AN INVESTIGATION INTO THE MECHANISM AND FUNCTION OF CYSTEINE OXIDATION IN THE PLANT N-END RULE PATHWAY

Daniel Joseph Rooney, BSc

September 2016

Submitted to the University of Nottingham for the degree of doctor of philosophy

Plant and Crop Sciences, School of Biosciences

The University of Nottingham

A**cknowledgements**

I would like to thank my supervisors Prof Mike Holdsworth and Prof Neil Oldham for their continued support and guidance over the duration of my project. Thanks also go to the BBSRC for funding this research.

I would like to thank all of the past and present colleagues at South lab and Chemistry for their help and support, but most importantly for making those long days in the lab/office that little bit shorter. Holdsworth lab: Dr Jorge Vicente Conde, Dr Sophie Berckhan, Dr Guillermina Mendiondo, Dr Cristina Sousa Correia, Dr Tinne Boeckx, Dr Mohamad Abbas, Dr Daniel Gibbs, Dr Ken Davis, Dr Geeta Prasad, Dr Melania Ghita, Dr Kerstin Mueller, Chris Till, Natasha Gladstone, Charlene Dambire, Julietta Marquez, Fiona Wilkinson; Oldham lab: Dr Daniel Scott, Dr Jose Correia Afonso, Dr Matthew Jenner, Lucio Manzi, Nkazi Tshuma, Jed Bellamy-Carter, Vivek Badiani.

Special thanks to Dr Daniel Gibbs who supervised me at the start of my PhD. Special thanks also to Dr Susan Liddell for generously offering her help, time and MS machine and all the useful discussions. Special thanks to both Julietta Marquez and Dr Ken Davis for their brilliant management of both labs and making the lab a conducive environment for work with a smile and a laugh (and on one too many times a song).

Lastly I would like to thank my family for their support and encouragement, especially my parents. The last mention goes to my wonderful wife Maryam, your love and encouragement has helped me keep going on numerous times over the past 4 years, I am lucky to have you.

Abstract

Flooding events are becoming more common throughout the world as a result of climate change, resulting in reduced crop yields. It was recently discovered that plants sense low oxygen (O_2) (associated with flooding) through regulated proteolysis of the group VII Ethylene Response Factor transcription factors (ERFVIIs), via the Cys-Arg/N-end rule pathway of ubiquitin mediated proteolysis, which also senses another gas, nitric oxide (NO). The N-terminal (Nt) Cys of physiological (e.g. ERFVIIs) and artificial substrates was shown to be key for N-end rule function, and work in mammalian systems suggested that oxidation of Nt-Cys by O_2 and NO was a required prerequisite for subsequent Nt arginylation by arginyl tRNA transferases (ATEs). However the exact mechanism of Nt-Cys oxidation has not been discovered. The primary aim of this thesis was to define the mechanism of *in vivo* Nt-Cys oxidation essential in determining the stability of Arg/N-end rule protein substrates, including ERFVIIs.

In this study a novel approach was developed to investigate the oxidation of Nt-Cys *in vivo* using transgenic Arabidopsis expressing Cys-2 reporter proteins (derived from the transgenes *35S::MC-polyG-HA-GUS* and *Ubi1::MCGGAIL-GUS*). Cys-2 of the reporter proteins is made Nt constitutively by co-translational Methionine Aminopeptidase (MAP) activity. Biochemical techniques were combined with analytical chemistry to investigate the *in vivo* oxidation of Nt-Cys using liquid chromatography mass spectrometry (LC-MS). In addition, synthetic peptides representing

the Nt-sequence of the *in vivo* reporter protein were used to define the oxidative and nitrosylative modifications occurring at Nt-Cys *in vitro.*

Findings of this study include the successful development of two new *in vivo* O_2 sensor artificial N-end rule substrates, in the plant genetic model *Arabidopsis thaliana*. A further outcome revealed that the ERFVII RELATED TO APETALA 2.12 (RAP2.12) is stabilised in the shoot and root apical meristem to a greater extent than in other regions of seedlings in response to hypoxia, indicating that meristematic cells could be important oxygen sensory zones. Cys-2 reporter proteins were used in *in vivo* studies to attempt to identify the nature of the Nt residues following N-end rule action. Although Nt-peptides derived from affinity purified Cys-2 reporter protein were not identified by LC-MS, it was possible to demonstrate that ubiquitination is required on Cys-2 reporter proteins before 26S proteasome degradation. Combined results from Nt-Cys *in vitro* synthetic peptides and Cys-2 reporter proteins substrates of the Arg/N-end rule pathway, provide indirect evidence that post translational modifications (PTMs) not defined before occur *in vivo*. Using synthetic peptides it was possible to show evidence for Nt-Cys-sulfenamide formation after oxidation of Nt-Cys. This finding suggests that Cys-sulfenamide formation, a previously recognised reversible modification preventing irreversible oxidation to Cys-sulfonic acid, could occur at Nt-Cys of *in vivo* protein substrates. An important finding of this study was the observed non reaction between the nitrosylated Nt-Cys and H_2O_2 or conversely oxidised Nt-Cys and NO. As Plant Cysteine Oxidases (PCOs) do not require NO to oxidise Nt-Cys, this result raises further uncertainty as to how NO is involved in the oxidation of Nt-Cys, suggesting that NO

may be involved in the enzymatic activation of PCOs or the arginylation of Nt-Cys by ATEs.

The findings of this study did not identify the Nt-peptide of the Cys-2 reporter protein, and hence the Nt-Cys oxidation state of an *in vivo* substrate of the Arg/N-end rule pathway remains unidentified. Despite this the result that Nt-Cys can be oxidised to Cys-sulfenamide is novel and an important discovery to the field of reactive Cys biology and chemistry.

Contents

1. Introduction 1

1.1 Protein degradation and the Ubiquitin Proteasome System 2

1.1.1 The role of protein degradation 2

1.1.2 Lysosomal protein degradation 3

1.1.3 The discovery of Ubiquitin 4

1.1.4 The role of Ubiquitin in protein degradation 5

1.1.5 The specificity and target of the Ubiquitin enzymes 6

1.2 Selective protein degradation via the N-end rule pathway 10

1.2.1 The discovery of the N-end rule pathway 10

1.2.2 The components and working of the Arg/N-end rule pathway 11

1.2.3 The N-end rule pathway in bacteria 13

1.2.4 The N-end rule pathway in *Saccharomyces cerevisiae* 16

1.2.5 The N-end rule pathway in mammals 17

1.2.6 The N-end rule pathway in plants 18

1.3 Oxygen and nitric oxide sensing in the N-end rule pathway 20

1.3.1 The relevance of Cys in biological systems 20

1.3.2 The oxidative degradation of Nt-Cys substrates in the N-end rule pathway 23

1.3.3 Cysteine oxidation in the mammalian Arg/N-end rule pathway 24

1.3.4 Cysteine oxidation in the plant Arg/N-end rule pathway 26

1.4 The role of oxygen and nitric oxide as plant signalling molecules 32
1.4.1 Reactive oxygen and nitrogen species 32
1.4.2 NO production strategies in plants 33
1.4.3 Function of NO and ROS sensing in plant development 34
1.4.4 S-nitrosylation and plant response to biotic and abiotic stress 36
1.5 Detection of post translational modifications in the N-end rule pathway using mass spectrometry 40
1.5.1 An introduction to mass spectrometry 40
1.5.2 Detecting *in vivo* post translational modifications to substrates of the N-end rule pathway 45
1.5.3 Monitoring PTMs of the N-end rule pathway *in vitro* using synthetic polypeptides 48
1.6 Aims and objectives of this investigation 50
2. Materials and methods 51
2.1 Plants and plant growth conditions 51
2.1.1 Plant material 51
2.1.2 Plant growth media 51
2.1.3 Seed plating and stratification 54
2.1.4 Growth conditions 54
2.1.5 Cross pollination of Arabidopsis flowers 55
2.2 DNA Amplification, Cloning and Transformations 56
2.2.1 Genomic DNA extraction 56
2.2.2 Oligonucleotide primers 56
2.2.3 Polymerase chain reaction (PCR) 58

2.2.4 DNA gel electrophoresis 59
2.3 Cloning and Transformations 61
2.3.1 Restriction digestions, ligations and Gateway cloning 61
2.3.2 Lysogeny broth media 62
2.3.3 Transformations – *Escherichia coli* 62
2.3.4 Plasmid DNA isolation from *E. coli* 63
2.3.5 DNA sequencing 63
2.3.6 Transformations – *Agrobacterium tumefaciens* 63
2.3.7 Transformations – *Arabidopsis thaliana* 64
2.3.8 Selection of transgenic Arabidopsis 65
2.4 Plant treatments and GUS staining 66
2.4.1 Hypoxia treatment of Arabidopsis 66
2.4.2 Analysis of protein stabilisation using the 26S proteasome inhibitor bortezomib 66
2.4.3 Analysis of expression and stability of β-glucuronidase protein in transgenic plants 67
2.5 Protein extraction, purification and identification 69
2.5.1 Protein extraction from plant material 69
2.5.2 Affinity purification of HA-tagged protein from Arabidopsis seedlings 70
2.5.3 Polyacrylamide gel electrophoresis (PAGE) 71
2.5.4 Identifying affinity purified protein using colloidal Coomassie stain 72
2.5.5 PAGE gel band excision, processing, digestion and extraction 73
2.5.6 Western blots 76

2.5.7 Monitoring equal protein loading after Western blot using Ponceau stain 77

2.6 Electrospray Ionisation Mass Spectrometry (ESI-MS) of *in vitro* synthetic peptides 78

2.6.1 Synthetic peptides 78

2.6.2 Analysis of *in vitro* synthetic peptides 78

2.6.3 Peptide oxidation – H_2O_2 78

2.6.4 Peptide nitrosylation – SNAP/SNP 78

2.6.5 Two stage treatments – H_2O_2 and SNAP 79

2.6.6 LTQ FT Ultra Mass spectrometer and ESI-MS/MS run conditions 81

2.7 *In vitro* analysis of CR dipeptide by LC-MS/MS 82

2.7.1 CR dipeptide 82

2.7.2 Analysis of CR dipeptide by LC-MS/MS 82

2.8 Nanospray Ionisation Mass Spectrometry (NSI-MS) of digested *in vivo* peptides from CaMV *35S::MC-GUS* in *prt6* 84

2.8.1 Liquid Chromatography Mass Spectrometry using LTQ FT Ultra MS 84

2.8.2 MS/MS conditions and peptide identification 85

3. Developing transgenic sensors of oxygen of the N-end rule pathway in in *Arabidopsis thaliana* 86

3.1 Introduction 86

3.2 Developing transgenic plants providing expression of RAP2.12 89

3.2.1 Generation of *promRAP2.12::MA-RAP2.12-GUS* 89

3.2.2 *In vivo* expression of *RAP2.12* in Arabidopsis seedlings and monitoring stabilisation under hypoxia 90

3.2.3 Stabilisation of MA-RAP2.12-GUS and comparison to MC-RAP2.12-GUS 95

3.2.4 RAP2.12 and the other ERFVIIs are important regulators of photomorphogenesis 97

3.3 *In vivo* expression of reporter proteins bearing the conserved Nt motif of plant ERFVIIs in Arabidopsis 100

3.3.1 Conserved N-terminal motif of the ERFVIIs 100

3.3.2 Transforming *Ubi1::M(C/A)GGAIL-GUS* in Arabidopsis and the *in vivo* expression of M(C/A)GGAIL-GUS 102

3.3.3 Stabilsation of M(C/A)GGAIL-GUS under hypoxia and treatment with 26S inhibitor bortezomib 105

3.4 Developing a Cys-2 reporter of the Arg/N-end rule pathway to be used for protein purification 109

3.4.1 Analysis of expression and accumulation levels of existing Met-Cys reporter proteins 109

3.4.2 Identifying differences of *in vivo* GUS reporter protein accumulation and comparing in Col-0 and *prt6* backgrounds 111

3.4.3 Stabilisation of GUS reporter protein in the intermediate line A25 after exposure to hypoxia and proteasome inhibitors 115

4. Identification of GUS from affinity purified seedlings by LC-MS/MS reveals various possible Nt-Cys oxidation states 119

4.1 Introduction 119

4.2 Affinity purification of GUS *in vivo* derived from the *35S::MC-GUS* transgene to identify PTMs involved in N-end rule processing 122

4.2.1 Anti-HA antibody detection of affinity purified GUS protein increases following treatment with bortezomib, a 26S proteasome inhibitor 122

4.2.2 Ubiquitinated GUS protein increases after inhibition of the 26S proteasome using bortezomib 124

4.2.3 Anti-sumoylation antibody detection of affinity purified sumoylated GUS protein unchanged after stabilisation of the 26S proteasome using bortezomib 127

4.3 LC-MS identification of GUS reporter protein from affinity purified *35S::MC-GUS* Arabidopsis seedlings 130

4.3.1 Matching profiles of GUS protein bands on a Coomassie gel and Western film 130

4.3.2 Bulked extractions of GUS protein derived from *35S::MC-GUS* for gel extractions 132

4.3.3 LC-MS/MS detection of GUS reporter protein and varied digestive performance 134

4.4 Predicted N-terminal MC-GUS peptides and possible oxidative Cys PTMs 139

4.4.1 Edman sequencing of GUS reporter protein 139

4.4.2 Trypsin, Asp-N and Lys-C Nt-peptide prediction of GUS reporter protein derived from *35S::MC-GUS* 142

4.4.3 Using the RCR-49 synthetic peptide to mimic the predicted Nt Asp-N peptide of the *in vivo* GUS reporter protein 146

5. Analysis of Nt-Cys oxidation and nitrosylation *in vitro* using synthetic peptides reveals both expected and unexpected oxidative routes 152

5.1 Introduction 152

5.2 Identification of the tryptic dipeptide Cys-Arg by LC-MS/MS and its Hydrogen peroxide driven oxidation 155

5.2.1 Detection of CR dipeptide as a mimic for Nt RCR tryptic peptide from the *in vivo* GUS reporter protein 155

5.2.2 Oxidation of CR dipeptide with hydrogen peroxide 158

5.3 *In vitro* oxidation of Nt-Cys 40 residue synthetic peptide reveals sulfenamide oxidative route and NO-induced nitrosylation 162

5.3.1 ESI-MS/MS identification of CR-40 synthetic peptide 162

5.3.2 Oxidation of the Nt-Cys of CR-40 using hydrogen peroxide 165

5.3.3 SNAP-induced Nt-Cys nitrosylation of CR-40 172

6. General Discussion 178

6.1 Cys-Arg/N-end rule regulated sensors of oxygen in Arabidopsis 178

6.2 PTMs occurring to *in vivo* GUS reporter protein and its digestion efficiency 182

6.3 Possible PTMs occurring on missing Nt-Asp-N MC-GUS peptide using *in vitro* RCR-49 synthetic peptide 187

6.4 Oxidation and nitrosylation of Nt-Cys CR-40 synthetic peptide *in vitro* suggests numerous possible oxidative routes *in vivo* 188

6.5 Conclusions and Further work 195

7. References 198

8. Appendix 231

A. Custom GUS MASCOT database 231

B. Root tip normoxia treated images of *promRAP2.12::MA-RAP2.12-GUS* seedlings 233

C. Root tip images of normoxia/hypoxia *Ubi1::M(C/A)GGAIL* seedlings 234

D. Identified tryptic GUS reporter peptides from Figure 4.5 235

E. Predicted masses of Nt Acetylated Asp-N peptide 236

F. Cys-nitrosothiol MS2 and MS3 spectra 237

G. MS spectra of dual treatment of CR-40 synthetic peptide if H_2O_2 used after SNAP 238

H. MS spectra of dual treatment of CR-40 synthetic peptide if SNAP used after H_2O_2 239

I. MS spectra after incubation of CR-40 with SNP 240

List of Figures

Figure 1.1 The Ubiquitin proteasome system 9

Figure 1.2 Residues degraded by the Arg/Ac N-end rule pathway 15

Figure 1.3 Diversity of oxidative Cys post translational modifications 22

Figure 1.4 Structural similarity of Cys-sulfinic acid and Aspartic acid 24

Figure 1.5 Cysteine oxidation branch of the Arg/N-end rule pathway 31

Figure 1.6 Arabidopsis seeds show enhanced germination when exposed to NO than when compared to pure air 35

Figure 1.7 Modifications to metabolism of reactive nitrogen species in response to abiotic stress in plants 38

Figure 1.8 Components of a mass spectrometer 41

Figure 1.9 Electrospray ionisation emission 42

Figure 1.10 Collision induced dissociation and electron capture dissociation fragmentation sites on a peptide backbone 45

Figure 2.1 Constructs used in this study which have been transformed into Arabidopsis 53

Figure 2.2 Reduction, alkylation and digestion process for excised Coomassie stained protein bands 75

Figure 3.1 Structure of *promRAP2.12::M(A/C)-RAP2.12-GUS* 89

Figure 3.2 Sequence comparison of *promRAP2.12::MA-RAP2.12-GUS* and *promRAP2.12::MC-RAP2.12-GUS* 90

Figure 3.3 Expression of *RAP2.12* in SAM, hypocotyl and root 91

Figure 3.4 Comparison of expression and stabilisation of MC-RAP2.12-GUS in normoxia/hypoxia 92

Figure 3.5 Expression of *RAP2.12* (AT1G53910) in total and polysomal RNA compared in normoxic and hypoxic conditions. 95

Figure 3.6 Comparison of expression and stabilisation of MC/MA-RAP2.12-GUS 97

Figure 3.7 Influence of oxygen and ERFVIIs in controlling shoot apical hook opening 99

Figure 3.8 *Ubi1::M(C/A)GGAIL-GUS* construct 101

Figure 3.9 Expression and stabilisation of M(C/A)GGAIL-GUS in barley 103

Figure 3.10 Expression and stabilisation of M(C/A)GGAIL-GUS in Arabidopsis 104

Figure 3.11 Expression and stabilisation of M(C/A)GGAIL-GUS in normoxia/hypoxia 106

Figure 3.12 Anti-GUS antibody Western blot of protein extracted from *Ubi1::MCGGAIL-GUS* 108

Figure 3.13 Anti-Haemmaglutin antibody Western blot of GUS reporter lines in *prt6* 110

Figure 3.14 Genotyping to identify intermediate GUS reporter Arabidopsis lines 111

Figure 3.15 Expression and stabilisation of GUS line A2/1 and A25 113

Figure 3.16 Anti-Haemmaglutinin antibody Western blot of GUS reporter lines A2/1 and A25 in Col-0 and *prt6* 114

3.17 Expression and stabilisation of MA/MC-GUS reporter lines in normoxia/hypoxia in the shoot 116

3.18 Expression and stabilisation of MA/MC-GUS reporter lines in normoxia/hypoxia in the root 117

3.19 Anti-Haemmaglutinin antibody Western blot of GUS reporter line A25 in hypoxia and bortezomib treated seedlings 118

4.1 Anti-Haemmaglutinin antibody Western blot showing affinity purified GUS protein after DMSO and bortezomib treatment 123

4.2 Anti-Ubiquitin antibody Western blot of affinity purified ubiquitinated GUS protein after DMSO and bortezomib treatment 126

4.3 Anti-sumoylation antibody Western blot of affinity purified sumoylated GUS protein after DMSO and bortezomib treatment 129

4.4 Anti-Haemmaglutinin antibody Western blot from affinity purified protein extracted from wild-type and *35S::MC-GUS* Arabidopsis seedlings in *prt6* 131

4.5 Coomassie gel showing affinity purified protein extracted from wild-type and *35S::MC-GUS* Arabidopsis seedlings in *prt6* 132

4.6 Bulked Anti-Haemmaglutinin antibody Western blot from affinity purified protein extracted from *35S::MC-GUS* Arabidopsis seedlings in *prt6* 133

4.7 Protein sequence of the GUS reporter protein 134

4.8 Tryptic peptides of GUS reporter protein identified by LC-MS/MS 136

4.9 Asp-N peptides of GUS reporter protein identified by LC-MS/MS 137

4.10 Total residues of the GUS reporter protein identified by LC-MS/MS after digestion with trypsin, Lys-C and Asp-N 138

4.11 Edman sequencing results of GUS reporter protein 141

4.12 Predicted trypsin, Lys-C and Asp-N cleavage sites of MC/RC-GUS 143

4.13 MS/MS spectra of tryptic peptide closest to N-terminus of the GUS reporter protein 144

4.14 MS/MS spectra for Asp-N peptide of the GUS reporter protein closest to N-terminus 145

4.15 Ion charge states of synthetic RCR-49 peptide 147

4.16 MS/MS spectra of RCR-49 synthetic peptide 148

4.17 MS/MS spectra of alkylated RCR-49 synthetic peptide 149

5.1 Chmoratogram elution peaks of CR dipeptide 155

5.2 MS and MS/MS analysis of CR dipeptide 157

5.3 Ion chromoatogram of untreated CR dipeptide elution peaks 158

5.4 LC-MS spectra of CR dipeptide after H_2O_2 incubation 159

5.5 Ion chromatogram of H_2O_2 treated CR dipeptide 160

5.6 MS/MS spectra of Cys-sulfenic acid and Cys-sulfonic acid 161

5.7 LC-MS of of the untreated CR-40 synthetic peptide 163

5.8 CID and ECD fragmentation ions of CR-40 synthetic peptide 164

5.9 ESI-MS of CR-40 before and after oxidation of H_2O_2 166

5.10 LC-MS and ECD fragmentation pattern of Cys-sulfonamide 167

5.11 Cysteine oxidation routes possible after oxidation with H_2O_2 169

5.12 Time course of oxidation of CR-40 synthetic peptide after H_2O_2 incubation 171

5.13 LC-MS of SNAP species 173

5.14 LC-MS of nitrosylated CR-40 and SNAP dimer species 174

5.15 Reaction mechanism of SNAP induced nitrosylation of CR-40 176

5.16 Time course of nitrosylation of the Nt-Cys of the CR-40 synthetic peptide 177

6.1 Hypothetical NO-induced nitrosylation mechanism of *in vivo* protein bearing Nt-Cys 194

List of Tables

Table 1.1 Cysteine oxidtative modifications and their anticipated mass change 48

Table 2.1 Antibiotics and herbicides used in this study 52

Table 2.2 Soil composition for plant culture 55

Table 2.3 Genomic DNA extraction buffer 56

Table 2.4 Oligonucleotide primers used for site directed mutagenesis 57

Table 2.5 Oligonucleotide primers used for sequencing 57

Table 2.6 Oligonucleotide primers used for genotyping 58

Table 2.7 Polymerase chain reaction (PCR) mix 59

Table 2.8 PCR thermocyler conditions 59

Table 2.9 Components of 0.8% agarose gel used for DNA electrophoresis 60

Table 2.10 Components of restriction digestion enzyme reation mix 61

Table 2.11 Components of the X-gluc working solution used to stain plant tissue 68

Table 2.12 Components of the Hoyers mounting medium used to mount slides 68

Table 2.13 Components of the protein extraction buffer 69

Table 2.14 Components of the affinity purification wash buffer 71

Table 2.15 Components required to cast an 8% 1.5 mm Bis/Acrylamide gel 72

Table 2.16 Endoroteinases used in this study and their site of action 74

Table 2.17 Western blot transfer buffer 77

Table 2.18 Synthetic peptides used in this study 80

Table 3.1 The conserved Nt motif conserved in Arabidopsis and Rice ERFVIIs 102

Table 4.1 Predicted oxidative reactions occurring to the Nt-Cys of the Asp-N Nt peptide of the GUS reporter protein and its resultant mass shift 150

Table 4.2 Predicted oxidative reactions occurring to the Nt-Cys of the Asp-N Nt peptide of the GUS reporter protein, and its resultant mass shift if Nt arginylation did not occur 151

Abbreviations

µm	micrometre
Å	Angstrom
β-gal	β-galactosidase
A/Ala	Alanine
ABA	abscisic acid
Ac	acetylated
Alk	alykylated
Asn/N	Asparagine
Asp/D	Aspartic acid
ATE	arginyl transferase
ATP	adenine triphosphate
Bp	base pairs
BSA	bovine serum albumin
C/Cys	Cysteine
CaMV	Cauliflower mosaic virus
CID	collision induced dissociation
d	days
D/Asp	Aspartic acid
Da	Dalton
DDA	data dependent acquisition
DTT	dithiotheritol
ECD	electron capture dissociation
ERFVIIs	group VII ethylene response factor
ESI	Electrospray ionisation

eV	electronvolt
F/Phe	Phenylalanine
FTICR	Fourier transform ion cyclotron resonance
g	gram
Gln/Q	Glutamine
Glu/E	Glutamic acid
G/Gly	Glycine
GA	gibberellin
GSH	glutathione
H_2O_2	Hydrogen peroxide
HA	Haemmaglutinin
His/H	Histidine
HPLC	high performance liquid chromatography
I	input (crude protein)
i.d.	internal diameter
IAA	iodoacetamide
ICAT	isotope coded affinity tags
Ile/I	Isoleucine
K/Lys	Lysine
kDa	kilodalton
kV	kilovolt
L	Litre
L/Leu	Leucine
LB	Lysogeny broth
LC	liquid chromatography
M	molar

m	metre
m/z	mass/charge
Ma	millon years ago
MALDI	matrix assisted laser desorption ionisation
MAP	Methionine aminopeptidase
mBar	millibar
MeCN	acetonitrile
mg	milligram
ml	millilitre
mm	millimetre
mM	millimolar
MS	mass spectrometry
MS/MS	tandem mass spectrometry
msec	millisecond
MSk	Murashige Skoog
MW	molecular weight
N/Asn	Asparagine
NB	not bound
ng	nanogram
NR	nitrate reductase
Nt	N-terminal
NTA	N-terminal amidase
OD	optical density
Ox.	oxidised
P/Pro	Proline
PAGE	polyacrylamide gel electrophoresis

PCO	Plant Cysteine Oxidase
pp	partial pressure
PRT1	PROTEOLYSIS1
PRT6	PROTEOLYSIS6
PTM	post translational modification
Q/Gln	Glutamine
R	release
R/Arg	Arginine
RAM	root apical meristem
REZ	root elongation zone
RF	radio frequency
RHDZ	root hair development zone
RMZ	root maturation zone
RNS	reactive nitrogen species
ROS	reactive oxygen species
rpm	revolutions per minute
SAM	shoot apical meristem
SDM	site directed mutagenesis
SDS	Sodium dodecyl sulfate
Ser/S	Serine
SNO	nitrosothiol
SUMO	small ubiquitin like modifiers
TAILS	terminal amine isotopic labelling of substrates
Thr/T	Threonine
TOF	time of flight
Ub	ubiquitin

UBR	ubiquitin box recognin
UPS	ubiquitin proteasome system
v	volume
V	Volt
Val/V	Valine
w	weight
W/Trp	Tryptophan
W5	fifth wash
WT	wild-type

1. Introduction

This study focusses on the oxidation of N-terminal (Nt) cysteine (Cys) in proteins whose stability depends on this reaction. The discovery that an amino terminal residue can control the stability of proteins is a principle of the N-end rule pathway, a biochemical system which regulates proteolysis, and that is a component of the ubiquitin proteasome system. The ability of plant and mammalian N-end rule substrates bearing Nt-Cys to sense both oxygen and nitric oxide in the cellular environment was an important physiological discovery in mammalian and plant systems (Gibbs et al., 2014b; Hu et al., 2005; Licausi et al., 2011). The developmental and physiological importance of Nt-Cys substrates of the plant N-end rule pathway have been identified, showing their involvement in seed dormancy, photomorphogenesis and stomatal closure (Abbas et al., 2015; Gibbs et al., 2014b). The group of Nt-Cys substrates responsible for this were the group VII Ethylene Response Factor transcription factors (ERFVIIs), these proteins are conserved in plants and when stabilised increase their tolerance to flooding and other abiotic stresses. The first aim of the research described in this thesis (chapters 3 and 4) was to investigate the *in vivo* oxidation of Nt-Cys in reporter proteins expressed in Arabidopsis, using a variety of protein detection techniques including mass spectrometry (MS). The second aim of this research was to investigate the controlled oxidation of Ntl-Cys synthetic peptides *in vitro* (chapter 5).

1.1 Protein degradation and the Ubiquitin Proteasome System

1.1.1 The role of protein degradation

Protein degradation is an essential and selective regulatory process occurring within all organisms. Individual protein concentrations vary within cells due to their differing *in vivo* half-lives and expression rates. Total protein concentrations within a cell are controlled by their rate of synthesis and degradation. Attempts to quantify the rate of protein production were made initially with *Escherichia coli*, with findings indicating how the production rate of individual proteins was related to growth and functionality (Li et al., 2014). The degradation of proteins into their constituent peptides, otherwise known as proteolysis, traditionally was viewed as only involved in the breakdown of misfolded, damaged and unwanted proteins (Goldberg, 1972). The small polypeptides produced during proteolysis could then be recycled and used to synthesise new proteins in a charged tRNA dependent manner (Goldberg, 1971). This is now known not to be true, with protein degradation also being an important regulatory mechanism controlling cellular functions. There are two dominant pathways involved with protein degradation in eukaryotic cells, lysosomal proteolysis and ubiquitin driven proteolysis.

1.1.2 Lysosomal protein degradation

Lysosomes, a term coined by de Duve et al. (1955) after studies into rat livers, are membrane enclosed organelles which engulf cellular substances such as sugars, nucleic acids and proteins using digestive enzymes to facilitate degradation. The discovery of these lysosomes in eukaryotes and their capabilities was the first evidence for intracellular protein degradation. Initially the discovery of the lysosome, and more specifically the inner separation of the digestive enzymes from potential substrates by a membrane, was seen to support the hypothesis that the rate of both protein synthesis and degradation is maintained at a constant rate (Ciechanover, 2005). This theory was further strengthened by associated works which detailed the different mechanisms of autophagy (see review Mortimore and Poso, 1987).

The hypothesis that proteins could be in a 'dynamic state' within cells preceded the discovery of the lysosome (Schoenheimer, 1942), and resurfaced afterwards challenging traditional thinking. As the working of the lysosome became more studied scientists began to question how protein degradation rates varied between minutes to a number of days for different proteins, contradicting the thinking that the lysosome facilitated all intracellular degradation (Schimke and Doyle, 1970). Additionally, numerous findings began to show how cells growing in sub-optimal conditions prevented or altered the degradation rate of selected proteins (e.g. Pine, 1973). The discovery that energy, in the form of ATP, was required for releasing amino acids after intracellular protein degradation and which was not purely hydrolytic (Mego and Farb, 1978),

raised questions about the apparent steady state lysosomal controlled proteolysis. Follow up work using energy inhibitors, inhibiting the production of ATP, prevented the degradation of Tyrosine aminotransferase in Hepatoma cell cultures (Hershko and Tomkins, 1971). Subsequent findings of ATP-dependent proteolytic activity in nuclear-free reticulocyte cells which lack lysosomes (Etlinger and Goldberg, 1977), clearly indicated that non-lysosomal proteolysis was occurring.

1.1.3 The discovery of Ubiquitin

The discovery of ubiquitin (Ub), a heat-stable 76 amino acid peptide was made after reticulocyte cells were fractionated on anion exchange resin (Ciechanover et al., 1978). Two distinct fractions from the lysate were discovered, fraction 1 contained Ub and other unadsorbed flow through material, whilst fraction 2 was the eluate of the adsorbed substances which was bound to the diethylaminoethyl (DEAE) cellulose (Ciechanover et al., 1978). To find out more about the proteolytic capacity of these reticulocytes, globin was exposed to both fractions individually and combined either with/without ATP addition. Individually only fraction 2 displayed very low proteolytic activity without ATP which increased in its presence (Ciechanover et al., 1978). When both fractions were mixed in the presence of ATP the degradation of globin increased six-fold (Ciehanover et al., 1978). Initial analysis of fraction 1 led to the identification of Ub, which was then postulated to play a key energy dependent role in interacting with other reticulocyte components to

facilitate protein degradation. Ubiquitin had been purified previously when Goldstein (1974) isolated thymopoietin and later incorrectly reported that this polypeptide with 'Lymphocyte-differentiating properties' was present in both prokaryotes and eukaryotes (Goldstein et al., 1975). Later additional studies into archaea and eubacteria would show that Ub was not 'ubiquitous' in these organisms, and thus the name ubiquitin was retrospectively inaccurate.

1.1.4 The role of Ubiquitin in protein degradation

Following the discovery of Ub the focus shifted to discovering its role in protein degradation. The conjugation of Ub to protein substrates was discovered by radiolabelling purified Ub with Iodine and observing its interactions with other reticulocyte components (Ciechanover et al., 1980). The covalent attachment of one or more Ub subunits indicated that proteins required modification or 'labelling' for its degradation to be triggered (Hershko et al., 1980). Once Ub had been identified as an essential component involved in protein degradation across eukaryotes, patterns regarding its structural binding and site of attachment were discovered.

Previous studies had found that a small protein (later to be identified as Ub) conjugated to Histones (Goldknopf and Busch, 1975), this attachment was then found to consistently occur between the C-terminal Gly of Ub and the amino (NH_2) group of an internal Lys of the histone (Hunt and Dayhoff, 1977). The isopeptide bond formed between Ub and

its substrate was also recognised to be identical to those between polyubiquitin subunits (Chau et al., 1989; Hershko and Heller, 1985). The specific nature of Ub attachment at Lys^{48} and then the subsequent polyubiquitination at this site was confirmed when Lys^{48} was mutated to Arg^{48}, preventing the formation of a polyubiquitin chain, thus only monoubiquitination was observed (Chau et al., 1989). The half-life of polyubiquitinated β-galactosidase (β-gal) was short, and differed from the monoubiquitinated mutant β-gal which did not degrade (Chau et al., 1989). Poylubiquitination, an energy intensive process, had thus been identified as a requisite for intracellular protein degradation. Although the causal modification, the addition of Ub to a substrate, for protein degradation had been identified, the mechanisms controlling conjugation of Ub and the subsequent recognition for degradation remained elusive.

1.1.5 The specificity and targets of the ubiquitin enzymes

To find other components required for ubiquitination, Ciechanover and Hershko returned to the initial two reticulocyte lysate fractions extracted. The fraction containing Ub (Fraction 1) required the high salt eluate (fraction 2) to drive any subsequent proteolytic activity (Ciechanover et al., 1978). Closer analysis of the reaction between Ub and fraction 2 in the presence of ATP, highlighted how Ub was first adenylated and 'activated' by an enzyme later termed E1 before conjugation could occur (Ciechanover et al., 1981; Haas et al., 1982). This E1 activator was later purified and found to bind to Ub through a thiolester bond (Ciechanover et al., 1982). The involvement of two additional enzymes involved in this

Ub-driven protein degradation was hinted by Ciechanover et al. (1981) and these were also found in fraction 2 of the reticulocyte lysate.

The following two enzymes found in the reticulocyte lysate work in conjunction with Ub. Once activated Ub is transferred to a specific Cys residue on small proteins, E2 conjugating enzymes, this E2 transfers the Ub to protein substrates (Hershko et al., 1983). The E2 conjugating enzyme requires the E1 activating enzyme to maintain its functionality, similarly it also requires ATP to accept Ub through formation of a high energy thiolester bond (Hershko et al., 1983). Once ubiquitinated proteins were found to degrade via an ATP assisted protease, later referred to as the 26S proteasome (Hough et al., 1986). Later work on E2s by Jentsch et al. (1987) highlighted the first physiological function of an E2 enzyme, when RAD6 a well-known DNA repair protein was isolated from *Sacchromyces cerevisiae* and discovered to catalyse the covalent attachment of Ub to yeast protein substrates. Using sequence similarities other E2s were discovered in *S. cerevisiae* including the cell cycle regulator CDC34 (Goebl et al., 1988), all E2s were later discovered to contain a similar 150 amino acid UBC domain and more than 40 have now been discovered in the human genome (Valimberti et al., 2015; van Wijk and Timmers, 2010). In Arabidopsis 37 proteins contain a Cys residue within a possible UBC protein domain (Bachmair et al., 2001), and recently one of these UBC22 was discovered to be involved in gametophyte development in female plants (Wang et al., 2016).

The final enzymatic group involved in the Ubiquitin Proteasome System (UPS) are the E3 enzymes. In addition to the E2 enzyme, conjugation of

Ub to a protein substrate requires the presence of an E3 enzyme (Hershko et al., 1983). The collective presence of E1, E2 and E3 enzymes along with Ub and ATP were shown to be imperative for protein degradation to occur (Figure 1.1). E3 ligases bind with specific protein substrates, a process which is associated with the conjugation of Ub to the substrate protein (Hershko et al., 1983). More information regarding the specific role of the E3 enzymes was later revealed by Hershko et al. (1986), who showed a correlation between the rate of degradation of a protein substrate and the binding efficiency of the E3 ligase. It was also discovered that modifications at the amino group of residues could render proteins more or less likely to undergo degradation (Hershko et al., 1986). This therefore demonstrated that these E3 ligases regulate which proteins are broken down dependent on their amino acid conformation. Later Hershko and Ciechanover (1998) would describe the E3s 'as an enzyme that binds, directly or indirectly, specific protein substrates and promotes the transfer of ubiquitin, directly or indirectly, from a thiolester intermediate to amide linkages with proteins or polyubiquitin chains' .

This discovery that the fate of a protein could in effect be dependent on its amino acid sequence raised new questions. The search was now on for two things; first, the discovery of more E3 ligases and their similarities and differences between organisms and, secondly, the rules dictating the specificity between an E3 ligase and a protein substrate which determines its degradative fate.

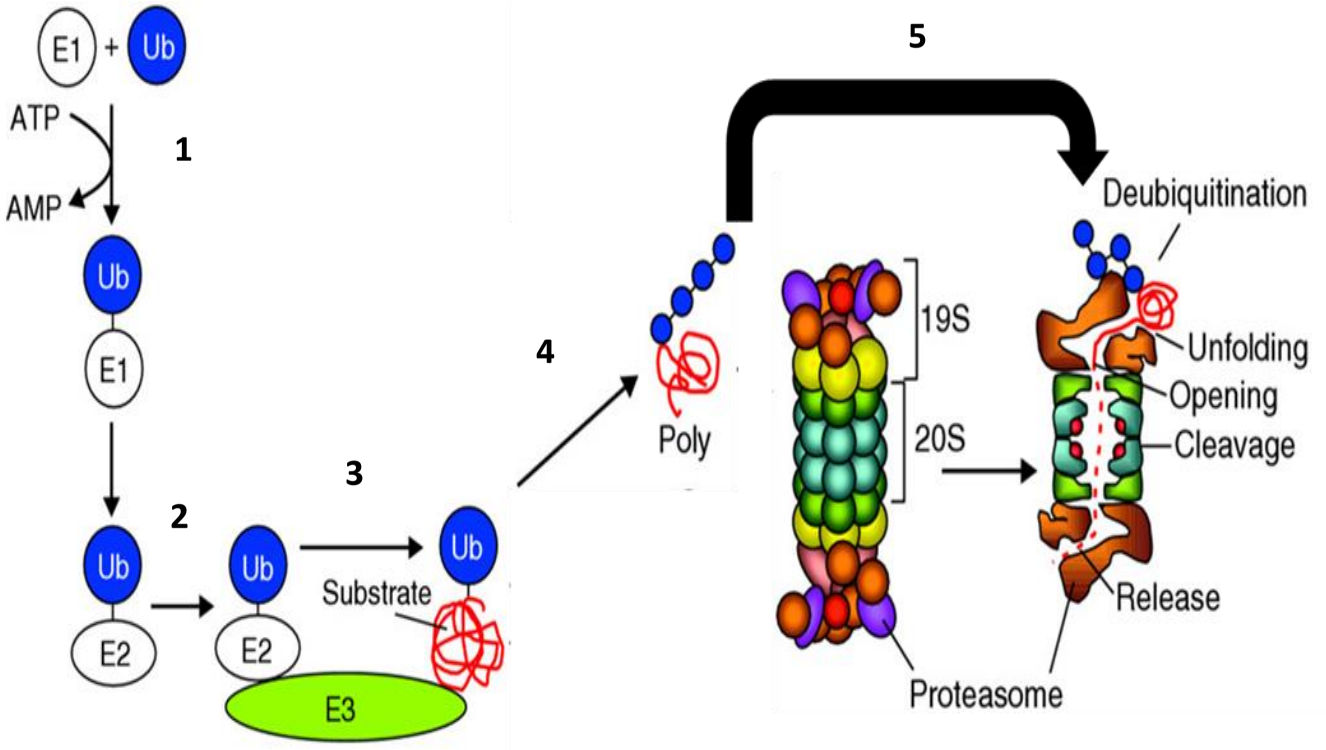


Figure 1.1 Schematic of the ubiquitin proteasome system. 1. Ubiquitin (Ub) is activated by the E1 activating enzyme in an ATP dependent reaction before; 2. Ub is conjugated to a protein substrate in another high energy reaction; 3. Conjugation requires the presence of an additional E3 recognition enzyme which recognises destabilising N-terminal residues; 4. Multiple Ub sub-units can bind forming a poly-Ub tail which is recognised by the; 5. 26S proteasome which initiates degradation and the substrate protein unfolds as it enters. The 26S proteasome is composed of a catalytic 20S and a regulatory 19S component (Adapted from Kaiser and Huang, 2005).

1.2 Selective protein degradation via the N-end Rule Pathway

1.2.1 The discovery of the N-end Rule Pathway

The discovery of Ub and its significance accelerated research into proteolysis in subsequent years. A major breakthrough made in the Varshavsky lab would answer the unsolved question regarding the specificities controlling the recognition of substrates by an E3 ligase. Following the discovery of Ub and its cascading enzymes the identity of only a few E3 ligases were available, therefore an approach to investigate this relationship would focus on the protein substrates themselves. It was proposed that perhaps proteins which share similar amino acid sequences might have a similar lifespan *in vivo* (Bachmair et al., 1986). The development of Nt fusion proteins with differing Nt amino acid residues highlighted the specific control of protein half-life *in vivo* by this determinant (Bachmair et al., 1986). Yeast ubiquitin conjugated to β-galactosidase (β-gal) creates a Gly-Met junction between each protein, commonly observed also when sub-units of Ub aggregate forming a poly-Ub chain (Ozkaynak et al., 1984). Deubiquitination efficiently occurs after the last Gly residue of Ub in yeast exposing the Nt-Met of β-gal, in an attempt to prevent deubiquitnation the Nt amino acid was mutated (Bachmair et al., 1986). Regardless of the different Nt residue (except Pro) deubiquitination still occurred, however the resultant stability of these mutant β-gal proteins varied from minutes to several hours (Bachmair et al., 1986).

The discovery that a single amino acid at the N-terminus of a protein can dramatically alter the fate of a protein (i.e. 'N-end rule'), brought some unexpected simplicity to the complexity of protein degradation studies. When an Nt residue requires no modification for its recognition by an E3 ligase it is regarded as a primary destabilising residue, these are either positively charged (Type 1) including Arg, Lys and His or bulky and hydrophobic (Type 2) including Phe, Leu, Trp, Ile and Tyr (e.g. Sriram et al., 2011). Two branches of the N-end rule pathway have now been discovered, the 'classic' or Arg N-end rule pathway where destabilising residues undergo modification with the addition of Arg or are recognised directly by E3 ligases. The second and more recent is the Acetylation or Ac-N-end rule pathway, where some Nt residues are acetylated (Figure 1.2) (Hwang et al., 2010; Shemorry et al., 2013).

1.2.2 The components and working of the Arg/N-end rule pathway

The Nt amino acid of all nascent proteins begins with translation initiator Met, which on initial experimental data was deemed to be a stabilising residue (Bachmair et al., 1986). The removal of this Nt-Met can be facilitated by Methionine Aminopeptidase (MAP), an enzyme first purified in prokaryotes but later in yeast a eukaryote (Ben-Bassat et al., 1987; Li and Chang, 1995), when the residues adjacent to the N-terminus have small side chains (Sherman et al., 1985). These residues with small side chains (excluding Met) were identified as being stabilising, therefore indicating that the removal of Nt-Met does not in itself destabilise a

protein (Li and Chang, 1995). This finding and the discovery that MAPs are conserved across kingdoms (Giglione et al., 2004), indicated that modifications to the N-terminus of proteins could be a stimulus for degradation in both prokaryotes and eukaryotes. Bachmair et al. (1986) postulated correctly that some Nt amino acids could be sites of action for aminoacyl-tRNA-transferases, or sites for amino acid modifications, which could enable these differences in half-life between proteins with different 'N-degrons' (Varshavsky, 1991). N-degrons comprise of a destabilising amino acid at the N-terminus, a Lys amino acid in the vicinity of the N-terminus which acts as a host for the attachment of a multi-Ub chain (Bachmair and Varshavsky, 1989; Chau et al., 1989), and the presence of an unstructured sequence which initiates degradation (Prakash et al., 2004).

A wide range of post translational modifications (PTMs) occur to amino acid residues which change their reactivity and the functionality of the protein they compose. A significant PTM was discovered in *E. coli*, this was facilitated by the aminoacyl tRNA transferases, capable of transferring specific amino acids (Phe and Leu) to an acceptor protein in the absence of ribosomes (Kaji et al., 1965).The aminoacyl tRNA transferases were later found to be important in protein recycling and then the N-end rule pathway (Goldberg, 1971; Kwon et al., 1999).

The diversity of these tRNA transferases was explored in different organisms including mammals, where Arg, Glu, Met, Trp and Gly tRNA transferases were discovered in rat liver cells (Kaji, 1968). Interestingly Arg, a highly destabilising Nt residue, was noted to be fused to protein N-

termini more than other residues (Kaji, 1968). The discovery of the propensity of Arg tRNA transferases to add Arg (a primary destabilising residue) to N-termini bearing Asp or Glu (secondary destabilising residues) (Soffer, 1980), highlighted the possible hierarchical nature of the N-end rule pathway. It was later discovered that Nt-Asp and Nt-Glu could be formed by the deamidation of Asn and Gln (Tertiary destabilising residues) (Baker and Varshavsky, 1995; Grigoryev et al., 1996) (Figure 1.2). This necessity for some Nt residues to be modified before substrates can be targeted for proteasomal destruction, partly explains why the variation in half-life differs between different destabilising residues with a tertiary destabilising which requires two PTMs degrading slower than a primary destabilising residue which requires no PTMs. N-terminal Cys is also a tertiary destabilising residue in mammals and plants, this residue will be the focus of this study so will be covered individually in section 1.3.

The N-end rule pathway was initially discovered in yeast, however it has also been found in bacteria, mammals and plants.

1.2.3 The N-end rule pathway in bacteria

The bacterial N-end rule pathway differs from those in eukaryotes due to the absence of Ub (Finley et al., 1988; Tobias et al., 1991). Despite the lack of Ub, the presence of the N-end rule and its similar structure in prokaryotes and eukaryotes indicated that this system is ancient, and later acquired a UPS in eukaryotes (Gibbs et al., 2014a). A ClpS protein

is involved in stimulating the degradation of substrates by modulating the ClpAP protease, a ring-like complex similar to the 26s proteasome which recognises substrates and facilitates degradation (Roman-Hernandez et al., 2009; Wang et al., 2007). A combination of the ClpA and ClpS proteins recognize Leu, Phe, Trp and Tyr as primary N-degrons, Asp and Glu can be treated as secondary degrons undergoing leucylation, whilst Arg, Lys and Met undergo Leu/phenylalanation through action of aminoacyltransferases (Erbse et al., 2006; Kaji et al., 1965; Wang et al., 2007). The hierarchical nature of bacterial N-end rule pathway was also demonstrated by Ninnis et al. (2009), who found that the first substrate discovered, putrescine aminotransferase (PAT), requires modification (attachment of a primary destabilising residues to Nt-Met) by Leu/Phe aminotransferase before ClpS-AP driven degradation (Graciet et al., 2006).

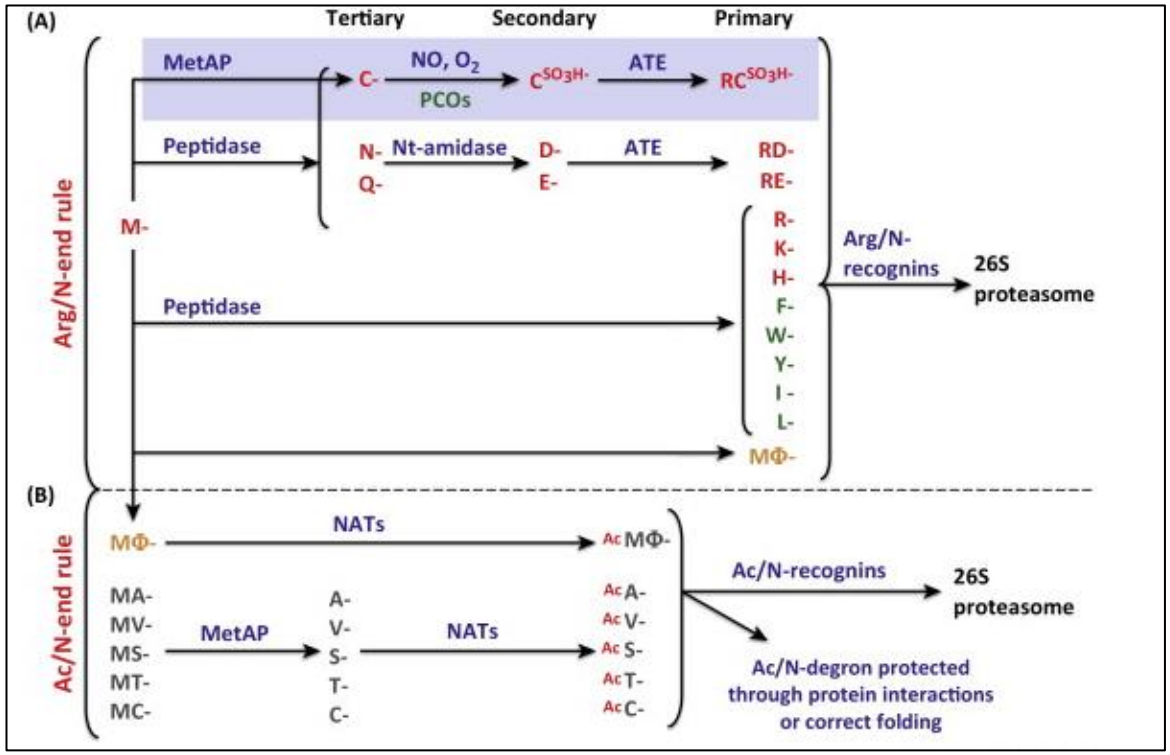


Figure 1.2 N-terminal (Nt) residues degraded by the Arg and Ac branch of the N-end rule pathway. (A) The Arg-N end rule pathway comprises of tertiary, secondary and primary destabilising residues which are either type 1 – basic (red) or type 2 – hydrophobic (green). Nascent proteins initiating with Nt-Methionine (M) can be cleaved by methionine aminopeptidases (MetAP), or by other peptidases which cleave the protein internally, creating a new Nt residue. Tertiary destabilising residues of the Arg-N-end rule include Cys (C) which is oxidised by oxygen and nitric oxide (NO) and in plants Plant Cysteine Oxidases (PCOs), Asn (N) and Gln (Q) which require specific deamidation (NTAN/NTAQ). Secondary destabilising residues, oxidised Cys (C^{SO_3H}), Asp (D) or Glu (E) can then be arginylated by arginyl tRNA transferases (ATEs). Primary destabilising residues are directly recognised by Arg-N-recognins (E3 ligases) for destruction by the 26S proteasome. This branch of degradation occurs in eukaryotes except yeast. Nt-Met (M) has also been discovered as destabilsing when followed by a type 2 residue (Φ). (B) The Ac-N-end rule pathway can also recognise MΦ motifs when acetylated (Ac) by Nt acetyltransferases (NAT's). Similarly to the Arg N-end rule other residues require Met excision by MetAPs before Nt acetylation and recognition by AC-N-recognins for degradation (Gibbs et al., 2014a).

1.2.4 The N-end rule pathway in Sacchromyces cerevisiae

The initial enzymatic components of the N-end rule pathway discovered in *Saccharomyces cerevisiae* include the arginylation enzyme arginyl tRNA transferase (ATE1) (Balzi et al., 1990), and the N-recognin Ubr1 (E3 ligase) (Bartel et al., 1990). Ubr1 binds arginylated Nt-Asp and Nt-Glu, created through deamidation of tertiary destabilising residues Asn and Gln by N-terminal amidase (NTA1) (Baker and Varshavsky, 1995; Bartel et al., 1990). Conjugation of Ub is previously facilitated by Rad6 (E2 enzyme) allowing N-degron recognition to occur (Dohmen et al., 1991). As well as the arginylation dependent N-end rule pathway (Nt-arginylated), the acetylation (Ac) dependent N-end rule pathway (Ac/N-end rule pathway) also exists in yeast (Hwang et al., 2010).

S. cerevisiae is thought to conduct Nt-acetylation on 50% of its proteins (Lee et al., 1989), compared with over 80% in higher eukaryotes, including humans (Arnesen et al., 2009; Shemorry et al., 2013). Previously it was considered that acetylation of proteins protected them against degradation (Persson et al., 1985). However, Hwang et al. (2010) demonstrated how Nt-acetylation of certain amino acids creates degradation signals accessible to the UPS, such as Ala, Val, Ser, Thr and Cys recognised by the N-recognin Doa10 (Figure 1.2). Recently it was discovered that there was some cross-over between the Arg and the Ac branches of the N-end rule pathway, proteins bearing a non-acetylated Nt-Met proceeded by a hydrophobic residue (excluding Met) were targets for Ubr1 in the Arg/N-end rule pathway and Doa10 in the Ac/N-end rule

pathway (Kim et al., 2014b). This discovery greatly increases the potential substrate diversity of the N-end rule pathway in yeast.

1.2.5 The N-end rule pathway in mammals

E3 ligase N-recognins of the mammalian N-end rule pathway incorporate a 70 residue zinc finger domain with two zinc fingers, which together target the binding site of positively charged N-degrons (Tasaki et al., 2009). The N-end rule machinery within mammals is similar to that in yeast but includes two specific N-terminal amidases NTAQ1 and NTAN1 for Gln and Asn respectively (Grigoryev et al., 1996; Wang et al., 2009), a single Arg tRNA transferases encoded by ATE1 which exists in two forms created by alternative splicing, ATE1-1 and ATE1-2 differing by a 43 residue region (Kwon et al., 1999) and lastly a UBR box described as the 'substrate recognition domain' (Tasaki et al., 2012). The deletion of ATE1 gene in mouse embryos was found to cause cardiovascular failure, highlighting the physiological importance of these components (Kwon et al., 2002). Seven individual UBR proteins are known (UBR1-UBR7), however only four of these UBR proteins function as N-recognins (UBR1, UBR2, UBR4 and UBR5) (Sriram et al., 2011). The identified physiological substrates which are targeted by N-recognins include three regulators of G protein signalling (RGS4, RGS5 and RGS16 (Lee et al., 2005), which when combined with the E2 enzymes HR6A and HR6B trigger polyubiquitination and degradation (Kwon et al., 2003).

1.2.6 The N-end rule pathway in plants

Arabidopsis thaliana encodes two almost identical Arg tRNA transferases, AtATE1 and AtATE2 (Yoshida et al., 2002). AtATE1 was originally discovered in a study examining why mutants lacking this gene had delayed leaf senescence, an important physiological process controlled by the N-end rule pathway (Yoshida et al., 2002). Recently ATE was investigated in *Physcomitrella patens* (moss) where it was found to be encoded by a single gene (Schuessele et al., 2016). Arginylated protein substrates have been identified in moss by mass spectrometry using ATE interacting proteins (Hoernstein et al., 2016). In plants there are two established N-recognins, PRT1, which shares no homology to yeasts Ubr1 or mammalian UBR1, lacking the Zinc-finger domain (UBR box) (Potuschak et al., 1998; Stary et al., 2003). PRT1 was identified in a genetic screen for Arabidopsis genes involved in Ub-dependent proteolysis, dihydrofolate reductase (DHFR) in mutant Arabidopsis *prt1* lines was found to be stabilised (Bachmair et al., 1993). PRT1 was found to only recognise hydrophobic primary destabilising residues (Stary et al., 2003). A second N-recognin PRT6, shows significant 'sequology' (Varshavsky, 2004) to Ubr1 in yeast, containing a UBR box and recognising basic residues (Garzon et al., 2007) (Figure 1.2).

Mutants of *Arabidopsis thaliana* lacking PRT1 and PRT6 prevent the respective degradation of both hydrophobic and basic amino residues (Garzon et al., 2007). Nothing is known about substrates of PRT1 in plants, but it has been shown to compensate for Ubr1 in yeast (Stary et al., 2003). In the absence of the two N-recognins PRT1 and PRT6, some

proteins bearing the Nt amino residue Leu continues to be degraded, suggesting that other E3 ligases are present in plants (Garzon et al., 2007). Several proteins within plants contain UBR-like sequences similar to known N-recognins in yeast, mammals and bacteria, which could have a dual function in plants. One example is a large 560 kDa protein, the aptly named BIG (originally Dark Over Expression of CAB1 (DOC1)), which was suggested by Graciet and Wellmer (2010) as potentially functioning as an N-recognin besides numerous other roles in auxin transport (Gil et al., 2001) and photomorphogenesis (Kanyuka et al., 2003).

Following the discovery of the plant E3 ligases it was important to characterise the phenotype of these mutants, revealing their role in plant development and functions. Seed germination is influenced by the N-end rule pathway, with seed-after ripening and ABA sensitivity being reduced in *prt6* and *ate1ate2* mutants (Holman et al., 2009). Additionally both shoot development and leaf senescence are repressed in mutants lacking ATE and/or PRT6, highlighting their involvement in these processes (Graciet et al., 2009; Yoshida et al., 2002). Leaves of *prt6* mutants appear more lobed and curvy in comparison to those of the wild-type Col-0 ecotype but similar to *ate1ate2* leaves (Graciet et al., 2009). The increased tolerance of *prt6* mutants to flooding has been reported in Arabidopsis (Gibbs et al., 2011; Licausi et al., 2011) and barley (Mendiondo et al., 2016).

1.3 Oxygen and nitric oxide sensing in the N-end rule pathway

1.3.1 The relevance of Cys in biological systems

Cysteine is regarded as one of the least abundant amino acids but one of the most reactive and biologically important. Thought to be one of the last amino acids to develop, Cys has been reported to be accumulating since the Archaean period (>3000 Ma) in all domains of life (Brooks, 2001; Jordan et al., 2005). This consistent accumulation of this residue within proteins indicates that its role in proteins has continued to grow in importance as organisms become more complex. The reactivity of any individual Cys residue will depend on its position in the amino sequence. For example whether it is in a hydrophobic or hydrophilic region, and whether the Cys is exposed or buried according to tertiary structure (Marino and Gladyshev, 2012).

The plasticity of reactions which can occur to the thiol (SH) group of Cys is large, including forming disulfide bonds (Cys-SS-Cys), reacting to form metal complexes and undergoing nitrosylation (SNO) and/or oxidation (Figure 1.3). The low pK_a of the thiol including strong nucleophicity of the sulfur (S) atom makes Cys prone to oxidation, making it a common site of post translational modification (PTM) in proteins (Kim et al., 2014a). In biological systems redox reactions involve the oxidation of a Cys thiol up to sulfenic acid (SOH), however the existence of sulfenic acid is unstable and can be reduced back to a thiol, further oxidised to more stable sulfinic/sulfonic (SO_2H/SO_3H) acid or form a disufhide by reacting with a glutathione (GSH) or another thiol group (Akter et al., 2015;

Gupta and Carroll, 2014). The redox or nitrosylative oxidation of Cys-thiols within a cell is largely determined by the abundance of reactive oxygen or nitrogen species (ROS/RNS), and concomitantly the efficiency of the antioxidant defence (Murray and Van Eyk, 2012).

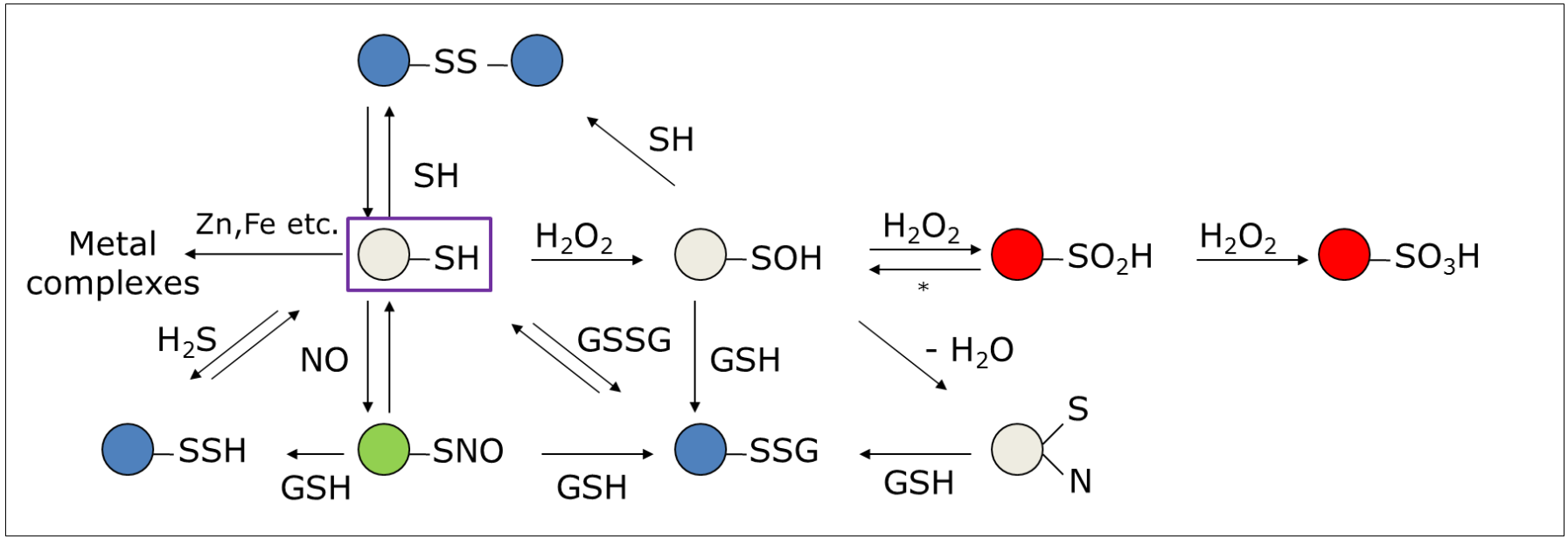


Figure 1.3 The diversity of oxidative post translational reactions occurring to cysteine. Cysteine (Cys, spheres) has a reactive thiol (SH) side chain (in purple box). This thiol group can react with another Cys-thiol to form a disulfide (SS) (all disulfides are highlighted in blue). Cys can also react with hydrogen peroxide (H_2O_2) to initally form a Cys-sulfenic acid (SOH) which can react with more H_2O_2 forming Cys-sulfinic (SO_2H) and subsequently Cys-sulfonic acid (SO_3H) (red spheres indicate highly oxidised species). * Sulfiredoxins a group of Peroxiredoxins have recently been shown to reduce Cys-sulfinic to Cys-sulfenic acid (Chang et al., 2004; Rey et al., 2007). Cys-sulfenic acid can be reduced by glutathione (GSH) to form a mixed gutathionylated thiol disulfide (SSG), alternatively Cys-sulfenic acid can lose water in a condensation reaction to form a Cys-sulfenamide (N-Cys-S) a protective oxidised species initially identified in mammalian protein tyrosine phosphatase 1B (PTP1B) (Salmeen et al., 2003). The cyclised Cys-sulfenamide can react with GSH to form SSG. Alternatively Cys can be nitrosylated by NO forming a Cys-nitrosothiol (SNO) (nitrosylated species highlighted in green), which can react with GSH to form a sulfhydrated thiol (SSH) which can also be formed from reaction of Cys with hydrogen sulfide (H_2S). Cys can also react with metals to form various oxidised metal complexes. Image adapted from Chung et al. (2013).

1.3.2 The oxidative degradation of Nt-Cys substrates in the Arg/N-end rule pathway

In plants and mammals, proteins bearing Nt-Cys can be tertiary destabilising residues. This differs from yeast where Cys has been shown to be stabilising (Kwon et al., 1999). Oxidation of Cys in the presence of both oxygen and Nitric oxide (NO) permits arginylation by Arg tRNA transferases, which is subsequently recognised by an N-recognin E3 ligase (Figure 1.2) (Gibbs et al., 2011; Gibbs et al., 2014b; Hu et al., 2005). It remains unclear how Cys is oxidised *in vivo* and whether Cys changes from bearing a thiol (Cys-SH) group to a Cys-sulfinic (Cys-SO_2H) or Cys-sulfonic acid (Cys-SO_3H) group (Hu et al., 2005; Gibbs et al., 2014). Initial studies investigating how proteins bearing an Nt-Cys were degraded used the mammalian substrate RGS4. Davydov and Varshavsky (2000) suggested that Cys requires extra modification or the existence of a specific Cys Arg tRNA transferase for arginylation to occur. Subsequently hypotheses were made that oxidised Cys is recognised by the existing Arg tRNA transferases as it mimics Asp (Figure 1.4) (Kwon et al., 2002). Edman degradation results from RGS4 purified from rabbit reticulocyte cells could not identify Cys at position 2, indicating that the Cys has been modified before the sample was alkylated and prepared for sequencing (Kwon et al., 2002).

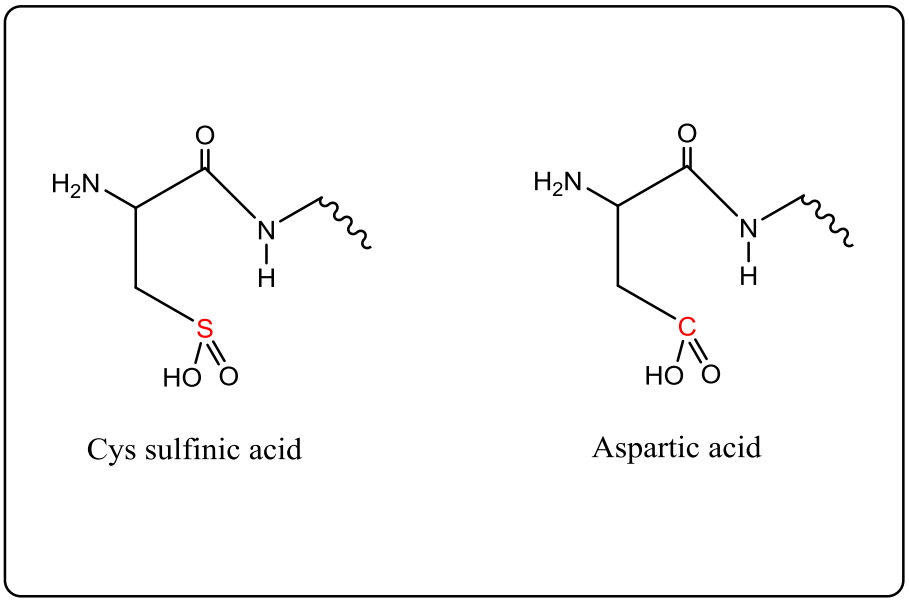


Figure 1.4 The structural similarity of Cys-sulfinic acid and aspartic acid. These two structures only differ by one atom the sulfur (S) and carbon (C) atoms highlighted in red.

1.3.3 Cysteine oxidation in the mammalian Arg/N-end rule pathway

The development and use of N-end rule mutants lacking key enzymes such as ATE and E3 ligases, has helped to reveal the biochemical and physiological importance of Cys oxidation and nitrosylation in mammals. Hu et al. (2005) demonstrated that oxidation of Cys is required before arginylation, three synthetic peptides produced with differing N-termini of either Cys, Asp or Cys-sulfonic acid were incubated with mouse ATE1. Peptides bearing Nt-Cys could not be arginylated, but when Cys was oxidised (Cys-SO_3H) it was efficiently arginylated like Asp in a NO-dependent manner (Hu et al., 2005). Oxygen-dependent proteolysis of RGS proteins was later discovered to be a cellular function of the Arg N-end rule pathway in cardiomyocytes and developing hearts (Lee et al., 2012). Embryonic failure was observed when mouse embryo cells were

cultured in an anoxic (0.1% O_2) environment due to the stabilisation of RGS4, similarly to the result in ATE mutant cells (Lee et al., 2012).

Nitric oxide (NO) is a reactive nitrogen species (RNS) which has only been studied in detail in the past few decades. The discovery that NO was the endothelium-derived relaxing factor, produced as a signalling molecule involved in vasodilation (Palmer et al., 1988; Palmer et al., 1987) led many researchers to investigate if NO was involved in signalling other cellular processes. The reactivity of NO with proteins, particularly those bearing Sulfur (S) such as Cys, can alter their function, and in the context of the N-end rule pathway also their stability. Experiments exogenously supplying NO or an inhibitor of NO production enabled the necessary presence of NO during Cys oxidation to be realised (Hu et al., 2005). In the presence of the NO synthase inhibitor N^G-monomethyl-L-arginine (LMMA) and the NO scavenger 2-(4-carboxyphenyl)-4,4,5,5-tetramethylimidazoline-1-oxyl-3-oxide (cPTIO), there was a correlative increase in the mammalian N-end rule substrates RGS4, RGS5 and RGS16 in mouse embryos (Hu et al., 2005). RGS4 is degraded in the presence of NO, inducing signalling increasing myocardial growth, conversely in the presence of NO synthase inhibitor *N*G-nitro-l-arginine methyl ester (L-NAME) RGS4 degradation was slowed (Jaba et al., 2013). Both of these studies report that the degradation of regulatory proteins bearing Nt-Cys is inhibited in the absence of NO, thus NO is pivotal in the oxidation of Cys. Although the necessity of NO was identified for efficient arginylation and degradation of Nt-Cys substrates, the results from Hu et al. (2005) don't reveal the specifics of NO

involvement in the oxidation process, or exclude the possibility that a Cys-nitrosothiol could be the substrate of ATE.

1.3.4 Cysteine oxidation in the plant Arg/N-end rule pathway

The first identified physiological substrates of the arginylation branch of the N-end rule pathway in plants, with a reactive Cys at the N-terminus, are the group VII ethylene response factor transcription factors (ERFVIIs). The ERFVIIs belong to the APETALA 2 ERF family which includes the Rice (*Oryza sativa*) transcription factors SUB1-A, SNORKEL1 and SNORKEL2 (Hattori et al., 2009; Xu et al., 2006). The expression of SUB1-A in rice has been demonstrated to improve survival of rice when submerged, SUB1-A prompts the rice to adopt quiescent metabolism strategy which involves the suppression of Gibberellin-driven shoot elongation (Fukao et al., 2006). Alternatively Hattori et al. (2009) reported how the ERFVIIs SNORKEL 1 and 2 initiate the rapid elongation of rice shoots to raise aerial shoot organs out of the water. The rice transcription factors SUB1-A, SNORKEL1 and SNORKEL 2 similarly to the Arabidopsis ERFVIIs possess a highly conserved initial seven residue domain at the N-terminus (MCGGAIL), which raised questions over whether the evolutionary divergent transcription factors in Arabidopsis may have a similar function to their rice counterparts (Nakano et al., 2006).

The Arabidopsis ERFVIIs consist of five members, three of which are constitutively expressed, called RELATED TO APETALA (RAP) 2.12, 2.2

and 2.3 (Hinz et al., 2010; Nakano et al., 2006), and HYPOXIA RESPONSIVE ELEMENT 1 and 2 (HRE1/2) whose expression is induced by hypoxia (Licausi et al., 2010). Stabilisation of the ERFVIIs occurs when Cys at position two is replaced with the stabilising residue Ala, or in mutants lacking essential components of the N-end rule pathway (Gibbs et al., 2011; Licausi et al., 2011). Papers published by Gibbs et al. (2011) and Licausi et al. (2011) revealed that the Arg/N-end rule pathway in plants, through the stabilisation of the ERFVII substrates, acts as sensors of oxygen in the environment. The stabilisation of HRE1, HRE2 and RAP2.12, after subjecting plants to hypoxia or submergence, was shown to increase the recovery thus survival of the plant (Gibbs et al., 2011; Licausi et al., 2011). Interestingly *in vitro* experiments testing the stability of the rice submergence gene SUB1-A showed that this was not a substrate of the N-end rule pathway (Gibbs et al., 2011), subsequently flood tolerance of rice species was shown to depend on the induction of SUB1-A mRNA upon submergence (Singh et al., 2010).

The N-end rule mutant plants *prt6* and *ate1ate2* both show increased up-regulation of hypoxia (<10% O_2) response genes (e.g. PYRUVATE DECARBOXYLASE, PDC1; ALCOHOL DEHYDROGENASE, ADH1 etc.), caused by the stabilisation of ERFVIIs (Gibbs et al., 2011; Hinz et al., 2010; Licausi et al., 2011). The ERFVIIs are stabilised when the plant suffers hypoxia (<10% O_2), as ATE does not recognise un-oxidised Cys. Both HRE2 and RAP2.12 have been shown to induce the expression of some hypoxic genes when stabilised, including the anaerobic enzymes ADH1 and PDC1 (Gibbs et al., 2011; Licausi et al., 2011). Upon re-oxygenation the same ERFVIIs were quickly degraded and PDC1 and

ADH1 levels reduced, indicating the oxygen dependent stability of these genes (Gibbs et al., 2011; Licausi et al., 2011). Under normoxic conditions RAP2.12 interacts with the plasma membrane Acyl-CoA binding proteins (ACBP), during hypoxia the ACBP is unbound from the membrane and is relocated in the nucleus up-regulating the transcription of hypoxic response genes (Kosmacz et al., 2015; Licausi et al., 2011; Weits et al., 2014). It is hypothesised that the oxygen dependent binding of RAP2.12 and ACBP protects the Nt-Cys from oxidation under normoxic conditions, highlighting a mechanism of Nt shielding enabling the stabilisation of N-end rule substrates under normoxia (Gibbs et al., 2014a; Licausi et al., 2013).

In addition to oxygen the ERFVIIs can also sense NO and N-end rule mutants are insensitive to several NO-regulated processes (Gibbs et al., 2014b). Seed dormancy, stomatal closure and hypocotyl elongation are some of the NO regulated processes in plants (Yu et al., 2014). In Arabidopsis the N-end rule pathway mutant *prt6* is insensitive to these NO-driven processes unlike the wild-type Col-0. In the presence of NO *prt6* seeds remained dormant, *prt6* seedlings elongated their hypocotyl in the dark and leaf stomata remained closed (Gibbs et al., 2014b). Scavenging NO using cPTIO resulted in the stabilisation of the Met-Cys Nt-β-glucuronidase (MC-GUS) reporter protein in Arabidopsis seedlings, indicating that NO is required for the degradation of Nt-Cys N-end rule substrates (Gibbs et al., 2014b). A nitrate reductase mutant (*nia1nia2*) in a overexpressing HRE2 background stabilised HRE2, due to the lack of the predominant source of NO required for degradation (Gibbs et al., 2014b). When the same mutant was exposed to an NO donor *S*-nitroso-

*N*acetyl-DL-penicillamine (SNAP) HRE2 was destabilised (Gibbs et al., 2014b).

Both NO and oxygen are also important environmental cues controlling photomorphogenesis and seedling establishment (Abbas et al., 2015). Wild type etiolated Arabidopsis seedlings treated with hypoxia or cPTIO retain a closed apical hook when studied on agar, with N-end rule mutants behaving similarly (Abbas et al., 2015). Previously it was assumed that the absence of light in sub-soil environments was the cause of this positive adaptive trait which protects the stem-cell niche until emergence. Interestingly hook opening is not repressed in a sextuple mutant lacking all ERFVIIs and the E3 ligase *prt6* (*prt6erfvll)*, indicating that the ERFVIIs are responsible for this protective characteristic. An additional function of the ERFVIIs, and their ability to sense oxygen is the ability of seedlings to survive prolonged periods in the dark which is enhanced in hypoxic soils (Abbas et al., 2015).

An enzymatic route for oxidising the Nt-Cys of N-end rule substrates in plants followed the discovery of a family of five Plant Cysteine Oxidase (PCOs) enzymes (Figure 1.5) (Weits et al., 2014). Chemically considered as Cys dioxygenases, a member of the non-heme iron dependent oxygenases, the PCOs catalyse the oxidation of Cys to Cys-sulfinic acid although their kinetic characterisation has yet to be completed (White and Flashman, 2016). Oxygen is used as a co-substrate by the PCOs to facilitate the reaction from a Cys-thiol to Cys-sulfinic acid (Weits et al., 2014). The PCO proteins have been described as oxygen 'sensors', which are induced by RAP2.12 and operate as feedback regulators dependent

on the extent of hypoxia (Weits et al., 2014). Recently molecular evidence for the oxidation of Nt-Cys from synthetic peptides by PCO to Cys-sulfinic acid has been presented (White et al., 2016). Using mass spectrometry (MS) a ten residue peptide modelling the initial ten residues of Cys N-terminal RAP2.12, was shown to undergo a mass shift of +32 Da in the presence of PCO1 and/or PCO4 (corresponding to 2x O atoms) (White et al., 2016). Following this confirmation White et al. (2016) showed that arginylation occurred on Cys-sulfinic acids produced after PCO exposure, but also the further oxidised Cys-sulfonic acid when incubated with ATE.

Despite an understanding that PCOs, NO and oxygen play a role in Cys oxidation, their interactions and their behaviour in different environmental conditions remain unclear. It is possible that both nitrosylation and PCO driven systems controlling the degradation of Nt-Cys N-end rule substrates could co-exist. However, the role of NO in relation to PCO-catalysed oxidation of Cys to Cys-sulfinic acid is unknown (Weits et al., 2014; White et al., 2016). The influence of NO on the PCO-Cys interaction needs to be examined *in vivo*, in order to define their relative contribution to Cys oxidation.

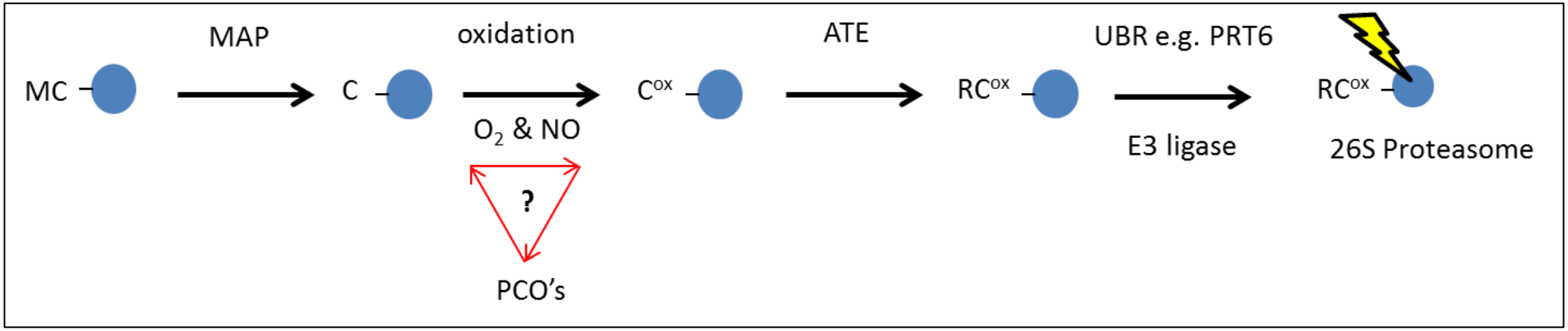


Figure 1.5 A schematic showing the cysteine oxidation branch of the Arg/N-end rule pathway in plants. Proteins (blue spheres) bearing Nt Met-Cys (MC) residues are tertiary destabilising residues for the Arg/N-end rule pathway. Nt-Met can be cleaved by methionine aminopetidases (MAP) to create an Nt-Cys which is oxidised to an unknown oxidised state through the interaction of oxygen, nitric oxide (NO) and the Plant Cysteine Oxidases (PCOs). The oxidised Cys (C^{ox}) can be arginylated by Arg t-RNA transferases (ATE) adding a primary destabilising Arg to the N-terminus (RC^{ox}) which is recognised by an E3 ligase such as PROTEOLYSIS6 (PRT6) which enables ubiquitination and subsequent degradation by the 26S proteasome

1.4 The role of oxygen and nitric oxide as plant signalling molecules

1.4.1 Reactive oxygen and nitrogen species

Molecular oxygen and NO are crucial gases which are both used and produced in plants and often present together. Within plants oxygen is present in numerous states; in the diatomic form O_2, in one of many reactive oxygen species (ROS), combined with other elements in an oxidative form (e.g. NO) and locked in the biochemical structural materials of the plant (e.g. carbohydrates, proteins etc.). Respiration and photosynthesis are the predominant biochemical and metabolic processes involving oxygen in plants. ROS species such as hydrogen peroxide (H_2O_2), superoxide (O_2^-) and hydroxyl radicals ($HO\cdot$) act as important signalling molecules in plants, for example they are involved in responses to symbiotic events (Marino et al., 2011), environmental stress (Swanson and Gilroy, 2010) and facilitating programmed cell death (PCD) (Torres et al., 2005). NO is a reactive nitrogen species (RNS) which is involved in similar processes to ROS, also acting as a signalling molecule in plants. The role of NO within plants has been researched in detail over the last two decades, but much remains unclear (Gupta et al., 2011; Mur et al., 2013; Yu et al., 2014). The short half-life of RNS due to their reactivity makes the exact roles they play difficult to define. The short-lived presence or RNS at important developmental stages (Airaki et al., 2015) and infection/stress sites is sensed by the plant (Corpas et al., 2009),

triggering a response in biochemical and physiological activity which has been widely reviewed (e.g Gupta et al., 2011; Yu et al., 2014) .

1.4.2 NO production strategies in plants

Production of plant NO may occur through at least seven possible mechanisms, divided into the largely unknown oxidative processes and the partially documented reductive strategies (Gupta et al., 2011). In mammals NO production is oxidative, involving nitric oxide synthase (NOS). Production of NO involves three different isoforms of nitric oxide synthase (NOS), located in the cytosol, mitochondria and in the endothelial tissues (Alderton et al., 2001). Arg oxidation through the activity of NOS produces citrulline during a two-step process which also releases NO (Palmer et al., 1988). A homologue of NOS has not yet been identified in plants, despite the observed decrease in NO production when animal inhibitors of NOS are present (Corpas et al., 2006). NO production via Arg-driven polyamine synthesis and a ROS-dependent hydroxylamine reaction have been identified (Rumer et al., 2009; Tun et al., 2006). However, much of the biochemical workings and physiological significance of these NO producing processes remain unknown.

The most characterised reductive nitrogen strategy, involves nitrate reductase (NR) and occurs in the cytosol (Yamasaki and Sakihama, 2000). NO production rate via NR differs between nitrate and nitrite depending on the NADH availability. Nitrite is the more efficient substrate requiring less NADH oxidation for NO production (Yamasaki and

Sakihama, 2000). Capacity for NO production in plants can be organ specific, with root mitochondria capable of reducing nitrite to NO through electron transfer, an ability absent in leaf mitochondria without NR activity (Gupta et al., 2005). Additional NO producing enzymes are active in the roots, nitrite-NO reductase (Ni-NOR) is active in the plasma membrane (Stöhr et al., 2001). Meyer and Stöhr (2002) postulated how Ni-NOR may be able to sense nitrate levels in the roots, supported by evidence documenting its involvement in sensing root infection by mycorrhizal fungi (Moche et al., 2010). However, the molecular cloning of this enzyme has not been achieved.

1.4.3 Function of NO and ROS sensing in plant development

Nitric oxide bioactivity plays an important role initiating developmental processes in plants. The accumulation of RNS has been observed in germinating seeds (El-Maarouf-Bouteau and Bailly, 2008). Germination is triggered by NO in the presence of light, which also initiates de-etiolation and inhibits hypocotyl elongation (Beligni and Lamattina, 2000). Interactions between NO and ROS during germination are important in regulating seed dormancy, including regulating the activity and biosynthesis of the important plant hormones Abscisic acid (ABA) and Gibberelic acid (GA) (El-Maarouf-Bouteau and Bailly, 2008). H_2O_2 mediates changes in the levels of both plant hormones upon signals from NO, acting to increase catabolism of ABA whilst promoting GA synthesis (Liu et al., 2010). NO participates in this shift from the dormant state of high ABA:GA to a high GA:ABA ratio, higher GA levels enable endosperm

weakening and promote embryo growth causing germination (Groot and Karssen, 1987). The aleurone cell layer in Arabidopsis represents the endosperm within the seed coat. The endosperm has different functions acting as a barrier to prevent radicle growth and to provide sugars to the growing seedling (Penfield et al., 2004). This endosperm tissue shows an ability to sense NO, discovered when the testa was removed from the seed and dormancy breaking was accelerated, this result was subsequently enhanced when in an NO rich environment (Figure 1.6) (Bethke et al., 2007). The NO-driven biochemical pathway regulating germination was later discovered to be controlled by the ERFVIIs. Gibbs et al. (2014b) found that developmental responses such as germination, stomatal closure and hypocotyl elongation are coordinated through the degradation of the ERFVIIs in Arabidopsis.

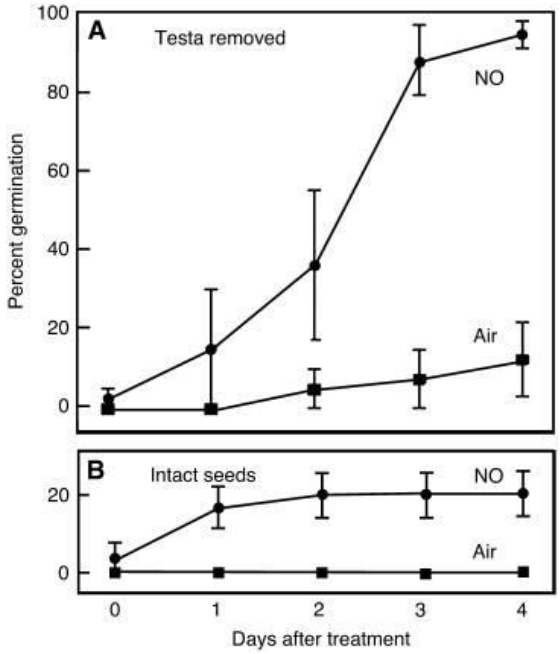


Figure 1.6 Arabidopsis seeds show enhanced germination when exposed to nitric oxide (NO) when compared to pure air. (A) Germination of testa removed seeds exposed to NO was significantly different to seeds exposed to purified air for the same duration after 2 days. (B) Germination of intact seeds (with testa) was observed when exposed to NO but did not occur in the presence of pure air. Germination is reduced for intact seeds when compared to seeds without testa, indicating that the testa slows the diffusion of NO and consequently the ability of aleurone cells to sense NO (Bethke et al., 2007).

Additional roles of NO sensing include its involvement in pollen germination (Sirova et al., 2011), with NO showing involvement in regulating pollen tube growth and ROS required to sustain normal growth at the growing tip (Potocky et al., 2007). Unsuccessful fertilisation of flowers when the pollen fails to grow on the stigma leads to controlled pollen cell death, another process involving ROS and NO (Serrano et al., 2010). Flowering initiation is also repressed by NO, which can repress the expression of *CONSTANS* and *GIGANTEA* which are involved in bolting/flowering induction (He et al., 2004). Therefore NO plays an important role in controlling the photoperiod of plants. NO is likely sensed by the plant and perceived to be a stress signal, the reproductive stages of development are then slowed in an attempt to delay flowering, for example during a period of environmental stress.

1.4.4 S-nitrosylation and plant response to biotic and abiotic stress

Oxygen is a relatively stable molecule which alone is not highly reactive, however NO is volatile and is involved in numerous reactions within cells. S-nitrosylation (SNO), or the reaction of NO with a Cys-thiol (SH) group to form an S-nitrosothiol is a naturally occurring reaction in plants (Lamotte et al., 2015). The highly electrophilic nature of NO makes it attract to the nucleophilic S atom causing deprotonation (Gaston et al., 2003). S-nitrosylation of Cys-thiol groups is a ubiquitous modification within different organisms, which has an important role regulating protein and gene expression. Proteome-wide studies combining both the

biotin switch method and isotope-coded affinity tags (ICAT) have identified candidate proteins for S-nitrosylation in Arabidopsis (Fares et al., 2011).

Plant cell death, caused by a bacterial infection, can be regulated through the S-nitrosylation of protein Cys-thiol groups (Yun et al., 2011). Increased numbers of S-nitrosylated thiol groups trigger the hypersensitive response, localising infection by killing the cells surrounding the source (Yun et al., 2011). Excess nitrosylation of Cys-thiols is prevented and regulated by NO, which limits the extent of the hypersensitive response by preventing ROS production by causing the S-nitrosylation of NADPH oxidases (Yun et al., 2011). Under stress conditions glutathione (GSH) can react with nitrosylated thiol groups forming a nitrosoglutathione (GSNO) (Astier et al., 2012). GSNO plays a messenger role in the wounding response in plants. Injured leaves increased GSNO in a uniform manner whilst systemic leaves increased GSNO in the vascular tissue, indicating that GSNO is correlated with the wound signal transmission (Espunya et al., 2012).

Abiotic stresses can also induce S-nitrosylation of plant Cys-thiols. NO lies in the centre of a complex regulatory and feedback mechanism when a plant encounters abiotic stress (Corpas et al., 2011, Figure 1.7). The oxidation and reduction of RNS including SNO and GSNO also act to prevent nitrosative stress, cellular damage caused by excess RNS (Corpas et al., 2011). Salt stress is regarded as the most characterised abiotic stress in regards to S-nitrosylation (Romero-Puertas et al., 2013). An effect of osmotic stress created by drought or salinity is NO-driven

stomatal closure, subsequently increasing NO concentrations in the leaf mesophyll cells. Increased salt concentration was observed to increase S-nitrosothiol occurrence correlating with oxidative damage in *Citrus aurantium* leaves (Tanou et al., 2009). However S-nitrosylation was prevented when leaves were pre-treated with H_2O_2 or the NO donor SNP (sodium nitroprusside), suggesting that both of these reactive species lead to the detection of salt stress and priming against it.

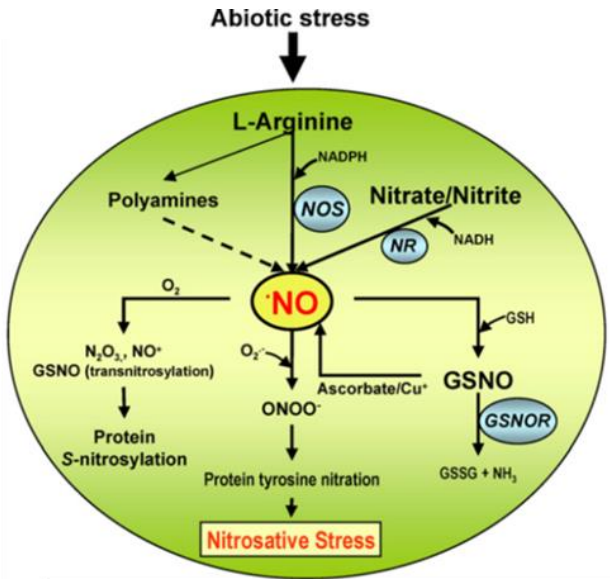


Figure 1.7 Schematic of the modifications to the metabolism of reactive nitrogen species (RNS) in response to abiotic stress. Abbreviations: NADPH, nicotinamide adenine dinucleotide phosphate; NOS, nitric oxide synthase; NR, nitrate reductase; NO, nitric oxide; O_2, oxygen; GSH, glutathione; GSNO, nitrosoglutathione; GSNOR, nitrosoglutathione reductase; GSSG, oxidised glutathione; $ONOO^-$, peroxynitrite ion; N_2O_3, dintirogen trioxide; NO^+, nitrosonium (Corpas et al., 2011).

Under hypoxic stress the Arabidopsis haemoglobin gene (AHb1) acts to detoxify NO by forming a S-nitrosohaemoglobin, thus reducing the damaging emission of NO caused by hypoxic stress (Perazzolli et al., 2004). In animals Hypoxia-inducible factor 1 (HIF-1) is an important transcription factor regulating over 60 genes and is involved in cancer biology (Li et al., 2007). HIF-1 is a heterodimer consisting of HIF-1β

which is constitutively expressed and HIF1-α which is regulated by hypoxia (Wang and Semenza, 1993). In normoxic conditions HIF1-α is degraded however under conditions of hypoxia or after S-nitrosylation of a single Cys-thiol in the protein sequence HIF-1α cannot be degraded and HIF-1 is stabilised and gene expression can continue (Li et al., 2007). This shows the ability of a mammalian protein to sense both oxygen and NO in the cell, and their presence or absence can control important biological functions such as tumour growth or death. Although the HIF-1 proteins are not substrates of the N-end rule pathway they like plant N-end rule substrates require oxygen and NO to initiate PTMs which initiate degradation. In HIF-1 this is the hydroxylation of Pro residues by oxygen sensitive Prolyl 4-hydroxylases (P4Hs) (Myllyharju, 2013), which is analogous to the oxygen dependent PCO oxidative PTM of Cys to Cys-sulfinic acid in plants (Weits et al., 2014; White et al., 2016).

1.5 Detection of post translational modifications in the N-end rule pathway using mass spectrometry

1.5.1 An introduction to mass spectrometry

Mass spectrometry (MS) is a method of determining a substance's molecular mass and structure by measuring its mass to charge (*m/z*) ratio. The first step of MS involves the production of ions, which can be achieved by various methods. The resultant ions are then separated by a mass analyser, allowing individual ions to be detected separately giving specific information regarding their molecular mass and potentially the structure of the original substance (Figure 1.8). The basics of MS involve the ionisation of a substance of interest and then its time or mass dependent detection. MS is now a widely used technique to identify substances and answer scientific questions from many research fields.

The forefathers of modern MS are Francis W. Aston and Arthur Dempster who both built on the work of J. J. Thomson, a British physicist who sparked ideas regarding subatomic particles in the late nineteenth century after discovering the electron (Thomson, 1897). Both Dempster and Aston made huge breakthroughs in 1918 and 1919 respectively, with Dempster developing the first MS which still resembles modern machines (Dempster, 1918); and Aston generating the first mass spectrograph allowing the mass differentiation of elemental isotopes (Aston, 1919). This ability to differentiate two different substances from a single input material would evolve into being able to separate substances from gases and liquids chromatographically (GC/LC). Using these instrumental

techniques Aston and Dempster would go onto discover numerous elemental isotopes.

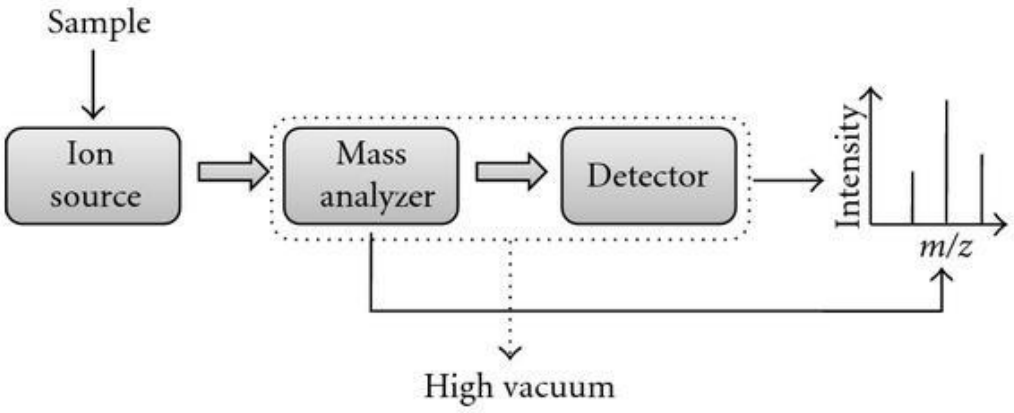


Figure 1.8 Schematic showing components of a mass spectrometer (Banerjee and Mazumdar, 2012)

Modern mass spectrometers employed in protein analysis generally use one of two ionisation methods. Matrix assisted laser desorption ionisation (MALDI) uses laser irradiation to desorb substances from a mixed solid or liquid matrix (Karas et al., 1985). The more widely used ionisation technique is electrospray ionisation (ESI). The principle of ESI involves spraying the analyte of interest from solution creating highly charged droplets which desolvate and enter into the inlet of the MS (Dole et al., 1968; Whitehouse et al., 1989). The capillary tip of the ESI source is kept at a high voltage (1-5 kV), the sample is therefore dispersed into a spray of similarly charged droplets under a strong electric field, and often in a dry nebulizing N_2 environment, which directs the spray towards the MS. The charged droplets are emitted from the ESI tip in a distinctive Taylor cone formation (Taylor, 1964), where the emitted droplets remain

stable in the electric field (Figure 1.9). One of the most popular spray emitters for small quantities of biological material is the nano-electrospray source, which allows a controlled low spray rate (e.g. 20 – 50 nl/min) allowing small quantities of the same sample to be used for different experiments. ESI is regarded as a soft ionisation technique which usually results in little or no fragmentation of the analyte ions.

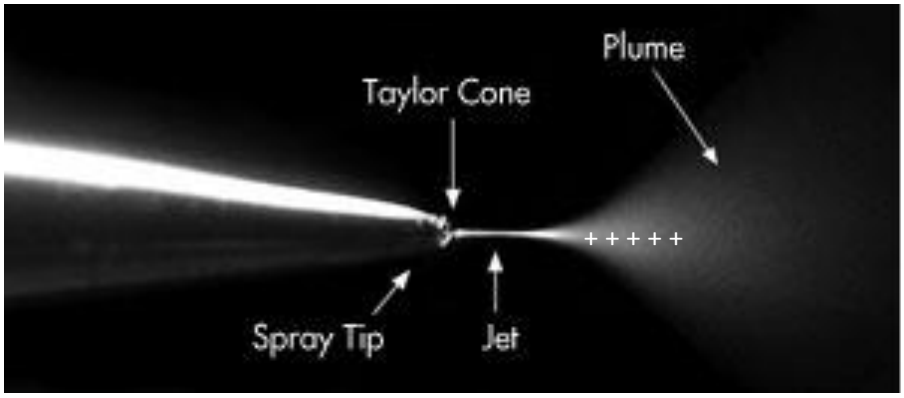


Figure 1.9 Electrospray ionisation emission. Electrically charged particles are emitted from a spray tip forming a Taylor cone (Adapted from New Objective, 2005).

Once ions have been generated the mass of each can be analysed to help reveal its original source identity. Mass analysers can be sub-divided into two main classes: 'beam analysers' where a beam of ions passes through the spectrometer to the detector and, 'trapping analysers' which hold ions whilst determining their mass (Bier et al., 1993; Stafford, 2002). All mass analysers either use electric or magnetic fields to manipulate ions allowing *m/z* measurements to be taken, and require a high vacuum (10^{-5} – 10^{-10} mbar) to be maintained throughout operation. The methods of manipulation in 'beam analysers' can be simple, such as measuring an ions time of flight (TOF) over a fixed distance, or more complex, through analysing the motion of ions in the *x* and *y* direction as they pass over

four charged parallel rods (quadrupole). Fourier transform ion cyclotron resonance (FTICR) is an example of a 'trapping' mass analyser, where ions are trapped in a magnetic field whilst absorbing radiofrequency (RF) radiation at a frequency resonant with the cyclotron (Marshall, 1985). The ICR unit is composed of trapping and detector plates, if the ions resonate with the RF their radius in the cyclotron increases bringing them in contact with the detectors.

In addition to detecting the mass of an ion, the constituent components of the ion can also be probed if it is isolated and entered into a second round of MS, known as tandem MS or MS/MS. Tandem MS can fragment the ions using a variety of methods, allowing the mass of the constituent components to also be detected (McLafferty, 1981). For example if a protein ion was entered into MS/MS it would be selected upon its *m/z* and then dissociated revealing its unique make-up of peptides and amino acids. The two most common fragmentation methods currently used during MS/MS are collision-induced dissociation (CID) also known as collisionally activated dissociation (CAD), and electron capture dissociation (ECD). CID is the most commonly used fragmentation technique building on the ground-breaking work in the late 1950's and in the 1960's where the structural impact of electrons on various natural products was revealed (Budzikiewicz et al., 1964; McLafferty, 1959). CID introduces inert gas molecules (Ar, He, N_2 etc.) into a confined environment held at a relatively high pressure, causing these molecules to collide with the ions of interest resulting in fragmentation.

Fragmented ions reveal the structural details of the parent structure, therefore enabling the identification of which precursor a fragment originated from, in case that the sample contained numerous substances. In protein analysis the vibrational energy created through CID encourages fragmentation at the amide bond on the peptide backbone, producing *b* and *y* ions giving sequence information from the N and C terminus respectively (Mitchell Wells and McLuckey, 2005). ECD produces both *c* and *z* ions from the N and C terminus respectively after a protein poly-cation accepts a high orbit electron, the charge is then neutralised causing radical bond dissociation which occurs at different rates depending on the bond (Zubarev et al., 1998). Both *b* and *y*, and *c* and *z* ion fragmentation is demonstrated in Figure 1.10.

In mixed samples ions generated and analysed by MS can belong to different proteins, making it difficult to identify which ions come from what protein. This is especially difficult if the ions from different samples happen to have the same mass, in these situations it is necessary to separate them chromatographically. Liquid chromatography (LC) is widely coupled to ESI-MS in contemporary protein studies, bolstering the ability to confidently identify individual proteins from crude extracts. Current MS instruments often use high performance liquid chromatography (HPLC), which involves the injection of a small volume of liquid into a packed column tube under high pressure (Engelhardt, 2012). Ions interact and elute off the column dependent on their physiochemical properties. In reverse phase chromatography column attachment by an ion involves hydrophobic adsorption to the non-polar column material, where one part of an ion binds to the hydrophobic

column surface while the non-bound part is still in contact with the mobile aqueous phase creating a strong interaction (Löchmuller and Wilder, 1979). This is overcome and the protein desorbs when the organic phase of the column reaches a specific concentration causing elution of the ion. Elution is followed by detection of the ions by MS with the ions creating a peak on a chromatogram at the time of elution. Monitoring these elution peaks on a chromatogram in conjunction with MS ion intensity enables ions with the same masses to be separated, and the relative abundance of each to be quantified.

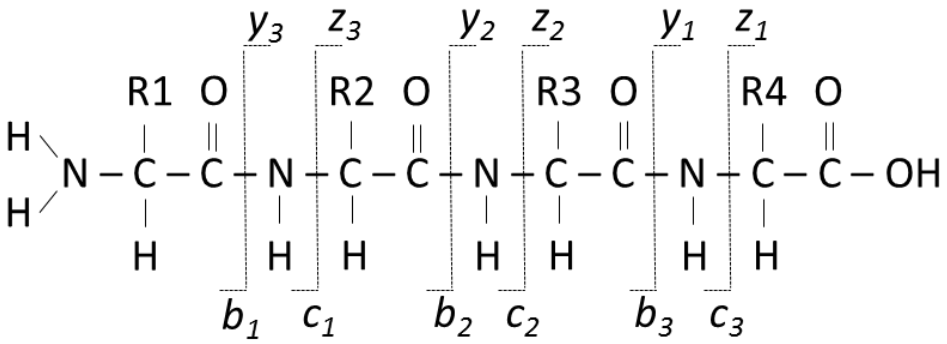


Figure 1.10 Collision induced dissociation and electron capture dissociation cleavage sites on a peptide. Collision induced dissociation (CID) cleaves at amides bonds whilst electron capture dissociation (ECD) cleaves at amine bonds along the peptide backbone creating *b* and *c* ions initiating from the N-terminus (*b1* and *c1*) and *y* and *z* ions from the C-terminus (*y1* and *z1*).

1.5.2 Detecting in vivo post translational modifications to substrates of the N-end rule pathway

Changes to protein structure after translation in the ribosomes are regarded as post translational modifications (PTMs). Such modifications are known to occur via the N-end rule pathway, for example the oxidation and arginylation of Nt-Cys. Edman sequencing was the first successful technique for identification of amino acid sequences, involving

repeat cyclical derivatization of Nt residues enabling residue identification (Edman, 1950). In relation to studies of the N-end rule pathway, Edman has been used to identify arginylation of Cys in mammalian cells but has failed to identify the oxidation state of Cys (Hu et al., 2005; Kwon et al., 2002). Using MS/MS to separate protein fragments by mass could potentially allow the identification of Cys oxidation in the N-end rule pathway. Observing differences in the *m/z* of peptides with a reactive Cys residue of +16 *m/z* (corresponding to the molecular mass of oxygen), in the spectrum is likely to correspond to an oxidation event. A comprehensive list of observed Cys modifications is shown in Table 1.1.

Recently in Arabidopsis targeted MS approaches have been used to monitor proteolytic events *in vivo*. Mutants of the Arg/N-end rule pathway either lacking ATE or PRT6 were used for tandem mass tag (TMT) analysis, when compared to WT this identified proteins which differed in abundance (Zhang et al., 2015). Root proteins were subsequently subject to terminal amine isotopic labelling of substrates (TAILS), revealing the positional identification of PTMs at the N-terminus as well as cleavage events which create new N-termini. Surprisingly in the *prt6* sample no Nt arginylation was detected perhaps indicating that arginylation is not a common modification in plants (Zhang et al., 2015). Similarly to TAILS combined or charged fractional diagonal chromatography (CO/ChaFRADIC) is a powerful MS approach where proteins of interest can be enriched and others diminished (Staes et al., 2008; Venne et al., 2013). The difference between ChaFRADIC and COFRADIC is the use of cation exchange or reverse phase chromatography respectively. Using ChaFRADIC over 2000 unique 'neo'

N-termini from 1270 Arabidopsis proteins were discovered, the under representation of destabilising Nt residues was also observed (Venne et al., 2015). These 'neo' N-termini generated by internal cleavage of proteins by caspases etc. may include some N-end rule substrates, although this would only occur if the protein was cleaved in a nuclear or cytosolic environment encoding the N-end rule machinery.

Targeting a specific residue where modification is known to occur such as the Nt oxidation of Cys from an *in vivo* protein is challenging, requiring a targeted proteomics approach. Isolation of enough target protein to enable MS identification following digestion is the first challenge. Affinity purification of a target protein from a crude mass of proteins involves incubating with a complementary antibody fixed to a resin membrane, purifying the target from the bulk protein. The target protein can then be identified by performing gel electrophoresis or Western blot. A further issue with studying the N terminus of a protein is the potential degradation of the protein *in vivo*, and the modification of this reactive site when extracted during processing for MS (Halliwell and Whiteman, 2004). Site specific quantitation of Cys oxidation can be identified through use of isotope coded affinity tags (ICAT) (Lindemann and Leichert, 2012; Sethuraman et al., 2004). The ICAT approach for Cys oxidation works on the premise that free thiols can be labelled by iodoacetamide, therefore treating a Cys residue with ROS/RNS prior to this labelling, will allow identification of oxidised Cys by monitoring thiols which are/are not labelled (Sethurman et al., 2004).

Table 1.1 Cysteine oxidative modifications and their anticipated mass change

PTM at Nt-Cys	Mass shift	Chemical adduct
Disulfide	-2	Cys-S-S-Cys
Cys-sulfenic acid	+16	Cys-SOH
Cys-sulfinic acid	+32	Cys-SO_2H
Cys-sulfonic acid	+48	Cys-SO_3H
Cys-thiosulfonic acid	+64	Cys-SO_2-SH
Selenylation	+80	Cys-S-Se
Cyclic Cys-sulfenamide	-2	Cys-S-N(R)
Cyclic Cys-sulfinamide	+14	Cys-SO-N(R)
Cyclic Cys-sulfonamide	+30	Cys-SO_2N(R)
Cys-nitrosothiol	+29	Cys-SNO

1.5.3 Monitoring PTMs of the N-end rule pathway in vitro using synthetic polypeptides

There are many uncontrollable variables which can interfere when trying to monitor PTMs occurring to proteins *in vivo*. The synthesis of polypeptides *in vitro* has enabled protein sequences of interest to be studied in a controlled environment. Initially synthesising long chains of amino acids was difficult, although technological improvements throughout the twentieth century has enabled high throughput synthetic peptide based research to become common (Lloyd-Williams et al., 1997). The utilisation of synthetic peptides has also been used in the N-end rule field. The covalent attachment of an 11 mer peptide bearing different N-terminal residues to a bead was used to try and bind mammalian N-recognins from mouse cell extracts (Tasaki et al., 2005). Using this method they were able to pulldown and identify two large 570 and 300

kDa proteins later termed UBR4 and UBR5 (Tasaki et al., 2005). Hu et al. (2005) utilised MALDI-MS to identify the arginylation of oxidised synthetic Cys Nt peptides when incubated with mammalian ATE. This was an important finding in mammals to show substrate-selectivity of ATEs.

Recently two papers have been published which use synthetic peptides to investigate Nt arginylaton by ATE (Wadas et al., 2016; White et al., 2016). White et al. (2016) show that plant PCO enzymes can effectively oxidise Nt-Cys synthetic peptides to Cys-sulfinic acid *in vitro*. Subsequently the incubation of Arabidopsis ATE1 with peptides bearing an Nt Asp or Cys-sulfinic/sulfonic acid was demonstrated to successfully cause arginylation, whilst Nt-Cys or Nt-Gly did not (White et al., 2016). Mouse ATE1 was shown similarly to arginylate the secondary destabilising residues of the Arg/N-end rule pathway Asp, Glu and Cys-sulfinic/sulfonic acid *in vitro* (Wadas et al., 2016). Interestingly the residues proceeding the N-terminus were shown to have some influence on the rate of degradation of the synthetic peptide, the presence of Lys at position 2 was shown to increase arginylation slightly when the N-terminus was unoxidised Cys (Wadas et al., 2016). Although it was not clear if the presence of a basic residue such as Lys next to Nt-Cys increased the chemical oxidation rate or helped ATE arginylate unoxidised Cys (Wadas et al., 2016).

1.6 Aims and objectives of this investigation

The aim of this study was to investigate the oxidation of Nt-Cys in protein substrates of the plant Arg/N-end rule pathway.

The objectives of this thesis were:

- To generate transgenic Nt-Cys environmental sensors of the Arg/N-end rule pathway in *Arabidopsis thaliana*

- To successfully extract and purify Nt-Cys reporter proteins from Arabidopsis, and identify the *in vivo* reporter protein and the Nt-Cys oxidation state using mass spectrometry

- To investigate *in vitro* oxidative reactions occurring at Nt-Cys synthetic peptides

2. Materials and Methods

2.1 Plants and plant growth conditions

2.1.1 Plant material

The plant species used for all experiments was *Arabidopsis thaliana* and the genetic lines were created by the Holdsworth group unless stated. Lines which ectopically express a MC-polyG-HA-GUS (MC-GUS) protein driven by the Cauliflower mosaic virus (CaMV 35S) promoter, *35S::MC-GUS* in *prt6-5* and *35S::MA-GUS* in Col-0 were supplied by Freddie Theodoulou (Rothamsted Research). The N-terminal extension of the MC-GUS protein construct including the Poly-Gly linker was based on the Ub-X-DHFR fusion proteins used in previous N-end rule pathway studies (Bachmair and Varshavsky, 1989; Chau et al., 1989). Figure 2.1 highlights the structure of both the gene and protein of the following constructs used in this study; (A) *promRAP2.12::MC-RAP2.12-GUS*, (B) *promRAP2.12::MA-RAP2.12-GUS*, (C) *Ubi1::MCGGAIL-GUS*, (D) *Ubi1::MAGGAIL-GUS*, (E) *35S::MC-GUS*, (F) *35S::MA-GUS*.

2.1.2 Plant growth media

Arabidopsis was grown on Murashige Skoog (MSk, (Murashige and Skoog, 1962)) media at half concentration. The ½ MSk media contained 1% Agar-Agar and 0.25% (Sigma Aldrich, Gillingham, UK) dissolved in distilled water adjusted to pH 6.2 by adding potassium hydroxide before

being autoclaved. Antibiotics or herbicides were added when the media had cooled but was still warm to the hand. Antibiotics and herbicides used are listed in Table 2.1. The ½ MSk media was then poured into square agar plates (10 x 10 mm) (Fisher Scientific, Loughborough, UK), that were then were used for seed plating directly or stored at 4 °C until use.

Table 2.1 – Antibiotics and herbicides used in this study

Antibiotic	Working concentration (mg/ml)
Ampicillin (Melford)	100
Gentamycin (Melford)	25
Hygromycin (Melford)	50
Kanamycin (Sigma Aldrich)	50
Spectinomycin (Sigma Aldrich)	100
Rifampicin (Melford)	50
Herbicide	
Glufosinate ammonium (Sigma Aldrich)	50

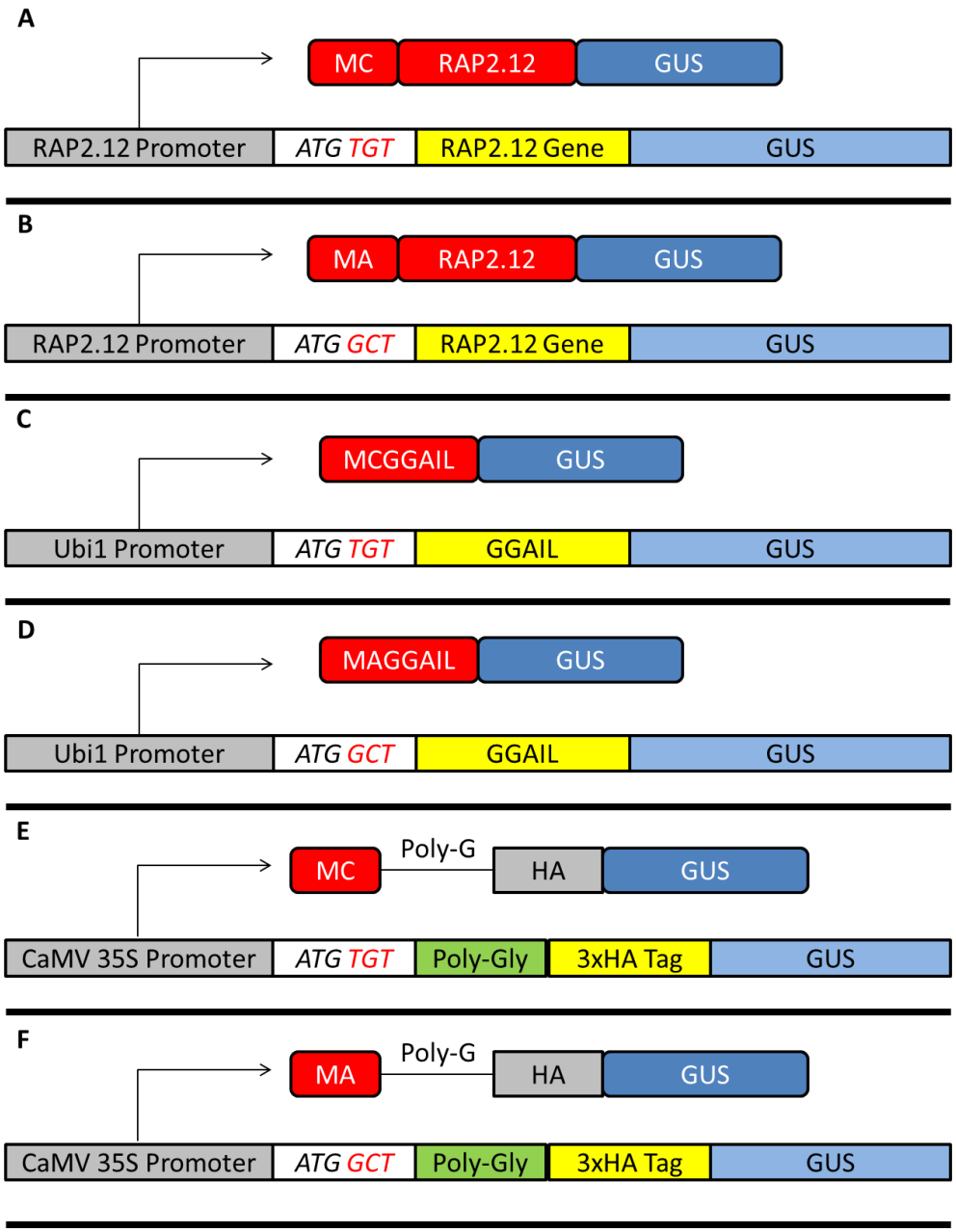


Figure 2.1 Constructs used in this study which were/have been transformed into Arabidopsis. (A) *promRAP2.12::MC-RAP2.12-GUS* – Native RAP2.12 promoter driving a protein consisting of the RAP2.12 gene/protein followed by a C-terminal β-glucuronidase (GUS) gene/protein. (B) *promRAP2.12::MA-RAP2.12-GUS* – Same as A but with a Met-Ala (MA) N-terminus encoded by ATG GCT. (C) *Ubi1::MCGGAIL-GUS* – Maize ubiquitin (Ubi1) promoter driving a protein consisting of a Met-Cys (MC) N-terminus encoded by ATG TGT this is followed by the residues GGAIL and then a C-terminal GUS gene/protein. (D) Ubi1::MAGGAIL-GUS – Same as C but with a Met-Ala (MA) N-terminus encoded by ATG GCT. (E) *35S::MC-GUS* - Cauliflower mosaic virus (CaMV) 35S promoter driving a protein consisting of an Met-Cys (MC) N-terminus encoded by the ATG TGT sequence connected to a 3xHA (HA) tag by a polyglycine (Poly-G) linker with a C-terminal GUS gene/protein. (F) *35S::MA-GUS* the same as E but with a Met-Ala (MA) N-terminus encoded by ATG GCT.

2.1.3 Seed plating and stratification

All seeds used for plating were transferred to 1.5 ml Eppendorf tubes and sterilized by soaking in 10% (v/v) bleach (Parazone, Jeyes, Cambridge) for 5 minutes before being washed twice with deionised water. Seeds were then plated using an Ultipipette tip (Barky Instruments International, Folkestone, UK) onto ½ MSk plates, the plates were then left to dry and sealed with micropore tape (Bracknell, UK). The plates (with seeds) were then left to stratify for 2-4 days in the dark at 4 °C.

2.1.4 Growth conditions

Stratified seeds plated on ½ MSk were placed in an indoor growth room operating with continuous (24 hour) lighting and at 22 °C, lighting was set at 150 μmol^{-1} m^2. Seedlings to be grown to maturity were transplanted from media to soil after 10 days. The components of the soil used are listed in Table 2.2. Trays of potted seedlings were grown in a controlled environment growth room, under the following conditions, 16 hour days, light intensity of 110 μmol^{-1} m^2, and a relative humidity of ~90 %.

Table 2.2 - Soil composition for plant culture

Component	% of total	Manufacturer
Levington M3 Compost	57	Scotts, Godalming, UK
Medium grade 2mm-5mm Vermiculite	29	Sinclair, Lincoln, UK
Medium grade 2mm-5mm Perlite	14	Sinclair, Lincoln, UK

2.1.5 Cross pollination of Arabidopsis flowers

To cross two different genetic lines pollen was transferred from the stamen from one parent flower (father) to the mature unfertilised stigma of another parent (mother) flower. Emerging flowers from the mother plant and mature stamen from the father plant were chosen for crossing. Under a microscope using sterile forceps the petals and stamen were carefully removed from a flower on the mother plant, leaving only the stigma. Mature open stamens from the father plant were then gently rubbed against the emasculated stigma on the mother plant until specs of yellow pollen were deposited. Seeds from siliques of successfully cross-pollinated inflorescences were collected once they matured and turned yellow.

2.2 DNA: Amplification, Cloning and Transformations

2.2.1 Genomic DNA extraction

Young plants transferred to soil were used for DNA extraction. A small leaf disk of roughly 2 mm^2 (~100 mg) was macerated in 40 µl extraction buffer (Table 2.3) using a pipette tip and then heated to 95 °C for 10 minutes to denature the DNA. The denatured DNA was left on ice for 5 minutes to cool before it was diluted by addition of 40 µl of 3% (w/v) Bovine serum albumin (Sigma Aldrich, Gillingham, UK) then the samples were centrifuged at 16,600 g for 30 seconds.

Table 2.3 – Genomic DNA Extraction Buffer

Component	Concentration
Potassium chloride	0.25 M
EDTA	10 mM
Tris-HCl (pH 9.5)	100 mM

2.2.2 Oligonucleotide primers

Oligonucleotide primers were designed to perform site directed mutagenesis (SDM) following the protocol of the Quikchange Site directed Mutagenesis Kit (Agilent Technologies, CA, USA). All primers were ordered from Eurofins (Eurofins Genomics, Ebersberg, Germany). The SDM primers used to produce the *promRAP2.12::MA-RAP2.12::GUS*

from *promRAP2.12::MC-RAP2.12::GUS* DNA are listed in Table 2.4. Primers were also used to confirm the sequence of transformed plasmid DNA through Sanger sequencing (Sanger et al., 1977) (Source Bioscience, Nottingham, UK) and to confirm the genotype and homozygosity of different Arabidopsis accessions, these primers are listed in Table 2.5 and 2.6 respectively.

Table 2.4 – Site directed mutagenesis oligonucleotide primers. Vectors containing DNA encoding RAP2.12 (At1G53910) were modified to change the resultant protein from Cys to Ala. The superscript number after an amino acid residue indicates the position where this modification occurred e.g. Cys^2 → Ala^2 indicates that the Cys residue at position 2 was changed to an Ala.

Accession	SDM (direction)	Sequence (5′ → 3′)
At1G53910	Cys^2 →Ala^2 (Fwd)	C-GCC-GCC-ATG-Gct-GGA-GGA-GC
	Cys^2 →Ala^2 (Rev)	GC-TCC-TCC-agC-CAT-GGC-GGC-G

Table 2.5 – Oligonucleotide primers used for sequencing. Primers designed roughly 100 bp upstream (U100) or downstream (D100) from the ATG start codon. Primers sent to Source Bioscience for Sanger sequencing along with plasmid DNA from the anticipated accession.

AGI code	Sequence (5′ → 3′)
At1G53910	GCGTTGAAGAGAGAAGAAGA
	GCTTCGAACTCAGCGTCA

Table 2.6 – Oligonucleotide primers used for genotyping Arabidopsis transgene and accessions.

Transgene	Sequence (5′ → 3′)
GUS	CTGGATCGCGAAAACTGTGG CCAAAGCCAGTAAAGTAGAACG
Mutant allele	
prt6-5	CAACATAAGAATCTGCGGGAG TGTGATGGAGATGGAGAAAGG

2.2.3 Polymerase chain reaction (PCR)

Standard components of a PCR reaction and the thermodynamic cycling conditions for amplifying DNA are listed in table 2.7 and table 2.8 respectively. Once the thermocycling had finished the amplified DNA was either be stored at 4 °C or used the same day for cloning, gel electrophoresis, digestions etc.

Table 2.7 – Polymerase Chain Reaction mix

Reagents	Volume for a single reaction (µl)
DNA template	2 (100 ng)
Water	35.9
Phusion buffer (NEB)	10
Forward Primer (10 mM)	0.5 (0.2 µM)
Reverse Primer (10 mM)	0.5 (0.2 µM)
DNTP's	1 (200 µM)
Taq Phusion (NEB)	0.1 (1.0 units)
Final Volume	50

Table 2.8 – PCR thermocycler conditions

PCR stage	Temperature (°C)	Time	
Denaturation	98	5 min	
Denaturation cycle	98	30 sec	x 40
Amplification cycle	50-60	30 sec	
Extension cycle	72	60 – 90 sec	
Extension	72	7 min	
End	4	-	

2.2.4 DNA gel electrophoresis

The reagents required to make an ethidium bromide agarose gel are listed in Table 2.9. Agarose and TBE buffer (45 mM Tris-Borate and 1 mM

EDTA) were mixed and heated in the microwave until all agarose dissolved. When the agarose solution had cooled until warm to the hand ethidium bromide was added. The gel was poured into a gel plate with an attached comb to create wells to load your sample, once the gel polymerised it was transferred to a gel tank containing TBE buffer. Before loading the gel the PCR reaction samples were mixed with 6x loading dye (60% (v/v) glycerol, 6% (v/v) 1M Tris HCl, 1.2% (v/v) 0.5 M EDTA, 0.06% (w/v) Bromophenol blue) and when loaded the gel was run at 120 V for 60 minutes (8 V/cm).

Table 2.9 – Components of a 0.8% (w/v) Agarose gel used for DNA electrophoresis

Component	Concentration
Agarose	0.8% (w/v)
TBE Buffer	99.1075% (v/v)
Ethidium Bromide (10 mg/ml)	0.0025% (v/v)

2.3 Cloning and Transformations

2.3.1 Restriction digestions, ligations and Gateway cloning

DNA amplified by PCR, after undergoing site directed mutagenesis, was digested to enable ligation of the amplified insert into a donor vector (pDONR221) via the Invitrogen Gateway cloning method (Thermo Fisher Scientific, MA, USA). Targeted enzymatic restriction using the enzyme BSRGI cut the DNA at specific points creating strands with an expected size (New England Biolabs, Hitchin, USA). The restriction digestion reaction mix (Table 2.10) was incubated at 37 °C for 2.5 hours before the digested DNA product was separated on an agarose gel. Control digestions were performed using plasmid DNA from transformed *Escherichia coli* (*E. coli*) cells. To recombine the target sequence after SDM and ligate into pDONR221 both the Gateway BP clonase II enzyme mix and the Gateway LR clonase II enzyme mix kits were used (Thermo Fisher Scientific, MA, USA).

Table 2.10 – Components of restriction digestion enzyme reaction mix

Component	Volume (µl)
Bovine Serine Albumin 10x (NEB)	2
Buffer 2 (NEB)	2
BsRGI (NEB)	0.5
DNA (1 µg)	10
H_2O	5.5

2.3.2 Lysogeny broth media

E. coli and *Agroacterium tumefaciens* were cultured using Lysogeny broth (LB). The solid LB media consisted of 1.5% (w/v) Agar-Agar, 1% (w/v) NaCl, 1% (w/v) Tryptone and 0.5% (w/v) yeast extract, this mixture was dissolved in distilled water when autoclaved. Agar-Agar was omitted from this mix when liquid LB was made. Antibiotics were added when the media had cooled and but was still warm to the hand (see Table 2.1 from section 2.1.2). The LB media was poured into round petri dishes if Agar-Agar had been added otherwise it was stored as a liquid culture.

2.3.3 Transformations – Escherichia coli

To transform the ligated SDM product 1 µl (1 ng) was used with 50 µl of high efficiency 5-alpha competent *E. coli* cells (New England Biolabs, Hitchin, UK). This mixture was left on ice for 30 minutes then heat shocked in water at 42 °C for 30 seconds, subsequently 950 µl of SOC media was added and the mix was incubated at 37 °C for 45 minutes whilst shaking at 200 rpm. The mix was then centrifuged at 12,000 g for 5 mins at room temperature, once the supernatant was discarded the pellet was resuspended in 100 µl of fresh SOC media. This resuspension was then plated on antibiotic plates corresponding to the respective resistance of the *E. coli* vector, the plates were then left overnight at 37 °C for resistant transformed *E.coli* to grow.

2.3.4 Plasmid DNA isolation from E. coli

Liquid LB cultures (3 ml, containing resistant antibiotic) of transformed E. coli were generated by scraping a single colony which had been cultured overnight on an antibiotic plate and placing into LB liquid. The LB was shaken at 200 rpm overnight at 37 °C before being centrifuged at 12,000 g for 5 minutes at room temperature, the supernatant was removed and plasmids were isolated using the QIAprep Spin Miniprep Kit (QIAgen, Crawley, UK). The concentration of isolated plasmid DNA was measured on a nanodrop ND1000 spectrophotometer (Nanodrop, Wilmington, USA), before being frozen -20 °C.

2.3.5 DNA sequencing

Isolated plasmid DNA was sent to Source Bioscience (Nottingham, UK) for sequencing, this either used in-house M13 forward and reverse primers or custom forward and reverse primers were designed so the complete sequence could be verified (See Table 2.3, section 2.1.3). Analysis of the sequencing was then completed using Geneious software (Biomatters, Auckland, NZ)

2.3.6 Transformations – Agrobacterium tumefaciens

1 µl (500 ng) of plasmid DNA with the correct/desired sequence was transformed by mixing with 200 µl (OD_{600} 1.5) of competent *Agrobacterium tumefaciens* cells (strain GV3101 pM-P90, (Koncz and

Schell, 1986)). The mix was left on ice for 10 minutes then transferred to a cold electroporation cuvette in 1 ml of LB media and electroporated using standard conditions (1.5 kV/cm for 5 msec) (Bio-Rad, Berkley, USA). The mixture was then transferred to a 1.5 ml Eppendorf tube and then shaken at 200 rpm, 28 °C for 4 hours. The Agrobacterium cells were then centrifuged, the supernatant was discarded and the pellet resuspended in 100 µl LB media and plated on LB plates containing Gentamycin, Rifampicin and Spectinomycin. The plates were incubated at 28 °C until colonies were visible (48 – 72 hours).

2.3.7 Transformations of Arabidopsis thaliana

The floral dipping method was used to transform Arabidopsis plants which were just developing flowers, or had been cut back and left to regrow for a week (Clough and Bent, 1998). One Agrobacterium colony resistant to the selective LB plate was picked and was added to a 5 ml pre-culture of LB containing Gentamycin, Rifampicin and Spectinomycin and incubated whilst shaking at 200 rpm overnight at 28 °C. This culture was used to inoculate a larger 250 ml of LB media with the same added antibiotics, this was then shaken at 200 rpm at 28 °C for 20 – 30 hours until the culture was turbid and pungent. The cells within the large culture were pelleted by centrifuging at 6300 g for 10 minutes at room temperature, the supernatant was removed and the pellet was resuspended in 5% (w/v) sucrose and 0.05% (v/v) Silwet L-77 (Lehle Seeds, Round Rock, USA). The Arabidopsis shoots were dipped into the Agrobacterium solution and immersed for 20 seconds. After dipping the

plants were placed in a propagator and kept out of direct light overnight before being placed back into an un-covered tray under lights until seeds had matured.

2.3.8 Selection of transgenic Arabidopsis plants

T1 seeds were harvested from plants transformed with Agrobacterium containing the *promRAP2.12::MA-RAP2.12-GUS* plasmid and selected on ½ MSk plates containing hygromycin. Sterilized seeds were spread on the Hygromycin selection plates evenly and left to stratify for 2 – 4 days in the dark. After chilling the plates were placed under light for 6 hours before covering with foil for 48 hours, the foil was removed and the plants were left in continuous light until resistant seedlings began to 'green' and could be differentiated from non-resistant seedlings which remained stunted and appeared yellow and failed to turn green. Established resistant seedlings were then transferred under sterile conditions to a ½ MSk plate with no antibiotics and left to recover before being transferred to soil.

2.4 Plant treatments and GUS staining

2.4.1 Hypoxia treatment of Arabidopsis

Seedlings aged 4 – 14 days grown on ½ MSk were placed in a large glass desiccator with an inlet/outlet valve connected to a gas mixing system (Photon Systems Instruments, Drasov, Czech Republic), which maintained a controlled flow of 4 L/min N_2 into the chamber. The desiccator lid was sealed with grease and the O_2 concentration of air in the outlet valve was monitored using a Kane 250 flue gas analyser (Kane, Herts, UK). The O_2 concentration was monitored until it dropped below 1% before the seedlings were left for 3 hours in the dark. As a control the same seedlings of the same genotypes were placed in another desiccator in ambient O_2 conditions in the dark for the same duration as the hypoxic treatment. After the treatment time had elapsed the seedlings were immersed in X-gluc (5-bromo-4-chloro-3-indolyl-beta-D-glucuronic acid) or flash frozen in liquid N_2 for protein extraction.

2.4.2 Analysis of protein stabilisation using the 26S proteasome inhibitor bortezomib

Bortezomib (Santa Cruz, TX, USA) was dissolved in DMSO to a stock solution of 10 mM and stored at -20 °C, when used this was diluted to a working concentration of bortezomib (50 μM). Seedlings aged 7 – 14 days were placed in a 12-well plate containing 1.5 ml of 50 μM. The 12-well plate was placed under lights and shaken at 70 rpm for 5 hours.

Seedlings were similarly immersed in 1.5 ml of 0.5% (v/v) DMSO to act as a control reaction. At the end of the treatment, seedlings were removed from the solution, dried and then immersed in X-gluc or flash frozen in liquid nitrogen.

2.4.3 Analysis of expression and stability of β-glucuronidase protein in transgenic plants

X-gluc solution (Table 2.11) was made by mixing X-gluc powder, sodium phosphate and Triton X-100 and mixing until all X-gluc had dissolved. This solution was aliquoted and stored at -20 °C until used, before use potassium ferricyanide and potassium ferrocyanide were added. Transgenic plant material expressing β-glucuronidase (GUS) protein was immersed in X-gluc solution and incubated at 37 °C until blue staining was visible in tissue/organs of interest, or staining differed between treated and untreated plant material. The X-gluc was then removed and 75% (v/v) ethanol was added and left shaking until the chlorophyll had been bleached and the tissue was destained. The samples were then washed twice with deionised water and were mounted on slides using Hoyers media (Table 2.12) for microscopic analysis.

Table 2.11 – Components of the X-gluc working solution used to stain plant tissue.

Component	Concentration
Sodium phosphate	100 mM
Triton X-100	0.1% (v/v)
X-Gluc	2 mM
Potassium ferricyanide	1 mM
Potassium ferrocyanide	1 mM

Table 2.12 – Components of the Hoyers mounting medium used to mount slides.

Component	Grams to add to 50 ml water
Gum Arabic	30
Chloral Hydrate	200
Glycerol	20

2.5 Protein extraction, purification and identification

2.5.1 Protein extraction from plant material

Protein was extracted from 7 – 10d seedlings flash frozen with liquid nitrogen using a mortar and pestle and extraction buffer (~150 µl, Table 2.13). Crude protein lysate was centrifuged in microcentrifuge tubes at 20,800 g for 20 minutes at 4 °C and the clear supernatant was removed. 2 x Laemmli buffer (Bio-Rad, Berkley, USA) was added to the protein and samples were boiled at 95 °C for 5 minutes.

Table 2.13 – Components of the protein extraction buffer.

Component	Concentration
Tris-HCl (pH 7.5)	50 mM
NaCl	150 mM
$MgCl_2$	5 mM
Glycerol	10% (v/v)
EDTA	1 mM
NP-40 (Sigma Aldrich)	0.5% (v/v)
Phenylmethanesulfonyl fluoride (PMSF) (Sigma Aldrich)	1 mM
ULTRA™ protease inhibitor (Roche)	1 tablet per 10 ml of buffer

2.5.2 Affinity Purification of HA-tagged protein from crude plant protein extract

Protein from 14 plates of 10 – 14d seedlings was extracted as previously described (see section 2.5.1). 50 µl of EZview anti-HA Affinity Gel beads (Sigma Aldrich, Gillingham, UK) were deposited into a 1.5 ml Eppendorf tube and washed twice in extraction buffer by adding 750 µl and inverting the tube for 20 seconds before centrifuging at 16,600 g for 20 seconds to pellet the gel beads. 1 ml of protein supernatant was added to the beads and the eppendorfs were incubated at 4 °C for 3 hours on an orbital shaker, encouraging binding between the anti-HA antibody beads and HA-tagged target protein. After 3 hours incubating the eppendorfs containing the beads and supernatant were centrifuged at 16,600 g for 20 seconds to pellet the gel beads. The supernatant was then removed and 1 ml of wash buffer (Table 2.14) was added before the eppendorfs were inverted for 1 minute centrifuged at 16,600 g for 20 seconds. This was repeated 4 times to remove all traces of the crude protein, once all wash supernatant had been removed the beads were stored at 4°C until 2 x Laemmli buffer (Bio-Rad, Berkley, USA) was added and the beads were boiled at 95 °C for 5 minutes to release the target HA-tagged protein.

Table 2.14 – Components of the Affinity Purification wash buffer.

Component	Concentration
Tris-HCl (pH 7.5)	50 mM
NaCl	300 mM
$MgCl_2$	5 mM
Glycerol	10% (v/v)
EDTA	1 mM
NP-40 (Sigma Aldrich)	0.5% (v/v)
Phenylmethanesulfonyl fluoride ((PMSF) Sigma Aldrich)	1 mM
ULTRA™ protease inhibitor (Roche)	1 tablet per 10 ml of buffer

2.5.3 Polyacrylamide gel electrophoresis (PAGE)

Polyacrylamide gels were cast using the reagents listed in Table 2.15. The percentage of acrylamide added varied, depending on the size of the target protein to be analysed. 40 µg of protein was loaded in comparative experiments either investigating protein expression or stability from untreated/treated plant samples, these gels would later be transferred for Western blots (see section 2.5.6). 70 µl of affinity purified protein was loaded onto gels which would later be coomassie stained. After loading the protein samples 4 µl Seeblue Prestained2 Plus Ladder (NEB, UK) was loaded in an empty well and the gels were run in running buffer containing 10% (v/v) TGS Buffer (Tris-HCl, Glycine, SDS; Fisher

Scientific, Loughborough, UK) starting at 80 V for 20 mins when protein was migrating from the stacking to running gel, then at 120 V for 100 mins or when the gel front had just migrated off the running gel.

Table 2.15 – Components required to cast an 8% (v/v) 1.5 mm Bis/Acrylamide gel.

Component	Volume	
	Running Gel	Stacking Gel
Water	8.4 ml	4.5 ml
Bis/Acrylamide	4.67 ml	1.11 ml
Tris	(pH 8.8) 4.4 ml	(pH 6.8) 0.84 ml
SDS	177.3 ul	66.7 ul
Ammonium persulfate	177.3 ul	66.7 ul
Tetramethylethylenediamine	7.07 ul	5.3 ul

2.5.4 Identifying affinity purified protein using colloidal Coomassie stain

After gel electrophoresis a polyacrylamide gel was fixed with fixing solution (79% (v/v) high performance liquid chromatography (HPLC) grade H_2O, 1% (v/v) phosphoric acid and 20% (v/v) methanol) for 1 hour. After the fix solution was removed the Roti-Blue colloidal coomassie stain (Karl Groth, Karlsruhe, Germany) was added (20% (v/v) Roti-Blue, 60% (v/v) HPLC H_2O and 20% (v/v) methanol. The gel was shaken at 50 rpm in coomassie stain until protein bands appeared. The

coomassie stain was removed and a destain solution (25% (v/v) methanol, 75% (v/v) High Performance Liquid Chormatography (HPLC) H_2O was added and the gel was washed until the background stain had been removed.

2.5.5 PAGE gel band excision, processing, digestion and extraction

Visible coomassie gel bands were excised using a sterile scalpel, the individual bands were diced into 1-2 mm sized pieces and transferred to a 96-well Corning Costar cell culture plate (Sigma Aldrich, Gillingham, UK) containing 50 µl HPLC water. Processing was performed robotically using the MASSprep™ station liquid handling robot (Waters, Massachusetts, USA). The robot was used to destain, reduce, alkylate and dehydrate the gel bands, a schematic of the gel processing, digestion and extraction method is shown in figure 2.2. Once processed the gel was digested with a selected endoproteinase in order to cleave protein into its constituent peptides with predicted N-terminal reisidues (Table 2.16). The concentration of protease used was dependent on the protein concentration estimated from the stain intensity of the gel band (varied between 1:20 – 1:100). The duration of digestion varied between 30 minutes and 16 hours, this depended if complete digestion (longer duration) or partial digestion (shorter duration) was desired and the digest efficiency of each protease. At the end of the desired digestion period 10 µl HPLC water was added to the digestion reaction mix and left for 10 mins, then all liquid from each well was moved to a clean 1.5 ml

Eppendorf microcentrifuge tube. The gel pieces were then washed twice for 10 minutes with 30 µl 50% (v/v) Aceotnitrile (MeCN)/5% (v/v) Formic acid and all liquid was moved to the respective 1.5 ml Eppendorf containing the digestion mix from that well. Finally 10 µl of 100% MeCN (v/v) was added to each well and left for 5 minutes to dehydrate the gel. The peptide mix collected was concentrated in a Savant DNA Speedvac (Thermo Scientific, Cramlington, UK) until only a few µl of the digest mix remained before being reconstituted (5 – 50 µl) in aqueous LC-MS solvent (95% (v/v) HPLC H_2O/5% (v/v) MeCN/ 0.1% (v/v) Formic acid).

Table 2.16 – Endoproteinases used and their site of action

Protease	Site of Action	Supplier
Trypsin	C-terminal of R and K (Huber and Bode, 1978)	Promega
Lys-C	C-terminal of K (Jekel et al., 1983)	Promega
Asp-N	N-terminal of D (Drapeau, 1980)	Promega

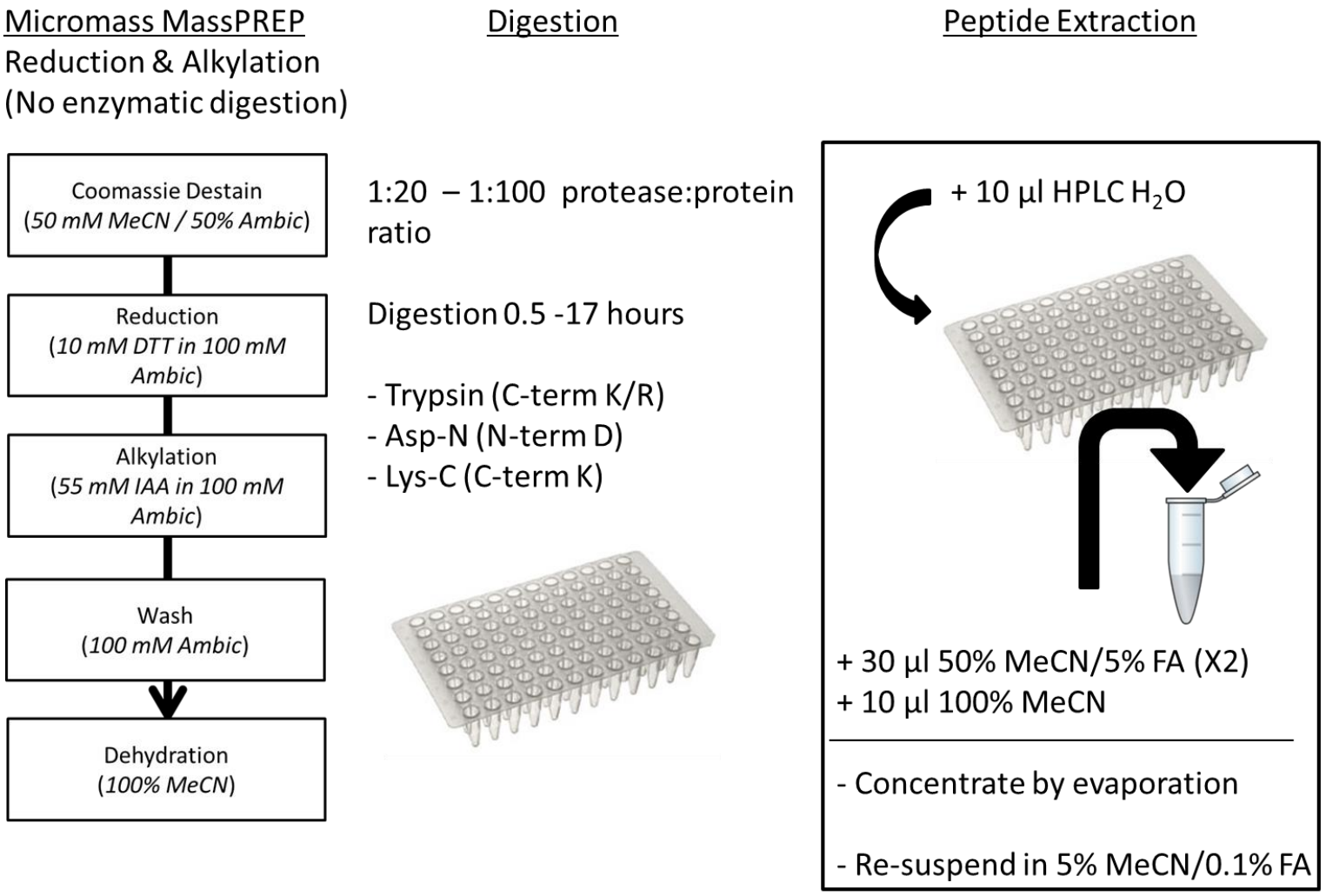


Figure 2.2 Schematic of the reduction, alkylation and digestion process for excised coomassie stained protein bands

2.5.6 Western blots

After PAGE (see section 2.5.3) gels used for Western blot analysis (Towbin et al., 1979) were sandwiched in a cassette with Immune Blot nylon membrane (Bio-Rad, Berkley, USA). The gel was transferred at 80 v for 2 hours in transfer buffer (Table 2.17). After transfer the membrane was blocked for 1 hour with Phospate Buffered Saline (PBS) (Fisher Bioreagents, Loughborough, UK) with 5% (w/v) milk powder and 0.1% (v/v) Tween (Fisher Bioreagents, Loughborough, UK). The primary anti-Hemagglutinin (HA) antibody produced in mouse (Sigma Aldrich, Gillingham, UK) was then incubated at a dilution of 1:2000 in the milk/PBS/tween solution with the membrane overnight shaking at at 4 °C. When anti-α-tubulin (Tub) antibody (Sigma Aldrich, Gillingham, UK) was used the dilution made was 1:10000. The following morning the primary antibody was removed and the membrane was washed in PBS. Additional antibodies used were anti-ubiquitin (Ub) (Santa Cruz, Germany) used at dilution 1:1000, anti-sumoylation (Sumo) (Abcam, Cambridge, UK) used at a dilution 1:5000 and anti-GUS (Thermo Fisher Scientific, USA) used at dilution 1:1000. The secondary antibody used for anti-HA, anti-Tub, anti-Ub and anti-Sumo was Goat-anti-mouse, (Sigma Aldrich, Gillingham, UK) diluted 1:10000 in milk/PBS/tween solution and incubated with the membrane shaking at 50 rpm for 1 hour at room temperature. The secondary antibody used for anti-GUS was also diluted (1:10000). The membrane was then washed with PBS-Tween three times before the target proteins were visualised, after development in a dark room following the addition of Pierce EVL Western Blotting Substrate (Thermo Scientific, Cramlington, UK) to the membrane, and exposure to

High Performance Chemiluminescence Film (GE Healthcare, Buckinghamshire, UK).

Table 2.17 – Western blot Transfer Buffer

Component	Volume (ml)
H_2O	700
Tris-Glycine (TG) Buffer (10x) (Flowgen)	200
Methanol	100

2.5.7 Monitoring equal protein loading after Western blot using Ponceau Stain

After developing on film the Immune-Blot membrane was stained with Ponceau S solution (Sigma Aldrich, Gillingham, UK) to analyse loading on the gel and analyse the efficiency of the transfer. After shaking in ponceau for 5 minutes bands appear on the membrane, the membrane was then washed with deionised water and left to dry at room temperature.

2.6 Electrospray Ionisation Mass Spectrometry (ESI-MS) of *in vitro* synthetic peptides

2.6.1 Synthetic peptides

Synthetic polypeptides were ordered from commercial suppliers to conduct *in vitro* oxidation and nitrosylation experiments. Peptides used in this study are listed in Table 18

2.6.2 Analysis of in vitro synthetic peptides

Each solid peptide was divided into 1 mg fractions and transferred to an Eppendorf tube. HPLC H_2O was added and stock peptide solutions were aliquoted and frozen for future use. The peptide was diluted to 1 μM and this concentration was infused into the inlet of an ESI source using a syringe pump.

2.6.3 Peptide oxidation – H_2O_2

To oxidise the Nt-Cys of a peptide it was diluted in to 0.027 % (v/v) solution of aqueous H_2O_2 (Fisher Scientific, Leicestershire, UK). The final concentrations were 1 μM peptide and 2.7 μM H_2O_2, the products were then observed over an initial 2 hour time course using ESI-MS/MS.

2.6.4 Peptide nitrosylation – SNAP and SNP

The NO donor SNAP (S-Nitroso-N-acetyl-DL-penicillamine, Sigma Aldrich, Gillingham, UK) was dissolved in methanol following the manufacturer's instructions and then diluted to a working concentration of 10-50 μM in

HPLC H_2O and incubated with 1 μM of peptide. The reaction was kept in the dark as SNAP is light sensitive. A second NO donor SNP (Sodium nitroprusside dehydrate, Sigma Aldrich, Gillingham, UK) was also used this was dissolved in water and used at the same concentrations as SNAP and also similarly kept in the dark. This reaction was then observed over a 1 hour time course using ESI-MS/MS.

2.6.5 Two stage treatments – H_2O_2 and SNAP

The CR-40 peptide was incubated initially with either H_2O_2 or SNAP as described above, and then when the Cys had become oxidised or nitrosylated this was mixed with either SNAP or H_2O_2 to initiate any subsequent oxidative modifications on the Nt-Cys, which were analysed by ESI-MS/MS.

Table 2.18 – Synthetic peptides used in this study. Showing the name, sequence, purity and supplier.

Name	Sequence	Purity (%)	Supplier
CR-40	CRSGGGGGGGGGGGGRGGSGAWLLPVSLVKRKTTLAPNTQ	≥ 98	*WuHan Moon Biosciences Limited*
RCR-49	RCRSGGGGGGGGGGGGRGGSGAWLLPVSLVKRKTTLAPNTQTASPRALA	92	*Pepscan Presto BV*

2.6.6 LTQ FT Ultra Mass Spectrometer and ESI-MS/MS run conditions

To analyse the *in vitro* experiments the synthetic peptides were identified using an LTQ FT Ultra Mass Spectrometer (Thermo Fisher Scientific, MA, USA). The inlet source to inject the peptides into LTQ FT Ultra was via electrospray ionisation (ESI).

The integrated syringe pump on the LTQ FT Ultra was used to directly inject a 95:5 HPLC H_2O:MeCN + 0.1% FA solvent from a 250 µl syringe (Hamilton, Bonaduz, Switzerland). The LTQ FT Ultra was fitted with a commercial electrospray source and the solvent was injected at a constant rate of 5 µl/min. The synthetic peptides were infused manually through injection of 1-2 µl of peptide diluted in HPLC H_2O or solvent to 1 µM. Data acquisition for the synthetic peptides was optimised on the mass of the CR-40 peptide when injected, to select an appropriate source current, voltage and tube lens voltage. The source voltage used for the CR-40 peptide was 4.95 kV with a 4.55 µA current, whilst the capillary voltage and temperature were 35 V and 275 °C respectively and the tube lens was 110 V. Data was acquired initially in a full scan window (400 – 2000 m/z) and then subsequently in MS/MS mode using collision induced dissociation (CID) or electron capture dissociation (ECD).

Predicted fragment peptide ions were generated using the Xcalibur protein calculator tool. The sequence of the peptide injected was inputted with any modifications and then the MS/MS residue ions were generated and these masses were manually searched for in the MS/MS data.

2.7 *In vitro* analysis of CR dipeptide by LC-MS/MS

2.7.1 CR dipeptide

The dipeptide CR (Cys-Arg)ws synthesised by Vivek Badiani using the Fmoc removal method following standard procedures (Fields, 1994). This involved the deprotection of Arg from Fmoc-Arg(Pbf)-Wang Resin (Sigma Aldrich, Gillingham, UK), before removing the Fmoc group and then coupling to Cys. The CR dipeptide was treated similarly to the synthetic peptides with H_2O_2 (See 2.6.3) and injected at 1 µM concentrations.

2.7.2 Analysis of CR dipeptide by LC-MS/MS

The CR dipeptide was analysed on an Agilent 1100 Series Liquid Chromatograph (Agilent Technologies, CA, USA) coupled to a Waters Synapt HDMS (Waters, Herts, UK) with an electrospray ionisation source. The LC was fitted with an Atlantis Silica HILIC column (100 Å, 5 µm, 2.1 mm x 50 mm, Waters, Herts, UK). The liquid mobile phases consisted of HPLC H_2O, Acetonitrile and Formic acid, A = 95:5 HPLC H_2O:MeCN + 0.1% FA and B = 5:95 HPLC H_2O:MeCN + 0.1% FA. The flow rate was set at 0.5 ml/min and 5 µl of 1 µM CR peptide was injected. A 20 minute gradient run was used with an initial isocratic 10 minute period with 100% B, then from 10 – 11 minute the mobile phases were set to 50:50 A:B, then from 11 minutes solvent A was set to 100% and this remained until the run finished at 20 minutes.

The capillary voltage was set to 3.2 kV, the sampling cone 20 V and the extraction voltage was 6 V. Positive polarity was used. The source temperature was fixed at 150 °C with a desolvation gas flow of 6.67 ml/min and a temperature of 500 °C. Backing pressure was held at 1.75 mBar. The collision energy in the trap was 5 eV similar to the transfer collision energy. The mass acquisition range was fixed at 100 – 600 m/z.

The instrument was controlled by using MassLynx 4.1 software (Waters, Herts, UK), which translated the raw data into spectra which could be smoothed and centred as required.

2.8 Nanospray Ionisation Mass Spectrometry (NSI-MS) of digested *in vivo* peptides from CaMV *35S::MC-GUS* in *prt6*

2.8.1 Liquid Chromatography and Mass Spectrometry

In vivo protein gel bands digested with endoproteinases (see table 2.16 in section 2.5.5) were analysed on a Dionex Ultimate 3000 Nano LC system coupled to the LTQ FT Ultra fitted with a nanoelectrospray ionisation source. The Dionex was fitted with a commercial C4 trapping column and either a custom packed WP C4 column (5 µM, 100 Å, Phenomenex) with a packed PicoFrit column (360 µm OD x 75 µm i.d. x 150 mm, New Objective, MA, USA) or an Acclaim Pepmap300 C18 Protein and Peptide column (5 µm, 300 Å) with fused silica column outlet (280 µm o.d. x 75 µm i.d. x 150 mm). 5 µl of digested sample was injected and conditions were maintained at 4°C throughout the run. The liquid mobile phases consisted of HPLC H_2O, Acetonitrile and Formic acid, A = 95:5 HPLC H_2O:MeCN + 0.1% FA and B = 5:95 HPLC H_2O:MeCN + 0.1% FA. The respective mobile phases were mixed and the peptides were eluted using a 30 – 50 minute linear gradient of B from 0 – 70% at a flow rate of 0.3 µl/min followed by column equilibration (100% A) for 10 minutes until the next sample was injected. The nanoelectrospray source emitted charged ions when a 1.7 kV voltage was applied to a coated PicoTip (New Objective, MA, USA). The capillary temperature was 275 °C with the inner capillary voltage set at 45 V and a tube lens value of 145 V. Spectra were acquired in positive ion mode with a 1.8 kV

source voltage and 100 µA current in a 400-2000 m/z mass window at a nominal resolution of 100,000 (at m/z 400)

2.8.2 MS/MS conditions and peptide identification

Peptides of the GUS reporter protein were analysed using a data dependent acquisition (DDA) method. The 5 most intense ions (with a charge of 2^+ or higher) from each full MS scan were subject to CID using a nominal energy of 25 - 35 eV or ECD using a nominal energy of 3.00 eV. Once the acquisition was complete the MS/MS data was searched on the MASCOT search engine (version 2.3.01, Matrix Science Ltd, London, UK) using the MS/MS ions search tool against the public databases NCBInr (2015_10_16 database, 73055898 sequences, 26613539447 residues) and Swissprot (2015_11 database, 549832 sequences, 196078138 residues). Searches were also performed against a custom database comprising of variations of sequences for the MC-GUS construct (Appendix A). A series of variable modifications were included in the searches to account for post translational modifications, this included the following; carbamidomethylation (cysteine), oxidation (cysteine and methionine), dioxidation (cysteine), trioxidation (cysteine), acetylation (N-terminal residue), formylation (N-terminal residue). Search parameters were also set to allow up to one missed cleavage during digestion.

3. Developing transgenic oxygen sensors of the N-end rule pathway in *Arabidopsis thaliana*

3.1 Introduction

The APETALA 2 (AP2 or ethylene responsive element binding protein (EREBP) is one of the three largest families of transcription factors in Arabidopsis and is specific to plants (Riechmann et al., 2000). The AP2 domain has been postulated to have evolved in bacteria with genes discovered in cyanobacteria, ciliates and viruses, indicating that it may have been acquired through lateral gene transfer (Magnani et al., 2004). The ERFVIIs are a subgroup of transcription factors in the AP2 domain and are present in a range of plant species including rice (*Oryza sativa*) (Nakano et al., 2006), barley (*Hordeum vulgare*) (Mendiondo et al., 2016) and other species of the *Rumex* and *Rorippa* families (van Veen et al., 2014). The physiological importance of the ERFVIIs in helping to sustain energy generation during periods of drought or flood has been demonstrated in Arabidopsis, rice and barley (Gibbs et al., 2011; Mendiondo et al., 2016; Voesenek and Bailey-Serres, 2013). The cellular response of RAP2.12 during the onset of hypoxia and re-oxygenation has been reported, with its translocation from the plasma membrane to the nucleus during hypoxia and then its degradation when oxygen levels increase (Kosmacz et al., 2015; Licausi et al., 2011; Weits et al., 2014).

Although the collective function of the ERFVIIs has been identified, less is known regarding the expression and stability of the individual ERFVIIs *in vivo*. To obtain the accurate expression of inserted transgenes *in vivo* it

is important that they are under the control of their own promoters. Expression under foreign promoters such as the CaMV 35S promoter are useful because of their constitutive strong expression (Benfey and Chua, 1990), however the 'over-expression' of genes of interest may not accurately reflect what the real expression of the gene is. The inclusion of the gene of interest and its native promoter in a destination vector translationally linked to a reporter protein, should highlight where this gene is naturally expressed. The expression of the ERFVIIs by their individual promoters in Arabidopsis could help identify where they are expressed which could indicate their individual functions. Creating stabilised versions of the ERFVIIs, by substituting the tertiary destabilising Nt-Cys at positon 2 with a stabilising Ala, would allow the examination of where accumulation of each ERFVII protein can occur (MA) and where it does occur (MC) (section 3.2).

Current research into the N-end rule pathway largely focusses on the search for substrates and the E3 ligase which detect them, or the physiological and biochemical function of known components. However less is known about the impact of the residues proceeding the Nt-Cys, and what effect they may have in determining the stability of a substrate. It has been reported that successful recognition of an N-end rule substrate requires that a free Lys residue must be in the vicinity of the N-terminus, this Lys is the site of Ub attachment essential for degradation (Bachmair and Varshavsky, 1989; Chau et al., 1989). The conservation of Nt residues of ERFVIIs in plants indicates that it is not just the Nt residue (after Met excision) which has a role in protein degradation. Section 3.3 investigates the expression of the Maize

ubiquitin promoter (*Ubi1*) M(C/A)GGAIL-GUS construct, containing the seven Nt residues of the ERFVIIs, when transformed into Arabidopsis and compared to its reported expression and stabilisation in barley.

Strong constitutive expression of a gene of interest *in vivo* is useful when trying to visualise the protein on a coomassie gel. The CaMV 35S promoter is the most extensively used promoter in plant biology, largely a result of its strong promotion of gene expression (Pauli et al., 2004). The strength of expression of a 35S gene transformed into a plant also depends on where in the chromosome the transgene is integrated. If adjacent genes are highly active increased expression may be visible and similarly if they are low reduced expression may occur (Zheng et al., 2007). For some applications the high expression of CaMV 35S driven genes has its draw-backs, this is especially true when investigating sensitive biological systems where strong expression of a gene may perturb the functionality. In section 3.4 an intermediate expressing CaMV 35S driven MC-GUS line is identified and its ability as an oxygen sensor investigated.

3.2 Developing *transgenic plants* providing expression of ERFVIIs under their own promoter

3.2.1 Generation of promRAP2.12::MA-RAP2.12-GUS

The *promRAP2.12::MC-RAP2.12* DNA was cloned into pDONR221 by BP reaction (see section 2.3.1) by previous Holdsworth lab member Dr George Bassel (now University of Birmingham, UK). A diagram of the *promRAP2.12::MA/MC-RAP2.12* gene regions is presented in Figure 3.1. To perform site directed mutagenesis (SDM) cDNA was extracted from the *promRAP2.12::MC-RAP2.12* BP vector to change Cys at position 2 to Ala using SDM primers (see table 2.4). The resultant PCR product was sequenced to confirm as MA (Figure 3.2) before being re-cloned using BP and then again using the LR reaction to recombine this new fragment in the destination vector PGWB533 which contained a GUS tag (C-terminal of gene). The construct was then transformed into Arabidopsis and the resultant T1 generation were selected on Hygromycin.

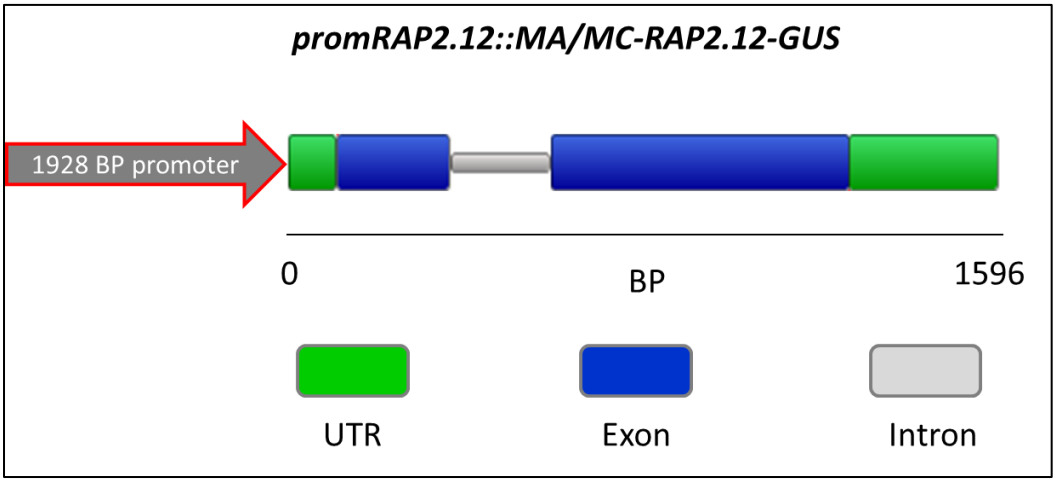


Figure 3.1 The structure of *promRAP2.12::M(A/C)-RAP2.12-GUS* before ligation into a destination GUS vector. Genes were cloned with a 1928 base pair (BP) native promoter. Gene diagram obtained from Plaza 3.0 (Proost et al., 2014).



Figure 3.2 Sequence comparison of *promRAP2.12::MC-RAP2.12-GUS* and *promRAP2.12::MA-RAP2.12-GUS* after performing site directed mutagenesis. Mutation highlighted in red showing change from TGT (encoding Cys) to GCT (encoding Ala). Comparative analysis performed using Clustal Omega (Sievers et al., 2011).

3.2.2 In vivo expression of RAP2.12 in Arabidopsis seedlings and monitoring stabilisation under hypoxia

The construct *promRAP2.12::MC-RAP2.12-GUS* was transformed into Arabidopsis by Dr Sophie Berckhan, a former member of the Holdsworth lab. This resource was valuable to use to indicate where RAP2.12 would accumulate under normoxic and hypoxic conditions. Hypoxia is expected to stabilise RAP2.12 in *promRAP2.12::MC-RAP2.12-GUS* seedlings by preventing the oxidation of Cys-2 and thus its subsequent degradation. Whilst mutant RAP2.12 with the stabilising Ala residue at positon 2 in *promRAP2.12::MA-RAP2.12-GUS* seedlings will be expected to be constitutively stabilised under normoxia. RNA microarray data extracted from different regions of wild-type Arabidopsis seedlings, has identified that expression of *RAP2.12* is higher in above ground shoots than below

ground roots (Figure 3.3) (Birnbaum et al., 2003; Nawy et al., 2005; Yadav et al., 2009). Expression of *RAP2.12* is high in the shoot apical meristem (SAM) of 7 day old (7d) seedlings, especially in the rib meristem but also in the peripheral and central zones (Yadav et al., 2009). Expression of *RAP2.12* is lower in the roots compared to shoots with a gradient reduction towards the root tip.

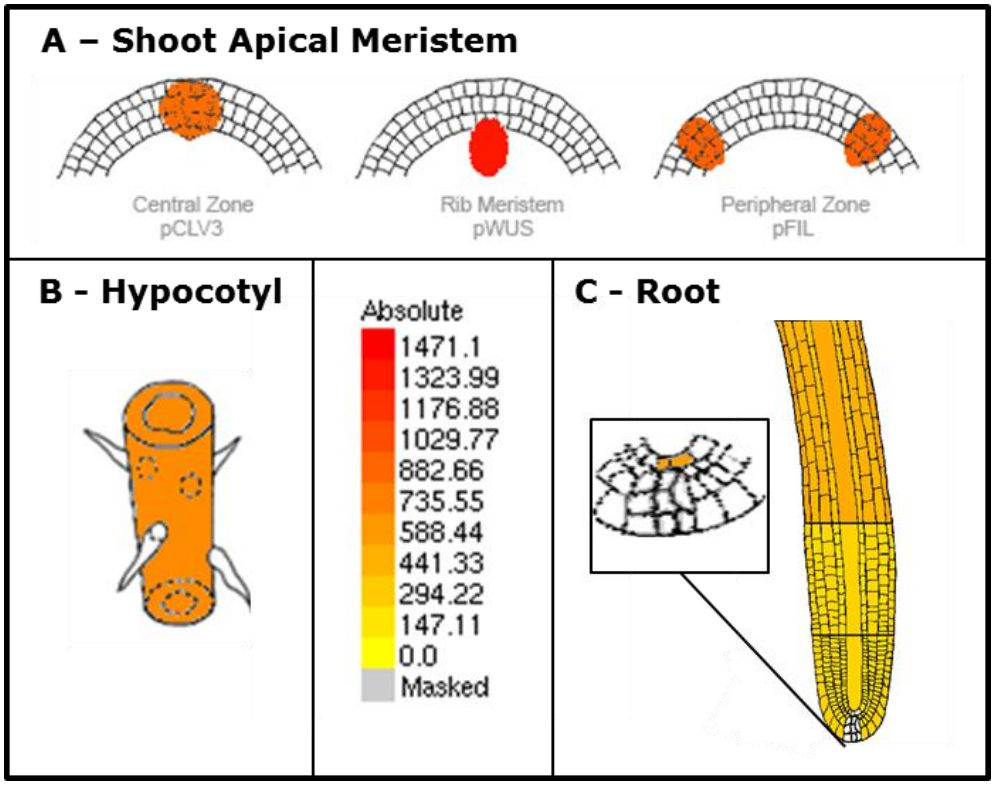


Figure 3.3 Schematic representation of the expression of *RAP2.12* (AT1G53910) in the shoot apical meristem, hypocotyl and root. The scale bar in the centre reveals that darker reds show area of high gene expression. Data taken from the eFP-browser (Winter et al., 2007), based on original work by Birnbaum et al. (2003), Nawy et al. (2005) and (Yadav et al., 2009).

Arabidopsis 7d seedlings containing *promRAP2.12::MC-RAP2.12-GUS* were exposed to either severe hypoxic conditions (<1 % O_2 for 3 hours) in the dark or kept in the dark under 'normoxic' ambient oxygen. Following treatment the seedlings were incubated in X-gluc to locate the

GUS-tagged RAP2.12 protein. GUS-tagged RAP2.12 was visible in seedlings from both treatments in the hypocotyl, shoot apical meristem (SAM) and first true leaves (Figure 3.4). The identification of stabilised RAP2.12 in both the SAM and hypocotyl under normoxia, where substrates such as RAP2.12 would be expected to be oxidised and degraded, may be a result of low physiological oxygen conditions which occur in quickly developing regions of young seedlings. The understanding of the spatial resolution of oxygen dynamics within plants is poor, but oxygen does play an important role regulating stem cell biology (see review Mohyeldin et al., 2010).

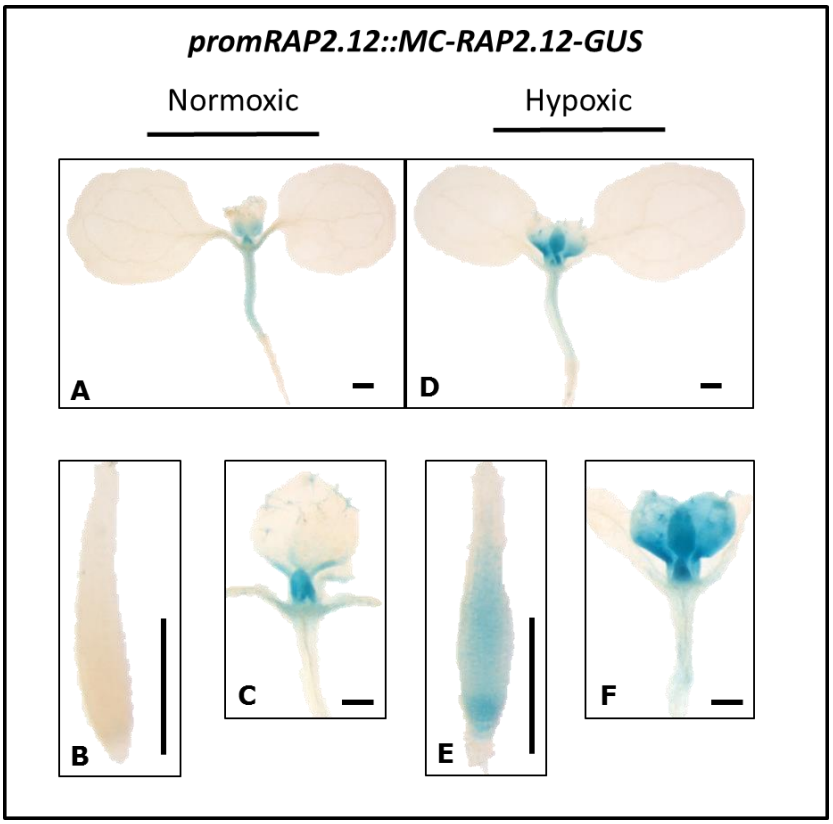


Figure 3.4 Comparison of expression and stabilisation of MC-RAP2.12-GUS in 7 day old *promRAP2.12::MC-RAP2.12-GUS* Arabidopsis seedlings in normoxic (dark, ambient oxygen) or treated with hypoxia (dark <1 % oxygen for 3 hours). Increased stabilisation of MC-RAP2.12-GUS is observed in the shoot apical meristem after hypoxia (D and F) compared to normoxia (A and C). Stabilisation was also observed in the root cap after hypoxia (E) when none was observed in normoxia (B). Scale bar = 250 µm.

An important feature of meristematic cells is their ability to use oxygen as a redox signal controlling dormancy, for example developing grape (*Vitis vinifera*) buds transition from a state of quiescence under hypoxia to development coinciding with an oxidative burst when favourable environmental conditions return (Meitha et al., 2015). It has also been shown previously that the quiescent centre of the Arabidopsis root apical meristem (RAM), a dividing cell zone, is maintained at a strongly oxidising environment (Jiang et al., 2003) hence this also has been postulated for the SAM (Considine and Foyer, 2014). However the oxidising RAM environment may be specific to the root due to the increased tendency for these organs to be submerged, with water limiting diffusion by 10,000 fold compared to air (Armstrong and Drew, 2002). Figure 3.4 shows no stabilisation of RAP2.12 in the RAM under normoxic conditions perhaps supporting this strong oxidising environment postulated by Jiang et al. (2003).

The enhanced expression of *RAP2.12* in the SAM as presented in figure 3.3 would make this area an ideal oxygen sensing region which could help regulate growth of the shoot. This indicates that RAP2.12 may be involved in some aspects of shoot establishment and SAM development, during normoxic conditions reduced oxygen levels have been recorded in the hypocotyl and SAM tissue of Arabidopsis (van Dongen et al., 2003). During seedling development these regions of the plant are undergoing cell expansion which consumes oxygen, additionally as shoot tissue thickens oxygen also declines *in vivo*. As hypocotyl tissue grows it becomes thicker and more anoxic with oxygen levels dropping to 14% and 6% in 6d and 10d Arabidopsis seedlings respectively (van Dongen et

al., 2003). Therefore this high expression of *RAP2.12* in both the hypocotyl and SAM could occur in conjunction with increased stabilisation caused by reduced oxygen-driven degradation. Papdi et al. (2008) initially identified how *RAP2.12* activates transcription of the anaerobic respiration gene *ALCOHOL DEHYDROGENASE* (*ADH1*) which is induced when the plant is suffering abiotic stress such as hypoxia or in the presence of abscisic acid (ABA). The overexpression of *RAP2.12* or the other ERFVII RAP's individually was later found to also confer increased tolerance to abiotic pressures such as extreme osmotic potentials and oxidative stress (Papdi et al., 2015).

When seedlings of *promRAP2.12::MC-RAP2.12-GUS* were subjected to severe hypoxia (<1 % O_2 for 3 hours) the intensity of GUS staining increased (Figure 3.4D-F). This was especially prominent in the SAM and in the RAM. This increased accumulation after the seedlings were exposed to hypoxia indicates that the RAP2.12 is not being degraded, caused by the lack of oxidation of Nt-Cys. This supports recent observations by Kosmacz et al. (2015) who used a GFP-tagged RAP2.12 construct to monitor translocation of RAP2.12 to the nucleus and increased stabilisation under hypoxia. The accumulation of RAP2.12 in the RAM (3.4E) but not in other regions, indicates that RAP2.12 could be sensing oxygen in the soil through the RAM and potentially determining the root growth patterns. Interestingly, microarray data from Arabidopsis treated with hypoxia showed slightly reduced *RAP2.12* expression throughout wild-type seedlings after hypoxia (Figure 3.5) (Branco-Price et al., 2005).

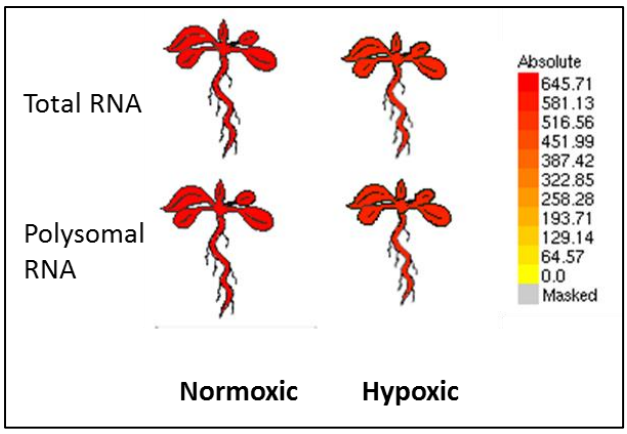


Figure 3.5 Schematic representation of the expression of *RAP2.12* (AT1G53910) in total and polysomal RNA compared in normoxic and hypoxic conditions. The scale bar in the centre reveals that darker reds show area of high gene expression. Data taken from the eFP-browser (Winter et al., 2007), with data from Branco-Price et al. (2005).

3.2.3 Stabilisation of MA-RAP2.12-GUS and comparison to MC-RAP2.12-GUS

Proteins bearing a destabilizing residue at their N-terminus can be made stabilizing if this is changed to a stabilising residue. When untreated Arabidopsis seedlings of *promRAP2.12::MA-RAP2.12-GUS* were stained just like their MC counterpart, the intensity of staining at the SAM was greater (Figure 3.6). Expression of *RAP2.12* in both the MA and MC construct should in theory be the same, as they are promoted by the same native promoter however this can vary dependent on the activity of the chromosomal region at which each construct was integrated. A sequence of repetitive DNA in the region of integration, or the post transcriptional methylation of the foreign gene has been shown to influence gene activity in petunia (Pröls and Meyer, 1992). The increased intensity of staining of the SAM tissue in the MA construct is a result of the Nt-Ala (after Met excision) not being recognised as destabilising, thus

this protein was not degraded by the N-end rule machinery. Stabilisation of RAP2.12 in the SAM was greater in the MA construct than *promRAP2.12::MC-RAP2.12-GUS* seedlings treated with hypoxia. Hypoxia-induced stabilization of *promRAP2.12::MC-RAP2.12-GUS* increased its prominence in both the hypocotyl and RAM (Figure 3.4). Seedlings of *promRAP2.12::MA-RAP2.12-GUS* did not show staining in the RAM at the same intensity of the hypoxia treated MC construct (Appendix B). It is therefore possible that RAP2.12 production had been reduced or silenced in these regions in the MA construct, preventing the build-up of stable MA-RAP2.12-GUS allowing more resources to be invested elsewhere. It has been reported that if the transcript level of a transgene, such as the mutant *promRAP2.12::MA-RAP2.12-GUS*, exceeded the gene-threshold RNA silencing can occur (Schubert et al., 2004). Although why this would only occur in the hypocotyl and RAM and not in the SAM would have to depend on different thresholds existing in different cells, which in relation to RAP2.12 is still unknown.

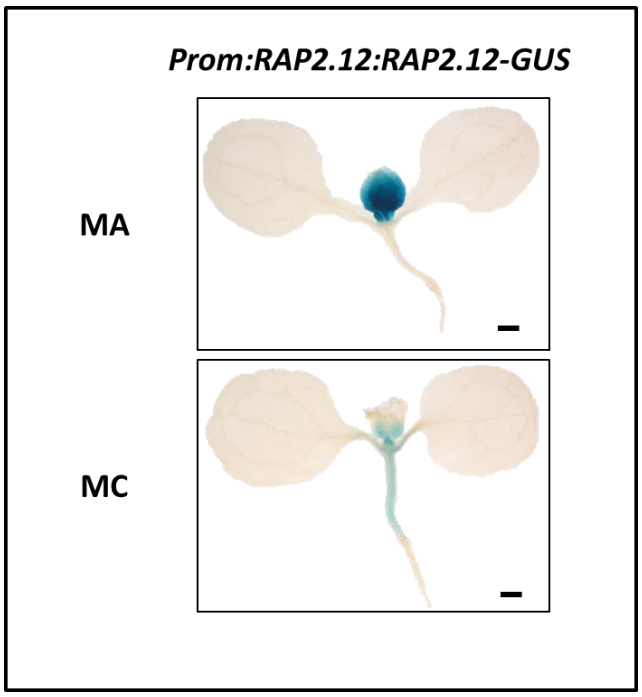


Figure 3.6 Comparison of expression and stabilisation of MC-RAP2.12-GUS and MA-RAP2.12-GUS in 7d old Arabidopsis seedlings of *promRAP2.12::MC-RAP2.12-GUS* and *promRAP2.12::MA-RAP2.12-GUS* respectively in normoxic conditions (dark, ambient oxygen). Increased stabilisation of MA-RAP2.12-GUS is observed in the shoot apical meristem compared to MC-RAP2.12-GUS. Scale bar = 250 µm.

3.2.4 RAP2.12 and the other ERFVIIs are important regulators of photomorphogenesis

Recently the ERFVIIs have been identified to be involved in early seedling development. Etiolated wild-type Arabidopsis (Col-0) fail to completely open their apical hook, a protective structure shielding the SAM and cotyledons, when oxygen is reduced below 14% (Abbas et al., 2015). An increase in hook angle correlates with decreasing oxygen conditions (Figure 3.7A), indicating that the apical hook has the ability to sense oxygen in the surrounding environment. Therefore, in addition to the

absence of light, reduced oxygen conditions in the soil help to prevent damage as young seedlings emerge towards the soil surface. The ERFVIIs have previously been reported to sense oxygen and NO (Gibbs et al., 2011; Gibbs et al., 2014b; Licausi et al., 2011), and their possible involvement in controlling opening of the apical hook was investigated.

Using a pentuple mutant *erfVII* which lacks the five ERFVIIs, apical hook development was compared with Col-0 during the normoxic (0 – 48h) and hypoxic (48 – 144h) phases of the experiment. The Col-0 seedlings which possess the ERFVIIs failed to complete hook opening during hypoxia, however the *erfVII* seedlings continue to develop and complete hook formation similarly to Col-0 in normoxia (Figure 3.7B). This indicates that the absence of the ERFs and hence the ability for seedlings to sense oxygen, removes an important protective component of seedling development. To confirm that the ERFVIIs are integral to regulating the inhibition of apical hook opening under hypoxia, the individual ERF's with a modified Met-Ala N-terminus expressed under their native promoter were monitored under normoxic conditions. Similarly to Col-0 in hypoxia the individual ERFs driven by their own promoter including *promRAP2.12::MA-RAP2.12-GUS*, all displayed inhibited hook opening suggesting that the native Met-Cys ERFVIIs are substrates of the Arg/N-end rule pathway (Figure 3.7C) (Abbas et al., 2015). In prolonged periods of darkness low oxygen conditions (<14% O_2) increase seedling survival when entering a light environment, conversely seedlings which remain in normoxic darkness for a long period before entering a light environment have reduced survival. This response is controlled entirely by the ERFVIIs which maintain seedling viability in the dark and in their

absence cotyledon greening is inhibited in response to light (Abbas et al., 2015).

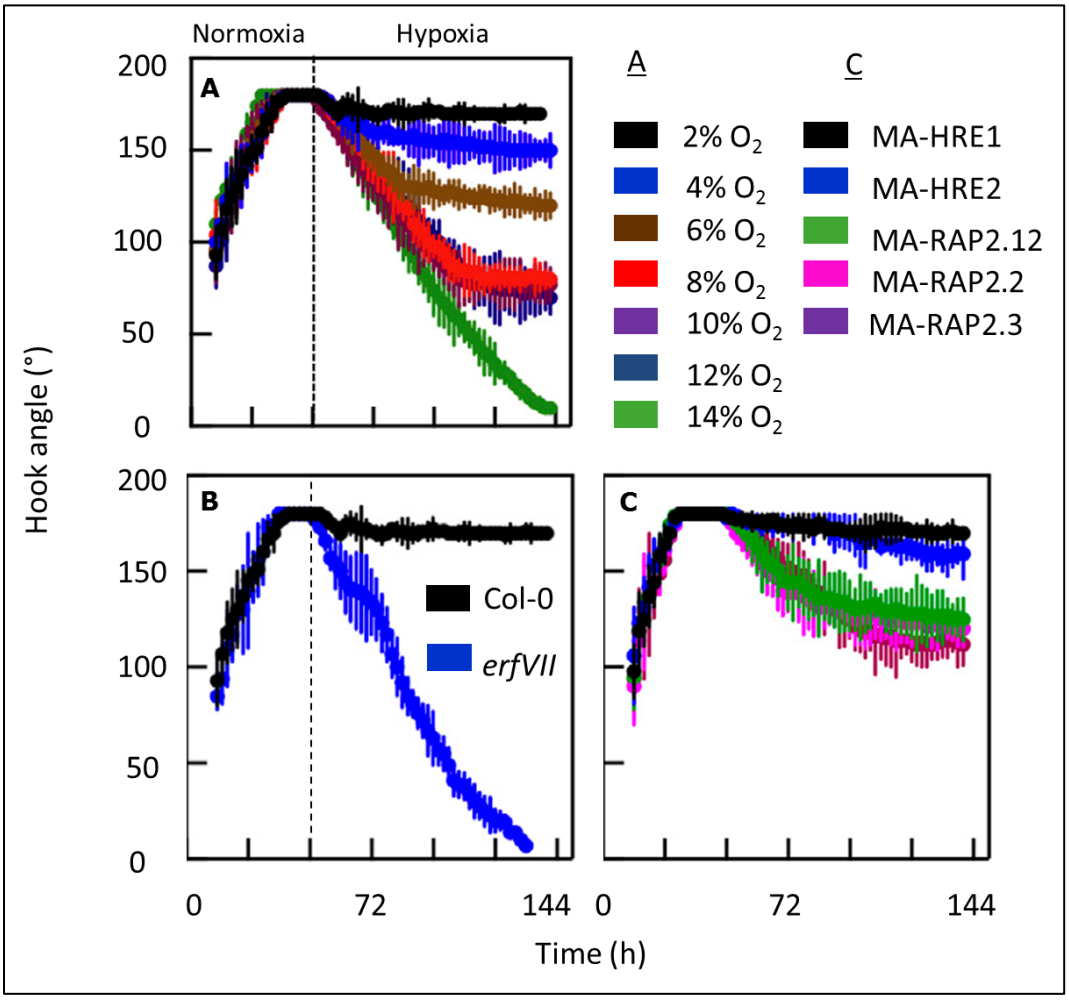


Figure 3.7 The influence of oxygen and the group VII ethylene response factor transcription factors (ERFVIIs) in controlling apical hook opening. (A) There is a linear correlation between hook opening angle and oxygen concentration, with oxygen levels below 14% inhibiting opening. (B) Pentuple mutants (*erfVII*) lacking all five Arabidopsis ERFs (RAP2.12, RAP2.2, RAP2.3, HRE1 and HRE2) can complete hook opening under hypoxic conditions providing unequivocal evidence that the ERFVIIs sense oxygen to control hook opening. (C) The ability of each individual ERF to repress hook opening was demonstrated by creating Met-Ala N-termini versions (including *promRAP2.12::MA-RAP2.12-GUS*) (Adapted from Abbas et al., 2015). Experiments performed by Mohamad Abbas.

3.3 *In vivo* expression of reporter proteins bearing the conserved N-terminal motif of plant ERFVIIs in Arabidopsis

3.3.1 Conserved N-terminal motif of the ERFVIIs

A conserved feature amongst the ERFVIIs in both Arabidopsis and rice is their sequence similarity at the N-terminus. Nakano et al. (2006) conducted a study looking at the phylogenetic relationship between different groups of ERFs in Arabidopsis and rice, finding that there was high conservation between them. The MCGGAI(I/L) motif common to ERFVIIs was identified in all Arabidopsis ERFVIIs. The rice ERFVII sub-family is larger and has similar Nt sequences were identified in *SUB1* and *SNORKEL* genes (Table 1). The presence of ERFVIIs in both Arabidopsis (dicot) and rice (monocot) is important evidence indicating that genes of this ERF sub-group developed before the divergence of these species (Nakano et al., 2006).

The maize Ubiquitin (*Ubi1*) promoter has been demonstrated to be effective in successfully promoting transgenes in cereals (Bartlett et al., 2008). A modified β-glucuronidase (GUS) reporter protein bearing the evolutionary conserved MCGGAI(I/L) or a mutant stabilising MAGGAI(I/L) Nt motif, driven by the *Ubi1* promoter was previously developed by the Holdsworth lab (*Ubi1::*M(C/A)GGAIL-GUS) (Figure 3.8).

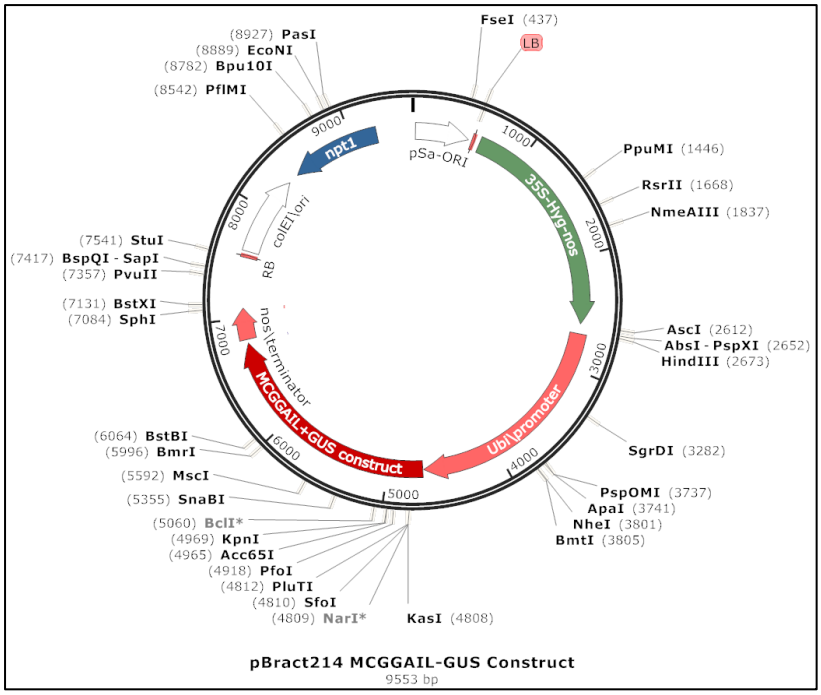


Figure 3.8 Schematic of the *Ubi1::MCGGAIL-GUS* construct. Showing the Hygromycin resistance gene (green), the maize ubiquitin (*Ubi1*) promoter (pink) and the MCGGAIL residues and GUS construct (red). Restriction sites for numerous enzymes are labelled on the outside of the construct.

Table 3.1 The conserved N-terminal motif across Arabidopsis and Rice (*Oryza sativa*) group VII ethylene response factor transcription factors (ERFVIIs).

Locus	Name	Sequence (first 12 residue)	Plant
AT1G53910	RAP2.12	MCGGAIISDFIP	Arabidopsis
AT3G14230	RAP2.2	MCGGAIISDFIP	Arabidopsis
AT3G16770	RAP2.3	MCGGAIISDYAP	Arabidopsis
AT1G72360	HRE1	MCGGAVISDYIA	Arabidopsis
AT2G47520	HRE2	MCGGAIISDFIW	Arabidopsis
SUB1	SUB1A-1	MCGGEVIPADMP	Rice
	SUB1B-1	MCGGALIPNDYG	Rice
	SUB1C-1	MRRRVSSSPSSS	Rice
SNORKEL	SNORKEL1	MCGGCLIPDELV	Rice
	SNORKEL2	MCGENDNNGAAA	Rice

3.3.2 Transforming Ubi1::M(C/A)GGAIL-GUS in Arabidopsis and the in vivo expression of M(C/A)GGAIL-GUS

The *Ubi1::M(C/A)GGAIL-GUS* constructs have previously been transformed into Barley and demonstrated to be oxygen and nitric oxide sensors. In the absence of Nitric oxide (NO) MCGGAIL-GUS is stabilised

in the embryos of barley (Gibbs et al., 2014b), highlighting the neccessity of NO-driven Cys oxidation in the Arg/N-end rule pathway in this species, and the validity of this construct. The untreated MAGGAIL-GUS construct behaved similarly to MCGGAIL-GUS in the absence of NO, indicating that Ala is a stabilising residue. Waterlogged barley seedlings, where the roots of the plant were completely submerged for 15 days thus limiting oxygen diffusion, show accumulation of MCGGAIL-GUS compared to when in control conditons where stabilisation was not oberved (Figure 3.9) (Mendiondo et al., 2016). Greater stabilisation was observed in both treatments compared to MCGGAIL-GUS in barley containing MAGGAIL-GUS.

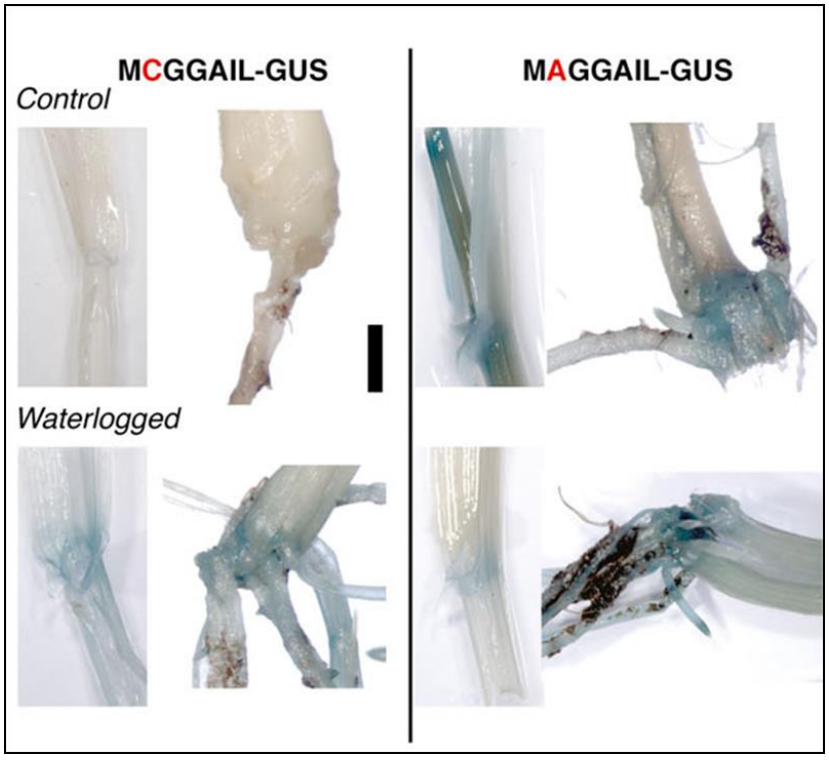


Figure 3.9 Expression and stabilisation of the GUS reporter protein in transgenic *Ubi1::M(C/A)GGAIL-GUS* barley (*Hordeum vulgare*) when in control or waterlogged conditions for 15 days. The MCGGAIL-GUS is stabilised in barley roots and shoots after GUS staining waterlogged seedlings, highlighting the requirement for Cys oxidation. The MAGGAIL-GUS construct containing the stabilising Ala N-termini was stabilised in root and shoots in both treatment conditions (image taken from Mendiondo et al., 2016).

The *Ubi1::M(C/A)GGAIL-GUS* constructs were transformed into WT Arabidopsis to observe whether the functionality of both constructs would be the same as in Barley. Once transformed and selected 7d homozygous lines of Arabidopsis containing the MCGGAIL-GUS and MAGGAIL-GUS constructs were stained with X-gluc. MCGGAIL-GUS was weakly expressed in the hypocotyl, petioles and SAM of 7d seedlings (Figure 3.10), an area where cells are expanding as the seedling grows which could be consuming oxygen and hence reducing the O_2 partial pressure (ppO_2) causing stabilisation (van Dongen et al., 2003). In contrast to the weak expression in MCGGAIL-GUS the MAGGAIL-GUS was expressed throughout the seedling.

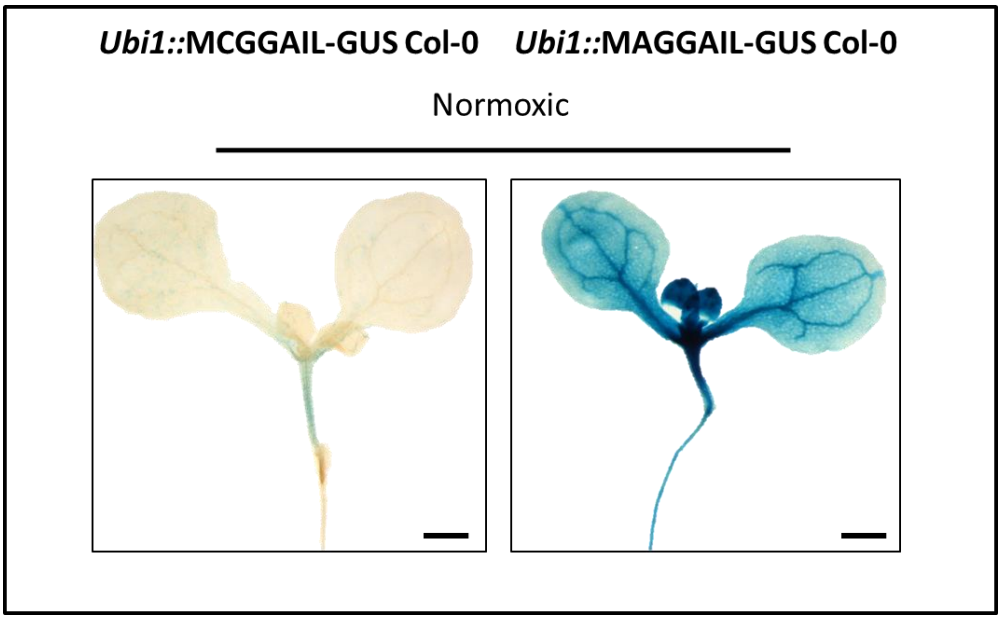


Figure 3.10 Expression of the GUS reporter protein in transgenic *Ubi1*::MCGGAIL-GUS 7 day old Arabidopsis seedlings. MCGGAIL-GUS is stabilised in the hypocotyl and petioles seedlings. MAGGAIL-GUS was stabilised throughout the seedling. Scale bar = 250 µm.

3.3.3 Stabilsation of M(C/A)GGAIL-GUS under hypoxia and treatment with the 26S proteasome inhibitor bortezomib

To determine whether MCGGAIL-GUS is a sensor of oxygen in Arabidopsis, 7d *Ubi1::MCGGAIL-GUS* seedlings were exposed to severe hypoxic conditions (<1 % O_2 for 3 hours) and its stability was observed. Increased stabilisation of MCGGAIL-GUS was visble in the hypocotyl, petioles, SAM and cotyledons as a result of hypoxia treatment, supporting the results of Gibbs et al. (2011) and Mendiondo et al. (2016) highlighting how the Nt-Cys residue of this substrate can sense oxygen *in vivo* (Figure 3.11). In disagreement with the results for barley where stabilisation of MCGGAIL-GUS is visible throughout sections of the crown root (Mendiondo et al., 2016), stabilisation of MCGGAIL-GUS in Arabidopsis roots was only visible in the root elongation zone (REZ) and to a greater extent in the adjoining cells in the maturation zone. Unlike the shoots of Arabidopsis the onset of hypoxia did not cause increased stabilisation in the roots, however MAGGAIL-GUS was stabilised to a greater extent in the RAM and REZ (Appendix C). The accumulation pattern of MCGGAIL-GUS in barley and Arabidopsis in figures 3.9 and 3.11 is not directly comparable as the treatments were different. Waterlogging is more prolonged than the hypoxic treatment but oxygen levels would not reach the severe low levels created in the hypoxic chamber. A further variable here is the promoter, which has previously been used predominantly in cereals and therefore may not have the same activity when active in Arabidopsis. The activity of monocot promoters in dicots and vice versa has been found to be lower than when

within their native flowering plant group (Christensen et al., 1992; Wilmink et al., 1995).

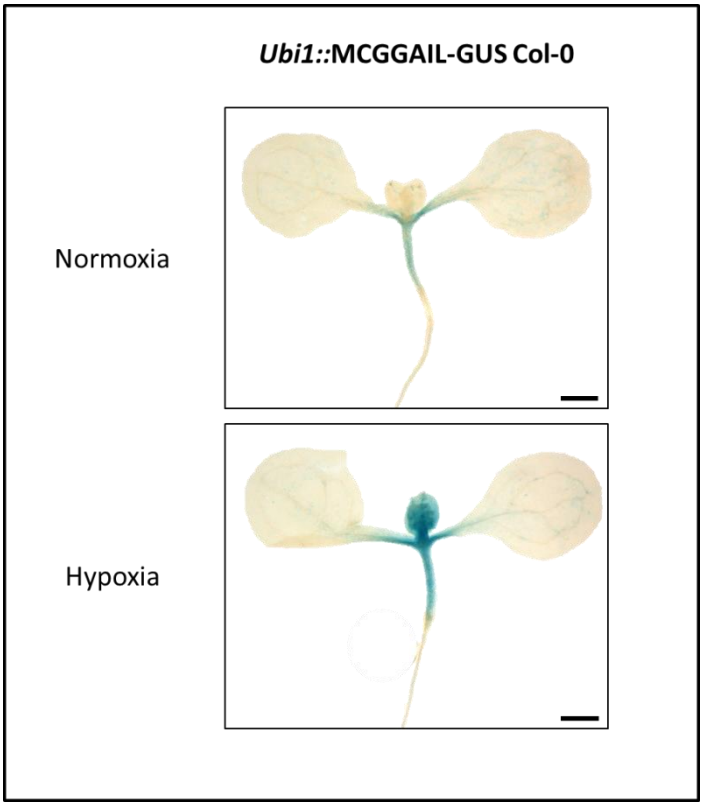


Figure 3.11 Expression and stabilisation of the GUS reporter protein in transgenic *Ubi1*::MCGGAIL-GUS 7 day old Arabidopsis seedlings when untreated and exposed to hypoxia. Increased stabilisation of MCGGAIL-GUS was observed after hypoxia in the hypocotyl, petioles and shoot apical meristem compare to the untreated normoxic control. Scale bar = 250 µm.

To quantify the results presented in figure 3.11 an anti-GUS antibody was used in Western blot analyses to probe GUS levels in samples. In addition to hypoxia treatment 7d seedlings of Arabidopsis containing transgenes expressing either MC or MAGGAIL-GUS were also incubated in bortezomib, an inhibitor of the 26S proteasome (control DMSO). In support of figure 3.11 increased stabilisation of MCGGAIL-GUS was observed under hypoxia which was not observed with MAGGAIL-GUS, that is constitutively stabilised (Figure 3.12). This result in Arabidopsis is

in agreement with previous anti-GUS antibody Western blots presented in Barley, when increased stabilisation of MCGGAIL-GUS was observed when waterlogged (Mendiondo et al., 2016). Bortezomib effectively inibited the degradation of MCGGAIL-GUS in respect to the DMSO control, this increased stabilisation was not observed in MAGGAIL-GUS highlighting how this protein is not a substrate of the Arg/N-end rule pathway. There was also a noticeable difference between the control samples normoxia and DMSO, where incubating seedlings with DMSO causes slightly increased stabilisation. This difference could be caused by numerous factors, the partial waterlogging caused by incubating seedlings in liquid DMSO, which may hinder oxygen diffusion, slowing *in vivo* Cys oxidation. The chemical effects of DMSO could also influence protein stability, Tjernberg et al. (2006) revealed that at low concentrations DMSO can change the properties of a protein, recoding possible DMSO-induced denaturation, degradation and aggregation as well as other surface-binding related changes. In addition to this an early study on lipase from wheat germ reported that DMSO had a stabilising effect on the enzyme (Rajeshwara and Prakash, 1994).

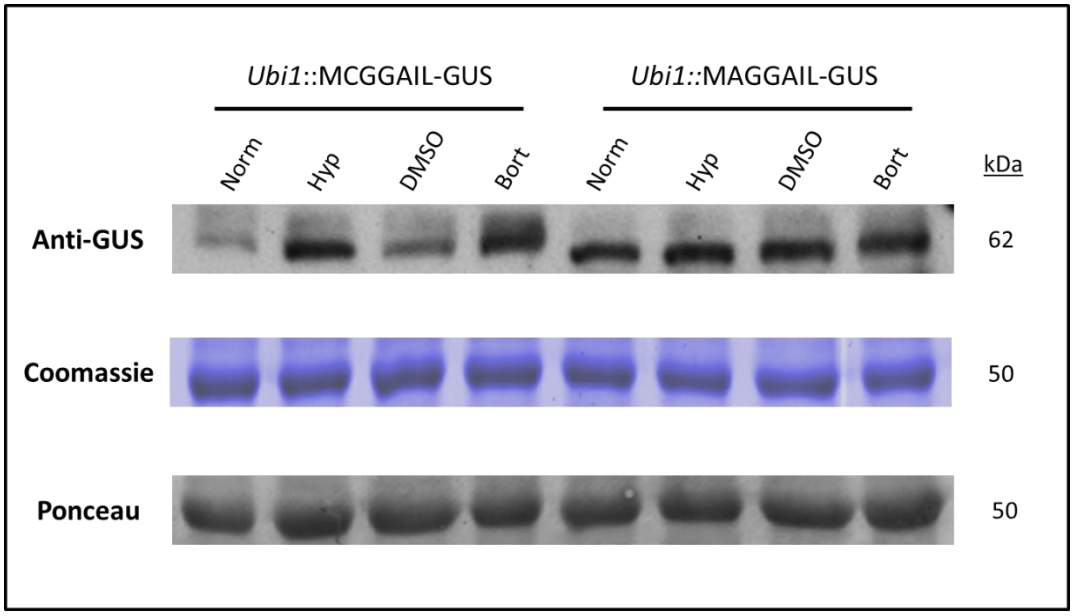


Figure 3.12 Anti-GUS antibody Western blot showing the abundance of the GUS reporter protein extracted from 7d transgenic *Ubi1::M(C/A)GGAIL-GUS* Arabidopsis seedlings either untreated, exposed to hypoxia, incubated in dilute DMSO or in bortezomib dissolved in DMSO (26S proteasome inhibitor). Increased stabilisation of MCGGAIL-GUS was observed after hypoxia and bortezomib treatment compared to the normoxic and DMSO control. Neither of the treatments affected stabilisation of MAGGAIL-GUS which remained the same throughout. Coomassie loading control and Ponceau transfer controls (protein present on membrane used for Western analysis) for each sample also visible.

3.4 Developing a Cys-2 reporter of the Arg/N-end rule pathway to be used for protein purification

3.4.1 Analysing of expression and accumulation levels of existing Met-Cys reporter proteins

Transgenic Nt Met-Cys-polyG-HA-GUS reporter lines (MC-GUS, see figure 2.1 section 2.1.1) in a *prt6* background were created by Professor Freddie Theodoulou (Rothamsted Research, Harpenden, UK), these were analysed to determine their expression levels. The Met-Cys reporters (*35S::MC-GUS*) were expressed under the control of the CaMV 35S promoter and contained an N-terminal Met-Cys (MC) motif extended from the rest of the protein by a poly-glycine chain, followed by a further extension which comprises part of the lac repressor from bacteria (resembling early N-end rule constructs Ub-β-gal and X-β-gal, (Bachmair et al., 1986; Bachmair and Varshavsky, 1989), proceeded by a 3x HA tag and a β-glucuronidase (GUS) protein (See section 2.1.1). Two of these transgenic lines, A2/1 and 13d, have very high and low expression respectively visible after analysis by Western blot and GUS staining. These lines had also previously been crossed into the WT (Col-0) background by former Holdsworth lab member Dr Daniel Gibbs (now University of Birmingham, UK). To quantify expression of three other MC-GUS transformants protein was extracted from 7d Arabidopsis seedlings, before performing SDS-PAGE and a Western blot incubating with an anti-HA antibody (Figure 3.13).

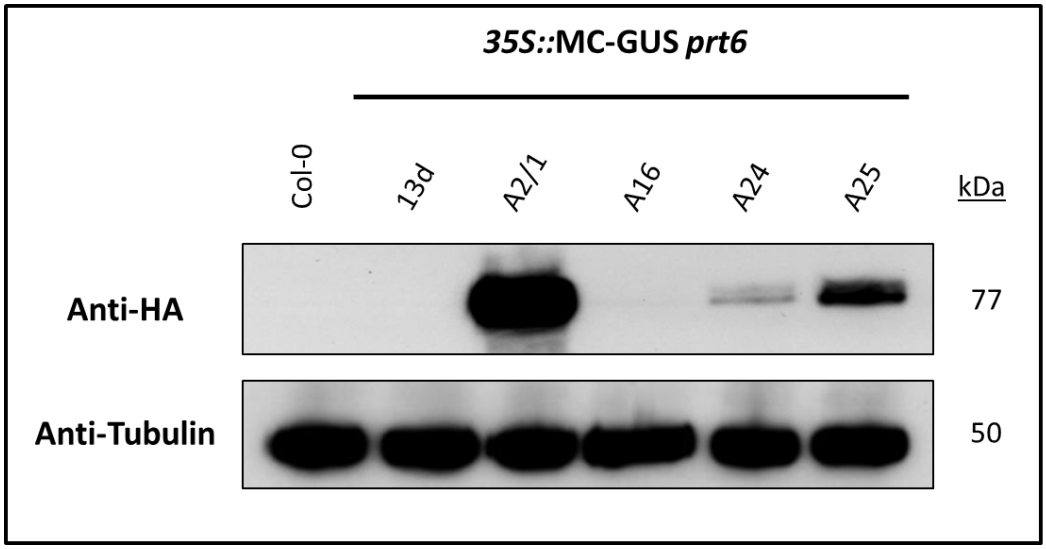


Figure 3.13 Anti-HA antibody Western blot showing the stabilisation of five GUS reporter protein lines extracted from 7d old *35S::MC-GUS* transgenic Arabidopsis seedlings in the *prt6* background when untreated. Line A2/1 showed the highest GUS accumulation indicating this line has the strongest expression, followed by A25, A24, A16 then 13d. An Anti-Tubulin loading control for each sample is also visible.

There was noticebable variation in GUS accumulation visible from different transgenic seedlings with line A2/1 having the strongest expression, followed by A25, A24, A16 and then 13d. Lines A24 and A25 were selected to be taken forwards to cross into Col-0 as their intermediate exression was predicted to show a noticeable difference if treated with normoxia/hypoxia. However, only results for the A25 line will be presented, as this was the stronger expresser of the two and deemed more suitable to maximise protein detection by MS (see Section 4). The very weak expression of A16 would likely make it diffcult to detect the GUS protein on a coomassie gel, hence this would require more protein to purify and was discounted to take forward into Col-0. After crossing, the F1 seeds were collected and selfed to F2 before being genotyped to identify MC-GUS in the wild-type Col-0 background (Figure 3.14).

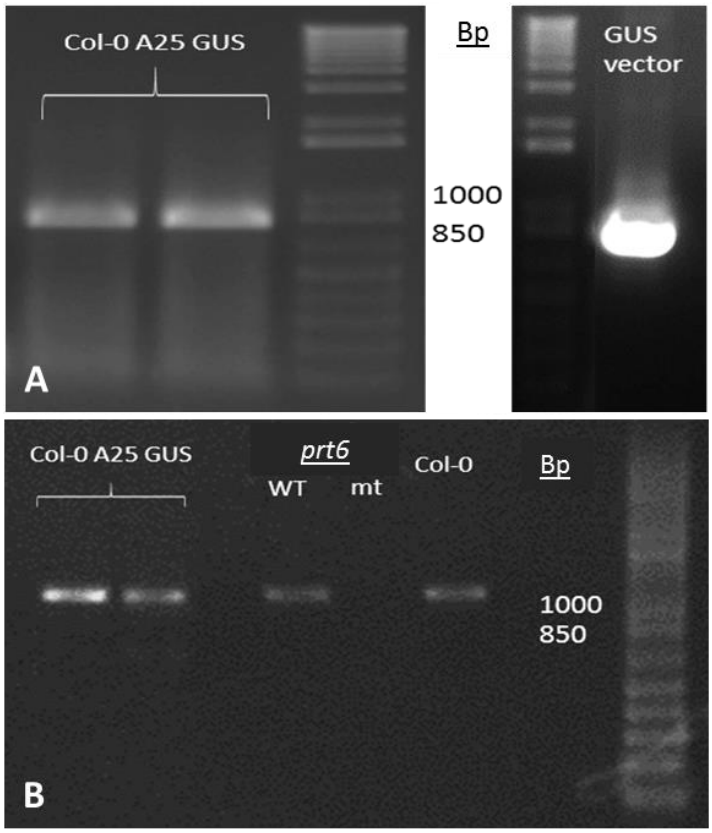


Figure 3.14 Genotyping to identify the intermediate expressing *35S::MC-GUS* A25 Col-0 plants. (A) F2 seedlings from a cross between *35S::MC-GUS* in *prt6-5* background (mt)and wild-type (WT) Col-0 were genotyped to select for seedlings containing GUS by performing a polymerase chain reaction (PCR) using GUS primers. Seedling DNA which transcribes GUS amplified a band at ~850 Bp. (B) Seedlings containing GUS were grown to the F3 generation before being screened for *prt6-5*, using a combination of primers. Homozygous WT lines were identified by ~1050 Bp band.

3.4.2 Identifying differences of in vivo GUS reporter protein accumulation and comparing in Col-0 and prt6 backgrounds

Once homozygous, 7d seedlings of *35S::MC-GUS* line A25 in both the Col-0 and *prt6* background were stained to compare expression to its counterparts of the A2/1 line. Seedlings were stained in X-gluc for 90 minutes, as both lines stain quickly and the staining becomes less

discernible between Col-0 and *prt6* backgrounds if stained overnight. Results of staining presented in figure 3.15 clearly show that line A2/1 accumulated more GUS than A25 in the *prt6* background (supporting Figure 3.13), but also that the mutant *prt6* background for both lines stabilises MC-GUS more than in Col-0. The accumulation of GUS is consistently strongest in all lines and backgrounds at the root hair development zone (RHDZ), in between the hypocotyl and the REZ. The accumulation of GUS generally diminshes from the RHDZ towards the root cap, and this is much more prominent in the Col-0 background. Regarding the shoot tissue no GUS accumulation was observed after 90 mins in the A25 line when in the Col-0 or *prt6* background, however accumulation was observed in the A2/1 line with the petioles staining in the Col-0 background and the petioles and cotyledons in *prt6*.

In contrast to the *Ubi1* promoter used with the MCGGAIL GUS constructs (Section 3.3), the expression of the GUS reporter protein driven by the CaMV 35S promoter is concentrated in the roots and not the SAM and hypocotyl of 7d Arabidopsis seedlings. The use of CaMV 35S promoter, compared to transgenes under a native or weaker promoter, has been found to potentially alter where the transgene is expressed and ultimately the phenotype in Arabidopsis (Yoo et al., 2005). The site of integration of a foreign construct on the chromosome after transformation, can have a potential spatial and temporal effect on the expression of other promoters in its vicinity (Zheng et al., 2007). Therefore it can change the specificity of expression for native genes making them constitutive expressers all over the plant (Zheng et al., 2007). It is therefore likely that the accumulation of GUS in the roots of

Arabidopsis Col-0, despite it being weaker in the A25 line, is due to the strong expression of the CaMV 35S promoter where GUS may saturate the actve site of the PCOs, ATEs or PRT6 causing stabilisation. However it is also possible that high growth activity at the RHDZ is very oxygen intensive, creating an hypoxic environment preventing the oxidation of the Nt-Cys of MC-GUS (after Met excision) reporter protein, thus stabilsing the protein.

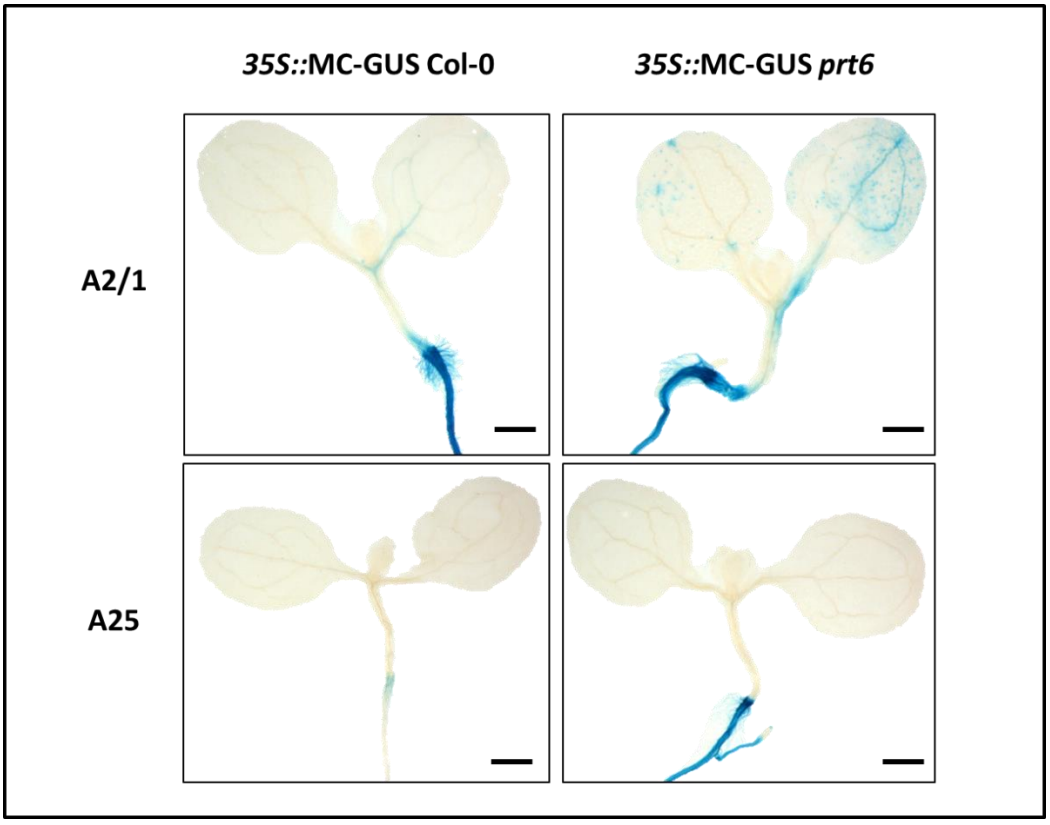


Figure 3.15 Expression and stabilisation of GUS in lines A2/1 and A25 in Col-0 and *prt6* backgrounds of 7d old *35S::MC-GUS* Arabidopsis. The strong expressing line A2/1 shows MC-GUS accumulation throughout the roots and in the petioles when in the Col-0 background. Line A25 only shows that GUS was accumulated in the root hair development zone. In the *prt6* background expression of MC-GUS in line A2/1 is similar to Col-0 albeit accumulation of MC-GUS was visible in the cotyledons. In the intermediate A25 line MC-GUS was accumulated throughout the root although no MC-GUS was stabilised in the cotyledons. Scale bar = 250 µm.

The results obtained from GUS staining experiments correlate well with Western blot analysis (Figure 3.16). The intermediate GUS reporter line A25 was not stable in the Col-0 background from 7d Arabidopsis seedlings, hence no band was observed in the Western blot with anti-HA antibody. This differed from the A2/1 line when in the Col-0 background which appears stabilising despite the presence of the E3 ligase *prt6*, indicating that the strong expression of the MC-GUS transgene is saturating the N-end rule machinery and hence this line is not ideal to use for stabilisation experiments. Line A25 was stabilised in the *prt6* background as expected but at significantly lower levels compared to its A2/1 counterpart.

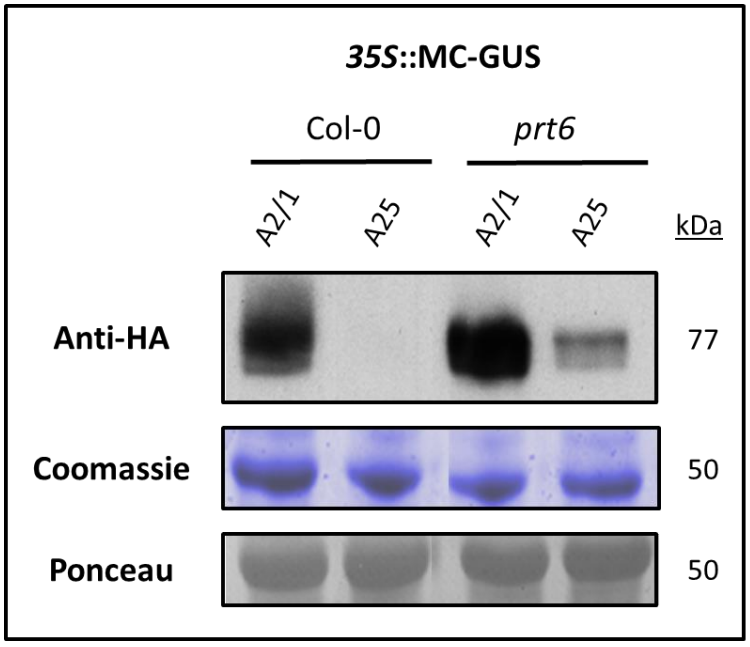


Figure 3.16 Anti-HA antibody Western blot showing the stabilisation of MC-GUS extracted from lines A2/1 and A25 in Col-0 and *prt6* backgrounds of 7d old *35S::MC-GUS* Arabidopsis seedlings. The strong expressing A2/1 line accumulates more MC-GUS in the mutant *prt6* compared to when in the Col-0 background. The intermediate expressing A25 line shows no MC-GUS accumulation when in the Col-0 background (functional N-end rule). However, when the E3 ligase is missing in the *prt6* background stabilised MC-GUS is visible. Coomassie loading control and Ponceau transfer controls (protein present on membrane used for Western analysis) for each sample also visible.

3.4.3 Stabilisation of MC-GUS in the intermediate line A25 after exposure to hypoxia and proteasome inhibitors

Having established MC-GUS line A25 having a more suitable transgene insertion in which to investigate Cys oxidation and stabilisation *in vivo*, 10d *35S::MC-GUS* seedlings of this line were treated with hypoxia (<1 % O_2 for 3 hours) to investigate stabilistion of GUS. This increased accumulation of GUS in the RHDZ but also in the nearby REZ (Figure 3.17). As a control an MA-GUS construct where Cys had been substituted for Ala at position 2 was stained and showed increased accumulation of GUS in the root similar to MC-GUS in the *prt6* background, both showed increased GUS accumulation compared to MC-GUS in Col-0 untreated and treated. This is expected as Ala is a stabilsing residue and thus MA-GUS is only slowly degraded. A similar pattern of stabilisation was observed in regards to the shoot material. In normoxia faint GUS staining was visible in the root cap of the root tip of *35S::MC-GUS* seedlings in the Col-0 background, this was intensified as GUS accumulated after hypoxia (<1 % O_2 for 3 hours) with staining also visible in the REZ and root maturation zone (RMZ) (Figure 3.18). In the *prt6* background *35S::MC-GUS* seedlings exhibited greater stabilisation of GUS throughout the RMZ, REZ, RAM and the root cap when kept under normoxic conditions.

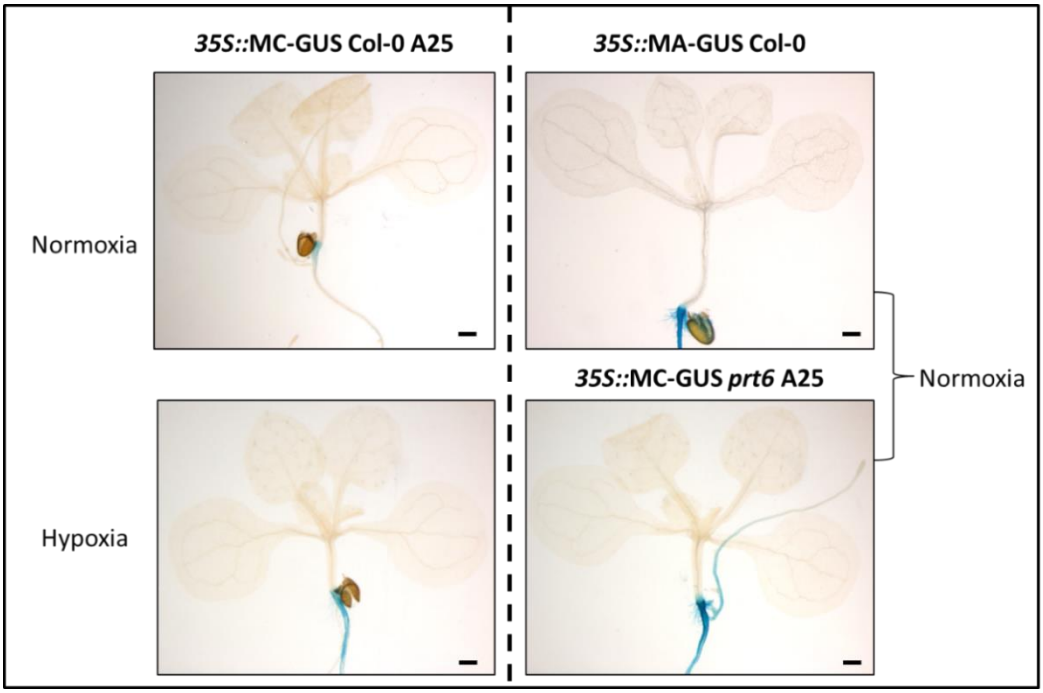


Figure 3.17 Expression and stabilisation of GUS in under normoxia and hypoxia in 10d old *35S::MC-GUS* line A25 Arabidopsis seedlings in Col-0 background, and the constituent stabilisation of GUS in the same line under normoxia in the *prt6* background or when the destabilising Nt Cys-2 is changed to Ala (MA). Increased GUS accumulation occurred in the root hair development zone after hypoxia compared to normoxia, and stabilisation also occurred in the root elongation zone of hypoxia treated 10d seedlings. N-end rule mutants lacking *prt6* had increased levels of stabilised GUS and similarly in the already stabilising Met-Ala initiating MA-GUS construct, which both accumulated high quantities of GUS in the roots. Scale bar = 250 µm.

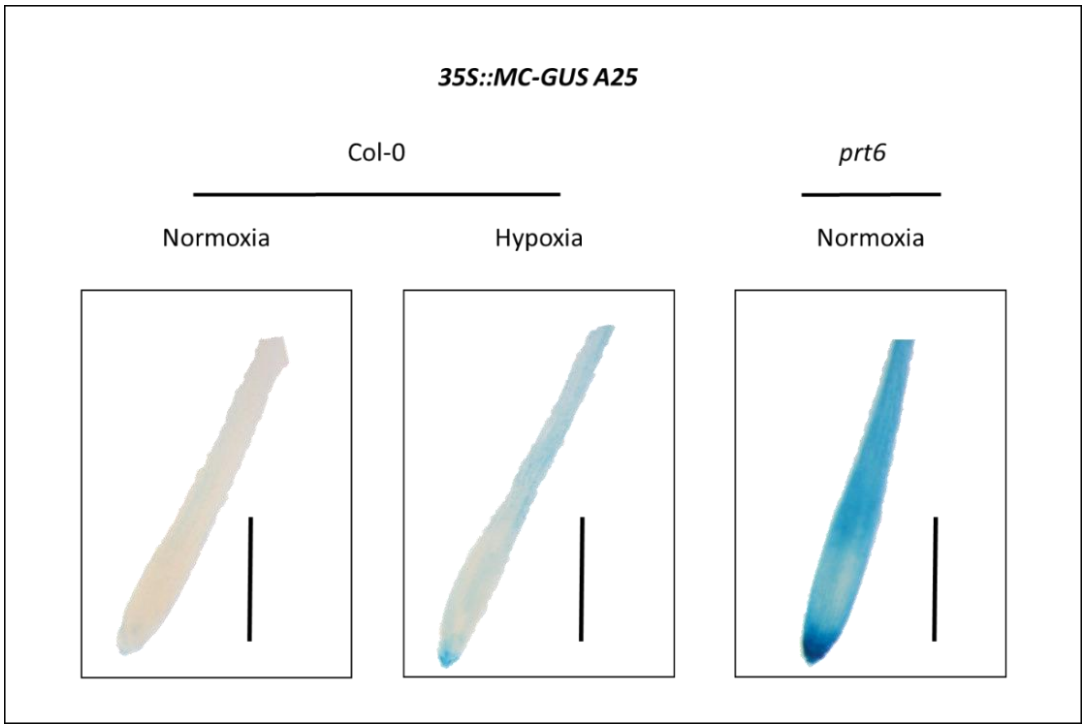


Figure 3.18 Expression and stabilisation of GUS in under normoxia and hypoxia in 10d old *35S::MC-GUS* line A25 Arabidopsis seedlings in Col-0 background, and the constituent stabilisation of GUS in the same line under normoxia in the *prt6* background. Increased GUS accumulation is observed in the root cap and root maturation zone after hypoxia and throughout the root in *prt6* background lacking the E3 ligase. Scale bar = 250 µm.

To quantitatively assess the GUS protein stabilised in the 10d old *35S::MC-GUS* seedlings, crude protein extracts were made and anti-HA antibody was used for Western blot analyses. Using seedlings treated with hypoxia or bortezomib from the WT Col-0 and E3 ligase mutant *prt6*, the effect of a period of hypoxic stress and the inhibition of the 26S proteasome could be observed in the presence of functional and disabled N-end rule machinery. The results for the Western blot with anti-HA antibody, show that MC-GUS line A25 (Col-0) is an oxidatively susceptible substrate of the N-end rule pathway, who's degradation is facilitated by the 26S proteasome (Figure 3.19). In the *prt6* background

there is no significant stabilisation effect of either treatment on GUS, although bortezomib appears to cause marginally more accumulation. This slight increased accumulation in *prt6* lines treated with bortezomib may represent the stabilsation of components of unidentified auxillary pathways, which continue to degrade stabilised proteins such as the GUS reporter protein in the absence of a functioning N-end rule. In support of what was observed and discussed with *Ubi1::MCGGAIL-GUS* seedlings (Figure 3.12) immersed in DMSO, the control treatment for bortezomib, appears to have an additive stabilisation effect on GUS compared to normoxia.

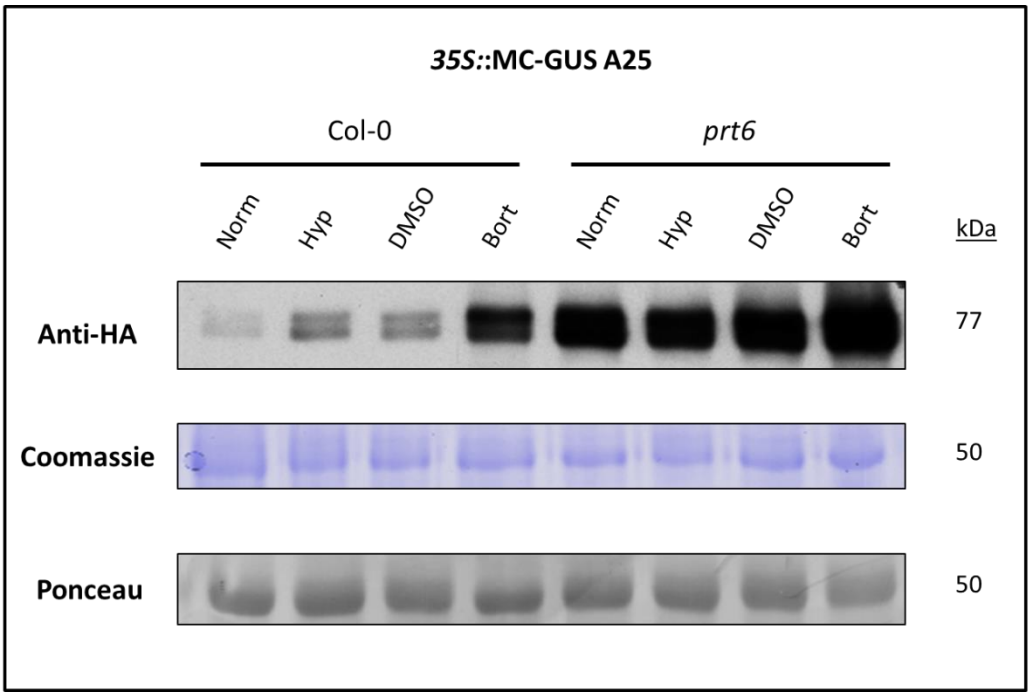


Figure 3.19 Anti-HA antibody Western blot showing the abundance of the GUS reporter protein extracted from 10d old *35S::MC-GUS* line A25 transgenic Arabidopsis seedlings, in the mutant *prt6* or WT Col-0 background when either untreated, exposed to hypoxia, incubated in dilute DMSO or in bortezomib dilutes in DMSO (26S proteasome inhibitor). Increased accumulation of GUS was observed after hypoxia and bortezomib treatment compared to the normoxic and DMSO control in the Col-0 background. Neither of the treatments affected the stabilisation of GUS in *prt6* which remained the same throughout. Coomassie loading control and Ponceau transfer controls (protein present on membrane used for Western analysis) for each sample also visible.

4. Identification of GUS from affinity purified seedlings by LC-MS/MS reveals various possible N-terminal cysteine oxidation states

4.1 Introduction

Previous work with mammalian cells used mass spectrometry (MS) to show that synthetic peptides bearing Nt Cys-sulfinic or Cys-sulfonic acid are arginylated when incubated with Arg tRNA transferase (Hu et al., 2005). A later study in plants discovered both PCOs and their Cys-sulfinic acid synthetase activity, and showed their ability to mediate the formation of Cys-sulfinic acid from an Nt-Cys peptide in the presence of oxygen, but without the need for NO (Weits et al., 2014). Recently a study using synthetic peptides and PCO enzymes revealed that Nt-Cys residues are oxidised to Cys-sulfinic acid, and both Cys-sulfinic and Cys-sulfonic acid are targets for arginylation by ATE. The issue with these experiments is that both were conducted *in vitro* and neither conclusively outlines the mechanism of Cys oxidation *in vivo* (White et al., 2016).

Affinity purification (AP) is a common technique used to enrich immuno-tagged proteins from a crude protein extract. The detection of ubiquitinated Arabidopsis proteins has been achieved by AP, by incubating crude protein extract with Agarose beads carrying the Ub binding domain and then releasing the bound protein and performing an anti-ubiquitin antibody Western blot (Manzano et al., 2008). An additional PTM common in plant proteins is sumoylation (small ubiquitin-like modifiers), sumoylation (SUMO) has been recently found to occur to

SnRK1 a component of many important developmental and other networks, this sumoylation is essential for its recognition by the E3 ligase SIZ1 (Crozet et al., 2016). In section 4.2 anti-HA antibody resin beads were used to affinity purify the HA-tagged GUS protein extracted from *35S::MC-GUS* Arabidopsis seedlings. The purification of GUS after AP was analysed with anti-HA antibody, then anti-Ubiquitin antibody to identify ubiquitinated GUS protein and lastly anti-Sumo antibody to investigate if sumoylated GUS protein is detectable.

The detection of GUS protein from the *35S::MC-GUS* transgene and the search for its Nt amino acid are presented in section 4.3. This work involved the bulk extraction of Arabidopsis seedlings and comparative visualisation of GUS protein on Western blot and coomasie gel. Numerous endoproteinases are available to digest protein for identification of their constituent peptides by liquid chromatography mass spectrometry (LC-MS). Three proteases were chosen (Trypsin, Asp-N and Lys-C) for digestion, each of which can generate differing sized peptides dependent on their specific site of cleavage. Digested samples were analysed by LC-MS followed by tandem mass spectrometry (MS/MS) to fragment peptides in order to identify individual amino acid residues. Data generated were searched both manually and using MASCOT™ against online databases to identify peptides belonging to the GUS reporter protein.

In section 4.4 affinity purified bands containing the GUS reporter protein from *35S::MC-GUS* seedlings are sent for Edman degradation, in an attempt to identify the oxidation state of Nt-Cys. The Nt peptides

expected to be produced after digestion with proteases were investigated more rigorously using MS/MS, and the peptide proceeding the Nt Asp-N peptide from the GUS reporter protein is identified. A synthetic peptide replicating the 49 residue Asp-N peptide, of GUS derived from the *35S::MC-GUS* transgene, was used to investigate the charge states it ionises at and its fragmentation pattern, which will help when trying to identify the same peptide from *in vivo* material. In addition to this the predicted mass of the alkylated Nt peptide of the GUS reporter protein produced after Asp-N digestion is presented, this includes the possible oxidation and nitrosylation modifications which could occur.

4.2 Affinity purification of GUS *in vivo* derived from the *35S::MC-GUS* transgene to identify PTMs involved in N-end rule processing

4.2.1 Anti-HA antibody detection of affinity purified GUS protein increases following treatment with bortezomib, a 26S proteasome inhibitor

Ten day old line A25 Arabidopsis seedlings in Col-0 containing the *35S::MC-GUS* transgene either immersed in DMSO or bortezomib (5 hours) were used for protein extractions. Protein was also extracted from line A25 in the *prt6* background. The crude protein from each respective treatment was affinity purified using anti-haemagglutnin antibody (anti-HA) resin beads (See section 2.5.2). Purified protein was released from the beads and used to perform SDS-PAGE and then an anti-HA antibody Western blot. GUS protein extracted from *35S::MC-GUS* seedlings in the Col-0 background, then detected by anti-HA antibody Western had increased after AP when released from the HA-beads (R), compared to the crude input protein (I) when untreated (Figure 4.1). The absence of any bands is expected in the 'not bound' (NB) crude protein which did not bind to the HA beads (NB), or what was washed off the HA beads before boiling (W5). As previously shown in figure 3.19 more GUS protein was stabilised after bortezomib treatment than in the control DMSO, and this reflected in both the crude input and the protein released from the HA beads after AP.

In the mutant *prt6* background substantially more GUS protein was stabilised in the control DMSO treatment, compared to after bortezomib

treatment in the Col-0 background. This result correlates well with previous studies, which demonstrated that all of the Arabidopsis ERFVIIs were stabilised in *in vitro* studies similarly when treated with another 26S proteasome inhibitor MG132 (Gibbs et al., 2011). Interestingly GUS protein was visible in the NB samples when in the *prt6* background, indicating that the HA-beads have been saturated due to the strong expression of *35S::MC-GUS* line A25 and the lack of PRT6 *in vivo*.

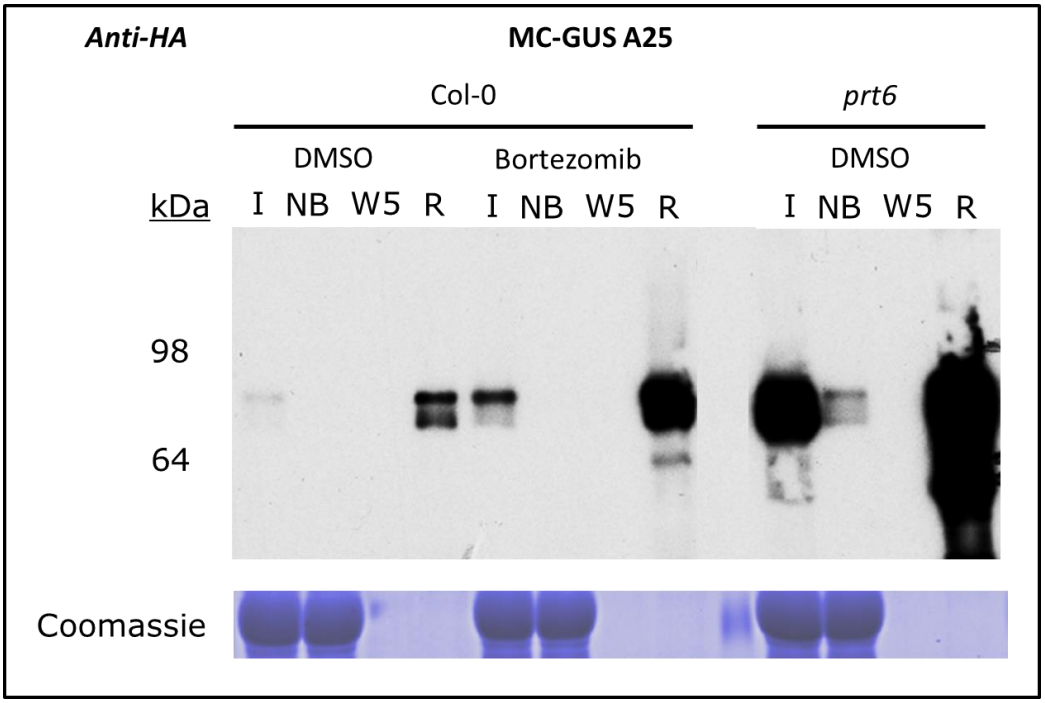


Figure 4.1 Anti-HA antibody Western blot showing the abundance of the ~77 kDa GUS protein extracted from 10d *35S::MC-GUS* line A25 Arabidopsis seedlings. The *35S::MC-GUS* seedlings were either in the mutant *prt6* or WT Col-0 background and immersed in DMSO or the 26S inhibitor bortezomib before affinity purification (AP) with anti-HA antibody beads. Increased stabilisation of GUS was observed after bortezomib treatment compared to the DMSO control in the Col-0 background crude input (I) and in samples containing the released GUS protein captured on the anti-HA beads (R). These results show that the AP worked with an accumulation of GUS released from the beads under either treatment. In the *prt6* background GUS was accumulated to an even higher extent than when in Col-0 after bortezomib treatment. Presence of GUS in the not bound (NB) samples indicates that GUS saturated the binding capacity of the beads. W5 = The 5th HA-bead wash after incubation of crude protein and HA-beads. Coomassie loading control for each sample also visible.

4.2.2 Ubiquitinated GUS protein increases after inhibition of the 26S proteasome using bortezomib

Substrates of the Arg/N-end rule pathway including the GUS reporter protein used in this study require ubiquitination for their degradation. Following the same procedure as section 4.2.1, but using anti-ubiquitin antibody after affinity purification, should allow detection of purified ubiquitinated GUS protein. Ubiquitination is a common PTM for many plant proteins, therefore it was not surprising to find ubiquitinated proteins in the crude input (I) or 'not bound' (NB) samples from *35S::MC-GUS* seedlings in Col-0 after DMSO immersion (Figure 4.2). However no ubiquitination (disregarding the anti-HA bead heavy antibody artefact ~50 kDa), was detectable on the GUS protein after affinity purification, indicating that this protein is unstable and in a state of constant degradation expected when the N-end rule is active. Ubiquitinated protein increased in general from the crude input and not bound samples of MC-GUS after bortezomib treatment, highlighting how the 26S proteasome facilitates the degradation of many Arabidopsis proteins.

Increased affinity purified GUS protein after bortezomib treatment is reflected in the amount binding to the anti-ubiquitin antibody, a 'Ub ladder' visible between 98 – ~130 kDa likely reflects the covalent attachment of the GUS protein (77 kDa) with Ub subunits. Poly-ubiquitin ladders have previously been identified using human cell lines treated with MG132 (Strachan et al., 2013) and in plants (Callis and Bedinger, 1994; Stary et al., 2003). The increased stabilisation of the GUS protein

after bortezomib treatment but not after DMSO indicates that the GUS protein normally has a short half-life. Unsurprisingly in the absence of the E3 ligase PRT6, a specific component of the Ub-cascading enzymes, identification of ubiquitinated proteins is greatly reduced in crude proteins (I and NB). Ubiquitinated GUS protein resembling that observed from *35S::MC-GUS* seedlings in Col-0 was not observed after affinity purification. It is important to refer to both figure 4.1 and 4.2 in tandem, as the results generally correlate for instance high quantities of GUS protein detected by anti-HA antibody in *prt6* with low levels of ubiquitinated GUS detected by anti-ubiquitin antibody in the same *prt6* background

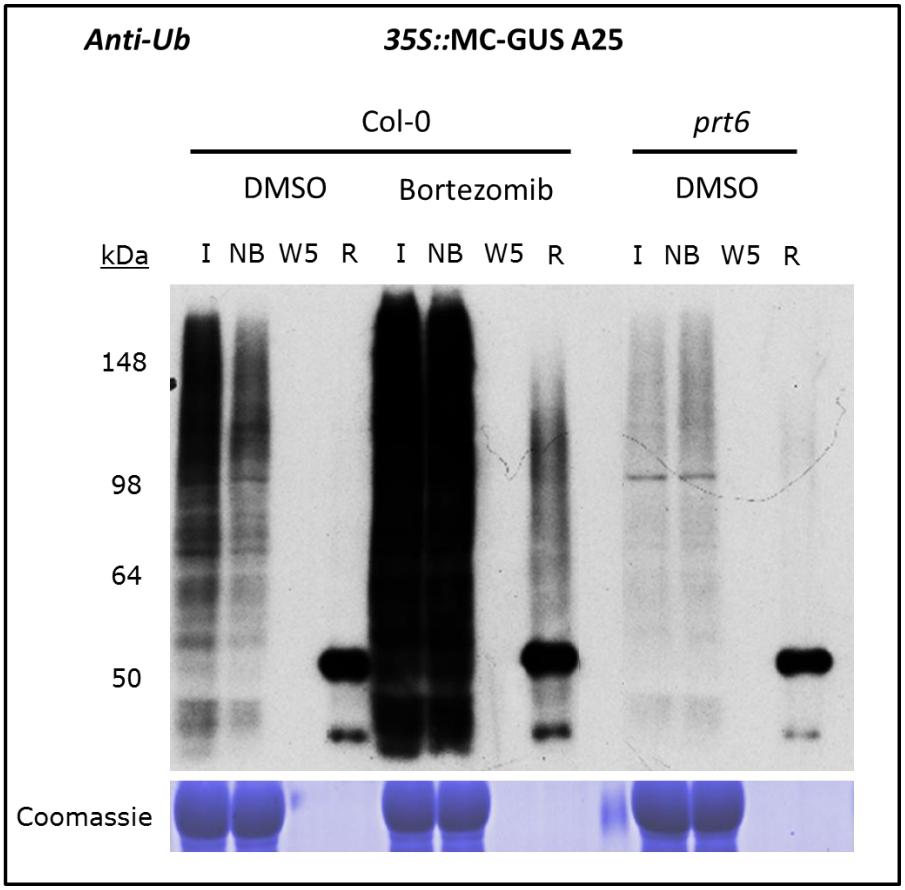


Figure 4.2 Anti-ubiquitin antibody Western blot showing the abundance of ubiquitinated protein extracted from 10d *35S::MC-GUS* line A25 Arabidopsis seedlings. The *35S::MC-GUS* seedlings were present in the mutant *prt6* or WT Col-0 background when immersed in DMSO or the 26S inhibitor bortezomib before affinity purification (AP) with anti-HA beads. Increased stabilisation of ubiquitinated protein was observed after bortezomib treatment compared to the DMSO control in the Col-0 background samples containing the released GUS protein captured on the anti-HA antibody beads (R). The increased intensity of the smeared background around 98 – 130 Kda in the R lane of bortezomib treated *35S::MC-GUS* seedlings indicates the potential GUS (~77 kDa) ubiquitin ladder. Increased ubiquitinated crude protein (I and NB) was also observed in bortezomib treated samples compared to DMSO. In *prt6* mutants relatively little ubiquitination occurred in the crude protein and no ubiqutinated GUS was observed when released from the beads. W5 = 5th HA-bead wash after incubation of crude extract and HA-beads. Note HA-antibody artefact present after boiling the HA-resin beads present at ~52 kDa in all R lanes. Coomassie loading control for each sample also visible.

4.2.3 Anti-sumoylation antibody detection of affinity purified sumoylated GUS protein unchanged after stabilisation of the 26S proteasome using bortezomib

Sumoylation is a PTM commonly occurring to many proteins, it has been noted as an important regulator of protein stability during periods of plant stress (Crozet et al., 2016; Knobbe et al., 2015). SUMO (small ubiquitin-like modifier) is a 12 kDa protein required for growth, which is involved in many biological processes. Upon periods of stress SUMO has been shown to accumulate across eukaryotes, in what is deemed a protective response (Guo and Henley, 2014). To detect whether GUS from *35S::MC-GUS* seedlings is a target for SUMO it was incubated with anti-sumoylation antibody (anti-SUMO) following similar procedures to section 4.2.1. No evidence of sumoylation was observed in the crude input (I) or 'not bound' (NB) samples after incubation with anti-HA antibody, indicating that sumoylation is not as prevalent as other PTMs like ubiquitination (Figure 4.3). However recently an analysis study on SUMO proteomics with human proteins has predicted that sumoylation likely occurs to thousands of proteins making it a frequent PTM (Hendriks and Vertegaal, 2016).

Sumoylated protein bands were detected from affinity purified protein from *35S::MC-GUS* seedlings from all treatments. A faint group of bands in the mass region of the expected GUS protein (77 kDa) produced by the transgene *35S::MC-GUS* line A25 were visible indicating that sumoylation may have occurred. However, no noticeable difference in the intensity of the possible GUS band was observed with either DMSO or bortezomib treatment in the Col-0 background. This reveals that unlike ubiquitination, sumoylation is not affected by the bortezomib induced

inhibition of the 26S proteasome. Previous studies have shown that there is close crosstalk between Ub and SUMO with them often competing for the same free Lys binding sites on proteins (Desterro et al., 1998). However, their effect once bound can be contrasting, such as when bound to mammalian Huntingtin (Htt) where SUMO increases neurodegeneration and Ub prevents it (Steffan et al., 2004). In the *prt6* ecotype the intensity of the possible GUS band after DMSO immersion of *35S::MC-GUS* seedlings is more intense than when in Col-0. Indicating that the absence of the E3 ligase, preventing ubiquitination, has reduced competition for SUMO attachment to the Lys amine group of proteins.

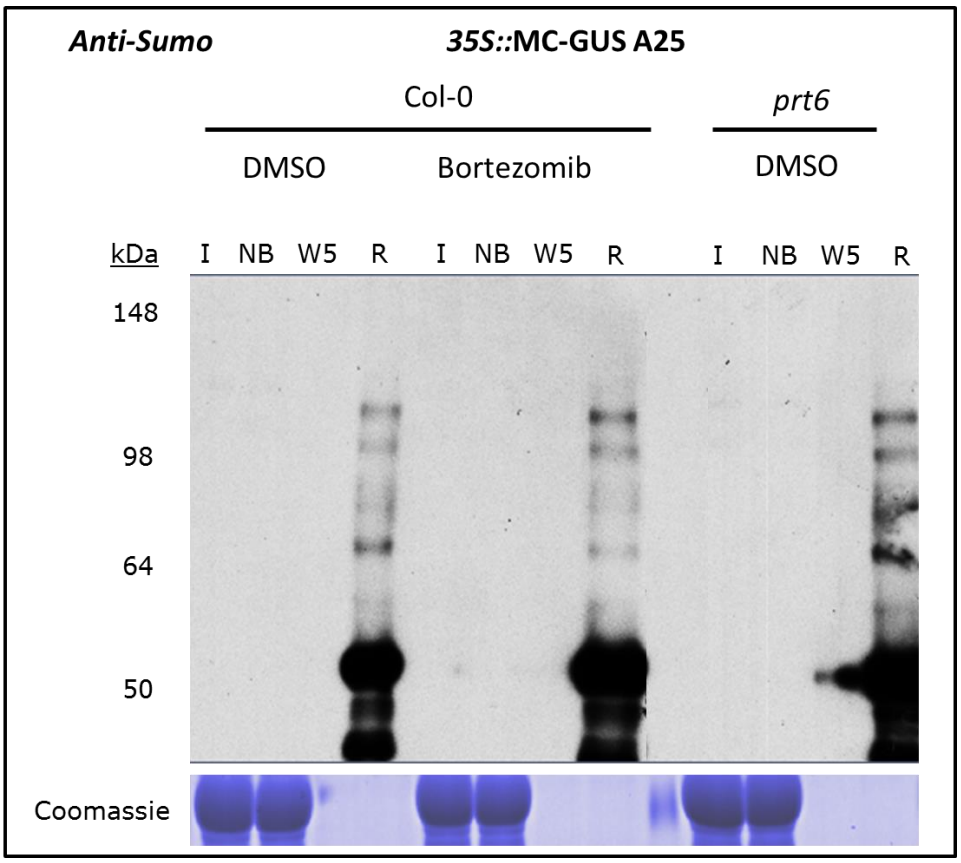


Figure 4.3 Anti-Sumoylation antibody Western blot showing sumoylated protein extracted and affinity purfied *35S::MC-GUS* line A25 10 days old transgenic Arabidopsis seedlings. *35S::MC-GUS* seedlings in the mutant *prt6* or wild-type (WT) Col-0 background were immersed in DMSO or the 26S proteasome inhibitor bortezomib before affinity purification (AP) with anti-HA antibody beads. No sumoylated protein was detected in the crude input (I), protein not bound to anti-HA beads after incubation with crude extract (NB), 5th wash HA-bead wash after incubation crude extract. Sumoylation appeared to occur to affinity purified protein extracted from *35S::MC-GUS* with numerous bands present from protein released form the HA- beads (R). These bands could be sumolyated GUS, sumoylated GUS degradation products or sumoylated foreign contaminants. This was similar in DMSO or bortezomib treated MC-GUS seedlings in Col-0 and when in the *prt6* background. Note HA-antibody artefact present after boiling the HA-resin beads present at ~52 kDa in all R lanes. Coomassie loading control for each sample also visible.

4.3 LC-MS identification of GUS reporter protein from affinity purified *35S::MC-GUS* Arabidopsis seedlings

4.3.1 Matching profiles of GUS protein bands on a coomassie gel and Western film

In order to visualise the GUS reporter protein from *35S::MC-GUS* seedlings on a coomassie gel to allow excision for LC-MS analysis, more protein needs to be loaded than for Western analysis. Protein extracted (20 µg/ul) from *35S::MC-GUS* Arabidopsis *prt6* seedlings was used for affinity purification, this would be stabilised *in vivo*, therefore the amount of GUS reporter protein purified and extracted would be maximised. In a control experiment protein from WT Col-0 seedlings was purified and extracted and used for anti-HA antibody Western blot, as expected no protein was visible unlike GUS protein extracted from *35S::MC-GUS* seedlings (Figure 4.4). Seedlings of *35S::MC-GUS* line A2/1 was used to maximise the possibility of purifying HA-tagged GUS, as this line has previously been shown to have the highest expression of the GUS reporter protein (Figure 3.14). To visualise the GUS protein on a coomassie gel much higher volumes of protein were loaded (50 µl compared to 5 µl), to overcome the insensitivity of this method compared to the immuno-targeting Western blot. Three individual bands appear to comprise the GUS reporter protein, visible on a coomassie at around 77 kDa (Figure 4.5), similar to the profile on the Western blot film (Figure 4.4). There are two lower molecular weight (MW) bands which sit close together then a ~4 kDa distance to the higher MW band.

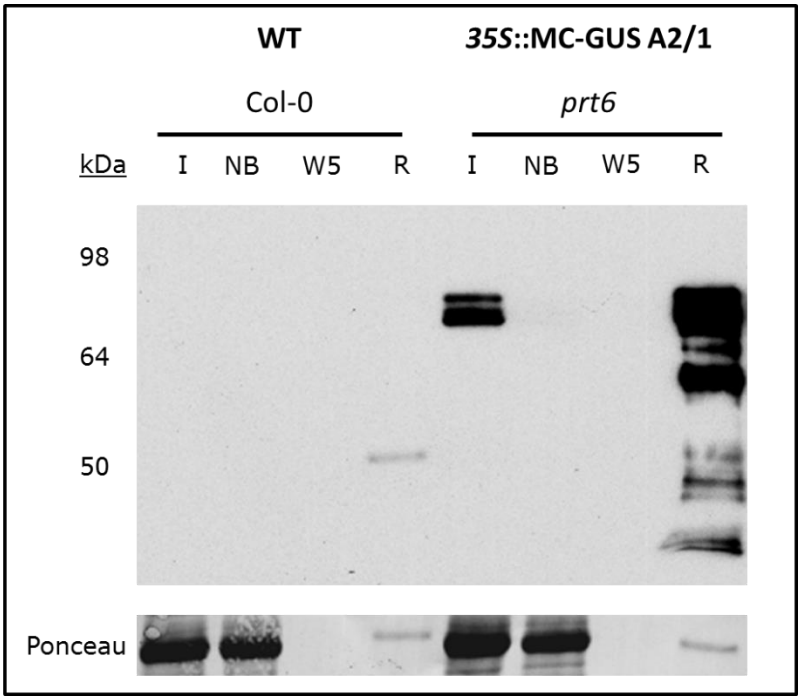


Figure 4.4 Anti-HA antibody Western blot showing the abundance of the GUS protein extracted from 10 days old *35S::MC-GUS* line A2/1 Arabidopsis seedlings in the mutant *prt6* background and wild-type Col-0 when affinity purified (AP) with anti-HA antibody beads. The AP worked an accumulated concentration of GUS was released from the beads (R) than present in the crude input (I). The GUS appears as a triplet band. The anti-HA heavy antibody chain from the anti-HA beads was the only thing detected from WT samples. Note HA-antibody artefact present after boiling the HA-resin beads present at ~52 kDa in both R lanes. Ponceau transfer control visible for each sample.

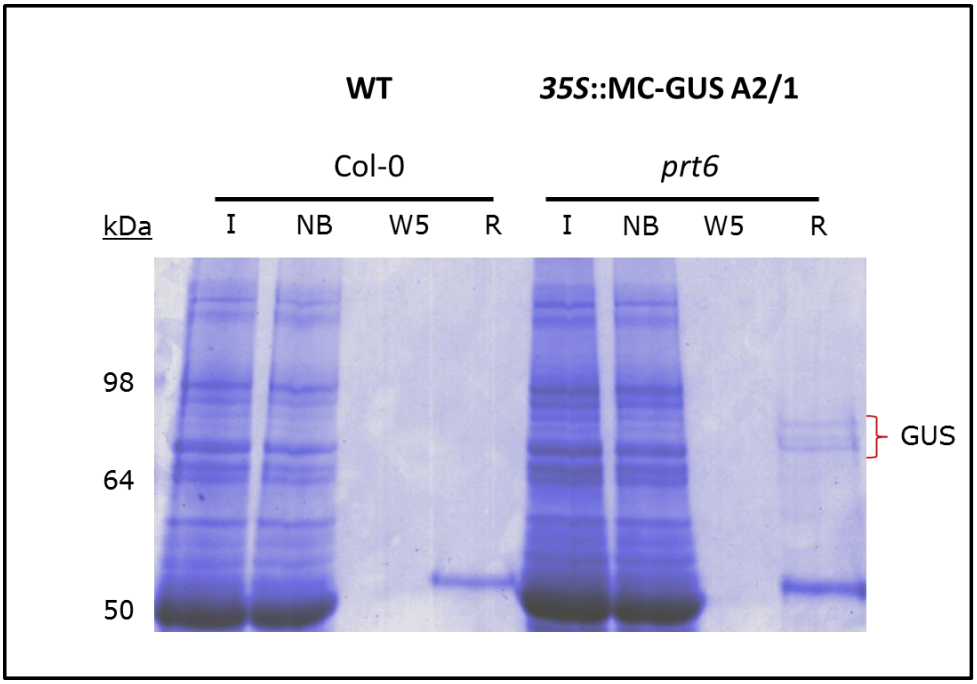


Figure 4.5 Coomassie gel showing protein extracted from 10d *35S::MC-GUS* line A2/1 Arabidopsis seedlings in the mutant *prt6* background and wild-type Col-0 when affinity purified (AP) with anti-HA beads. The coomassie detects total protein, hence the band profile is the same for the crude protein input (I) and protein not bound (NB) to anti-HA antibody beads for both *35S::MC-GUS* and WT seedlings. After incubation of beads and crude protein the beads were washed five times to remove unspecific bound protein and this was loaded (W5) and the protein released from the anti-HA beads (R). Three bands are visible around 77 kDa in lane R for *prt6* seedlings, matching the anticipated mass of the GUS reporter protein. The anti-HA heavy antibody chain (~52 kDa) from the anti-HA beads was the only thing released from the beads from WT samples.

4.3.2 Bulked extractions of GUS protein derived from 35S::MC-GUS for gel extractions

The three bands excised from the gel in figure 4.5 all contained peptides from the MC -GUS derived from *35S::MC-GUS* seedlings, although coverage was poor (~15% of protein, Appendix D). Interestingly no unique peptides which might suggest a PTM causing the mass difference between the three bands were identified. In order to maximise the

peptide recovery and subsequent coverage of the GUS reporter protein after excision and digestion of the purified gel bands, larger scale experiments were established. Larger 1.5 mm gel electrophoresis plates replaced the 1.0 mm plates originally used, these could hold increased volumes of crude protein extract. In addition to this the quantity of seedlings used in extractions was also increased, maximising the intensity of the GUS reporter protein visualised by coomassie which could be extracted for LC-MS analysis. Figure 4.6 shows the optimised coomassie stained gels where only crude input (I) and released (R) protein from *35S::MC-GUS* seedlings were loaded.

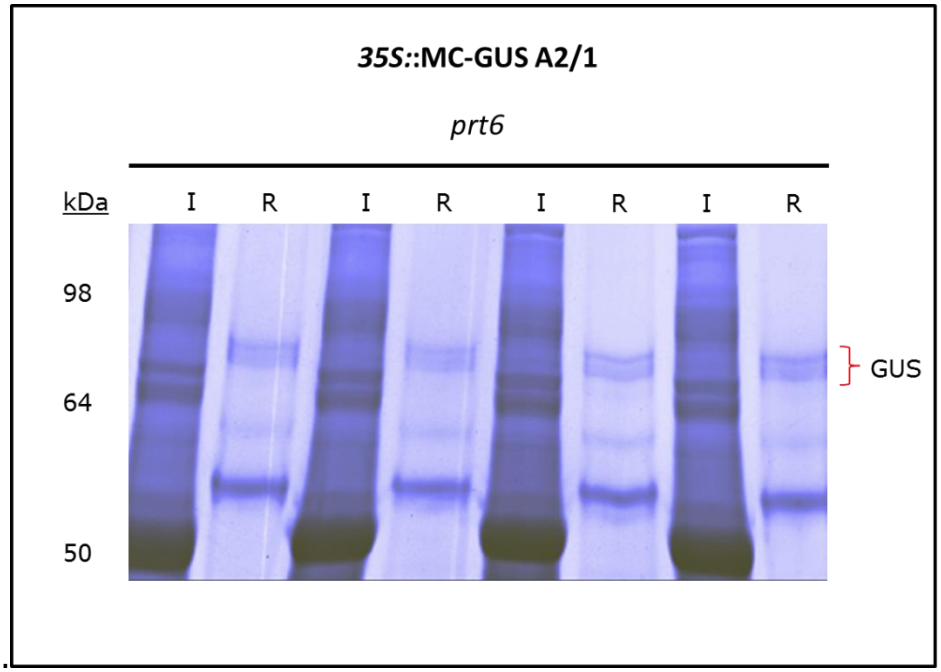


Figure 4.6 Bulked extraction and purification of protein from *35S::MC-GUS* line A2/1 10 day seedlings. Triplet bands previously identified as containing GUS are present in each of the lanes were protein was released from HA-beads (77 kDa, lane R). I = crude protein input; R= Released MC-GUS protein after affinity purification. Note HA antibody artefact present after boiling the HA-resin beads at ~52 kDa in both R lanes.

4.3.3 LC-MS/MS detection of GUS reporter protein and varied digestive performance

Endoproteinases cleave proteins producing smaller polypeptides which can be identified by mass spectrometry (MS). The sequence of the MC-GUS construct contains three recognisable components, the poly-Gly linker near the N-terminus, a 3x HA tag for pulldowns and a β-glucuronidase (GUS) protein for *in vivo* localisation (Figure 4.7). Due to the predominance of the GUS protein in the MC-GUS reporter the majority of the peptides generated through proteolysis belong to GUS protein.

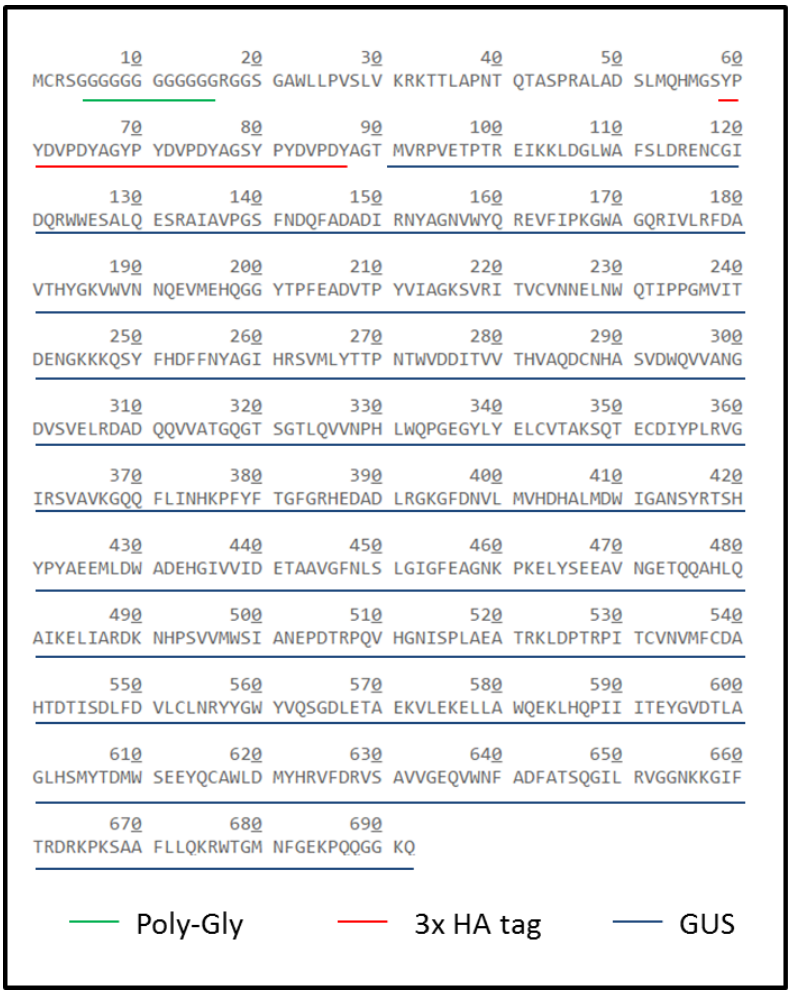


Figure 4.7 MC-polyG-HA-GUS (MC-GUS) protein sequence including annotated poly-Gly linker, 3x HA tag and β-glucuronidase (GUS).

Trypsin is the most common endoproteinase used for digesting proteins before MS analysis (Gupta et al., 2010). Trypsin cleaves after Arg (R) and Lys (K) at the C-terminal side of protein residues. Trypsin was used to digest HA-purified GUS reporter protein before formic extraction, when analysed by LC-MS/MS 27% of total residues belonging to the GUS protein were identified (Figure 4.8). The peptides detected were well distributed over the entire sequence of the GUS reporter protein, although there was a region between residues 220 and 365 where no peptides were identified. However, the Nt peptide, a small RCR tripeptide, or the proceeding N-terminal peptide containing the poly-Gly linker was not detected. The three residue Arg-Cys-Arg N-terminal peptide (expected after excision of Met and addition of Arg during N-end rule processing) is likely to elute out early because of its low hydrophobicity and charge, looking at the size of other tryptic peptides the smallest which were identified were only 7-8 residues. Due to the sequence of the GUS reporter, trypsin was not ideal to use when trying to find identify the oxidation state of Nt-Cys. However, the closest peptide to the N-terminus was identified using trypsin, located only 17 residues from the N-terminus.

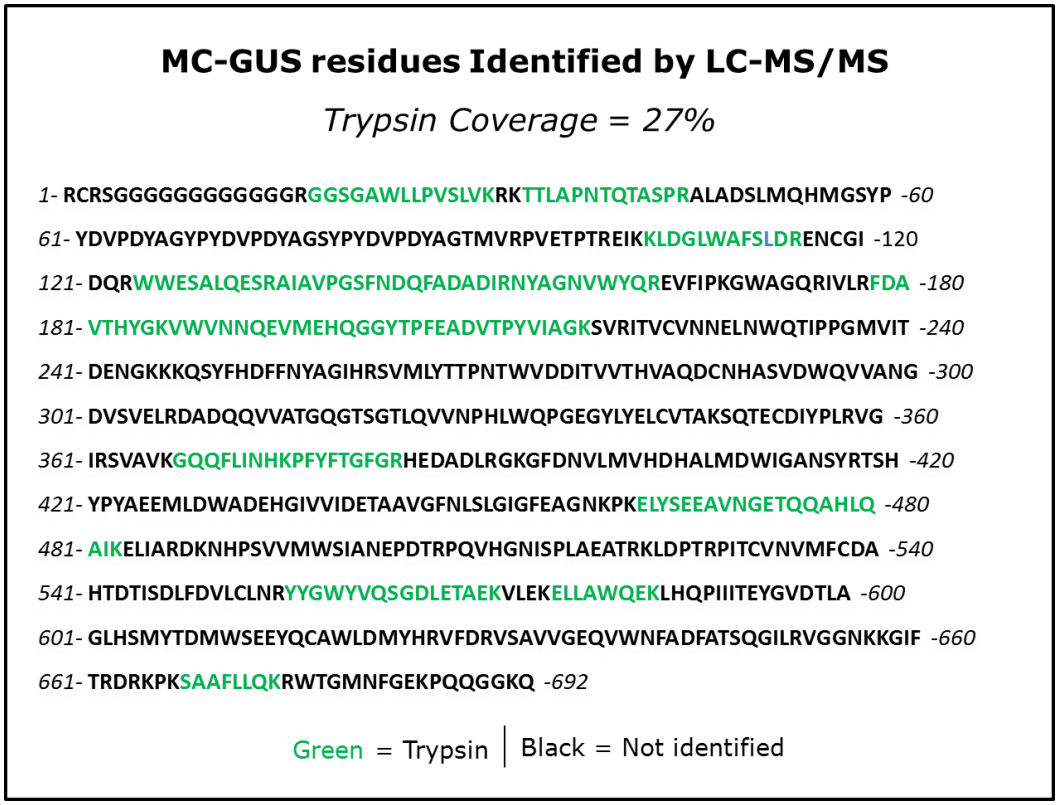


Figure 4.8 Residues identified from the MC-GUS reporter protein expressed in *35S::MC-GUS* Arabidopsis seedlings by LC-MS/MS after digestion with trypsin. Residues identified matching the MC-GUS sequence after trypsin digestion are green, unidentified residues are black.

When Asp-N, an endoproteinase which cleaves Nt to Asp (D), was used to digest the HA-purified GUS reporter gel bands more residues were identified in comparison to trypsin (Figure 4.9). In total 59% of all residues were identified to match the sequence of the MC-GUS protein and therefore Asp-N was the most efficient protease used in this study. Importantly the Nt peptide, a large 49 residue (4.6 kDa) peptide, was not detected therefore the oxidation state of Cys could not be identified. In contrast to the tryptic tripeptide, this 49 residue Asp-N derived polypeptide is large and if detected would likely be present as a high charge state ion. However the peptide adjacent to this Nt peptide was

identified (Section 4.4.2), indicating that the Nt peptide is present in some form in the digested gel protein extract.

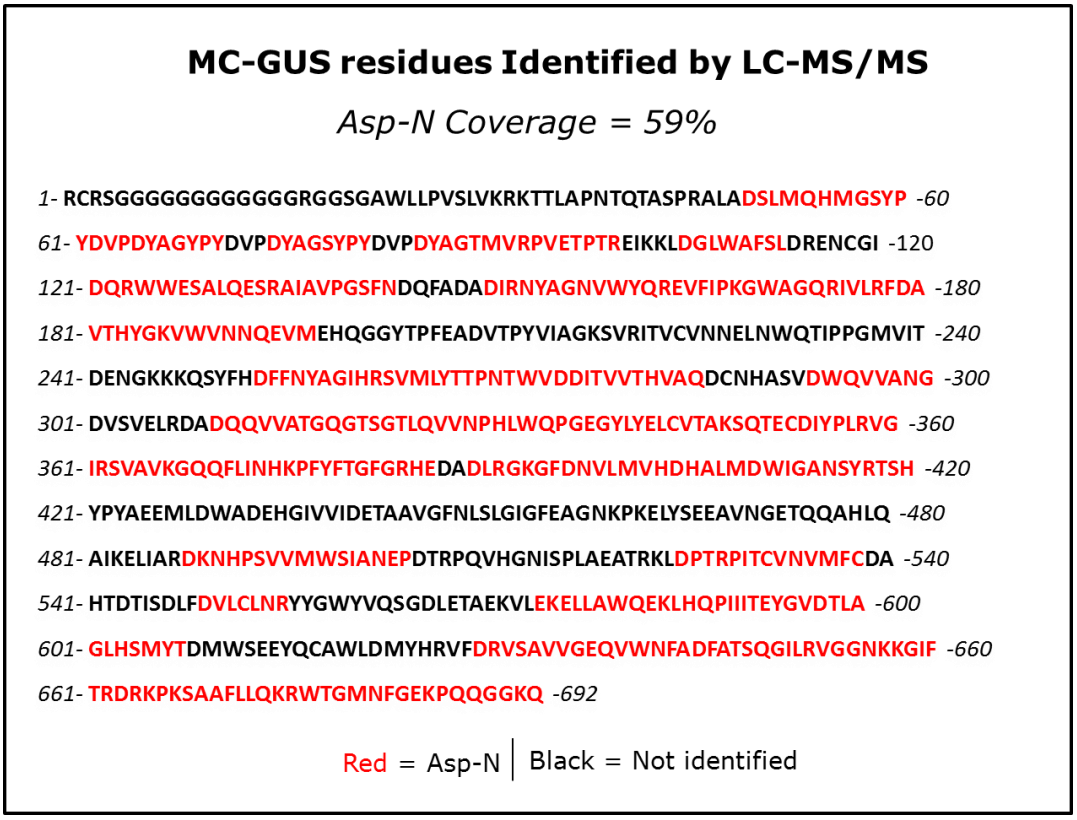


Figure 4.9 MC-GUS residues identified by LC-MS/MS after Asp-N digestion of the MC-GUS construct expressed by *35S::MC-GUS* Arabidopsis seedlings. Residues of MC-GUS detected after Asp-N digestion are highlighted in red, unidentified residues are in black.

The total combined proportion of MC-GUS residues identified observed after separate digestions with trypsin and Asp-N was 75% (Figure 4.10). Residues identified after digestion with both trypsin and Asp-N accounted for 17% of the MC-GUS sequence. A third protease, Lys-C, which cleaves exclusively at the C-terminus of Lys residues was also used multiple times to digest MC-GUS, however only one peptide was ever identified. This was unfortunate as efficient Lys-C digestion would create a 32 residue Nt peptide, which should be identifiable by LC-MS. The absence of this Nt peptide indicates that the N-terminus may be modified in an

unexpected manner. The fixed modifications for MC-GUS peptide used in database searches are discussed in section 2.8.2, however Cys oxidation, dioxidation and trioxidation were all variable modifications searched.

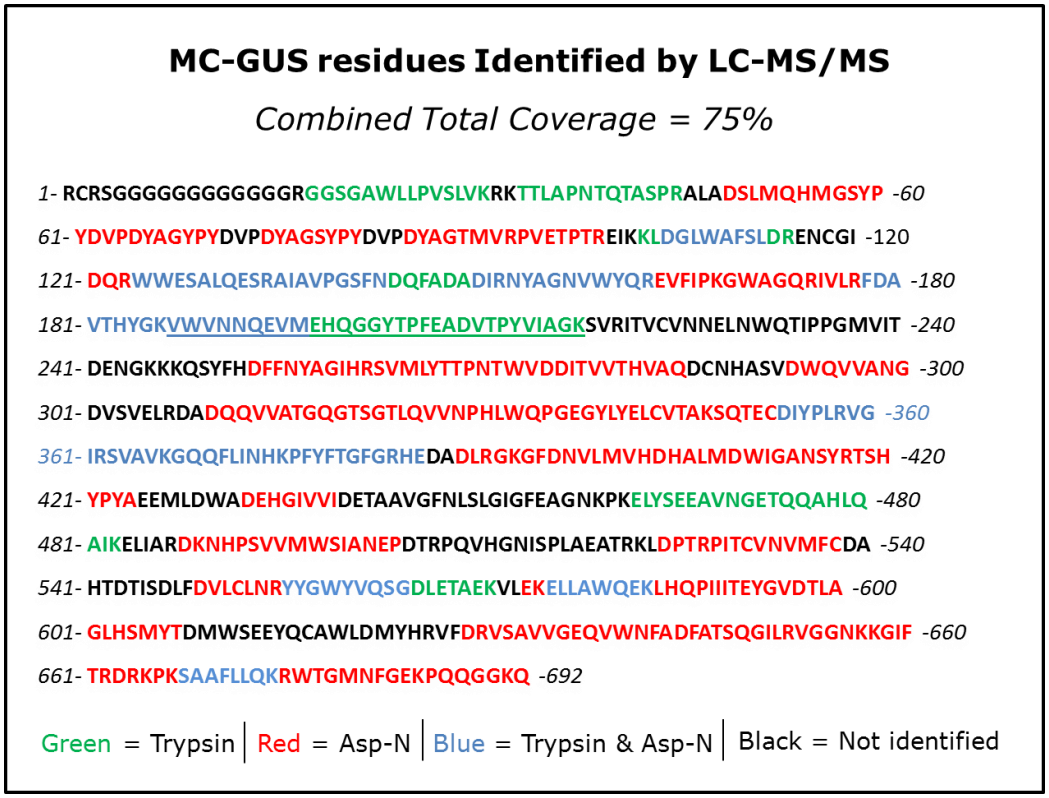


Figure 4.10 Total MC-GUS amino acid residues identified by LC-MS/MS after digestion with three proteases. MC-GUS residues sequenced specific to trypsin (green) or Asp-N (red) are highlighted. Residues commonly identified by both proteases are highlighted in blue. Unidentified sequences are in black. A third protease, Lys-C, which cuts C-terminal of Lys (K) was also used but this only returned one peptide (underlined).

4.4 Predicted N-terminal MC-GUS peptides and possible Cys-ox PTMs

4.4.1 Edman sequencing of GUS reporter protein

Digesting of the GUS reporter protein with various proteases before analysing by LC-MS/MS did not produce an identifiable Nt peptide. This could possibly have been caused by an unexpected cleavage which wouldn't be detected by the search parameters, additionally if a number of unexpected PTMs occurred to this peptide this would also change its predicted mass. A method which focusses on identifying Nt protein residues is Edman sequencing, where residues at the N-terminus are labelled and sequentially cleaved. Affinity purified GUS reporter protein gel bands from eight separate affinity purifications were excised, pooled and sent to Altabioscience to be subjected to Edman sequencing (Birmingham, UK).

The results returned showed that no signal was identifiable in any sample but Altabioscience suggested that the N-terminus may be 'blocked', by modifications which can block the Edman technique, which is not uncommon in Nt proteins (Figure 4.11). Typical PTMs which can obscure the Nt sequence identification of proteins include acetylation and formylation, which if occurring to the GUS protein may occur on the Nt-Arg residue (after Nt-Met excision *in vivo*). Previous studies looking at the oxidation of the mammalian Nt substrate RGS4 have similarly failed to identify the Cys residue at position 2 post alkylation, suggesting that there is some unknown modification occurring before this step (Hu et al.,

2005; Kwon et al., 2002). Hu et al. (2005) also reported that using the NO scavenger cPTIO prevented RGS4 arginylation occurring and this in turn had an effect which blocked the N-terminus, however the causative agent blocking the N-terminus was not discovered.

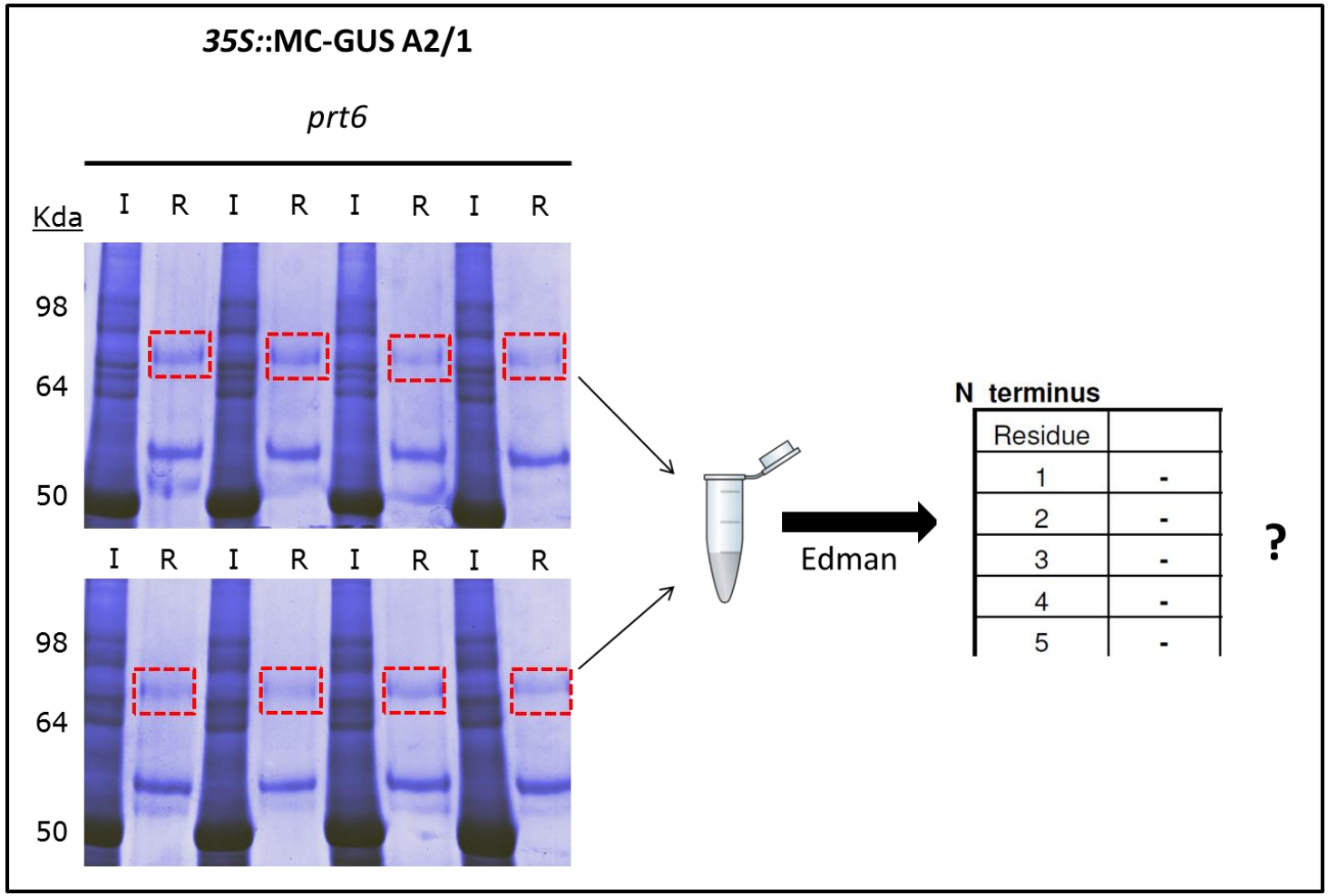


Figure 4.11 Schematic showing eight purified GUS band sections excised from coomassie gel, pooled in an Eppendorf and sent for Edman sequencing. Edman sequencing returned no distinguishable residues and suggested that on account of the high protein content sent the N-terminus of the protein was likely blocked.

4.4.2 Trypsin, Asp-N and Lys-C Nt peptide prediction of GUS protein derived from 35S::MC-GUS

The detection of peptides from the GUS reporter protein after affinity purification highlights how the technique was efficient in maximising target capture. However the elusivity of the Nt peptide from three different proteases required closer analysis. The cleavage sites of trypsin, Lys-C and Asp-N are highlighted in figure 4.12. Following what is currently known about N-end rule processing, the purified *in vivo* HA-bound protein would be expected to remain arginylated in *prt6* mutant seedlings. Hence, the N-terminus of what formerly was MC-GUS would now be RC_{ox}-GUS. The addition of an Arg residue at the N-terminus potentially adds another tryptic cleavage site (Figure 4.12), however when Cys is proceeding Arg cleavage this site is not cleaved hence the tryptic Nt peptide expected to remain should be RCR. The large 49 residue Asp-N may be difficult to detect due to its size, with the additional possible issue of potential PTMs occurring at one of the 48 residues other than Cys. If Nt-Cys is oxidised to Cys-sulfinic acid this has also been reported to act as a site for Asp-N cleavage due to its previously mentioned similarity to Asp (see section 1.3.2). A Lys-C N-terminal RC_{ox}-GUS peptide appears the most amenable size (31 residues) but unfortunately in my experiments with Lys-C did not digest MC-GUS efficiently (only generating one identifiable peptide, Figure 4.10).

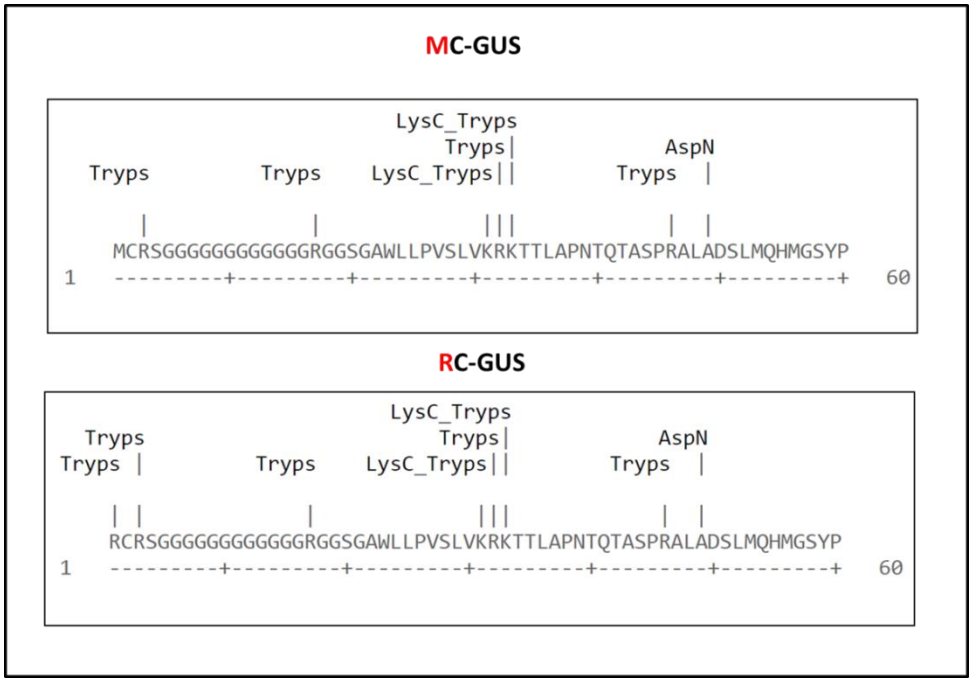


Figure 4.12 Predicted trypsin, Lys-C and Asp-N cleavage sites in the first 60 N-terminal residues of MC-GUS and RC-GUS

The closest identified peptide to the N-terminus is located just after the poly-Gly linker and was cleaved by trypsin (Figure 4.13). However, due to the small size of the Nt tryptic tripeptide and the poor digestion efficiency of Lys-C, Asp-N was identified as the protease to pursue the investigation identifying the Nt Cys oxidation state. The identification of the peptide after the first Asp-N cleavage site means that the Nt peptide must be present in the MC-GUS Asp-N digest solution, although the oxidation state of Cys and any other PTMs are unknown. The collision induced dissociation (CID) MS/MS fingerprint data of peptide DSLMQHMGSYPY (adjacent to Nt-Asp-N peptide) is presented in Figure 4.14

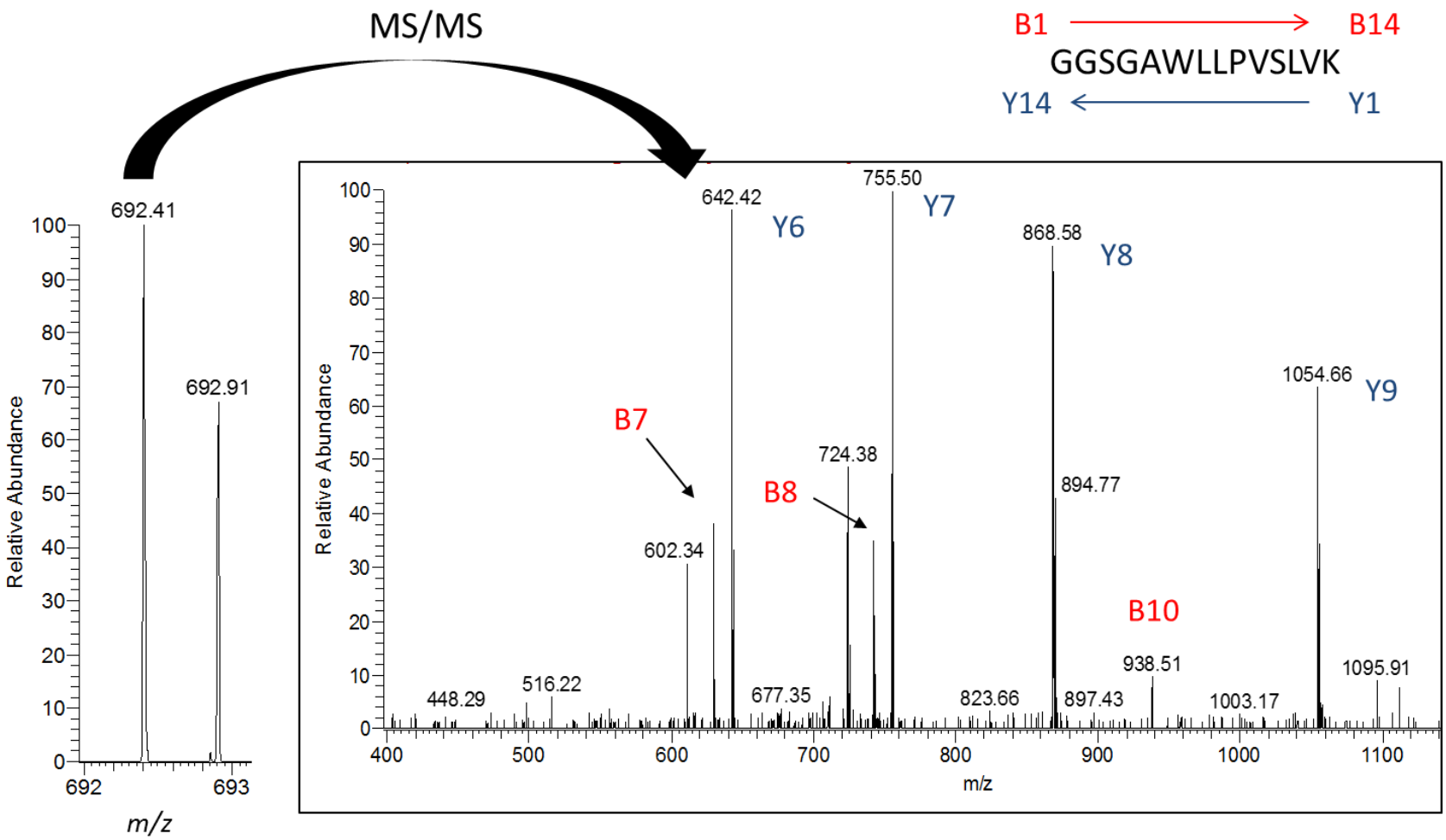


Figure 4.13 Fourier transform ion cyclotron resonance (FTICR) LC-MS and MS/MS spectra for the 14 residue tryptic peptide which was the closest peptide identified to the N-terminus of the GUS protein extracted from *35S::MC-GUS* Arabidopsis seedlings. Collision induced dissociation (CID) of a 2^+ *m/z* 692.4 mass peak revealed its identity as GGSAWLLPVSLVK through the detection of its constituent B and Y ions fragmenting from the N-terminus and C-terminus respectively of the peptide backbone.

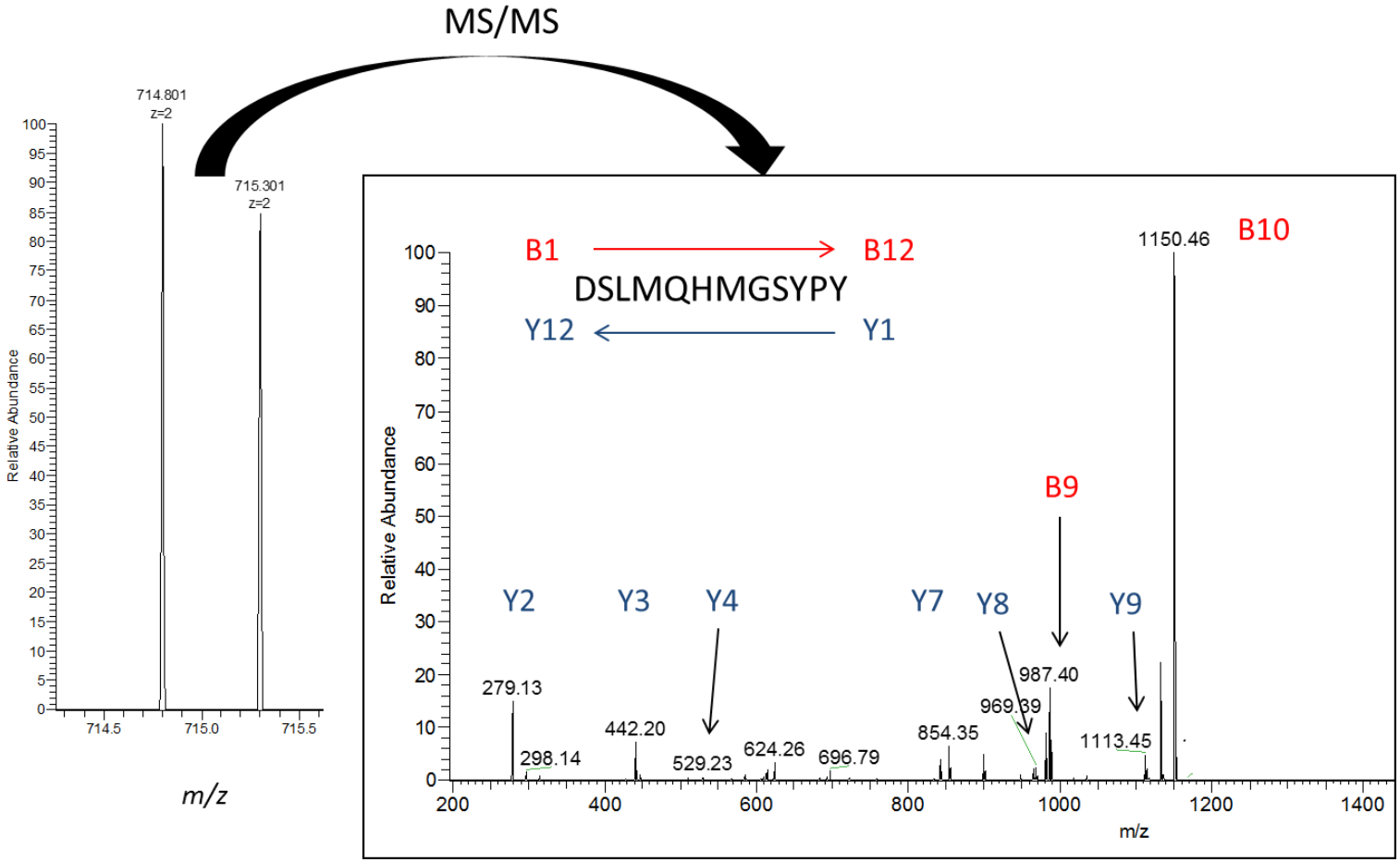


Figure 4.14 Fourier transform ion cyclotron resonance (FTICR) LC-MS and MS/MS spectra for the 12 residue Asp-N peptide from the second cleavage site of Asp-N and neighbouring the unidentified 49 residue Nt-peptide of the GUS protein extracted from *35S::MC-GUS* Arabidopsis seedlings. Collision induced dissociation (CID) of a 2^{+} *m/z* 714.8 mass peak revealed its identity as DSLMQHMGSYPY through the detection of its constituent B and Y ions fragmenting from the N-terminus and C-terminus respectively of the peptide backbone.

4.4.3 Using the RCR-49 synthetic peptide to mimic the predicted Nt Asp-N peptide of the in vivo GUS reporter protein

To maximise the chances of identifying the Nt peptide if the GUS reporter protein after digestion with Asp-N, a synthetic peptide replicate was purchased. The 49 residue Arg-Cys Nt Asp-N peptide (RCR-49) would provide data regarding the charge states typically observed for this polypeptide and hence what to look for in samples derived from the *in vivo* HA-purified GUS. Synthetic peptides replicating Nt peptides from proteins of interest have been used previously when investigating the oxidation and arginylation of Nt-Cys (Hu et al., 2005; Kwon et al., 2002; White et al., 2016). LC-MS data from the RCR-49 peptide unsurprisingly identifed that RCR-49 was detected at high charge states (4^+ to 7^+) because of the high MW of the peptide (4604 Da) (Figure 4.15). In addition to this the ability to perform MS/MS on the RCR-49 at similar energy conditions to what would be used with *in vivo* material, would give indications to how the Nt Asp-N peptide of MC-GUS would fragment. This CID fragmentation generated *b* and *y* ions which covered the whole RCR-49 residue sequence (Figure 4.16). The identification of a sequence of *y* ions would be useful, as these should remain the same regardless of any modification which could occur *in vivo* at the N-terminus. Therefore these are reference ions to look for as an indication that *in vivo* their parent species is the Nt Asp-N peptide, which could then be interrogated more rigorously to identify PTMs at Nt-Cys and elsewhere.

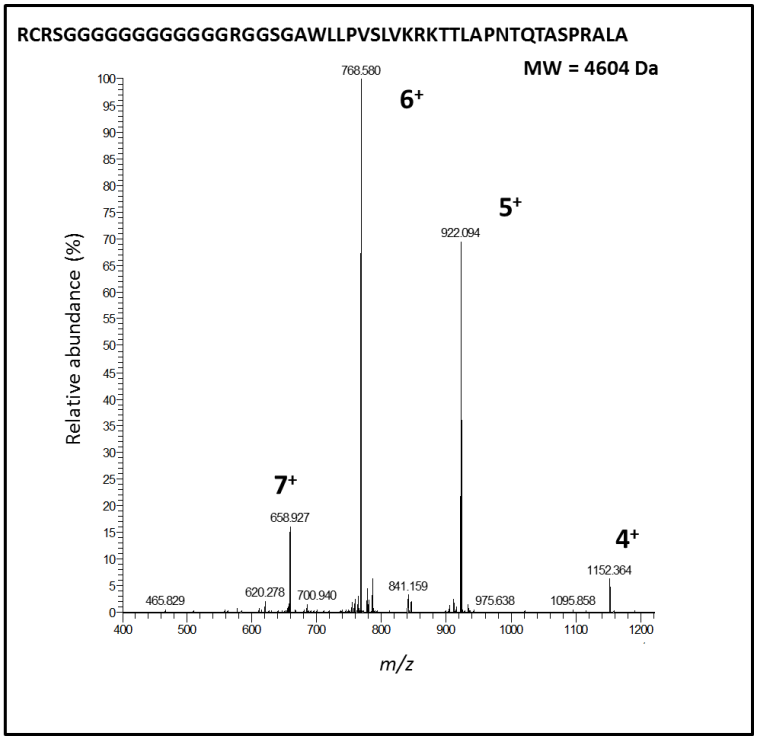


Figure 4.15 Fourier transform ion cyclotron resonance (FTICR) LC-MS spectra of a 1 µM solution of the synthetic peptide RCR-49 which replicates an arginylated Nt Asp-N peptide of MC-GUS. The 4^+, 5^+, 6^+ and 7^+ charge states of the 4604 Da protein were visible.

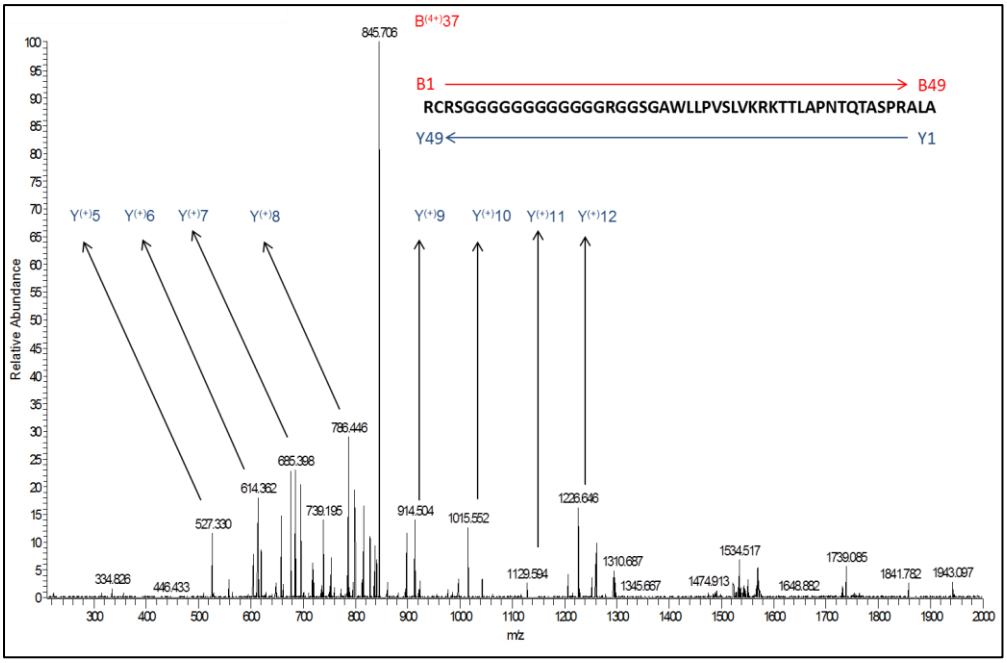


Figure 4.16 Fourier transform ion cyclotron resonance (FTICR) MS/MS spectra of the *m/z* 768 (6^+) ion from 1 µM solution of the synthetic peptide RCR-49. A series of *y* ions from the C-terminus of RCR-49 were identified but also a *b* ion 37 was identified. The presence of the *b* ions as the most abundant peak is not surprising as it represents a fragmentation at a labile Pro, but importantly it confirms that there have been no modifications to the Cys at the N-terminus. The *y* ions will be useful to look for from *in vivo* MC-GUS material, as regardless of what modifications are occurring at the N-terminus these should remain the same.

When searching through GUS MS/MS data generated after Asp-N digestion looking for the Nt peptide, results from this study have identifed a multitude of possible PTMs at Nt-Cys which need to be accounted for. Table 4.1 lists the possible modifications which could be occuring at the Nt-Cys. The excised GUS bands were all reduced and alkylated using iodocetamide (IAA) therefore the Nt-Cys will already contain a carbamidomethyl group, adding +57 Da to the original MW of any *in vivo* GUS material. To help determine whether alkylation does occur as exprected to the GUS protein bands excised, and whether any additonal PTMs occur during the Asp-N digestion, the RCR-49 peptide was added to a blank SDS gel piece after Coomassie staining. An

increase of +57 Da on the unmodified RCR-49 was the only modiifcation observed indicating the alkylation of Cys-2. MS/MS interrogation of the +57 Da species sharing the same charge states as RCR-49 matched those of a carbamidomethlyated Cys (Figure 4.17)

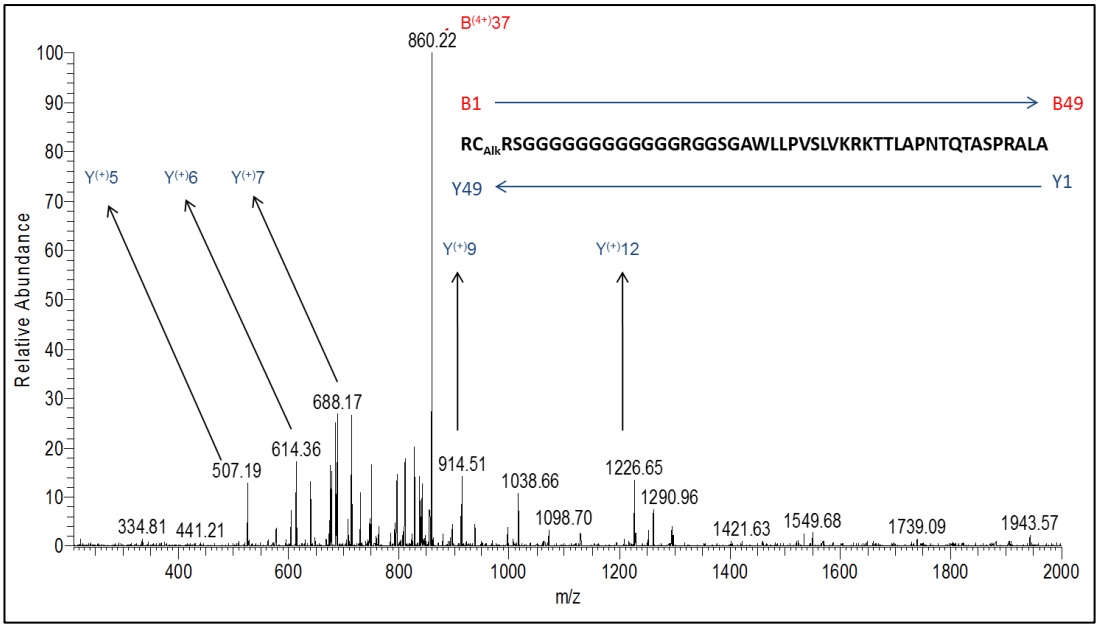


Figure 4.17 Fourier transform ion cyclotron resonance (FTICR) MS/MS spectra of a blank gel band with 2.5 µM solution of the synthetic peptide RCR-49 added before reduction, alkylation and AspN digestion. A series of *y* ions from the C-terminus of R_{Alk}CR-49 were identified but also a *b* ions from the N-terminus was identified. The presence of this *b* ions as the most abundant peak is not surprising as it represents a fragmentation at a labile Pro, but importantly it confirms that the Nt-Arg (R) of the RCR-49 has been alkylated (+57 Da). The *y* ions will be useful to look for from *in vivo* GUS material, as regardless of what modifications are occurring at the N-terminus these should remain the same.

Oxidation, dioxidation and trioxidation of Cys to Cys-sulfenic, Cys-sulfinic and Cys-sulfonic acid respectively is the recognised oxidative route of Cys (e.g. Hu et al., 2005). However, as demonstrated in the upcoming chapter another route of oxidation formed by the closing of the Cys-ring structure is less common but also possible, forming a Cys-sulfenamide then Cys-sulfinamide or Cys-sulfonamide respectively. In addition to this we know that NO is also involved in N-end rule processing (Gibbs et al.,

2014b), thus NO action could also lead to a possible PTM. The MW shift of these modifications is predicted for the common charge states of the Asp-N digest Nt peptide. In addition to this Table 4.2 presents the masses for the same modifications but if Nt-Arg was missing so that Cys was the true Nt residue. A Table of masses for an acetylated N-terminus were also searched (Appendix E)

Table 4.1 Predicted oxidative reactions occurring at the Nt-Cys and their resultant mass shift are predicted for the Nt Asp-N peptide of the GUS reporter protein for the charge states identified using the RCR-49 synthetic peptide

Fixed Cys Mods. = Alkylation (+57 Da)	RC-GUS Mass = 4604 Da	Charge state			
Ox./NO Modification at Cys	Mass shift (Da)	4^+	5^+	6^+	7^+
None	*+ 0*	1162.611	930.289	775.407	664.778
Oxidation	*+ 16*	1170.114	933.488	778.073	667.063
Dioxidation	*+ 32*	1170.609	936.687	780.739	669.348
Trioxidation	*+ 48*	1174.607	939.886	783.405	671.633
Cyclic Sulfenamide	*- 2*	1162.107	929.886	775.071	664.490
Cyclic Sulfinamide	*+ 14*	1166.106	933.085	777.737	666.775
Cyclic Sulfonamide	*+30*	1170.105	936.284	780.403	669.060
Nitrosylation	*+ 29*	1169.859	936.087	780.239	668.919

Table 4.2 Predicted oxidative reactions occurring at the Nt-Cys and their resultant mass shift are predicted for the Nt Asp-N peptide of GUS reporter protein if Nt arginylation does not occur for the charge states identified using the RCR-49 synthetic peptide

Fixed Cys Mods. = Alkylation (+57 Da)	C-GUS Mass = 4491 Da	Charge state			
Ox./NO Modification at Cys	Mass shift (Da)	4^+	5^+	6^+	7^+
None	*+ 0*	1134.611	907.889	756.741	648.778
Oxidation	*+ 16*	1138.610	911.088	759.407	651.063
Dioxidation	*+ 32*	1142.609	914.287	762.072	653.348
Trioxidation	*+ 48*	1146.607	917.486	764.738	655.633
Cyclic Sulfenamide	*- 2*	1134.107	907.486	756.405	648.490
Cyclic Sulfinamide	*+ 14*	1138.106	910.685	759.071	650.775
Cyclic Sulfonamide	*+30*	1142.105	913.884	761.737	653.060
Nitrosylation	*+ 29*	1141.859	913.687	761.572	652.919

5. Analysis of N-terminal cysteine oxidation and nitrosylation *in vitro* using synthetic peptides reveals both expected and unexpected oxidative routes

5.1 Introduction

Oxidation reactions on amino acid residues of proteins are common signalling cues for various developmental and physiological responses in plants including stress responses (Job et al., 2005; Johansson et al., 2004; Møller and Kristensen, 2004). The preferential target of these oxidative modifications to proteins is the Sulfur atom (S) present on both Cys and Met residues, although Tyr and Trp are other less common oxidative PTMs (Rinalducci et al., 2008). The oxidative PTMs occurring to the S atom of both Cys and Met residues and their importance has been reviewed extensively (e.g. Jacques et al., 2013; Waszczak et al., 2015). Identifying oxidation and nitrosylation reactions *in vivo* is difficult due to the volatility of ROS and RNS species but also their low abundance, with NO concentrations ranging from ~100 pM to 5 nM (Hall and Garthwaite, 2009). To monitor the reactivity of residues such as Nt-Cys it would be useful to control conditions, this cannot be achieved *in vivo*. Therefore, synthetic peptides with an identical sequence to the full length protein of interest are now commonly used. Synthetic peptides have reduced in price (per amino acid residue) and increased in length and purity over the past decade. In relation to investigations of the N-end rule pathway, synthetic peptides have been used to highlight that Arg tRNA transferase is active upon Cys-sulfonic acid in mammals (Hu et al., 2005; Kwon et

al., 2002) and sulfinic/sulfonic in plants (White et al., 2016). Nt-Cys synthetic peptides bearing the conserved plant CGGAI motif of the ERFVIIs, have been shown to oxidise upon incubation with the PCOs forming Cys-sulfinic acid (White et al., 2016).

The detection of small polar molecules, including di/tripeptides, by MS is difficult using conventional C4 and C18 LC columns. Detecting small peptides using reverse phase chromatography using a C18 column works on the principle of hydrophobic adsorption, where one part of an analyte molecule (protein etc.) creates a strong interaction through binding the hydrophobic column surface while the unbound portion is still in contact with the mobile aqueous phase. This is overcome and the protein desorbs when the organic phase of the column reaches a specific concentration causing elution of the peptide. Small di/tripeptides are generally polar due to their small size, and thus do not bind to the hydrophobic surface of the column and are not efficiently retained. The increased adoption of hydrophilic interaction chromatography (HILIC) (Alpert, 1990) has been used since the 1970's, utilising a biphasic system where a low water mobile phase and a high water stationary phase is formed a polar molecule will preferentially be retained on the aqueous stationary phase thus observed by MS. HILIC methods and techniques are covered well in the review by (Hemström and Irgum, 2006). In section 5.2 a HILIC column is used for the LC-MS/MS identification of a CR dipeptide, similar to what what would be produced if the GUS reporter protein was digested with trypsin. Oxidation studies using hydrogen peroxide (H_2O_2) were performed to monitor the oxidative route of this dipeptide.

Controlled environments are ideal to perform oxidative reactions involving the reactive Cys-thiol. Due to the pM to nM cellular concentrations of ROS and RNS *in vivo* they are hard to detect because of their short half-life caused by their high reactivity. *In vitro* reactions allow substances such as Cys-thiols, ROS and RNS to be mixed in the absence of the milieu of potential other cellular components. In section 5.3 a 40 residues Nt-Cys synthetic peptide (CR-40) which replicates 40 Nt residues of the GUS reporter protein was analysed by ESI-MS/MS. The CR-40 was exposed to H_2O_2 and NO separately and together and the modifications occurring to Cys were monitored over a time course.

5.2 Identification of the tryptic dipeptide Cys-Arg by LC-MS/MS and its Hydrogen peroxide driven oxidation

5.2.1 Detection of CR dipeptide as a mimic for N-term RCR tryptic peptide from the in vivo GUS reporter protein

Coinciding with the search for the Nt-Cys oxidation state of MC-GUS *in vivo,* synthetic peptides were used to examine the chemistry of Cys oxidation *in vitro*. The tryptic Nt MC-GUS peptide is predicted to be RCR (see section 4.4.1), however before the expected arginylation of the Nt-Cys it is required to be oxidised. Therefore a small dipeptide Cys-Arg would be a useful model to look at the oxidation of Cys as its small size minimises other potential PTMs. The CR dipeptide was synthesised and was injected for LC-MS which would tell us information regarding its purity and retention time (Figure 5.1).

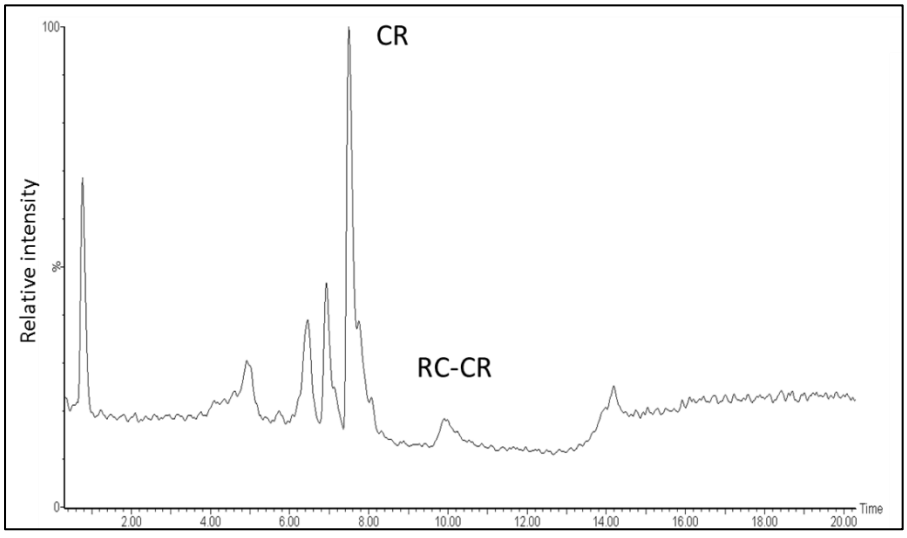


Figure 5.1 Chromatogram showing the elution peaks when the Cys-Arg dipeptide was injected into the LC-MS. 20 µM of CR was injected into the LC-MS and the different species were separated over time. The CR dipeptide eluted at 7.49 minutes and the dimerised CR dipeptide (RC-CR) eluted at 9.90 minutes.

Waters Synapt LC-MS analysis using QTOF MS showed that the CR dipeptide did not solely consist of one species, with 6 noticeable retention time peaks. The MS spectra from the elution peaks at 0.78, 4.89, 6.47 and 6.91 minutes were obscure and no species related to the CR dipeptide could be found. At 9.90 minutes the elution peak contained an ion with *m/z* 544.4 matching that for a disulfide of the CR dipeptide (RC-CR). The predominant peak at 7.49 minutes did contain the CR dipeptide however this did not represent the largest ion peak. Free Arg with mass *m/z* 175.1 was approximately nine times more abundant than the CR dipeptide, likely a result of insufficient removal of Arg during the F-moc synthesis (see section 2.7.1). To confirm the identification of our CR dipeptide its predicted mass *m/z* 278.2 was selected for MS/MS and fragmented (Figure 5.2). The resultant fragment ions were predominantly those for a single Arg residue (*y* ion, *m/z* 175.2 and *y** ion *m/z* 158.1) and the resultant *b*2 ion confirming the presence of both an Nt-Cys and C-terminal Arg (*m/z* 278.2).

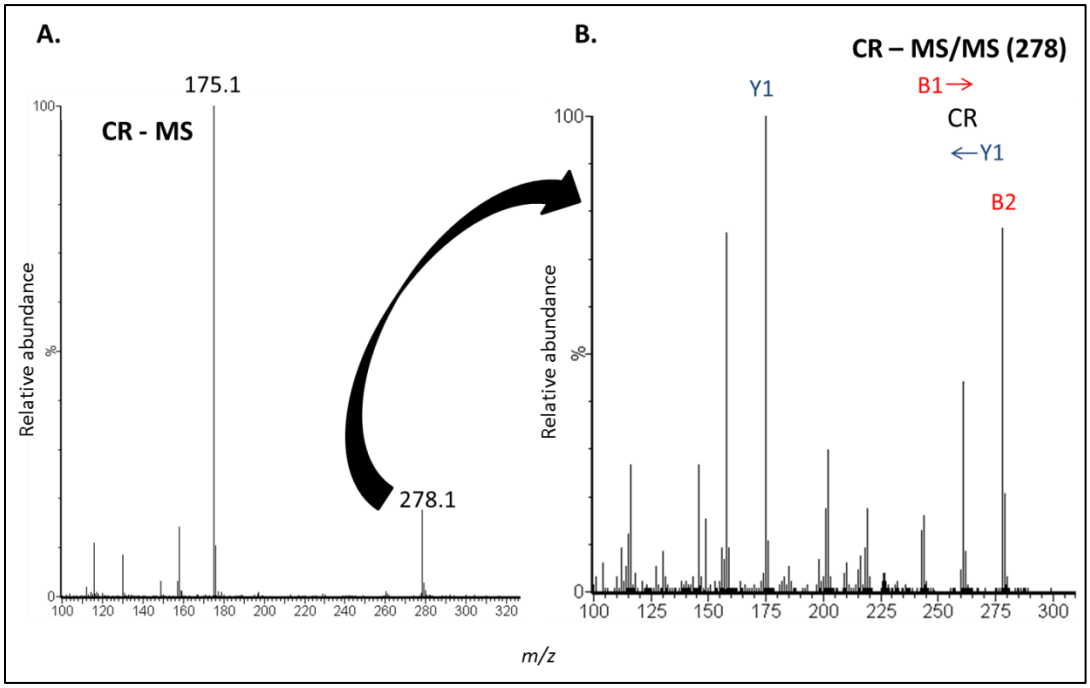


Figure 5.2 Waters Synapt LC-MS and QTOF MS/MS of the Cys-Arg dipeptide. A. 20 µM of CR dipeptide was analysed by MS and the parent *m/z* 278 ion peak (matching the mass of CR) was chosen for MS/MS. B. Fragmentation of the *m/z* 278 species generated *Y* ions and *B* ions, the *m/z* 175 peak represents Arg (*Y*1). The presence of Arg in the initial MS scan indicates that surplus Arg was present during synthesis of the dipeptide.

The CR dipeptide was initially left to react for 10 minutes in water before LC-MS injection. The reactive nature of the Cys-thiol can cause it to react quickly even in non-reactive neutral solvents such as water. The mass of the expectant oxidative Cys products was searched, including Cys-sulfenic (*m/z* 294), Cys-sulfinic (m/z *310*) and Cys-sulfonic (*m/z* 326). Expectedly the ion count (IC) for these species was very low and compared to unoxidised Cys which was 50-100x more abundant (Figure 5.3). Interestingly a species with the mass of Cys-sulfinic acid co-eluted with the unoxidised Cys 7.7 minutes after injection, this came roughly a minute after the co-elution of species matching the mass of both Cys-sulfenic and Cys-sulfonic acid. Due to the low abundance of the oxidative species MS/MS failed to prove the identification of the species with matching masses.

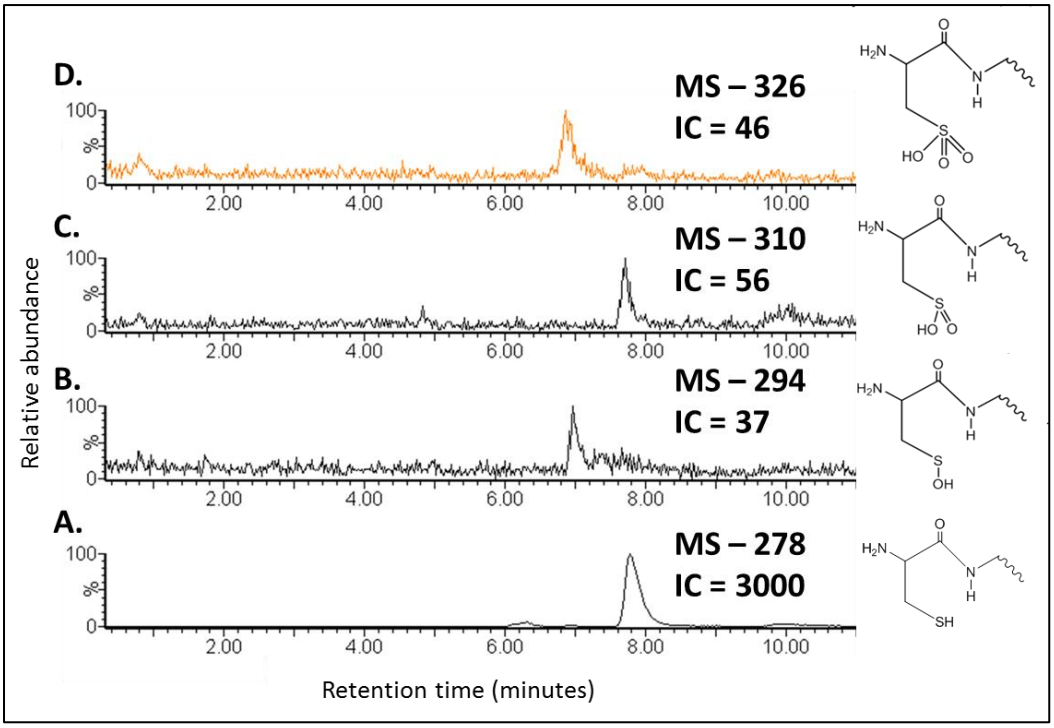


Figure 5.3 Ion chromatograms of untreated Cys-Arg (CR) dipeptide elution peaks. 20 µM of CR was injected in water and the retention time (RT) of unoxidised Cys and oxidised Cys species were monitored. Targeted isolation and their respective IC (ion count) is shown for A. Cys (unoxidsed)-Arg = *m/z* 278; B. Cys(sulfenic acid)-Arg = *m/z* 294; C. Cys(sulfinic acid)-Arg = *m/z* 310; D. Cys(sulfonic acid)-Arg = *m/z* 326. The RT of unoxidised Cys appears to co-elute with Cys-sulfinic acid just before 8 minutes whist Cys-sulfenic and Cys-sulfonic acid appear to co-elute earlier at 7 minutes. The smooth peak of the unoxidised Cys represents the higher abundance (IC > 3000) of this species compared to the low IC for the other oxidised species.

5.2.2 Oxidation of CR dipeptide with hydrogen peroxide

In an attempt to forcibly oxidise the synthetic CR dipeptide, CR was incubated with hydrogen peroxide (H_2O_2). In biological systems H_2O_2 is a natural by-product of respiration and other metabolic reactions which is particularly prone to react with Cys-thiols (Cys-SH) (Winterbourn, 2013). The initial oxidation of Cys forms a disulfide (Cys-S-S-Cys) which can then break upon additional reaction with the oxidative source forming a Cys-thiolate ($Cys\text{-}S^-$), which is more reactive than the initial Cys-thiol (Trivedi et al., 2009). After incubating CR with H_2O_2 for 45 minutes oxidative products were observed, the most abundant was the disulfide

(RC-CR) which was present in its 2^+ charge state (*m/z* 277.2) (Figure 5.4).

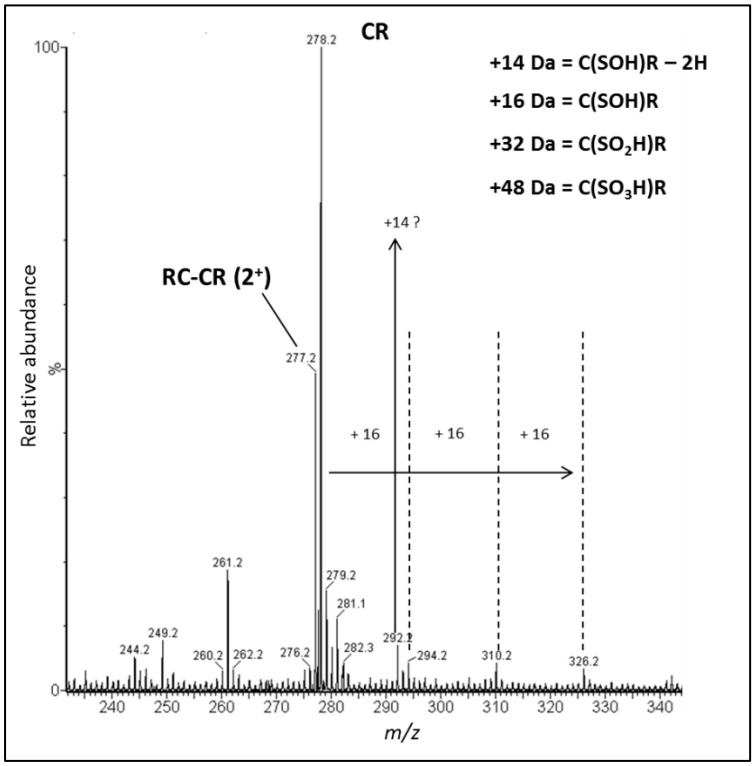


Figure 5.4 Waters Synapt QTOF LC-MS mass spectra of the Cys-Arg dipeptide after incubation with H_2O_2. 20 uM of CR dipeptide was incubated with 0.027% H_2O_2 and left to react for 45 minutes before injecting and performing LC-MS. The parent unoxidised CR was the predominant peak (*m/z* 278.2), although the disulfide (RC-CR, 2^+) had formed (*m/z* 277.2) and had reached 50% abundance of CR. Further oxidised species were also visible with the possible Cys-sulfenamide (CysSOH -2H, *m/z* 292.2) +14 Da from *m/z* 278.2. A sequence of +16 Da oxidations were also visible from the unoxidised CR likely representing the Cys-sulfenic (CysSOH, *m/z* 294.2), Cys-sulfinic ($CysSO_2H$, *m/z* 310.2) and Cys-sulfonic ($CysSO_3H$, *m/z* 326.2) acids respectively.

Although low in relative abundance (<10%) there were peaks that matched expected masses for Cys oxidation species. The most abundant of these was +14 Da on the mass of CR matching a Cys-sulfinamide (Cys-SO-N). A sequence of peaks were also observed matching the addition of oxygen each spaced +16 Da apart and would match the mass

for the Cys-sulf(en/in/on)ic acid. The oxidation reaction with CR was left to continue for 24 hours and the elution peak for CR and the Cys-sulf(en/in/on)ic acid was analysed (Figure 5.5)

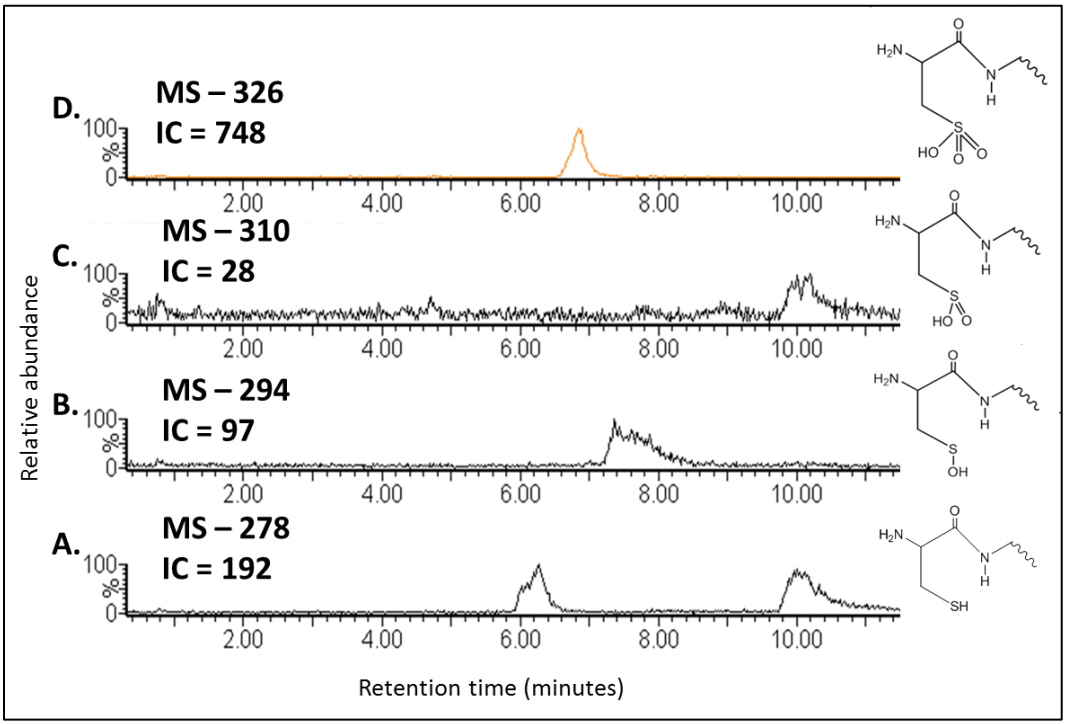


Figure 5.5 Ion chromatograms presenting the elution peaks from the oxidised Cys-Arg dipeptide. 20 µM of CR was incubated for 24 hours in 0.027% H_2O_2 and the retention time (RT) of unoxidised Cys and oxidised Cys species were monitored. The targeted RT's of A. Cys (unoxidsed)-Arg = *m/z* 278; B. Cys(sulfenic acid)-Arg = *m/z* 294; C. Cys(sulfinic acid)-Arg = *m/z* 310; D. Cys(sulfonic acid)-Arg = *m/z* 326 are also presented. The RT of unoxidised Cys has changed from when unoxidised (figure 5.3) and is split with two different species at ~6 and ~10 mins respectively both with a much lower ion count (IC) than when unoxidised. The RT peak for masses matching Cys-sulfenic acid has also changed, eluting later but still has a weak IC. Masses matching Cys-sulfinic acid has also changed but still appears to co-elute with some *m/z* 278 species. A mass representing Cys-sulfonic acid now has the highest IC indicating that CR has been oxidised to the maximum state, the RT remains the same for this species after oxidation.

In comparison to figure 5.3 the elution of CR and its anticipated oxidised products is changed after incubation with H_2O_2. Overall the total ion count (TIC) for the oxidised CR sample was similar to when unoxidised, however the most striking change is the absence of the elution peak

matching CR previously observed after 7.8 minutes. Although other peaks are RT peaks containing species of this mass the ion count (IC) has fallen from 3000 to under 200. The most abundant species was Cys-sulfonic acid (*m/z* 326.1), indicating that after 24 hours most of the initial Cys had oxidised to its highest stable form. A mass addition of 151 Da was observed from the Arg peak (*m/z* 175.1), matching the addition of Cys-sulfonic acid from the MS/MS data acquired for the 326 species (Figure 5.6). Unlike the Cys-sulfinic version of the dipeptide which couldn't be identified by MS/MS, strangely the Cys-sulfenic (*m/z* 294) was more abundant and MS/MS data for this species (although with a weak signal) showed a mass addition of 119 Da from the Arg peak.

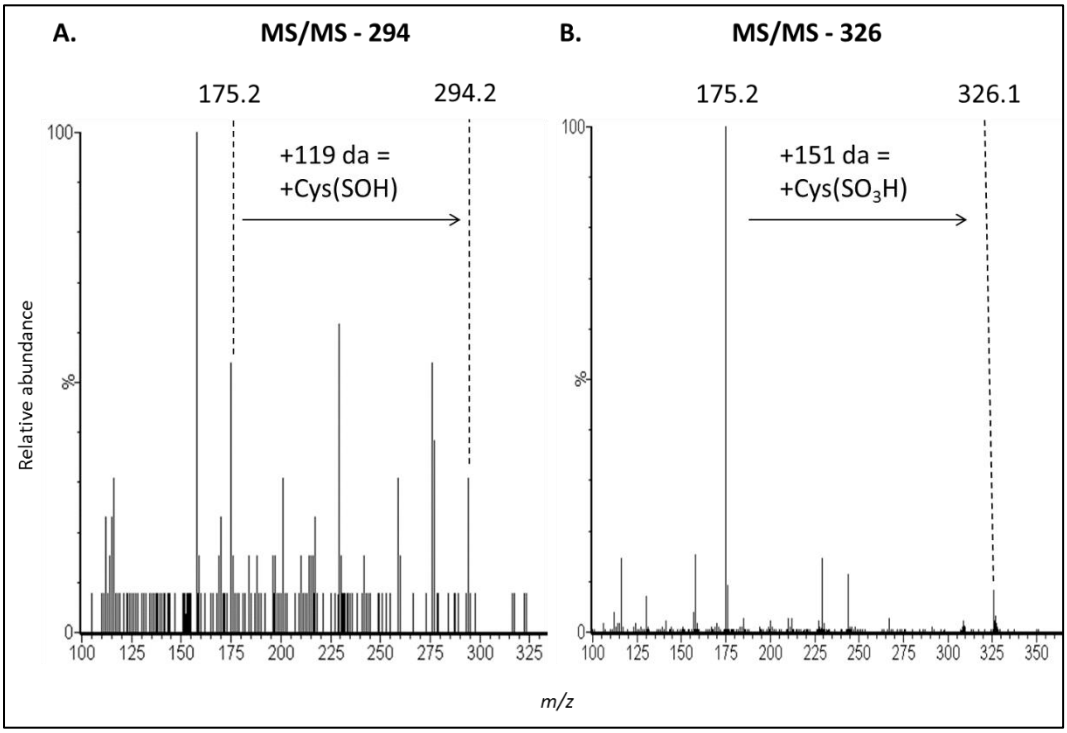


Figure 5.6 Waters Synapt QTOF LC-MS/MS spectra of A. Cys-sulfenic acid and B. Cys-sulfonic acid. 20 µM of CR dipeptide was incubated with 0.027% H_2O_2 and left to react for 2 hours before MS/MS was performed on the a *m/z* 294 species. A. This was identified to be Cys-sulfenic acid indicated by the +119 Da increase from the *m/z* 175 Arg peak corresponding to Cys(SOH). B. When the CR dipeptide was left to incubate in H_2O_2 for 24 hours MS/MS was performed on a *m/z* 326 species, this was identified to be Cys-sulfonic acid due to the +151 Da increase corresponding to Cys(SO_3H) upon Arg.

5.3 *In vitro* oxidation of cysteine N-terminal 40 residue synthetic peptide reveals sulfenamide oxidative route and NO-induced nitrosylation

5.3.1 ESI-MS/MS identification of CR-40 synthetic peptide

To monitor the oxidation of Cys in a long polypeptide replicating what would be created after Nt cleavage of MC-GUS with Asp-N, a 40 residue Nt-Cys protein was ordered. The 40 residue Nt-Cys peptide replicating the *in vivo* GUS reporter protein was termed CR-40 and had a mass of 3.68 kDa. The CR-40 peptide was initially analysed using LC-MS, an elution peak containing the CR-40 peptide was the only eluting substance observed, indicating high purity of the peptide (> 98%) (Figure 5.7). CR-40 was present in 4 charge states, 3^+ to 6^+, with 5^+ the most abundant.

Due to the purity of CR-40 electrospray ionisation mass spectrometry (ESI-MS) was the ionisation source used to perform MS/MS. ESI-MS is a faster analysis method performed manually and not involving unnecessary prolonged chromatographic separation. Direct infusion of 1 µM CR-40 with aqueous solvent before MS identified a *m/z* 737.39 (5^+) ion peak (Figure 5.8). Two different methods for fragmenting the *m/z* 737.39 CR-40 ion into its constituent residue ions were used, electron capture dissociation (ECD) and collision induced dissociation (CID). The C-ions produced during ECD match those predicted if Cys and the other Nt residues were unmodified (Figure 5.8). The CID fragmentation of CR-40 produced *b* and *y* ions, which together matched masses which covered all residues of unmodified CR-40.

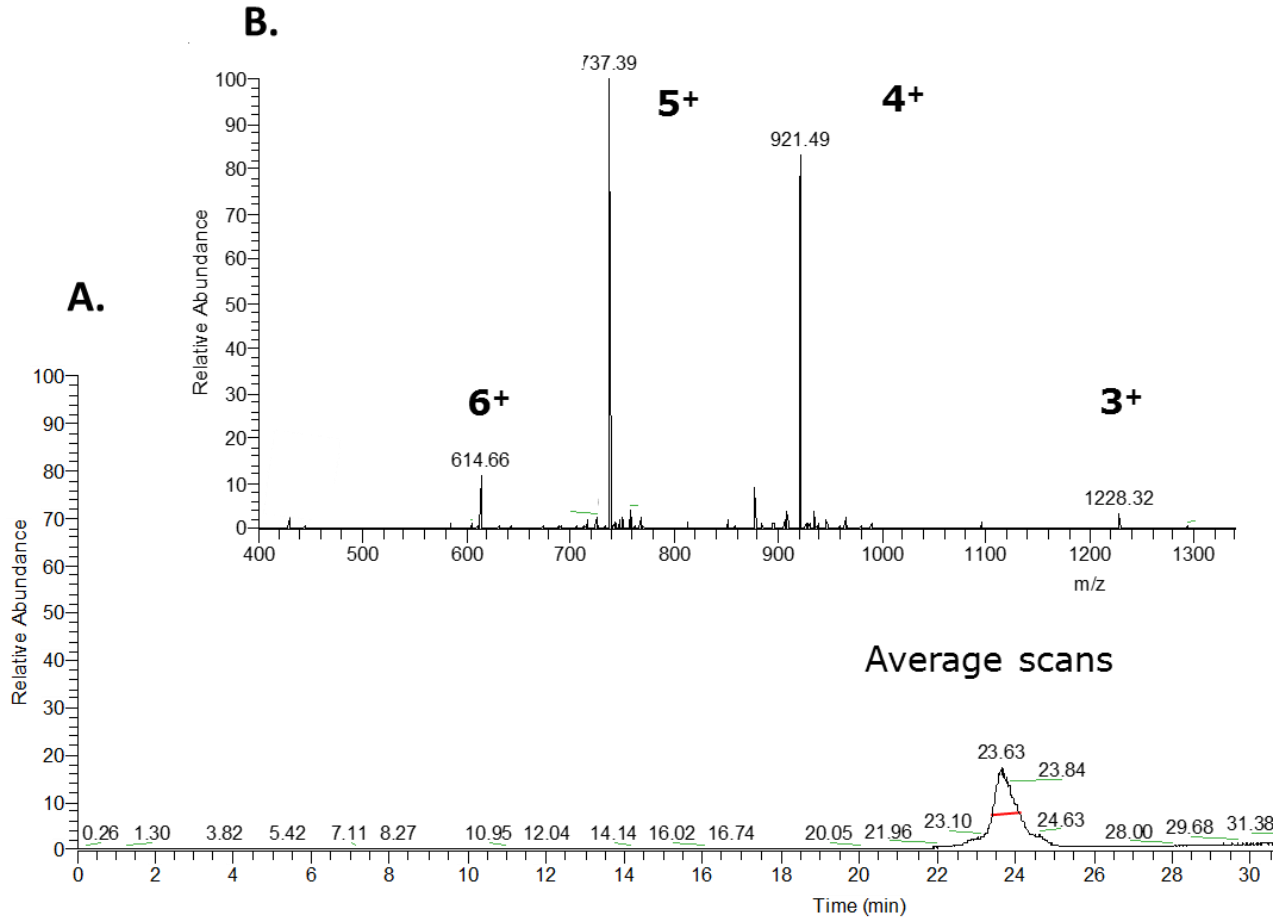


Figure 5.7 Fourier transform ion cyclotron resonance (FTICR) LC-MS A. chromatogram and ion spectra of the CR-40 synthetic peptide displaying the retention time and charge states. B. The CR-40 peptide carries four visible charge states 3^+ to 6^+. 1 µM of CR-40 solubilised in water was directly infused (95:5 H_2O:MeCN) before injecting into the MS.

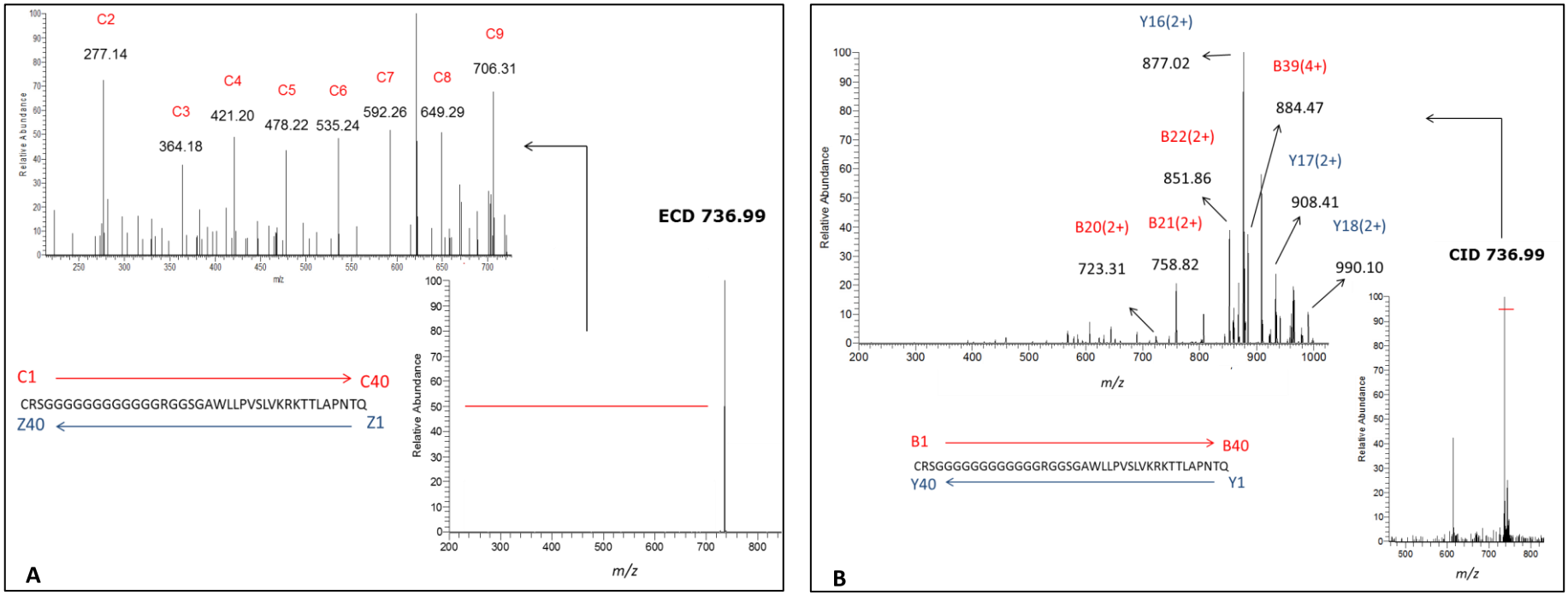


Figure 5.8 Fourier transform ion cyclotron resonance (FTICR) LC-MS and MS/MS spectra showing the electron capture dissociated and collision induced dissociated fragment residue ions of the CR-40 synthetic peptide. 1 µM of the CR-40 peptide was dissolved in water and left to react at room temperature for 10 minutes before direct infusion (95:5 H_2O:MeCN) into the ESI-source. (A) Electron capture dissociation (ECD) performed on the *m/z* 737.39 5^+ peak of CR-40 generating *c* ions from the N-terminus matching those expected for an unoxidised Nt-Cys. (B) Collision induced dissociation (CID) performed on the *m/z* 737.39 5^+ peak of CR-40 generating *b* and *y* ions from the N and C-terminus covering the whole peptide sequence and matching an unmodified peptide.

5.3.2 Oxidation of the N-terminal Cys of CR-40 using hydrogen peroxide

In order to track experimental Nt oxidation of CR-40, ECD was used as the fragmentation method. The *c* ions identified from the untreated CR-40 all include the Nt-Cys, therefore using *c* ions such as *c2* or *c3* are useful in finding modifications which are only occurring at the Nt-Cys. There were noticeable differences between the MS spectra when CR-40 was unoxidised or oxidised with H_2O_2 for two hours. Most strikingly was the decline in the ion peak for the unoxidised CR-40 peak (monoisotopic (mono) mass = *m/z* 736.99) and the increase in the Cys-sulfonic acid (mono mass = *m/z* 746.59) as the predominant species after exposure to H_2O_2 (Figure 5.9). The other noticeable change is the emergence of the Cys-sulfenamide CR-40 (mono mass = *m/z* 736.59) which is 2 Da lighter than the original CR-40. In both the unoxidised and oxidised MS spectra there appears to be Cys-sulfonamide CR-40 (mono mass = *m/z* 742.99) which increases from 10% relative abundance to over 40% relative abundance when oxidised. To confirm the identity of the Cys-sulfonamide ECD was used to fragment this species after incubating CR with H_2O_2, producing C-ions matching a +30 Da increase on the Nt-Cys thus confirming the identity of the Cys-sulfonamide (Figure 5.10).

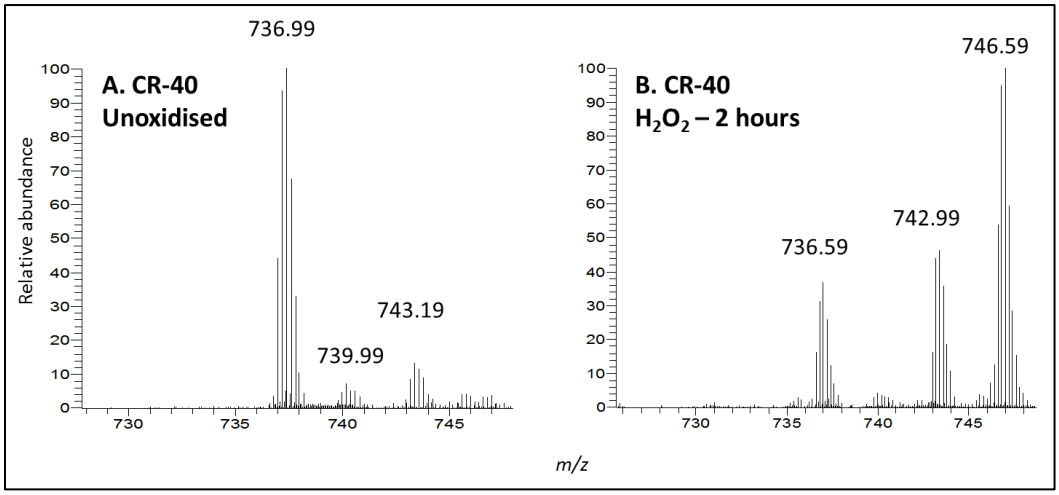


Figure 5.9 Fourier transform ion cyclotron resonance (FTICR) ESI-MS spectra of CR-40 before and after oxidising with Hydrogen peroxide. A. Unoxidised CR-40 is dominated by a species matching that of unmodified CR-40 (Monoisotopic (mono) mass peak = *m/z* 736.99), this having an unoxidised N-term Cys. A series of oxidised CR-40 species are visible at abundance < 10% of the level of unoxidised CR-40, including Cys-thiolate (*m/z* 739.99), Cys-sulfinic acid (*m/z* 743.19). B. After oxidation CR-40 with Cys-sulfonic acid (*m/z* 746.59) had become the predominant species. The Cys-sulfonamide (*m/z* 742.99) had formed and was the second most abundant species after 2 hours in H_2O_2. The unoxidised Cys CR-40 was now not visible with the increased formation of the Cys-sulfenamide (*m/z* 736.59).

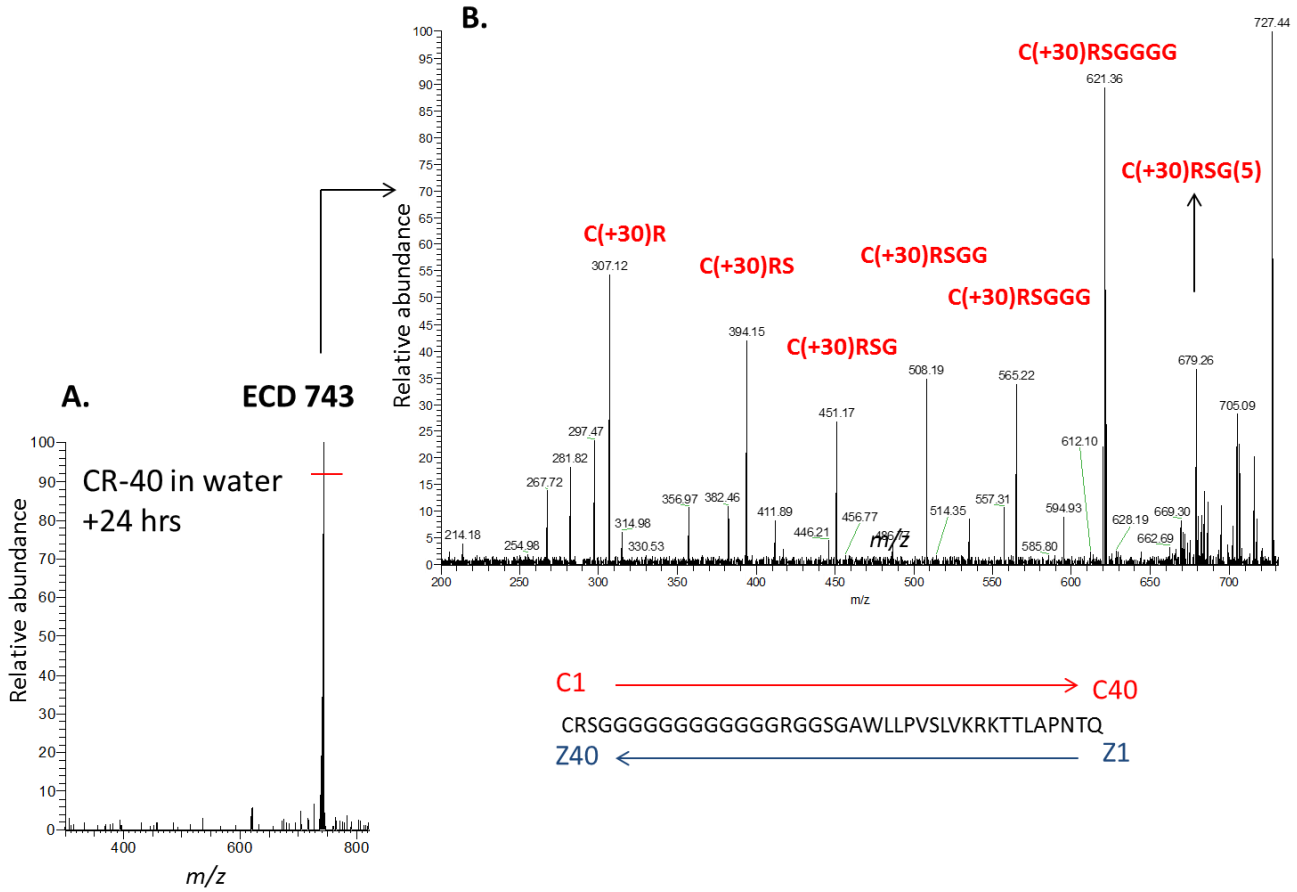


Figure 5.10 Fourier transform ion cyclotron resonance (FTICR) A. LC-MS and B. MS/MS spectra showing the electron capture dissociated fragment residue ions of an oxidised CR-40 product after reacting with water for 24 hours. A. 1 µM of CR-40 was solubilised in HPLC water and left at room temperature for 24 hours before direct infusion (95:5 H_2O:MeCN) and injection for MS. B. Electron capture dissociation performed on the predominant *m/z* 743.39 5^+ peak of CR-40 generating C-ions from the N-terminus matching a species with a +30 Da modification to the Nt-Cys matching the formation of a Cys-sulfonamide.

The oxidation of the Nt-Cys of CR-40 appears to involve two oxidative routes. The first follows the conventional oxidation of the thiol forming a Cys-sulfenic acid, with a corresponding mass addition of an oxygen atom (+16 Da) (Figure 5.11). Subsequent oxidation of Cys-sulfenic acid creates Cys-sulfinic acid (+32 Da) which can be oxidised further to the highest Cys oxidation state, Cys-sulfonic acid (+48 Da). Alternatively Cys-sulfenic acid can undergo a dehydration reaction which causes ring closure of Cys, forming Cys-sulfenamide (-2 Da). Oxidation of Cys-sulfenamide produces Cys-sulfinamide (+14 Da) and then if oxidised further Cys-sulfonamide (+30 Da). The hydrolysis of Cys-sulfinamide can break the pentose ring and form Cys-sulfinic acid. The formation of Cys-sulfenamide species is not conventional, initially discovered in kidney cells where protein tyrosine phosphatase 1B (PTP1B), a negative regulator of the insulin signalling pathway, is inactivated through reaction with H_2O_2 and superoxide (O_2^-) (Salmeen et al., 2003). Oxidation of the Cys-215 residue of PTP1B can render it non-functional, the formation of the reversible intermediate Cys-sulfenamide species was suggested to be a protective mechanism, preventing the irreversible inactivation of the enzyme (Salmeen et al., 2003; van Montfort et al., 2003).

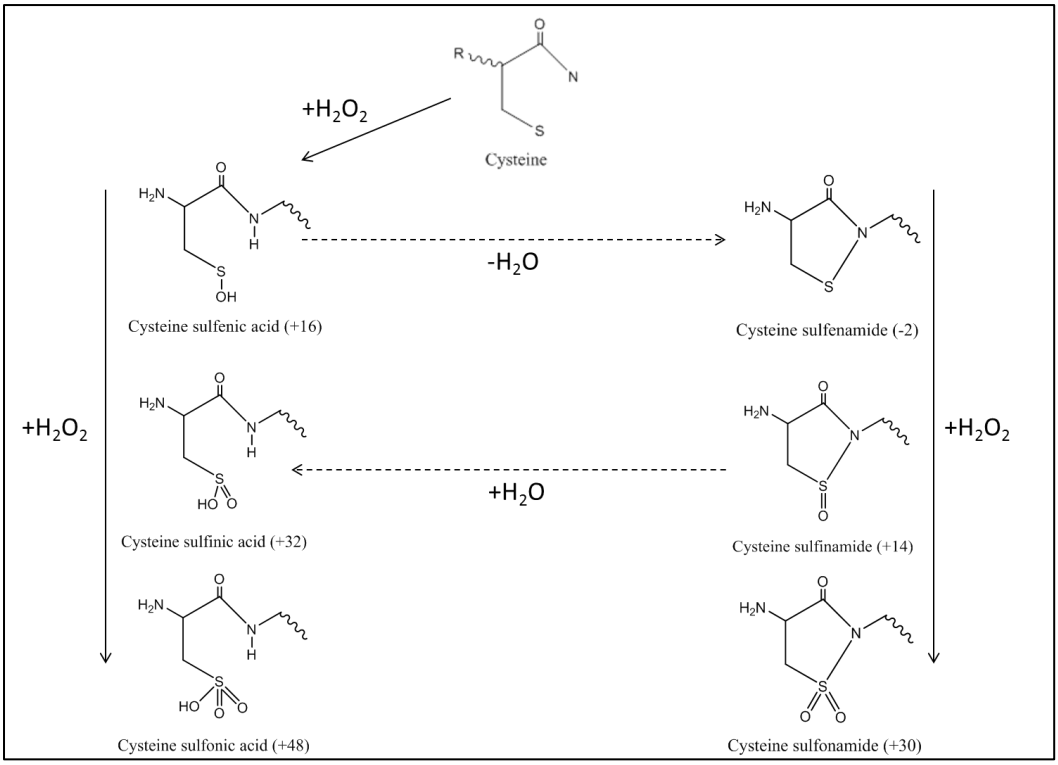


Figure 5.11 Cysteine oxidative modifications which can be driven by Hydrogen peroxide exposure. Highlighting the conventional oxidation of cysteine to Cys-sulfenic (+16 Da), Cys-sulfinic (+32 Da) and then Cys-sulfonic acid (+48 Da). Cysteine can also undergo dehydration when oxidised, closing the cysteine ring forming a Cys-sulfenamide (-2 Da), which can then be further oxidised to Cys-sulfinamide (+14 Da) which undergo hydrolysis to break the ring structure forming Cys-sulfonic acid or be further oxidised to Cys-sulfonamide (+30 Da).

The oxidation of CR-40 was monitored over a time course to investigate the rate of oxidation and the transition between CR-40 species. An advantage of using ESI-MS for pure samples such as synthetic peptides is the quick acquisition of data, without the need to use long chromatographic gradient programmes to separate different species within samples. The effect of exposure to H_2O_2 was monitored after 1 minute reaction with CR-40, and within this minute oxidation had already occurred, with Cys oxidising to the Cys-sulfenamide species (figure 5.12). The abundance of unoxidised Cys had dropped from 85% of total Cys species when in water to being undetectable after exposure to H_2O_2.

The formation of Cys-sulfenamide as the predominant species supports previous work highlighting the high reactivity of the Cys-sulfenic acid (Salmeen et al., 2003; van Montfort et al., 2003) which immediately undergoes a condensation reaction to form the more stable Cys-sulfenamide.

After 5 minutes the levels of Cys-sulfenamide and Cys-sulfinamide had decreased compared to their relative abundance after one minute, these decreases coincided with an increase in both the Cys-sulfonic and Cys-sulfonamide species. The Cys-sulfenamide levels remained stable up until 20 minutes whilst Cys-sulfinamide and Cys-sulfonamide begin to decline slowly, this coincides with a slow increase in Cys-sulfonic acid which had an 8% rise in total abundance between 5 and 20 minutes. After 40 minutes of H_2O_2 oxidation Cys-sulfonic acid was the predominant Cys product, which continued to increase in abundance until the end of the time course after two hours. The increase of Cys-sulfonic acid may be as a result of the hydrolysis of the Cys-sul(fe/fi/fo)namide species which all decrease in abundance.

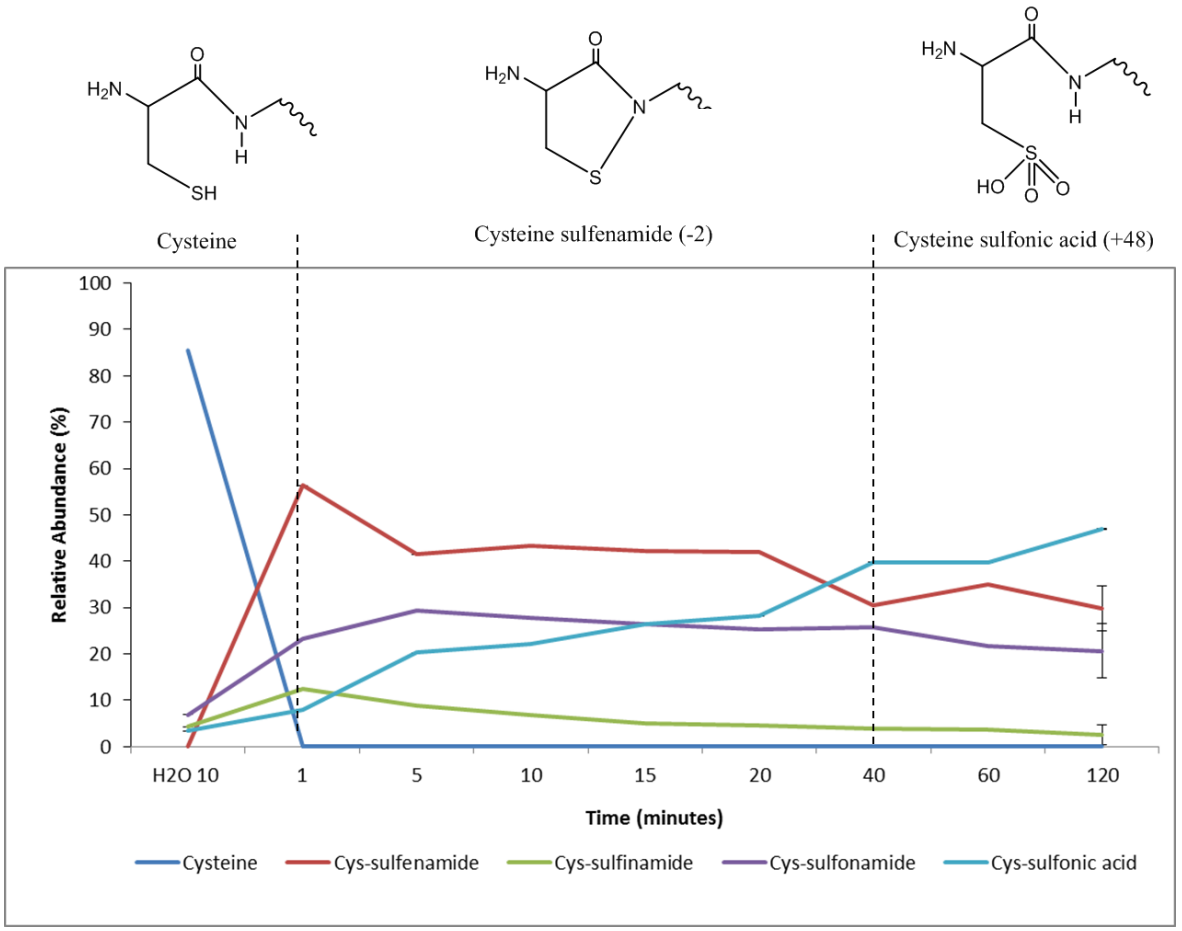


Figure 5.12 Reaction time course showing the oxidative modifications occurring the Nt-Cys of the CR-40 synthetic peptide when incubated with Hydrogen peroxide. 1 µM of CR-40 peptide was incubated with 0.027% H_2O_2 and reacted for 1, 5, 10, 15, 20, 40, 60 and 120 minutes before direct infusion (95:5 H_2O:MeCN) and injection. The initial data point as a baseline was the CR-40 peptide incubated in water for 10 minutes. The dashed lines are used to separate periods over the time course with a different predominating species. SEM error bars included at last time point (120 mins).

5.3.3 SNAP-induced CR-40 Nt-Cys nitrosylation

Nitric oxide (NO) a ubiquitously cellular reactive species is the causative agent in Cys-nitrosylation, a PTM which regulates protein function and stability. The S-NO bond of the Cys-nitrosothiol (Cys-SNO) is extremely labile and, similarly to the Cys-sulfenic acid is also redox sensitive (Hare and Stamler, 2005). Endogenous production of NO can be substituted *in vitro* using *S*-Nitroso-*N*-acetyl-DL-penicillamine (SNAP), an NO donor. SNAP was visible by LC-MS when diluted in water as a 221 Da ion peak, present alongside its disulfide dimer (Figure 5.13). When SNAP was incubated with CR-40 nitrosylation of the Cys-thiol did occur with the formation of a peak matching CR-40 but with a Cys-nitrosothiol (+29 Da; mono mass = m/z 742.79) (Figure 5.14). The reaction of SNAP with CR-40 was slower than when oxidising with H_2O_2. Dimerization of SNAP with CR-40 was also visible by LC-MS, three peaks matched masses for the SNAP-(S-S)-CR-40 disulfide before SNAP had donated NO (m/z 781), after NO had been donated (m/z 774.8) and after SNAP had received NO (m/z 787). Despite identifying the mass of Cys-SNO via LC-MS its identity was not confirmed by MS/MS, using CID only *y* ions from the C-terminus of CR-40 were detected (Appendix F).

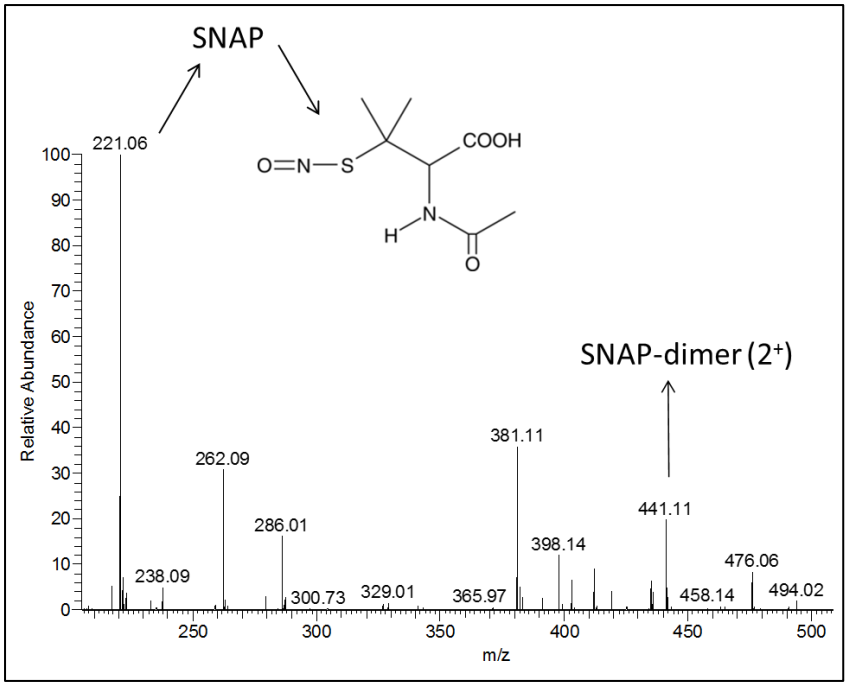


Figure 5.13 Fourier transform ion cyclotron resonance (FTICR) LC-MS profile of *S*-Nitroso-*N*-acetyl-DL-penicillamine (SNAP). 10 µM of SNAP diluted in water was directly infused (95:5 H_2O:MeCN) and injected. SNAP was detected in both its monomeric and dimerised form. The structure of SNAP is presented.

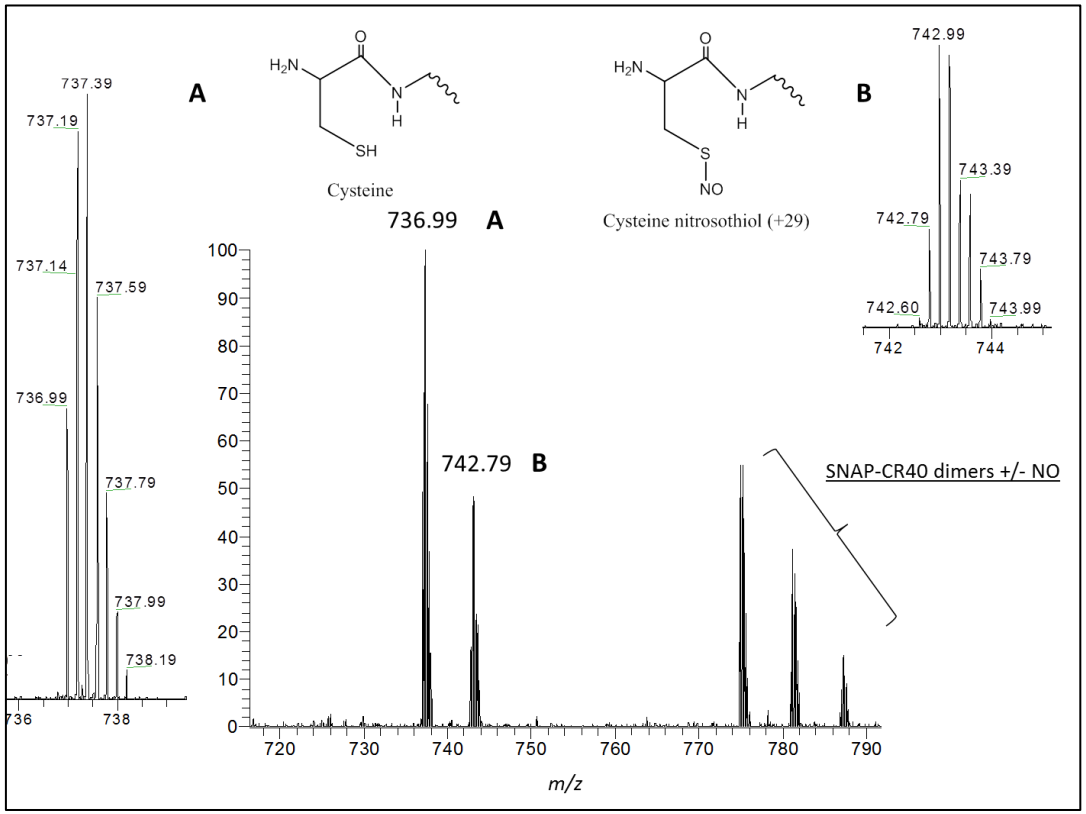


Figure 5.14 Fourier transform ion cyclotron resonance (FTICR) LC-MS spectra of the product ion peaks after incubating the CR-40 synthetic peptide with the NO-donor S-Nitroso-N-acetyl-DL-penicillamine (SNAP). 1 µM of CR-40 was incubated with 50 µM SNAP for 10 minutes before direct infusion (95:5 H2O:MeCN) and injection into the MS. A. The predominant species after 10 minutes was unmodified cysteine (*m/z* 736.99) B. However at ~45% of the unmodified cysteine relative abundance a modified peak with a representative +29 Da mass shift was visible (*m/z* 742.79) matching the formation of a Cys-nitrosothiol (Cys-SNO). Additionally a series of SNAP dimers where also found matching CR-40 bound to SNAP which had donated NO (*m/z* 775), remained unmodified (*m/z* 781.20) or acquired NO (*m/z* 787.20).

The reaction mechanism for the SNAP induced nitrosylation of CR-40 is outlined in figure 5.15. Using ESI-MS a time course for the reaction between CR-40 and SNAP was monitored. The two species monitored were the unmodifed CR-40 and the nitrosylated Cys-SNO CR-40. Incubation of CR-40, with 20 μM SNAP caused the gradual nitrosylation of the Nt-Cys thiol of CR-40 which became the predominant species after 60 minutes reaction (Figure 5.16). This lag-time before nitrosylation of CR-40 occurs may be related to the decomposition of SNAP releasing NO into solution, therefore as reaction time increases more NO will be released into solution increasing reaction rates with CR-40. In proteins *in vivo* there may be two different routes of Cys-SNO formation. The first could involve a disulfide bond being formed between the Nt-Cys of two CR-40 peptides which then undergoes homolytic scission (R1), the resultant thiolate radical can react with NO forming a Cys-SNO (R2) (Jourd'heuil et al., 2003).

R1: Cys-S$^{\cdot}$S-Cys = Cys-S$^{\cdot}$

R2: Cys-S$^{\cdot}$ + NO = Cys-SNO

The second reaction may involve the inital oxidation of NO forming nitrogen dioxide (NO_2) (R3) (Lewis and Deen, 1994). The reaction of NO_2 with excess NO produced by SNAP would then form dinitrogen trioxide (N_2O_3) (R4), which then go on to nitrosylate a Cys-thiol (R5) (Keshive et al., 1996).

R3: $2NO + O_2 = 2NO_2$

R4: $NO_2 + NO = N_2O_3$

R5: Cys-SH + N_2O_3 = Cys-SNO + H^+ + NO_2^-

Interestingly when CR-40 was nitrosylated there was no evidence for subsequent oxidation to either oxidised Cys-acids or Cys-amides. Decomposition of the Cys-nitrosothiol S-NO bond in biologcal systems can occur in the presence of white light, heat and interaction with heavy metals (Cu, Hg etc.). This indicates that the S-N bond of nitrosothiols is weak, measured between 23 and 32 kcal/mol using crystallography analysis (Bartberger et al., 2000).

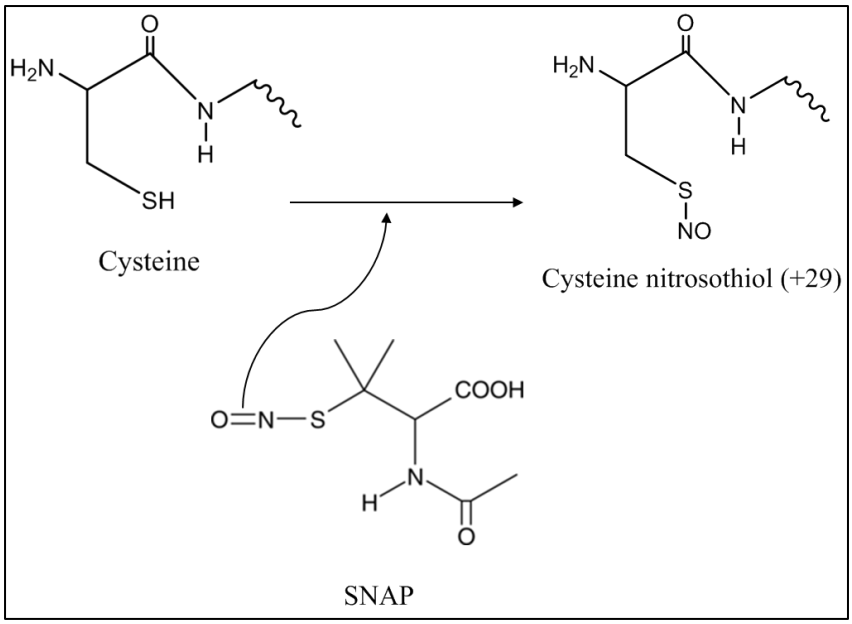


Figure 5.15 Reaction mechanism for Cysteine and S-Nitroso-N-acetyl-DL-penicillamine (SNAP) forming a Cys-nitrosothiol

To investgate whether the S-N bond could be further oxidised H_2O_2 was introduced after SNAP with CR-40, however no subsequent modifications occurred to the Cys-SNO, although increased Cys-dimer species were formed after reacting with unoxidised CR-40 (Appendix G). Similarly whether the species created after CR-40 was exposed to H_2O_2 could be modified if SNAP was added was also investigated. After the H_2O_2-

induced formation of the Cys-sulfenamide, Cys-sulfonamide and Cys-sulfonic acid, SNAP was added to the reaction mix. Addition of SNAP did not cause any further reactions to the exisiting oxidised Cys species (Appendix H). This is an interesting finding as *in vivo* oxidation of Nt-Cys substrates of the N-end rule pathway requires both NO and oxygen (Gibbs et al., 2011; Gibbs et al., 2014b; Hu et al., 2005), so although no modifications were observed there could be an interaction between these two species *in vivo*. A second NO source was also used in this study, sodium niroprusside (SNP), which releases NO into solution, however no nitrosylative effect was observed when 20 µM was incubated with 1 µM CR-40 in solution. Despite this SNP did have an apparent oxidative effect, similar although slower than H_2O_2, with increased formation of oxidised species such as Cys-sulfonamide and Cys-sulfonic acid after 3 hours incubation (Appendix I).

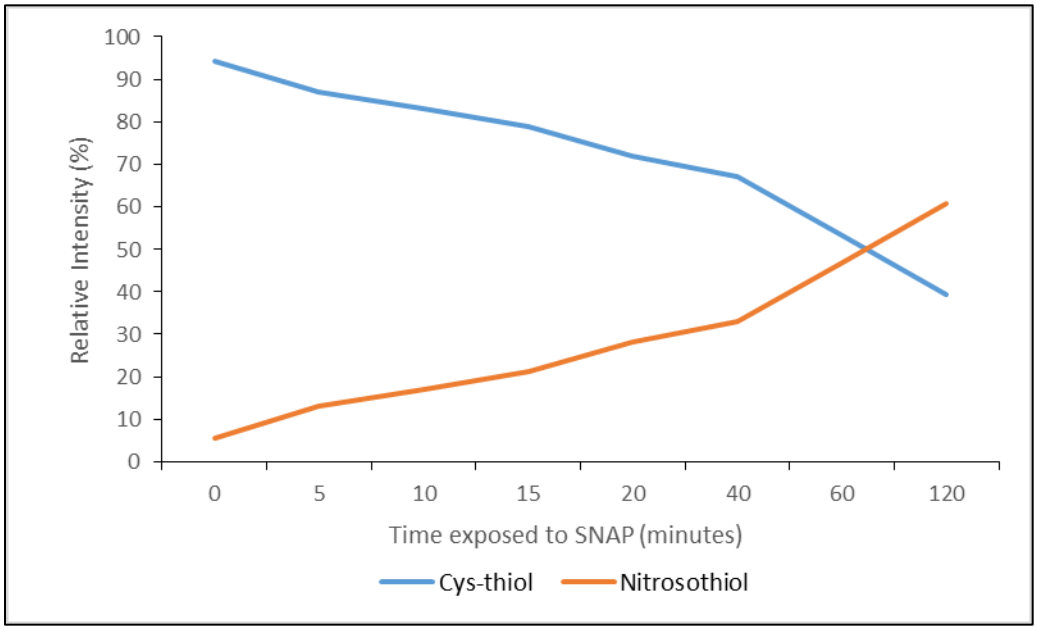


Figure 5.16 Time course following the reaction between Cysteine and *S*-Nitroso-*N*-acetyl-DL-penicillamine (SNAP) over a two hour period. 1 µM CR-40 synthetic peptide was incubated with 20 µM SNAP and the relative abundance of Cysteine (Cys-thiol) and the nitrosylated cysteine (nitrosothiol) were monitored at 0 (before SNAP addition), 5, 10 15, 20, 40, 60 and 120 minutes after SNAP addition. The Cys-nitrosothiol became the abundant species after one hour of incubation.

6. General discussion

The creation of hypoxic zones in soils is becoming a more common occurrence as a result of climate change, flooding of agricultural lands causes crops to be starved of oxygen, resulting in reduced yields or death. With the increasing global population and the increased weather extremities forecast there is a need to improve the resilience of our crops to flooding. The discovery that plants sense oxygen and nitric oxide through ERFVII protein substrates of the Arg/N-end rule pathway (Gibbs et al., 2011; Gibbs et al., 2014b; Licausi et al., 2011), revealed new opportunities to improve flooding crop tolerance in plants. The identification of how Nt-Cys is oxidised and what oxidation state is required for arginylation by ATE, may help identify new substrates of the N-end rule pathway and novel oxidation routes of Nt-Cys.

6.1 Cys-Arg/N-end rule regulated sensors of oxygen in Arabidopsis

This study has highlighted the localised expression of *RAP2.12* in Arabidopsis seedlings, and the importance of the Nt-Cys in the stabilisation of this protein. Enhanced stabilisation of GUS was observed in the SAM from transgenic GUS reporter lines of Nt-Met-Ala (MA) RAP2.12, when driven by the *RAP2.12* promoter. Developing a transgenic GUS reporter line of RAP2.12 driven by its native promoter, with a mutant Ala-2 N-terminus of RAP2.12 enhanced stabilisation of GUS thus RAP2.12 in the SAM in normoxia. This stabilising effect of the Nt-MA

residues on RAP2.12 is consistent with Licausi et al. (2011), which reported increased stabilisation of green fluorescent protein (GFP) fused to MA-RAP2.12 from leaves of *35S::MA-RAP2.12-GFP* Arabidopsis seedlings compared to MC-RAP2.12.

Stabilisation of RAP2.12 increased in the SAM and hypocotyl after hypoxia treatment in *promRAP2.12::MC-RAP2.12-GUS* seedlings in Col-0. The hypoxia induced stabilisation of RAP2.12 observed in this study is in agreement with results reported in Licausi et al. (2011) and Kosmacz et al. (2015), where the GFP and Western blot protein signal for RAP2.12 increased after subjecting Arabidopsis to hypoxia (1% O_2 for 4 hours) and subsequently decreased when re-oxygenated. These studies confirmed the subcellular localisation of RAP2.12 from the plasma membrane to the nucleus in leaves upon hypoxia, however it took 3-4 hours of re-oxygenation to revert to the normoxic state (Kosmacz et al., 2015). It would be interesting to see if the high expression of RAP2.12 in the SAM of 7d seedlings of Arabidopsis seedlings as shown in this study, would result in increased nuclear RAP2.12 stabilisation and thus perhaps a longer re-oxygenation period to translocate it out of the nucleus in the SAM.

In the roots the results from this study show that RAP2.12 was only stabilised in the RAM and to a lesser extent in the REZ after *promRAP2.12::MC-RAP2.12-GUS* seedlings were subject to hypoxia. Previously expression data of RAP2.12 at the root tip has found expression of *RAP2.12* to be similar in the REZ than the RAM (Birnbaum et al., 2003; Bui et al., 2015), indicating that the RAM cells may be more

sensitive to oxygen. MA-RAP2.12 appeared to be specifically stable at locations in the REZ, indicating a different expression pattern to *MC-RAP2.12*. The stability of RAP2.12 in apical meristem tissues in the shoot (Abbas et al., 2015) and now also the root after hypoxia suggests that these tissues may form specialised oxygen sensory zones in the plant. Recently it has been discovered that the oxygen sensing mechanism of RAP2.12 in the roots is linked to potassium uptake (Shahzad et al., 2016), it may therefore be possible that the RAM is also an important nutrient sensor in roots. As the experiments in this study were conducted on ½ MS plates the roots grew along the surface of the media and had constant access to environmental oxygen, it would be interesting to see if GUS from *promRAP2.12::MC-RAP2.12-GUS* seedlings would be visible in the RAM in plants grown in soil, in an environment where oxygen availability would be reduced. Perhaps in these conditions RAP2.12 would be stabilised more because of reduced oxygen availability.

A GUS reporter protein carrying the seven residue Nt motif conserved in ERFVIIs (MCGGAIL) has previously been observed as an *in vivo* sensor of hypoxia in barley (Mendiondo et al., 2016). Using the same construct this study found that GUS protein derived from this construct can also sense oxygen in Arabidopsis through its Nt-Cys, these findings are congruent with Nt-Cys oxygen sensing previously described as a component of the Arg/N-end rule pathway (Gibbs et al., 2011; Licausi et al., 2011). In accordance with the results from Mendiondo et al. (2016) Arabidopsis seedlings showed stabilisation of GUS in the hypocotyl but also in the first true leaves. The high level of accumulated GUS in the *Ubi1::MAGGAIL-GUS* seedlings was visible throughout the plant,

therefore the absence of visible staining in the cotyledons and root of *Ubi1::MCGGAIL-GUS* suggests that although GUS is expressed it does not accumulate. Therefore the much increased stability of MAGGAIL-GUS is because of the single amino acid change of Cys-2 to Ala-2, supporting previous work where untreated MAGGAIL-GUS expressing barley embryos was stabilised compared to untreated MCGGAIL-GUS (Gibbs et al., 2014b).

Although the results found in this study demonstrate that the degradation of MCGGAIL-GUS is dependent on oxygen and a functioning 26S proteasome, the role of the proceeding residues GGAIL after Cys-2 has not been identified. To investigate whether the residues GGAIL have any influence on the stability of the MCGGAIL-GUS protein, a new construct could be made with these residues omitted (e.g. *Ubi1::MC-GUS*) or with the substitution of these for different residues, so that the significance of these residues and their effect on expression/stabilisation pattern could be examined. It could be possible that the residues proceeding the Nt-Cys of ERFVIIs have a function coordinating the oxidation of Cys, the recognition of a substrate by an E3 ligase or the Ub attachment to a nearby Lys side chain.

GUS protein in transgenic Arabidopsis reporter lines containing a *35S::MC-GUS* construct was also demonstrated to be an oxygen sensor in this study. Results from this study identify a new, intermediate expressing, *35S::MC-GUS* Arabidopsis line (A25) which is more sensitive to oxygen than existing A2/1 lines. In addition to hypoxia, stabilisation of GUS from *35S::MC-GUS* seedlings was caused by inhibition of the 26S

proteasome using bortezomib, matching what has previously been observed with RGS4 treated with MG115 (Davydov and Varshavsky, 2000) The accumulation of GUS from *35S::MC-GUS* seedlings was strongest in the root unlike the *Ubi1::MCGGAIL-GUS* which showed stronger GUS accumulation in shoot material. Although GUS from both constructs was accumulated in response to hypoxia, the pattern of GUS staining is almost opposite; for example in the root tip where GUS localisation in the REZ, RAM and root cap are different. The reason for this difference is not clear, although further experiments should investigate the RNA stability of both constructs to check they are similar. If not this it may be a result of their different promoters or Nt sequences. To definitively distinguish between the effects of the sequence on the behaviour of the construct *in vivo* they would have to be expressed under the same promoter.

6.2 PTMs occurring to *in vivo* GUS reporter protein and its digestion efficiency

The findings of this study show that the GUS reporter protein from *35S::MC-GUS* in Col-0 Arabidopsis seedlings is likely to be ubiquitinated *in vivo* (Figure 4.2). The absence of the E3 ligase PRT6 prevents ubiquitination of affinity purified GUS protein, and correlates well with the increased stability of this protein (Figure 4.1). In addition to this the increased stability of the GUS reporter protein when treated with bortezomib, inhibiting the 26S proteasome, provides indirect evidence that this protein is a substrate for the 26S proteasome and thus requires

ubiquitination. These results with the GUS reporter protein also match those for the mammalian Cys-Arg/N-end rule pathway substrate RGS4 when treated with the proteasome inhibitor MG115. Davydov and Varshavsky (2000) showed increased stabilisation of this protein and an apparent multi-Ub ladder visible at higher MW. The smeared appearance of ubiquitinated protein released after purification of bortezomib treated *35S::MC-GUS* seedlings with anti-HA antibody beads, indicates that although Ub is present there has been some protein degradation occurring. This may be a result of some contaminants in the gel but interestingly the degradation would have to occur via another degradative pathway. Recently the degradative ability of the E3 ligase PRT1 has been demonstrated in real-time, Mot et al. (2016) have shown the sequential generation of a Ub-ladder using *in vitro* substrate reporters with a fluorescent probe.

The MC-GUS protein used in this study contains a lac repressor sequence between the poly-Gly linker and the HA-tag, this linker contains an Lys-Arg-Lys sequence which has been shown to be the site of attachment of multi-ubiquitin to Ub-X-β-gal fusion proteins (Chau et al., 1989). The deletion of both of these two Lys residues from Ub-X-DHFR fusion proteins has also been shown to increase the half-life of the protein, due to the removal of the site of the attachment for Ub (Bachmair and Varshavsky, 1989). In the MC-GUS protein sequence these Lys residues lie 31 and 33 residues from the N-terminus and are the only Lys residues in the first 100 from the N-terminus. It is likely therefore that these are the site of ubiquitination *in vivo*. The tryptic peptide containing Lys-31 was identified from *35S::MC-GUS* seedlings in the *prt6* background in

this study (Figure 4.13). In the *prt6* line ubiquitination does not occur and an unmodified Lys-31 was identified.

To unequivocally prove that ubiquitination occurs to the GUS reporter protein, LC-MS could be used to identify this PTM. It would be interesting and informative to affinity purify protein from *35S::MC-GUS* seedlings in the Col-0 background treated with bortezomib. The GUS protein would be expected to be stabilised after ubiquitination, therefore if visible on a Coomassie gel this stabilised protein could be digested with trypsin and analysed by LC-MS. If the Lys-31 residue had been ubiquitinated the mass of the non-ubiquitinated peptide would have increased, the expected mass of different length polyubiquitin chains could be defined and the MS data manually searched for a corresponding ion peak. Subsequent MS/MS analysis of a ubiquitinated tryptic peptide would reveal a characteristic K-GG motif, initially observed by Goldknopf and Busch (1977) when the C-terminal diglycine of Ub was found to extend from the Lys-119 of the Histone substrate.

The detection of sumoylated protein released from affinity purified protein of *35S::MC-GUS* seedlings, indicated that sumoylation does occur to HA-purified GUS reporter protein in a treatment unspecific manner. This experiment should be repeated with a higher concentration of anti-SUMO antibody, the results of this study show that no sumolyated protein was detected in the crude protein extract. Sumoylation has to be considered as a potential PTM which could perturb the degradation of substrates of the N-end rule pathway, through competing with Ub for free Lys side chains or modifying the activity of interacting components

(e.g. MAPs, ATEs, E3s etc.), but may also promote degradation as demonstrated in Crozet et al (2016) for SnRK1. More research is required to elucidate the interaction between sumoylation and the N-end rule pathway.

Results from this study show that excised GUS reporter protein bands from *35S::MC-GUS* seedlings sent for Edman degradation returned no identifiable protein sequences. This was strange as affinity purified protein bands sharing the same profile have been identified as belonging to the GUS reporter protein by LC-MS/MS. Therefore presuming that the experiment was performed successfully it is likely that the N-terminus of the GUS reporter protein could be blocked. This could also explain why the Nt Asp-N peptide has not been identified, as this unknown modification would alter its predicted mass. It is therefore also currently not clear when this 'blocking' would occur during Nt processing, although due to evidence obtained in this study that ubiquitination occurs to affinity purified GUS protein, it is likely after the arginylation of an oxidised Cys. However it is unlikely that this blocking could have occurred during protein extraction and processing, as no PTM's other than carbamidomethylation were observed when the Nt-Cys synthetic peptide RCR-49 was reductively alkylated and digested with Asp-N. Performing Edman degradation with the CR-40 synthetic peptide used in this study would be an interesting next step, the peptide could be sequenced with Nt-Cys unoxidised and then oxidised to Cys-sulfonic acid or Cys-sulfonamide to determine whether blocking can be caused by the oxidation of Nt-Cys.

It would be interesting to use some of the various Nt 'unblocking' techniques prior to Edman degradation, in an attempt to successfully identify the Nt residues of the GUS reporter protein. Numerous unblocking protocols have been outlined by Leone et al. (2011), these are generally chemical or enzymatic and are specific to the blocking group. It is important here to remember that one of the most common Nt-PTMs in plants is acetylation and is expected to occur at a similar extent to humans (Bienvenut et al., 2012), where total or partial Nt Ac-proteins are estimated at over 80% of all proteins (Arnesen et al., 2009). Nt-Ac proteins can be substrates for the Ac branch of the N-end rule pathway (Hwang et al., 2010), although the Nt-acetylation of the GUS reporter protein is unlikely given that this protein was shown to be stabilised in the absence of PRT6. To determine whether this apparent blocking of the N-terminus of GUS is related to the modifications occurring at the Nt-Cys, affinity purification and subsequent Edman degradation of protein derived from *35S::MA-GUS* seedlings would help elucidate more information. This is not a substrate of the N-end rule pathway, therefore if the Nt-residue generated by Edman sequencing was identified as Ala this would infer that the suspected blocking occurring to affinity purified protein of *35S::MC-GUS* is related to Cys-2.

Peptides from the GUS reporter protein were successfully identified after digesting protein affinity purified from *35S::MC-GUS* seedlings. Extracting protein from *35S::MC-GUS* seedlings in *prt6* was expected to bear RC_{ox} N-terminus after Arg/N-end rule processing. Interestingly the digestion of the GUS reporter protein was most efficient with Asp-N and not trypsin, which contradicts a recent study looking at the digestion

efficiency of numerous proteases on bovine serum albumin (BSA), where coverage following trypsin digestion was 78% compared to 38% for Asp-N (Giansanti et al., 2016). The difference in digestion efficiency between trypsin and Asp-N for the GUS reporter could be related to sequence of the GUS reporter. Over 50% of Arg and Lys cleavage sites of the GUS reporter sequence are in close proximity to negatively charged residues e.g. Asp, Glu, Ser, Tyr or other cleavage sites (Arg and Lys), which often results in missed cleavages (Giansanti et al., 2016). However this issue is not known for Lys-C so it remains unclear why this protease performed inefficiently. In an attempt to create a more amenable sized Nt peptide, an alternative to Lys-C is Lys-N, which could be investigated in future studies.

6.3 Possible PTMs occurring on the missing Nt-Asp-N peptide of the GUS reporter protein using *in vitro* RCR-49 synthetic peptide

Utilising a 49 residue synthetic peptide (RCR-49) modelling the predicted N-terminal peptide of the GUS reporter protein generated by Asp-N digestion, this study presents the expected charge states and MS/MS fragmentation pattern of the RCR-49 synthetic peptide. This data would help detect the Asp-N digested *in vivo* GUS Nt-peptide from *35S::MC-GUS* seedlings because the sequences should be identical. Using the RCR-49 peptide enabled the discovery that other than alkylation, no additional PTMs would be anticipated at the N-terminus of the *in vivo* GUS reporter protein during reduction, alkylation and digestion with Asp-N. Using these findings a list of potential masses for the *in vivo* Nt GUS

protein after Asp-N digestion were made, based on the residues 3 to 49 remaining unmodified and then the possible Nt PTMs; the presence/absence of the Nt Arg, the unknown variable oxidative modification to the alkylated Cys-2 and the possible acetylation of Nt-Arg.

The ion peaks in the MS data for the *in vivo* GUS reporter protein did not match any of the expected masses for a modified *in vivo* Nt-Asp-N peptide (Table 4.1 and 4.2). Therefore the *in vivo* Cys oxidation state required for arginylation and subsequent degradation remains unknown. From the digestion studies of the GUS reporter protein presented in this study, the probability that the Nt-Cys and Nt-Asp-N peptide of the GUS reporter protein is in the pool of Asp-N digested peptides is high. Evidence of alkylation being the sole PTM recorded for RCR-49, combined with the identification of the peptide belonging to the GUS reporter protein produced by Asp-N cleavage at the first and second site, provides indirect evidence that the Nt peptide is present in the digestion mix and that the PTM(s) which have occurred have likely occurred *in vivo*.

6.4 Oxidation and nitrosylation of Nt-Cys CR-40 synthetic peptide *in vitro* suggests numerous possible oxidative routes *in vivo*

The forcible oxidation and nitrosylation of Nt-Cys synthetic peptides *in vitro* using H_2O_2 and the NO donor SNAP, were used to predict how the Nt-Cys of N-end rule substrates may oxidise *in vivo*. The work presented highlights how oxidation occurs differently in the two Nt-Cys peptides used, with the CR dipeptide following the Cys-sulfenic to Cys-sulfonic

acid route, and the CR-40 peptide forming a protective Cys-sulfenamide species possibly preventing irreversible oxidation to Cys-sulfonic acid. The formation of Cys-sulfenamide at the Nt-Cys of CR-40 is likely a result of the nucleophilic attack of the backbone N atom of the proceeding Arg-2 residue onto the electrophilic S of the Cys-sulfenic acid. The results presented for the oxidation of the Nt-Cys of CR-40 match what was observed with protein tyrosine phosphatase 1B (PTP1B), where Cys in the enzymatic catalytic active site formed a sulfenamide species when exposed to H_2O_2 (Salmeen et al., 2003). Prolonged or high exposure of the PTP1B to H_2O_2 caused the irreversible formation of Cys-sulfonic acid (Salmeen et al., 2003), which is what was observed with CR-40 in this study.

This is the first report, in regards to known reactive Cys biology, thus studies of the oxidation of Nt-Cys synthetic peptides in the N-end rule pathway, where oxidation of Nt-Cys apparently occurs via a sulfenamide route. Previous studies using synthetic peptides as a proxy to investigate the oxidation state of Nt-Cys required for arginylation in the N-end rule pathway, have found that Cys-sulfinic/sulfonic acid was the Cys oxidative modification. Hu et al. (2005) used eight residue Nt-Cys synthetic peptides to show that successful arginylation by mammalian ATE1 requires the Nt-Cys to be oxidised to the Cys-sulfonic acid. In plants the PCO enzymes were discovered as dioxygenases which can catalyse the conversion of Cys to Cys sulfinic acid (Weits et al., 2014). Synthetic peptides bearing the ten Nt residues of three ERFVIIs; RAP2.12, RAP2.2 and HRE2 (CGGAIISDFI), exposed to either PCO1 or PCO4 show that the Nt-Cys oxidises to Cys-sulfinic acid (White et al., 2016). The Cys-sulfinic

acid formed after PCO exposure was then found to be successfully arginylated by plant ATE1 along with peptides bearing Nt-Asp and Nt-sulfonic acid (White et al., 2016). It therefore appears that acidic N-termini are substrates for arginylation *in vitro*. However, it is yet unproven whether ATE could successfully arginylate a Cys-sulf(en/in/on)amide species.

The determining factor promoting the different oxidative routes observed with Nt-Cys synthetic peptides remains unknown, although it is likely related to their respective size and sequence composition. Not only did the CR dipeptide and CR-40 produce different oxidative products, their reactivity was also markedly different with CR-40 oxidising faster with H_2O_2 than the dipeptide. It is interesting to speculate how the length of the amino acid residue sequence proceeding the Nt-Cys of CR-40 could alter its reactivity? Previously it has been shown that ubiquitination efficiency of X-β-gal fusion proteins appears to depend on the sequence length between the destabilising N-terminus and free Lys residue (Bachmair and Varshavsky, 1989; Chau et al., 1989). Bachmair and Varshavsky (1989) reported that there was some protein half-life relationship between the distance between a destabilising Nt-residue and a free Lys site for ubiquitination, with increased degradation rates the longer the unstructured sequence between them, which was further increased when a poly-Gly linker was near the N-terminus. Therefore due to the disordered poly-Gly and lac repressor sequence of the *in vivo* MC-GUS protein used in this study, the half-life could be expected to be less than other native Nt-Cys substrates such as the ERFVIIs. This could explain why no ubiquitinated affinity purified GUS protein was detected in

the Col-0 background (Figure 4.2) despite the identification of GUS using anti-HA antibody (Figure 4.1).

In addition to affecting its reactivity it would be interesting to discover if the sequences proceeding the Nt-Cys residue of CR-40 also influence the oxidative products formed, and may this explain why CR-40 followed the Cys-sulfenamide oxidative route? In relation to this Bachmair and Varshavsky (1989) demonstrated that changes in extension length closer to the N-terminus had a greater impact on degradation rate of fusion proteins, indicating that residues in the vicinity of the N-terminus may hold greater influence on protein stability. Later Davydov and Varshavsky (2000) mutated the Lys residue proceeding the Nt-Cys (after Met excision) of the Met-Cys initiating protein RGS4, if Lys was replaced with Arg (another basic residue) no change in the degradability of RGS4 occurred *in vitro*, whilst if Ser was present RGS4 was stabilised. A recent study has found that residues on the C-terminal side of an Nt destabilising residue can influence the arginylation rate of Nt-Cys (Wadas et al., 2016). If a basic Lys residue follows Nt-Cys in the presence of ATE some evidence of arginylation was observed, although it remains unclear if the presence of Lys actually increases arginylation efficiency or makes the Nt-Cys more susceptible to oxidation in air (Wadas et al., 2016). A further interesting finding of this study highlighted how if a Lys residue lies C-terminal of Nt destabilising Asp, the rate of degradation of fusion proteins was increased compared to if either Asp or Pro were the following residue (Wadas et al., 2016).

The effect of the residues in the vicinity of a destabilising Nt residues has been a neglected area of study. As well as influencing the degradation rate of proteins, these Nt residues may also help coordinate oxidation of Nt-Cys either chemically or enzymatically (e.g. PCOs in plants). The oxidation of Nt-Cys of the conserved ERFVII Nt sequence by PCOs to Cys-sulfinic acid does not mean this is the case for all Nt-Cys substrates in the plant Arg/N-end rule pathway. The oxidative route for other Nt-Cys Arg/N-end rule substrate such as VERNALIZATION2 (VRN2), which does not contain the same Nt-sequence as the ERFVIIs (CRQNCRAKSS) remains unknown and it is currently unproven that PCOs can oxidise this protein. Therefore it may be possible that proteins such as VRN2 and MC-GUS in this study are oxidised via another route, which could involve reversible Cys-sulfenamide formation.

Reacting an NO donor with the CR-40 synthetic peptide caused the nitrosylation of Cys to Cys-nitrosothiol (Cys-SNO). Gibbs et al. (2014b) and Hu et al. (2005) highlighted that NO is required in conjunction with oxygen for the *in vivo* degradation of Nt-Cys Arg/N-end rule substrates such as the ERFVIIs and RGS proteins respectively. Exposing seedlings expressing individual ERFVII proteins or the MC-GUS construct to the NO scavenger cPTIO caused stabilisation, this was then reversed after subsequent exposure to the NO donor SNAP (Gibbs et al., 2014b). This discovery followed the finding that when exposed to an NO donor *in vitro*, recombinant RGS4 proteins bearing an Nt-Cys were efficiently arginylated (Hu et al., 2005). One function of NO sensing in mammals was identified by Jaba et al. (2013) who highlighted how RGS4, and its interaction with NO, regulates cardiovascular development in mice. These

studies provide the only knowledge we have on the role of NO in destabilising Nt-Cys Arg/N-end rule substrates in plants and mammals. In plants the interaction of NO and PCO enzymes has yet to be elucidated, with studies involving PCO driven oxidation of Nt-Cys overlooking the role of NO, as PCOs alone were sufficient to oxidise Nt-Cys *in vitro* (Weits et al., 2014; White et al., 2016). The lack of information regarding the role of NO in the destabilisation of Nt-Cys substrates of the Arg/N-end rule pathway leaves much uncertainty in the field. Currently it is not clear if NO interacts with oxygen facilitating a non-enzymatic oxidation mechanism, or if it is required to activate either an enzymatic oxidative step (e.g. PCOs) or the arginylation of oxidised Cys. In this study our results indicate that NO and oxidative species do not interact when oxidising Nt-Cys, incubating CR-40 with SNAP and then H_2O_2 or vice versa do not subsequently change the Cys-nitrosothiol or Cys-sulfenamide initially formed (Appendix G and H).

The NO source used in this study to nitrosylate the Nt-Cys was SNAP, an NO donor which intramolecularly donates NO rather than releasing NO into solution. When SNP was used no nitrosylation of CR-40 was observed, however this should be repeated with the reaction taking place in an anoxic environment. The release of NO from SNP would be more analogous to NO production *in vivo*, a hypothetical route for a potential NO-induced destabilisation of an Nt-Cys residue is outlined in Figure 6.1.

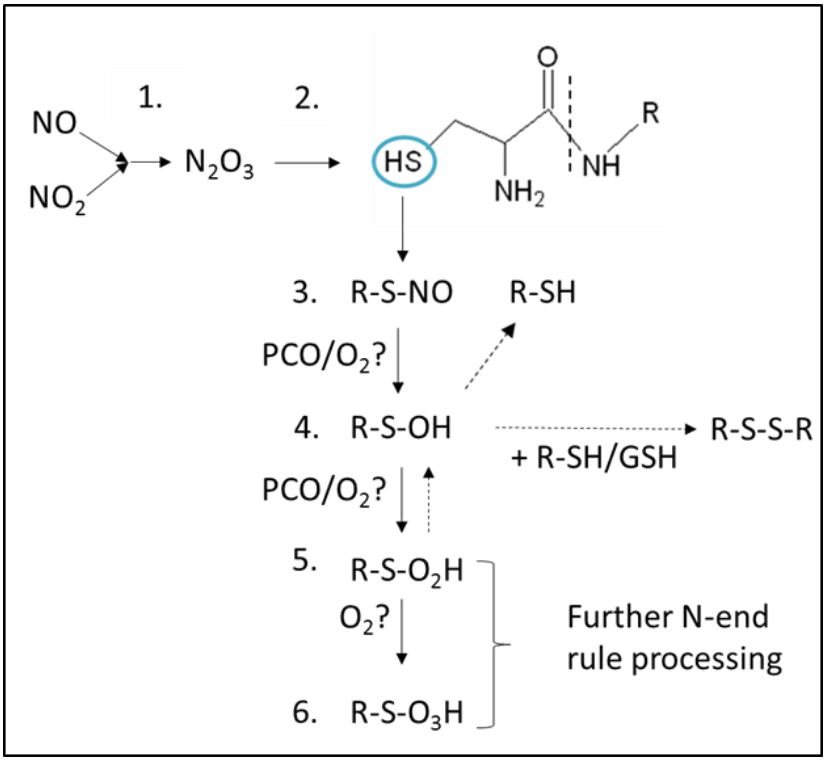


Figure 6.1 Schematic of a hypothetical nitrosylation-driven degradation mechanism for an *in vivo* Nt-Cys Arg/N-end rule substrate. 1. Dinitrogen trioxide (N_2O_3) formed from reacting nitric oxide and nitrogen dioxide; 2. Electrophillic N_2O_3 reacts with the thiol (SH) group on the Cys causing deprotonation and forming a S-nitrosothiol (R-S-NO); 3. The R-S-NO is oxidised via a combination of Plant Cysteine Oxidase (PCO) and/or oxygen to Cys-sulfeinc acid (R-S-OH); 4. R-S-OH is unstable and may be reduced and return to a thiol group (R-SH), it may reach with another R-SH or glutathione (GSH) forming a stable disulfide bond (R-S-S-R), or it could oxidise further to Cys-sulfinic acid (R-S-O_2H); 5. R-S-O_2H can undergo further modifications in the N-end rule pathway, or can be reduced back to R-S-OH by peroxiredoxins or it can be oxidised to Cys-sulfonic acid (R-S-O_3H); 6. Generating R-S-O_3H is an irreversible reaction and this species indicates oxidative stress, however this could possibly undergo further N-end rule processing.

6.5 Conclusions and Further work

The aim of this study was to investigate the oxidation of Nt-Cys using both *in vivo* protein substrates of the Arg/N-end rule pathway and *in vitro* synthetic peptides. In this study new transgenic lines of Arabidopsis expressing an Nt-Cys GUS protein which sense oxygen via the Arg/N-end rule pathway were developed. Using transgenic Arabidopsis lines expressing RAP2.12 coupled to GUS driven by the RAP2.12 promoter, this study discovered that RAP2.12 is stabilised in both the root and shoot apical meristem by hypoxia, indicating that meristem cells are specialised oxygen sensory zones in plants. The results of this study show the successful affinity purification and identification of the Nt-Cys GUS protein reporter protein using LC-MS/MS, and presents strong evidence that this protein is ubiquitinated *in vivo* before degradation via the Arg/N-end rule pathway.

Using a synthetic peptide bearing an Nt-Cys sequence representing that of the *in vivo* GUS protein, indirect evidence is presented indicating that unexpected PTMs occur *in vivo* preventing the identification of the Nt-Cys oxidation state. This study is also the first to report the non-enzymatic oxidation of Nt-Cys of a synthetic peptide to a Cys-sulfenamide, a species previously reported as a reversible protective modification *in vivo*, if occurring to *in vivo* Nt-Cys substrates of the Arg/N-end rule pathway it is unknown if this oxidative route would be a protective stabilising or destabilising degradative mechanism. Finally exposing Nt-Cys synthetic peptides to H_2O_2 then NO and vice versa, ESI-MS results show no interaction between the nitrosylated Cys and H_2O_2 or oxidised Cys and NO. This raises further uncertainty as to how NO is involved in the

oxidation of Nt-Cys, therefore suggests that NO is likely involved in the enzymatic activation of PCOs or the arginylation of Nt-Cys by ATE.

The impact of the results presented in this study are important to the reactive Cys biology field and for the future development of flood tolerant crops. The discovery that Nt-Cys of a synthetic peptide can be oxidised to Cys-sulfenamide is unreported in the literature and definitely a first in relation to the N-end rule pathway. The protective effect of Cys-sulfenamide *in vivo* could prevent the irreversible oxidation of Nt-Cys to Cys-sulfonic acid, therefore this could maintain the functionality of the protein under oxidative stress. If the Nt-Cys of an *in vivo* protein can be oxidised to Cys-sulfenamide the effect of this would have to be discovered across all fields related to reactive Cys biology and chemistry. In relation to the plant N-end rule pathway, the formation of Cys-sulfenamide at Nt-Cys of Arg/N-end rule substrates would have to be investigated to determine if this is a degradative or stabilising oxidative signal by incubating this species with ATE. The discovery that the SAM and RAM of seedlings appears to be a sensory zone for oxygen, and consequently RAP2.12 stabilisation upon hypoxia, identifies the meristematic regions of plants as regions which deserve further study in regards to discovering other Nt-Cys substrates of the Arg/N-end rule, and thus resources which may be used in developing flood tolerant crops.

Further work arising from the results presented in this study should include a proteomic screen of the affinity purified Nt-Cys GUS protein digested with Asp-N. The 49 residue Nt-peptide predicted to be generated should be screened so all possible PTMs to each individual

residue are searched, this would include sumoylation and Nt-acetylation identified as potential modification in this study. Unexpected cleavages from the N-terminus or C-terminus would also have to be accounted for, therefore a database of potential masses corresponding to variable oxidising states of Nt-Cys, with all potential PTMs to each of the 49 residues would have to be searched against ion peaks from a full LC-MS scan which would hopefully find a match. To determine what oxidised Cys species are substrates for ATE-driven arginylation, they could be incubated with ATE. Using the Nt-Cys of the CR-40 synthetic peptide various oxidised and nitrosylated Nt-Cys species could be formed using methods presented in this study. Incubating the synthetic peptide species with ATE and then subsequently analysing the peptide after incubation, would indicate what oxidised species are suitable substrates for ATE when analysed by LC-MS and a mass addition corresponding to the mass of Arg added to the original oxidised species.

7. References

Abbas, M., Berckhan, S., Rooney, Daniel J., Gibbs, Daniel J., Vicente Conde, J., Sousa Correia, C., Bassel, George W., Marín-de la Rosa, N., León, J., Alabadí, D.*, et al.* (2015). Oxygen Sensing Coordinates Photomorphogenesis to Facilitate Seedling Survival. *Current Biology 25*, 1483-1488.

Airaki, M., Leterrier, M., Valderrama, R., Chaki, M., Begara-Morales, J.C., Barroso, J.B., del Río, L.A., Palma, J.M., and Corpas, F.J. (2015). Spatial and temporal regulation of the metabolism of reactive oxygen and nitrogen species during the early development of pepper (Capsicum annuum) seedlings. *Annals of Botany 116*, 679-693.

Akter, S., Huang, J., Waszczak, C., Jacques, S., Gevaert, K., Van Breusegem, F., and Messens, J. (2015). Cysteines under ROS attack in plants: a proteomics view. *Journal of Experimental Botany 66*, 2935-2944 .

Alderton, W., Cooper, C., and Knowles, R. (2001). Nitric oxide synthases: structure, function and inhibition. *Biochemical Journal 357*, 593-615.

Alpert, A.J. (1990). Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds. *Journal of Chromatography A 499*, 177-196.

Armstrong, W., and Drew, M. (2002). Root growth and metabolism under oxygen deficiency. Plant roots: the hidden half *3*, 729-761.

Arnesen, T., Van Damme, P., Polevoda, B., Helsens, K., Evjenth, R., Colaert, N., Varhaug, J.E., Vandekerckhove, J., Lillehaug, J.R., Sherman, F.*, et al.* (2009). Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. *Proceedings of the National Academy of Sciences 106*, 8157-8162.

Astier, J., Kulik, A., Koen, E., Besson-Bard, A., Bourque, S., Jeandroz, S., Lamotte, O., and Wendehenne, D. (2012). Protein S-nitrosylation: what's going on in plants? *Free Radical Biology and Medicine 53*, 1101-1110.

Aston, F.W. (1919). LXXIV. A positive ray spectrograph. *Philosophical Magazine Series 6 38*, 707-714.

Bachmair, A., Becker, F., and Schell, J. (1993). Use of a reporter transgene to generate arabidopsis mutants in ubiquitin-dependent protein degradation. *Proceedings of the National Academy of Sciences 90*, 418-421.

Bachmair, A., Finley, D., and Varshavsky, A. (1986). In vivo half-life of a protein is a function of its amino-terminal residue. *Science 234*, 179-186.

Bachmair, A., Novatchkova, M., Potuschak, T., and Eisenhaber, F. (2001). Ubiquitylation in plants: a post-genomic look at a post-translational modification. *Trends in Plant Science 6*, 463-470.

Bachmair, A., and Varshavsky, A. (1989). The degradation signal in a short-lived protein. *Cell 56*, 1019-1032.

Baker, R.T., and Varshavsky, A. (1995). Yeast N-terminal Amidase: A NEW ENZYME AND COMPONENT OF THE N-END RULE PATHWAY. *Journal of Biological Chemistry 270*, 12065-12074.

Balzi, E., Choder, M., Chen, W.N., Varshavsky, A., and Goffeau, A. (1990). Cloning and functional analysis of the arginyl-tRNA-protein transferase gene ATE1 of Saccharomyces cerevisiae. *Journal of Biological Chemistry 265*, 7464-7471.

Banerjee, S., and Mazumdar, S. (2012). Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *International Journal of Analytical Chemistry 2012*, 40.

Bartberger, M.D., Houk, K.N., Powell, S.C., Mannion, J.D., Lo, K.Y., Stamler, J.S., and Toone, E.J. (2000). Theory, Spectroscopy, and Crystallographic Analysis of S-Nitrosothiols: Conformational Distribution Dictates Spectroscopic Behavior. *Journal of the American Chemical Society 122*, 5889-5890.

Bartel, B., and Citovsky, V. (2012). Focus on Ubiquitin in Plant Biology. *Plant Physiology 160*, 1.

Bartel, B., Wünning, I., and Varshavsky, A. (1990). The recognition component of the N-end rule pathway. *The EMBO Journal 9*, 3179.

Bartlett, J.G., Alves, S.C., Smedley, M., Snape, J.W., and Harwood, W.A. (2008). High-throughput Agrobacterium-mediated barley transformation. *Plant Methods 4*, 1-12.

Beligni, M.V., and Lamattina, L. (2000). Nitric oxide stimulates seed germination and de-etiolation, and inhibits hypocotyl elongation, three light-inducible responses in plants. *Planta 210*, 215-221.

Ben-Bassat, A., Bauer, K., Chang, S.Y., Myambo, K., Boosman, A., and Chang, S. (1987). Processing of the initiation methionine from proteins: properties of the Escherichia coli methionine aminopeptidase and its gene structure. *Journal of Bacteriology 169*, 751-757.

Benfey, P.N., and Chua, N.-H. (1990). The cauliflower mosaic virus 35S promoter: combinatorial regulation of transcription in plants. *Science 250*, 959-966.

Bethke, P.C., Libourel, I.G., Aoyama, N., Chung, Y.Y., Still, D.W., and Jones, R.L. (2007). The Arabidopsis aleurone layer responds to nitric oxide, gibberellin, and abscisic acid and is sufficient and necessary for seed dormancy. *Plant Physiology 143*, 1173-1188.

Bienvenut, W.V., Sumpton, D., Martinez, A., Lilla, S., Espagne, C., Meinnel, T., and Giglione, C. (2012). Comparative Large Scale Characterization of Plant versus Mammal Proteins Reveals Similar and Idiosyncratic N-α-Acetylation Features. *Molecular & Cellular Proteomics 11*.

Bier, M.E., Winkler, P.C., and Herron, J.R. (1993). Coupling a particle beam interface directly to a quadrupole ion trap mass spectrometer. *Journal of the American Society for Mass Spectrometry 4*, 38-46.

Birnbaum, K., Shasha, D.E., Wang, J.Y., Jung, J.W., Lambert, G.M., Galbraith, D.W., and Benfey, P.N. (2003). A Gene Expression Map of the Arabidopsis Root. *Science 302*, 1956-1960.

Branco-Price, C., Kawaguchi, R., Ferreira, R.B., and Bailey-Serres, J. (2005). Genome-wide analysis of transcript abundance and translation in Arabidopsis seedlings subjected to oxygen deprivation. *Annals of Botany 96*, 647-660.

Brooks, D.J. (2001). Increased Frequency of Cysteine, Tyrosine, and Phenylalanine Residues Since the Last Universal Ancestor. *Molecular & Cellular Proteomics 1*, 125-131.

Budzikiewicz, H., Djerassi, C., Williams, D.H., and Williams, D.H. (1964). Structure elucidation of natural products by mass spectrometry, Vol 2 (Holden-Day San Francisco).

Bui, L.T., Giuntoli, B., Kosmacz, M., Parlanti, S., and Licausi, F. (2015). Constitutively expressed ERF-VII transcription factors redundantly activate the core anaerobic response in Arabidopsis thaliana. *Plant Science 236*, 37-43.

Callis, J., and Bedinger, P. (1994). Developmentally regulated loss of ubiquitin and ubiquitinated proteins during pollen maturation in maize. *Proceedings of the National Academy of Sciences 91*, 6074-6077.

Chang, T.-S., Jeong, W., Woo, H.A., Lee, S.M., Park, S., and Rhee, S.G. (2004). Characterization of Mammalian Sulfiredoxin and Its Reactivation of Hyperoxidized Peroxiredoxin through Reduction of Cysteine Sulfinic Acid in the Active Site to Cysteine. *Journal of Biological Chemistry 279*, 50994-51001.

Chau, V., Tobias, J., Bachmair, A., Marriott, D., Ecker, D., Gonda, D., and Varshavsky, A. (1989). A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science 243*, 1576-1583.

Christensen, A.H., Sharrock, R.A., and Quail, P.H. (1992). Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Molecular Biology 18*, 675-689.

Chung, H.S., Wang, S.-B., Venkatraman, V., Murray, C.I., and Van Eyk, J.E. (2013). Cysteine Oxidative Posttranslational Modifications: Emerging Regulation in the Cardiovascular System. *Circulation Research 112*, 382-392.

Ciechanover, A. (2005). Intracellular protein degradation: from a vague idea thru the lysosome and the ubiquitin-proteasome system and onto human diseases and drug targeting[ast]. *Cell Death and Differentiation 12*, 1178-1190.

Ciechanover, A., Elias, S., Heller, H., and Hershko, A. (1982). "Covalent affinity" purification of ubiquitin-activating enzyme. *Journal of Biological Chemistry 257*, 2537-2542.

Ciechanover, A., Heller, H., Elias, S., Haas, A.L., and Hershko, A. (1980). ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. *Proceedings of the National Academy of Sciences of the United States of America 77*, 1365-1368.

Ciechanover, A., Heller, H., Katz-Etzion, R., and Hershko, A. (1981). Activation of the heat-stable polypeptide of the ATP-dependent proteolytic system. *Proceedings of the National Academy of Sciences of the United States of America 78*, 761-765.

Ciehanover, A., Hod, Y., and Hershko, A. (1978). A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. *Biochemical and Biophysical Research Communications 81*, 1100-1105.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method forAgrobacterium-mediated transformation ofArabidopsis thaliana. *The Plant Journal 16*, 735-743.

Considine, M.J., and Foyer, C.H. (2014). Redox Regulation of Plant Development. *Antioxidants & Redox Signaling 21*, 1305-1326.

Corpas, F., Barroso, J., Carreras, A., Valderrama, R., Palma, J., León, A., Sandalio, L., and del Río, L. (2006). Constitutive arginine-dependent nitric oxide synthase activity in different organs of pea seedlings during plant development. *Planta 224*, 246-254.

Corpas, F.J., Hayashi, M., Mano, S., Nishimura, M., and Barroso, J.B. (2009). Peroxisomes are required for in vivo nitric oxide accumulation in the cytosol following salinity stress of Arabidopsis plants. *Plant Physiology 151*, 2083-2094.

Corpas, F.J., Leterrier, M., Valderrama, R., Airaki, M., Chaki, M., Palma, J.M., and Barroso, J.B. (2011). Nitric oxide imbalance provokes a nitrosative response in plants under abiotic stress. *Plant Science 181*, 604-611.

Crozet, P., Margalha, L., Butowt, R., Fernandes, N., Elias, C.A., Orosa, B., Tomanov, K., Teige, M., Bachmair, A., Sadanandom, A.*, et al.* (2016).

SUMOylation represses SnRK1 signaling in Arabidopsis. *The Plant Journal 85*, 120-133.

Davydov, I.V., and Varshavsky, A. (2000). RGS4 Is Arginylated and Degraded by the N-end Rule Pathway in Vitro. *Journal of Biological Chemistry 275*, 22931-22941.

de Duve, C., Pressman, B.C., Gianetto, R., Wattiaux, R., and Appelmans, F. (1955). Tissue fractionation studies. 6. Intracellular distribution patterns of enzymes in rat-liver tissue. *Biochemical Journal 60*, 604-617.

Dempster, A.J. (1918). A new Method of Positive Ray Analysis. *Physical Review 11*, 316-325.

Desterro, J.M., Rodriguez, M.S., and Hay, R.T. (1998). SUMO-1 modification of IκBα inhibits NF-κB activation. *Molecular cell 2*, 233-239.

Dohmen, R.J., Madura, K., Bartel, B., and Varshavsky, A. (1991). The N-end rule is mediated by the UBC2 (RAD6) ubiquitin-conjugating enzyme. *Proceedings of the National Academy of Sciences 88*, 7351-7355.

Dole, M., Mack, L.L., Hines, R.L., Mobley, R.C., Ferguson, L.D., and Alice, M.B. (1968). Molecular Beams of Macroions. *The Journal of Chemical Physics 49*, 2240-2249.

Drapeau, G.R. (1980). Substrate specificity of a proteolytic enzyme isolated from a mutant of Pseudomonas fragi. *Journal of Biological Chemistry 255*, 839-840.

Edman, P. (1950). Method for determination of the amino acid sequence in peptides. *Acta Chemica Scandinavica 4*, 34.

El-Maarouf-Bouteau, H., and Bailly, C. (2008). Oxidative signaling in seed germination and dormancy. *Plant Signaling & Behavior 3*, 175.

Engelhardt, H. (2012). High performance liquid chromatography (Springer Science & Business Media).

Erbse, A., Schmidt, R., Bornemann, T., Schneider-Mergener, J., Mogk, A., Zahn, R., Dougan, D., and Bukau, B. (2006). ClpS is an essential component of the N-end rule pathway in Escherichia coli. *Nature 439*, 753-756.

Espunya, M.C., De Michele, R., Gomez-Cadenas, A., and Martinez, M.C. (2012). S-Nitrosoglutathione is a component of wound- and salicylic acid-induced systemic responses in Arabidopsis thaliana. *Journal of Experimental Botany 63*, 3219-3227.

Etlinger, J.D., and Goldberg, A.L. (1977). A soluble ATP-dependent proteolytic system responsible for the degradation of abnormal proteins in reticulocytes. *Proceedings of the National Academy of Sciences of the United States of America 74*, 54-58.

Fares, A., Rossignol, M., and Peltier, J.B. (2011). Proteomics investigation of endogenous S-nitrosylation in Arabidopsis. *Biochemical and Biophysics Research Communications 416*, 331-336.

Finley, D., Özkaynak, E., Jentsch, S., McGrath, J.P., Bartel, B., Pazin, M., Snapka, R.M., and Varshavsky, A. (1988). Molecular Genetics of the Ubiquitin System. In Ubiquitin, M. Rechsteiner, ed. (Boston, MA: Springer US), pp. 39-75.

Fukao, T., Xu, K., Ronald, P.C., and Bailey-Serres, J. (2006). A Variable Cluster of Ethylene Response Factor–Like Genes Regulates Metabolic and Developmental Acclimation Responses to Submergence in Rice. *The Plant Cell 18*, 2021-2034.

Garzon, M., Eifler, K., Faust, A., Scheel, H., Hofmann, K., Koncz, C., Yephremov, A., and Bachmair, A. (2007). PRT6/At5g02310 encodes an

Arabidopsis ubiquitin ligase of the N-end rule pathway with arginine specificity and is not the CER3 locus. *FEBS Letter 581*, 3189-3196.

Gaston, B.M., Carver, J., Doctor, A., and Palmer, L.A. (2003). S-nitrosylation signaling in cell biology. *Molecular interventions 3*, 253.

Giansanti, P., Tsiatsiani, L., Low, T.Y., and Heck, A.J. (2016). Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. *Nature protocols 11*, 993-1006.

Gibbs, D.J., Bacardit, J., Bachmair, A., and Holdsworth, M.J. (2014a). The eukaryotic N-end rule pathway: conserved mechanisms and diverse functions. *Trends in Cell Biology 24,* 603-611

Gibbs, D.J., Lee, S.C., Isa, N.M., Gramuglia, S., Fukao, T., Bassel, G.W., Correia, C.S., Corbineau, F., Theodoulou, F.L., Bailey-Serres, J.*, et al.* (2011). Homeostatic response to hypoxia is regulated by the N-end rule pathway in plants. *Nature 479*, 415-418.

Gibbs, D.J., Md Isa, N., Movahedi, M., Lozano-Juste, J., Mendiondo, G.M., Berckhan, S., Marin-de la Rosa, N., Vicente Conde, J., Sousa Correia, C., Pearce, S.P.*, et al.* (2014b). Nitric oxide sensing in plants is mediated by proteolytic control of group VII ERF transcription factors. *Molecular Cell 53*, 369-379.

Giglione, C., Boularot, A., and Meinnel, T. (2004). Protein N-terminal methionine excision. *Cellular and Molecular Life Sciences CMLS 61*, 1455-1474.

Gil, P., Dewey, E., Friml, J., Zhao, Y., Snowden, K.C., Putterill, J., Palme, K., Estelle, M., and Chory, J. (2001). BIG: a calossin-like protein required for polar auxin transport in Arabidopsis. *Genes and Development 15*, 1985-1997.

Goebl, M.G., Yochem, J., Jentsch, S., McGrath, J.P., Varshavsky, A., and Byers, B. (1988). The yeast cell cycle gene CDC34 encodes a ubiquitin-conjugating enzyme. *Science 241*, 1331-1335.

Goldberg, A.L. (1971). A Role of Aminoacyl-tRNA in the Regulation of Protein Breakdown in Escherichia coli. *Proceedings of the National Academy of Sciences 68*, 362-366.

Goldberg, A.L. (1972). Degradation of abnormal proteins in Escherichia coli. *Proceedings of the National Academy of Sciences 69*, 422-426.

Goldknopf, I.L., and Busch, H. (1975). Remarkable similarities of peptide fingerprints of histone 2A and nonhistone chromosomal protein A24. *Biochemical and Biophysical Research Communications 65*, 951-960.

Goldknopf I.L., and Busch, H. (1977). Isopeptide linkage between nonhistone and histone 2A polypeptides of chromosomal conjugate-protein A24. *Proceedings of the National Academy of Sciences of the United States of America 74*, 864–868.

Goldstein, G. (1974). Isolation of bovine thymin: a polypeptide hormone of the thymus. *Nature 247*, 11-14.

Goldstein, G., Scheid, M., Hammerling, U., Schlesinger, D.H., Niall, H.D., and Boyse, E.A. (1975). Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proceedings of the National Academy of Sciences of the United States of America 72*, 11-15.

Graciet, E., Hu, R.-G., Piatkov, K., Rhee, J.H., Schwarz, E.M., and Varshavsky, A. (2006). Aminoacyl-transferases and the N-end rule pathway of prokaryotic/eukaryotic specificity in a human pathogen. *Proceedings of the National Academy of Sciences of the United States of America 103*, 3078-3083.

Graciet, E., Walter, F., Maoileidigh, D.O., Pollmann, S., Meyerowitz, E.M., Varshavsky, A., and Wellmer, F. (2009). The N-end rule pathway controls multiple functions during Arabidopsis shoot and leaf development. *Proceedings of the National Academy of Sciences of the United States of America 106*, 13618-13623.

Graciet, E., and Wellmer, F. (2010). The plant N-end rule pathway: structure and functions. *Trends in Plant Science 15*, 447-453.

Grigoryev, S., Stewart, A.E., Kwon, Y.T., Arfin, S.M., Bradshaw, R.A., Jenkins, N.A., Copeland, N.G., and Varshavsky, A. (1996). A Mouse Amidase Specific for N-terminal Asparagine: THE GENE, THE ENZYME, AND THEIR FUNCTION IN THE N-END RULE PATHWAY. *Journal of Biological Chemistry 271*, 28521-28532.

Groot, S., and Karssen, C. (1987). Gibberellins regulate seed germination in tomato by endosperm weakening: a study with gibberellin-deficient mutants. *Planta 171*, 525-531.

Guo, C., and Henley, J.M. (2014). Wrestling with stress: Roles of protein SUMOylation and deSUMOylation in cell stress response. *IUBMB Life 66*, 71-77.

Gupta, K.J., Fernie, A.R., Kaiser, W.M., and van Dongen, J.T. (2011). On the origins of nitric oxide. *Trends in Plant Science 16*, 160-168.

Gupta, K.J., Stoimenova, M., and Kaiser, W.M. (2005). In higher plants, only root mitochondria, but not leaf mitochondria reduce nitrite to NO, in vitro and in situ. *Journal of Experimental Botany 56*, 2601-2609.

Gupta, N., Hixson, K.K., Culley, D.E., Smith, R.D., and Pevzner, P.A. (2010). Analyzing protease specificity and detecting in vivo proteolytic events using tandem mass spectrometry. *Proteomics 10*, 2833-2844.

Gupta, V., and Carroll, K.S. (2014). Sulfenic acid chemistry, detection and cellular lifetime. *Biochimica et Biophysica Acta (BBA)-General Subjects 1840*, 847-875.

Haas, A.L., Warms, J., Hershko, A., and Rose, I.A. (1982). Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation. *Journal of Biological Chemistry 257*, 2543-2548.

Hall, C.N., and Garthwaite, J. (2009). What is the real physiological NO concentration in vivo? *Nitric Oxide 21*, 92-103.

Halliwell, B., and Whiteman, M. (2004). Measuring reactive species and oxidative damage in vivo and in cell culture: how should you do it and what do the results mean? *British Journal of Pharmacology 142*, 231-255.

Hare, J.M., and Stamler, J.S. (2005). NO/redox disequilibrium in the failing heart and cardiovascular system. *Journal of Clinical Investigation 115*, 509-517.

Hattori, Y., Nagai, K., Furukawa, S., Song, X.-J., Kawano, R., Sakakibara, H., Wu, J., Matsumoto, T., Yoshimura, A., and Kitano, H. (2009). The ethylene response factors SNORKEL1 and SNORKEL2 allow rice to adapt to deep water. *Nature 460*, 1026-1030.

He, Y., Tang, R.H., Hao, Y., Stevens, R.D., Cook, C.W., Ahn, S.M., Jing, L., Yang, Z., Chen, L., Guo, F.*, et al.* (2004). Nitric oxide represses the Arabidopsis floral transition. *Science 305*, 1968-1971.

Hemström, P., and Irgum, K. (2006). Hydrophilic interaction chromatography. *Journal of Separation Science 29*, 1784-1821.

Hendriks, I.A., and Vertegaal, A.C.O. (2016). A comprehensive compilation of SUMO proteomics. *Nature Reviews Molecular Cell Biology 17*, 581-595.

Hershko, A., and Ciechanover, A. (1998). THE UBIQUITIN SYSTEM. *Annual Review of Biochemistry 67*, 425-479.

Hershko, A., Ciechanover, A., Heller, H., Haas, A.L., and Rose, I.A. (1980). Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis. *Proceedings of the National Academy of Sciences 77*, 1783-1786.

Hershko, A., and Heller, H. (1985). Occurrence of a polyubiquitin structure in ubiquitin-protein conjugates. *Biochemical and Biophysical Research Communications 128*, 1079-1086.

Hershko, A., Heller, H., Elias, S., and Ciechanover, A. (1983). Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. *Journal of Biological Chemistry 258*, 8206-8214.

Hershko, A., Heller, H., Eytan, E., and Reiss, Y. (1986). The protein substrate binding site of the ubiquitin-protein ligase system. *Journal of Biological Chemistry 261*, 11992-11999.

Hershko, A., and Tomkins, G.M. (1971). Studies on the Degradation of Tyrosine Aminotransferase in Hepatoma Cells in Culture INFLUENCE OF THE COMPOSITION OF THE MEDIUM AND ADENOSINE TRIPHOSPHATE DEPENDENCE. *Journal of Biological Chemistry 246*, 710-714.

Hinz, M., Wilson, I.W., Yang, J., Buerstenbinder, K., Llewellyn, D., Dennis, E.S., Sauter, M., and Dolferus, R. (2010). Arabidopsis RAP2.2: an ethylene response transcription factor that is important for hypoxia survival. *Plant Physiology 153*, 757-772.

Hoernstein, S.N.W., Mueller, S.J., Fiedler, K., Schuelke, M., Vanselow, J.T., Schuessele, C., Lang, D., Nitschke, R., Igloi, G.L., Schlosser, A.*, et al.* (2016). Identification of Targets and Interaction Partners of Arginyl-

tRNA Protein Transferase in the Moss Physcomitrella patens. *Molecular & Cellular Proteomics 15*, 1808-1822.

Holman, T.J., Jones, P.D., Russell, L., Medhurst, A., Ubeda Tomas, S., Talloji, P., Marquez, J., Schmuths, H., Tung, S.A., Taylor, I.*, et al.* (2009). The N-end rule pathway promotes seed germination and establishment through removal of ABA sensitivity in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America 106*, 4549-4554.

Hough, R., Pratt, G., and Rechsteiner, M. (1986). Ubiquitin-lysozyme conjugates. Identification and characterization of an ATP-dependent protease from rabbit reticulocyte lysates. *Journal of Biological Chemistry 261*, 2400-2408.

Hu, R.G., Sheng, J., Qi, X., Xu, Z., Takahashi, T.T., and Varshavsky, A. (2005). The N-end rule pathway as a nitric oxide sensor controlling the levels of multiple regulators. *Nature 437*, 981-986.

Huber, R., and Bode, W. (1978). Structural basis of the activation and action of trypsin. *Accounts of Chemical Research 11*, 114-122.

Hunt, L.T., and Dayhoff, M.O. (1977). Amino-terminal sequence identity of ubiquitin and the nonhistone component of nuclear protein A24. *Biochemical and Biophysical Research Communications 74*, 650-655.

Hwang, C.S., Shemorry, A., and Varshavsky, A. (2010). N-terminal acetylation of cellular proteins creates specific degradation signals. *Science 327*, 973-977.

Jaba, I.M., Zhuang, Z.W., Li, N., Jiang, Y., Martin, K.A., Sinusas, A.J., Papademetris, X., Simons, M., Sessa, W.C., and Young, L.H. (2013). NO triggers RGS4 degradation to coordinate angiogenesis and cardiomyocyte growth. *The Journal of Clinical Investigation 123*, 1718.

Jacques, S., Ghesquière, B., Van Breusegem, F., and Gevaert, K. (2013). Plant proteins under oxidative attack. *Proteomics 13*, 932-940.

Jekel, P.A., Weijer, W.J., and Beintema, J.J. (1983). Use of endoproteinase Lys-C from Lysobacter enzymogenes in protein sequence analysis. *Analytical biochemistry 134*, 347-354.

Jentsch, S., McGrath, J.P., and Varshavsky, A. (1987). The yeast DNA repair gene RAD6 encodes a ubiquitin-conjugating enzyme. *Nature 329*, 131-134.

Jiang, K., Meng, Y.L., and Feldman, L.J. (2003). Quiescent center formation in maize roots is associated with an auxin-regulated oxidizing environment. *Development 130*, 1429-1438.

Job, C., Rajjou, L., Lovigny, Y., Belghazi, M., and Job, D. (2005). Patterns of protein oxidation in Arabidopsis seeds and during germination. *Plant Physiology 138*, 790-802.

Johansson, E., Olsson, O., and Nyström, T. (2004). Progression and specificity of protein oxidation in the life cycle of Arabidopsis thaliana. *Journal of Biological Chemistry 279*, 22204-22208.

Jordan, I.K., Kondrashov, F.A., Adzhubei, I.A., Wolf, Y.I., Koonin, E.V., Kondrashov, A.S., and Sunyaev, S. (2005). A universal trend of amino acid gain and loss in protein evolution. *Nature 433*, 633-638.

Jourd'heuil, D., Jourd'heuil, F.L., and Feelisch, M. (2003). Oxidation and Nitrosation of Thiols at Low Micromolar Exposure to Nitric Oxide: EVIDENCE FOR A FREE RADICAL MECHANISM. *Journal of Biological Chemistry 278*, 15720-15726.

Kaiser, P., and Huang, L. (2005). Global approaches to understanding ubiquitination. *Genome Biology 6*, 1-8.

Kaji, A., Kaji, H., and Novelli, G.D. (1965). Soluble Amino Acid-incorporating System: I. PREPARATION OF THE SYSTEM AND NATURE OF THE REACTION. *Journal of Biological Chemistry 240*, 1185-1191.

Kaji, H. (1968). Soluble amino acid incorporating system from rat liver. *Biochemistry 7*, 3844-3850.

Kanyuka, K., Praekelt, U., Franklin, K.A., Billingham, O.E., Hooley, R., Whitelam, G.C., Halliday, K.J., Kanyuka, K., Praekelt, U., Franklin, K.A.*, et al.* (2003). Mutations in the huge Arabidopsis gene BIG affect a range of hormone and light responses. *The Plant Journal 35*, 57-70.

Karas, M., Bachmann, D., and Hillenkamp, F. (1985). Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Analytical Chemistry 57*, 2935-2939.

Keshive, M., Singh, S., Wishnok, J.S., Tannenbaum, S.R., and Deen, W.M. (1996). Kinetics of S-Nitrosation of Thiols in Nitric Oxide Solutions. *Chemical Research in Toxicology 9*, 988-993.

Kim, H.J., Ha, S., Lee, H.Y., and Lee, K.J. (2014a). ROSics: Chemistry and proteomics of cysteine modifications in redox biology. *Mass Spectrometry Reviews 34,* 184-204

Kim, H.K., Kim, R.R., Oh, J.H., Cho, H., Varshavsky, A., and Hwang, C.S. (2014b). The N-terminal methionine of cellular proteins as a degradation signal. *Cell 156*, 158-169.

Knobbe, A.R., Horken, K.M., Plucinak, T.M., Balassa, E., Cerutti, H., and Weeks, D.P. (2015). SUMOylation by a Stress-Specific Small Ubiquitin-Like Modifier E2 Conjugase Is Essential for Survival of Chlamydomonas reinhardtii under Stress Conditions. *Plant Physiology 167*, 753-765.

Koncz, C., and Schell, J. (1986). The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type

of Agrobacterium binary vector. *Molecular and General Genetics MGG 204*, 383-396.

Kosmacz, M., Parlanti, S., SchwarzlÄNder, M., Kragler, F., Licausi, F., and Van Dongen, J.T. (2015). The stability and nuclear localization of the transcription factor RAP2.12 are dynamically regulated by oxygen concentration. *Plant, Cell & Environment 38*, 1094-1103.

Kwon, Y.T., Kashina, A.S., Davydov, I.V., Hu, R.G., An, J.Y., Seo, J.W., Du, F., and Varshavsky, A. (2002). An essential role of N-terminal arginylation in cardiovascular development. *Science 297*, 96-99.

Kwon, Y.T., Kashina, A.S., and Varshavsky, A. (1999). Alternative Splicing Results in Differential Expression, Activity, and Localization of the Two Forms of Arginyl-tRNA-Protein Transferase, a Component of the N-End Rule Pathway. *Molecular and Cellular Biology 19*, 182-193.

Kwon, Y.T., Xia, Z., An, J.Y., Tasaki, T., Davydov, I.V., Seo, J.W., Sheng, J., Xie, Y., and Varshavsky, A. (2003). Female Lethality and Apoptosis of Spermatocytes in Mice Lacking the UBR2 Ubiquitin Ligase of the N-End Rule Pathway. *Molecular and Cellular Biology 23*, 8255-8271.

Lamotte, O., Bertoldo, J.B., Besson-Bard, A., Rosnoblet, C., Aimé, S., Hichami, S., Terenzi, H., and Wendehenne, D. (2015). Protein S-nitrosylation: specificity and identification strategies in plants. *Frontiers in Chemistry 2*.

Lee, F.-J.S., Lin, L.-W., and Smith, J.A. (1989). Nα-acetyltransferase deficiency alters protein synthesis in Saccharomyces cerevisiae. *FEBS letters 256*, 139-142.

Lee, M.J., Kim, D.E., Zakrzewska, A., Yoo, Y.D., Kim, S.H., Kim, S.T., Seo, J.W., Lee, Y.S., Dorn, G.W., 2nd, Oh, U.*, et al.* (2012). Characterization of arginylation branch of N-end rule pathway in G-

protein-mediated proliferation and signaling of cardiomyocytes. *Journal of Biological Chemistry 287*, 24043-24052.

Lee, M.J., Tasaki, T., Moroi, K., An, J.Y., Kimura, S., Davydov, I.V., and Kwon, Y.T. (2005). RGS4 and RGS5 are in vivo substrates of the N-end rule pathway. *Proceedings of the National Academy of Sciences of the United States of America 102*, 15030-15035.

Leone, J.W., Hampton, B., Fowler, E., Moyer, M., Krishna, R.G., and Chin, C.C. (2011). Removal of N-Terminal Blocking Groups from Proteins. Current *Protocols in Protein Science*, 11.17. 11-11.17. 20.

Lewis, R.S., and Deen, W.M. (1994). Kinetics of the reaction of nitric oxide with oxygen in aqueous solutions. *Chemical Research in Toxicology 7*, 568-574.

Li, F., Sonveaux, P., Rabbani, Z.N., Liu, S., Yan, B., Huang, Q., Vujaskovic, Z., Dewhirst, M.W., and Li, C.Y. (2007). Regulation of HIF-1alpha stability through S-nitrosylation. *Molecular Cell 26*, 63-74.

Li, G.-W., Burkhardt, D., Gross, C., and Weissman, Jonathan S. (2014). Quantifying Absolute Protein Synthesis Rates Reveals Principles Underlying Allocation of Cellular Resources. *Cell 157*, 624-635.

Li, X., and Chang, Y.H. (1995). Amino-terminal protein processing in Saccharomyces cerevisiae is an essential function that requires two distinct methionine aminopeptidases. *Proceedings of the National Academy of Sciences 92*, 12357-12361.

Licausi, F., Kosmacz, M., Weits, D.A., Giuntoli, B., Giorgi, F.M., Voesenek, L.A., Perata, P., and van Dongen, J.T. (2011). Oxygen sensing in plants is mediated by an N-end rule pathway for protein destabilization. *Nature 479*, 419-422.

Licausi, F., Pucciariello, C., and Perata, P. (2013). New role for an old rule: N-end rule-mediated degradation of ethylene responsive factor proteins governs low oxygen response in plants(F). *Journal of Integrated Plant Bio*logy *55*, 31-39.

Licausi, F., van Dongen, J.T., Giuntoli, B., Novi, G., Santaniello, A., Geigenberger, P., and Perata, P. (2010). HRE1 and HRE2, two hypoxia-inducible ethylene response factors, affect anaerobic responses in Arabidopsis thaliana. *Plant Journal 62*, 302-315.

Lindemann, C., and Leichert, L.I. (2012). Quantitative redox proteomics: the NOxICAT method. *Methods in Molecular Biology 893*, 387-403.

Liu, Y., Ye, N., Liu, R., Chen, M., and Zhang, J. (2010). H2O2 mediates the regulation of ABA catabolism and GA biosynthesis in Arabidopsis seed dormancy and germination. *Journal of Experimental Botany 61*, 2979-2990.

Lloyd-Williams, P., Albericio, F., and Giralt, E. (1997). Chemical approaches to the synthesis of peptides and proteins, Vol 10 (CRC Press).

Löchmuller, C., and Wilder, D. (1979). The sorption behavior of alkyl bonded phases in reverse-phase, high performance liquid chromatography. *Journal of Chromatographic Science 17*, 574-579.

Magnani, E., Sjölander, K., and Hake, S. (2004). From Endonucleases to Transcription Factors: Evolution of the AP2 DNA Binding Domain in Plants. *The Plant Cell 16*, 2265-2277.

Manzano, C., Abraham, Z., López-Torrejón, G., and Del Pozo, J.C. (2008). Identification of ubiquitinated proteins in Arabidopsis. *Plant Molecular Biology 68*, 145-158.

Marino, D., Andrio, E., Danchin, E.G., Oger, E., Gucciardo, S., Lambert, A., Puppo, A., and Pauly, N. (2011). *A Medicago truncatula NADPH oxidase is involved in symbiotic nodule functioning*. New Phytologist *189*, 580-592.

Marino, S.M., and Gladyshev, V.N. (2012). Analysis and functional prediction of reactive cysteine residues*. Journal of Biological Chemistry 287*, 4419-4425.

Marshall, A.G. (1985). Fourier transform ion cyclotron resonance mass spectrometry. *Accounts of Chemical Research 18*, 316-322.

McLafferty, F. (1981). Tandem mass spectrometry. *Science 214*, 280-287.

McLafferty, F.W. (1959). Mass spectrometric analysis. Molecular rearrangements. *Analytical Chemistry 31*, 82-87.

Mego, J.L., and Farb, R.M. (1978). An energy requirement for the degradation of intravenously injected 125I-labeled albumin in mouse liver and kidney slices. *Biochemical Journal 172*, 233-238.

Meitha, K., Konnerup, D., Colmer, T.D., Considine, J.A., Foyer, C.H., and Considine, M.J. (2015). Spatio-temporal relief from hypoxia and production of reactive oxygen species during bud burst in grapevine (Vitis vinifera). *Annals of Botany.*

Mendiondo, G.M., Gibbs, D.J., Szurman-Zubrzycka, M., Korn, A., Marquez, J., Szarejko, I., Maluszynski, M., King, J., Axcell, B., Smart, K.*, et al.* (2016). Enhanced waterlogging tolerance in barley by manipulation of expression of the N-end rule pathway E3 ligase PROTEOLYSIS6. *Plant Biotechnology Journa*l *14*, 40-50.

Meyer, C., and Stöhr, C. (2002). Soluble and plasma membrane-bound enzymes involved in nitrate and nitrite metabolism. In Photosynthetic

nitrogen assimilation and associated carbon and respiratory metabolism (Springer), pp. 49-62.

Mitchell Wells, J., and McLuckey, S.A. (2005). Collision-Induced Dissociation (CID) of Peptides and Proteins. *In Methods in Enzymology (Academic Press*), pp. 148-185.

Moche, M., Stremlau, S., Hecht, L., Gobel, C., Feussner, I., and Stohr, C. (2010). Effect of nitrate supply and mycorrhizal inoculation on characteristics of tobacco root plasma membrane vesicles. *Planta 231*, 425-436.

Mohyeldin, A., Garzón-Muvdi, T., and Quiñones-Hinojosa, A. (2010). Oxygen in Stem Cell Biology: A Critical Component of the Stem Cell Niche. *Cell Stem Cell 7*, 150-161.

Møller, I.M., and Kristensen, B.K. (2004). Protein oxidation in plant mitochondria as a stress indicator. *Photochemical & Photobiological Sciences 3*, 730-735.

Mortimore, G.E., and Poso, A.R. (1987). INTRACELLULAR PROTEIN CATABOLISM AND ITS CONTROL DURING NUTRIENT DEPRIVATION AND SUPPLY. *Annual Reviews of Nutrition*. *7*, 539-564.

Mot, A.C., Prell, E., Klecker, M., Naumann, C., Faden, F., Westermann, B., and Dissmeyer, N. (2016). Real-time detection of PRT1-mediated ubiquitination via fluorescently labeled substrate probes. *bioRxi*v.

Mur, L.A., Mandon, J., Persijn, S., Cristescu, S.M., Moshkov, I.E., Novikova, G.V., Hall, M.A., Harren, F.J., Hebelstrup, K.H., and Gupta, K.J. (2013). Nitric oxide in plants: an assessment of the current state of knowledge. *Annals of Botany plants 5*.

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiology Plantarum 15*.

Murray, C.I., and Van Eyk, J.E. (2012). Chasing cysteine oxidative modifications: proteomic tools for characterizing cysteine redox status. *Circulation: Cardiovascular Genetics 5*, 591.

Myllyharju, J. (2013). Prolyl 4-hydroxylases, master regulators of the hypoxia response. *Acta Physiologica 208*, 148-165.

Nakano, T., Suzuki, K., Fujimura, T., and Shinshi, H. (2006). Genome-wide analysis of the ERF gene family in Arabidopsis and rice. *Plant Physiology 140*, 411-432.

Nawy, T., Lee, J.-Y., Colinas, J., Wang, J.Y., Thongrod, S.C., Malamy, J.E., Birnbaum, K., and Benfey, P.N. (2005). Transcriptional Profile of the Arabidopsis Root Quiescent Center. *The Plant Cell 17*, 1908-1925.

Ninnis, R.L., Spall, S.K., Talbo, G.H., Truscott, K.N., and Dougan, D.A. (2009). Modification of PATase by L/F-transferase generates a ClpS-dependent N-end rule substrate in Escherichia coli. *EMBO* Journal *28*, 1732-1744.

Ozkaynak, E., Finley, D., and Varshavsky, A. (1984). The yeast ubiquitin gene: head-to-tail repeats encoding a polyubiquitin precursor protein. *Nature 312*, 663-666.

Palmer, R.M., Ashton, D.S., and Moncada, S. (1988). Vascular endothelial cells synthesize nitric oxide from L-arginine. *Nature 333*, 664-666.

Palmer, R.M.J., Ferrige, A.G., and Moncada, S. (1987). Nitric oxide release accounts for the biological activity of endothelium-derived relaxing factor. *Nature 327*, 524-526.

Papdi, C., Ábrahám, E., Joseph, M.P., Popescu, C., Koncz, C., and Szabados, L. (2008). Functional Identification of Arabidopsis Stress

Regulatory Genes Using the Controlled cDNA Overexpression System. *Plant Physiology 147*, 528-542.

Papdi, C., Pérez-Salamó, I., Joseph, M.P., Giuntoli, B., Bögre, L., Koncz, C., and Szabados, L. (2015). The low oxygen, oxidative and osmotic stress responses synergistically act through the ethylene response factor VII genes RAP2.12, RAP2.2 and RAP2.3. *The Plant Journal 82*, 772-784.

Pauli, S., Rothnie, H.M., Chen, G., He, X., and Hohn, T. (2004). The Cauliflower Mosaic Virus 35S Promoter Extends into the Transcribed Region. *Journal of Virology 78*, 12120-12128.

Penfield, S., Rylott, E.L., Gilday, A.D., Graham, S., Larson, T.R., and Graham, I.A. (2004). Reserve mobilization in the Arabidopsis endosperm fuels hypocotyl elongation in the dark, is independent of abscisic acid, and requires PHOSPHOENOLPYRUVATE CARBOXYKINASE1. *Plant Cell 16*, 2705-2718.

Perazzolli, M., Dominici, P., Romero-Puertas, M.C., Zago, E., Zeier, J., Sonoda, M., Lamb, C., and Delledonne, M. (2004). Arabidopsis nonsymbiotic hemoglobin AHb1 modulates nitric oxide bioactivity. *Plant Cell 16*, 2785-2794.

Persson, B., Flinta, C., von Heijne, G., and JÖRnvall, H. (1985). Structures of N-terminally acetylated proteins. *European Journal of Biochemistry 152*, 523-527.

Pine, M.J. (1973). Stringent Control of Intracellular Proteolysis in Escherichia coli. *Journal of Bacteriology 116*, 1253-1257.

Potocky, M., Jones, M.A., Bezvoda, R., Smirnoff, N., and Zarsky, V. (2007). Reactive oxygen species produced by NADPH oxidase are involved in pollen tube growth. *New Phytologist 174*, 742-751.

Potuschak, T., Stary, S., Schlögelhofer, P., Becker, F., Nejinskaia, V., and Bachmair, A. (1998). PRT1 of Arabidopsis thaliana encodes a component of the plant N-end rule pathway. *Proceedings of the National Academy of Sciences 95*, 7904-7908.

Prakash, S., Tian, L., Ratliff, K.S., Lehotzky, R.E., and Matouschek, A. (2004). An unstructured initiation site is required for efficient proteasome-mediated degradation. *Nature Structural & Molecular Biology 11*, 830-837.

Pröls, F., and Meyer, P. (1992). The methylation patterns of chromosomal integration regions influence gene activity of transferred DNA in Petunia hybrida. *The Plant Journal 2,* 465-475.

Proost, S., Van Bel, M., Vaneechoutte, D., Van de Peer, Y., Inzé, D., Mueller-Roeber, B., and Vandepoele, K. (2014). PLAZA 3.0: an access point for plant comparative genomics. *Nucleic Acids Research.*

Rajeshwara, A., and Prakash, V. (1994). Structural stability of lipase from wheat germ. *International Journal of Peptide and Protein Research 44*, 435-440.

Rey, P., Bécuwe, N., Barrault, M.-B., Rumeau, D., Havaux, M., Biteau, B., and Toledano, M.B. (2007). The Arabidopsis thaliana sulfiredoxin is a plastidic cysteine-sulfinic acid reductase involved in the photooxidative stress response. *The Plant Journal 49*, 505-514.

Riechmann, J.L., Heard, J., Martin, G., Reuber, L., Jiang , C.-Z., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O.J., Samaha, R.R.*, et al.* (2000). Arabidopsis Transcription Factors: Genome-Wide Comparative Analysis Among Eukaryotes. *Science 290*, 2105-2110.

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

Roman-Hernandez, G., Grant, R.A., Sauer, R.T., and Baker, T.A. (2009). Molecular basis of substrate selection by the N-end rule adaptor protein ClpS. *Proceedings of the National Academy of Sciences of the United States of America 106*, 8888-8893.

Romero-Puertas, M.C., Rodriguez-Serrano, M., and Sandalio, L.M. (2013). Protein S-nitrosylation in plants under abiotic stress: an overview. *Frontiers in Plant Science 4*, 373.

Rumer, S., Gupta, K.J., and Kaiser, W.M. (2009). Plant cells oxidize hydroxylamines to NO. *Journal of Experimental Botany 60*, 2065-2072.

Salmeen, A., Andersen, J.N., Myers, M.P., Meng, T.-C., Hinks, J.A., Tonks, N.K., and Barford, D. (2003). Redox regulation of protein tyrosine phosphatase 1B involves a sulphenyl-amide intermediate. *Nature 423*, 769-773.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America 74*, 5463-5467.

Schimke, R.T., and Doyle, D. (1970). CONTROL OF ENZYME LEVELS IN ANIMAL TISSUES. *Annual Review of Biochemistry 39*, 929-&.

Schoenheimer, R. (1942). The Dynamic State of Body Constituents. Harvard Univ. Monogr. Med. publ. Hlth.

Schubert, D., Lechtenberg, B., Forsbach, A., Gils, M., Bahadur, S., and Schmidt, R. (2004). Silencing in Arabidopsis T-DNA Transformants: The Predominant Role of a Gene-Specific RNA Sensing Mechanism versus Position Effects. *The Plant Cell 16*, 2561-2572.

Schuessele, C., Hoernstein, S.N.W., Mueller, S.J., Rodriguez-Franco, M., Lorenz, T., Lang, D., Igloi, G.L., and Reski, R. (2016). Spatio-temporal

patterning of arginyl-tRNA protein transferase (ATE) contributes to gametophytic development in a moss. *New Phytologist 209*, 1014-1027.

Serrano, I., Pelliccione, S., and Olmedilla, A. (2010). Programmed-cell-death hallmarks in incompatible pollen and papillar stigma cells of Olea europaea L. under free pollination. *Plant Cell Reports 29*, 561-572.

Sethuraman, M., McComb, M.E., Huang, H., Huang, S., Heibeck, T., Costello, C.E., and Cohen, R.A. (2004). Isotope-coded affinity tag (ICAT) approach to redox proteomics: identification and quantitation of oxidant-sensitive cysteine thiols in complex protein mixtures. *Journal of Proteome Researc*h *3*, 1228-1233.

Shahzad, Z., Canut, M., Tournaire-Roux, C., Martinière, A., Boursiac, Y., Loudet, O., and Maurel, C. (2016). A Potassium-Dependent Oxygen Sensing Pathway Regulates Plant Root Hydraulics. *Cell.*

Shemorry, A., Hwang, C.S., and Varshavsky, A. (2013). Control of protein quality and stoichiometries by N-terminal acetylation and the N-end rule pathway. *Molecular Cell 50*, 540-551.

Sherman, F., Stewart, J.W., and Tsunasawa, S. (1985). Methionine or not methionine at the beginning of a protein. *BioEssays 3*, 27-31.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J.*, et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *In Molecular Systems Biology,* p. 539.

Singh, N., Dang, T.T.M., Vergara, G.V., Pandey, D.M., Sanchez, D., Neeraja, C.N., Septiningsih, E.M., Mendioro, M., Tecson-Mendoza, E.M., Ismail, A.M.*, et al.* (2010). Molecular marker survey and expression analyses of the rice submergence-tolerance gene SUB1A. *Theoretical and Applied Genetics 121*, 1441-1453.

Sirova, J., Sedlarova, M., Piterkova, J., Luhova, L., and Petrivalsky, M. (2011). The role of nitric oxide in the germination of plant seeds and pollen. *Plant Science 181*, 560-572.

Soffer, R.L. (1980). Biochemistry and biology of aminoacyl-tRNA-protein transferases. Cold Spring Harbor Monograph Archive *9*, 493-505.

Sriram, S.M., Kim, B.Y., and Kwon, Y.T. (2011). The N-end rule pathway: emerging functions and molecular principles of substrate recognition. *Nature Reviews in Molecular Cell Biology 12*, 735-747.

Staes, A., Van Damme, P., Helsens, K., Demol, H., Vandekerckhove, J., and Gevaert, K. (2008). Improved recovery of proteome-informative, protein N-terminal peptides by combined fractional diagonal chromatography (COFRADIC). *Proteomics 8*, 1362-1370.

Stafford, G. (2002). Ion trap mass spectrometry: a personal perspective. *Journal of the American Society for Mass spectrometry 13*, 589-596.

Stary, S., Yin, X.J., Potuschak, T., Schlogelhofer, P., Nizhynska, V., and Bachmair, A. (2003). PRT1 of Arabidopsis is a ubiquitin protein ligase of the plant N-end rule pathway with specificity for aromatic amino-terminal residues. *Plant Physiology 133*, 1360-1366.

Steffan, J.S., Agrawal, N., Pallos, J., Rockabrand, E., Trotman, L.C., Slepko, N., Illes, K., Lukacsovich, T., Zhu, Y.-Z., Cattaneo, E.*, et al.* (2004). SUMO Modification of Huntingtin and Huntington's Disease Pathology. *Science 304*, 100-104.

Stöhr, C., Strube, F., Marx, G., Ullrich, W.R., and Rockel, P. (2001). A plasma membrane-bound enzyme of tobacco roots catalyses the formation of nitric oxide from nitrite. *Planta 212*, 835-841.

Strachan, J., Shaw, B., Krishna, V.G., Scott, D., del Pozo, J.C., Hill, K., Gautam, M., Jacobson, A.D., Liu, C.-W., and Oldham, N.J. (2013). Broad

Utility of an Affinity-enrichment Strategy for Unanchored Polyubiquitin Chains. *Journal of Proteomics & Bioinformatics 2014*.

Swanson, S., and Gilroy, S. (2010). ROS in plant development. *Physiologia Plantarum 138*, 384-392.

Tanou, G., Job, C., Rajjou, L., Arc, E., Belghazi, M., Diamantidis, G., Molassiotis, A., and Job, D. (2009). Proteomics reveals the overlapping roles of hydrogen peroxide and nitric oxide in the acclimation of citrus plants to salinity. *Plant Journal 60*, 795-804.

Tasaki, T., Mulder, L.C.F., Iwamatsu, A., Lee, M.J., Davydov, I.V., Varshavsky, A., Muesing, M., and Kwon, Y.T. (2005). A Family of Mammalian E3 Ubiquitin Ligases That Contain the UBR Box Motif and Recognize N-Degrons. *Molecular and Cellular Biology 25*, 7120-7136.

Tasaki, T., Sriram, S.M., Park, K.S., and Kwon, Y.T. (2012). The N-end rule pathway. *Annual Reviews in Biochemistry 81*, 261-289.

Tasaki, T., Zakrzewska, A., Dudgeon, D.D., Jiang, Y., Lazo, J.S., and Kwon, Y.T. (2009). The substrate recognition domains of the N-end rule pathway. *Journal of Biological Chemistry 284*, 1884-1895.

Taylor, G. (1964). Disintegration of water drops in an electric field. In Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences (The Royal Society), pp. 383-397.

Thomson, J.J. (1897). XL. Cathode Rays. *Philosophical Magazine Series 5 44*, 293-316.

Tjernberg, A., Markova, N., Griffiths, W.J., and Hallén, D. (2006). DMSO-Related Effects in Protein Characterization. *Journal of Biomolecular Screening 11*, 131-137.

Tobias, J.W., Shrader, T.E., Rocap, G., and Varshavsky, A. (1991). The N-end rule in bacteria. *Science 254*, 1374-1377.

Torres, M.A., Jones, J.D., and Dangl, J.L. (2005). Pathogen-induced, NADPH oxidase-derived reactive oxygen intermediates suppress spread of cell death in Arabidopsis thaliana. *Nature Genetics 37*, 1130-1134.

Towbin, H., Staehelin, T., and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences of the United States of America 76*, 4350-4354.

Trivedi, M.V., Laurence, J.S., and Siahaan, T.J. (2009). The role of thiols and disulfides in protein chemical and physical stability. *Current Protein & Peptide Science 10*, 614-625.

Tun, N.N., Santa-Catarina, C., Begum, T., Silveira, V., Handro, W., Floh, E.I., and Scherer, G.F. (2006). Polyamines induce rapid biosynthesis of nitric oxide (NO) in Arabidopsis thaliana seedlings. *Plant Cell Physiology 47*, 346-354.

Valimberti, I., Tiberti, M., Lambrughi, M., Sarcevic, B., and Papaleo, E. (2015). E2 superfamily of ubiquitin-conjugating enzymes: constitutively active or activated through phosphorylation in the catalytic cleft. *Scientific Reports 5*, 14849.

van Dongen, J.T., Schurr, U., Pfister, M., and Geigenberger, P. (2003). Phloem Metabolism and Function Have to Cope with Low Internal Oxygen. *Plant Physiology 131*, 1529-1543.

van Montfort, R.L.M., Congreve, M., Tisi, D., Carr, R., and Jhoti, H. (2003). Oxidation state of the active-site cysteine in protein tyrosine phosphatase 1B. *Nature 423*, 773-777.

van Veen, H., Akman, M., Jamar, D.C.L., Vreugdenhil, D., Kooiker, M., van Tienderen, P., Voesenek, L.A.C.J., Schranz, M.E., and Sasidharan, R. (2014). Group VII Ethylene Response Factor diversification and

regulation in four species from flood-prone environments. *Plant, Cell & Environment 37*, 2421-2432.

van Wijk, S.J.L., and Timmers, H.T.M. (2010). The family of ubiquitin-conjugating enzymes (E2s): deciding between life and death of proteins. *The FASEB Journal 24*, 981-993.

Varshavsky, A. (1991). Naming a targeting signal. *Cell 64*, 13-15.

Varshavsky, A. (2004). 'Spalog' and 'sequelog': neutral terms for spatial and sequence similarity. *Current Biology 14*, R181-183.

Venne, A.S., Solari, F.A., Faden, F., Paretti, T., Dissmeyer, N., and Zahedi, R.P. (2015). An improved workflow for quantitative N-terminal charge-based fractional diagonal chromatography (ChaFRADIC) to study proteolytic events in Arabidopsis thaliana. *Proteomics 15*, 2458-2469.

Venne, A.S., Vögtle, F.N., Meisinger, C., Sickmann, A., and Zahedi, R.P. (2013). Novel Highly Sensitive, Specific, and Straightforward Strategy for Comprehensive N-Terminal Proteomics Reveals Unknown Substrates of the Mitochondrial Peptidase Icp55. *Journal of Proteome Research 12*, 3823-3830.

Voesenek, L.A., and Bailey-Serres, J. (2013). Flooding tolerance: O2 sensing and survival strategies. *Current Opinions in Plant Biology 16*, 647-653.

Wadas, B., Piatkov, K., Brower, C., and Varshavsky, A. (2016). Analyzing N-terminal Arginylation Through the Use of Peptide Arrays and Degradation Assays. *Journal of Biological Chemistry*.

Wang, G.L., and Semenza, G.L. (1993). General involvement of hypoxia-inducible factor 1 in transcriptional response to hypoxia. *Proceedings of the National Academy of Sciences 90*, 4304-4308.

Wang, H., Piatkov, K.I., Brower, C.S., and Varshavsky, A. (2009). Glutamine-specific N-terminal amidase, a component of the N-end rule pathway. *Molecular Cell 34*, 686-695.

Wang, K.H., Sauer, R.T., and Baker, T.A. (2007). ClpS modulates but is not essential for bacterial N-end rule degradation. *Genes and Development 21*, 403-408.

Wang, S., Cao, L., and Wang, H. (2016). Arabidopsis ubiquitin-conjugating enzyme UBC22 is required for female gametophyte development and likely involved in Lys11-linked ubiquitination. *Journal of Experimental Botany*.

Waszczak, C., Akter, S., Jacques, S., Huang, J., Messens, J., and Van Breusegem, F. (2015). Oxidative post-translational modifications of cysteine residues in plant signal transduction. *Journal of Experimental Botany*.

Weits, D.A., Giuntoli, B., Kosmacz, M., Parlanti, S., Hubberten, H.M., Riegler, H., Hoefgen, R., Perata, P., van Dongen, J.T., and Licausi, F. (2014). Plant cysteine oxidases control the oxygen-dependent branch of the N-end-rule pathway. *Nature Communications 5*, 3425.

White, M.D., and Flashman, E. (2016). Catalytic strategies of the non-heme iron dependent oxygenases and their roles in plant biology. *Current Opinion in Chemical Biology 31*, 126-135.

White, M.D., Klecker, M., Hopkinson, R., Weits, D., Mueller, C., Naumann, C., O Neill, R., Wickens, J., Grossmann, T.N., Dissmeyer, N., *et al.* (2016). Plant Cysteine Oxidases are Dioxygenases that Directly Enable Arginyl Transferase-Catalyzed Arginylation of N-End Rule Targets. *bioRxiv*.

Whitehouse, C., Dreyer, R., Yamashita, M., and Fenn, J. (1989). Electrospray ionization for mass-spectrometry of large biomolecules. *Science 246*, 64-71.

Wilmink, A., Van de Ven, B., and Dons, J. (1995). Activity of constitutive promoters in various species from the Liliaceae. *Plant Molecular Biology 28*, 949-955.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V., and Provart, N.J. (2007). An "Electronic Fluorescent Pictograph" Browser for Exploring and Analyzing Large-Scale Biological Data Sets. *PLoS ONE 2*, e718.

Winterbourn, C.C. (2013). Chapter One - The Biological Chemistry of Hydrogen Peroxide. In Methods in Enzymology, C. Enrique, and P. Lester, eds. (Academic Press), pp. 3-25.

Xu, K., Xu, X., Fukao, T., Canlas, P., Maghirang-Rodriguez, R., Heuer, S., Ismail, A.M., Bailey-Serres, J., Ronald, P.C., and Mackill, D.J. (2006). Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature 442*, 705-708.

Yadav, R.K., Girke, T., Pasala, S., Xie, M., and Reddy, G.V. (2009). Gene expression map of the Arabidopsis shoot apical meristem stem cell niche. *Proceedings of the National Academy of Sciences 106*, 4941-4946.

Yamasaki, H., and Sakihama, Y. (2000). Simultaneous production of nitric oxide and peroxynitrite by plant nitrate reductase: in vitro evidence for the NR-dependent formation of active nitrogen species. *Febs Letters 468*, 89-92.

Yoo, S.Y., Bomblies, K., Yoo, S.K., Yang, J.W., Choi, M.S., Lee, J.S., Weigel, D., and Ahn, J.H. (2005). The 35S promoter used in a selectable marker gene of a plant transformation vector affects the expression of the transgene. *Planta 221*, 523-530.

Yoshida, S., Ito, M., Callis, J., Nishida, I., and Watanabe, A. (2002). A delayed leaf senescence mutant is defective in arginyl-tRNA:protein arginyltransferase, a component of the N-end rule pathway in Arabidopsis. *The Plant Journal 32*, 129-137.

Yu, M., Lamattina, L., Spoel, S.H., and Loake, G.J. (2014). Nitric oxide function in plant biology: a redox cue in deconvolution. *New Phytologist 202*, 1142-1156.

Yun, B.W., Feechan, A., Yin, M., Saidi, N.B., Le Bihan, T., Yu, M., Moore, J.W., Kang, J.G., Kwon, E., Spoel, S.H.*, et al.* (2011). S-nitrosylation of NADPH oxidase regulates cell death in plant immunity. *Nature 478*, 264-268.

Zhang, H., Deery, M.J., Gannon, L., Powers, S.J., Lilley, K.S., and Theodoulou, F.L. (2015). Quantitative proteomics analysis of the Arg/N-end rule pathway of targeted degradation in Arabidopsis roots. *Proteomics 15*, 2447-2457.

Zheng, X., Deng, W., Luo, K., Duan, H., Chen, Y., McAvoy, R., Song, S., Pei, Y., and Li, Y. (2007). The cauliflower mosaic virus (CaMV) 35S promoter sequence alters the level and patterns of activity of adjacent tissue- and organ-specific gene promoters. *Plant Cell Reports 26*, 1195-1203.

Zubarev, R.A., Kelleher, N.L., and McLafferty, F.W. (1998). Electron capture dissociation of multiply charged protein cations. A nonergodic process. *Journal of the American Chemical Society 120*, 3265-3266.

8. Appendix

A. Custom GUS MASCOT database. Comprising of the native sequence of the *in vivo* GUS reporter protein expressed in 35S::MC-GUS Arabidopsis seedlings (MC-GUS), an Nt-Cys version of this sequence possible after N-terminal Met excision of MC-GUS (C-GUS), and lastly and Arg-Cys Nt version of the GUS reporter protein expected to occur to *in vivo* GUS reporter protein after arginylation of oxidised Cys (RC-GUS).

MC-GUS

MCRSGGGGGGGGGGGRGGSGAWLLPVSLVKRKTTLAPNTQTASPRALADSLMQHMGSYPYDV
PDYAGYPYDVPDYAGSYPYDVPDYAGTMVRPVETPTREIKKLDGLWAFSLDRENCGIDQRWWESA
LQESRAIAVPGSFNDQFADADIRNYAGNVWYQREVFIPKGWAGQRIVLRFDAVTHYGKVWVNNQ
EVMEHQGGYTPFEADVTPYVIAGKSVRITVCVNNELNWQTIPPGMVITDENGKKKQSYFHDFFNYA
GIHRSVMLYTTPNTWVDDITVVTHVAQDCNHASVDWQVVANGDVSVELRDADQQVVATGQGTS
GTLQVVNPHLWQPGEGYLYELCVTAKSQTECDIYPLRVGIRSVAVKGQQFLINHKPFYFTGFGRHED
ADLRGKGFDNVLMVHDHALMDWIGANSYRTSHYPYAEEMLDWADEHGIVVIDETAAVGFNLSLGI
GFEAGNKPKELYSEEAVNGETQQAHLQAIKELIARDKNHPSVVMWSIANEPDTRPQVHGNISPLAEA
TRKLDPTRPITCVNVMFCDAHTDTISDLFDVLCLNRYYGWYVQSGDLETAEKVLEKELLAWQEKLHQ
PIIITEYGVDTLAGLHSMYTDMWSEEYQCAWLDMYHRVFDRVSAVVGEQVWNFADFATSQGILRV
GGNKKGIFTRDRKPKSAAFLLQKRWTGMNFGEKPQQGGKQ

C-GUS

CRSGGGGGGGGGGGRGGSGAWLLPVSLVKRKTTLAPNTQTASPRALADSLMQHMGSYPYDVPD
YAGYPYDVPDYAGSYPYDVPDYAGTMVRPVETPTREIKKLDGLWAFSLDRENCGIDQRWWESALQ
ESRAIAVPGSFNDQFADADIRNYAGNVWYQREVFIPKGWAGQRIVLRFDAVTHYGKVWVNNQEV

MEHQGGYTPFEADVTPYVIAGKSVRITVCVNNELNWQTIPPGMVITDENGKKKQSYFHDFFNYAGI

HRSVMLYTTPNTWVDDITVVTHVAQDCNHASVDWQVVANGDVSVELRDADQQVVATGQGTSGT

LQVVNPHLWQPGEGYLYELCVTAKSQTECDIYPLRVGIRSVAVKGQQFLINHKPFYFTGFGRHEDADL

RGKGFDNVLMVHDHALMDWIGANSYRTSHYPYAEEMLDWADEHGIVVIDETAAVGFNLSLGIGFE

AGNKPKELYSEEAVNGETQQAHLQAIKELIARDKNHPSVVMWSIANEPDTRPQVHGNISPLAEATRK

LDPTRPITCVNVMFCDAHTDTISDLFDVLCLNRYYGWYVQSGDLETAEKVLEKELLAWQEKLHQPIIIT

EYGVDTLAGLHSMYTDMWSEEYQCAWLDMYHRVFDRVSAVVGEQVWNFADFATSQGILRVGGN

KKGIFTRDRKPKSAAFLLQKRWTGMNFGEKPQQGGKQ

RC-GUS

RCRSGGGGGGGGGGGGRGGSGAWLLPVSLVKRKTTLAPNTQTASPRALADSLMQHMGSYPYDVP

DYAGYPYDVPDYAGSYPYDVPDYAGTMVRPVETPTREIKKLDGLWAFSLDRENCGIDQRWWESAL

QESRAIAVPGSFNDQFADADIRNYAGNVWYQREVFIPKGWAGQRIVLRFDAVTHYGKVWVNNQE

VMEHQGGYTPFEADVTPYVIAGKSVRITVCVNNELNWQTIPPGMVITDENGKKKQSYFHDFFNYAG

IHRSVMLYTTPNTWVDDITVVTHVAQDCNHASVDWQVVANGDVSVELRDADQQVVATGQGTSG

TLQVVNPHLWQPGEGYLYELCVTAKSQTECDIYPLRVGIRSVAVKGQQFLINHKPFYFTGFGRHEDA

DLRGKGFDNVLMVHDHALMDWIGANSYRTSHYPYAEEMLDWADEHGIVVIDETAAVGFNLSLGIG

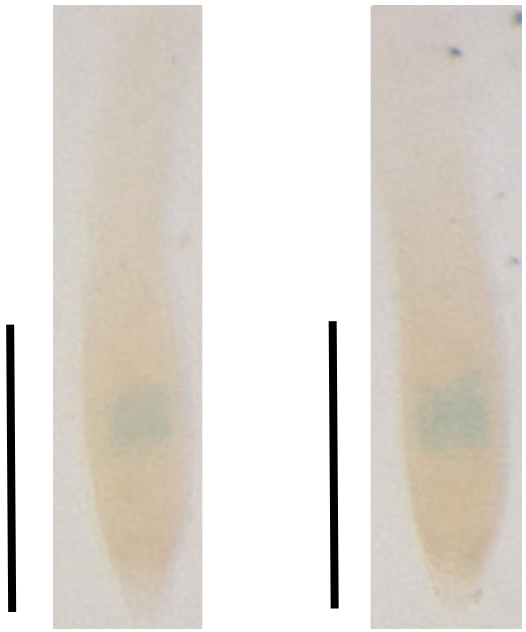
FEAGNKPKELYSEEAVNGETQQAHLQAIKELIARDKNHPSVVMWSIANEPDTRPQVHGNISPLAEAT

RKLDPTRPITCVNVMFCDAHTDTISDLFDVLCLNRYYGWYVQSGDLETAEKVLEKELLAWQEKLHQPI

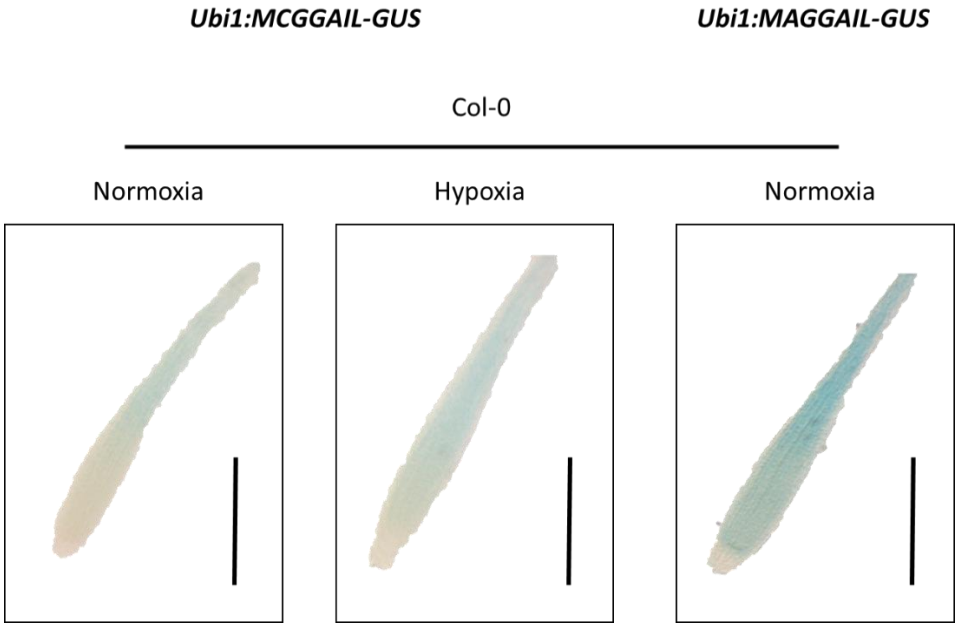
IITEYGVDTLAGLHSMYTDMWSEEYQCAWLDMYHRVFDRVSAVVGEQVWNFADFATSQGILRVG

GNKKGIFTRDRKPKSAAFLLQKRWTGMNFGEKPQQGGKQ

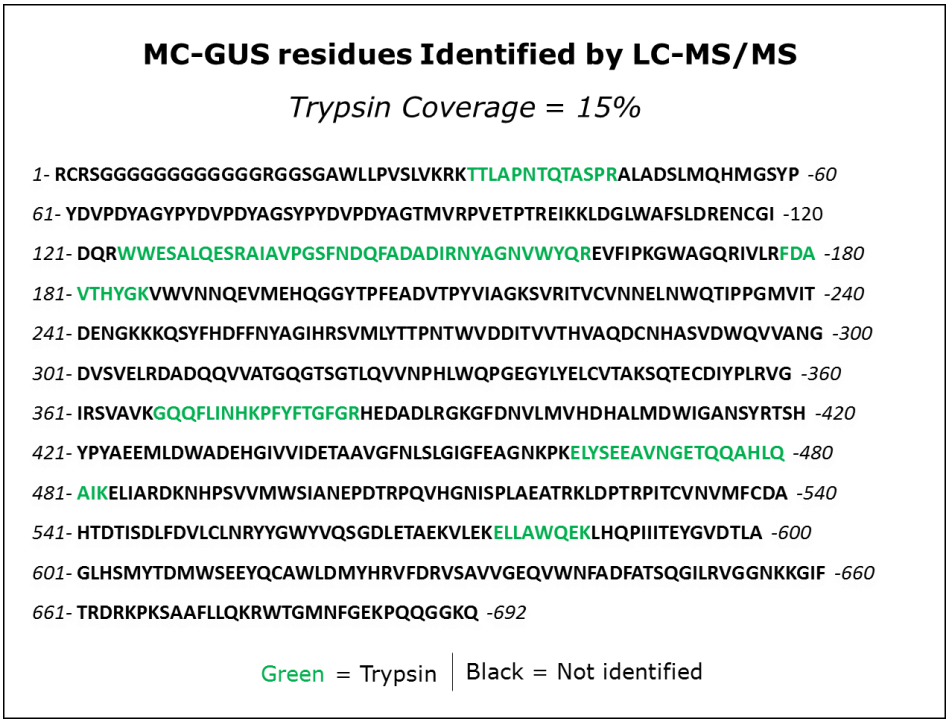
B. Images of root tip after GUS staining of normoxia treated *promRAP2.12::MA-RAP2.12-GUS*. Scale bar = 250 μm.



C. Stabilisation of GUS in the root tip of *Ubi1::M(C/A)GGAIL-GUS* after normoxia/hypoxia treatment. Scale bar = 250 μm.

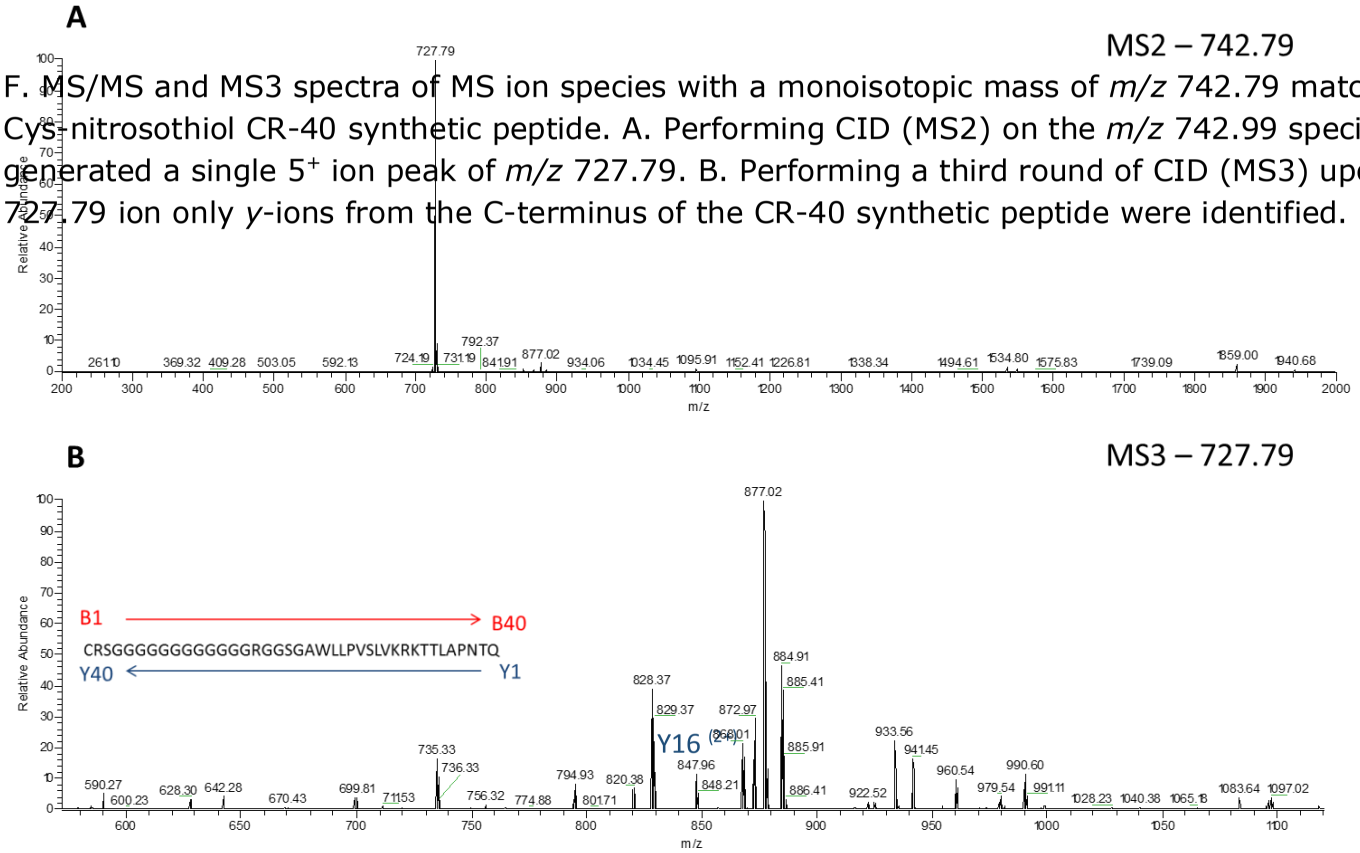


D. Identified tryptic peptides from the GUS reporter protein from *35S::MC-GUS* seedlings. Tryptic peptides identified from gel shown in Figure 4.5.



E. Anticipated mass of Nt-Asp-N peptide of GUS reporter protein if Nt-Arg of peptide is acetylated. The mass shift caused by the fixed acetylation modification and the variable oxidative modifications are shown for each of the charge states observed for the RCR-49 synthetic peptide.

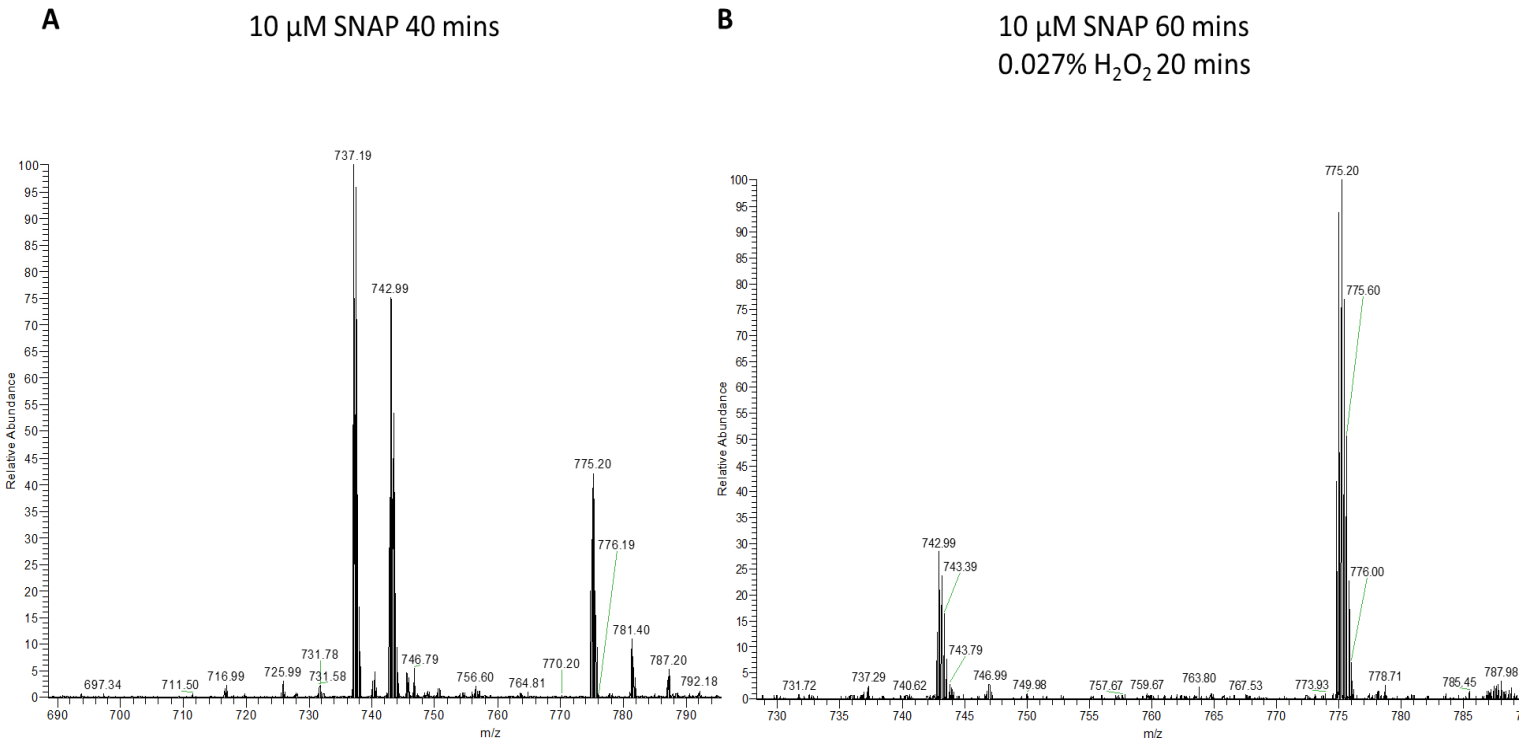
Fixed Cys Mods. = Alkylation (+57 da)	Ac/RC-GUS Mass = 4645 da	Charge state			
Ox./NO Modification at Cys	Mass shift (da)	4^+	5^+	6^+	7^+
None	*+ 0*	1176.615	941.4918	784.7432	672.7799
Oxidation	*+ 16*	1180.614	944.6908	787.409	675.0649
Dioxidation	*+ 32*	1184.612	947.8898	790.0748	677.3499
Trioxidation	*+ 48*	1188.611	951.0888	792.7407	679.6349
Cyclic Sulfenamide	*- 2*	1176.111	941.0886	784.4072	672.4919
Cyclic Sulfinamide	*+ 14*	1180.11	944.2876	787.073	674.7769
Cyclic Sulfonamide	*+30*	1184.109	947.487	789.7392	677.0621
Nitrosylation	*+ 29*	1183.862	947.2898	789.5748	676.9213



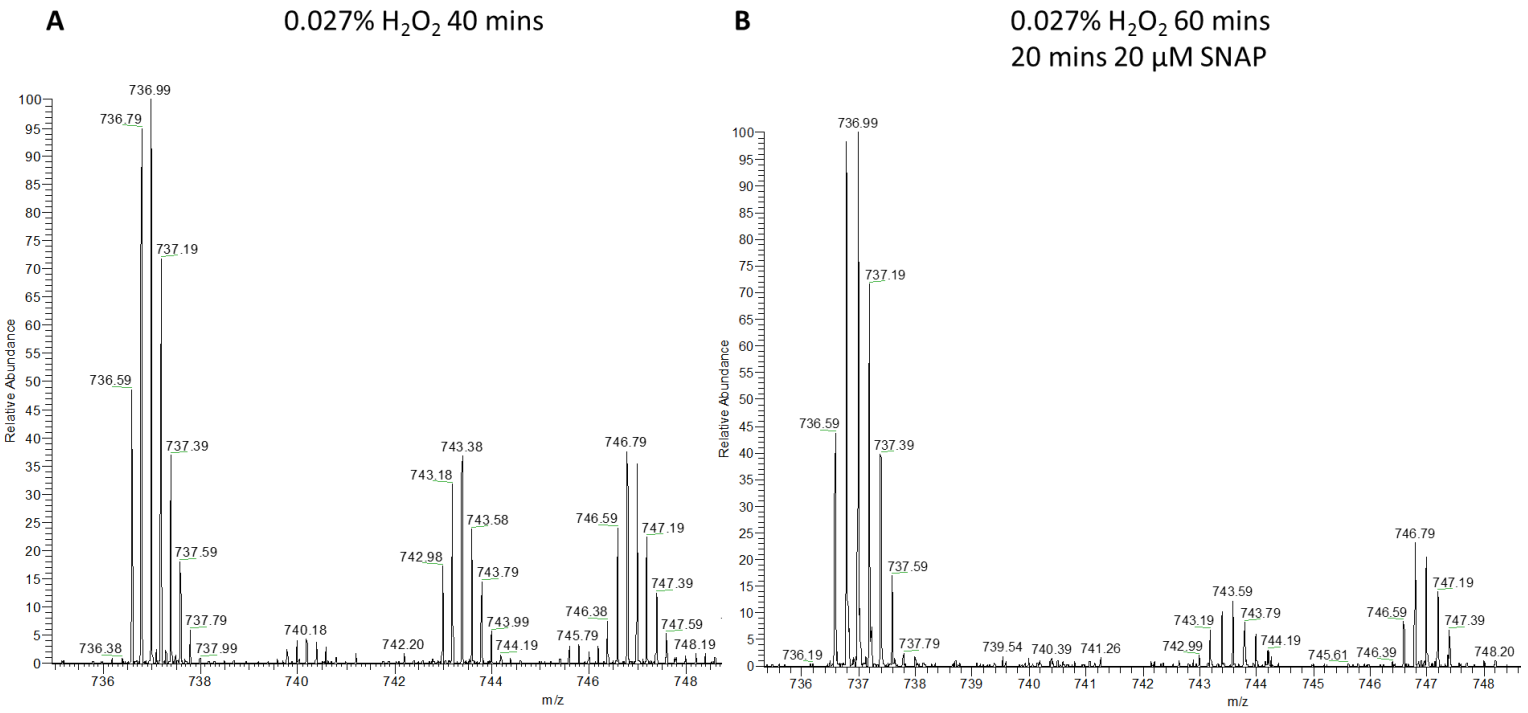
F. MS/MS and MS3 spectra of MS ion species with a monoisotopic mass of *m/z* 742.79 matching that of Cys-nitrosothiol CR-40 synthetic peptide. A. Performing CID (MS2) on the *m/z* 742.99 species generated a single 5^+ ion peak of *m/z* 727.79. B. Performing a third round of CID (MS3) upon the m/z 727.79 ion only *y*-ions from the C-terminus of the CR-40 synthetic peptide were identified.

Y18 (2+)

G. Dual treatment of the synthetic CR-40 peptide with first H_2O_2 and then SNAP. A. Initial treatment of 1 µM CR-40 with 10 µM SNAP for 40 minutes which appeared to cause the nitrosylation of the N-terminal Cys and the m/z 742.79 ion peaks matches that for the Cys-nitrosothiol. B. After 40 minutes of SNAP incubation with CR-40 H_2O_2 was added to a concentration of 0.027% but no noticeable changes to the Cys-nitrosothiol were observed despite the formation of SNAP-CR-40 dimer.



H. Dual treatment of the synthetic CR-40 peptide with first H_2O_2 and then SNAP. A. Initial treatment of 1 µM CR-40 with 0.027% H_2O_2 for 40 minutes induced oxidation of Cys to Cys-sulfenamide (*m/z* 736.59), Cys-sulfonamide (*m/z* 742.99) and Cys-sulfonic acid (*m/z* 746.59). B. After incubating CR-40 with H_2O_2 for 40 minutes the NO donor SNAP was added to a concentration of 20 µM and after 20 minutes no nitrosylative reaction had occurred.



I. Incubation of 1 μM CR-40 synthetic peptide with 100 μM of the NO donor SNP for 3 hours under normoxic conditions appeared to cause oxidation of Cysteine to Cys-sulfonamide and some Cys-sulfonic acid. It is important to note that no nitrosylative species was identified and it is likely that NO released by SNP reacted in solution before reacting with CR-40.

