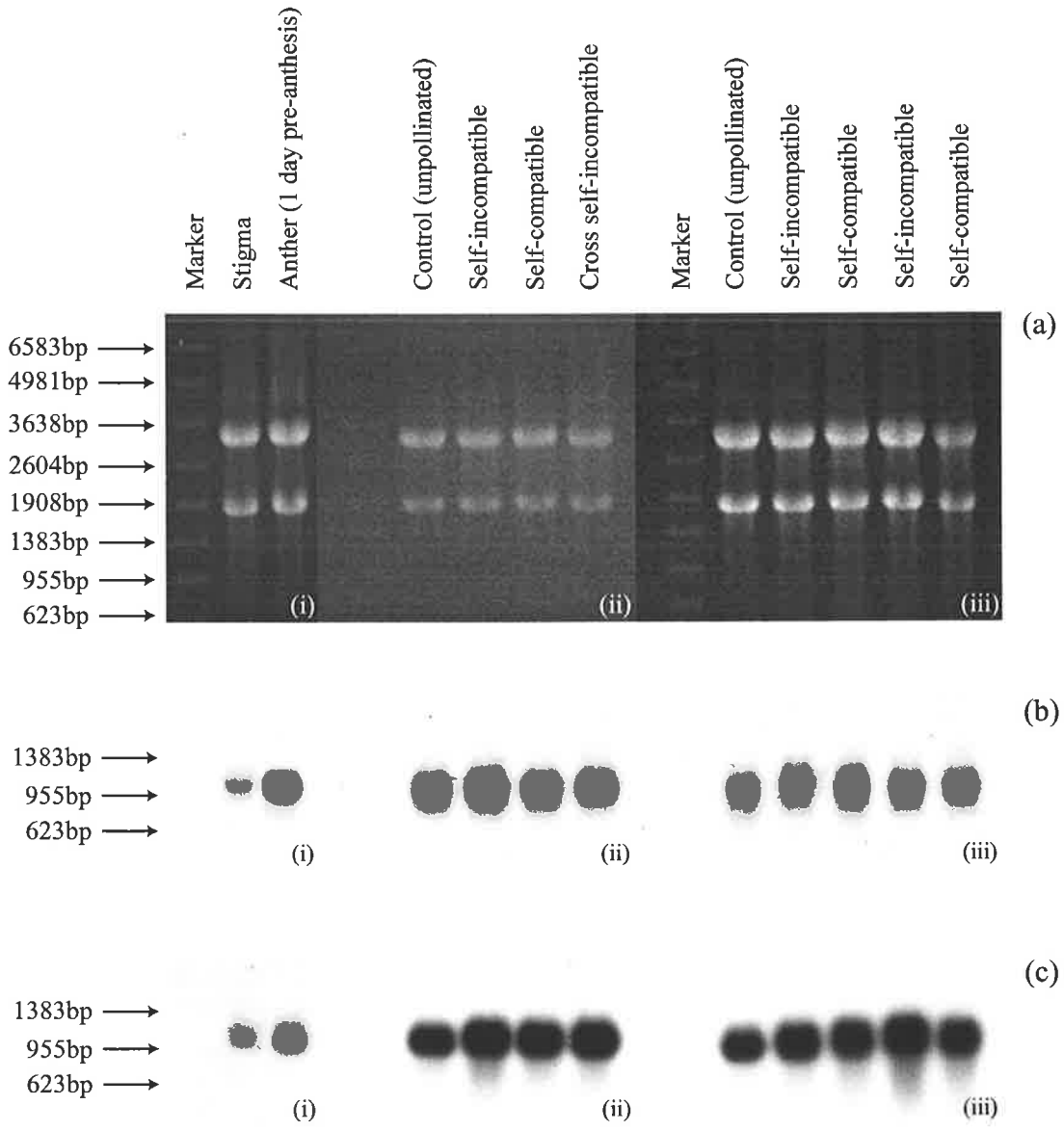


Figure 3.10
The expression characteristics of *PTrx1* and *PcPrx* in *Phalaris* stigmas in response to self-incompatible and self-compatible pollinations. (a) Denaturing agarose gel of the total RNA extracted from (Panel i) mature stigmas and anthers excised from the florets of glasshouse grown plants (Panels ii and iii) stigmas planted onto support media and pollinated with either genetically identical *Phalaris* pollen (self-incompatible), genetically different *Phalaris* pollen (self-compatible) or pollen derived from a different species (Cross-incompatible). Sizes of RNA marker bands are indicated. (b) Northern blot with the *Phalaris PTrx1* cDNA probe. All blots were exposed for 5 days. (c) Northern blot with the *PcPrx* cDNA probe. Exposure times were: blot (i), 5 days; blots (ii) and (iii), 2 days.

3.4 Discussion

Previous studies aimed at the identification and characterisation of SI genes in the grasses (Li et al. 1994; 1997) have yielded incomplete and in some cases conflicting data. In particular, the predicted structure of the *Bm2* gene has yet to be confirmed for *Phalaris* or any other self-incompatible grass. Additionally, the spatial expression pattern of *Bm2* is inconsistent with that of a male SI determinant. This study sought to clarify the structure and expression profile of *Bm2* and homologues in other self-incompatible grasses using the putative pollen *S*-gene of *Phalaris* as a starting point.

The data presented here conflict with previously published results. The cDNA's isolated have an open reading frame which translates into a protein of 131 amino acids, less than half the size of the predicted S_1 and S_2 protein sequences (Li et al. 1994) which are reported to be 282 and 281 amino acids respectively. The ORFs identified represent the first encountered in all frames of all sequences and the only one common to the sequences of all four species. The cDNA sequences in this study are almost identical to the proposed S_1 and S_2 alleles of *Phalaris*. Indeed, *PTrx1* and *PTrx2* are actually identical to the thioredoxin domains of S_1 and S_2. However, from the number and, in several cases, conserved position of upstream stop codons it is clear that, the thioredoxin-like proteins isolated here do not constitute a component of larger proteins. Furthermore, when the sequences of *PTrx1* and *PTrx2* are compared with the published *Phalaris coerulescens* S_1 and S_2 alleles it is clear that the proposed allelic domain almost certainly represents the 5' untranslated region of these genes. The misinterpretation of the allelic domain in the S_1 and S_2 sequences undoubtedly resulted from the prediction of the first intron and exon. The current evidence also enables the results of the RT-PCR study reported by Li et al. (1997) to be re-interpreted. The primers designed against the thioredoxin domain successfully amplified homologous products from other grasses, as they were complementary to sequences within the coding region of the actual gene. By contrast, the primers to the proposed allelic domain were targeted to the variable non-coding 5'UTR sequence and consequently were ineffective. Thus the sequence variability of the 5'UTR was incorrectly construed as representing the variability inherently found in different *S*-alleles.

From sequence analysis it appears that these genes are thioredoxins belonging to the plant *h* class, as they not only show highest homology to published thioredoxin *h* genes but also possess all structurally important amino acids. These include the active centre consensus sequence (WCGPC) and structurally important residues given as Phe50, Ala52, Pro63, Asp84, Pro99 and Gly 115 corresponding to Phe27, Ala29, Pro40, Asp61, Pro76 and Gly92 in the *E. coli* thioredoxin sequence (Eklund et al. 1991). Furthermore, these sequences are not preceded by transit peptides as has been described for the other thioredoxin classes (*m, f, x* and *o*) from higher plants (Mestres-Ortega and Meyer, 1999; Lalio et al. 2001).

The genes described here show 56.0-97.7% identity with one another at the nucleotide level. Interestingly, the predicted proteins of several of these genes display higher levels of interspecific amino acid identity than that calculated for thioredoxin *h* genes isolated from a single species. For example, proteins encoded by the five thioredoxin *h* genes of *Arabidopsis* share 60-70% identity (Rivera-Madrid et al. 1995). By comparison, the grass sequences reported here share a minimum of 95.4% amino acid identity and the database homologues are between 54% and 70% identical with one another at the protein level. Excluding amino acids of known structural importance, the position of no less than 33 residues is completely conserved across these proteins and can be used to distinguish them from other thioredoxins. In addition, chemically similar residues are maintained at a further 35 positions. Based upon these results, the genes isolated would appear to represent a highly similar subclass of plant thioredoxins *h*. To enable members of this subclass to be readily identifiable throughout this thesis, genes belonging to the subclass will be entitled thioredoxins *n* (for new thioredoxins). Members of the thioredoxin *n* subclass are present in a diverse array of species including a member of the gymnosperms, a result that implies the *n* subclass is of ancient origin.

Detection of the *Bm2* mRNA by *in-situ* hybridisation has shown that *Bm2* is expressed in stigma tissue in response to a self-incompatible pollination (Baumann et al. 2000). Given that *Bm2* appears to encode a thioredoxin the *in-situ* findings raise the question of the role of thioredoxins in the SI response of grasses. In the *Brassicaceae* thioredoxins appear to be an important component of the SI

reaction. Yeast two-hybrid and mutational analyses have demonstrated that thioredoxins interact with the kinase domain of the SRK stigma protein and that the interaction is dependant upon a functional thioredoxin active site (Bower et al. 1996; Mazzurco et al. 2001). The *Brassica* thioredoxins are proposed to act as phosphorylation inhibitors preventing the spontaneous phosphorylation of SRK and activation of the SI signalling pathway (Cabrillac et al. 2001). Furthermore, a biochemical study of SI in lily has found that SI is associated with oxidative stress (Tezuka et al. 1997). The lily study poses an additional question of whether SI in *Phalaris* is also associated with oxidative stress and if so could this explain the accumulation of *Bm2* detected by the *in-situ* experiment?

From the Northern hybridisations conducted in this study it is evident that *Bm2* (*PTrx1*) is not specifically expressed in response to a self-incompatible pollination. Both self-incompatible and self-compatible pollinated stigmas showed a virtually identical level of expression. Consequently, the most plausible explanation for the previously reported accumulation of *Bm2* mRNA at the site of a self-incompatible pollination is that the mRNA was released from the ruptured incompatible pollen tube and it is transcript from this source that the *in-situ* RNA probe detected. *PcPrx* expression levels revealed that stigmas subjected to the *in-vitro* pollination assay experienced oxidative stress. Interestingly, *PTrx1* expression was also upregulated in stigma tissue under oxidative stress. *PTrx1* expression was induced by both mechanical damage and the *in-vitro* assay procedure suggesting that this gene may be involved in the oxidative stress response of grass stigmas. Since *PcPrx* was highly expressed in stigma tissue pollinated with both incompatible and compatible pollen, it was impossible to determine whether an incompatible pollination results in oxidative stress as has been found in lily. It is possible that incompatibility *in-planta* results in a slight upregulation of genes involved in oxidative protection, but that this was masked by the sizeable expression induced by the *in-vitro* assay method. To address whether SI in *Phalaris* is associated with oxidative stress, future expression studies will have to use stigmas pollinated *in-planta*.

The thioredoxin *n* genes isolated here display all of the sequence characteristics of thioredoxin *h* genes. Whether they encode functional thioredoxin proteins is the focus of the following chapter.

Chapter 4 Assessing the Activity of Thioredoxins *n*

4.1 Introduction

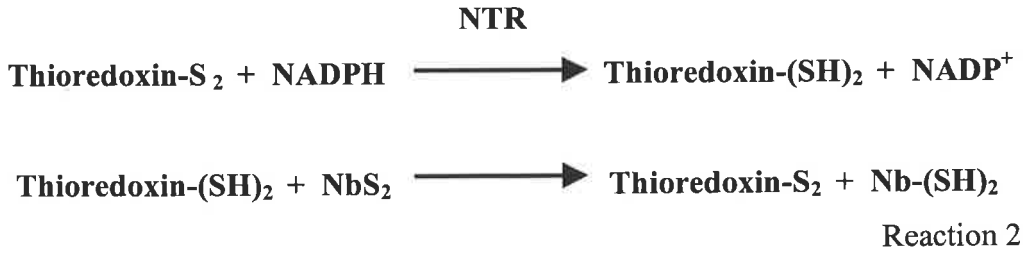
Thioredoxin contains a redox active cysteine disulfide that is reduced to a dithiol by thioredoxin reductase and NADPH (Holmgren 1977). In the reduced form thioredoxin is capable of reducing disulfides in a large number of target proteins including proteins that, *in-planta*, thioredoxin would not interact with. This feature of thioredoxins has been used to develop several protein assays for thioredoxin activity. Two assays most commonly employed by researchers are the insulin reduction assay (Holmgren 1979) and the 5,5'-dithio-bis-(2-nitrobenzoic acid) (NbS_2) reduction assay (Slaby and Holmgren, 1975).

The insulin reduction assay is a non-specific test that measures the capacity of thioredoxin to catalyse the reduction of insulin by dithiothreitol (DTT) (Reaction 1). Insulin comprises two chains, the α and β-chains, joined by two interchain disulfide bonds. At neutral pH the reduction of the interchain disulfides result in the precipitation of the two free chains. The free β-chain forms a white precipitate that can be measured spectrophotometrically at 650 nm. Since reduced thioredoxin reduces insulin with a second order rate constant that is 10^4 times that measured for insulin reduced with DTT alone (Holmgren 1979), the catalytic effect of thioredoxin is easily discernible.

$$\mathbf{DTT\text{-}(SH)_2 + Insulin\text{-}S_2} \xrightarrow{\textbf{Thioredoxin}} \mathbf{DTT\text{-}S_2 + Insulin\text{-}(SH)_2}$$

Reaction 1 (from Holmgren 1979)

The NbS_2 assay is a more specific assay that incorporates the other components of the thioredoxin system (NADPH and NADPH-dependant thioredoxin reductase (NTR)). This assay is designed to measure the capacity of a thioredoxin protein to serve as a substrate for NTR (Reaction 2). If the assayed thioredoxin is able to act as a substrate it is reduced by NTR and the reduced form can in turn reduce NbS_2. The reduction of NbS_2 is accompanied by the development of colour that can be measured as a change in absorbance at 412 nm.



Reaction 2

Both assays were used to determine the thioredoxin activity of the LpTrx and PTrx proteins (hereafter referred to as Lpn and Pcn). The assays were also applied to a tobacco (*Nicotiana tabacum*) thioredoxin *n* homologue (Ntn) kindly provided by David Olde. The *Ntn* cDNA was amplified from leaf RNA, using primers designed against the tomato (*Lycopersicon esculentum*) thioredoxin *n* EST described in the previous chapter, and was included so that both monocotyledon and dicotyledon representatives of the *n* group were tested.

4.2 Materials and Methods

4.2.1 Cloning and Construction of the Recombinant Expression Vector

The QIAexpress (Qiagen) recombinant protein expression system was employed to express and purify the thioredoxin *n* proteins from *Lolium, Phalaris* and *Nicotiana*. The system generates recombinant proteins with an N-terminal 6-histidine tag that can be purified by affinity chromatography. All cloning, transformation, expression and purification was conducted according to the manufacturers instructions.

The *Pcn* sequence was excised from the sequencing vector (pGemT, Promega) by a *HindIII-DraI* double digest. The *HindIII* site was blunt ended and the fragment cloned into the *SmaI* site of the pQE-31 expression vector. The fragment consisted of the entire coding region of the gene less the initiation codon and the first 39 nucleotides (13 amino acids). The *Lpn* gene was cloned into the *HindIII* site of the pQE-30 expression vector in two stages. Firstly, the entire coding region of the gene was released from the sequencing vector by an *SphI-PstI* double digest and cloned into a pQE-31 vector that had been digested with same enzymes. Secondly, the recombinant pQE vector was digested with *HindIII* and the resultant fragment cloned into the *HindIII* site of pQE-30. Again this constituted the entire coding region less the initiation codon and the first 13 amino acids. The *Ntn* expression

construct was kindly supplied by Dr Ute Baumann and David Olde. The construct consisted of the entire coding region cloned into pQE-30 double digested with *BamHI* and *HindIII*.

4.2.2 Expression and Purification of Recombinant Proteins

Recombinant plasmids were transformed into competent M15 *E. coli* cells (Qiagen, USA) by heat shock treatment (2.6). Cells were plated on LB plates containing 25 μg/ml kanamycin and 100 μg/ml ampicillin and incubated overnight at 37°C. Individual colonies were selected and inoculated into 5 ml of LB containing both antibodies and grown at 37°C with constant shaking for approximately 12 hrs. 500 μl of the starter culture was removed and inoculated into 10 ml of pre-warmed (37°C) LB (containing 25 μg/ml kanamycin and 100 μg/ml ampicillin) and the culture grown at 37°C with shaking until the OD_{600} reached 0.8. Protein expression was induced by the addition of IPTG to a final concentration of 2 mM. Three hours after induction, cells were harvested by centrifugation and the pellet resuspended in 1 ml of lysis solution (50 mM NaH_2PO_4, 300 mM NaCl, 1% Triton, 5mM imidazole, pH 8.0) containing 1 mg, 0.3 mg and 0.3 mg of Lysozyme, RNase and DNase, respectively. The solution was then left on ice for 30 min. Cells were lysed by a combination of rapid freeze-thawing (in liquid nitrogen) followed by sonication (6 x 6 s) at 40 W in a Branson b-12 Sonifier and the cellular debris removed by centrifugation at 10,000 rpm for 10 min. A 50% slurry of Ni/nitriloacetic acid resin (Qiagen, USA) in lysis buffer was added to the supernatant and the recombinant proteins separated from endogenous proteins by virtue of their histidine tag. Contaminating proteins were removed by a series of three individual washing steps: Step1, four washes with 50 mM NaH_2PO_4, 300 mM NaCl, 5mM imidazole, pH 8.0; Step 2, three washes with 50 mM NaH_2PO_4, 300 mM NaCl, pH 6.0; Step 3, three washes with 100 mM NaH_2PO_4, pH 6.0. The purified protein was eluted from the resin by the addition of 100 mM NaH_2PO_4, 2 mM EDTA, pH 3.0. The eluate containing the recombinant protein was then titrated to pH 7.0 by the addition of 100 mM NaH_2PO_4, 2 mM EDTA, pH 10.0. Protein concentration was determined spectrophotometrically (Shimadzu UV-160 A) at 280 nm and by comparison with a BSA standard curve. The purified protein was visualised on a 12.5%

polyacrylamide gel with protein markers in the 4 to 250 kDa range (SeeBlue, Novex USA).

4.2.3 Thioredoxin Activity Assays

4.2.3.1 Reduction of Disulfides in Insulin

The thioredoxin catalysis of insulin (Sigma) reduction was measured spectrophotometrically (Shimadzu UV-160 A) at 650 nm and 25 °C as an increase in turbidity resulting from the precipitation of the free insulin β-chain (Holmgren 1979). The assay mixture was 100 mM NaH_2PO_4, 2 mM EDTA, pH 7.0 containing 1 mg insulin and 0.4 μM to 4.0 μM of the Pcn, Lpn and Ntn proteins. The reaction was initiated by the addition of dithiothreitol, to a final concentration of 0.33 mM, to the reaction cuvette only. The blank contained all components less the dithiothreitol.

4.2.3.2 Thioredoxin *n* Reduction by *E. coli* NTR

The three thioredoxin-*n* proteins were tested as substrates for *E. coli* NADPH-dependant thioredoxin reductase (Sigma). The reduction of *n* proteins was determined by measuring the reduction of NbS_2 in a reaction mixture containing 100 mM NaH_2PO_4, 2 mM EDTA, pH 7.0, 150 mM NADPH and 200 mM NbS_2. The reaction was initiated by the addition of 10 ρmol of *E. coli* NTR and the *n* protein reduction assessed as a function of NbS_2 reduction. The reduction of NbS_2 was measured directly as the change in absorbance at 412 nm. The thioredoxin protein of the prokaryotic organism *Spirulina sp* (Sigma) was used as a positive control.

4.3 Results

4.3.1 Preparation of Expression Constructs

The *Lolium, Phalaris* and *Nicotiana* thioredoxin *n* cDNAs were cloned into pQE expression vectors in order to generate recombinant proteins containing an N-terminal histidine tag, which was subsequently used to facilitate purification. The sequence of the *Ntn* cDNA is given in Figure 4.1. A restriction fragment containing the entire coding region of the *Pcn* cDNA, less 13 N-terminal amino acids, was cloned into the *SmaI* site of pQE-31, whilst a two step procedure was employed to clone an identical region of the *Lpn* cDNA into the *HindIII* site of

pQE-30 (Figure 4.2). The endogenous translational start codon was removed in each case to ensure that the translational products would contain the 6xHis tag at the N-terminus. Internal, in frame, initiation codons are present in all sequences (positions 68 and 76 for *Pcn*, 68, 76 and 86 for *Lpn* and 6, 73 and 89 for *Ntn*). However, translation initiating from these points generates polypeptides lacking the histidine tag and consequently these are removed by subsequent purification steps.

4.3.2 Purification of Recombinant Thioredoxin *n* Proteins

Recombinant *Lolium, Phalaris* and *Nicotiana* thioredoxin *n* proteins were purified by affinity chromatography on a Nickel-resin column. The inclusion of low concentrations of imidazole to the initial binding and first washing step was found to result in a highly stringent purification protocol (Figure 4.3). Despite the loss of recombinant protein at each purification step, the protocol resulted in a pure protein preparation and yielded sufficient recombinant protein for later assays.

The presence of the recombinant thioredoxin *n* proteins was verified by SDS/PAGE (Figure 4.4). For all species a protein of the expected size (approximately 12 kDa for the *Lolium* and *Phalaris* proteins and 13.5 kDa for the *Nicotiana* protein) was purified. The purification regime effectively removed all endogenous *E. coli* proteins without substantially reducing the yield of the recombinant proteins. The yield of recombinant protein was found to be identical (approximately 5mg/ml) regardless of whether the proteins were purified under native or denaturing conditions suggesting that the majority of the protein is present in the soluble fraction.

4.3.3 Thioredoxin Assays

Thioredoxin has been shown to catalyse the non-specific reduction of insulin disulfides by the reducing agent DTT (Holmgren, 1979). The capacity of the *Lpn, Pcn* and *Ntn* thioredoxins to reduce insulin disulfides following reduction by DTT was assessed. The proteins of all species could reduce disulfides in insulin after the addition of DTT (Figure 4.5). As in other studies (Gautier et al. 1998), the rate of insulin reduction was found to increase as a function of thioredoxin concentration. Not surprisingly, at the same concentrations (0.9 μM *Lpn* and *Pcn*,

2.0 μM *Lpn* and *Ntn*), the activities of the thioredoxins *n* were similar. The similarity in activity reflects the high level of identity (63%-95%) between the proteins. A Lineweaver-Burk plot was generated from the *Lolium* assay data (Figure 4.5) and the V_{max} (0.05 ΔA_{650}/min) and K_m (0.54μM) values calculated. The values are consistent with those that have been previously reported for wheat thioredoxin *h* (0.05ΔA_{650}/min and 0.65μM) (Gautier et al. 1998).

The *Lolium*, *Phalaris, Nicotiana* and *Spirulina* proteins were tested as substrates for *E. coli* NTR by measuring their ability to catalyse the NADPH-dependent reduction of NbS_2 (5,5'-dithiobis(2-nitrobenzoic acid)). The plant proteins were poor substrates for *E. coli* NTR, a finding reflected in the high concentrations of thioredoxins required to reduce measurable quantities of Nbs_2 (Figure 4.6). The low substrate recognition is seen in the high K_m (80μM) and V_{max} (167μM reduced thioredoxin/min) values calculated for *Lolium* (Figure 4.7). Conversely, the *Spirulina* thioredoxin, a protein more closely related to *E. coli* thioredoxin, was rapidly reduced by *E. coli* NTR. This result is congruent to the K_m of 81μM recorded by Jaquot et al. (1994) when they tested *E. coli* thioredoxin as a substrate for *Arabidopsis* NTR (Table 4.1). The low level of activity between the plant proteins and *E. coli* NTR demonstrates that the purification procedure effectively removed all endogenous *E. coli* thioredoxin, whilst the rapid reduction of the *Spirulina* thioredoxin confirmed the activity of the commercially acquired NTR protein. The similarity of activity seen in the insulin assay is confirmed in the NbS_2 reduction assay where proteins show comparable activities at the same concentration (8 μM *Lolium* and *Phalaris*, 20 μM *Lolium* and *Nicotiana*).

Table 4.1 K_m values for different thioredoxin and NTR interactions

Thioredoxin	NTR	K_m (μM)
A. thaliana	*A. thaliana*	1.1
E. coli	*A. thaliana*	81.0
E. coli	*E. coli*	2.8
L. perenne	*E. coli*	80.0
T. aestivum	*T. aestivum*	3.1

```
Pcn1                    MGGCVGKDRGIVEDKLDFKGGNVHVITTKEDWD  33
Len    MGISDTVRSLFPCIKSHSTS-GDDSTHNVE-A----SL-----S--  46
Ntn    MGITDMVHSLFPCIKSRSTNND-DSSHNVK-A----SL-----S--  46

Pcn1   QKIAEANKDGKIVVANFSASWCGPCRVIAPVYAEMSKTYPQLMFLT  79
Len    --L---K-E----I-----------M-S-F-C-L-EK-LS-----  92
Ntn    --L-----E----I-----------M---F-C-L-EK-LS-----  92

Pcn1   IDVDDLVDFSSTWDIRATPTFFFLKNGQQIDKLVGANKPELEKKVQ 125
Len    V---E-TE---S---K---------D-E-------------Q--IT 138
Ntn    V---E-TE---S---K---------D-E-------------Q--IT 138

Pcn1   ALGDGS            131
Len    TVA-RHVACEPQPQPQ  154
Ntn    -IA-TQVVCETQPQ    152
```

Figure 4.1
Alignment of the *Phalaris coerulescens* (Pcn1) amino acid sequence with the homologous expressed sequence tag (EST) sequence from tomato (*Lysopersicon esculentum;* Len) and the corresponding sequence amplified from tobacco (*Nicotiana tabaccum*; Ntn). Dashes indicate residues in the tomato and tobacco sequences that are identical to the *Phalaris* protein. The thioredoxin active site sequence is in red.

Figure 4.2

The cloning strategy used to generate the *Lpn* expression construct. A two step cloning procedure was employed so that the endogenous *Lpn* initiation codon was removed and the remaining coding sequence was cloned in frame with the 6-histidine tag:

a) The entire *Lpn* cDNA was excised from the sequencing vector (pGemT) by a *SphI/PstI* double digest.

b) The resulting 814bp fragment was cloned into a pQE-31 expression vector (cut with the same enzymes).

c) The endogenous ATG-start codon was removed by digestion with *HindIII*. The resultant 508bp fragment contained the entire *Lpn* coding sequence less the first 39 nucleotides.

d) The *Lpn-HindIII* fragment was cloned into a *HindIII* digested pQE-30 expression vector. In this vector the *Lpn* fragment was cloned in frame with the upstream sequence encoding the 6-histidine tag, enabling the recombinant protein to be purified by affinity chromatography.

e) Since the *Lpn* fragment was not directionally cloned into pQE-30 the orientation was confirmed by a *PstI* digest. Fragments that ligated in the correct orientation produced a 508bp restriction product (Lanes 1 and 2) whilst fragments in the reverse orientation produced a 20bp product (Lane 3).

f) The identity of the products from Lanes 1 and 2 was further confirmed by a *SacI* digest. The *Lpn* cDNA contains a single *SacI* site resulting in two bands of 303 and 205 bp.

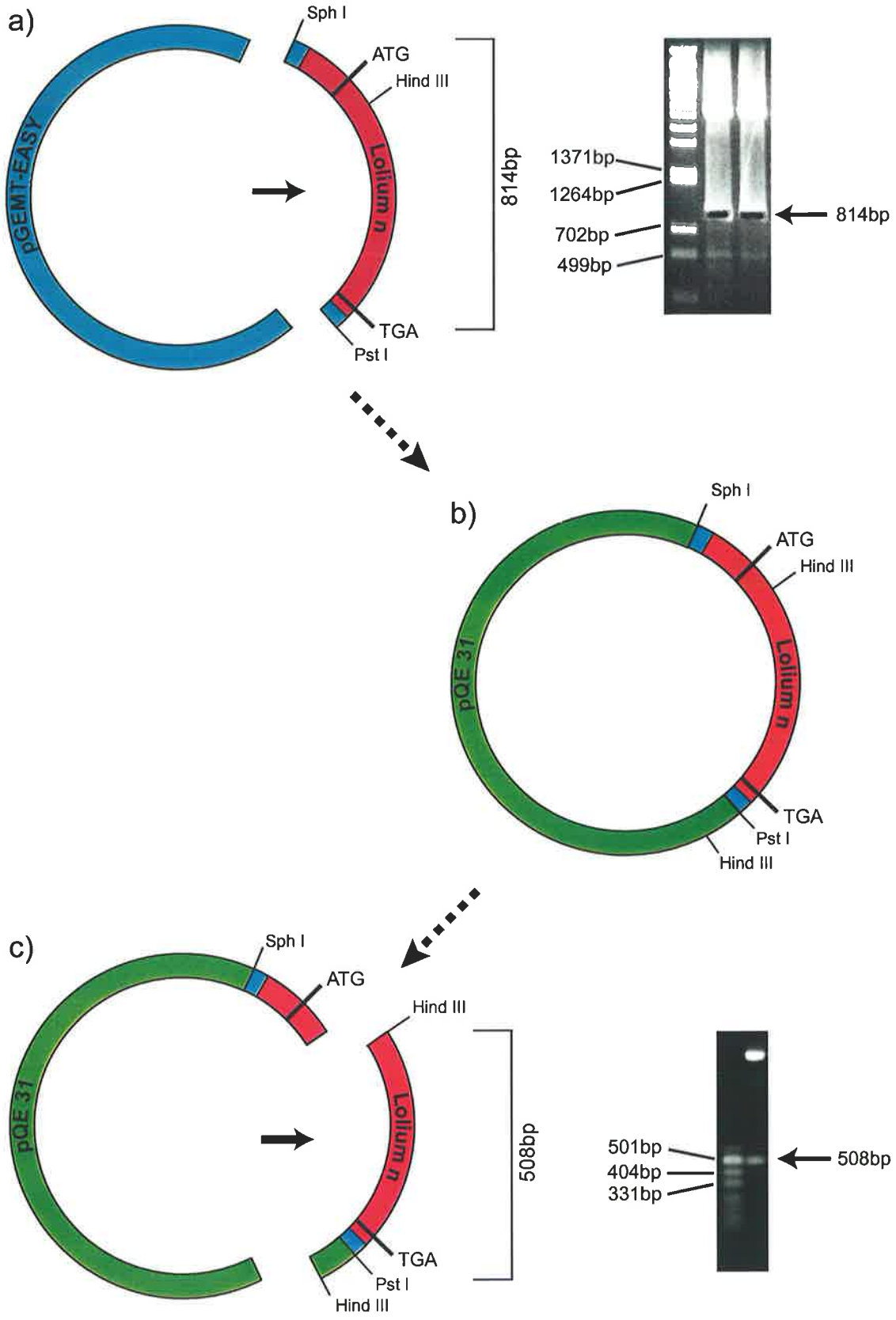


Figure 4.2 Cloning the *Lolium perenne* thioredoxin *n* cDNA into the pQE-30 expression vector.

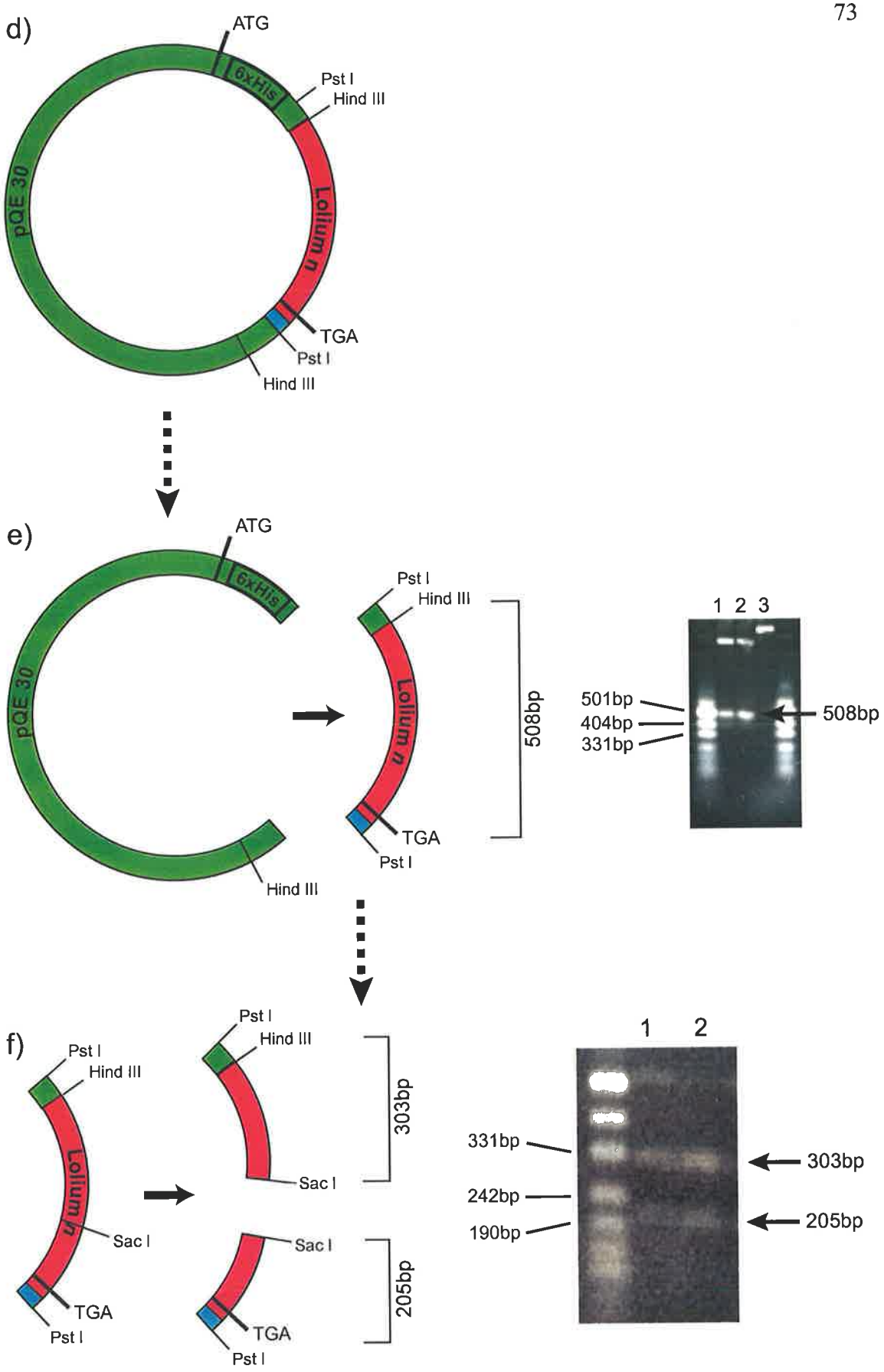


Figure 4.2 Continued Cloning the *Lolium perenne* thioredoxin *n* cDNA into the pQE-30 expression vector.

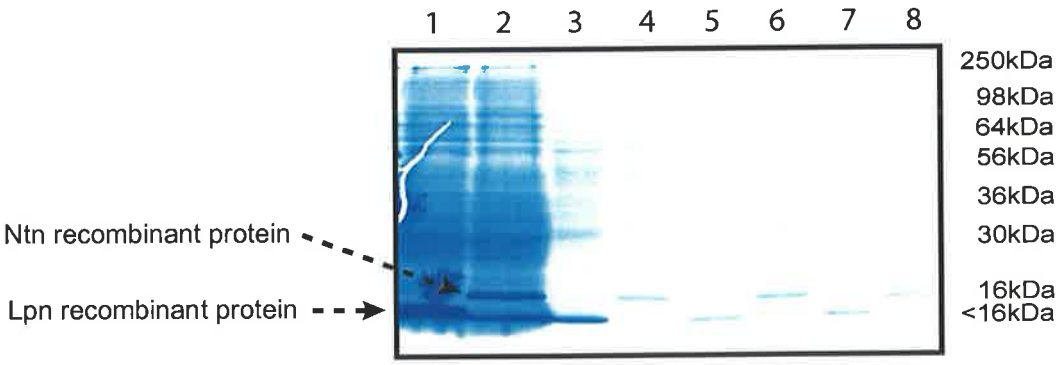


Figure 4.3 Monitoring of protein washing steps
The removal of proteins by each sucessive washing step was determined by SDS/PAGE gel electrophoresis. Lanes are:
1) Lolium perenne (Lpn) total protein
2) Nicotian tabacum (Ntn) total protein
3) Lpn proteins removed by Step1 washes
4) Ntn proteins removed by Step1 washes
5) Lpn proteins removed by Step2 washes
6) Ntn proteins removed by Step2 washes
7) Lpn proteins removed by Step3 washes
8) Ntn proteins removed by Step3 washes
The positions of the recombinant Lpn and Ntn proteins are given.

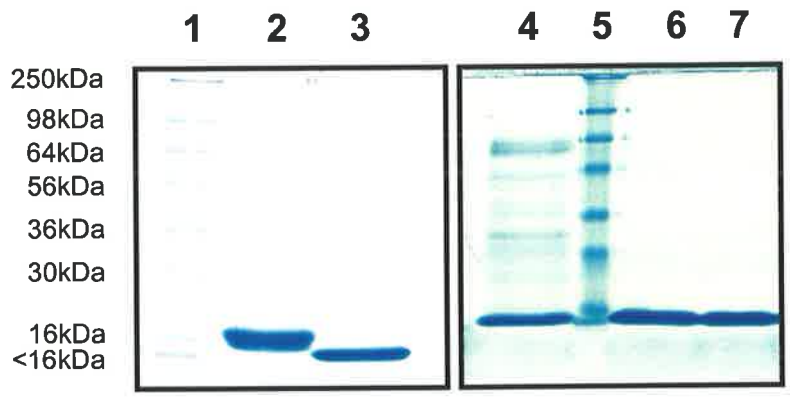


Figure 4.4
SDS/PAGE gel of the purified *Lolium, Phalaris* and *Nicotiana* n proteins. Lanes 1 and 5, SeeBlue (Novex, USA) protein size markers with the band sizes indicated: Lane 2, Ni-resin-purified Ntn protein: Lanes 3 and 6, Ni-resin-purified Lpn recombinant protein: Lane 4, Lpn total protein after lysis: Lanes 7, Ni-resin-purified Pcn recombinant protein.

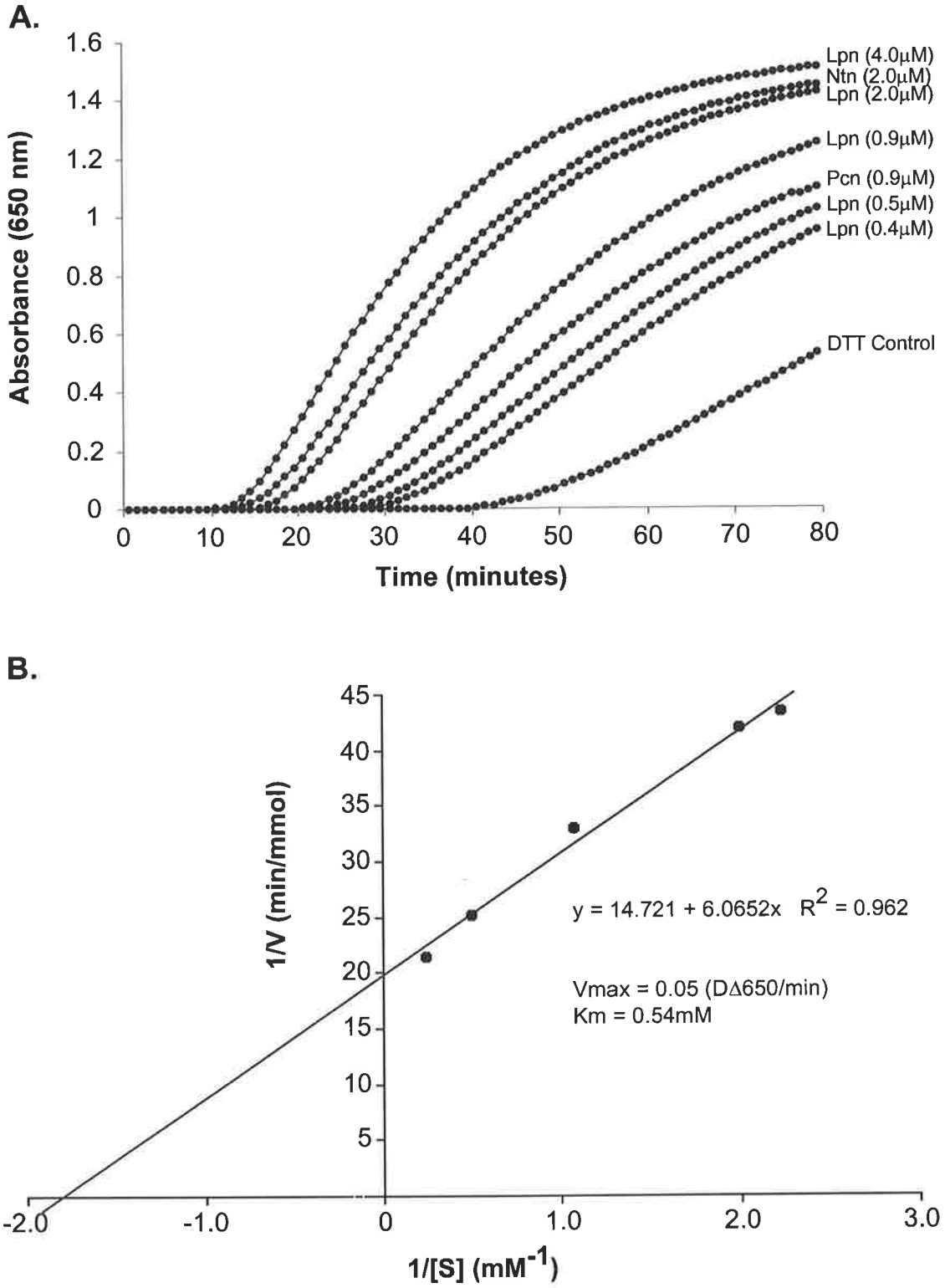


Figure 4.5
The reduction of insulin by recombinant Lpn, Pcn and Ntn proteins (A), and Lineweaver-Burk plot calculated using the *Lolium* protein concentrations (B). Reactions in (A) were initiated by the addition of dithiothreitol (DTT) and insulin reduction was measured spectrophotomerically (650 nm) as the increase in turbidity associated with the precipitation of the insulin β-chain. The reduction of insulin by DTT in the absence of thioredoxin served as a control.

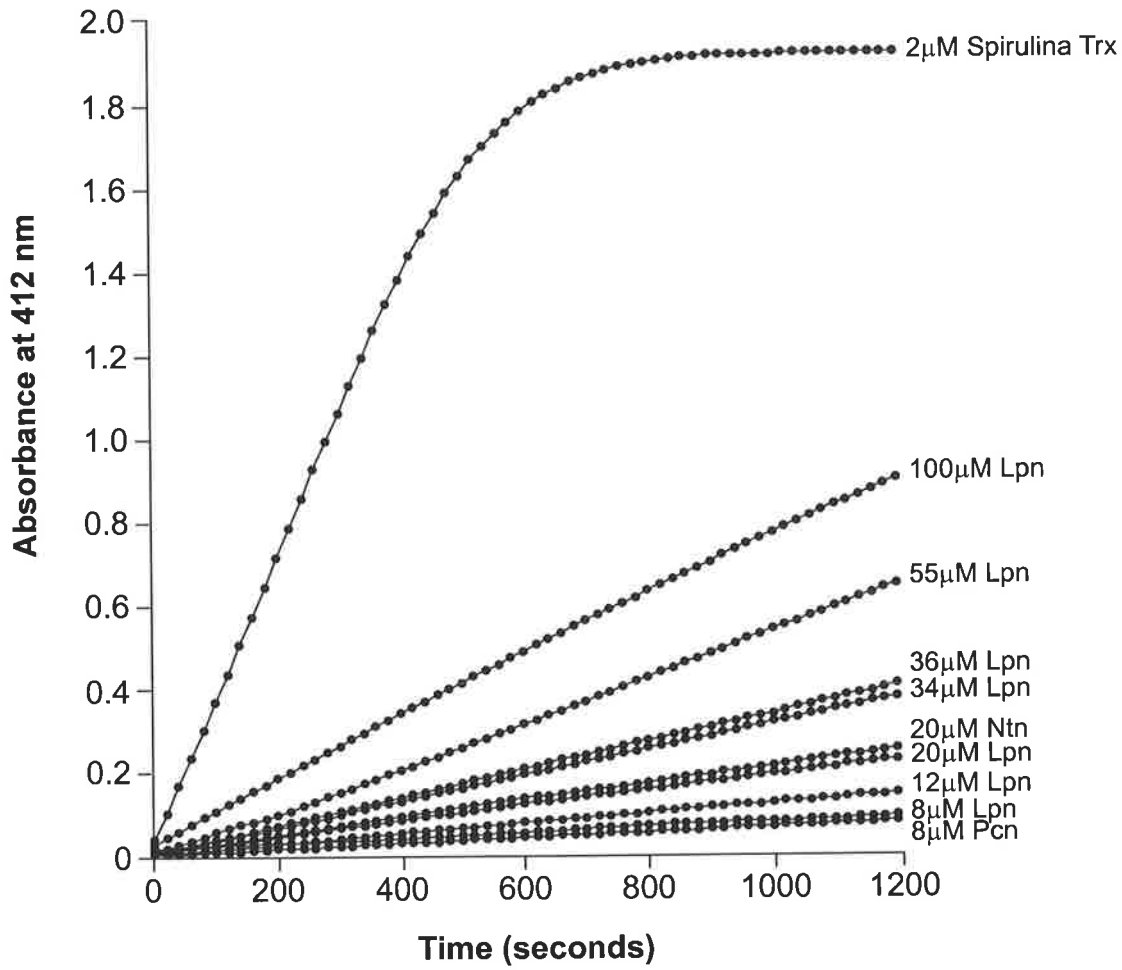


Figure 4.6
The reduction of Nbs_2 by Lpn, Pcn and Ntn proteins in combination with *E. coli* NTR. Reactions were initiated by the addition of NTR and the reduction of Nbs_2 was determined spectrophotomerically (412 nm). The *Spirulina* thiorexoxin is included as a positive control.

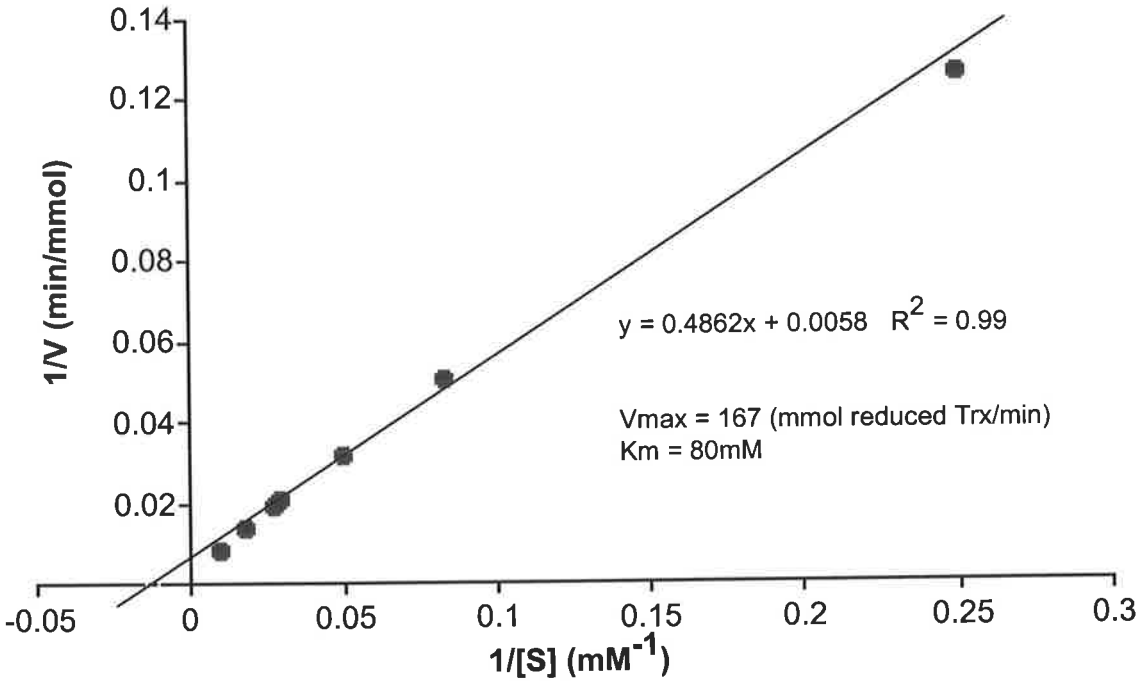


Figure 4.7
Lineweaver-Burk plot of the Nbs_2 reduction assay using the *Lolium* protein concentrations.

4.4 Discussion

This study sought to determine whether the proteins encoded by members of the thioredoxin *n* subclass did indeed possess thioredoxin activity. The results presented here indicate that the *n* subclass proteins do represent active thioredoxins. The *Lolium perenne, Phalaris coerulescens* and *Nicotiana tabacum* thioredoxin *n* proteins were able to reduce insulin disulfides following reduction by DTT. The K_m of the *Lolium* protein is similar to that reported for wheat thioredoxin *h* (Gautier et al. 1998) and *E. coli* thioredoxin (Holmgren 1979) suggesting a similar level of activity for this non-specific test. The catalytic nature of these proteins is confirmed by the observation that thioredoxin catalysed reduction occurred faster than insulin reduction by DTT alone.

In contrast to the insulin assay, all plant proteins were found to be poor substrates for *E. coli* NADPH-dependant thioredoxin reductase (NTR) (*Lolium n* K_m 80μM). Previous studies have also found poor interactions between NTRs and thioredoxins from highly divergent organisms. Decottignies et al. (1991) reported that the *Chlamydomonas* thioredoxin *Ch1* could not serve as a substrate for *E. coli* NTR at concentrations of up to 6μM. Similarly, wheat thioredoxin *h* could not be reduced by *E. coli* NTR although the range of thioredoxin concentrations tested was not reported (Gautier et al. 1998). It is possible that at higher thioredoxin concentrations low levels of reducing activity may have been recorded not unlike those reported here. Indeed the K_m of 81μM calculated for the reduction of *E. coli* thioredoxin by *Arabidopsis* NTR (Jacqout et al. 1994) is very similar to the results achieved in this assay. Future analyses with the *n* group will undoubtedly require the use of an NTR from a eukaryotic, and preferably plant, source.

Several publications have provided evidence for the existence of different levels of specificity in the interaction between NTRs and thioredoxins. Rivera-Madrid et al. (1995) report that *Arabidopsis* NTR displays a much lower affinity for *Arabidopsis* thioredoxin *h4* than any of the other *Arabidopsis* thioredoxins *h*. Similarly, *Chlamydomonas Ch1* has a higher affinity for *Arabidopsis* NTR *in vitro* than *N. tabacum h1*, a more closely related species (Jacquot et al. 1994). A subsequent study has identified multiple NTR sequences in the genome of

Arabidopsis including an organelle specific form (Laloi et al. 2001). Moreover, one NTR gene was found to encode multiple transcription initiation sites, resulting in the translation of two distinct proteins *in vitro*. Assuming that most plants also contain multiple NTR proteins it is conceivable that specific groups, such as the *n* subclass described in this thesis, interact with, or are preferentially reduced by, a particular NTR. Testing this proposition would require an inclusive set of biochemical data for interactions between all thioredoxins *h* and all non-organellar NTRs from a single species. Currently, *Arabidopsis* represents the only plant species for which the entire complement of these genes is known and for which many of the corresponding proteins have been expressed and tested (Jacqout et al. 1994; Rivera-Madrid et al. 1995; Laloi et al. 2001).

Chapter 5 Characterisation of the Thioredoxin *h* Class in Wheat and Other Grasses

5.1 Introduction

In contrast to many other organisms, the molecular and genetic profile of thioredoxins in most plants has yet to be determined. *Arabidopsis* is currently the only plant species for which there exists a comprehensive set of data for all gene members of a given thioredoxin class (see 1.3.4). In the grasses, molecular information regarding thioredoxins is restricted to five *h*-class sequences isolated from two cereal species, rice (Ishiwitari et al. 1995) and wheat (Gautier et al. 1998; Serrato et al. 2001).

The previous chapters have demonstrated that grass thioredoxin *n* genes encode active thioredoxin proteins and potentially constitute a subclass within the *h*-class. The next logical step in the investigation of thioredoxins *n* would be their molecular characterisation in a model grass species. For example, the determination of gene copy number, chromosomal location and expression patterns. However, since thioredoxins *n* are members of the multigenic *h*-class it was deemed necessary to characterise all members of this class in the grasses. The rationale for this decision being:

1) Investigating all members of the *h*-class would ensure that hybridisation results (eg Southerns and Northerns) reflected the profile of an individual gene and did not arise by cross-hybridisation to distinct yet sequentially related members of the class.

2) The study would provide the first comprehensive molecular profile of all members of any thioredoxin class for a grass species.

3) The results obtained would be a valuable tool for future studies investigating the function of grass thioredoxins *h*.

Since the majority of published data relating to grass thioredoxins *h* has arisen from studies in wheat, the wheat variety Chinese Spring was chosen as the grass species on which this investigation would be conducted.

The involvement of thioredoxins in the response to oxidative stress, either by direct reduction of an oxidised protein or as electron donors for a reactive-oxygen scavenging system, has been well documented in animal, yeast and bacterial cells (Fernando et al. 1992; Grant 2001; Takemoto et al. 1998). In contrast, no study has shown a direct involvement of plant thioredoxins in the response to oxidative stress although *Arabidopsis* thioredoxins *h3* and *h4* have been demonstrated to restore H_2O_2 tolerance to an H_2O_2-sensitive yeast thioredoxin-mutant. To determine whether grass thioredoxins *h* might participate in oxidative protection, the transcript expression pattern of wheat thioredoxins in plants subjected to varying forms of abiotic stress was also investigated.

5.2 Materials and Methods

5.2.1 Plant Material

Wheat (*Triticum aestivum* cv Chinese Spring), rice (*Oryza sativa* cv Amaroo) and maize (*Zea mays*) plants were grown in the greenhouse under natural light supplemented with sodium vapour lamps.

5.2.2 Nulli-tetrasomic Lines

A complete set of nulli-tetrasomic lines was obtained from Dr Adam Lukaszewski (Department of Botany and Plant Science, University of California, Riverside, CA). Southern-blot membranes carrying restriction-digested DNA from each line were kindly provided by Margaret Pallotta (University of Adelaide).

5.2.3 Database Searches and Sequence Analysis

Genbank (http://www.ncbi.nlm.nih.gov) cereal EST databases were screened for thioredoxin *h* cDNAs by BLAST searches using the published thioredoxin *h* sequences of rice (Ishiwitari et al. 1995) and wheat (Gautier et al. 1998) and the *Phalaris* thioredoxin *n* as the initial query sequences. All thioredoxin *h* sequences identified that were not identical to the query sequence were subsequently used to conduct further BLAST searches. This process was repeated until no new sequences were identified. The deduced proteins of all grass thioredoxin *h* sequences were aligned using the ClustalW program (Thompson et al. 1994) employing the Dayhoff PAM 250 matrix. Possible transmembrane sequences were analysed using the prediction programs:

TMHMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/)

TMPRED (http://www.ch.embnet.org/software/TMPRED)

TMAP (http://www.mbb.ki.se/tmap/)

SOSUI (http://sosui.proteome.bio.tuat.ac.jp/sosuiframe0.html)

The phylogeny of grass sequences was determined using the neighbour joining algorithm implemented in PAUP (Swofford 1998) with vertebrate thioredoxin sequences as an outgroup.

5.2.4 PCR Amplification of Thioredoxin *n* cDNAs from Wheat, Rice and Maize

Total RNA was extracted from mature anthers of wheat and rice and pollen of maize and was purified by CsCl gradient centrifugation (2.2.3). cDNAs were amplified by RT-PCR and 5'RACE using PCR Reactions 3, 4, 5 and 6 as described previously (Chapter 3) except that PCR Protocol 5 (Appendix B) was used for Reaction 6. PCR products were subcloned and sequenced according to the method outlined earlier.

5.2.5 RT-PCR Amplification of Wheat cDNA Probes

Total RNA was extracted from young leaves of wheat plants and purified by CsCl gradient centrifugation (2.2.3). First strand cDNA was synthesised with oligo-dT as the initiation primer and used as the template for PCR amplification of wheat cDNAs. Partial length cDNA sequences of wheat thioredoxin *h1, h2, h3* and *n* were amplified using the Ta1F/R, Ta2F/R, Ta3F/R and R1/T2 primer pairs (Appendix A), respectively. Wheat EST sequences were used as the templates for the design of the Ta1F/R and Ta3F/R primer sets whilst the Ta2F/R primer pair was designed against the published wheat thioredoxin *h* sequence (Gautier et al. 1998). Wheat peroxiredoxin (*TaPrx*) was amplified with the PrxF/R primer set (Appendix A) described earlier (Chapter 3). The wheat CDSP32 homologue (*TaCDSP32*) was amplified using the CDSP32F/R primer pair (Appendix A). The sequences of barley (Genbank accession number BE421930) and maize (BG841177, BM074832) CDSP32 homologues (Appendix G) were used as templates for the design of the CDSP32F/R primers.

All cDNAs were amplified using PCR protocol 5 (Appendix B) except wheat thioredoxin *n* (PCR Protocol 2). PCR products were confirmed by sequencing using the procedure described previously (Chapter 3).

5.2.6 Isolation of DNA and Southern Blotting Analysis

Southern-blot membranes carrying restriction-digested DNA (digested with *BamHI, HindIII* and *EcoRI*) from each nulli-tetrasomic line were hybridised with wheat thioredoxin *h* probes. The wheat thioredoxin *h* cDNAs, amplified by RT-PCR were radioactively labelled (2.9.3) and used as the probes for hybridisation (2.12.3). Membranes were washed four times to a final stringency of 0.2X SSC, 0.1% SDS (2.12.3).

5.2.7 Stress Treatments

Stress treatments were applied to four-week-old wheat plants (Chinese Spring). Plants were grown in the glasshouse under controlled temperature and (18-22°C) and 12 h day length (natural light supplemented by sodium vapour lamps).

5.2.7.1 Drought Stress

Wheat seeds were surface-sterilised in 10% sodium hypochlorite (20 min) and rinsed 10 times in autoclaved H_2O then germinated on moist, sterilised Whatmann paper. Five days after germination seedlings were transferred to a rockwool support medium (Grodan™) and grown in aerated hydroponic tanks containing Hoaglands solution (Hoagland and Arnon, 1950) that was replenished weekly. Plants were grown under controlled growth room conditions (25°C day/20°C night, 12 h photoperiod, 80% relative humidity, ~500μmol $m^{-2}s^{-1}$ photon flux density) and stressed at four weeks. Drought stress was applied by the addition of 18% (w/v) polyethylene glycol (PEG molecular weight 8000kDa) to the nutrient solution creating an osmotic potential of –3.0 MPa.

5.2.7.2 Methyl Viologen Treatment

Methyl viologen (Sigma, USA) was applied by spraying 10 ml per plant of a 10 μM solution in 25% Tween 20. Control plants were treated with 10 ml of 25% Tween 20 only. All treatments were applied in the glasshouse.

5.2.7.3 Ultraviolet Light Treatment

Pots containing glasshouse grown wheat plants were placed in a laminar flow cabinet (Gelman) and subjected to 2 hrs of UV irradiation ($1200mW/m^2$) then placed back into the glasshouse. Control plants were subjected to the same conditions except that they were exposed to 2 hrs of fluorescent light in the cabinet.

5.2.7.4 Hydrogen Peroxide (H_2O_2) Treatment

Young leaves of glasshouse grown plants were cut and floated in petri dishes containing a solution of 1 mM H_2O_2 for 1, 2 or 3 hrs. For the control treatment leaves were floated in autoclaved nanopure H_2O for 3 hrs. The treatment was performed in the glasshouse to ensure that light and temperature conditions were unchanged.

5.2.7.5 Wounding Treatment

The epidermal layer of young leaves was gently abraded with wet acid-washed sand by very light rubbing between the thumb and forefinger. Plants were glasshouse grown and the treatment was applied in the glasshouse on the youngest fully expanded leaf. For the control, leaves were gently rubbed between wet thumb and forefinger without sand.

5.2.8 Isolation of RNA and Northern Blotting Analysis

Total RNA was extracted from mature anthers (one day pre-anthesis), mature stigmas (one day pre-anthesis), young roots (from 7 day old germinated seeds), young leaves and stress-treated leaves (5.2.7) using the small-scale RNA isolation method (2.2.2). RNA concentrations were determined spectrophotometrically (2.2.4) and either 12μg or 20 μg of RNA fractionated through a 1.2% denaturing agarose gel (2.9.1). RNA was transferred to nylon membranes by capillary

blotting (2.9.2) and hybridised with radiolabelled wheat thioredoxin *h1, h2, h3, n, TaPrx* and *TaCDSP32*. The washing conditions were the same as for Southern blots.

5.3 Results

Characterising all members of the grass thioredoxin *h*-class required the identification of a large number of grass thioredoxins. Thioredoxin sequences where either amplified using the RT-PCR based procedure outlined in Chapter 3 or located in public grass EST databases. The stress-responsive expression profile of each thioredoxin *h* was then analysed in wheat.

5.3.1 Determining the Thioredoxin *h* Compliment in the Grasses

5.3.1.1 PCR Amplification of Thioredoxin *n* cDNAs from Wheat, Rice and Maize

Wheat (*Tan*), maize (*Zmn*) and rice (*Osn*) thioredoxin *n* cDNAs were amplified from pollen RNA by RT-PCR using the same approach described in Chapter 3. Comparison of the coding region of each sequence with that of *Phalaris* thioredoxin *n* (*Pcn*) revealed a high level of sequence conservation (Figure 5.1). *Tan, Zmn* and *Osn* were 95%, 86% and 87% similar to *Pcn* at the nucleotide level and shared 96%, 86% and 89% residue identity at the protein level respectively. The predicted-protein of Osn contains an additional glutamic acid residue in the N-terminal region of the protein whilst Zmn contains 2 additional residues (threonine and serine) at the carboxyl terminus.

5.3.1.2 Identification of Thioredoxin *h* ESTs

The 803,206 grass ESTs present in the NCBI EST-database were screened for *h*-class thioredoxins. 368 ESTs were identified as having high sequence similarity to the query thioredoxin sequences used. The details of these ESTs are summarised in Table 5.1.

Thioredoxin sequences including published grass *h* sequences, EST database clones and all grass *n*s cloned in this study were aligned using the ClustalW multiple alignment program (Appendix E). Sequences were found to form four distinct groups and were labelled subclasses 1, 2, 3 and n. A comparative

alignment of the deduced protein sequences of each thioredoxin h subclass from wheat, rice, sorghum and maize is given in Figure 5.2.

Subclass 1 comprised homologues of a rice thioredoxin *h*, which was identified in sieve elements (Ishiwitari et al. 1995). Mutational studies of the rice thioredoxin h have shown that the removal of the five N-terminal amino acids abolishes the ability of the protein to mediate its own transport across plasmodesmata (Ishiwitari et al. 1998). The deduced protein sequences of other grass members of subclass 1 are very similar (70-83% identity) to that of rice (Os1) and all have virtually identical N-terminal sequences.

Subclass 2 is composed of homologues of wheat seed thioredoxins *h* that have been located to the scutellum, aleurone and endosperm of germinating grain (Gautier et al. 1998; Serrato et al. 2001). The sequences of this group are also highly conserved (71-83% protein identity) with most differences restricted to the N-terminal region of the predicted proteins. The N-terminus of subclass 2 proteins is predominantly hydrophobic with a high proportion of alanine and threonine residues, a composition characteristic of protein presequences. However, it is unlikely that these proteins are post-translationally processed since Western-blot analysis of native protein extracts, with wheat subclass 2 thioredoxin *h* antibodies, has only identified a wheat band of size consistent with the full-length protein (Serrato et al. 2001). Both maize and sorghum subclass 2 proteins were found to encode an active site with the consensus WCPPC as has been described for *Arabidopsis* thioredoxins *h3, h4* and *h5* (Rivera-Madrid et al. 1995). These two sequences represent the only grass thioredoxins identified in this study that display this variant form of the more common WCGPC consensus.

Subclass 3 was found to be composed of sequences as yet undescribed in the literature. The predicted proteins of this group are the longest of all the grass thioredoxins *h* and, like the other 3 subclasses, are highly conserved (72%-90% identity). Similar to the subclass 2 thioredoxins, the N-terminus of subclass 3 proteins are predominantly hydrophobic (10 of the first 15 residues). In addition, proteins of all subclass 3 thioredoxins contain a typical endoplasmic reticulum (ER) retention signal (KDEL) near the C-terminus suggesting that these proteins

may reside in the ER of plant cells. To determine whether the hydrophobic N-termini represent a membrane bound domain, as has been described for two soybean thioredoxins *h* (Shi and Bhattacharyya, 1996), all subclass 3 sequences were analysed using the transmembrane prediction programs TMHMM, TMPRED, TMAP and SOSUI. Irrespective of the program used, the N-terminus of all subclass 3 representatives were predicted not to be anchored within a membrane.

Thioredoxin subclass n was the most conserved of the four subclasses with at least 86% amino acid identity between thioredoxin n proteins of different grass species. Grass thioredoxins n could be readily distinguished from other grass thioredoxins by presence of a conserved cysteine at position 4 in the mature protein and 3 conserved phenylalanines towards the carboxyl end.

5.3.1.3 Phylogenetic Analysis of Grass Thioredoxins *h*

All grass thioredoxin *h* sequences amplified and identified were used for phylogenetic analysis (Figure 5.3). Each thioredoxin *h* subclass formed a monophyletic clade with strong support among bootstrap replicates [(100% support for all subclasses except subclass 1 (93%)] indicating that the grass thioredoxin *h* class is indeed composed of four distinct subclasses. In addition, since each subclass clade is composed of sequences representing three different subfamilies of the grasses (the *Pooideae, Oryzoideae* and *Panicoideae*) it is clear that the four grass *h* subclasses evolved prior to the divergence of these subfamilies. In this analysis, subclass *n* is a sister group to all other grass thioredoxins *h* (78% bootstrap agreement) and more closely related to the vertebrate outgroup sequences, a topology suggesting that subclass *n* may represent a more ancient subclass. This interpretation is supported by the observation that other non-grass plant species contain thioredoxin *n* homologues (Chapter 3).

The position of the root in this phylogeny is strongly supported (100% of bootstrap replicates), however, the relative positions of subclass clades are less clear (note the lower bootstrap support for the basal node of subclasses 1 and 2). Since there is no variation in the position of the root in replicates and the

individual subclass clades are strongly supported, the uncertainty of basal node position must be due to variation in the relative positions of subclass clades during bootstrap replicates.

5.3.2 Characterisation of Wheat Thioredoxins *h*

Since an objective of this study was investigation of thioredoxin *h* expression patterns in response to oxidative stress, a wheat type-C peroxiredoxin (*TaPrx*) and a wheat CDSP32 homologue (*TaCDSP32*) were also amplified for use as markers of oxidative stress. *TaPrx* was employed for the same reasons as outlined in Chapter 3. *TaCDSP32* was chosen for two reasons. Primarily because the gene encoding CDSP32, a thioredoxin-like 32 kDa chloroplastic protein identified in potato, was found to be induced in response to several oxidative stresses (Broin et al. 2000). Secondly, this study provided the opportunity to determine whether the transcriptional response of grass homologues, such as *TaCDSP32*, to oxidative stress mirrored that of potato.

5.3.2.1 Amplification of Wheat cDNA Probes

cDNA probes of all four wheat thioredoxin *h* subclasses, a wheat type-C peroxiredoxin (*TaPrx*) and a wheat CDSP32 homologue (*TaCDSP32*) were amplified from leaf RNA by RT-PCR. Visualisation of the PCR products in an agarose gel (Figure 5.4a and b) revealed bands of expected size: 327 bp for *TaPrx*, 394bp for *TaCDSP32* and 226bp, 155bp, 260bp and 171bp for wheat thioredoxins *h1*, *h2*, *h3* and *n* respectively. The identity of each PCR product was confirmed by sequencing and comparison with the sequences against which the PCR primers were designed. The sequences of the wheat thioredoxins *h1*, *h2*, *h3* and *n* and *TaPrx* are given in Appendix F. The *TaCDSP32* PCR sequence was aligned with the corresponding region of the potato CDSP32 cDNA (Rey et al. 1998) and the barley and maize CDSP32 homologues. *TaCDSP32* displayed 95%, 85% and 70% nucleotide similarity with the barley, maize and potato sequences respectively (Appendix G) indicating that this gene is highly conserved.

5.3.2.2 Thioredoxin *h* Copy Number and Chromosomal Location in Wheat

The chromosomal location of each thioredoxin *h* gene in wheat was determined by RFLP analysis of a set of nulli-tetrasomic lines. Each line is missing an individual

pair of chromosomes that have been replaced by an additional homoeologous pair (eg nulli 2A/tetra 2B has lost both copies of chromosome 2A yet contains two additional copies of 2B) Consequently, the absence of a RFLP band indicates that the sequence corresponding to the hybridisation probe is located on the missing chromosome. Southern blots of a complete set of nulli-tetrasomic wheat lines probed with all wheat thioredoxins *h* are given in Figure 5.5. Based upon the RFLP pattern of each blot it is clear that each thioredoxin gene is located on a single group of chromosomes with at least a single copy on each homoeologous member of the group. The *Ta1* gene maps to chromosome group two, *Ta2* and *Tan* to group one and *Ta3* to group five.

Each probe resulted in relatively simple restriction fragment length polymorphism (RFLP) pattern of between 4 (*HindIII*/*Tan*) and 8 bands. As none of the probes is cut by the restriction enzymes chosen, the presence of multiple bands has two possible explanations. Firstly, there are restriction enzyme sites in the genomic sequence corresponding to each cDNA. Alternatively, the wheat genome carries multiple copies of thioredoxins *h1, h2* and *h3*. Given that wheat is a hexaploid and potentially carries three copies of each *h*-subclass gene, the simplicity of RFLP patterns suggests that each gene is likely to be present as a single copy per genome

5.3.2.3 Expression Patterns of Wheat Thioredoxins *h*

Transcript levels of wheat thioredoxins *h* in mature anthers, mature stigmas, roots and leaves were determined by RNA gel blot analysis. Transcripts of each thioredoxin were detected in all tissues examined, but varied in abundance (Figure 5.6). Based upon the strength of the hybridisation signal and the duration each blot was exposed to X-ray film, thioredoxin *h1* was found to be the most abundant transcript in the tissues investigated followed by *h2*, *h3* and lastly *n*. For thioredoxins *h1, h2* and *h3* the lowest level of transcript was detected in leaf. mRNAs of both thioredoxins *h1* and *h3* were present in similar quantities in anther, stigma and root whilst *h2* expression was highest in stigma and root. By contrast, the wheat thioredoxin *n* transcript was most abundantly expressed in anther tissue, and at a lower and similar level in other tissues. In addition to the four tissues investigated here the transcript of *h2* has been detected in germinating

seed (Gautier et al. 1998) whilst EST data also indicates that *h3* is present in seed and that all thioredoxins are expressed in the spike (Table 5.1).

5.3.2.4 Expression of Wheat Thioredoxins *h* in Response to Oxidative Treatments

Wheat plants were subjected to five abiotic stresses: i) spraying with methyl viologen, a chemical that disrupts the flow of electrons in chloroplast photosystem I and results in the generation of superoxide (Babbs et al. 1989); ii) direct exposure of whole leaves to 2mM H_2O_2; iii) exposure of whole plants to very high levels of ultraviolet light (UV); iv) mechanical damage designed to stimulate defence responses in plants; v) drought stress sufficient to reduce leaf relative water content to 70% in 48 hrs.

The expression response of each thioredoxin and the control probes *TaPrx* and *TaCDSP32* to stress treatments was determined by Northern blotting. All thioredoxins displayed changes in expression pattern in response to at least one of the abiotic stresses applied. The most pronounced changes detected were strong inductions of thioredoxins *h1* and *h2* in response to UV treatment (Figure 5.6). Thioredoxin *h1* displayed the most rapid induction response with a detectable increase in transcript abundance 2 hours after the initiation of UV treatment and an even higher level after 24 hours. By contrast, the quantity of *h2* mRNA appeared unchanged after 2 hours yet by 24 hours was approximately an order of magnitude more abundant than in the control (non-stressed) leaf sample. Similarly, thioredoxin *h3* transcript accumulation was only evident 24 hours after UV treatment, however, unlike *h1* and *h2* the level of induction was significantly lower. Interestingly, the initial response of thioredoxin *n* to several stresses appears to be a reduction in transcript levels. 2 hours after UV treatment thioredoxin *n* transcript was less abundant than the unstressed control level yet within 24 hours was present at normal levels. A similar pattern was recorded for methyl viologen treated plants with a small reduction in thioredoxin *n* message levels detected in leaves 6 hours after spraying and an apparent recovery to endogenous levels by 24 hours. By comparison, transcript levels of all other thioredoxins were found to increase in response to methyl viologen application.

Thioredoxin *h1* was again found to be most rapidly induced with a detectable increase within 6 hours of spraying.

Wounding appeared to induce the rapid expression of thioredoxins *h2*, *h3* and *n* since Northerns of all three detected a greater quantity of message in leaves 2 hours after wounding. However, the quantity of RNA present in the 2 and 24 hour post wounding sample lanes was also higher than the control. Consequently, these results are probably a function of unequal loading or transfer of RNA during the preparation of the Northern membranes.

Thioredoxin *n* was found to be the only gene to clearly respond to the H_2O_2 treatment. Thioredoxin *n* transcript levels were significantly lower in leaf tissue 1, 2 and 3 hours after H_2O_2 stress was applied than in the control sample. Although lower, the quantity of thioredoxin *n* increased with time indicating a rapid depletion of *n* message upon treatment followed by the transcription of replacement RNA.

Drought stress elicited a different expression response from each thioredoxin in wheat leaves (Figure 5.7). As expected, the *TaCDSP32* control probe was upregulated in response to drought stress confirming that the plants used in the study were indeed under stress. Thioredoxin *h2* transcript levels were unaltered by drought stress or plant rehydration whilst *h3* expression was only induced in recovered plants. By contrast, thioredoxin *h1* expression was strongly upregulated with the level of *h1* transcript increasing over time. Thioredoxin *n* was similarly induced by drought however, unlike *h1*, the *n* transcript level only increased initially and by 48 hours had fallen to less than unstressed levels. Significantly, all mRNAs except *h2* were also upregulated in plants recovered after 48 hours of drought stress. This expression pattern suggests that the cellular processes of recovered plants had not, after 24 hours rehydration, returned to normal. Indeed, the expression profiles of thioredoxin *n* and *TaCDSP32* would suggest that rehydration resulted in the recovery of the transcriptional apparatus and that h1, h3, n and TaCDSP32 proteins were required for cellular repair.

Interestingly, in wheat leaves *TaPrx* expression was only induced by UV and to a very slight degree by drought stress. This expression profile not consistent for a gene with a generalised role in oxidative protection and suggests that *TaPrx* may only be involved in the response to specific forms of oxidative stress (eg light stress).

5.3.2.5 Examining the Promoters of Rice Thioredoxins *h1* and *h2*

The rapid induction of thioredoxin *h1* mRNA in response to methyl viologen, UV and drought stress treatments and the large accumulation of *h2* mRNA in response to UV raises questions as to the transcriptional regulation of these genes. Numerous plant genes are induced at the transcriptional level in response to various abiotic stresses including those examined in this study (Chandler and Robertson, 1994; O'Donnell et al. 1996; Shinozaki and Yamaguchi-Shinozaki, 1997; Thomashow 1999). For many of these genes stress-responsive expression has been shown to be modulated by transcription factors interacting with specific *cis*-elements present in the gene promoter (Martin and Paz-Ares, 1997; Stockinger et al. 1997; Busk and Pagés, 1998; Liu et al. 1998). Since the core sequence motifs of *cis*-elements targeted by specific transcription factors or families of transcription factors are often highly conserved (Giuliano et al. 1988; Izawa et al. 1993), their presence in the promoter of a new gene provides clues as to the potential regulation of the gene. The recent sequencing of the entire rice genome (Goff et al. 2002) provided the opportunity to examine the promoters of rice thioredoxins *h1* and *h2* for such regulatory-element motifs. Assuming that the gene representing each thioredoxin *h* subclass in rice is performing a similar function to its wheat homologue, the results of promoter analyses in rice may provide clues as to the transcriptional regulation of wheat thioredoxins.

Rice thioredoxins *h1 (Os1), h2 (Os2), h3 (Os3)* and *n (Osn)* were identified on P1 artificial chromosome (PAC) contigs of rice chromosomes 7, 5, 3 and 5, respectively. 1.5kb of genomic sequence upstream of the *Os1* and *Os2* initiation codons was analysed for the presence of putative *cis*-acting regulatory elements, recognised by stress-responsive plant transcription factors, using the search program PLACE (http://www.dna.affrc.go.jp/htdocs/PLACE/). The search identified TATA-box motifs 126 bp and 309 bp upstream of the translational start

codon of *Os1* and *Os2*, respectively, and numerous sequence motifs previously shown to act as binding sites for a range of plants transcription factors. Of these, five motifs in the *Os1* sequence and three motifs in *Os2* sequence represented potential stress-responsive regulatory elements (Figure 5.8).

Upstream of the *Os1* TATA-box, the first potential element was a motif that represents the core binding sequence of two cereal Dof proteins, maize PBF and barley BPBF (Vicente-Carbajosa et al. 1997; Mena et al. 1998). Dof proteins are a class of DNA-binding proteins that contain a highly conserved Cys_2-Cys_2 zinc-finger motif (Yanagisawa 1996). The second element was a telomere motif, the *telo*-box, a sequence conserved in all known plant translation elongation factor (*eEF1A*) genes (Tremousaygue et al. 1999). The remaining three elements were sequence motifs demonstrated to be targeted by the maize Dof1 and Dof2 DNA-binding proteins, proteins proposed to act as transcriptional activators or repressers (Yanagisawa and Sheen, 1998). Of the three *Os2* motifs detected the first represented the core sequence motif of the *Arabidopsis* C-repeat/dehydration response element (C/DRE), an element essential for transcriptional activation in response to cold, drought and/or salt stress (Yamaguchi-Shinozaki and Shinozaki, 1994). The second motif CCGAAA is identical to a hexanucleotide sequence found to be involved in the low-temperature response of the barley *blt4.9* gene (Dunn et al. 1998). The third putative element was a GCC-box, the core sequence of the ethylene-responsive element in tobacco and the target of ethylene-responsive transcription factors (Ohme-Takagi and Shinshi, 1995).

Table 5.1 Summary of the database search for *h* class thioredoxin in grass species.

[a] The Genebank accession number of a representative EST clone encoding each thioredoxin *h* subclass sequence is provided for each species. In the case of thioredoxin *h* sequences that have been published the reference is given.

[b] A summary of the tissues in which each thioredoxin mRNA is present (derived from EST database sequences). The tissues are abbreviated and are: A anther; Ap apex; C caryopsis; Ca callus; E ear; Em embryo; En endosprm; G glume; L leaf; O ovary; P pollen; Pa panicle; R root; Ra rachis; S seedling; Sc sperm cells; Sd seed; Sh shoot; Sk silk; Sp spike; T tassel; Te testa; Wp whole plant. The presence of (vs) after the tissue abbreviation indicates that the mRNA has been extracted from various developmental stages of that particular tissue. Tissue subjected to a form of abiotic stress is indicated by (as) after the tissue abbreviation whilst (bs) indicates biotic stress.

Table 5.1 Summary of the database search for *h* class thioredoxins in grass species.

Thioredoxin subclass	Species	Number of database entries	GenBank accession number [a]	Tissues in which mRNA is present [b]
Subclass 1	*Aegilops speltoides (As1)*	1	BF291690	A
	Hordeum vulgare (Hv1)	41	BM376549	C, Em, L(vs), R, R(as), S(bs), Sp
	Oryza sativa (Os1)	49	(Ishiwitari et al. 1995)	Ca, L(vs), Pa, R, S, Sc, Sh
	Sorghum bicolor (Sb1)	12	AW923038	L(bs), S, Wp(as)
	Triticum aestivum (Ta1)	6	BE404630	R, Sp(vs)
	Zea mays (Zm1)	26	AI770787	A, E, En, L(vs), P, S, Sh, Sk, T
Subclass 2	*Hordeum vulgare (Hv2)*	47	AV930350	C, Em, En, L(bs), L(vs), R(as), R(vs), Sh, Sh(as), Sp, Sp(bs), Te
	Oryza sativa (Os2)	17	AU032254	L(as), L(vs), Pa(vs), Sd, Sh
	Secale cereale (Sc2)	4	BE586970	A, R, R(as)
	Sorghum bicolor (Sb2)	12	AW924685	Em, O, Wp(as)
	Triticum aestivum (Ta2)	16	(Gautier et al. 1999)	En, L, R(vs), S(as), Sp
	Triticum durum (Td2)	1	(Gautier et al. 1999)	En
	Triticum monococcum (Tm2)	2	BG608088	Ap(vs)
	Zea mays (Zm2)	28	BE510268	A, E(vs), En, G, P, S, Sh, Sk, T, T(vs)
Subclass 3	*Aegilops speltoides (As3)*	1	BF292230	A
	Hordeum vulgare (Hv3)	8	BG418853	L, R, Ra, Sp, Te
	Oryza sativa (Os3)	14	AU069205	Ca, En, L, Pa, Sh
	Secale cereale (Sc3)	1	BE637204	A
	Sorghum bicolor (Sb3)	3	AW678266	Wp, Wp(as)
	Triticum aestivum (Ta3)	3	BE402773	En, Sp, Sp(bs)
	Triticum monococcum (Tm3)	1	BG607107	A
	Zea mays (Zm3)	28	AW566291	Ap, E, L, S, Sh(vs), Sk, T
Subclass n	*Hordeum vulgare (Hvn)*	29	BE437654	C, L(bs), L(vs), R(as), R(vs), Sh(as), Sh(vs)
	Oryza sativa (Osn)	2	AU082255	L, Pa
	Sorghum bicolor (Sbn)	4	BG047831	O
	Triticum aestivum (Tan)	1	BI750469	Sp(bs)
	Zea mays (Zmn)	11	BM269264	A, G, P, S, Sk

(a)

```
Pcn ATGGGAGGCTGTGTGGGCAAGGATCGTGGCATTGTGGAA---GACAAGCTTGATTTCAAAGGTGGGAA  65
Tan ATGGGGGGCTGTGTGGGCAAGGATCGTAGCATTGTGGAA---GAAAAGCTTGATTTCAAAGGTGGAAA  65
Osn ATGGGAGGTTGTGTGGGCAAGGGACGACGACATATTGAGGAAGACAAGCTTGATTTCAAAGGTGGAAA  68
Zmn ATGGGAGGCTGTGCGGGAAAGGTACGTCGTGATGACGAA---GAAAAGCTTGATTTTAAAGGTGGAAA  65

Pcn TGTGCATGTCATAACTACCAAAGAGGACTGGGACCAGAAAATTGCAGAAGCAAACAAGGATGGGAAAA 133
Tan TGTGCATGTCATAACAACCAAAGAGGACTGGGACCAGAAGATTGAAGAAGCAAACAAGGATGGGAAAA 133
Osn TGTTCATGTCATAACAAGCAAAGAGGACTGGGATAGGAAGATTGAAGAAGCAAACAAGGATGGGAAAA 136
Zmn TGTTCATATTATAACAAGCAATGAGGGCTGGGACCAGAAGATTGCAGAAGCAAACAGAGATGGGAAAA 133

Pcn TTGTTGTGGCAAATTTCAGTGCTTCCTGGTGTGGGCCATGCCGTGTCATTGCACCTGTTTATGCTGAG 201
Tan TTGTTGTAGCAAACTTAAGTGCTTCGTGGTGTGGGCCATGCCGTGTCATTGCACCTGTTTATGCCGAG 201
Osn TTGTTGTGGCAAATTTCAGTGCTTCCTGGTGTGGACCATGTCGTGTTATTGCACCTATTTATGCTGAG 204
Zmn CTGTTGTTGCAAATTTCAGCGCTTCCTGGTGTGGGCCATGCCGTGTCATTGCTCCTGTCTATGCTGAA 201

Pcn ATGTCAAAGACGTATCCTCAACTCATGTTCTTGACAATTGATGTTGATGACCTAGTGGATTTCAGCTC 269
Tan ATGTCCAAGACTTATCCTCAACTCATGTTCTTGACAATTGATGTTGATGACCTAATGGATTTCAGCTC 269
Osn ATGTCAAAGACTTATCCTCAACTCATGTTCTTGACAATAGATGTTGATGACCTAATGGATTTCAGCTC 272
Zmn ATGTCAAAGACTTATCCTCAACTCATGTTCTTGACAATAGACGTTGATGACCTGATGGACTTCAGCTC 269

Pcn AACATGGGACATCCGTGCAACCCCAACGTTCTTCTTCCTCAAGAATGGCCAGCAGATCGACAAGCTCG 337
Tan TACATGGGACATCCGTGCAACCCCGACGTTCTTCTTCCTCAAAAACGGCCAGCAGATCGAAAAGCTCG 337
Osn ATCATGGGATATCCGCGCAAAGCCGACCTTTTTCTTCATAAAAAACGAGAAGCAGGTCGACAAGCTTG 340
Zmn TTCATGGGACATCCGTGCAACCCCGACATTCTTCTTCCTCAAGAACGGGCAGCAGATCGACAAGCTCG 337

Pcn TTGGCGCCAACAAGCCTGAGCTCGAAAAGAAAGTACAAGCTCTTGGCGATGGCAGT------TGA 396
Tan TCGGCGCCAACAAACCTGAGCTCGAAAAGAAAGTACAAGCTCTTGGCGATGGCAGT------TGA 396
Osn TGGGCGCCAACAAACCTGAGCTCGAAAAGAAGGTTCAGGCACTTGCTGATGGCAGC------TGA 399
Zmn TGGGCGCGAACAAACCTGAGCTCGAGAAGAAAGTGCTAGCAGCCGCTGATGCCAGTACGTCCTAG 402
```

(b)

```
Pcn MGGCVGKDRGIV-EDKLDFKGGNVHVITTKEDWDQKIAEANKDGKIVVANFSASWCGPCR  59
Tan MGGCVGKGRSIV-EEKLDFKGGNVHVITTKEDWDQKIEEANKDGKIVVANFSASWCGPCR  59
Osn MGGCVGKDRRHIEEDKLDFKGGNVHVITSKEDWDRKIEEANKDGKIVVANFSASWCGPCR  60
Zmn MGGCAGKVRRDD-EEKLDFKGGNVHIITSNEGWDQKIAEANRDGKTVVANFSASWCGPCR  59

Pcn VIAPVYAEMSKTYPQLMFLTIDVDDLVDFSSTWDIRATPTFFFLKNGQQIDKLVGANKPE 119
Tan VIAPVYAEMSKTYPQLMFLTIDVDDLMDFSSTWDIRATPTFFFLKNGQQIDKLVGANKPE 119
Osn VIAPIYAEMSKTYPQLMFLTIDVDDLMDFSSSWDIRAKPTFFFIKNEKQVDKLVGANKPE 120
Zmn VIAPVYAEMSKTYPQLMFLTIDVDDLMDFSSSWDIRATPTFFFLKNGQQIDKLVGANKPE 119

Pcn LEKKVQALGDGS   131
Tan LEKKVQALGDGS   131  (96 %/ 98%)
Osn LEKKVQALADGS   132  (89 %/ 95%)
Zmn LEKKVLAAADASTS 133  (86 %/ 90%)
```

Figure 5.1 Comparative alignments of the nucleotide (a) and deduced amino-acid (b) sequences of thioredoxins *n* from *Phalaris coerulescens* (Pcn), *Triticum aestivum* (Tan), *Oryza sativa* (Osn) and *Zea mays* (Zmn). Dashes indicate gaps introduced into the alignment. Nucleotides and residues different to those of the Pcn sequence are in bold and the active site motif is in blue. The percentage protein identity and similarity of each gene relative to the *Phalaris* n protein is given in brackets at the end of the sequence (% identity/%similarity).

Grass thioredoxin *h* subclass 1

```
Ta1 MAAEEGAVIACHTKQEFDTHMANGKETGKLVIIDFTASWCGPCRVIAPVFAEYAKKFPGAI  61
Os1 MAAEEGVVIACHNKDEFDAQMTKAKEAGKVVIIDFTASWCGPCRFIAPVFAEYAKKFPGAV  61
Zm1 MASEQGVVIACHSKAEFDAHMTKAQEAGKLVVIDFTAAWCGPCRAIAPLFVEHAKKFTQVV  61
Sb1 MASEEGVVIACHTKAEFDAQMAKAKEAGKLVVIDFTASWCGPCRAIAPLFVEHAKKYTQAV  61

Ta1 FLKVDVDELKDVAEAYNVEAMPTFLFIKDGAKVDTVVGGRKDDIHTKIVALMGSASA      118
Os1 FLKVDVDELKEVAEKYNVEAMPTFLFIKDGAEADKVVGARKDDLQNTIVKHVGATAASASA 122
Zm1 FLKVDVDEVKEVTAAYEVEAMPTFHFVKNGKTVATIVGARKDELLAQIEKHAAPAPASASA 122
Sb1 FLKVDVDELKEVTAEYKIEAMPTFHFIKNGETVETIVGARKDELLALIQKHTASASA      118
```

Grass thioredoxin *h* subclass 2

```
Ta2 MAASAATATATAAV-GA-GEVISVHSLEQWTMQIEEANAAKKLVVVDFTASWCGPCRIMAP  59
Os2 MAAASAAA-------QAEGTVIAIHSLDEWTIQIEEANSAKKLVVIDFTASWCGPCRIIAP  54
Zm2 MAASEAAAAAATPVTPTEGTVIAIHSLEEWSIQIEEANSAKKLVVIDFTATWCPPCRAMAP  61
Sb2 MATTEAAAA--TPVAPAEGSVIAIHSLDEWSIQIEEANSAKKLVVIDFTATWCPPCRMIAP  59

Ta2 VFADLAKKFPAAVFLKVDVDELKPIAEQFSVEAMPTFLFMKEGDVKDRVVGAIKEELTTKV 120
Os2 VFADLAKKHTNAVFLKVDVDELKPIAEQFSVEAMPTFLFMKEGDVKDRVVGAMKDELASKV 115
Zm2 IFADMAKKSPNVVFLKVDVDEMKTIAEQFSVEAMPTFLFMREGDVKDRVVGAAKEELARKL 122
Sb2 VFAELAKKHPNVVFLKVDVDEMKTIAEQFSVEAMPTFLFMREGDVKDRVVGAAKEELANKL 120

Ta2 GLHAAQ 126
Os2 QLHMA  120
Zm2 ELHMAS 128
Sb2 QLQMAQ 126
```

Grass thioredoxin *h* subclass 3

```
Ta3 MGSFLSSLWTPPPLPGDDPDSAVVAVHSKPAWDRHWEAHR-NACKLMVIDFSASWCGPCRF  60
Os3 MGSFFSTMFTPPPAADDGGDSRVVAVHSTATWDEQWGAHKSNPNKLIVIDFSATWCGPCRF  61
Zm3 MGSFLSSLVTPPPAADD-PNCAVVAAHSKATYDEQWAAHK-SSSKLMVIDFSASWCGPCRF  59
Sb3 MGNFISSLVTPPPAADD-PNCAVVAAHSKTTYDEQWAAHK-NGGKLMVIDFSASWCGPCRF  59

Ta3 IEPAFKEMASRFADALFVKIDVDELAEVAKTFRVEAMPTFVLVKGGQEVSRVVGAKKDELD 121
Os3 IEPAFKDMAGHFADAVFFKIDVDELSEVARQWKVEAMPTFVLIKGGKEVSRVVGAKKDELE 122
Zm3 IEPAFKELASRFSDAIFVKVDVDELAEVARTWKVEAMPTFVLVKDGKEVSRVVGAKKDELE 120
Sb3 IEPAFKELASRFTDAIFVKIDVDELAEVARTWKVEAMPTFVLVKDGKEVSRVIGAKKDELE 120

Ta3 RKIKTFISSS    131
Os3 RKVNMFISSSSS  134
Zm3 RKIRMFTSSSSS  132
Sb3 RKIQMFIMSSSSS 133
```

Grass thioredoxin *h* subclass n

```
Tan MGGCVGKGRSIV-EEKLDFKGGNVHVITTKEDWDQKIEEANKDGKIVVANLSASWCGPCRV  60
Osn MGGCVGKGRRHIEEDKLDFKGGNVHVITSKEDWDRKIEEANKDGKIVVANFSASWCGPCRV  61
Zmn MGGCAGKVRRDD-EEKLDFKGGNVHIITSNEGWDQKIAEANRDGKTVVANFSASWCGPCRV  60
Sbn MGGCIAKQHADD-EDKIDFGGGNVHVVTSKEDWDQKIAEANKDGKIVVANFSASWCGPCRV  60

Tan IAPVYAEMSKTYPQLMFLTIDVDDLMDFSSTWDIRATPTFFFLKNGQQIDKLVGANKPELE 121
Osn IAPIYAEMSKTYPQLMFLTIDVDDLMDFSSSWDIRAKPTFFFIKNEKQVDKLVGANKPELE 122
Zmn IAPVYAEMSKTYPQLMFLTIDVDDLMDFSSSWDIRATPTFFFLKNGQQIDKLVGANKPELE 121
Sbn ISPVYAEMSQTYPQLMFLTIDVDELMEFSSSWDIRATPTFFFLKNGQQVDKLVGANKPELE 121

Tan KKVGALGDGS    131
Osn KKVQALADGS    132
Zmn KKVLAAADASTS 133
Sbn KKVAAIAGAS    131
```

Figure 5.2 Deduced amino acid sequences of wheat (Ta), rice (Os), maize (Zm) and sorghum (Sb) representatives of each grass thioredoxin *h* subclass. Identifiers of sequences derived from published sources are highlighted in red. Dashes indicate gaps introduced to maximise alignments and the thioredoxin active site sequence is in bold. Residues of each subclass that are discussed in the text are in blue.

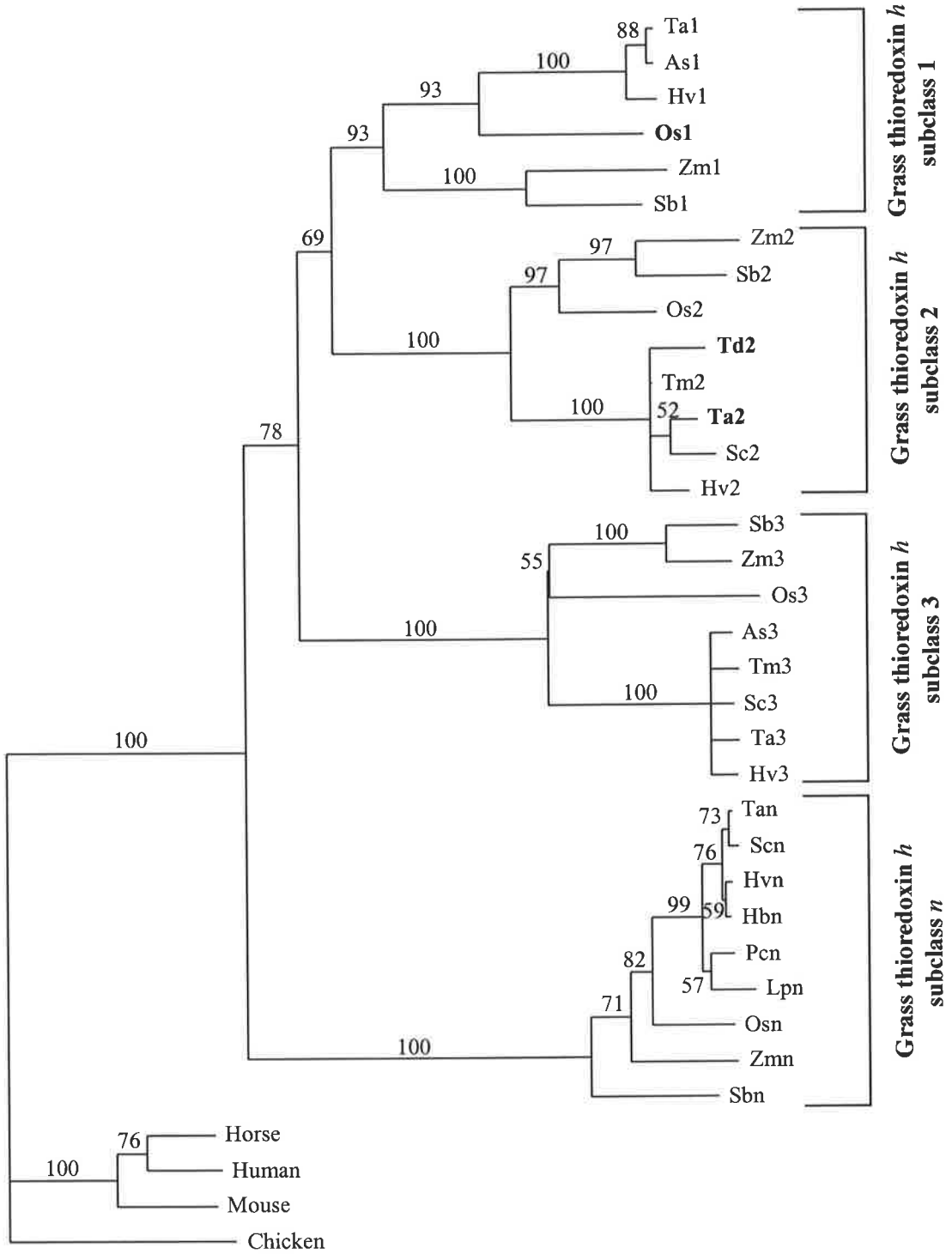


Figure 5.3 Phylogeny of grass *h*-class thioredoxins

Deduced amino acid sequences of EST clones and published thioredoxin *h* protein sequences were aligned using ClustalW (Appendix E). The phylogeny was determined with the neighbour joining algorithm implemented in PAUP (Swafford 1998) using distances calculated as mean percentage amino acid difference. Numbers indicate the percentage of 1000 bootstrap replicates in which a group was found (numbers <50% not shown). Species are identified by their initials (refer to Table 5.1). The initials of the published rice Os1 (Ishiwitari et al. 1995) and wheat Ta2, Td2 (Gautier et al. 1998) sequences are highlighted in bold. The vertebrate cytosolic-thioredoxin sequences of horse (GenBank acession number AB022431), human (AF276919), mouse (NM_011660) and chicken (J03882) were specified as outgroup sequences.

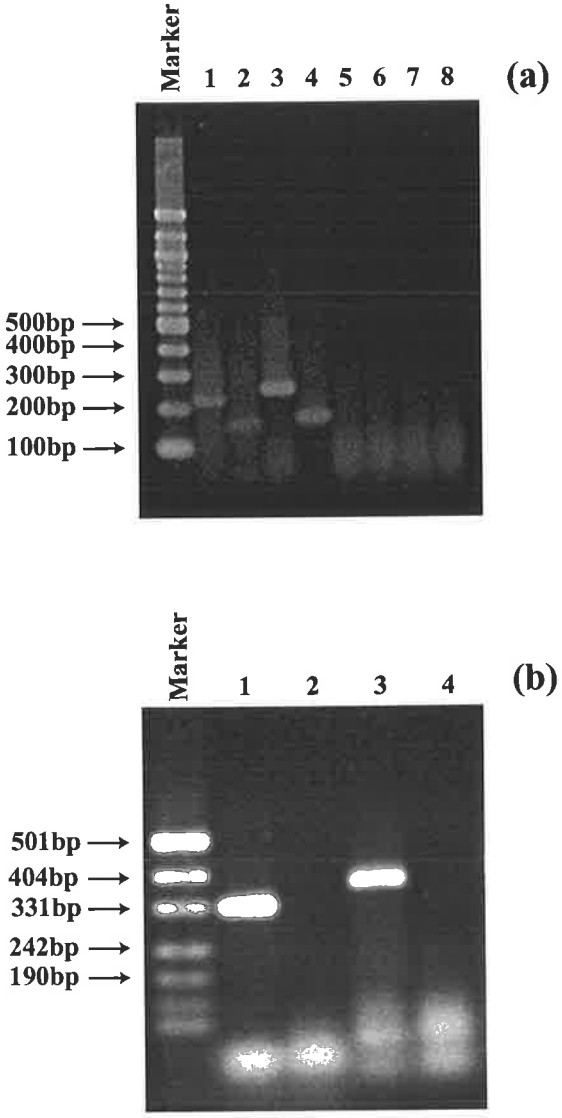


Figure 5.4 PCR amplification of wheat cDNA probes
(a) 2% agarose gel of the PCR fragments (cDNA) amplified using primer sets complimentary to each wheat thioredoxin *h* sequence. Lane 1, wheat thioredoxin *h1*; Lane 2, *h2*; Lane 3, *h3*; Lane 4, *n*; Lanes 5-8, control PCRs (no template cDNA) of wheat *h1, h2, h3* and *n* respectively. (b) 2% agarose gel of the PCR amplification of the wheat C-type peroxiredoxin (*TaPrx*) and CDSP32 homologue (*TaCDSP32*) cDNA probes. Lane 1, RT-PCR product using the PrxF/R primer set; Lane 2, control (no template cDNA) with the PrxF/R primer set; Lane 3, RT-PCR product using the TaCDSP32F/R primer set; Lane 2, control (no template cDNA) with the TaCDSP32F/R primer set. Size marker bands are given in base pairs.

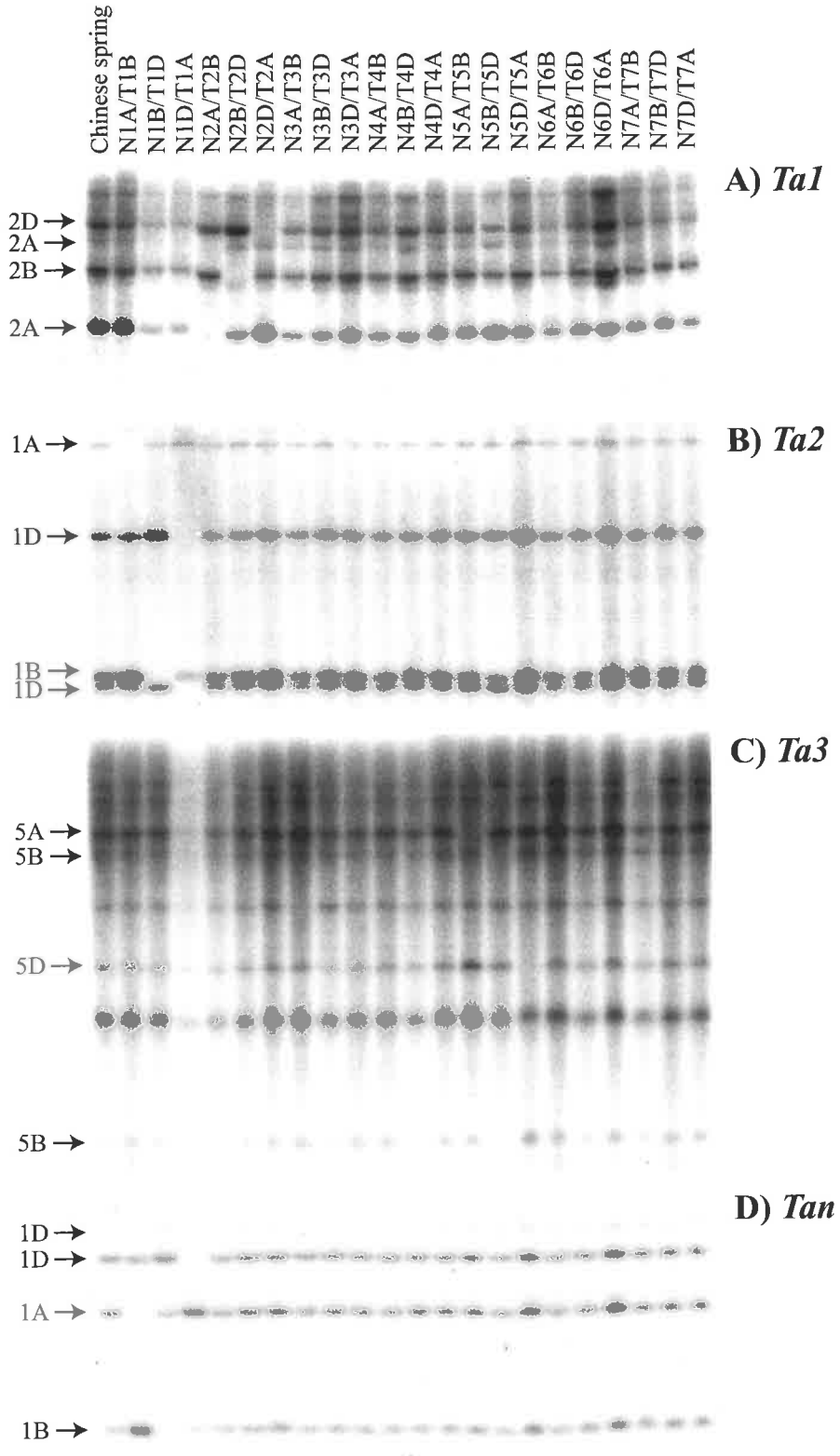


Figure 5.5 The chromosomal locations of wheat thioredoxins *h*.
Southerns are nulli-tetrasomic lines: A) digested with *BamHI* and probed with wheat thioredoxin *h1* cDNA; B) digested with *HindIII* and probed with thioredoxin *h2* cDNA; C) digested with *EcoRI* and probed with thioredoxin *h3* cDNA; D) digested with *HindIII* and probed with thioredoxin *n* cDNA. The chromosome pair missing in each line is indicated with an N and the additional copy with a T. RFLP bands associated with specific chromosomes are indicated.

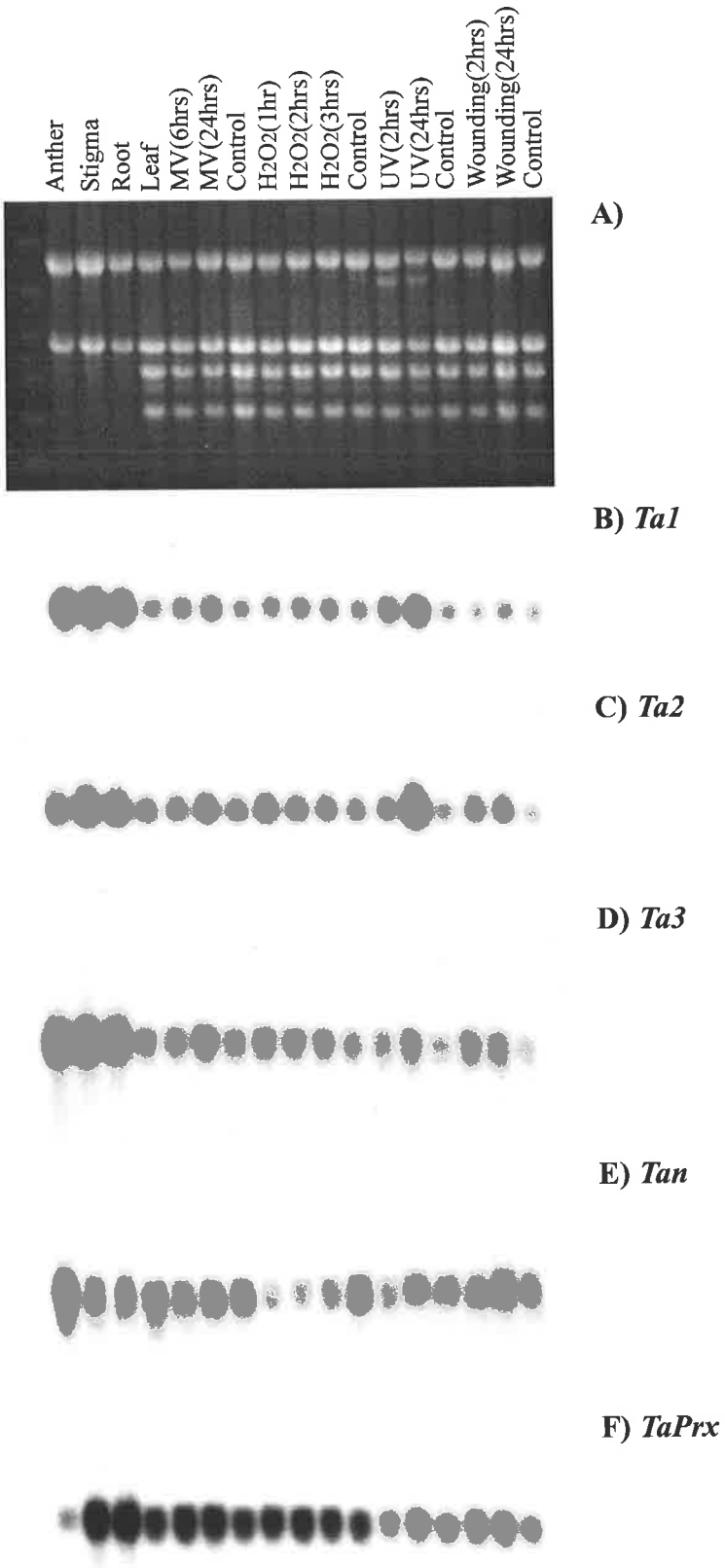


Figure 5.6 Expression profiles of wheat thioredoxins *h* in anther, stigma, root and leaf tissue and in leaves of plants under abiotic stresses.
A) Denaturing agarose gel of total RNA (12 ug per lane) derived from different tissues of unstressed plants (lanes 1-4) and leaves of abiotically stressed plants (lanes 5-17). B-F) Northern blots with wheat thioredoxins *h* and *TaPrx* cDNAs as probes. Abbreviations: MV, methyl viologen treatment; H_2O_2, hydrogen peroxide treatment; UV, ultraviolet light exposure. Probe exposure times are: *Ta1* 1 day, *Ta2* 3 days, *Ta3* 9 days, *Tan* 8 days, *TaPrx* 2 days.

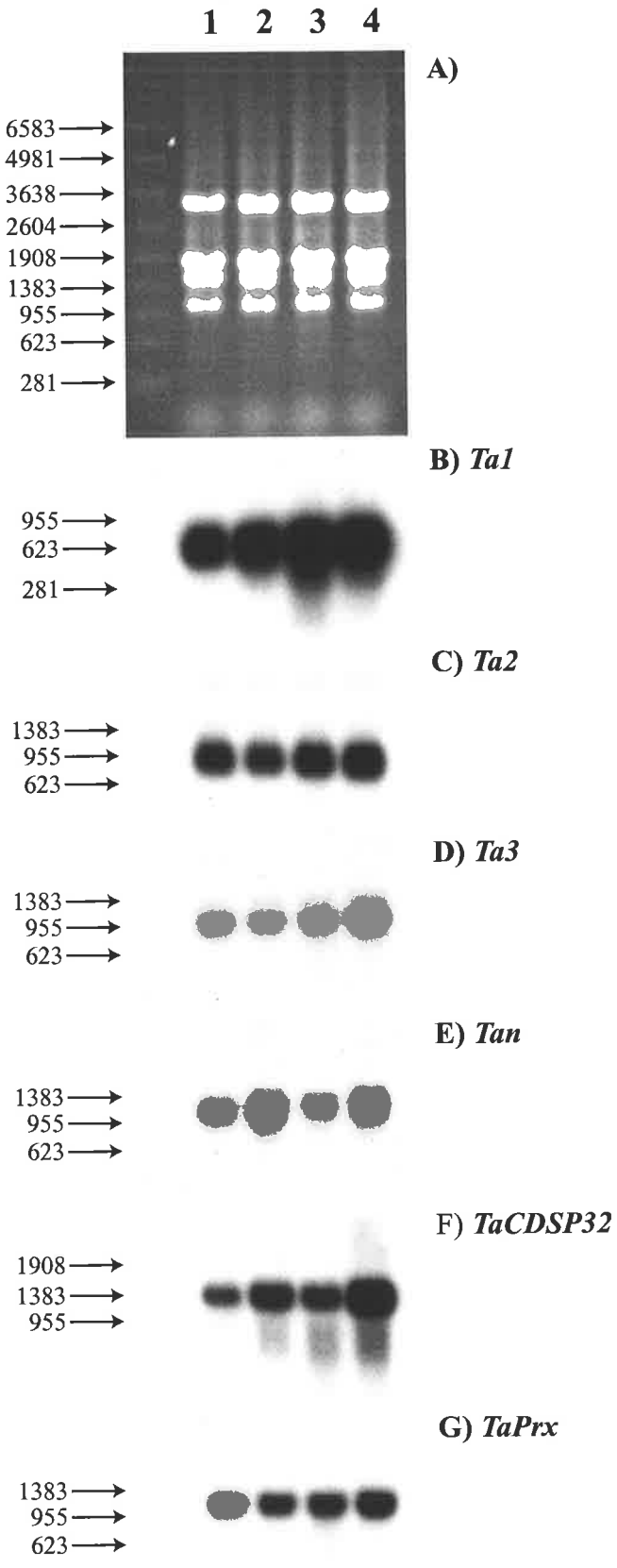


Figure 5.7 Expression of thioredoxin *h* genes in wheat plants subjected to drought stress.

A) Denaturing agarose gel of total leaf RNA (20 ug per lane) derived from: 1) Control (unstressed) plants; 2) Plants subjected to 24 hrs drought stress; 3) Plants subjected to 48 hrs drought stress; 4) Plants subjected to 48 hrs drought stress then recovered for 24 hrs. B-G) Northern blots with wheat thioredoxins *h, TaPrx* and *TaCDSP32* cDNAs as probes. Sizes of RNA marker bands (in base pairs) are given.

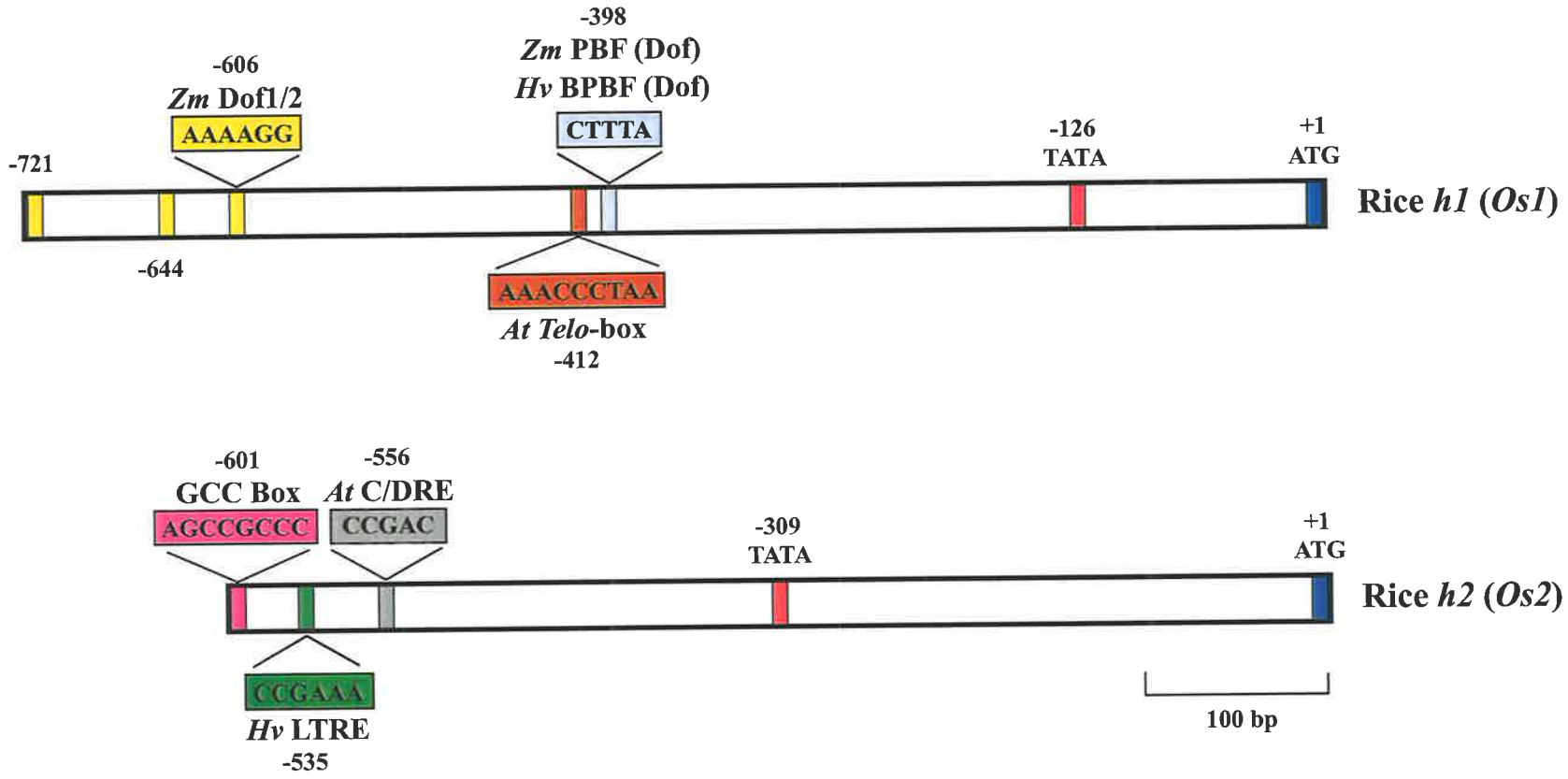


Figure 5.8 Schematic illustration of the sequence upstream of rice thioredoxins *h1 (Os1)* and *h2 (Os2)*.
Putative *cis*-elements are highlighted in colour with the corresponding transcription factor and binding-motif provided. Positions of TATA-box motifs are also indicated. Numbers indicate the nucleotide position of each element relative to the thioredoxin translation start codon.

5.4 Discussion

In the study reported in this chapter, the identification and molecular characterisation of grass thioredoxin *h* genes was undertaken, and their expression characteristics in response to abiotic stress examined.

The starting point for these analyses was the identification of all grass thioredoxin *h* sequences in the NCBI sequence database and the amplification of additional grass thioredoxins *n*. The collection and sequence analysis of grass thioredoxins *h* revealed that the *h*-class is composed of four distinct subclasses: two subclasses (subclasses 1 and 2) comprising homologues of previously published grass thioredoxins, one subclass (subclass *n*) representing the cDNAs cloned in this thesis whilst the fourth subclass (subclass 3) is a new group of grass thioredoxins *h*. It is likely that this is an accurate representation of the total number of thioredoxins *h* present in grasses since at the time database searches were conducted (August 2001) the total number of EST entries for several species (eg barley, wheat, maize and rice) was substantial (652,225 clones) and represented mRNAs from many tissues. The presence of distinct subclasses and the high level of sequence similarity among subclass members hints at a conservation of function for subclasses.

The genomic organization of each subclass was determined for wheat. From RFLP patterns of nulli-tetrasomics it is clear that each wheat subclass *h* gene is restricted to a single homeologous chromosome group. Although the total number of copies of each gene was not determined, the simple RFLP banding pattern detected in Southern blots suggests that genes are probably only present as a single copy on the individual chromosomes of each group. Comparative mapping and genetic synteny studies have demonstrated that gene content and order are highly conserved between different species of grasses (Devos and Gale, 1997; Gale and Devos, 1998). The chromosomal locations of wheat and rice thioredoxin *h* genes reported here, as determined by nulli-tetrasomic RFLP patterns and sequence analysis of PAC clones respectively, are consistent with the evidence derived from comparative mapping studies (Moore et al. 1995). There is significant conservation of linkage groups between wheat chromosome 1 and rice chromosome 5 indicating that these chromosomes are largely syntenous (Ahn et

al. 1993; Van Deynze et al. 1995). Correspondingly, the rice homologues (*Os2* and *Osn*) of wheat thioredoxins *h2* and *n* (present on chromosome group 1) are found on PACs representing chromosome 5. Similarly, conserved linkage groups between wheat chromosome 2 and 5 and rice chromosomes 7 and 3 respectively are supported by the data presented here.

In contrast to all other grasses, the screening of wheat EST libraries identified several, virtually identical, sequences in each subclass. Similar results are present in the literature with Serrato et al. (2001) and Gautier et al. (1998) cloning cDNAs of three almost identical (96-98.4% protein identity) subclass 2 thioredoxins from grain libraries of wheat. The nucleotide sequence of the 3 cDNAs was highly conserved within the coding region and only differed significantly in the 5' and 3' UTRs. Given that wheat is a hexaploid it is likely that these genes represent homeologues encoded by each genome. Such an explanation is also supported by the nulli-tetrasomic Southern results that found thioredoxin *h* genes are present on each homeologous member of a chromosome group.

The expression analyses of wheat thioredoxins *h* reported here revealed differences in the relative abundance of each transcript within the tissues examined. All thioredoxins were expressed in the four tissues examined with thioredoxin *h1* representing the most abundant message overall. Although endogenous expression levels were only determined for four tissues in this study, the compilation of EST data revealed that mRNA of thioredoxins representing each subclass have a broad spatial distribution including photosynthetic and non-photosynthetic tissues. Each thioredoxin in barley, the grass species with the greatest number of database entries, is represented by ESTs derived from no less than five distinct tissues (*h3*) and as many as nine (*h2*). Moreover, ESTs were identified in several libraries of the same tissue at different stages of development indicating that grass thioredoxins *h* also have a broad temporal expression profile.

The exposure of wheat plants to abiotic stresses revealed that the transcript levels of all thioredoxins were influenced by at least one form of stress. Thioredoxin *h1* was strongly induced in leaf tissue exposed to methyl viologen treatment, drought and UV exposure whilst *h2* expression was upregulated most dramatically by UV

exposure. Moreover, the induction of *h1* was rapid with a significant accumulation of transcript detectable within 2 hrs of UV stress and 6 hrs of methyl viologen application. The rapid induction of *h1* in response to methyl viologen is interesting given that methyl viologen affects the electron transport chain in chloroplasts yet *h1* is cytosolic. This result demonstrates that electron transport disruption can rapidly effect the transcription of genes whose products are not targeted to the chloroplast. In contrast to *h1*, the thioredoxin *n* expression response to many stresses was a rapid (1 hr in the case of H_2O_2 treatment) initial reduction in transcript level followed by a return to control levels. Such an expression profile might indicate a high sensitivity of the thioredoxin *n* transcriptional apparatus to oxidative stress. Conversely, the reduction in message levels could have been due to stress-induced mRNA degredation. However, following an initial reduction, thioredoxin *n* transcript can be seen to accumulate even when the plant tissue is still under stress (eg gradual accumulation in H_2O_2 treated tissue). An alternative explanation is that upon oxidative stress the thioredoxin *n* mRNA is rapidly translated to protein in order to protect other proteins from oxidation and it takes the cellular transcriptional apparatus time to restore *n* message quantities to endogenous levels. Future analyses will be required to test these hypotheses.

The large, and in some cases rapid, alteration of wheat thioredoxin *h* transcript levels in plants exposed to abiotic stress implies that *h*-class thioredoxins are an integral component of the stress-response mechanism of grasses. If indeed grass thioredoxins are involved in the stress response, the proteins encoded by these genes have the potential to act at the protective and/or signalling level. Since all of the stresses applied in this study result in oxidative stress, thioredoxin proteins may have an antioxidant function similar to that described for thioredoxins in other organisms (Chae et al. 1994; Kang et al. 1998). In this capacity thioredoxins could either directly reduce oxidised target proteins or act as hydrogen donors for other antioxidant proteins such as peroxiredoxin. Alternatively, thioredoxin proteins may function as messengers or regulatory proteins involved in signal transduction. A potential long-distance signalling function has been suggested for the rice *Os1* protein (Ishiwitari et al. 1995; 1998) since this protein is present in large quantities and moves with the phloem translocation stream. Such a suggestion is appealing in light of the observation that the wheat homologue (*h1*)

of rice *Os1* is strongly induced in response to drought stress. However, it is equally plausible that the wheat h1 proteins are involved in oxidative protection of sieve-element proteins and that the *h1* gene is induced for this reason. Grass thioredoxins could also be involved in the response to stress at the level of transcriptional modulation. In mammalian cells, cytosolic thioredoxin has been clearly shown to promote the DNA-binding of transcription factors such as AP-1 (Hirota et al. 1997), PEBP2 (Akamatsu et al. 1997), NF-κB (Hirota et al. 1999) and p53 (Ueno et al. 1999) whilst a redox-dependant mechanism has been implicated in the modulation of plant homeodomain transcription factors (Tron et al. 2002). In the case of NF-κB, oxidative stress (UVB) was found to promote the translocation of cytosolic thioredoxin into the nucleus where it activated the transcription factor. Presently, thioredoxins of subclass 2 are the grass thioredoxins most likely to interact with transcription factors since immunolocalization experiments have identified wheat subclass 2 proteins in the nucleus of scutellum and aleurone cells (Serrato et al. 2001). Interestingly, this study found there was a substantial induction of wheat h2 in leaves responding to UV stress. Although there is no evidence for a NF-κB homologue in cereals it is conceivable that h2 induced by UV irradiation is rapidly translated and transported into the nucleus where it interacts with a stress-responsive transcription factor. To assess such hypotheses, future transformation experiments employing sense thioredoxin h2 constructs linked to a reporter gene such as green fluorescent protein (GFP) could be used to track the intracellular movement of recombinant proteins in UV stressed leaf tissue. Moreover, if the thioredoxin construct employed an active site with the 2nd cysteine mutated to a serine (will form a stable mixed disulfide with interacting proteins: see Meyer et al. (1999) for a review of the method) it may be possible to isolate the proteins that interact with thioredoxin h2 *in-vivo* and determine whether any represent transcription factors. Such a strategy has been successful tested *in-vitro* with *E. coli* and human thioredoxins (Wynn et al. 1995; Qin et al. 1995) and *in-vivo* with yeast (Verdoucq et al. 1999). Clearly this will require further analysis.

Examination of the transcriptional response of wheat thioredoxins to abiotic stress revealed differences in the expression profile of each transcript. In addition, it was

observed that no thioredoxin was induced/repressed by all of the stresses applied, a result implying some level of specificity in transcriptional regulation. Stress-regulated gene expression in plants is typically modulated by stress-responsive transcription factors that bind to specific sequence elements in the gene promoter. Working on the assumption that thioredoxin homologues of each subclass are performing similar functions in different grasses, the transcriptional regulatory mechanism of thioredoxins in all grasses could be determined from promoter analyses of genes in one species. As a first step towards understanding how thioredoxins *h1* and *h2* are transcriptionally regulated, putative cis-acting regulatory elements were identified upstream of the rice subclass 1 (*Os1*) and subclass 2 (*Os2*) thioredoxins. Of the five elements located upstream of the *Os1* TATA-box, four represented known binding motifs of cereal 'Dof' transcription factors. Dof proteins are proteins found only in plants that contain a highly conserved single zinc-finger DNA binding domain. In maize Dof proteins have been shown to function as transcriptional activators or repressers of tissue-specific and light-regulated gene expression (Vincente-Carbajosa et al. 1997; Mena et al. 1998; Yanagisawa and Sheen, 1998). In addition, transformation experiments in tobacco have demonstrated that the tobacco Dof protein, NtBBF1, is involved in the vascular expression of the *rolB* oncogene (Baumann et al. 1999). The fifth sequence motif identified in *Os1* was a *telo*-box, a sequence motif found to be essential for the activation of expression of the *eEF1A* gene in root primordia of *Arabidopsis* (Tremousaygue et al. 1999). Given that *Os1* is a predominant protein of rice phloem sap (Ishiwitari et al, 1995) and wheat *h1* was upregulated in response to drought and UV irradiation, Dof proteins may also be involved in modulating the transcription of grass subclass 1 genes. *Os2* was found to contain three potential stress-associated sequence motifs, the core sequence of the *Arabidopsis* C/DRE element, an element essential for the low-temperature response of the barley *blt4.9* gene and a typical GCC-box. Transformation experiments in *Arabidopsis* have revealed that the C/DRE element is essential for transcriptional activation in response to stresses including cold, drought and salt stress (Yamaguchi-Shinozaki and Shinozaki, 1994). Significantly, recent research (Kim et al. 2002) demonstrated that light was required to activate cold-induced gene expression through the *C/DRE*. Although the response of wheat thioredoxins to cold stress was not investigated in this study, the involvement of light in stress-

induced gene expression is important given that wheat *h2* was most strongly upregulated following UV irradiation. The GCC-box represents a sequence motif shown to be targeted by ethylene-responsive transcription factors, a large multigene family with members in dicots and monocots (Fujimoto et al. 2000). Since many plant stresses, including those applied in this study, induce ethylene biosynthesis (Ecker 1995; O'Donnell et al. 1996; Penninckx et al. 1998) it is possible that wheat *h2* induction was also mediated by ethylene-responsive transcription factors acting through an upstream GCC-box.

Despite the detection of *cis*-elements in rice that may be recognised by transcription factors previously shown to regulate expression under oxidative stress, it is possible that this analysis failed to identify the true regulatory elements involved in thioredoxin transcriptional control. For example, transcription factors controlling thioredoxin gene expression may bind to *cis*-acting elements as yet undescribed in the literature. Alternatively, the sequence motifs recognised by such factors may be located further upstream than the 1.5 kb of sequence examined. Moreover, this analysis assumed that the function and expression profile of each thioredoxin subclass is identical irrespective of the species examined. This assumption has yet to be tested and requires the future determination of the stress-responsive expression pattern of all thioredoxins *h* in several grass species.

Phylogenetic analysis of grass thioredoxins *h* supports the view that the *h* class is composed of four distinct subclasses that evolved prior to the divergence of the *Pooideae, Oryzoideae* and *Panicoideae* subfamilies. Moreover, the grass *n* genes represent a subclass of plant thioredoxins *h* that clearly evolved prior to the divergence of monocotyledonous and dicotyledonous plants. Unlike the thioredoxin *n* subclass, a brief screening of plant sequence databases with the three other grass *h* subclasses failed to identify clear non-grass homologues. This result could be interpreted as indicating that the other grass subclasses evolved by duplication of an *n*-like ancestral sequence after the divergence of monocot from dicotyledonous plants. Alternatively, it is possible that the progenitors of these subclasses evolved prior to the monocot-dicot split and the sequences have diverged substantially since. To clarify which interpretation more closely reflects

the likely path of thioredoxin evolution requires a detailed phylogenetic analysis of all plant thioredoxins encompassing gymnosperms and angiosperms including both monocotyledonous and dicotyledonous representatives. This is the focus of the following chapter.

Chapter 6 Phylogenetic Analysis of the Thioredoxin *h*-class of Higher Plants

6.1 Introduction

Plants have the most complex thioredoxin profile of all organisms studied comprising five distinct and often multigenic classes. Of these classes the *h*-class is the largest, with as many as eight thioredoxins *h* having been identified in the genome of *Arabidopsis*. Results presented in Chapter 5 indicated that there is a level of subdivision in the *h*-class of grasses, with each thioredoxin *h* gene representing one of four phylogenetically distinct subclasses. In addition, results in Chapter 3 showed that in the case of subclass *n* the level of subdivision extends beyond the grass family (*Pooideae*) and probably encompasses all plants. The corollary is that at least one thioredoxin *h* subclass evolved early in plant evolution and has remained highly conserved since this event. The high sequence similarity of thioredoxins *n* in plant species as diverse as angiosperms and gymnosperms implies a conservation of gene function. Moreover, this observation raises the question of whether the subdivision of the *h*-class reported for grasses is representative of all plants.

Several phylogenetic analyses have examined relationships between different classes of plant thioredoxins and have uncovered evolutionary histories for most classes (Hartman et al. 1990; Sahrawy et al. 1996; Jacquot et al. 1997; Mestres-Ortega and Meyer, 1999); for example, whether the class has a prokaryotic or eukaryotic origin. However, to date no comprehensive study has examined the evolutionary relationships of thioredoxins within a class. Given that most thioredoxin classes are multigenic, such a study requires the use of a large number of sequences encompassing plant species from a diverse range of families. The recent increase of plant entries in EST sequence databases provided an opportunity to collect sufficient thioredoxin *h* sequences to perform the first detailed analysis of evolutionary relationships within the *h*-class. This study details the identification of all plant thioredoxins *h* present in the Genbank public database (as of January 2002) and an investigation of the evolutionary relationship of *h*-class thioredoxins in plants.

6.2 Materials and Methods

6.2.1 Identification of Thioredoxin *h* Sequences in the Genbank Database

The Genbank non-redundant and EST databases were screened for thioredoxin *h* sequences by BLAST searches using the published thioredoxin *h* cDNAs of *Arabidopsis* (Rivera-Madrid et al. 1995) as the initial query sequences. Screening of ESTs was performed systematically by, firstly identifying all plant species with entries listed in the Genbank EST database and secondly specifying BLAST-searches of individual species. All cDNAs containing a thioredoxin active-site motif were collected for each species. Thioredoxins m, f, o and x were identified on the basis of conserved residues (ie C-terminal cysteine of thioredoxins f) and motifs (N-terminal signal sequences of thioredoxins m, f and o) and higher sequence similarity to other plant thioredoxins of these classes. Thioredoxins representing these classes were removed and the remaining *h*-class thioredoxins used for subsequent analyses. In this way all plant thioredoxins present in the EST database at the time of searching were identified and those representing the *h*-class collected.

6.2.2 Sequence Alignment

The deduced protein sequences of all plant thioredoxins were aligned using the ClustalW software package employing the PAM 250 matrix. The alignment was then assessed by eye to ensure that the active site motif and structurally important residues were maintained at the same position in all sequences. Amino acid sequences were chosen for alignment since all thioredoxin share structurally and functionally conserved residues. These residue positions were therefore used as anchor points and ensured that intervening variant residues were aligned more accurately.

6.2.3 Assessing the Phylogenetic Signal in the Thioredoxin Data Set

To estimate the amount of phylogenetic signal in the plant thioredoxin *h* data, the Hillis and Huelsenbeck (1992) skewness test was implemented using the random trees option (20,000) implemented in PAUP. The test provides a measure of whether the data under analysis is or is not more structured than a random data set of the same number of sequences and characters. Since plant thioredoxins *h*

contain a mixture of highly conserved (eg the active site) and variable residue positions (eg the N-terminal region) phylogenetic relationships are naturally constructed on the basis of sequence similarity at variant positions. The test enabled the randomness of variant positions to be assessed. Skewness is assessed by examining the tree-length distribution resulting from the parsimony analysis of a random subset of trees (in this case 20,000). Normal (symmetrical) tree-length distributions indicate little or no phylogenetic signal is present (the variant positions are essentially randomised) whilst a strongly left-skewed distribution is an indicator of the presence of correlated characters (Hillis et al. 1993). Skewness of a data set is measured by the g_1 statistic in comparison with tables of critical g_1 values derived from random data sets (Hillis and Huelsenbeck, 1992). As the tables of critical values only extend to 25 taxa (or sequences), tree length distributions were determined by parsimony analysis of a subset of 25 thioredoxin sequences chosen at random.

6.2.4 Phylogenetic Analyses

Phylogenies were reconstructed using both distance-based (Neighbour Joining) and character-based (Parsimony) methods implemented in PAUP. Distance-based methods make pairwise comparisons of whole sequences, thus evolutionary divergence is determined on the basis of a single coefficient of sequence similarity/difference. By contrast, the character-based method of parsimony constructs phylogenies using sequence characters (amino acids), however, only relies upon "phylogenetically informative" amino acid positions. Both methods were chosen for this analysis to determine whether the phylogeny constructed was robust (ie was not compromised by the method employed). For a comprehensive comparison of different phylogenetic algorithms refer to Hillis et al. (1993).

Parsimony trees were found using the heuristic search algorithm, generating 100 replicates using random stepwise addition of sequences. Multiple equal-length parsimony trees were collapsed into 50% majority rule consensus trees. Bootstrap scores for all resolved nodes in the consensus tree were obtained from 100 bootstrapping permutations. The bootstrapping runs were produced using a heuristic search via random stepwise addition and under the tree-bisection-reconnection (TBR) algorithm for branch swapping.

Distance matrices were derived from the thioredoxin *h* sequence dataset under a mean character difference criterion. Amino acid substitutions were weighted in accordance with the BLOSUM substitution matrix. Trees were subsequently generated via the Neighbour-Joining algorithm as implemented in PAUP. Bootstrap values for resolved nodes were derived from 1000 runs.

6.3 Results

6.3.1 Identification of Plant Thioredoxin *h* Sequences

All plant entries (a total of 1,935,959) in the Genbank EST database were screened for thioredoxin *h* sequences. In addition to the grass sequences described in Chapter five, a further 569 plant ESTs were found to encode *h*-class thioredoxins. These represented 74 distinct thioredoxin *h* genes of which 62 were as yet undescribed in the literature. The genes were derived from a diverse range of plant species encompassing angiosperms, gymnosperms and bryophytes. A summary of all non-grass thioredoxin sequences identified during this study is given in Table 6.1.

Multiplicity of thioredoxin *h* genes was found to be a common feature of plants with as many as 11 distinct genes distinguishable in a single species (soybean). As was found for grass thioredoxins *h*, the mRNAs of *h*-class thioredoxins in other plant species had a wide spatial and temporal distribution. The predicted proteins of all plant thioredoxins *h*, including wheat, rice, maize and sorghum *h*-subclass representatives, were aligned by the ClustalW alignment program (Appendix H). Thioredoxin *h* protein sequences displayed between 30% and 98% amino acid identity to each another with most variation found in the N and C-terminal regions of aligned sequences. Excluding the active site, 24 residue positions were found to be either invariant or highly conserved across all sequence (Figure 6.1). 17 positions contained residues reported previously as being essential for the maintenance of thioredoxin structure and/or active site conformation (Ecklund et al. 1991). The remaining residues were a tryptophan, threonine and serine found 22, 3 and 1 amino acids upstream of the active site respectively and a phenylalanine, tryptophan, phenylalanine and lysine located 19, 34, 42 and 59 amino acids downstream of the active site. Analysis of prokaryote and other

eukaryotic thioredoxins revealed that the conservation of chemically similar residues at these seven positions was not unique to plants. Vertebrate cytosolic-thioredoxin sequences were found to have identical or chemically similar residues at each position, with the exception of Trp34 where vertebrates contain a cysteine. Therefore, several of these conserved amino acids may be representative of higher eukaryotes in general. The variant active site consensus WCPPC was found to be uncommon in plants with only three proteins from *Populus balsamifera* (Pb1), *Thellungiella salsuginea* (Ts1) and a hybrid poplar (Ph1) containing this form.

6.3.2 Assessing the Phylogenetic Signal in the Thioredoxin Data Set

To determine whether plant thioredoxins contained sufficient correlated characters at variable residue positions to proceed with phylogenetic analyses, the Hillis and Huelsenbeck skewness test (6.2.3) was performed on a random subset (25) of sequences. The distribution of lengths for the 20,000 trees evaluated was strongly skewed to the left (Figure 6.2) with a g_1 value of –0.839 compared to the critical value of –0.12 ($P < 0.01$) for 25 taxa and 100 characters. This value indicates that the thioredoxin data is significantly more structured than a random data set of identical sequence number and similar character size and implies the presence of strong phylogenetic signal.

6.3.3 Phylogenetic Analysis of Plant Thioredoxins *h*

Phylogenetic analyses were performed on aligned amino acid sequences using both distance (neighbour joining) and character (parsimony) based algorithms. The resultant phylogenies constructed with both methods were found to be virtually identical (Figures 6.3 and 6.4) differing only in the placement of bryophyte (mosses and liverwort) and fern sequences and the levels of bootstrap support for major nodes. Thus the method of tree construction had little effect on the plant thioredoxin *h* phylogeny. In the neighbour joining (NJ) tree bryophytes and ferns were grouped into a single, although weakly supported (56% of bootstrap replicates), monophyletic clade that also contained all *n*-type thioredoxins. Conversely, parsimony (PAR) analysis placed bryophytes, ferns and thioredoxins *n* into three separate clades with higher bootstrap support of 84%, 100% and 93% respectively. With the exception of thioredoxin *n* and fern

sequences, all other vascular plant thioredoxins were grouped into three distinct monophyletic clades comprising only angiosperm or gymnosperm representatives.

In both analyses the thioredoxin *n* clade was well supported among bootstrap replicates (NJ 84% and PAR 93%) and represented the only major grouping to contain both angiosperm and gymnosperm sequences. The result implies that the *n*-subclass is of ancient origin and predates the evolution of the angiosperms at least 200 million years ago (Bold 1977). This conclusion is also supported by the observation that in both analyses thioredoxins *n* are the *h*-class sequences that share greatest similarity with ancient plant forms (bryophytes and ferns). By contrast, no clear gymnosperm homologues can be found for angiosperm groups 1 and 2 nor is there an angiosperm representative for the gymnosperm clade. There are however sequences from the same species, including monocots and dicots, in both groups 1 and 2. These results indicate that the evolution of groups 1 and 2 predates the divergence of monocots and dicots approximately 180-220 million years ago (Soltis et al. 1999; Yang et al. 1999).

The angiosperm group 1 clade contained the shortest thioredoxin *h* sequences including both grass subclasses 1 and 2. The two plant thioredoxins isolated from sieve tubes, Os1 (Ishiwitari et al. 1995) and Rc1 (Schobert et al. 1998), are represented in this group, as are very similar proteins from 25 other species. 31 members of this group, encompassing 23 species, have N-terminal amino acids virtually identical to those identified in rice Os1 as being essential for the protein to mediate its own transport across plasmodesmata (see 5.3.1.2). Therefore it is probable that this clade contains all angiosperm sieve-element thioredoxins.

Group 1 contains multiple genes for most species indicating widespread gene duplication. However, many genes do not form associations with other thioredoxins from the same or closely related species. For example Rc1, Ph1 and Pb1 are all representatives of the same family yet are not grouped in the tree. Similarly, Mt3, Gm5/8 and Gm 4/7 are not grouped with other legume sequences. The uncertain placement of many group 1 sequences reflects the similarity of sequence distance of these proteins and indicates a high level of variability in their cladal location in bootstrap replicates. Furthermore, the similarity of sequence

variation suggests that these genes evolved prior to the divergence of their respective plant families and that they have not been under strong sequence constraint since. By contrast, some sequences displayed a higher level of correlated characteristics and formed subclades consistent with known familial relationships. For example, Mt4, Lj1, Sros1 and Gm3/9 are from species of the same legume subfamily. Such associations were also found for grass sequences and certain solanaceous sequences (St2, Le2, Nt1 and Ib1). This topology could indicate that these group 1 genes evolved later by duplication after the divergence of specific family lineages.

Unlike the group1 clade, the relationship of group 2 sequences was more clearly defined in both the neighbour-joining and parsimony topologies. In the parsimony tree, group 2 sequences were divided into monocot (grass subclass 3 sequences) and dicot sister clades with 100% and 59% bootstrap support respectively. The dicot clade comprised three distinct subclades supported in 63%, 52% and 100% of bootstrap replicates. The neighbour-joining analysis identified group 2 subclades virtually identical to those present in the parsimony tree, however was unable to resolve the topology of these clades. The separation of monocot and dicot sequences and the presence of distinct dicot subclades in angiosperm group 2 indicates that dicot subclades arose by duplication after the monocot-dicot split. In addition, since two dicot subclades contain sequences of identical species representing the *solanaceae* and *fabaceae* families, the duplication that gave rise to these subclades must predate the divergence of the *solanaceae* and *fabaceae.*

A common feature of angiosperm group 2 thioredoxins was the presence of a long N-terminus. This group included the published Gm1 and Gm2 proteins of soybean (Shi and Bhattacharyya, 1996). Gm1/2 were reportedly plasma membrane-bound proteins, anchored to the membrane through 16 predominantly hydrophobic N-terminal amino acids. All sequences represented in the same dicot subclade as Gm1 and Gm2 were found to have an N-terminal sequence very similar to the proposed transmembrane region of the soybean proteins. Therefore it possible that plasma membrane bound homologues from other angiosperm species are represented in the group 2 clade.

Table 6.1 Summary of the database search for *h* class thioredoxin in other plant (non-grass) species.

[a] The GeneBank accession number of a representative EST clone encoding each thioredoxin *h* subclass sequence is provided for each species. In the case of thioredoxin *h* sequences that have been published the reference is given.

[b] A summary of the tissues in which each thioredoxin mRNA is present (derived from EST database sequences and published Northern data). The tissues are abbreviated and are: B bark; Bf boll fibres; C carpel; Ca callus; Cb cambia; Cc cell culture; Cg crown gall; Ch chloronemata; Co cotyledon; Cu caulonemata; E epicotyl; Em embryo; F fruit; Fb flower bud; Fl flower; Fm floral meristem; G gametophore; H hypocotyl; K kernel; L leaf; M mycorrhiza; Mb malformed bud; P protoplast; Pc pollen cone; Pe pericarp; Pr protonemata; R root; Ra radicle; Rn root nodule; S seedling; Sc suspension culture; Sd seed; Sdc seed coat; Se somatic embryo; Sh shoot; So sexual organs; Sp spore; T tuber; Tr trichome; Wp whole plant; X xylem. The presence of (vs) after the tissue abbreviation indicates that the mRNA has been extracted from various developmental stages of that particular tissue. Tissue subjected to a form of abiotic stress is indicated by (as) after the tissue abbreviation whilst (bs) indicates biotic stress.

Table 6.1

Family	Species	Common name	Identifier	Number of database entries	Source[a]	Tissues[b]
Brassicaceae	*A. thaliana*	Arabidopsis	At1 - At5	47	Rivera-Madrid et al. 1995	Sc, Ca, Wp, R, L, F, Fb
Brassicaceae	*A. thaliana*	Arabidopsis	Atn	1	AC012562	ND
Brassicaceae	*B. napus*	Oilseed rape	Bn1, Bn2	4	Bower et al. 1996	
Brassicaceae	*T. salsuginea*	Thellungiella	Ts1	11	BI698880	S
Fabaceae	*G. max*	Soybean	Gm1	16	Shi and Bhattacharyya, 1996	Fl, H, Co(a/s), R, L,
			Gm2	16	Shi and Bhattacharyya, 1996	H, Se, S, R, L(v/s), E
			Gm3	21	BQ627739	R(v/s), Se, Fm, S, L(a/s)(v/s), Sdc, Sh, F
			Gm4	29	BE807885	Fl(v/s), H, L(v/s), Sh, S(a/s), Sd, Sdc, Sh, R, Se, Fb
			Gm5	2	AW569018	S, Rn
			Gm6	3	BF066575	Sh, H, Co
			Gm7	10	AI441505	R, L(v/s), S
			Gm8	9	BQ627882	S, R, H, Fm, Sp, Sc, Fl
			Gm9	4	BI967602	R, Sh
			Gm10	13	BG652681	Fm, Se, R, H, Co(a/s), Fl, S, L
			Gmn	7	BQ298430	S(a/s), R, H, L(v/s), Sd
Fabaceae	*L. japonica*	Lotus	Lj1	3	BI421029	Rn
Fabaceae	*M. truncatula*	Barrel medic	Mt1	2	AW560796	R, L
			Mt2	22	BE943230	R(v/s), L, M, Cc
			Mt3	11	AL386021	M, R, L, Co
			Mt4	44	BE320702	R, R(a/s), M, L(a/s)(b/s), Rn, Sd, Fl, Cc, S
			Mt5	2	AW686237	R, Rn
			Mtn	7	BG647951	R, M, Rn, Fl
Fabaceae	*S. rostrata*	Tropical legume	Sros1	1	AJ301738	R
Salicaceae	*P. balsamifera*	Poplar	Pb1	2	BI136756	Fl
Salicaceae	*Populus* hybrid	Poplar	Ph1	17	AI161830	Cb, L
Euphobiaceae	*R. communis*	Castor bean	Rc1	1	Z70677	Co

Table 6.1 continued

Family	Species	Common name	Identifier	Number of database entries	Source[a]	Tissues[b]
Malvaceae	*G. arboreum*	Cotton	Ga1	8	BG446507	Bf
			Ga2	2	BE053246	Bf
			Ga3	1	BE053835	Bf
			Ga4	1	BQ408049	Bf
			Ga5	2	BQ412293	Bf
Vitaceae	*V. vinifera*	Wine grape	Vv1	8	BQ800442	F, L(a/s)
Rosaceae	*P. persica*	Peach	Pper1	2	BU044549	F
Rutaceae	*C. sinensis*	Orange	Cs1	2	BQ625213	S
Chenopodiaceae	*B. vulgaris*	Sugar beet	Bv1	1	BQ489103	Sh, R
Polygonaceae	*F. esculentum*	Buckwheat	Fe1	1	D87984	Sd
Aizoaceae	*M. crystallinum*	Iceplant	Mc1	6	BG269748	L
			Mc2	2	BM301185	L
			Mc3	1	BG269520	L
			Mc4	2	BE034818	Fl, Sp, R
Asteraceae	*H. annus*	Sunflower	Ha1	8	BQ913020	P, Sh(a/s), K
Asteraceae	*S. rebaudiana*	Stevia	Sr1	7	BG524210	L
			Sr2	1	BG522391	L
Lamiaceae	*M. piperita*	Peppermint	Mp1	2	AW255457	Tr
Convolvulaceae	*I.batatas*	Sweet potato	Ib1	1	BM878771	Wp

Table 6.1 continued

Family	Species	Common name	Identifier	Number of database entries	Source[a]	Tissues[b]
Solanaceae	*L. esculentum*	Tomato	Le1	8	AW094467	L, Ra
			Le2	49	AI776110	L, Sc, R, Ra, F, Sh, S, Cg, F(v/s)
			Le3	11	BI208855	Sc, Fb, L
			Le4	5	AW626141	Ra, L, C, Cg, Ca
			Le5	2	AW036157	Sd, Sh
			Len	12	AW092362	L, Pe, C, Sc, Ca
Solanaceae	*N. tabacum*	Tobacco	Nt1, Nt2	ND	Brugidou et al. 1993	
Solanaceae	*S. tuberosum*	Potato	St1	5	BM111010	R, L(b/s)
			St2	4	BQ507668	T
			St3	2	BI434632	L(b/s)
			St4	2	BQ509563	T
			Stn	5	BG888653	T, L(b/s)
Cupressaceae	*C. japonica*	Japanese cedar	Cj1	2	AU084606	B
Pinaceae	*P. mariana*	Black spruce	Pm1	1	AF051206	Em
Pinaceae	*P. taeda*	Loblolly pine	Pt1	43	BQ290672	X, Sh, Pc,
			Pt2	32	AW043379	X, Sh, Pc,
			Ptn	9	AW011514	X, Sh, Pc,
Pteridaceae	*C. richardii*	Fern	Cr1	1	BE642843	Sp
			Cr2	2	BE642761	Sp
Marchantiaceae	*M. polymorpha*	Liverwort	Liv1	1	AU081772	So
Funariaceae	*P. patens*	Moss	Pp1	12	BJ187518	Ch, Cu, Pr, G, Mb
			Pp2	8	BQ041803	Ch, Cu, Pr, Mb

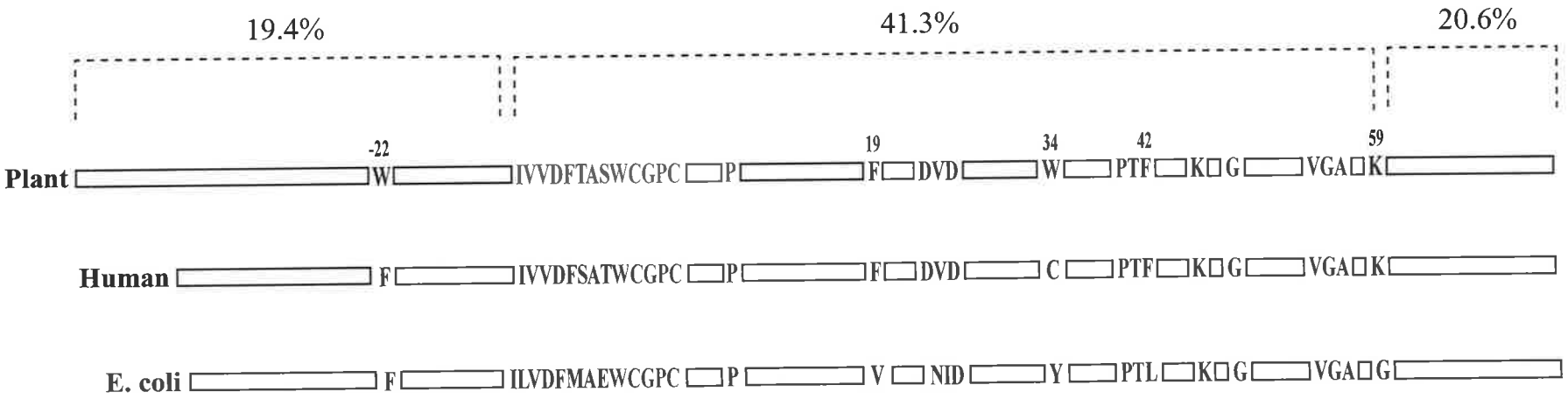


Figure 6.1 Schematic illustration of the position of chemically conserved residues in plant thioredoxins *h*.
The positions of conserved amino acids are indicated with the consensus residue. The active site amino acids are in red whilst structurally important residues are in blue. Residues in black are those discussed in the text and their position relative to the active centre motif is given above. Bars represent the intervening chemically variant amino acids. The average amino acid identity of the N-terminal, C-terminal and central regions of plant thioredoxins *h* is indicated above each region. The residues of human and *E. coli* cytosolic thioredoxins at each of the plant positions are provided for comparison.

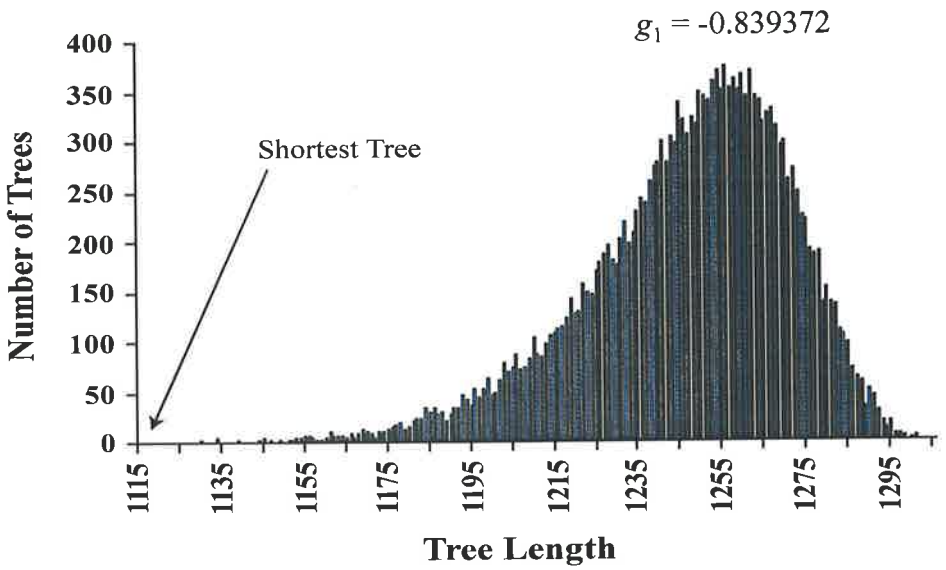


Figure 6.2 Assessing phylogenetic signal from the distribution of tree-lengths.

The barchart illustrates the distribution of tree lengths among 20,000 randomly derived trees. Trees were generated by parsimony analysis of 25 plant *h*-class thioredoxins chosen at random from the 95 sequences used in later phylogenetic analyses. The g_1 statistic and the position of the shortest (most parsimonous) tree are indicated.

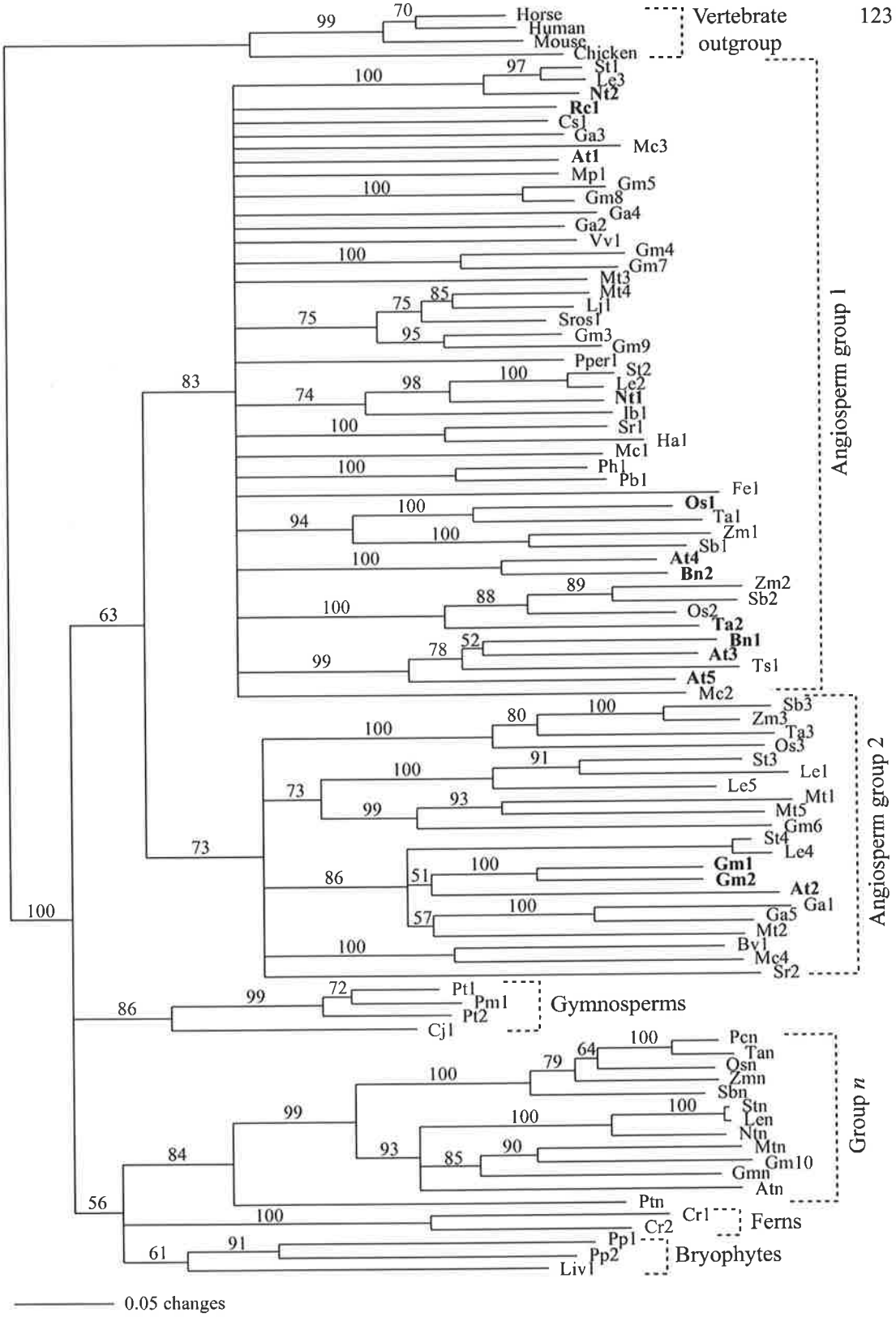


Figure 6.3 Phylogeny of plant thioredoxins *h* determined by Neighbour-Joining. Phylogeny of all plant thioredoxin *h* protein sequences determined with the neighbour-joining algorithm implemented in PAUP (Swafford 1998). Distances were calculated as mean percentage amino acid difference. Numbers indicate the percentage of 1000 bootstrap permutations in which a group was found (numbers <50% not shown). Species are identified by their initials (refer to Table 6.1) and initials of published sequences are highlighted in bold. Major clades discussed in the text are indicated. The vertebrate thioredoxin h sequences of horse, human, mouse and chicken were specified as outgroup sequences (see Figure 5.3 for accession numbers).

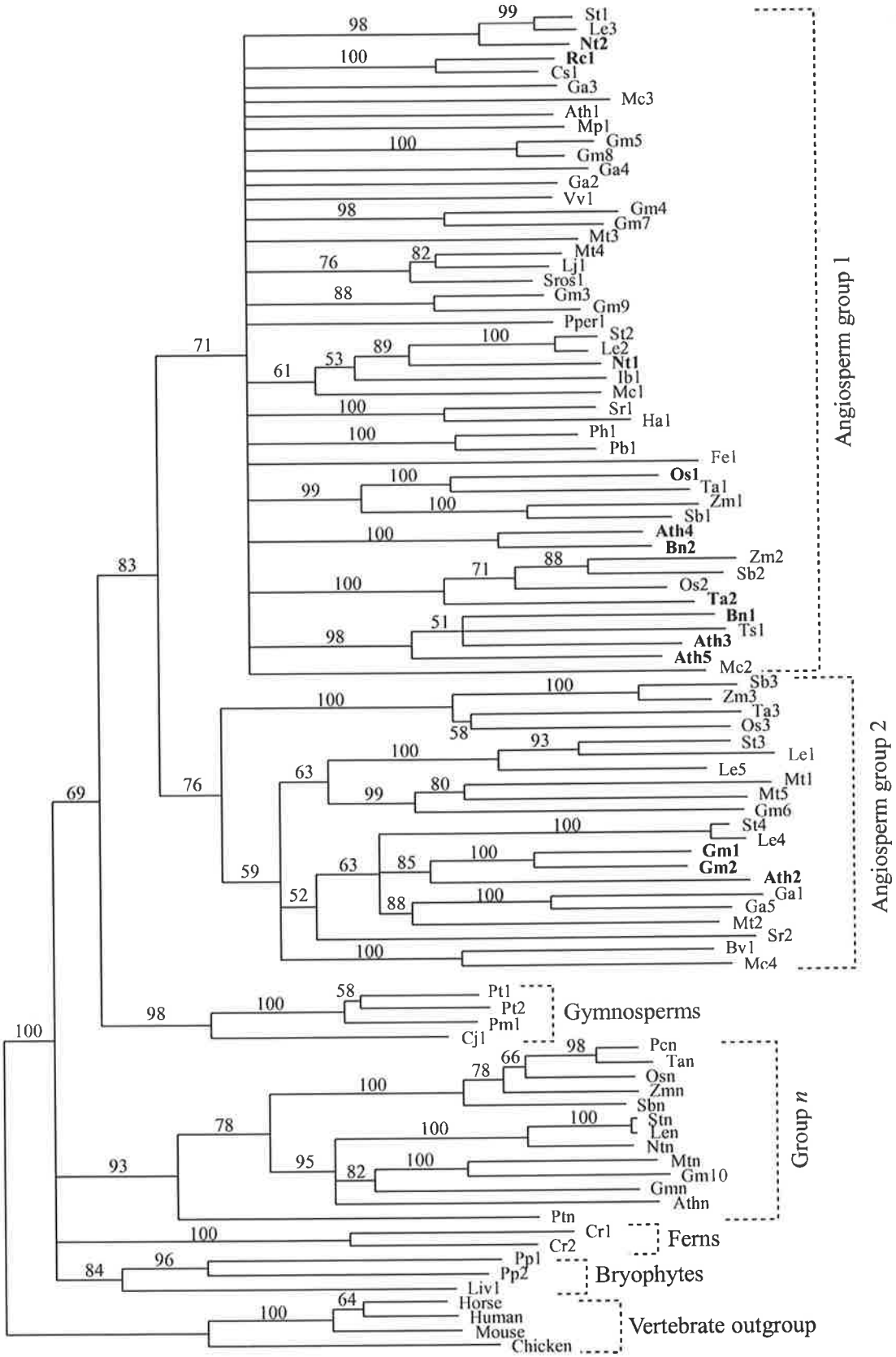


Figure 6.4 Phylogeny of plant thioredoxins *h* determined by Parsimony.
A consensus (50% majority rule) of the most parsimonous trees from the analysis of all plant thioredoxin *h* protein sequences using PAUP (Swafford 1998). Numbers indicate the percentage of 100 bootstrap permutations in which a group was found (numbers <50% not shown). Species are identified by their initials (refer to Table 6.1) and initials of published sequences are highlighted in bold. Major clades discussed in the text are indicated. The vertebrate thioredoxin h sequences of horse, human, mouse and chicken were specified as outgroup sequences (see Figure 5.3 for accession numbers).

6.4 Discussion

In this chapter the phylogeny of plant *h*-class thioredoxins was investigated. The analysis commenced with the identification of all plant thioredoxin *h* sequences in the Genbank EST sequence database and concluded with the finding that the *h*-class of flowering plants is composed of three distinct and monophyletic subclasses. The presence of genes representing all subclasses in such a diverse range of species implies that the subclasses arose early in angiosperm evolution and raises the question of the mechanisms that may have generated the number of thioredoxins *h* present in modern plants. Genomic studies in plant species such as *Arabidopsis* and soybean provide some insight into the possible mechanisms that may have produced multiple thioredoxins *h*. In both species, large segmental duplications cover much of their genomes and have been shown to be due to a combination of whole genome duplication followed by diploidisation and smaller localised duplication events (Blanc et al. 2000; Shoemaker et al. 1996). In *Arabidopsis* at least four major duplication and whole genome polyploidisation events are estimated to have occurred between 100 to 200 million years ago (Vision et al. 2000). Evidence in support of ancient duplication events has come from comparative mapping studies that have detected conserved linkage blocks between tomato and *Arabidopsis* (Ku et al. 2000) and between *Arabidopsis* and soybean (Grant et al. 2000; Lee et al. 2001), species whose ancestors are estimated to have diverged approximately 90 million (Gandolfo et al. 1998) and 112 million (Yang et al. 1999) years ago, respectively. Since several duplication events are predicted to predate important angiosperm evolutionary landmarks, including the divergence of monocots and dicots (Soltis et al. 1999) and that of major family lineages, they are likely to have contributed to the generation of multiple thioredoxins *h* found in all modern angiosperms.

In addition to the aforementioned ancestral duplication events, the genome of the soybean ancestor appears to have undergone a more recent polyploidisation event followed by reversion to a diploid state. Evidence for this comes from observation that linkage groups between legumes are represented twice in soybean (Menancio-Hautea et al. 1993; Boutin et al. 1995; Lee et al. 2001) and several multigene families in soybean contain 2 distinct subgroups of closely related genes (Lee and Verma 1984; Neilsen et al. 1989; Yamamoto and Knap, 2001). The results

presented here also support the proposition of a polyploidisation event in soybean. Soybean EST libraries were found to encode at least 11 distinct thioredoxin *h* genes, almost twice as many as were identified in libraries of *Medicago truncatula*. Furthermore, the thioredoxin *h* multigene family in soybean can be seen to contain highly related sequence pairs (Gm1/Gm2, Gm3/Gm9, Gm4/Gm7, Gm5/Gm8 and Gmn/Gm10). Phylogenetic analysis demonstrated that each pair of thioredoxins except Gmn/Gm10 were more similar to one another than to thioredoxins derived from the related legumes *Medicago truncatula* (Mt), *Lotus japonica* (Lj) and *Sesbania rostrata* (Sros). Since all of these legumes are members of the same subfamily (the Papilionoideae) the implication is that the genome polyploidisation event occurred after the soybean linage diverged from other tribes of the subfamily.

Thioredoxins *n* appear to represent the most ancient subclass of *h* thioredoxins since homologous *n* genes are clearly distinguishable in angiosperms and gymnosperms. Thioredoxin *n* genes from species as diverse as pine and wheat are more similar to one other than to other pine or wheat thioredoxin *h* sequences. This indicates that thioredoxins *n* evolved early in terrestrial plant evolution and that sequences changes have been largely constrained over a long period of time. One possible explanation for the conservation of sequence identity is stringent functional constraint at the protein level. If indeed this is the case, the function must be one common to higher plants that has yet to be identified. All thioredoxins *n* contain several conserved residues that distinguish them from other plant *h*-class thioredoxins (see 3.4). Such amino acids could be essential for the functional specificity of thioredoxins *n* or the interaction with target proteins and as such represent potential targets for future mutational and functional studies.

The absence of gymnosperm representatives in angiosperm Groups 1 and 2 raises the question of when the progenitor genes of these clades evolved. During the evolution of the angiosperms would infer that there was only a single *n*-like thioredoxin gene present in both ancestral gymnosperms and angiosperms and that multiple sequences in each arose from duplications and divergence after the split. However, such a model is not supported by the observation that modern relatives of ancient plants (eg *Ceratopteris richardii* (Cr1/2) and *Physcomitrella patens*

(Pp1/2)) also contain multiple *h*-class thioredoxins. Moreover the green algae *Chlamydomonas*, the closest living relative to the progenitor of all terrestrial plants, contains two cytosolic (*h*-class) thioredoxins (Stein et al. 1995). Thus ancient vascular plants are likely to have contained multiple thioredoxins *h*. An alternative explanation is that the sequences of the gymnosperm group and angiosperm Groups 1 and 2 share a common ancestry yet have diverged substantially since. To determine whether this hypothesis is probable would require the inclusion of a greater number of gymnosperm sequences than were used in this study.

The topology of angiosperm Groups 1 and 2 and the low bootstrap scores for some subclades within these groups reflected the uncertain placement of sequences during bootstrap replicates. Although a large number of thioredoxins were used in this analysis, these sequences were derived primarily from the dicot families *Brassicaceae, Fabaceae, Malvaceae, Aizoaceae* and *Solanaceae*. With the exception of the grasses (*Poaceae*) no other monocotyledon thioredoxins were available on the public database. The lack of sequence representatives from other monocot families undoubtedly contributed to the inability to determine the relationships between monocot and dicot sequences in angiosperm group 1. Improving the resolution of cladal relationships between monocots and dicots, will require the future isolation and utilisation of Group 1 and 2 thioredoxins from other monocot species such as orchids, lilies, bromiliads and palms. Such an analysis would also provide an insight into the total number of *h*-class thioredoxins likely to have been in existence at different stages of angiosperm evolution and the chronology of gene duplication events.

In the phylogenies constructed using both the neighbour joining and parsimony methods the topology of the Group 1 clade was the most poorly defined reflecting the similar level of amino acid difference between members of this clade. This finding suggests that the function of these proteins does not place a strong constraint upon sequence evolution and may indicate a generalised role for members of Group1 subclass. Such a proposition is supported by the observation that wheat thioredoxin *h1* expression was induced in response to a range of oxidative stresses. Given that rice (*Os1*) and *Ricinus* (*Rc1*) homologues of wheat

h1 are abundant proteins in phloem sap (Ishiwitari et al. 1995; Schobert et al. 1998) it is possible that these proteins have a broad antioxidant function in the plant translocation system. Should group 1 thioredoxins have a generalised antioxidant role with many target proteins it may also explain the apparent redundancy of the Group. However, it must be noted that not all group 1 sequences fit this explanation. Wheat *h2* mRNAs are not upregulated in response to some of the same oxidative stresses, nor have their proteins been isolated from phloem exudate. Since most dicot Group 1 sequences resemble grass subclass 1 genes it is plausible that grass subclass 2 thioredoxins evolved by duplication from subclass 1 genes after the monocot-dicot split. If such a hypothesis were true there should be subclass 2 homologues in the genomes of non-grass monocot species. Hence it is essential that future investigations designed to resolve topological relationships in angiosperm Group 1 include Group 1 genes from non-grass monocots.

The compilation of all ESTs encoding plant thioredoxins *h* provides valuable data on the spatial and temporal expression profiles of thioredoxin mRNAs. In species with a large number of database entries, thioredoxins *h* were identified in many tissues and at a broad range of developmental stages suggesting that thioredoxins *h* are expressed in most plant tissues. Typically, the greatest number of ESTs, of a particular gene, came from cDNA libraries of young or developing tissue (eg shoot tips). This finding is consistent with Northern hybridisation results that found the expression levels of *h*-class thioredoxins in *Arabidopsis* and tobacco were highest in young or rapidly growing tissue (Marty et al. 1993; Rivera-Madrid et al. 1995). The increased expression of thioredoxin *h* mRNA in rapidly growing cells may be related to changes in the redox environment of cells that occur during the cell cycle since studies in mammalian cells have found that the protein level of the major cellular redox compound, glutathione, is highest during the exponential growth phase (Post et al. 1983; Atzori et al. 1990). EST analyses also revealed that thioredoxins representing each of the three thioredoxin *h* subclasses had overlapping spatial and temporal expression patterns. The presence of multiple thioredoxins expressed in the same tissue raises the question of functional redundancy. A possible explanation may lie in the subcellular location of proteins of each subclass. Although the intracellular and intercellular distribution of all

thioredoxin h proteins from a single plant species has yet to be determined, representatives from Groups 1 and 2 have been identified in different locations. For example, Os1 and Rc1 proteins (Group 1) are present in high concentrations in the phloem sap, wheat h2 (Group1) has been localised to the nucleus of scutellum and aleurone cells and soybean thioredoxins Gm1 and Gm2 (Group 2) are plasma membrane bound. If members of each thioredoxin Group are performing functions in different locations of a plant it may explain the high number of apparently similar thioredoxins *h* present in angiosperms. Alternatively, the functions of thioredoxins in plants may overlap with one another and with related redox systems such as the glutathione-glutaredoxin system as has been described for yeast (Draculic et al. 2000). Cumulative gene knockout experiments could provide insights into the minimum thioredoxin requirements of plant cells and the degree of redundancy within and between thioredoxin *h* subclasses.

Chapter 7 General Discussion

Most plant nuclear genes are members of multigene families; multiple genes that have a common evolutionary origin and encode proteins of related function (Zhang et al. 2000). These families are a significant component of plant genomes and vary in size from a few to several hundred genes (Clegg et al. 1997; Meyers et al. 1999). Despite the importance of plant gene families, few studies have comprehensively addressed the evolutionary genomic processes that influence gene family structure and dynamics. An understanding of these processes is important in the context of identifying mechanisms of gene family evolution and essential to an understanding of mechanisms that drive molecular adaptation and functional divergence in plants. In this thesis the dynamics of plant gene family evolution and functional relationships were investigated using thioredoxins *h* as a model system within the phylogenetically well-characterised framework of the grasses and other plants.

The starting point for this analysis was *Bm2*, the putative pollen *S*-gene of *Phalaris coerulescens* (Li et al. 1994). *Bm2* provided a unique opportunity to study the relationship between gene function, genomic position and gene evolution since information relating to each of these components already existed. Firstly, the function of *Bm2* was thought to be known (the male SI determinant). Secondly, there existed data from other SI systems on the evolutionary forces at work in the region of SI loci. Therefore the effect of genomic position on the evolution of *Bm2* in the grasses could be addressed and compared with other SI loci. Thirdly, initial studies on *Bm2* had reported the gene to encode a highly conserved thioredoxin *h*-like domain (Li et al. 1995). Thus, the evolution of a thioredoxin-like protein under the positional and functional constraints of an SI system could be compared with that of other plant thioredoxins *h*. Early in these analyses it became clear that *Bm2* did not represent the male determinant of *Phalaris* SI but rather encoded an active thioredoxin (termed thioredoxin *n*) belonging to a highly conserved subclass of plant thioredoxins *h* (Baumann et al. 2000; Juttner et al. 2000). Consequently, the focus of this thesis shifted to an

evolutionary and functional investigation of the thioredoxin *h* multigene family in plants. Whilst *Bm2* does not represent a SI recognition molecule, the results presented here do not discount the involvement of *Bm2* in the SI response *per se*. Indeed, redox mechanisms appear to play a role in at least two SI systems (Tezuka et al. 1997; Bower et al. 1998). Moreover, thioredoxins have been reported to act as negative regulators of SI (Cabrillac et al. 2001) and apoptosis (Saitoh et al. 1998; Hashimoto et al. 1999) signal-regulating kinases in plants and mammals, respectively. The involvement of thioredoxins in signal transduction pathways leading to cell death in such divergent organisms raises the question of whether this is a function common to thioredoxins in other eukaryotes including self-incompatible grasses.

The change in scale of study from a single locus within one family of plants (the grasses) to an entire gene family across a broad phylogenetic spectrum (angiosperms) presented several problems. Primarily, the number of thioredoxin *h* genes published in the literature at the commencement of this project (14) was insufficient to accurately resolve phylogenetic relationships in a multigenic class such as the *h*-class. As a result the new electronic resources of EST databases and whole-genome sequences were used as a source of additional thioredoxin *h* genes. This approach culminated in the screening of greater than 1.9 million ESTs and the identification of 86 new plant thioredoxins *h*. Sequence assembly and phylogenetic analyses of these genes provided the first detailed profile of the copy number complexity and evolutionary dynamics of the thioredoxin *h*-class in plants. The central finding of this work was the identification of three primary clades (n, Group 1 and Group 2) in the thioredoxin *h* class of flowering plants. Sequences from a diverse range of angiosperm families were located to each clade indicating the clades are ancient and arose either prior to, or early in, the evolution of flowering plants. The presence of these clades in all flowering plants studied to date may indicate a similar function for representatives from each clade. Determining the intracellular location of proteins from each clade will be the first logical step in addressing such a question and remains an important objective for future investigations.

Of all thioredoxin *h* clades the n clade was found to be particularly interesting since it was highly conserved and was the only clade to contain a gymnosperm representative. The implication of this finding is that thioredoxin *n* genes evolved very early in plant evolution but were then conserved. A possible explanation for such sequence conservation is a common function for thioredoxin n proteins. If this is true, then thioredoxin n proteins must be involved in a reaction or series of related reactions that are yet to be defined. In this analysis, thioredoxin *n* representatives were not detected in ESTs of mosses, liverworts or ferns despite the screening of as many as 67,738 ESTs from the moss *Physcomitrella patens*. It is possible that plants representing these divisions do encode thioredoxin *n* homologues, however the tissue(s) in which the gene is transcribed were not used for cDNA construction. Alternatively, this result may indicate that the function of thioredoxins n is not one common to bryophytes or ferns. One significant difference between bryophytes, ferns and the pollen-producing plants examined here is their method of fertilisation. The sperm of ferns and bryophytes actively swim from the organ of production (the antheridium) to the egg (contained within the archegonium) to effect fertilisation (Richardson 1981; Bell and Woodcock, 1983). By contrast, the sperm of gymnosperms and angiosperms are carried to the egg by the developing pollen tube. Considering that many liverwort, moss and fern ESTs in the database were derived from reproductive tissue (and contained no thioredoxin *n* gene), the reproductive processes of gymnosperms and angiosperms may represent a good focal point for functional analyses of thioredoxins n. Alternatively, thioredoxin *n* sequence conservation may be related to the genomic location of thioredoxin *n* genes. In self-incompatible grass species thioredoxins *n* are closely linked to the *S*-locus (Baumann et al. 2000), regions of low recombination (Matton et al. 1997; Ishimizu et al. 1998). In addition, mapping studies in self-compatible species such as wheat and barley have detected very low levels of polymorphism in the region of the thioredoxin *n* locus (A. Hay pers. comm. 2000). Together these observations suggest that the thioredoxin *n* locus in grasses may be subject to different intragenic evolutionary pressures than other regions of the genome. This raises the question of whether thioredoxin *n* loci in other plant families are similarly associated with regions of low recombination. Determining the placement of thioredoxins *n* in the genomes of other well studied

species such as *Arabidopsis* and soybean will be an important step in addressing this question.

In contrast to the situation in animals, few of the putative functions ascribed to thioredoxins in plants have been verified. This is particularly evident within the *h* class, where no single biological function has been unambiguously attributed to a thioredoxin *h*. The difficulties associated with determining plant thioredoxin function are compounded by the fact that plants contain approximately four to five times as many thioredoxins as have been identified in animals. In addition, data collected during the course of this study indicate that most thioredoxins *h* have spatially and temporally overlapping expression profiles. The presence of multiple genes for each thioredoxin class in plants has raised two important questions regarding their function. Firstly, do the functions of thioredoxins overlap with one another and with other redox-active proteins as has been found in *E. coli* (Gallardo-Madueno et al 1998) and yeast (Draculic et al. 2001). Secondly, is there specificity in the interaction between thioredoxins and target proteins? These question can be addressed in part by the analysis of gene expression patterns in plants under specific conditions. Expression profiling was used as an indicator of the involvement of thioredoxins *h* in the stress responses of wheat. These studies demonstrated that the transcript levels of all wheat thioredoxins *h* underwent stress-responsive modulation. Moreover, differences in the expression profile suggest that the biochemical processes incorporating each thioredoxin were affected in a different manner and are indicative of each thioredoxin h having a separate function. However, the function of each thioredoxin h and the identity of interacting proteins still remain undefined.

Currently, the most promising tools for defining the function of each thioredoxin and the level of redundancy that exists in the *h*-class are those based upon transformation techniques. New technologies such as RNAi may provide a means of addressing isoform redundancy since this technique will enable the silencing of individual and multiple genes within a family of related sequences. In addition, transformation experiments employing a thioredoxin with an active site containing a second cysteine mutated to a serine (to form a stable mixed disulfide with interacting proteins) could be used to isolate thioredoxin-target protein complexes.

The data presented here suggest that such a study is likely to be less complex if attempted using a grass species since the level of thioredoxin *h* gene redundancy appears to be lowest in grasses.

Molecular evolutionary investigations in plants have been dominated by studies focusing on the chloroplast genome. By contrast, an understanding of the molecular evolutionary processes shaping nuclear genes in plants, particularly those belonging to gene families, is still in its infancy. The results presented in this study reveal that the primary phylogenetic structure of the thioredoxin *h* class is the same in all flowering plants. The defined nature of gene relationships in the thioredoxin *h*-class suggest that this class of genes may represent a good model family for studies on the evolution of nuclear multigene families and the levels of functional redundancy between family members.

Appendix A

PCR Primer Sequences

AAP	5' GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG 3'
(I = Inosine)	
A1	5' CGGACATGAGCAACACAAAGA 3'
PrxF	5' ACCCCCACCTGCAGCAATCAG 3'
PrxR	5' CTCGGCACCAGAGATTGTGAA 3'
R1	5' AAGCAAACAAGGATGGGAAAA 3'
R2	5' GCGGAAAAGACACGGAAACTG 3'
T1	5' CGATCCTTGCCCACACAG 3'
T2	5' ACGGATGTCCCATGTTGAG 3'
$_{CDSP32}$F	5' TGCCGCACTTCTCCTTCTACAA 3'
$_{CDSP32}$R	5' TAGCGGCGGCAGATCTTGTCGT 3'
Ta1F	5' GACCGGCAAGCTGGTGATCAT 3'
Ta1R	5' CCTGCCACCGACAACAGTGTC 3'
Ta2F	5' GGGCGGGGGAGGTGATCTC 3'
Ta2R	5' CTTGGCGAGATCAGCGAAAA 3'
Ta3F	5' CCGCCGTCGTCGCCGTCCACTC 3'
Ta3R	5' CTGCCCGCCCTTCACCAGTAC 3'

Appendix B

PCR Protocols

All PCR reactions were performed in a programmable thermal controller (MJ research, USA). A typical reaction composition was as follows; 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 0.01%(w/v) gelatin, 200 μM of each dNTP, 1 mM $MgCl_2$, 15 ng of each primer, approximately 50 ng of template cDNA and 2.5 units of *Taq* DNA polymerase in a reaction volume of 50μl.

Protocol 1

Step 1; 94^0C for 4 min

Step 2; 94^0C for 1 min

Step 3; 52^0C for 1 min

Step 4; 72^0C for 1 min

Step 5; return to step 2 for 30 cycles

Step 6; 72^0C for 10 min

Step 7; end

Protocol 2

As for Protocol 1 except the annealing temperature (step 3) was 55^0C

Protocol 3

As for Protocol 1 except the annealing temperature (step 3) was 58^0C

Protocol 4

Step 1; 94^0C for 4 min

Step 2; 94^0C for 1 min

Step 3; 61^0C for 1 min

Step 4; 72^0C for 1.5 min

Step 5; return to step 2 for 5 cycles

Step 6; 94^0C for 1 min

Step 7; 60^0C for 1 min

Step 8; 72^0C for 1.5 min

Step 9; return to step 6 for 5 cycles

Step 10; 94^0C for 1 min

Step 11; 58^0C for 1 min

Step 12; 72^0C for 1.5 min

Step 13; return to step 9 for 25 cycles

Step 14; 72^0C for 10 min

Step 15; end

Protocol 5

Step 1; 94^0C for 4 min

Step 2; 94^0C for 1 min

Step 3; 60^0C for 1 min

Step 4; 72^0C for 1 min

Step 5; return to step 2 for 5 cycles

Step 6; 94^0C for 1 min

Step 7; 58^0C for 1 min

Step 8; 72^0C for 1 min

Step 9; return to step 6 for 5 cycles

Step 10; 94^0C for 1 min

Step 11; 56^0C for 1 min

Step 12; 72^0C for 1 min

Step 13; return to step 9 for 25 cycles

Step 14; 72^0C for 10 min

Step 15; end

Appendix C

cDNAs of Bm2 Homologues

Alignment of the cDNA sequences from *Phalaris coerulescens* PTrx1 (GenBank accession number AF159388) and PTrx2 (AF159389), *Hordeum bulbosum* HbTrx (AF159385), *Lolium perenne* LpTrx (AF159387) and *Secale cereale* ScTrx (AF159386). Sequences are numbered from the translation initiation codon which is underlined. In frame stop codons are given in red and are in italics. Nucleotides that are identical within the alignment are in bold. Dashes represent gaps introduced to maximise the alignment.

```
PTrx1                                  CGCACGGCACGAGCGGCGCCGGCCGGGCCAA-----GTGGTGCGACGCCGAC
PTrx2                                  CGCACGGCACGAGCGGCGCCGGCCGGGCCAA-----GTGGTGCGACGCCGAC
HbTrx          AAAGCCGTTTCTTTCCGCAGGCGCACGCAAACACCAGCGGC-GCCGG-CCAACCCAGGTGGTGCAACCCC---
LpTrx     AAAAAAAGCCGTTTCTTTCCGCAGGCGCACGCGAACACCAGCGGC-GCCGG-CCGAGCCAGGTGGTGCGACGCC---
ScTrx                                                                  GTGCGACGC-GA-

PTrx1     CGACGCGGTTAATTCACGGGCTCGC----CCACGCACGCCGGCCGTCCGTCCGACG--CGG----------TTAAT-
PTrx2     CGACGCGGTTAATTCACGGGCTCGC----CCACGCACGCCGGCCGTCCGTCCGACG--CGG----------TTAAT-
HbTrx     -AACGCGGTTAATCCACGGGCTCGCTCACGCACGCACGCCGGCCGCCCGCCCGAC-----------TCCATCCAAT-
LpTrx     -GACGCGGTTAATCCACGGGCTCACTCACGCACGCACGCCGGCCGCCCGCCCGAC-----------TCCATCCAAT-
ScTrx     -GACGC-GT-----CACGGGCCCGCTCACAC-CG---GCCGGCCGGGAGC-GGACGGACGGGCCGATCGATCCA-TC

PTrx1     -----CCACCGCGTCGTCGCCGTCGCCGCCGTCGTCGTCGTCG---CCTCCAGAAACACGAGCCGGCCATAGCACGG
PTrx2     -----CCACCGCGTCGTCGCCGTCGCCGCCGTCGTCGTCGTCG---CCTCCAGAAACACGAGCCGGCCATAGCACGG
HbTrx     ----CCCACCGGGCCGTCGCCATCGCCGTCGTCGTCCTCGTCGTC-CCTCCAAAAACACAAACCGGGCATACCACGG
LpTrx     ----CCCACCGGGCCGTCGCCGTCGCCGTCGTCGTCCTCGTCGTCGCCTCCAGAAACACGAGCCGGGCATAGCACGG
ScTrx     CAGCCCCACCGGGTCGTCCCCGTCGTCCCC---GTCGTCGTCGTCTCCTCCAGAAACACGAGCCCAGCATAGCACGG

PTrx1     CCGCGGAATATTCCACGTCCC-TCCCCTCCCCCCTCCTCCGTCCTCATCCACGGGACCCCCCAGCACCACTACAAAA
PTrx2     CCGCGGAATATTCCACGTCCC-TCCCCTCCCCCCTCCTCCGTCCTCATCCACGGGACCCCCCAGCACCACTACAAAA
HbTrx     CCGCAAAATATTCCACTTCCCTTCCCTTCCACCCTC-TCCGTCCTCATCCACGGGA-CCCCCAGCACCAATACAAAA
LpTrx     CCGCAGAATATTCCACGTCCCTTCCCTTCCACCCTT-TCCGTCCTCATCCACGGGA-CCCCCAGCAGCAATACAAAA
ScTrx     CCGCAGAATATTCCACGTCCCTTCACTT---------TCCG-GCT-------------CCCCAGCAACAATA-AAAA

PTrx1     AAGAAAGAAAACCAATCCTCCTTCCCCCTCCTCCTCCGATTAATCAAACCA--CA--CGCTGCGGCGGCGCCCAGCG
PTrx2     AAGAAAGAAAACCAATCCTCCTTCCCCCTCCTCCTCCGATTAATCAAACCA--CA--CGCTGCGGCGGCGCCCAGCG
HbTrx     --CAAA-AAAACAAATCCTCCT-CCTCCTCCACCTCCGATTAATCAATCAAACCACTCCCTGCGGCGGCGGC-----
LpTrx     --CAA-GAAACCAAATCCTCCT-CCTCCTCCACCTCCGATTAATCAATCGAACCACTCGCTGCGGCGGCGGCG----
ScTrx     AGCAA-GC---CAA-TCCCCCA-CCTCTT--TCCTCCGATTAATCAA-------AC-CGC-GCGGCGGCGGCG-GCG

PTrx1     AAGCGAGCAACCTACGACCAGGAGAGGAGACC-------------GTCCACTGACCCCCACCACCACCACCACCGGA
PTrx2     AAGCGAGCAACCTACGACCAGGAGAGGAGACC-------------GTCCACTGACCCCCACCACCACCACC---GGA
HbTrx     AAA-GACCAACCTACAACCACCAGAGGAAACCCAACCAA------GTCCTCTGACCCCC-TC-ACCACCGC---CGGA
LpTrx     -AG-GAGCAACCTACGACCAGGAGAGGAGACCCAACCAACCACTGGTCGTCTGACCCCCTC-ACCACCGGC--CGGA
ScTrx     ACGTGCG--ACCTGCGACCACCAGA--CGACCCAACCGAC-------CCACTGACCC--AC-ACCACC------GGA

PTrx1     GACAGCGTTGCCTTGCACGTTAATCCACCCAGCC-GAGTTAAGGAGCCTCTTAATTGCCC----GGAGATCCGCCAG
PTrx2     GACAGCGTTGCCTTGCACGTTAATCCACCCAGCC-GAGTTAAGGAGCCTCTTAATTGCCC----GGAGATCCGCCAG
HbTrx     GAGCTC-TT-CCTTGTACGTTAATCAATCCAACCCGAGTTAAGGAACCTCTTAATTGCCC----GGAGATCCGCCAG
LpTrx     GAGCT--TTGCCTTCTACGTCAATCAATCCAGCCCGAGTTAAGGGACCTCTTAATTGCCC----GGAGATCCGCCAG
ScTrx     GAGCTC--TCCCTTTTACGTTAATCAATCCAACC-CAGTTAAAGAACCTCTTAATTGCCCGCCAGGAGATCCGCCAG

PTrx1     GCTCGTCGTCTTCGTTTC----------------------CCTT--CCCGCGG-CT------CTTGGCGCCAAAATCC
PTrx2     GCTCATCGTCTTCGTCTC----------------------CCTT--CCCGCGG-CT------CTTGGCGCCAAAATCC
HbTrx     GCTTACCGTCTTCGTCTCCTCTC-----------CTCCGCCTC--CCCGCGG-CT------CTTGGCGCCAAAATCC
LpTrx     GCTTATCGTCTTCGTCTCCTTCTC----------GTCCCCCTC--CCCGCGG-CT------CTTGGCGCCAAAATCC
ScTrx     GCTTACCTTCTCCGTCTCCTCCCAACCCAGCCCCCGCCCCCTCGCCCCGCGGTCTGGGTTCCTTGGCGCCAAAATCG
```

Appendix C - continued

```
PTrx1   CCGCTCCCCATCCCAGGGCCTTCAGG-------GGGCCTTTTG--CTCAGTGCTGCGTGC-TAGTTT-GTCGATTGA
PTrx2   CCGCTCCCCATCCCAGGGCCTTCAGG-------GGGCCTTTTG--CTCAGTGCTGCGTGC-TAGTTT-GTCGATTGA
HbTrx   CCCCTTCCCATCTCAGGGCCTTCAGG-------GGGCCCTTCG--TTCGGTG--GTGTGCCTAGTTTTGTCGACGGA
LpTrx   CCCTTTCCCATCTCAGGGCCTTCAGG-------GGGCCCTTCG--TTCGGTG--GTGTGCCTAGTTCGTCGACGGA
ScTrx   TCGCTTCCCATCCCAGG-CCTTCAGGAATCAGGGGGCGTTTTTTATTCAG----------CTAGTATTGTTGATTGA

                +1                                                              69
PTrx1   AGTTTACAATGGGAGGCTGTGTGGGCAAGGATCGTGGCATTGTGGAAGACAAGCTTGATTTCAAAGGTGGGAATGTG
PTrx2   AGTTTACAATGGGAGGCTGTGTGGGCAAGGATCGTGGCATTGTGGAAGACAAGCTTGATTTCAAAGGTGGGAATGTG
HbTrx   AGTTCACAATGGGAGGCTGTGTGGGCAAGGACCGTAGCATTGTAGAAGATAAGCTTGATTTCAAAGGTGGGAATGTG
LpTrx   AGTTCACAATGGGAGGCTGTGTGGGCAAGGACCGTAGCATTGTGGAAGATAAGCTTGATTTCAAAGGTGGGAATGTG
ScTrx   AGTTCAAGATGGGGGGCTGTGTGGGGAAGGGTCGTAGCATTGTGGAAGAAAAGCTTGATTTCAAAGGTGGGAATGTG

        70                                                                      146
PTrx1   CATGTCATAACTACCAAAGAGGACTGGGACCAGAAAATTGCAGAAGCAAACAAGGATGGGAAAATTGTTGTGGCAAA
PTrx2   CATGTCATAACTACCAAAGAGGACTGGGACCAGAAAATTGCAGAAGCAAACAAGGATGGGAAAATTGTTGTGGCAAA
HbTrx   CATGTCATCACAACCAAAGAGGACTGGGACCAAAAGGTTGCAGAAGCAAACAAGGATGGGAAAATTGTTGTTGCAAA
LpTrx   CATGTCATCACAACCAAAGAGGACTGGGACCAAAAGGTTGCAGAAGCTAACAAAGATGGGAAAATTGTTGTGGCGAA
ScTrx   CATGTCATAACAACCAAAGAGGACTGGGACCAGAAGATTGAAGAAGCAAACAAGGATGGGAAAATTGTTGTAGCAAA

        147                                                                     223
PTrx1   TTTCAGTGCTTCCTGGTGTGGGCCATGCCGTGTCATTGCACCTGTTTATGCTGAGATGTCAAAGACGTATCCTCAAC
PTrx2   TTTCAGTGCTTCCTGGTGTGGGCCATGCCGTGTCATTGCACCTGTTTATGCTGAGATGTCAAAGACGTATCCTCAAC
HbTrx   CTTCAGTGCTTCGTGGTGTGGGCCATGCCGCGTCATTGCACCTGTTTATGCTGAGATGTCCAAGACTTATCCTCAAC
LpTrx   TTTCAGCGCTTCCTGGTGTGGCCCATGCCGTGTCATTGCACCTGTTTATGCTGAGATGTCAAAGACTTATCCCCAAC
ScTrx   CTTCAGTGCTTCGTGGTGTGGGCCATGCCGTGTCGTTGCACCTGTTTATGCTGGGATGTCCAAGACTTATCCTCAAC

        224                                                                     300
PTrx1   TCATTCTTGACAATTGATGTTGATGACCTAGTGGATTTCAGCTCAACATGGGACATCCGTGCAACCCCAACGTTCTT
PTrx2   TCATTCTTGACAATTGATGTTGATGACCTAGTGGATTTCAGCTCAACATGGGACATCCGTGCAACCCCAACGTTCTT
HbTrx   TCATTCTTGACAATTGATGTTGATGACCTAATGGATTTCGGCTCAACATGGGACATCCGCGCAACGCCGACGTTTTT
LpTrx   TCATTCTTGACAATTGACGTTGATGACCTAATGGATTTCAGCTCAACATGGGACATCCGTGCAACCCCGACGTTCTT
ScTrx   TCATTCTTGACAATTGATGTTGATGACCTAATGGATTTCAGCTCAACATGGGACATCCGTGCAACCCCGACGTTCTT

        301                                                                     377
PTrx1   GTCTTCCTCAAGAATGGCCAGCAGATCGACAAGCTCGTTGGCGCCAACAAGCCTGAGCTCGAAAAGAAAGTACAAGC
PTrx2   GTCTTCCTCAAGAATGGCCAGCAGATCGACAAGCTCGTCGGCGCCAACAAGCCTGAGCTCGAAAAGAAAGTACAAGC
HbTrx   GTCTTCCTCAAAAACGGCCAGCAGATCGACAAGCTCGTCGGCGCCAACAAACCTGAGCTCGAGAAGAAAGTGCAAGC
LpTrx   GTCTTCCTCAAGAACGGCCAGCTGATTGACAAGCTCGTCGGCGCCAACAGACCTGAGCTCGAAAAGAAAGTACAAGC
ScTrx   GTCTTCCTGAAAAACGGCCAGCAGATCGACAAGCTCGTCGGCGCCAACAAACCTGAGCTCGAGAAGAAAGTACAAGC

        378             396
PTrx1   TCTTGGCGATGGCAGTTGA
PTrx2   TCTTGGCGATGGCAGTTGA
HbTrx   TCTTGGTGATGGCAGTTGA
LpTrx   TATTGGCGATGGCAGCTGA
ScTrx   TCTTGGTGATGGCAGTTGA
```

Appendix D

Grass Peroxiredoxins

Alignment of the *Phalaris* peroxiredoxin sequence (PcPrx) with the type-C peroxiredoxin EST sequences of barley (HvPrx, GenBank accession number BQ468113), maize (ZmPrx, BI430757), rice (OsPrx, AU031625) and wheat (TaPrx, BE427212). The sequences of the PrxF and PrxR primers are underlined and nucleotides differing from the consensus are in bold.

```
HvPrx ATGGCTCCGATTGGCGTGGGCAGCACCCTCCCCGACGGGCAGCTCGGGTGGTTCGACGA  59
ZmPrx ATGGCTCCCATCGCTGTCGGCGACGCCCTCCCCGACGGCCAGCTCGGGTGGTTCGACGA
OsPrx ATGGCCCCGGTTGCCGTGGGCGACACCCTCCCCGACGGCCAGCTGGGGTGGTTCGACGG
TaPrx ATGGCTCCGATTGGCGTGGGCAGCACCCTCCCCGACGGCCAGCTCGCGTGGTTCGACGA

HvPrx GAACGACCAGCTGCAGCAGGTCTCGATCCACTCCCTGGCCGCCGGCAAGAAGGTCATCC 118
ZmPrx GAACGACCAGCTGCAGCAGGTCTCCGTCCACGCGCTCGCCGCCGGCAAGAAGGTCATCC
OsPrx GGAGGACAAGCTGCAGCAGGTCTCCGTCCACGGCCTCGCCGCCGGCAAGAAGGTCGTCC
TaPrx GAGCGAACAGATGCAGCAGGTCTCCATCCACTCCCTGGCCGCCGGCAAGAAGGTCATCC

HvPrx TCTTCGGCGTCCCCGGCGCCTTCACCCCCACCTGCAGCAATCAGCATGTACCAGGCTTC 177
ZmPrx TCTTCGGCGTCCCCGGCGCCTTCACCCCGACCTGCAGCAACCAGCATGTGCCAGGCTTC
OsPrx TCTTCGGCGTCCCCGGTGCCTTCACCCCGACCTGCAGCAATCAGCATGTGCCAGGATTC
TaPrx TCTTCGGCGTCCCTGGCGCCTTCACCCCCACCTGCAGCAATCAGCATGTACCAGGCTTC
PcPrx                      ACCCCCACCTGCAGCAATCAGCATGTACCAGGCTTC

HvPrx ATTACTCAAGCTGAGGATCTCAAAGCCAAGGGTGTAGAGGAGATTCTTCTTGTCAGCGT 236
ZmPrx ATCACACAGGCTGAGCAGCTCAAAGCCAAGGGTGTAGACGAGATCCTGCTTATCAGCGT
OsPrx ATAAATCAGGCTGAGCAGCTCAAAGCCAAGGGTGTAGACGACATCTTGCTTGTCAGTGT
TaPrx ATTACTCAGGCCGAGGAGCTCAAAGCCAAGGGTGTAGATGAGATCCTGCTTGTCAGCGT
PcPrx ATTACTCAAGCCGAGGAGCTCAAAGCCAAGGGTGTAGACGAGATTCTGCTTGTCAGCGT

HvPrx TAATGACCCCTTTGTCATGAAGGCATGGGCAAAGACATACCCAGAGAACAAGCATGTGA 295
ZmPrx TAACGACCCCTTCGTCATGAAGGCGTGGGCGAAGACCTACCCCGAGAACAAGCACGTGA
OsPrx TAACGACCCCTTTGTCATGAAGGCGTGGGCAAAGTCATACCCTGAGAATAAGCATGTGA
TaPrx TAATGACCCCTTTGTCATGAAAGCATGGGCGAAGACATACCCAGATAACAAGCATGTGA
PcPrx TAATGACCCCTTTGTCATGAAGGCATGGGCAAAGACATAGCCAGAGAACAAGCATGTGA

HvPrx AGTTCCTTGCTGATGGAGCGGCAGCATACACAAAAGCACTTGGTCTTGAGCTTGATCTC 354
ZmPrx AGTTCCTCGCCGACGGATCTGGAGCGTACACCAAGGCGCTCGACCTCGAGCTCGATCTC
OsPrx AATTCCTTGCCGATGGTTTGGGAACATACACCAAGGCACTTGGTCTTGAGCTTGACCTT
TaPrx AGTTCCTTGCTGATGGAGCGGCAGCATACACAAAAGCACTTGGTCTTGAGCTTGATCTT
PcPrx AGTTCCTTGCTGATGGAGCGGCAGCATACACAAAAGCACTTGGTCTTGAGCTTGATCTC

HvPrx ACGGAGAAGGGATTGGGTCTTCGTTCGAAGCGCTTTGCTCTCCTTGCTGACGACCTCAA 413
ZmPrx ACGGACAAAGGGCTGGGCGTGCGCTCGAAGAGGTTCGCTCTCCTGGCCGACGACCTCAC
OsPrx TCGGAGAAAGGGCTTGGTATTCGTTCGAGACGGTTTGCTCTCCTTGCTGACAACCTCAA
TaPrx AGTGAGAAAGGATTGGGTCTCCGTTCGAGGCGGTTTGCTCTCCTTGCTGACGACCTCAA
PcPrx ACTGAGAAGGGATTGGGTCTTCGTTCGAAGCGCTTTGCTCTCCTTGCTGACGACCTCAA

HvPrx GGTCACCGTCGCAAACATCGAGGAAGGAGGCCAGTTCACAATCTCTGGTGCTGAAGAGA 472
ZmPrx GGTCACCGTCGCAAACATCGAGGAAGGCGGGCAGTTCACGATCTCCGGCGCTGAGGAGA
OsPrx GGTTACTGTTGCAAACATTGAGGAAGGTGGCCAATTCACAATCTCTGGTGCTGAGGAGA
TaPrx GGTCACCGTCGCAAACGTCGAGGAAGGTGGCCAGTTCACAATCTCTGGTGCCGAGGAGA
PcPrx GGTCACCGTCGCAAACATCGAGGAAGGTGGCCAGTTCACAATCTCTGGTGCCGAG

HvPrx TCCTCAAGGCACTGTAG 489
ZmPrx TCCTGAAGGCGCTTTAG
OsPrx TCCTCAAGGCACTGTAA
TaPrx TCCTCAAGGCACTGTAG
```

Appendix E

ClustalW alignment of the deduced protein sequences of grass *h*-class thioredoxins.

```
Ta1 MAAEE.................GAVIACHTKQEFDTHMANGK.ETGKLVIIDFTASWCGPCRVIAPVFAEYA
As1 MAAAE.................GAVIACHTKQEFDTHMANGK.ETGKLVIIDFTASWCGPCRVIAPVFAEYA
Hv1 MAAEE.................GAVIACHTKQEFDTHMANGK.DTGKLVIIDFTASWCGPCRVIAPVFAEYA
Os1 MAAEE.................GVVIACHNKDEFDAQMTKAK.EAGKVVIIDFTASWCGPCRFIAPVFAEYA
Zm1 MASEQ.................GVVIACHSKAEFDAHMTKAQ.EAGKLVVIDFTAAWCGPCRAIAPLFVEHA
Sb1 MASEE.................GVVIACHTKAEFDAQMAKAK.EAGKLVVIDFTASWCGPCRAIAPLFVEHA
Zm2 MAASE..AA.AAAATPVTPTEGTVIAIHSLEEWSIQIEEAN.SAKKLVVIDFTATWCPPCRAMAPIFADMA
Sb2 MATTE..AA.A..ATPVAPAEGSVIAIHSLDEWSIQIEEAN.SAKKLVVIDFTATWCPPCRMIAPVFAELA
Os2 MAAAS..AA.A.......QAEGTVIAIHSLDEWTIQIEEAN.SAKKLVVIDFTASWCGPCRIIAPVFADLA
Td2 MAAAATATT.TAAATAAAVGPGEVISVHSLEQWTMQIEEAN.AAKKLVVIDFTASWCGPCRIMAPIFADLA
Tm2 .......................EVISVHSLEQWTMQIEEAN.AAKKLVVIDFTASWCGPCRIMAPIFADLA
Ta2 MAASA...A.TATAT.AAVGAGEVISVHSLEQWTMQIEEAN.AAKKLVVVDFTASWCGPCRIMAPVFADLA
Sc2 MAASAT..A.AATATTAAVGAAEVISVHTLEQWTMQIEEAN.AAKKLVVVDFTASWCGPCRIMAPVFADLA
Hv2 MAASAT..........AAAVAAEVISVHSLEQWTMQIEEAN.TAKKLVVIDFTASWCGPCRIMAPVFADLA
Tan MGGCVGKGR.SIVEEKLDFKGGNVHVITTKEDWDQKIEEAN.KDGKIVVANFSASWCGPCRVIAPVYAEMS
Scn MGGCVGKGR.SIVEEKLDFKGGNVHVITTKEDWDQKIEEAN.KDGKIVVANFSASWCGPCRVIVPVYAEMS
Hvn MGGCVGKGR.GVVEEKLDFKGGNVHVITTKEDWDQKIEEAN.KDGKIVVANFSASWCGPCRVIAPVYAEMS
Hbn MGGCVGKGR.GIVEEKLDFKGGNVHVITTKEDWDQKIEEAN.KDGKIVVANFSASWCGPCRVIAPVYAEMS
Pcn MGGCVGKDR.GIVEDKLDFKGGNVHVITTKEDWDQKIAEAN.KDGKIVVANFSASWCGPCRVIAPVYAEMS
Lpn MGGCVGKDR.SIVEDKLDFKGGNVHVITTKEDWDQKVAEAN.KDGKIVVANFSASWCGPCRVIAPVYAEMS
Osn MGGCVGKGRRHIEEDKLDFKGGNVHVITSKEDWDRKIEEAN.KDGKIVVANFSASWCGPCRVIAPIYAEMS
Zmn MGGCAGKVR.RDDEEKLDFKGGNVHIITSNEGWDQKIAEAN.RDGKTVVANFSASWCGPCRVIAPVYAEMS
Sbn MGGCIAKQH.ADDEDKIDFGGGNVHVVTSKEDWDQKIAEAN.KDGKIVVANFSASWCGPCRVISPVYAEMS
Sb3 MGNFIS.SLVTPPP.AADDPNCAVVAAHSKTTYDEQWAAHK.NGGKLMVIDFSASWCGPCRFIEPAFKELA
Zm3 MGSFLS.SLVTPPP.AADDPNCAVVAAHSKATYDEQWAAHK.SSSKLMVIDFSASWCGPCRFIEPAFKELA
As3 MGSFLS.SLWTPPPLPGDDPDSAVVAVHSKPAWDKHWEEHR.NASKLMVIDFSASWCGPCRFIEPAFKEMA
Tm3 MGSFLS.SLWTPPPLPGDDPDSAVVAVHSKPAWDRHWEAHR.NASKLMVIDFSASWCGPCRFIEPAFNEMA
Ta3 MGSFLS.SLWTPPPLPGDDPDSAVVAVHSKPAWDRHWEAHR.NACKLMVIDFSASWCGPCRFIEPAFKEMA
Sc3 MGSFLS.SLWTPPPLAGDDPDSAVVAVHSKPAWDRHWEAHR.NASKLMVIDFSASWCGPCRFIEPAFKEMA
Hv3 MGSFLS.SLFTPPPLPGDDPDSAVVAVHSKPAWDQHWEAHR.NASKLMVIDFSASWCGPCRFIEPAFKEMA
Os3 MGSFFS.TMFTPPPAADDGGDSRVVAVHSTATWDEQWGAHKSNPNKLIVIDFSATWCGPCRFIEPAFKDMA

Ta1 KKFPGAIFLKVDVDELKDVAEAYNVEAMPTFLFIKDGAKVDTVVGGRKDDIHTKIVALMG....SASA....
As1 KKFPGAIFLKVDVDELKDVAEAYNVEAMPTFLFIKDGAKVDTVVGGRKDDIHTKIVALMG....SASA....
Hv1 KKFPGAIFLKVDVDELKDVAEAYNVEAMPTFLFIKDGEKVDSVVGGRKDDIHTKIVALMG....SAST....
Os1 KKFPGAVFLKVDVDELKEVAEKYNVEAMPTFLFIKDGAEADKVVGARKDDLQNTIVKHVGATAASASA....
Zm1 KKFTQVVFLKVDVDEVKEVTAAYEVEAMPTFHFVKNGKTVATIVGARKDELLAQIEKHAA....PAPASASA
Sb1 KKYTQAVFLKVDVDELKEVTAEYKIEAMPTFHFIKNGETVETIVGARKDELLALIQKHTA....SASA....
Zm2 KKSPNVVFLKVDVDEMKTIAEQFSVEAMPTFLFMREGDVKDRVVGAAKEELARKLELHMAS..........
Sb2 KKHPNVVFLKVDVDEMKTIAEQFSVEAMPTFLFMREGDVKDRVVGAAKEELANKLQLQMAQ..........
Os2 KKHTNAVFLKVDVDELKPIAEQFSVEAMPTFLFMKEGDVKDRVVGAMKDELASKVQLHMA..........
Td2 KKFPAAVFLKVDVDELKPIAEQFSVEAMPTFLFMKEGDVKDRVVGAIKEELTTKVGLHAAA..........
Tm2 KKFPAAVFLKVDVDELKPIAEQFSVEAMPTFLFMKEGDVKDRVVGAIKEELTTKVGLHAAA..........
Ta2 KKFPAAVFLKVDVDELKPIAEQFSVEAMPTFLFMKEGDVKDRVVGAIKEELTTKVGLHAAQ..........
Sc2 KKFPNAVFLKVDVDELKPIAEQFSVEAMPTFLFIKEGDVKDRVVGAIKEELTTKVGLHAAAAQ.........
Hv2 KKFPNAVFLKVDVDELKPIAEQFSVEAMPTFLFMKEGDVKDRVVGAIKEELTAKVGLHAAAQ..........
TaN KTYPQLMFLTIDVDDLMDFSSTWDIRATPTFFFLKNGQQIDKLVGANKPELEKKVQALGDGS..........
ScN KTYPQLMFLTIDVDDLMDFSSTWDIRATPTFFFLKNGQQIDKLVGANKPELEKKVQALGDGS..........
HvN KTYPQLMFLTIDVDDLMDFSSTWDIRATPTFFFLKNGQQIDKLVGANKPELEKKVQALGDGS..........
HbN KTYPQLMFLTIDVDDLMDFSSTWDIRATPTFFFLKNGQQIDKLVGANKPELEKKVQALGDGS..........
PcN KTYPQLMFLTIDVDDLVDFSSTWDIRATPTFFFLKNGQQIDKLVGANKPELEKKVQALGDGS..........
LpN KTYPQLMFLTIDVDDLMDFSSTWDIRATPTFFFLKNGQLIDKLVGANRPELEKKVQAIGDGS..........
OsN KTYPQLMFLTIDVDDLMDFSSSWDIRAKPTFFFIKNEKQVDKLVGANKPELEKKVQALADGS..........
ZmN KTYPQLMFLTIDVDDLMDFSSSWDIRATPTFFFLKNGQQIDKLVGANKPELEKKVLAAADASTS........
SbN QTYPQLMFLTIDVDELMEFSSSWDIRATPTFFFLKNGQQVDKLVGANKPELEKKVAAIAGAS..........
Sb3 SRFTDAIFVKIDVDELAEVARTWKVEAMPTFVLVKDGKEVSRVIGAKKDELERKIQMFIMSSSSS.......
Zm3 SRFSDAIFVKVDVDELAEVARTWKVEAMPTFVLVKDGKEVSRVVGAKKDELERKIRMFTSSSSS........
As3 SRFADALFVKIDVDELAEVAKTFRVEAMPTFVLVKG.....................................
Tm3 SRFADALFVKIDVDELAEVAKTFRVEAMPTFVLV.......................................
Ta3 SRFADALFVKIDVDELAEVAKTFRVEAMPTFVLVKGGQEVSRVVGAKKDELDRKIKTFISSS..........
Sc3 SRFADALFVKIDVDELAEVAKTFRVKAMPTFVLMKGG.....................................
Hv3 SRFADALFVKIDVDELAEVAKTFRVEAMPTFVLVKGGQEVSRVVGAKKDELDRKIKTFISSS..........
Os3 GHFADAVFFKIDVDELSEVARQWKVEAMPTFVLIKGGKEVSRVVGAKKDELERKVNMFISSSSS........
```

Appendix F

Sequences of wheat cDNA probes amplified by RT-PCR. Primer sequences are underlined.

Wheat thioredoxin *h1* probe (Ta1F/R primer set)

GACCGGCAAGCTGGTGATCATTGACTTCACTGCTTCCTGGTGCGGTCCTTGTCGTGTCATAGCCCC
AGTCTTTGCTGAGTACGCCAAGAAGTTCCCTGGCGCCATTTTCCTGAAGGTGGACGTTGACGAGCT
GAAGGACGTCGCTGAAGCATACAACGTTGAGGCAATGCCGACCTTCCTGTTTATCAAGGATGGTGC
GAAGGTGGACACTGTTGTCGGTGGCAGG

Wheat thioredoxin *h2* probe (Ta2F/R primer set)

GGGCGGGGGAGGTGATCTCCGTCCACAGCCTGGAGCAGTGGACCATGCAGATCGAGGAGGCCAACG
CCGCCAAGAAGCTGGTGGTGATTGACTTCACTGCATCATGGTGCGGACCATGCCGCATTATGGCTC
CAATTTTCGCTGATCTCGCCAAG

Wheat thioredoxin *h3* probe (Ta3F/R primer set)

CCGCCGTCGTCGCCGTCCACTCCAAGCCCGCCTGGGACCGGCACTGGGAGGCGCACCGGAACGCGT
GCAAGCTGATGGTGATCGACTTCTCCGCCTCCTGGTGCGGGCCCTGCCGCTTCATCGAGCCCGCCT
TCAAGGAGATGGCCTCCCGCTTCGCCGACGCGCTCTTCGTCAAGATCGACGTCGACGAGCTCGCGG
AGGTTGCAAAGACATTCCGCGTAGAGGCGATGCCGACCTTCGTACTGGTGAAGGGCGGGCAG

Wheat thioredoxin *n* probe (R1/T2 primer set)

AAGCAAACAAGGATGGGAAAATAGTTGTGGCAAACTTCAGCGCTTCGTGGTGTGGGCCATGCCGTG
TCATTGCACCTGTTTATGCTGAGATGTCCAAGACTTACCCTCAACTCATTCTTGACAATTGATGTT
GATGACCTAATGGATTTCAGCTCAACATGGGACATCCGT

Wheat C-type peroxiredoxin probe (PrxF/R primer set)

ACCCCCACCTGCAGCAATCAGCATGTACCAGGCTTCATTACTCAGGCCGAGGAGCTCAAAGCCAAG
GGTGTAGATGAGATCCTGCTTGTCAGCGTTAATGACCCCTTTGTCATGAAAGCATGGGCGAAGACA
TACCCAGATAACAAGCATGTGAAGTTCCTTGCTGATGGAGCGGCAGCATACACAAAAGCACTTGGT
CTTGAGCTTGATCTTAGTGAGAAAGGATTGGGTCTCCGTTCGAGGCGGTTTGCTCTCCTTGCTGAC
GACCTCAAGGTCACCGTCGCAAACGTCGAGGAAGGTGGCCAGTTCACAATCTCTGGTGCCGAG

Appendix G

Alignment of the wheat CDSP32 RT-PCR fragment sequence (*TaCDSP32*) with the corresponding region of the potato CDSP32 cDNA (*StCDSP32*; Rey et al. 1998) and EST sequences of CDSP32 homologues from barley (*HvCDSP32*, Genbank accession number BE421930) and maize (*ZmCDSP32*, BG841177). The sequences of the WCDSP32F and WCDSP32R primers are underlined and nucleotides differing from the consensus are in bold. Dashes indicate gaps introduced to maximise the alignment.

```
StCDSP32 TACCACATTTCAACTTCTACAAAAGCATGGAGAAGATCCACGAGGAAGAAGGCATT
HvCDSP32 TGCCGCACTTCTCCTTCTACAAGGGCACCGAGAAGGTGCACGAGGAGGAAGGCATC
TaCDSP32 TGCCGCACTTCTTCTTCTACAAGGGCACCGACAAGGTGCACGAGGAGGAGGGCATC  56

StCDSP32 GGCCCGGACCTACTAGCCGGTGATGTACTCTACTACGGTGACAGCCACTCTGAAGT
ZmCDSP32                                          CAGCCACTCGGCGGT
HvCDSP32 GGCCCCGACCAGCTCGCCGGCGACGTGCTCTACTACGGCGACAACCACGCCGGCGT
TaCDSP32 GGCCCCGACCAGCTCGCCGGCGACGTGCTCTACTACGGCGACAACCACGCCGGCGT 112

StCDSP32 AGTGCAGCTCCACAGCAGAGAAGACGTTGAAAAGGTAATCCAAGATCACAAAATCG
ZmCDSP32 GGTGCAGCTGCACTCGCGGGAGGACGTGGAGGCGCTCATCGACGAGCACCCGCGGCG
HvCDSP32 CGTGCAGCTGCACAGCCGGGCCGACGTGGAGGCGCTCATGGCGGAGAACAGCGGCG
TaCDSP32 CGTGCAGCTGCACAGCCGCACGGACGTGGAGGCGCTCATGGCGGAGCACAGCGGCG 168

StCDSP32 ATAAA---AAGTTAATAGTCCTCGATGTTGGATTGAAGCATTGTGGACCATGTGTG
ZmCDSP32 ACAAGGGCAAGCTCGTCGTGCTGGACGTGGGCCTCAAGCACTGCGGGCCCTGCGTC
HvCDSP32 AGGGCGGGAAGCTGCTGGTGCTGGACGTCGGGCTCAAGCATTGCGGGCCCTGCGTC
TaCDSP32 AGGGCGGCAAGCTGCTGGTGCTGGACGTCGGGCTCAAGCACTGCGGGCCATGCGTC 224

StCDSP32 AAAGTTTATCCAACAGTGATCAAGCTATCGAAGCAGATGGCTGATACAGTCGTGTT
ZmCDSP32 AAGGTGTACCCCACCGTGCTGAAGCTGTCGCGGTCCATGGTCGACAACACCGTCTT
HvCDSP32 AAGGTCTACCCCACCGTGGTCAAGCTGTCCCGCTCCATGGCCGACACAACCGTCTT
TaCDSP32 AAGGTCTACCCCACCGTCGTCAAGCTCTCCCGCTCCATGGCCGACACCGCCGTCTT 280

StCDSP32 CGCGAGAATGAATGGCGATGAGAATGATAGTTGTATGCAGTTTTTGAAAGACATGG
ZmCDSP32 CGCGCGCATGAACGGCGACGAGAACGACAGCTGCATGGAGTTCCTCAGGGCCATGA
HvCDSP32 CGCGCGCATGAACGGCGACGAGAACGACGCCTGCATGGACTTCCTCAAGGACATGG
TaCDSP32 CGCGCGCATGAACGGCGACGAGAACGACGCATGCATGCAGTTCCTCAAGGACATGG 336

StCDSP32 ATGTTATTGAAGTGCCTACATTTTTGTTTATAAGAGATGGTGAGATTTGTGGAAGG
ZmCDSP32 AGATCGTGGAGGTGCCCACTTTCGTCTTCATCAGGGACGGCCAGATCGTCGGCCGC
HvCDSP32 AGGTCGTGGAGGTGCCCACCTTCCTCTTCATCAGGGACAACAAGATCTGTCGCCGC
TaCDSP32 AGGTCGTCGAGGTGCCCACCTTCCTCTTCATCAGGGACGTCAAGATCGTCGGCCGC 392

StCDSP32 TA
ZmCDSP32 TA
HvCDSP32 TA
TaCDSP32 TA 394
```

Appendix H

ClustalW alignment of all plant *h*-class thioredoxins present on the Genbank database. Species identifiers are the same as listed in Table 6.1. Dots indicate gaps in the alignment. Residues at invariant or highly conserved positions are in bold.

```
Horse     .....MVKQIESKSAFQEALNSAG..EKLVVVDFSAT
Human     .....MVKQIESKTAFQEALDAAG..DKLVVVDFSAT
Mouse     .....MVKLIESKEAFQEALAAAG..DKLVVVDFSAT
Chicken   .....MVKSVGNLADFEAELKAAG..EKLVVVDFSAT
St1       .....MAEE.....GQVFGVHNVDEWNQHLQKGIDNKKLIVVDFTAS
Le3       .....MAEG.....GQVFGVHTVDEWHQHLQKGIDNKKLLVVDFTAS
Nt2       .....MAEE.....GQVIGVHTVDAWNEHLQKGIDDKKLIVVDFTAS
Rc1       .....MAAEE.....GQVIGCHTVEAWNEQLQKGNDTKGLIVVDFTAS
Cs1       .....MAAAEE.....GQVIGCHTVEAWNEQLQKSNETKQLVVVDFTAS
Ga3       .....MAAEE.....GQVYGCHTIESWNEQLQKGNESEKLVVVDFTAS
Mc3       .....MAAEE.....GQVISCHTTQAWEEQLNKANESKKLMVVDFTAS
At1       .....MASEE.....GQVIACHTVETWNEQLQKANESKTLVVVDFTAS
Mp1       .....MASSES.....EGQVIGCHTTDTWNEQLQKANDNKKLVVVDFTAS
Gm5       .....MAGSSE.....EGQVISCHTVDAWNDQLQKGNQSKKLIVVDFTAS
Gm8       .....MAGSSE.....EGQVISCHTVEEWNDQLQKGNESKKLIVVDFTAS
Ga4       .....MAEE.....GQVIACHTVESWQQQLQMGNESNKLVVVDFTAS
Ga2       .....MAAEE.....GQVISCHTLDSWKEQLQKGGQSKKLVVVDFTAS
Vv1       .....MAEE.....GQVVGCHSVESWKEQFQHGIESKKLVVVDFTAS
Gm4       .....MAEVE.....EGQVIGVHTVDEWKLQLQNAKDSKKLIVVDFTAS
Gm7       .....MAE.....EGQVVGVHTVDAWNQQLQNGKDSQKLIVVDFTAS
Mt3       .....MAAE.....EGQVIGVHTVEQWKEEIQKGNDSKKLIVVDFTAS
Mt4       .....MAAE.....EGHVIGVHTVEAWKEHLEKGNGSKKLIVVDFTAS
Lj1       .....MAAE.....EGTVIGVHTVEAWKEHLEKGNVSKKLIVVDFTAS
Sros1     .....MAE.....EGQVIGVHTVDAWKQHLEKGNENKKLIVVDFTAS
Gm3       .....MAE.....EGQVIGVHSVEEWKEHLKKGEESKKLIVVDFTAS
Gm9       .....MAE.....EGQVIGVHSVEEWEEHLKKGQESKKLIVVDFTAS
Pper1     .....MAE.....ENQVIGCHTTQAWEEQLHKGNENKKLVVVDFTAS
St2       ....MATSSEE.....GQVIGCHKVEEWEVQLQKGVETKKLVVVDFTAS
Le2       ....MATSSEE.....GQVIGCHKVEEWKVQLQKGVETKKLVVVDFTAS
Nt1       MAANDATSSEE.....GQVFGCHKVEEWNEYFKKGVETKKLVVVDFTAS
Ib1       ..MAAATSSEE.....GQVIACHTVDHWKEQFAKGVETKKLVVVDFTAS
Sr1       .....MAEE.....GVVIGCHSTDQWKEQLEKHKSTQKLVVVDFTAS
Ha1       .....MAEE.....GIVVGCHNADQWKEHFDKHKASQKLVVVDFTAS
Mc1       .....MAEQ.....NQVIGCHNVAQWEEHFNKGKDAKRLVVVDFTAS
Ph1       .....MAEE.....GQVIACHTVDTWKEHFEKGKGSQKLIVVDFTAS
Pb1       .....MAEE.....GQVIACHTVDVWKEQFEKGKGTQKLIVVDFTAS
Fe1       .....MAEE.....AQVIACHTVQEWNEKFQKAKDSGKLIVIDFTAS
Os1       .....MAAEE.....GVVIACHNKDEFDAQMTKAKEAGKVVIIDFTAS
Ta1       .....MAAEE.....GAVIACHTKQEFDTHMANGKETGKLVIIDFTAS
Zm1       .....MASEQ.....GVVIACHSKAEFDAHMTKAQEAGKLVVIDFTAA
Sb1       .....MASEE.....GVVIACHTKAEFDAQMAKAKEAGKLVVIDFTAS
At4       .....MAAEE.....GQVIGCHTNDVWTVQLDKAKESNKLIVIDFTAS
Bn2       .....MAAEE.....GQVIGCHEIDVWAVQLDTAKQSNKLIVIDFTAS
Zm2       .....MAASEAAAA.....AATPVTPTEGTVIAIHSLEEWSIQIEEANSAKKLVVIDFTAT
Sb2       .....MATTEAAA.....ATPVAPAEGSVIAIHSLDEWSIQIEEANSAKKLVVIDFTAT
Os2       .....MAAASAAA.....QAEGTVIAIHSLDEWTIQIEEANSAKKLVVIDFTAS
Ta2       .....MAASAATA.....TATAAVGAGEVISVHSLEQWTMQIEEANAAKKLVVVDFTAS
Bn1       .....MAATAEVI.....PAGEVIACHTVEDWNNKLKAAKESNKLIVIDFTAV
Ts1       .....MAG.....DGEVIACQTVEDWNEKLKAAQESKKLIVIDFTAV
At3       .....MAA.....EGEVIACHTVEDWTEKLKAANESKKLIVIDFTAT
At5       .....MAG.....EGEVIACHTLEVWNEKVKDANESKKLIVIDFTAS
Mc2       .....MAA.....DAVIACHTVDVWNEKVKDANESKKLIVIDFTAS
Sb3       .....MGNFIS.....SLVTPPP.AADDPNCAVVAAHSKTTYDEQWAAHKNGGKLMVIDFSAS
Zm3       .....MGSFLS.....SLVTPPP.AADDPNCAVVAAHSKATYDEQWAAHKSSSKLMVIDFSAS
Ta3       .....MGSFLS.....SLWTPPPLPGDDPDSAVVAVHSKPAWDRHWEAHRNACKLMVIDFSAS
Os3       .....MGSFFS.....TMFTPPPAADDGGDSRVVAVHSTATWDEQWGAHKNPNKLIVIDFSAT
St3       .....MGGYYY.....PTWPEFNKKPTTPQIKGSQVIAFHSSTKWKLHFDSLKDTNKLVVIDFTAT
Le1       .....MSGYNY.....PTWPEFNKKPITQQIKRCQVITFNSSTKWMIHLNFLKDTNKLVVIDFTAT
Le5       .....S.....VTLIIEAIINMPITPQFKRSQVIAFHSSTKWKLHFDSLKNTNKLVVIDFTAT
Mt1       .....MG.....GNLSNMEHGHTSAKSLSSHILTFHSTAKWKAYFDASKETNKLMVIEFTAA
Mt5       .....MG.....ANFSNLES.VVEKSSQPSLILTFHSTAKWKAHFEASKVTNKLMVIDFTAT
Gm6       .....MAMAYVARSSESSQVLNFHSTAKWNAHFDALKQTNKLMVVDFTAS
St4       .....ASFLG.....GGEAQAATEVSGSPSEPSRVIAFHSSNRWQLHFNSSKQLNKLIVVDFAAA
Le4       .....MGSFVSTLLG.....GGEAQAATEVSGSPSEPSRVIAFHSSNRWQLHFNSSKQLNKLIVVDFAAA
```

```
Gm1     .......MGAILSALTG......GAATAAT...SSPESSASRVQSFHSSARWQLHPNELKETNKLVVIDFSAS
Gm2     .......MGGILSGLLG......SDAAAAA...SSPESSTSRVSSFHSSPRWQLYFNEIKDTDKLVVIDFSAS
At2     .......MGGALSTVFG......SGEDATA...AGTESEPSRVLKFSSSARWQLHFNEIKESNKLLVVDFSAS
Ga1     .......MGSFFS..........SDSTPEK..SSSSSEHSGIQTFHSSPRWQLHFNSVKDSSQLMVIDFSAS
Ga5     .......MGSFLSSLLG......SSGSPSED..SPSSSESSRVSTFHSAPRWQLHFNSVKESPKLMVIDFSAS
Mt2     .......MGSFLSSLVG......GDSATASQ..SSESSENSSVKTFHSSARWQLHFNELKDSPRLVVIDFSAT
Bv1     .......MGNRFA.....................RGKPPAVVAIHSTEQWNAHFELSKSSAKLMVVDFSAS
Mc4     .......MGNNCS.....................RRKLPAVLVFHSADQWKTYFESSKYSEKLMVIDFSAS
Sr2     .......MGAQVSS....................QQRSDHGEVLSVHSLDNWNTRFQNSKTSHKLMVIDFSAA
Pt1     .......MAD..........................GNVFACHSTEAWRSKLQEAIDTKRLVVVDFTAT
Pt2     .......MAE..........................GNVFACHSTEAWRSKLQEAIDTKRLVVVDFTAT
Pm1     .......MAE..........................GNVFACHSTEGWRSKLQEAIDTKRLVAVDFTAT
Cj1     .......MAD..........................GLVVACHSVDTWRSKLQEAKTSGKLVVVDFTAT
Pcn     .......MGG.........CVG.KDRGIV...EDKLDFKGGNVHVITTKEDWDQKIAEANKDGKIVVANFSAS
Tan     .......MGG.........CVG.KGRSIV...EEKLDFKGGNVHVITTKEDWDQKIEEANKDGKIVVANFSAS
Osn     .......MGG.........CVG.KGRRHI...EDKLDFKGGNVHVITSKEDWDRKIEEANKDGKIVVANFSAS
Zmn     .......MGG.........CAG.KVRRDD...EEKLDFKGGNVHIITSNEGWDQKIAEANRDGKTVVANFSAS
Sbn     .......MGG.........CIA.KQHADD...EDKIDFGGGNVHVVTSKEDWDQKIAEANKDGKIVVANFSAS
Stn     .......MGISDTVRSLFPCIKSHSTSDGDDSTHNVEFAGGNVSLITTKESWDQKLAEAKKEGKIVIANFSAS
Len     .......MGISDTVRSLFPCIKSHSTSDGDDSTHNVEFAGGNVSLITTKESWDQKLAEAKKEGKIVIANFSAS
Ntn     .......MGITDMVHSLFPCIKSRSTNNDGDSSHNVKFAGGNVSLITTKESWDQKLAEANKEGKIVIANFSAS
Mtn     .......MGN.........CLA.KSRDRDNDSDQHVEFAAGNVALITTKEAWDQKLEEAKKDGKIVIANFSAV
Gm10    .......MGS.........CVS.KNHAKDNDSDHDVDFAAGNVKLITTKEAWDQSLEEARRGGKIVIANFSAT
Gmn     .......MGN.........CLR.KAHADD.DSDHIVELASGNVQLITTKESWDQKLDQARKESKIVIANFSAT
Atn     .......MGS.........CVS.KGKGDD.DSVHNVEFSGGNVHLITTKESWDDKLAEADRDGKIVVANFSAT
Ptn     ..............................MAVAHGNMHVITSKQEWDAKIFEANTNGKIIVVDFTAS
Cr1     ...................................MSHGNVHVIDSKNAFENKLGEGKSSNKVVVVDFTAT
Cr2     ...................................MAHGNVHVVHSQSFWDDKLAEGKSTNKVVVVDFTAT
Pp1     ..................................MAADHGNIHIVNNTVEWKTKLDEATSSGKIVVVDFTAT
Pp2     ....................................MDHGKVHVINNSAAWDAKLAEATSTGKIVLVDFTAT
Liv1    ..................................MADTGNVHVIDSKESWSSILADGATHKKTIVVDFTAT

Horse   WCGPCKMIKPFFHSLSEKYS.NVVFLEVDVDDCQDVAAECEVKCMPTFQFFKKGQKVDEFSGANK.EKLEATI
Human   WCGPCKMIKPFFHSLSEKYS.NVIFLEVDVDDCQDVASECEVKCMPTFQFFKKGQKVGEFSGANK.EKLEATI
Mouse   WCGPCKMIKPFFHSLCDKYS.NVVFLEVDVDDCQDVAADCEVKCMPTFQFYKKGQKVGEFSGANK.EKLEASI
Chicken WCGPCKMIKPFFHSLCDKFG.DVVFIEIDVDDAQDVATHCDVKCMPTFQFYKNGKKVQEFSGANK.EKLEETI
St1     WCGPCKFIAPFLAELAKKIP.TVTFLKVDVDELKSVATDWAVEAMPTFMFIKEGKIVDKVVGAKK.EELQQTI
Le3     WCGPCKFIAPFLAELAKKIP.TVTFLKVDVDELKSVATDWAVEAMPTFMFIKEGKIVDKVVGAKK.DELQQTI
Nt2     WCGPCKFIAPFYAELAKKMP.TVTFLKVDVDELKSVATDWAVEAMPTFMFLKEGKIVDKVVGAKK.DELQQTI
Rc1     WCGPCRFIAPFLAELAKKLP.NVTFLKVDVDELKTVAHEWAVESMPTFMFLKEGKIMDKVVGAKK.DELQQTI
Cs1     WCGPCRFIAPFLAELAKKLP.NVLFLKVDVDELKSVATDWAVEAMPTFMFLKEGKIVDKVVGAKK.EELQQTI
Ga3     WCGPCRFIAPFLAELARKFP.SVTFLKVDVDELKEVATDWAVEAMPTFMFLKEGKIVDKVVGAKK.EELQQTL
Mc3     WCGPCRFIAPFLAELAKKLP.QVIFVKVDVDELKQLAADWAIEAMPTFMFLKEGKIVDKVIGAKK.DELQQTV
At1     WCGPCRFIAPFFADLAKKLP.NVLFLKVDTDELKSVASDWAIQAMPTFMFLKEGKILDKVVGAKK.DELQSTI
Mp1     WCGPCRFIAPFFAELAKKFP.NVTFLKVDVDELKSVASDWAVEAMPTFIFLKEGKILDRVVGAKK.EELQANI
Gm5     WCGPCRFIAPFLAELAKKFT.SVVFLKVDVDELKSVSQDWAIEAMPTFVFVKEGTLLSKVVGAKK.DELQQTI
Gm8     WCGPCRFIAPFLAELAKKFT.SVIFLKVDVDELKSVSQDWAIEAMPTFVFVKEGTLLDKVVGAKK.DELQQKI
Ga4     WCGPCRFISPILVDLAKKLP.NVIFLKVDVDELKSVAQSWAIEAMPTFVFLKGGTIIDKLVGARK.DELQQKI
Ga2     WCGPCRFISPFLAELAKKFP.NVMFLKVDVDELKEVAAEWDVDAMPTFLFLKEGKAVEKVVGAKK.EELQQAV
Vv1     WCGPCRVISPFLAELAKKMP.NVIFLKVDVDELETVAKEWEVEAMPTFLFVKEGNVVDKVVGAKK.EELVQKT
Gm4     WCGPCRFMAPVLAEIAKKTP.ELIFLKVDVDEVRPVAEEYSIEAMPTFLFLKDGEIVDKVVGASK.DDLQATI
Gm7     WCGPCRFIAPVLAEIARHTP.QVIFLKVDVDEVRPVAEEYSIEAMPTFLFLKDGKIVDKVVGAKK.EELQLTI
Mt3     WCGPCRFIAPILAEIAKKIP.EVIFLKVDIDEVKSVAKEWSVEAMPTFLFLKEGKEVDKVVGARK.EELENAI
Mt4     WCGPCRFIAPILAEIAKKLP.NVTFLKVDVDELKTVAEEWAVDAMPTFLFLKEGKLVDKVVGAQK.EQLQAAI
Lj1     WCGPCRFIAPILAEIAKKLP.HVVFLKVDVDELKTVAEEWNVEAMPTFLFLKEGKAVDKVVGAKK.EELQLTI
Sros1   WCGPCRFIAPILAEIAKKLP.QVTFLKVDVDELKTVAEEWGVEAMPTFLFLKEGKLVDKVVGAKK.EELQLTI
Gm3     WCGPCRFIAPILAEFAKKLP.NVTFLKVDVDELETVSKEWGIEAMPTFLFLKEGKLVDKVVGAKK.EELQLTI
Gm9     WCGPCRFIAPILADMAKKLP.NVTFLKVDVDELATVSREWEVEAMPTFLFLKEGKLVKKLVGARK.EELQDII
Pper1   WCGPCRLIAPILAELAKKTP.EVTFLKVDVDELRTVSEEWGVEAMPTFLFLKEGKIVDKVVGAKK.DELQIKV
St2     WCGPCRFIAPILADIAKKMP.HVIFLKVDVDELKKVAEDWTVEAMPTFVFIKEGKEVDRVVGANK.DGLLQTI
Le2     WCGPCRFIAPILADIAKKMP.HVMFLKVDVDELKKVAEEWNVEAMPTFVFIKEGKEVDRVVGANK.DGLLQTI
Nt1     WCGPCRFIAPILADIAKKMP.HVIFLKVDVDELKTVSAEWSVEAMPTFVFIKDGKEVDRVVGAKK.EELQQTI
Ib1     WCGPCRMIAPILAEIAKKMT.HVIFVKVDVDELQAVAEEYKVEAMPTFVFLKDGKEVDRMVGAKK.DDLQNCI
Sr1     WCGPCRVIAPILADFAKKIP.HVTFLKVDVDELESVAQEYSVEAMPTFLFFKNGEVIDKVVGAKK.DDLHACI
Ha1     WCGPCRVIAPILADYAKKFP.HVAFLKVDVDEVESVAQEYSVEAMPTFLFFKNGVKVDTVVGAKK.DELLACI
Mc1     WCGPCRFIAPILAEVAKKTP.HVMFLKVDVDELKSVAEQYKVSAMPTFVFLKDGEEIDRVVGAKK.DELIVAI
Ph1     WCPPCKMIAPIFAELAKKFP.NVTFLKVDVDELKAVAEEWNVEAMPTFIFLKDGKLVDKTVGADK.DGLPTLV
Pb1     WCPPCKFIAPVFADLAKKFT.NVTFLKVDVDELKPVAAEWEVEAMPTFIFLKDGKLVDKIVGADK.DGLPALV
Fe1     WCGPCRVITPYVSELAKKFP.HVAFFKVDVDDLKDVAEEYKVEAMPSFVILKEGQEVERIVGARK.DELLHKI
Os1     WCGPCRFIAPVFAEYAKKFP.GAVFLKVDVDELKEVAEKYNVEAMPTFLFIKDGAEADKVVGARK.DDLQNTI
Ta1     WCGPCRVIAPVFAEYAKKFP.GAIFLKVDVDELKDVAEAYNVEAMPTFLFIKDGAKVDTVVGGRK.DDIHTKI
Zm1     WCGPCRAIAPLFVEHAKKFT.QVVFLKVDVDEVKEVTAAYEVEAMPTFHFVKNGKTVATIVGARK.DELLAQI
Sb1     WCGPCRAIAPLFVEHAKKYT.QAVFLKVDVDELKEVTAEYKIEAMPTFHFIKNGETVETIVGARK.DELLALI
At4     WCPPCRMIAPIFNDLAKKFMSSAIFFKVDVDELQSVAKEFGVEAMPTFVFIKAGEVVDKLVGANK.EDLQAKI
Bn2     WCPPCRMIAPVFADLAKKFMSSAIFFKVDVDELQNVAQEFGVEAMPTFVLIKDGNVVDKVVGARK.EDLHATI
Zm2     WCPPCRAMAPIFADMAKKSP.NVVFLKVDVDEMKTIAEQFSVEAMPTFLFMREGDVKDRVVGAAK.EELARKL
```

```
Sb2       WCPPCRMIAPVFAELAKKHP.NVVFLKVDVDEMKTIAEQFSVEAMPTFLFMREGDVKDRVVGAAK.EELANKL
Os2       WCGPCRIIAPVFADLAKKHT.NAVFLKVDVDELKPIAEQFSVEAMPTFLFMKEGDVKDRVVGAMK.DELASKV
Ta2       WCGPCRIMAPVFADLAKKFP.AAVFLKVDVDELKPIAEQFSVEAMPTFLFMKEGDVKDRVVGAIK.EELTTKV
Bn1       WCPPCRFIAPIFVELAKKHL.DVVFFKVDVDELATVAQEFDVQAMPTFVYMKGEEKLDKVVGAAK.EEIEAKL
Ts1       WCPPCRFIAPIFEEYAKKFL.NVVFFKVDVDKLSDVAKAFEVEAMPTFIFMREEAILDKIVGAAK.DEIHAKL
At3       WCPPCRFIAPVFADLAKKHL.DVVFFKVDVDELNTVAEEFKVQAMPTFIFMKEGEIKETVVGAAK.EEIIANL
At5       WCPPCRFIAPVFAEMAKKFT.NVVFFKIDVDELQAVAQEFKVEAMPTFVFMKEGNIIDRVVGAAK.DEINEKL
Mc2       WCGPCRLVAPIFEELAKKFT.ETIFLKVDVDELGSVAEEFKVKAMPTFVYLKEGREVERIVGAVKKDDMIKAI
Sb3       WCGPCRFIEPAFKELASRFT.DAIFVKIDVDELAEVARTWKVEAMPTFVLVKDGKEVSRVIGAKK.DELERKI
Zm3       WCGPCRFIEPAFKELASRFS.DAIFVKVDVDELAEVARTWKVEAMPTFVLVKDGKEVSRVVGAKK.DELERKI
Ta3       WCGPCRFIEPAFKEMASRFA.DALFVKIDVDELAEVAKTFRVEAMPTFVLVKGGQEVSRVVGAKK.DELDRKI
Os3       WCGPCRFIEPAFKDMAGHFA.DAVFFKIDVDELSEVARQWKVEAMPTFVLIKGGKEVSRVVGAKK.DELERKV
St3       WCGPCKNMDPIINDFASKYT.DVEFVKIDVDELVDVALEYEVQAMPTFTLMKRGKVVDKIVGADK.DGLKMKI
Le1       WCGPCRNMDPIINDFAAKYT.NVEFVKIDVDELVDVAEKYGVQAMPTFVLMKKGEVVDQIVGADK.DGLKMKI
Le5       WCGPCKYMEPILNDFAAKYI.DVEFVKIDVDELDDVAQEYGVQAMPTFVLIKKGKVVDKVVGADK.DGLKMKI
Mt1       WCGPCKYMDPIIQDFAAKYI.KVDFIKIDVDELMSISQEFQVQAMPTFILMKKGKVVDKVVGAKK.EELEKLI
Mt5       WCGPCKYMDPIIKELAAKYK.DVEFIKIDVDELMDVASAFQVQAMPTFILLKKGKVVEKVVGAKK.EQLQKLI
Gm6       WCGPCKLMDPVIQEFATKYR.DVEFVKIDVDELMEVSQHYQVQGMPTFMLIKKGNVADKVVGVRK.DELQRLI
St4       WCGPCKFMEPAINAMASKYT.DVDFVKIDVDELSDVAKEFGVQAMPTFLLLKQGKEVERVVGAKK.DELEKKI
Le4       WCGPCKFMEPAINAMASKYT.DVDFVKIDVDELSDVAKEFGVQAMPTFLLLKQGKEVERVVGVKK.DELEKKI
Gm1       WCGPCKFIEPAIHAMSEKFT.DVDFVKIDVDELPDVAKEFNVEAMPTFVLCKKGKEVDKVVGAKK.DELEKKI
Gm2       WCGPCKFIEPAIHAMADKFN.DVDFVKIDVDELPDVAQEFQVQAMPTFVLWKKGKEVDKVVGAKK.DELEKKI
At2       WCGPCRMIEPAIHAMADKFN.DVDFVKLDVDELPDVAKEFNVTAMPTFVLVKRGKEIERIIGAKK.DELEKKV
Ga1       WCGPCKFMEPVLNAMAAKFT.DVEFVKLDVDELPDVAQEFGVQGMPTFVLLKKGKEVDRVVGAQK.NDLEKKI
Ga5       WCGPCKFMEPVLNGMAAKFT.DVDFVKLDVDELPDVAQEFGVQAMPTFVLVKQGKEVDRVVGAQK.NELEKKV
Mt2       WCGPCKMMEPILQAMANEFT.DVEFIKIDVDELSDVAQEFKVQAMPTFLLLKNGKEVDKVVGAKK.DELKNKV
Bv1       WCGPCQYMEPIIKDLSTKYG.DVEFVNIDVDELSDVASEFGVDAMPTFVFFKKGKEVERLVGANK.GELKEKI
Mc4       WCGPCQYMEPILKDFSIKYG.DVEFIKIDVDELPDVASEFKVDAMPTFIFVKKGKEIDKLVGANK.AELKDKI
Sr2       WCGPCNFIEPAVHDLAVEFS.DVDFIKIDVDELPDVAKDFEVKAMPTFVLVKKGKERERIVGVKK.DELHRMI
Pt1       WCGPCRVISPVFVELSRRFP.EIFFLKVDVDELRDVAQEWDVEAMPTFIFIKEGKAVDKVVGAKK.DELERKV
Pt2       WCGPCRIISPVFVELSKKFS.EIFFLKVDVDELRDVAQEWDVEAMPTFIFIKDGRAVDKFVGANK.AELERKV
Pm1       WCGPCRVIGPVFVELSKKFP.EIFFLKVDVDELRDVAQEWDVEAMPTFIFIKDGKAVDKVVGAKK.DDLERKV
Cj1       WCGPCRIISPVFVELSKKYT.NVVFLKVDVDELQDVTAEWEVQAMPTFIFIKDGKAIDKIVGAKK.DELERKV
Pcn       WCGPCRVIAPVYAEMSKTYP.QLMFLTIDVDDLVDFSSTWDIRATPTFFFLKNGQQIDKLVGANK.PELEKKV
Tan       WCGPCRVIAPVYAEMSKTYP.QLMFLTIDVDDLMDFSSTWDIRATPTFFFLKNGQQIDKLVGANK.PELEKKV
Osn       WCGPCRVIAPIYAEMSKTYP.QLMFLTIDVDDLMDFSSSWDIRAKPTFFFIKNEKQVDKLVGANK.PELEKKV
Zmn       WCGPCRVIAPVYAEMSKTYP.QLMFLTIDVDDLMDFSSSWDIRATPTFFFLKNGQQIDKLVGANK.PELEKKV
Sbn       WCGPCRVISPVYAEMSQTYP.QLMFLTIDVDELMEFSSSWDIRATPTFFFLKNGQQVDKLVGANK.PELEKKV
Stn       WCGPCRMISPFYCELSEKYL.SLMFLTVDVDELTEFSSSWDIKATPTFFFLKDGEQIDKLVGANK.PELQKKI
Len       WCGPCRMISPFYCELSEKYL.SLMFLTVDVDELTEFSSSWDIKATPTFFFLKDGEQIDKLVGANK.PELQKKI
Ntn       WCGPCRMIAPFYCELSEKYL.SLMFLTVDVDELTEFSSSWDIKATPTFFFLKDGEQIDKLVGANK.PELQKKI
Mtn       WCGPCKVIAPYYCEMSEKYT.SMMFLLVDVDELTDFSTSWDIKATPTFFFLKDGQQLDKLVGANK.PELEKKL
Gm10      WCGPCKMIAPYYSELSEKYT.SMMFLLVDVDDLTDFSTSWDIKATPTFFFLKDGQQLDKLVGANK.PELLKKI
Gmn       WCGPCKVIAPHYCELSVKYP.SIMFLLVDVDELADFSTSWDIKATPTFFFLKDGKEVDKLVGANK.PELEKKI
Atn       WCGPCKIVAPFFIELSEKHS.SLMFLLVDVDELSDFSSSWDIKATPTFFFLKNGQQIGKLVGANK.PELQKKV
Ptn       WCGPCKMIAPFYAELSEKYP.QLVFLKVDVDEMAEVSAEWDVRAMPTFIFIKDGKQIDKLVGLNQ.PELEKKV
Cr1       WCGPCRMIAPFYEELSEKYS.NLIFLKVDVDQLSDISAAYEVRAMPTFMFIENGAVVEKIVGADR.TKLEAMV
Cr2       WCGPCRMIAPLFEELSEKYS.NLIFLKVDVDQMNEISNAYGVRAMPTFMFIENGSVVDKIVGADK.YKLEAKV
Pp1       WCGPCRMMAPIFADLSKKFE.KLLFLKVDVDAVQEVTQEWEVRAMPTFLFIKDGKLLDKIVGANK.DELEKKC
Pp2       WCGPCRLMAPIFVELSKKYE.NIIFLKVDVDEVKDVTSQWEVRAMPTFIFIKDGKSIHKVVGANK.DELEKKC
Liv1      WCGPCRYMSPIFTELSKKHE.DIVFLKVDVDEFQEIASEWEVRAMPTFLFIKDGKIVEKIVGANK.DELESKV


Horse     KGLI.....................
Human     NELV.....................
Mouse     TEYA.....................
Chicken   KSLV.....................
St1       AKHISSTSSA...............
Le3       AKHISSTSSA...............
Nt2       AKHISSTSTA...............
Rc1       AKHMATAST................
Cs1       AKHLATASA................
Ga3       SKHMAVPSTSSA.............
Mc3       AKHAASASQ................
At1       AKHLA....................
Mp1       AKHLNTATSTA..............
Gm5       EKYVASASA................
Gm8       QKHVASASA................
Ga4       AFHSSTSVQTA..............
Ga2       TKHMATSSTSA..............
Vv1       EKHA.....................
Gm4       AKHASAVAAASSS............
Gm7       AKHVSAAAASS..............
Mt3       TKHKDATVATA..............
Mt4       TKHATTDA.................
Lj1       TKHAEAPGTSTA.............
Sros1     NKHAATASA................
```

```
Gm3       AKHAAIAAA.....
Gm9       VKLAAIDAA.....
Pper1     AKHVAAAAAASATSASATAATATATASA.....
St2       EKHGAAPAIVTA.....
Le2       EKHGAAPAVVTA.....
Nt1       VKH.AAPATVTA.....
Ib1       TKH..AAAVMTA.....
Sr1       VKHAGAASA.....
Ha1       VKHAGEAAATVSA.....
Mc1       TKHA.....
Ph1       AKHATA.....
Pb1       EKHSVYTA.....
Fe1       AVHAPITA.....
Os1       VKHVGATAASASA.....
Ta1       VALMG....SASA.....
Zm1       EKHAAPAPASASA.....
Sb1       QKHTASASA.....
At4       VKHTGVTTVVNQFEA.....
Bn2       AKHTGVATA.....
Zm2       ELHMAS.....
Sb2       QLQMAQ.....
Os2       QLHMA.....
Ta2       GLHAAQ.....
Bn1       LKHSQVAAA.....
Ts1       EKHSQAVAAA.....
At3       EKHKTVVAAA.....
At5       MKHGGLVASA.....
Mc2       EPNL.....
Sb3       QMFIMSSSSS.....
Zm3       RMFTSSSSS.....
Ta3       KTFISSS.....
Os3       NMFISSSSS.....
St3       EKHKASLY.....
Le1       EKHKGSLY.....
Le5       EKHKAMFI.....
Mt1       EKHQN.....
Mt5       EKRRV.....
Gm6       EQHR.....
St4       LKHREAPKYAA.....
Le4       LKHREAPKYAA.....
Gm1       EKHRSQS.....
Gm2       EKHRS.....
At2       SKLRA.....
Ga1       EKHRALVAAA.....
Ga5       EKNRC.....
Mt2       QKHKA.....
Bv1       TKHRV.....
Mc4       LEHRVS.....
Sr2       EKHRF.....
Pt1       TALA.....TEAALTAKA.....
Pt2       AALA.....TQEALVTQA.....
Pm1       AALAAAATTTEATLPAQA.....
Cj1       VTLS.....QA.....
Pcn       QALGDGS.....
Tan       GALGDGS.....
Osn       QALADGS.....
Zmn       LAAADAS.....
Sbn       AAIAGAS.....
Stn       TTVADTHVACEPQPQPQ.....
Len       TTVADRHVACEPQPQPQ.....
Ntn       TAIADTQVVCETQPQ.....
Mtn       VAIADSVPQISSDLFLFMFVDATAFLPSQIANS
Gm10      VVINDSLPEYKQ.....
Gmn       VVVNDVVPHKQ.....
Atn       TSIIDSVPESPQRP.....
Ptn       LNCVAMTQSA.....
Cr1       KHYSEKCASLAA.....
Cr2       KQFSEKCAGLAA.....
Pp1       NQYASQPVVATA.....
Pp2       QQFASLPSTV.....
Liv1      RF.....
```

References

Ahn S., Anderson J. A., Sorrells M. E. and Tanksley S. D. (1993) Homeologous relationships of rice, wheat and maize chromosomes. Mol. Gen. Genet. **241**: 483-490.

Akamatsu Y., Ohno T., Hirota K., Kagoshima H., Yodoi J. and Shigesada K. (1997) Redox regulation of the DNA binding activity in transcription factor PEBP2. The roles of two conserved cysteine residues. J. Biol. Chem. **272**: 14497-14500.

Altschul S. F., Madden T. L., Schaffer A. A., Zhang J., Zhang Z., Miller W. and Lipman D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**: 3389-3402.

Anderson M. A., Cornish E. C., Mau S. –L., Williams E. G., Hoggart R., Atkinson A., Bönig I., Grego B., Simpson R., Roche P. J., Haley J. D., Penschow J. D., Niall H. D., Tregear G. W., Coghlan J. P., Crawford R. J. and Clarke A. E. (1986) Cloning of a cDNA stylar glycoprotein associated with expression of self-incompatibility in *Nicotiana alata*. Nature. **321**: 38-44.

Andersson M., Holmgren A. and Spyrou G. (1996) Nk-lysin, a disulfide-containing effector peptide of T-lymphocytes, is reduced and inactivated by human thioredoxin reductase. Implication for a protective mechanism against Nk-lysin cytotoxicity. J. Biol. Chem. **271**: 10116-10120.

Araki M., Nanri H., Ejima K., Murasato Y., Fujiwara T., Nakashima Y. and Ikeda M. (1999) Antioxidant function of the mitochondrial protein Sp-22 in the cardiovascular system. J. Biol. Chem. **274**: 2271-2278.

Arnér E. S. J. and Holmgren A. (2000) Physiological functions of thioredoxin and thioredoxin reductase. Eur. J. Biochem. **267**: 6102-6109.

Atzori L., Dypbukt J. M., Sundqvist K., Cotgreave I., Edman C. C., Moldeus P. and Grafstrom R. C. (1990) Growth associated modifications of low-molecular-weight thiols and protein sulfhydryls in human bronchial fibroblasts. J. Cell Physiol. **143**: 165-171.

Baalmann E., Backhausen J. E., Rak C., Vetter S. and Scheibe R. (1995) Reductive modification and nonreductive activation of purified spinach

chloroplast NADP-dependent glyceraldehyde-3-phosphate dehydrogenase. Arch. Biochem. Biophys. **324**: 201-208.

Babbs C. F., Pham J. A. and Coolbaugh R. C. (1989) Lethal hydroxyl radical production in paraquat-treated plants. Plant Physiol. 90: 1267-1270.

Baier M. and Dietz K. J. (1997) The plant 2-Cys peroxiredoxin BAS1 is a nuclear-encoded chloroplast protein: its expressional regulation, phylogenetic origin, and implications for its specific physiological function in plants. Plant J. **12**: 179-190.

Baier M. and Dietz K. J. (1999) Protective function of chloroplast 2-cysteine peroxiredoxin in photosynthesis: evidence from transgenic *Arabidopsis*. Plant Phyiol. **119**: 1407-1414.

Balachandran S., Xiang Y., Schobert C., Thompson G. A. and Lucas W. J. (1997) Phloem sap proteins from *Curcurbita maxima* and *Ricinus communis* have the capacity to traffic cell to cell through plasmadesmata. Proc. Natl. Acad. Sci. USA **94**: 14150-14155.

Ballicora M. A., Frueauf J. B., Fu Y. B., Schürmann P. and Preiss J. (2000) Activation of the potato tuber ADP-glucose pyrophosphorylase by thioredoxin. J. Biol. Chem. **275**: 1315-1320.

Banze M. and Follmann H. (2000) Organelle-specific NADPH thioredoxin reductase in plant mitochondria. J. Plant Physiol. **156**: 126-129.

Bateman A. J. (1955) Self-incompatibility systems in Angiosperms III. Cruciferae. Heredity **9**:53-68.

Baumann U., Juttner J., Bian X. and Langridge P. (2000) Self-incompatibility in the grasses. Ann. Bot. **85 (A)**: 203-209.

Baumann U. and Juttner J. (2002) Plant thioredoxins: the multiplicity conundrum. Cell. Mol. Life Sci. **59**: 1042-1057.

Beddows A. R. (1931) Seed setting and flowering in various grasses. University College of Wales. Series H. **12**: 5-99.

Bell P. R. and Woodcock C. L. F. (1983) The diversity of green plants, 3rd edn., Edward Arnold (publishers), Ltd., London.

Berstermann A., Vogt K. and Follmann F. (1983) Plant seeds contain several thioredoxins of regular size. Eur. J. Biochem. **131**: 339-344.

Besse I., Wong J. H., Kobrehel K. and Buchanan B. B. (1996) Thiocalsin - a thioredoxin-linked, substrate-specific protease dependent on calcium. Proc. Natl. Acad. Sci. USA **93**: 3169-3175.

Birnboim H. C. and Doly J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513-1522.

Blanc G., Barakat A., Guyot R., Cooke R. and Delseny M. (2000) Extensive duplication an reshuffling in the *Arabidopsis* genome. Plant Cell **12**: 1093-1101.

Bold H. C. (1977) *The plant kingdom*. Prentice-Hall, Engelwood Cliffs, NJ.

Bower M. S., Matias D. D., Fernandes-Carvalho E., Mazzurco M., Gu T., Rothstein S. J. and Goring D. R. (1996) Two members of the thioredoxin-h family interact with the kinase domain of a *Brassica* S locus receptor kinase. Plant Cell **8**: 1641-1650.

Boyes D. C., Nasrallah M. E., Vrebalov J. and Nasrallah J. B. (1997). The self-incompatibility (*S*) haplotypes of *Brassica* contain highly divergent and rearranged sequences of ancient origin. Plant Cell **9**: 237-247.

Boutin S., Young N., Olson T., Yu Z. –H., Shoemaker R. and Vallejos C. E. (1995) Genome conservation among three legume genera detected with DNA markers. Genome **38**: 928-937.

Brandes H. K., Hartman F. C., Lu T. Y. S. and Larimer F. W. (1996) Efficient expression of the gene for spinach phosphoribulokinase in *Pichia pastoris* and utilization of the recombinant enzyme to explore the role of regulatory cysteinyl residues by site-directed mutagenesis. J. Biol. Chem. **271**: 6490-6496.

Bredemeijer G. M. and Blaas J. (1981) S-specific proteins in styles of self-incompatible *Nicotiana alata*. Theoret. Appl. Genet. **59**: 185-190.

Bréhélin C., Mouaheb N., Verdoucq L., Lancelin J. –M. and Meyer Y. (2000) Characterisation of determinants for the specificity of *Arabidopsis* thioredoxin h in yeast complementation. J. Biol. Chem. **275**: 31641-31647.

Broin M., Cuiné S., Peltier G. and Rey P. (2000) Involvement of CDSP 32, a drought-induced thioredoxin, in the response to oxidative stress in potato plants. FEBS Lett. **467**: 245-248.

Broothaerts W., Janssens G. A., Proost P. and Broekaert W. F. (1995) cDNA cloning and molecular analysis of two self-incompatibility alleles from apple. Plant Mol. Biol. **27**: 499-511.

Brugidou C., Marty I., Chartier Y. and Meyer Y. (1993) The *Nicotiana tabacum* genome encodes two cytoplasmic thioredoxin genes which are differently expressed. Mol. Gen. Genet. **238**: 285-293.

Buchanan B. B. (1991) Regulation of CO_2 assimilation in oxygenic photosynthesis: The ferredoxin/thioredoxin system. Perspective on its discovery, present status, and future development. Arch. Biochem. Biophys. **288**: 1-9.

Bunik V., Follmann H. and Bisswanger H. (1997) Activation of mitochondrial 2-oxoacid dehydrogenases by thioredoxin. Biol. Chem. **378**: 1125-1130.

Bunik V., Shoubnikova A., Loeffelhardt S., Bisswanger H., Borbe H. O. and Follmann H. (1995) Using lipoate enantiomers and thioredoxin to study the mechanism of the 2-oxoacid-dependent dihydrolipoate production by the 2-oxoacid dehydrogenase complexes. FEBS Lett. **371**: 167-170.

Busk P. K. and Pagés M. (1998) Regulation of abscisic acid-induced transcription. Plant Mol. Biol. **37**: 425-435.

Cabrillac D., Cock J. M., Dumas C. and Gaude T. (2001) The S-locus receptor kinase is inhibited by thioredoxins and activated by pollen coat proteins. Nature **410**: 220-223

Capitani G., Markovic-Housley Z., del Val G., Morris M., Jansonius J. N. and Schürmann P. (2000) Crystal structures of two functionally different thioredoxins in spinach chloroplasts. J. Mol. Biol. **302**: 135-154.

Chae H. Z., Kim I. –H., Kim K. and Rhee S. G. (1993) Cloning, sequencing and mutation of a thiol-specific antioxidant gene of *Saccharomyces cerevisiae*. J. Biol. Chem. **268**: 16815-16821.

Chae H. Z., Robison K., Poole L. B., Church G., Storz G. and Rhee S. G. (1994) Cloning and sequencing of a thiol-specific antioxidant from mammalian brain: Alkyl hydroperoxide reductase and thiol-specific antioxidant define a large family of antioxidant enzymes. Proc. Natl. Acad. Sci. USA **91**: 7017-7021.

Chandler P. M. and Robertson M. (1994) Gene expression regulated by abscisic acid and its relation to stress tolerance. Annu. Rev. Plant Physiol. Plant Mol. Biol. **45**: 113-141.

Chen C-H. and Nasrallah J. B. (1990) A new class of *S* sequences defined by a pollen recessive self-incompatibility allele of *Brassica oleracea*. Mol. Gen. Genet. **222**: 241-248.

Chiadmi M., Navaza A., Migniac-Maslow M., Jacquot J.-P. and Cherfils J. (1999) Redox signalling in the chloroplast: Structure of oxidized pea fructose-1,6-bisphosphate phosphatase. EMBO J. **18**: 6809-6815.

Cho M. J., Wong J. H., Marx C., Jiang W., Lemaux P. G. and Buchanan B. B. (1999) Overexpression of thioredoxin h leads to enhanced activity of starch debranching enzyme (pullulanase) in barley grain. Proc. Natl. Acad. Sci. USA **96**: 14641-14646.

Choi Y. O., Cheong N. E., Lee K. O., Jung B. G., Hong C. H., Jeong J. H., Chi Y. H., Kim K., Cho M. J. and Lee S. Y. (1999) Cloning and expression of a new isotype of the peroxiredoxin gene of Chinese cabbage and its comparison to a 2-Cys peroxiredoxin isolated from the same plant. Biochem. Biophys. Res. Commun. **258**: 768-771.

Clegg M. T., Cummings M. P. and Durbin M. L. (1997) The evolution of plant nuclear genes. Proc. Natl. Acad. Sci. USA **94**: 7791-7798.

Conner J. A., Tantikanjana T., Stein J. C., Kandasamy M. K., Nasrallah J. B. and Nasrallah M. E. (1997) Transgene induced silencing of S-locus genes and related genes in *Brassica*. Plant J. **11**: 809-824.

Connor H. E. (1979) Breeding systems in the grasses: a survey. New Zealand J. Bot. **17**: 547-574.

Cornish M. A., Hayward M. D. and Lawrence M. J. (1980) Self-incompatibility in ryegrass. III. The joint segregation of S and PGI-2 in *Lolium perenne* L. Heredity **44**: 55-62.

Cornish E. C., Pettitt J. M., Bönig I. And Clarke A. E. (1987) Developmentally controlled expression of a gene associated with self-incompatibility in *Nocotiana alata*. Nature **326**: 99-102.

Crawford N. A., Yee B. C., Hutcheson S. W., Wolosiuk R. A. and Buchanan B. B. (1986) Enzyme regulation in C4 photosynthesis: Purification, properties, and activities of thioredoxins from C4 and C3 plants. Arch. Biochem. Biophys. **244**: 1-15.

Dai S., Saarinen M., Ramaswamy S., Meyer Y., Jacquot J. P. and Eklund H. (1996) Crystal structure of *Arabidopsis thaliana* NADPH dependent thioredoxin reductase at 2.5 Å resolution. J. Mol. Biol. **264**: 1044-1057.

Dai S. D., Schwendtmayer C., Johansson K., Ramaswamy S., Schürmann P. and Eklund H. (2000a) How does light regulate chloroplast enzymes? Structure-function studies of the ferredoxin/thioredoxin system. Quart. Rev. Biophys. **33**: 67-108.

Dai S., Schwendtmayer C., Schürmann P., Ramaswamy S. and Eklund H. (2000b) Redox signaling in chloroplasts: Cleavage of disulfides by an iron- sulfur cluster. Science **287**: 655-658.

Decottignies P., Schmitter J. M., Dutka S., Jacquot J. –P. and Miginiac-Maslow M. (1991) Characterisation and primary structure of a second thioredoxin from the green alga *Chlamydomonas reinhardtii*. Eur. J. Biochem. **198**: 505-512.

Delorme V., Giranton J-L., Hatzfeld Y., Friry A., Heizmann P., Ariza M. J., Dumas C., Gaude T. and Cock J. M. (1995) Characterization of the *S* locus genes, *SLG* and *SRK*, of the *Brassica* S^3 haplotype: identification of a membrane-localised protein encoded by the *S* locus receptor kinase. Plant J. **7**: 429-440.

Devos K. M. and Gale M. D. (1997) Comparative genetics in the grasses. Plant Mol. Biol. **35**: 3-15.

Dixit R., Nasrallah M. E. and Nasrallah J. B. (2000) Post-transcriptional maturation of the S receptor kinase of *Brassica* correlates with co-expression of the S-locus glycoprotein in the stigmas of two *Brassica* strains and in transgenic tobacco plants. Plant Phys. **124**: 297-311.

Draculic T., Dawes I. W. and Grant C. M. (2000) A single glutaredoxin or thioredoxin gene is essential for viability in the yeast *Saccharomyces cerevisiae*. Molec. Microbiol. **36**: 1167-1174.

Droux M., Jacquot J. P., Miginac-Maslow M., Gadal P., Huet J. C., Crawford N. A., Yee B. C. and Buchanan B. B. (1987) Ferredoxin-thioredoxin reductase, an iron-sulfur enzyme linking light to enzyme regulation in oxygenic photosynthesis: Purification and properties of the enzyme from C3, C4, and cyanobacterial species. Arch. Biochem. Biophys. **252**: 426-439.

Duek P. D. and Wolosiuk R. A. (2001) Rapeseed chloroplast thioredoxin-m - modulation of the affinity for target proteins. Biochim. Biophys. Acta **1546**: 299-311.

Dunford R. P., Durrant M. C., Catley M. A. and Dyer T. A. (1998) Location of the redox-active cysteines in chloroplast sedoheptulose-1,7-bisphosphatase indicates that its allosteric regulation is similar but not identical to that of fructose-1,6-bisphosphatase. Photosyn. Res. **58**: 221-230.

Dunn M. A., White A. J., Vural S. and Hughes M. (1998) Identification of promoter elements in a low-temperature-responsive gene (*blt4.9*) from barley. Plant Mol. Biol. **38**: 551-564.

Ecker J. R. (1995) The ethylene signal transduction pathway in plants. Science **268**: 667-675.

Eklund H., Gleason F. K. and Holmgren A. (1991) Structural and functional relations among thioredoxins of different species. Proteins **11**: 13-28.

Fernando M. R., Nanri H., Yoshitake S., Nagata-Kuno K. and Minakami S. (1992) Thioredoxin regenerates proteins inactivated by oxidative stress in endothelial cells. Eur. J. Biochem. **209**: 917-922.

Fisher D. B., Wu Y. and Ku M. S. B. (1992) Turnover of soluble proteins in the wheat sieve tube. Plant Physiol. **100**: 1433-1441.

Follmann H. and Häberlein I. (1995) Thioredoxins: Universal, yet specific thiol-disulfide redox cofactors. Biofactors **5**: 147-156.

Foote H. C. C., Ride J. P., Franklin-Tong V. E., Walker E. A., Lawrence M. J. and Franklin F. C. H. (1994) Cloning and expression of a distinctive class of self-incompatibility (*S*) gene from *Papaver rhoeas* L. Proc. Natl. Acad. Sci. USA **91**: 2265-2269.

Franklin-Tong V. E. (1999a) Signaling and the modulation of pollen tube growth. Plant Cell. **11**: 727-738.

Franklin-Tong V. E. (1999b) Signaling in pollination. Curr. Opin. Plant Biol. **2**: 490-495.

Franklin-Tong V. E. and Franklin F. C. H. (1992) Gametophytic self-incompatibility in *Papaver rhoeas* L. Sex. Plant Reprod. **5**: 1-7.

Franklin-Tong V. E., Lawrence M. J. and Franklin F. C. H. (1988) An *in vitro* bioassay for the stigmatic product of the self-incompatibility gene in *Papaver rhoeas* L. New Phytol. **110**: 109-118.

Franklin-Tong V. E., Lawrence M. J. and Franklin F. C. H. (1989) Characterization of a stigmatic component from *Papaver rhoeas* L. which exhibits the specific activity of a self-incompatibility (S-) gene product. New Phytol. **112**: 307-315.

Franklin-Tong V. E., Lawrence M. J. and Franklin F. C. H. (1990) Self-incompatibility in *Papaver rhoeas* L.: inhibition of incompatible pollen is dependent on pollen gene expression. New Phytol. **116**: 319-324.

Franklin-Tong V. E., Lawrence M. J. and Franklin F. C. H. (1991) Self-incompatibility in *Papaver rhoeas* L.: there is no evidence for the involvement of stigmatic ribonuclease activity. Plant Cell Environ. **14**: 423-429.

Franklin-Tong V. E., Ride J. P., Read N. D., Trewavas A. J. and Franklin F. C. H. (1993) The self-incompatibility response in *Papaver rhoeas* is mediated by cytosolic free calcium. Plant J. **4**: 163-177.

Fujimoto S. Y., Ohta M., Usui A., Shinshi H. and Ohme-Takagi M. (2000) *Arabidopsis* ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. Plant Cell **12**: 393-404.

Gale M. D. and Devos K. M. (1998) Comparative genetics in the grasses. Proc. Natl. Acad. Sci. USA. **95**: 1971-1974.

Gallardo-Madueno, R., Leal J. F. M., Dorado G., Holmgren A., Lopez-Barea J. and Pueyo C. (1998) *In vivo* transcription of *nrdAB* operon and of *grxA* and *fgp* genes is triggered in *Eschericia coli* lacking both thioredoxin and glutaredoxin 1 or thioredoxin and glutathione, respectively. J. biol. Chem. **273**: 18382-18388.

Gandolfo M., Nixon K. and Crepet W. (1998) A new fossil flower from the Turonian of New Jersey: *Dressiantha bicarpellata* gen. Et sp. Nov. (Capparales). Am. J. Bot. **85**: 964-974.

Gautier M. –F., Lullien-Pellerin V., Lamotte-Guery F., Guirao A. and Joudrier P. (1998) Characterisation of wheat thioredoxin *h* cDNA and production of an active *Triticum aestivum* protein in *Escherichia coli*. Eur. J. Biochem. **252**: 314-324.

Geck M. K. and Hartman F. C. (2000) Kinetic and mutational analyses of the regulation of phosphoribulokinase by thioredoxins. J. Biol. Chem. **275**: 18034-18039.

Geck R. K., Larimer F. W. and Hartman F. C. (1996) Identification of residues of spinach thioredoxin f that influence interactions with target enzymes. J. Biol. Chem. **271**: 24736-24740.

Geitmann A., Snowman B. N., Emons A-M. C. and Franklin-Tong V. E. (2000) Alterations in the actin cytoskeleton of pollen tubes are induced by the self-incompatibility reaction in *Papaver rhoeas*. Plant Cell. **12**: 1239-1251.

Genot G., Wintz H., Houlné G. and Jamet E. (2001) Molecular characterisation of a bean chloroplastic 2-Cys peroxiredoxin. Plant Physiol. Biochem. **39**: 449-459.

Giuliano G., Pichersky E., Malik V. S., Timko M. P., Scolnik P. A. and Cashmore A. R. (1988) An evolutionary conserved protein binding sequence upstream of a plant light-regulated gene. Proc. Natl. Acad. Sci. USA **85**: 7089-7093.

Goff S. A., Ricke D., Lan T. H., Presting G., Wang R. L., Dunn M., Glazebrook J., Sessions A., Oeller P., Varma H., Hadley D., Hutchinson D., Martin C., Katagiri F., Lange B. M., Moughamer T., Xia Y., Budworth P., Zhong J. P., Miguel T., Paszkowski U., Zhang S. P., Colbert M., Sun W. L., Chen L. L., et al. (2002) A draft sequence of the rice genome (Oryza sativa L. ssp japonica). Science. **296**:92-100.

Goring D. R. and Rothstein S. J. (1992) The *S*-locus receptor kinase gene in a self-incompatible *Brassica napus* line encodes a functional serine/threonine kinase. Plant Cell **4**: 1273-1281.

Goring D. R., Glavin T. L., Schafer U. and Rothstein S. J. (1993) An *S* receptor kinase gene in self-compatible *Brassica napus* has a 1-bp deletion. Plant Cell **5**: 531-539.

Goyer A., Haslekås C., Miginiac-Maslow M., Klein U., Le Marechal P., Jacquot J. –P. and Decottignies P. (2002) Isolation and characterization of a thioredoxin-dependent peroxidase from *Chlamydomonas reinhardtii*. Eur. J. Biochem. **269**: 272-282.

Grant C. M. (2001) Role of the glutathione/glutaredoxin and thioredoxin systems in yeast growth and response to stress conditions. Mol. Microbiol. **39**: 533-541.

Grant D., Cregan P. and Shoemaker R. C. (2000) Genome organisation in dicots: Genome duplication in *Arabidopsis* and synteny between soybean and *Arabidopsis*. Proc. Natl. Acad. Sci. USA **97**: 4168-4173.

Gray J. E., McClure B. A., Bönig I., Anderson M. A. and Clarke A. E. (1991) Action of the style product of the self-incompatibility gene of *Nicotiana alata* (S-RNase) on *in-vitro* grown pollen tubes. Plant Cell **3**: 271-283.

Gu T., Mazzurco M., Sulaman W., Matias D. D. and Goring D. R. (1998) Binding of an arm repeat protein to the kinase domain of the *S*-locus receptor kinase. Proc. Natl. Acad. Sci. USA **95**: 382-387.

Häberlein I. (1991) Separation of the complete thioredoxin pattern of soybean leaves (*Glycine max*) by high-performance anion exchange chromatography on mono Q. J. Chromatography **587**: 109-115.

Häberlein I., Wolf M., Mohr L. and Follmann H. (1995) Differentiation of six distinct thioredoxins in seeds of the soybean. J. Plant Physiol. **146**: 385-392.

Hartman H., Syvanen M. and Buchanan B. B. (1990) Contrasting evolutionary histories of chloroplast thioredoxins f and m. Mol. Biol. Evol. **7**: 247-254.

Hashimoto S., Matsumoto K., Gon Y., Furuichi S., Maruoka S., Takeshita I., Hirota K., Yodie J. and Horie T. (1999) Thioredoxin negatively regulates p38 MAP kinase activation and IL-6 production by tumour necrosis factor-α. Biochem. Biophys. Res. Commun. **258**: 443-447

Hayman D. L. (1956) The genetic control of incompatibility in *Phalaris coerulescens* DESF. Aust. J. Biol. Sci. **9**: 321-331.

Hayman D. L. (1992) The S-Z incompatibility system. In *Grass Evolution and Domestication*, G. P. Chapman (ed.) Cambridge University Press, 117-137.

Hillis D. M., Allard M. W. and Miyamoto M. M. (1993) Analysis of DNA sequence data: Phylogenetic inference. Methods Enzym. **224**: 456-487.

Hillis D. M. and Huelsenbeck J. P. (1992) Signal, noise and reliability in molecular phylogenetic analyses. J. Hered. **83**: 189-195.

Hirasawa M., Brandes H. K., Hartman F. C. and Knaff D. B. (1998) Oxidation-reduction properties of the regulatory site of spinach phosphoribulokinase. Arch. Biochem. Biophys. **350**: 127-131.

Hirasawa M., Ruelland E., Schepens I., Issakidis-Bourguet E., Miginiac-Maslow M. and Knaff D. B. (2000) Oxidation-reduction properties of the regulatory disulfides of sorghum chloroplast nicotinamide adenine dinucleotide phosphate-malate dehydrogenase. Biochemistry **39**: 3344-3350.

Hirota K., Matsui M., Iwata S., Nishiyama A., Mori K. and Yodoi J. (1997) AP-1 transcriptional activity is regulated by a direct association between thioredoxin and Ref-1. Proc. Natl. Acad. Sci. USA **94**: 3633-3638.

Hirota K., Murata M., Sachi Y., Nakamura H., Takeuchi J., Mori K. and Yodoi J. (1999) Distinct roles of thioredoxin in the cytoplasm and in the nucleus - a two-step mechanism of redox regulation of transcription factor NF-kappa B. J. Biol. Chem. **274**: 27891-27897.

Hoagland D.R. and Arnon D.I. (1950). The water-culture method for growing plants without soil. Circ. 347. Univ. of Calif. Agric. Exp. Station, Berkley.

Holmgren A. (1977) Bovine thioredoxin system. Purification of thioredoxin reductase from calf liver and thymus and studies of its function in disulfide reduction. J. Biol. Chem. **252**: 4600-4606.

Holmgren A. (1979) Reduction of disulfides by thioredoxin. J. Biol. Chem. **254**: 9113-9119.

Holmgren A. (1979) Thioredoxin catalyzes the reduction of insulin disulfides by dithiothreitol and dihydrolipoamide. J. Biol. Chem. **254**: 9627-9632.

Holmgren A. (1989) Thioredoxin and glutaredoxin systems. J. Biol. Chem. **264**: 13963-13966.

Holmgren A. (1995) Thioredoxin structure and mechanism: Conformational changes on oxidation of the active-site sulfhydryls to a disulfide. Structure **3**: 239-243.

Huppe H. C., de Lamotte F., Jacquot J.P., Buchanan B.B. (1990) The ferredoxin-thioredoxin system of a green alga, *Chlamydomonas reinhardtii*. Identification and characterization of thioredoxins and ferredoxin-thioredoxin reductase components. Planta **186**: 341-351.

Ioerger T. R., Gohlke J. R., Xu B. and Kao T-H. (1991) Primary structural features of the self-incompatibility protein in Solanaceae. Sex. Plant Reprod. **4**: 81-87.

Ishimizu T., Shinkawa T., Sakiyama F. and Norioka S. (1998) Primary structural features of rosaceous S RNases associated with gametophytic self-incompatibility. Plant Mol. Biol. **37**: 931-941.

Ishiwatari Y., Fujiwara T., McFarland K. C., Nemoto K., Hayashi H., Chino M. and Lucas W. J. (1998) Rice phloem thioredoxin h has the capacity to mediate its own cell-to-cell transport through plasmodesmata. Planta **205**: 12-22.

Ishiwatari Y., Honda C., Kawashima I., Nakamura S., Hirano H., Mori S., Fujiwara, T, Hayashi, H. and Chino M. (1995) Thioredoxin h is one of the major proteins in rice phloem sap. Planta **195**: 456-463.

Issakidis-Bourguet E., Mouaheb N., Meyer Y. and Miginiac-Maslow M. (2001) Heterologous complementation of yeast reveals a new putative function for chloroplast *m*-type thioredoxin. Plant J. **25**: 127-135.

Izawa S., Maeda K., Sugiyama K. –I., Mano J., Inoue Y. and Kimura A. (1999) Thioredoxin deficiency causes the constitutive activation of Yap1, an AP-1-like transcription factor in *Saccharomyces cerevisiae.* J Biol. Chem. **274**: 28459-28465.

Izawa T., Foster R. and Chua N.-H. (1993) Plant bZIP protein DNA binding specificity. J. Mol. Biol. **230**: 1131-1144.

Jacquot J. P., Lancelin J. M. and Meyer Y. (1997) Tansley review no 94 - thioredoxins - structure and function in plant cells. New Phytol. **136**: 543-570.

Jacquot J. P., López-Jaramillo J., Miginiac-Maslow M., Lemaire S., Cherfils J., Chueca A. and Lopez-Gorge J. (1997) Cysteine-153 is required for redox regulation of pea chloroplast fructose-1,6-bisphosphatase. FEBS Lett. **401**: 143-147.

Jacquot J. P., Rivera-Madrid R., Marinho P., Kollarova M., Le Marechal P., Miginiac-Maslow M. and Meyer Y. (1994) *Arabidopsis thaliana* NADPH thioredoxin reductase. cDNA characterization and expression of the recombinant protein in *Escherichia coli*. J. Mol. Biol. **235**: 1357-1363.

Jacquot J. P., Vidal J. and Gadal P. (1976) Identification of a protein factor involved in dithiothreitol activation of NADP malate dehydrogenase from french bean leaves. FEBS Lett. **71**: 223-227.

Jacquot J. P., Vidal J., Gadal P. and Schürmann P. (1978) Evidence for the existence of several enzyme-specific thioredoxins in plants. FEBS Lett. **96**: 243-246.

Jahnen W., Lush W. M. and Clarke A. E. (1989) Inhibition of *in vitro* pollen tube growth by isolated S-glycoproteins of *Nicotiana alata*. Plant Cell **1**: 493-499.

Jeng M. F., Campbell A. P., Begley T., Holmgren A., Case D. A., Wright P. E. and Dyson H. J. (1994) High-resolution solution structures of oxidized and reduced *Escherichia coli* thioredoxin. Structure **2**: 853-868.

Jeong J. S., Kwon S. J., Kang S. W., Rhee S. G. and Kim K. (1999) Purification and characterization of a second type of thioredoxin peroxidase (type II Tpx) from *Saccharomyces cerevisiae*. Biochem. **38**: 776-783.

Jiao J., Yee B. C., Kobrehel K. and Buchanan B. B. (1992) Effect of thioredoxin-linked reduction on the activity and stability of the Kunitz and Bowman-Birk soybean trypsin inhibitor proteins. J. Agric. Food Chem. **40**: 2333-2336.

Johansson K., Ramaswamy S., Saarinen M., Lemaire-Chamley M., Issakidis-Bourguet E., Miginiac-Maslow M. and Eklund H. (1999) Structural basis for light activation of a chloroplast enzyme: The structure of sorghum NADP-malate dehydrogenase in its oxidized form. Biochemistry **38**: 4319-4326.

Jordan N. D., Franklin F. C. H. and Franklin-Tong V. E. (2000). Evidence for DNA fragmentation triggered in the self-incompatibility response in pollen of *Papaver rhoeas*. Plant J. **23**: 471-479.

Juttner J., Olde D., Langridge P. and Baumann U. (2000) Cloning and expression of a distinct subclass of plant thioredoxins. Eur. J. Biochem. **267**: 7109-7117.

Kang S. W., Chae H. Z., Seo M. S., Kim K., Baines I. C. and Rhee S. G. (1998) Mammalian peroxiredoxin isoforms can reduce hydrogen peroxide generated in response to growth factors and tumour necrosis factor-α. J. Biol. Chem. **273**: 6297-6302.

Kao T-H. and McCubbin A. G. (1996) How flowering plants discriminate between self and non-self pollen to prevent inbreeding. Proc. Natl. Acad. Sci. USA **93**: 12059-12065.

Kao T-H. and McCubbin A. G. (2000) A social stigma. Nature **403**: 840-841.

Kawata Y., Sakiyama F. and Tamaoki H. (1988) Amino-acid sequence of ribonuclease T2 from *Aspergillus oryzae*. Eur. J. Biochem. **176**: 683-697.

Kim H. –J., Kim Y. –K., Park J. –Y. and Kim J. (2002) Light signalling mediated by phytochrome plays an important role in cold-induced gene expression through the C-repeat/dehydration responsive element (C/DRE) in *Arabidopsis thaliana*.

Kobrehel K., Wong J. H., Balogh A., Kiss F., Yee B. C. and Buchanan B. B. (1992) Specific reduction of wheat storage proteins by thioredoxin h. Plant Physiol. **99**: 919-924.

Kobrehel K., Yee B. C. and Buchanan B. B. (1991) Role of the NADP/thioredoxin system in the reduction of alpha-amylase and trypsin inhibitor proteins. J. Biol. Chem. **266**: 16135-16140.

Konrad A., Banze M. and Follmann F. (1996) Mitochondria of plant leaves contain two thioredoxins - completion of the thioredoxin profile of higher plants. J. Plant Physiol. **149**: 317-321.

Kozaki A., Kamada K., Nagano Y., Iguchi H. and Sasaki Y. (2000) Recombinant carboxyltransferase responsive to redox of pea plastidic acetyl-CoA carboxylase. J. Biol. Chem. **275**: 10702-10708.

Ku H. –M., Vision T., Liu J. and Tanksley S. D. (2000) Comparing sequence segments of the tomato and *Arabidopsis* genomes: large scale duplication followed by selective gene loss creates a network of syteny. Proc. Natl. Acad. Sci. USA **97**: 9121-9126.

Kuge S. and Jones N (1994) *YAP1* dependent activation of *TRX2* is essential to the response of *Saccharomyces cerevisiae* to oxidative stress by hydroperoxides. EMBO J. **13**: 655-664.

Kumar V. and Trick M. (1994) Expression of the *S*-locus receptor kinase multigene family in *Brassica oleracea*. Plant J. **6**: 807-813.

Kurup S., Ride J. P., Jordan N. D., Fletcher G., Franklin-Tong V. E. and Franklin F. C. H. (1998) Identification and cloning of related self-incompatibility S genes in *Papaver rhoeas* and *Papaver nudicaule*. Sex. Plant Reprod. **11**: 192-198.

Kusaba M., Nishio T., Satta Y., Hinata K. and Ockendon D. (1997) Striking sequence similarity in inter- and intra-specific comparisons of class I SLG alleles from *Brassica oleracea* and *Brassica campestris*: Implications for the evolution and recognition mechanism. Proc. Natl. Acad. Sci. USA **94**: 7673-7678.

Laloi C., Rayapuram N., Chartier Y., Grienenberger J. –M., Bonnard G. and Meyer Y. (2001) Identification and characterisation of a mitochondrial thioredoxin system in plants. Proc. Natl. Acad. Sci. USA **98**: 14144-14149.

Langlotz P., Wagner W. and Follmann F. (1986) Green algae (*Scendesmus obliquus*) contain three thioredoxins of regular size. Z. Naturforsch. **41**: 979-987.

Larsen K. (1977) Self incompatibility in *Beta vulgaris* L. I. Four gametophytic, complementary *S*-loci in sugar beet. Hereiditas **85**: 93-100.

Lawrence M. J., Afzal M. and Kenrick J. (1978) The genetical control of self-incompatibility in *Papaver rhoeas* L. Heredity **40**: 239-285.

Leach C. R. and Hayman D. L. (1987) The incompatibility loci as indicators of conserved linkage groups in the Poaceae. Heredity **58**: 303-305.

Lee H. S., Huang S. and Kao T-H. (1994) S proteins control rejection of incompatible pollen in *Petunia inflata*. Nature **367**: 560-563.

Lee J. M., Grant D., Vallejos c. E. and Shoemaker R. C. (2001) Genome organisation in dicots. II. *Arabidopsis* as a 'bridging species' to resolve genome evolution events among legumes. Theor. Appl. Genet. **103**: 765-773.

Lee J. S. and Verma D. P. S. (1984) Structure and chromosomal arrangement of leghemoglobin genes in kidney bean suggest divergence in soybean leghemoglobin gene loci following tetraploidisation. EMBO J. **3**: 2745-2752.

Lee S. R., Kim J. R., Kwon K. S., Yoon H. W., Levine R. L., Ginsburg A. and Rhee S. G. (1999) Molecular cloning and characterization of a mitochondrial selenocysteine-containing thioredoxin reductase from rat liver. J. Biol. Chem. **274**: 4722-4734.

Lee J., Spector D., Godon C., Labarre J. and Toledano M. B. (1999) A new antioxidant with alkyl hydroperoxide defense properties in yeast. J. Biol. Chem. **274**: 4537-4544.

Lennon B. W., Williams C. H., Jr. and Ludwig M. L. (2000) Twists in catalysis: Alternating conformations of *Escherichia coli* thioredoxin reductase. Science **289**: 1190-1194.

Lepiniec L., Hodges M., Gadal P. and Cretin C. (1992) Isolation, characterization and nucleotide sequence of a full-length pea cDNA encoding thioredoxin-f. Plant Mol. Biol. **18**: 1023-1025.

Lewis M. L., Miki K. and Ueda T. (2000) FePer 1, a gene encoding an evolutionarily conserved 1-Cys peroxiredoxin in buckwheat (*Fagopyrum esculentum Moench*) is expressed in a seed-specific manner and induced during seed germination. Gene **246**: 81-91.

Li X., Nield J., Hayman D. and Langridge P. (1994) Cloning a putative self-incompatibility gene from the pollen of the grass *Phalaris coerulescens*. Plant Cell **6**: 1923-1932.

Li X., Nield J., Hayman D. and Langridge P. (1995) Thioredoxin activity in the C terminus of *Phalaris* S protein. Plant J. **7**: 133-138.

Li X., Paech N., Nield J., Hayman D. and Langridge P. (1997) Self-incompatibility in the grasses: evolutionary relationship of the *S* gene from *Phalaris coerulescens* to homologous sequences in other grasses. Plant Mol. Biol. **34**: 223-232.

Liu Q., Kasuga M., Sakuma Y., Abe H., Miura S., Yamaguchi-Shinozaki K. and Shinozaki K. (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain, separate two cellular signal transduction pathways in drought- and low temperature-responsive gene expression, respectively in *Arabidopsis*. Plant Cell **10**: 1391-1406.

López Jaramillo J., Chueca A., Jacquot J. P., Hermoso R., Lázaro J. J., Sahrawy M. and Gorge J. L. (1997) High-yield expression of pea thioredoxin m and assessment of its efficiency in chloroplast fructose-1,6-bisphosphatase activation. Plant Physiol. **114**: 1169-1175.

Lozano R. M., Wong J. H., Yee B. C., Peters A., Kobrehel K. and Buchanan B. B. (1996) New evidence for a role for thioredoxin h in germination and seedling development. Planta **200**: 100-106.

Lundqvist A. (1954) Studies on self-sterility in rye, *Secale cereale* L. Hereditas **40**: 278-294.

Marcus F., Chamberlain S. H., Chu C., Masiarz F. R., Shin S., Yee B. C. and Buchanan B. B. (1991) Plant thioredoxin h: An animal-like thioredoxin occurring in multiple cell compartments. Arch. Biochem. Biophys. **287**: 195-198.

Martin C. and Paz-Ares J. (1997) MYB transcription factors in plants. Trends Genet. **13**: 67-73.

Martin J. L. (1995) Thioredoxin-a fold for all reasons. Structure **3**: 245-250.

Marty I., Brugidou C., Chartier Y. and Meyer Y. (1993) Growth-related gene expression in *Nicotiana tabacum* mesophyll protoplasts. Plant J. **4**: 265-278.

Matton D. P., Luu D. T., Xike Q., Laublin G., O'Brien M., Maes O., Morse D. and Cappadocia M. (1999) Production of an S RNase with dual specificity suggests a novel hypothesis for the generation of new *S* slleles. Plant Cell **11**: 2087-2097.

Matton D. P., Maes O., Laublin G., Xike Q., Bertrand C., Morse D. and Cappadocia M. (1997) Hypervariable domains of self-incompatibility RNases mediate allele-specific pollen recognition. Plant Cell **9**: 1757-1766.

May J. M., Mendiratta S., Hill K. E. and Burk R. F. (1997) Reduction of dehydroascorbate to ascorbate by the selenoenzyme thioredoxin reductase. J. Biol. Chem. **272**: 22607-22610.

Mazzurco M., Sulaman W., Elina H., Cock J. M. and Goring D. R. (2001) Further analysis of the interactions between the *Brassica* S receptor kinase and three interacting proteins (Arc1, Thl1 and Thl2) in the yeast two-hybrid system. Plant Mol. Biol. **45**: 365-376.

McClure B. A., Gray J. E., Anderson M. A. and Clarke A. E. (1990) Self-incompatibility in *Nicotiana alata* involves degredation of pollen rRNA. Nature **347**: 757-760.

McClure B. A., Haring V., Ebert P. E., Anderson M. A., Simpson R. J., Sakiyama F. and Clarke A. E. (1989) Style self-incompatibility gene products of *Nicotiana alata* are ribonucleases. Nature. **342**: 955-957.

McCubbin A. G. and Kao T. –H. (1999) The emerging complexity of self-incompatibility (S-) loci. Sex. Plant. Reprod. **12**: 1-5.

Mena M., Vincente-Carbajosa J., Schmidt R. and Carbonero P. (1998) An endosperm-specific Dof protein from barley, highly conserved in wheat, binds to and activates transcription from the prolamin-box of a native B-hordein promoter in barley endosperm. Plant J. **16**: 53-62.

Menancio-Hautea D., Fatokun C., Kumar L., Danesh D. and Young N. D. (1993) Comparative genome analysis of mungbean (*Vigna radiata* l. Wilczek) and cowpea (*V. unguiculata* L. Walpers) using RFLP mapping data. Theor. Appl. Genet. **86**:797-801.

Mestres-Ortega D. and Meyer Y. (1999) The *Arabidopsis thaliana* genome encodes at least four thioredoxins m and a new prokaryotic-like thioredoxin. Gene **240**: 307-316.

Meyer Y., Verdoucq L. and Vignols F. (1999) Plant thioredoxins and glutaredoxins: identity and putative roles. Trends Plant Sci. **4**: 388-394.

Meyers B. C., Dickerman A. W., Michelmore R. W., Sivaramakrishnan S., Sobral B. W. and Young N. D. (1999) Plant disease resistance genes encode members

of an ancient and diverse protein gene family within the nucleotide-binding superfamily. Plant J. **20**: 317-332.

Minakuchi K., Yabushita T., Masumura T., Ichihara K. and Tanaka K. (1994) Closing and sequence analysis of a cDNA encoding rice glutaredoxin. FEBS Lett. **337**: 157-160.

Miranda-Vizuete A., Damdimopoulos A. E., Pedrajas J. R., Gustafsson J. A. and Spyrou G. (1999) Human mitochondrial thioredoxin reductase - cDNA cloning, expression and genomic organization. Eur. J. Biochem. **261**: 405-412.

Mittard V., Blackledge M. J., Stein M., Jacquot J. P., Marion D. and Lancelin J. M. (1997) NMR solution structure of an oxidised thioredoxin h from the eukaryotic green alga *Chlamydomonas reinhardtii*. Eur. J. Biochem. **243**: 374-383.

Moore G., Devos K. M., Wang Z. and Gale M. D. (1995) Grasses, line up and form a circle. Current Biol. **5**: 737-739.

Mouaheb N., Thomas D., Verdoucq L., Monfort P. and Meyer Y. (1998) *In vivo* functional discrimination between plant thioredoxins by heterologous expression in the yeast *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **95**: 3312-3317.

Muller E. G. D. (1996) A glutathione reductase mutant of yeast accumulates high levels of oxidised glutathione and requires thioredoxin for growth. Mol. Biol. Cell **7**: 1805-1813.

Murfett J., Atherton T. L., Mou B., Gasser C. S. and McClure B. A. (1994) S-RNase expressed in transgenic *Nicotiana* causes S-allele-specific pollen rejection. Nature **367**: 563-566.

Nasrallah J. B. (1997). Evolution of the *Brassica* self-incompatibility locus: A look into *S* locus gene polymorphisms. Proc. Natl. Acad. Sci. USA **94**: 9516-9519.

Nasrallah J. B., Kao T-H., Chen C-H., Goldberg M. L. and Nasrallah M. E. (1987) Amino-acid sequence of glycoproteins encoded by three alleles of the *S*-locus of *Brassica oleracea*. Nature **326**: 617-619.

Nasrallah J. B., Kao T-H., Goldberg M. L. and Nasrallah M. E. (1985) A cDNA clone encoding an *S*-locus specific glycoprotein from *Brassica oleracea*. Nature **318**: 263-267.

Nasrallah J. B., Rundle S. J. and Nasrallah M. E. (1994) Genetic evidence for the requirement of the *Brassica* *S*-locus receptor kinase gene in the self-incompatibility response. Plant J. **5**: 373-384.

Nasrallah J. B., Yu S-M. and Nasrallah M. E. (1988) Self-incompatibility genes of *Brassica oleracea*: expression, isolation and structure. Proc. Natl. Acad. Sci. USA **85**: 5551-5555.

Nasrallah M. E., Kandasamy M. K. and Nasrallah J. B. (1992) A genetically defined *trans*-acting locus regulates *S*-locus function in *Brassica*. Plant J. **2**: 497-506.

Nielsen N. C., Dickinson C. D., Cho T. –J., Thanh V. C., Scallon B. J., Fischer R. L., Sims T. L., Drews G. N. and Goldberg R. B. (1989) Characterisation of the glycinin gene family in soybean. Plant Cell **1**: 313-328.

O'Donnell P. J., Calvet C., Atzorn R., Wasternack C., Leyser H. M. O. and Bowles D. J. (1996) Ethylene as a signal mediating the wound response of tomato plants. Science **274**: 1914-1917.

Ohme-Takagi M. and Shinshi H. (1995) Ethylene-iducible DNA binding proteins that interact with an ethylene-responsive element. Plant Cell **7**: 173-182.

Paxman G. J. (1963) The maximum likelihood estimation of the number of self-sterility alleles in a population. Genetics **48**: 1029-1032.

Pedrajas J. R., Kosmidou E., Miranda-Vizuete A., Gustafsson J. A., Wright A. P. H. and Spyrou G. (1999) Identification and functional characterization of a novel mitochondrial thioredoxin system in *Saccharomyces cerevisiae*. J. Biol. Chem. **274**: 6366-6373.

Penninckx I. A. M. A., Thomma B. P. H. J., Buchala A., Metraux J. –P. and Broekaert W. F. (1998) Concomitant activation of jasmonate and ethylene respons pathways is required for induction of a plant defensin gene in *Arabidopsis*. Plant Cell **10**: 2103-2113.

Post G. B., Keller D. A., Connor K. A. and Menzel D. B. (1983) Effects of culture conditions on glutathione content in A549 cells. Biochem. Biophys. Res. Commun. **114**: 737-742.

Powis G. and Montfort W. R. (2001) Properties and biological activities of thioredoxins. Ann. Rev. Pharmacol. Toxicol. **41**: 261-295.

Qin J., Clore G. M. and Gronenborn A. M. (1994) The high-resolution three-dimensional solution structures of the oxidized and reduced states of human thioredoxin. Structure **2**: 503-522.

Qin J., Clore G. M., Kennedy W. M., Huth J. R. and Groneborn A. M. (1995) Solution structure of human thioredoxin in a mixed disulfide intermediate complex with its target peptide from the transcription factor NF-κB. Structure **15**: 289-297.

Reichert A., Baalmann E., Vetter S., Backhausen J. E. and Scheibe R. (2000) Activation properties of the redox-modulated chloroplast enzymes glyceraldehyde 3-phosphate dehydrogenase and fructose-1,6-bisphosphatase. Physiol. Plant. **110**: 330-341.

Rey P., Pruvot G., Becuwe N., Eymery F., Rumeau D. and Peltier G. (1998) A novel thioredoxin-like protein located in the chloroplast is induced by water deficit in *Solanum tuberosum* L. plants. Plant J. **13**: 97-107.

Reynolds A. E., Chesnick J. M., Woolford J. and Cattolico R. A. (1994) Chloroplast encoded thioredoxin genes in the red algae *Porphyra yezoensis* and *Griffithsia pacifica*: Evolutionary implications. Plant Mol. Biol. **25**: 13-21.

Rhee S. G., Kang S. W., Netto L. E., Seo M. S. and Stadtman E. R. (1999) A family of novel peroxidases, peroxiredoxins. Biofactors **10**: 207-209.

Richardson D. H. S. (1981) The biology of mosses. John Wiley and Sons Inc., New York.

Richman A. D. and Kohn J. R. (1996) Learning from rejection: the evolutionary biology of single-locus self incompatibility. Trends Ecol. Evol. **11**: 497-502.

Richman A. D., Uyenoyama M. K. and Kohn J. R. (1996) Allelic diversity and gene genealogy at the self-incompatibility locus in the Solanaceae. Science **273**: 1212-1216.

Rintamäki E., Martinsuo P., Pursiheimo S. and Aro E. M. (2000) Cooperative regulation of light-harvesting complex II phosphorylation via the plastoquinol and ferredoxin-thioredoxin system in chloroplasts. Proc. Natl. Acad. Sci. USA **97**: 11644-11649.

Ritz D. and Beckwith J. (2001) Roles of thiol-redox pathways in bacteria. Ann. Rev. Microbiol. **55**: 21-48.

Ritz D., Patel H., Doan B., Zheng M., Åslund F., Storz G. and Beckwith J. (2000) Thioredoxin 2 is involved in the oxidative stress response in *Escherichia coli*. J. Biol. Chem. **275**: 2505-2512.

Rivera-Madrid R., Mestres D., Marinho P., Jacquot J. P., Decottignies P., Miginiac-Maslow M. and Meyer Y. (1995) Evidence for five divergent thioredoxin h sequences in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. U S A **92**: 5620-5624.

Rodriguez-Suarez R. J., Mora-Garcia S. and Wolosiuk R. A. (1997) Characterization of cysteine residues involved in the reductive activation and the structural stability of rapeseed (*Brassica napus*) chloroplast fructose-1,6-bisphosphatase. Biochem. Biophys. Res. Commun. **232**: 388-393.

Rouhier N., Gelhaye E., Sautiere P. –E., Brun A., Laurent P., Tagu D., Gerard J., de Faÿ E., Meyer Y. and Jacquot J. –P. (2001) Isolation and characterisation of a new peroxiredoxin from poplar sieve tubes that uses either glutaredoxin or thioredoxin as a proton donor. Plant Physiol. **127**: 1299-1309.

Rudd J. J., Lord J. M., Franklin F. C. H. and Franklin-Tong V. E. (1996) Increased phosphorylation of a 26-kD pollen protein is induced by the self-incompatibility response in *Papaver rhoeas*. Plant Cell **8**: 713-724.

Rudd J. J., Franklin F. C. H. and Franklin-Tong V. E. (1997) Ca^{2+}-independent phosphorylation of a 68-kD pollen protein is stimulated by the self-incompatibility response in *Papaver rhoeas*. Plant J. **12**: 507-514.

Ruelland E. and Miginiac-Maslow M. (1999) Regulation of chloroplast enzyme activities by thioredoxins: Activation or relief from inhibition? Trends Plant Sci. **4**: 136-141.

Saba-El-Leil M., Rivard S., Morse D. and Cappadocia M. (1994) The *S11* and *S13* self-incompatibility alleles in *Solanum chacoense* Britt. Are remarkably similar. Plant Mol. Biol. **24**: 571-583.

Sahrawy M., Hecht V., López Jaramillo J., Chueca A., Chartier Y. and Meyer Y. (1996) Intron position as an evolutionary marker of thioredoxins and thioredoxin domains. J. Mol. Evol. **42**: 422-431.

Saitoh M., Nishitoh H., Fujii M., Takeda K., Tobiume K., Sawada Y., Kawabata M., Miyazono K. and Ichijo H. (1998) Mammalian thioredoxin is a direct inhibitor of apoptosis signal-regulating kinase (ASK) 1. EMBO J. **17**: 2596-2606.

Sakuth T., Schobert C., Pecsvaradi A., Eichholz A., Komor E. and Orlich G. (1993) Specific proteins in the sieve tube exudate of *Ricinius communis* L. seedlings: Separation, characterisation and *in vivo* labelling. Planta **191**: 207-213.

Sambrook J., Fritsch E. F. and Maniatis T. (1989) *Molecular Cloning: a Laboratory Manual.* 2nd edn. Cold Spring Harbour Laboratory Press. Cold Spring Harbour, New York, USA.

Sasaki Y., Kozaki A. and Hatano M. (1997) Link between light and fatty acid synthesis - thioredoxin-linked reductive activation of plastidic acetyl- CoA carboxylase. Proc. Natl. Acad. Sci. USA **94**: 11096-11101.

Sassa H., Hirano H. and Ikeshana H. (1992) Self-incompatibility-related RNases in styles of Japanese pear (*Prunus serotina* Rehd.). Plant Cell Physiol. **33**: 811-814.

Sassa H., Hirano H., Nishio T. and Koba T. (1997) Style-specific self-compatible mutation caused by the deletion of the S-RNase gene in Japanese pear (*Prunus serotina*). Plant J. **12**: 223-227.

Scheibe R. (1991) Redox modulation of chloroplast enzymes. Plant Physiol. **96**: 1-3.

Scheibe R. and Anderson L. E. (1981) Dark modulation of NADP-dependent malate dehydrogenase and glucose-6-phosphate dehydrogenase in the chloroplast. Biochim. Biophys. Acta **636**: 58-64.

Schenk H., Klein M., Erdbrugger W., Droge W. and Schulze-Osthoff K. (1994) Distinct effects of thioredoxin and antioxidants on the activation of transcription factors NF-kappa-B and AP-1. Proc. Natl. Acad. Sci. USA **91**: 1672-1676.

Schobert C., Baker L., Szederkényi J., Groβmann P., Komor E., Hayashi H., Chino M. and Lucas W. J. (1998) Identification of immunologically related proteins in sieve-tube exudate collected from monocotyledonous and dicotyledonous plants. Planta **206**: 245-252.

Schopfer C. R., Nasrallah M. E. and Nasrallah J. B. (1999) The male determinant of self-incompatibility in *Brassica*. Nature **286**: 1697-1700.

Schürmann P. and Jacquot J. P. (2000) Plant thioredoxin systems revisited. Ann. Rev. Plant Physiol. Plant Mol.Biol. **51**: 371-400.

Schürmann P., Maeda K. and Tsugita A. (1981) Isomers in thioredoxins of spinach chloroplasts. Eur. J. Biochem. **116**: 37-45.

Schürmann P., Wolosiuk R. A., Breazeale V. D. and Buchanan B.B. (1976) Two proteins function in the regulation of photosynthetic CO_2 assimilation in chloroplasts. Nature **263**: 257-258.

Schwarz O., Schürmann P. and Strotmann H. (1997) Kinetics and thioredoxin specificity of thiol modulation of the chloroplast H^+-ATPase. J. Biol. Chem. **272**: 16924-16927.

Serrato A. J., Crespo J. L., Florencio F. J. and Cejudo F. J. (2001) Characterization of two thioredoxin h with predominant localization in the nucleus of aleurone and scutellum cells of germinating wheat seeds. Plant Mol. Biol. **46**: 361-371.

Shi J. R. and Bhattacharyya M. K. (1996) A novel plasma membrane-bound thioredoxin from soybean. Plant Mol. Biol. **32**: 653-662.

Shiba H., Hinata K., Suzuki A. and Isogai A. (1995) Breakdown of self-incompatibility in *Brassica* by the antisense RNA of the *SLG* gene. Proc. Jpn. Acad. **71**: 81-83.

Shinozaki K. and Yamaguchi- Shinozaki K. (1997) Gene expression and signal transduction in water-stress response. Plant Physiol. **115**: 327-334.

Shoemaker R. C., Polzin K., Labate J., Specht J., Brummer E. C., Olson T., Young N., Concibido V., Wilcox J., Tamulonis J. P., Kochert G. and Boerma H. R. (1996) Genome duplication in soybean (*Glycine* subgenus *soja*). Genetics **144**: 329-338.

Sims T. L. (1993) Genetic regulation of self-incompatibility. Critical Reviews in Plant Sciences. **12**: 129-167.

Singh A. and Kao T-H. (1991) Characterization of ribonuclease activity of three S-allele-associated proteins of *Petunia inflata*. Plant Physiol. **96**: 61-68.

Sjölund R. D. (1997) The phloem sieve element: A river runs through it. Plant Cell **9**: 1137-1146.

Slaby I. And Holmgren A. (1975) Reconstitution of Escherichia coli thioredoxin from complementing peptide fragments obtained by cleavage at methionine-37 or arginine-73. J. Biol. Chcm. **250**: 1340-1347.

Soltis P. S., Soltis D. E. and Chase M. W. (1999) *Arabidopsis* is a member of the rosid lineage of dicot angiosperms. Nature **402**: 402-404.

Southern E. M. (1975) Detection of specific sequences among DNA fragments separated by gel elctrophoresis. J. Mol. Biol. **98**: 503-517.

Stahl R. J., Arnoldo M-A., Glavin T. L., Goring D. R. and Rothstein S. J. (1998) The self-incompatibility phenotype in Brassica is altered by the transformation of a mutant S locus receptor kinase. Plant Cell **10**: 209-218.

Stein J. C., Howlett B., Boyes D. C., Nasrallah M. E. and Nasrallah J. B. (1991) Molecular cloning of a putative receptor protein kinase gene encoded at the self-incompatibility locus of *Brassica oleracea*. Proc. Natl. Acad. Sci. USA **88**: 8816-8820.

Stein J. C. and Nasrallah J. B. (1993) A plant receptor-like gene, the *S*-locus receptor kinase of *Brassica oleracea*, encodes a functional serine/threonine protein kinase. Plant Physiol. **101**: 1103-1106.

Stein M., Jacquot J. P., Jeannette E., Decottignies P., Hodges M., Lancelin J. M., Mittard V., Schmitter J. M. and Migniac-Maslow M. (1995) *Chlamydomonas reinhardtii* thioredoxins: structure of the genes coding for the chloroplastic *m* and cytosolic *h* isoforms; expression in *Escherichia coli* of the recombinant proteins, purification and biochemical properties. Plant Mol. Biol. **28**: 487-503.

Stockinger E. J., Gilmour S. J. and Thomashow M. F. (1997) *Arabidopsis thaliana CBF1* encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a *cis*-acting DNA regulatory element that stimulates transcription in response to low teperature and water deficit. Proc. Natl. Acad. Sci. USA **94**: 1035-1040.

Stone S. L., Arnoldo M-A. and Goring D. R. (1999) A breakdown of *Brassica* self-incompatibility in ARC1 antisense transgenic plants. Science. **286**: 1729-1731.

Suzuki G., Kai N., Fukui K., Nishio T., Takayama S., Isogai A., Watanabe M. and Hinata K. (1999) Genomic organisation of the *S* locus: Identification and characterisation of genes in *SLG/SRK* region of S^9 haplotype of *Brassica campestris* (syn. *rapa*). Genetics **153**: 391-400.

Swafford D. L. (1998) PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4, Sinauer Associates, Sunderland, Mass.

Szederkenyi J., Komor E. and Schobert C. (1997) Cloning of the cDNA for glutaredoxin, an abundant sievé-tube exudate protein from *Ricinus communis*

L. and characterisation of the glutathione-dependent thiol-reduction system in sieve tubes. Planta **202**: 349-356.

Takahata N. (1990) A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism. Proc. Natl. Acad. Sci. USA **87**: 2419-2423.

Takasaki T., Hatakeyama K., Suzuki G., Watanabe M., Isogai A. and Hinata K. (2000) The *S* receptor kinase determines self-incompatibility in *Brassica* stigma. Nature **403**: 913-916.

Takayama S., Shiba H., Iwano M., Shimosato H., Che F-S., Kai N., Watanabe M., Suzuki G., Hinata K. and Isogai A. (2000) The pollen determinant of self-incompatibility in *Brassica campestris*. Proc. Natl. Acad. Sci. USA **97**: 1920-1925.

Takemoto T., Zhang Q. –M. and Yonei S. (1998) Different mechanisms of thioredoxin in its reduced and oxidized forms in defense against hydrogen peroxide in *Escherichia coli*. Free Radic. Biol. Med. **24**: 556-562.

Tezuka T., Tsuruhara A., Suzuki H. and Takahashi S. (1997) A connection between the self-incompatibility mechanism and the stress response in lily. Plant cell Physiol. **38**: 107-112.

Thomashow M. F. (1999) Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. Ann. Rev. Plant Phys. Plant Mol. Biol. **50**: 571-599.

Thompson J. D., Higgins D. G. and Gibson T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**: 4673-4680.

Tremousaygue D., Manevski A., Bardet C., Lescure N. and Lescure B. (1999) Plant interstitial telomere motifs participate in the control of gene expression in root meristems. Plant J. **20**: 553-561.

Tron A. E., Bertoncini C. W., Chan R. L. and Gonzalez D. H. (2002) Redox regulation of plant homeodomain transcription factors. J. Biol. Chem. **277**: 348000-34807.

Tsai D-S., Lee H-S., Post L. C., Kreiling K. M. and Kao T-H. (1992) Sequence of an S-protein of *Lycopersicon peruvianum* and comparison with other solanaceous S-proteins. Sex. Plant Reprod. **5**: 256-263.

Ueno M., Masutani H., Arai R. J., Yamauchi A., Hirota K., Sakai T., Inamoto T., Yamaoka Y., Yodoi J. and Nikaido T. (1999) Thioredoxin-dependant redox regulation of p53-mediated p21 activation. J. Biol. Chem. **274**: 35809-35815.

Ushijima K., Sassa H., Tao R., Yamane H., Dandekar A. M., Gradziel T. M. and Hirano H. (1998) Cloning and characterization of cDNAs encoding S-RNases from almond (*Prunus dulcis*): primary structural features and sequence diversity of the S-RNases in Rosaceae. Mol. Gen. Genet. **260**: 261-268.

Val –del. G., Maurer F., Stutz E. and Schürmann P. (1999) Modification of the reactivity of spinach chloroplast thioredoxin f by site-directed mutagenesis. Plant Sci. **149**: 183-190.

Van Deynze A. E., Nelson J. C., Yglesias E. S., Harrington S. E., Braga D. P., McCouch S. R. and Sorrels M. E. (1995) Comparative mapping in the grasses. Wheat relationships. Mol. Gen. Genet. **248**: 744-754.

Vanlerberghe G. C. and McIntosh L. (1997) Alternative oxidase: From gene to function. Ann. Rev. Plant Physiol. Plant Mol.Biol. **48**: 703-734.

Vekemans X. and Slatkin M. (1994) Gene and allelic genealogies at a gametophytic self-incompatibility locus. Genetics **137**: 1157-1165.

Verdoucq L., Vignols F., Jacquot J. –P., Chartier Y and Meyer Y. (1999) In vivo characterization of a thioredoxin h target protein defines a new peroxiredoxin family. J. Biol. Chem. **274**: 19714-19722.

Villeret V., Huang S. H., Zhang Y. P., Xue Y. F. and Lipscomb W. N. (1995) Crystal structure of spinach chloroplast fructose-1,6-bisphosphatase at 2.8 Angstrom resolution. Biochemistry **34**: 4299-4306.

Vincente-Carbajosa J., Moose S. P., Parson R. and Schmidt R. J. (1997) A maize zinc-finger protein binds the prolamine box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2. Proc. Natl. Acad. Sci. USA **94**: 7685-7690.

Vision T. J., Brown D. G. and Tanksley S. D. (2000) The origins of genomic duplications in *Arabidopsis*. Science **290**: 2114-2117.

Vogt K. and Follmann F. (1986) Characterization of three different thioredoxins in wheat. Biochem. Biophys. Acta **873**: 415-418.

Waksman G., Krishna T. S., Williams C. H., Jr. and Kuriyan J. (1994) Crystal structure of *Escherichia coli* thioredoxin reductase refined at 2 Å resolution.

Implications for a large conformational change during catalysis. J. Mol. Biol. **236**: 800-816.

Wang P. F., Veine D. M., Ahn S. H. and Williams C. H., Jr. (1996) A stable mixed disulfide between thioredoxin reductase and its substrate, thioredoxin: Preparation and characterization. Biochemistry **35**: 4812-4819.

Wangensteen O. S., Chueca A., Hirasawa M., Sahrawy M., Knaff D. B. and López Gorgé J. (2001) Binding features of chloroplast fructose-1,6-bisphosphatase-thioredoxin interaction. Biochim. Biophys. Acta **1547**: 156-166.

Watabe S., Hiroi T., Yamamoto Y., Fujioka Y., Hasegawa H., Yago N., Takahashi S. Y. (1997) Sp-22 is a thioredoxin-dependent peroxide reductase in mitochondria. Eur. J. Biochem. **249**: 52-60.

Watanabe M., Ito A., Takada Y., Ninomiya C., Kakizaki T., Takahata Y., Hatakeyama K., Hinata K., Suzuki G., Takasaki T., Satta Y., Shiba H., Takayama S. and Isogai A. (2000) Highly divergent sequences of the pollen self-incompatibility (*S*) gene in class-I *S* haplotypes of *Brassica campestris* (syn. *rapa*) L. FEBS Lett. **473**: 139-144.

Watanabe M., Takasaki T., Toriyama K., Yamakawa S., Isogai A., Suzuki A. and Hinata K. (1994) A high degree of homology exists between the protein encoded by *SLG* and the *S* receptor domain encoded by *SRK* in self-incompatible *Brassica campestris* L. Plant Cell Phys. **35**: 1221-1229.

Wedel N., Clausmeyer S., Herrmann R. G., Gardet-Salvi L. and Schürmann P. (1992) Nucleotide sequence of cDNAs encoding the entire precursor polypeptide for thioredoxin m from spinach chloroplasts. Plant Mol. Biol. **18**: 527-533.

Weichsel A., Gasdaska J. R., Powis G. and Montfort W. R. (1996) Crystal structures of reduced, oxidized, and mutated human thioredoxins: Evidence for a regulatory homodimer. Structure **4**: 735-751.

Wenderoth I., Scheibe R. and von Schaewen A. (1997) Identification of the cysteine residues involved in redox modification of plant plastidic glucose-6-phosphate dehydrogenase. J. Biol. Chem. **272**: 26985-26990.

Williams C. H., Jr. (1995) Mechanism and structure of thioredoxin reductase from *Escherichia coli*. FASEB J. **9**: 1267-1276.

Williams C. H., Arscott L. D., Muller S., Lennon B. W., Ludwig M. L., Wang P. -F., Veine D. M., Becker K. and Schirmer R. H. (2000) Thioredoxin reductase - two modes of catalysis have evolved. Eur. J. Biochem. **267**: 6110-6117.

Wolosiuk R. A., Crawford N. A., Yee B. C. and Buchanan B. B. (1979) Isolation of three thioredoxins from spinach leaves. J. Biol. Chem. **254**: 1627-1632.

Wright S. (1939) The distribution of self-sterility alleles in populations. Genetics **24**: 538-552.

Wynn R., Cocco M. J. and Richards F. M. (1995) Mixed disulfide intermediates during the reduction of disulfides by *Escherichia coli* thioredoxin. Biochem. **34**: 11807-11813.

Yamaguchi-Shinozaki K. and Shinozaki K. (1994) A novel *cis*-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low temperature, or high-salt stress. Plant Cell **6**: 251-264.

Yamamoto E. and Knap H. T. (2001) Soybean receptor-like protein kinase genes: Paralogous divergence of a gene family. Mol. Biol. Evol. **18**: 1522-1531.

Yanagisawa S. (1996) Dof DNA-binding proteins contain a novel zinc-finger motif. Trends Plant Sci. **1**: 213-214

Yanagisawa S. and Sheen J. (1998) Involvement of maize Dof zinc finger proteins in tissue-specific and light-regulated gene expression. Plant Cell **10**: 75-89.

Yang Y. –W., Lai K. –N., Tai P. –Y. and Li W. –H. (1999) Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. J. Mol. Evol. **48**: 597-604.

Yano H., Wong J. H., Lee Y. M., Cho M. J. and Buchanan B. B. (2001) A strategy for the identification of proteins targeted by thioredoxin. Proc. Natl. Acad. Sci. USA **98**: 4794-4799.

Yokoyama S. and Hetherington L. E. (1982) The expected number of self-incompatibility alleles in finite plant populations. Heredity **48**: 299-303.

Yu K., Schafer U., Glavin T. L., Goring D. R. and Rothstein S. J. (1996) Molecular characterisation of the *S* locus in two self-incompatible *Brassica napus* lines. Plant Cell **8**: 2369-2380.

Zang L., Pond S. K. and Gaut B. S. (2001) A survey of the molecular evolutionary dynamics of twenty-five multigene families from four grass taxa. J. Mol. Evol. **52**: 144-156.

Zhang N. and Portis A. R. (1999) Mechanism of light regulation of rubisco: A specific role for the larger rubisco activase isoform involving reductive activation by thioredoxin-f. Proc. Natl. Acad. Sci. USA **96**: 9438-9443.

Zhou Y., Kok K. H., Chun A. C., Wong C. M., Wu H. W., Lin M. C., Fung P. C., Kung H. and Jin D. Y. (2000) Mouse peroxiredoxin V is a thioredoxin peroxidase that inhibits p53-induced apoptosis. Biochem. Biophys. Res. Commun. **268**: 921-927.